Reliable Low-Cost Co-Kriging Modeling of Microwave Devices

Slawomir Koziel[#], *Senior Member*, *IEEE*, Ivo Couckuyt[$], and Tom Dhaene[$]

[#]Engineering Optimization & Modeling Center, School of Science and Engineering,
Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland

[$]Department of Information Technology, Ghent University - IBBT, Gent, Belgium

***Abstract*—A reliable methodology for accurate modeling of microwave devices is presented. Our approach exploits co-kriging which utilizes low- and high-fidelity EM simulation data and combines them into a single surrogate model. Densely sampled low-fidelity data determines a trend function which is further corrected by sparsely sampled high-fidelity simulations. Low-fidelity EM data is also enhanced by using a frequency scaling. With our method, accurate models can be obtained at a fraction of cost required by conventional approximation models that are exclusively based on high-fidelity simulations. Two cases of microstrip bandpass filters are considered. Comparisons with conventional approximation models as well as application examples are also given.**

***Index Terms*—Microwave modeling, response-surface modeling, co-kriging, electromagnetic simulation.**

I. Introduction

Accurate and fast models (surrogates) are indispensable in the design of microwave structures and components. Many design tasks, such as parametric optimization, statistical analysis or yield-driven optimization, require numerous evaluations of a structure of interest and the use of high-fidelity electromagnetic (EM) simulations may be prohibitive because of unacceptably high computational cost.

Cheap models can be obtained using response surface approximation techniques such as polynomial regression [1], radial basis functions [2], kriging [2], support vector regression [3], or artificial neural networks [4]. However, for good accuracy, these techniques require a large number of training points, which exponentially grows with the dimensionality of the design space. This high initial setup cost may be justifiable for multiple-use library models but not quite for one-time design and analysis of a specific structure.

Low-cost microwave modeling can be realized using space mapping (SM) [5]. Reasonably accurate SM surrogate model can be created using a limited number of high-fidelity EM simulations by applying suitable correction to the underlying low-fidelity (or coarse) model, e.g., equivalent circuit. A drawback of SM models is that increasing the number of training points may have little effect of the model's quality [6]. Also, SM requires that the coarse model is very fast, as each evaluation of the SM surrogate requires evaluation of the underlying coarse model.

In this paper, we consider models constructed using both high- and low-fidelity EM simulations. Simulation of coarsely-discretized structure is less accurate but it is much faster than the high-fidelity one. By means of co-kriging [7], densely sampled low-fidelity simulations are combined with limited number of high-fidelity ones. The resulting model is as accurate as conventional approximation based on high-fidelity data only but uses much smaller number of training points. In order to further improve the co-kriging model accuracy, we precondition low-fidelity model data using frequency scaling. Also, a co-kriging covariance function [7] is suitably selected to be best suited to model sharp responses typical for many microwave devices, e.g., filters. The efficiency of the proposed approach is demonstrated using two microstrip filters. A comparison with conventional kriging interpolation [8] used here as a benchmark technique, as well as application of co-kriging models to parametric optimization is also given.

This work was supported in part by the Icelandic Centre for Research (RANNIS) Grant 110034021. Ivo Couckuyt is funded by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

II. Co-Kriging-Based Filter Modeling

A. EM-Simulation Models

Our goal is to create a computationally cheap surrogate of the high-fidelity EM-simulated filter model $\boldsymbol{R}_f(\boldsymbol{x})$ which is accurate but expensive to evaluate. Here, $\boldsymbol{x}$ is a vector of designable (e.g., geometry) parameters. The components of $\boldsymbol{R}_f$ may represent, e.g., $|S_{21}|$ over the frequency band of interest. We also consider an auxiliary (low-fidelity) model $\boldsymbol{R}_c$ which is evaluated using the same EM solver, however, with coarser discretization. $\boldsymbol{R}_c$ is much faster than $\boldsymbol{R}_f$ but lacks accuracy.

B. Kriging Interpolation

Kriging, otherwise known as a Gaussian Process, is a compact and cheap interpolation technique frequently used to solve computational expensive design problems [1], [2], [8]. In particular, kriging is well-suited to handle deterministic noise-free data, though kriging can also be applied to stochastic simulation [12].

