Shape Synthesis Based on Topology Sensitivity

Miloslav Capek, Senior Member, IEEE, Lukas Jelinek, and Mats Gustafsson, Senior Member, IEEE

Abstract—A method evaluating the sensitivity of a given parameter to topological changes is proposed within the method of moments paradigm. The basis functions are used as degrees of freedom which, when compared to the classical pixeling technique, provide important advantages, one of them being impedance matrix inversion free evaluation of the sensitivity. The devised procedure utilizes port modes and their superposition which, together with only a single evaluation of all matrix operators, leads to a computationally effective procedure. The proposed method is approximately one hundred times faster than contemporary approaches, which allows the investigation of the sensitivity and the modification of shapes in real-time. The method is compared with known approaches and its validity and effectiveness is verified using a series of examples. The procedure can be implemented in up-to-date EM simulators in a straightforward manner. It is shown that the iterative repetition of the topology sensitivity evaluation can be used for gradient-based topology synthesis. This technique can also be employed as a local step in global optimizers.

Index Terms—Antennas, optimization methods, structural topology design, numerical methods, shape sensitivity analysis.

I. INTRODUCTION

ANTENNA synthesis, in its simplest form, attempts to extract a specific shape of the radiator and its feeding with respect to predetermined behavior, which is specified in terms of antenna metrics, such as radiation pattern, antenna bandwidth, efficiency, or input impedance. Often, additional constraints on form factor or maximal area occupied by the radiator exist [1], [2]. The problem of the synthesis remains unsolved, despite the remarkable progress and many achievements made in the field of computer science [3], numerical methods [4], and optimization theory [5], [6]. The final resolution is also immune from the dramatic growth of computation power. It can be shown that, in its entirety, the problem of synthesis originates from integer programming [7], [8], a discipline commonly understood as one of NP-hard complexity [9].

To date, attempts on synthesis have been based on a mixture of empirical understanding and the utilization of robust optimization algorithms, typically of heuristic nature. Emphasizing empirical knowledge can significantly reduce the number of free parameters, e.g., restricting the potential solution to a meanderline antenna, only the number of meanders, aspect ratio, and, potentially, the width of the strip must be found. On the other hand, there is a serious risk that such a parametrization will be unable to find the optimal (or close to the optimal) solution. Therefore, an alternative is to equally divide the solution space without exerting the designers preference into small domains – pieces of metal known here as pixels [10]. While popular and well-suited for, e.g., evaluation via powerful genetic algorithms [11], [12], this approach is known to suffer from complexity explosion [13].

Topology (shape) optimization [14] renders itself important in an electrically small region, mainly due to the existence of strict bounds on its operation, mutually conflicting design criteria and limited size [15]. It is the limited electrical size which also improves the number of potential solutions (considering the fixed number of discretization elements per wavelength), albeit, at the expense of available degrees-of-freedom (DOF), which can create problems with premature convergence into local minima.

However, no matter which approach is followed, any additional knowledge in the behavior of the optimized function(s) is helpful [16]. In this respect, the inspiration mostly stems from mechanics [17] in an attempt to adopt the concept of topology sensitivity and the related topological derivative [18]–[21]. Unfortunately, the incorporation of topology sensitivity is commonly aligned with the introduction of surface modulated material parameters, including conductivity or permittivity [18] into the finite element method [22]. Recent works also deal with method of moments (MoM) [23] formulation [24], [25]. However, in all cases, it is the designer who has to decide which threshold of the optimized material parameter will be chosen to distinguish the conducting part and free space [21], i.e., to tell the difference between solids and voids. This is a heuristic decision, producing non-unique, always sub-optimal results, which are often relatively far from the performance of optimal material distribution before rounding.

Thus, this paper proposes an efficient way how to easily incorporate the evaluation of topology sensitivity directly into MoM which is one of the most popular methods for solving the radiation (open) problems. The key step is the utilization of the Sherman-Morrison-Woodbury formula [26], which has already shown its usefulness in other applications dealing with compression of electrically large problems [27], [28] via the characteristic basis functions method [29]. All derivation steps are done with operators and quantities discretized into their algebraic forms, i.e., with matrices and vectors. Therefore, the implementation is straightforward. A novel form of the pixeling paradigm is proposed, where basis functions are removed rather than small pieces of metal. This procedure is justified by full-wave verification and explained both from

Manuscript received August 9, 2018, 2018; revised August 9, 2018, 2018.
This work was supported by the Swedish Foundation for Strategic Research (SSF) under the Program Applied Mathematics and the Project Complex Analysis and Convex Optimization for EM Design.
M. Capek and L. Jelinek are with the Department of Electromagnetic Field, Faculty of Electrical Engineering, Czech Technical University in Prague, 166 27 Prague, Czech Republic (e-mail: miloslav.capek@fel.cvut.cz, lukas.jelinek@fel.cvut.cz).
M. Gustafsson is with the Department of Electrical and Information Technology, Lund University, 221 00 Lund, Sweden (e-mail: mats.gustafsson@eit.lth.se).
a geometrical and physical point of view. It is shown that within this paradigm, the original number of degrees of freedom, given by the discretization and application of local basis functions, is preserved. No modification of the MoM kernel is needed and, significantly, the calculation does not require matrix inversion nor an adjoint sensitivity parameter. Consequently, its determination is extremely fast and makes it possible to utilize gradient search in every step of the heuristic optimizer. In such a way, the global optimizer moves in a solution space consisting of local minima only. As compared to a naive (gradient-based) greedy algorithm [3] utilizing matrix inversion, the proposed technique produces the same results with a speed-up of hundreds for mid-sized problems.

The paper is organized as follows. The key methods used throughout the paper are briefly reviewed in Section II, an attempt is made to highlight the aspects and interpretations necessary for the derivation of the basis function-removal technique in Section III where a new technique is introduced, explained and compared with the classical pixeling technique, and validated against full-wave results calculated with conventional procedure. It is shown in Section IV that the derived formulas can be easily adopted towards the evaluation of topology sensitivity through a deterministic gradient greedy algorithm. Consequently, its determination is extremely fast and makes it possible to utilize gradient search in every step of the heuristic optimizer. In such a way, the global optimizer moves in a solution space consisting of local minima only. As compared to a naive (gradient-based) greedy algorithm [3] utilizing matrix inversion, the proposed technique produces the same results with a speed-up of hundreds for mid-sized problems.

The paper is organized as follows. The key methods used throughout the paper are briefly reviewed in Section II, an attempt is made to highlight the aspects and interpretations necessary for the derivation of the basis function-removal technique in Section III where a new technique is introduced, explained and compared with the classical pixeling technique, and validated against full-wave results calculated with conventional procedure. It is shown in Section IV that the derived formulas can be easily adopted towards the evaluation of topology sensitivity through a deterministic gradient greedy algorithm. Consequently, its determination is extremely fast and makes it possible to utilize gradient search in every step of the heuristic optimizer. In such a way, the global optimizer moves in a solution space consisting of local minima only. As compared to a naive (gradient-based) greedy algorithm [3] utilizing matrix inversion, the proposed technique produces the same results with a speed-up of hundreds for mid-sized problems.

The paper is organized as follows. The key methods used throughout the paper are briefly reviewed in Section II, an attempt is made to highlight the aspects and interpretations necessary for the derivation of the basis function-removal technique in Section III where a new technique is introduced, explained and compared with the classical pixeling technique, and validated against full-wave results calculated with conventional procedure. It is shown in Section IV that the derived formulas can be easily adopted towards the evaluation of topology sensitivity through a deterministic gradient greedy algorithm. Consequently, its determination is extremely fast and makes it possible to utilize gradient search in every step of the heuristic optimizer. In such a way, the global optimizer moves in a solution space consisting of local minima only. As compared to a naive (gradient-based) greedy algorithm [3] utilizing matrix inversion, the proposed technique produces the same results with a speed-up of hundreds for mid-sized problems.

II. METHOD OF MOMENTS AND PIXELING TECHNIQUE

The classical MoM procedure [23] applied on the electric field integral equation (EFIE) [30] is briefly introduced in this section, including lumped elements and resistive sheets [31]. All quantities are considered within the time-harmonic steady state [30]. This framework is utilized for the analysis and synthesis of optimal radiators. The second part of this section reviews the classical pixeling technique [11], extensively used in topology optimization [10].

