Postprint of: Kozieł S., Pietrenko-Dąbrowska A., Reliable EM-driven size reduction of antenna structures by means of adaptive penalty factors, IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION (2021), pp. 1-13, DOI: 10.1109/TAP.2021.3111285

Reliable EM-Driven Size Reduction of Antenna Structures by Means of Adaptive Penalty Factors

Slawomir Koziel, *Senior Member*, *IEEE*, and Anna Pietrenko-Dabrowska, *Senior Member*, *IEEE*

***Abstract*—Miniaturization has become of paramount importance in the design of modern antenna systems. In particular, compact size is essential for emerging application areas such as internet of things, wearable and implantable devices, 5G technology, or medical imaging. On the other hand, reduction of physical dimensions generally has a detrimental effect on antenna performance. From the perspective of numerical optimization, miniaturization is a heavily constrained problem, with most constraints being expensive to evaluate due to involving full-wave electromagnetic (EM) simulation. A convenient way of handling such a task is a penalty function approach where constraint violations contribute, upon suitable scaling, to the primary objective, i.e., the antenna size. The penalty coefficients determining the aforementioned contributions are normally adjusted through engineering experience, which does not allow for a reliable control of antenna performance figures. This paper proposes an efficient management scheme for adaptive adjustment of penalty coefficients, which eliminates the need for objective function setup by trial and error, and ensures precise control of the design constraints. Our approach is demonstrated using three broadband antennas optimized for minimum size with acceptance thresholds imposed on the in-band matching level. The adaptive adjustment of penalty coefficients is shown to outperform experience-driven setups in terms of constraint control precision and the final design quality.**

***Index Terms*— Antenna design; compact antennas; EM-driven optimization; size reduction; constraint handling; penalty functions.**

I. Introduction

Design of modern antennas is a highly intricate endeavour that has to account for stringent performance requirements imposed on the structure characteristics, e.g., broadband [1] or multi-band matching [2], high gain [3], circular polarization [4], multiple-input-multiple-output (MIMO) operation [5], low sidelobe levels [6], or pattern diversity [7]. Many of these are dictated by the demands pertinent to the emerging areas such as 5G communications [8], medical imaging [9], or the internet of things (IoT) [10]. For the majority of these but also other space-limited applications (e.g., wireless sensing [11], wearable [12], or implantable devices [13]), one of the critical prerequisites is a compact size of the radiator. Ensuring sufficient miniaturization levels adds another layer of difficulty to an already challenging design process because reducing physical dimensions of the antenna may negatively affect its electrical performance parameters, including efficiency [14], impedance bandwidth [15], or pattern stability [16]. Therefore, in practice, a trade-off solution has to be sought. The latter is usually understood as minimum-size design for which acceptance thresholds for relevant performance figures are still satisfied.

The design process of compact antennas normally starts at the level of topology development [17], [18]. Therein, appropriate modifications are incorporated into the base geometries. These alterations may be applied to the feeding structures [19], ground plane [20], but also the radiator itself [21]. Some of the most popular techniques include L-shape stubs [22], ground-plane slits [23], embedded cross slots [24], radiator slots [25], or stepped-impedance feeding lines [26]. Other methods include application of defected ground structures (DGS) [27], [28], shorting pins between the radiator and the ground plane [29], utilization of dielectric resonators [30], the employment of tailored shapes of radiating components [31], meandering the radiating elements (in dipole antennas [32]), or exploiting the symmetric field distribution and cutting techniques (e.g., [33]). All of the aforementioned approaches often allow for achieving considerable miniaturization rates as compared to the base designs; however, they also lead to the increase of the number of geometry parameters that need to be meticulously tuned to maintain satisfactory levels of the relevant performance figures. An additional challenge is that the relationships between certain parameters and antenna characteristics may become counter-intuitive, which essentially rules out traditional parameter adjustment methods such as supervised parameter sweeping as a reliable design tool. As a matter of fact, some of the geometry alterations that appear to be beneficial from the miniaturization perspective at the stage of parametric studies, may turn out to be irrelevant when rigorous optimization is applied [34].

Numerical optimization methods have become widely accepted tool of choice for reliable parameter adjustment of

The manuscript was submitted on March 3, 2021. This work was supported in part by the Icelandic Centre for Research (RANNIS) Grant 217771, and by 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.

antenna structures regardless of particular design goals (e.g., matching improvement [35], gain maximization [36], pattern synthesis [37], isolation enhancement in MIMO systems [38], wideband near-field correction of antennas [39], design of magnetic conductor surfaces [40], and others, e.g., [41]-[43]), in both single- [44] and multi-objective setups [45]. Notwithstanding, EM-driven optimization is challenging for several reasons, which include high computational cost incurred by repetitive simulations of the antenna at hand, a typically large number of parameters that need to be tuned, but also the necessity of handling several performance figures. Over the years, a number of methodologies and algorithmic approaches have been developed to alleviate these difficulties. One possibility is acceleration of gradient-based procedure through the involvement of adjoint sensitivities [46]. However, this is an intrusive approach of limited accessibility. Another option is the utilization of sparse sensitivity updates (e.g., [47], [48]), or dedicated EM solvers combined with CAD procedures [49], [50]. Recently, a rapid growth of interest in surrogate-based optimization (SBO) methods has been observed [51]-[53], which can be attributed to unprecedented computational speedup that can be achieved by the use of fast metamodels [54]. The two popular branches of SBO are physics-based [55] and data-driven techniques [56]. The former rely on an underlying low-fidelity representations (e.g., equivalent networks in the case of microwave devices, or coarse-mesh simulations for antenna systems [57]), which make these methods less generic but also more immune to dimensionality issues and (potentially) very efficient. Nevertheless, the performance of physics-based SBO procedures is highly dependent on the setup (and, consequently, the user experience). Popular methods include space mapping [58], response correction routines [59], [60], cognition-driven design [61], or feature-based optimization [62]. Because in general, physics-based models lack the universal approximation capability, they are mostly employed for local optimization [63]. On the other hand, data-driven techniques are more versatile and accessible [64] but also severely limited by the curse of dimensionality and nonlinearity of antenna responses. A construction of reliable metamodels for antenna structures described by more than a few parameters over broad ranges thereof is hardly possible. Yet approximation methods can be combined with machine learning algorithms [65] or nature-inspired metaheuristics [66], [67], to enable globalized search. The most popular modelling techniques include kriging [68], radial basis function [69], support vector regression [70], Gaussian process regression [71], neural networks [72], or polynomial chaos expansion [73] (especially suited for solving uncertainty quantification tasks [74], [75]).

In the case of compact antennas, the major role is played by local optimization techniques because reasonable initial solutions are often available upon the accomplishment of the topology evolution stage followed by the initial parametric studies. Nevertheless, optimization-based size reduction is challenging due to being a constrained task with expensive constraints (evaluated through EM simulation). Consequently, in practice, an implicit approach is often assumed, where the initial design obtained through parametric studies and featuring satisfactory size is optimized to satisfy the acceptance thresholds concerning electrical performance figures (e.g., the impedance bandwidth, etc.). This is an implicit approach. Explicit size reduction, where the antenna footprint is the primary minimization objective, is much more difficult in numerical terms because the optimum design is normally located at the boundary of the feasible region [76] (so that at least one constraint, typically related to antenna matching, is active [77]). This boundary is difficult to be explored due to numerical noise issues as well as the aforesaid expensive to evaluate constraints, although some mitigation techniques have been developed (e.g., the objective relaxation method [78], or adaptive acceptance thresholds [79]). A more generic workaround is the penalty function approach, where the constraints are handled implicitly by contributing to the primary objective in case of their violations [78]. This alleviates the major difficulties, in particular, turns the design task into a formally unconstrained one, apart from the usual lower and upper bounds set on the design parameters. At the same time, the balance between the contribution of the main objective (here, the antenna size) and constraint violations is controlled by the proportionality coefficients (referred to as penalty factors). When dealing with smooth functions, the penalty factor adjustment is not critical, yet it becomes important when handling EM-simulated antenna responses. On the one hand, setting these too relaxed fosters more significant size reduction at the expense of just as significant constraint violation. On the other hand, increasing the penalty factors limits violations but makes the optimization problem more challenging due to high nonlinearity of the objective function in the vicinity of the feasible region boundary. Consequently, setting up the penalty coefficient using engineering experience (i.e., to ensure noticeable contributions of the penalty terms upon unacceptable constraint violations) often leads to poor constraint control or inferior results in terms of achievable values of the primary objective, let alone the fact that optimum setup is very much problem dependent.

In this paper, a technique for adaptive adjustment of penalty coefficients in EM-driven size reduction of antenna structures is proposed. According to the presented methodology, the penalty terms are controlled using the currently observed levels of constraint violation as well as the convergence status of the optimization algorithm. At the initial stages of the process, the penalty coefficients are kept at low values to facilitate exploration of the parameter space. The tolerance for constraint violations is gradually reduced, and the coefficients are increased or decreased accordingly, depending on the relative values of the violations (with respect to the current tolerance levels). The three major benefits of the developed mechanisms include a more precise control over the constraint violation, eliminating the need for 'manual' setup of the objective function, and, improved reliability of the optimization process. The latter is understood as a possibility of rendering antenna designs that feature smaller footprints as compared to the conventional approach with fixed coefficient. The presented procedure is demonstrated using three examples

of broadband monopole antennas, optimized for minimum size with the constraints imposed on their reflection characteristics. Comprehensive comparison to the reference algorithm, involving multiple optimization runs from random initial designs, conclusively corroborates the aforementioned advantages.

The novelty and the technical contributions of this work can be highlighted as follows: (i) the development of a management scheme for automated adjustment of penalty coefficients in EM-driven size reduction of antenna structures, (ii) corroborating a possibility of a precise control of constraint violation while eliminating the necessity of manual objective function setup, (iii) demonstrating the improvement of the performance and reliability of the optimization process, which manifests itself in yielding smaller antenna footprints while retaining the target values of electrical performance figures. To the best knowledge of the authors, none of these have been reported so far in the literature. The latter statement does not only apply to the methodological aspects but also to the demonstrated improvements of the design quality rendered with the proposed procedure with respect to the engineering-insight-based optimization problem setup.

