Demystifying Surrogate Modeling for Circuits and Systems

Mustafa Berke Yelten, Ting Zhu, Slawomir Koziel, Paul D. Franzon, and Michael B. Steer

1. Introduction

Surrogate models are used in grey-box or black-box modeling of a wide variety of systems including electromagnetic modeling of complex structures, geological distributions of minerals, interaction of airflows with airfoils, chemical processes, to name just a few. Only recently have surrogate models been used explicitly in electronics. Surrogate modeling is a macromodeling technique which has solid mathematical foundations and most often, has a strong intuitive linkage to the physical world. So when did surrogate modeling start? One could say that it has always been with us but the term in its current usage was coined by Osio and Amon in 1996 [1]. This work was preceded in 1994 in a paper by Bandler *et al.* in which

Digital Object Identifier 10.1109/MCAS.2011.2181095
Date of publication: 21 February 2012

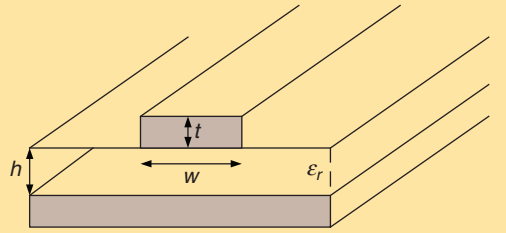

Figure 1. A microstrip transmission line with the strip having thickness t, and width w on a substrate with relative permittivity ϵ_r and thickness h.

a specialized form of surrogate modeling called space mapping was introduced [2]. What the practitioners of surrogate modeling and surrogate model-based design particularly like is what is often described as 'capturing the way engineers think'. Osio and Amon put it this way, surrogate modeling is:

> "an adaptive . . . engineering design methodology for the efficient use of numerical simulations of physical models. These surrogates are nonlinear regression models fitted with data obtained from deterministic numerical simulations using optimal sampling. . . . the formulation of surrogates . . . support[s] the evolutionary nature of engineering design. Information from computer simulations of different levels of accuracy and detail is integrated, updating surrogates sequentially to improve their accuracy."

This formulation enables mathematicians to directly support engineering design with engineers focusing on intuitive insight, intuition which is directed at developing low-fidelity models.

Consider how humans interact with the world [3], [4]—these ideas are called Theories of Mind. One theory is that the human brain is forever trying to fit [surrogate] models to experiences, continuously refining the model fit. We anticipate, make model projections, usually without being consciously aware. So what does this have to do with electronic design and in particular modeling circuits? The engineer's brain uses the model fitting and model projection process that works so well in the rest of our lives. Design of an electronic system begins with a crude model of what we want to accomplish, perhaps the crudest of models materializes on the back of an envelope. Until the very end, the model in our head, on paper, and in computer-aided design tools, is a surrogate model that is continuously refined.

Bandler *et al.* is given credit for introducing a formal surrogate modeling process called Space Mapping [2]. They initially applied it to developing good models of structures that require detailed and exhaustive electromagnetic (EM) modeling. The EM model is called a high-fidelity model since it represents the actual system with great accuracy. Conversely a far more computationally efficient model, but necessarily less accurate, is called the low-fidelity model. The high-fidelity model has also been called the true model [5] or the fine model [6], mainly due to the exactness of the underlying simulation methodology which leads correspondingly to higher computational cost.

Consider the modeling of a microstrip transmission line. The model parameters are the width and thickness of the strip and the height of the substrate as well as material parameters, see Fig. 1. Equation based formulas for the characteristics impedance and propagation constant, which we will call the low-fidelity model, are commonly used for microstrip and these are surprisingly good over a defined range of dimensions and material parameters [7]. However accuracy of this model degrades unacceptably for extremes of parameters. It only takes a small modification of the input parameters, and perhaps a small modification of the low-fidelity model outputs to obtain very good agreement between the high-fidelity model and the surrogate model. The surrogate model now comprises the low-fidelity model and input and output space mappings, which are transformations to correct the low-fidelity model. Crucially the input and output space mappings depend on the input parameters themselves. The mappings are often quite simple, and the parameters controlling the mappings are usually orthogonal, e.g. here the thickness of the substrate is independent of the strip thickness. It is this orthogonality that make this approach so useful and the fitting of the surrogate model to the high-fidelity model such a tractable problem, even for very high dimensionality situations.

The space mapping-based surrogate model described is a grey-box surrogate model in that physical insight is used to develop the physically-based low-fidelity model. Black box surrogate models largely remove the need for physically-based low-fidelity model development. However in a grey-box model features

Mustafa Berke Yelten was with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, 27695 USA; he is now with Intel Corporation, Hillsboro, OR, 97124, USA, e-mail: (mbyelten@ieee.org). Ting Zhu, Paul D. Franzon, and Michael B. Steer are with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, 27695 USA. Slawomir Koziel is with the School of Science and Engineering, Reykjavik University, Iceland.

can be captured such as intermodulation distortion levels that could not be captured with sufficient fidelity by a black-box model.

In this article we describe grey-box and black-box surrogate modeling and present the key findings. The important point is that surrogate modeling has a solid mathematical basis leading to what has become a dramatic increase in our ability to develop engineering models and to engineer systems. In Section 2, a systematic approach to constructing surrogate models is provided. Each step is explained using published methods. Section 3 presents surrogate modeling examples from the domain of circuits and systems.

2. Surrogate Modeling

In this section surrogate modeling is described in terms of black box modeling using a response surface fitting approach. The same approach can be used in developing the input and output space mapping in the grey box form of surrogate modeling.

Surrogate modeling is a concept which is loosely used interchangeably with other concepts such as metamodeling, and response surface modeling. All these concepts refer in essence to modeling a complex system at less computational cost. A complex system has responses which require computer simulations utilizing large amounts of memory and taking excessive time to complete. In surrogate modeling, a 'surrogate' for the high-fidelity model is sought which will be used instead of the high-fidelity model. Obviously, the surrogate should be easier to evaluate compared to the high-fidelity model whereby the accuracy and range of applicability are traded off.

2.1. Surrogate Model Construction

Consider an operational transconductance amplifier with a large number of transistors. The metrics that define the performance of this amplifier can only be found by using SPICE transistor-level simulations that can be computationally expensive. As a remedy engineers mostly use a simple representation of the amplifier by using a transconductance element along with resistors representing the input and output resistances of the amplifier and capacitors that indicate frequency dependence. It is a macromodel that encapsulates the information required for further design steps. However this macromodeling approach can only yield an approximate result with respect to the actual outcome of the complete amplifier circuit. At the expense of increased computation cost more accuracy could be obtained. Surrogate modeling attempts to balance the simplistic macromodels with precise yet computationally intensive computer simulations.

A surrogate model takes a vector of input variables and provides the corresponding output. Each of these vectors is called a sample vector. Sampling is an essential part of the surrogate modeling process and includes two important modeling criteria. The first is the number of elements contained in the sample vector and the second is the range of variables. The number of elements of the sample vectors can substantially change the complexity of the final surrogate models. "The curse of dimensionality" results in surrogate modeling needing more samples as dimensionality increases to maintain the desired accuracy level. This, in turn, affects the second modeling criterion, the ranges of the input variables. A surrogate model representing a high dimensional problem can only describe a certain portion of the entire response surface which translates to smaller ranges of the input variables. Conversely, for problems with lower dimensionality, surrogate models can still accurately represent most of the entire response surface of the actual system.

The trade-off described above leads to the conclusion that a reduced-order scheme is beneficial for the surrogate modeling flow. Reduced-order modeling has been largely investigated in the context of model order reduction of passive networks, interconnect networks, and nonlinear integrated circuits and systems [8]. The main goal of these model-order applications is expressing complex circuits and systems of high order as multiple linear or low-order structures. Various techniques including Krylov subspace projections, implicit moment-matching projections and dimension reduction methods involving Padé approximation, and Volterra and Taylor series have been proposed to achieve the reduction of order [8]. These, however, are not used in surrogate modeling. The main distinction is that in surrogate modeling, reduced-order modeling reduces the input variable space, i.e. uses fewer variables, and the surrogate model is developed as accurately as possible.

Once the number of input variables is chosen then the ranges are determined. Global surrogate modeling that aims to describe the whole possible ranges of the inputs is more computationally expensive to develop and use, and requires a large number of sample vectors. However this number can be optimized by using a well-defined sampling plan. For smaller dimensionality problems classical 2-factorial design of experiments plans can be used that gives the responses of the extreme input values though this approach misses the interactions of parameters covering large ranges. On the contrary uniform sampling plans where the ranges of each input variable are divided in to several smaller units can alleviate this problem. Here the resolution of the sampling plan is determined by the range of each small unit. The final distribution of the sample vectors across the input

Accuracy of the surrogate models should be carefully quantified as this ultimately decides the usefulness of the models developed.

space should be as homogeneous as possible so that the whole response surface of the high-fidelity model can be described by a surrogate model using as few sample vectors as possible. Various sampling techniques have been proposed in the literature that try to achieve this goal using different approaches [9].

After the sample vectors are determined then the surrogate model can be established. The core of the model can take several forms including basis functions, neural networks, and machine learning techniques such as support vector machines. In one approach basis functions are introduced for each input variable with each of the basis functions having one or multiple parameters to be optimized. The optimization is based on the available sample vectors. Choices for basis functions range from polynomial to radial basis functions [10]. Nonlinearity of the basis functions strongly affects the quality of the fit to the response surface of the high-fidelity model. This effectively means that some basis functions can be more successful in representing certain response surfaces than others. It should be remembered that the optimization of basis function parameters requires a global search that becomes computationally expensive as the dimensionality of the problem increases.

A second form of surrogate model has neural network at its core. Neural networks are used in many fields of electrical engineering such as modeling the semiconductor process control [11] where it has been noted that the strong fitting capability of neural networks can enable complex input-output relations to be represented. Neural networks strongly depend on the number of neurons contained in different layers of the network and their connections, as well as on the evaluation methodology for determining the weight coefficients of the connections. The third most common core of surrogate models uses machine learning techniques. These also employ weight vectors that are updated with each new sample. In particular, more attention is focused on support vector machines (SVMs) that are successfully used in regression problems such as the performance modeling of analog circuit blocks [12]. SVMs are constructed as an optimization problem in which a positive definite kernel such as the radial basis function is used to transform the sample vectors to a higher dimensionality [13]. Thus SVMs are also called kernel-based methods.

Normally the high-fidelity model which the surrogate model attempts to fit is computationally expensive to evaluate. This forces the number of sample vectors to be small in order to keep the associated computational cost at a reasonable level. However this number may not be sufficient to attain the required modeling accuracy. In that case an equation-based, approximate physical model can be incorporated in the sample vector evaluation. Such a model is often called a low-fidelity model and is easy to evaluate. It needs to be modulated to faithfully represent the high-fidelity model. In most cases the high-fidelity model and the approximate model are used together. This choice leads to the concept of multifidelity simulation. However it is also possible to include mapping functions along with the low-fidelity model. These mapping functions try to map the input or output variables of the model so that they yield the same input-output relationship as the high-fidelity model input-outputs when applied to the low-fidelity model. The mapping functions can have a variety of forms ranging from a simple linear equation to a highly nonlinear expression or be trained using a neural network. This approach is called the space mapping and often used in the design process of microstrip transmission lines, microwave filters and waveguides [6], [14], where reasonably good low-fidelity models are available.