A. Electric Field Integral Equation

Let us assume an arbitrarily shaped surface $\Omega$ with surface resistance $R_\Omega (r')$, see Fig. 1a. The surface current density $J (r')$, $r' \in \Omega$ flowing on this surface is related to the total electric field via the EFIE [23]

$$\hat{n} \times E = \hat{n} \times (E_i + E_s) = R_\Omega \hat{n} \times J,$$

(1)

where $E_i (r)$ is the incident field and

$$E_s (r) = -j k Z_0 \int_\Omega G (r, r') \cdot J (r') \, dS'$$

(2)

is the scattered field with $k$ denoting the wavenumber, $Z_0$ denoting the free-space impedance and $G$ denoting the free-space dyadic Green’s function defined as [30]

$$G (r, r') = \left(1 + \frac{\nabla \nabla}{k^2}\right) e^{-j k |r-r'|}/4\pi |r-r'|,$$

(3)

Symbol $I$ abbreviates the identity dyadic.

The formulation (1) operates with continuous quantities and is, with the exception of the canonical examples, analytically unsolvable. Its solution is, therefore, commonly approached via MoM [23].

B. Discretization of EFIE via MoM

The MoM approach assumes that a continuous geometry $\Omega$ is discretized into surface patches of the prescribed shape. Throughout this paper, a discretization into triangles $\Omega \rightarrow \Omega_T = \bigcup_{i=1}^T T_i $ is used, where $T$ denotes the number of triangles, see Fig. 1. Within this discretization a set of basis functions $\{\psi_n (r')\}$ is used to expand the surface current density as

$$J (r) \approx \sum_{n=1}^N I_n \psi_n (r').$$

(4)

The subsequent application of Galerkin’s testing procedure [32], [33] recasts the EFIE (1) into its matrix form

$$(Z_G + Z_L + R_s \Psi) I = Z I = V,$$

(5)

where $Z_G$ is the impedance matrix characterizing the perfect electric conductor (PEC) surface, $Z_L$ is a diagonal matrix of lumped elements [34], $R_s$ is surface resistivity, and $\Psi$ is the basis function Gram matrix [35]. The matrices of lumped elements $Z_L$ and surface resistivity $R_s \Psi$ are discretized counterparts to the right-hand side of (1), i.e., $R_\Omega (r') = \text{Re} \{Z_L\} + R_s \Psi$, and are introduced in order to deal with the pixeling methods in Sections II-C and III-A appropriately.

Explicitly, the aforementioned matrix operators are defined as

$$Z_{G, mn} = j k Z_0 \int_\Omega \int_\Omega \psi_m (r) \cdot G (r, r') \cdot \psi_n (r') \, dS \, dS'$$

(6)

and

$$\Psi_{mn} = \int_\Omega \psi_m (r) \cdot \psi_n (r) \, dS,$$

(7)

where, without loss of generality, the Rao-Wilton-Glisson (RWG) basis functions are employed in this paper [36].

Using (5), the system described by (1) and schematically depicted in Fig. 1a is reduced to a system having a finite,
however typically large, number of DOF, see Fig. 1b. All subsequent operations, e.g., calculation of current density or evaluation of far-field radiation pattern, are performed over this discretized system. The matrix $\mathbf{Z}$ is, therefore, often called the impedance matrix.

C. Triangle Removal (Pixeling)

Let us start with an arbitrarily shaped radiator made of a PEC, i.e., $\mathbf{Z}_L = \mathbf{0}$, $R_u = 0$. Considering a problem of topology optimization, discretization in Fig. 1b suggests the potential domains to be used as optimized variables. Traditionally, the triangles [37] (or rectangles [10]) are utilized as pixels whose presence is subject to a boolean (integer) optimization. A fixed discretization $\Omega_T$ is used during the entire optimization process which tremendously accelerates the solution as the discretization and evaluation of matrix operators are done only once. On the other hand, it restricts the achievable details.

As a simple use-case, imagine a situation depicted in Fig. 2. Triangle $T_1$, $t \in T = \{1\}$ and three associated basis functions $\psi_n$, $n \in \mathcal{E} = \{1, 2, 3\}$, are removed, i.e., a triangular pixel (made of PEC) is replaced by a triangular hole (vacuum). This is equivalent to the removal of the three columns and rows from the impedance matrix corresponding to basis functions $\mathcal{E}$. This operation can be formally written as

$$\hat{\mathbf{Z}} = \mathbf{C}_T^T \mathbf{Z} \mathbf{C}_T,$$

where the rectangular projection matrix $\mathbf{C}_T$ is defined as

$$\mathbf{C}_{T,nn} = \begin{cases} 
1 & \Leftrightarrow \psi_n (r) \cap T_t = \emptyset, \\
0 & \text{otherwise,}
\end{cases}$$

and where all its columns containing only zeros are removed, i.e., $\mathbf{C}_T \in \mathbb{R}^{N \times (N - D)}$, where $D$ is the number of removed basis functions, $D = |\mathcal{E}|$. The subscript “$T$” of the system matrix in (8) reflects that the triangular pixels from set $T$ were removed and the hat symbol indicates that the size of the matrix was reduced. Vectors of the expansion coefficients and excitation coefficients are reduced analogously, i.e.,

$$\hat{\mathbf{I}} = \mathbf{C}_T^T \mathbf{I},$$

$$\hat{\mathbf{V}} = \mathbf{C}_T^T \mathbf{V}.$$ 

Apart from the system matrix reduction, the triangle removal sketched above also allows for another interpretation in which the system matrix $\mathbf{Z}_G$ retains its original size, i.e., no triangle is removed, though the surface resistivity $R_u$ of the underlying triangle is changed to

$$R_u (r) = R_\infty \Leftrightarrow r \in T_1,$$

where we can anticipate $R_\infty \rightarrow \infty$. It will be later shown that inversion of matrix $\mathbf{Z}$ with surface resistivity (12) can be formally performed via the Sherman-Morrison-Woodbury formula. However, before that, we present an alternative way to modify the topology of structure $\Omega_T$.

III. REMOVAL OF BASIS FUNCTIONS

A detailed study of Fig. 1b reveals that, apart from the triangular domains, there is another possible partitioning to be used in conjunction with the pixeling technique. The basis functions form a natural set of DOF whose presence can be optimized. Their utilization simplifies the pixeling technique and brings computational benefits, but also introduces some questions regarding the physical interpretation.

A. Performing Basis Functions Removal

Following the introductory discussion from the previous section, let us start with a small perturbation of system $\Omega_T$, this time solely through the removal of set $\mathcal{E}$ of the basis functions. Using an analogy with (12), we claim that the basis function removal corresponds to the introduction of lumped elements

$$\mathbf{Z}_{L,n} = R_\infty, \quad \Leftrightarrow \quad n \in \mathcal{E},$$

or, alternatively, to

$$\mathbf{Z}_L = \mathbf{C}_E R_\infty \mathbf{C}_E^T,$$

where the projection matrix $\mathbf{C}_E$ is defined as

$$\mathbf{C}_{E,n} = \begin{cases} 
0 & \Leftrightarrow n \notin \mathcal{E}, \\
1 & \text{otherwise,}
\end{cases}$$

and where all columns of $\mathbf{C}_E$ containing only zeros are removed. Consequently, system matrix $\mathbf{Z}$ becomes

$$\hat{\mathbf{Z}} = \mathbf{Z}_G + \mathbf{Z}_L = \mathbf{Z}_G + \mathbf{C}_E R_\infty \mathbf{C}_E^T.$$ 

The above-mentioned removal procedure for $n \in \mathcal{E} = \{1\}$ is illustrated in Fig. 3. As compared to the classical pixeling, the radiator’s body remains unchanged, however, the number of DOF of the system is reduced by one. Figure 3 also shows that the removal of a single basis function $n \in \mathcal{E} = \{1\}$ leads to admittance matrix $\hat{\mathbf{Y}} = \hat{\mathbf{Z}}^{-1}$ in which row $n$ and column $n$ have been zeroed (no current flow is associated with the removed basis function) and in which the other entries have been modified according to the algorithm described in Section IV-B.

