

Expedited Design Closure of Antennas By Means Of Trust-Region-Based Adaptive Response Scaling

Slawomir Koziel, *Senior Member, IEEE*, Sigmar D. Unnsteinsson

Abstract—In the letter, a reliable procedure for expedited design optimization of antenna structures by means of trust-region adaptive response scaling (TR-ARS) is proposed. The presented approach exploits two-level electromagnetic (EM) simulation models. A predicted high-fidelity model response is obtained by applying nonlinear frequency and amplitude correction to the low-fidelity model. The surrogate created this way is iteratively rebuilt and optimized within the trust region framework. Utilization of the correlations between the EM models of various fidelities allows for significant reduction of the design optimization cost. The main contributions of the work are twofold: (i) application of ARS for antenna optimization (in particular, making it work with coarse-discretization EM models as low-fidelity models), and (ii) integration of ARS with TR optimization framework. The operation and performance of the algorithm are demonstrated using two antenna designs optimized for several scenarios. A comparative study reveals computational benefits of TR-ARS over direct optimization of the high-fidelity EM model. Reliability of the optimization process is further confirmed by experimental validation of the fabricated antenna prototypes.

Index Terms—Antenna optimization, EM-driven design, surrogate modeling, adaptive response scaling, trust-region framework, variable-fidelity simulations.

I. INTRODUCTION

DESIGN of contemporary antennas and antenna arrays is heavily based on full-wave electromagnetic (EM) simulations. EM analysis is mandatory to ensure reliable evaluation of antenna performance. Unfortunately, high-fidelity simulation is computationally expensive particularly for electrically large structures or when electromagnetic interactions with antenna environment (e.g., housing or connectors) has to be accounted for. Furthermore, increasing topological complexity of modern antenna structures leads to a large number of geometry and/or material parameters that need to be adjusted. These reasons make EM-driven design closure a challenging task. In particular, parameter sweeping (which is still the most widely used design tuning approach) fails to find optimum designs. On the other hand, direct numerical optimization of EM-models may be very expensive

[1]-[4].

Design speedup can be achieved by exploiting adjoint sensitivities [5], [6], or, attracting more and more attention, surrogate-assisted techniques [7]-[10]. One of the most popular methods of this class in high-frequency design is space mapping, which is however of limited use in case of antennas because of the lack of fast low-fidelity models (e.g., equivalent circuits) [11]. For certain classes of antennas, such as narrow-band or multi-band, reduction of the design optimization cost can be achieved by exploiting a particular structure of the antenna responses as realized in shape-preserving response prediction [12] or feature-based optimization [13]. Another option is a utilization of auxiliary data-driven surrogates (e.g., [14]), however, this is normally limited to antennas described by a small number of parameters.

An adaptive response scaling (ARS) is one of generic techniques (i.e., not relying on a particular structure of the system response) [15]. It exploits correlations between the low- and high-fidelity models and response correction in the form of nonlinear frequency and amplitude scaling. It has been introduced for optimization of microwave structures and requires a very fast low-fidelity model (e.g., equivalent circuit). In this letter, ARS is adopted for antenna design. In order to address the issue of an expensive low-fidelity model (coarse-discretization EM simulation one), the ARS surrogate is updated during its optimization process (i.e., after each successful iteration), which considerably limits the number of low-fidelity model evaluations and makes ARS a feasible approach to antenna optimization. The ARS model construction and optimization are embedded within the trust-region framework [16]. The major novelty of the work is in making ARS applicable for antenna design optimization, which is achieved by combining it with the trust-region framework. Operation and performance of TR-ARS is demonstrated using two antenna examples, a dual-band patch and a wideband monopole. Comparison with the benchmark methods as well as experimental validation of the optimized designs are also provided.

II. ANTENNA DESIGN WITH ADAPTIVE RESPONSE SCALING

Here, a formulation of adaptive response scaling (ARS) [15] is recalled. We also outline the proposed ARS enhancements oriented towards making the technique applicable for antenna optimization.

