Postprint of: Pietrenko-Dąbrowska A., Kozieł S., Expedited Gradient-Based Design Closure of Antennas Using Variable-Resolution Simulations and Sparse Sensitivity Updates, IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, Vol. 70 No 6 (2022), pp. 4925 - 4930, DOI: 10.1109/TAP.2021.3138487

Expedited Gradient-Based Design Closure of Antennas Using Variable-Resolution Simulations and Sparse Sensitivity Updates

Anna Pietrenko-Dabrowska and Slawomir Koziel

***Abstract*—Numerical optimization has been playing an increasingly important role in the design of contemporary antenna systems. Due to the shortage of design-ready theoretical models, optimization is mainly based on electromagnetic (EM) analysis, which tends to be costly. Numerous techniques have evolved to abate this cost, including surrogate-assisted frameworks for global optimization, or sparse sensitivity updates for speeding up local search. In the latter, CPU-heavy updates of the system response sensitivity through finite differentiation are suppressed based on, e.g., the magnitude of design variability during the optimization run. Another approach is to incorporate variable-resolution simulations. Recently, a technique exploiting a continuous spectrum of admissible model fidelity levels has been reported, thereby allowing for a considerable reduction of the computational expenditures. Seeking further savings, this work introduces an accelerated gradient-based algorithm with sparse sensitivity updates and variable-resolution EM simulations. Our technique is validated using four broadband antennas, and demonstrated to offer substantial (around eighty percent) savings over the benchmark while maintaining acceptable design quality.**

***Index Terms*—Antenna design; parameter tuning; EM-driven optimization; design specification adaptation; multi-band antennas**

I. Introduction

Modern antennas feature increasingly complex topologies, partially resulting from stringent performance requirements implied by emerging application areas such as 5G technology [1], [2], body-centric communication solutions [3], the internet of things (IoT) [4], or wearable [5] and implantable devices [6]. Functionality demands (multi-band operation [7], pattern diversity [8], circular polarization [9], or band-notches [10]) only redound to this complexity. Reliable evaluation of m antennas requires full-wave electromagnetic (EM) analysis. At the same time, rigorous numerical optimization has been playing increasingly important role in fine tuning of antenna parameters. The reasons are manifold: the need for a simultaneous refinement of numerous parameters, the necessity of handling multiple objectives, and, clearly, boosting the performance as much as possible. Thus, EM-driven optimization appears to be imperative in the design of contemporary antenna structures. Yet, it is an expensive procedure, even for local algorithms (both gradient-based [11], and derivative-free [12]). For global optimization, the expenditures may be exorbitant (thousands of EM simulations [13]). This makes applicability of otherwise popular global algorithms, e.g., particle swarm optimizers [14], [15], or evolutionary algorithms [16], [17], significantly limited to the cases described by few parameters and within restricted ranges thereof when directly handing EM simulation models. The same pertains to multi-objective (MO) optimization [18]. The cost entailed by MO versions of population-based metaheuristic algorithms [19], [20] is unacceptably high when carried out at the level of EM simulations. Solving tasks related to uncertainty quantification [21] or tolerance-aware design [22] is similarly expensive.

Over the last two decades, a considerable research effort has been devoted to alleviating high cost of EM-driven design. This includes utilization of adjoint sensitivities [23], as well as an employment of fast replacement models (surrogates or metamodels).

Manuscript received May 28, 2021. This work was supported in part by the Icelandic Centre for Research (RANNIS) Grant 217771, by the National Science Centre of Poland Grant 2020/37/B/ST7/01448.

S. Koziel is with Engineering Optimization and Modeling Center of Reykjavik University, Reykjavik, Iceland (e-mail: koziel@ru.is). A. Pietrenko-Dabrowska and also S. Koziel are with Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, 80-233 Gdansk, Poland.

The efficacy of the surrogate-assisted frameworks is subject to the surrogate being considerably faster than EM analysis while providing adequate accuracy. For local optimization or robust design, the surrogates are normally constructed along the optimization path, whereas global surrogate-based optimization requires the metamodel to be valid over the entire parameter space. There are two main categories of surrogate models: data-driven [24] and physics-based ones [25]. The former are versatile and easily accessible [26], [27]. Popular methods include kriging [28], radial basis functions [29], support vector regression [30], polynomial chaos expansion [31], and neural networks [32]. Despite the advantages, the curse of dimensionality and nonlinearity of system responses hinders the employment of data-driven surrogates for antenna design. Many attempts have been made to mitigate these difficulties. A popular approach is sequential sampling [33] offering adaptive allocation of training points that accounts for the shape of the functional landscape being modeled. Other techniques include high-dimensional model representation (HDMR) [34], model order reduction (MOR) [35], or orthogonal matching pursuit [36]. Physics-based models (space mapping [37], response correction methods [38], or feature-based optimization [39]) are less prone to the dimensionality issues due to the involvement of the underlying low-fidelity models. Yet, unavailability of low-fidelity models makes the physics-based approach of limited use in antenna design.

Practical EM-driven antenna design largely relies on local optimization algorithms. Local optimizers are significantly cheaper than the global routines (nature-inspired algorithms [40], [41], or machine learning methods [42], [43]). Also, local search is normally sufficient if a reasonable initial design is at hand, e.g., when re-designing an antenna for somewhat different target operating frequencies, or to boost performance (e.g., improve antenna matching or gain). Several methods exist for accelerating local search routines, including sparse sensitivity updates [44]-[47], and variable-fidelity methods [48]-[53]. The former rely on reducing the amount of expensive finite-differentiation-based evaluations of the response gradients, thereby allowing for cost savings exceeding fifty percent at the cost of a moderate deterioration of the design quality. The methods for lowering the optimization cost include various decision-making schemes for omitting some finite differentiation (FD) sensitivity updates during the optimization run (e.g., design relocation change [44] or gradient variability [45] between iterations). Other strategies include replacing costly FD updates with a Broyden updating formula [46], or a combination of the above mentioned approaches [47].

Another method for expediting local search is the usage of variable-fidelity simulations. Most of the reported variable-fidelity frameworks exploit two levels of resolution: low- and high-fidelity. In antenna design, due to the lack of parameterized equivalent network models [48], relatively expensive coarse-discretization models seem to be the only option [49]. A properly corrected low-fidelity model allows for predicting high-fidelity responses throughout the optimization process. The efficacy of these methods depends on the setup of the optimization framework, including the appropriate selection of the low-fidelity model [50], and the correction technique [51]. Another issue is the model management scheme, i.e., the selection of the discretization levels, and the criteria for switching between them [52]. In a recent technique of [53], a discretization level at each stage of the optimization run is selected from the continuous spectrum of admissible resolutions: from the lowest acceptable accuracy up to the high-fidelity model decided upon by the designer. The selection

criterion is based on the algorithm convergence status. The optimization process is launched with the coarsest model, which allows for a low-cost exploration of parameter space. Upon convergence, the fidelity is increased to ensure reliability.

