scientific reports

Check for updates

OPEN

Machine-learning-based global optimization of microwave passives with variable-fidelity EM models and response features

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

Maximizing microwave passive component performance demands precise parameter tuning, particularly as modern circuits grow increasingly intricate. Yet, achieving this often requires a comprehensive approach due to their complex geometries and miniaturized structures. However, the computational burden of optimizing these components via full-wave electromagnetic (EM) simulations is substantial. EM analysis remains crucial for circuit reliability, but the expense of conducting rudimentary EM-driven global optimization by means of popular bio-inspired algorithms is impractical. Similarly, nonlinear system characteristics pose challenges for surrogate-assisted methods. This paper introduces an innovative technique leveraging variable-fidelity EM simulations and response feature technology within a kriging-based machine-learning framework for cost-effective global parameter tuning of microwave passives. The efficiency of this approach stems from performing most operations at the low-fidelity simulation level and regularizing the objective function landscape through the response feature method. The primary prediction tool is a co-kriging surrogate, while a particle swarm optimizer, guided by predicted objective function improvements, handles the search process. Rigorous validation demonstrates the proposed framework's competitive efficacy in design quality and computational cost, typically requiring only sixty high-fidelity EM analyses, juxtaposed with various state-of-the-art benchmark methods. These benchmarks encompass nature-inspired algorithms, gradient search, and machine learning techniques directly interacting with the circuit's frequency characteristics.

Over the years, the performance requirements formulated for high-frequency systems, including microwave passive components, have increased significantly. Fields like mobile communications[1], radio-frequency identification[2], implantable systems[3], internet of things[4], or energy harvesting[5], require structures that can be reconfigured[6], enable broadband[7] or multi-band operation[8], harmonic suppression[9], but also feature small physical dimensions[10–12]. The circuits delivering the aforementioned and other functionalities exhibit complex topologies, typically parameterized by large numbers of variables, mainly geometrical ones. Their layouts are often densely arranged, especially in the case of miniaturized components, obtained using approaches like transmission line meandering[13], slow-wave phenomena[14], or integrating extra components (e.g., resonators[15], stubs[16], etc.). Precisely characterizing geometrically intricate microwave circuits necessitates full-wave electromagnetic (EM) analysis. Unlike conventional models (analytical[17], equivalent networks[18]), EM simulations can accurately capture phenomena like cross-coupling, dielectric and radiation losses, and the effects of environmental components (e.g., SMA connectors). As a result, the use of electromagnetic tools has become indispensable today. They are extensively utilized across all design phases, from shaping the geometry[19], conducting parametric studies[20], to finalizing the design[21].

One of the most prevalent simulation-driven processes involves adjusting geometry parameters, yet it is a costly endeavour. Numerical optimization algorithms demand numerous system evaluations, making it an expensive task. Even with local optimization methods such as gradient-[22] or stencil-based search[23], the expenses

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

can quickly accumulate, often requiring many dozens or even hundreds of EM analyses. Interactive methods (experience-driven parametric studies[24]) are cheaper, yet unable to handle more than one or two parameters at a time, let alone several design goals and constraints. On the other hand, for a growing number of cases, it is at least recommended if not necessary, to employ global search methods. Examples include optimization of frequency selective surface[25], coding metasurfaces[26], metamaterial-based components[27], antenna pattern synthesis[28,29], design of conformal[30] or sparse arrays[31], as well as multi-objective design[32]. Other instances include, among various others, redesigning microwave components across extensive ranges of operating conditions[33], and optimization of compact circuits[34]. For the latter, redundancy of geometry parameters introduced by using various miniaturization strategies (line folding[35], CMRCs[36]) makes the relations between the circuit dimensions and its electrical characteristics rather unintuitive.

Currently, population-based algorithms inspired by nature[37–40] have taken the lead in global optimization. They are implemented to mimic various biological[41] or social[42] phenomena, including biological evolution[43], feeding or hunting strategies of various species[44,45], etc. The fundamental mechanisms involve information exchange among the individuals (particles, agents, members)[46] within the pool of potential solutions (swarm, population, pack)[47] for the specific problem. This exchange can impact the individuals by altering their composition (through crossover, mutation[48]), or by repositioning them within the search space. For instance, this might involve introducing a randomized bias towards the best solution discovered thus far, whether individual or overall[49]. Practical observations confirm global search capability achieved using such means although there is little of no supporting theory. Further, the computational efficiency of population-based algorithms is poor. A typical algorithm run requires thousands of objective function evaluations, which essentially rules out direct nature-inspired optimization of EM simulation models unless the evaluation time is short (e.g., up to 10–30 s per analysis). Common techniques within this category comprise genetic[50] and evolutionary algorithms[51], particle swarm optimization (PSO)[52], differential evolution[53], firefly algorithms[54], ant systems[55], grey wolf optimization[56]. Although every year witnesses new development[57–60], it seems that—despite fancy names—the new algorithms offer incremental adjustments of the existing techniques.

Enabling nature-inspired EM-driven optimization requires acceleration, which can be achieved using surrogate modelling methods[61,62]. Several widely used techniques in this context involve Gaussian process regression (GPR)[63], kriging[64], neural networks[65], and polynomial chaos expansion[66]. Typically, the surrogate model serves as a predictor, determining optimal design locations, and continuously updates using full-wave simulation results gathered in the course of the search process[67]. The criteria for infilling depend on the objective, such as enhancing model accuracy (maximizing mean squared error reduction[68]), identifying the global optimum (minimizing predicted objective function[69]), or balancing exploration and exploitation[70]. The search procedures of this class are often referred to as machine learning methods[71–73]. The fundamental issue with surrogate-assisted methods is building the metamodel. Modelling the behaviour of microwave circuit responses, usually represented as frequency characteristics, becomes challenging due to their highly nonlinear nature. Covering extensive frequency ranges and accounting for diverse geometry/material parameters demands substantial training datasets. Many of machine learning frameworks in the realm of high-frequency design are therefore demonstrated for systems defined over small parameters spaces, either in terms of dimensionality or parameter ranges[74,75]. These issues can be mitigated using variable-fidelity simulations[76]. Other potential acceleration techniques include constrained modelling[77–79], and the response feature technology[80]; however, no machine learning approach incorporating these methods has been described in the literature thus far. Yet another approach that might be useful for enhancing dependability of data-driven modelling is data normalization, which is normally applied to transform the features of the system at hand to be on a similar scale. This contributes to the improvement of the performance and model training stability. Some of popular normalization and data augmentation techniques include min–max normalization, Z-score normalization[81], decimal scaling, or class-balancing (e.g., synthetic minority oversampling technique, SMOTE[82], oriented towards oversampling minority classes thereby improving the balance of the dataset).

This study introduces a machine-learning framework aimed at accelerating the global optimization of microwave passive components, also miniaturized structures. The approach tackles the design challenge by focusing on response features—a discrete set of points derived from EM-simulated outputs that quantify the circuit's performance specifications. Leveraging the weakly-nonlinear dependence of feature point coordinates on geometry/material parameters enables efficient surrogate modelling. The algorithm's initial phases, i.e., parameter space pre-screening and surrogate model construction, utilize low-fidelity EM models. The core iteration involves generating infill points and refining the surrogate model, primarily based on the co-kriging approach. This method combines low-fidelity data with high-fidelity EM analyses carried out at this stage. By employing characteristic points and multi-resolution EM analysis, the computational efficiency is noteworthy: the average optimization cost is around sixty high-fidelity EM simulations for the circuit under design. Concurrently, the quality and consistency of designs compete favourably against benchmarks, which encompass multiple-start gradient search, nature-inspired optimization, and kriging-based machine learning directly working with circuit frequency responses. This highlights the significance of response features in enhancing both reliability and computational efficiency within the proposed algorithm.

Global machine-learning microwave optimization using response features and multi-resolution computational models

This section delves into the intricacies of the proposed optimization approach. "Microwave design optimization. Problem statement", "Response feature methodology", "Variable-fidelity computational models" sections revisit the formulation of the design task, the concept of characteristic points, and variable-fidelity simulation models, respectively. A concise overview of kriging and co-kriging surrogates is presented in "Surrogate modelling using

kriging and co-kriging" section. The optimization process begins with the pre-screening stage, detailed alongside the construction of the initial (kriging) surrogate model in "Parameter space pre-screening. Initial surrogate model construction" section. "Generating infill by means of nature-inspired optimization. Co-kriging surrogate" section expounds on the machine-learning framework involving co-kriging models and the infill criterion based on predicted objective function improvement. Lastly, "Complete optimization framework" section encapsulates the entire procedure using the pseudocode and a flow diagram.

Microwave design optimization. Problem statement

In this study, optimizing microwave circuits involves adjusting their independent variables, typically geometry parameters like component widths, lengths, and their spacings, consolidated within the parameter vector x (as illustrated in Fig. 1). The objective vector F_t comprises the design goals, encompassing target operating parameters such as center frequency, power split ratio, and bandwidth. Evaluating the design x in relation to the target vector F_t is accomplished through a scalar objective function $U(x, F_t)$. Figure 2 provides several examples of common design scenarios along with the corresponding definitions of U. The primary type of microwave circuit responses are scattering parameters (cf. Fig. 1)[83], along with the quantities that can be derived therefrom (e.g., the phase characteristic).

It is important to highlight that most real-world design challenges involve multiple objectives; that is, there is a need to enhance or regulate more than just one parameter or quantity. As most of available optimization algorithms are single-objective ones, multi-criterial problems are normally reformulated, e.g., by casting all but the primary objective into constraints[84]. Another popular option is scalarization using, e.g., a linear combination of goals[85]. Genuine multi-objective design is outside the scope of this paper.

Now, we can express the microwave optimization task as a minimization task represented by the following form:

$$x^* = \arg \min_{x \in X} U(x, F_t) \tag{1}$$

where x^* represents the sought-after optimal design, and X denotes the parameter space, typically an interval determined by the lower and upper variable bounds for x_k, $k = 1, \ldots, n$.

Response feature methodology

Full-wave computational models ensure reliable evaluation of microwave components, but they are CPU-intensive. This is a major hindrance to optimization procedures. The nonlinearity of circuit responses poses additional obstacles to globally exploring the parameter space. Figure 3 shows examples of frequency characteristics of a microstrip coupler at a number of randomly generated designs. Circuit optimization for a centre frequency of 1.5 GHz requires global search as local tuning initiated from most of the shown designs would not succeed.

The response feature techniques[86] were introduced to address the mentioned challenges by redefining the design problem using a discrete set of characteristic points (or features) from the system outputs[87]. This method leverages straightforward connections between the circuit's geometry/material parameters and the frequency and level coordinates of these points[80,86–89]. Consequently, it regularizes the objective function, hastening optimization processes[87], and simplifying behavioural modelling[89].

The characteristics points are tailored to address specific design objectives[80]. Take, for example, a microwave coupler like the rat-race structure, and its responses depicted in Fig. 3. To regulate the center frequency, bandwidth of isolation characteristics, and power division ratio, the feature points can be defined as illustrated in Fig. 4, further detailed in the figure caption. Figure 4b demonstrates the straightforward connections between these feature points and the adjustable parameters. For further insights into this topic, additional discussions can be found in references such as[80] and[86]. In this work, we exploit these properties to facilitate a construction of surrogate models employed to enable and accelerate the global search process.

Symbol	Meaning	Comment
$x = [x_1 \ldots x_n]^T$	Vector of design parameters	Typically, the variables are geometry parameters (dimensions in mm)
$S_{ij}(x, f)$	Scattering parameters at the design x and frequency f	Subscripts i and j stand for the circuit ports
$F_t = [F_{t.1} \ldots F_{t.K}]^T$	Vector of target operating parameters	Examples: operating frequency (or frequencies for multi-band structures, power split ratio (in the case of couplers), bandwidth, substrate parameters (permittivity, thickness)
$U(x, F_t)$	Objective function to be minimized in the design process	Function quantifying the design quality, cf. Fig. 2

Figure 1. Microwave design optimization: notation and terminology.

