Kriging, Co-Kriging and Space Mapping for Microwave Circuit Modeling

Ivo Couckuyt*, Slawomir Koziel*, and Tom Dhaene*

*Surrogate Modeling Lab, Department of Information Technology, Gent University - IBBT, B-9050 Gent, Belgium
ivo.couckuyt@ugent.be, tom.dhaene@ugent.be
# Engineering Optimization & Modeling Center, School of Science and Engineering, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland
koziel@ru.is

Abstract—Space mapping (SM) is a popular technique that allows creating computationally cheap and reasonably accurate surrogates of EM-simulated microwave structures (so-called fine models) using underlying coarse models, typically equivalent circuits. Here, we consider various ways of enhancing SM surrogates by exploiting additional training data as well as two function approximation methodologies, kriging and co-kriging. To our knowledge, it is the first application of co-kriging for microwave circuit modeling. Based on the three examples of microstrip filters, we present a comprehensive numerical study in which we compare the accuracy of the basic SM models as well as SM enhanced by kriging and co-kriging. Direct kriging interpolation of fine model data is used as a reference.

I. INTRODUCTION

Reliable evaluation of microwave components and devices can be obtained through full-wave electromagnetic (EM) simulation. High-fidelity simulation is CPU intensive, which is a serious bottleneck for EM-simulation-based design tasks such as parametric optimization or statistical analysis. Therefore, accurate and computationally cheap models of microwave structures are indispensable.

Computationally cheap models can be implemented using various approximation techniques such as polynomial regression [1], radial basis functions [2], kriging [3], support vector regression [4], fuzzy systems [5], rational approximation [6], or artificial neural networks [7]. For good accuracy, all of these techniques require a large number of training points, and are therefore suitable for creating multiple-use library models.

Physical surrogate modeling techniques, particularly space mapping (SM) [8], [9], allow creating reasonably-accurate and fast models without using excessive amount of training data. An SM surrogate is constructed using a simplified representation (coarse model) of the microwave structure in question (fine model), e.g., an equivalent circuit, and some auxiliary mappings that “reshape” the parameter space and/or response of the coarse model. The enhancement of the coarse model is typically realized through suitable analytical formulas, which allows the surrogate model to be almost as computationally cheap as the coarse model. Because the coarse model is supposedly physics-based, however, the accuracy of the space mapping surrogate is considerably better than the accuracy of possible function approximation models using a comparable amount of fine model data [10].

A downside of SM models is that increasing the number of training points may have little effect on the model’s quality [11]. This problem can be overcome in various ways. SM modeling with variable weight coefficients [11] provides efficient utilization of available fine model data, however, at the expense of computational overhead related to a separate parameter extraction process required for each evaluation of the surrogate. SM modeling enhanced by fuzzy systems [10] offer accuracy comparable with [11] without compromising computational cost, however, the model of [10] may not be differentiable (hence difficult to optimize); also, it works best if the base set is a rectangular grid; otherwise its performance may be degraded.

One of the most successful improvements of space mapping, leading to a full utilization of available fine model data has been proposed in [12], where kriging interpolation was used as an output-SM-like correction layer improving the accuracy of the surrogate. In this paper, we consider this and other SM enhancements based on kriging [13] and co-kriging [14], the latter never used for microwave circuit modeling before. We carry out a comprehensive numerical study where—using the three microstrip filter examples—we compare the accuracy of basic SM models, direct kriging interpolation of fine model data, as well as SM combined with kriging and co-kriging. Discussion and recommendations are also presented.

II. SURROGATE MODELS EXPLOITING SPACE MAPPING, KRIGING, AND CO-KRIGING

A. Surrogate Modeling Using Space Mapping

Space mapping (SM) [9] is a simple way of constructing the surrogate of a microwave structure by correcting the underlying coarse model. The coarse model should be computationally cheap and physically-based so that good alignment between the surrogate and the fine model can be obtained by using a limited number of fine model evaluations [9]. For these reasons, a preferred choice for the coarse model is an equivalent circuit.