This work proposes a gradient-based optimization technique, in which the variable-resolution technique of [53] is accelerated by incorporating sparse sensitivity updates [45]. The FD updates are suppressed when stable patterns of the system parameter sensitivities are detected. The merger of the two acceleration mechanism allows for a significant reduction of CPU expenses while retaining satisfactory design quality. Our approach has been benchmarked against the conventional reference gradient-based procedure (trust-region (TR) gradient search), its accelerated version with sensitivity change monitoring, and variable-resolution algorithm with full-FD sensitivity updates. As demonstrated using four broadband antenna examples, the proposed approach is considerably more efficient than the benchmark procedures (average speedup is almost eighty percent as compared to the TR algorithm, and around fifty percent over multi-resolution algorithm with full-FD sensitivity updates) without compromising the design quality.

The originality and the technical contributions of the paper can be summarized as follows: (i) the development of an algorithm for direct optimization of antenna structures, capitalizing on variable-resolution EM simulations and sparse sensitivity updates, (ii) demonstrating the possibility of accomplishing the optimization process at remarkably low cost corresponding to about two dozens of high-fidelity EM analyses, which, according to the authors knowledge is the lowest cost reported for direct (i.e., not surrogate-based) methods, (iii) providing the procedure that is fully automated and globally convergent. The computational complexity of our algorithm is weakly dependent on the parameter space dimensionality, and offers the same level or reliability as classical gradient-based methods. Consequently, it can be considered a practical alternative to existing methods, and can be readily incorporated into EM simulation environments.

II. Expedited Antenna Optimization with Sparse Sensitivity Updates and Variable-Resolution Simulations

The purpose of this section is to introduce the proposed optimization procedure. It exploits trust-region gradient-based search as an optimization engine, which is expedited through incorporating the two following acceleration mechanisms: (i) sparse sensitivity updates [45], and (ii) variable-resolution EM simulations [53]. The former utilizes gradient variability tracking to decide upon omitting costly FD updates, whereas the latter exploits convergence-based decision-making scheme for appointing appropriate model fidelity level at each stage of the optimization run: starting from the coarsest, through the intermediate ones, up to the high-fidelity representation when closer to convergence.

A. EM-driven Antenna Design by Numerical Optimization

To fulfil increasingly stringent design specifications imposed on modern antenna structures, it is often necessary to resort to numerical optimization. A final tuning of geometry parameters is referred to as design closure [24]. It requires a definition of a design quality metric, which is typically a scalar function of design variables, unless multi-objective optimization is of interest [54]. At the presence of multiple goals, a possible approach is to select a primary objective and handle the remaining ones implicitly through constraints. A single-objective design task is defined as

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} U(\boldsymbol{x}) \tag{1}$$

where $\boldsymbol{x}$ denotes the vector of geometry parameters, and U stands for the objective function. In addition, (1) might be subject to inequality constraints $g_k(\boldsymbol{x}) \le 0$, $k = 1, \ldots, n_g$, and equality constraints $h_k(\boldsymbol{x}) = 0$, $k = 1, \ldots, n_h$. Evaluation of electrical/field-related constraints involves costly EM analysis. A convenient way of dealing with them is offered by a penalty function approach [55], where the problem is reformulated into

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} U_P(\boldsymbol{x}) \tag{2}$$

with the merit function U_P taking the form

$$U_P(\boldsymbol{x}) = U(\boldsymbol{x}) + \sum_{k=1}^{n_g+n_h} \beta_k c_k(\boldsymbol{x}) \tag{3}$$

In (2), constraint violations are quantified by penalty functions $c_k(\boldsymbol{x})$, $k = 1, \ldots, n_g + n_h$, with β_k being the penalty coefficients.

A specific formulation of the merit function in (2) or (3) depends on the particular antenna, the imposed design constraints, as well as the assumed performance figures. Here, we focus on ultra-wideband (UWB) antennas, with the design objective defined as minimization of the maximum reflection level within the frequency range from 3.1 GHz to 10.6 GHz, denoted as F. For this task, we have

$$U(\boldsymbol{x}) = \max\{f \in F : |S_{11}(\boldsymbol{x}, f)|\} \tag{4}$$

B. Trust-Region-Based Gradient Search

The proposed optimization routine utilizes a standard trust-region (TR)-based gradient search algorithm [55] as an underlying search engine. The conventional TR gradient-based algorithm iteratively renders the consecutive approximations $\boldsymbol{x}^{(i+1)}$, $i = 0, 1, \ldots$, to the optimum design $\boldsymbol{x}^*$ as

$$\boldsymbol{x}^{(i+1)} = \arg\min_{x_k;\, -d^{(i)}{}_k \le x_k - x^{(i)}{}_k \le d^{(i)}{}_k;\, k=1,\ldots,n} U(\boldsymbol{L}_S^{(i)}(\boldsymbol{x})) \tag{5}$$

where n stands for a number of design variables, and $d^{(i)}$, $k = 1, \ldots, n$, denote the entries of the trust region size vector $\boldsymbol{d}^{(i)}$. Whereas $\boldsymbol{L}_S^{(i)}(\boldsymbol{x}) = S_{11}(\boldsymbol{x}^{(i)}, f) + \boldsymbol{G}_S(\boldsymbol{x}^{(i)}, f)\cdot(\boldsymbol{x} - \boldsymbol{x}^{(i)})$ refers to a linear approximation of S_{11} at the current iteration $\boldsymbol{x}^{(i)}$, and $\boldsymbol{G}_S(\boldsymbol{x}^{(i)}, f)$ represents the gradient of S_{11} at $\boldsymbol{x}^{(i)}$ and frequency f. $\boldsymbol{L}_S^{(i)}(\boldsymbol{x})$ has to be formulated taking into account the particular definition of the objective function, i.e., it is to be a linear approximation of relevant antenna response. Throughout the optimization run, the size vector $\boldsymbol{d}^{(i)}$ is being modified in accordance with the standard rules based on the gain ratio $\rho = [U(S_{11}(\boldsymbol{x}^{(i+1)})) - U(S_{11}(\boldsymbol{x}^{(i)}))]/[U(\boldsymbol{L}_S^{(i)}(\boldsymbol{x}^{(i+1)})) - U(\boldsymbol{L}_S^{(i)}(\boldsymbol{x}^{(i)}))]$ [55]. If $\rho > 0$, the current iteration is considered successful, and the candidate design rendered by (5) is accepted.