It is important to realize that matrix $\mathbf{C}_T$ is fundamentally different from matrix $\mathbf{C}_E$ in that matrix $\mathbf{C}_T$ denotes basis functions to be kept, while matrix $\mathbf{C}_E$ denotes basis functions to be removed. However, the triangle removal procedure described in Sec. II-C can also be described as removal of
certain basis functions via an appropriate $\mathbf{C}_E$ matrix and, with respect to later developments, we will solely use $\mathbf{C}_E$ matrices for any kind of pixeling.

The physical realization of a single basis function removal is more intricate than the removal of a triangle. The problem is that for a two-dimensional setup, the addition of a general lumped load via $\mathbf{Z}_L$ does not have a simple interpretation in terms of the physical modification of the underlying surface. Nevertheless, the addition of significantly high lumped resistance can be seen as carving a slot into the original resistive sheet alongside the edge associated with the removed basis function. This forbids electric current flow across the slot and the corresponding basis function. The displacement current still flows across the slot via the slot capacitance.

**B. Numerical Verification**

The basis function removal paradigm derived in the previous section deserves to be numerically verified in a careful manner before its practical utilization. It remains to be answered which slot thickness in the resistive sheet corresponds to the removal of a basis function in the numerical model. Here, the possible issues are here mostly related to the charge accumulation on the slot periphery, i.e., to the correct model of the capacitance of the slot.

The structure depicted in Fig. 4 is used for a qualitative answer. The setup is a capacitor made mostly of an open-ended coplanar strip line of length $\ell_0$, with a strip width of $w$ and the distance between the strips equal to $g$. The left side of the structure shows an adapter with a delta gap feed [38] whose capacitance is negligible with respect to the strip line.

The normalized per-unit-length capacitance $\hat{C}_\ell = C_\ell/\varepsilon_0$ of the coplanar strips is known to be [40]

$$\hat{C}_\ell = \frac{K(\sqrt{1 - \kappa^2})}{K(\kappa)}, \quad (17)$$

with

$$\kappa = \frac{g}{g + 2w}, \quad (18)$$

and $K(\cdot)$ being the complete elliptic integral of the first kind. Its sweep over parameter $g/w$ is depicted via the solid line in Fig. 5 and is compared with the normalized capacitance $\hat{C}_n$ calculated from the input reactance $X_{in}$ seen by the delta gap in the MoM simulation of structure from Fig. 4 for a case of $g/w = 0$, i.e., for a closed gap, in which, however, there are no basis functions carrying current across the gap between the strips. This case corresponds to the basis function removal paradigm described in Sec. III-A. Several mesh densities were investigated, see horizontal dashed lines in Fig. 5. The crossing points of these horizontals, with a solid line representing the analytical result, shows that removal of basis functions corresponds to a physical gap of certain non-zero thickness. The physical interpretation of the basis function removal is, thus, explicitly dependent on the mesh density of the discretized model and the exact correspondence of the basis function removal and slot carving into a metal should not be expected. As a rule of thumb it is advised to choose the mesh density so that triangle edge-lengths do not exceed the width of the physical gap.

**C. Triangle Pixeling vs. Basis Function Removal**

A comparison of Figs. 2a and 3a reveals that the removal of triangles changes the geometry of the shape $\Omega_T$, however, removal of basis functions only changes the topology\textsuperscript{1} of the

\textsuperscript{1}With topology, we mean the connectivity of the region $\Omega_T$ which describes both the shape of the outline as well as holes in the design domain.
studied object. Hence, a small loop is created in Fig. 3a, while such a loop cannot be created in Fig. 2a — if any triangle is removed, a dipole-type structure is created. In other words, classical pixeling from Section II-C does not treat all DOF of the discretized system $\Omega_T$ independently.

As an example, a rectangular area of side ratio $1:2$ is studied in Fig. 6. The optimized region is triangularized based on given granularity (20 triangles per longer side, resulting in $T = 100$ triangles and $N = 135$ basis functions). Let us anticipate that the optimal shape is a meanderline, as is common for the minimal radiation Q-factor of a single-fed antenna [41]. Potential solutions of the highest possible number of meanders are depicted in Fig. 6 and it can be seen that the more compact antenna (therefore potentially electrically smaller) is achieved via the removal of basis functions.

IV. EFFICIENT EVALUATION OF SMALL TOPOLOGICAL PERTURBATIONS

When studying design optimality, it is often desired to know how the design parameter $p$ changes if the geometry is slightly perturbed$^2$. This parameter can, e.g., be an antenna metric, such as input impedance, fractional bandwidth or directivity, and will be a function of the current flowing on the antenna $p(\mathbf{I})$.

In order to evaluate the current $\mathbf{I}$ on the perturbed structure, an inversion of the perturbed system matrix $\mathbf{Z}$ is needed to evaluate

$$\mathbf{I} = \mathbf{Z}^{-1} \mathbf{V} = \mathbf{YV},$$  

(20)

where

$$\mathbf{Z} = \mathbf{Z}_G + \mathbf{C}_G R_\infty \mathbf{C}_G^T$$  

(21)

and where, for the sake of simplicity, but without loss of generality, we have omitted potential surface resistance $R_n$ introduced in (5). In order to evaluate smallest $N$ perturbations of region $\Omega_T$ one by one, we have to perform (20) $N$-times which is computationally demanding.

In this section, we will show how to accelerate the evaluation of (20), and how to completely avoid the inversion of (21) if just one basis function is to be removed. A subsequent section will define topology sensitivity $\mathbf{r}(p, \Omega_T)$ of parameter $p$ for discretized system $\Omega_T$ and derive its evaluation as a simple matrix product, using the original admittance matrix $\mathbf{Y}_G$.

$^2$In the discretized case, slightly means removing a few basis functions.

A. Efficient Inversion of the Perturbed System

The perturbation (21) is a low rank correction to the original matrix $\mathbf{Z}_G$ and the inversion of $\mathbf{Z}$ can advantageously be approached [42] via the Sherman-Morrison-Woodbury formula [26] which reads

$$(\mathbf{A} + \mathbf{EBF})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{E} (\mathbf{B}^{-1} + \mathbf{FA}^{-1} \mathbf{E})^{-1} \mathbf{FA}^{-1},$$  

(22)

for $\mathbf{A} \in \mathbb{C}^{N \times N}$, $\mathbf{B} \in \mathbb{C}^{D \times D}$. Its application to (21) readily gives

$$\mathbf{Y} = \mathbf{Z}^{-1} = \mathbf{Z}_G^{-1} - \mathbf{Z}_G^{-1} \mathbf{C} \left( \frac{1}{R_{\infty}} \mathbf{I}_D + \mathbf{C}^T \mathbf{Z}_G^{-1} \mathbf{C} \right)^{-1} \mathbf{C}^T \mathbf{Z}_G^{-1},$$  

(23)

where $\mathbf{I}_D$ is an identity matrix of size $D \times D$. Substituting $\mathbf{Y}_G = \mathbf{Z}_G^{-1}$ and using the limit $R_\infty \to \infty$, the final formula reads

$$\mathbf{Y} = \mathbf{Y}_G - \mathbf{Y}_G \mathbf{C} (\mathbf{C}^T \mathbf{Y}_G \mathbf{C})^{-1} \mathbf{C}^T \mathbf{Y}_G.$$  

(24)

This formula still requires a matrix inversion, however, the admittance matrix $\mathbf{Y}_G$ can be calculated only once at the beginning and, then, the inversion of only a small $D \times D$ matrix has to be done separately. Below, it is shown how to further reduce the formula (24) using successive basis function removals. Notice that considering the implementation of (24), e.g., in MATLAB [43], the outer matrix multiplications are reduced to computationally cheap indexing since matrix $\mathbf{C}$ contains only a single non-zero element in each row/column. Note also that this methodology can, in general, be used in connection with arbitrary reactive loading [44].

B. Removal of a Single Basis Function

If a single basis function is removed ($D = 1$), the formula (24) can be further simplified to

$$\mathbf{Y} = \mathbf{Y}_G - \mathbf{Y}_G \mathbf{n} \frac{\mathbf{Y}_G \mathbf{n}^T}{\mathbf{Y}_{nn}},$$  

(25)

where $\mathbf{Y}_{G,n}$ is the $n$-th column of the admittance matrix $\mathbf{Y}_G$ and $\mathbf{Y}_{nn}$ is the $n$-th diagonal element of the admittance matrix $\mathbf{Y}_G$. One can easily verify that (25) modifies entries in the admittance matrix $\mathbf{Y}$ so that the current flowing through the $n$-th basis function is zeroed irrespective of the impressed voltage and all interactions with this basis function are eliminated (the $n$-th row and the $n$-th column of matrix $\mathbf{Y}$ is zeroed).

