scientific reports

Check for updates

OPEN

Fast machine-learning-enabled size reduction of microwave components using response features

Slawomir Koziel[1,2]✉ & Anna Pietrenko-Dabrowska[2]

Achieving compact size has emerged as a key consideration in modern microwave design. While structural miniaturization can be accomplished through judicious circuit architecture selection, precise parameter tuning is equally vital to minimize physical dimensions while meeting stringent performance requirements for electrical characteristics. Due to the intricate nature of compact structures, global optimization is recommended, yet hindered by the excessive expenses associated with system evaluation, typically conducted through electromagnetic (EM) simulation. This challenge is further compounded by the fact that size reduction is a constrained problem entailing expensive constraints. This paper introduces an innovative method for cost-effective explicit miniaturization of microwave components on a global scale. Our approach leverages response feature technology, formulating the optimization problem based on a set of characteristic points derived from EM-analyzed responses, combined with an implicit constraint handling approach. Both elements facilitate handling size reduction by transforming it into an unconstrained problem and regularizing the objective function. The core search engine employs a machine-learning framework with kriging-based surrogates refined using the predicted improvement in the objective function as the infill criterion. Our algorithm is demonstrated using two miniaturized couplers and is shown superior over several benchmark routines, encompassing both conventional (gradient-based) and population-based procedures, alongside a machine learning technique. The primary strengths of the proposed framework lie in its reliability, computational efficiency (with a typical optimization cost ranging from 100 to 150 EM circuit analyses), and straightforward setup.

Keywords Microwave circuits, Compact circuits, Simulation-driven design, Size reduction, Numerical optimization, Machine learning, Surrogate modeling

The significance of maintaining compact size has been consistently increasing in the design of contemporary passive high-frequency devices, primarily driven by various application areas such as implantable systems[1], energy harvesting[2], mobile communications[3], the internet of things[4], or RFID[5]. Due to the dependence of the physical dimensions of passives components on the guided wavelength, achieving miniaturization is a complex endeavour. It is typically accomplished through a combination of methods, including transmission line (TL) meandering[6], exploiting slow-wave phenomena[7] (e.g., compact microstrip resonant cells, CMRCs[8]), integrating defected grounds[9], utilizing metamaterials[10], employing substrate-integrated waveguides[11], or incorporating auxiliary structures (stubs[12], slots[13], shorting pins[14]). While often unavoidable, this approach results in intricate geometries characterized by numerous parameters and the presence of strong cross-coupling effects. Both aspects pose significant design challenges. On the one hand, precise rendition of circuit characteristics demands computationally expensive electromagnetic (EM) simulation, as neither analytical nor circuit-theory-based models provide sufficient accuracy[15,16]. On the other hand, geometry parameters must be meticulously tuned in a synchronized manner, preferably utilizing rigorous optimization methods[17,18].

Optimizing microwave components at the EM level poses challenges primarily due to its computational cost. The associated expenses are considerable, even for local parameter tuning (utilizing methods like gradient-based[19] or stencil-based algorithms[20]), not to mention global search[21], statistical analysis (e.g., yield estimation[22]),

[1]Engineering Optimization & Modeling Center, Reykjavik University, 102, Reykjavik, Iceland. [2]Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Gdansk 80-233, Poland. ✉email: koziel@ru.is

design centering[23], or multi-criterial design[24]. Numerous efforts have been made in the literature to streamline EM-driven design procedures[25–30]. Some available methods include cost reduction in sensitivity estimation for gradient-based algorithms (employing adjoint sensitivities[31], restricted gradient updates[32,33]), surrogate-based techniques utilizing both behavioural[34] and physics-based metamodels[35], and machine learning frameworks[36]. Surrogate modelling methods applied in these procedures encompass kriging[37], radial basis functions[38], Gaussian process regression[39], support vector machines[40], as well as artificial neural networks[41–43]. Physics-based routines often rely on space mapping[44] or various response correction methods[45–47].

In the realm of compact components, parameter tuning is typically geared towards achieving miniaturization, making the reduction of circuit size a primary objective[48]. Simultaneously, stringent conditions are imposed on electrical performance metrics, such as the required return loss level, port isolation across a specified frequency range, power split ratio, or the phase response[49,50]. Numerically, this results in highly constrained optimization problems, where the constraints are computationally heavy to assess, relying on EM simulation. Explicit constraint handling is generally challenging, although recent developments have shown promise (e.g[51]). , . Another approach involves implicit constraint control using penalty functions[52], with the main goal (diminishing size) supplemented by penalty components measuring potential constraint violations[53]. While this enables reformulation of the problem into an unconstrained task, penalty coefficient adjustment is crucial for optimization process performance. Manual setup is non-trivial and usually requires multiple attempts, each followed by a test optimization run. Recently, adaptive approaches have been proposed with penalty terms automatically adjusted using actual constraint levels and/or the algorithm's convergence status[54,55]. However, these methods are suitable for local parameter tuning, whereas the geometric complexity of compact circuits necessitates global optimization. The most prevalent solution methods are bio-inspired algorithms[56–59]. Unfortunately, their global search capability comes at the cost of poor computational efficiency. In fact, direct EM-driven nature-inspired optimization is impractical if not prohibitive. Mitigation methods primarily rely on surrogate modelling techniques[25,26,60,61]. Typical surrogate-assisted frameworks take the form of iterative procedures, where the metamodel serves as a predictor, facilitating the identification of the optimum design, and is progressively refined using accumulated EM simulation data. Infill points are generated using various rules favouring the improvement of the model's dependability[62], allocation of the optimum[63], or balanced exploration and exploitation[64]. Algorithms of this kind are often referred to as machine learning procedures[65,66].

The bottleneck in surrogate-assisted methods lies in constructing the metamodel itself, hindered by the nonlinearity of microwave circuit responses and dimensionality issues[67]. Consequently, many algorithms are only showcased using test cases that involve a limited number of variables and/or within restricted parameter ranges[68,69]. This study introduces a novel surrogate-based methodology for globally simulating-driven miniaturization of microwave components. Our approach is a machine learning framework that operates by incorporating response features of the circuit of interest, enabling the creation of a reliable surrogate model with small training datasets. The infill criterion is involves the enhancement of the predicted merit function, and the infill points are generated using a particle swarm optimizer (PSO)[70]. This is facilitated by the employing penalty functions for controlling design constraints, which allow us to reformulate the problem into an unconstrained task, aside from the lower/upper parameter bounds defining the search space. Our technique has been showcased using two couplers and juxtaposed against several benchmark methods, including gradient-based optimization, population-based procedures, and a machine learning framework directly processing complete frequency characteristics of the circuit. The results demonstrate consistent and competitive operation with respect to achievable size reduction and low running costs. The average expenses incurred by the search process are merely about 100 EM circuit analyses.

The paper encompasses several original components and technical contributions, including: (i) the creation of an innovative machine learning procedure designed for explicit reduction of microwave device's size; (ii) the implementation of mechanisms that facilitate the search process, such as conducting optimization using response features and implicit constraint handling; (iii) the demonstration of the framework's competitive operation in terms of design quality and consistency of produced solutions; (iv) the exhibition of excellent computational efficiency in the proposed procedure. As far as the authors are aware, there have been no comparable algorithms reported in the literature to date.

Simulation-based size reduction. Constraint handling

This part of the work revisits the formulation of EM-based miniaturization for microwave components as a nonlinear constrained minimization problem. We also explore explicit and implicit constraint handling and provide an overview of existing solution methodologies. The global search procedure will be detailed in Sect. 3.

Microwave circuit size reduction as optimization problem

Compact structures have gained significant importance across various application areas, including the IoT, RFID, and wearable/implantable devices. The design of these structures involves selecting appropriate circuit architectures, employing techniques like transmission line meandering[6], incorporating metamaterial components[10], or leveraging the slow-wave phenomenon[7]. To achieve the necessary electrical performance while keeping a small size, concurrent adjustment of all circuit dimensions is essential. Traditional parametric studies are no longer sufficient, necessitating rigorous numerical optimization instead.

It should be emphasized that design of compact microwave components differs somehow from the development of other types of high-frequency structures, in particular, compact antennas. In antenna design, a possible size reduction is impeded by physical constraints, e.g., the required current length path to ensure sufficient impedance matching at the lower end of the operating bandwidth. In the case of conventional microwave components, the situation is similar. For example, the transmission line lengths are related to the operating frequency and expressed in terms of the guided wavelength. However, using technique such as

Symbol	Meaning	Comment
$\boldsymbol{x} = [x_1 \ldots x_n]^T$	Vector of design parameters	Typically, the variables are the circuit geometry parameters (dimensions in mm)
X	Parameter space	Typically, X is an interval $[\boldsymbol{l}\ \boldsymbol{u}]$, where $\boldsymbol{l} = [l_1 \ldots l_n]^T$ and $\boldsymbol{u} = [u_1 \ldots u_n]^T$ are lower and upper bounds on design parameters so that $l_k \le x_k \le u_k$, for $k = 1, \ldots, n$
$A(\boldsymbol{x})$	Circuit size	For planar circuits, $A(\boldsymbol{x})$ is typically a footprint area in mm^2
$S_{jk}(\boldsymbol{x},f)$	Scattering parameters at the design $\boldsymbol{x}$ and frequency f; j and k refer to the circuit port numbers	Scattering parameters are complex numbers; in the design process we handle their moduli $\lvert S_{jk}\rvert$, expressed in decibels
$g_k(\boldsymbol{x}) \le 0, k = 1, \ldots, n_g$,	Inequality constraints	Typically, expressed using acceptance thresholds for selected circuit characteristics, e.g., $\lvert S_{11}(\boldsymbol{x},f)\rvert \le -10$ dB for f within the frequency range of interest
$h_k(\boldsymbol{x}) = 0, k = 1, \ldots, n_h$,	Equality constraints	Typically, expressed using target values for selected circuit characteristics
X_f	Feasible space	$X_f \subset X$ contains parameter vectors $\boldsymbol{x}$ for which all constraints are satisfied, i.e., $g_k(\boldsymbol{x}) \le 0$ for $k = 1, \ldots, n_g$, and $h_k(\boldsymbol{x}) = 0$ for $k = 1, \ldots, n_h$

Table 1. Simulation-based microwave size reduction. Notation and terminology.

Constraint	Type	Analytical description[$]
Input matching $\lvert S_{11}\rvert$ not exceeding −20 dB over the operating bandwidth $[f_1 f_2]$	Inequality	$\lvert S_{11}(\boldsymbol{x},f)\rvert \le -20$ dB for $f \in [f_1 f_2]$
Port isolation $\lvert S_{41}\rvert$ not exceeding −20 dB over the operating bandwidth $[f_1 f_2]$	Inequality	$\lvert S_{41}(\boldsymbol{x},f)\rvert \le -20$ dB for $f \in [f_1 f_2]$
In-band transmission ripple not exceeding 0.2 dB over the operating bandwidth $[f_1 f_2]$	Inequality	$\lvert S_{21}(\boldsymbol{x},f)\rvert \ge -0.2$ dB for $f \in [f_1 f_2]$
Power split ratio between output ports 2 and 3 equal to K_P at the center frequency f_0	Equality	$\lvert S_{31}(\boldsymbol{x},f)\rvert - \lvert S_{21}(\boldsymbol{x},f)\rvert = K_P$ at $f = f_0$;
Phase difference between output ports 2 and 3 equal to 90° at the center frequency f_0	Equality	$\angle S_{31}(\boldsymbol{x},f) - \angle S_{21}(\boldsymbol{x},f) = 90°$ at $f = f_0$;

Table 2. Example constraints in size-reduction of microwave components [$]The symbol $\lvert S_{jk}(\boldsymbol{x},f)\rvert$ stands for the modulus of the S-parameter S_{jk} at the design $\boldsymbol{x}$, and frequency f.

those mentioned in the previous paragraph allows to bypass these physical limitations. At the same time, these techniques result in topologically complex designs described by the increased number of parameters, and the need to use full-wave EM simulations to ensure accurate circuit evaluations. These factors contribute to making the optimization process (including explicit size reduction) a challenging endeavour.