The gradient $\boldsymbol{G}_S$ is usually estimated through finite differentiation (FD) at the computational cost of n additional full-wave EM simulations per algorithm iteration. If an iteration turns out to be unsuccessful (i.e., $\rho < 0$), a new search for a candidate design is carried out by solving (5), yet, with a reduced size vector $\boldsymbol{d}^{(i)}$ [55].

C. Expedited Trust-Region Algorithm with Sparse Sensitivity Update Management

In [45], a technique for accelerating the conventional TR algorithm has been proposed, in which the fluctuations of the antenna response sensitivities have been monitored to enable detection of the antenna parameters featuring stable gradient behaviour. For these design variables, costly FD-based sensitivity updates were suppressed for a predefined number of iterations. This allowed for a significant reduction of the overall optimization cost with only slight deterioration of the design quality as compared to the standard TR algorithm with full-FD gradient updates. This section briefly recalls the technique of [45] as one of the major components of the optimization framework presented in this work.

The entries of the gradient vector of the antenna reflection characteristics $\boldsymbol{G}_S = [G_1 \ldots G_n]^T$ are the sensitivities $G_k = \partial S_{11}/\partial x_k$ w.r.t the k-th parameter, $k = 1, \ldots, n$. Also, $G_k^{(i)}(f)$ and $G_k^{(i-1)}(f)$ are the k-th entries of $\boldsymbol{G}_S$ in ith and (i–1)th iteration, respectively (with the dependence of the gradient components on the frequency f shown explicitly). In [45], a metric for quantifying the difference between $G_k^{(i)}$ and $G_k^{(i-1)}$ in two subsequent iterations has been defined as

$$\delta_k^{(i)} = \underset{f\in F}{mean}\left(2\cdot\left(\left|G_k^{(i)}(f)\right| - \left|G_k^{(i-1)}(f)\right|\right) / \left(\left|G_k^{(i)}(f)\right| + \left|G_k^{(i-1)}(f)\right|\right)\right) \tag{6}$$

In each iteration, the parameters for which the gradient change factors assumed the minimum and the maximum value are pinpointed, i.e., $\delta_{\min}^{(i)} = \min\{k = 1, \ldots, n : \delta_k^{(i)}\}$ and $\delta_{\max}^{(i)} = \max\{k = 1, \ldots, n : \delta_k^{(i)}\}$. The response sensitivities for the parameters featuring the smallest values of $\delta_k^{(i)}$ are not updated through FD (for the maximum allowable number of iterations $N_{\max}$), whereas for the variables of the highest $\delta_k^{(i)}$ the FD update is mandatory. For each parameter, the actual number $N_k^{(i+1)}$ of upcoming iterations without FD is assessed through the following affine conversion function

$$N_k^{(i+1)} = \left[\!\left[N_{\max} + A^{(i)}(\delta_k^{(i)} - \delta_{\min}^{(i)}) \right]\!\right] \quad (7)$$

In (7), the slope factor $A^{(i)} = (N_{\max} - N_{\min})/(\delta_{\min}^{(i)} - \delta_{\max}^{(i)})$, whereas $[[.]]$ is the nearest integer function. $N_{\min}$ and $N_{\max}$ are the algorithm control parameters: the minimum and the maximum number of iterations for which FD is to be omitted, respectively.

D. Convergence-Based Adjustment of EM-Simulations Fidelity

In this work, the TR algorithm is expedited by employing the sparse sensitivity updates scheme of Section II.*C*, along with multi-fidelity EM simulations with convergence-based model management scheme [53], briefly outlined below.

Appropriate assessment of the range of the admissible discretization levels: from the coarsest one up to that corresponding to the high-fidelity model, can be performed by visual examination of the simulated antenna responses. The discretization levels are parameterized using LPW (lines per wavelength), which governs the mesh density in CST Microwave Studio, utilized in this work to evaluate the antenna structures. Let $L_{\min}$ be the lowest value usable in practice, and $L_{\max}$ be the highest one, ensuring an adequate rendition of the characteristics of the antenna at hand. In [53], the optimization process has been expedited by using models from the range $L_{\min} \le L \le L_{\max}$, where L refers to model fidelity. The model management scheme has been based on its convergence status. The optimization process has been initiated with $L = L_{\min}$. As the algorithm converges, the model discretization level is gradually increased, to assume the highest value $L_{\max}$ in the concluding phase, to ensure reliability.

The algorithm is terminated if either $\|\boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}\| < \varepsilon_x$ (convergence in argument), $\|\boldsymbol{d}^{(i)}\| < \varepsilon_x$ (TR region shrinking), or $|U_P(\boldsymbol{x}^{(i+1)}) - U_P(\boldsymbol{x}^{(i)})| < \varepsilon_U$ (convergence in the objective function value). The following auxiliary variable is defined

$$Q^{(i)}(\varepsilon_x, \varepsilon_U) = \max\left\{\varepsilon_x / \|\boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}\|, \varepsilon_U / |U_P(\boldsymbol{x}^{(i+1)}) - U_P(\boldsymbol{x}^{(i)})|\right\} \quad (8)$$

$Q^{(i)}$ is associated with the aforementioned thresholds, and it is used to determine the current discretization level $L^{(i)}$ with the use of the following (monotonic) function

$$L^{(i+1)} = \begin{cases} L_{\min} & \text{if } Q^{(i)}(\varepsilon_x, \varepsilon_U) \le M \\ \max\left\{L^{(i)}, L_{\min} + (L_{\max} - L_{\min})\left[Q^{(i)}(\varepsilon_x, \varepsilon_U) - M\right]^{\frac{1}{\alpha}}\right\} & \end{cases} \quad (9)$$

For reliability, the discretization level is switched to the highest one in the final stage of the optimization run. This is because using solely (9) does not ensure that $L^{(i+1)}$ eventually assumes $L_{\max}$. More specifically, when the optimization process is close to convergence, the condition $L^{(i+1)} = L_{\max}$ is enforced when needed (i.e., if $L^{(i)} < L_{\max}$ upon convergence) with simultaneous increase of the trust region size by a factor M_d (to increase room for possible further design improvement after setting $L^{(i+1)} = L_{\max}$) [53].