II. EM-Driven Size Reduction with Implicit Constraint Handling

This section formulates the problem of explicit simulation-driven size reduction of antenna structures, and explains the concept of implicit constraint handling by means of penalty functions. Furthermore, we motivate and introduce the proposed adaptive penalty factor adjustment scheme. The section is complemented by the formulation of the entire optimization framework, here, involving gradient-based algorithm as the underlying optimization engine. Numerical verification and benchmarking will be presented in Section III.

A. Antenna Miniaturization by EM-Driven Optimization

Despite ongoing popularity of hands-on procedures, primarily based on parameter sweeping, rigorous numerical optimization has become a widely accepted tool in antenna design. Its importance is especially pronounced at the late stages of the design process, when the structure topology has been already established, and geometry parameter values have to be tuned to improve the performance figures as much as possible. In this work, we are particularly interested in size reduction, which—as elaborated on in Section I—is becoming a prerequisite for a growing number of application areas that include wireless communications, internet of things, remote sensing, or wearable/implantable devices.

One of the challenges of optimization-driven miniaturization is that size reduction is only one of the performance figures to be taken into account in the process. Other requirements are pertinent to the electrical and field properties of the antenna at hand, e.g., the impedance bandwidth, axial ratio bandwidth, or gain variability. Given a specific application, those demands can normally be expressed in terms of the acceptance thresholds. From this perspective, miniaturization is a constrained optimization problem. An additional challenge is that most of the constraints are expensive to evaluate as the knowledge of antenna characteristics comes from EM simulation. This is a primary reason why a penalty function approach is perhaps the most convenient way of addressing antenna size reduction.

In the following, the vector of antenna adjustable parameters will be denoted as $\boldsymbol{x} = [x_1 \ \ldots \ x_n]^T$, where n is the number of independent variables. The simulation-based antenna miniaturization can be formulated as

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

subject to the inequality constraints $g_k(\boldsymbol{x}) \le 0$, $k = 1, \ldots, n_g$, and equality constraints $h_k(\boldsymbol{x}) = 0$, $k = 1, \ldots, n_h$. In (1), U stands for the merit function, which is a metric of the design quality. Here, the merit function will take a form of

$$U(\boldsymbol{x}) = A(\boldsymbol{x}) \tag{2}$$

where A is the antenna footprint area (in the case of planar antennas). Alternatively, it may be any other measure of the antenna size such as the volume in the case of, e.g., dielectric resonator antennas.

B. Implicit Constraint Handling through Penalty Functions

As mentioned before, in this work, the constraints are handled in an implicit manner, using the penalty function approach. This requires reformulation of the optimization task as follows

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} U_P(\boldsymbol{x}) \tag{3}$$

The objective function U_P in (3) takes the form of

$$U_P(\boldsymbol{x}) = U(\boldsymbol{x}) + \sum_{k=1}^{n_g+n_h} \beta_k c_k(\boldsymbol{x}) \tag{4}$$

The penalty functions $c_k(\boldsymbol{x})$ measure violations of the respective requirements; β_k are the penalty coefficients. Note that the inequality and equality constraints are handled together in (4). The total number of constraints is $n_c = n_g + n_h$. It should also be mentioned that for antenna systems, the majority of constraints are of the inequality type. Several examples of possible constraints are discussed below for the sake of clarification:

- Given an operating bandwidth F, ensure that the antenna reflection $|S_{11}(\boldsymbol{x},f)|$ does not exceed –10 dB within F, i.e. $|S_{11}(\boldsymbol{x},f)| \le -10$ dB for $f \in F$. Here, as a default, relative constraint violations are considered, therefore, the penalty function is of the form

$$c(\boldsymbol{x}) = \left[\frac{\max\{S(\boldsymbol{x}) + 10, 0\}}{10}\right]^2 \tag{5}$$

  where $S(\boldsymbol{x}) = \max\{f \in F : |S_{11}(\boldsymbol{x},f)|\}$. Note that $c(\boldsymbol{x}) \ne 0$ if and only if the constraint is violated. Note also that using the second power $[.]^2$ ensures smoothness of U_P at the boundary of the feasible region, which is essential to facilitate exploration of this region (most constraints are active at the optimum design).
- Ensure that the axial ratio $AR(\boldsymbol{x},f)$ of a CP antenna does not exceed 3 dB within the operating range F, i.e., $AR(\boldsymbol{x},f) \le 3$ dB for $f \in F$. In this case, the penalty function takes a similar form as in (5), i.e.,

$$c(\boldsymbol{x}) = \left[\frac{\max\{AR(\boldsymbol{x}) - 3, 0\}}{3}\right]^2 \tag{6}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

where $A_R(\boldsymbol{x}) = \max\{f \in F : AR(\boldsymbol{x},f)\}$.

- Ensure that variability of realized gain $G(\boldsymbol{x},f)$ is below 2 dB within the antenna operating range F, i.e., $\Delta G(\boldsymbol{x},f) \leq 2$ dB for $f \in F$, where $\Delta G(\boldsymbol{x},f) = \max\{f \in F : G(\boldsymbol{x},f)\} - \min\{f \in F : G(\boldsymbol{x},f)\}$. The penalty function would be of the form

$$c(\boldsymbol{x}) = \left[\frac{\max\{G(\boldsymbol{x}) - 2, 0\}}{2}\right]^2 \quad (7)$$

where $G(\boldsymbol{x}) = \max\{f \in F : \Delta G(\boldsymbol{x},f)\}$.

C. Adaptive Penalty Factors: Why and How

Formulation (4) of the EM-driven miniaturization problem allows for incorporating all relevant performance figures into a scalar objective function. The major benefit is that the task can be solved using standard, e.g., gradient-based optimization routines. However, the values of the penalty terms β_k play an important role in controlling the balance between the primary objective (here, size reduction of the antenna), and satisfying the acceptance thresholds for other figures of interest. When the penalty coefficients are too small, considerable violations of the design constraints will be observed. On the other hand, using large coefficient values may lead to numerical problems because of high nonlinearity of the objective function in the vicinity of the feasible region boundary, which is where the optimum design is normally located. The standard procedure is to set up the coefficients to ensure that the contributions of the penalty terms are noticeable when constraint violations become unacceptably high; however, all of this is very much subjective. In particular, this way of handling the problem does not guarantee that the acceptance thresholds will be satisfied within the prescribed tolerance. Furthermore, the optimum setup is clearly problem dependent. Some of these issues have been illustrated in Fig. 1. In particular, the picture shows the dependence of the optimum design according to the objective function U_P on the penalty coefficient setup, including the specific cases of too relaxed or too stiff establishment of the penalty factors.

In this work, the aim is to develop, implement, and validate a strategy for adaptive adjustment of the penalty factors so that the entire process of setting up the objective function becomes automated. At the same time, the penalty terms should be altered in the course of the optimization process to ensure satisfaction of the design constraints at the final solution (within the user-decided tolerances), while facilitating identification of the constrained optimum. Having these in mind, the adjustment scheme is developed by taking into account the following assumptions:

- The initial value of the penalty terms should be low in order to aid the parameter space exploration in the vicinity of the feasible region boundary;
- For the same reason, at the early stages of the search process, the tolerances for constraint violations should be increased with respect to their target values;
- The penalty coefficient values should gradually increase over time in order to ensure that constraint violations do not exceed the user-defined tolerances upon the algorithm convergence;
- For the same reason, the tolerances should be tightened to eventually reach their original (target) values;
- The increase of the penalty coefficients should be coupled with the constraint violations, and the adjustments should be realized adaptively by taking into account the violation status at any given iteration of the optimization algorithm.

Implementation of the penalty coefficient adjustment scheme according to the above guidelines requires a measure of the algorithm convergence status, i.e., an indicator of whether the optimization process is in the early, middle, or the late stage.

Here, the following two metrics are employed:

- $\|\boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}\|$ – convergence in argument;
- $U_P(\boldsymbol{x}^{(i+1)}) - U_P(\boldsymbol{x}^{(i)})$ – convergence in objective function value.

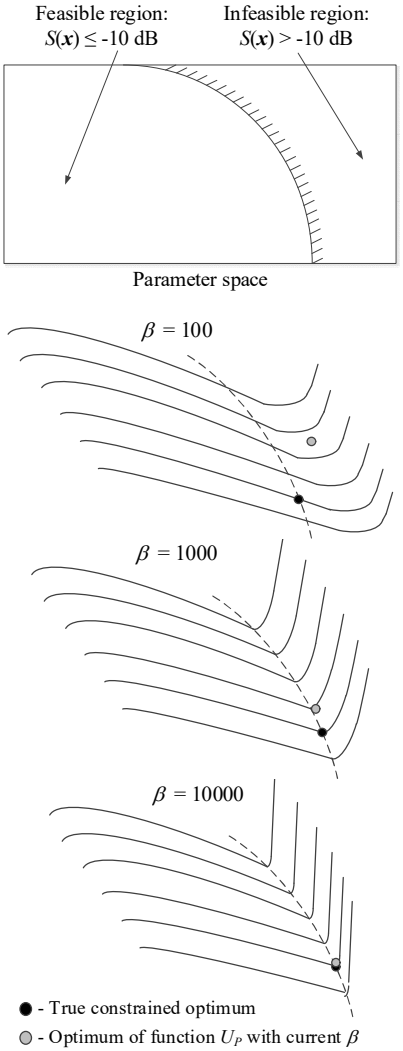


Fig. 1. Graphical illustration of the parameter space and feasible set using the example of antenna reflection constraint $S(\boldsymbol{x}) \leq -10$ dB with $S(\boldsymbol{x}) = \max\{f \in F : |S_{11}(\boldsymbol{x},f)|\}$ (cf. (5)): (a) feasible and infeasible sets, (b) landscapes of the objective function U_P (4) for the various values of the penalty coefficient β. The true optimum is located at the boundary of the feasible region. For small β, the objective function optimum is infeasible and may exhibit considerable constraint violation; for large values of β, the optimum approaches the true optimum; however, the optimization problem becomes numerically demanding due to very steep landscape in the vicinity of the feasible set boundary. The optimization result may become increasingly dependent on the initial design as traversing the boundary is hindered by the objective function nonlinearity.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

The termination conditions of the optimization algorithm are based on the same metrics. More specifically, the process is terminated if either $\|\boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}\| < \varepsilon_x$, or $|U_P(\boldsymbol{x}^{(i+1)}) - U_P(\boldsymbol{x}^{(i)})| < \varepsilon_U$, where ε_x and ε_U are user-defined thresholds. In the numerical experiments of Section II, we use $\varepsilon_x = \varepsilon_U = 10^{-3}$.