C. Topology Changes with Feeding Included

Consider an initial structure $\Omega_T$ (with $N$ inner edges) fed by a delta gap source at the $f$-th basis function, i.e.,

$$\mathbf{V}_f = V_0 \begin{bmatrix} 0 & \cdots & l_f & \cdots & 0 \end{bmatrix}^T,$$  

(26)

where $V_0$ is the feeding voltage and $l_f$ is the length of the edge associated with the $f$-th basis function. This excitation generates a corresponding current vector (further on called “port mode” [44]), defined as

$$\mathbf{I}_f = \mathbf{Y}_G \mathbf{V}_f.$$  

(27)
Let us now investigate the removal of the $n$-th basis function. The perturbed port mode current $\mathbf{I}_{fn}$ (feeding at the $f$-th basis function, removal of the $n$-th basis function) can be evaluated using (27) and (25), explicitly

$$
\mathbf{I}_{fn} = \left( \mathbf{Y}_G - \frac{\mathbf{Y}_Gn \mathbf{Y}_G \mathbf{n}^T}{Y_{nn}} \right) \mathbf{V}_f = \left( \mathbf{Y}_G, f - \frac{\mathbf{Y}_fn}{Y_{nn}} \mathbf{Y}_G, n \right) \mathbf{V}_0, f_n
$$

(28)

$$
= \mathbf{I}_f - \left( \frac{l_1 l_2 f_n}{Y_{nn}} \right) V_{0n} \mathbf{Y}_G, n = \mathbf{I}_f + \zeta_{fn} \mathbf{I}_n,
$$

(29)

where

$$
\zeta_{ij} = -\frac{l_1 l_2}{l_j} \frac{Y_{ij}}{Y_{jj}} = -\frac{Z_{in,ij}}{Z_{in,in}},
$$

and where $Z_{in,ij}$ denotes the mutual impedance between the $i$-th and $j$-th basis functions.

Formula (28) shows that basis function removal in a single-feed scenario is equivalent to the linear combination of two port modes of the original structure, or, in other words, that the feed scenario is equivalent to the linear combination of two removals of all basis functions into a matrix $\mathbf{Z}$ which forces zero current through the $n$-th basis function.

A form of (28) allows the accumulation of individual removals of all basis functions into a matrix

$$
[\mathbf{I}_{fS}] = \left[ \mathbf{I}_f + \zeta_{f1} \mathbf{I}_1 \cdots \mathbf{I}_f + \zeta_{fN} \mathbf{I}_N \right],
$$

(30)

where $S = \{1, \ldots, f-1, f+1, \ldots, N\}$ denotes a set of basis functions to be removed (one at a time). Since $\mathbf{I}_{fS}$ is identically zero, it has been dropped from the set. For the sake of the compact notation we will only use $\mathbf{I}_{fS} \equiv [\mathbf{I}_{fS}]$ in the following text.

D. Topology Sensitivity

Within the MoM paradigm, a common electromagnetic metric, say $x$, can be defined via quotients of the quadratic form [23], [35], [45] as

$$
x(\mathbf{I}) = \frac{\mathbf{I}^H \mathbf{A} \mathbf{I}}{\mathbf{I}^H \mathbf{B} \mathbf{I}},
$$

(32)

where, generally, $\mathbf{A} = \sum_a \alpha_a \mathbf{A}_a$ and $\mathbf{B} = \sum_b \beta_b \mathbf{B}_b$. Examples of metric $x$ of interest for antenna problems are shown in Sec. V-A.

An immediate question in connection with topology changes is how much metric $x$ varies when individual basis functions are removed. The formulation (31) offers an extremely efficient answer. The effect of all individual single-basis function removals can be evaluated as

$$
x(\mathbf{I}_{fS}) = \text{diag}(\mathbf{I}_{fS}^H \mathbf{A} \mathbf{I}_{fS}) \odot \text{diag}(\mathbf{I}_{fS}^H \mathbf{B} \mathbf{I}_{fS}),
$$

(33)

where the symbol $\odot$ denotes Hadamard division [46]. It is also plausible to define a topology derivative [47] as

$$
\tau_{fS}(x, \Omega_T) = x(\mathbf{I}_{fS}) - x(\mathbf{I}_f),
$$

(34)

which measures the sensitivity of the studied metric with regard to small topology changes. It is important to stress that using (31) and (33), the topology derivative is calculated via matrix products between $N \times N$ and $N \times |\mathcal{E}|$ matrices, while, classically, it would be evaluated using $|\mathcal{E}|$ matrix inversions (of size $N \times N$) followed by $|\mathcal{E}|$ substitutions into (32). The computational effort is thus tremendously reduced.

V. EXAMPLES

In this section, the effectiveness of the proposed scheme is demonstrated on examples covering the investigation of topology sensitivity with respect to selected antenna parameters and fast shape optimization based on the greedy (gradient) algorithm.

A. Metrics for Electrically Small Antennas

This subsection lists several power quantities used in antenna theory written as quadratic forms (32):

- The input impedance seen by a source connected to the the $f$-th basis function is defined as

$$
Z_{in}(\mathbf{I}) = \frac{\mathbf{I}^H \mathbf{Z} \mathbf{I}}{||\mathbf{I}_f||^2},
$$

(35)

where $\mathbf{I}_f$ is the source current at $f$-th basis function.

- The radiation $Q$-factor is defined as [34], [48]

$$
Q(\mathbf{I}) = \frac{\max \left\{ \mathbf{I}^H \mathbf{X}_m \mathbf{I}, \mathbf{I}^H \mathbf{X}_e \mathbf{I} \right\}}{\mathbf{I}^H \mathbf{R} \mathbf{G} \mathbf{I}},
$$

(36)

where the matrices $\mathbf{X}_{m/e}$ are defined as

$$
\mathbf{X}_{m/e} = \frac{1}{2} \left( \omega \frac{\partial \mathbf{X}}{\partial \omega} \pm \mathbf{X} \right)
$$

(37)

with a plus sign (+) for the magnetic energy matrix $\mathbf{X}_m$ and with a minus sign (−) for the electric energy matrix $\mathbf{X}_e$.

- The dissipation factor is defined as [49]

$$
\delta(\mathbf{R}_e, \mathbf{I}) = \mathbf{R}_e \frac{\mathbf{I}^H \mathbf{\Psi} \mathbf{I}}{\mathbf{I}^H \mathbf{R} \mathbf{G} \mathbf{I}},
$$

(38)

- The partial directivity is defined as

$$
D(\mathbf{\hat{e}}, \mathbf{\hat{r}}, \mathbf{I}) = 4\pi \frac{\mathbf{I}^H \mathbf{U}(\mathbf{\hat{e}}, \mathbf{\hat{r}}) \mathbf{I}}{\mathbf{I}^H \mathbf{R} \mathbf{G} \mathbf{I}},
$$

(39)

with $\mathbf{\hat{e}}$ denoting polarization, $\mathbf{\hat{r}}$ denoting direction, and with the far-field density matrix defined in [35]. Notice that other quantities, such as antenna gain $G$ [50] or radiation efficiency $\eta$ [50] can be introduced in the same way.

B. Topology Sensitivity – Input Reactance of a Dipole

Let us start with the investigation of a thin strip dipole of length $\ell$, width $w = \ell/40$, which is discretized into $N = 79$ basis functions. The dipole is fed at its center via a delta gap generator [38], see the red line in Fig. 7. The topology sensitivity of the absolute value of the input reactance, $|\tau_{fS}(||\mathcal{X}_{in}||, \Omega_T)|$, is studied first.