To further reduce the optimization costs, the antenna response gradients are estimated through FD using lower value of the discretization parameter, referred to as L_{FD}, instead of the current one $L^{(i)}$ (utilized for the evaluation of the antenna response), with $L_{FD} = \max\{L_{\min}, \lambda L^{(i)}\}$, where $0 \le \lambda \le 1$ is the control parameter. The usage of L_{FD} rather than $L^{(i)}$ allows for achieving sufficient accuracy of the system gradients at a reduced cost.

E. Proposed Optimization Framework

The proposed optimization framework is an expedited version of the TR algorithm recollected in Section II.*A*. In our approach, the following two acceleration mechanisms are employed: sparse sensitivity updating scheme of Section II.*C*, and convergence-based model management routine using multi-fidelity simulations delineated in Section II.*D*. The control parameters include the control parameters of the aforementioned two component procedures, as well as the termination criteria for the TR algorithm. The control parameters pertaining to the multi-fidelity EM simulations are briefly described in Section II.*D* (a more thorough discussion on their adjustment is provided in [53]). Whereas the control parameters pertaining to the sparse sensitivity updates procedure are only the minimum and the maximum numbers of algorithm iterations for which FD is not to be carried out, $N_{\min}$ and $N_{\max}$, respectively (here, we set $N_{\min} = 1$ and $N_{\max} = 5$). Whereas the set of the control parameters for multi-fidelity model management scheme comprises six parameters. The first two parameters, $L_{\min}$ and $L_{\max}$, are to be established by the user prior to optimization through grid convergence analysis. Here, $L_{\max}$ refers to the discretization level ensuring satisfactory accuracy, and $L_{\min}$ denotes the discretization level of the coarsest usable model. The termination thresholds of the entire optimization process are set by the user and have to reflect the required resolution level.

In the initial two iterations of the proposed procedure, the entire sensitivity matrix $\boldsymbol{G}_S(\boldsymbol{x})$ is estimated using finite differentiation. Whereas in the subsequent iterations, $\boldsymbol{G}_S(\boldsymbol{x})$ is a merger of the selected parameter sensitivities evaluated using FD, as well as those for which FD-update has been omitted and their prior values form the preceding iterations have been retained.

III. Verification Case Studies

This section provides numerical verification of the optimization framework delineated in Section II. The proposed optimization algorithm (Algorithm IV) has been benchmarked against three gradient-based local search procedures: (i) the reference trust-region gradient search routine (Algorithm I [55]), (ii) the accelerated algorithm with sparse sensitivity updates based on gradient change tracking using high-fidelity EM simulations (Algorithm II [45]), and (iii) the multi-fidelity TR procedure with full-FD sensitivity update (Algorithm III [53])). The benchmark procedures have been recalled in Sections II.*B*, II.*C* and II.*D*, respectively. Since all of the benchmark antenna structures have been already validated (both in the source papers [56]-[59] and in our previous work, e.g., [45], [47]), the experimental validation of the optimized designs has not been provided, as being immaterial to the scope of the paper.

A. Results

The performance of all algorithms (both benchmark Algorithms I through III, and the proposed Algorithm IV) is assessed with respect to the computational overhead they incur, as well as their reliability (among others, the ability to render quality optimal designs meeting the design specifications).

The verification examples include four ultra-wideband antennas that have been optimized to ensure best in-band matching (see Fig. 1). Table I gathers information about the geometry parameters (both designable and fixed ones), as well as the details on the substrate the structure has been implemented on. The computational models of all antennas are implemented in CST Microwave Studio and simulated using its time-domain solver. All models incorporate the SMA connectors. The antennas are to operate within the frequency range 3.1 GHz to 10.6 GHz (UWB frequency range). The design task is to minimize the maximum in-band reflection within the said band, and the merit function is formulated as follows $U(\boldsymbol{x}) = \max\{3.1 \text{ GHz} \le f \le 10.6 \text{ GHz} : |S_{11}(\boldsymbol{x},f)|\}$.

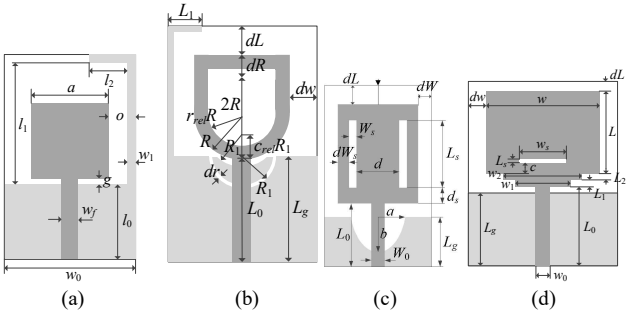

Fig. 1. Antenna structures used as verification case studies: (a) Antenna I [56], (b) Antenna II [57], (c) Antenna III [58], (d) Antenna IV [59]. Light-grey shade indicates ground plane.

TABLE I
ANTENNA STRUCTURES USED AS VERIFICATION CASE STUDIES

Antenna	Substrate	Designable Parameters [mm]	Other
I [56]	RF-35 (ε_r = 3.5, h = 0.762 mm)	$\boldsymbol{x} = [l_0\ g\ a\ l_1\ l_2\ w_1\ o]^T$	$w_0 = 2o + a$, w_f = 1.7 mm
II [57]	RF-35 (ε_r = 3.5, h = 0.762 mm)	$\boldsymbol{x} = [L_0\ dR\ R\ r_{rel}\ dL\ dw\ Lg\ L_1\ R_1\ dr\ c_{rel}]^T$	w_0 = 1.7 mm
III [58]	FR4 (ε_r = 4.3, h = 1.55 mm)	$\boldsymbol{x} = [L_g\ L_0\ L_s\ W_s\ d\ dL\ d_s\ dW_s\ dW\ a\ b]^T$	W_0 = 3.0 mm
IV [59]	RO4350 (ε_r = 3.48, h = 0.762 mm)	$\boldsymbol{x} = [L_0\ L_1\ L_2\ L\ dL\ L_g\ w_1\ w_2\ w\ dw\ L_s\ w_s\ c]^T$	w_0 = 1.7 mm

TABLE II
DISCRETIZATION DENSITY RANGES FOR ANTENNAS [53]

Antenna	Lowest-fidelity model		High-fidelity model	
	L_{min}	Simulation time [s]	L_{max}	Simulation time [s]
I	10	42	21	150
II	11	41	24	424
III	10	46	20	265
IV	10	37	25	97

Table II gathers the lowest and the highest discretization levels for all antennas. The time evaluation ratio (the ratio between the simulation times of the high-resolution model and the coarsest one) is problem specific: below three in the case of Antenna IV, and around ten for Antenna II. Thus, potential computational speedup due to the involvement of multi-fidelity simulations may be sizeable.