Accuracy of the surrogate models should be carefully quantified as this ultimately decides the usefulness of the models developed. Test vectors, different from the sample vectors, must be created that will be used to evaluate both the high-fidelity and the surrogate models. The error between the outcomes can be quantified using the root mean square error (RMSE) and root relative square error (RRSE) metrics. More advanced techniques such as cross validation requires building several surrogate models based on a portion of the available sample vectors and use the remaining as a group of test vectors [13]. In case the model does not meet the accuracy criteria, additional samples should be chosen. This process is called sequential sampling. The choice of these samples is specific to the nature of the described response surface as well as the error distribution among the test vectors.

Surrogate modeling has been a focal point for computationally efficient simulation of large-scale systems, in particular in the modeling and design applications of aerospace sciences [15]. Initial attempts at using surrogate modeling in the domain of circuits date back to the space mapping approach [6] and in the domain of

The first step in constructing a surrogate model is choosing the input parameters for the sample vectors.

systems, generating surrogates to detect determinism in time series [16]. The remainder of this section will separately investigate each step of the modeling process based on their effect to the final surrogate model. First the dimensionality reduction techniques will be explained in conjunction with input parameter selection techniques in Subsection 2.2. Subsequently different sampling methodologies will be introduced and their differences are indicated in Subsection 2.3. Core modeling approaches and their accuracy evaluation are described in Subsections 2.4 and 2.5, respectively. The flow of all these steps leading to the construction of a surrogate model is shown in Fig. 2.

2.2. Selection of Input Parameters and Dimensionality Reduction

The first step in constructing a surrogate model is choosing the input parameters for the sample vectors to be used in the evaluation of the high-fidelity/low-fidelity model. Although it is straightforward to directly employ all variables associated with the original problem this can be infeasible if the number of variables is large. A reduced-order approach is then necessary. Several methods have been proposed to deal with this problem and are considered below.

2.2.1. Problem-Specific Approaches

In a problem-specific approach to dimensionality reduction two main topics must be considered. Firstly, which parameters are of interest for the model development. In most cases, although the system to be modeled consists of many parameters, the variation of only a few may need to be included in the surrogate model. This by itself reduces the dimensionality of the problem.

Secondly, the nature and the type of the input parameters can be used to differentiate among them. In a system where both physics-based and empirical parameters are employed concurrently, modeling the system in terms of the physics-based parameters is preferred over the empirical parameters since they do not have a well-defined range and they strictly depend on previous data. This differentiation of the available input parameters leads to the desired dimensionality reduction.

2.2.2. Sensitivity Analysis

A different method of selecting the input parameters is to apply sensitivity analysis on the initial variables. The sensitivity analysis aims to determine how much the output changes to a unit change in each of the input variables. The variables to which the output is more sensitive can be chosen for further analysis. This will also lead to the desired dimensionality reduction [15].

One of the most popular sensitivity analysis methods is the Morris screening analysis (MSA) [17]. MSA does not make any assumptions on the objective function or the possible values that the model parameters can take and it is based on the elementary effects of individual variables. Elementary effects specify how much the output values change in response to an increase of a certain variable given a standardized step size. It is in that sense that it is very similar to the concept of derivative. Algorithmically, for each variable two operations need to be repeated: first calculating the output with all variables at their original values; then the same operation with the variable under investigation increased by the step size keeping the other parameters at their previous levels. Owing to this property, MSA is called one-at-a-time (OAT) design technique. Multiple elementary effects can be utilized by incorporating a randomized sampling matrix.

2.2.3. Principal Component Analysis

A new reduced-order set of parameters can be created by eliminating the correlation between initial variables. This method is called principal component analysis (PCA). PCA can be described as a projection technique where a n-dimensional vector is projected on to a n-dimensional space with $m < n$ [13]. The n-dimensional vector, $\mathbf{x}$ is assumed to be a random vector, and can follow a probability density distribution (pdf) such as a Gaussian distribution. This is an important property and it has formed the basis of the usage of PCA in many fields of electrical engineering including circuit analysis and communications systems with many parameters are distributed with a Gaussian pdf.

The method continues with developing the covariance matrix of the random vector $\mathbf{x}$ which is computed by $\mathbf{R} = \mathrm{E}[\mathbf{x}\mathbf{x}^{\mathrm{T}}]$, i.e. the expected value of $\mathbf{x}\mathbf{x}^{\mathrm{T}}$. The eigenvalues of $\mathbf{R}$ are determined by introducing the unit vectors $\mathbf{q}$ and solving the equation $\mathbf{R}\mathbf{q}_j = \lambda_j \mathbf{q}_j$ with $j = 1, 2, \ldots, n$ [13]. Corresponding to each eigenvalue λ_j, the vector $\mathbf{q}_j$ becomes the eigenvector. When the λ_j s are ordered with respect to their magnitude, the largest

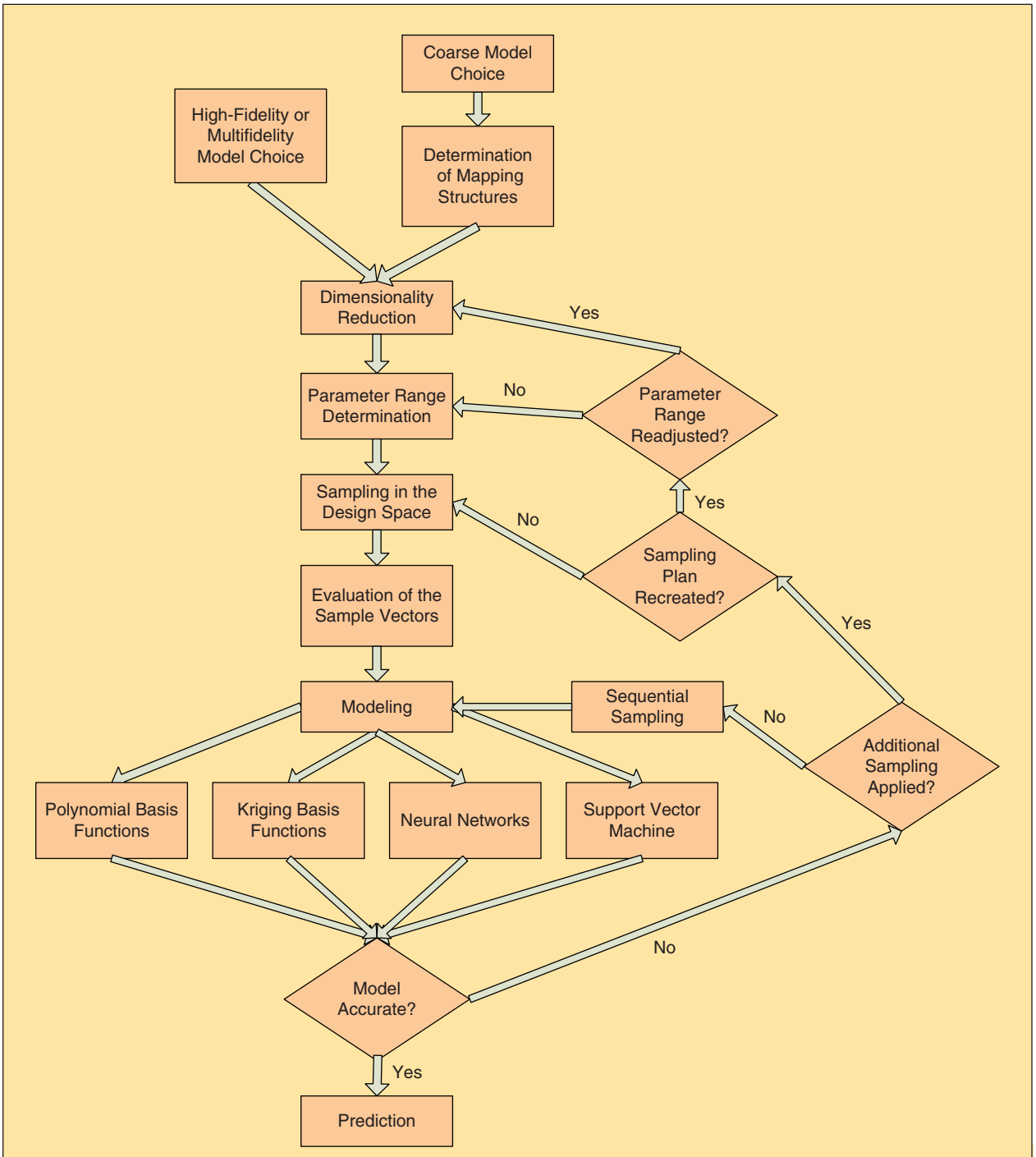

Figure 2. A generic flowchart for the development of a surrogate model.

m of them can be chosen for the dimensionality reduction. So, the projected vector $\tilde{\mathbf{x}}$ of $\mathbf{x}$ can be given as [13].

$$\tilde{\mathbf{x}} = \sum_{j=1}^{m} \mathbf{q}_j^T \mathbf{x} \mathbf{q}_j. \tag{1}$$

This approach is based on the fact that each dimension of the reduced space is orthogonal to the others thus the correlations of the original dimensions are eliminated. Furthermore, dimensions with larger eigenvalues have a bigger impact so that the outcome difference between the original random vector $\mathbf{x}$ and the reduced (or projected vector) $\tilde{\mathbf{x}}$ is minimized. A similar method to PCA is common factor analysis which focuses on the common variance of $\mathbf{x}$ rather than the total variance [18].

2.2.4. Performance-Based Dimensionality Reduction

Performance-based dimensionality reduction can be thought as of a variation to PCA. PCA does not consider the performance metrics of a circuit or system in dimension reduction and thus PCA can eliminate important variables for a certain metric. One approach to dealing with this problem is Principle Hessian Direction-based parameter reduction [19]. Here, the Hessian matrix $\mathbf{H}_\phi(\mathbf{p})$ of a given performance metric ϕ is computed with respect to random sample vector $\mathbf{p}$. Then the eigenvalue problem of

$$\mathbf{R}_\mathbf{p} \mathrm{E}[\mathbf{H}_\phi(\mathbf{p})]\, \mathbf{R}_\mathbf{p}\, \mathbf{q}_j = \lambda_j \mathbf{q}_j, \tag{2}$$

is solved. The rest of the algorithm proceeds similarly to PCA. In this way the sensitivity of ϕ to the input vector $\mathbf{p}$ is also taken into consideration thus making the dimensionality reduction performance-aware.

A similar approach is reduced rank regression [20]. Here the first-order sensitivity matrix $\mathbf{S}$ of the performance metric vector $\mathbf{Y}$ with respect to the input vector $\mathbf{p}$ is calculated using a linear regression such that $\mathbf{Y} = \mathbf{Sp}$. Then, the eigenvalue problem of

$$\mathbf{SR}(\mathbf{p})\mathbf{S}^T \mathbf{q}_j = \lambda_j \mathbf{q}_j \tag{3}$$

is solved. This procedure also follows the PCA approach after this step. Again the impact on performance is accounted for. Depending on the nature of the dimensionality reduction problem, other methods might yield better fitting results. Thus, approaches such as singular value decomposition [21], projection-based performance modeling [22], and slice inverse regression [23] are brought to the attention of the interested reader.

2.3. Sampling

Creating sample and test vectors by sampling across the design space is important to determining the accuracy of the models developed. The samples chosen should uniformly cover the design space. This property is described as the space-filling characteristics of a sampling plan. Various sampling techniques have been suggested to achieve this goal and several will be described here.

2.3.1. Monte Carlo Sampling

Monte Carlo sampling is a straightforward method which only requires a random number generator for each dimension of the input vectors [15]. This property makes it desirable from the computational efficiency perspective. However, for the very same reason, Monte Carlo sampling fails sometimes to faithfully represent the design space as the samples might cluster in certain regions of the design space. Thus more advanced sampling techniques are often preferred.