D. Adaptive Penalty Factors: Implementation

Based on the prerequisites of Section II.*C*, the management of penalty coefficients proposed in this work will be implemented with the use of the following two mechanisms: (i) adaptive adjustment of the constraint violation tolerance thresholds, and (ii) adaptive adjustment of the penalty coefficient values.

Current Constraint Violation

A violation of the kth constraint at the design $\boldsymbol{x}$ will be denoted as $D_k(\boldsymbol{x})$. It is defined to be non-negative. For example, in the case of the reflection constraint $|S_{11}(\boldsymbol{x},f)| \le -10$ dB, it would be $D_k(\boldsymbol{x}) = \max\{S(\boldsymbol{x}) + 10, 0\}$, where $S(\boldsymbol{x}) = \max\{f \in F : |S_{11}(\boldsymbol{x},f)|\}$ (cf. (5)).

Let $D_{k.\max}$ be a (user-defined) tolerance for the kth constraint violation, and $M_k > 1$ be a multiplication factor (also user-defined), which determines the maximum tolerance $M_k D_{k.\max}$ for the kth constraint violation that may occur during the optimization process. As mentioned in Section II.*C*, one of the assumptions of the proposed procedure is to allow for larger violations (here, $M_k D_{k.\max}$) at the beginning of the optimization process, and tighten the tolerance (here, to $D_{k.\max}$) towards the convergence. The transition between $M_k D_{k.\max}$ and $D_{k.\max}$ is realized using the convergence metric $C^{(i)}$, defined as

$$C^{(i)}(\varepsilon_x, \varepsilon_U) = \max\left\{\frac{\varepsilon_x}{\|\boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}\|}, \frac{\varepsilon_U}{|U_P(\boldsymbol{x}^{(i+1)}) - U_P(\boldsymbol{x}^{(i)})|}\right\} \quad (9)$$

It can be observed that $C^{(i)}$ estimates the convergence stage of the optimization process with respect to the termination criteria discussed in Section II.*C*. At the beginning of the optimization run, i.e., when the design relocations are large, $C^{(i)}$ is small. On the other hand, it approaches the unity when close to convergence.

Current Violation Tolerance

The current violation tolerance $D_{k.tol}^{(i+1)}$ of the kth constraint for the $(i+1)$th iteration of the optimization algorithm is computed as

$$D_{k.tol}^{(i+1)} = \begin{cases} M_k D_{k.\max} & \text{if } C^{(i)}(\varepsilon_x, \varepsilon_U) \le M_C \\ \max\left\{D_{k.\max}, d_{k.tol}^{(i+1)}\right\} & \end{cases} \quad (10)$$

where

$$d_{k.tol}^{(i+1)} = M_k D_{k.\max} + D_{k.\max}\left(1 - M_k\right)\left[1 - \frac{\log(C^{(i)}(\varepsilon_x, \varepsilon_U))}{\log M_C}\right] \quad (11)$$

Note that $D_{k.tol}^{(i+1)}$ is set to the maximum value of $M_k D_{k.\max}$ if $C^{(i)}(\varepsilon_x, \varepsilon_U) \le M_C$, where M_C is a user-defined threshold determining the initialization of tolerance adjustment. In the numerical experiments of Section III, we use $M_C = 10^{-2}$. Upon convergence, $D_{k.tol}^{(i+1)}$ is reduced to $D_{k.\max}$. Figure 2 shows the dependence of $D_{k.tol}$ on the C-factor (9) for exemplary values of $D_{k.\max}$, M_k, and M_C.

Current Penalty Coefficients

Using $D_{k.tol}^{(i+1)}$, and the actual constraint violation $D_k(\boldsymbol{x}^{(i)})$ at the current iteration point $\boldsymbol{x}^{(i)}$, the kth penalty coefficient $\beta_k^{(i+1)}$ for the $(i+1)$th iteration is calculated as

if $D_k(\boldsymbol{x}^{(i)}) < D_{k.tol}^{(i+1)} / M_D$

$$\beta_k^{(i+1)} = \max\{\beta_k^{(i)} / m_{dec}, \beta_{k.\min}\}$$

elseif $D_k(\boldsymbol{x}^{(i)}) > D_{k.tol}^{(i+1)}$

$$\beta_k^{(i+1)} = \min\{\beta_k^{(i)} m_{inc}, \beta_{k.\max}\}$$

else

$$\beta_k^{(i+1)} = \beta_k^{(i)}$$

end

where the control parameters have the following meaning:

- M_D – multiplication factor determining the threshold for penalty coefficient reduction;
- m_{dec}, m_{inc} – decrease and increase factors for penalty coefficient adjustment, respectively;
- $\beta_{k.\min}$, $\beta_{k.\max}$ – minimum and maximum penalty coefficient values, respectively.

It can be noticed that the thresholds for a reduction and an increase of the penalty coefficients are different ($D_{k.t}^{(i+1)} / M_D$ and $D_{k.t}^{(i+1)}$, respectively), which is to improve the stability of the optimization process; in this work, $M_D = 2$ is used, but this value is not critical. The coefficients m_{dec} and m_{inc} are set to 2 and 3, respectively, and the reason to keep them different is the same as before. Finally, the minimum and maximum penalty coefficient values are introduced to avoid the situation when the penalty factors are adjusted beyond reasonable limits. For size reduction with antenna footprint expressed in mm^2, these are set to $\beta_{k.\min} = 1$, and $\beta_{k.\max} = 10^6$.

E. Trust-Region Gradient Search Algorithm

The adjustment scheme of Section II.*D*. can be incorporated into any iterative search procedure of a descent type, in particular, gradient-based algorithms. In this work, for illustration purposes, it is coupled with trust-region (TR) algorithm [80], which is briefly recalled below. Furthermore, the TR procedure using fixed (user-defined) penalty coefficients will be employed as a reference algorithm.

The trust-region gradient-search algorithm finds a series of approximations to the optimum solution $\boldsymbol{x}^*$ of (3), (4), denoted as $\boldsymbol{x}^{(i)}$, $i = 0, 1, \ldots$, where $\boldsymbol{x}^{(0)}$ is the starting point, and

$$\boldsymbol{x}^{(i+1)} = \arg\min_{\boldsymbol{x};\, -\boldsymbol{d}^{(i)} \le \boldsymbol{x} - \boldsymbol{x}^{(i)} \le \boldsymbol{d}^{(i)}} U_L^{(i)}(\boldsymbol{x}) \quad (12)$$

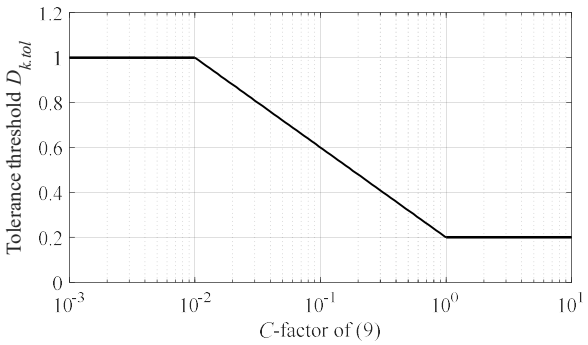


Fig. 2. Constraint violation tolerance $D_{k.t}$ as calculated using (10) as a function of the C-factor (9). For illustration, the plots were created for $D_{k.\max} = 0.2$, $M_k = 5$, and $M_C = 10^{-2}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

Here, $U_L^{(i)}$ is a temporary objective function constructed the same way as that of (4); however, using linear expansion models of the relevant antenna responses, collectively denoted as $\boldsymbol{R}(\boldsymbol{x})$, instead of EM-simulated characteristics.

The local model $\boldsymbol{L}^{(i)}$ at the design $\boldsymbol{x}^{(i)}$ is defined as

$$\boldsymbol{L}^{(i)}(\boldsymbol{x}) = \boldsymbol{R}(\boldsymbol{x}^{(i)}) + \boldsymbol{J}(\boldsymbol{x}^{(i)}) \cdot (\boldsymbol{x} - \boldsymbol{x}^{(i)}) \quad (13)$$

where $\boldsymbol{J}(\boldsymbol{x}^{(i)})$ is a Jacobian matrix of $\boldsymbol{R}$ at $\boldsymbol{x}^{(i)}$, estimated using finite differentiation. At the level of individual components, the objective function is defined as

$$U_L^{(i)}(\boldsymbol{x}) = A(\boldsymbol{x}) + \sum_{k=1}^{n_c} \beta_k^{(i)} c_{L.k}(\boldsymbol{x}) \quad (14)$$

where constraint functions $c_{L.k}$ have the same form as c_k (cf. (4)), except that they are computed from the expansion model $\boldsymbol{L}^{(i)}$ of the antenna responses rather than directly from $\boldsymbol{R}$.

The sub-problem (12) is solved within the interval $[\boldsymbol{x}^{(i)} - \boldsymbol{d}^{(i)}, \boldsymbol{x}^{(i)} + \boldsymbol{d}^{(i)}]$, referred to as a trust region. The size thereof is normally proportional to the parameter space size (i.e., it is a fraction of the interval $[\boldsymbol{l}, \boldsymbol{u}]$, where $\boldsymbol{l}$ and $\boldsymbol{u}$ are the lower and upper parameter bounds, respectively), and adjusted using the standard TR rules [80]. The candidate design $\boldsymbol{x}^{(i+1)}$ is accepted if $U_P(\boldsymbol{x}^{(i+1)}) < U_P(\boldsymbol{x}^{(i)})$; otherwise, the vector $\boldsymbol{d}^{(i)}$ is reduced [81], and the iteration is re-launched.

