Variable-Fidelity Optimization of Microwave Filters Using Co-Kriging and Trust Regions

Slawomir Koziel

School of Science and Engineering
Reykjavik University
Reykjavik, Iceland
koziel@ru.is

Ivo Couckuyt, Tom Dhaene

Department of Information Technology
Ghent University - IBBT
Gent, Belgium
ivo.couckuyt@ugent.be, tom.dhaene@ugent.be

Abstract—**In this paper, a variable-fidelity optimization methodology for simulation-driven design optimization of filters is presented. We exploit electromagnetic (EM) simulations of different accuracy. Densely sampled but cheap low-fidelity EM data is utilized to create a fast kriging interpolation model (the surrogate), subsequently used to find an optimum design of the high-fidelity EM model of the filter under consideration. The high-fidelity data accumulated during the optimization process is combined with the existing surrogate using the co-kriging technique. This allows us to improve the surrogate model accuracy while approaching the optimum. The convergence of the algorithm is ensured by embedding it into the trust region framework that adaptively adjusts the search radius based on the quality of the predictions made by the co-kriging model. Three filter design cases are given for demonstration and verification purposes.**

Keywords-filter design; kriging; co-kriging; design optimization;trust regions

I. INTRODUCTION

Design optimization of microwave structures is more and more dependent on electromagnetic (EM) simulation. While the initial designs can be obtained—in some cases—using simplified models, e.g., analytical formulas or circuit equivalents, they need to be verified and further refined, typically by repetitive simulations. Often, such adjustment of geometry or material parameters is carried out through tedious parameter sweeps. Design automation by employing numerical optimization techniques may be capable of shortening the design cycle and making it more robust.

Accurate EM analysis is, however, computationally expensive so that the use of conventional optimization methods requiring a large number of simulations may be prohibitive. One way of reducing the design cost is through the use of so-called surrogate models, i.e., cheap and reasonably accurate representations of the structure under consideration. The surrogate can be created by approximating high-fidelity EM data using, e.g., low-order polynomials [1], kriging [1], support vector regression [2], or neural networks [3]. Unfortunately, obtaining an accurate model requires dense sampling of the design space (hundreds or thousands of sampled may be necessary), which can be justified for multiple-use library models but not for ad-hoc optimization.

Surrogate-based optimization methods exploiting physically-based models, such as space mapping (SM) [4] or simulation-based tuning [5] proved to be much more efficient alternatives to conventional methods. However, these methods are either not sufficiently robust or have other limitations (such as relatively complex implementation) to be widely accepted by engineering community. Another way of reducing the cost of EM-based design is the use of adjoint sensitivities [6], which allow us to obtain derivative information at little or no extra computational cost. Adjoint sensitivities have been successfully used in conjunction with gradient-based optimization [7], [8]. Although, this technique has recently become commercially available (e.g., [9]), the development of robust algorithm that exploits adjoint sensitivity is a relatively complex task for an inexperienced user. On the other hand, even when using derivative information, a gradient-based routine may still require considerable number of EM analyses.

Here, we introduce a robust algorithm for microwave design optimization exploiting co-kriging technology [10]. Co-kriging allows us to create the surrogate using mostly coarse-discretization EM simulations (much cheaper than the high-fidelity ones) and limited amount of high-fidelity EM data that is accumulated during the iterative process of optimizing and improving the surrogate. Co-kriging is a natural way to blend EM data of different fidelity, which allows us to yield an optimized design at a low cost corresponding to a few high-fidelity EM simulations. Good convergence properties are ensured by embedding the algorithm into the trust region framework [11]. The operation and performance of our approach is demonstrated using several design cases of microstrip filters.