A. Adaptive Response Scaling Formulation

The foundation of adaptive response scaling [15] is to construct a mapping between the low- (\mathbf{R}_c) and high-fidelity (\mathbf{R}_f) models of the structure under design. As opposed to many other physics-based surrogate-assisted optimization methods (e.g., space mapping [9]), ARS can be considered as non-parametric, therefore, more flexible and more generic. In case of antennas, we have two EM-simulated models, with the low-fidelity model being a coarse-discretization one. Here, the ARS concept is explained using a narrow-band antenna example. Fig. 1 shows the high- and low-fidelity reflection responses at a certain reference design $\mathbf{x}^{(i)}$ and a test design \mathbf{x} . ARS exploits the correlation between the models, and, in case of complex characteristics, the scaling process is conducted independently for the real and imaginary parts of the response(s). ARS modeling is realized in several stages. At the first stage, a frequency relationship between \mathbf{R}_c and \mathbf{R}_f is established at the design $\mathbf{x}^{(i)}$ (in the optimization context this would be the current iteration point found by the algorithm). The frequency scaling function F is obtained by solving a nonlinear regression problem of the form

$$F^{(i)}(\omega) = \arg \min_F \int_{\omega_{\min}}^{\omega_{\max}} |r_f(\mathbf{x}^{(i)}, \omega) - r_c(\mathbf{x}^{(i)}, F(\omega))| d\omega \quad (1)$$

In our implementation, the F is realized by means of cubic splines; r_f/r_c are the high- and low-fidelity responses of interest (e.g., $\text{Re}(S_{11})$); the purpose of $F^{(i)}$ is to reduce the (“horizontal”) discrepancies between the responses within the frequency range of interest ω_{\min} to ω_{\max} .

In the next stage, a frequency relationship between the low-fidelity model responses at the design $\mathbf{x}^{(i)}$ and at the (evaluation) design \mathbf{x} is established. This is done similarly as in (1), i.e.,

$$F(\mathbf{x}, \omega) = \arg \min_F \int_{\omega_{\min}}^{\omega_{\max}} |r_c(\mathbf{x}, \omega) - r_c(\mathbf{x}^{(i)}, F(\omega))| d\omega \quad (2)$$

The third stage is to quantify the level (amplitude) relationship between the frequency-scaled low-fidelity model at the reference design and the low-fidelity model at the evaluation design (here, \div denotes component-wise division)

$$A(\mathbf{x}, \omega) = [r_c(\mathbf{x}, \omega) + M] \div [r_c(\mathbf{x}^{(i)}, F(\mathbf{x}, \omega)) + M] \quad (3)$$

The shift by M (we use $M = 1$) is necessary to avoid division by zero (for frequencies for which $r_c(\mathbf{x}^{(i)}, F(\mathbf{x}, \omega)) = 0$).

At the prediction stage, ARS surrogate model r_s estimated the high-fidelity model response at the evaluation design \mathbf{x} . We have

$$r_s(\mathbf{x}, \omega) = A(\mathbf{x}, F^{(i)}(\omega)) \circ [r_f(\mathbf{x}^{(i)}, F(\mathbf{x}, \omega)) + M] - M \quad (4)$$

where \circ denotes component-wise multiplication.

The above concepts are explained in Figs. 1(b) and 2. Fig. 1(b) shows $\text{Re}(S_{11})$ of the responses shown in Fig. 1(a) as

well as the response of the ARS surrogate. The prediction power of ARS is very good also at the level of $|S_{11}|$ (cf. Fig. 2). This is not the case for conventional correction (output space mapping) which fails to account for frequency changes of the response. For practical applications, ARS is realized at the level of real and imaginary parts of the reflection response which makes it more accurate [15], and further recalculated to whatever form is required by calculation of the design objectives (here, the absolute value).

B. ARS for Antenna Design

In original implementation of ARS [16], construction the the surrogate model $\mathbf{R}_s^{(i)}(\mathbf{x})$ at the current design $\mathbf{x}^{(i)}$ is followed by its optimization to yield the next approximation

$$\mathbf{x}^{(i+1)} = \arg \min \{\mathbf{x} : U(\mathbf{R}_s^{(i)}(\mathbf{x}))\} \quad (5)$$

of the optimum design $\mathbf{x}^* = \arg \min \{\mathbf{x} : U(\mathbf{R}_f(\mathbf{x}))\}$ (U is the objective function). Subsequently, the high-fidelity model is evaluated and the surrogate is reset (both at $\mathbf{x}^{(i+1)}$). This procedure works fine assuming that the low-fidelity model is very fast (e.g., an equivalent circuit as in [15]). In case of coarse-discretization EM antenna models, iterative solving of (5) is just too expensive.