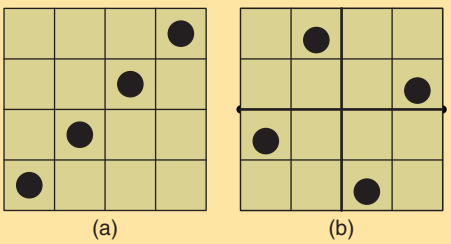

Figure 3. Space-filling: (a) 2D LHS example; and (b) 2D OLHS example introducing orthogonality.

2.3.2. Latin Hypercube Sampling (LHS)

Latin hypercube sampling (LHS) enforces only one sample selected per column per row in a 2D design space [24]. This rule results in improved space filling compared to traditional Monte Carlo sampling. However, it does not guarantee uniformity of sampling across the whole design space. Consider Fig. 3(a) where a 2D example of LHS is depicted. As can be seen, all samples are chosen along one of the diagonals of the square leaving a large design space without a representative sample. An improvement over LHS is accomplished by imposing additional constraints such as partitioning the design space into subspaces and enforcing a certain number of samples in each subspace. This sampling is called orthogonal Latin hypercube sampling (OLHS) where the orthogonality refers to the separation of the design space into smaller subspaces [25]. An example of OLHS in 2D is shown in Fig. 3(b). OLHS can be extended into multiple dimensions [26].

2.3.3. Hammersley Sampling

Hammersley sampling is based on a mathematical formulation. Given that any integer $r \in \mathbb{Z}^+$ can be expressed in terms of a prime number s as $r = b_0 + b_1 s + b_2 s^2 + \ldots + b_z s^z$ where $0 \le b_j \le s-1$, a function φ is defined so that [27]

$$\varphi_s(r) = \frac{b_0}{s} + \frac{b_1}{s^2} + \frac{b_2}{s^3} + \cdots + \frac{b_z}{s^{z+1}}. \tag{4}$$

Using (4), sample points in d dimensional design space are selected as $(r/N_s, \varphi_{s_1}(i), \varphi_{s_2}(i), \ldots, \varphi_{s_{(d-1)}}(i))$ for $s_1 < s_2 < \cdots < s_{(d-1)}$ and $s = 0, 1, \ldots, N_s$ where N_s is the total number of sample points [27]. For a 2D design space optimum space-filling sampling is accomplished using $s = 2$. For larger choices of s, samples tend to be aligned and have an undesirable periodic pattern [27].

The basic reason for the popularity of the Kriging basis functions is their ability to fit nonlinear response surfaces.

2.3.4. Voronoi Tesselation

Voronoi tesselation starts with a small number of samples that are called centers. These are assigned a certain volume of design space for which the boundaries are found on the basis of the distances of encapsulated samples to nearby centers. All sample points in the design space that are closer to one center than any other are contained in the volume of that center. These volumes are called Voronoi cells [28].

Once the Voronoi cells are determined, new samples are chosen adaptively. For this purpose neighboring samples to one center sample are taken and the ones that are as close as possible to the center and as far from each other as possible are chosen. Then the gradient from the center sample to these neighbors is individually estimated. Based on this information the sample with the highest gradient will be chosen as the next sample since a high gradient value indicates a large change in the shape of the response surface. Successful examples of this methodology are in the domain of microwave circuits and systems [28], [29]. However these examples are based on a few dimensions and in general, computational cost associated with gradient estimation might increase disproportionately with higher dimensionality.

Voronoi tesselation with adaptive sampling based on local gradient information has been included in the MATLAB surrogate modeling toolbox SUMO [30].

2.4. Core Modeling Approaches

Various methods have been proposed for the core model in constructing the surrogate model. Here the most prominent of them will be introduced with their associated examples in the circuits and system area.

2.4.1. Polynomial Basis Functions

Polynomial basis functions form the basis of many response surface modeling approaches. The suitability of these basis functions to a particular problem depends on the maximum degree of the chosen polynomial. In general, a second-order multidimensional polynomial basis function is employed [15]:

$$\tilde{y} = c_o + \sum_{i=1}^{k} c_i x_i + \sum_{i=1}^{k} c_{k-1+i+j} x_i x_j, \tag{5}$$

where $j \geq i$ and $\tilde{y}$ is the polynomial surrogate model output for the input vector $\mathbf{x}$. In applying (5), the coefficients are optimized for accuracy. It should be noted that for a k-dimensional input vector space there are $(k+1)(k+2)/2$ coefficients to be estimated. With increased dimensionality of the problem higher order polynomials are needed to increase the fit since quadratic equations generally cannot capture strongly nonlinear response surfaces. However this leads to a large number of coefficients to be optimized. Thus polynomial basis functions have limited use in surrogate modeling. Examples employing polynomial-based surrogate models include design of substrate integrated waveguide interconnect with transitions to conductor backed coplanar waveguides [31], and modeling of certain performance metrics of analog circuits [32].

2.4.2. Kriging Basis Functions

Kriging basis functions are frequently used in surrogate modeling. Their mathematical properties are explained in several review papers [33], [34]. The basic reason for their popularity is their ability to fit nonlinear response surfaces [10]. A Kriging function calculates the correlation C between two sample vectors $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(\ell)}$ [10]:

$$C[\mathbf{x}^{(i)}, \mathbf{x}^{(\ell)}] = \exp\left(-\sum_{j=1}^{k} \theta_j |\mathbf{x}_j^{(i)} - \mathbf{x}_j^{(\ell)}|^{p_j} \right), \tag{6}$$

where p_j and θ_j are parameters to be optimized for each of the k elements of the sample vector j with $j \leq k$. Kriging basis functions are similar to Gaussian basis functions if $p_j = 2\ \forall j$ where the variance of the exponential functions is optimized to realize the greatest accuracy for the given samples and their corresponding outputs. This means that the optimum θ_j values should be found commonly using a global search algorithm such as a genetic algorithm-based search or a pattern search. The latter is successfully applied using the Hooke-Jeeves method in the MATLAB toolbox DACE (Design and Analysis of Computer Experiments) [35].

Given that the total number of sample vectors is N_s, a correlation matrix of size $(N_s \times N_s)$ must be stored with the Kriging-based surrogate model. This can be an important drawback as N_s becomes larger since the model construction and evaluation times also increase with a bigger correlation matrix [26].

Kriging models have been employed in several circuit and system applications. These include the modeling of the input referred noise and admittance matrix elements of a low-noise amplifier [36], circuit performance

Support vectors can be used as basis functions and are reported to be computationally efficient.

modeling of analog circuits in terms of circuit component properties [37], and optimization of integrated circuits to increase their yield by considering process variations [38].

2.4.3. Neural Network-Based Models

Neural networks are one of the most popular function approximation and data classification tools [13]. The first of several layers of neurons is the input layer composed of the sample vectors. The next layer is the hidden layer with each neuron having a nonlinear activation function. There can be several such hidden layers. Finally, the last hidden layer is connected to the output neuron(s). Each connection in the network has a different weight the value of which must be optimized based on the sample inputs and outputs of the high-fidelity model. This optimization can be done using a back-propagation algorithm [13], several of which have been suggested in the literature. The Levenberg-Marquardt back-propagation algorithm is noted to be one of the most computationally efficient algorithms in this category and it is the default back-propagation algorithm in the MATLAB Neural Network Toolbox [13]. Neural networks have been reported to have sufficient accuracy even for high dimensionality problems [39].

There is a variety of surrogate modeling applications that use neural networks in the circuits and systems field. In [40] a high-frequency transformer is synthesized in conjunction with machine learning techniques with the inner diameters and widths of the primary and secondary inductors as inputs. In [41] the large-signal characteristics of microwave devices, e.g. a high electron mobility transistors (HEMTs), are modeled. Finally, variability analysis of analog circuits in terms of important process parameters, transistor temperature, and operating voltages uses neural network-based surrogate models in [26].

2.4.4. Machine Learning-Based Models

Machine learning methods are also an option for the construction of a surrogate model. In particular, support vectors can be used as basis functions and are reported to be computationally efficient [42]. In this case, the aim is to estimate the target response within a given maximum tolerance ϵ using support vectors that are essentially sample vectors. Thus support vector machines (SVMs) convert the regression problem into an optimization problem. In order to solve this optimization problem, a kernel is introduced which increases the dimensionality of the problem, solves it in this new space, and maps the solution back to the actual dimensionality of the problem. This is basically Cover's Theorem [13]. Kernels can take different forms but they should be positive definite [13]. One popular kernel is the radial basis function network that is similar to Gaussian basis functions.

The optimization problem considers a linear regression between the input vector $\mathbf{x}_j$ and the output y_j where $j \leq k$ and k is the dimensionality of the problem. The solution to this problem minimizes the magnitude of the weight vectors for a predefined level of error ϵ. This has been achieved in a computationally efficient way using Least Squares SVM [43] and implemented as a MATLAB toolbox [44].

Applications of SVM-based surrogate modeling in circuits and systems include interconnect modeling for microwave packaging structures [45], performance modeling and design automation of analog circuits [12], and design of microwave antenna structures [46].

2.5. Accuracy Evaluation of Generated Surrogate Models

The accuracy of surrogate models can be tested in several ways. One method develops a separate distinct set test vectors evaluated using both the high-fidelity model and the surrogate model. Like the sample vectors, test vectors should be a representative of the design space to make a correct estimation of the modeling error. Elements of test vectors can be applied to the high-fidelity model e.g. a circuit simulation and to the surrogate model. Differences between these two characterize the error as, typically, the root means square error (RMSE) and the root relative square error (RRSE) [5]:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{n=1}^{N_T}\left[y_{\mathrm{model,n}} - y_{\mathrm{true,\,n}}\right]^2}{N_T}}, \tag{7}$$

$$\mathrm{RRSE} = \sqrt{\frac{\sum_{n=1}^{N_T}\left[y_{\mathrm{model,n}} - y_{\mathrm{true,n}}\right]^2}{\sum_{n=1}^{N_T}\left[y_{\mathrm{true,n}} - \frac{1}{N_T}\sum_{n=1}^{N_T} y_{\mathrm{true,n}}\right]^2}}. \tag{8}$$

RMSE expresses the error in terms of the unit of the modeled quantity whereas RRSE determines it relatively. Depending on the context of the accuracy evaluation, both are useful metrics.

It should be remembered that many high-fidelity models are expensive to evaluate and thus having a separate set of test vectors to the sample vectors might not be feasible. Considering this, the available sample vectors can be divided in two and one subgroup used to create the model whereas the other subgroup is used to test the model. This is called the cross validation method [13]. Cross validation can be repeated several times while building a model with different selections for the training subgroup and testing subgroup. It has been reported that using cross validation in neural network-based modeling significantly improves the generalization of the model over the whole design space [13].

A further approach to calculate the accuracy of the surrogate models is the hypothesis testing [16]. Using a common statistic measure such as the correlation dimension estimation the difference between the high-fidelity and surrogate-based time series can be found. The main advantage of this method is the incorporation of transient effects into the accuracy evaluation of surrogate models. This is in contrast with the RMSE-based (or RRSE-based) accuracy evaluation, which considers the models at the same time instant. Thus hypothesis testing was employed particularly in nonlinear dynamic systems [16]. It should be emphasized that the choice of the statistic is crucial for a correct accuracy estimation.

3. Application of Surrogate Modeling to Circuits

In this section, three examples of surrogate modeling drawn from across the broad field of circuits and systems are explained.

3.1. Surrogate-Based Optimization and Space Mapping

3.1.1. Surrogate-Based Optimization Concept

The generic design problem in electrical engineering can usually be formulated as a nonlinear minimization problem of the form

$$\mathbf{x}^* = \arg[\min_{\mathbf{x}} f(\mathbf{x})], \tag{9}$$

where $f(\mathbf{x})$ denotes the objective function to be minimized evaluated at the point $\mathbf{x} \in R^n$, a vector of designable parameters. An optimal design vector is denoted by $\mathbf{x}^*$. In many problems f is of the form $f(\mathbf{x}) = U(R_f(\mathbf{x}))$, where $R_f \in R^m$ denotes the response vector of the system (the high-fidelity, fine, or true model) [47], whereas U is a given scalar merit function.

