Reliable Design Closure of Sonnet-Simulated Structures Using Co-Kriging

Slawomir Koziel¹, Ivo Couckuyt², and Tom Dhaene²

¹Engineering Optimization & Modeling Center, School of Science and Engineering, Reykjavik University, 101 Reykjavik, Iceland
koziel@ru.is

²Department of Information Technology, Ghent University - IBBT Gent, Belgium
ivo.couckuyt@ugent.be, tom.dhaene@ugent.be

Abstract: A simple and reliable algorithm for design optimization of structures simulated using Sonnet em is presented. Our approach exploits coarse-discretization electromagnetic (EM) simulation data (low-fidelity model) of the structure of interest for creating a fast surrogate model through kriging interpolation. The surrogate is utilized to predict the optimum design of the structure, verified through high-fidelity EM simulations. The verification data is fed back to the surrogate through a co-Kriging technique, which allows continuous improvement of its accuracy while the optimization process progresses. The presented approach yields a satisfactory design at a low computational cost and is simple to implement.

Keywords: Computer-Aided Design (CAD), Sonnet em, Electromagnetic Simulation, Design Closure, Co-Kriging

1. Introduction

Electromagnetic (EM) simulation is nowadays a primary design tool in microwave engineering. EM-based design closure [1] in an important step of the design process where geometry and/or material parameters of the device of interest are adjusted in an iterative process involving repetitive simulations so that given performance requirements can be satisfied. In practice, such an adjustment process is often performed through parameters sweeps (typically, one parameter at a time), guided by expert knowledge. While automation of this process through numerical optimization is highly desirable, it is also quite challenging, with the fundamental difficulty being high computational cost of accurate EM evaluation. In particular, most conventional optimization techniques (both gradient-based and derivative free) require large number of EM simulations, which may be prohibitive.

Probably the most promising way of reducing the cost of EM-based design closure is by using surrogate models. In surrogate-based optimization (SBO) [2], a direct optimization of the structure under consideration (so-called high-fidelity model) is replaced by iterative updating and re-optimization of its cheap representation, the surrogate [3]. There are various ways of constructing the surrogate, from approximating sampled high-fidelity model data [2], [4] to by suitable correction of a physically-based low-fidelity (or “coarse”) model, e.g., an equivalent circuit [5].

The most successful techniques in microwave engineering exploiting physically-based surrogates are (SM) [5]-[7] and various forms of tuning [1], [8], [9] and tuning SM [10]. The tuning approaches are particularly suited to be used with Sonnet em [11] because of its co-calibrated ports technology [1].
methods include various response correction techniques such as manifold mapping [12], adaptive response correction [13] or shape-preserving response prediction [14].

Space mapping is probably the most generic approach but its efficiency heavily depends on the quality of the coarse model [15]. Also, SM normally requires that the coarse model is very fast. These requirements are often contradictory. In particular, fast coarse models (e.g., equivalent circuits) are usually not quite accurate, whereas accurate models (e.g., coarse-discretization EM simulations) are relatively expensive. In [16], an algorithm was proposed that uses SM as well as coarse-discretization Sonnet simulations and shape-preserving response prediction (SPRP) [14] to create the coarse model. This methodology proved very efficient; unfortunately, SPRP assumes that the low- and high-fidelity model response shapes must be similar (in terms of specifically defined characteristic points) for all designs considered during the optimization run. This limits the range of applications of SPRP and requires that the set of characteristic points is individually defined on case to case basis. In [17], space mapping using the coarse model constructed from coarse-discretization Sonnet simulations has been proposed which overcomes the limitations of [15].

Here, we adopt co-Kriging [18] for optimization of Sonnet-simulated structures. Co-Kriging allows us to create the surrogate using mostly coarse-discretization Sonnet simulations (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 Sonnet simulation data of different fidelity, which allows us to yield an optimized design at a low cost corresponding to a few high-fidelity simulations. Our technique is demonstrated through the design of two microstrip filters. While its efficiency is comparable [17], it is easier to implement and does not require user interaction with respect to setting up the SM surrogate nor implementing parameter extraction.

2. Design Optimization Using Co-Kriging

A. Design Optimization Problem

The design problem is formulated as a nonlinear minimization problem of the following form:

\[
x^{*}_f \in \arg \min_x U(R_f(x)),
\]

Here, \( R_f(x) \in \mathbb{R}^m \) is a response vector of a structure of interest, e.g., \([S_{21}]\) at \( m \) frequencies; \( x \in \mathbb{R}^n \) is a design variable vector; \( U \) is a scalar merit function, e.g., a minimax function with upper/lower specifications; \( x_f^{*} \) is the optimal design to be determined. Here, \( R_f \) is evaluated using Sonnet em with a \( g_{hf} \times g_{vf} \) grid.