II. MICROWAVE DESIGN OPTIMIZATION USING CO-KRIGING AND TRUST REGIONS

A. Design Problem Formulation

Let $\boldsymbol{R}_f \in R^m$ denote the response vector of a high-fidelity model of the microwave structure of interest (e.g., $|S_{21}|$ evaluated at m different frequencies), $\boldsymbol{x} \in R^n$ be a vector of design variables (e.g., geometry parameters), and U be a given objective function, e.g., minimax. We want to solve the following problem

$$\boldsymbol{x}_f^* = \arg \min_{\boldsymbol{x}} U\left(\boldsymbol{R}_f(\boldsymbol{x})\right) \qquad (1)$$

The high-fidelity model is assumed to be computationally expensive, typically obtained by CPU-intensive EM simulation, so that optimizing $U(\boldsymbol{R}_f(\boldsymbol{x}))$ directly may be prohibitive.

B. Surrogate Modeling Using Kriging and Co-Kriging

Kriging is an interpolation technique for deterministic noise-free data [12]. Let $X_{B.c} = \{x_c^{\ 1}, x_c^{\ 2}, \ldots, x_c^{\ N.c}\}$ be the training set and $R_c(X_{B.c})$ the associated coarse-discretization model responses. The kriging interpolant with a constant mean α is derived as,

$$R_{s.KR}(x) = \alpha + r(x) \cdot \Psi^{-1} \cdot (R_f(X_{B.c}) - 1\alpha) \tag{1}$$

where 1 is a column vector of ones. The coefficient vector α is determined by generalized least squares. $r(x)$ is an $1 \times N_{KR}$ vector of correlations between the point x and the base set $X_{B.KR}$, where the entries are $r_i(x) = \psi(x, x_c^{\ i})$, and Ψ is a $N_c \times N_c$ correlation matrix, with the entries given by $\Psi_{i,j} = \psi(x_c^{\ i}, x_c^{\ j})$. In this work, the exponential correlation function is used, i.e., $\psi(x,y) = \exp(\sum_{k=1,\ldots,n} -\theta_k |x^k - y^k|)$, where the parameters $\theta_1, \ldots, \theta_n$ are identified by Maximum Likelihood Estimation (MLE).

Co-kriging [10] is an extension to kriging where the fine data R_f and the coarse data R_c are used to make an accurate prediction. In particular, in this work, the autoregressive co-kriging model of Kennedy et al. [10] is applied as follows.

In a first step, a kriging model $R_{s.KRc}$ of the coarse data $(X_{B.c}, R_c(X_{B.c}))$ is constructed. Subsequently, in a second step, a separate kriging model $R_{s.KRd}$ is constructed on $(X_{B.f}, R_d)$, where $R_d = R_f(X_{B.f}) - \rho \cdot R_c(X_{B.f})$. The parameter ρ of the second kriging model $R_{s.KRd}$ is included in the MLE. Note that if the response values $R_c(X_{B.f})$ are not available, they can be approximated by using the first kriging model $R_{s.KRc}$, namely, $R_c(X_{B.f}) \approx R_{s.KRc}(X_{B.f})$. The resulting co-kriging interpolant is defined as

$$R_s(x) = M\alpha + r(x) \cdot \Psi^{-1} \cdot (R_d - F\alpha) \tag{3}$$

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

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

$$\Psi = \begin{bmatrix} \sigma_c^2 \Psi_c & \rho \cdot \sigma_c^2 \cdot \Psi_c(X_{B.c}, X_{B.f}) \\ 0 & \rho^2 \cdot \sigma_c^2 \cdot \Psi_c(X_{B.f}, X_{B.f}) + \sigma_d^2 \cdot \Psi_d \end{bmatrix} \tag{4}$$

$$F = \begin{bmatrix} 1 & 0 \\ \rho \cdot 1 & 1 \end{bmatrix}, \quad M = [\rho \quad 1]$$

where (σ_c, Ψ_c) and (σ_d, Ψ_d) are matrices obtained from the kriging models $R_{s.KRc}$ and $R_{s.KRd}$, respectively. 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 with the optimized $\theta_1, \ldots, \theta_n$ parameters and correlation function of the kriging models $R_{s.KRc}$ and $R_{s.KRd}$, respectively.