Three electrical sizes $k \ell = (3, 4, 6) \pi/4$ of the dipole are used to evaluate topology sensitivity with respect to resonance, i.e., $|\mathcal{X}_{in}| \rightarrow 0$. The first size, $k \ell = 3\pi/4$, corresponds to an electrically short dipole with operating frequency below the
first resonance. In order to approach the resonance, \textit{i.e.}, to lower the absolute value of input reactance, the dipole must be enlarged. This is, however, not possible within the paradigm of this paper. On the other hand, removal of any basis function increases the studied parameter, \textit{cf.} (34). This is confirmed by the inspection of Fig. 7 where all values of topology sensitivity are positive.

At the second electrical size, \( k_\ell = \pi = \lambda/2 \), the dipole is driven slightly above the first resonance [50]. Therefore, topology sensitivity starts to be negative close to the ends of the dipole arms suggesting that the removal of these “most negative” basis functions will decrease the absolute value of the input reactance the most, see (34). Basically the same commentary holds for the last electrical size, \( k_\ell = 3\pi/2 \), where a minima of topology sensitivity shows that the dipole should be considerably shortened.

C. Topology Sensitivity – Minimization of Q-factor

In this example, the same dipole and its excitation, as in the previous subsection, is assumed. The topology sensitivity of the Q-factor (36) is evaluated and normalized to its fundamental bound [51] which corresponds to a Q-factor of the optimal current density. Two cases are studied in Fig. 8. The first case (depicted by the solid black line) considers the original dipole with no basis function removed. The normalized Q-factor \( Q/Q_{lb} \) is minimal in the vicinity of the resonance, \( k_\ell \approx \pi \). The second case evaluates topology sensitivity \( \tau_{fs}(Q, \Omega) \) and adjusts the Q-factor performance of the dipole by the successive removal of basis functions with \( \max_{n \in S} \{ -\tau_{fs} \} \). This removal is performed as long as any basis function with a negative value of topology sensitivity \( \tau_{fs} \) exists. The values of topology sensitivity at the initial step of this procedure is shown in Fig. 9 for \( k_\ell = 4 \) (marker A) and \( k_\ell = 8 \) (marker B). For \( k_\ell = 4 \), we can see that no modifications are made since the topology sensitivity at all basis functions is positive. The corresponding current density is depicted as an inset in Fig. 8. For \( k_\ell = 8 \), however, four basis functions can successively be removed dramatically reducing the initial Q-factor, see the shape modifications (depicted by the solid black lines) and the initial topology sensitivity in Figs. 8 and 9, respectively.

The procedure for adjusting shape \( \Omega \) to get the minimum Q-factor can be understood as a discrete version of a gradient algorithm performing

\[
-\nabla Q(I) \approx Q(I_f) - Q(I_{fs})
\]

(40)

Since the steepest descent is followed in each iteration, the procedure is recognized as a greedy algorithm [9]. This technique does not ensure convergence to a global optimum (this is demonstrated further on in Section V-E), but its evaluation is appealing since its computational cost is minimal (compared to other contemporary techniques). Furthermore, in typical scenarios, the gradient method is capable of delivering results of high quality (see the next section), and, finally, for its fast convergence into a local minimum, it acts as a perfect candidate to be evaluated after each step of the global optimizer.

D. Shape Synthesis – Rectangular Plate

The effectiveness of the greedy algorithm based on the evaluation of topology sensitivity is further demonstrated on an example of a rectangular plate with an aspect ratio of 2 : 1, electrical size \( ka = 0.5 = \sqrt{5}k_\ell/4 \), discretized into \( N = 180 \) basis functions, see Fig. 10. The feeding edge is placed in the middle of the structure and highlighted by the red color. The Q-factor is minimized using \( \tau_{fe}(Q) \), successively removing the worst basis functions. At the beginning a vertical slot is cut to eliminate the shorts and to initially separate

\[
Q/Q_{lb}
\]

Fig. 7. Topology sensitivity \( \tau_{fe}(|X_{in}|) \) of a center-fed dipole \( \Omega_{dip} \), discretized into \( N = 79 \) basis functions, of three different electrical lengths \( k_\ell \). The parameter \( \xi \) denotes a coordinate along the dipole. The studied parameter is the absolute value of the input reactance, \( |X_{in}| \), seen by the delta gap, \textit{i.e.}, the ability to self-resonate. Negative values of topology sensitivity indicate the possibility to decrease the value of the input reactance \( |X_{in}| \). The dipole and its discretization is depicted true-to-scale above the graph. Data corresponding to the basis function associated with the feeding edge are not depicted as this basis function is not supposed to be removed.

Fig. 8. Radiation Q-factor of center-fed dipole \( \Omega_{dip} \), discretized into \( N = 79 \) basis functions, normalized to the lower Q-factor bound [51], \( Q_{lb} \). The solid black line corresponds to the original dipole \( \Omega_{dip} \), the dashed red line corresponds to the dipole modified via the greedy algorithm. Current densities and the resulting shapes are depicted for \( k_\ell = 4 \) (case A) and \( k_\ell = 8 \) (case B). The basis functions indicated by the red edges are fed, the basis functions indicated by the black edges are removed.
charge [52]. Then, the structure is perturbed further, adjusting the shape closer to its self-resonance. The final structure reaches $Q/Q_{lb} = 1.57$ and a resulting current density is depicted in Fig. 11.

The final modification involves the evaluation of the Q-factor for 10332 different shapes in $I = 71$ iterations, reaching a local minimum. Thanks to the formulation based on (28) and (33), the entire calculation takes $3\frac{43}{100}$ s. In order to provide a comparison, the same shape modifications are repeated via classical pixel removal based on repetitive matrix inversion and the calculation took 10 s.

Fig. 9. Topology sensitivity $\tau_{fE}(Q)$ of a center-fed dipole $\Omega_{dip}$, discretized into $N = 79$ basis functions, of two different electrical lengths $k\ell$. The parameter $\xi$ denotes a coordinate along the dipole. The studied parameter is the radiation Q-factor, $Q$. The negative numbers indicate the possibility to improve (lower) the value of the Q-factor. Data corresponding to the basis function associated with the feeding edge are not depicted as this basis function is not supposed to be removed.

Fig. 10. Topology sensitivity $\tau_{fE}(Q)$ of a rectangular plate with aspect ratio 2 : 1, electrical size $ka = 0.5$, discretized into $N = 180$ basis functions. Delta gap excitation is placed at the edge highlighted by the red color (horizontally centered, in upper row). Pink edges correspond to a basis function with negative sensitivity, i.e., their removal decreases the value of the Q-factor. The green edges correspond to a basis function with positivity sensitivity, i.e., they should be kept.

$E$. Shape Synthesis – Spherical Shell

The next example reveals a well-known shortage of gradient-based methods – their convergence into the nearest local minimum – where the algorithm remains trapped. Let us consider a spherical shell of electrical size $ka = 0.5$, discretized into $N = 900$ basis functions. One edge is fed by a delta gap (the exact position is irrelevant thanks to the symmetry of the spherical shell). The Q-factor is once again minimized and normalized to the fundamental bound $Q_{lb}$, which is realized by a resonant combination of dominant TM and TE spherical harmonics [53], [54].

The greedy removal of basis function ran for 380 iterations, 270129 antenna candidates were explored, and the resulting normalized Q-factor is $Q/Q_{lb} = 1.51$, see Table I. This value is considerably higher than the lower bound on the mixed TM+TE modes case ($Q/Q_{lb} = 1$), but higher than the TM mode bound only ($Q_{TM}^{lb}/Q_{lb} = 1.32$, [54]). The reason is the termination at the local minimum which is not the lowest one. A better result would be achieved if the algorithm could find a spherical helix geometry [55]. This is, however, challenging since the number of local minima is huge and a spherical helix is an extremely specific design. A combination of the Greedy algorithm and some heuristic technique would probably be necessary for this step and will be a topic of future research.

Fig. 11. Current density with (local) minimum Q-factor on a rectangular plate from Fig. 10 modified by a greedy algorithm. Electrical size and feeding placement is the same as in Fig. 10.

Fig. 12. Current density on a spherical shell modified by a greedy algorithm. Electrical size $ka = 0.5$ was used and the spherical shell was discretized into $N = 900$ basis functions.

charge [52]. Then, the structure is perturbed further, adjusting the shape closer to its self-resonance. The final structure reaches $Q/Q_{lb} = 1.57$ and a resulting current density is depicted in Fig. 11.