Design task	Target vector	Objective function*
Circuit: Impedance matching transformer **Objectives:** Minimize maximum in-band reflection, operating band $F = [f_1\,f_2]$	$\boldsymbol{F}_t = [f_1\,f_2]^T$	$U(\boldsymbol{x}, \boldsymbol{F}_t) = U(\boldsymbol{x}, [f_1\,f_2]^T) = \max\{f \in [f_1\,f_2] : \lvert S_{11}(\boldsymbol{x}, f)\rvert\}$
Circuit: Microstrip coupler **Objectives:** • Minimize matching and isolation at the target operating frequency f_0 • Ensure the required power split K_P at f_0	$\boldsymbol{F}_t = [f_0\,K_P]^T$	$U(\boldsymbol{x}, \boldsymbol{F}_t) = U(\boldsymbol{x}, [f_0\,K_P]^T) =$ $= \max\{\lvert S_{11}(\boldsymbol{x}, f_0)\rvert, \lvert S_{41}(\boldsymbol{x}, f_0)\rvert\} +$ $+ \beta\big[\lvert S_{21}(\boldsymbol{x}, f_0)\rvert - \lvert S_{31}(\boldsymbol{x}, f_0)\rvert - K_P\big]^2$
Circuit: Dual-band coupler **Objectives:** • Minimize matching and isolation at the operating frequencies f_1 and f_2 • Ensure equal power split at f_1 and f_2	$\boldsymbol{F}_t = [f_1\,f_2]^T$	$U(\boldsymbol{x}, \boldsymbol{F}_t) = U(\boldsymbol{x}, [f_1\,f_2]^T) = \max\{\lvert S_{11}(\boldsymbol{x}, f_1)\rvert, \lvert S_{11}(\boldsymbol{x}, f_2)\rvert,$ $\lvert S_{41}(\boldsymbol{x}, f_1)\rvert, \lvert S_{41}(\boldsymbol{x}, f_1)\rvert\}$ $+ \beta[(\lvert S_{21}(\boldsymbol{x}, f_1)\rvert - \lvert S_{31}(\boldsymbol{x}, f_1)\rvert)^2$ $+ (\lvert S_{21}(\boldsymbol{x}, f_2)\rvert - \lvert S_{31}(\boldsymbol{x}, f_2)\rvert)^2]$
Circuit: Compact rat-race coupler **Objectives:** • Minimize the footprint area $A(\boldsymbol{x})$ • Ensure equal power split at the operating frequency f_0 • Maintain $\lvert S_{11}\rvert$, $\lvert S_{41}\rvert \leq -20$ dB over the bandwidth $[f_0 - B, f_0 + B]$	$\boldsymbol{F}_t = [f_0\,B]^T$	$U(\boldsymbol{x}, \boldsymbol{F}_t) = U(\boldsymbol{x}, [f_0\,B]^T) = A(\boldsymbol{x}) +$ $+ \beta_1\big[\max\{c(\boldsymbol{x}) + 20, 0\}/20\big]^2$ $+ \beta_2\big[\lvert S_{21}(\boldsymbol{x}, f_0)\rvert - \lvert S_{31}(\boldsymbol{x}, f_0)\rvert\big]^2$ where $c(\boldsymbol{x}) = \max\{f \in [f_0 - B, f_0 + B] :$ $\max\{\lvert S_{11}(\boldsymbol{x}, f)\rvert, \lvert S_{41}(\boldsymbol{x}, f)\rvert\}$
Circuit: Dual-band coupler **Objectives:** • Implement the circuit on the substrate of relative permittivity ε_r • Minimizes $\lvert S_{11}\rvert$ and $\lvert S_{41}\rvert$ at the operating frequencies f_1 and f_2 • Ensure equal power split at f_1 and f_2	$\boldsymbol{F}_t = [f_{0.1}\,f_{0.2}\,\varepsilon_r]^T$	$U(\boldsymbol{x}, \boldsymbol{F}_t) = U(\boldsymbol{x}, [f_1\,f_2\,\varepsilon_r]^T) =$ $= \max\{\lvert S_{11}(\boldsymbol{x}, f_1)\rvert, \lvert S_{41}(\boldsymbol{x}, f_1)\rvert, \lvert S_{11}(\boldsymbol{x}, f_2)\rvert, \lvert S_{41}(\boldsymbol{x}, f_2)\rvert\}$ $+ \beta\big[\lvert S_{21}(\boldsymbol{x}, f_1)\rvert - \lvert S_{31}(\boldsymbol{x}, f_1)\rvert\big]^2$ $+ \beta\big[\lvert S_{21}(\boldsymbol{x}, f_2)\rvert - \lvert S_{31}(\boldsymbol{x}, f_2)\rvert\big]$
Circuit: Triple-band power divider **Objectives:** • Ensure equal power split at the operating frequencies f_1, f_2, and f_3 • Minimize input matching $\lvert S_{11}\rvert$, output matching $\lvert S_{22}\rvert$ and $\lvert S_{33}\rvert$, and isolation $\lvert S_{23}\rvert$ at f_1, f_2, and f_3	$\boldsymbol{F}_t = [f_1\,f_2\,f_3]^T$	$U(\boldsymbol{x}, \boldsymbol{F}_t) = U(\boldsymbol{x}, [f_1\,f_2\,f_3]^T) = \max\{\max_{k,l\in\{1,2,3\}} \lvert S_{kk}(\boldsymbol{x}, f_l)\rvert,$ $\max_{l\in\{1,2,3\}} \lvert S_{23}(\boldsymbol{x}, f_l)\rvert\}$ $+ \beta\sum_{l=1}^{3}(\lvert S_{21}(\boldsymbol{x}, f_l)\rvert - \lvert S_{31}(\boldsymbol{x}, f_l)\rvert)^2$

* The terms followed by β are penalty factors introduced to enforce the design constraints (e.g., required power split or matching/isolation bandwidth).

Figure 2. Examples of microwave design optimization problems. Verbal description of the task (left) is followed by a definition of the target vector $\boldsymbol{F}_t$ (middle), and a possible definition of the objective function (right).

At this point, it should be emphasized that utilization of response feature adds a layer of complexity in algorithm implementation, which is due to the necessity of defining and extracting the feature points. In practice, this requires implementation of a separate algorithm for scanning the frequency characteristics and identifying the feature points, here, realized in Matlab. On the other hand, especially when the parameter space is very large (broad ranges of geometry parameters) and highly-dimensional, the features might be difficult to identify. The latter normally occurs for heavily distorted responses at poor-quality designs. In practice, designs like these may be assigned high value of the objective function U_F to mark them as low-quality. Nonetheless, the aforementioned issue constitutes a limitation of the proposed approach, yet can be mitigated to a certain extent by appropriate selection of the parameter space (e.g., based on prior experience with a given microwave circuit).

To integrate the response feature approach into the optimization process, we must establish suitable notation. Specifically, we will denote the feature points of the circuit under design as $\boldsymbol{f}_P(\boldsymbol{x}) = [\boldsymbol{f}_f(\boldsymbol{x})^T \boldsymbol{f}_L(\boldsymbol{x})^T]^T$, with $\boldsymbol{f}_f(\boldsymbol{x}) = [f_{f.1}(\boldsymbol{x}) \ldots f_{f.K}(\boldsymbol{x})]^T$ and $\boldsymbol{f}_L(\boldsymbol{x}) = [f_{L.1}(\boldsymbol{x}) \ldots f_{L.K}(\boldsymbol{x})]^T$ being the vector of frequency and level coordinates, respectively. Table 1 provides a selection of example characteristics points for a microwave coupling circuit, focusing on the device's operating frequency and power split ratio.
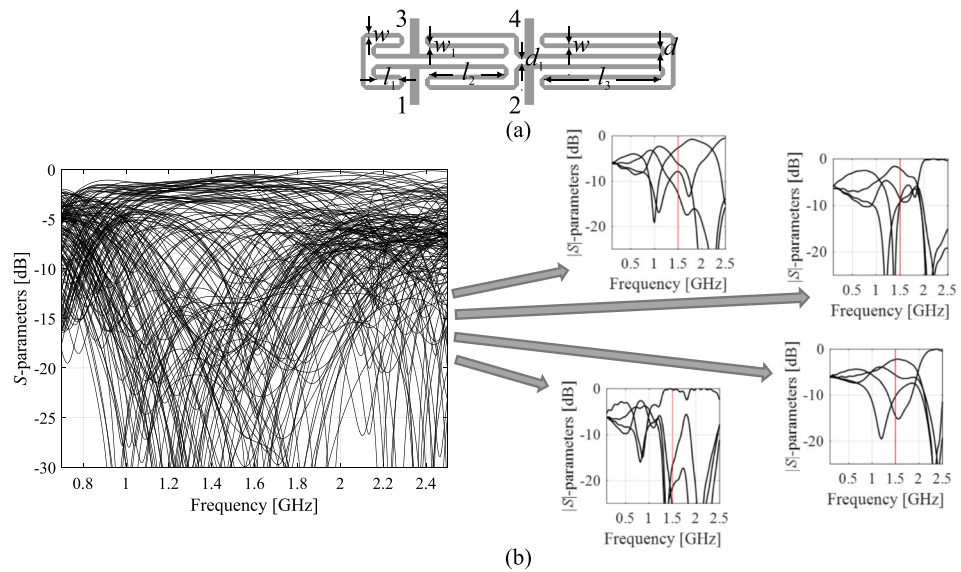
Figure 3. Miniaturized rat-race coupler and its scattering parameters: (**a**) circuit architecture, (**b**) scattering parameters at randomly selected designs. If local search based on the objective function (as illustrated in Fig. 2, with the target frequency set at 1.5 GHz and marked using vertical lines) were initiated from many of the shown designs, it would prove unsuccessful due to significant misalignment between the target and the actual operating conditions.
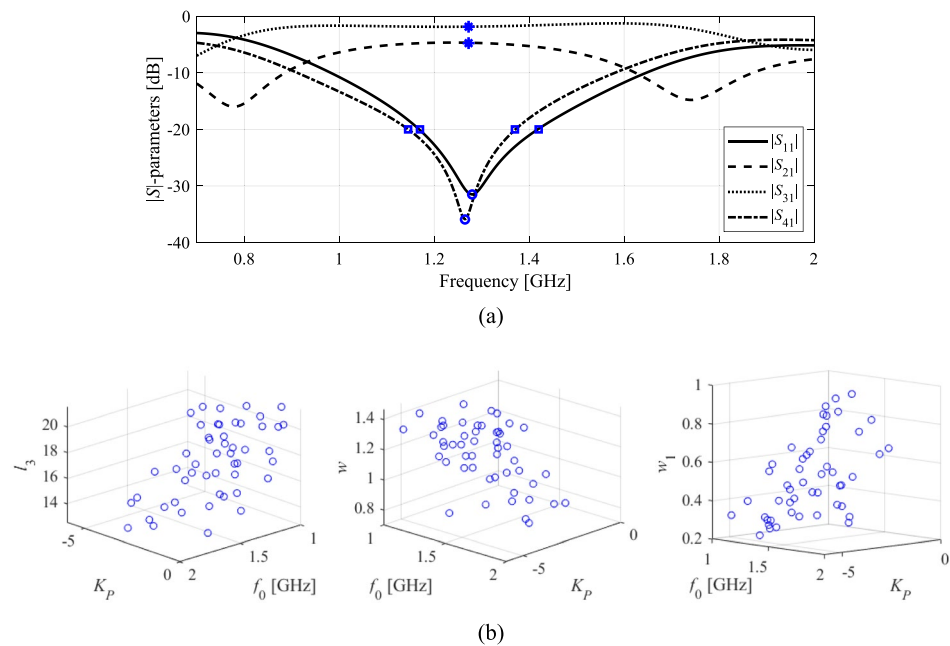

Figure 4. Characteristic points selection for a microwave coupler: (**a**) possible feature point choices: open circle—points corresponding to the minima of the matching and isolation characteristics, *—points corresponding to the power split ratio (evaluated at the frequency f_0 being the average of the frequencies of $|S_{11}|$ and $|S_{41}|$ minima), open square—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 geometry parameter 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 shown. Clear patterns are visible even though the trial points were not optimized whatsoever.