F. Optimization with Adaptive Penalty Factors. Complete Procedure

This section summarizes the operating flow of the trust-region gradient-based algorithm with adaptively adjusted penalty coefficients as described in Section II.*D*. The control parameters of the procedure are as follows:

- ε_x, ε_U – termination thresholds (cf. Section II.*C*);
- $M_k > 1$ – multiplication factor determining the maximum tolerance $M_k D_{k.\max}$ for the kth constraint violation allowed to occur during the optimization process (cf. Section II.*D*);
- M_C – threshold determining the initialization of tolerance adjustment (cf. Section II.*D*);
- M_D – multiplication factor determining the threshold for penalty coefficient reduction (cf. Section II.*D*);
- m_{dec}, m_{inc} – decrease and increase factors for penalty coefficient adjustment, respectively;
- $\beta_{k.\min}$, $\beta_{k.\max}$ – minimum and maximum penalty coefficient values; $k = 1, \ldots, n_c$;
- $D_{k.\max}$ – maximum allowed constraint violations; $k = 1, \ldots, n_c$;
- $\beta_{k.0}$ – initial penalty coefficient values, respectively; $k = 1, \ldots, n_c$.

The termination thresholds are set by the user and depend on the required resolution of the optimization process (here, we use 10^{-3} for both ε_x and ε_U). The values for other parameters were discussed in the earlier sections and are summarized here: $M_k = 5$, $M_C = 10^{-2}$, $M_D = 2$, $m_{dec} = 2$, $m_{inc} = 3$, $\beta_{k.\min} = 1$, and $\beta_{k.\max} = 10^6$. The initial coefficient values $\beta_{k.0}$ will be set to 10^2, which corresponds to relatively relaxed conditions in terms of enforcing constraint satisfaction (assuming that the penalty functions are defined in a relative sense as in (5) or (6), and the antenna footprint is in mm^2). The thresholds $D_{k.\max}$ are set by the user depending on the required accuracy of satisfying the constraints.

The algorithm operation has been summarized in Fig. 3. For additional clarification, Fig. 4 shows the flow diagram of the optimization procedure. For the convenience of the reader, the following provides a concise summary of the penalty coefficient adaptation process. Starting from the initially relaxed conditions (low values of penalty coefficients), the penalty terms are tightened when the constraints are violated beyond the prescribed tolerance levels, and relaxed otherwise. This allows for exploring the boundary of the feasible region boundary and facilitates the antenna size reduction process. In order to improve the constraint control, the mentioned tolerance levels are gradually reduced upon the algorithm convergence, which generally leads to increasing the penalty coefficient values. When close to the conclusion of the optimization process, the design is expected to be close to the feasible space boundary, and the penalty coefficients might be related, typically reaching the levels close to their optimum values for the respective antenna structures.

1. Set the iteration counter $i = 0$;
2. Set $\beta_k^{(i)} = \beta_{k.0}$, $k = 1, \ldots, n_c$;
3. Evaluate antenna response $\boldsymbol{R}(\boldsymbol{x}^{(i)})$;
4. Evaluate antenna sensitivities $\boldsymbol{J}(\boldsymbol{x}^{(i)})$ using finite differentiation;
5. Construct a linear model $\boldsymbol{L}^{(i)}(\boldsymbol{x}) = \boldsymbol{R}(\boldsymbol{x}^{(i)}) + \boldsymbol{J}(\boldsymbol{x}^{(i)}) \cdot (\boldsymbol{x} - \boldsymbol{x}^{(i)})$ (cf. (13));
6. Construct objective function $U_L^{(i)}(\boldsymbol{x}) = A(\boldsymbol{x}) + \sum_{k=1}^{n_c} \beta_k^{(i)} c_{L.k}(\boldsymbol{x})$ (cf. (14));
7. Solve $\boldsymbol{x}^{(i+1)} = \arg \min_{\boldsymbol{x};\, -\boldsymbol{d}^{(i)} \le \boldsymbol{x} - \boldsymbol{x}^{(i)} \le \boldsymbol{d}^{(i)}} U_L^{(i)}(\boldsymbol{x})$;
8. Evaluate antenna response $\boldsymbol{R}(\boldsymbol{x}^{(i+1)})$;
9. Update trust-region size vector $\boldsymbol{d}^{(i)}$ [80];
10. **if** $U_P(\boldsymbol{x}^{(i+1)}) < U_P(\boldsymbol{x}^{(i)})$
    Compute $D_{k.tol}^{(i+1)}$, $k = 1, \ldots, n_c$, using (10);
    Evaluate constraint violations $D_k(\boldsymbol{x}^{(i+1)})$, $k = 1, \ldots n_c$;
    if $D_k(\boldsymbol{x}^{(i)}) < D_{k.tol}^{(i+1)} / M_D$
    $\beta_k^{(i+1)} = \max\{\beta_k^{(i)} / m_{dec}, \beta_{k.\min}\}$
    elseif $D_k(\boldsymbol{x}^{(i)}) > D_{k.tol}^{(i+1)}$
    $\beta_k^{(i+1)} = \min\{\beta_k^{(i)} m_{inc}, \beta_{k.\max}\}$
    else
    $\beta_k^{(i+1)} = \beta_k^{(i)}$
    end
    Set $i = i + 1$;
    end
11. **if** $||\boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}|| < \varepsilon_x$ OR $|U_P(\boldsymbol{x}^{(i+1)}) - U_P(\boldsymbol{x}^{(i)})| < \varepsilon_U$
    Go to 12;
    else
    Go to 4;
    end
12. END.

Fig. 3. Operation of the trust-region gradient-based algorithm with numerical derivatives and adaptively adjusted penalty coefficients.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

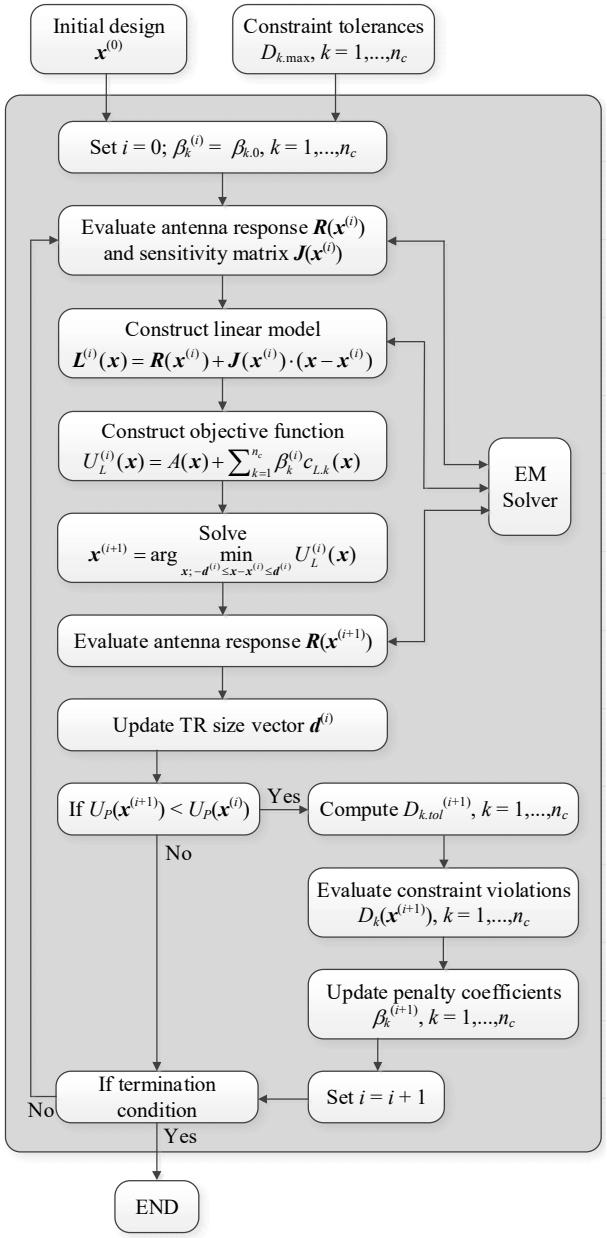


Fig. 4. Flow diagram of the prosed optimization algorithm with adaptively adjusted penalty coefficients.

III. Demonstration Examples

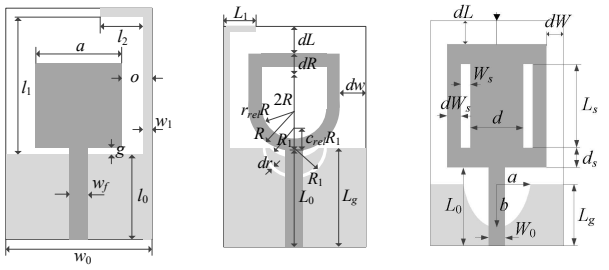
This section provides numerical verification of the adaptive adjustment scheme introduced in Section II to control the penalty coefficients. The procedure is demonstrated using three examples of broadband monopole antennas optimized for minimum size with the constraint imposed on their reflection response. The results obtained using adaptive adjustment are juxtaposed against the fixed setup featuring different coefficient values, corresponding to both the relaxed and tight conditions concerning constraint satisfaction. The considered benchmark structures and the experimental setup are described in Section III.*A*. Section III.*B* gathers the numerical data, whereas Section III.*C* provides a result discussion.

A. Verification Cases and Experimental Setup

Figure 5 shows the geometries of the three benchmark antennas employed to carry out numerical verification of the penalty coefficient adaptation procedure of Section II. All structures are broadband monopoles supposed to operate within UWB band (3.1 GHz to 10.6 GHz). The details concerning antenna substrates and design variables can be found in Table I. More information about each structure can be found in the cited references. The computational models are implemented and simulated in CST Microwave Studio using the time-domain solver; the models incorporate the SMA connectors.