This section presents the formalization of the miniaturization task as a nonlinear minimization problem. The pertinent notation is detailed in Table 1, and the design task at hand is expressed as follows:

The feasible space X_f in Eq. (1) is defined as the subset of X such that design constraints are satisfied for all $\boldsymbol{x} \in X_f$.

$$\mathbf{x}^* = \arg\min_{\mathbf{x} \in X_f} A(\mathbf{x}) \tag{1}$$

Explicit and implicit treatment of design constraints

Managing design constraints represents a challenging aspect of optimization-driven circuit miniaturization. In addition to geometrical conditions such as lower and upper parameter bounds, constraints are applied to electrical characteristics and necessitate electromagnetic (EM) analysis for evaluation. Explicitly addressing these costly constraints is troublesome, particularly in the context of global optimization. An alternative approach is implicit treatment using a penalty function method[52], where the problem (1) is formulated as:

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

The merit function U_P in (2) is defined as

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

In this formulation, the main design goal (miniaturization) is accompanied by contributions from penalty functions $c_k(\boldsymbol{x})$, which quantify constraint violations. The weighting factors β_k (penalty coefficients) dictate the influence of individual terms. Examples of constraints are outlined in Table 2. Table 3 offers potential definitions for the penalty functions, wherein some cases measure relative constraint violations concerning assumed acceptance thresholds (e.g., − 20 dB).

Note that the power factor of two is used, $[.]^2$, so that U_p is a smooth function at the boundary feasible space X_f. This alleviates the difficulties pertinent to the exploration of that region, otherwise necessary as one or more constraints are normally active at the minimum-size design.

A practical problem of implicit constraint handling is a selection of the coefficients β_k. Their values cannot be too low, as this would impede constraint control. If β_k are too high, numerical issues arise due to high nonlinearity of U_p in the region adjacent to the boundary of X_f. Finding the optimum levels of β_k is non-trivial. The adaptive constraint handling strategies proposed recently[54,55], address this issue to a great extent. At the same

Constraint	Penalty function
Input matching $\lvert S_{11}\rvert$ not exceeding –20 dB over the operating bandwidth $[f_1 f_2]$	$c(\mathbf{x}) = \left[\frac{\max\{\max\{f_1 \le f \le f_2 : \lvert S_{11}(\mathbf{x},f)\rvert\}+20,0\}}{20}\right]^2$
Port isolation $\lvert S_{41}\rvert$ not exceeding –20 dB over the operating bandwidth $[f_1 f_2]$	$c(\mathbf{x}) = \left[\frac{\max\{\max\{f_1 \le f \le f_2 : \lvert S_{41}(\mathbf{x},f)\rvert\}+20,0\}}{20}\right]^2$
In-band transmission ripple not exceeding 0.2 dB over the operating bandwidth $[f_1 f_2]$	$c(\mathbf{x}) = \left[\frac{\max\{-\min\{f_1 \le f \le f_2 : \lvert S_{21}(\mathbf{x},f)\rvert\}-0.2,0\}}{0.2}\right]^2$
Power split ratio between output ports 2 and 3 equal to K_P at the center frequency f_0	$c(\mathbf{x}) = [\lvert S_{31}(\mathbf{x}, f_0)\rvert - \lvert S_{21}(\mathbf{x}, f_0)\rvert - K_P]^2$
Phase difference between output ports 2 and 3 equal to 90° at the center frequency f_0	$c(\mathbf{x}) = [\angle S_{31}(\mathbf{x}, f_0) - \angle S_{21}(\mathbf{x}, f_0) - 90°]^2$

Table 3. Possible formulation of penalty functions for constraints of Table 2.

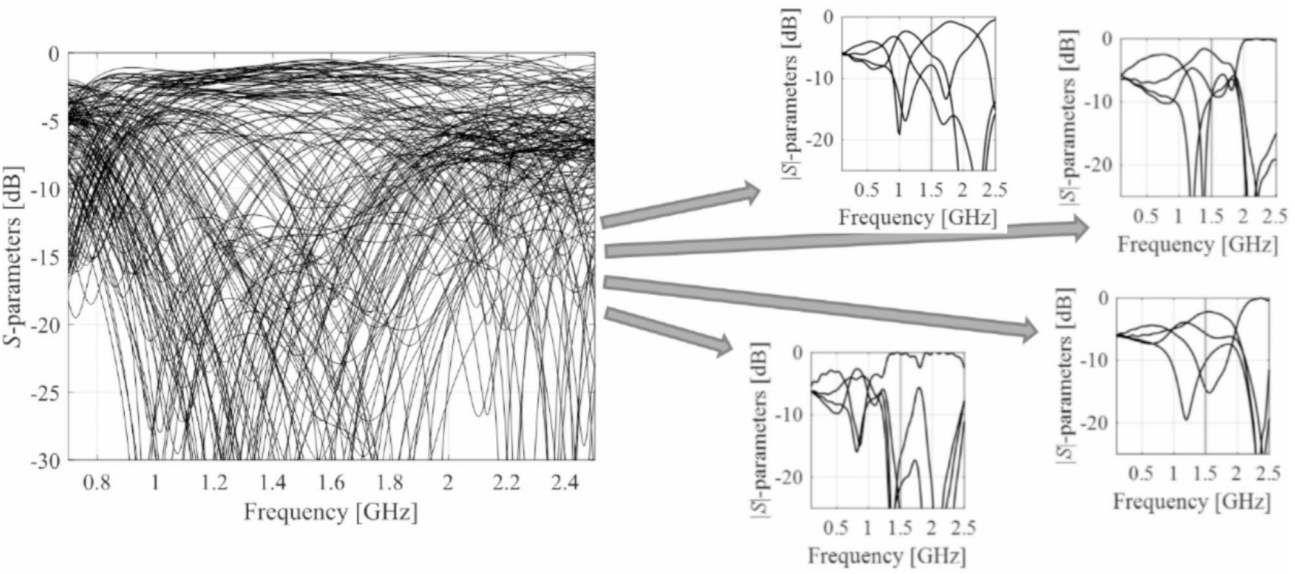


Fig. 1. Scattering parameters of a microstrip coupler at selected random parameter vectors. The characteristics are highly nonlinear, therefore, difficult to be represented using data-driven surrogates. At the same time, given a target center frequency (here, 1.5 GHz), coupler optimization with local (e.g., gradient-based) routines would fail when initiated from majority of the shown designs.

time, improved explicit constraint handling methods have been proposed[51,71]. Both types of approaches offer reasonable trade-offs between constraint control and possible size reduction; however, they are only applicable to local optimization.

Global size reduction by response features and parameter space pre-screening

This section offers a comprehensive overview of the proposed global size reduction method. The technique utilizes a surrogate-assisted machine learning approach, integrating an iterative prediction-correction scheme. In this process, the surrogate model guides the search towards the optimal design and refines itself iteratively with accumulated electromagnetic (EM) simulation data. The size reduction task is reformulated concerning the characteristic points of the circuit, as outlined in Eq. (1). Subsequently, we revisit the response feature method, present an analytical formulation of the feature-based merit function, discuss the pre-screening step of the optimization process, detail the generation of infill points using nature-inspired algorithms, and provide a summary of the operational flow of the complete framework.

Response feature approach. Size reduction task reformulation

The size reduction procedure proposed in this study relies on efficient surrogate models designed to accelerate the exploration of the parameter space. The primary challenge lies in the high nonlinearity of microwave component characteristics, making them challenging to represent accurately using data-driven methods, especially across broad ranges of frequencies and geometry parameters. Figure 1 illustrates exemplary responses of a miniaturized coupler, which is considered as one of the verification circuits in Sect. 4. The high nonlinearity of the frequency characteristics poses difficulty in being accurately represented by data-driven surrogates. Consequently, optimization for any specific objectives (e.g., size reduction) concerning a given target operating frequency must be conducted in a global sense, as local tuning is likely to converge to an inferior local optimum due to the inherent nonlinearity.

To tackle these challenges, we leverage the response feature technology[72]. The underlying concept involves reformulating the problem using suitably defined characteristic (feature) points extracted from the frequency responses, utilizing EM simulation data[73]. As observed in existing literature, the dependence between the feature points and geometry variables exhibits less nonlinearity compared to the corresponding relationship for frequency characteristics[74–76]. This regularization of objective functions leads to faster convergence in optimization processes[75], enabling global search ability even with formally local algorithms, and reducing the training set size for metamodel rendition[76]. The feature point selection should align with the designated design objectives, making them problem-dependent[73]. Specific examples are illustrated in Fig. 2. The characteristic points identified in Fig. 2(a) correspond to the minima of the matching and isolation responses (relevant for managing the operating frequency of the circuit), −20 dB levels of $|S_{11}|$ and $|S_{41}|$ (pertinent for the circuit's matching/isolation bandwidth), and the levels of the transmission responses at the center frequency (helpful for controlling the power split ratio).

The most important advantage of response features is their weakly-nonlinear dependence on the circuit designable parameters[72–76]. Figure 2(a) demonstrates this using the coupler of Fig. 1, in the form of the dependency between the feature-based-evaluated operating parameters and selected design variables. Needless to say, representing such dependencies through behavioral models is considerably simpler than behavioral modeling of the frequency characteristics. The size reduction framework proposed in this study leverages the properties of response features illustrated here.

For the purpose of restating the design task using characteristic points, we need appropriate notation, which has been introduced in Fig. 3. Figure 3(a) describes a general notation, whereas Fig. 3(b) elaborates on the specific choice of the feature points for coupling structures. These particular points allow us to account for the −20 dB bandwidth for $|S_{11}|$ and $|S_{41}|$, as well as the power split ratio K_P, which can be calculated as $K_P = f_{L.5} - f_{L.6}$.

Let, as before, f_1 and f_2 define the frequency range over which both $|S_{11}|$ and $|S_{41}|$ are to be not higher than −20 dB, and K_P stands for the target power split ratio. These conditions become design constraints (cf. Tables 2 and 3) from the points of view of circuit size reduction $A(\boldsymbol{x})$. The size reduction problem, equivalent to (1), but expressed using response features takes the form of