In order to make ARS feasible for antenna design, a modified procedure is proposed here. It generates a sequence of designs $\mathbf{x}^{(i)}$, $i = 0, 1, \dots$, (approximations to \mathbf{x}^*) as

$$\mathbf{x}^{(i+1)} = \arg \min_{\mathbf{x}, \|\mathbf{x} - \mathbf{x}^{(i)}\| \leq \delta^{(i)}} U(\mathbf{R}_s(\mathbf{x}^{(i)}) + \mathbf{J}_{\mathbf{R}_s^{(i)}}(\mathbf{x}^{(i)}) \cdot (\mathbf{x} - \mathbf{x}^{(i)})) \quad (6)$$

where $\mathbf{J}_{\mathbf{R}_s^{(i)}}$ is a Jacobian of the ARS surrogate estimated using finite differentiation. The cost of each iteration of (6) is one evaluation of the high-fidelity model and n evaluations of \mathbf{R}_c (n being the number of design variables).

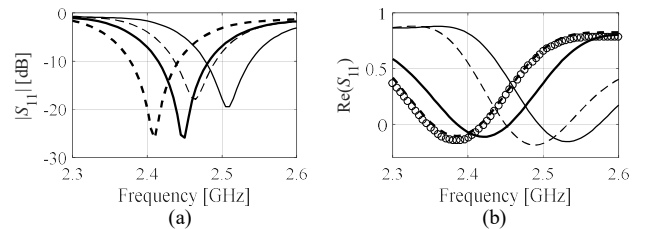


Fig. 1. Example narrow-band antenna responses: (a) $|S_{11}|$ at a reference design $\mathbf{x}^{(i)}$ (—) and at evaluation design \mathbf{x} (- - -); High- and low-fidelity models shown using thick and thin lines, respectively. Designs are taken arbitrarily for illustration purposes; (b) responses at $\mathbf{x}^{(i)}$ (—) and \mathbf{x} (- - -) (cf. Fig. 1(a)); \mathbf{R}_f (thick lines) and \mathbf{R}_c (thin lines); ARS surrogate model response shown using circles.

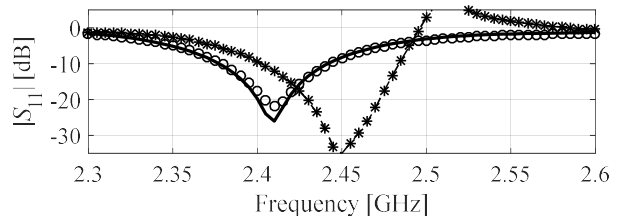


Fig. 2. High-fidelity model response at the evaluation design \mathbf{x} (thick line), ARS surrogate model response (o), and conventional response correction (output space mapping) (*). Note severe distortion of the latter due to inability to account for frequency changes of the low-fidelity model response.

The entire process is embedded in the trust-region (TR) framework [16] in which δ^j is the TR radius, updated in each iteration. The design $\mathbf{x}^{(1)}$ is obtained through direct optimization of the low-fidelity model.

The differences between (6) and the original implementation of the ARS algorithm should be explained. In (5), a complete optimization process of the ARS surrogate is carried out, which incurs a considerable number of low-fidelity model evaluations. In (6), the surrogate model is updated upon each successful TR iteration. In practice, a larger number of iterations (6) may be necessary to converge, however, the overall computational cost of the optimization process is lower.

It should also be clarified that the ARS algorithm can work with any objective function U (cf. (5)). The purpose of ARS is to improve alignment between the corrected low-fidelity model and the high-fidelity EM model, whereas objective function only translated design requirements imposed on the antenna structure into a scalar value to be minimized. In the examples shown in Section III, optimization of input characteristics is considered as a representative situation in antenna design.