C. Optimization Algorithm

The co-kriging-based optimization algorithm follows the generic structure of surrogate-based approaches, i.e., produces a sequence $x^{(i)}$ of approximate solutions to (1) as follows:

$$x^{(i+1)} = \arg \min_{\|x - x^{(i)}\| \leq \delta^{(i)}} U\left(R_s^{(i)}(x)\right) \tag{5}$$

where $R_s^{(i)}$ is a co-kriging surrogate model (3) created at iteration i. The surrogate model optimization step is constrained to the vicinity of the current design $x^{(i)}$ defined by a trust region radius $\delta^{(i)}$ [13]. The value of $\delta^{(i)}$ is updated based on the gain ratio $\rho^{(i)} = [U(R_f(x^{(i+1)})) - U(R_f(x^{(i)}))]/[U(R_s^{(i)}(x^{(i+1)})) - U(R_s^{(i)}(x^{(i)}))]$ that measures that quality of the objective function improvement prediction made by the surrogate model. The trust region radius is increased if $\rho^{(i)}$ is sufficiently large and decreased if it is too small. If $U(R_f(x^{(i+1)})) \geq U(R_f(x^{(i)}))$, the new design $x(i+1)$ is rejected and the search starts again from $x^{(i)}$ using the reduced $\delta^{(i)}$. The trust region approach ensures good convergence properties of the algorithm.

The design optimization flow can be summarized as follows.

1. Set the initial design x^{init}; Optimize R_c to find $x^{(0)}$ – initial design for the co-kriging optimization;
2. Sample R_c in the vicinity of $x^{(0)}$ to obtain $(X_{B.c}, R_c(X_{B.c}))$;
3. Set $i = 0$;
4. Create a co-kriging model $R_s^{(i)}$ as in (3) using $(X_{B.c}, R_c(X_{B.c}))$ and $(X_{B.f}, R_f(X_{B.f}))$ with $X_{B.f} = \{x^{(0)}, \ldots, x^{(i)}\}$;
5. Find $x^{(i+1)}$ by optimizing $R_s^{(i)}$ as in (5);
6. Calculate gain ratio $\rho^{(i)} = [U(R_f(x^{(i+1)})) - U(R_f(x^{(i)}))]/[U(R_s^{(i)}(x^{(i+1)})) - U(R_s^{(i)}(x^{(i)}))]$;
7. If $\rho^{(i)} > \rho_{incr}$ set $\delta^{(i+1)} = \delta^{(i)} \cdot m_{incr}$, elseif $\rho^{(i)} < \rho_{decr}$ set $\delta^{(i+1)} = \delta^{(i)}/m_{decr}$, else $\delta^{(i+1)} = \delta^{(i)}$;
8. If $\|x^{(i)} - x^{(i-1)}\| < \varepsilon$ or $\delta^{(i)} < \varepsilon$ terminate;
9. If $U(R_f(x^{(i+1)})) < U(R_f(x^{(i)}))$ set $i = i + 1$ and go to 4, else set $\delta^{(i)} = \delta^{(i+1)}$ and go to 5;

Note that the co-kriging model is created in the vicinity of the R_c optimum, which is the best approximation of the optimal design we can get at a low cost. This allows us to use a limited number of R_c samples while creating the surrogate. The size of the vicinity is typically 5 to 20 percent of the design space. The initial co-kriging surrogate is created using only one evaluation of R_f and then updated using the designs obtained by optimizing the surrogate. By definition $R_s^{(i)}(x^{(k)}) = R_f(x^{(k)})$ for $k = 0, \ldots, i$, so that the surrogate accuracy improves in the vicinity of the expected optimum upon the algorithm convergence.

Here, we use the following values for the trust region radius updating parameters: $\rho_{incr} = 0.5$, $\rho_{decr} = 0.01$, $m_{incr} = 2$, $m_{decr} = 5$. The termination condition parameter ε is set to 10^{-3}. As indicated above, the algorithm is terminated either upon convergence or if the trust region size is sufficiently small.