B. Coarse-Discretization Model and Initial Optimization Stage

The optimization technique introduced here exploits a coarse-discretization model \( R_c \), also evaluated using Sonnet em. The model \( R_c \) exploits a grid \( g_{hc} \times g_{vc} \) so that \( g_{hc} > g_{hf} \) and \( g_{vc} > g_{vf} \).

The model \( R_c \) is optimized on the grid \( g_{hc} \times g_{vc} \) using a pattern search algorithm [19] in order to find a design \( x^{(0)} \) that will be used as a starting point for the next optimization stage. The resolution of this initial optimization stage is limited by the coarseness of the grid \( g_{hc} \times g_{vc} \), however, for the same reason, the computational cost of finding \( x^{(0)} \) is low and typically corresponds to a few evaluations of the fine model \( R_f \).

C. Kriging and Co-Kriging Interpolation

Kriging is a popular technique to interpolate deterministic noise-free data [20]. Let \( X_{hc} = \{ x_1^c, x_2^c, \ldots, x_{N_c}^c \} \) be the training set and \( R_c(X_{hc}) \) the associated coarse-discretization model responses. The kriging interpolant is derived as,

\[
R_{x,kr}(x) = M \alpha + r(x) \cdot \Psi^{-1} \cdot (R_f(X_{hc}) - F \alpha)
\]

where \( M \) and \( F \) are Vandermonde matrices of the test point \( x \) and the base set \( X_{hb} \), 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 \( X_{b,kr} \), where the entries are \( r_i(x) = \psi(x, x_i) \), and \( \Psi \) is a \( N_c \times N_c \) correlation matrix, with the entries given by \( \psi_{ij} = \psi(x_i^c, x_j^c) \). Here, the exponential correlation
function is used, i.e., \( \psi(x,y) = \exp(\sum_{k=1}^{n} \alpha_k x_k y_k) \), where the parameters \( \alpha_1, ..., \alpha_n \) are identified by Maximum Likelihood Estimation (MLE). The regression function is chosen constant, \( F = [1 ... 1]^T \) and \( M = (1) \).

Co-Kriging [18] is a type of kriging where the \( R_f \) and \( R_c \) model data are combined to enhance prediction accuracy. Co-Kriging is a two-steps process: first a kriging model \( R_{s.KR_c} \) of the coarse data \((X_{B_c}, R_c(X_{B_c}))\) is constructed and on the residuals of the fine data \((X_{B_f}, R_f(X_{B_f}))\) a second kriging model \( R_{s.KR_f} \) is applied, where \( R_c = R_c(X_{B_c}) - \rho R_c(X_{B_c}) \). The parameter \( \rho \) 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 \( R_{s.KR_c} \), namely, \( R_c(X_{B_f}) = R_{s.KR}(X_{B_f}) \). The resulting co-Kriging interpolant is defined as

\[
R_c(x) = M \alpha + r(x) \cdot \Psi^{-1}(R_f - F \alpha)
\]

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

\[
r(x) = [\rho \cdot \sigma^2_c \cdot r_c(x), \rho^2 \cdot \sigma^2_c \cdot r_c(x, X_{B_c}) + \sigma^2_d \cdot r_d(x)]
\]

\[
\Psi = \begin{bmatrix}
\sigma^2_c \Psi_c & \rho \cdot \sigma^2_c \cdot \Psi_c(X_{B_c}, X_{B_f}) \\
0 & \rho^2 \cdot \sigma^2_c \cdot \Psi_c(X_{B_f}, X_{B_f}) + \sigma^2_d \cdot \Psi_d
\end{bmatrix}
\]

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

where \( (F_c, \sigma_c, \Psi_c, M_c) \) and \( (F_d, \sigma_d, \Psi_d, M_d) \) are matrices obtained from the kriging models \( R_{s.KR_c} \) and \( R_{s.KR_f} \) respectively. In particular, \( \sigma^2_c \) and \( \sigma^2_d \) are process variances, while \( \Psi_c(\cdot, \cdot) \) and \( \Psi_d(\cdot, \cdot) \) denote correlation matrices of two datasets with the optimized \( \alpha_1, ..., \alpha_n \) parameters and correlation function of the kriging models \( R_{s.KR_c} \) and \( R_{s.KR_f} \) respectively.