In practice, evaluation of f is often computationally expensive (e.g., when f is obtained through CPU-intensive computer simulation) and its analytical properties are poor (e.g. due to discontinuity or numerical noise). In such cases, the use of conventional optimization algorithms is prohibitive because numerical optimization typically requires a large number of function evaluations and, in many cases (e.g. gradient-based algorithms), require that f is sufficiently smooth.

Surrogate-based optimization (SBO) [6], [48] has been suggested as an effective approach for solving challenging design problems as described above. The basic concept of SBO is that that the direct optimization of the high-fidelity model is replaced by an iterative process that involves the creation, optimization and updating of a fast and analytically tractable surrogate model. The generic SBO algorithm can be formulated as the iterative procedure [47]

$$\mathbf{x}^{(i+1)} = \arg\{\min_{\mathbf{x}}[s^{(i)}(\mathbf{x})]\}, \tag{10}$$

where $s^{(i)}$ is the surrogate model at iteration i, whereas $\mathbf{x}^{(i)}$, $i = 0, 1, \ldots$ is a sequence of approximate solutions to (9). The surrogate should be a reasonably accurate representation of the high-fidelity model, at least locally. The design obtained through optimizing the surrogate model is verified by evaluating the high-fidelity model at each new design $\mathbf{x}^{(i+1)}$. The high-fidelity model data obtained in this verification process is then used to update the surrogate. SBO proceeds in this predictor-corrector fashion iteratively until some termination criterion is met. Because most of the operations are performed on the surrogate model, SBO reduces the computational cost of the optimization process when compared to optimizing the high-fidelity model directly (without resorting to surrogate). Convergence of $\{\mathbf{x}^{(i)}\}$ to $\mathbf{x}^*$ can be guaranteed assuming sufficient consistency between the high-fidelity model and the surrogate model [49], or when the surrogate model is a sufficiently good representation of f [47]. In practice the algorithm, (10), is often executed until a satisfactory design is found.

3.1.2. Approximation and Physics-Based Surrogate Models

The surrogate model is a key component of any SBO algorithm. Practical SBO techniques differ in the way the surrogate model is set up and then exploited in the optimization process. There are two main approaches to constructing the surrogate model: by approximating the sampled high-fidelity model data; or through suitable correction of a physically-based low-fidelity (or "coarse") model. The drawback of approximation models in the SBO context is that the computational cost of creating approximation-based models may be quite substantial, particularly when the number of design variables is large, which may be prohibitive in practice.

The space mapping surrogate model is constructed by suitable correction of the low-fidelity model $\mathbf{R}_c$.

Another way of creating the surrogate, by using the physics-based low-fidelity model, is typically much more efficient [6]. The low-fidelity model $\mathbf{R}_c$ represents the same structure or device as $\mathbf{R}_f$. It is less accurate but normally much cheaper to evaluate. A good example of $\mathbf{R}_f$ and $\mathbf{R}_c$ is a microwave filter evaluated using full-wave EM simulation ($\mathbf{R}_f$) and its equivalent circuit ($\mathbf{R}_c$). The knowledge about the system embedded in $\mathbf{R}_c$ allows the creation of a relatively accurate surrogate model by using a limited amount of $\mathbf{R}_f$ data. Specific ways of exploiting low-fidelity models to create the surrogate are discussed in the next subsection.

3.1.3. Surrogate-Based Optimization Algorithms

The simplest realization of the SBO algorithm, (10), is a one-step process where the surrogate model is constructed and optimized only once with $\mathbf{x}^{(1)}$ obtained from (10) considered to be the final design. This approach requires that the surrogate is (usually globally) accurate. It is typically performed in conjunction with approximation surrogates, e.g. [50]. One of the most popular techniques used in SBO is space mapping (SM) [6], [47]. The space mapping surrogate model is constructed by suitable correction of the low-fidelity model $\mathbf{R}_c$. The generic form of the SM surrogate is

$$s^{(i)}(\mathbf{x}) = U(\mathbf{R}_s(\mathbf{x};\, \mathbf{p}^{(i)})), \tag{11}$$

where $\mathbf{p}^{(i)}$ are model parameters that describe the type of the correction performed. Space mapping low-fidelity model corrections can be categorized into four groups [47]:

1) Input SM [6]. The correction is based on an affine transformation on the low-fidelity model parameter space. Example: $\mathbf{R}_s(\mathbf{x};\mathbf{p}) = \mathbf{R}_s(\mathbf{x};\mathbf{B},\, \mathbf{c}) = \mathbf{R}_c(\mathbf{Bx} + \mathbf{c})$).
2) Output SM [47]. The correction is based on an affine transformation on the low-fidelity model response. Example: $\mathbf{R}_s(\mathbf{x};\mathbf{p}) = \mathbf{R}_s(\mathbf{x};\mathbf{A},\mathbf{d}) = \mathbf{AR}_c(\mathbf{x}) + \mathbf{d}$.
3) Implicit SM. In some cases, there are additional parameters $\mathbf{x}_p \in R^{np}$ in the low-fidelity model response $\mathbf{R}_c(\mathbf{x};\, \mathbf{x_p})$ that can be tuned for better aligning of the high-fidelity and low-fidelity model responses. Example: $\mathbf{R}_s(\mathbf{x};\, \mathbf{p}) = \mathbf{R}_s(\mathbf{x};\, \mathbf{x}_p) = \mathbf{R}_c(\mathbf{x};\, \mathbf{x}_p)$. These additional parameters are known as preassigned parameters, and are in general different from the optimization variables $\mathbf{x}$.
4) Custom corrections that exploit the structure of the given design problem [6]. In many occasions the model responses are obtained through the sweeping of some parameter t:

$$\mathbf{R}_f(\mathbf{x}) = \mathbf{R}_f(\mathbf{x};\, t) = [R_f(\mathbf{x};\, t_1) R_f(\mathbf{x};\, t_2) \ldots R_f(\mathbf{x};\, t_m)]^{\mathrm{T}} \tag{12}$$

$$\mathbf{R}_c(\mathbf{x}) = \mathbf{R}_c(\mathbf{x};\, t) = [R_c(\mathbf{x};\, t_1) R_c(\mathbf{x};\, t_2) \ldots R_c(\mathbf{x};\, t_m)]^{\mathrm{T}}. \tag{13}$$

Examples of this situation appear when the parameter t represents time or frequency [6]. The response correction considered in this case could be based on an affine transformation on the sweeping parameter space $\mathbf{R}_s(\mathbf{x},\, \mathbf{p}) = \mathbf{R}_s(\mathbf{x};\, r_0;\, r_1) = \mathbf{R}_c(\mathbf{x};\, r_0 + r_1 t)$.

In Fig. 4 this is illustrated by means of block diagrams of the four SM-based correction strategies introduced above, together with a combination of three of them.

The surrogate response is usually optimized with respect to the SM parameters $\mathbf{p}$ in order to reduce the model discrepancy for all or part of the data available $\mathbf{R}_f(\mathbf{x}(0))$, $\mathbf{R}_f(\mathbf{x}(1)))$, . . . , $\mathbf{R}_f(\mathbf{x}(i))$:

$$\mathbf{p} = \arg \min_{\mathbf{P}} \sum_{k=1}^{p} \omega^{(k)} \| \mathbf{R}_f(\mathbf{x}^{(k)}) - \mathbf{R}_s(\mathbf{x}^{(k)};\, \mathbf{p}) \|, \tag{14}$$

where $0 \le \omega^{(k)} \le 1$ are weights for each of the samples. The corrected surrogate $\mathbf{R}_s(\mathbf{x};\, \mathbf{p})$ can be used as an approximation to the high-fidelity response $\mathbf{R}_f(\mathbf{x})$ in the vicinity of the sampled data. The minimization in (14) is known in SM literature as parameter extraction [6].

A recently introduced specialized version of space mapping is tuning SM [51], where the surrogate model is a combination of EM-simulated data and circuit-based tuning components. Tuning SM is extremely efficient, however, it is an invasive method that requires "cutting" of the high-fidelity model structure (e.g., through the co-calibrated ports technology of Sonnet ***em*** [52]). Its application range includes mostly planar microwave filters [51].

In general, the surrogate model for the SBO algorithm, (10), can be obtained by correcting the response of the low-fidelity model, e.g. by composing $\mathbf{R}_c$ with a suitable function $\mathbf{C} : R^m \to R^m$:

$$\mathbf{R}_s(\mathbf{x}) = \mathbf{C}(\mathbf{R}_c(\mathbf{x})). \tag{15}$$

Output SM [47] is a simple example of such a procedure. Other examples include Approximation Model

Management Optimization [53], popular in aerospace engineering. An elegant way of realizing response correction for vector-valued models is manifold mapping [54], where the surrogate model is constructed as

$$\mathbf{R}_s^{(i)}(\mathbf{x}) = \mathbf{R}_f(\mathbf{x}^{(i)}) + \mathbf{S}^{(i)}(\mathbf{R}_c(\mathbf{x}) - \mathbf{R}_c(\mathbf{x}^{(i)})), \tag{16}$$

where $\mathbf{S}^{(i))} = \Delta\mathbf{F}\Delta\mathbf{C}^{\dagger}$ is the $m \times m$ matrix with

$$\Delta\mathbf{F} = [\mathbf{R}_f(\mathbf{x}^{(i)}) - \mathbf{R}_f(\mathbf{x}^{(i-1)}) \cdots - \mathbf{R}_f(\mathbf{x}^{\max\{i-n,\,0\}})], \tag{17}$$

$\Delta\mathbf{C}$ is defined analogously,

$$\Delta C^{\dagger} = \mathbf{V}_{\Delta\mathbf{C}}\,\Sigma_{\Delta\mathbf{C}^{\dagger}}\mathbf{U}_{\Delta\mathbf{C}^{T}}, \tag{18}$$

and where $\mathbf{V}_{\Delta\mathbf{C}}$, $\Sigma_{\Delta\mathbf{C}}$, and $\mathbf{U}_{\Delta\mathbf{C}^{T}}$ are the factors in the singular value decomposition of $\Delta\mathbf{C}$. The matrix $\Sigma_{\Delta\mathbf{C}^{\dagger}}$ is the result of inverting the nonzero entries in $\Sigma_{\Delta\mathbf{C}}$, leaving the zeros invariant [54]. The surrogate, (12), ensures that $\mathbf{R}_s^{(i)}(\mathbf{x}^{(k)}) = \mathbf{R}_f(\mathbf{x}^{(k)})$ for $k = \max\{i - n, 0\}, \ldots, i$.

Recently several response correction techniques were introduced in electrical engineering that are focused on more efficient exploitation of system knowledge embedded in the low-fidelity model. One of these approaches is shape-preserving response prediction (SPRP) [55]. SPRP may be more efficient than SM, however it is not as generic because additional assumptions regarding the similarity between the low- and high-fidelity model responses are necessary for it to work properly.

It should be emphasized that the SBO framework, (10), relies on updating the surrogate model using the high-fidelity data at the new design found by optimizing the current surrogate (typically using a local search method). This is only one possible way of finding the so-called infill points and corresponds to exploitation of a given region of the design space. In general, exploration of the space implies a global search by performing global sampling, e.g., by selecting those points that maximize some estimation of the error associated with the surrogate considered [56]. In practice, pure exploration may not be efficient (particularly if the number of design variables is large) and there should be a balance between exploitation and exploration. As suggested in [56], this trade-off could be formulated in the context of surrogate-based optimization, for example, by means of a bi-objective optimization problem (with a global measure of the error associated with the surrogate as a second objective function) by maximizing the probability of improvement of the best observed objective function value, or through the maximization of the expected cost function improvement.