Given a base (training) set $X_{B.KR} = \{\boldsymbol{x}_{KR}^1, \boldsymbol{x}_{KR}^2, \ldots, \boldsymbol{x}_{KR}^{NKR}\} \subset X_R$ and the associated fine model responses $\boldsymbol{R}_f(X_{B.KR})$, then, the kriging interpolant is defined by,

$$\boldsymbol{R}_{s.KR}(\boldsymbol{x}) = M\alpha + r(\boldsymbol{x}) \cdot \Psi^{-1} \cdot (\boldsymbol{R}_f(X_{B.KR}) - F\alpha) \qquad (1)$$

where $r(\boldsymbol{x})$ is an $1\times N_{KR}$ vector of correlations between the point $\boldsymbol{x}$ and the base set $X_{B.KR}$, the entries are $r_i(\boldsymbol{x}) = \psi(\boldsymbol{x}, \boldsymbol{x}_{KR}^i)$, and Ψ is a $N_{KR}\times N_{KR}$ correlation matrix, with the entries given by $\Psi_{i,j} = \psi(\boldsymbol{x}_{KR}^i, \boldsymbol{x}_{KR}^j)$. Furthermore, M and F are design matrices (=model matrices) of the test point $\boldsymbol{x}$ and the base set $X_{B.KR}$, respectively. The coefficient vector α is determined by Generalized Least Squares (GLS), namely, $\alpha = (F'\Psi^{-1}F)^{-1}F'\Psi^{-1}R_f(X_{B.KR})$.

In this work, we use kriging as a benchmark technique for comparison with the co-kriging approach of Section II.*C*. The kriging surrogate model is configured with the exponential correlation function, i.e., $\psi(\boldsymbol{x},\boldsymbol{y}) = \exp(\sum_{k=1,\ldots,n} -\theta_k|\boldsymbol{x}^k - \boldsymbol{y}^k|)$, where the parameters $\theta_1, \ldots, \theta_n$ are identified by Maximum Likelihood Estimation (MLE). The regression function is chosen constant, $F = [1 \ldots 1]^T$ and $M = (1)$.

C. Co-Kriging Surrogate Modeling

Kriging has been extended in literature to handle different types of prior knowledge to enhance the prediction accuracy, e.g., gradient information, amount of noise on the data [12], etc. Of particular interest is the ability of kriging to properly integrate the $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ model data into its prediction.

This technique, known as co-kriging [7], is actually the combination of two standard kriging models. The first kriging model $\boldsymbol{R}_{s.KRc}$ interpolates the coarse data $(X_{B.KRc},\boldsymbol{R}_c(X_{B.KRc}))$, while the second kriging model $\boldsymbol{R}_{s.KRd}$ is applied on the residuals of the fine data $(X_{B.KRf},\boldsymbol{R}_d)$, where $\boldsymbol{R}_d = \boldsymbol{R}_f(X_{B.KRf}) - \rho\cdot\boldsymbol{R}_c(X_{B.KRf})$.

The resulting co-kriging interpolant is defined as

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

where the block matrices M, F, $r(\boldsymbol{x})$ and Ψ can be written in function of the two separate kriging models $\boldsymbol{R}_{s.KRc}$ and $\boldsymbol{R}_{s.KRd}$:

$$r(\boldsymbol{x}) = [\rho\cdot\sigma_c^2\cdot r_c(\boldsymbol{x}), \rho^2\cdot\sigma_c^2\cdot r_c(\boldsymbol{x}, X_{B.KR_f}) + \sigma_d^2\cdot r_d(\boldsymbol{x})]$$

$$\Psi = \begin{bmatrix} \sigma_c^2\Psi_c & \rho\cdot\sigma_c^2\cdot\Psi_c(X_{B.KR_c}, X_{B.KR_f}) \\ 0 & \rho^2\cdot\sigma_c^2\cdot\Psi_c(X_{B.KR_f}, X_{B.KR_f}) + \sigma_d^2\cdot\Psi_d \end{bmatrix} \quad (3)$$

$$F = \begin{bmatrix} F_c & 0 \\ \rho\cdot F_d & F_d \end{bmatrix}, \qquad M = [\rho\cdot M_c \;\; M_d]$$

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 $\boldsymbol{R}_{s.KRc}$ and $\boldsymbol{R}_{s.KRd}$, respectively (see Section II.*B*). In particular, σ_c^2 and σ_d^2 are process variances, while $\Psi_c(\cdot,\cdot)$ and $\Psi_d(\cdot,\cdot)$ denote correlation matrices of two datasets using the optimized $\theta_1, \ldots, \theta_n$ parameters and correlation function of the kriging models $\boldsymbol{R}_{s.KRc}$ and $\boldsymbol{R}_{s.KRd}$, respectively. The parameter ρ is estimated as part of the MLE of the second kriging model. Furthermore, if not available, the response values $\boldsymbol{R}_c(X_{B.KRf})$ can be approximated by using the prediction of the first kriging model $\boldsymbol{R}_{s.KRc}$, namely, $\boldsymbol{R}_c(X_{B.KRf}) \approx \boldsymbol{R}_{s.KRc}(X_{B.KRf})$.