The considered optimization tasks are, in general, multimodal, and the procedures employed to solve them are local ones. So, the search process launched from different initial designs typically yields different local optimum solutions. Therefore, statistical verification of the considered algorithms has been conducted as follows: each procedure has been re-run ten times starting from random initial designs in order to assess both the reliability and efficiency of the optimization process. The considered algorithms are compared with respect to the following aspects: computational expenditures (averaged over the set of ten initial designs), design quality (quantified as the average value of the merit function), as well as the result repeatability (measured as standard deviation of the objective function across the entire set). Due to the problem multimodality, the said standard deviation does not assume zero value even for the reference TR algorithm with high-fidelity simulations. Thus, the standard deviation values obtained for each algorithm should be considered with regard to the reference technique instead of a zero value.

The control parameters of Algorithm II and IV (with sparse sensitivity updates) have been set to: N_{min} = 1 and N_{max} = 5 (cf. Section II.*C*). Moreover, for Algorithm III and IV (employing multi-fidelity model management scheme) (see Section II.*D*) we have: M = 10^{-2}, α = 3, λ = 2/3, and M_d = 10. The termination thresholds have been set to: $\varepsilon_x = \varepsilon_U = 10^{-3}$ (see Section II.*E*).

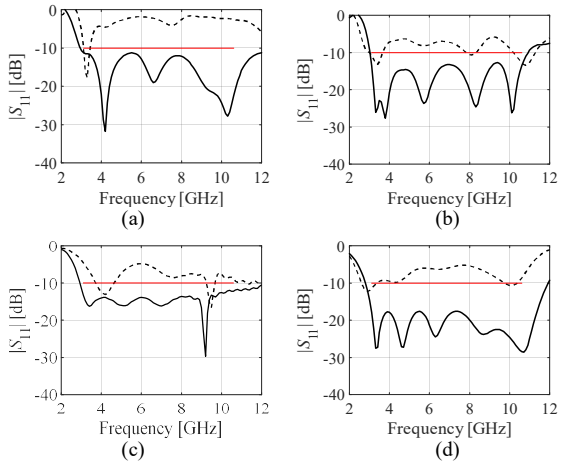

Fig. 2. Antenna input characteristics for the selected runs of the proposed algorithm: (a) Antenna I, (b) Antenna II, (c) Antenna III, (d) Antenna IV (obtained for the maximum discretization level for the respective antenna structures). Horizontal lines indicate design specifications. Initial and optimal designs are marked with dashed (- - -), and solid lines (—), respectively.

TABLE III
OPTIMIZATION RESULTS FOR ANTENNAS I THROUGH IV

Antenna	Algorithm		Performance figure: Cost[1]	Cost savings[2]	Max $\lvert S_{11}\rvert$[3]	Δmax $\lvert S_{11}\rvert$[4]	Std max $\lvert S_{11}\rvert$[5]
I	I	Conventional TR search	97.6	–	–11.9	–	0.4
	II	Accelerated TR search [45]	44.3	55 %	–11.5	0.4	0.8
	III	Multi-fidelity [53]	48.2	51 %	–11.2	0.7	0.7
	IV	This work	24.7	75 %	–10.9	1.0	0.9
II	I	Conventional TR search	111.2	–	–14.9	–	0.6
	II	Accelerated TR search [45]	66.3	40%	–14.7	0.2	0.8
	III	Multi-fidelity [53]	25.8	77%	–13.8	1.1	1.0
	IV	This work	23.7	79 %	–13.9	1.0	1.0
III	I	Conventional TR search	111.0	–	–13.9	–	1.0
	II	Accelerated TR search [45]	68.7	57 %	–13.5	0.4	1.1
	III	Multi-fidelity [53]	42.3	62 %	–11.3	2.6	1.0
	IV	This work	26.1	76 %	–12.3	1.6	1.7
IV	I	Conventional TR search	139.7	–	–17.6	–	1.2
	II	Accelerated TR search [45]	41.1	71%	–13.5	4.1	4.1
	III	Multi-fidelity [53]	97.2	31%	–17.0	0.6	2.1
	IV	This work	35.2	75 %	–16.3	1.4	1.8

[1] Number of equivalent high-fidelity EM simulations averaged over 10 algorithm runs.
[2] Relative computational savings in percent w.r.t. the reference algorithm.
[3] Objective function value (max. in-band reflection in dB), averaged over 10 algorithm runs.
[4] Degradation of max$|S_{11}|$ w.r.t. the TR algorithm in dB, averaged over 10 algorithm runs.
[5] Standard deviation of max$|S_{11}|$ in dB across the set of 10 algorithm runs.

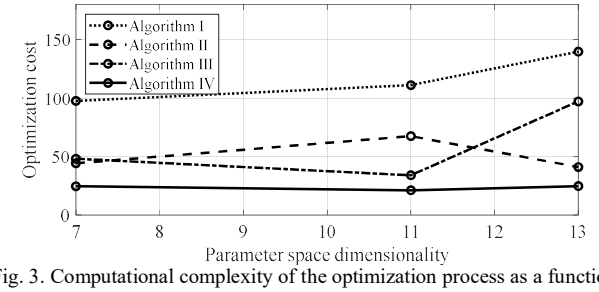

Fig. 3. Computational complexity of the optimization process as a function of parameter space dimensionality (of all benchmark antennas) for the proposed algorithm (Algorithm IV) and the reference procedures (Algorithms I through III). The parameter space dimensionalities (i.e., the numbers of the design variables) are: seven (Antenna I), eleven (Antennas II and III), and thirteen (Antenna IV). The optimization cost expressed in the equivalent number of high-fidelity model evaluations. The costs for Antenna II and III are averaged as the number of parameters is eleven in both cases.