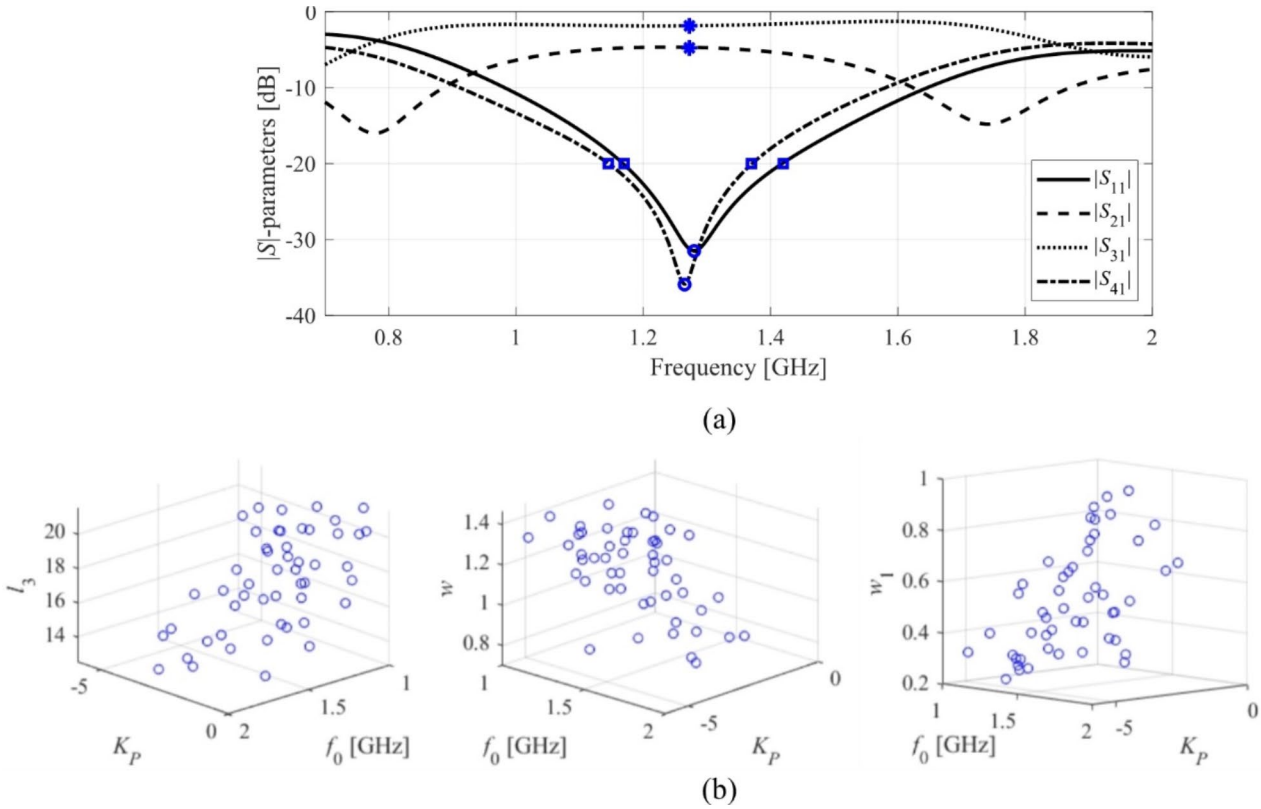


Fig. 2. Response features for a coupler structure: (a) possible characteristic point choices: o – points associated with the minima of $|S_{11}|$ and $|S_{41}|$, ∗ - points corresponding to the power division ratio computed at the frequency f_0 being the average of the frequencies of $|S_{11}|$ and $|S_{41}|$ minima, ▢- points corresponding to −20 dB levels of $|S_{11}|$ and $|S_{41}|$; (b) relationship between operating conditions (extracted from the response features, here the center frequency f_0 and power split ratio K_P) and selected design variables of the circuit. The plots are created using a set of randomly-generated designs. Only the points for which the corresponding characteristics allow for extracting the approximated operating parameters, as indicated above, are illustrated. Clear patterns can be observed although the shown designs were not subject to optimization.

Symbol	Explanation
$\boldsymbol{f}_P(\boldsymbol{x}) = [\boldsymbol{f}_f(\boldsymbol{x})^T \boldsymbol{f}_L(\boldsymbol{x})^T]^T$	An aggregated vector of K response features
$\boldsymbol{f}_f(\boldsymbol{x}) = [f_{f.1}(\boldsymbol{x}) \ldots f_{f.K}(\boldsymbol{x})]^T$	A vector of frequency coordinates of the feature points from 1 to K
$\boldsymbol{f}_L(\boldsymbol{x}) = [f_{L.1}(\boldsymbol{x}) \ldots f_{L.K}(\boldsymbol{x})]^T$	A vector of level coordinates of the feature points from 1 to K

(a)

Symbol	Explanation
$\boldsymbol{f}_f(\boldsymbol{x}) = \left[f_{f.1}(\boldsymbol{x})\, f_{f.2}(\boldsymbol{x})\, f_{f.3}(\boldsymbol{x})\, f_{f.4}(\boldsymbol{x})\, f_{f.5}(\boldsymbol{x})\, f_{f.6}(\boldsymbol{x})\right]^T$	Frequency coordinates of the feature points for a microstrip coupler
$\boldsymbol{f}_L(\boldsymbol{x}) = \left[f_{L.1}(\boldsymbol{x})\, f_{L.2}(\boldsymbol{x})\, f_{L.3}(\boldsymbol{x})\, f_{L.4}(\boldsymbol{x})\, f_{L.5}(\boldsymbol{x})\, f_{L.6}(\boldsymbol{x})\right]^T$	Level coordinates of the feature points for a microstrip coupler
$f_{f.1}$ and $f_{f.2}$	Frequencies corresponding to –20 dB levels of $\lvert S_{11}\rvert$; if the minimum of $\lvert S_{11}\rvert$ is above –20 dB, then both frequencies correspond to the frequency of the minimum
$f_{f.3}$ and $f_{f.4}$	Frequencies corresponding to –20 dB levels of $\lvert S_{41}\rvert$; if the minimum of $\lvert S_{41}\rvert$ is above –20 dB, then both frequencies correspond to the frequency of the minimum
$f_{f.5} = f_{f.6}$	Average value of the previous frequencies, i.e., $f_{f.5} = f_{f.6} = [f_{f.1} + f_{f.2} + f_{f.3} + f_{f.4}]/4$
$f_{L.1}$ and $f_{L.2}$	Both equal to –20 dB, or equal to the value of $\lvert S_{11}\rvert$ at its minimum (if higher than –20 dB)
$f_{L.3}$ and $f_{L.4}$	Both equal to –20 dB, or equal to the value of $\lvert S_{41}\rvert$ at its minimum (if higher than –20 dB)
$f_{L.5}$ and $f_{L.6}$	Equal to $\lvert S_{21}\rvert$ and $\lvert S_{31}\rvert$ at $f_{f.5} = f_{f.6}$

(b)

Fig. 3. Notation pertinent to response features as utilized in this work: (a) general notation, (b) an example selection of features for the microstrip coupler design problem (cf. (5)).

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} U_F(\mathbf{x}, \mathbf{f}_P(\mathbf{x})) \tag{4}$$

in which the merit function U_F is determined using the feature vector $\boldsymbol{f}_P(\boldsymbol{x})$. For a particular coupler design task as discussed above, it may be defined as

$$U_F(\mathbf{x}, \mathbf{f}_P(\mathbf{x})) = A(\mathbf{x}) + \beta_1 \left\| \begin{bmatrix} \max\{f_{f.1} - f_1, 0\} \\ \max\{f_{f.3} - f_1, 0\} \\ \max\{f_2 - f_{f.2}, 0\} \\ \max\{f_2 - f_{f.4}, 0\} \end{bmatrix} \right\|^2 + \\ + \beta_2 \left[f_{f.5} - \tfrac{f_1+f_2}{2} \right]^2 + \beta_3 [(f_{L.5} - f_{L.6}) - K_P]^2 \tag{5}$$

As before, minimization of the circuit footprint area $A(\boldsymbol{x})$ is the main goal. Then, we have three penalty terms. The first one is to ensure that the bandwidth constraint imposed on $\lvert S_{11}\rvert$ and $\lvert S_{41}\rvert$ is fulfilled.

Note that this term is only contributing to U_F if there is a bandwidth violation either at f_1 or f_2. The second term is to enforce that the center frequency of the coupler is in the middle of the prescribed bandwidth, whereas the last term controls the power split constraint. The coefficients β_k can be set to relatively high values because the search process will be conducted by means of the surrogate model. In this paper, we set $\beta_k = 10^4$ for $k = 1, 2, 3$.

Parameter space pre-screening. Constructing initial surrogate model

The machine-learning-based framework proposed in this paper utilizes a surrogate model to expedite the optimization process. The procedure begins by a pre-screening of the search space and constructing the initial metamodel $\boldsymbol{s}^{(0)}(\boldsymbol{x})$. For efficiency and reliability, $\boldsymbol{s}^{(0)}(\boldsymbol{x})$ is built using characteristic points (cf. Section 3.1). It represents the components of the feature vector, i.e.,

$$\mathbf{s}^{(0)}(\mathbf{x}) = \left[\left[s_{f.1}^{(0)}(\mathbf{x}) \ \ldots \ s_{f.K}^{(0)}(\mathbf{x}) \right]^T \ \left[s_{L.1}^{(0)}(\mathbf{x}) \ \ldots \ s_{L.K}^{(0)}(\mathbf{x}) \right]^T \right]^T \tag{6}$$

The modelling method employed here is kriging interpolation[37], yet, this choice is not critical. The model is identified based on the dataset $\{\boldsymbol{x}_B^{(j)}, \boldsymbol{f}_P(\boldsymbol{x}_B^{(j)})\}$, $j = 1, \ldots, N_{init}$, with the feature vectors $\boldsymbol{f}_P(\boldsymbol{x}_B^{(j)})$ determined using EM simulation results at $\boldsymbol{x}_B^{(j)}$.

The parameter N_{init} is not determined beforehand. The samples are generated sequentially until the metamodel's accuracy achieves the prescribed level. Here, the metric of choice is the relative root-mean square (RMS) error[77]. The target predictive power is $E_{\max}$. The samples $\boldsymbol{x}_B^{(j)}$ are generated to satisfy the following conditions:

- The circuit size $A(\boldsymbol{x}_B^{(j)}) \le A_{\max}$ (a user-defined parameter);
- The circuit size $A(\boldsymbol{x}_B^{(j)}) \ge A_{\min}$ (a user-defined parameter);
- The feature points are extractable at $\boldsymbol{x}_B^{(j)}$;

The first two conditions are optional and introduced to avoid EM analysis for designs at which the circuit size is clearly too large or too small; the necessary limits $A_{\min}$ and $A_{\max}$ can be deduced from prior experiments with the circuit. For the last condition, at certain designs the of the response features may not be possible, e.g., due to a severe distortion of the circuit outputs, in which case the design would not be included into the training dataset.

The pseudocode of the pre-screening and training data generation procedures has been shown in Fig. 4. Owing to a relatively simple relation between the characteristic points and circuit's design variables, it is possible to construct usable models using small numbers of samples N_{init}, typically less than a hundred. However, as the circuit outputs at some of the random designs are degenerated, the actual number of trial points is considerably larger than N_{init}, usually by a factor between 1.5 and 3.

1. Input parameters:
   - Design space X (interval [$\boldsymbol{l}$ $\boldsymbol{u}$], where $\boldsymbol{l}$ and $\boldsymbol{u}$ are lower and upper bounds for designable parameters; other constraints are possible per designer's needs);
   - Definition of the response features $\boldsymbol{f}_P$;
   - Required modelling error $E_{\max}$;
   - Circuit size bounds $A_{\min}$ and $A_{\max}$ (optional);
2. Set $j = 0$;
3. Generate a sample $\boldsymbol{x}_{tmp} \in X$ using uniform probability distribution;
4. **if** $A(\boldsymbol{x}_{tmp}) \ge A_{\min}$ AND $A(\boldsymbol{x}_{tmp}) \le A_{\max}$
   Evaluate circuit characteristics at $\boldsymbol{x}_{tmp}$ using EM analysis;
   else
   Go to 3;
   end
5. Extract the feature vector $\boldsymbol{f}_P(\boldsymbol{x}_{tmp})$; assign $\boldsymbol{f}_P(\boldsymbol{x}_{tmp}) = 0$ if features are not extractable (cf. Fig. 3);
6. **if** $||\boldsymbol{f}_P(\boldsymbol{x}_{tmp})|| > 0$
   Set $j = j + 1$;
   Set $\boldsymbol{x}_B^{(j)} = \boldsymbol{x}_{tmp}$;
   Construct surrogate model $\boldsymbol{s}_{tmp}(\boldsymbol{x})$ using dataset $\{\boldsymbol{x}_B^{(k)}, \boldsymbol{f}_P(\boldsymbol{x}_B^{(k)})\}_{k = 1, \ldots, j}$;
   Estimate model error E_{tmp} using K-fold cross-validation [93], $K = \min\{j, 10\}$;
   if $E_{tmp} < E_{\max}$
   Go to 8;
   end
   end
7. Go to 3;
8. Return $\boldsymbol{s}^{(0)}(\boldsymbol{x}) = \boldsymbol{s}_{tmp}(\boldsymbol{x})$;

Fig. 4. Parameter space pre-screening and training data generation for initial surrogate model construction.