The antennas are to be optimized for minimum size, which is understood here as the substrate area $A(\boldsymbol{x})$. For the sake of simplicity, we only consider one constraint, which is $|S_{11}(\boldsymbol{x},f)| \le -10$ dB for the frequency range from 3.1 GHz to 10.6 GHz. Owing to this setup, there is only one penalty coefficient, which will be referred to as β (note that the subscript k has been dropped). The penalty function c is defined as in (5), i.e., based on the relative constraint violation. However, for the purpose of penalty coefficient adjustment, the absolute violation is measured. The maximum allowed constraint violation $D_{k.\max}$ is set to 0.2 dB.

Compact antennas are often described by many parameters, which is a result of various topological alterations introduced to facilitate miniaturization of the structure. This, along with strictly numerical issues, in particular, the presence of numerical noise inherent to EM simulation results, make the optimization task a multimodal one. This means, that the final result may depend on the starting point, and the algorithm performance evaluated based on a single run may not be representative. To address this issue, here, the algorithm is assessed in a statistical sense, based on ten independent runs, executed using random initial designs.



(a) (b) (c)

Fig. 5. Benchmark antenna structures: (a) Antenna I, (b) Antenna II, (c) Antenna III. Ground planes marked using light gray shade.

TABLE I Benchmark antenna structures

Antenna	Substrate	Designable Parameters [mm]	Other Parameters [mm]
I [82]	RF-35 ($\varepsilon_r = 3.5$, $h = 0.762$ mm)	$\boldsymbol{x} = [l_0\ g\ a\ l_1\ l_2\ w_1\ o]^T$	$w_0 = 2o + a$, $w_f = 1.7$
II [83]	RF-35 ($\varepsilon_r = 3.5$, $h = 0.762$ mm)	$\boldsymbol{x} = [L_0\ dR\ R\ r_{rel}\ dL\ dw\ Lg\ L_1\ R_1\ dr\ c_{rel}]^T$	$w_0 = 1.7$
III [84]	FR4 ($\varepsilon_r = 4.3$, $h = 1.55$ mm)	$\boldsymbol{x} = [L_g\ L_0\ L_s\ W_s\ d\ dL\ d_s\ dW_s\ dW\ a\ b]^T$	$W_0 = 3.0$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

There are two performance metrics of interest: (i) the average antenna footprint, and (ii) the average constraint violation, both at the final design obtained through optimization.

The algorithm proposed in this work is compared to the standard TR procedure utilizing fixed penalty coefficient with the experiments performed for the following values: $\beta = 10^p$, with $p = 2$, 3, 4, 5, and 6. The termination thresholds are set to $\varepsilon_x = \varepsilon_U = 10^{-3}$, other control parameters are as given in Section II.*F*.

B. Results

The numerical results have been gathered in Tables II through IV. Figures 6, 7 and 8 show the reflection characteristics of Antennas I through III at the selected initial and optimized designs obtained using the proposed algorithm. The same figures also illustrate the evolution of the penalty factor throughout the optimization process. The data in Tables II, II, and IV include the average antenna footprint and constraint violation, the standard deviations thereof, as a measure of repeatability of solutions, as well as the computational cost of the optimization process expressed in terms of the required number of EM simulations of the respective antenna structure. The detailed discussion of the results will be provided in Section III.*C* should be noted that the experimental validation is not provided as it is irrelevant to the paper topic and meaningless due to statistical evaluation of the algorithm performance. Also, all of the antenna structures considered here have been validated in the earlier works (e.g., [82]-[84]).

C. Discussion

The analysis of the results presented in Section III.C lead to several conclusions concerning both the importance of the penalty coefficient adjustment, and the performance of the proposed adaptive procedure when compared to that of the conventional approach. To facilitate the analysis, the data from Tables II through IV has been presented graphically in Fig. 9 showing the average antenna footprint and constraint violation versus the penalty coefficient, obtained for fixed-β optimization runs, along with the horizontal lines marking the footprint areas and constraint violation for the proposed adaptive procedure.

TABLE II OPTIMIZATION RESULTS FOR ANTENNA I

Penalty function setup	Performance figure				
	Antenna footprint A [mm^2] [1]	Std(A) [2]	Constraint violation D [dB] [3]	Std(D) [dB] [4]	CPU cost[5]
$\beta = 10^2$ (fixed)	305.4	49.7	6.7	1.7	40.2
$\beta = 10^3$ (fixed)	318.1	42.6	1.2	0.4	43.8
$\beta = 10^4$ (fixed)	317.7	42.3	0.4	0.7	42.2
$\beta = 10^5$ (fixed)	318.8	43.3	0.05	0.2	41.4
$\beta = 10^6$ (fixed)	320.9	45.8	0.06	0.3	42.2
Adaptive β (this work)	314.1	42.3	0.3	0.2	50.0

[1] Optimized antenna footprint averaged over ten algorithm runs.
[2] Standard deviation of the optimized antenna footprint averaged over ten algorithm runs.
[3] Constraint violation, defined as $D = \{3.1 \text{ GHz} \le f \le 10.6 \text{ GHz} : \max|S_{11}(f)|\} + 10$, averaged over ten algorithm runs.
[4] Standard deviation of the constraint violation D, averaged over ten algorithm runs.
[5] CPU cost expressed in terms of the average number of EM simulations of the antenna.

TABLE III OPTIMIZATION RESULTS FOR ANTENNA II

Penalty function setup	Performance figure				
	Antenna footprint A [mm^2] [1]	Std(A) [2]	Constraint violation D [dB] [3]	Std(D) [dB] [4]	CPU cost[5]
$\beta = 10^2$ (fixed)	113.7	9.07	8.4	0.53	124.2
$\beta = 10^3$ (fixed)	250.4	24.0	1.2	0.5	180.3
$\beta = 10^4$ (fixed)	318.6	60.0	0.14	0.1	133.2
$\beta = 10^5$ (fixed)	331.6	63.4	0.10	0.14	119.2
$\beta = 10^6$ (fixed)	367.6	51.9	0.05	0.11	168.3
Adaptive β (this work)	281.6	37.1	0.23	0.15	181.7

[1] Optimized antenna footprint averaged over ten algorithm runs.
[2] Standard deviation of the optimized antenna footprint averaged over ten algorithm runs.
[3] Constraint violation, defined as $D = \{3.1 \text{ GHz} \le f \le 10.6 \text{ GHz} : \max|S_{11}(f)|\} + 10$, averaged over ten algorithm runs.
[4] Standard deviation of the constraint violation D, averaged over ten algorithm runs.
[5] CPU cost expressed in terms of the average number of EM simulations of the antenna.

TABLE IV OPTIMIZATION RESULTS FOR ANTENNA III

Penalty function setup	Performance figure				
	Antenna footprint A [mm^2] [1]	Std(A) [2]	Constraint violation D [dB] [3]	Std(D) [dB] [4]	CPU cost[5]
$\beta = 10^2$ (fixed)	56.1	3.8	8.6	0.60	175.6
$\beta = 10^3$ (fixed)	212.8	14.3	1.0	0.40	164.9
$\beta = 10^4$ (fixed)	255.0	25.1	0.15	0.10	138.1
$\beta = 10^5$ (fixed)	280.1	47.4	0.05	0.07	154.0
$\beta = 10^6$ (fixed)	285.8	29.6	0.0	0.01	131.0
Adaptive β (this work)	215.6	3.6	0.25	0.14	189.9

[1] Optimized antenna footprint averaged over ten algorithm runs.
[2] Standard deviation of the optimized antenna footprint averaged over ten algorithm runs.
[3] Constraint violation, defined as $D = \{3.1 \text{ GHz} \le f \le 10.6 \text{ GHz} : \max|S_{11}(f)|\} + 10$, averaged over ten algorithm runs.
[4] Standard deviation of the constraint violation D, averaged over ten algorithm runs.
[5] CPU cost expressed in terms of the average number of EM simulations of the antenna.

The observations can be summarized as follows.

- The optimum value of penalty coefficient is clearly problem dependent. For Antenna I, it is between $\beta = 10^4$ and 10^5, about 10^4 for Antenna II, and between 10^3 to 10^4 for Antenna III. This means that setting up any specific value of β up front is unlikely to turn out optimal;
- Using lower than optimum penalty coefficient results in a rapid increase of the constraint violation, which may readily exceed the assumed tolerance thresholds by a factor of five or more. On the other hand, using too high values generally leads to quality degradation in terms of achievable footprint area. This degradation may be significant as in the case of Antennas II and III;
- Adaptive adjustment of penalty coefficients allows for considerably more precise control over constraint violation; in all considered cases, the average violation is close to the assumed tolerance threshold of 0.2 dB;

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

- At the same time, adaptive adjustment of β results in improved design quality. In all considered cases, the obtained footprint areas are smaller than those rendered with fixed-β versions, assuming comparable levels of constraint violation (see Fig. 9 for graphical illustration);
- Equivalently, constraint violations for the optimization runs with fixed β, resulting in antenna sizes comparable to those obtained with the adaptive scheme, are significantly higher (at the level of 2 dB or more).
- The computational cost of the optimization process is comparable for both fixed-beta setups and the adaptive scheme, with the slightly higher expenses for the latter. This is expected because more precise control over the constraint violation require more elaborative search in the vicinity of the feasible region boundary. As a matter of fact, this can be observed by comparing the CPU cost for fixed-beta setups: the cost of optimization process is higher whenever the penalty coefficient is close to its optimum value.

Figures 10 through 12 show the realized gain, total efficiency and radiation patterns for Antennas I through III, respectively, for at the initial and final designs corresponding to the same algorithm runs as visualized in Figs. 6 through 8. It should be emphasized that none of these characteristics were handled in the optimization process; consequently, whatever is observed is merely a by-product of the size reduction procedure. Because the initial designs generally violate reflection specifications for most of the operating bandwidth, the efficiency is degraded as compared to the optimized designs. The antenna gain is evaluated at the broadside direction. The radiation patterns do not change considerably at the final designs and—as expected from the monople antennas—they exhibit omnidirectional characteristics in the H-plane and dipole-like patterns in the E-plane.