D. Design Optimization Flow

The co-Kriging-based design optimization procedure can be summarized as follows [21]:

1. Set the initial design \( x^{(0)} \); Optimize \( R_c \) to find \( x^{(1)} \) – 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. Evaluate \( R_f \) at \( x^{(i)} \); Create a co-Kriging model \( R_c^{(i)} \) as in (3) using \((X_{B_c}, R_c(X_{B_c}))\) and \((X_{B_f}, R_f(X_{B_f}))\) with \( X_{B_c} = \{x^{(0)}, ..., x^{(i)}\} \);
5. Find \( x^{(i+1)} \) by optimizing \( R_c^{(i)} \); Set \( i = i + 1 \);
6. If \( \|x^{(i)} - x^{(i-1)}\| < \varepsilon \) (here, \( \varepsilon = 10^{-2} \)) terminate, else go to 5;

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

3. Illustration Examples

A. Compact Stacked Slotted Resonators Microstrip Bandpass Filter [22]

Consider the stacked slotted resonators bandpass filter [22] shown in Fig. 1(a). The design parameters are \( x = [L_1, L_2, W_1, S_1, S_2, d]^T \) mm. The filter is simulated in Sonnet [11] using a grid of 0.05 mm \times 0.05 mm (model \( R_f \)). The design specifications are \( |S_{21}| \geq -3 \) dB for 2.35 GHz \( \leq \omega \leq 2.45 \) GHz, and \( |S_{21}| \leq -20 \) dB for 1.9 GHz \( \leq \omega \leq 2.3 \) GHz and 2.6 GHz \( \leq \omega \leq 2.9 \) GHz. The initial design is \( x^{(0)} = [7 10 0.6 1 2 1]^T \) mm. The low-fidelity model \( R_c \) is also evaluated in Sonnet [11] using a grid of 0.2 mm \times 0.2 mm. The evaluation times for \( R_c \) and \( R_f \) are 25s and 12min, respectively.
The filter was optimized using the co-Kriging-based algorithm of Section 2. The optimum of \( \mathbf{R}_c, \mathbf{x}^{(0)} = [6.0 \ 9.6 \ 1.0 \ 1.0 \ 2.0 \ 2.0]^T \) mm. is obtained at the cost of 38 evaluations of \( \mathbf{R}_c \) using a pattern search algorithm [19] working on a grid corresponding to the simulation grid of the low-fidelity model. The co-Kriging surrogate is created in the region \( [\mathbf{x}^{(0)} - \mathbf{d}, \mathbf{x}^{(0)} + \mathbf{d}] \), with \( \mathbf{d} = [0.4 \ 0.8 \ 0.5 \ 0.5 \ 0.5 \ 0.4]^T \) mm, using 63 \( \mathbf{R}_c \) samples (13 samples of the star distribution [23] and 50 samples allocated with Latin Hypercube Sampling (LHS) [24]). The co-Kriging optimization process is accomplished in 4 iterations with the optimized design \( \mathbf{x}^{(4)} = [5.95 \ 9.5 \ 1.0 \ 0.95 \ 2.0 \ 2.15]^T \) mm. Figure 1(b) shows the responses of \( \mathbf{R}_c \) at \( \mathbf{x}^{(0)}, \mathbf{x}^{(3)} \) and \( \mathbf{x}^{(4)} \). The total design cost (Table 1) corresponds to about 8 evaluations of \( \mathbf{R}_f \). Direct optimization of the high-fidelity model has not been performed, however, the cost of such a process would be much higher as indicated by the cost of optimizing the low-fidelity model (38 \( \times \mathbf{R}_c \), cf. Table 1), where only approximate optimum was found – finding the optimum more accurately would require around 80 to 100 high-fidelity model evaluations. This indicates that the proposed approach is capable to reduce the design cost by a factor of 10.