III. CASE STUDY AND RESULTS

The proposed optimization approach is demonstrated using two antenna structures. The first test case is a dual-band planar antenna shown in Fig. 3(a). The structure is implemented on RF-35 substrate ($h = 0.762$ mm, $\epsilon_r = 3.5$). The design variables are $\mathbf{x}^I = [L \ l_1 \ l_2 \ l_3 \ W \ w_1 \ w_2 \ g]^T$; $o = 7$, $l_0 = 10$, $s = 0.5$, and $w_0 = 1.7$, are fixed (all dimensions in mm). The EM models are implemented in CST Microwave Studio: \mathbf{R}_f (~3,300,000 cells, 90 minutes) and \mathbf{R}_c (~82,000, ~138,000, and ~154,000 mesh cells, simulation time 88, 139, and 169 seconds, for the three mesh densities corresponding to the global mesh parameter LPW (lines per wavelength) of 12, 14, and 16). Some discussion concerning the low-fidelity model selection can be found in [7] and [17]. The second structure is a UWB monopole shown in Fig. 4(b), implemented on FR4 substrate ($\epsilon_r = 4.4$, $h = 1.55$ mm), with design variables $\mathbf{x}^{II} = [L_g \ L_0 \ L_s \ W_s \ d \ dL \ d_s \ dW_s \ dW \ a \ b]^T$. The EM models are implemented in CST: \mathbf{R}_f (~5,500,000 cells, 22 minutes) and \mathbf{R}_c (~195,000, ~360,000, and ~500,000 mesh cells, simulation time 102, 138, and 236 seconds, for the three mesh densities corresponding to the global mesh parameter LPW of 12, 14, and 16). The high-fidelity models are equipped with the SMA connector.

The dual-band patch antenna has been optimized using ARS to maximize the symmetric bandwidth at the operating frequencies of 2.45 GHz and 5.3 GHz. The objective function was defined as follows. Let f_{k1} and f_{k2} , $k = 1, 2$, denote the frequencies corresponding to the beginning and end of the operating band for the lower and upper band, respectively. This information can be extracted from the simulated antenna response. We have $U(\mathbf{R}_f(\mathbf{x})) = -\min\{B_1, B_2\}$, where $B_k = 2 \cdot \min\{f_{k0} - f_{k1}, f_{k2} - f_{k0}\}$, $k = 1, 2$, is a symmetric part of the antenna bandwidth. In case the resonance is not sufficiently

deep (i.e., above -10 dB), a penalty term is used instead proportional to the difference between -10 dB and the resonance level at its minimum.

The design objective for the UWB monopole was to minimize the maximum in-band reflection with the frequency range of 3.1 GHz to 10.6 GHz. Here, the objective function is simply defined as $U(\mathbf{R}_f(\mathbf{x})) = \max\{3.1 \text{ GHz} \leq f \leq 10.6 \text{ GHz} : |S_{11}(\mathbf{x}, f)|\}$ (minimization of the maximum reflection within the UWB frequency range).

The ARS-optimized designs for low-fidelity models with LPW = 16 are $\mathbf{x}^* = [15.83 \ 3.57 \ 12.63 \ 5.54 \ 16.43 \ 1.57 \ 9.00 \ 6.70]^T$, and $\mathbf{x}^{II*} = [9.67 \ 14.09 \ 9.36 \ 0.49 \ 4.03 \ 7.64 \ 0.99 \ 1.15 \ 3.29 \ 0.29 \ 0.60]^T$, see also Fig. 4.

ARS has been compared to build-in CST optimizers (a trust region algorithm, a Nelder-Mead algorithm, genetic algorithm and particle swarm optimizer, cf. Tables I and II) as well as direct optimization of the high-fidelity model. It can be observed that ARS offers significant cost savings without compromising accuracy. This is despite the fact that the time evaluation of the high- and low-fidelity model is quite low. It should also be noted that in some cases, CST optimizers fail or provide relatively poor results, especially the stochastic methods, which were terminated upon reaching five hundred EM simulations of the antenna at hand. On the other hand, the number of high-fidelity model evaluations during ARS optimization is very low (typically a few) which translates into low overall cost. Finally, ARS performance for various low-fidelity model discretization densities has been investigated. It can be observed that the results are consistent although slight degradation is observed when discretization is reduced, which can be attributed to the increased level of numerical noise pertinent to low-fidelity EM analysis.