III. DESIGN EXAMPLES

A. Half-Wavelength Stepped Impedance Resonator Filter

Consider the half-wavelength stepped impedance resonator (SIR) bandpass filter [14] in Fig. 1(a). The high-fidelity model R_f of the filter is evaluated using FEKO [15] (total mesh number is 1090, simulation time 22 min). The low-fidelity model R_c is also evaluated in FEKO (total mesh number 128, simulation time 30 s). The design variables are $x = [L_1 \ L_2 \ L_3 \ L_4 \ S \ W_1 \ W_2]^T$. The design specifications are $|S_{21}| \geq -1$ dB for 2.35 GHz to 2.45 GHz, and $|S_{21}| \leq -20$ dB for 1.5 GHz to 2.15 GHz and 2.65 GHz to 3.2 GHz.

The initial design is x^{init} = [2.5 1.5 11.6 2.95 0.328 0.305 1.613]^T mm. The filter was optimized using the co-kriging based algorithm of Section II. The approximate optimum of R_c, $x^{(0)}$ = [2.96 1.54 10.41 3.76 0.59 0.35 1.81]^T mm, is obtained at the cost of 100 evaluations of R_c. The co-kriging surrogate is created in the region $[x^{(0)} - \delta, x^{(0)} + \delta]$, with δ = [0.5 0.1 0.2 0.2 0.1 0.1 0.1]^T mm, using 100 R_c samples. The co-kriging optimization process is accomplished in 6 iterations with the optimized design $x^{(6)}$ = [2.68 1.56 10.38 3.72 0.61 0.38 1.85]^T mm. Figure 2 shows the responses of R_c and R_f at x^{init}, and the response of R_c at $x^{(0)}$. Fig. 3 shows high-fidelity model response at $x^{(6)}$. The total design cost (Table I) corresponds to about 13 evaluations of R_f. The convergence plot of the algorithm is shown in Fig. 4.

B. Fourth-Order Ring Resonator Bandpass Filter

Consider the fourth-order ring resonator bandpass filter [16] shown in Fig. 1(b). The design parameters are x = [L_1 L_2 L_3 S_1 S_2 W_1 W_2]^T mm. Both R_f (total mesh number 978; evaluation time 20 min) and R_c (mesh number 174; evaluation time 30 s) models are simulated in FEKO [15]. The design specifications are $|S_{21}| \geq -1$ dB for 1.75 GHz $\leq \omega \leq$ 2.25 GHz, and $|S_{21}| \leq -20$ dB for 1.0 GHz $\leq \omega \leq$ 1.5 GHz and 2.5 GHz $\leq \omega \leq$ 3.0 GHz. The initial design is x^{init} = [25 20 25 0.1 0.1 1.2 0.8]^T mm.

The filter was optimized using the co-kriging based algorithm of Section II starting from the approximate optimum of R_c, $x^{(0)}$ = [22.9 20.4 26.6 0.12 0.05 1.2 0.72]^T mm, obtained at the cost of 80 evaluations of R_c. The co-kriging surrogate is created in the region $[x^{(0)} - \delta, x^{(0)} + \delta]$, with δ = [0.5 0.5 0.5 0.05 0.05 0.1 0.1]^T mm, using 100 R_c samples. The final design, $x^{(7)}$ = [22.5 20.2 26.5 0.169 0.061 1.16 0.72]^T mm, is found in 7 iterations of the co-kriging optimization process. The results are shown in Figs 5 and 6 as well as in Table II. The total design cost corresponds to about 14 evaluations of R_f.
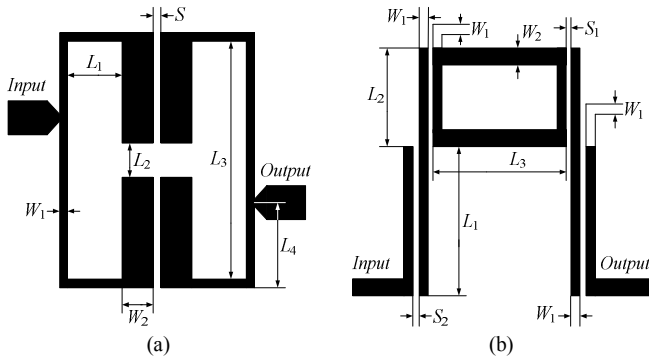

Fig. 1. Geometry of (a) half-wavelength SIR bandpass filter [14], and (b) fourth-order ring resonator bandpass filter [16].