Verbal description	Four feature points:
	    Two points corresponding to the minima of the $|S_{11}|$ and $|S_{41}|$ responses
	    Two points corresponding to the power split ratio (i.e., the evaluations of $|S_{31}|$ and $|S_{21}|$ at the centre frequency of the circuit)

Vector notation	$f_f(x) = \begin{bmatrix} f_{f.1}(x) & f_{f.2}(x) & f_{f.3}(x) & f_{f.4}(x) \end{bmatrix}^T$ $f_L(x) = \begin{bmatrix} f_{L.1}(x) & f_{L.2}(x) & f_{L.3}(x) & f_{L.4}(x) \end{bmatrix}^T$
Explanation of terms	$f_{f.1}$ and $f_{f.2}$—frequencies corresponding to the minima of $|S_{11}|$ and $|S_{41}|$ $f_{L.1}$ and $f_{L.2}$—corresponding minima levels $f_{f.3} = f_{f.4} = [f_{f.1} + f_{f.2}]/2 = f_0$—approximated operating frequency of the coupler $f_{L.3}$ and $f_{L.4}$—levels of $|S_{31}|$ and $|S_{41}|$ at f_0 The power split ratio can be computed as $K_P = f_{L.3} - f_{L.4}$

Table 1. Example response features for a microwave coupler (cf. Fig. 4).

Having f_P, we denote by $F_o(x) = F_o(f_P(x))$ the circuit's operating parameters, computed from $f_P(x)$. Following this, the design task posed in terms of characteristics points appears as:

$$x^* = \arg\min_x U_F(x, f_P(x), F_t) \tag{2}$$

The analytical form of the function U_F is similar to that shown in Fig. 2 but employs $F_o(x)$ rather than scattering parameters S_{kl}. It can also be written in a compact manner as

$$U_F(x, f_P(x), F_t) = U_0(f_P(x)) + \beta||F_o(f_P(x)) - F_t||^2 \tag{3}$$

Here, the function U_0 accounts for the main objective. For clarification, consider a scenario, under which we seek to achieve the following: (i) align the coupler center frequency at the target value f_t, (ii) establish the target power division ratio, and (iii) reduce $|S_{11}|$ (impedance matching) and $|S_{41}|$ (port isolation) at f_t.

In this case, the main goal is to reduce $f_{L.1}$ and $f_{L.2}$, cf. Table 1, i.e., we have $U_0(f_P(x)) = \max\{f_{L.1}(x), f_{L.2}(x)\}$. Further, given $F_t = [f_t \ K_P]^T$, we define $F_o(f_P(x)) = [(f_{f.1} + f_{f.2})/2 \ f_{L.3} - f_{L.4}]^T$. The importance of the operating condition alignment is controlled by the scalar factor β. Typically, we set β = 100. It should be noted that conventional formulation (as shown in Fig. 2) is equivalent to (3) assuming that the optimum is attainable. Further, the second term in (3) (not present in standard formulation) acts as a regularization factor that facilitates identification of the optimum.

Variable-fidelity computational models

Reducing the EM analysis resolution—by decreasing the structural discretization density (alternative options discussed in[90])—expedites the simulation process but compromises accuracy, as shown in Fig. 5. This accuracy loss can be rectified by appropriately adjusting the low-fidelity model, forming the basis of physics-based surrogate-assisted methodologies. One prominent example is space mapping[91,92]. The degree of acceleration achievable varies with the problem. For most microwave passive components, it ranges between 2.5 and six for the low-fidelity model, ensuring adequate representation of crucial circuit output details.

In this study, the low- and high-resolution models are denoted as $R_c(x)$ and $R_f(x)$, respectively. R_c will be utilized to:
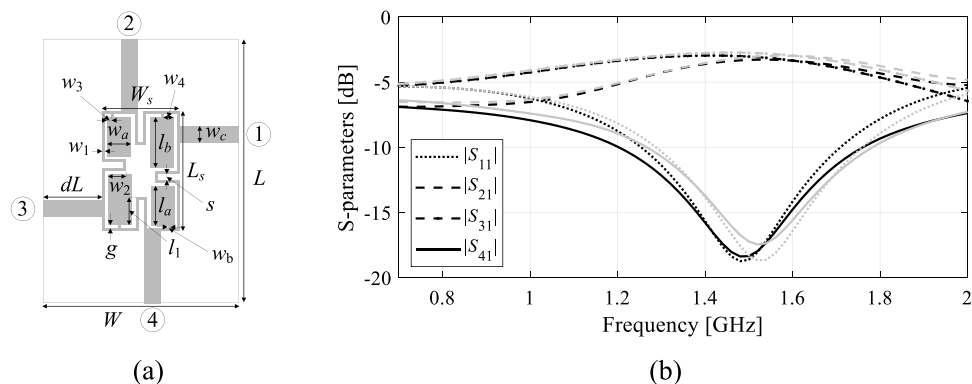

Figure 5. Variable-fidelity models: (**a**) geometry of an exemplary compact branch-line coupler, (**b**) scattering parameters evaluated using the low-fidelity EM model (gray) and the high-fidelity one (black). At the design shown, the simulation time of the high-fidelity model is about 250 s, whereas the evaluation of the low-fidelity model takes about 90 s.

- Pre-screen the parameter space, i.e., to generate a set of random observables, from which those with extractable feature points will be selected for further processing;
- Construct the initial surrogate model using kriging (cf. "Surrogate modelling using kriging and co-kriging" section).

Subsequent optimization steps will employ the low-fidelity samples into a co-kriging surrogate (cf. "Surrogate modelling using kriging and co-kriging" section), which will also incorporate R_f data acquired during the algorithm run.

At this point, it should be emphasized that the pre-screening process in large parameter spaces (broad ranges of geometry parameters, high dimensionality) may be a limiting factor in terms of the computational efficiency. This is because is large spaces, the percentage of random designs with extractable features might be low, meaning that acquisition of a required number of accepted design would incur considerable expenses. One way of mitigating this issue is appropriate selection of the search space, based on engineering experience and prior knowledge about the circuit at hand, so that reasonably narrow parameter ranges can be established. On the other hand, the aforementioned issue would be also detrimental to any global search algorithm, whether it is a direct method of surrogate-assisted one.

Surrogate modelling using kriging and co-kriging

Kriging and co-kriging interpolation[93,94] will be used as the modelling methods of choice within the machine learning framework proposed in this work. Both techniques are briefly recalled below.

We consider two datasets:

- The low-fidelity one $\{x_{Bc}^{(k)}, R_c(x_{Bc}^{(k)})\}_{k=1,\dots,NBc}$, which consists of the parameter vectors $x_{Bc}^{(k)}$ and the corresponding circuit outputs;
- The high-fidelity set $\{x_{Bf}^{(k)}, R_f(x_{Bf}^{(k)})\}_{k=1,\dots,NBf}$, acquired through evaluation of the high-fidelity model R_f at parameter vectors $x_{Bf}^{(k)}$.

Tables 2 and 3 outline both kriging and co-kriging models $s_{KR}(x)$ and $s_{CO}(x)$, respectively. Note that the co-kriging surrogate incorporates the kriging model s_{KRc} set up based on the low-fidelity data $(X_{Bc}, R_c(X_{Bc}))$, and s_{KRf} generated on the residuals (X_{Bf}, r); here $r = R_f(X_{Bf}) - \rho \cdot R_c(X_{Bf})$. The parameter ρ is included in the Maximum Likelihood Estimation (MLE) of the second model[94]. If the low-fidelity data at X_{Bf} is unavailable, one can use an approximation $R_c(X_{Bf}) \approx s_{KRc}(X_{Bf})$. The same correlation function is utilized by both models (cf. Tables 2 and 3).

Model component	Analytical form
Model formulation	$s_{KR}(x) = M\gamma + r(x) \cdot \Psi^{-1} \cdot (R_c(X_{Bc}) - F\gamma)$ where M is a $N_{Bc} \times t$ model matrix of X_{Bc}, whereas F is a $1 \times t$ vector of the evaluation point x (t is the number of terms used in the regression function[91])
Regression function coefficients	$\gamma = (X_{Bc}^T \Psi^{-1} X_{Bc})^{-1} X_{Bc} \Psi^{-1} R_f(X_{Bc})$
$1 \times N_{Bc}$ vector of correlations between x and X_{Bc}	$r(x) = (\psi(x, x_{Bc}^{(1)}), \dots, \psi(x, x_{Bc}^{(N_{Bc})}))$
Correlation matrix	$\Psi = [\Psi_{i,j}]$ is a correlation matrix, where $\Psi_{i,j} = \psi(x_{Bf}^{(i)}, x_{Bf}^{(j)})$
Correlation function	$\psi(x, x') = \exp\left(\sum_{k=1}^{n} -\theta_k |x^k - x'^k|^P\right)$
Model identification: finding hyperparameters θ_k, $k = 1, \dots, n$, using Maximum Likelihood Estimation (MLE[94])	$(\theta_1, \dots, \theta_n) = \arg\min_{\theta_1,\dots,\theta_n} \left[-(N_{Bf}/2)\ln(\hat{\sigma}^2) - 0.5\ln(|\Psi|) \right]$ where $\hat{\sigma}^2 = (R_f(X_{Bf}) - F\alpha)^T \Psi^{-1} (R_f(X_{Bf}) - F\alpha)/N_{Bf}$ and $|\Psi|$ is the determinant of Ψ. In practice, a Gaussian correlation function ($P = 2$) is often employed, as well as $F = [1 \dots 1]^T$ and $M = 1$

Table 2. An outline of the kriging surrogate modeling. More information can be found in the literature, e.g.,[94].

Model component	Analytical form
Model formulation	$s_{CO}(x) = M\gamma + r(x) \cdot \Psi^{-1} \cdot (r - F\gamma)$
Vector of correlations	$r(x) = [\rho \cdot \sigma_c^2 \cdot r_c(x), \rho^2 \cdot \sigma_c^2 \cdot r_c(x, X_{Bf}) + \sigma_d^2 \cdot r_d(x)]$
Correlation matrix	$\Psi = \begin{bmatrix} \sigma_c^2 \Psi_c(X_{Bc}, X_{Bc}) & \rho\,\sigma_c^2 \Psi_c(X_{Bc}, X_{Bf}) \\ \rho\,\sigma_c^2 \Psi_c(X_{Bf}, X_{Bc}) & \rho^2\sigma_c^2 \Psi_c(X_{Bf}, X_{Bf}) + \sigma_d^2 \Psi_d \end{bmatrix}$ where and $M = [\rho M_c\ M_d]$ where $(F_c, \sigma_c, \Psi_c, M_c)$ and $(F_d, \sigma_d, \Psi_d, M_d)$ are matrices obtained from s_{KRc} and s_{KRf}, respectively[91]; parameter ρ is included in the MLE during model identification

Table 3. An outline of the co-kriging surrogate modeling. More information can be found in the literature, e.g.,[94].

Parameter space pre-screening. Initial surrogate model construction

The initial stage of the machine learning process described in this paper involves screening the parameter space, conducted at the level of low-fidelity EM models. Subsequently, an initial surrogate model is developed, utilizing the low-fidelity data collected during this phase. The pre-screening stage operates at the response feature level (refer to "Response feature methodology" section), enabling a sizeable curtailment of the number of samples compared to analysing the complete frequency characteristics of the designed circuit.