Let \( \mathbf{R}_f^{(j)} : X_f \rightarrow \mathbb{R}^n_+ \subseteq \mathbb{R}^m, \) and \( \mathbf{R}_c^{(j)} : X_c \rightarrow \mathbb{R}^m_+ \subseteq \mathbb{R}^n, \) denote the fine and coarse model response vectors. For example, \( R_f(x) \) and \( R_c(x) \) may represent the magnitude of a transfer function at \( m \) chosen frequencies. Let \( X_f \subseteq X_f \) be a region of interest where we want enhanced matching between the surrogate and \( R_f. \) Here, \( X_f \) is an \( n \)-dimensional interval in \( \mathbb{R}^n \) with center at...
reference point $x^0 = [x_{0,1} \ldots x_{0,n}]^T \in \mathbb{R}^n$ and size $\mathbf{d} = [\delta_1 \ldots \delta_1]^T$ [9]. Let $X_{SW} = \{x_{SW1}, x_{SW2}, \ldots, x_{SWN}\} \subset X_R$ be the base set, such that the fine model response is known at all points $x^j$, $j = 1, 2, \ldots, N_{SW}$.

In this work, the base set is chosen to be a factorial design, so-called star distribution [9], consisting of $2n+1$ points located at the center and all the faces of $X_R$.

The SM surrogate $R_{SM}$ is defined as

$$R_{SM}(x) = A \cdot R(x, B \cdot x + c)$$

(1)

where $A$, $B$, and $c$ are $m \times m$, $n \times n$, and $n \times 1$ matrices determined through the parameter extraction process

$$[A, B, c] = \arg \min_{c} \sum_{i=1}^{N_{SW}} || R(x^{SW}_i) - \alpha \cdot R(x, \beta \cdot x + \gamma) ||$$

(2)

Other types of SM surrogates can be found in [15]. Space mapping typically ensures reasonable accuracy while using a limited amount of fine model data. Moreover, as the parameter extraction process (2) is independent of the evaluation point $x$ of the surrogate model, the accuracy of the model (1)-(2) is barely dependent on the number of the base points $N_{SW}$.

B. Kriging Interpolation

Kriging is a popular technique to interpolate deterministic noise-free data [3], [13], [16]. These Gaussian Process based surrogate models are compact and cheap to evaluate, and have proven to be very useful for tasks such as optimization, design space exploration, visualization, prototyping, and sensitivity analysis. A first use case for kriging is to rely solely on the fine model.

Let $X_{KR} = \{x_{KR1}, x_{KR2}, \ldots, x_{KRN}\} \subset X_R$ be the base (training) set and $R(X_{KR})$ the associated fine model responses. Then, the kriging interpolant, also known as the Best Linear Unbiased Predictor (BLUP), is derived as

$$R_{K}(x) = M \alpha + r(x) \cdot \Psi^{-1} \cdot (R_{f} - F \alpha)$$

(3)

where $M$ and $F$ are Vandermonde matrices of the test point $x$ and the base set $X_{KR}$, respectively. The coefficient vector $\alpha$ is determined by Generalized Least Squares (GLS). $r(x)$ is an $1 \times N_{KR}$ vector of correlations between the point $x$ and the base set $X_{KR}$, where the entries are $r_j(x) = \phi(x, x_{KR})$, and $\Psi$ is a $N_{KR} \times N_{KR}$ correlation matrix, where the entries are given by $\Psi_{ij} = \phi(x_{KRi}, x_{KRj})$.

In this work, the correlation function is chosen Gaussian, i.e., $\psi(z, y) = \exp\left(\sum_{i=1}^{n} - \theta_i |z^i - y^i|^2\right)$, where the parameters $\theta_1, \ldots, \theta_n$ are identified by Maximum Likelihood Estimation (MLE). The regression function is chosen constant, i.e., $F = [1 1 \ldots 1]^T$ and $M = (1)$.

C. Kriging-Enhanced Space Mapping Modeling

Kriging can be utilized to enhance the SM surrogate model (1), (2) by interpolating the residuals $R_{K}(x) = R(X_{KR}) - R_{SM}(x)$, where $x_{KR} \in X_{KR}$ [12]. The kriging-enhanced SM model can be formulated as

$$R_{SM,K}(x) = R_{SM}(x) + R_{K}(x)$$

(4)

where $R_{SM,K}$ is defined by (3) except that it interpolates $R_{K}$ instead of $R_{KH}(x)$. In general, the base set of the SM model, $X_{SM}$, can be a subset of $X_{KR}$. Here, we use a star-distribution $X_{BSM}$ as described in Section II.A.

D. Co-Kriging Enhanced Space Mapping Modeling

Enhancement of the SM surrogate model can also be realized through co-kriging [14]. Co-kriging is a type of kriging where fine and coarse model data are combined to enhance the prediction accuracy.