The final modification involves the evaluation of the Q-factor for 10332 different shapes in $I = 71$ iterations, reaching a local minimum. Thanks to the formulation based on (28) and (33), the entire calculation takes $3\frac{43}{100}$ s. In order to provide a comparison, the same shape modifications are repeated via classical pixel removal based on repetitive matrix inversion and the calculation took 10 s.

$E$. Shape Synthesis – Spherical Shell

The next example reveals a well-known shortage of gradient-based methods – their convergence into the nearest local minimum – where the algorithm remains trapped. Let us consider a spherical shell of electrical size $ka = 0.5$, discretized into $N = 900$ basis functions. One edge is fed by a delta gap (the exact position is irrelevant thanks to the symmetry of the spherical shell). The Q-factor is once again minimized and normalized to the fundamental bound $Q_{lb}$, which is realized by a resonant combination of dominant TM and TE spherical harmonics [53], [54].

The greedy removal of basis function ran for 380 iterations, 270129 antenna candidates were explored, and the resulting normalized Q-factor is $Q/Q_{lb} = 1.51$, see Table I. This value is considerably higher than the lower bound on the mixed TM+TE modes case ($Q/Q_{lb} = 1$), but higher than the TM mode bound only ($Q_{TM}^{lb}/Q_{lb} = 1.32$, [54]). The reason is the termination at the local minimum which is not the lowest one. A better result would be achieved if the algorithm could find a spherical helix geometry [55]. This is, however, challenging since the number of local minima is huge and a spherical helix is an extremely specific design. A combination of the Greedy algorithm and some heuristic technique would probably be necessary for this step and will be a topic of future research.

$E$. Shape Synthesis – Spherical Shell

The next example reveals a well-known shortage of gradient-based methods – their convergence into the nearest local minimum – where the algorithm remains trapped. Let us consider a spherical shell of electrical size $ka = 0.5$, discretized into $N = 900$ basis functions. One edge is fed by a delta gap (the exact position is irrelevant thanks to the symmetry of the spherical shell). The Q-factor is once again minimized and normalized to the fundamental bound $Q_{lb}$, which is realized by a resonant combination of dominant TM and TE spherical harmonics [53], [54].

The greedy removal of basis function ran for 380 iterations, 270129 antenna candidates were explored, and the resulting normalized Q-factor is $Q/Q_{lb} = 1.51$, see Table I. This value is considerably higher than the lower bound on the mixed TM+TE modes case ($Q/Q_{lb} = 1$), but higher than the TM mode bound only ($Q_{TM}^{lb}/Q_{lb} = 1.32$, [54]). The reason is the termination at the local minimum which is not the lowest one. A better result would be achieved if the algorithm could find a spherical helix geometry [55]. This is, however, challenging since the number of local minima is huge and a spherical helix is an extremely specific design. A combination of the Greedy algorithm and some heuristic technique would probably be necessary for this step and will be a topic of future research.
F. Advanced Version of Greedy Algorithm

One final example shows the performance of the greedy algorithm based on topology sensitivity. As presented, the algorithm is, with the exception of the initial precalculation, free of the computationally expensive matrix inversion $Y_G = Z_G^{-1}$. However, for an increased number of iterations $I$, the number of evaluations of the formula (28) is huge and the number of matrix multiplications (32) is proportionally high as well. To reduce the computational burden, it is advantageous to periodically reduce the size of the initial admittance matrix $Y_G$ and the operators $A$ and $B$ in (32) by removing the columns and rows which correspond to the already removed basis functions. This paper refers to this operation as a paper matrix reduction and its impact on the performance of the procedure is depicted in Fig. 13. Periodicity $p = \infty$ means no reduction is made; periodicity $p = 1$ means the matrices are reduced in each iteration.

It can be seen in Fig. 13 that neither of these two extremal options represent an optimal choice of $p$. The optimal reduction period is problem dependent (the factors are: the hardware used, the number of unknowns $N$, the number of iterations $I$) and, for the particular example of the spherical shell with $N = 900$ is $p \approx 50$, reducing the computational time further by a further 20%. Figure 14 shows that, generally, lower period $p$ implies more matrix inversions, but also a monotonic decrease in computational time needed to calculate the next iteration. The second extreme is no reduction, $p = \infty$, in which no time is spent with matrix inversion, but the computational time decreases slowly.

As the last study, Table I summarizes the computational demands and compares them with the same study done with the contemporary technique of pixeling. It is obvious that the novel method is approximately one hundred times faster being capable to modify shapes in real-time.

![Graph showing comparison of computational time needed to calculate $J = 380$ iterations of a greedy algorithm with different compression schemes $p$. The impedance matrix is reduced every $p = \{1, 5, 50, \infty\}$-th iteration. Periodicity $p = \infty$ means that the compression technique is not active. The total computational time of the entire process is depicted in the third column of the legend. The tested object was a spherical shell of $N = 900$ basis functions, $ka = 0.5$, with one delta gap, optimized with respect to a minimum $Q$-factor. For comparison with the classical pixel removal, see Table I.](image)

**Fig. 13.** Comparison of computational time needed to calculate $I = 380$ iterations of a greedy algorithm with different compression schemes $p$. The impedance matrix is reduced every $p = \{1, 5, 50, \infty\}$-th iteration. Periodicity $p = \infty$ means that the compression technique is not active. The total computational time of the entire process is depicted in the third column of the legend. The tested object was a spherical shell of $N = 900$ basis functions, $ka = 0.5$, with one delta gap, optimized with respect to a minimum $Q$-factor. For comparison with the classical pixel removal, see Table I.

| $p$ | Total comp. time, $t(1:i)$ [s] |
|-----|--------------------------------|
| $\infty$ | 79.7 s |
| 1 | 84.9 s |
| 5 | 66.4 s |
| 50 | 63.3 s |

![Graph showing computational time needed to evaluate $i$-th iteration of greedy algorithms with different compression schemes $p$. The same object, setup and metrics as in Fig. 13 were used. Peak values (denoted by the markers) for periodicity $p = \{5, 50\}$ correspond to the computational time needed for a matrix inverse.](image)

**Fig. 14.** Computational time needed to evaluate $i$-th iteration of greedy algorithms with different compression schemes $p$. The same object, setup and metrics as in Fig. 13 were used. Peak values (denoted by the markers) for periodicity $p = \{5, 50\}$ correspond to the computational time needed for a matrix inverse.

| $p$ | comp. time, $t(i)$ [s] |
|-----|------------------------|
| $\infty$ | 79.7 s |
| 1 | 84.9 s |
| 5 | 66.4 s |
| 50 | 63.3 s |

**Table I**

| Basis functions ($N$) | 180 | 567 | 900 |
|-----------------------|-----|-----|-----|
| number of iterations ($I$) | 71 | 279 | 380 |
| realized $Q/Q_{ll0}$ | 1.57 | 1.45 | 1.31 |
| edge removal ($p = \infty$) | 0.30 s | 23.5 s | 79.7 s |
| edge removal ($p = 50$) | 0.28 s | 19.4 s | 63.6 s |
| edge removal ($p = 1$) | 0.43 s | 23.3 s | 84.9 s |
| classical pixel removal | 10 s | 1437 s | 10500 s |

**VI. Discussion**

The topology sensitivity derived in this paper for fixed discretization, i.e., for matrix operators, embodies many similarities with other known methods. These relations and other observations are discussed in this section.

Topology sensitivity assumes a fixed number of DOF, $N$. For $N \rightarrow \infty$, the method can be seen as a direct analogy to the topology derivative extensively studied and used not only in structural optimization [14] but recently also in antenna design [18]. The concept of taking advantage of fixed discretization is not new and is, for example, utilized in constrained surrogate modeling [56].

If the investigation of topology sensitivity is accompanied by the removal of the worst-case basis function, a gradient-based shape synthesis is effectively performed. Thanks to the definition via port modes (33), the impedance matrix inversion is theoretically needed only once at the beginning of the procedure. The rest of the procedure is inversion free. Another advantage is that the evaluation of the topology sensitivity for the entire object ($N \rightarrow \infty$) different shapes are investigated.
at once) is reduced to a sole matrix multiplication and a Hadamard product. This minimal number of multiplications yields evaluations approximately one hundred times faster as compared to contemporary approaches.