The verification designs have been manufactured and measured for the sake of additional validation. Fig. 5 shows the photographs of the fabricated prototypes, whereas Figs. 6 and 7 show the reflection characteristics and the E- and H-plane patterns, respectively. The agreement between the simulations and measurements is acceptable which additionally validates the reliability of the optimization procedure. Discrepancies for E-plane radiation pattern result from the measurement setup (shadowing effect of the 90-degree bend used to mount the structure as well as a piece of cable extending from the measurement tower, neither of which have been included in the computational model of the antennas).

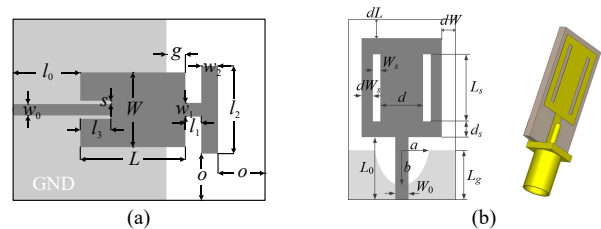


Fig. 3. Benchmark antennas: (a) dual-band patch antenna; ground plane marked using the lighter shade of gray, (b) UWB monopole.

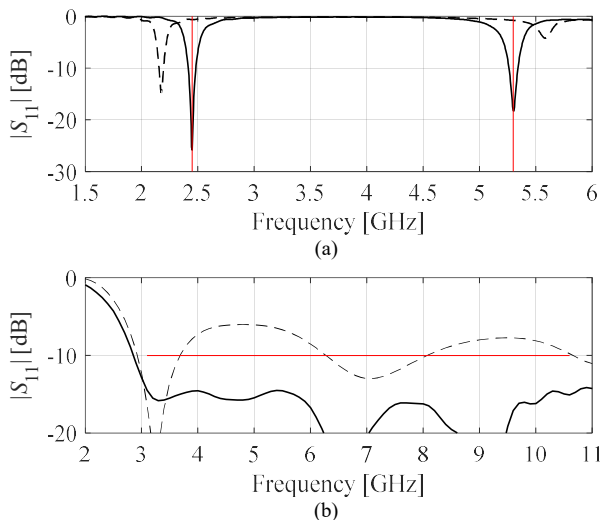


Fig. 4. Initial (---) and optimized (—) antenna responses obtained using ARS: (a) dual-band patch, (b) UWB monopole.

TABLE I DUAL-BAND ANTENNA OPTIMIZATION RESULTS

| Optimization algorithm | Optimization cost | | | Objective function [§] |
|-----------------------------|-------------------|---------|--------------------|---------------------------------|
| | # R_c | # R_f | Total [#] | |
| CST built-in (trust-region) | - | 22 | 22 | failed |
| CST built-in (Nelder-Mead) | - | 488 | 488 | 0.071 |
| CST built-in (GA) | - | 500 | 500 | 0.027 |
| CST built-in (PSO) | - | 500 | 500 | 0.006 |
| Direct optimization | - | 75 | 75 | 0.068 |
| ARS (R_c with LPW = 12) | 112 | 8 | 9.7 | 0.050 |
| ARS (R_c with LPW = 14) | 103 | 8 | 10.5 | 0.036 |
| ARS (R_c with LPW = 16) | 117 | 7 | 10.4 | 0.065 |

[§] Minimum of the symmetric bandwidth for the two operating frequencies at the final design.
[#] Optimization cost expressed in the equivalent number of high-fidelity number evaluations.