Infill point generation

The core stage of the size reduction process is launched upon constructing the initial surrogate $\boldsymbol{s}^{(0)}$, as described in Sect. 3.2. It iteratively produces infill points $\boldsymbol{x}^{(i+1)}$, $i=0, 1, 2, \ldots$, and the updated surrogate models $\boldsymbol{s}^{(j)}$, $j=1, 2, \ldots$.

The role of the surrogates is to provide predictions about the minimum-size design, which are obtained by optimizing the current metamodel in a global sense, as

$$\mathbf{x}^{(i+1)} = \arg\min_{\mathbf{x}\in X} U_F(\mathbf{x}, \mathbf{s}^{(i)}(\mathbf{x})) \tag{7}$$

The problem (7) is equivalent to (5), except that the vector $\boldsymbol{f}_P(\boldsymbol{x})$ at $\boldsymbol{x}$ is predicted using the surrogate $\boldsymbol{s}^{(i)}$. Further, (7) is solved using particle swarm optimization (PSO)[78]. At this juncture, the specific choice of the global optimization routine is of secondary significance, as the surrogate model is cost-effective to evaluate. PSO (or any other nature-inspired algorithm) can be configured with a substantial computational budget. Furthermore, the optimization problem is formulated in terms of response features, rendering it numerically more manageable than the conventional approach (refer to Sect. 2).

From a machine learning standpoint, the procedure outlined in Eq. (7) aligns with the infill criterion being grounded on the predicted improvement in the objective function[79]. The rationale for selecting this criterion is twofold: (i) due to the feature-based formulation, the initial surrogate model typically exhibits a sufficiently high quality, and (ii) the promising subset of the search space has already been identified through the pre-screening process detailed in Sect. 3.2. The surrogate model undergoes refinement after each iteration.

In particular, the infill point $\boldsymbol{x}_B^{(i+1)}$ generated by solving (7) is added, along with the corresponding response feature vector $\boldsymbol{f}_P(\boldsymbol{x}_B^{(i+1)})$, to the training dataset, which becomes $\{\boldsymbol{x}_B^{(k)}, \boldsymbol{f}_P(\boldsymbol{x}_B^{(k)})\}_{k=1,\ldots,Ninit+j}$, where $\boldsymbol{x}_B^{(Ninit+j)} = \boldsymbol{x}^{(i)}$ for $j=1, 2, \ldots$. The surrogate model is rebuilt using this set and employed as a predictor in the next iteration.

The following termination conditions are utilized (as logical alternatives): (i) convergence in argument, i.e., $||\boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}|| < \varepsilon$, (ii) no objective function improvement over the last $N_{no\_improve}$ iterations. The default control parameter values are $\varepsilon = 10^{-3}$ and $N_{no\_improve} = 10$.

Optimization algorithm

This section provides an overview of the procedural workflow for the proposed method of global feature-based size reduction of microwave components. We commence by delving into the control parameters, which are consolidated in Table 4. It is crucial to underscore that there are only three parameters, with two being associated with the termination condition. These parameters primarily govern the requisite resolution of the search process. For typical microwave components featuring geometry parameters expressed in millimeters, a resolution of 0.001 mm is more than sufficient. The last parameter determines the necessary predictive power of the initial surrogate. A 10% relative error (default value) is a mild condition; however, given that the model is rendered using characteristic points of the system outputs, this condition can typically be achieved with less than a hundred samples (also due to the pre-screening procedure), resulting in computational efficiency for the entire framework.

Figure 5 shows the pseudocode of the complete algorithm. For auxiliary elucidation, Fig. 6 provides its flow diagram. Steps 2 and 3 constitute the initial stage of the process that include parameter space pre-screening and a rendition of the initial metamodel. Both were elaborated on in Sect. 3.2. The core part of the size reduction procedure are Steps 5 through 8. Therein, a series of optimum design approximations are generated using the assumed infill criterion (predicted objective function improvement), and the surrogate model is re-built based on the EM analysis data acquired thus far. The process is continued until the termination condition has been fulfilled.

At this point, a few comments should be made concerning the implementation and the complexity of the proposed procedure. The underlying programming environment is Matlab. As indicated earlier, there are several algorithmic components that include randomized pre-screening, definition and extraction of the feature points, surrogate model construction and its (global) optimization, as well as interfacing EM simulation software. Most of these components are straightforward to implement (pre-screening, surrogate model rendition, model optimization). In particular, in this work, kriging interpolation is employed as an underlying modelling approach. It is widely available through third-party toolboxes[82–84] (here, we use the Matlab toolbox SUMO[82]). Surrogate model optimization is carried out using PSO, which is one of the most popular bio-inspired techniques. This algorithm is easy to implement, and although a plethora of implementations are available, our own Matlab implementation is employed here. There are two algorithmic components which are less generic: (i) the interface between the programming language (here, Matlab) and EM simulation software (here, CST Microwave Studio), which is implemented using Visual Basic to allow batch simulations and automated data acquisition from the EM solver; (ii) a procedure for extracting response features. The last component is the only one which is problem-dependent, and generally changes for a different arrangement of the response features.

Parameter	Meaning	Default value
E_{max}	Maximum value of relative root-mean square error of the initial surrogate (error estimated using cross-validation), cf. Section 3.2	10%
ε	Termination threshold for convergence in argument, cf. Section 3.3	10^{-2}
$N_{no\_improve}$	Termination threshold for no objective function value improvement, cf. Section 3.3	10

Table 4. Control parameters of the presented algorithm for globalized size reduction procedure.

1. Input parameters:
   - Design specifications;
   - Definition of the characteristic points $\boldsymbol{f}_P$ and the merit function U_F (cf. Section 3.1);
   - Design space X (interval $[\boldsymbol{l}\ \boldsymbol{u}]$, with $\boldsymbol{l}$ and $\boldsymbol{u}$ being lower and upper designable variable bounds);
   - Required modelling error E_{max};
   - Circuit size bounds A_{min} and A_{max} (optional);
   - Termination thresholds ε and $N_{no\_improve}$;
2. Generate the set of initial samples $\{\boldsymbol{x}_B^{(k)}, \boldsymbol{f}_P(\boldsymbol{x}_B^{(k)})\}_{k = 1,\ldots,Ninit}$, as described in Section 3.2 (cf. Fig. 4);
3. Construct initial surrogate model $\boldsymbol{s}^{(0)}(\boldsymbol{x})$;
4. Set $i = 0$;
5. Obtain infill point $\boldsymbol{x}^{(i+1)}$ by solving (7) using the PSO algorithm:
$$\boldsymbol{x}^{(i+1)} = \arg\min_{\boldsymbol{x}\in X} U_F(\boldsymbol{x}, \boldsymbol{s}^{(i)}(\boldsymbol{x}))$$
6. Update the dataset: $\{\boldsymbol{x}_B^{(k)}, \boldsymbol{f}_P(\boldsymbol{x}_B^{(k)})\}_{k = 1, \ldots, Ninit + j}$, with $\boldsymbol{x}_B^{(Ninit+j)} = \boldsymbol{x}^{(j)}$ for $j = 1, 2, \ldots$.
7. Set $i = i + 1$;
8. Construct the metamodel model $\boldsymbol{s}^{(i)}(\boldsymbol{x})$ based on the updated dataset;
9. **if** $||\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(i-1)}|| < \varepsilon$ OR no objective function improvement for $N_{no\_improve}$ iterations
   Go to 11;
   end
10. Go to 5;
11. Return $\boldsymbol{x}^* = \boldsymbol{x}^{(i)}$;

Fig. 5. Operating flow of the suggested global feature-based size reduction algorithm.

However, the feature selection and extraction only vary between different types of microwave circuits (e.g., couplers, filters, power dividers). Consequently, a library of extraction procedures may be implemented and employed as necessary because the number of different types of circuits is rather limited. Apart from this aspect, the remaining part of the procedure is generic and does not require re-implementation for different design scenarios.

Results and benchmarking

In this part of the paper, we examine the characteristics of the size reduction framework outlined in Sect. 3. The algorithm is deployed to obtain minimum-size designs for two microstrip couplers. Additionally, it is juxtaposed against several state-of-the-art methods, encompassing nature-inspired optimization, local (gradient-based) tuning, and a machine learning procedure directly handling the frequency characteristics of the circuits under consideration. The aim of the numerical experiments is to evaluate the operation of both the proposed method and selected state-of-the-art techniques in terms of achievable miniaturization rate, precision in controlling design constraints, and computational efficiency.

Test cases

Our numerical experiments are conducted using a compact rat-race coupler[80] (Circuit I), and a compact branch-line coupler with CMRCs[81] (Circuit II). The circuit geometries as well as the essential parameters have been shown in Figs. 7 and 8, respectively. The EM simulation models are carried out using CST Microwave Studio. The primary design goal is minimization of the footprint area $A(\boldsymbol{x})$. Furthermore, we impose the following constraints: (i) $h_1(\boldsymbol{x}) = |S_{31}(\boldsymbol{x},f_0) - S_{21}(\boldsymbol{x},f_0)|$, and (ii) $g_1(\boldsymbol{x}) = \max\{f \in F : \max\{|S_{11}(\boldsymbol{x},f)|, |S_{41}(\boldsymbol{x},f)|\}\} + 20$ dB; therein, f_0 stands for the center frequency, whereas F is the target bandwidth. The first constraint is introduced to maintain equal power division, whereas the second is to secure the assumed -20 dB bandwidth for $|S_{11}|$ and $|S_{41}|$. Both constraints are controlled using penalty functions, cf. (2), (3).

The penalty coefficients are set to 100 and 1,000 for the equality and inequality constraint, respectively. These values provide a good trade-off between the miniaturization rate and the quality of constraint handling. In particular, given that the typical circuit size is a few hundred mm^2, the contribution of the penalty term (cf. (3)) is 2.5 for inequality constraint violation of 1 dB (corresponding to relative violation of 0.05), and quickly goes up if the violation increases. This setup is generally sufficient to ensure that the constraint violation at the optimized

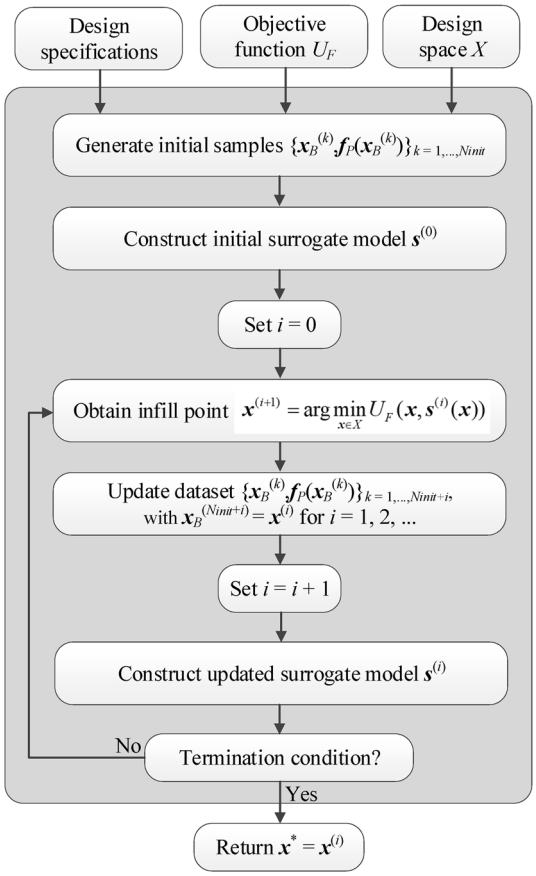


Fig. 6. Flow diagram of the proposed algorithm for global size reduction of microwave passive components.