D. Low-Fidelity Model Preconditioning. Frequency Scaling

In order to improve the accuracy of the co-kriging model, the low-fidelity model $\boldsymbol{R}_c$ is preconditioned by means of frequency scaling to improve its alignment with the high-fidelity one. Typically, $\boldsymbol{R}_c(\boldsymbol{x}) = \boldsymbol{R}_c(\boldsymbol{x},\boldsymbol{f}) = [R_c(\boldsymbol{x},f_1)\ R_c(\boldsymbol{x},f_2)\ \ldots\ R_c(\boldsymbol{x},f_m)]^T$, where $\boldsymbol{f} = [f_1\ \ f_2\ \ f_m]^T$ are frequencies at which the low-fidelity model is evaluated. Preconditioning of $\boldsymbol{R}_c$ is realized as follows

$$\boldsymbol{R}_{c.prec}(\boldsymbol{x}) = \boldsymbol{R}_c(\boldsymbol{x}, p(\boldsymbol{f})) = [R_c(\boldsymbol{x}, p(f_1))\ \ldots\ R_c(\boldsymbol{x}, p(f_m))]^T \quad (4)$$

where p is a scaling function determined to minimize misalignment between $\boldsymbol{R}_c$ and $\boldsymbol{R}_f$ at points from $X_{B.KRf}$ (high-fidelity model data is available at these points anyway):

$$p = \arg\min_p \sum_{\boldsymbol{x}\in X_{B.KRf}} \|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_c(\boldsymbol{x}, p(\boldsymbol{f}))\|^2 \quad (5)$$

Here, as use third-order polynomial $p(f) = p_0 + p_1 f + p_2 f^2 + p_3 f^3$. In practice, the values of $R_c(\boldsymbol{x},p(\boldsymbol{f}))$ are obtained by interpolating $R_c(\boldsymbol{x},\boldsymbol{f})$ to frequencies $p(\boldsymbol{f})$.

III. Verification Examples

A. CCDBR Filter

Consider a second-order capacitively-coupled dual-behavior resonator (CCDBR) microstrip filter [9] shown in Fig. 1. The design parameters are $\boldsymbol{x} = [L_1\ L_2\ L_3]^T$; $S = 0.05$ mm is fixed. The high- and low-fidelity models are simulated in FEKO [10]. Total mesh numbers for $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ are 1134 (evaluation time 19 min) and 130 (evaluation time 20 s), respectively. The surrogate model is set up in the interval $[\boldsymbol{x} - \boldsymbol{\delta}, \boldsymbol{x} + \boldsymbol{\delta}]$ with $\boldsymbol{x}^0 = [3\ 5\ 1.5]^T$ mm, and $\boldsymbol{\delta} = [1\ 1\ 0.5]^T$ mm.

The kriging and co-kriging models ($\boldsymbol{R}_{s.KR}$, $\boldsymbol{R}_{s.CO}$) are constructed using various numbers of training points (from $N_{KR} = 20$ to $N_{KR} = 400$). Co-kriging models are configured using 400 $\boldsymbol{R}_c$ samples (the CPU cost of which corresponds to about 6 evaluations of $\boldsymbol{R}_f$). The modeling accuracy has been verified using 50 test points allocated randomly in the region of interest. We use the relative error measure $\|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x})\|/\|\boldsymbol{R}_f(\boldsymbol{x})\|$ expressed in percent.

The modeling errors are shown in Table I (see also Fig. 2). The co-kriging model accuracy obtained with 20 $\boldsymbol{R}_f$ samples is as good as that of the kriging model obtained for 100 samples. In other words, our approach allows us to significantly reduce the computational cost of creating surrogate model compared to conventional approximation based on high-fidelity data only.