B. Discussion

The performance statistics for the proposed and the benchmark algorithms are provided in Table III. Figure 2 shows the reflection responses at the respective algorithm runs: for the initial and optimal designs of Antennas I through IV. The data of Table III include: the cost of the optimization process (equivalent number of high-fidelity EM simulations of the respective antenna), the cost savings w.r.t. the reference TR algorithm using high-fidelity simulations (Algorithm I), along with: the average objective function value across the set of ten algorithm runs (i.e., the maximum antenna reflection in UWB frequency range), its deterioration with respect to Algorithm I, and the standard deviation of the merit function (indicator of the results repeatability, to be compared to that of Algorithm I, as discussed above). The optimization cost of Algorithm II and IV algorithms which exploit multi-fidelity simulations is computed by taking into account the time evaluation ratios between the low- and the high-fidelity models utilized throughout the optimization run in the following manner. In each iteration, the equivalent number of high-fidelity EM simulations is assessed as $N_c^{(i)}/m^{(i)}$, where $N_c^{(i)}$ is the actual number of simulations executed at a current resolution $L^{(i)}$, whereas $m^{(i)}$ denotes the ratio of the simulation time between the high- and (current) low-fidelity model. The overall optimization cost constitutes a sum of these numbers throughout all iterations.

At the same time, the obtained savings are achieved without a significant loss of design quality. The deterioration of the average objective function value in comparison to Algorithm I ranges from around 1.0 dB for Antenna I and II, and it is around 1.5 dB for Antennas III and IV (i.e., around 1.3 dB on average). The result repeatability, measured as the standard deviation of the objective function value across the set of ten optimal designs, is somewhat worse than the reference Algorithm I (slightly over 0.5 dB across the benchmark antenna set). Moreover, it is comparable to that of Algorithm II and III on average.

C. Computational Complexity

This section investigates and compares computational complexity of the considered algorithms. The relevant dependencies, i.e., the optimization costs versus parameter space dimensionality, have been shown in Fig. 3. Despite relatively limited data, it can be observed that the reference trust-region procedure (Algorithm I), and the multi-fidelity procedure (Algorithm III), exhibit the complexity typical for gradient-based algorithms involving first-order sensitivities, i.e., the dependence between the number of antenna parameters and is slightly higher than linear. The cost of procedures involving sparse sensitivity updates (here, gradient tracking), i.e., Algorithms II and IV, is only weakly dependent on the parameter space dimensionality. At the same time, Fig. 3, clearly indicates that the procedure proposed in this work capitalizes on both mechanisms (incorporation of variable-resolution simulations and sparse gradient updates) to the fullest extent. It features by far the best computational efficiency (as reported in Table III, the savings are as high as 80 percent as compared to Algorithm I), with the cost being almost independent of the number of antenna parameters within the considered range (7 to 13 variables). In absolute terms, the average cost is only about 27 high-fidelity EM analyses of the respective antenna structures, which makes the presented procedure one of the fastest algorithms for direct (i.e., not surrogate-based) optimization of antenna structures ever presented in the literature. Its important advantage is that it is fully automated. This is in contrast to surrogate-assisted methods, especially physics-based ones, which require considerable experience and problem-specific knowledge to properly setup and handle the algorithm.

IV. Conclusion

This work introduced a novel algorithm for expedited trust-region gradient-based design optimization of antenna structures. Our methodology involves two separate acceleration mechanisms: sparse sensitivity updates and variable-resolution simulations. The former allows for omitting a certain number of the CPU-heavy finite-differentiation-based evaluations of antenna response gradients. The decision about abstaining from some updates is based on the assessment of the variability of the antenna response sensitivities pattern throughout the optimization run. The second acceleration mechanism, i.e., variable-resolution model management scheme, controls a continuous adaptation of the model discretization level taking into account the algorithm convergence status. The procedure is initiated with the lowest acceptable fidelity, and the resolution is gradually increased towards its conclusion, with the high-fidelity model only utilized as the algorithm is close to convergence.

The employed acceleration mechanisms allow for achieving dramatic computational savings of around eighty percent. Further, our framework outperforms the recently proposed accelerated optimization routines of around fifty percent in terms of computational efficiency. Moreover, the computational cost of the proposed procedure is only weakly dependent on the parameter space dimensionality. This is not the case for gradient-based algorithms involving first-order sensitivities, such as the reference trust region algorithm, whose dependence between the number of antenna parameters is slightly stronger than linear. The proposed algorithm is capable of yielding high-quality designs at the cost of around two dozens of full-wave antenna analyses. This makes it one of the most efficient algorithms for direct (i.e., not surrogate-based) optimization of antenna structures ever presented in the literature so far. Finally, the proposed framework is fully automated, which is in contrast to surrogate-assisted methods, normally requiring problem-specific knowledge to be properly set up.

References

[1] M. Ciydem and E. A. Miran, "Dual-polarization wideband sub-6 GHz suspended patch antenna for 5G base station," *IEEE Ant. Wireless Prop. Lett.*, vol. 19, no. 7, pp. 1142-1146, 2020.

[2] D. Serghiou, M. Khalily, V. Singh, A. Araghi, and R. Tafazolli, "Sub-6 GHz dual-band 8 × 8 MIMO antenna for 5G smartphones," *IEEE Ant. Wireless Prop. Lett.*, vol. no. 9, pp. 1546-1550, 2020.

[3] I. M. Danjuma, M. O. Akinsolu, C. H. See, R. A. Abd-Alhameed, and B. Liu, "Design and optimization of a slotted monopole antenna for ultra-wide band body centric imaging applications," *IEEE J. Electromagn., RF and Microw. Medicine Biol.*, vol. 4, no. 2, pp. 140-147, 2020.

[4] T. Houret, L. Lizzi, F. Ferrero, C. Danchesi, and S. Boudaud, "DTC-enabled frequency-tunable inverted-F antenna for IoT applications," *IEEE Ant. Wireless Prop. Lett.*, vol. 19, no. 2, pp. 307-311, 2020.

[5] A. Ruaro, J. Thaysen, and K. B. Jakobsen, "Wearable shell antenna for 2.4 GHz hearing instruments," *IEEE Trans. Ant. Propag.*, vol. 64, no. 6, pp. 2127-2135, 2016.

[6] D. Nikolayev, W. Joseph, A. Skrivervik, M. Zhadobov, L. Martens, and R. Sauleau, "Dielectric-loaded conformal microstrip antennas for versatile in-body applications," *IEEE Ant. Wireless Prop. Lett.*, vol. 18, no. 12, pp. 2686-2690, 2019.

[7] Y. Xie, F. Chen, and J. Qian, "Design of integrated duplexing and multi-band filtering slot antennas," *IEEE Access*, vol. 8, pp. 126119-126126, 2020.