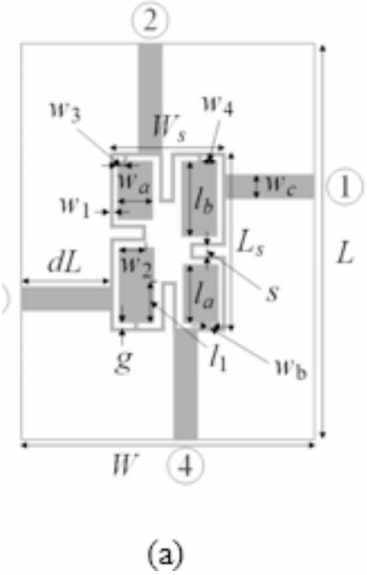
(a)

Substrate	RO4003 ($\varepsilon_r = 3.38$, $h = 0.76$ mm)
Design variables	$\boldsymbol{x} = [l_1\ l_2\ l_3\ d\ w\ w_1]^T$
Other variables	$d_1 = d + \lvert w - w_1 \rvert$, $d = 1.0$, $w_0 = 1.7$, $l_0 = 15$
Design goal	Minimize circuit footprint area $A(\boldsymbol{x})$
Design constraints	• $h_1(\boldsymbol{x}) = \lvert S_{31}(\boldsymbol{x}, f_0) - S_{21}(\boldsymbol{x}, f_0) \rvert$ dB; • $g_1(\boldsymbol{x}) = \max\{f \in F : \max\{\lvert S_{11}(\boldsymbol{x}, f) \rvert, \lvert S_{41}(\boldsymbol{x}, f) \rvert\}\} + 20$ dB, where $f_0 = 1$ GHz is the center frequency, and $F = [0.95\ 1.05]$ GHz is the intended circuit bandwidth.
Parameter space	$\boldsymbol{l} = [0.5\ 5.0\ 5.0\ 0.2\ 0.2\ 0.2]^T$ $\boldsymbol{u} = [15.0\ 30.0\ 50.0\ 2.0\ 2.0\ 2.0]^T$

(b)

Fig. 7. Circuit I[80]: (a) geometry, (b) essential parameters.

design does not exceed one dB or so, which is practically acceptable. Similarly, the penalty coefficient of 100 for the equality constraints brings the penalty of one for absolution power split ration violation of 0.1 dB, which is again sufficient for practical purposes. It should also be noted that the parameters spaces are large for both circuits: the average upper-to-lower variable bound ratio is as high as about thirteen for Circuit I, and almost seven for Circuit II.



(a)

Substrate	AD300 ($\varepsilon_r = 2.97$, $h = 0.76$ mm)
Design variables	$\boldsymbol{x} = [g\ l_{1r}\ l_a\ l_b\ w_1\ w_{2r}\ w_{3r}\ w_{4r}\ w_a\ w_b]^T$
Other variables	$L = 2dL + L_s$, $L_s = 4w_1 + 4g + s + l_a + l_b$, $W = 2dL + W_s$, $W_s = 4w_1 + 4g + s + 2w_a$, $l_1 = l_b l_{1r}$, $w_2 = w_a w_{2r}$, $w_3 = w_{3r} w_a$, $w_4 = w_{4r} w_a$, $w_c = 1.9$ mm
Design goal	Minimize circuit footprint area $A(\boldsymbol{x})$
Design constraints	• $h_1(\boldsymbol{x}) = \lvert S_{31}(\boldsymbol{x}, f_0) - S_{21}(\boldsymbol{x}, f_0)\rvert$ dB; • $g_1(\boldsymbol{x}) = \max\{f \in F : \max\{\lvert S_{11}(\boldsymbol{x}, f)\rvert, \lvert S_{41}(\boldsymbol{x}, f)\rvert\}\} + 20$ dB, where $f_0 = 1.5$ GHz is the center frequency, and $F =$ [1.45 1.55] GHz is the intended circuit bandwidth.
Parameter space	$\boldsymbol{l} = [0.4\ 0.1\ 3.0\ 3.0\ 0.4\ 0.1\ 0.1\ 0.1\ 2.0\ 0.2]^T$ $\boldsymbol{u} = [1.0\ 0.99\ 15.0\ 25.0\ 1.5\ 0.99\ 0.9\ 0.9\ 12.0\ 1.0]^T$

(b)

Fig. 8. Circuit II[81]: (a) geometry, (b) essential parameters.

Algorithm		Algorithm type	Setup
I		Trust-region gradient based optimizer	Algorithm setup: • Random initial design; • Response gradients estimated using finite differentiation; • Termination criteria based on convergence in argument and reduction of the trust region size.
II	Particle swarm optimizer (PSO)	Algorithm setup: • Swarm size $N = 10$, • Standard control parameters ($\chi = 0.73$, $c_1 = c_2 = 2.05$); • Number of iterations set to 100.	
III	Differential evolution (DE)[85]	Algorithm setup: • Population size $N = 10$; • Standard control parameters (crossover probability $CR =$ 0.9, differential weight $F = 0.8$); • Number of iterations set to 100.	
IV		Machine learning procedure	Algorithm similar to that of Sect. 3: • Initial surrogate set up to ensure relative RMS error not higher than 10% with the maximum number of training samples equal to 400; • Optimization based on processing the circuit frequency characteristics (unlike response features in the proposed procedure); • Infill criterion: minimization of the projected objective function improvement[79].

Table 5. Benchmark algorithms.

Results

The presented methodology has been employed to optimize Circuits I and II using the setup of Table 4, i.e., $E_{\max} = 10\%$, $\varepsilon = 10^{-2}$, and $N_{no\_improve} = 10$. For the sake of comparison, three benchmark techniques have been employed, as specified in Table 5. These include:

- A gradient-based optimizer, utilized to demonstrate multimodality of the considered design task, the latter justifying the need for global search. The algorithm is initiated from random starting points.
- A bio-inspired algorithm, utilized to demonstrate the challenges of nature-inspired size reduction of microwave components. Particle swarm optimizer (PSO) is chosen as a representative procedure. The computational budget is kept at 1,000 objective function evaluations. This number is low for this class of algorithms, yet, it is high from the perspective of EM-driven design (the typical algorithm run would take two to three days to complete).
- A population-based search procedure, differential evolution (DE), selected as one of the most successful methods for continuous optimization. Here, the computational budget is also kept at 1,000 objective function evaluations for the same reasons as in the case of PSO.
- A machine learning procedure that employs the same infill criterion and surrogate model (kriging) as the algorithm of Sect. 3. However, it works at the level of complete circuit characteristics. This algorithm is incorporated into the benchmark set to illustrate the benefits of integrating the response feature technology into the optimization framework proposed in this work. The values of control parameters of the PSO and DE

Optimization algorithm	Performance figure						
	Circuit size A [mm^2] [1]	Sth(A) [mm^2] [2]	Inequality constraint		Equality constraint		CPU cost[7]
			Violation g_1 [dB] [3]	Sth(g_1) [dB] [4]	Violation h_1 [dB] [5]	Sth(h_1) [dB] [6]	
Algorithm I	378.0	59.3	4.5	4.3	0.2	0.2	63 [2.8 h]
Algorithm II	543.1	86.8	−1.0	1.6	0.1	0.1	1,000 [44.4 h]
Algorithm III	528.3	77.2	2.1	1.8	0.3	0.2	1,000 [44.4 h]
Algorithm IV	500.9	25.6	12.3	4.8	6.6	4.4	417.4 [18.5 h]
Machine learning with response features (this work)	354.3	8.0	0.4	0.2	0.04	0.03	154 [6.8 h]

Table 6. Circuit I: optimization results. [1]Optimized footprint area of the circuit averaged over ten algorithm runs.[2]Standard deviation of the optimized footprint area averaged over ten algorithm runs.[3]Violation of inequality constraint, defined as $g_1=\max\{f\varepsilon F: \max\{|S_{11}(x,f)|,|S_{41}(x,f)|\}\}+20$ dB averaged over ten algorithm runs.[4]Standard deviation of the constraint violation g1, averaged over ten algorithm runs.[5]Violation of equality constraint, defined as $h_1=|S_{31}(x,f_0)|- |S_{21}(x,f_0)|$ dB averaged over ten algorithm runs.[6]Standard deviation of the constraint violation h1, averaged over ten algorithm runs.[7]Cost expressed in terms of equivalent number of EM analyses. Numbers in bracket correspond to the running time in hours.

Optimization algorithm	Performance figure						
	Circuit size A [mm^2] [1]	Sth(A) [mm^2] [2]	Inequality constraint		Equality constraint		CPU cost[7]
			Violation g_1 [dB] [3]	Std(g_1) [dB] [4]	Violation h_1 [dB] [5]	Std(h_1) [dB] [6]	
Algorithm I	295.1	24.7	3.6	1.9	0.2	0.1	77 [5.2 h]
Algorithm II	541.5	240.4	5.5	6.8	0.7	0.1	1,000 [66.7 h]
Algorithm III	573.2	112.3	3.8	3.6	0.5	0.3	1,000 [66.7 h]
Algorithm IV	270.3	13.3	6.3	3.6	1.0	2.2	398.4 [26.6 h]
Machine learning with response features (this work)	293.9	15.3	0.7	0.2	0.0	0.06	110.0 [7.3 h]

[1]Optimized footprint area of the circuit averaged over ten algorithm runs.[2]Standard deviation of the optimized footprint area averaged over ten algorithm runs.[3]Violation of inequality constraint, defined as $g_1=\max\{f\varepsilon F: \max\{|S_{11}(x,f)|,|S_{41}(x,f)|\}\}+20$ dB averaged over ten algorithm runs.[4]Standard deviation of the constraint violation g1, averaged over ten algorithm runs.[5]Violation of equality constraint, defined as $h_1=|S_{31}(x,f_0)|- |S_{21}(x,f_0)|$ dB averaged over ten algorithm runs.[6]Standard deviation of the constraint violation h1, averaged over ten algorithm runs.[7]Cost expressed in terms of equivalent number of EM analyses. Numbers in bracket correspond to the running time in hours

Table 7. Circuit II: optimization results.

algorithms are standard as found in the literature. The assumption is that a typical used does not tune the algorithm for a particular task at hand but employs the most conventional parameter values. Regarding Algorithm IV, the acceptance threshold of 10% is chosen because this level of relative RMS error normally corresponds to almost decent visual alignment between surrogate-predicted and EM-simulated circuit responses and can therefore be treated as a sufficient starting point for further ML-based search process. The computational budget of 400 samples for initial surrogate model construction is selected for purely practical reasons (to avoid excessive costs of model rendition given a relatively long EM simulation time).

The investigation into the repeatability of solutions involved conducting ten independent runs of each algorithm. Tables 6 and 7 present numerical results, offering average values and standard deviations of the circuit size, design constraint violations, and the average computational cost of the optimization process. Figures 9 and 10 visually depict the circuit responses for the optimized designs obtained in selected runs of the proposed procedure, highlighting the evolution of the circuit footprint area and the values of design constraints.

Discussion

The results gathered in Sect. 4.2 provide an extensive performance assessment of the explicit circuit miniaturization framework, as well as how it compares to the three benchmark algorithms outlined earlier.