The primary aim of the initial surrogate $s^{(0)}(x)$ is to capture the behaviour exhibited by the frequency and level coordinates of the feature points delineated for the optimized circuit. We have

$$s^{(0)}(x) = \left[\left[s_{f,1}^{(0)}(x) \ldots s_{f,K}^{(0)}(x) \right]^T \left[s_{L,1}^{(0)}(x) \ldots s_{L,K}^{(0)}(x) \right]^T \right]^T \qquad (4)$$

The model is rendered by means of kriging[95] (cf. "Surrogate modelling using kriging and co-kriging" section, Table 2). The training data pairs are denoted as $\{x_{Bc}^{(j)}, f_P(x_B^{(j)})\}$, $j = 1, \ldots, N_{init}$. The points are generated randomly in the parameter space X using independent joint uniform probability distributions; only the observables with extractable response features are included into the training set. The dataset size N_{init} is adjusted to secure a sufficient surrogate model accuracy. For quantification, we utilize the relative RMS error, wherein the acceptance threshold E_{max} serves as a controlling parameter within the process. Given the characteristics of the response feature, the demand for training points to establish a dependable model remains minimal, typically ranging from fifty to one hundred. Depending on the parameter space's dimensionality and ranges, the ratio of accepted observables falls between twenty to seventy percent. Consequently, the actual count of low-fidelity EM simulations required to compile the dataset $\{x_{Bc}^{(j)}, f_P(x_B^{(j)})\}$ varies from $1.5N_{init}$ to $5N_{init}$. In practice, the process of rejecting (possibly significant) subset of the random samples acts as a pre-selection mechanism: it enables the identification of promising regions within the parameter space, concentrating the search process on these areas while excluding others. The pre-screening procedure has been summarized in Table 4.

Generating infill by means of nature-inspired optimization. Co-kriging surrogate

The core part of the optimization run consists of iterative generation of candidate solutions $x^{(i)}$, $i = 1, 2, \ldots$, to the problem (2), (3), and construction of the refined surrogate models $s^{(j)}$, $j = 1, 2, \ldots$. This stage is carried out using the high-resolution model R_f. Each iteration produces a new design

$$x^{(i+1)} = \arg \min_{x \in X} U_F(x, s^{(i)}(x), F_t) \qquad (5)$$

using the current surrogate $s^{(i)}$, which is a co-kriging model (cf. "Parameter space pre-screening. Initial surrogate model construction" section) constructed using the low-fidelity dataset $\{x_{Bc}^{(j)}, f_P(x_{Bc}^{(j)})\}$, $j = 1, \ldots, N_{init}$, and the high-fidelity dataset $\{x_f^{(j)}, f_P(x_f^{(j)})\}$, $j = 1, \ldots, i$, consisting of the samples accumulated until iteration i.

The formulation of the sub-problem (5) is the same as for the task (2), except that the response features are obtained from $s^{(i)}$. Optimization is conducted in a global sense using the particle swarm optimization (PSO) algorithm[97]. PSO was chosen as a representative nature-inspired method, which also belongs to the most popular ones. Yet, any other population-based technique can be used instead because operating at the level of fast surrogate does not imposes any practical constraints on the computational budget expressed as the number

---

1. Input parameters:
   - Design space X (interval [l u], where l and u are lower and upper bounds for designable parameters; other constraints are possible per designer's needs);
   - Required modelling error E_{max};
2. Set $j = 0$;
3. Generate a sample $x_{tmp} \in X$ using uniform probability distribution;
4. Evaluate circuit characteristics at x_{tmp} using low-fidelity EM analysis;
5. Extract the feature vector $f_P(x_{tmp})$; assign $f_P(x_{tmp}) = 0$ if features are not extractable (cf. Fig. 3);
6. **if** $\|f_P(x_{tmp})\| > 0$
      Set $j = j + 1$;
      Set $x_{Bc}^{(j)} = x_{tmp}$;
      Construct surrogate model $s_{tmp}(x)$ using dataset $\{x_{Bc}^{(k)}, f_P(x_{Bc}^{(k)})\}_{k = 1, \ldots, j}$;
      Estimate model error E_{tmp} using K-fold cross-validation[96], $K = \min\{j, 10\}$;
      if $E_{tmp} < E_{max}$
           Go to 8;
      end
   end
7. Go to 3;
8. Return $s^{(0)}(x) = s_{tmp}(x)$;

Table 4. Parameter space pre-screening and initial surrogate model construction. A pseudocode.

of objective function calls. Furthermore, finding the global optimum of $U_F(x,s^{(i)}(x),F_t)$ is considerably easier handling the original merit function $U(x,F_t)$ due to inherent regularization that comes with the employment of response features.

In view of machine learning, generating candidate designs using (5) corresponds to the infill criterion being the predicted objective function improvement[98]. Upon completing the pre-screening stage, the parameter space subset containing the optimum has been presumably identified, therefore, the search process can now be focused on its exploitation, rather than enhancing the overall metamodel's accuracy. The algorithm is stopped either due to convergence in argument, i.e., $||x^{(i+1)} - x^{(i)}|| < \varepsilon$ or if no improvement of the objective function has been detected over the last $N_{no\_improve}$ iterations, whichever occurs first. The default values of the termination thresholds are $\varepsilon = 10^{-2}$ and $N_{no\_improve} = 10$.

Complete optimization framework

The proposed global optimization framework is outlined in Table 5 as a pseudocode and depicted in Fig. 6 as a flow diagram. The pre-screening stage unfolds within Steps 2 and 3, while the heart of the search process—creating infill points and refining the surrogate model—is executed across Steps 4 through 8. Step 9 verifies the termination criteria.

Let us discuss the control parameters compiled in Table 6. It is essential to note that there are only three parameters, with two determining the termination conditions. Their primary function is to govern the resolution of the optimization process. The third parameter manages the accuracy of the initial surrogate. The default value of E_{max} corresponds to ten percent of the relative RMS error, which is a mild condition. A small number of control parameters is an advantage of the method as it eliminates the need for tailoring the setup to the particular being solved. In fact, identical arrangement will be used for all verification experiments described in "Verification experiments and benchmarking" section.

The full-wave electromagnetic simulations were performed on Intel Xeon 2.1 GHz dual-core CPU with 128 GB RAM, using CST Microwave Studio. The optimization framework has been implemented in MATLAB. The particle swarm optimizer and CST simulation software communicate through a Matlab-CST socket. The kriging and co-kriging surrogate models were set up using the SUMO toolbox of[99]. As mentioned earlier, the underlying optimization engine is particle swarm optimizer PSO[100], which is one of the most representative nature-inspired population-based algorithms. PSO processes a swarm of N particles (parameter vectors) x_i and velocity vectors v_i, which stand for the position and the velocity of the ith particle, respectively. These are updated as follows:

---

1. Input parameters:
   - Target operating frequencies F_t (cf. "Microwave design optimization. Problem statement" section);
   - Definition of the response features f_P and the objective function U_F (cf. "Response feature methodology" section);
   - Design space X (interval $[l\ u]$, where l and u are lower and upper bounds for designable parameters);
   - Required modelling error E_{max};
   - Termination thresholds ε and $N_{no\_improve}$;
2. Generate the set of initial samples $\{x_{Bc}^{(k)}, f_P(x_{Bc}^{(k)})\}_{k=1,\ldots,Ninit}$, as described in "Parameter space pre-screening. Initial surrogate model construction" section (cf. Fig. 4), using low-fidelity simulation data;
3. Construct (initial) kriging surrogate model $s^{(0)}(x)$;
4. Set $i = 0$;
5. Use the PSO algorithm to obtain infill point $x^{(i+1)}$ by solving (5):
$$x^{(i+1)} = \arg \min_{x \in X} U_F(\ ,\ ^{(i)}(\ ),\ _t)$$
6. Update the dataset: $\{x_{Bc}^{(k)}, f_P(x_{Bc}^{(k)})\}_{k=1,\ldots,Ninit} \cup \{x_i^{(k)}, f_P(x_i^{(k)})\}_{k=1,\ldots,i}$;
7. Set $i = i + 1$;
8. Construct the co-kriging surrogate model $s^{(i)}(x)$ using the updated dataset;
9. **if** $||x^{(i)} - x^{(i-1)}|| < \varepsilon$ OR no objective function improvement for $N_{no\_improve}$ iterations
      Go to 11;
   end
10. Go to 5;
11. Return $x^* = x^{(i)}$;

Table 5. Pseudocode of the ML framework for feature-based global parameter tuning of microwave components introduced in this study.
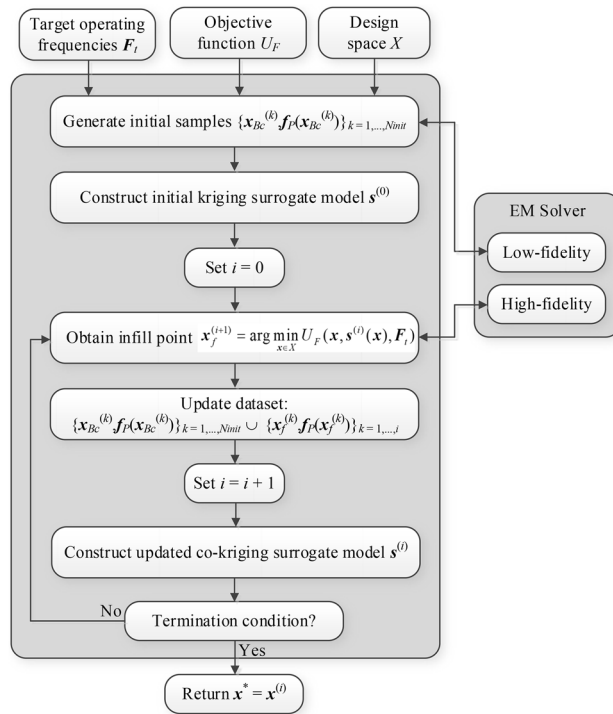
Figure 6. Flow diagram of the proposed ML algorithm for global optimization of microwave passive components.

Parameter	Meaning	Default value
E_{max}	Maximum value of relative RMS error of the initial surrogate model (error estimated using cross-validation), cf. "Parameter space pre-screening. Initial surrogate model construction" section	10%
ε	Termination threshold for convergence in argument, cf. "Generating infill by means of nature-inspired optimization. Co-kriging surrogate" section	10^{-2}
$N_{no\_improve}$	Termination threshold for no objective function value improvement, cf. "Generating infill by means of nature-inspired optimization. Co-kriging surrogate" section	10

Table 6. Control parameters of the proposed ML algorithm for global optimization microwave passives.

$$v_i \leftarrow \chi [v_i + c_1 r_1 \cdot (x_i^* - x_i) + c_2 r_2 \cdot (g - x_i)] \tag{6}$$

$$x_i \leftarrow x_i + v_i \tag{7}$$

where r_1 and r_2 are vectors whose components are uniformly distributed random numbers between 0 and 1; · denotes component-wise multiplication. In our numerical experiments we use a standard setup:

- Size of the swarm $N = 10$,
- Maximum number of iterations $k_{max} = 100$,
- Control parameters, $\chi = 0.73$, $c_1 = c_2 = 2.05$, cf.[100].

As indicated in (7), the first step of altering the positions x_i of the particles is the adjustment of the velocity vector, which is partially stochastic. There are three components therein, one being the current velocity, the second fostering particle relocation towards its local best position x_i^*, and the third one pushing the particle towards global best position g found so far by the swarm. The mentioned setup of control parameters is the most widely used one, typically recommended in the literature[99].

