Postprint of: Pietrenko-Dabrowska A. and Koziel S., Computationally-efficient design optimisation of antennas by accelerated gradient search with sensitivity and design change monitoring, in IET Microwaves, Antennas & Propagation, vol. 14, no. 2, pp. 165-170, 5 2 2020
DOI: 10.1049/iet-map.2019.0358

Computationally-Efficient Design Optimization of Antennas by Accelerated Gradient Search with Sensitivity and Design Change Monitoring

Anna Pietrenko-Dabrowska[1] and Slawomir Koziel[1,2]

[1] Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdansk Poland
anna.dabrowska@pg.edu.pl

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

Abstract: Electromagnetic (EM) simulation tools are of primary importance in the design of contemporary antennas. The necessity of accurate performance evaluation of complex structures is a reason why the final tuning of antenna dimensions, aimed at improvement of electrical and field characteristics, needs to be based on EM analysis. Design automation is highly desirable and can be achieved by coupling EM solvers with numerical optimization routines. Unfortunately, its computational overhead may be impractically high for conventional algorithms. This paper proposes an efficient gradient search algorithm with numerical derivatives. The acceleration of the optimization process is obtained by means of the two mechanisms developed to suppress some of finite-differentiation-based updates of the antenna response sensitivities that involve monitoring and quantifying the gradient changes as well as design relocation between the consecutive algorithm iterations. Both methods considerably reduce the need for finite differentiation, leading to significant computational savings. At the same time, excellent reliability and repeatability is maintained, which is demonstrated through statistics over multiple algorithm runs with random initial designs. Our approach is validated using a benchmark set of wideband antennas. The proposed algorithm is competitive to both the reference trust-region algorithm as well as its recently reported accelerated versions.

1. Introduction

Design of contemporary antenna systems is a challenging endeavour. It involves several steps including a conceptual design, development of antenna topology (usually aided by a parameter sweeping), as well as design optimization [1]. The importance of this last stage has been steadily increasing as a result of a growing complexity of antenna structures which can no longer be reliably evaluated using design-ready theoretical models. Instead, computationally expensive full-wave electromagnetic (EM) analysis is required. It is especially critical to account for the phenomena that exhibit non-negligible effects on the system performance such as mutual radiator coupling, feed radiation, the effects of connectors or radomes. The need for using EM simulation for the antennas described by large number of geometry parameters makes the design closure a time consuming process. On the top of this, several performance figures have to be handled at the same time (e.g., impedance matching, radiation patterns, gain, axial ratio, as well as the physical size of the structure). The last issue is particularly problematic for the traditional tuning methods, primarily based on parameter sweeping. Nowadays, rigorous numerical optimization has been gaining popularity as it is capable of simultaneous adjustment of all relevant parameters. A practical problem is the high computational cost of conventional algorithms, which is the case even for local methods, e.g., gradient algorithms [2] or pattern search techniques [3]. Clearly, it is much more pronounced for global procedures, especially population-based metaheuristics (genetic algorithms [4], differential evolution [5], particle swarm optimizers [6]); making them virtually unusable for direct optimization of high-fidelity EM simulation models.

Improvement of computational efficiency of numerical optimization has been investigated over the last years leading to a number of interesting techniques and promising results. In the case of gradient-based algorithms, a considerable speedup can be achieved using the adjoint sensitivity techniques [7], [8]. Unfortunately, the adjoint technology is not widely accessible through commercial solvers [9]. An alternative approach are surrogate-based optimization (SBO) methods [10]. As a matter of fact, SBO is a large and still growing class of techniques whose common feature is a utilisation of a faster representation (referred to as the surrogate model) of the original simulation model in order to accelerate the process of the optimum design identification. The surrogate models fall into one of the two categories: physics-based models (e.g., space mapping [11], response correction methods [12], [13], feature-based optimization [14], adaptive response scaling [15]) or approximation (or data-driven) ones (e.g., kriging [16], Gaussian process regression [17], combinations with machine learning techniques [18]). The principal distinction between these two groups stems from their very foundations. The data-driven models are merely approximations of the sampled simulation data, therefore they are fast and generic. However, they are affected by the curse of generality and are mainly used for local applications such as statistical analysis [19] or for handling low-dimensional parameter spaces [20]. On the other hand, physics-based surrogates are constructed using underlying low-fidelity models which make them more immune to the dimensionality issues. At the same time, they are of limited use for antenna optimization, primarily due to unavailability of such models.

One of the prerequisites for speeding up local search algorithms, both direct EM-based design and variable-fidelity SBO algorithms [11], is a reduction of the number of antenna simulations at any level of fidelity that is relevant to the optimization framework at hand. Recently, some variations of the trust-region (TR) gradient search framework have been proposed with this in mind [21], [22]. The methods aim at reducing the number of EM analyses through omission of the gradient updates under certain conditions. These include small relative changes of the design between the algorithm iterations in particular directions [21], and sufficient alignment of the design relocation direction with the coordinate system vectors [22]. In both cases, the computational speedup has been achieved at the expense of noticeable loss of design quality.