3.1.4. Example: Space Mapping Optimization of a Fourth-Order Ring-Resonator Bandpass Filter

Consider the fourth-order ring resonator bandpass filter [57] shown in Fig. 5. The design parameters are $\mathbf{x} = [L_1\, L_2\, L_3\, S_1\, S_2\, W_1\, W_2]^{\mathrm{T}}$ mm. The high-fidelity model $\mathbf{R}_f$ is simulated in FEKO [58]. The low-fidelity model, Fig. 6, is implemented in Agilent ADS [59]. The design specifications are $|S_{21}| \geq -1$ dB for 1.75 GHz $\leq \omega \leq$ 2.25 GHz,

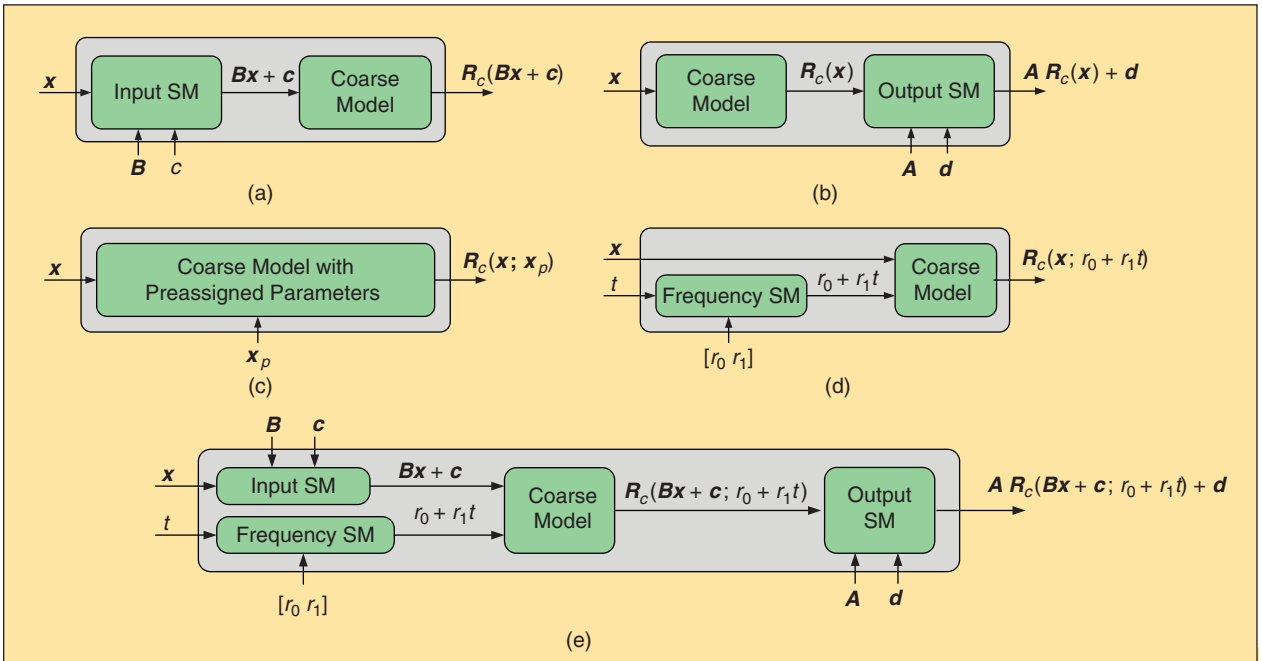

Figure 4. Basic space-mapping surrogate correction types: (a) input SM, (b) output SM, (c) implicit SM, (d) frequency SM, and (e) composite using input, output and frequency SM.

Surrogate models are also used in developing accurate flexible macromodels for digital input-output (IO) buffers.

and $|S_{21}| \leq -20$ dB for $1.0\ \text{GHz} \leq \omega \leq 1.5\ \text{GHz}$ and $2.5\ \text{GHz} \leq \omega \leq 3.0\ \text{GHz}$.

The initial design,

$$\mathbf{x}^{(0)} = [24.74\ 19.51\ 24.10\ 0.293\ 0.173\ 1.232\ 0.802]^{\text{T}}\ \text{mm},$$

is the optimum design using the low-fidelity model. The filter is optimized using the space mapping algorithm working with the input SM surrogate of the form $\mathbf{R}_s^{(i)}(\mathbf{x}) = \mathbf{R}_c(\mathbf{x} + \mathbf{c}^{(i)})$ enhanced by implicit SM. Substrate heights and relative permittivities of the microstrip line segments corresponding to L_1, L_2 and L_3 are used as preassigned parameters for implicit SM [60]. The standard SM algorithm is endowed with constrained parameter extraction and adaptively constrained surrogate model optimization [60], which improves convergence properties of the SM optimization process. The final design

$$\mathbf{x}^{(*)} = [22.5\ 19.68\ 27.00\ 0.206\ 0.100\ 1.232\ 0.812]^{\text{T}}\ \text{mm},$$

is obtained after 5 SM iterations (6 evaluations of the high-fidelity model). The minimax specification error at this design is -0.45 dB. The final design is compared to the design based on the uncorrected low-fidelity model in Fig. 7.

3.2. PVT-Aware Surrogate Model-Based IBIS Macromodels

In this section, surrogate models are used in developing accurate and flexible macromodels for digital input-output (IO) buffers. The most popular IO buffer model is the table-based input-output buffer information specification (IBIS) [61]. IBIS models are simple, portable, intellectual property (IP)-protected, and fast in simulation. However they are not able to capture

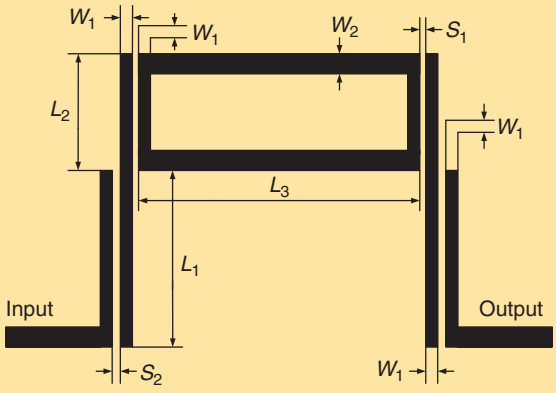

Figure 5. Fourth-order ring-resonator bandpass filter: geometry [57].

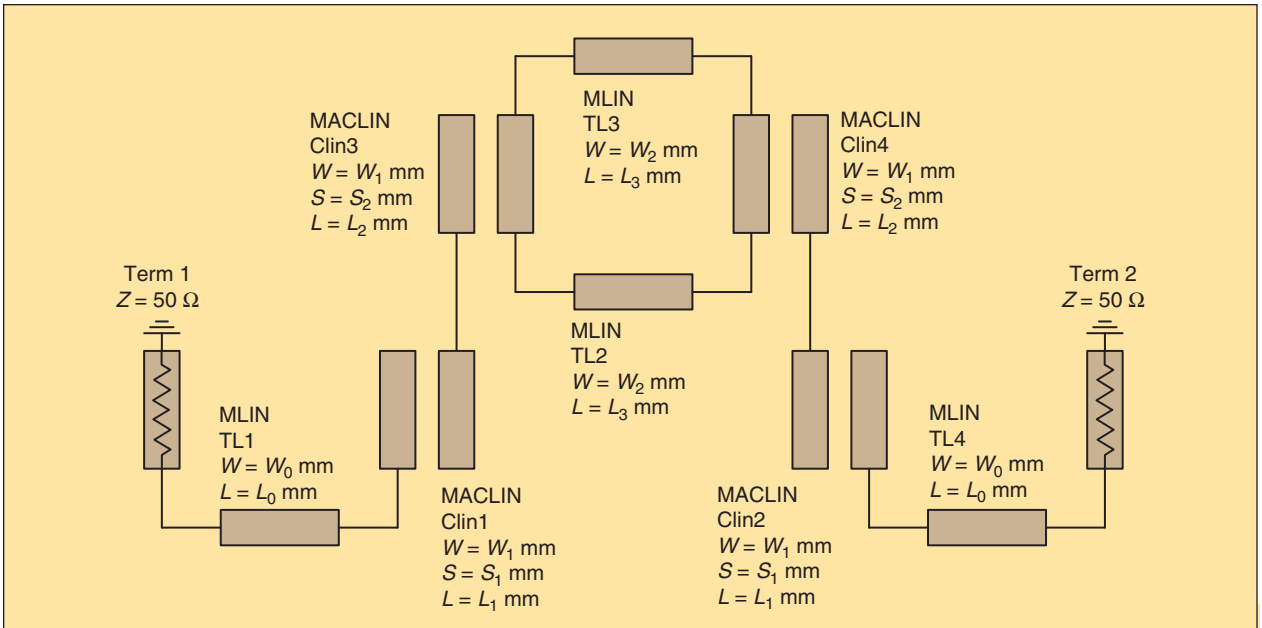

Figure 6. Fourth-order ring-resonator bandpass filter: low-fidelity model (Agilent ADS).

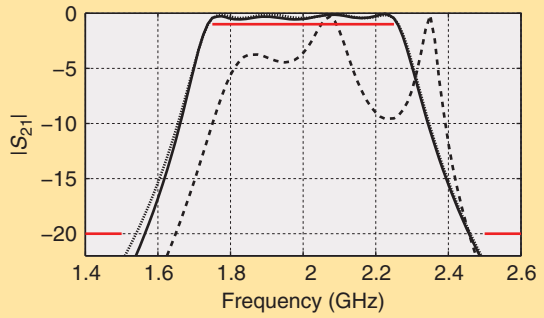

Figure 7. Fourth-order ring resonator filter: response of the high-fidelity model (---) and the low-fidelity model (···) at the initial design, and the response of the high-fidelity model at the final design found by the constrained SM algorithm (–) [60].

process variations and are unsuitable for statistical analysis. To solve these problems, a surrogate IBIS model was developed [62]. In this model a physically-based equivalent circuit structure, the low-fidelity model, is used to capture static and dynamic circuit behaviors, while surrogate modeling is used to approximate each element over a range of Process-Voltage-Temperature (PVT) parameters.

The key benefit of the surrogate model is that it is able to handle multi-dimensional parameters and highly nonlinear circuit characteristics. Also, it is scalable over continuous PVT parameters and all of the parameters are directly accessible in simulation. In addition, the model stores data in numerical expressions which are portable and IP-protected, reducing both data size and the uncertainty caused by *ad hoc* interpolations.

3.2.1. Surrogate Model Structure

Fig. 8 shows the structure of the surrogate IBIS model. Elements I_{pu} and I_{pd} represent the nonlinear output current. The time-variant K_{pu} and K_{pd} functions determine the partial turn-on of the pull-up/down networks during switching transitions. C_{power} and C_{gnd} represent the nonlinear parasitic capacitances between the output and the supply rails.

In this example the basic structure of the IBIS model is fixed by the standard and enables efficient noniterative circuit simulation with buffers separated by the time-delay inherent to interconnects. As well the structure enables physical measurements at external terminals, the response of the high-fidelity model. Similarly exterior numerical measurements of transistor-level simulations must be employed since often the circuit-level details of the buffer are hidden because of IP concerns. Here, data from transistor-level SPICE circuit simulations, with the test benches of Fig. 9, are used as the high-fidelity response.

