
The material cannot be used for any other purpose without further permission of the publisher and is for private use only.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.gla.ac.uk/221994/

Deposited on 11 August 2020

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk
Antenna design optimization continues to attract a lot of interest. This is mainly because traditional antenna design methodologies are exhaustive and have no guarantee of yielding successful outcomes due to the complexity of contemporary antennas in terms of topology and performance requirements. Though design automation via optimization complements conventional antenna design approaches, antenna design optimization still presents a number of challenges. The major challenges in antenna design optimization include the efficiency and optimization capability of available methods to address a broad scope of antenna design problems considering the growing stringent specifications of modern antennas. This paper presents a review of the most recent progress in antenna design optimization with a focus on methods which address the challenges of efficiency and optimization capability via machine learning techniques. The methods highlighted in this paper will likely have an impact on the future development of antennas for a multiplicity of applications.

Index Terms—antenna optimization, machine learning, surrogate model-based optimization.

I. INTRODUCTION

Over the last few decades, antennas and their associated systems have evolved rapidly due to unprecedented changes in their geometric and material profiles to meet modern applications such as body-centric communications [1] and multiband operations for 2G/3G/4G/5G [2]. Generally, antennas can be designed by following rules of thumb which often premise on design experience [3]. Due to increasingly stringent specifications and the realization of additional performance requirements, present-day antenna structures are usually topologically and electromagnetically complex with a large number of sensitive design parameters [2]. Even though experience-based rules of thumb can provide a practical guide to antenna designers, they are mostly suitable for simple antenna structures and applying them correctly often yield sub-optimal designs (even for simple antenna structures [4]). Consequently, finding the best designs, which fulfill the desired performance for contemporary antennas, could be very challenging.

To address the above bottleneck, it is a common practice to fine-tune the geometric and/or material parameters of antenna structures for performance improvement. The most popular approaches still revolve around experience-based parameter sweeping of a few parameters at a time [1]. For modern antennas with many sensitive parameters, this process could be time-consuming without any guarantee of successful outcomes - it is often a process of trial and error. Thus, the need for design automation via optimization. To improve the performance of antennas via optimization, local and/or global numerical optimization methods are chiefly employed [4]. Even though numerical optimization is evidently more superior than experience-driven parameter sweeping, there are still some challenges.

Local optimization methods often require a good initial design or starting point (which is typically not available in practice for modern antenna structures) to obtain good results [4]. Global optimization methods on the other hand are more attractive because of their robustness and non-requirement of an initial design, but they often require a very large (sometimes unaffordable) number of electromagnetic (EM) simulations to obtain near-optimum designs [4]. For a thorough characterization of antennas, numerical technique-based full-wave EM simulations are inevitable. Full-wave EM simulations are inherently computationally expensive. A one-time design characterization via an EM simulation does not constitute a problem, but the massive amounts of such EM simulations required by global optimization methods constitutes an unaffordable computational overhead.

To lower the computational overhead of antenna synthesis, machine learning methods are often used to aid numerical optimization methods by integrating them in the optimization kernel a priori and/or a posteriori [5], [6]. Surrogate modelling tends to be the most popular machine learning method used to aid numerical optimization methods for antenna synthesis [6]. In the optimization process, surrogate modelling mainly works by replacing computationally exact function evaluations (i.e., computationally expensive EM simulations) with computationally cheap approximation models. These approximation models are called surrogate models and they are usually constructed using statistical learning techniques [7]. Though a number of surrogate modelling techniques have been shown to be very successful for the machine learning-assisted optimization of EM designs [8], Gaussian process (GP) or kriging [9] tends to be popular in the antenna design domain [10]–[12]. The conjunctive use of surrogate modelling techniques and numerical optimization methods in a single optimization framework is referred to as surrogate-based optimization (SBO) [5], [6].
Several SBO methods employing local optimization methods and/or global optimization methods as their search engines have been proposed for the single-objective, multi-fidelity, multi-objective and process variation-aware or yield optimization of antennas [11], [13]–[16]. As a result, there is a variety of SBO paradigms for the machine learning-assisted optimization of antennas. In this review, state-of-the-art SBO methods for the machine learning-assisted optimization of antennas are discussed in an attempt to establish their use cases. Particularly, these methods are evaluated in terms of generality, efficiency, optimization capability and ease-of-use according to existing literature.