In this work, a novel trust-region-based algorithm with numerical derivatives for computationally-efficient antenna design optimization is proposed. There are two mechanisms introduced to control (or, more specifically, suppress in some cases) the gradient updates using finite differentiation. The first one is the monitoring of the antenna sensitivity change between the algorithm iterations; the second involves analysing the amount of design relocation in relation to the trust region size. A combination of these mechanisms makes it possible to considerably reduce the number of EM analyses with respect to the reference TR algorithm, without compromising the design quality in a significant manner. Furthermore, the proposed approach outperforms the methods of [21] and [22] both in terms of the computational efficiency and reliability.

The main contributions of the work are the following: (i) establishing a computationally-efficient procedure for antenna optimization by adaptation and fusion of previously reported acceleration mechanisms; the procedure yields the designs of quality comparable to those produced by reference TR-region algorithm at less than half of its computational cost, (ii) guiding the optimization process in a novel manner by combining various sensitivity update techniques in an adaptive fashion, and (iii) demonstrating the relevance of the approach for solving real-world antenna design tasks. One of possible applications of the framework is a direct EM-driven design closure of antennas. The algorithm can also be used within variable-fidelity SBO routines for optimizing surrogates constructed from coarse-mesh EM simulations.

2. Acceleration Procedures for Trust-Region Gradient-Based Antenna Optimization

In this section, the formulations of the design closure task as well as the conventional trust-region algorithm are recalled, followed by an exposition of the proposed optimization procedure. Specifically, the two adopted acceleration mechanisms and their integration into the trust-region framework are discussed in detail.

2.1. Formulation of the Design Optimization Task

An essential part of design automation by means of numerical optimization is a properly defined performance measure. Typically, scalar cost functions are utilized to assess the quality of the design. This is the most convenient approach, in which single-objective search routines are employed (e.g., gradient-based algorithms). Multi-objective design [23], handling vector-valued objective functions is out of the scope of this work.

Here, the antenna design closure task is formulated as a nonlinear minimization problem of the form

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

where U stands for the scalar objective function and $\boldsymbol{x}$ denotes a vector of antenna geometry parameters. The actual definition of the objective function depends on the performance figures of interest. Probably the most popular type of optimization problems in antenna design is an improvement of the impedance matching over a frequency range of interest F. In this case, the cost function U can be defined as

$$U(\boldsymbol{x}) = \max_{f \in F} |S_{11}(\boldsymbol{x}, f)| \tag{2}$$

where $|S_{11}(\boldsymbol{x},f)|$ is the reflection coefficient and it depends both on the vector of antenna parameters $\boldsymbol{x}$ and the frequency f. For ultra-wideband (UWB) antennas considered in Section 3, the frequency range F corresponds to the UWB band from 3.1 GHz to 10.6 GHz.

The design optimization task is formulated in a minimax sense as a nonlinear minimization problem of the form (1), whereas the objective function U is defined by (2). The antenna characteristics, for the sake of reliability, are obtained from full-wave EM simulation. One of widely used and reliable (local) routines for solving such tasks is a trust-region gradient search recalled in Section 2.2. This is the starting point of the proposed accelerated algorithm introduced in Section 2.3.

2.2. Reference Optimization Algorithm

The reference method is the conventional local trust-region (TR)-based algorithm (e.g., [24]) that solves the problem (1) by yielding approximations $\boldsymbol{x}^{(i)}$, i = 0, 1, …, to the optimum design $\boldsymbol{x}^*$. The consecutive designs are obtained by optimizing a local linear model $U(\boldsymbol{L}^{(i)})$ of U at $\boldsymbol{x}^{(i)}$

$$\boldsymbol{x}^{(i+1)} = \arg \min_{\boldsymbol{x};\, -\boldsymbol{d}^{(i)} \le \boldsymbol{x}-\boldsymbol{x}^{(i)} \le \boldsymbol{d}^{(i)}} U\left(\boldsymbol{L}^{(i)}\left(\boldsymbol{x}, f\right)\right) \tag{3}$$

where $\boldsymbol{L}^{(i)}(\boldsymbol{x})$ is a first-order Taylor expansion of the adopted cost function at the current iteration point $\boldsymbol{x}^{(i)}$. For the reflection characteristic it takes the form of

$$\boldsymbol{L}^{(i)}(\boldsymbol{x}, f) = S_{11}(\boldsymbol{x}^{(i)}, f) + \boldsymbol{G}_S(\boldsymbol{x}^{(i)}, f) \cdot (\boldsymbol{x} - \boldsymbol{x}^{(i)}) \tag{4}$$