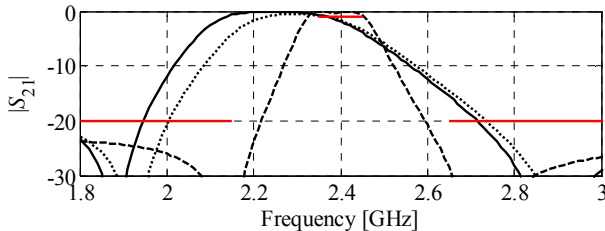Fig. 2. Half-wavelength SIR bandpass filter: responses of the low- (- - -) and high-fidelity (—) models at the initial design x^{init}.
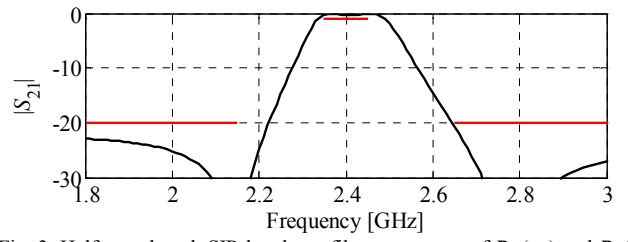

Fig. 3. Half-wavelength SIR bandpass filter: responses of R_c (····) and R_f (- - -) at $x^{(0)}$ as well as the response of R_f (—) at the final design $x^{(6)}$.
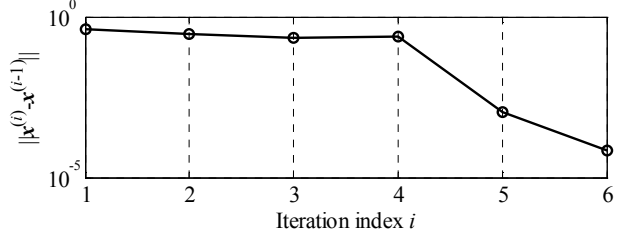

Fig. 4. Half-wavelength SIR bandpass filter: convergence of the co-kriging-based optimization algorithm.

TABLE I
HALF-WAVELENGTH SIR FILTER: OPTIMIZATION RESULTS

Algorithm Component	Number of Model Evaluations*	CPU Time	
		Absolute	Relative to R_f
Evaluation of R_c $^\$$	200	100 min	4.5
Evaluation of R_f $^\#$	9	198 min	9.0
Total cost*	N/A	298 min	**13.5**

$^\$$ Includes 100 evaluation necessary to optimize R_c and 100 evaluations to set up the co-kriging model.
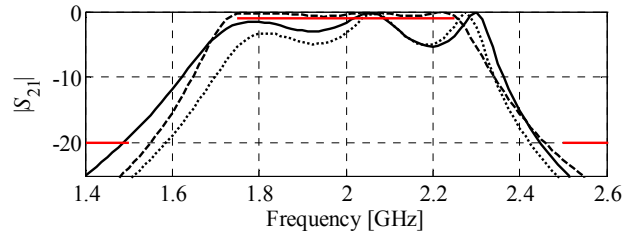$^\#$ Excludes R_f evaluation at the initial design.