II. OPTIMIZATION OF ANTENNAS

A. Single-objective optimization

Very often antenna optimization problems are modeled as single-objective optimization problems. A typical example is the maximization of the isotropic gain [4]. Typically, a single-objective optimization can be mathematically described as shown in (1)

$$\begin{align*}
\text{minimize} & \quad f(x) \\
\text{s.t.} & \quad g_i(x) \leq 0, \ i = 1, 2, \ldots, k. \\
\end{align*}$$

(1)

where $x$ is the decision variable (typically, a set of parametric values which describe an antenna design), $d$ is the dimension of $x$ and $[a, b]^d$ are the search ranges of the design variable, $f(x)$ is the objective function such as the in-band maximum reflection coefficient to be minimized, $g_i(x)$ are the constraints such as the specifications on the design, and $k$ is the number of constraints. Considering a minimization problem, the optimal solution for (1) will be the value of the value of $x_o \in [a, b]^d$ where there is no other point $x_i \in [a, b]^d$ with $f(x_i) < f(x_o)$.

Typically, antenna optimization problems are constrained optimization problems. This is because antennas specifications are often two or more as described mathematically in (1). For most constrained optimization problems, a weighted sum of the constraints and the objective is used to derive a penalty function which then becomes the single objective function value for the optimization. This approach is called the penalty method [17] and it is the method mostly employed for the aggregation of the multiple requirements of antenna problems.

In recent times, several SBO methods have been proposed for the single-objective and/or constrained optimization of antennas [15], [18], [19]. In general, a critical trade-off between the surrogate model quality and the efficiency of the optimization have been proposed in [11], [12], [23]. This class of SBO methods employing global search and online surrogate models (i.e., surrogate models which are updated continuously throughout the optimization process) is called the surrogate model-assisted differential evolution for antenna optimization (SADEA) family of algorithms. SADEA family of algorithms use the state-of-the-art surrogate model-aware evolutionary search framework [24] for their surrogate model management and they offer three to 20 times speed improvement compared to standard global optimization methods [4]. They are robust and very generic because they have no limitations on the dimension space and no ad-hoc processes are required in their modus operandi.

B. Multi-fidelity optimization

Multi-fidelity optimization of antennas can easily be categorized as a class of antenna optimization efficiency improvement method on its own. Basically, the general idea of multi-fidelity optimization of antennas is to use cheaper and less accurate low-fidelity models to filter out non-promising solutions, and to use expensive but accurate high-fidelity models to propose to address this challenge amongst others in SBO methods are discussed as follows.

To address the well-known wrong convergence issue in standard space-mapping optimization of antennas [20], a local search mechanism (the trust region (TR) approach) where a surrogate model is optimized in a restricted neighbourhood (i.e., a local region) is employed in [21]. For most practical cases, the Jacobian matrix which characterizes the linear model of the TR gradient search is evaluated through finite differentiation at the cost of additional EM simulations per algorithmic iteration [22]. The number of the additional EM simulations corresponds to the dimensionality of the design space; consequently, the Jacobian updates primarily decide the computational overhead of the optimization process. To reduce the computational cost of the TR gradient search in [21], a method which uses an adaptive scheme for sparse Jacobian updates is proposed in [15]. In [15], the Jacobian matrix is only updated in part by considering the relationship between the subsequent design vectors and TR region size.

To enhance the convergence speed, reduce the dimensionality of the search space and improve the starting point of the local search mechanism (TR gradient search in [21]), a multi-stage optimization method using large scale sensitivity analysis and local optimization routines is proposed in [19]. The method in [19] iteratively locates a feasible region in the antenna’s multi-dimensional parameter space through a sequence of constrained optimization runs (considering a goal at a time). The optimization procedure is then initialized by focusing on the subspace of the most influential parameters (selected via a large scale sensitivity analysis) first to improve the starting point of the TR gradient search.

To circumvent the requirement of a starting point or an initial design, while ensuring high efficiency and good convergence speed, a class of SBO methods which offer a good balance between the quality of the surrogate modelling and the efficiency of the optimization have been proposed in [11], [12], [23]. This class of SBO methods employing global search and online surrogate models (i.e., surrogate models which are updated continuously throughout the optimization process) is called the surrogate model-assisted differential evolution for antenna optimization (SADEA) family of algorithms. SADEA family of algorithms use the state-of-the-art surrogate model-aware evolutionary search framework [24] for their surrogate model management and they offer three to 20 times speed improvement compared to standard global optimization methods [4]. They are robust and very generic because they have no limitations on the dimension space and no ad-hoc processes are required in their modus operandi.
perform a search around "promising" solutions obtained by the low-fidelity model. Depending on the optimization framework, these models could be surrogate and/or EM models [25], [26]. To achieve this co-working, very often data-driven surrogate models and EM (i.e., antenna) models of varying accuracies or fidelities are integrated into a single optimization framework to lower the overall computational cost of optimization [27].

