Efficient Simulation-Driven Design Optimization of Antennas Using Co-Kriging

Slawomir Koziel, Stanislav Ogurtsov
School of Science and Engineering
Reykjavik University
Reyjkavik, Iceland
koziel@ru.is, stanislav@ru.is

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

***Abstract*—We present an efficient technique for design optimization of antenna structures. Our approach exploits coarse-discretization electromagnetic (EM) simulations of the antenna of interest that are used to create its fast initial model (a surrogate) through kriging. During the design process, the predictions obtained by optimizing the surrogate are verified using high-fidelity EM simulations, and this high-fidelity data is used to enhance the surrogate through co-kriging technique that accommodates all EM simulation data into one surrogate model. The co-kriging-based optimization algorithm is simple, elegant and is capable of yielding a satisfactory design at a low cost equivalent to a few high-fidelity EM simulations of the antenna structure. To our knowledge, this is a first application of co-kriging to antenna design. An application example is provided.**

I. Introduction

Antenna design is a challenging process that involves the adjustment of dimensional and material parameters in order to satisfy various, often conflicting objectives concerning antenna figures [1]. In many cases, interactions with antenna environment (e.g., housing, connectors, etc.) can be influential so they have to be taken into account. Due to this, as well as because of the lack of reliable theoretical (closed form) models for many antenna structures (e.g., DRAs [2]), electromagnetic (EM)-simulation-driven design may be the only option.

The major bottleneck of EM-driven design is its high computational cost. High-fidelity antenna simulation may take a few hours so that straightforward approaches employing the EM solver directly in an optimization loop are impractical. Efficient design can be realized using surrogate models, fast and yet reasonably accurate representations of the structure of interest. The surrogate can be created by approximating high-fidelity EM data using, e.g., polynomials [3], kriging [3] or neural networks [4]. However, obtaining an accurate model requires dense sampling of the design space (hundreds or thousands of sampled may be necessary), which makes sense for multiple-use library models but not for ad-hoc optimization. Techniques such as space mapping (SM) [5] and tuning [6] are much more efficient because they construct the surrogate using an underlying (physically-based) coarse model. Unfortunately, applicability of these techniques for antenna design is limited. SM relies on a fast coarse model, typically, circuit equivalent [5]. Reliable circuit models are not available for many important types of antenna. On the other hand, tuning is not directly applicable for radiating structures.

Here, for the first time, we consider application of co-kriging [7] for antenna design. 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 antenna simulations. The operation and performance of our technique is demonstrated using a wideband monopole antenna.

II. Antenna Design Using Co-Kriging

A. Design Problem and Antenna Models

The antenna design task can be formulated as a nonlinear minimization problem of the form $\boldsymbol{x}^* = \text{argmin}\{\boldsymbol{x} : U(\boldsymbol{R}_f(\boldsymbol{x}))\}$, where $\boldsymbol{R}_f(\boldsymbol{x}) \in R^m$ is the response vector of a high-fidelity antenna model; U is a given objective function; $\boldsymbol{x} \in R^n$ is a vector of design variables. In this work, we use coarse-discretization EM simulations (low-fidelity model denoted as $\boldsymbol{R}_c$) to construct the antenna surrogate. $\boldsymbol{R}_c$ is faster than $\boldsymbol{R}_f$ but not as accurate.

B. Kriging and Co-Kriging Interpolation

Kriging is a popular technique to interpolate deterministic noise-free data [8]. Let $X_{B.c} = \{\boldsymbol{x}_c^1, \boldsymbol{x}_c^2, \ldots, \boldsymbol{x}_c^{N.c}\}$ be the training set and $\boldsymbol{R}_c(X_{B.c})$ the associated coarse-discretization model responses. The kriging interpolant is derived as,

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

where M and F are Vandermonde matrices of the test point $\boldsymbol{x}$ and the base set X_{BK}, respectively. The coefficient vector α is determined by Generalized Least Squares (GLS). $r(\boldsymbol{x})$ is an $1{\times}N_{KR}$ vector of correlations between the point $\boldsymbol{x}$ and the base set $X_{B.KR}$, where the entries are $r_i(\boldsymbol{x}) = \psi(\boldsymbol{x},\boldsymbol{x}_c^i)$, and Ψ is a $N_c{\times}N_c$ correlation matrix, with the entries given by $\Psi_{i,j} = \psi(\boldsymbol{x}_c^i,\boldsymbol{x}_c^j)$. In this work, the exponential correlation function is used, 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 $M = (1)$.