where $\boldsymbol{G}_s$ denotes the gradient vector. In (3), the inequalities $-\boldsymbol{d}^{(i)} \le \boldsymbol{x} - \boldsymbol{x}^{(i)} \le \boldsymbol{d}^{(i)}$ are understood component-wise and they define the interval-type trust region. This allows us to handle antenna parameters of substantially different ranges: ranging from fractions of millimeters for, e.g., spacings between the lines, up to tens of millimeters in the case of, e.g., transmission line lengths. The initial size vector $\boldsymbol{d}^{(0)}$ is proportional to the design space size. During the optimization process, the TR region size is adjusted in adherence with the standard rules based on the gain ratio defined as $\rho = [U(\boldsymbol{x}^{(i+1)}) - U(\boldsymbol{x}^{(i)})]/\,[U(\boldsymbol{L}^{(i)}(\boldsymbol{x}^{(i+1)})) - U(\boldsymbol{L}^{(i)}(\boldsymbol{x}^{(i)}))]$. The new design $\boldsymbol{x}^{(i+1)}$ is accepted if $\rho > 0$.

In vast majority of applications, the gradient is estimated through finite differentiation (FD), which implies an additional computational burden of n EM simulations of the antenna per algorithm iteration, where n denotes the number of parameters. Thus, the overall CPU overhead of the optimization process is primarily determined by the number of EM simulations induced by the gradient updates, which is further increased in the case of unsuccessful iterations (i.e., $\rho < 0$) when a new candidate design has to be found by solving (3) with a reduced $\boldsymbol{d}^{(i)}$ [24].

2.3. Reduced-Cost Algorithm: Component Procedures

Here, the issue of the high computational overhead of antenna optimization is addressed by introducing two independent acceleration mechanisms: sensitivity variation tracking and design change monitoring. The FD-based gradient update is carried out if recommended by either of the component procedures.

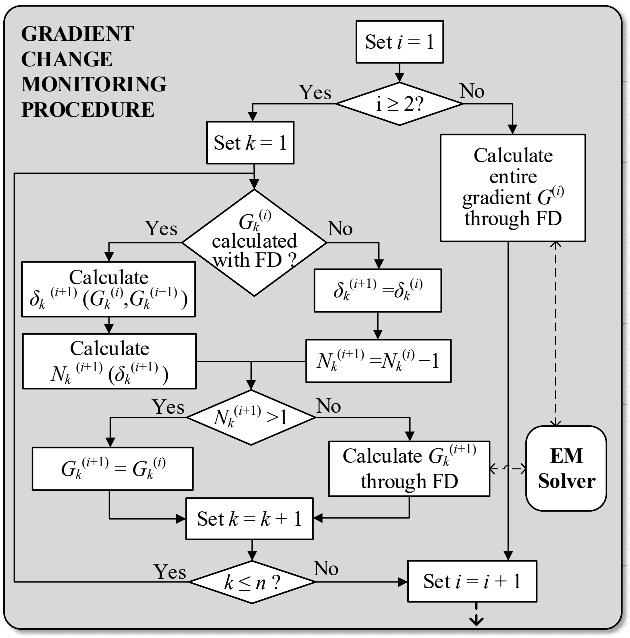


Fig. 1. *Flow diagram of gradient change monitoring procedure (GCMP). The following notation is used: $G_k^{(i)}(f)$ – the k-th column of the antenna response gradient $\boldsymbol{G}_S$; $\delta_k^{(i)}$ – a gradient difference factor; $N_k^{(i)}$ – a number of the future iterations without FD; n – number of antenna parameters; the indices k and i refer to the parameter and algorithm iteration number, respectively.*

In the first procedure (Gradient Change Monitoring Procedure, GCMP), the response gradient differences throughout the algorithm run are monitored in order to detect the parameters characterized by stable sensitivity patterns according to appropriately developed metric. For those parameters, CPU costly FD-based gradient updates are not performed for a certain number of iterations that is proportional to the gradient change magnitude. The flow diagram of GCMP is presented in Fig. 1.

The second acceleration mechanism (Design Change Monitoring Procedure, DCMP) is based on monitoring relative design relocations with respect to the trust region size, as well as a history of the optimization process. The gradient estimation through FD is suppressed for the variables exhibiting minor alterations. The procedure is shown graphically in Fig. 2.

Let us introduce the following notation for the GCMP procedure:

- $G_k^{(i)}(f) = \partial S_{11}(\boldsymbol{x},f)/\partial x_k$ – the k-th column of $\boldsymbol{G}_S$ in the ith iteration of the algorithm;
- $\Delta G_k^{(i)}(f) = |G_k^{(i)}(f)| - |G_k^{(i-1)}(f)|$ – the sensitivity change of the k-th parameter between iterations;
- $M_{G_k^{(i)}(f)} = (|G_k^{(i)}(f)| + |G_k^{(i-1)}(f)|)\ /\ 2$ – the mean sensitivity for the k-th parameter in consecutive iterations;
- $\delta_k^{(i)}$ – a gradient difference factor for the k-th parameter in the i-th iteration; $\delta_{\min}^{(i)} = \min\{k = 1, \ldots, n : \delta_k^{(i)}\}$, $\delta_{\max}^{(i)} = \max\{k = 1, \ldots, n : \delta_k^{(i)}\}$;
- $N_k^{(i)}$ – a number of the future iterations without FD;
- $N_{\min}$, $N_{\max}$ – the minimum and maximum number of iterations without FD, respectively (GCMP control parameter).