Verification experiments and benchmarking

The validation of the optimization algorithm outlined in "Global machine-learning microwave optimization using response features and multi-resolution computational models" section involves numerical testing with two microstrip circuits. For comparison, these devices are also optimized with the use of several benchmark methods that include a gradient-based search with random starting point, a particle swarm optimizer (PSO), but also two machine-learning frameworks: (i) a procedure that directly handles frequency characteristics of the

circuit, and (ii) the algorithm of "Global machine-learning microwave optimization using response features and multi-resolution computational models" section exclusively using high-fidelity EM simulations. We chose these specific methods to showcase the multimodal nature of the design challenges and to validate the significance of the algorithmic tools embedded in the proposed procedure—particularly the use of response features and variable-fidelity simulations. The performance evaluation criteria encompass the optimization process's reliability (quantified as a success rate, i.e., the proportion of algorithm runs yielding acceptable outcomes), design quality, and computational efficiency.
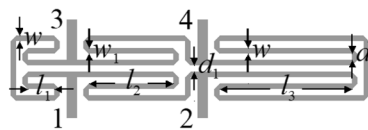
The material in this section is arranged as follows: "Verification circuits" section offers insights into the verification structures. The experimental setup and results are outlined in "Experimental setup. Numerical results" section, while "Discussion" section delves into the characteristics of the techniques considered and encapsulates the overall performance of the proposed approach.

Verification circuits

The machine learning procedure introduced in this study is demonstrated using two planar structures:

- A miniaturized rat-race coupler (RRC) with meandered transmission lines (Circuit I)[101];
- A dual-band power divider with equal division ratio (Circuit II)[102].

The circuit topologies are depicted in Figs. 7a and 8a, respectively. Essential data for both circuits, including substrate parameters, design variables, and design goals, are detailed in Figs. 7b and 8b. The computational models are simulated in CST Microwave Studio, employing the time-domain solver. The low-fidelity models are representations with coarser discretization compared to the high-fidelity versions. For specific details regarding the number of mesh cells and simulation times for R_c and R_f, please refer to Table 7.
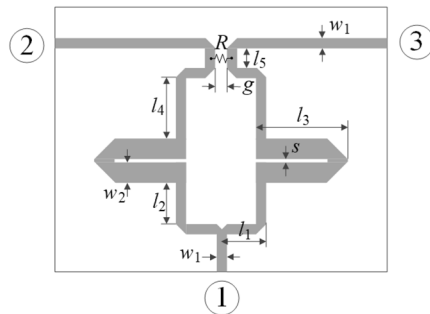

Substrate	RO4003 ($\varepsilon_r = 3.38$, $h = 0.76$ mm)
Design variables	$\mathbf{x} = [l_1\ l_2\ l_3\ d\ w\ w_1]^T$
Other variables	$d_1 = d + |w - w_1|$, $d = 1.0$, $w_0 = 1.7$, $l_0 = 15$
Design goals	• Minimize matching $|S_{11}|$ and isolation $|S_{41}|$ at the operating frequency f_0 • Ensure power split ratio K_P

(a)　　　　　　　　　　　　　　　(b)

Figure 7. Circuit I (rat-race coupler with folded transmission lines)[101]: (**a**) geometry, (**b**) essential parameters.

Substrate	AD250 ($\varepsilon_r = 2.5$, $h = 0.81$ mm)
Design variables	$\mathbf{x} = [l_1\ l_2\ l_3\ l_4\ l_5\ s\ w_2]^T$
Other variables	$w_1 = 2.2$, $g = 1$
Design goals	• Minimize input matching $|S_{11}|$, output matching $|S_{22}|$, $|S_{33}|$, and isolation $|S_{23}|$ at two operating frequencies f_1 and f_2 • Ensure equal power split ratio

(a)　　　　　　　　　　　　　　　(b)

Figure 8. Circuit II (dual-band equal-split power divider)[102]: (**a**) geometry, (**b**) essential parameters.

	EM simulation model			
	Low-fidelity R_c		**High-fidelity R_f**	
Circuit	**Discretization density (# of mesh cells)**	**Simulation time [s]**	**Discretization density (# of mesh cells)**	**Simulation time [s]**
I	~ 50,000	50	~ 180,000	120
II	~ 130,000	60	~ 570,000	310

Table 7. EM simulation models for Circuit I and II.

The design goals for have been set as follows:

- *Circuit I* (i) enhance the matching and isolation responses, $|S_{11}|$ and $|S_{41}|$, at the required frequency f_0; (ii) realize required power division $|S_{31}| - |S_{21}| = K_P$ [dB] (also at f_0).
- *Circuit II* (i) enhance input impedance matching $|S_{11}|$, output matching $|S_{22}|$, $|S_{33}|$, and isolation $|S_{23}|$ at the frequencies f_1 and f_2 corresponding to the two operating bands of the circuit, (ii) ensure equal power division at f_1 and f_2. Here, the second condition automatically follows from the circuit symmetry.

In order to expand the scope of the verification experiments, two design scenarios are considered for each circuit as detailed in Table 8. One can also notice that the search spaces, determined by the bounds on design variables, are extensive: the (mean) upper-to-lower bound ratio is 13 and 5 for Circuit I and II, respectively.

Experimental setup. Numerical results

Circuits I and II were optimized by means of the procedure proposed in this study. The setup uses default values of control parameters, as specified in Table 6: $E_{max} = 10\%$, $\varepsilon = 10^{-2}$, and $N_{no\_improve} = 10$. It is kept identical for all considered design cases. For the purpose of comparison, our verification structures were also optimized using several benchmark methods. The technical details on these techniques can be found in Table 9. Below, we explain the reasons for selecting this particular testbed:

- Algorithm I: particle swarm optimizer (PSO), chosen as one of the most popular nature-inspired algorithms to date. Its inclusion in the benchmark set aims to showcase the difficulties associated with direct EM-driven population-based optimization. For practical reasons, PSO is run at low computational budgets of 500 (version I) and 1000 objective function evaluations (version II), yet, the computational costs entailed even with these budgets are at the edge of being prohibitive;
- Algorithm II: A gradient-based search with random initial design. This method is selected to illustrate multi-modality of our optimization tasks, thereby indicating the need for employing global search methodologies;
- Algorithm III: Machine-learning procedure processing complete frequency responses of the circuit at hand. The algorithm utilizes the similar components as those described in "Global machine-learning microwave optimization using response features and multi-resolution computational models" section (kriging surrogates, predicted objective function improvement) but works exclusively with the high-resolution EM model. This method is selected to investigate the benefits of incorporating response features.
- Algorithm IV: Feature-based machine-learning procedure similar to that described in "Global machine-learning microwave optimization using response features and multi-resolution computational models" section, but only using high-fidelity EM models. Consequently, kriging surrogate is employed throughout the

Circuit	Target operating parameters		Parameter space X (lower bounds l and upper bounds u)
	Symbols	Specific values for numerical experiments	
I	$F_t = [f_0\ K_P]^T$	Case 1: $f_0 = 1.8$ GHz, $K_P = -3$ dB	$l = [0.5\ 5.0\ 5.0\ 0.2\ 0.2\ 0.2]^T$ $u = [15.0\ 30.0\ 50.0\ 2.0\ 2.0\ 2.0]^T$
		Case 2: $f_0 = 1.2$ GHz, $K_P = 0$ dB	
II	$F_t = [f_1\ f_2]^T$	Case 1: $f_1 = 3.0$ GHz, $f_2 = 4.8$ GHz	$l = [10.0\ 1.0\ 10.0\ 0.5\ 1.0\ 0.1\ 1.5]^T$ $u = [40.0\ 20.0\ 40.0\ 15.0\ 6.0\ 1.5\ 8.0]^T$
		Case 2: $f_1 = 2.0$ GHz, $f_2 = 3.3$ GHz	

Table 8. Design goals and parameter spaces for Circuits I and II.

Algorithm	Algorithm type	Setup
I	Particle swarm optimizer (PSO)	Swarm size $N = 10$, standard control parameters ($\chi = 0.73$, $c_1 = c_2 = 2.05$); number of iterations set to 50 (version I) and 100 (version II)
II	Trust-region gradient based optimizer[103]	Random initial design, response gradients estimated using finite differentiation, termination criteria based on convergence in argument and reduction of the trust region size[103]
III	Machine learning algorithm (cf. "Global machine-learning microwave optimization using response features and multi-resolution computational models" section)	Algorithm similar to that of "Global machine-learning microwave optimization using response features and multi-resolution computational models" section 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 antenna frequency characteristics (unlike response features in the proposed procedure); Infill criterion: minimization of the projected objective function improvement[98]
IV	Feature-based machine learning algorithm utilizing high-fidelity EM simulations only	Algorithm highlights Surrogate model constructed at the level of response features; Optimization process only uses high-fidelity EM simulations; Infill criterion: minimization of the predicted objective function[98]

Table 9. Benchmark algorithms: the outline.

entire optimization run instead of co-kriging. This approach serves to highlight the benefits of integrating variable-fidelity models into the method presented in this work.

Each algorithm underwent ten executions, and the average outcomes (merit function value, computational expenses) are documented in Tables 10 and 11 for Circuit I and II, respectively. The success rate denotes the number of runs (out of ten) generating designs that sufficiently match their target operating parameters. Figures 9, 10, 11, and 12 exhibit the circuit responses and the progression of the objective function value across both design cases and chosen algorithm runs.

Optimization algorithm		Performance figure		
		Average objective function value [dB]	Computational cost[a]	Success rate[b]
Case 1 $f_0 = 1.8$ GHz, $K_P = -3$ dB	Algorithm I: PSO (50 iterations)	− 24.8	500	9/10
	Algorithm I: PSO (100 iterations)	− 34.0	1000	10/10
	Algorithm II: Trust-region gradient-based algorithm	− 18.7	102.8	6/10
	Algorithm III: Machine learning procedure handling complete circuit responses	− 34.1	435.7	10/10
	Algorithm IV: Feature-based machine learning procedure using high-fidelity model only	− 34.8	89.9	10/10
	Proposed algorithm (feature-based machine learning with variable-fidelity simulation models)	− 34.5	48.2	10/10
Case 2 $f_0 = 1.2$ GHz, $K_P = 0$ dB	Algorithm I: PSO (50 iterations)	− 23.7	500	9/10
	Algorithm I: PSO (100 iterations)	− 36.2	1000	10/10
	Algorithm II: Trust-region gradient-based algorithm	48.3	68.7	5/10
	Algorithm III: Machine learning procedure handling complete circuit responses	− 45.6	433.6	10/10
	Algorithm IV: Feature-based machine learning procedure using high-fidelity model only	− 43.9	106.3	10/10
	Proposed algorithm (feature-based machine learning with variable-fidelity simulation models)	− 44.1	59.3	10/10

Table 10. Optimization of Circuit I: results. [a]The cost expressed in terms of the number of high-fidelity EM simulations of the circuit under design. [b]Number of algorithms runs at which the operating frequency was allocated near the target value.