Fig. 5. Forth-order ring-resonator bandpass filter: responses of the low- (- - -) and high-fidelity (—) models at the initial design x^{init}.
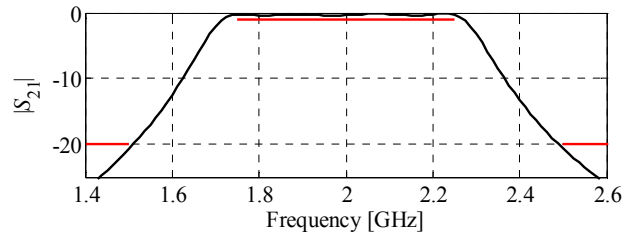

Fig. 6. Forth-order ring-resonator bandpass filter: responses of R_c (····) and R_f (- - -) at $x^{(0)}$ as well as the response of R_f (—) at the final design $x^{(???)}$.

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

Algorithm Component	Number of Model Evaluations*	CPU Time	
		Absolute	Relative to R_f
Evaluation of R_c $^\$$	180	90	4.5
Evaluation of R_f $^\#$	10	200	10.0
Total cost*	N/A	290	**14.5**

$^\$$ Includes 80 evaluation necessary to optimize R_c and 100 evaluations to set up the co-kriging model.
$^\#$ Excludes R_f evaluation at the initial design.

C. Capacitively-Coupled Dual-Behavior Resonator Microstrip Filter [17]

Consider the second-order capacitively-coupled dual-behavior resonator (CCDBR) microstrip filter [17] shown in Fig. 7. The design variables are $x = [L_1\ L_2\ L_3\ S]^T$. Both high- (total mesh number 1134; evaluation time 30 min) and low-fidelity (mesh number 130; evaluation time 36 s) models are simulated in FEKO [15]. The design specifications are $|S_{21}| \geq -3$ dB for 3.8 GHz $\leq \omega \leq$ 4.2 GHz, and $|S_{21}| \leq -20$ dB for 2.0 GHz $\leq \omega \leq$ 3.2 GHz and 4.8 GHz $\leq \omega \leq$ 6.0 GHz. The initial design is $x^{(0)} = [3.0\ 5.0\ 1.0]^T$ mm.

The co-kriging optimization starts from the approximate optimum of R_c, $x^{(0)} = [3.2\ 4.96\ 1.2]^T$ mm, obtained at the cost of 50 evaluations of R_c. The size of the region for setting up the co-kriging surrogate is $\delta = [0.25\ 0.25\ 0.25]^T$ mm. The final design, $x^{(3)} = [3.2\ 4.98\ 1.22]^T$ mm, is found in 3 iterations. The results are shown in Figs 7 and 8 as well as in Table III. The total design cost corresponds less than 8 evaluations of R_f.
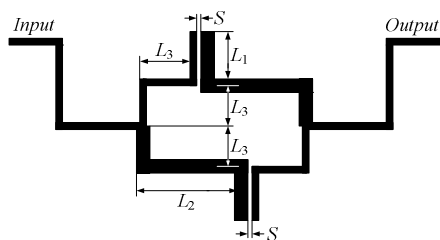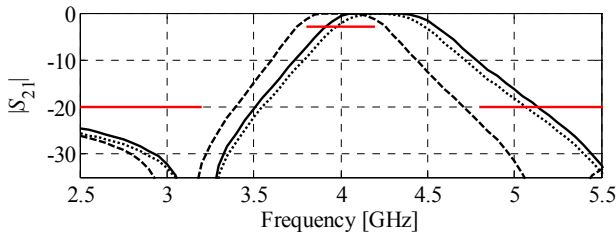

Fig. 7. Second-order CCDBR filter: geometry [17].

Fig. 8. CCDBR filter: responses of the low- (·····) and high-fidelity (——) models at the initial design x^{init}, and the response of the low-fidelity model at $x^{(0)}$ (- - -).
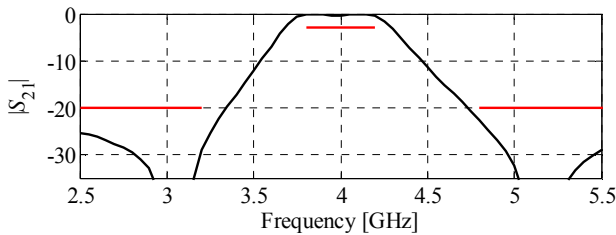

Fig. 9. CCDBR filter: responses of the high-fidelity model R_f at the final design found by the co-kriging algorithm.