[8] Y. Zheng, G. A. E. Vandenbosch, and S. Yan, "Low-profile broadband antenna with pattern diversity," *IEEE Ant. Wireless Prop. Lett.*, vol. 19, no. 7, pp. 1231-1235, 2020.

[9] G. Bogdan, P. Bajurko, and Y. Yashchyshyn, "Time-modulated antenna array with dual-circular polarization," *IEEE Ant. Wireless Prop. Lett.*, vol. 19, no. 11, pp. 1872-1875, 2020.

[10] S. Fu, Z. Cao, X. Quan, and C. Xu, "A broadband dual-polarized notched-band antenna for 2/3/4/5G base station," *IEEE Ant. Wireless Prop. Lett.*, vol. 19, no. 1, pp. 69-73, 2020.

[11] V. Gongal-Reddy, S. Zhang, C. Zhang, and Q. Zhang, "Parallel computational approach to gradient based EM optimization of passive microwave circuits," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 1, pp. 44-59, 2016.

[12] X. Peng, L. Kong, X. Sun, and H. Lyu, "Design and analysis of optical receiving antenna for LED visible light communication based on Taguchi method," *IEEE Access*, vol. 7, pp. 186364-186377, 2019.

[13] H. M. Torun and M. Swaminathan, "High-dimensional global optimization method for high-frequency electronic design," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 6, pp. 2128-2142, 2019.

[14] C. Jarufe, R. Rodriguez, V. Tapia, P. Astudillo, D. Monasterio, R. Molina, F.P. Mena, N. Reyes, and L. Bronfman, "Optimized corrugated tapered slot antenna for mm-wave applications," *IEEE Trans. Ant. Propag.*, vol. 66, no. 3, pp. 1227-1235, 2018.

[15] A. Dadgarpour, G. Dadashzadeh, M. Naser-Moghadasi, F. Jolani. and B.S. Virdee, "PSO/FDTD optimization technique for designing UWB in-phase power divider for linear array antenna application," *IEEE Ant. Wireless Prop. Lett.*, vol. 9, pp. 424-427, 2010.

[16] X. Li, S. Ma, and G. Yang, "Synthesis of difference patterns for monopulse antennas by an improved cuckoo search algorithm," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, pp. 141-144, 2017.

[17] B. Liu, H. Aliakbarian, Z. Ma, G. A. E. Vandenbosch, G. Gielen, and P. Excell, "An efficient method for antenna design optimization based on evolutionary computation and machine learning techniques," *IEEE Trans. Ant. Propag.*, vol. 62, no. 1, pp. 7-18, 2014.

[18] S. K. Goudos, K. Siakavara, E. Vafiadis, and J. N. Sahalos, "Pareto optimal Yagi-Uda antenna design using multi-objective differential evolution," Progr. Electromag. Research, vol. 105, pp. 231-251, 2010.

[19] A. Lalbakhsh, M. U. Afzal, and K. P. Esselle, "Multiobjective particle swarm optimization to design a time-delay equalizer metasurface for an electromagnetic band-gap resonator antenna," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, pp. 912-915, 2017.

[20] S. K. Goudos, K.A. Gotsis, K. Siakavara, E.E. Vafiadis, and J. N. Sahalos, "A multi-objective approach to subarrayed linear antenna arrays design based on memetic differential evolution," *IEEE Trans. Ant. Propag.*, vol. 61, no. 6, pp. 3042-3052, 2013.

[21] A. Kouassi, N. Nguyen-Trong, T. Kaufmann, S. Lalléchère, P. Bonnet, and C. Fumeaux, "Reliability-Aware Optimization of a Wideband Antenna," *IEEE Trans. Ant. Prop.*, vol. 64, no. 2, pp. 450-460, 2016.

[22] A. Pietrenko-Dabrowska, S. Koziel, and M. Al-Hasan, „Expedited yield optimization of narrow- and multi-band antennas using performance-driven surrogates," *IEEE Access*, pp. 143104-143113, 2020.

[23] S. Koziel, S. Ogurtsov, Q.S. Cheng, and J.W. Bandler, "Rapid EM-based microwave design optimization exploiting shape-preserving response prediction and adjoint sensitivities," IET Microwaves, Ant. Prop., vol., 8, no. 10, pp. 775-781, 2014.

[24] J. A. Easum, J. Nagar, P. L. Werner, and D. H. Werner, "Efficient multi-objective antenna optimization with tolerance analysis through the use of surrogate models," *IEEE Trans. Ant. Prop.*, vol. 66, no. 12, pp. 6706-6715, 2018.

[25] J. C. Cervantes-González, J. E. Rayas-Sánchez, C. A. López, J. R. Camacho-Pérez, Z. Brito-Brito, and J. L. Chávez-Hurtado, "Space mapping optimization of handset antennas considering EM effects of mobile phone components and human body," *Int. J. RF Microwave CAE*, vol. 26, no. 2, pp. 121-128, 2016.

[26] J. van der Herten, I. Couckuyt, D. Deschrijver, and T. Dhaene, "Adaptive classification under computational budget constraints using sequential data gathering," *Adv. Eng. Softw.*, vol. 99, pp. 137-146, 2016.

[27] S. Marelli and B. Sudret, "UQLab: A framework for uncertainty quantification in MATLAB," *Proc. 2nd Int. Conf. Vulnerability Risk Anal. Manage. (ICVRAM)*, pp. 2554-2563, 2014.

[28] A.K.S.O. Hassan, A.S. Etman, and E.A. Soliman, "Optimization of a novel nano antenna with two radiation modes using kriging surrogate models," *IEEE Photonic J.*, vol. 10, no. 4, art. no. 4800807, 2018.

[29] P. Barmuta, F. Ferranti, G.P. Gibiino, A. Lewandowski, and D. M. M. P. Schreurs, "Compact behavioral models of nonlinear active devices using response surface methodology," *IEEE Trans. Microwave Theory and Tech.,* vol. 63, no. 1, pp. 56-64, 2015.

[30] J. Cai, J. King, C. Yu, J. Liu, and L. Sun, "Support vector regression-based behavioral modeling technique for RF power transistors," *IEEE Microwave and Wireless Comp. Lett.,* vol. 28, no. 5, pp. 428-430, 2018.

[31] A. Petrocchi, A. Kaintura, G. Avolio, D. Spina, T. Dhaene, A. Raffo, and D. M. P. P. Schreurs, „Measurement uncertainty propagation in transistor model parameters via polynomial chaos expansion," *IEEE Microwave Wireless Comp. Lett.*, vol. 27, no. 6, pp. 572-574, 2017.