### B. Dual-Band Bandpass Filter with Stub-Loaded Resonators [25]

Consider the dual-band bandpass with stub-loaded resonators filter [25] shown in Fig. 2(a). The design parameters are \( \mathbf{x} = [L_1 \ L_2 \ L_3 \ g_1 \ g_2 \ d \ W_1]^T \) mm. Other variables are fixed: \( W_1 = 0.5 \) mm, and \( W_2 = 1.0 \) mm. The filter is simulated in Sonnet em [11] using a grid of 0.05 mm \( \times \) 0.05 mm (high-fidelity model \( \mathbf{R}_f \)). The design specifications are \( |S_{21}| \geq -3 \) dB for 1.7 GHz \( \leq \omega \leq 1.8 \) GHz and for 3.1 GHz \( \leq \omega \leq 3.2 \) GHz, and \( |S_{21}| \leq -20 \) dB for 1.0 GHz \( \leq \omega \leq 1.5 \) GHz, 2.1 GHz \( \leq \omega \leq 2.7 \) GHz, and 3.4 GHz \( \leq \omega \leq 4.0 \) GHz. The initial design is \( \mathbf{x}^{(0)} = [12 \ 16 \ 2 \ 1 \ 1 \ 1]^T \) mm. The low-fidelity model \( \mathbf{R}_c \) is also evaluated in Sonnet em using a grid of 0.2 mm \( \times \) 0.2 mm. The evaluation times for \( \mathbf{R}_c \) and \( \mathbf{R}_f \) are 50s and 12 min, respectively.

![Fig. 1. Stacked slotted resonators filter: (a) geometry [22], (b) responses of the high-fidelity model \( \mathbf{R}_f \) at the initial design \( \mathbf{x}^{(0)} \) (dotted line), at the optimized design of \( \mathbf{R}_c, \mathbf{x}^{(0)} \), (dashed line), and at the final design obtained using co-Kriging-based algorithm (solid line).](image)

**Table 1. Optimization cost of the stacked slotted resonators bandpass filter**

<table>
<thead>
<tr>
<th>Algorithm Component</th>
<th>Number of Model Evaluations</th>
<th>Evaluation Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Absolute [min]</td>
</tr>
<tr>
<td>Optimization of the low-fidelity model ( \mathbf{R}_c )</td>
<td>( 38 \times \mathbf{R}_c )</td>
<td>16</td>
</tr>
<tr>
<td>Setting up initial kriging surrogate(^1)</td>
<td>( 63 \times \mathbf{R}_c )</td>
<td>26</td>
</tr>
<tr>
<td>Evaluation of the high-fidelity model ( \mathbf{R}_f )^2</td>
<td>( 5 \times \mathbf{R}_f )</td>
<td>60</td>
</tr>
<tr>
<td>Total optimization time</td>
<td>N/A</td>
<td>102</td>
</tr>
</tbody>
</table>

\(^1\)The base set included 13 points of so-called star distribution [23] and 50 points allocated using Latin Hypercube Sampling [24].

\(^2\)Includes evaluation of the high-fidelity model at \( \mathbf{x}^{(0)} \).
The filter was optimized using the co-Kriging-based algorithm of Section 2. The optimum of \( R_c \), \( x^{(0)} = [11.9 \ 11.8 \ 5.8 \ 1.0 \ 0.4 \ 0.3 \ 0.6]^T \) mm, is obtained at the cost of 48 evaluations of \( R_c \) using a pattern search algorithm [19] working on a grid corresponding to the simulation grid of the low-fidelity model. The co-Kriging surrogate is created in the region \( [x^{(0)} - \delta, x^{(0)} + \delta] \), with \( \delta = [0.2 \ 0.2 \ 0.2 \ 0.2 \ 0.2 \ 0.2 \ 0.2]^T \) mm, using 65 low-fidelity model samples and similar scheme as in the first example (15 samples of the star distribution [23] and 50 samples allocated with LHS [24]). The co-Kriging optimization process is accomplished in 4 iterations with the optimized design \( x^{(0)} = [11.9 \ 12.0 \ 6.0 \ 1.0 \ 0.4 \ 0.1 \ 0.6]^T \) mm. Figure 2(b) shows the responses of \( R_f \) at \( x^{\text{init}}, x^{(0)} \) and \( x^{(4)} \). The total design cost (Table 2) corresponds to about 8 evaluations of \( R_c \).

### 3. Conclusion

Simple and reliable procedure for microwave design optimization with Sonnet is discussed that utilizes coarse-discretization Sonnet simulations and co-Kriging as a way of creating a fast surrogate model of the structure under design, with sparsely referenced high-fidelity simulations used to correct the surrogate. Our technique is demonstrated through the design of two microstrip filters.

### Acknowledgement

The author thanks Sonnet Software, Inc., Syracuse, NY, for \emph{em}™. This work was supported in part by the Icelandic Centre for Research (RANNIS) Grants 110034021 and 120016021.

### References