The responses enable models of each of the boxes in Fig. 9 to be developed independently. Details of the extraction process are given in [62]. In summary the 2EQ/2UK algorithm is used to obtain model coefficients [63].

At the start of the process, a small set of initial sample vectors is generated. Transistor-level SPICE simulations are run using this initial set, and the corresponding responses are recorded. The surrogate models are then constructed and the model accuracy is evaluated. If the desired accuracy is not obtained, the adaptive sampling method creates new sample vectors. The process continues until the fit reaches the targeted accuracy. Modeling construction was performed using the SUMO Toolbox [30]. The surrogate models were implemented in Verilog-A based on the standard table look-up IBIS model [64]. The tables were replaced by the surrogate models implemented.

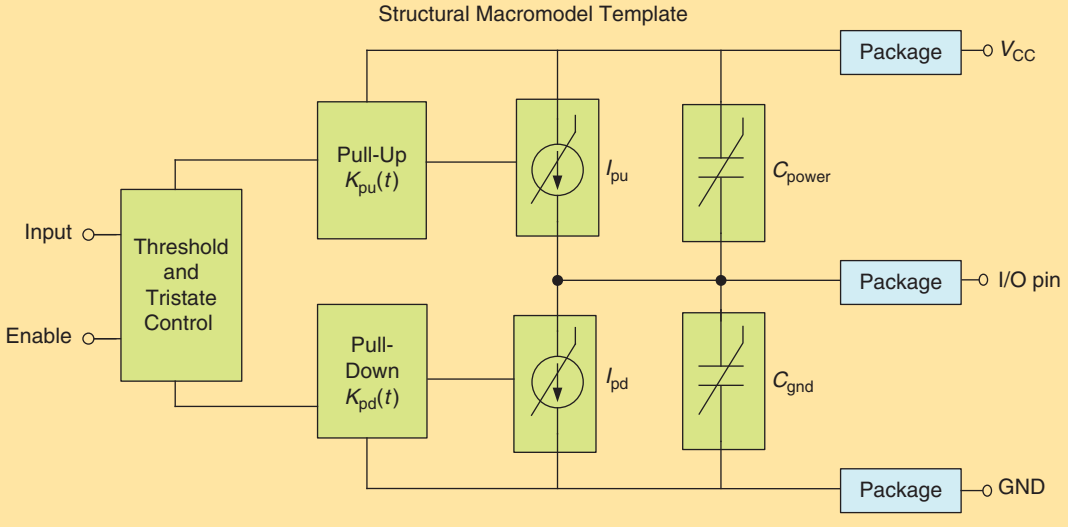

Figure 8. Structure of surrogate IBIS model [62].

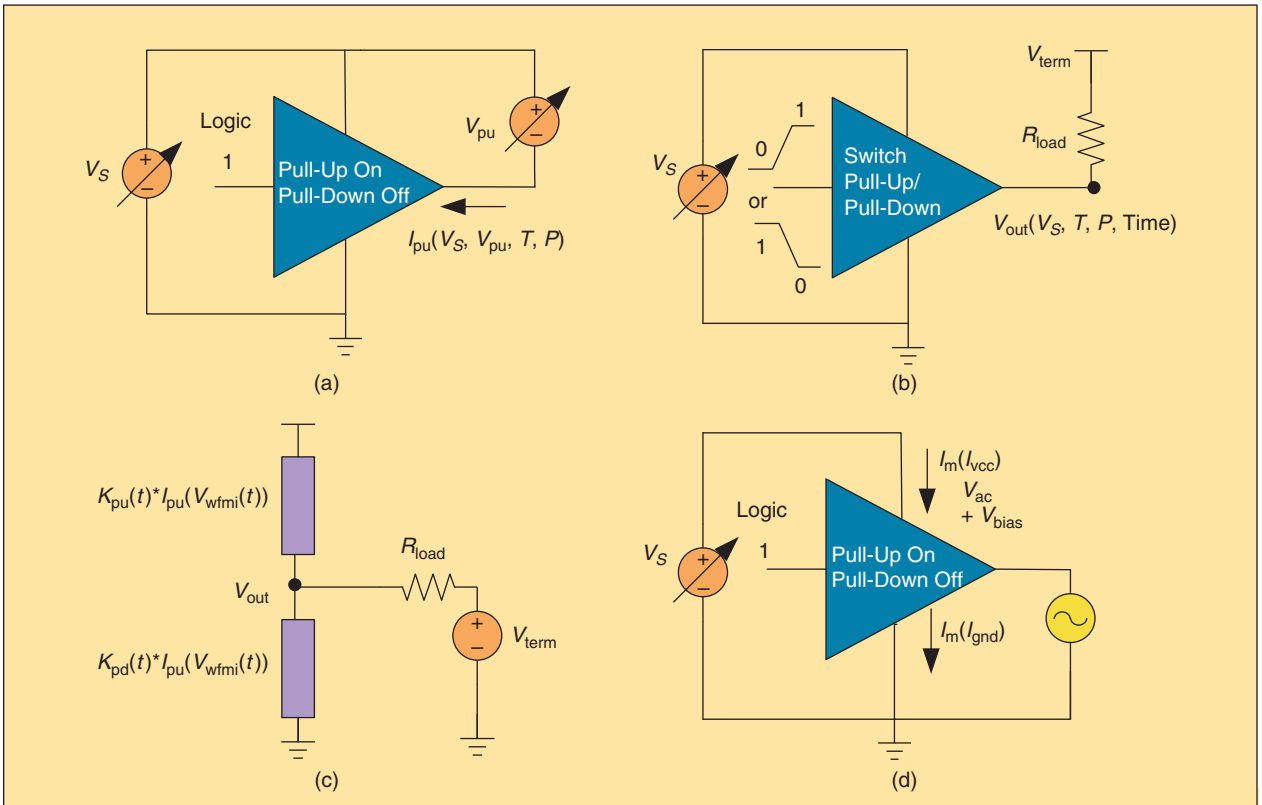

Figure 9. Test-benches for extracting model elements: (a) pull-up current I_{pu} (b) rising/falling transition waveforms for K_{pu} and K_{pd} (c) illustration of 2EQ/2UK algorithm (d) output capacitance C_{gnd} and C_{power} [62]. Here T is the temperature and P is the process parameter.

The proposed method is demonstrated on a single-ended output buffer circuit shown in Fig. 10. The circuit is designed in 180 nm CMOS process with a 3.3 V normal supply voltage (V_{CC}). For simplification, the threshold voltage variations of the MOS transistors are considered as the main process variations, thus using physical insight to reduce the dimensionality of the surrogate model. Other elements of the sample vectors are the applied voltages and the temperature.

The core of the surrogate model is chosen as rational functions model having the form:

$$f(\mathbf{x}) = \frac{P(\mathbf{x})}{Q(\mathbf{x})}, \tag{19}$$

where P and Q are polynomial functions in $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$. The model parameters are obtained by solving a nonlinear least squares fitting problem.

An adaptive sampling technique is used to achieve efficient sampling of the multi-dimensional variable space [28]. The RRSE error of the surrogate model is evaluated using the five-fold cross-validation technique [10]. The data pattern for this study is chosen to be a 1024 bit long pseudorandom bit sequence (PRBS) with 2 ns bit time. The circuit simulations are done in HSPICE. Fig. 11 shows the responses at the far-end of the transmission line under the nominal PVT condition. The maximum timing error is found to be 70 ps (3.5% of the bit-time) and the maximum relative voltage error has been calculated as 6.45%.

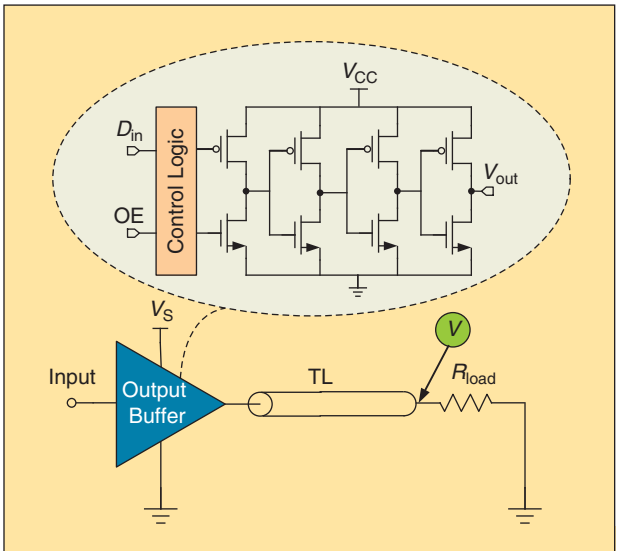

Figure 10. Buffer circuit with a 0.75 m transmission line (characteristic impedance 50 Ω) and a loading resistor of 75 Ω.

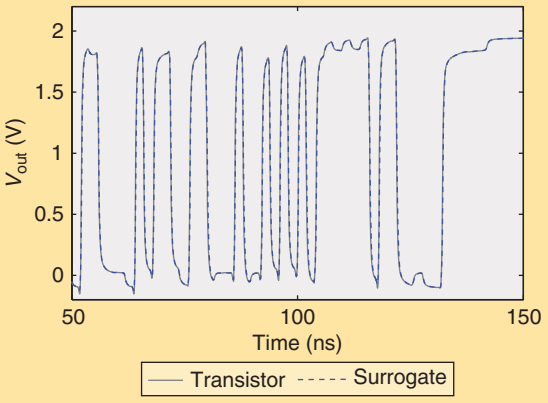

Figure 11. Output voltage at far end of transmission line when input 1024 bit-long PRBS (partial). Grey solid line: transistor, black dashed line: macromodel.

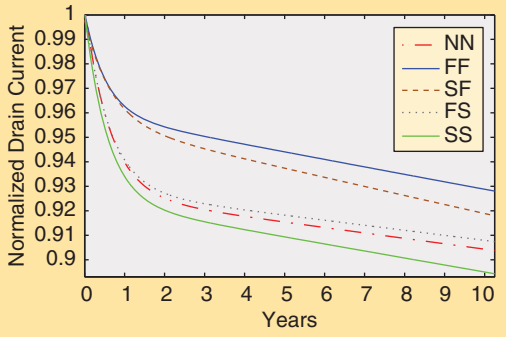

Figure 13. The degradation of I_1 for the circuit in Fig. 12 at different process corners over the time [72].

3.3. Surrogate Model-Based Variability and Reliability Analysis of Analog Circuits

Variability and reliability analysis is especially important for nanoscale circuits. Although the device sizes are continuously scaling down, the applied voltages are not reducing at the same pace. This increases the electric fields along and across the channel of transistors. Traps generated at the interface of the device channel and gate dielectric lead to degradation of effective mobility and increased threshold voltage over time, i.e. aging. Specifically the aging phenomena are called hot carrier injection (HCI) [65], [66] and negative bias temperature instability (NBTI) [67].

Furthermore manufacturing tolerance has not improved at the same rate as device sizes scaled down [68]. A direct consequence is the increased variability of process parameters and affecting the drain-source current I_{ds} and the small-signal transconductance g_m.

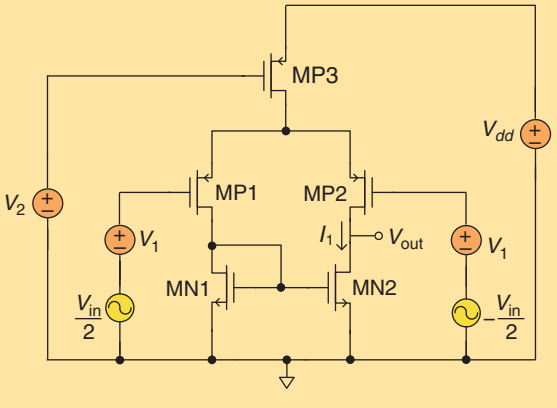

Figure 12. Actively loaded pMOS-input differential amplifier, circuit parameters are given in [72].