[32] A. Rawat, R. N. Yadav, and S. C. Shrivastava, "Neural network applications in smart antenna arrays: a review," *AEU – Int. J. Elec. Comm.*, vol. 66, no. 11, pp. 903-912, 2012.

[33] J. Tak, A. Kantemur, Y. Sharma, and H. Xin, "A 3-D-printed W-band slotted waveguide array antenna optimized using machine learning," *IEEE Ant. Wireless Prop. Lett.*, vol. 17, no. 11, pp. 2008-2012, 2018

[34] A. C. Yücel, H. Bağcı, and E. Michielssen, "An ME-PC enhanced HDMR method for efficient statistical analysis of multiconductor transmission line networks," *IEEE Trans. Comp., Packag. Manuf. Techn.*, vol. 5, no. 5, pp. 685-696, 2015.

[35] D. Spina, F. Ferranti, G. Antonini, T. Dhaene, and L. Knockaert, ''Efficient variability analysis of electromagnetic systems via polynomial chaos and model order reduction,'' *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 4, no. 6, pp. 1038-1051, 2014

[36] H. Li and G. Liu, "An improved analysis for support recovery with orthogonal matching pursuit under general perturbations," *IEEE Access*, vol. 6, pp. 18856-18867, 2018.

[37] F. Feng, J. Zhang, W. Zhang, Z. Zhao, J. Jin, and Q. J. Zhang, „Coarse- and fine-mesh space mapping for EM optimization incorporating mesh deformation," *IEEE Microwave Wireless Comp. Lett.*, vol. 29, no. 8, pp. 510-512, 2019.

[38] S. Koziel and S.D. Unnsteinsson "Expedited design closure of antennas by means of trust-region-based adaptive response scaling," *IEEE Antennas Wireless Prop. Lett.*, vol. 17, no. 6, pp. 1099-1103, 2018.

[39] S. Koziel, "Fast simulation-driven antenna design using response-feature surrogates," *Int. J. RF & Microwave CAE*, vol. 25, no. 5, pp. 394-402, 2015.

[40] A.A. Al-Azza, A.A. Al-Jodah, and F.J. Harackiewicz, "Spider monkey optimization: A novel technique for antenna optimization," *IEEE Antennas Wireless Prop. Lett.*, vol. 15, pp. 1016-1019, 2016, doi: 10.1109/LAWP.2015.2490103.

[41] X. Li and K. M. Luk, "The Grey Wolf optimizer and its applications in electromagnetics," *IEEE Trans. Ant. Prop.*, vol. 68, no. 3, pp. 2186-2197, 2020.

[42] Q. Wu, H. Wang, and W. Hong, "Multistage collaborative machine learning and its application to antenna modeling and optimization," *IEEE Trans. Ant. Prop.*, vol. 68, no. 5, pp. 3397-3409, 2020.

[43] Y. Sharma, H. H. Zhang, and H. Xin, "Machine learning techniques for optimizing design of double T-shaped monopole antenna," *IEEE Trans. Ant. Prop.*, vol. 68, no. 7, pp. 5658-5663, 2020.

[44] S. Koziel, "Improved trust-region gradient-search algorithm for accelerated optimization of wideband antenna input characteristics," *Int. J. RF Microwave CAE*, vol. 29, no. 4, e21567, 2019.

[45] A. Pietrenko-Dabrowska and S. Koziel, "Expedited antenna optimization with numerical derivatives and gradient change tracking," *Eng. Comp.*, vol. 37, no. 4, pp. 1179-1193, 2019.

[46] A. Pietrenko-Dabrowska and S. Koziel, "Numerically efficient algorithm for compact microwave device optimization with flexible sensitivity updating scheme," *Int. J. RF Microwave CAE*, vol. 29, no. 7, e21714, 2019.

[47] S. Koziel and A. Pietrenko-Dabrowska, "Reduced-cost design closure of antennas by means of gradient search with restricted sensitivity updates," *Metrology Meas. Syst.*, vol. 26, no. 4, pp. 595-605, 2019.

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

[49] F. Feng, J. Zhang, W. Zhang, Z. Zhao, J. Jin, and Q.J. Zhang, „Coarse- and fine-mesh space mapping for EM optimization incorporating mesh deformation," *IEEE Microwave Wireless Comp. Lett.*, vol. 29, no. 8, pp. 510-512, 2019.

[50] L. Leifsson and S. Koziel, *Simulation-driven aerodynamic design using variable-fidelity models*, Imperial College Press, London, 2015.

[51] S. Koziel and L. Leifsson, *Simulation-driven design by knowledge-based response correction techniques*, Springer, New York, 2016.

[52] S. Koziel and S. Ogurtsov, "Model management for cost-efficient surrogate-based optimization of antennas using variable-fidelity electromagnetic simulations," *IET Microwaves Ant. Prop.*, vol. 6, no. 15, pp. 1643-1650, 2012.

[53] S. Koziel and A. Pietrenko-Dabrowska, "Accelerated gradient-based optimization of antenna structures using multi-fidelity simulations and convergence-based model management scheme," *IEEE Trans. Ant. Prop.*, *Early view*, doi: 10.1109/TAP.2020.2966051.

[54] C. Zhang, X. Fu, X. Chen, S. Peng, and X. Min, "Synthesis of uniformly excited sparse rectangular planar array for sidelobe suppression using multi-objective optimization algorithm," *J. Eng.*, vol. 2019, no. 19, pp. 6278-6281, 2019.

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

[56] S. Koziel and A. Bekasiewicz, "Low-cost multi-objective optimization of antennas using Pareto front exploration and response features," *Int. Symp. Antennas Prop.*, Fajardo, Puerto Rico, 2016.

[57] M. G. N. Alsath and M. Kanagasabai, "Compact UWB monopole antenna for automotive communications," *IEEE Trans. Ant. Prop.*, vol. 63, no. 9, pp. 4204-4208, 2015.

[58] M. A. Haq and S. Koziel, "Simulation-based optimization for rigorous assessment of ground plane modifications in compact UWB antenna design," *Int. J. RF Microwave CAE*, vol. 28, no. 4, e21204, 2018.

[59] D. R. Suryawanshi and B. A. Singh, "A compact UWB rectangular slotted monopole antenna", *IEEE Int. Conf. Control, Instrumentation, Comm. Comp. Tech.* (ICCICCT), pp. 1130-1136, 2014.