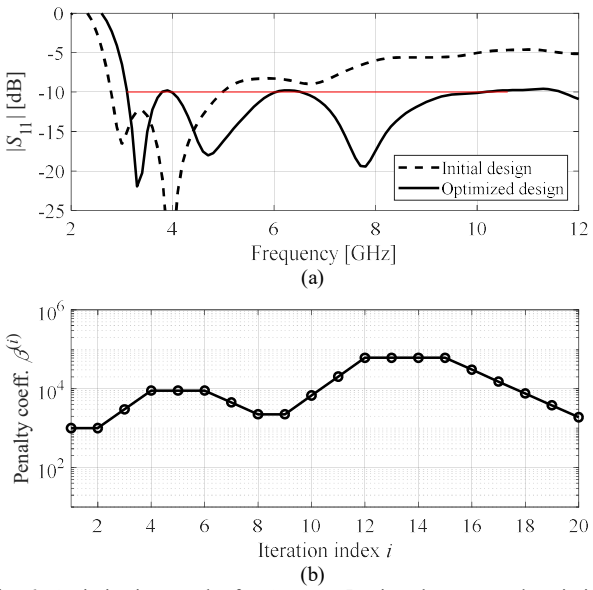


Fig. 6. Optimization results for Antenna I using the proposed optimization algorithm with adaptively adjusted penalty coefficients: (a) reflection characteristics for the representative run: initial design (- - -), and optimized design (—); design specifications marked using a horizontal line; (b) penalty coefficient versus iteration index.

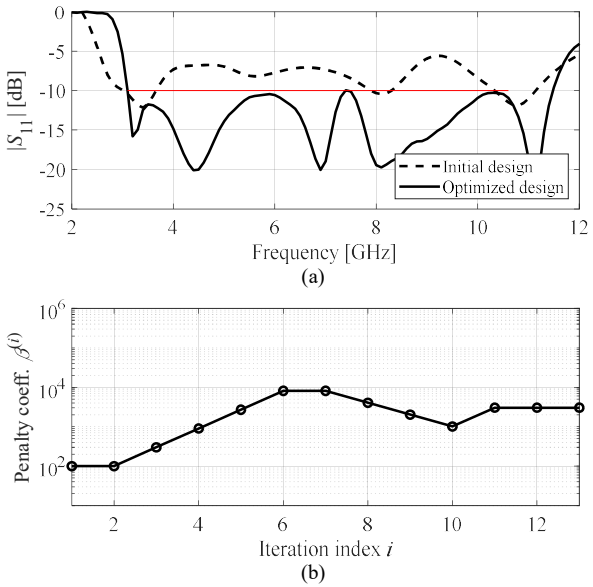


Fig. 7. Optimization results for Antenna II using the proposed optimization algorithm with adaptively adjusted penalty coefficients: (a) reflection characteristics for the representative run: initial design (- - -), and optimized design (—); design specifications marked using a horizontal line; (b) penalty coefficient versus iteration index.

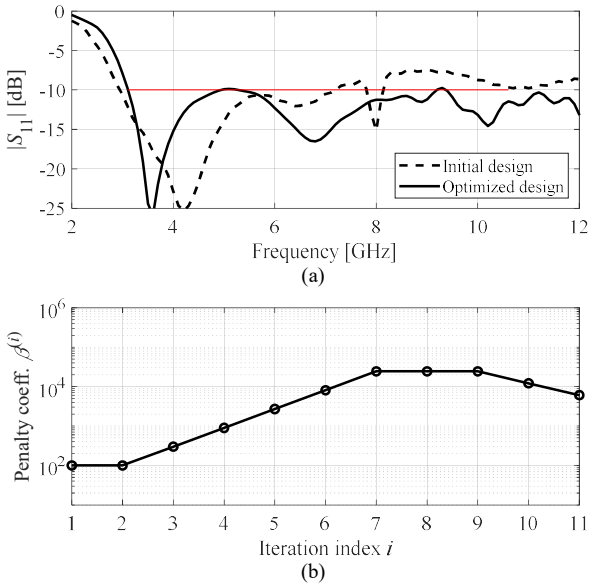


Fig. 8. Optimization results for Antenna III using the proposed optimization algorithm with adaptively adjusted penalty coefficients: (a) reflection characteristics for the representative run: initial design (- - -), and optimized design (—); design specifications marked using a horizontal line; (b) penalty coefficient versus iteration index.

Overall, the proposed adaptive scheme for penalty coefficient adjustment demonstrably improves the performance of the optimization process. On the one hand, it eliminates the need for trial-and-error-based objective function setup: as shown, the optimum values of penalty coefficients are indeed problem dependent. On the other hand, it ensures a more precise control of the constraint violation, while leading to smaller antenna footprints

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

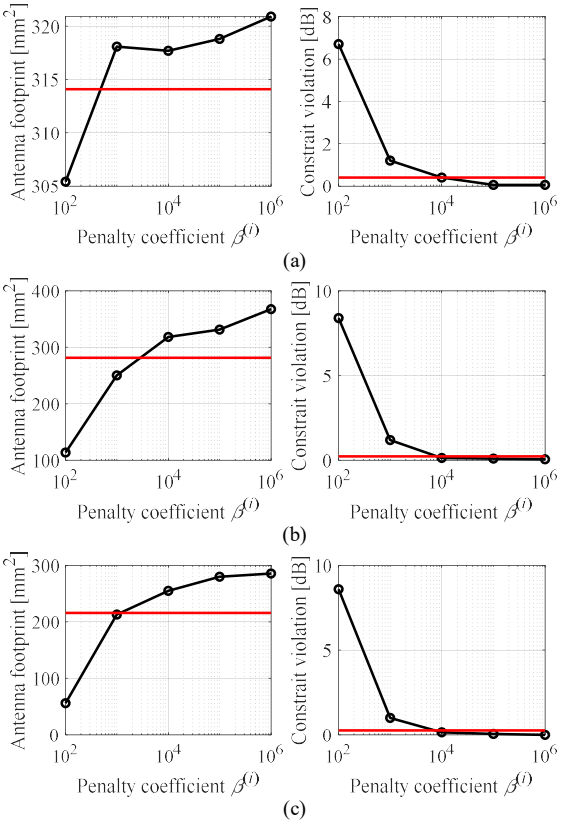


Fig. 9. Optimized antenna footprint area (top) and constraint violations (bottom), averaged over the ten optimization runs, versus the penalty coefficient value for fixed-β algorithm. The horizontal lines denote the average footprint area (top) and constraint violation (bottom) for the proposed adaptive procedure: (a) Antenna I, (b) Antenna II, (c) Antenna III. It can be observed that the miniaturization rate achieved using the adaptive penalty coefficient adjustment is better than for the fixed-β versions assuming the same constraint violation levels. Thus, the performance of adaptive procedure is better than the benchmark even assuming that the penalty coefficient of the latter was set at its optimum value, which is clearly impossible beforehand.

IV. Conclusion

This paper proposed a novel approach to simulation-driven size reduction of antenna structures using local optimization routines and adaptive adjustment of penalty factors for efficient handling of design constraints. The penalty coefficient adjustment is based on the convergence indicators of the search process as well as monitoring the constraint violation levels. It eliminates the need of a guess-work and ensures precise control of both the primary objective (antenna size) and the acceptance thresholds imposed on antenna electrical characteristics.

Our procedure has been coupled with the trust-region gradient search algorithm and comprehensively validated using three broadband antennas optimized for minimum footprint area with constraints imposed on their reflection characteristics. The procedure was benchmarked against optimization with fixed penalty coefficients of different values, corresponding to both relaxed and tight conditions regarding the constraint violation tolerances. The results indicate that the adaptive adjustment scheme enables the precise control of the antenna reflection levels while yielding the design of as small footprints as those obtained with optimum objective function setups for the respective test cases. Automation of the coefficient selection process turns out to be instrumental in maintaining the optimization process reliability as the suitable penalty coefficient arrangement is very much problem dependent. The future work will include more extensive validation of the method for handling multiple constraints along with its application for miniaturization of circular polarization and high-gain antennas.

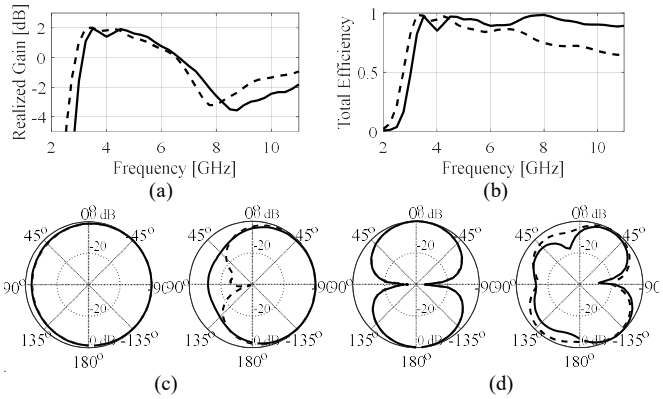


Fig. 10. EM-simulated characteristics of Antenna I at the initial (- - -) and the final design (—), corresponding to the algorithm run considered in Fig. 6: (a) realized gain, (b) total efficiency, (c) H-plane radiation patterns at 4 GHz and 8 GHz, (d) E-plane radiation patterns at 4 GHz and 8 GHz.