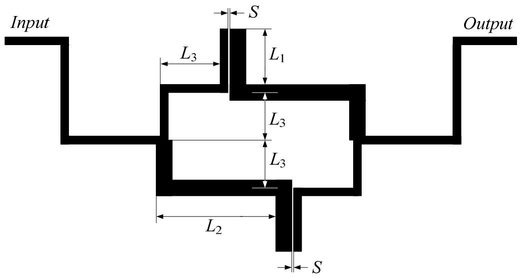

Fig. 1. CCDBR filter: geometry [9].

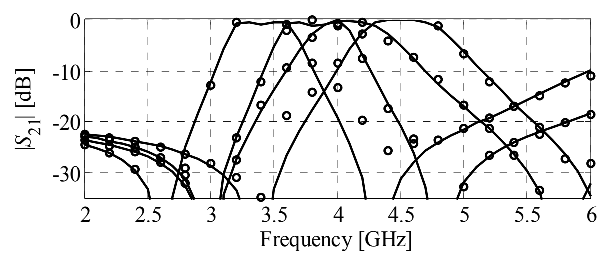

Fig. 2. CCDBR filter: responses of $\boldsymbol{R}_f$ (—) and co-kriging surrogate model (o) at selected test points. Co-kriging model created using 50 evaluations of $\boldsymbol{R}_f$ and 400 evaluations of $\boldsymbol{R}_c$.

TABLE I

CCDBR Filter: Modeling Results

Model	Average Modeling Error [%]				
	$N_{KR} = 20$	$N_{KR} = 50$	$N_{KR} = 100$	$N_{KR} = 200$	$N_{KR} = 400$
$\boldsymbol{R}_{s.KR}$	12.3	7.3	6.0	4.5	3.2
$\boldsymbol{R}_{s.CO}$	5.9	5.3	4.3	4.0	3.5

B. Fourth-Order Ring-Resonator Bandpass Filter

Consider a fourth-order ring resonator bandpass filter [11] shown in Fig. 1. The design parameters are $\boldsymbol{x} = [L_1\ L_2\ L_3\ S_1\ S_2\ W_1\ W_2]^T$. The high- and low-fidelity models are simulated in FEKO [10]. Total mesh numbers for $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ are 1344 (evaluation time 25 min) and 150 (evaluation time 22 s), respectively. The surrogate model is set up in the interval $[\boldsymbol{x} - \boldsymbol{\delta}, \boldsymbol{x} + \boldsymbol{\delta}]$ with $\boldsymbol{x}^0 = [24\ 20\ 26\ 0.1\ 0.1\ 1.2\ 0.8]^T$ mm, and $\boldsymbol{\delta} = [2\ 2\ 2\ 0.05\ 0.05\ 0.1\ 0.1]^T$ mm. The kriging and co-kriging models ($\boldsymbol{R}_{s.KR}$, $\boldsymbol{R}_{s.CO}$) are constructed as in the previous example. The CPU cost of 400 low-fidelity model evaluations used by $\boldsymbol{R}_{s.CO}$ corresponds to about 6 evaluations of $\boldsymbol{R}_f$.

IV. Application Examples

The co-kriging surrogate models have been used to carry out design optimization of the filters considered in Section III. The design specifications for the CCDBR filter are: $|S_{21}| \geq -3$dB for 3.8GHz to 4.2GHz and $|S_{21}| \leq -20$dB for 2GHz to 3.2GHz and for 4.8GHz to 6GHz; initial design is $\boldsymbol{x}_{init} = [3\ 5\ 1]^T$ mm. The specifications for the 4th-order ring resonator filter are: $|S_{21}| \geq -1$dB for 1.75GHz to 2.25GHz and $|S_{21}| \leq -20$dB for 1GHz to 1.5GHz and for 2.5GHz to 3GHz; initial design $\boldsymbol{x}_{init} = [25\ 20\ 25\ 0.1\ 0.1\ 1.2\ 0.8]^T$ mm. Figures 5 and 6 show the filter responses at the initial and optimized designs, with performance specifications satisfied in both cases.

V. Conclusion

Microwave filter modeling technique using variable-fidelity electromagnetic simulations, frequency scaling and co-kriging is presented. As demonstrated through examples, our approach allows creating fast and accurate models at low computational cost.

References

[1] A.I.J. Forrester and A.J. Keane, "Recent advances in surrogate-based optimization," *Prog. in Aerospace Sc.*, vol. 45, no. 1-3, pp. 50-79, 2009.

[2] T.W. Simpson, J. Peplinski, P.N. Koch, and J.K. Allen, "Metamodels for computer-based engineering design: survey and recommendations," *Engineering with Computers*, vol.17, no. 2, pp. 129-150, July 2001.