In the GCMP procedure, the gradient variation between the two consecutive iterations is quantified using the following metric, referred to as the gradient change factor

$$\delta_k^{(i+1)} = \mathrm{std}\left(\frac{\Delta G_k^{(i)}(f)}{M_{G_k^{(i)}(f)}}\right) \tag{5}$$

In (5), the standard deviation is calculated over the frequency range F. In each iteration, the gradient change factors are first arranged in an ascending order. For lower values of δ_k, a higher value $N_k^{(i)}$ of future iterations without FD is assigned to the k-th parameter. Hence. the estimation of the gradient through FD is performed less frequently. The number $N_k^{(i)}$ is calculated with the use of the following conversion function

$$N_k^{(i)} = \left[\!\left[a \exp\left(b\,\delta_k^{(i)}\right) \right]\!\right] \tag{6}$$

$$a = \exp\left(\frac{\delta_{\min}^{(i)} \ln\left(N_{\min}\right) - \delta_{\max}^{(i)} \ln\left(N_{\max}\right)}{\delta_{\min}^{(i)} - \delta_{\max}^{(i)}}\right) \tag{7}$$

$$b = \frac{\ln\left(N_{\max}\right) - \ln\left(N_{\min}\right)}{\delta_{\min}^{(i)} - \delta_{\max}^{(i)}} \tag{8}$$

where [[.]] denotes the nearest integer function. The control parameters of GCMP procedure: $N_{\min}$ and $N_{\max}$ refer to the minimum and the maximum number of iterations without FD, respectively. They govern the frequency of performing the sensitivity updates through FD. These occur at most once per $N_{\min}$ iterations and at least once per $N_{\max}$ iterations. For the parameters of the smallest, as measured by $\delta_k^{(i)}$, gradient change between subsequent iterations, the maximum number $N_{\max}$ is assigned. Conversely, the higher $\delta_k^{(i)}$, the smaller number $N_k^{(i)}$ is assigned to the k-th parameter indicating that more frequent FD is required.

The value of $N_k^{(i)}$ is based on: (i) the current difference factors $\delta_k^{(i)}$, determined with (5) for the variables with FD in the i-th iteration, and (ii) the previous difference factors, kept from the previous iterations if FD in the i-th iteration was omitted. Throughout the iterations without FD, the value δ_k holds valid, and it is used for selecting $\delta_{\min}$ and $\delta_{\max}$. Therefore, δ_k affects N_k for other parameters with FD update. In addition, for those parameters, the preceding number of iterations is decremented, i.e., $N_k^{(i+1)} = N_k^{(i)} - 1$. For the variables to be updated with FD, $N_k^{(i+1)}$ is calculated using (6).

Let us now introduce the following notation for the DCPM procedure:

- $\|\boldsymbol{d}^{(i)}\|$ – the size of TR region;
- d_{thr} – a user-specified trust-region size threshold (DCMP control parameter);
- $\varphi_k^{(i)}$ – a decision factor for the k-th parameter;
- φ_{thr} – a user-specified threshold for decision factors;
- $v_k^{(i)}$ – an update history count for the k-th parameter.

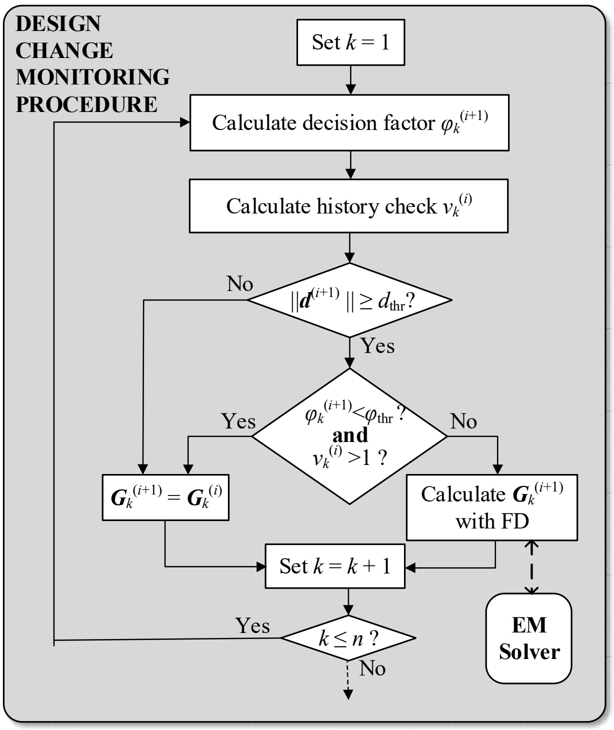


Fig. 2. *Flow diagram of design change monitoring procedure (DCMP). The following notation is used:* $G_k^{(i)}(f)$ *– the k-th column of the antenna response gradient* $\boldsymbol{G}_S$*;* $\|\boldsymbol{d}^{(i)}\|$ *– the size of TR region;* d_{thr} *– user-specified trust-region size threshold;* $\varphi_k^{(i)}$ *– the decision factor;* φ_{thr} *– user-specified threshold for decision factors;* $v_k^{(i)}$ *– update history count; the indices k and i refer to the parameter and algorithm iteration number, respectively.*