Optimization algorithm		Performance figure		
		Average objective function value [dB]	Computational cost[a]	Success rate[b]
Case 1 $f_1 = 3.0$ GHz, $f_2 = 4.8$ GHz	Algorithm I: PSO (50 iterations)	− 19.6	500	8/10
	Algorithm I: PSO (100 iterations)	− 18.8	1000	9/10
	Algorithm II: Trust-region gradient-based algorithm	− 12.3	95.1	2/10
	Algorithm III: Machine learning procedure handling complete circuit responses	− 32.0	446.8	10/10
	Algorithm IV: Feature-based machine learning procedure using high-fidelity model only	− 32.8	194.4	10/10
	Proposed algorithm (feature-based machine learning with variable-fidelity simulation models)	− 32.3	85.2	10/10
Case 2 $f_1 = 2.0$ GHz, $f_2 = 3.3$ GHz	Algorithm I: PSO (50 iterations)	− 18.8	500	8/10
	Algorithm I: PSO (100 iterations)	− 19.7	1000	9/10
	Algorithm II: Trust-region gradient-based algorithm	− 20.6	93.8	7/10
	Algorithm III: Machine learning procedure handling complete circuit responses	− 22.5	430.5	10/10
	Algorithm IV: Feature-based machine learning procedure using high-fidelity model only	− 21.1	153.2	10/10
	Proposed algorithm (feature-based machine learning with variable-fidelity simulation models)	− 21.9	52.4	10/10

Table 11. Optimization of Circuit II: results. [a]The cost expressed in terms of the number of high-fidelity EM simulations of the circuit under design. [b]Number of algorithms runs at which the operating frequency was allocated near the target value.
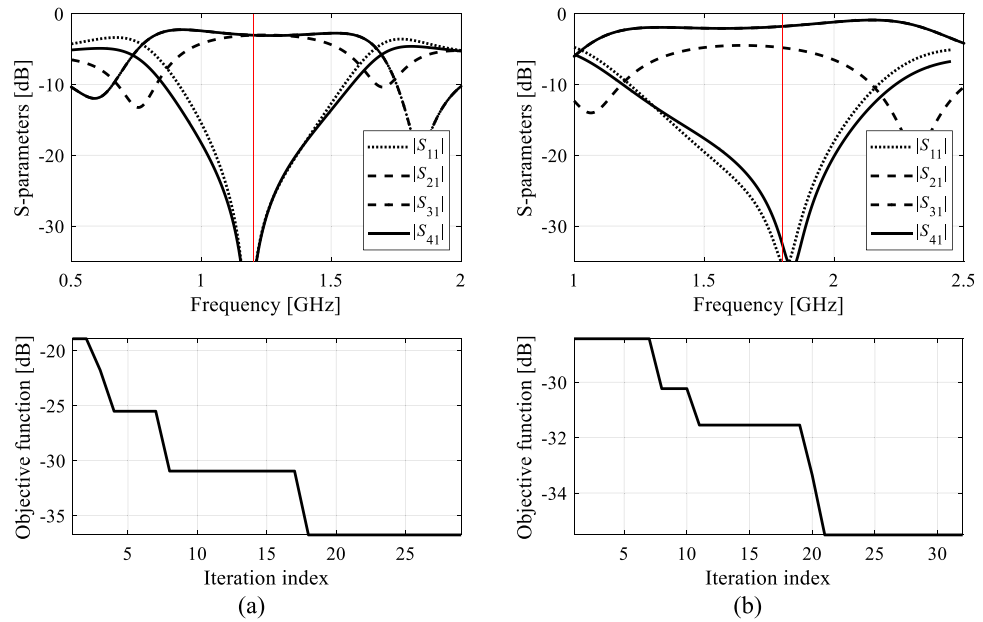
Figure 9. Exemplary runs of the proposed machine-learning framework. Shown are: *S*-parameters of Circuit I (Case 1) at the designs produced by the proposed technique (top), and the evolution of the objective function value (bottom): (**a**) run 1, (**b**) run 2. The iteration counter starts after constructing the initial surrogate model. Target operating frequency, here, 1.8 GHz, marked using the vertical lines.
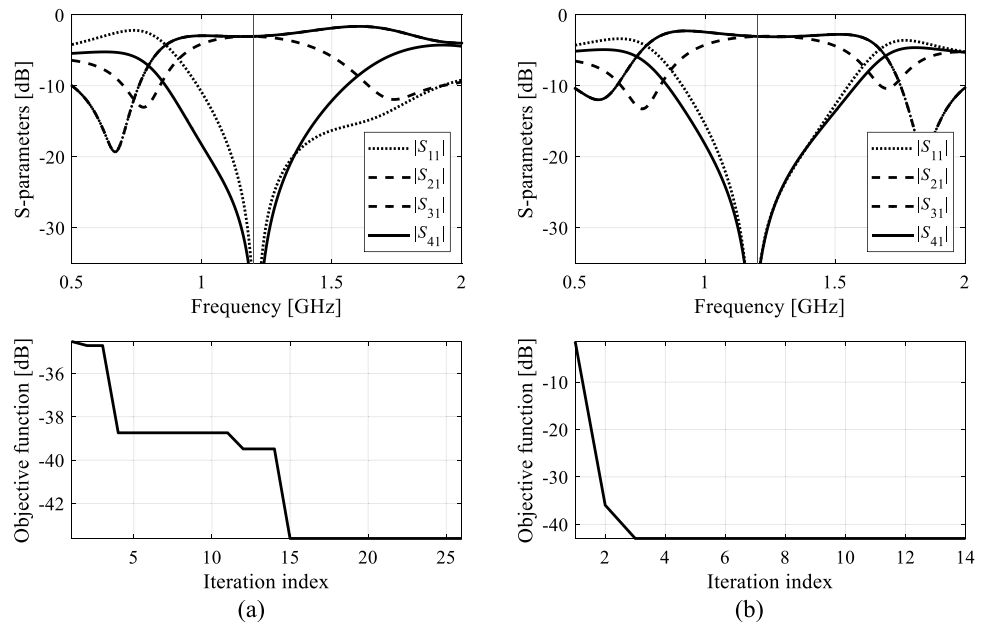

Figure 10. Exemplary runs of the proposed machine-learning framework. Shown are: *S*-parameters of Circuit I (Case 2) at the designs produced by the proposed technique (top), and the evolution of the objective function value (bottom): (**a**) run 1, (**b**) run 2. The iteration counter starts after constructing the initial surrogate model. Target operating frequency, here, 1.2 GHz, marked using vertical lines.

Discussion

The data presented in Tables 10 and 11 are examined here to evaluate the efficiency of the suggested machine-learning approach and to juxtapose it with benchmark techniques. We're particularly focused on evaluating the following metrics: the search process's reliability assessed through the success rate, the design quality evaluated using the merit function value, and the cost efficiency of the global search process.
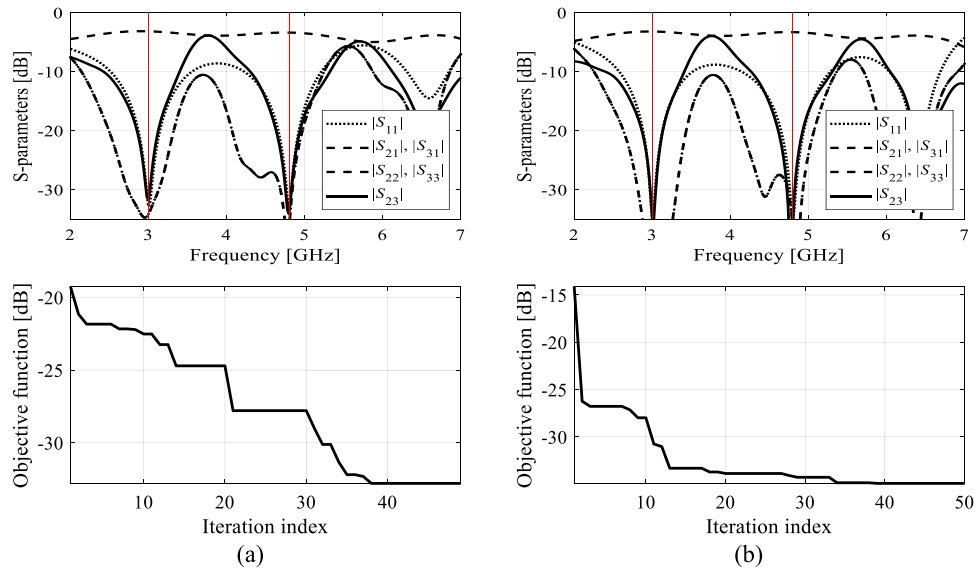
Figure 11. Exemplary runs of the proposed machine-learning framework. Shown are: *S*-parameters of Circuit II (Case 1) at the designs produced by the proposed technique (top), and the evolution of the objective function value (bottom): (**a**) run 1, (**b**) run 2. The iteration counter starts after constructing the initial surrogate model. Target operating frequencies, 3.0 GHz and 4.8 GHz, marked using vertical lines.
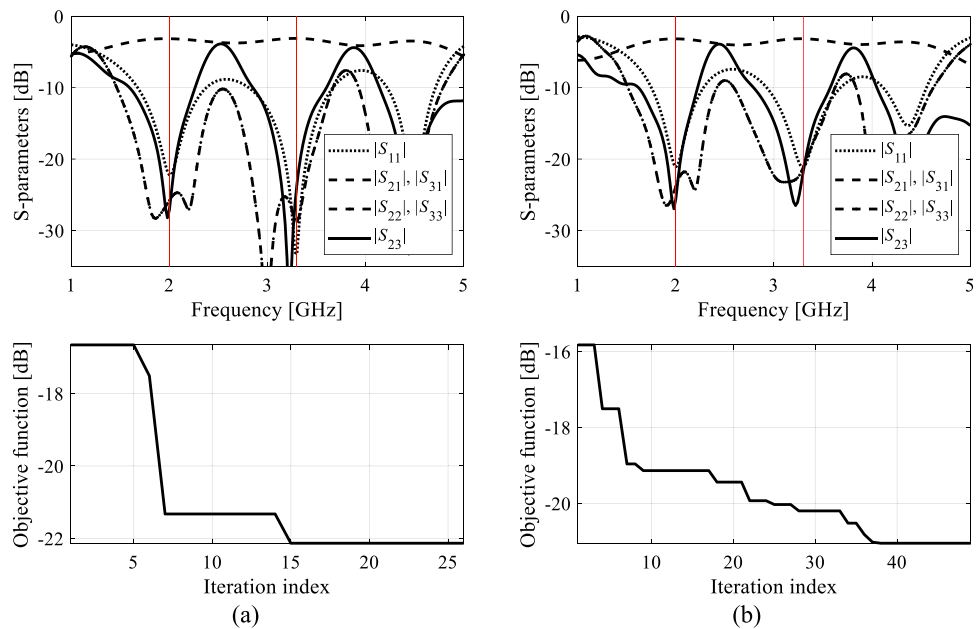

Figure 12. Exemplary runs of the proposed machine-learning framework. Shown are: *S*-parameters of Circuit II (Case 2) at the designs produced by the proposed technique (top), and the evolution of the objective function value (bottom): (**a**) run 1, (**b**) run 2. The iteration counter starts after constructing the initial surrogate model. Target operating frequencies, 2.0 GHz and 3.3 GHz, marked using vertical lines.

Another point for discussion is the effect of incorporating the response features and variable-fidelity simulation models. The observations are as follows:

- *Search process reliability* Reliability is measured using the success rate (right-hand-side columns of Tables 10 and 11), which is the number of algorithm runs yielding the designs that feature operating parameters being close to the targets. For the proposed method, the success rate is 10/10, just as it is for both benchmark machine learning methods (Algorithms III and IV) and PSO (Algorithm I) set up with the budget of 1000 objective function calls. However, PSO working with the computational budget of 500 function evaluations

does not perform as well. Also, for Circuit II, which is a more challenging of the two test problems, the success rate of PSO is 9/10 even the budget of 1000 function calls. This indicates that direct nature-inspired optimization needs higher budgets (e.g., > 2000) to ensure the perfect score. At the same time, gradient-based optimization routinely fails (e.g., the success rate is only 2/10 for Circuit II, Case 1), which corroborates multimodality of the considered design tasks.