In particular, within-die variations have emerged as the source of variability that is most difficult to deal with due to their random nature and contribution to the uncertainty of major process parameter values such as the effective gate length L_{eff} and intrinsic threshold voltage V_{th0} [69]. The well-established techniques of process corner analysis and Monte Carlo analysis [70] do not sufficiently cover the parameter space [5], [71], [72].

Surrogate models have been proposed to solve this problem where an eleven element input vector defines the surrogate model of each transistor [5]. The numerous other transistor parameters are fixed at their designed values as they are not subject to aging and the most significant process variations. The DC drain-source current I_{ds} was modeled based on BSIM device equations [73] and 65 nm process parameters (IBM 10SF). Six of these elements are the process parameters remaining following elimination of empirical parameters and the physical parameters whose variation have little impact on I_{ds}. So the dimensionality of the surrogate model is reduced from that of the high-fidelity (transistor) model. The process uses engineering insight as well as sensitivity analysis of a single transistor. While all transistor parameters are used (and important) in establishing a design, only a few impact I_{ds} and g_m variations from that of the nominal design. Additional surrogate model parameters to these are the three terminal voltages, the operating temperature T, and the device age t_{age} completing the input sample and test vector. Aging effects on I_{ds} are incorporated using analytical expressions for HCI and NBTI [66], [72]. Thus a surrogate model can capture more than the original high-fidelity model (here the transistor model) by including additional information (e.g. the aging model equations). The ranges of input parameters were chosen based on the foundry process file, and the typical limits of terminal voltages, temperature, and the expected device age [72]. The

surrogate model creation used 2880 sample vectors that are generated using OLHS. Following this, 3200 test vectors are used to evaluate the accuracy of the models with RRSE $< 3\%$ [72].

Thus surrogate model-based analysis combines electrical circuit simulation with variability and reliability analysis. As an example consider aging of the differential amplifier circuit in Fig. 12. The combined variability and reliability analysis of the circuit is shown in Fig. 13 where aging at different process corners is considered. Here NN, FF, SF, FS, and SS represent the nominal, fast-fast, slow-fast, fast-slow, and slow-slow process corners of nMOS and pMOS transistors, respectively. Each of the transistors in Fig. 12 ages at a different rate depending on biasing and signal history as well as the process parameters all of which is captured in surrogate model-based analysis. The variations of these responses over the process corners indicate the effect of process variability on aging degradation mechanisms.

4. Conclusion

Surrogate modeling can be an unnecessarily confusing topic largely because technical papers must focus on their precise contribution. Those who practice surrogate modeling in engineering design are committed advocates. Surrogate modeling closely matches the engineering design process. There are different styles of surrogate models, many with greatly reduced dimensionality compared to their corresponding high-fidelity or true model. In the design optimization process the surrogate model can be continuously refined providing increasingly accurate modeling, and thus design, as the final design point is converged upon. Surrogate models can combine information from various sources including the response of detailed high-fidelity models calculations, indirect observations (e.g. aging data), and engineering intuition. Demystifying surrogate modeling should encourage greater use of this technique in circuit design.

Acknowledgment

The paper has the distribution statement "A" (Approved for Public Release, Distribution Unlimited). The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. This work was supported by the Self-HEALing mixed-signal Integrated Circuits (HEALICs) program of the Defense Advanced Research Projects Agency (DARPA) and the prime contractor Raytheon Company (Contract number: FA8650-09-C-7925).

Mustafa Berke Yelten (S'03–M'12) was born in Istanbul, Turkey, in 1982. He received the B.S. degree in electrical engineering from Bogazici University, Istanbul, Turkey in 2006 and the M.S. and Ph.D. degrees in electrical engineering from North Carolina State University, Raleigh, NC, USA in 2008 and 2011, respectively. Since October 2011, he has been with Intel Corporation, Hillsboro, OR, USA.

During his M.S. thesis studies he worked on the design methodologies for tunable low noise amplifiers employing variable capacitors. In his Ph.D. dissertation, he focused on the computer-aided design and mathematical modeling techniques for the variability and reliability analysis of analog circuits. His research interests include the design, optimization, modeling and simulation of variability- and reliability-aware analog/RF as well as digital and mixed-signal integrated circuits.

Ting Zhu (S'09) received the B.S. and M.S. degrees in Electrical Engineering from Zhejiang University, Hangzhou, China, in 2003 and 2006, respectively. She is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC.

In the summer of 2005, she was an intern IC design engineer with ON Semiconductor Design Center, Shanghai, China. In 2006, she was a senior module development engineer at Intel, Shanghai, China. In 2008, she was a summer intern with Custom IC Division, Mentor Graphics, Wilsonville, OR. Her current research interests include high-speed IO circuit design, behavioral modeling, and signal integrity analysis, mixed-signal/RF circuit design and modeling with an emphasis on reliability optimization.

Slawomir Koziel (M'03–SM'07) received the M.Sc. and Ph.D. degrees in electronic engineering from Gdansk University of Technology, Poland, in 1995 and 2000, respectively. He also received the M.Sc. degrees in theoretical physics and in mathematics, in 2000 and 2002, respectively, as well as the PhD in mathematics in 2003, from the University of Gdansk, Poland.

He is currently an Associate Professor with the School of Science and Engineering, Reykjavik University, Iceland. His research interests include CAD and modeling of microwave circuits, surrogate-based optimization, space mapping, circuit theory, analog signal processing, evolutionary computation and numerical analysis.

Paul D. Franzon (M'88–SM'99–F'06) is currently a Distinguished Alumni Professor of Electrical and Computer Engineering at North Carolina State University. He earned his Ph.D. from the University of Adelaide, Adelaide, Australia in 1988. He has also worked at AT&T Bell Laboratories, DSTO Australia, Australia Telecom and two companies he cofounded, Communica and LightSpin Technologies. His current interests center on the technology and design of complex microsystems incorporating VLSI, MEMS, advanced packaging and nano-electronics. He has lead several major efforts and published over 200 papers in these areas. In 1993 he received an NSF Young Investigators Award, in 2001 was selected to join the NCSU Academy of Outstanding Teachers, in 2003, selected as a Distinguished Alumni Professor, and received the Alcoa Research Award in 2005. He is a Fellow of the IEEE.

Michael B. Steer (M'82–SM'90–F'99) received the B.E. degree in 1976 and Ph.D. in 1983 both from the University of Queensland, Australia. He is the Lampe Distinguished Professor of Electrical and Computer Engineering at North Carolina State University (NCSU). He is a Fellow of the IEEE (Institute of Electrical and Electronic Engineers). He has authored more than 430 refereed publications and 3 books Foundations of Interconnect and Microstrip Design (with T.C. Edwards); Multifunctional Adaptive Microwave Circuits and Systems (with W. D. Palmer); and the widely used textbook *Microwave and RF Design: A Systems Approach.* He was Secretary of the IEEE Microwave Theory and Techniques Society (MTT-S) in 1997 and was on the IEEE MTT-S AdCom from 1998 to 2001, and from 2003 to 2006. He is a former Editor-In-Chief of the *IEEE Transactions on Microwave Theory and Techniques*. He received the Alcoa Foundation Distinguished Research Award from NCSU's College of Engineering at NCSU in 2003; he was the Jack S. Kilby Lecturer in 2003; and in 1994 and 1996 received the Bronze Medallion from U.S. Army Research for Outstanding Scientific Accomplishment. In 2007 he received a Distinguished Service Recognition Award from the IEEE MTT-S. In 2009 he received a Military Medal, the "Commander's Award For Public Service" from the Commanding General of the U.S. Army Research, Development and Engineering Command (RDECOM). He received the 2010 Microwave Prize for the best paper on Microwave Engineering in any IEEE Publication in the preceding year, and received the 2011 Distinguished Educator Award from IEEE/MTT-S.

References

[1] I. G. Osio and C. H. Amon, "An engineering design methodology with multistage Bayesian surrogates and optimal sampling," *Res. Eng. Des.*, vol. 8, no. 4, pp. 189–206, Dec. 1996.

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

[3] R. Pally, "The predicting brain: Unconscious repetition, conscious reflection and therapeutic change," *Int. J. Psychoanal.*, vol. 88, no. 4, pp. 861–881, Aug. 2007.

[4] J. Hawkins and S. Blakeslee, *On Intelligence*. New York: Macmillan, 2005.

[5] M. B. Yelten, P. D. Franzon, and M. B. Steer, "Surrogate model-based analysis of analog circuits—Part I: Variability analysis," *IEEE Trans. Device Mater. Rel.*, vol. 11, no. 3, pp. 458–465, Sept. 2011.

[6] 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. Microwave Theory Tech.*, vol. 52, no. 1, pp. 337–361, Jan. 2004.

[7] T. C. Edwards and M. B. Steer, *Foundations of Interconnect and Microstrip Design*. West Sussex, U.K.: Wiley, 2000.

[8] P. Li and L. T. Pileggi, "Compact reduced-order modeling of weakly nonlinear analog and RF circuits," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 23, no. 2, pp. 184–203, Feb. 2005.

[9] A. A. Giunta, S. F. Wojtkiewicz, and M. S. Eldred, "Overview of modern design of experiments methods for computational simulations," in *Proc. 41st Aerospace Sciences Meeting and Exhibit,* Reno, NV, Jan. 2003, pp. 1–17.

[10] A. Forrester, A. Sobester, and A. Keane, *Engineering Design via Surrogate Modelling*. West Sussex, U.K.: Wiley, 2008.

[11] Z. N. Mevawalla, G. S. May, and M. W. Kiehlbauch, "Neural network modeling for advanced process control using production data," *IEEE Trans. Semicond. Manuf.*, vol. 24, no. 2, pp. 182–189, May 2011.

[12] T. Kiely and G. Gielen, "Performance modeling of analog integrated circuits using least-squares support vector machines," in *Proc. Design, Automation and Test in Europe Conf. Exhibition, DATE*, Paris, France, Feb. 2004, pp. 448–453.

[13] S. Haykin, *Neural Networks and Learning Machines,* 3rd ed. New Jersey: Pearson, 2009.

[14] S. Koziel and J. Bandler, "Recent advances in space-mapping-based modeling of microwave devices," *Int. J. Numer. Model. Electron. Netw. Devices Fields*, vol. 23, no. 6, pp. 425–446, Nov. 2010.

[15] A. J. Keane and P. B. Nair, *Computational Approaches for Aerospace Design*. West Sussex, U.K.: Wiley, 2005.

[16] M. Small and C. K. Tse, "Detecting determinism in time series: The method of surrogate data," *IEEE Trans. Circuits Syst. I*, vol. 50, no. 5, pp. 663–672, May 2003.

[17] M. D. Morris, "Factorial sampling plans for preliminary computational experiments," *Technometrics*, vol. 33, no. 2, pp. 161–174, May 1991.

[18] D. Child, *The Essentials of Factor Analysis*. New York: Continuum International Publishing, 2006.

[19] A. Mitev, M. Marefat, D. Ma, and J. M. Wang, "Principle Hessian direction-based parameter reduction for interconnect networks with process variation," *IEEE Trans. VLSI Syst.*, vol. 18, no. 9, pp. 1337–1347, Sept. 2010.

[20] Z. Feng and P. Li, "Performance-oriented parameter dimension reduction of VLSI circuits," *IEEE Trans. VLSI Syst.*, vol. 17, no. 1, pp. 137–150, Jan. 2009.

[21] Z. Li, X. Lu, and W. Shi, "Process variation dimension reduction based on SVD," in *Proc. IEEE Int. Symp. Circuits and Systems, ISCAS*, Bangkok, Thailand, May 2003, pp. IV(672)–IV(675).