In the DCMP procedure, design relocation between iterations is quantified in order to decide upon a possible omission of the gradient update through FD. Let us introduce the following decision factors $\varphi_k^{(i+1)}$ (defined as a relative change of the kth parameter, w.r.t. the TR region size $d_k^{(i)}$ in the ith iteration)

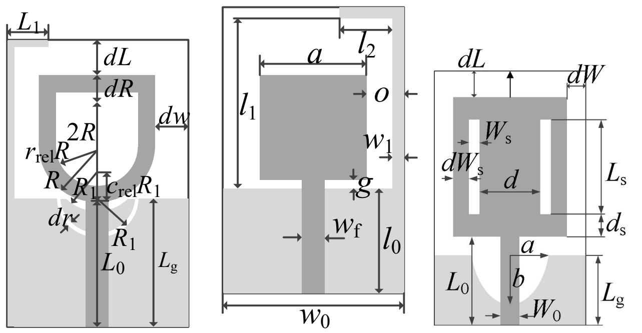
$$\varphi_k^{(i+1)} = \frac{\left| x_k^{(i+1)} - x_k^{(i)} \right|}{d_k^{(i)}} \tag{9}$$

where $x_k^{(i)}$ and $x_k^{(i+1)}$ refer to the k-th components of the vectors $\boldsymbol{x}^{(i)}$, $\boldsymbol{x}^{(i+1)}$, respectively. The maximum number N of iterations without FD is the DCMP procedure control parameter. It ensures that the FD-based gradient update is executed at least once per N iterations. The frequency of FD updates depends also on the size of TR region $\|\boldsymbol{d}^{(i)}\|$. If it is below a user-specified threshold d_{thr} (close to convergence), FD is discouraged for all parameters. However, if $\|\boldsymbol{d}^{(i+1)}\| > d_{\text{thr}}$, for a given parameter k, the FD is skipped if the following conditions are simultaneously fulfilled: (i) the decision factor $\varphi_k^{(i+1)}$ is lower than the user-specified threshold φ_{thr}, and (ii) G_k was updated through FD at least once in the last N iterations (an update history count $v_k^{(i)} \geq 1$). The count $v_k^{(i)}$ is calculated as the total number of gradient updates performed through FD within the last N iterations.

The major difference between the proposed and the reference algorithm lies in the frequency of performing gradient estimation through FD. In the conventional routine, the entire gradient FD is estimated through FD, whereas in the proposed algorithm FD-based gradient estimation is performed only in the two first iterations. Next, FD is omitted if the gradient stability is detected and the design change between iterations is small with respect to the corresponding variables. The combination of the two procedures leads to noticeable computational savings with minor design quality deterioration, which is confirmed by the results of Section 3.

3. Verification Case Studies

This section provides numerical verification of the proposed algorithm as well as comparisons with the reference TR algorithm along with its accelerated versions of [21] and [22].



a b c

Fig. 3. *Antenna structures used for verification of the proposed algorithm. Ground plane marked using light gray shade.*
(a) *Antenna I [25],*
(b) *Antenna II [26],*
(c) *Antenna III [27].*

3.1. Benchmark Cases

The benchmark set consists of three antennas shown in Fig. 3. Antennas I and II are implemented on Taconic RF-35 substrate ($\varepsilon_r = 3.5$, $\tan\delta = 0.0018$, $h = 0.762$ mm). Antenna I, shown in Fig. 3(a), utilizes a quasi-circular radiator with a modified ground plane for bandwidth enhancement [25]. Its independent design variables are: $\boldsymbol{x} = [L_0\ dR\ R\ r_{\rm rel}\ dL\ dw\ L_g\ L_1\ R_1\ dr\ c_{\rm rel}]^T$. Antenna II (cf. Fig. 3 (b)) is a standard rectangular monopole [26] described by the vector of seven independent parameters $\boldsymbol{x} = [l_0\ g\ a\ l_1\ l_2\ w_1\ o]^T$; with $w_0 = 2o + a$. The feeding line width $w_f = 1.7$ mm ensures 50 ohm input impedance. Antenna III is based on the structure of [27], see Fig. 3 (c), and it is implemented on FR4 substrate ($\varepsilon_r = 4.3$ and $h = 1.55$ mm). The structure geometry parameters are $\boldsymbol{x} = [L_g\ L_0\ L_s\ W_s\ d\ dL\ d_s\ dW_s\ dW\ a\ b]^T$. The computational models of all antennas are implemented in CST Microwave Studio and evaluated using its transient solver. The EM models incorporate SMA connectors.