TABLE II UWB MONOPOLE OPTIMIZATION RESULTS

| Optimization algorithm | Optimization cost | | | Objective function [§] |
|-----------------------------|-------------------|---------|--------------------|---------------------------------|
| | # R_c | # R_f | Total [#] | |
| CST built-in (trust-region) | - | 59 | 59 | -10.9 |
| CST built-in (Nelder-Mead) | - | 375 | 375 | -13.5 |
| CST built-in (GA) | - | 500 | 500 | -11.4 |
| CST built-in (PSO) | - | 500 | 500 | -12.9 |
| Direct optimization | - | 152 | 152 | -15.0 |
| ARS (R_c with LPW = 12) | 168 | 12 | 25.4 | -12.9 |
| ARS (R_c with LPW = 14) | 190 | 11 | 31.6 | -13.9 |
| ARS (R_c with LPW = 16) | 192 | 11 | 46.7 | -14.5 |

[§] Maximum in-band reflection level (operating frequency 3.1 GHz to 10.6 GHz).
[#] Optimization cost expressed in the equivalent number of high-fidelity number evaluations.

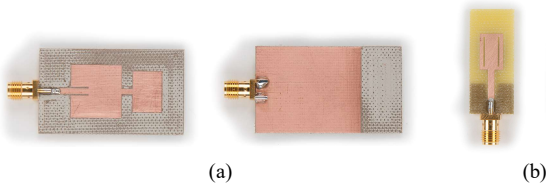


Fig. 5. Photographs of the fabricated prototypes of the ARS-optimized antennas: (a) dual-band patch, (b) UWB monopole.

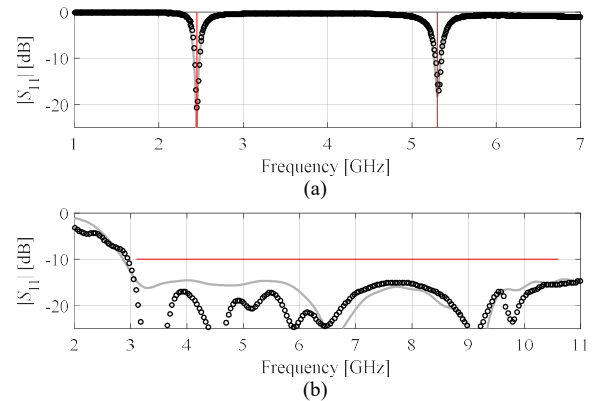


Fig. 6. Simulated (gray) and measured (o) reflection characteristics of the antennas of Fig. 4: (a) dual-band patch, (b) UWB monopole.

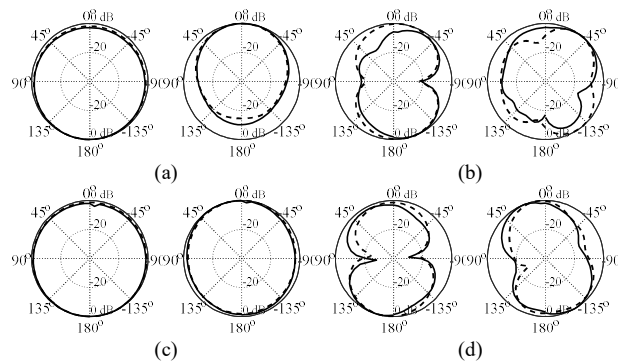


Fig. 7. Simulated (---) and measured (—) radiation patterns of the antennas of Fig. 5: (a) H-plane for the dual-band patch at 2.45 GHz and 5.3 GHz, (b) E-plane for the dual-band patch at 2.45 GHz and 5.3 GHz, (c) H-plane for the UWB monopole at 4 GHz and 8 GHz, (d) E-plane for the UWB monopole at 4 GHz and 8 GHz.

IV. CONCLUSION

A trust-region adaptive response scaling (TR-ARS) for antenna design optimization has been presented. Our approach builds upon the original ARS by interleaving the process of surrogate model construction and optimization (all embedded in the trust region framework). This combination is the major contribution of the work as it allows for efficient utilization of coarse-discretization EM antenna models, not possible within the original ARS implementation. Verification cases studies indicate considerable speedup compared to direct optimization of the high-fidelity model despite a relatively high cost of the coarse-mesh model. ARS technique is a generic one and can be utilized to handle a variety of designs, including antenna structures (as indicated in the paper) but also other high-frequency designs such as microwave filters or couplers. A particular ARS implementation proposed here is particularly suited for cases where the computational cost of the low-fidelity model is relatively high.