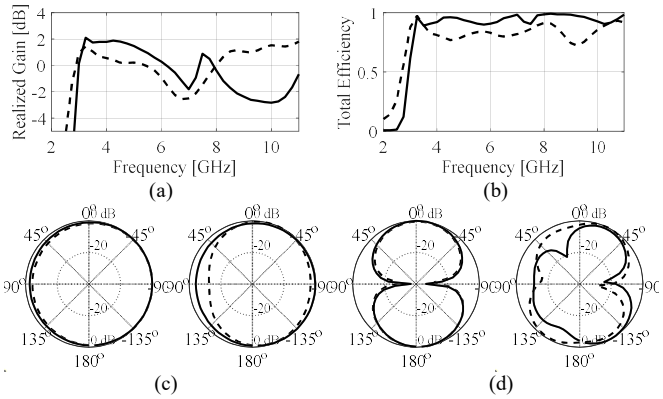


Fig. 11. EM-simulated characteristics of Antenna II at the initial (- - -) and the final design (—), corresponding to the algorithm run considered in Fig. 7: (a) realized gain, (b) total efficiency, (c) H-plane radiation patterns at 4 GHz and 8 GHz, (d) E-plane radiation patterns at 4 GHz and 8 GHz.

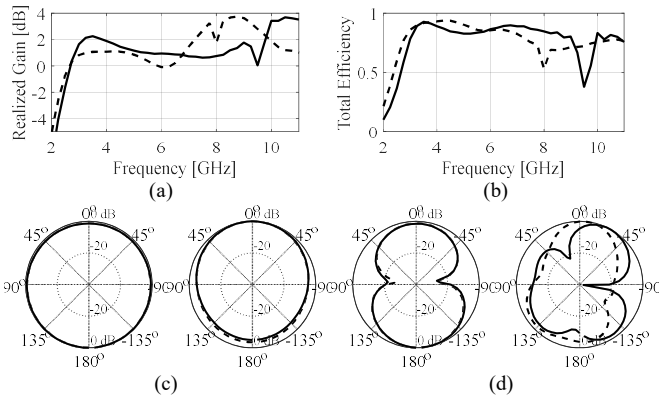


Fig. 12. EM-simulated characteristics of Antenna II at the initial (- - -) and the final design (—), corresponding to the algorithm run considered in Fig. 8: (a) realized gain, (b) total efficiency, (c) H-plane radiation patterns at 4 GHz and 8 GHz, (d) E-plane radiation patterns at 4 GHz and 8 GHz.

ACKNOWLEDGMENT

The authors would like to thank Dassault Systemes, France, for making CST Microwave Studio available.

REFERENCES

[1] M. Midya, S. Bhattacharjee, and M. Mitra, "Broadband circularly polarized planar monopole antenna with g-shaped parasitic strip," *IEEE Ant. Wireless Propag. Lett.*, vol. 18, no. 4, pp. 581-585, 2019.

[2] Y. Gong, S. Yang, B. Li, Y. Chen, F. Tong, and C. Yu, "Multi-band and high gain antenna using AMC ground characterized with four zero-phases of reflection coefficient," *IEEE Access*, vol. 8, pp. 171457-171468, 2020.

[3] Z. Liang, S. Lv, Y. Li, J. Liu, and Y. Long, "Compact folded slot antenna and its endfire arrays with high gain and vertical polarization," *IEEE Ant. Wireless Propag. Lett.*, vol. 19, no. 5, pp. 786-790, 2020.

[4] B. Ghosh, D. Bhattacharya, P. D. Sinha, and D. H. Werner, "Design of circular waveguide annular slot-coupled two-layer DRA for linear and circular polarizations," *IEEE Ant. Wireless Propag. Lett.*, vol. 19, no. 6, pp. 1012-1016, 2020.

[5] S. S. Jehangir and M. S. Sharawi, "A compact single-layer four-port orthogonally polarized yagi-like MIMO antenna system," *IEEE Trans. Ant. Propag.*, vol. 68, no. 8, pp. 6372-6377, 2020.

[6] Y. Geng, J. Wang, Y. Li, Z. Li, M. Chen, and Z. Zhang, "High-efficiency leaky-wave antenna array with sidelobe suppression and multibeam generation," *IEEE Ant. Wireless Propag. Lett.*, vol. 16, pp. 2787-2790, 2017.

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

[8] P. Mei, G. F. Pedersen, and S. Zhang, "A broadband and FSS-based transmitarray antenna for 5G millimeter-wave applications," *IEEE Ant. Wireless Propag. Lett.*, vol. 20, no. 1, pp. 103-107, 2021.

[9] S. Li, S. Wang, Q. An, G. Zhao, and H. Sun, "Cylindrical MIMO array-based near-field microwave imaging," *IEEE Trans. Ant. Propag.*, vol. 69, no. 1, pp. 612-617, 2021.

[10] Z. Cai, Y. Zhou, Y. Qi, W. Zhuang, and L. Deng, "A millimeter wave dual-lens antenna for IoT-based smart parking radar system," *IEEE Internet of Things J.*, vol. 8, no. 1, pp. 418-427, 2021.

[11] T. Wu, R. Li, and M. M. Tentzeris, "A scalable solar antenna for autonomous integrated wireless sensor nodes," *IEEE Ant. Wireless Propag. Lett.*, vol. 10, pp. 510-513, 2011.

[12] G. Gao, R. -F. Zhang, W. -F. Geng, H. -J. Meng, and B. Hu, "Characteristic mode analysis of a nonuniform metasurface antenna for wearable applications," *IEEE Ant. Wireless Propag. Lett.*, vol. 19, no. 8, pp. 1355-1359, 2020.

[13] A. Valanarasi and R. Dhanasekaran, "Optimum band ε shaped miniature implantable antennas for telemetry applications," *IEEE Trans. Ant. Propag.*, vol. 69, no. 1, pp. 55-63, 2021.

[14] S. Koziel, Q.S. Cheng and S. Li, "Optimization-driven antenna design framework with multiple performance constraints," *Int. J. RF Microwave CAE*, vol. 28, no. 4, 2018.

[15] T. K. Roshna, U. Deepak, V. R. Sajitha, K. Vasudevan, and P. Mohanan, "A compact UWB MIMO antenna with reflector to enhance isolation," *IEEE Trans. Ant. Propag.*, vol. 63, no. 4, pp. 1873-1877, 2015.

[16] J. Liu, K.P. Esselle, S.G. Hay and S. Zhong, "Effects of printed UWB antenna miniaturization on pulse fidelity and pattern stability," *IEEE Trans. Ant. Prop.*, vol. 62, no. 8, pp 3903-3910, 2014.

[17] L. Wen, S. Gao, Q. Luo, Q. Yang, W. Hu, Y. Yin, J. Wu, and X. Ren, "A wideband series-fed circularly polarized differential antenna by using crossed open slot-pairs," *IEEE Trans. Ant. Prop.*, vol. 68, no. 4, pp. 2565-2574, 2020.

[18] K. Wong, H. Chang, C. Wang, and S. Wang, "Very-low-profile grounded coplanar waveguide-fed dual-band WLAN slot antenna for on-body antenna application," *IEEE Ant. Wireless Propag. Lett.*, vol. 19, no. 1, pp. 213-217, 2020.

[19] M.A. Haq and S. Koziel, "Feed line alterations for optimization-based design of compact super wideband MIMO antennas in parallel configuration," *IEEE Ant. Wireless Prop. Lett.*, vol. 18, no. 10, pp. 1986-1990, 2019.

[20] M.A. Haq and S. Koziel, "Ground plane alterations for design of high-isolation compact wideband MIMO antenna," *IEEE Access*, vol. 6, pp. 48978-48983, 2018.

[21] V.V. Reddy and N.V.S.N. Sarma, "Compact circularly polarized asymmetrical fractal boundary microstrip antenna for wireless applications," *IEEE Ant. Wireless Prop. Lett.*, vol. 13, pp. 118-121, 2014.

[22] J. Ren, W. Hu, Y. Yin, and R. Fan, "Compact printed MIMO antenna for UWB applications," *IEEE Ant. Wireless Prop. Lett.*, vol. 13, pp. 1517-1520, 2014.

[23] G. Teni, N. Zhang, J. Qiu, and P. Zhang, "Research on a novel miniaturized antipodal Vivaldi antenna with improved radiation," *IEEE Ant. Wireless Prop. Lett.*, vol. 12, pp. 417-420, 2013.

[24] X. Qin and Y. Li, "Compact dual-polarized cross-slot antenna with colocated feeding," *IEEE Trans. Ant. Propag.*, vol. 67, no. 11, pp. 7139-7143, 2019.

[25] J. Tao and Q. Feng, "Compact ultrawideband MIMO antenna with half-slot structure," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, pp. 792-795, 2017.

[26] W.H. Tu, S.H. Hsu, and K. Chang, "Compact 5.8-GHz rectenna using stepped-impedance dipole antenna," *IEEE Ant. Wireless Prop. Lett.*, vol. 6, pp. 282-284, 2007.

[27] B.R.S. Reddy and D. Vakula, "Compact zigzag-shaped-slit microstrip antenna with circular defected ground structure for wireless applications," *IEEE Ant. Wireless Prop. Lett.*, vol. 14, pp. 678-681, 2015.

[28] P.R. Prajapati, G.G.K. Murthy, A. Patnaik, and M.V. Kartikeyan, "Design and testing of a compact circularly polarised microstrip antenna with fractal defected ground structure for L-band applications," *IET Microwave Ant. Prop.*, vol. 9, no. 11, pp. 1179-1185, 2015.

[29] A. Dahlan and M.R. Kamarudin, "Shorted microstrip patch antenna with parasitic element," *J. Electr. Waves. Appl.*, vol. 24, pp. 327-339, 2010.

[30] M.A. Kasgari, S.K.A. Rahim and M. Khalily, "Two segments compact dielectric resonator antenna for UWB application," *IEEE Ant. Wireless Prop. Lett.*, vol. 1, pp. 1533-1536, 2012.

[31] J.P. Park, S.M. Han, and T. Itoh, "A rectenna design with harmonic-rejecting circular-sector antenna," *IEEE Ant. Wireless Prop. Lett.*, vol. 3, pp. 52-54, 2004.

[32] K. Ding, C. Gao, D. Qu, and Q. Yin, "Compact broadband MIMO antenna with parasitic strip," *IEEE Ant. Wireless Propag. Lett.*, vol. 16, pp. 2349-2353, 2017.