Table 1 Initial Geometry Parameter Vectors of Antennas I through III for the Representative Algorithm Runs of Fig. 4, along with Their Lower and Upper Bounds.

Antenna		Geometry parameter values [mm]										
		L_0	dR	R	$r_{\rm rel}$	dL	dw	L_g	L_1	R_1	dr	$c_{\rm rel}$
I	initial	10.5	1.13	5.2	0.31	2.2	6.3	12.2	2.9	2.7	0.43	0.37
	lower	4.0	0	3	0.1	0	0	4	0	2	0.2	0.2
	upper	15	6	8	0.9	5	8	15	6	5	1.0	0.9
		l_0	g	a	l_1	l_2	w_1	o				
II	initial	24.5	17.3	10.7	7.4	8.6	3.9	2.5				
	lower	10	10	5	5	2	0.1	0.2				
	upper	35	20	15	12	15	10	3.0				
		L_g	L_0	L_s	W_s	d	dL	d_s	dW_s	dW	a	b
III	initial	10.0	12.6	9.2	0.95	3.2	2.9	1.58	2.7	3.1	0.39	0.57
	lower	2	2	2	0.2	2	0	0.1	0.1	0	0.01	0.01
	upper	15	15	20	3	15	15	3	8	5	0.80	0.80

Table 2 Optimization Results for Antennas I through III

Antenna	I			II			III		
Algorithm	CPU overhead[#]	max$\|S_{11}\|$[$]	std max$\|S_{11}\|$)[*]	CPU overhead[#]	max$\|S_{11}\|$[$]	std max$\|S_{11}\|$)[*]	CPU overhead[#]	max$\|S_{11}\|$[$]	std max$\|S_{11}\|$)[*]
Reference	111.2	−14.9	0.6	97.6	−11.9	0.4	111.0	−13.9	1.0
Algorithm of [21]	58.3 [47.6]	−13.7 [1.2]	1.3	45.1 [53.8]	−11.1 [0.8]	1.1	73.1 [34.1]	−12.8 [1.1]	1.3
Algorithm of [22]	37.5 [66.3]	−13.9 [1.0]	1.3	31.5 [67.7]	−11.0 [0.9]	1.6	36.1 [67.5]	−11.6 [2.3]	1.9
This work	40.9 [63.2]	−14.6 [0.3]	0.9	32.3 [66.9]	−11.4 [0.5]	1.0	41.4 [62.7]	−13.7 [0.2]	1.2

[#] Number of EM simulations averaged over 10 algorithm runs.
[$] Maximum in-band reflection max$|S_{11}|$ in dB averaged over 10 algorithm runs. In brackets below: relative computational savings in percent w.r.t. the reference algorithm.
[*] Standard deviation of max$|S_{11}|$ in dB across the set of 10 algorithm runs. In brackets below: degradation of max$|S_{11}|$ w.r.t. the reference algorithm in dB.

3.2. Numerical Results

The proposed algorithm has been applied to optimize Antennas I through III for the best in-band matching within the UWB frequency range of 3.1 GHz to 10.6 GHz. The objective function is defined in Section 2.1 (equation (2)). For all antennas, in order to gather statistical data on the algorithm performance, multiple algorithm runs have been executed with random initial designs. For each initial design, one algorithm run is executed because all considered optimization procedures are deterministic.

Table 1 provides the lower and upper bounds for geometry parameters of Antennas I through III. In addition, Table 1 contains the initial parameter vectors for the representative algorithm runs presented in Fig. 4, where both the initial and optimal antenna responses are shown. The numerical data has been gathered in Table 2. It should be noted that in each iteration of the algorithm, a new full-wave electromagnetic simulation is performed for a new candidate design. The antenna structures are parameterized in a way that virtually eliminates a possibility of yielding physically inconsistent designs (e.g., featuring negative values of certain dimensions), which could lead to failure of the simulation process.