A potential cost for the great speed-up described above is that it will not be possible to add DOF (basis functions), i.e., to enlarge the original structure or to revert the previous removal, see Fig. 15. The current modification of the object under study is, therefore, always considered as the initial shape for the next iteration. This operation can be formalized as a movement through the large directional graph where the allowed direction is along the basis function removal.

The reason for not being able to revert a previous removal of a basis function can be seen from (5). In order to remove the n-th basis function, an infinite value is added to the n-th diagonal term of impedance matrix $Z_G$. Numerically, this means that whatever value has been on that position is lost and cannot be recovered. If, however, a finite value is added, this possibility theoretically exists. Regarding this point, imagine that a basis function removal is performed in an approximate sense using $R_\infty \gg 1$. This allows to revert the basis function removal by simply subtracting $R_\infty$ at the corresponding diagonal position in $Z_G$. For a finite value of $R_\infty$, the formula (25) is changed to

$$Y = Y_G - \frac{y_{G,n}y_{G,n}^T}{\pm R_\infty^{-1} + Y_{nn}} \quad (41)$$

and coefficients $\zeta_{ij}$ in (30) are changed to

$$\zeta_{ij} = -\frac{l_i l_j}{l_j^2} \frac{Y_{ij}}{\pm R_\infty^{-1} + Y_{jj}} \quad (42)$$

where a plus sign (+) denotes removal and a minus sign (−) denotes reversal of a previous removal. Changing the evaluation of topology sensitivity to both basis function removal and reversal of a removal is minor and has been numerically tested to work. The problem of this procedure is that high values of $R_\infty$ make the removals more precise but the reversals ill-conditioned (and impossible for $R_\infty \rightarrow \infty$), while low values of $R_\infty$ make removals imprecise and reversals well-conditioned. This unpleasant trade-off is the main drawback of this idea.

Another important fact is the key role of mesh grid. Conventionally, discretization is used only to transform the integro-differential equations into their algebraic form. In this work, a mesh grid plays the extra role of structural parametrization for the investigation of topology changes. The smallest perturbation is given by the resolution of the mesh grid. A dramatic increase in the number of DOF $N$ makes the effective resolution computationally demanding.

Some parts of the derivation are recognized to be formally similar to several previous works typically used for different purposes. The column vectors $y_{G,n}$, extensively utilized for the inversion free evaluation (25), are the port modes from Harrington’s paper [44] where they were used for the optimization of antenna arrays. The univariate search method is devised by Harrington in [44] to effectively find optimal antenna gain. The formula (33) can be seen in a similar light. The key step undertaken in this paper, (28), can, thus, be interpreted as the addition of two port modes where one is fixed (represented by fixed feeding) and the second one is sought for. This operation is formally similar to the difference between primary and secondary characteristic basis functions [57], even though the physics is different.

The port modes [44], extensively utilized in this paper, have many interesting properties, the most remarkable being their immediate connection to the excitation. This feature is missing in the case of characteristic modes [58], therefore, the port modes offer a great alternative whenever feeding is assumed.

The basis function removal paradigm was compared in detail with the classical pixel removal in Section III. These two techniques both remove nodes of the underlying graph, but the corresponding graphs are dual to each other. Imagine the classical pixel removal as a graph described with nodes representing the triangles and edges connecting these nodes according to the map of the basis function. Then, on the contrary, a graph representing the basis function removal has the nodes representing the common edges of the basis functions and edges connecting the adjacent triangles. These two graphs can be anticipated from Figs. 2 and 3, respectively, and it is their difference which constitutes the different number of shape candidates and different topology properties for pixel and basis function removals.

### VII. Conclusion

A basis function removal in the method of moment paradigm has been devised and used for the fast evaluation of topology sensitivity. As compared to classical pixel removal, the proposed removal of basis functions offers more degrees of freedom which results in a significantly higher number of available topologies. Thanks to the formulation via port modes and using the Sherman-Morrison-Woodbury formula it has been possible to avoid repetitive impedance matrix inversions, common in up-to-date methods, and substitute them by matrix and Hadamard products only.

When applied iteratively, the proposed procedure can be used in connection with a greedy (gradient) algorithm synthesizing locally optimal shapes with respect to a given parameter almost in real-time. This approach offers approximately one
hundred times faster evaluation as compared to classical pixel removal. The versatility and effectiveness of the proposed shape synthesis has been demonstrated on several examples.

A future work will aim at incorporating the greedy algorithm powered by topology sensitivity into global optimization routines such as genetic algorithms and machine learning. This should significantly improve their capability to find the global optimum as the global method moves solely through local extremes. A related problem to be studied is the number and structure of local extremes and their position with respect to the global minimum. An important aspect to be studied further is also utilization of port modes which have been shown to be powerful tools due to their direct relation to feedable currents.

A long term goal is the unification of the topology sensitivity method with other known shape and feeding synthesis techniques. This formal unification would establish a solid theoretical background for the further investigation of the role and impact of the geometry and topology changes on the performance of electromagnetic devices.

REFERENCES

[1] F. Gross, Ed., *Frontiers in Antennas: Next Generation Design & Engineering*. McGraw-Hill Professional, 2011.

[2] C. A. Balanis, *Modern Antenna Handbook*. Wiley, 2008.

[3] M.-Y. Kao, Ed., *Encyclopedia of Algorithms*. Springer, 2008.

[4] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*. Wiley – IEEE Press, 1998.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, Great Britain: Cambridge University Press, 2004.

[6] G. C. Onwubolu and B. V. Babu, *New Optimization Techniques in Engineering*. Berlin, Germany: Springer, 2004.

[7] M. Conforti, G. Cornuèjols, and G. Zambelli, *Integer Programming*. Berlin, Germany: Springer, 2014.

[8] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*. John Wiley & Sons, 1999.

[9] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. Massachusetts Institute of Technology, 2009.

[10] Y. Rahmat-Samii, J. M. Kovitz, and H. Rajagopalan, “Nature-inspired optimization techniques in communication antenna design,” Proc. IEEE, vol. 100, no. 7, pp. 2132–2144, July 2012.

[11] Y. Rahmat-Samii and E. Michielssen, Eds., *Electromagnetic Optimization by Genetic Algorithms*. Wiley, 1999.

[12] R. L. Haupt and D. H. Werner, *Genetic Algorithms in Electromagnetics*. John Wiley & Sons, 2007.

[13] E. Lawler, *Combinatorial Optimization: Networks and Matroids*. Mineola, New York, United States: Dover, 2011.

[14] M. P. Bendsoe and O. Sigmund, *Topology Optimization*, 2nd ed. Berlin, Germany: Springer, 2004.

[15] J. L. Volakis, C. Chen, and K. Fujimoto, *Small Antennas: Miniaturization Techniques and Applications*. McGraw-Hill, 2010.

[16] J. Nocedal and S. Wright, *Numerical Optimization*. New York, United States: Springer, 2006.

[17] N. Aage, E. Andreassen, B. S. Lazarov, and O. Sigmund, “Giga-voxel computational morphogenesis for structural design,” *Nature*, vol. 530, pp. 84–88, Oct. 2017.

[18] A. Erentok and O. Sigmund, “Topology optimization of sub-wavelength antennas,” *IEEE Trans. Antennas Propag.*, vol. 59, no. 1, pp. 58–69, 2011.

[19] M. Ghassemi, M. Bakr, and N. Sangary, “Antenna design exploiting joint sensitivity-based geometry evolution,” *IET Microwaves, Antennas & Propagation*, vol. 7, no. 4, pp. 268–276, 2013.

[20] E. Hassan, E. Wadbro, and M. Berggren, “Topology optimization of metallic antennas,” *IEEE Trans. Antennas Propag.*, vol. 62, no. 5, pp. 2488–2500, May 2014.

[21] S. Liu, Q. Wang, and R. Gao, “MoM-based topology optimization method for planar metallic antenna design,” *Acta Mechanica Sinica*, vol. 32, no. 6, pp. 1058–1064, Dec. 2016.

[22] J. L. Volakis, A. Chatterjee, and L. C. Kempel, *Finite Element Method Electromagnetics: Antennas, Microwave Circuits, and Scattering Applications*. Wiley – IEEE Press, 1998.