[33] X. Zhu, Y. Guo, and W. Wu, "A compact dual-band antenna for wireless body-area network applications," *IEEE Ant. Wireless Propag. Lett.*, vol. 15, pp. 98-101, 2016.

[34] M.A. Haq and S. Koziel, "On topology modifications for wideband antenna miniaturization," *AEU - Int. J. Electr. Comm.*, vol. 94, pp. 215-220, 2018.

[35] D. Bianchi, S. Genovesi and A. Monorchio, "Fast optimization of ultra-broadband antennas with distributed matching networks," *IEEE Ant. Wireless Propag. Lett.*, vol. 13, pp. 642-645, 2014.

[36] L. Sang, S. Wu, G. Liu, J. Wang, and W. Huang, "High-gain UWB Vivaldi antenna loaded with reconfigurable 3-D phase adjusting unit lens," *IEEE Ant. Wireless Propag. Lett.*, vol. 19, no. 2, pp. 322-326, 2020.

[37] Y. Liu, M. Li, R. L. Haupt, and Y. J. Guo, "Synthesizing shaped power patterns for linear and planar antenna arrays including mutual coupling by refined joint rotation/phase optimization," *IEEE Trans. Ant. Propag.*, vol. 68, no. 6, pp. 4648-4657, 2020.

[38] Z. Niu, H. Zhang, Q. Chen, and T. Zhong, "Isolation enhancement in closely coupled dual-band MIMO patch antennas," *IEEE Ant. Wireless Propag. Lett.*, vol. 18, no. 8, pp. 1686-1690, 2019.

[39] A. Lalbakhsh, M. U. Afzal, K. P. Esselle, and S. L. Smith, "Wideband near-field correction of a Fabry–Perot resonator antenna," *IEEE Trans. Ant. Propag.*, vol. 67, no. 3, pp. 1975-1980, 2019.

[40] A. Lalbakhsh, M. U. Afzal, K. P. Esselle, and S. Smith, "Design of an artificial magnetic conductor surface using an evolutionary algorithm," *Int. Conf. Electr. Adv. Appl. (ICEAA)*, pp. 885-887, 2017.

[41] 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," in *IEEE Ant. Wireless Propag. Lett.*, vol. 16, pp. 912-915, 2017.

[42] A. Lalbakhsh, M. U. Afzal, and K. P. Esselle, "Simulation-driven particle swarm optimization of spatial phase shifters," *Int. Conf. Electr. Adv. Appl. (ICEAA)*, 2016.

[43] P. Lalbakhsh, B. Zaeri, and A. Lalbakhsh, "An improved model of ant colony optimization using a novel pheromone update strategy," *IEICE Trans. Inf. Syst.*, vol. E96, no. 11, pp. 2309-2318, 2013.

[44] R. Lehmensiek and D. I. L. de Villiers, "Optimization of log-periodic dipole array antennas for wideband omnidirectional radiation," *IEEE Trans. Ant. Propag.*, vol. 63, no. 8, pp. 3714-3718, 2015.

[45] S. Koziel and A. Pietrenko-Dabrowska, "Fast multi-objective optimization of antenna structures by means of data-driven surrogates and dimensionality reduction," *IEEE Access*, vol. 8, pp. 183300-183311, 2020.

[46] E. Hassan, D. Noreland, R. Augustine, E. Wadbro, and M. Berggren, "Topology optimization of planar antennas for wideband near-field coupling," *IEEE Trans. Ant. Prop.*, vol. 63, no. 9, pp. 4208-4213, 2015.

[47] S. Koziel and A. Pietrenko-Dabrowska, "Expedited optimization of antenna input characteristics with adaptive Broyden updates," *Eng. Comp.*, vol. 37, no. 3, 2019.

[48] S. Koziel and A. Pietrenko-Dabrowska, "Expedited feature-based quasi-global optimization of multi-band antennas with Jacobian variability tracking," *IEEE Access*, vol. 8, pp. 83907-83915, 2020.

[49] F. Arndt, "WASP-NET®: Recent advances in fast full 3D EM CAD of waveguide feeds and aperture antennas," *IEEE Int. Symp. Ant. Propag., APS-URSI*, Spokane, WA, pp. 2724-2727, 2011.

[50] B. B. Tierney and A. Grbic, "Designing anisotropic, inhomogeneous metamaterial devices through optimization," *IEEE Trans. Ant. Prop.*, vol. 67, no. 2, pp. 998-1009, 2019.

[51] 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.

[52] S. Koziel and S. Ogurtsov, *Antenna design by simulation-driven optimization. Surrogate-based approach*, Springer, New York, 2014.

[53] S. Koziel and L. Leifsson, "Simulation-driven design by knowledge-based response correction techniques," Springer, 2016.

[54] 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.

[55] 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.

[56] 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.

[57] S. Koziel, S. Ogurtsov, W. Zieniutycz, and L. Sorokosz, "Expedited design of microstrip antenna subarrays using surrogate-based optimization," *IEEE Ant. Wireless Prop. Lett.*, vol. 13, pp. 635-638, 2014.

[58] I. A. Baratta, C. B. de Andrade, R. R. de Assis, and E. J. Silva, "Infinitesimal dipole model using space mapping optimization for antenna placement," *IEEE Ant. Wireless Prop. Lett.*, vol. 17, no. 1, pp. 17-20, 2018.

[59] 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.

[60] Y. Su, J. Li, Z. Fan, and R. Chen, „Shaping optimization of double reflector antenna based on manifold mapping," *Int. Applied Comp. Electromagnetics Soc. Symp. (ACES)*, Suzhou, China, pp. 1-2, 2017.

[61] C. Zhang, F. Feng, V. Gongal-Reddy, Q. J. Zhang, and J. W. Bandler, "Cognition-driven formulation of space mapping for equal-ripple optimization of microwave filters," in *IEEE Trans. Microwave Theory Techn.*, vol. 63, no. 7, pp. 2154-2165, 2015.

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

[63] J. Xu, M. Li, and R. Chen, "Space mapping optimisation of 2D array elements arrangement to reduce the radar cross-scattering," *IET Microwaves Ant. Prop.*, vol. 11, no. 11, pp. 1578-1582, 2017.

[64] S. Marelli and B. Sudret, „UQLab: a framework for uncertainty quantification in Matlab," in *The 2nd Int. Conf. on Vulnerability and Risk Analysis and Management (ICVRAM 2014)*, University of London, UK, July 13-15, pp. 2554-2563, 2014.

[65] 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.

[66] B. Liu, S. Koziel, and N. Ali, "SADEA-II: a generalized method for efficient global optimization of antenna design," *J. Comp. Design Eng.*, vol. 4, no. 2, pp. 86-97, 2017.

[67] D. K. Lim, D. K. Woo, H. K. Yeo, S. Y. Jung, J. S. Ro, and H. K. Jung, "A novel surrogate-assisted multi-objective optimization algorithm for an electromagnetic machine design," *IEEE Trans. Magn.*, vol. 51, no. 3, paper 8200804, 2015.

[68] S. Koziel, A. Bekasiewicz, I. Couckuyt, and T. Dhaene, "Efficient multi-objective simulation-driven antenna design using co-kriging," *IEEE Trans. Antennas Prop.*, vol. 62, no. 11, pp. 5900-5905, 2014.

[69] 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.

[70] 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.

[71] J.P. Jacobs and S. Koziel, "Two-stage framework for efficient Gaussian process modeling of antenna input characteristics," *IEEE Trans. Antennas Prop.*, vol. 62, no. 2, pp. 706-713, 2014.

[72] 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.

[73] 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.

[74] J. Du and C. Roblin, "Statistical modeling of disturbed antennas based on the polynomial chaos expansion," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, p. 1843-1847, 2017.

[75] 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.

[76] U. Ullah, S. Koziel, and I.B. Mabrouk, "Rapid re-design and bandwidth/size trade-offs for compact wideband circular polarization antennas using inverse surrogates and fast EM-based parameter tuning," *IEEE Trans. Ant. Prop.*, vol. 68, no. 1, pp. 81-89, 2019.

[77] M.A. Haq and S. Koziel, "Feed line alterations for optimization-based design of compact super wideband MIMO antennas in parallel configuration," *IEEE Ant. Wireless Prop. Lett.*, vol. 18, no. 10, pp. 1986-1990, 2019.

[78] S. Koziel, "Objective relaxation algorithm for reliable simulation-driven size reduction of antenna structure," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, no. 1, pp. 1949-1952, 2017.

[79] D.O. Johannesson, S. Koziel, and A. Bekasiewicz, "EM-driven constrained miniaturization of antennas using adaptive in-band reflection acceptance threshold," *Int. J. Numerical Modeling*, vol. 32, no. 2, 2019.

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

[81] S. Koziel and A. Pietrenko-Dabrowska, "Variable-fidelity simulation models and sparse gradient updates for cost-efficient optimization of compact antenna input characteristics," *Sensors*, vol. 19, no. 8, 2019.

[82] 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.

[83] 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.

[84] 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.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3111285, IEEE Transactions on Antennas and Propagation

SLAWOMIR KOZIEL received the M.Sc. and Ph.D. degrees in electronic engineering from Gdansk University of Technology, Poland, in 1995 and 2000, respectively. He also received the M.Sc. degrees in theoretical physics and in mathematics, in 2000 and 2002, respectively, as well as the PhD in mathematics in 2003, from the University of Gdansk, Poland. He is currently a Professor with the Department of Engineering, Reykjavik University, Iceland. His research interests include CAD and modeling of microwave and antenna structures, simulation-driven design, surrogate-based optimization, space mapping, circuit theory, analog signal processing, evolutionary computation and numerical analysis.

ANNA PIETRENKO-DABROWSKA received the M.Sc. and Ph.D. degrees in electronic engineering from Gdansk University of Technology, Poland, in 1998 and 2007, respectively. Currently, she is an Associate Professor with Gdansk University of Technology, Poland. Her research interests include simulation-driven design, design optimization, control theory, modeling of microwave and antenna structures, numerical analysis.