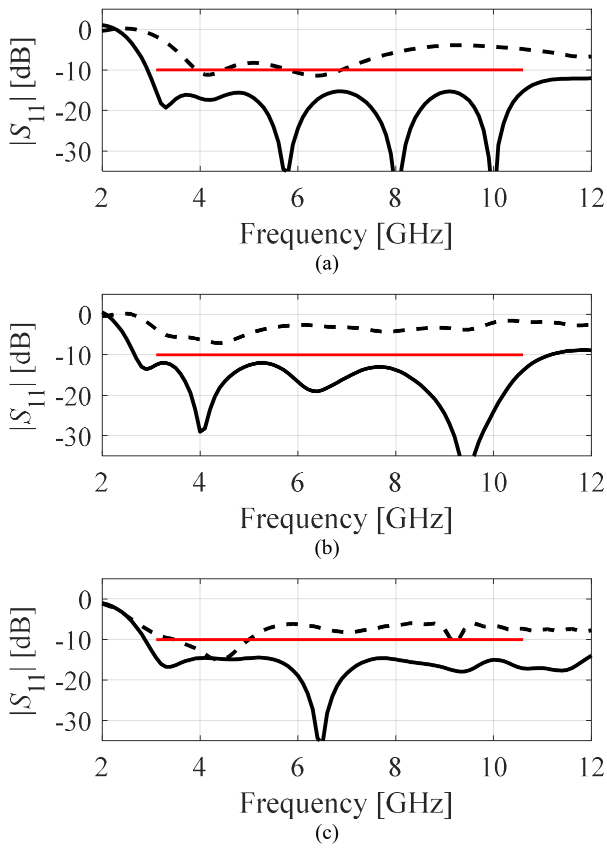


Fig. 4. *Responses for the representative algorithm runs. Horizontal lines indicate the design specifications; (- - -) initial design, (—) optimized design:*
(a) *reflection characteristics of Antenna I [25],*
(b) *reflection characteristics of Antenna II [26],*
(c) *reflection characteristics of Antenna III [27].*

3.3. Discussion

The obtained results are presented in Table 2, which contains: the optimization overhead, the maximum in-band reflection, the computational savings and the design quality measures. The results are obtained for the algorithm proposed in this work, as well as for several benchmark procedures: the reference conventional TR algorithm, and the previously reported routines of [21] and [22].

It can be observed that the proposed algorithm delivers excellent design quality. For all antennas, the degradation of the maximum in-band reflection does not exceed 0.5 dB in comparison to the reference algorithm. At the same time, for the algorithm of [21], the objective function value deteriorates more significantly (up to 1.2 dB for Antenna I). Whereas for the routine of [22], the design quality deterioration is around 1 dB for Antennas I and II; however, it reaches a considerable value of 2.3 dB for Antenna III.

The analysis of computational savings reveals that the number of EM simulations necessary by the proposed optimization routine to converge is notably reduced by 64 percent on average w.r.t. the reference algorithm. For the algorithm of [22], the savings are slightly higher (around 67 percent). This, however, comes at a price of a significant design quality degradation. The algorithm of [21], is not as fast as the two previously meant, it delivers the average computational savings of around 45 percent.

The result repeatability is quantified by the standard deviation of the maximum in-band reflection over the set of ten algorithm runs. It should be emphasized, that the use of the random initial designs generally leads to different local optima. Hence, even for the reference routine a nonzero standard deviation is observed. This should be taken into consideration when comparing the standard deviation for any of the accelerated algorithms with the reference procedure. The true indicator of the results repeatability is the difference between the standard deviations of the expedited algorithm and the reference one. For the proposed algorithm, it is the lowest among all compared accelerated routines, which confirms its reliability.

4. Conclusion

In the paper, a novel trust-region algorithm for expedited design optimization of antenna structures has been introduced. Simultaneous utilization of two mechanisms for suppressing finite-differentiation-based gradient updates allows us for achieving considerable computational savings as well as only minor quality degradation when compared to the reference procedure. In addition, the performance improvements over the recently reported accelerated algorithms have been demonstrated. Our methodology offers excellent result repeatability, which has been verified through statistical analysis involving multiple optimization runs from random initial designs. The proposed approach can be utilized to both accelerate direct EM-driven optimization of antennas and to improve efficiency of design frameworks exploiting variable-fidelity EM simulations.

5. Nomenclature

DCMP – design change monitoring procedure;
EM – electromagnetic;
FD – finite differentiation;
GCMP – gradient change monitoring procedure;
SBO – surrogate-based optimization;
TR – trust region;
UW – ultra-wideband.

6. Acknowledgments

The authors would like to thank Dassault Systemes, France, for making CST Microwave Studio available. This work is partially supported by the Icelandic Centre for Research (RANNIS) Grant 174114051 and by National Science Centre of Poland Grant 2017/27/B/ST7/00563.

6. References

[1] Koziel, S., Ogurtsov, S.: '*Antenna design by simulation-driven optimization. Surrogate-based approach*' (Springer, New York, 2014)

[2] Nocedal, J., Wright, S.J.: '*Numerical Optimization*' (Springer, New York, 2nd edn. 2006)

[3] Rios, L.M., Sahinidis, N.V.: 'Derivative-free optimization: a review of algorithms and comparison of software implementations', *J. Glob. Optim.*, 2013, **56**, (3), pp. 1247–1293

[4] Binelo, M.O., de Almeida, A.L.F., Cavalcanti, F.R.P.: 'A genetic algorithm for the optimization of MIMO antenna arrays' in Cavalcanti, F. (Ed.) '*Resource Allocation and MIMO for 4G and Beyond*' (Springer, New York, 2014)