To start, it is noteworthy to examine the outcomes obtained using the gradient search procedure (Algorithm I). This method demonstrates poor repeatability of solutions; for instance, the standard deviation of the footprint

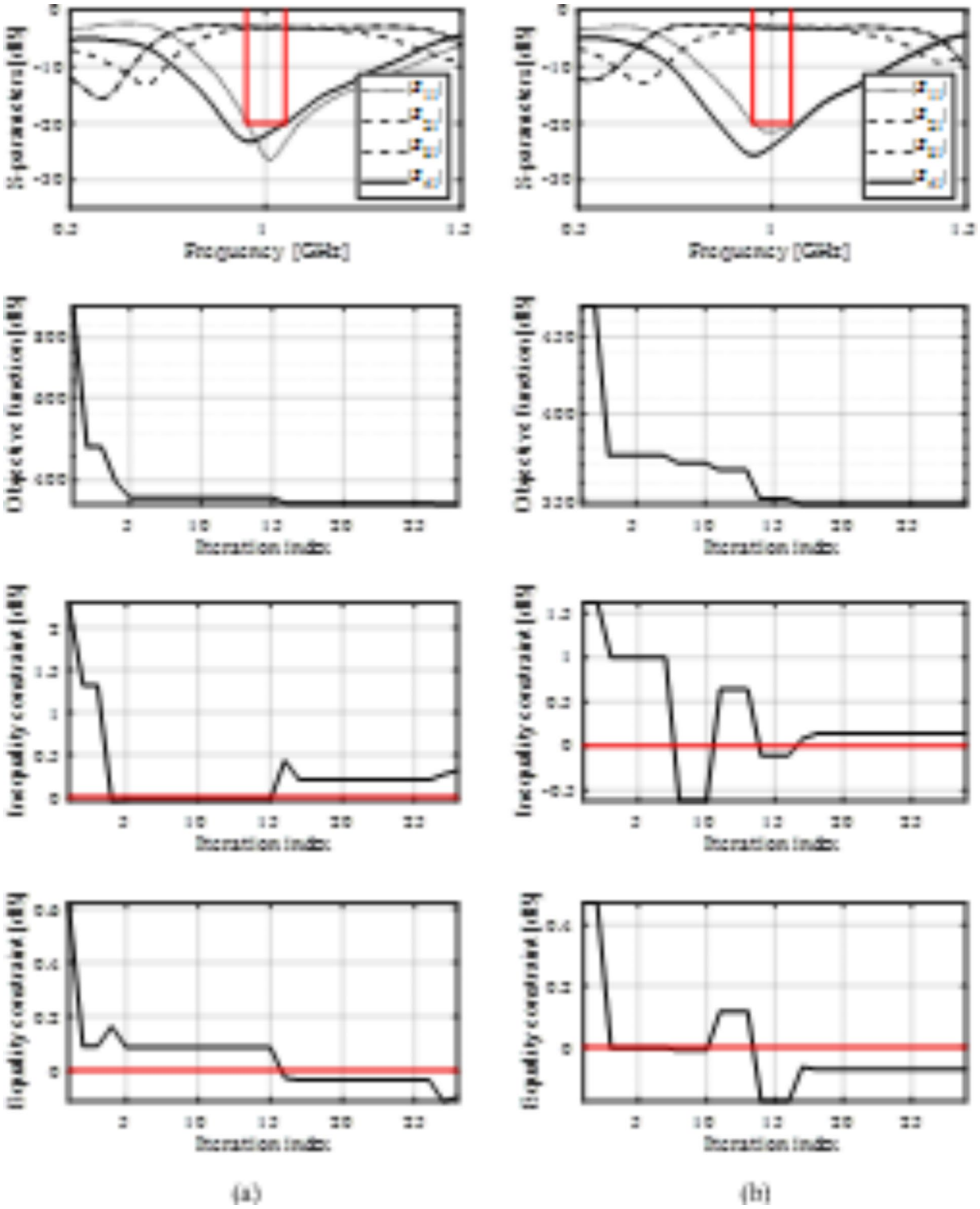


Fig. 9. *S*-parameters of Circuit I at the optimum designs obtained using our size reduction framework (top), circuit size versus iteration index (middle), and design constraint violations versus iteration index (bottom), shown for the chosen algorithm executions: (a) run 1, (b) run 2. The iteration counter starts after constructing the initial surrogate model. Vertical and horizontal lines indicate the intended operating frequency range (here, from 0.95 GHz to 1.05 GHz), and the acceptance level for $|S_{11}|$ and $|S_{41}|$ (here, − 20 dB).

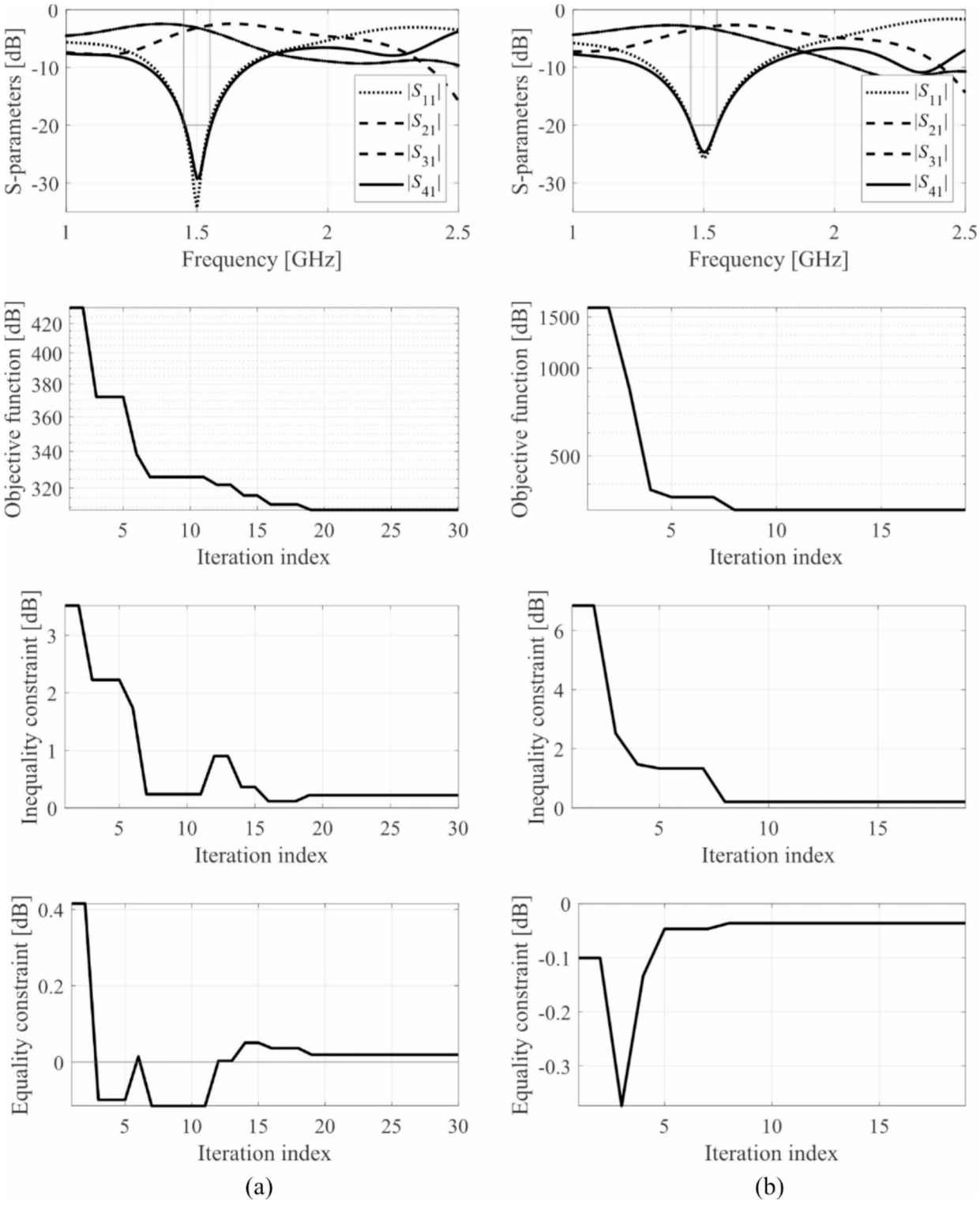


Fig. 10. *S*-parameters of Circuit II at the optimum designs obtained using our size reduction framework (top), circuit size versus iteration index (middle), and design constraint violations versus iteration index (bottom), shown for the chosen algorithm executions: (a) run 1, (b) run 2. The iteration counter starts after constructing the initial surrogate model. Vertical and horizontal lines indicate the intended operating frequency range (here, from 1.45 GHz to 1.55 GHz), and the acceptance level for $|S_{11}|$ and $|S_{41}|$ (here, − 20 dB).

area exceeds fifteen and ten% for Circuit I and II, respectively. This observation indicates the presence of multiple local optima for the considered test problems. Given that the results are heavily dependent on the initial design, the use of global optimizers becomes essential. Moreover, while the average circuit size produced by gradient search is small, the constraint control is subpar, especially for the inequality condition, with an average violation of about four decibels.

Figure 9. *S*-parameters of Circuit I at the optimum designs obtained using our size reduction framework (top), circuit size versus iteration index (middle), and design constraint violations versus iteration index (bottom), shown for the chosen algorithm executions: (a) run 1, (b) run 2. The iteration counter starts after constructing the initial surrogate model. Vertical and horizontal lines indicate the intended operating frequency range (here, from 0.95 GHz to 1.05 GHz), and the acceptance level for $|S_{11}|$ and $|S_{41}|$ (here, −20 dB).

Algorithms II (particle swarm optimizer) and III (differential evolution) perform even worse. On the one hand, the average circuit sizes are considerable larger than for the remaining methods, just as are their standard deviation. On the other hand, the constraint control is mediocre, although not as bad as for Algorithm I. This level of performance is—in a large part—a result of the assigned computational budget (only 1,000 objective function calls). Arguably, increasing the budget to five or ten thousands of objective function evaluations would improve the results of Algorithms II and III significantly. Unfortunately, this is impractical or even prohibitive due to a typical algorithm run taking several weeks to conclude.

The operation of Algorithm IV (ML working with the complete circuit responses) is similarly poor. Although the surrogate model accuracy threshold E_{max} of 10% seems reasonable, the actual predictions (especially concerning design constraints) are impeded by high nonlinearity of the circuit frequency characteristics. Consequently, the resulting constraint violations are the highest among the considered algorithms, while the obtained circuit sizes are the smallest. Meanwhile, the average number of data points needed to secure the error level of E_{max} is over 350, and it is the major contribution to the overall CPU expenses of the search process. It is evident that improving the design quality would require setting up considerably lower accuracy threshold (e.g., two or three%), which, however, would also lead to significantly higher CPU expenses.

The algorithm proposed in this work performs better than the benchmark methods. Its most important property is that it delivers truly minimum-size designs with excellent constraint control. The average violation of the inequality constraint is a fraction of decibel, whereas the violation of the equality constraint is essentially zero. This places the designs produced by the algorithm on the feasible region boundary. At the same time, the achieved footprint areas are competitive (it should be stressed that smaller sizes found by Algorithms I and IV are strongly infeasible), whereas repeatability of solutions is good: the footprint area standard deviation amounts to only two and five% on the average for Circuit I and II, respectively. It should be recalled that design feasibility refers to a situation where all design constraints are satisfied. For inequality constraints, the reflection and port isolation ($|S_{11}|$ and $|S_{41}|$, respectively) do not exceed −20 dB within the frequency range of interest. The equality constraints are satisfied if the power split ratio is precisely equal to the target value (here, 0 dB, i.e., equal power division). Rendering the exact required value of the power split is impossible, so no design can be formally feasible. However, power split errors at the level of 0.1 dB (or about 1% in relative terms, i.e., in terms of relative deviation from 50/50 power division) are considered feasible for practical purposes. For the inequality constraint, when the target level is set at −20 dB, violations of up to 1 dB are also acceptable. Keeping this in mind, "strong infeasibility" is understood by excessive violation of constraints, which in the case of the inequality constraint is more than 1–2 dB, and in the case of the equality constraint, it is more than 0.2–0.3 dB. The latter is because the power split ratio should be more precise than the impedance matching and port isolation levels.

A separate note should be made concerning standard deviations of both the achieved circuit size and constraint violations for the proposed and the benchmark algorithms. High standard deviation for Algorithms I, II, and III are related to two factors. For Algorithm I, the primary reason is that this method is a local one; consequently, it identified the local optimum which is typically nearest to the starting points (the latter being randomly selected). Most of these local optima are different from each other and many are of inferior quality. For Algorithms II and III, the principal issue is relatively limited computational budget, meaning that the algorithms were unable to properly converge. At the same time, this budget (1,000 EM simulations) is significant from practical perspective. Thus, the results indicate that direct population-based size reduction is not an attractive option. Algorithm IV performs better with this respect; however, its standard deviations are generally higher than for the proposed algorithm. Here, the underlying reason is limited accuracy of the surrogate model leading to limited-accuracy predictions of the circuit characteristics during the machine learning search process. On the other hand, the proposed approach constructs the model in a pre-screened parameter space region and leverages the properties of the response features, both leading to considerably better reliability. The latter translates into improved consistency of the results and lower standard deviation values.