Co-kriging [7] is a type of kriging where the $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ model data are combined to enhance the prediction accuracy. Co-kriging is a two-steps process: first a kriging model $\boldsymbol{R}_{s.KRc}$

of the coarse data ($X_{B.c}$,$\boldsymbol{R}_c(X_{B.c})$) is constructed and on the residuals of the fine data ($X_{B.f}$,$\boldsymbol{R}_d$) a second kriging model $\boldsymbol{R}_{s.KRd}$ is applied, where $\boldsymbol{R}_d = \boldsymbol{R}_f(X_{B.f}) - \rho\cdot\boldsymbol{R}_c(X_{B.f})$. The parameter ρ is included in the MLE. Note that if the response values $\boldsymbol{R}_c(X_{B.f})$ are not available, they can be approximated by using the first kriging model $\boldsymbol{R}_{s.KRc}$, namely, $\boldsymbol{R}_c(X_{B.f}) \approx \boldsymbol{R}_{s.KRc}(X_{B.f})$. The resulting co-kriging interpolant is defined as

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

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 $\boldsymbol{R}_{s.KRd}$:

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

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

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

where (F_c,σ_c,Ψ_c,M_c) and (F_d,σ_d,Ψ_d,M_d) are matrices obtained from the kriging models $\boldsymbol{R}_{s.KRc}$ and $\boldsymbol{R}_{s.KRd}$, 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 θ_1, ..., θ_n parameters and correlation function of the kriging models $\boldsymbol{R}_{s.KRc}$ and $\boldsymbol{R}_{s.KRd}$, respectively.

C. Design Optimization Procedure

The proposed design optimization procedure can be summarized as follows.

1. Set the initial design $\boldsymbol{x}^{init}$; Optimize $\boldsymbol{R}_c$ to find $\boldsymbol{x}^{(0)}$ – initial design for the co-kriging optimization;
2. Sample $\boldsymbol{R}_c$ in the vicinity of $\boldsymbol{x}^{(0)}$ to obtain ($X_{B.c}$,$\boldsymbol{R}_c(X_{B.c})$);
3. Set $i = 0$;
4. Evaluate $\boldsymbol{R}_f$ at $\boldsymbol{x}^{(i)}$; Create a co-kriging model $\boldsymbol{R}_s^{(i)}$ as in (2) using ($X_{B.c}$,$\boldsymbol{R}_c(X_{B.c})$) and ($X_{B.f}$,$\boldsymbol{R}_f(X_{B.f})$) with $X_{B.f} = \{\boldsymbol{x}^{(0)},\ldots,\boldsymbol{x}^{(i)}\}$;
5. Find $\boldsymbol{x}^{(i+1)}$ by optimizing $\boldsymbol{R}_s^{(i)}$; Set $i = i + 1$;
6. If $\|\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(i-1)}\| < \varepsilon$ (here, $\varepsilon = 10^{-2}$) terminate, else go to 5;

Note that the co-kriging model is created in the vicinity of the $\boldsymbol{R}_c$ 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 $\boldsymbol{R}_c$ 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 $\boldsymbol{R}_f$ and then updated using the designs obtained by optimizing the surrogate. By definition $\boldsymbol{R}_s^{(i)}(\boldsymbol{x}^{(k)}) = \boldsymbol{R}_f(\boldsymbol{x}^{(k)})$ for $k = 0,\ldots,i$, so that the surrogate accuracy constantly improves in the vicinity of the expected optimum upon the algorithm convergence.

III. DESIGN EXAMPLE

Consider a monopole antenna shown in Fig. 1. The design variables are $\boldsymbol{x} = [l_1\ l_2\ l_3\ w_1]^T$. The input microstrip of the monopole is fed through an edge mount SMA connector. Simulation time of $\boldsymbol{R}_c$ (223,728 mesh cells) is 3 min, and that of $\boldsymbol{R}_f$ (1,733,892 mesh cells) is 60 min. Both models are evaluated using the transient solver of CST Microwave Studio [9]. The design specifications for reflection are $|S_{11}| \leq -15$ dB for 3 GHz to 6 GHz. The initial design is $\boldsymbol{x}^{init} = [20\ 2\ 0\ 25]^T$ mm.