ACKNOWLEDGMENT

The authors would like to thank Dassault Systemes, France, for making CST Microwave Studio available. This work has been partially supported by the Icelandic Centre for Research (RANNIS) Grant 174114051 and by National Science Centre of Poland Grant 2015/17/B/ST6/01857.

REFERENCES

- [1] S. Chamaani, M.S. Abrishamian, S.A. Mirtaheeri, "Time-domain design of UWB Vivaldi antenna array using multiobjective particle swarm optimization," *IEEE Ant. Wireless Prop. Lett.*, vol. 9, pp. 666-669, 2010.
- [2] K. Choi, D.H. Jang, S.I. Kang, J.H. Lee, T.K. Chung, and H.S. Kim, "Hybrid algorithm combining genetic algorithm with evolution strategy for antenna design," *IEEE Trans. Magn.*, vol. 52, no. 3, 2016.
- [3] Z.D. Zaharis, I.P. Gravas, T.V. Yioultis, P.I. Lazaridis, I.A. Glover, C. Skeberis, and T.D. Xenos, "Exponential log-periodic antenna design using improved particle swarm optimization with velocity mutation," *IEEE Trans. Magn.*, vol. 53, no. 6, 2017.
- [4] J. Wang, X.S. Yang, X. Ding, and B.Z. Wang, "Antenna radiation characteristics optimization by a hybrid topological method," *IEEE Trans. Ant. Prop.*, vol. 65, no. 6, pp. 2843-2854, 2017.
- [5] M. Ghassemi, M. Bakr, and N. Sangary, "Antenna design exploiting adjoint sensitivity-based geometry evolution," *IET Microwaves Ant. Prop.*, vol. 7, no. 4, pp. 268-276, 2013.
- [6] S. Koziel, and A. Bekasiewicz, "Rapid design optimization of antennas using variable-fidelity EM models and adjoint sensitivities," *Eng. Comp.*, vol. 33, no. 7, pp. 2007-2018, 2016.
- [7] S. Koziel and S. Ogurtsov, "Antenna design by simulation-driven optimization. Surrogate-based approach," Springer, 2014.
- [8] S. Koziel and A. Bekasiewicz, "Multi-objective design of antennas using surrogate models," World Scientific, 2016.
- [9] M. Simsek and A. Aoad, "Multiple operating frequency selections for reconfigurable antenna design by SM based optimisation," *IET Microwaves, Ant. Prop.*, vol. 11, no. 13, pp. 1898-1908, 2017.
- [10] J.A. Easum, J. Nagar, and D.H. Werner, "Multi-objective surrogate-assisted optimization applied to patch antenna design," *Int. Symp. Ant. Prop.*, pp. 339-340, San Diego, USA, 2017.
- [11] S. Koziel, Q.S. Cheng, and J.W. Bandler, "Space mapping," *IEEE Microwave Magazine*, vol. 9, no. 6, pp. 105-122, Dec. 2008.
- [12] S. Koziel, S. Ogurtsov, and S. Szczepanski, "Rapid antenna design optimization using shape-preserving response prediction," *Bulletin of the Polish Academy of Sciences. Tech. Sc.*, vol. 60, no. 1, pp. 143-149, 2012.
- [13] S. Koziel, "Fast simulation-driven antenna design using response-feature surrogates," *Int. J. RF & Micr. CAE*, vol. 25, no. 5, pp. 394-402, 2015.
- [14] D.I.L. de Villiers, I. Couckuyt, and T. Dhaene, "Multi-objective optimization of reflector antennas using kriging and probability of improvement," *Int. Symp. Ant. Prop.*, pp. 985-986, San Diego, USA, 2017.
- [15] S. Koziel and A. Bekasiewicz, "Rapid microwave design optimization in frequency domain using adaptive response scaling," *IEEE Trans. Microwave Theory Tech.*, vol. 64, no. 9, pp. 2749-2757, 2016.
- [16] 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, 2010.
- [17] S. Koziel and A. Bekasiewicz, "Low-fidelity model considerations for EM-driven design of antenna structures," *J. Electromagnetic Waves App.*, vol. 30, no. 18, pp. 2444-2458, 2016.