In regard to computational efficiency, our methodology is evidently more time-consuming than local search; however, the differences are not substantial (seven hours versus three hours for Circuit I, and seven hours versus five hours for Circuit II). It is important to note that our technique provides global search capability. In comparison to global design procedures, our algorithm demonstrates an average optimization cost that is over six and about nine times faster than PSO/DE for Circuit I and II, respectively, and approximately three times faster than the machine learning procedure for both circuits. Therefore, the computational advantages are undeniable.

Given the comments formulated above, one can conclude that the suggested procedure exhibits properties that are attractive from the perspective of practical global optimization of microwave passives. It does not only offer quasi-global search capability, but as a result of incorporating the response features and the pre-screening stage, the computational efficiency is better than that of representative nature-inspired and machine learning methods.

It should also be emphasized that it is not possible to compare the circuit size before and after optimization. This is because the optimization process is global and there is no initial design that the optimized design might be compared to. In particular, the search process starts from randomized pre-screening used to acquire data for initial surrogate model construction. This data contains a mixture of feasible and infeasible designs (mostly infeasible), and it is not possible to find any reasonable point of reference. Only during the optimization process, the feasibility of designs is gradually improved and so is the size reduction. Consequently, we may only compare the circuit sizes obtained by different algorithms (the proposed one and the benchmark).

Conclusion

This paper proposed an innovative technique for explicit miniaturization of microwave passive components. Our method involves the response feature approach to perform a design space pre-screening, and to facilitate a rendition of dependable surrogate model. The latter is employed to identify the position of the optimum design. The search process in embedded in a machine learning framework that uses kriging interpolation as a surrogate modelling technique, particle swarm optimizer to generate the sequence of designs approximating the minimum-size design, and predicted merit function improvement as the infill criterion. The design constraints are controlled implicitly by means of penalty functions. Numerical experiments conducted using two compact microstrip couplers demonstrate competitive performance of our procedure w.r.t. the design quality, repeatability of solutions, and computational efficiency. In particular, our algorithm enables generation of small-size designs with improved control of constraint violations, and good consistency: the standard deviation of the circuit footprint area is lower than 5% of its mean value. The CPU cost associated with circuit optimization is slightly above 100 EM simulations (150 and 110 for the first and the second verification circuit), which is significantly less than for the benchmark methods.

A comparison with a machine-learning framework that processes complete circuit characteristics corroborates the benefits of incorporating response features, both w.r.t. the improved surrogate model accuracy but also computational efficiency of the search process. It should be mentioned that the employment of this technology is also a source of potential limitations of the method. On the one hand, the very definition and extraction of characteristic points is problem dependent. On the other hand, for larger parameter spaces, the likelihood of generating random observables with extractable feature points is diminished, which would lead to increasing the cost of the size reduction process, contributed to in a large part by the pre-screening step. Notwithstanding, this issue would not be pronounced if the parameter space is reasonably defined, e.g., using the designer's insight established at the stage of developing the circuit geometry. This means, among others, that the parameter bounds are set up to focus the optimization process on designs that are likely to be of decent quality, and to eliminate parameter combinations that evidently (i.e., according to the designer's experience) have no chance to fulfill the assumed specifications imposed on electrical characteristics (e.g., physically too small). For the specific test cases considered in this work, the design spaces were quite extensive with respect to the parameter ranges, yet, the algorithm managed to identify the minimum-size designs in a consistent manner.

In conclusion, the size-reduction framework introduced in this study emerges as a possible alternative to state-of-the-art methods, particularly when the CPU budget assigned to the optimization process is a key consideration. Aside from the necessity to individually define feature points in a problem-dependent manner, the algorithm is undemanding to implement and configure, requiring no tailoring of control parameters.

One of the objectives of the future work is to investigate the properties and performance of the proposed algorithm for high-dimensional parameter spaces ($n > 10$). Another topic of interest will be extending the range of applicability of the presented framework to other types of microwave components (e.g., filters, power dividers, multi-band structures). As the algorithm itself is generic, the main challenge will be appropriate definition and extraction of the response features, which must be carried out individually for each type of circuit.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 10 July 2024; Accepted: 16 September 2024
Published online: 28 September 2024

References

1. Zhang, H. et al. „A low-profile compact dual-band L-shape monopole antenna for microwave thorax monitoring. *IEEE Ant Wirel. Propag. Lett.***19** (3), 448–452 (2020).
2. Hu, Y. Y., Sun, S. & Xu, H. Compact collinear quasi-yagi antenna array for wireless energy harvesting. *IEEE Access.***8**, 35308–35317 (2020).
3. Zhu, F., Luo, G. Q., Liao, Z., Dai, X. W. & Wu, K. „Compact dual-mode bandpass filters based on half-mode substrate-integrated waveguide cavities. *IEEE Microw. Wirel. Comp. Lett.***31** (5), 441–444 (2021).
4. Matos, D., da Cruz, M. D., Jordão, R., Correia & Carvalho, N. B. „Millimeter-wave BiCMOS backscatter modulator for 5 G-IoT applications. *IEEE Microw. Wirel. Comp. Lett.***31** (2), 173–176 (2021).
5. Wu, C., Yuan, J. & Chen, Z. A UHF RFID tag antenna placeable on a metal surface without degraded performances. *IEEE Ant Wirel. Propag. Lett.***23** (7), 2101–2105 (2024).
6. Martinez, L., Belenguer, A., Boria, V. E. & Borja, A. L. Compact folded bandpass filter in empty substrate integrated coaxial line at S-Band. *IEEE Microw. Wirel. Comp. Lett.***29** (5), 315–317 (2019).
7. Chen, S. et al. A frequency synthesizer based microwave permittivity sensor using CMRC structure. *IEEE Access.***6**, 8556–8563 (2018).
8. Qin, W. & Xue, Q. Elliptic response bandpass filter based on complementary CMRC. *Electr. Lett.***49** (15), 945–947 (2013).
9. Sen, S. & Moyra, T. Compact microstrip low-pass filtering power divider with wide harmonic suppression. *IET Microwaves Ant Propag.***13** (12), 2026–2031 (2019).

10. Karimbu Vallappil, A., Rahim, M. K. A., Khawaja, B. A. & Iqbal, M. N. Compact metamaterial based 4×4 Butler matrix with improved bandwidth for 5G applications. *IEEE Access*.**8**, 13573–13583 (2020).
11. Liu, S. & Xu, F. Compact multilayer half mode substrate integrated waveguide 3-dB coupler. *IEEE Microw. Wirel. Comp. Lett.***28** (7), 564–566 (2018).
12. Wei, F., Jay Guo, Y., Qin, P. & Wei Shi, X. Compact balanced dual- and tri-band bandpass filters based on stub loaded resonators. *IEEE Microw. Wirel. Comp. Lett.***25** (2), 76–78 (2015).
13. Zhang, W., Shen, Z., Xu, K. & Shi, J. A compact wideband phase shifter using slotted substrate integrated waveguide. *IEEE Microw. Wirel. Comp. Lett.***29** (12), 767–770 (2019).
14. Yang, D., Zhai, H., Guo, C. & Li, H. A compact single-layer wideband microstrip antenna with filtering performance. *IEEE Antennas Wirel. Propag. Lett.***19** (5), 801–805 (2020).
15. Chen, Z., Xu, R. & Shen, Z. Design of a broadband antenna array with compact surface-wave antenna elements. *IEEE Ant Wirel. Propag. Lett.***21** (2), 337–340 (2022).
16. Chen, C. A compact wideband filtering omnidirectional dipole antenna without extra circuits. *IEEE Trans. Ant Propag.***70** (3), 1729–1739 (2022).
17. Hosseini, S. M. & Shishegar, A. A. Shape optimization for enhancing the registration of prior-based microwave imaging techniques and improving the dielectric property retrieval. *IEEE Trans. Ant Propag.***72** (4), 3620–3628 (2024).
18. Zhang, F., Li, J., Lu, J. & Xu, C. Optimization of circular waveguide microwave sensor for gas-solid two-phase flow parameters measurement. *IEEE Sens. J.***21** (6), 7604–7612 (2021).
19. Feng, F. et al. Parallel gradient-based EM optimization for microwave components using adjoint- sensitivity-based neuro-transfer function surrogate. *IEEE Trans. Microw. Theory Techn*. **68** (9), 3606–3620 (2020).
20. Kolda, T. G., Lewis, R. M. & Torczon, V. Optimization by direct search: new perspectives on some classical and modern methods. *SIAM Rev.***45**, 385–482 (2003).
21. Koziel, S., Pietrenko-Dabrowska, A. & Al-Hasan, M. Improved-efficacy optimization of compact microwave passives by means of frequency-related regularization. *IEEE Access*.**8**, 195317–195326 (2020).
22. Ochoa, J. S. & Cangellaris, A. C. Random-space dimensionality reduction for expedient yield estimation of passive microwave structures. *IEEE Trans. Microw. Theory Techn*. **61** (12), 4313–4321 (2013).
23. Yu, Y. et al. State-of-the-art: AI-assisted surrogate modeling and optimization for microwave filters. *IEEE Trans. Microw. Theory Techn*. **70** (11), 4635–4651 (2022).
24. Güneş, F., Uluslu, A. & Mahouti, P. Pareto optimal characterization of a microwave transistor. *IEEE Access*.**8**, 47900–47913 (2020).
25. Zhang, Z., Cheng, Q. S., Chen, H. & Jiang, F. An efficient hybrid sampling method for neural network-based microwave component modeling and optimization. *IEEE Microw. Wirel. Comp. Lett.***30** (7), 625–628 (2020).
26. Van Nechel, E., Ferranti, F., Rolain, Y. & Lataire, J. Model-driven design of microwave filters based on scalable circuit models. *IEEE Trans. Microw. Theory Techn*. **66** (10), 4390–4396 (2018).
27. Li, Y., Xiao, S., Rotaru, M. & Sykulski, J. K. A dual kriging approach with improved points selection algorithm for memory efficient surrogate optimization in electromagnetics. *IEEE Trans. Magn.*, **52**, 3, pp. 1–4, Art 7000504, 2016.
28. Spina, D., Ferranti, F., Antonini, G., Dhaene, T. & Knockaert, L. Efficient variability analysis of electromagnetic systems via polynomial chaos and model order reduction. *IEEE Trans. Comp. Packaging Manuf. Techn*. **4** (6), 1038–1051 (2014).
29. Zhao, P. & Wu, K. Homotopy optimization of microwave and millimeter-wave filters based on neural network model. *IEEE Trans. Microw. Theory Techn*. **68** (4), 1390–1400 (2020).
30. Liu, S., Pei, C., Khan, L., Wang, H. & Tao, S. Multiobjective optimization of coding metamaterial for low-profile and broadband microwave absorber. *IEEE Ant Wirel. Propag. Lett.***23** (1), 379–383 (2024).
31. Sabbagh, M. A. E., Bakr, M. H. & Bandler, J. W. Adjoint higher order sensitivities for fast full-wave optimization of microwave filters. *IEEE Trans. Microw. Theory Techn*. **54** (8), 3339–3351 (2006).
32. Pietrenko-Dabrowska, A. & Koziel, S. „Computationally-efficient design optimization of antennas by accelerated gradient search with sensitivity and design change monitoring. *IET Microwaves Ant Prop.***14** (2), 165–170 (2020).
33. Pietrenko-Dabrowska, A. & Koziel, S. „Expedited gradient-based design closure of antennas using variable-resolution simulations and sparse sensitivity updates. *IEEE Trans. Ant Propag.***70** (6), 4925–4930 (2022).
34. Liu, B., Yang, H. & Lancaster, M. J. Global optimization of microwave filters based on a surrogate model-assisted evolutionary algorithm. *IEEE Trans. Microw. Theory Techn*. **65** (6), 1976–1985 (2017).
35. Cervantes-González, J. C. et al. Space mapping optimization of handset antennas considering EM effects of mobile phone components and human body. *Int. J. RF Microw. CAE*. **26** (2), 121–128 (2016).
36. Na, W. et al. Efficient EM optimization exploiting parallel local sampling strategy and bayesian optimization for microwave applications. *IEEE Microw. Wirel. Comp. Lett.***31** (10), 1103–1106 (2021).
37. Zhao, Q. & Sarris, C. D. Space–time adaptive modeling and shape optimization of microwave structures with applications to metasurface design. *IEEE Trans. Microw. Theory Techn*. **70** (12), 5440–5453 (2022).
38. Sivaram, S. A. & Vinoy, K. J. Inverse multiquadric radial basis functions in eigenvalue analysis of a circular waveguide using radial point interpolation method. *IEEE Microw. Wirel. Comp. Lett.***30** (6), 537–540 (2020).
39. Jacobs, J. P. Characterization by gaussian processes of finite substrate size effects on gain patterns of microstrip antennas. *IET Microwaves Ant Prop.***10** (11), 1189–1195 (2016).
40. Cai, J., King, J., Yu, C., Liu, J. & Sun, L. Support vector regression-based behavioral modeling technique for RF power transistors. *IEEE Microw. Wirel. Comp. Lett.***28** (5), 428–430 (2018).
41. Zhang, W., Feng, F., Jin, J. & Zhang, Q. J. Parallel multiphysics optimization for microwave devices exploiting neural network surrogate. *IEEE Microw. Wirel. Comp. Lett.***31** (4), 341–344 (2021).
42. Feng, F. et al. Multifeature-assisted neuro-transfer function surrogate-based EM optimization exploiting trust-region algorithms for microwave filter design, *IEEE Trans. Microwave Theory Techn.*, vol. 68, no. 2, pp. 531–542, Feb. (2020).
43. Kim, D., Kim, M. & Kim, W. Wafer edge yield prediction using a combined long short-term memory and feed- forward neural network model for semiconductor manufacturing. *IEEE Access*.**8**, 215125–215132 (2020).
44. Li, S., Fan, X., Laforge, P. D. & Cheng, Q. S. Surrogate model-based space mapping in postfabrication bandpass filters' tuning. *IEEE Trans. Microw. Theory Tech.***68** (6), 2172–2182 (2020).
45. Koziel, S. Unnsteinsson Expedited design closure of antennas by means of trust-region-based adaptive response scaling. *IEEE Antennas Wirel. Prop. Lett.***17** (6), 1099–1103 (2018).
46. Su, Y., Li, J., Fan, Z. & Chen, R. „Shaping optimization of double reflector antenna based on manifold mapping, *Int. Applied Comp. Electromagnetics Soc. Symp. (ACES)*, Suzhou, China, pp. 1–2, (2017).
47. Koziel, S. & Leifsson, L. Efficient knowledge-based optimization of expensive computational models using adaptive response correction. *J. Comp. Sci.***11**, 1–11 (2015).
48. Ullah, U., Al-Hasan, M., Koziel, S. & Ben Mabrouk, I. EM-driven size reduction and multi-criterial optimization of broadband circularly-polarized antennas using Pareto front traversing and design extrapolation, *Scientific Reports*, vol. 12, paper no. 9877, (2022).
49. Zhu, H. & Abbosh, A. M. A compact tunable directional coupler with continuously tuned differential phase. *IEEE Microw. Wirel. Comp. Lett.***28** (1), 19–21 (2018).
50. Jiao, M. R., Zhu, F., Chu, P., Yu, W. & Luo, G. Q. Compact hybrid bandpass filters using substrate-integrated waveguide and stripline resonators. *IEEE Trans. Microw. Theory Techn*. **72** (1), 391–400 (2024).