TABLE III
CCDBR FILTER: OPTIMIZATION RESULTS

Algorithm Component	Number of Model Evaluations*	CPU Time Absolute	CPU Time Relative to R_f
Evaluation of R_c[$]	150	90	1.5
Evaluation of R_f[#]	6	180	6.0
Total cost*	N/A	270	**7.5**

[$]Includes 50 evaluation necessary to optimize R_c and 100 evaluations to set up the co-kriging model.
[#]Excludes R_f evaluation at the initial design.

IV. CONCLUSIONS

Robust and efficient design optimization procedure exploiting variable-fidelity electromagnetic simulations and co-kriging has been presented. The co-kriging interpolation blends together low- and high-fidelity EM data into a fast and reliable surrogate model that is iteratively optimized and updated using the high-fidelity simulation accumulated during the course of optimization. A trust region approach is used as a convergence safeguard. The performance of the algorithm is demonstrated using the filter design examples with satisfactory designs obtained at a low computational cost in each case.

REFERENCES

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

[2] G. Angiulli, M. Cacciola, and M. Versaci, "Microwave devices and antennas modeling by support vector regression machines," *IEEE Trans. Magn.*, vol. 43, no. 4, pp. 1589-1592, Apr. 2007.

[3] H. Kabir, Y. Wang, M. Yu, and Q.J. Zhang, "High-dimensional neural-network technique and applications to microwave filter modeling," *IEEE Trans. Microwave Theory Tech.*, vol. 58, no. 1, pp. 145-156, Jan. 2010.

[4] S. Koziel, Q.S. Cheng, and J.W. Bandler, "Space mapping," *IEEE Microwave Magazine*, vol. 9, no. 6, pp. 105-122, Dec. 2008.

[5] D. Swanson and G. Macchiarella, "Microwave filter design by synthesis and optimization," *IEEE Microwave Magazine*, vol. 8, no. 2, pp. 55-69, Apr. 2007.

[6] N. Georgieva, S. Glavic, M.H. Bakr, and J.W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, no. 12, pp. 2751-2758, 2002.

[7] D. Nair and J. P. Webb, "Optimization of microwave devices using 3-D finite elements and the design sensitivity of the frequency response," *IEEE Trans. Magn.*, vol. 39, no. 3, pp. 1325-1328, May 2003.

[8] M.A. El Sabbagh, M.H. Bakr, and J.W. Bandler, "Adjoint higher order sensitivities for fast full-wave optimization of microwave filters," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 8, pp. 3339-3351, 2006.

[9] CST Microwave Studio, ver. 2011, CST AG, Bad Nauheimer Str. 19, D-64289 Darmstadt, Germany, 2011.

[10] 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.

[11] A.R. Conn, N.I.M. Gould, and P.L. Toint, *Trust Region Methods*, MPS-SIAM Series on Optimization, 2000.

[12] 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.

[13] S. Koziel, J.W. Bandler, and Q.S. Cheng, "Robust trust-region space-mapping algorithms for microwave design optimization," *IEEE Trans. Microwave Theory and Tech.*, vol. 58, no. 8, pp. 2166-2174, Aug. 2010.

[14] R. Ma, G. Han, X. Chen, L. Li, L. Han, and W. Zhang, "Method for determining feed positions in coupled half-wavelength stepped impedance resonator bandpass filter," *IET Microwaves Ant. Prop.*, vol. 3, no. 5, pp. 856-862, 2009.

[15] FEKO, Suite 6.0, EM Software & Systems-S.A. (Pty) Ltd, 32 Techno Lane, Technopark, Stellenbosch, 7600, South Africa, 2010.

[16] 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.

[17] 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.