- *Design quality* The solutions produced using the presented procedure exhibit comparable quality to those produced by other benchmark methods, as indicated by the average objective function value with the exception of gradient-based search, where the failed algorithm runs degrade the average. In terms of absolute numbers, all machine learning methods deliver results that are sufficiently good for practical applications, i.e., matching and isolation below – 20 dB, and power division close to the required values of 0 dB and 3 dB for Circuits I and II, respectively

- *Computational efficiency* The efficiency of the proposed framework is by far the best across the entire benchmark set and the considered test problems. We omit comparison with gradient-based algorithms as they were included in the benchmark solely to highlight the necessity of global optimization for the test problems considered. The average computational cost, quantified in terms of equivalent high-fidelity EM simulations, is approximately sixty. This signifies considerable savings compared to PSO (with the budget of 1000 function calls) are between 90 and 95 percent, depending on the test case. The savings over Algorithm II (machine learning processing full circuit responses) are about 85 percent on the average, and about 50 percent over Algorithm III (feature-based machine learning working at single-fidelity EM level). Regarding the latter, the average savings are 45 and over sixty percent for Circuit I and II, respectively, because the time evaluation ratio between the high- and low-fidelity model is more advantageous for the latter (5.2 versus 2.4).

- The above numbers confirm that the employment of both response features and variable-fidelity are instrumental in expediting the search process without being detrimental to its reliability. The speedup obtained due to variable-fidelity modeling is significant also because the majority of the circuit evaluations are associated with the parameter space pre-screening, which, in the proposed methods, is conducted using the low-fidelity system representation. A better perspective of the computational benefits due to the mentioned mechanisms can be provided by considering the acceleration factors. For example, the proposed methods is over 17 times faster than PSO, almost eight times faster than Algorithm III, and over twice as fast as Algorithm IV.

The overall performance of the presented methodology is quite promising. The variable-fidelity feature-based machine learning yields consistent results at remarkably low computational cost. Consequently, it seems to be suitable for replacing less efficient global optimization methods in the field of high-frequency EM-driven design. At this point, we should also discuss its potential limitation, which is related to the response feature aspect of the procedure. On one hand, defining and extracting characteristic points from the system outputs for a specific structure and design task adds complexity to the implementation process, albeit not in its core segment, which remains problem independent.

On the other hand, for highly-dimensional problems and very broad parameter ranges, the number of random observables generated in the pre-screening stage may be large as compared to those that exhibit extractable features. Such factors would compromise the computational efficiency of the methods, impacting their overall performance. Naturally, these same factors would also impede the efficacy of all benchmark methods. Yet, assuming that the parameter space is defined by an experienced designer, i.e., it is not excessively large, the risk of the occurrence of the aforementioned issue is rather low.

Conclusion

This paper introduced a machine-learning framework designed for the efficient global optimization of passive microwave components. Our methodology integrates several crucial mechanisms pivotal for achieving competitive reliability and minimizing search process expenses. These mechanisms encompass the response feature approach, parameter space pre-screening, and the utilization of variable-fidelity EM simulations. The response feature method helps in regularizing the objective function landscape, consequently reducing the necessary dataset size for constructing precise surrogate models. Pre-screening of the parameter space aids in the initial identification of the most promising regions, while variable-resolution models contribute to additional computational acceleration. Within this framework, both low- and high-fidelity simulation data are combined into a unified surrogate model using co-kriging. The optimization process itself focuses on rapid identification of the optimum design, which is facilitated by the infill criterion applied in our framework (predicted objective function improvement). The particle swarm optimization (PSO) algorithm serves as the underlying search engine. Numerical verification experiments involving two microstrip components illustrate the superior performance of the proposed technique, showcasing its ability to achieve superior design quality, reliability, and computational efficiency. The CPU savings versus nature-inspired optimization are up to 95 percent (average acceleration factor of 17), 85 percent over the machine learning procedure working directly with circuit frequency characteristics (acceleration factor of eight), and 50 percent over the feature-based machine learning algorithm that only uses high-fidelity EM models (acceleration factor of two). The mean running cost corresponds to sixty high-resolution EM simulations. This level of expenses is comparable to local optimization. Consequently, the presented framework seems to be an attractive alternative to both conventional and surrogate-assisted global optimization procedures (including machine-learning algorithms) utilized so far in high-frequency engineering. A primary objective for future work involves extending the method to encompass other types of microwave components, such as filters. This extension would involve automating the procedures for defining and extracting feature points. Furthermore, utilization of alternative machine learning technique will be considered, oriented towards improving the reliability of surrogate model construction as well as computational efficiency of the modelling

process. Some of the tools to be employed include deep learning methods (e.g., convolutional neural networks, etc.). Finally, the properties of the presented technique will be investigated when applied to higher-dimensional problems (ten design variables and more), and enhancements wil be developed to facilitate operation of the framework under such challenging scenarios.

Data availability

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

References

1. 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. Compon. Lett.* **31**(5), 441–444 (2021).
2. Erman, F., Koziel, S., Hanafi, E., Soboh, R. & Szczepanski, S. Miniaturized metal-mountable U-shaped inductive-coupling-fed UHF RFID tag antenna with defected microstrip surface. *IEEE Access* **10**, 47301–47308 (2022).
3. 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).
4. Matos, D., da Cruz Jordão, M. D., Correia, R. & Carvalho, N. B. Millimeter-wave BiCMOS backscatter modulator for 5 G-IoT applications. *IEEE Microw. Wirel. Compon. Lett.* **31**(2), 173–176 (2021).
5. Hu, Y.-Y., Sun, S. & Xu, H. Compact collinear quasi-Yagi antenna array for wireless energy harvesting. *IEEE Access* **8**, 35308–35317 (2020).
6. Li, Q., Chen, X., Chi, P. & Yang, T. Tunable bandstop filter using distributed coupling microstrip resonators with capacitive terminal. *IEEE Microw. Wirel. Compon. Lett.* **30**(1), 35–38 (2020).
7. Liu, M. & Lin, F. Two-section broadband couplers with wide-range phase differences and power-dividing ratios. *IEEE Microw. Wirel. Compon. Lett.* **31**(2), 117–120 (2021).
8. Gómez-García, R., Rosario-De Jesus, J. & Psychogiou, D. Multi-band bandpass and bandstop RF filtering couplers with dynamically-controlled bands. *IEEE Access* **6**, 32321–32327 (2018).
9. Zhang, R. & Peroulis, D. Mixed lumped and distributed circuits in wideband bandpass filter application for spurious-response suppression. *IEEE Microw. Wirel. Compon. Lett.* **28**(11), 978–980 (2018).
10. He, Z. & Liu, C. A compact high-efficiency broadband rectifier with a wide dynamic range of input power for energy harvesting. *IEEE Microw. Wirel. Compon. Lett.* **30**(4), 433–436 (2020).
11. Jiang, Z. H., Gregory, M. D. & Werner, D. H. Design and experimental investigation of a compact circularly polarized integrated filtering antenna for wearable biotelemetric devices. *IEEE Trans. Biomedical Circuits Syst.* **10**(2), 328–338 (2016).
12. Kracek, J., Švanda, M., Mazanek, M. & Machac, J. Implantable semi-active UHF RFID tag with inductive wireless power transfer. *IEEE Ant. Wirel. Propag. Lett.* **15**, 1657–1660 (2016).
13. Firmansyah, T., Alaydrus, M., Wahyu, Y., Rahardjo, E. T. & Wibisono, G. A highly independent multiband bandpass filter using a multi-coupled line stub-SIR with folding structure. *IEEE Access* **8**, 83009–83026 (2020).
14. Chen, S. *et al.* A frequency synthesizer based microwave permittivity sensor using CMRC structure. *IEEE Access* **6**, 8556–8563 (2018).
15. Zhu, Y., Wang, J., Hong, J., Chen, J.-X. & Wu, W. Two- and three-way filtering power dividers with harmonic suppression using triangle patch resonator. *IEEE Trans. Circuits Syst. I Regul. Pap.* **68**(12), 5007–5017 (2021).
16. 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. Compon. Lett.* **25**(2), 76–78 (2015).
17. Koziel, S. & Bandler, J. W. Space mapping with multiple coarse models for optimization of microwave components. *IEEE Microw. Wirel. Compon. Lett.* **18**, 1–3 (2008).
18. Chione, G. & Pirola, M. *Microwave Electronics* (Cambridge University Press, 2018).
19. Ullah, U., Al-Hasan, M., Koziel, S. & Ben Mabrouk, I. Series-slot-fed circularly polarized multiple-input-multiple-output antenna array enabling circular polarization diversity for 5G 28-GHz indoor applications. *IEEE Trans. Ant. Prop.* **69**(9), 5607–5616 (2021).
20. Zhu, Y. & Dong, Y. A novel compact wide-stopband filter with hybrid structure by combining SIW and microstrip technologies. *IEEE Microw. Wirel. Compon. Lett.* **31**(7), 841–844 (2021).
21. Koziel, S., Pietrenko-Dabrowska, A. & Plotka, P. Reduced-cost microwave design closure by multi-resolution EM simulations and knowledge-based model management. *IEEE Access* **9**, 116326–116337 (2021).
22. 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).
23. 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).
24. Shen, Z., Xu, K., Mbongo, G. M., Shi, J. & Yang, Y. Compact balanced substrate integrated waveguide filter with low insertion loss. *IEEE Access* **7**, 126111–126115 (2019).
25. Li, Y., Ren, P. & Xiang, Z. A dual-passband frequency selective surface for 5G communication. *IEEE Antennas Wirel. Propag. Lett.* **18**(12), 2597–2601 (2019).
26. Abdullah, M. & Koziel, S. Supervised-learning-based development of multi-bit RCS-reduced coding metasurfaces. *IEEE Trans. Microw. Theory Tech.* **70**(1), 264–274 (2021).
27. Blankrot, B. & Heitzinger, C. Efficient computational design and optimization of dielectric metamaterial structures. *IEEE J. Multiscale Multiphysics Comp. Tech.* **4**, 234–244 (2019).
28. Ma, Y., Yang, S., Chen, Y., Qu, S.-W. & Hu, J. Pattern synthesis of 4-D irregular antenna arrays based on maximum-entropy model. *IEEE Trans. Antennas Propag.* **67**(5), 3048–3057 (2019).
29. Tang, M., Chen, X., Li, M. & Ziolkowski, R. W. Particle swarm optimized, 3-D-printed, wideband, compact hemispherical antenna. *IEEE Antennas Wirel. Propag. Lett.* **17**(11), 2031–2035 (2018).
30. Li, H., Jiang, Y., Ding, Y., Tan, J. & Zhou, J. Low-sidelobe pattern synthesis for sparse conformal arrays based on PSO-SOCP optimization. *IEEE Access* **6**, 77429–77439 (2018).
31. 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).
32. Rayas-Sanchez, J. E., Koziel, S. & Bandler, J. W. Advanced RF and microwave design optimization: A journey and a vision of future trends. *IEEE J. Microw.* **1**(1), 481–493 (2021).
33. Abdullah, M. & Koziel, S. A novel versatile decoupling structure and expedited inverse-model-based re-design procedure for compact single-and dual-band MIMO antennas. *IEEE Access* **9**, 37656–37667 (2021).

34. Jin, H., Zhou, Y., Huang, Y. M., Ding, S. & Wu, K. Miniaturized broadband coupler made of slow-wave half-mode substrate integrated waveguide. *IEEE Microw. Wirel. Compon. Lett.* **27**(2), 132–134 (2017).

35. 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. Compon. Lett.* **29**(5), 315–317 (2019).

36. Shum, K. M., Luk, W. T., Chan, C. H. & Xue, Q. A UWB bandpass filter with two transmission zeros using a single stub with CMRC. *IEEE Microw. Wirel. Compon. Lett.* **17**(1), 43–45 (2007).

37. Li, X. & Luk, K. M. The grey wolf optimizer and its applications in electromagnetics. *IEEE Trans. Antennas Propag.* **68**(3), 2186–2197 (2020).

38. Luo, X., Yang, B. & Qian, H. J. Adaptive synthesis for resonator-coupled filters based on particle swarm optimization. *IEEE Trans. Microw. Theory Tech.* **67**(2), 712–725 (2019).