[22] X. Li, J. Le, L. T. Pileggi, and A. Strojwas, "Projection-based performance modeling for inter/intra-die variations," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design, ICCAD*, San Jose, CA, Nov. 2005, pp. 721–727.

[23] A. Mitev, M. Marefat, D. Ma, and J. M. Wang, "Parameter reduction for variability analysis by slice inversion regression method," *IET Circuits, Devices Syst.*, vol. 2, no. 1, pp. 16–22, Feb. 2008.

[24] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, May 1979.

[25] B. Tang, "Orthogonal array-based latin hypercubes," *J. Amer. Statist. Assoc.*, vol. 88, no. 424, pp. 1392–1397, Dec. 1993.
[26] M. B. Yelten, P. D. Franzon, and M. B. Steer, "Comparison of modeling techniques for circuit variability analysis," *Int. J.Numerical Modelling: Electron. Netw., Devices Fields,* to be published.
[27] T.-T. Wong, W.-S. Luk, and P.-A. Heng, "Sampling with Hammersley and Holton points," *J. Graph. Tools*, vol. 2, pp. 1–16, 1997.
[28] D. Deschrijver, K. Crombecq, H. Nguyen, and T. Dhaene, "Adaptive sampling algorithm for macromodeling of parameterized S-parameter responses," *IEEE Trans. Microwave Theory Tech.*, vol. 59, no. 1, pp. 39 –45, Jan. 2011.
[29] K. Crombecq, L. D. Tommasi, D. Gorissen, and T. Dhaene, "A novel sequential design strategy for global surrogate modeling," in *Proc. Winter Simulation Conference WSC,* Austin, TX, Dec. 2009, pp. 731–742.
[30] D. Gorissen, K. Crombecq, I. Couckuyt, T. Dhaene, and P. Demeester, "A surrogate modeling and adaptive sampling toolbox for computer-based design," *J. Machine Learn. Res.*, vol. 11, no. 5, pp. 2051–2055, July 2010.
[31] J. E. Ráyas-Sanchez, J. Aguliar-Torrentera, and J. A. Jasso-Urzúa, "Surrogate modeling of microwave circuits using polynomial function interpolants," in *Proc. IEEE Int. Microwave Symp., IMS,* Anaheim, CA, May 2010, pp. 197–200.
[32] H. Filiol, I. O'Connor, and D. Morche, "A new approach for variability analysis of ICs," in *Proc. IEEE European Conf. Circuit Theory and Design, ECCTD,* Antalya, Turkey, Aug. 2009, pp. 237–240.
[33] J. P. C. Kleijnen, "Kriging metamodeling in simulation: A review," *Eur. J. Operation. Res.*, vol. 192, no. 3, pp. 707–716, Feb. 2009.
[34] J. D. Martin, "Computational improvements to estimating Kriging metamodel parameters," *ASME J. Mech. Des.*, vol. 131, no. 084501, pp. 1–7, Aug. 2009.
[35] S. N. Lophaven, H. B. Nielsen, and J. Søndergaard. (2002, Aug.). A MATLAB Kriging toolbox version 2.0. [Online]. Available: http://www2.imm.dtu.dk/~hbn/dace/dace.pdf
[36] D. Gorissen, L. D. Tommasi, W. Hendrickx, J. Croon, and T. Dhaene, "RF circuit block modeling via Kriging surrogates," in *Proc.*17th *Int. Conf. Microwaves, Radar and Wireless Communications, MIKON*, Wroclaw, Poland, May 2008, pp. 1–4.
[37] H. You, M. Yang, D. Wang, and X. Jia, "Kriging model combined with Latin hypercube sampling for surrogate modeling of analog integrated circuit performance," in *Proc. IEEE Int. Symp. Quality of Electronic Design, ISQED*, San Jose, CA, Mar. 2009, pp. 554–558.
[38] G. Yu and P. Li, "Yield-aware analog integrated circuit optimization using geostatics motivated performance modeling," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design, ICCAD*, San Jose, CA, Nov. 2007, pp. 464–469.
[39] D. Gorissen, L. D. Tommasi, K. Crombecq, and T. Dhaene, "Sequential modeling of a low noise amplifier with neural networks and active learning," *Neural Comput. Applicat.*, vol. 18, no. 5, pp. 485–494, June 2009.
[40] B. Liu, D. Zhao, P. Reynaert, and G. Gielen, "Synthesis of integrated passive components for high-frequency RF ICs based on evolutionary computation and machine learning techniques," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 30, no. 10, pp. 184–203, Oct. 2011.
[41] L. Zhang, Q.-J. Zhang, and J. Wood, "Statistical neuro-space mapping technique for large-signal modeling of nonlinear devices," *IEEE Trans. Microwave Theory Tech.*, vol. 56, no. 11, pp. 2453–2467, Nov. 2008.
[42] A. Díaz-Manríquez, G. Toscano-Pulido, and W. Gómez-Flores, "On the selection of surrogate models in evolutionary optimization algorithms," in *Proc. IEEE Congr. Evolutionary Computation, CEC*, New Orleans, LA, June 2011, pp. 2155–2162.
[43] J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, B. D. Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002.
[44] K. D. Brabanter, P. Karsmakers, F. Ojeda, C. Alzate, J. D. Brabanter, K. Pelckmans, B. D. Moor, J. Vandewalle, and J. A. K. Suykens. (2010, Sept.). LS-SVMlab toolbox version 1.7. [Online]. Available: http://www.esat.kuleuven.be/sista/lssvmlab/
[45] L. Xia, J. Meng, R. Xu, B. Yan, and Y. Guo, "Modeling of 3-D vertical interconnect using support vector machine regression," *IEEE Microwave Wireless Compon. Lett.*, vol. 16, no. 12, pp. 639–641, Dec. 2006.
[46] G. Angiulli, M. Cacciola, and M. Versaci, "Microwave devices and antennas modelling by support vector regression machines," *IEEE Trans. Magn.*, vol. 43, no. 4, pp. 1589–1592, Apr. 2007.
[47] S. Koziel, J. W. Bandler, and K. Madsen, "A space mapping framework for engineering optimization: Theory and implementation," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 10, pp. 3721–3730, Oct. 2006.
[48] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. K. Tucker, "Surrogate-based analysis optimization," *Progr. Aerospace Sci.*, vol. 41, pp. 1–28, Jan. 2005.
[49] N. M. Alexandrov, J. E. Dennis, R. M. Lewis, and V. Torczon, "A trust region framework for managing use of of approximation models in optimization," *Struct. Multidiscip. Optim.*, vol. 15, no. 1, pp. 16–23, Feb. 1998.
[50] H. Kabir, Y. Wang, M. Yu, and Q. J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microwave Theory Tech.*, vol. 56, no. 4, pp. 867–879, Apr. 2008.
[51] Q. S. Cheng, J. C. Rautio, J. W. Bandler, and S. Koziel, "Progress in simulator-based tuning—The art of tuning space mapping," *IEEE Microwave Mag.*, vol. 11, no. 4, pp. 96–110, June 2010.
[52] "em," Sonnet Software, Inc., North Syracuse, NY, 2010, version 12.54.
[53] N. M. Alexandrov and R. M. Lewis, "An overview of first-order model management for engineering optimization," *Optim. Eng.*, vol. 2, no. 4, pp. 413–430, Dec. 2001.
[54] D. Echeverria and P. W. Hemker, "Space mapping and defect correction," *CMAM Int. Math. J. Computat. Methods Appl. Math.*, vol. 5, pp. 107–136, 2005.
[55] S. Koziel, "Shape-preserving response prediction for microwave design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 58, no. 11, pp. 2829–2837, Nov. 2010.
[56] A. I. J. Forrester and A. J. Keane, "Recent advances in surrogate-based optimization," *Progr. Aerosci. Sci.*, vol. 45, no. 1–3, pp. 50–79, Jan. 2009.
[57] M. H. M. Salleh, G. Prigent, O. Pigaglio, and R. Crampagne, "Quarter-wavelength side-coupled ring resonator for bandpass filters," *IEEE Trans. Microwave Theory Tech.*, vol. 56, no. 1, pp. 156–162, Jan. 2008.
[58] "FEKO," EM Software & Systems-S.A. (Pty) Ltd., Stellenbosch, South Africa, 2010, suite 6.0.
[59] "Agilent ADS," Agilent Technologies, Santa Rosa, CA, 2008, version 2008.
[60] S. Koziel, "Role of constraints in surrogate-based design optimization of microwave structures," *IET Microwaves, Antennas Propagat.*, vol. 5, no. 5, pp. 588–595, Apr. 2011.
[61] IO buffer information specification (ibis). [Online]. Available: http://www.eda.org/ibis/
[62] T. Zhu, M. B. Steer, and P. D. Franzon, "Accurate and scalable IO buffer macromodel based on surrogate modeling," *IEEE Trans. Compon. Packag. Manufact. Technol.*, vol. 1, no. 8, pp. 1240–1249, Aug. 2011.
[63] A. Muranyi. (2006, Mar.). Accuracy of IBIS models with reactive loads. [Online]. Available: http://www.eda.org/ibis/summits/mar06/muranyi1.pdf
[64] M. LaBonte and A. Muranyi. (2007, Mar.). IBIS advanced technology modeling task group work-achievement: Verilog-a element library HSPICE test. [Online]. Available: http://vhdl.org/pub/ibis/macromodel˙wip/archive-date.html
[65] J. B. Bernstein, M. Gurfinkel, X. Li, J. Walters, Y. Shapira, and M. Talmor, "Electronic circuit reliability modeling," *Microelectron. Rel.*, vol. 46, pp. 1957–1979, 2006.
[66] W. Wang, V. Reddy, A. T. Krishnan, R. Vattikonda, S. Krishnan, and Y. Cao, "Compact modeling and simulation of circuit reliability for 65-nm CMOS technology," *IEEE Trans. Device Mater. Rel.*, vol. 7, no. 4, pp. 509–517, Dec. 2007.
[67] H. Kufluoglu and M. A. Alam, "Theory of interface-trap-induced NBTI degradation for reduced cross section MOSFETs," *IEEE Trans. Electron Dev.*, vol. 53, no. 5, pp. 1120–1130, May 2006.
[68] M. Orshansky, S. R. Nassif, and D. Boning, *Design for Manufacturability and Statistical Design: A Constructive Approach*. New York: Springer-Verlag, 2008.
[69] N. Drego, A. Chandrakasan, and D. Boning, "All-digital circuits for measurement of spatial variation in digital circuits," *IEEE J. Solid-State Circuits*, vol. 45, no. 3, pp. 640–651, Mar. 2010.
[70] "Virtuoso analog design environment XL user guide," product version 6.1.2, Cadence Design Systems, Oct. 2007.
[71] S. K. Saha, "Modeling process variability in scaled CMOS technology," *IEEE Des. Test. Comput.*, vol. 27, no. 2, pp. 8–16, 2010.
[72] M. B. Yelten, P. D. Franzon, and M. B. Steer, "Surrogate model-based analysis of analog circuits—Part II: Reliability analysis," *IEEE Trans. Device Mater. Rel.*, vol. 11, no. 3, pp. 466–473, Sept. 2011.
[73] T. H. Morshed, W. Yang, M. V. Dunga, X. Xi, J. He, W. Liu, K. Yu, M. Cao, X. Jin, J. J. Ou, M. Chan, A. M. Niknejad, and C. Hu. (2009, Apr.). BSIM4.6.4 MOSFET model—User's manual. [Online]. Available: http://www-device.eecs.berkeley.edu/~bsim3/BSIM4/BSIM464/BSIM464˙Manual.pdf/