Co-kriging is a two-steps process: first a kriging model $R_{K,KR}$ of the coarse data $X_{KR1} = R(X_{KR})$ is constructed and on the residuals of the fine data $(X_{KR2}, R_{f})$ a second kriging model $R_{K,KR}$ is applied, where $R_{f} = R(X_{KR}) - \rho R(X_{KR})$. The parameter $\rho$ is included in the MLE. Note that if the response values $R(X_{KR})$ are not available, they can be approximated by using the first kriging model $R_{K,KR}$, namely, $R(X_{KR}) = R_{K,KR}(X_{KR})$.

The resulting co-kriging interpolant is defined similarly as (3), i.e.,

$$R_{SM,Co}(x) = M \alpha + r(x) \cdot \Psi^{-1} \cdot (R_{f} - F \alpha)$$

(5)

where the block matrices $M$, $F$, $r(x)$ and $\Psi$ can be written in function of the two separate kriging models $R_{K,KR}$ and $R_{K,KR}:$

$$r(x) = [\rho \cdot \sigma_{\epsilon}^2 \cdot r(x), \rho \cdot \sigma_{\epsilon}^2 \cdot r(x, X_{KR}) + \sigma_{\epsilon}^2 \cdot r_j(x)]$$

$$\Psi = \begin{bmatrix} \sigma_{\epsilon}^2 & \rho \cdot \sigma_{\epsilon}^2 \cdot \Psi, & (X_{KR}, X_{KR}) \\ \rho \cdot \sigma_{\epsilon}^2 \cdot \Psi, & 0 & \sigma_{\epsilon}^2 \cdot \Psi_j \end{bmatrix}$$

(6)

$$F = \begin{bmatrix} F_c, 0 \\ F_c \\ F_c \end{bmatrix}$$

where $(F_c, \sigma_c, \Psi_c, M_c)$ and $(F_d, \sigma_d, \Psi_d, M_d)$ are matrices obtained from the kriging models $R_{K,KR}$ and $R_{K,KR}$, respectively (see Section II.B). In particular, $\sigma_c^2$ and $\sigma_d^2$ are process variances, while $\Psi_c(\cdot, \cdot)$ and $\Psi_d(\cdot, \cdot)$ denote correlation matrices of two datasets with the optimized $\theta_1, \ldots, \theta_n$ parameters and correlation function of the kriging models $R_{K,KR}$ and $R_{K,KR}$, respectively.

III. NUMERICAL RESULTS

A. Experimental Setup

Our numerical study uses the three examples of microstrip filters: the second- (Fig. 1) and the fourth-order (Fig. 2) ring resonator bandpass filters [17], as well as the bandpass microstrip filter with open stub inverter [18] (Fig. 3). We refer to them as Case 1, 2, and 3, respectively. All filters are simulated using FEKO solver [19]. The coarse models are equivalent circuits shown in Fig. 4(a) to 4(c).

The design variables and the region of interest for our examples are the following (all sizes in mm): $x = [L_1, L_2, W_1, W_2, S_1]^T$, $x^0 = [22 20 0.2 0.8 1.7]^T$, $\delta = [2 2 0.1 0.2 0.2]^T$ (Case 1), $x = [L_1, L_2, L_3, S_1, S_2, W_1, W_2]^T$, $x^0 = [24 26 20 0.1 0.1 1.2 0.8]^T$, $\delta = [2 2 0.05 0.05 0.1 0.1]^T$ (Case 2), and $x = [L_1, L_2, L_3, S_1, S_2, W_1, W_2]^T$, $x^0 = [22 25 0.5 0.2 0.2 0.5]^T$, $\delta = [2 1 1 0.2 0.1 0.2]^T$ (Case 3).

The space mapping model $R_{SM}$ is set up as described in Section II.A. The kriging and co-kriging models ($R_{K,KR}, R_{SM}$, $R_{SM,Co}$) are constructed using various numbers of training points (from $N_{KR} = 20$ to $N_{KR} = 400$). Co-kriging models are configured using the SM model as the coarse model, namely, $R_c = R_{SM}$, and use a fixed number of SM model evaluations $N_c = 400$. 


The quality of the surrogate is assessed using a relative error measure $\| R(x) - R_*(x) \| / \| R(x) \|$ expressed in percent, where $R(x)$ and $R_*(x)$ denote the fine and the respective surrogate model response at a given test point $(design)$ $x$.

B. Results and Discussion

Table I shows the average modeling error for the considered models and the three test cases, obtained for 100 test points randomly allocated in the region of interest. The errors corresponding to the coarse model itself are also shown for comparison.