The antenna was optimized using the co-kriging-based algorithm of Section II.*C*. The approximate optimum of $\boldsymbol{R}_c$, $\boldsymbol{x}^{(0)} = [22.1\ 3.44\ 0.0\ 19.1]^T$, is obtained at the cost of 80 evaluations of $\boldsymbol{R}_c$. The co-kriging surrogate is created in the region $[\boldsymbol{x}^{(0)} - \boldsymbol{\delta}, \boldsymbol{x}^{(0)} + \boldsymbol{\delta}]$, with $\boldsymbol{\delta} = [1\ 0.5\ 0.5\ 1]^T$, using 50 $\boldsymbol{R}_c$ samples. The co-kriging optimization process is accomplished in 5 iterations with the optimized design $\boldsymbol{x}^{(5)} = [21.56\ 3.08\ –0.099\ 20.11]^T$. Figure 2 shows the responses of $\boldsymbol{R}_c$ and $\boldsymbol{R}_f$ at $\boldsymbol{x}^{init}$, $\boldsymbol{x}^{(0)}$ and $\boldsymbol{x}^{(5)}$. The total design cost (Table I) corresponds to only 11 evaluations of $\boldsymbol{R}_f$.

REFERENCES

[1] J.L. Volakis (Ed.), *Antenna engineering handbook*, McGraw-Hill, 2007.

[2] A. Petosa, *Dielectric resonator antenna handbook*, Artech House, 2007.

[3] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidynathan, and P.K. Tucker, "Surrogate-based analysis and optimization," *Progress in Aerospace Sciences*, vol. 41, no. 1, pp. 1-28, Jan. 2005.

[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.

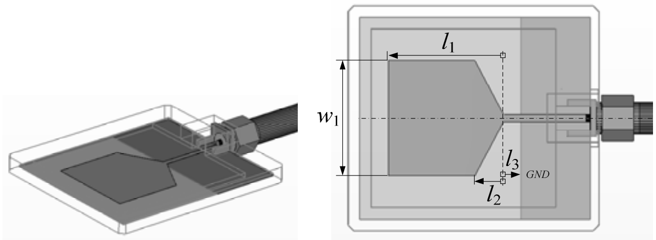
[5] J.W. Bandler, Q.S. Cheng, S.A. Dakroury, A.S. Mohamed, M.H. Bakr, K. Madsen, and J. Søndergaard, "Space mapping: the state of the art," *IEEE Trans. Microwave Theory Tech.*, vol. 52, no. 1, pp. 337-361, Jan. 2004.

[6] J.C. Rautio, "EM-component-based design of planar circuits," *IEEE Microwave Magazine*, vol. 8, no. 4, pp. 79-90, Aug. 2007.

[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.

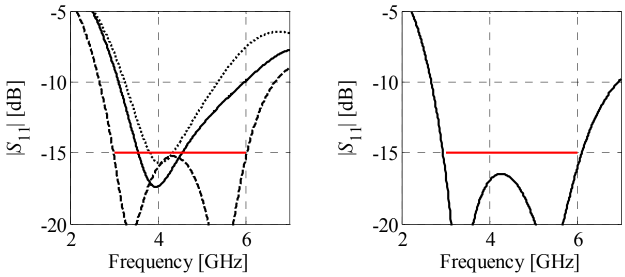
[8] D. Gorissen, K. Crombecq, I. Couckuyt, P. Demeester, and T. Dhaene, "A surrogate modeling and adaptive sampling toolbox for computer based design," *J. Machine Learning Research*, vol. 11, pp. 2051-2055, 2010.

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



(a) (b)

Fig. 1. Wideband monopole: (a) 3D view; (b) top view. The housing is shown transparent.



(a) (b)

Fig. 2. Wideband monopole: (a) responses of the low-fidelity model at $\boldsymbol{x}^{init}$ (···) and $\boldsymbol{x}^{(0)}$ (- - -) and the response of the high-fidelity model at $\boldsymbol{x}^{init}$ (—); (b) response of the high-fidelity model at the final design (—).

TABLE I. WIDEBAND MONOPOLE: DESIGN COST

Algorithm Component	Number of Model Evaluations	CPU Time	
		Absolute	*Relative to $\boldsymbol{R}_f$*
Evaluation of $\boldsymbol{R}_c$ [1]	130 × $\boldsymbol{R}_c$	390 min	6.5
Evaluation of $\boldsymbol{R}_f$ [2]	5 × $\boldsymbol{R}_f$	300 min	5.0
Total cost	N/A	690 min	**11.5**

[1] 80 evaluations to optimize $\boldsymbol{R}_c$ and 50 evaluations to set up the co-kriging model.

[2] Excludes evaluation of $\boldsymbol{R}_f$ at the initial design.