51. Koziel, S. & Pietrenko-Dabrowska, A. On EM-driven size reduction of antenna structures with explicit constraint handling. *IEEE Access.***9**, 165766–165772 (2021).
52. Koziel, S. & Pietrenko-Dabrowska, A. On computationally-efficient reference design acquisition for reduced-cost constrained modeling and re-design of compact microwave passives. *IEEE Access.***8**, 203317–203330 (2020).
53. Pietrenko-Dabrowska, A. & Koziel, S. Numerically efficient algorithm for compact microwave device optimization with flexible sensitivity updating scheme. *Int. J. RF Microw. CAE*, **29**, 7, (2019).
54. Koziel, S. & Pietrenko-Dabrowska, A. Reliable EM-driven size reduction of antenna structures by means of adaptive penalty factors. *IEEE Trans. Ant Propag.***70** (2), 1389–1401 (2021).
55. Mahrokh, M. & Koziel, S. Improved-efficacy EM-based antenna miniaturization by multi-fidelity simulations and objective function adaptation. *Energies*. **15** (2), 403 (2021).
56. Liang, S. et al. „Sidelobe reductions of antenna arrays via an improved chicken swarm optimization approach. *IEEE Access.***8**, 37664–37683 (2020).
57. Zhang, H., Bai, B., Zheng, J. & Zhou, Y. Optimal design of sparse array for ultrasonic total focusing method by binary particle swarm optimization. *IEEE Access.***8**, 111945–111953 (2020).
58. Li, X. & Luk, K. M. The grey wolf optimizer and its applications in electromagnetics. *IEEE Trans. Ant Prop.***68** (3), 2186–2197 (2020).
59. Luo, X., Yang, B. & Qian, H. J. Adaptive synthesis for resonator-coupled filters based on particle swarm optimization. *IEEE Trans. Microw. Theory Techn*. **67** (2), 712–725 (2019).
60. Akinwande, O., Erdogan, S., Kumar, R. & Swaminathan, M. Surrogate modeling with complex-valued neural nets for signal integrity applications. *IEEE Trans. Microw. Theory Techn*. **72** (1), 478–489 (2024).
61. Taran, N., Ionel, D. M. & Dorrell, D. G. Two-level surrogate-assisted differential evolution multi-objective optimization of electric machines using 3-D FEA, *IEEE Trans. Magn.*, vol. 54, no. 11, paper 8107605, Nov. (2018).
62. Couckuyt, I., Declercq, F., Dhaene, T., Rogier, H. & Knockaert, L. Surrogate-based infill optimization applied to electromagnetic problems. *Int. J. RF Microw. Computt -Aided Eng.***20** (5), 492–501 (2010).
63. Chen, C., Liu, J. & Xu, P. Comparison of infill sampling criteria based on Kriging surrogate model, *Sc. Rep.*, vol. 12, Art. No. 678, (2022).
64. Forrester, A. I. J. & Keane, A. J. Recent advances in surrogate-based optimization. *Prog Aerosp. Sci.***45**, 50–79 (2009).
65. Tak, J., Kantemur, A., Sharma, Y. & Xin, H. A 3-D-printed W-band slotted waveguide array antenna optimized using machine learning. *IEEE Ant Wirel. Prop. Lett.***17** (11), 2008–2012 (2018).
66. Wu, Q., Wang, H. & Hong, W. Multistage collaborative machine learning and its application to antenna modeling and optimization. *IEEE Trans. Ant Propag.***68** (5), 3397–3409 (2020).
67. Roy, C. & Wu, K. Surrogate model-based filter optimization by a field-circuit model mapping. *IEEE Trans. Microw. Theory Techn*. **72** (5), 3144–3157 (2024).
68. Lim, D. K., Yi, K. P., Jung, S. Y., Jung, H. K. & Ro, J. S. Optimal design of an interior permanent magnet synchronous motor by using a new surrogate-assisted multi-objective optimization. *IEEE Trans. Magn.***51** (11), 8207504 (2015).
69. Toktas, A., Ustun, D. & Tekbas, M. Multi-objective design of multi-layer radar absorber using surrogate-based optimization, *IEEE Trans. Microw. Theory Techn*., vol. 67, no. 8, pp. 3318–3329, Aug. (2019).
70. Greda, L. A., Winterstein, A., Lemes, D. L. & Heckler, M. V. T. Beamsteering and beamshaping using a linear antenna array based on particle swarm optimization. *IEEE Access.***7**, 141562–141573 (2019).
71. Pietrenko-Dabrowska, A., Koziel, S. & Mahrokh, M. „Optimization-based high-frequency circuit miniaturization through implicit and explicit constraint handling: recent advances. *Energies*, **15**, no. (2022). 19, paper 6955.
72. Koziel, S. & Pietrenko-Dabrowska, A. Expedited feature-based quasi-global optimization of multi-band antennas with jacobian variability tracking. *IEEE Access.***8**, 83907–83915 (2020).
73. Pietrenko-Dabrowska, A. & Koziel, S. „Generalized formulation of response features for reliable optimization of antenna input characteristics. *IEEE Trans. Ant Propag.***70** (5), 3733–3748 (2021).
74. Fu, J. et al. Feature-assisted neural network surrogate-based multiphysics optimization for microwave filters. *IEEE Microw. Wirel. Techn Lett.***34** (5), 474–477 (2024).
75. Na, W. et al. Advanced EM optimization using adjoint-sensitivity-based multifeature surrogate for microwave filter design. *IEEE Microwve Wirel. Techn Lett.***34** (1), 1–4 (2024).
76. Pietrenko-Dabrowska, A. & Koziel, S. Simulation-driven antenna modeling by means of response features and confined domains of reduced dimensionality. *IEEE Access.***8**, 228942–228954 (2020).
77. Cawley, G. C. & Talbot, N. L. C. „On over-fitting in model selection and subsequent selection bias in performance evaluation. *J. Mach. Learn.***11**, 2079–2107 (2010).
78. Vinod Chandra, S. S. & Anand, H. S. „Nature inspired meta heuristic algorithms for optimization problems, *Computing*, vol. 104, pp. 251–269, (2022).
79. Liu, J., Han, Z. & Song, W. „Comparison of infill sampling criteria in kriging-based aerodynamic optimization, *28th Int. Congress of the Aeronautical Sciences*, pp. 1–10, Brisbane, Australia, 23–28 Sept., (2012).
80. Lin, Z. & Chu, Q. X. „A novel approach to the design of dual-band power divider with variable power dividing ratio based on coupled-lines. *Prog Electromagn. Res.***103**, 271–284 (2010).
81. Koziel, S. & Pietrenko-Dabrowska, A. Reduced-cost surrogate modeling of compact microwave components by two-level kriging interpolation. *Eng. Opt.***52** (6), 960–972 (2019).
82. Gorissen, D., Crombecq, K., Couckuyt, I., Dhaene, T. & Demeester, P. „A surrogate modeling and adaptive sampling toolbox for computer based design. *J. Mach. Learn. Res.***11**, 2051–2055 (2010).
83. Marelli, S. & Sudret, B. „UQLab: a framework for uncertainty quantification in Matlab, *2nd Int. Conf. on Vulnerability and Risk Analysis and Management (ICVRAM 2014)*, University of London, UK, July 13–15, pp. 2554–2563, (2014).
84. Bouhlel, M. A. et al. A Python surrogate modeling framework with derivatives. *Adv. Eng. Soft*, p. 102662, (2019).
85. Storn, R. & Price, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces, *J. Global Opt.*, vol. 11, pp. 341–359, Dec. (1997).

Author contributions

Conceptualization, S.K., A.P.; methodology, S.K. and A.P.; data generation, S.K.; investigation, S.K. and A.P.; writing—original draft preparation, S.K. and A.P.; writing—review and editing, S.K.; visualization, S.K. and A.P.; supervision, S.K.; project administration, S.K. and A.P.

Declarations

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

Correspondence and requests for materials should be addressed to S.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.