Figures 5 to 8 show the responses of the SM model $R_{SM}$ as well as the kriging-enhanced SM model $R_{SM, KR}$ ($N_{KR} = 400$) for a few selected test points. The visual matching between the $R_{SM, KR}$ and the fine model is very good for all test cases.

The results of Table I indicate that the $R_{SM, KR}$ model is superior for all three test cases considered. In particular, by employing kriging interpolation as a response correction term for the SM surrogate allows us to reduce the modelling error by up to 60%, depending on the example. Enhancements based on co-kriging are not as good. In fact, the quality of $R_{SM, CO}$ is marginally better than that of $R_{SM, KR}$, which may suggest that co-kriging is not the best choice for the type of responses (here, filter transmission) considered in this work, at least in the form realized in Section II.D.

Obviously, the primary advantage of SM models is that only a handful of training points is necessary to create the surrogate (here, $2n + 1$, with $n$ being the number of design variables). Whether it is reasonable to improve the accuracy by, say, 50%, at the expense of additional two or four hundred samples depends on the application of the model. It is perfectly justified for multiple-use library models, particularly if the application is statistical analysis or yield-driven design [14]. It appears though that (global) accuracy may not be that critical in case of certain tasks such as parametric design optimization, because the surrogate can be locally corrected using, e.g., output SM [15] or manifold mapping [20].

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Model</th>
<th>Average Modeling Error [%]</th>
<th>$N_{KR} = 20$</th>
<th>$N_{KR} = 50$</th>
<th>$N_{KR} = 100$</th>
<th>$N_{KR} = 200$</th>
<th>$N_{KR} = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$R_c$</td>
<td>43.6</td>
<td>43.6</td>
<td>43.6</td>
<td>43.6</td>
<td>43.6</td>
<td>43.6</td>
</tr>
<tr>
<td></td>
<td>$R_{SM}$</td>
<td>18.5</td>
<td>18.5</td>
<td>18.5</td>
<td>18.5</td>
<td>18.5</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td>$R_{KR}$</td>
<td>22.6</td>
<td>21.5</td>
<td>17.6</td>
<td>14.5</td>
<td>11.2</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td>$R_{SM, KR}$</td>
<td>18.2</td>
<td>15.5</td>
<td>13.9</td>
<td>11.4</td>
<td>8.5</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td>$R_{SM, CO}$</td>
<td>18.2</td>
<td>17.6</td>
<td>17.0</td>
<td>14.9</td>
<td>11.6</td>
<td>9.2</td>
</tr>
<tr>
<td>2</td>
<td>$R_c$</td>
<td>29.1</td>
<td>29.1</td>
<td>29.1</td>
<td>29.1</td>
<td>29.1</td>
<td>29.1</td>
</tr>
<tr>
<td></td>
<td>$R_{SM}$</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>$R_{KR}$</td>
<td>13.7</td>
<td>13.4</td>
<td>12.8</td>
<td>12.1</td>
<td>10.3</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>$R_{SM, KR}$</td>
<td>3.1</td>
<td>3.0</td>
<td>2.8</td>
<td>2.6</td>
<td>2.3</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>$R_{SM, CO}$</td>
<td>13.8</td>
<td>11.9</td>
<td>10.4</td>
<td>10.4</td>
<td>9.2</td>
<td>6.9</td>
</tr>
<tr>
<td>3</td>
<td>$R_c$</td>
<td>36.1</td>
<td>36.1</td>
<td>36.1</td>
<td>36.1</td>
<td>36.1</td>
<td>36.1</td>
</tr>
<tr>
<td></td>
<td>$R_{SM}$</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>$R_{KR}$</td>
<td>31.1</td>
<td>28.3</td>
<td>26.9</td>
<td>24.3</td>
<td>18.4</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td>$R_{SM, KR}$</td>
<td>6.0</td>
<td>5.5</td>
<td>5.0</td>
<td>4.2</td>
<td>3.2</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>$R_{SM, CO}$</td>
<td>27.0</td>
<td>27.1</td>
<td>22.2</td>
<td>22.5</td>
<td>19.4</td>
<td>16.7</td>
</tr>
</tbody>
</table>
CONCLUSIONS

Several ways of improving accuracy of space mapping surrogate models by means of kriging and co-kriging have been investigated. Our results indicate that using kriging as the response correction technique is probably the best way of enhancing the model quality. Co-kriging does not seem to be as efficient, which suggests that this technique is not quite suitable for the type of responses considered in this paper. Our conjectures are supported by a comprehensive numerical study involving three microstrip filters and various combinations of space mapping, kriging, and co-kriging models.

REFERENCES