[3] A.J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199-222, Aug. 2004.

[4] 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, April 2008.

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

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

[7] M.C. Kennedy and A. O'Hagan, "Predicting the output from complex computer code when fast approximations are available", *Biometrika,* vol. 87, pp. 1-13, 2000.

[8] D. Gorissen, K. Crombecq, I. Couckuyt, P. Demeester, and T. Dhaene, "A surrogate modeling and adaptive sampling toolbox for computer based design," *J. Machine Learning Research*, vol. 11, pp. 2051-2055, 2010.

[9] A. Manchec, C. Quendo, J.-F. Favennec, E. Rius, and C. Person, "Synthesis of Capacitive-Coupled Dual-Behavior Resonator (CCDBR) Filters," *IEEE Trans. Microwave Theory Tech.,* vol. 54, no. 6, pp. 2346-2355, June 2006.

[10] FEKO, Suite 5.3, EM Software & Systems-S.A. (Pty) Ltd, 32 Techno Lane, Technopark, Stellenbosch, 7600, South Africa, 2008.

[11] 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.

[12] J. Staum, "*Better simulation metamodeling: The why, what, and how of stochastic kriging*", Proceedings of the 2009 Winter Simulation Conference (WSC), pp. 119-133, Dec. 2009.

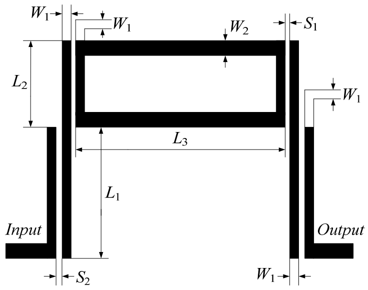

Fig. 3. Fourth-order ring-resonator bandpass filter: geometry [11]

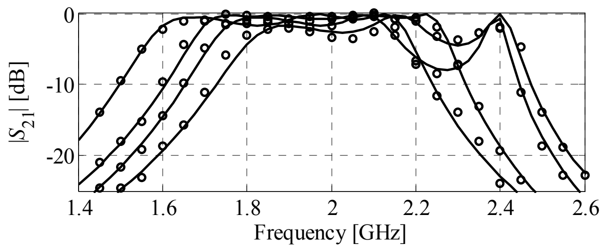

Fig. 4. CCDBR filter: responses of $\boldsymbol{R}_f$ (—) and co-kriging surrogate model (o) at selected test points. Co-kriging model created using 50 evaluations of $\boldsymbol{R}_f$ and 400 evaluations of $\boldsymbol{R}_c$.

TABLE II
4TH-ORDER RING RESONATOR FILTER: MODELING RESULTS

Model	Average Modeling Error [%]				
	$N_{KR} = 20$	$N_{KR} = 50$	$N_{KR} = 100$	$N_{KR} = 200$	$N_{KR} = 400$
$\boldsymbol{R}_{s.KR}$	13.7	10.8	4.1	3.4	2.5
$\boldsymbol{R}_{s.CO}$	3.4	3.3	2.8	2.5	2.4

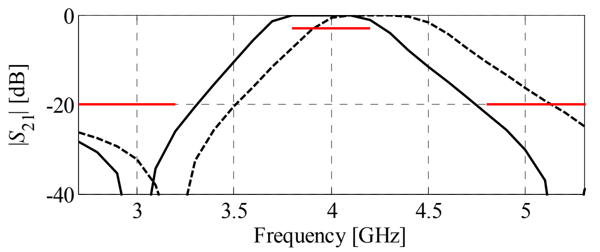

Fig. 5. CCDBR filter optimization using co-kriging surrogate: high-fidelity model responses at the initial (- - -) and at the optimized design (—).

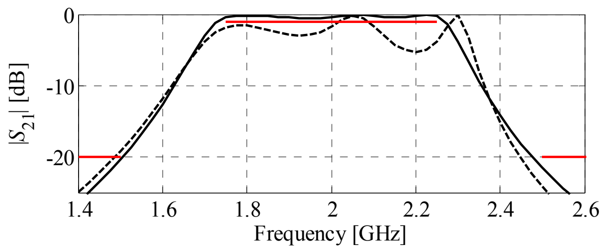

Fig. 6. Fourth-order ring-resonator filter optimization using co-kriging surrogate: high-fidelity model responses at the initial (- - -) and at the optimized design (—).