39. Majumder, A., Chatterjee, S., Chatterjee, S., Sinha Chaudhari, S. & Poddar, D. R. Optimization of small-signal model of GaN HEMT by using evolutionary algorithms. *IEEE Microw. Wirel. Compon. Lett.* **27**(4), 362–364 (2017).

40. Oyelade, O. N., Ezugwu, A.E.-S., Mohamed, T. I. A. & Abualigah, L. Ebola optimization search algorithm: A new nature-inspired metaheuristic optimization algorithm. *IEEE Access* **10**, 16150–16177 (2022).

41. Milner, S., Davis, C., Zhang, H. & Llorca, J. Nature-inspired self-organization, control, and optimization in heterogeneous wireless networks. *IEEE Trans. Mobile Comput.* **11**(7), 1207–1222 (2012).

42. Zhao, Q. & Li, C. Two-stage multi-swarm particle swarm optimizer for unconstrained and constrained global optimization. *IEEE Access* **8**, 124905–124927 (2020).

43. Jiacheng, L. & Lei, L. A hybrid genetic algorithm based on information entropy and game theory. *IEEE Access* **8**, 36602–36611 (2020).

44. Zhao, Z., Wang, X., Wu, C. & Lei, L. Hunting optimization: A new framework for single objective optimization problems. *IEEE Access* **7**, 31305–31320 (2019).

45. Zhang, Q. & Liu, L. Whale optimization algorithm based on Lamarckian learning for global optimization problems. *IEEE Access* **7**, 36642–36666 (2019).

46. Ismaeel, A. A. K., Elshaarawy, I. A., Houssein, E. H., Ismail, F. H. & Hassanien, A. E. Enhanced elephant herding optimization for global optimization. *IEEE Access* **7**, 34738–34752 (2019).

47. Wang, P., Rao, Y. & Luo, Q. An effective discrete grey wolf optimization algorithm for solving the packing problem. *IEEE Access* **8**, 115559–115571 (2020).

48. Liu, F., Liu, Y., Han, F., Ban, Y. & Jay Guo, Y. Synthesis of large unequally spaced planar arrays utilizing differential evolution with new encoding mechanism and cauchy mutation. *IEEE Trans. Antennas Propag.* **68**(6), 4406–4416 (2020).

49. Kovaleva, M., Bulger, D. & Esselle, K. P. Comparative study of optimization algorithms on the design of broadband antennas. *IEEE J. Multiscale Multiphysics Comput. Tech.* **5**, 89–98 (2020).

50. Ghorbaninejad, H. & Heydarian, R. New design of waveguide directional coupler using genetic algorithm. *IEEE Microw. Wirel. Compon. Lett.* **26**(2), 86–88 (2016).

51. Ding, D., Zhang, Q., Xia, J., Zhou, A. & Yang, L. Wiggly parallel-coupled line design by using multiobjective evolutionay algorithm. *IEEE Microw. Wirel. Compon. Lett.* **28**(8), 648–650 (2018).

52. 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).

53. Cui, C., Jiao, Y. & Zhang, L. Synthesis of some low sidelobe linear arrays using hybrid differential evolution algorithm integrated with convex programming. *IEEE Antennas Wirel. Propag. Lett.* **16**, 2444–2448 (2017).

54. Baumgartner, P. *et al.* Multi-objective optimization of Yagi-Uda antenna applying enhanced firefly algorithm with adaptive cost function. *IEEE Trans. Magn.* **54**(3), 8000504 (2018).

55. Zhu, D. Z., Werner, P. L. & Werner, D. H. Design and optimization of 3-D frequency-selective surfaces based on a multiobjective lazy ant colony optimization algorithm. *IEEE Trans. Antennas Propag.* **65**(12), 7137–7149 (2017).

56. Li, X. & Guo, Y.-X. Multiobjective optimization design of aperture illuminations for microwave power transmission via multi-objective grey wolf optimizer. *IEEE Trans. Antennas Propag.* **68**(8), 6265–6276 (2020).

57. Liang, S. *et al.* Sidelobe reductions of antenna arrays via an improved chicken swarm optimization approach. *IEEE Access* **8**, 37664–37683 (2020).

58. Li, W., Zhang, Y. & Shi, X. Advanced fruit fly optimization algorithm and its application to irregular subarray phased array antenna synthesis. *IEEE Access* **7**, 165583–165596 (2019).

59. Jiang, Z. J., Zhao, S., Chen, Y. & Cui, T. J. Beamforming optimization for time-modulated circular-aperture grid array with DE algorithm. *IEEE Ant. Wireless Propag. Lett.* **17**(12), 2434–2438 (2018).

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

61. 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. Compon. Lett.* **30**(7), 625–628 (2020).

62. Van Nechel, E., Ferranti, F., Rolain, Y. & Lataire, J. Model-driven design of microwave filters based on scalable circuit models. *IEEE Trans. Microw. Theory Tech.* **66**(10), 4390–4396 (2018).

63. Jacobs, J. P. Characterization by Gaussian processes of finite substrate size effects on gain patterns of microstrip antennas. *IET Microw. Antennas Propag.* **10**(11), 1189–1195 (2016).

64. 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), 1–4 (2016).

65. Ogut, M., Bosch-Lluis, X. & Reising, S. C. A deep learning approach for microwave and millimeter-wave radiometer calibration. *IEEE Trans. Geosci. Remote Sens.* **57**(8), 5344–5355 (2019).

66. Petrocchi, A. *et al.* Measurement uncertainty propagation in transistor model parameters via polynomial chaos expansion. *IEEE Microw. Wirel. Compon. Lett.* **27**(6), 572–574 (2017).

67. Na, W. *et al.* Efficient EM optimization exploiting parallel local sampling strategy and Bayesian optimization for microwave applications. *IEEE Microw. Wirel. Compon. Lett.* **31**(10), 1103–1106 (2021).

68. Couckuyt, I., Declercq, F., Dhaene, T., Rogier, H. & Knockaert, L. Surrogate-based infill optimization applied to electromagnetic problems. *Int. J. RF Microw. Comput. Aided Eng.* **20**(5), 492–501 (2010).

69. Chen, C., Liu, J. & Xu, P. Comparison of infill sampling criteria based on Kriging surrogate model. *Sc. Rep.* **12**, 678 (2022).

70. Forrester, A. I. J. & Keane, A. J. Recent advances in surrogate-based optimization. *Prog. Aerospace Sci.* **45**, 50–79 (2009).

71. Tak, J., Kantemur, A., Sharma, Y. & Xin, H. A 3-D-printed W-band slotted waveguide array antenna optimized using machine learning. *IEEE Ant. Wireless Propag. Lett.* **17**(11), 2008–2012 (2018).

72. Wu, Q., Wang, H. & Hong, W. Multistage collaborative machine learning and its application to antenna modeling and optimization. *IEEE Trans. Antennas. Propag.* **68**(5), 3397–3409 (2020).

73. 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.* **54**(11), 8107605 (2018).

74. 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).

75. Toktas, A., Ustun, D. & Tekbas, M. Multi-objective design of multi-layer radar absorber using surrogate-based optimization. *IEEE Trans. Microw. Theory Tech.* **67**(8), 3318–3329 (2019).
76. Pietrenko-Dabrowska, A. & Koziel, S. Antenna modeling using variable-fidelity EM simulations and constrained co-kriging. *IEEE Access* **8**(1), 91048–91056 (2020).
77. Koziel, S & Pietrenko-Dabrowska, A. *Performance-Driven Surrogate Modeling of High-Frequency Structures* (Springer, 2020).
78. Koziel, S. Low-cost data-driven surrogate modeling of antenna structures by constrained sampling. *IEEE Antennas Wirel. Propag. Lett.* **16**, 461–464 (2017).
79. Koziel, S. & Pietrenko-Dabrowska, A. Performance-based nested surrogate modeling of antenna input characteristics. *IEEE Trans. Antennas Propag.* **67**(5), 2904–2912 (2019).
80. Pietrenko-Dabrowska, A. & Koziel, S. Generalized formulation of response features for reliable optimization of antenna input characteristics. *IEEE Trans. Antennas Propag.* **70**(5), 3733–3748 (2021).
81. Brase, C. H. *Understanding Basic Statistics* 8th edn. (Centage Learning, 2018).
82. Alex, S. A., Jhanjhi, N. Z., Humayun, M., Ibrahim, A. O. & Abulfaraj, A. W. Deep LSTM model for diabetes prediction with class balancing by SMOTE. *Electronics* **11**(17), 2737 (2022).
83. Pozar, D. M. *Microwave Engineering* 4th edn. (Wiley, 2011).
84. Ullah, U., Koziel, S. & Mabrouk, I. B. Rapid re-design and bandwidth/size trade-offs for compact wideband circular polarization antennas using inverse surrogates and fast EM-based parameter tuning. *IEEE Trans. Antennas Propag.* **68**(1), 81–89 (2019).
85. Marler, R. T. & Arora, J. S. The weighted sum method for multi-objective optimization: new insights. *Structural Multidisc. Opt.* **41**, 853–862 (2010).
86. Koziel, S. Fast simulation-driven antenna design using response-feature surrogates. *Int. J. RF Microw. Comput. Aided Eng.* **25**(5), 394–402 (2015).
87. 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).
88. Koziel, S. & Bandler, J. W. Reliable microwave modeling by means of variable-fidelity response features. *IEEE Trans. Microw. Theory Tech.* **63**(12), 4247–4254 (2015).
89. 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).
90. Koziel, S. & Ogurtsov, S. *Simulation-Based Optimization of Antenna Arrays* (World Scientific, 2019).
91. 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. Comput. Aided Eng.* **26**(2), 121–128 (2016).
92. Koziel, S. & Bandler, J. W. A space-mapping approach to microwave device modeling exploiting fuzzy systems. *IEEE Trans. Microw Theory Tech.* **55**(12), 2539–2547 (2007).
93. Pietrenko-Dabrowska, A. & Koziel, S. Surrogate modeling of impedance matching transformers by means of variable-fidelity EM simulations and nested co-kriging. *Int. J. RF Microw. Comput. Aided Eng.* **30**(8), e22268 (2020).
94. Kennedy, M. C. & O'Hagan, A. Predicting the output from complex computer code when fast approximations are available. *Biometrika* **87**, 1–13 (2000).
95. Koziel, S. & Pietrenko-Dabrowska, A. Recent advances in high-frequency modeling by means of domain confinement and nested kriging. *IEEE Access* **8**, 189326–189342 (2020).
96. Cawley, G. C. & Talbot, N. L. C. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J. Machine Learning* **11**, 2079–2107 (2010).
97. Vinod Chandra, S. S. & Anand, H. S. Nature inspired meta heuristic algorithms for optimization problems. *Computing* **104**, 251–269 (2022).
98. Liu, J., Han, Z. & Song, W. Comparison of infill sampling criteria in kriging-based aerodynamic optimization. In *28th Int. Congress of the Aeronautical Sciences, Brisbane, Australia, 23–28 Sept.* 1–10 (2012).
99. 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).
100. Clerc, M. *Particle Swarm Optimization* 1st edn. (Wiley-ISTE, 2013).
101. 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).
102. 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).
103. Conn, A. R., Gould, N. I. M. & Toint, P. L. *Trust Region Methods*, MPS-SIAM Series on Optimization (2000).

Acknowledgements

Author contributions

Conceptualization, S.K. (Slawomir Koziel) and A.P.D. (Anna Pietrenko-Dabrowska); Data curation, S.K.; Formal analysis. S.K.; Funding acquisition, S.K. and A.P.D.; Visualization, S.K. and A.P.D.; Writing—original draft, S.K.; Writing—review and editing, A.P.D; Sofware and Resources, S.K. and A.P.D.; Supervision, S.K.

Competing interests

The authors declare no competing interests.

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.