[5] Bayraktar, Z., Komurcu, M., Bossard, J.A., Werner, D.H.: 'The wind driven optimization technique and its application in electromagnetics', *IEEE Trans. Antennas Propag.*, 2013, **61**, (5), pp. 2745–2757

[6] Vaz, A.I.F., Vicente, L.N.: 'A particle swarm pattern search method for bound constrained global optimization', *J. Glob. Optim.*, 2007, **39**, (2), pp. 197–219

[7] Khalatpour, A., Amineh, R.K., Cheng, Q.S., Bakr, M.H., Nikolova, N.K., Bandler, J.W.: 'Accelerating space mapping optimization with adjoint sensitivities', *IEEE Microw. Wireless Comp. Lett.*, 2011, **21**, (6), pp. 280–282

[8] Papadimitriou, D.I., Giannakoglou, K.C.: 'Aerodynamic shape optimization using first and second order adjoint and direct approaches', *Arch. Comput. Methods Eng.*, 2008, **15**, pp. 447–488

[9] Zhang, Y., Negm, M.H., Bakr, M.H.: 'An adjoint variable method for wideband second-order sensitivity analysis through FDTD', *IEEE Trans. Antennas Propag.*, 2016, **64**, (2), pp. 675–686, 2016

[10] Forrester, A.I.J., Keane, A.J.: 'Recent advances in surrogate-based optimization', *Prog. in Aerospace Sciences*, 2009, **45**, pp. 50−79

[11] Bandler, J.W., Koziel, S., Madsen, K.: 'Editorial—surrogate modeling and space mapping for engineering optimization', *Optim. Eng.*, 2008, **9**, (4), pp. 307–310

[12] Koziel, S., Leifsson, L.: '*Simulation-driven design by knowledge-based response correction techniques*' (Springer, New York, 2016)

[13] Koziel, S., Bandler, J.W., Madsen, K.: 'Space mapping with adaptive response correction for microwave design optimization', *IEEE Trans. Microwave Theory Tech.*, 2009, **57**, (2), pp. 478–486

[14] Koziel, S., Bandler, J.W.: 'Rapid yield estimation and optimization of microwave structures exploiting feature-based

statistical analysis', *IEEE Trans. Microwave Theory Tech.*, 2015, **63**, (1), pp. 107–114
[15] Koziel, S., Unnsteinsson, S.D.: 'Expedited design closure of antennas by means of trust-region-based adaptive response scaling', *IEEE Ant. Wireless Propag. Lett.*, 2018, **17**, (6), pp. 1099–1103.
[16] Kleijnen, J.P.C.: 'Kriging metamodeling in simulation: A review', *European J. Operat. Research*, 2009, **192**, (3), pp. 707–716
[17] Jacobs, J.P.: 'Characterization by Gaussian processes of finite substrate size effects on gain patterns of microstrip antennas', *IET Microwaves Ant. Prop.*, 2016, **10**, (11), pp. 1189–1195
[18] Bishop, C.: '*Pattern recognition and machine learning'* (Springer, New York, 2006)
[19] Rayas-Sanchez, J.E., Gutierrez-Ayala, V.: 'EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping', *IEEE Trans. Microw. Theory Techn.*, 2006, **54**, (12), pp. 4528–4537
[20] Wu, X., Peng, X., Chen, W., Zhang W.: 'A developed surrogate-based optimization framework combining HDMR-based modeling technique and TLBO algorithm for high-dimensional engineering problems', *Struct. Multidisc. Optim.*, 2019, **60**, (2), pp. 663–680
[21] Koziel, S., Pietrenko-Dabrowska, A.: 'Reduced-cost EM-driven optimization of antenna structures by means of trust-region gradient-search with sparse Jacobian updates', *IET Microwaves Ant. Prop.*, 2019, **13**, (10), pp. 1646–1652
[22] Koziel, S., Pietrenko-Dabrowska, A.: 'Accelerated antenna optimization using gradient search with selective Broyden updates', IEEE Ant. Prop. Symp., Atlanta, USA, July 2019
[23] Koziel, S., Bekasiewicz, A.: '*Multi-objective design of antennas using surrogate models'* (World Scientific, Singapore, 2016)
[24] Conn, A.R., Gould, N.I.M., Toint, P.L.: '*Trust region methods'* (MPS-SIAM Series on Optimization, 2000)
[25] Alsath M.G.N., Kanagasabai, M.: 'Compact UWB monopole antenna for automotive communications', *IEEE Trans. Ant. Prop.*, 2015, **63**, (9), pp. 4204–4208.
[26] Haq, M.A., Koziel, S., Cheng, Q.S.: 'EM-driven size reduction of UWB antennas with ground plane modifications', Int. Applied Comp. Electromagnetics Society (ACES China) Symp., Suzhou, China, August 2017, pp. 1–2
[27] Koziel, S., Bekasiewicz, A.: 'Low-cost multi-objective optimization of antennas using Pareto front exploration and response features', IEEE Int. Symp. Ant. Prop., Fajardo, Puerto-Rico, July 2016