[23] R. F. Harrington, *Field Computation by Moment Methods*. Piscataway, New Jersey, United States: Wiley – IEEE Press, 1993.

[24] J. J. Toivanen, R. A. E. Makeni, S. Jarvenpaa, P. Ylajälä, and J. Rahola, “Electromagnetic sensitivity analysis and shape optimization using method of moments and automatic differentiation,” *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 1, pp. 168–175, 1 2009.

[25] D. Li, J. Zhu, N. K. Nikolova, M. H. Bakr, and J. W. Bandler, “Electromagnetic optimisation using sensitivity analysis in the frequency domain,” *IET Microwaves, Antennas and Propagation*, vol. 1, pp. 852–859, 8 2007.

[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Johns Hopkins University Press, 2012.

[27] X. Chen, C. Gu, Z. Li, and Z. Niu, “Accelerated direct solution of electromagnetic scattering via characteristic basis function method with Sherman-Morrison-Woodbury formula-based algorithm,” *IEEE Trans. Antennas Propag.*, vol. 64, no. 10, pp. 4482–4486, 10 2016.

[28] X. Fang, Q. Cao, Y. Zhou, and Y. Wang, “Multiscalar compressed and spliced Sherman-Morrison-Woodbury algorithm with characteristic basis function method,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, no. 3, pp. 716–724, 6 2018.

[29] V. V. S. Prakash and R. Mittra, “Characteristic basis function method: A new technique for efficient solution of method of moments matrix equations,” *Microwave and Optical Technology Letters*, vol. 36, no. 2, pp. 95–100, 2003. [Online]. Available: http://dx.doi.org/10.1002/mop.10685

[30] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, 2nd ed. Wiley – IEEE Press, 2001.

[31] B. A. Senior and J. L. Volakis, *Approximate Boundary Conditions in Electromagnetics*. IEE, 1995.

[32] W. C. Chew, M. S. Tong, and B. Hu, *Integral Equation Methods for Electromagnetic and Elastic Waves*. Morgan & Claypool, 2009.

[33] W. C. Gibson, *The Method of Moments in Electromagnetics*, 2nd ed. Chapman and Hall/CRC, 2014.

[34] R. F. Harrington and J. R. Mautz, “Control of radar scattering by reactive loading,” *IEEE Trans. Antennas Propag.*, vol. 20, no. 4, pp. 446–454, July 1972.

[35] L. Jelinek and M. Capek, “Optimal currents on arbitrarily shaped surfaces,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 1, pp. 329–341, Jan. 2017.

[36] S. M. Rao, D. R. Wilton, and A. W. Glisson, “Electromagnetic scattering by surfaces of arbitrary shape,” *IEEE Trans. Antennas Propag.*, vol. 30, no. 3, pp. 409–418, May 1982.

[37] B. Yang and J. J. Adams, “Systematic shape optimization of symmetric mimo antennas using characteristic modes,” *IEEE Trans. Antennas Propag.*, vol. 64, no. 7, pp. 2668–2678, July 2016.

[38] C. A. Balanis, *Advanced Engineering Electromagnetics*. Wiley, 1989.

[39] (2017) Antenna Toolbox for MATLAB (AoTM). Czech Technical University in Prague. [Online]. Available: www.antennatoolbox.com

[40] G. Ghione and C. Naldi, “Analytical formulas for coplanar lines in hybrid and monolithic MICS,” *Electronics Letters*, vol. 20, no. 4, pp. 179–181, 2 1984.

[41] S. R. Best, “Electrically small resonant planar antennas,” *IEEE Antennas Propag. Mag.*, vol. 57, no. 3, pp. 38–47, June 2015.

[42] W. W. Hager, “Updating the inverse of a matrix,” *SIAM Review*, vol. 31, pp. 221–239, 1989.

[43] (2018) The Matlab. The MathWorks. [Online]. Available: www.mathworks.com

[44] R. Harrington, “Reactively controlled directive arrays,” *IEEE Trans. Antennas Propag.*, vol. 26, no. 3, pp. 390–395, 1978.

[45] M. Gustafsson, D. Tayli, C. Ehrenborg, M. Cismasu, and S. Norbedo, “Antenna current optimization using MATLAB and CVX,” *FERMAT*, vol. 15, no. 5, pp. 1–29, May–June 2016. [Online]. Available: http://www.e-fermat.org/articles/gustafsson-art-2016-vol15-mayjun-005/0

[46] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, United Kingdom: Cambridge University Press, 1990.

[47] A. A. Novotny and J. Sokolowski, *Topological Derivatives in Shape Optimization*, ser. Interaction of Mechanics and Mathematics, L. Truskinovsky, Ed. Berlin: Springer, 2013.

[48] M. Cismasu and M. Gustafsson, “Antenna bandwidth optimization with single frequency simulation,” *IEEE Trans. Antennas Propag.*, vol. 62, no. 3, pp. 1304–1311, 2014.

[49] R. F. Harrington, “Antenna association for maximum gain,” *IEEE Trans. Antennas Propag.*, vol. 13, no. 6, pp. 896–903, Nov. 1965.

[50] C. A. Balanis, *Antenna Theory Analysis and Design*, 3rd ed. Wiley, 2005.
[51] M. Capek, M. Gustafsson, and K. Schab, “Minimization of antenna quality factor,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 8, pp. 4115–4123, Aug. 2017.

[52] M. Gustafsson, M. Cismasu, and B. L. G. Jonsson, “Physical bounds and optimal currents on antennas,” *IEEE Trans. Antennas Propag.*, vol. 60, no. 6, pp. 2672–2681, June 2012.

[53] R. E. Collin and S. Rothschild, “Evaluation of antenna Q,” *IEEE Trans. Antennas Propag.*, vol. 12, no. 1, pp. 23–27, Jan. 1964.

[54] M. Capek and L. Jelinek, “Optimal composition of modal currents for minimal quality factor Q,” *IEEE Trans. Antennas Propag.*, vol. 64, no. 12, pp. 5230–5242, Dec. 2016.

[55] S. R. Best, “The radiation properties of electrically small folded helical antennas,” *IEEE Trans. Antennas Propag.*, vol. 52, no. 4, pp. 953–960, 2004.

[56] S. Koziel and A. T. Sigurdsson, “Triangulation-based constrained surrogate modeling of antennas,” *IEEE Transactions on Antennas and Propagation*, 2018.

[57] C. Craeye, J. Laviada, R. Maaskant, and R. Mittra, “Macro basis function framework for solving maxwell’s equations in surface integral equation form,” *Fermat J.*, vol. 3, pp. 1–16, 2014.

[58] R. F. Harrington and J. R. Mautz, “Theory of characteristic modes for conducting bodies,” *IEEE Trans. Antennas Propag.*, vol. 19, no. 5, pp. 622–628, Sept. 1971.

**Miloslav Capek** (SM’17) received his Ph.D. degree from the Czech Technical University in Prague, Czech Republic, in 2014. In 2017 he was appointed Associate Professor at the Department of Electromagnetic Field at the same university.

He leads the development of the AToM (Antenna Toolbox for Matlab) package. His research interests are in the area of electromagnetic theory, electrically small antennas, numerical techniques, fractal geometry and optimization. He authored or co-authored over 70 journal and conference papers.

Dr. Capek is member of Radioengineering Society, regional delegate of EurAAP, and Associate Editor of Radioengineering.

**Lukas Jelinek** received his Ph.D. degree from the Czech Technical University in Prague, Czech Republic, in 2006. In 2015 he was appointed Associate Professor at the Department of Electromagnetic Field at the same university.

His research interests include wave propagation in complex media, general field theory, numerical techniques and optimization.

**Mats Gustafsson** (SM’17) received the M.Sc. degree in Engineering Physics 1994, the Ph.D. degree in Electromagnetic Theory 2000, was appointed Docent 2005, and Professor of Electromagnetic Theory 2011, all from Lund University, Sweden.

He co-founded the company Phase holographic imaging AB in 2004. His research interests are in scattering and antenna theory and inverse scattering and imaging. He has written over 90 peer reviewed journal papers and over 100 conference papers. Prof. Gustafsson received the IEEE Schelkunoff Transactions Prize Paper Award 2010 and Best Paper Awards at EuCAP 2007 and 2013. He served as an IEEE AP-S Distinguished Lecturer for 2013-15.