The primary challenge with machine learning-assisted multi-fidelity antenna optimization methods has always been how to efficiently ensure the uniqueness of the parameter extraction from the low-fidelity design space to the high-fidelity design space [26]–[28]. In recent times, a number of methods have been proposed to adequately overcome this bottleneck for various use cases [26], [27], [29]. Some of the latest methods are discussed as follows based on their innovations.

To lower the computational overhead of the multi-fidelity optimization of antennas, a conjunctive use of varying-fidelity antenna models with data-driven surrogate models is proposed in [29]. In [29], a surrogate model realized in a constrained domain through a space reduction technique is used. The space reduction technique works by employing the simplex formed by designs from the low-fidelity EM model design space to determine (by way of estimation) the lateral spread of the solution domain. The designs obtained from the low-fidelity design space are then verified/validated using a few high-fidelity EM simulations. Typically, only one of the final designs is required and used.

To improve the convergence speed of the multi-fidelity optimization of antennas, the method proposed in [26] exploits a number of multi-fidelity coarse models with increasing discretization levels (with the discretization level of the last one closest to the fine model) for the iterative construction of a series of local surrogate models by means of polynomial interpolation. In [26], a judgement factor based on the overall degree of similarity between the current surrogate model and the corresponding model is used to provide information for the update of the local region size. The optimized design of the final local surrogate model is assumed to be a good estimation of the optimal design of the high-fidelity model and its accuracy is improved by means of input space mapping performed in its local region.

To reliably handle model discrepancies in the multi-fidelity optimization of antennas, while ensuring high efficiency and good convergence speed, a multi-stage SBO method which features data mining and a local search mechanism is proposed in [27] stemming from the algorithmic framework in [30]. In [27], a one-off coarse model is used to carry out an SADEA-based optimization to generate a pool of data designs at the first stage. The pool of data designs from the first stage are then clustered using an iterative clustering algorithm to form an initial database for the final stage. At the final stage, a one-off fine model is used to carry out an SADEA-based optimization aided by a surrogate-model-assisted local search starting from the initial database from the previous stage. The method proposed in [27] is very generic because there are no limitations on the dimension space and no ad-hoc processes are required for the update of the coarse and/or fine EM models (both are one-off in the entire optimization process).

C. Multi-objective optimization

Mathematically, a typical multi-objective optimization problem can be described according to (2). A typical case of an antenna design problem handled as a multi-objective optimization problem is the minimization of the reflection coefficient values within multiple bands and the minimization of the antenna structure for a planar antenna [31].

$$\text{minimize } \{ f_1(x), f_2(x), \ldots, f_m(x) \}$$

$$x \in [a, b]^d.$$  \hspace{1cm} (2)

where \{f_1(x), f_2(x), \ldots, f_m(x)\} are the optimization objectives and \(m\) is the total number of objectives.

Multi-objective optimization techniques can be broadly classified into a priori and a posteriori methods according to the decision-making processes [32]. In contrast to a priori methods, a posteriori methods do not require prior preference information from the decision maker [32]. Instead, they produce a number of well representative optimal trade-off candidate solutions for a decision maker to check on a Pareto front (PF) - an image of Pareto optimal solutions (called the Pareto set) in the objective space [32]. A Pareto optimal solution is a candidate solution that obtains the best trade-off.

The primary challenge associated with the design and optimization of antennas using conventional multi-objective optimization methods is the overly large (often impractical) amount of expensive EM simulations required for the completion of the optimization process [33]. Additionally, after a successful run, the set of alternative design solutions generated by the Pareto front is often redundant when the designer preferences are vivid and only one final design solution is selected and used [34]. To overcome this challenge, several machine learning-assisted multi-objective optimization methods have been proposed in recent times for the synthesis of antennas [31], [34]–[36]. Some of these methods are discussed as follows based on their innovations.

To lower the computational overhead of the multi-objective optimization of antennas via off-line surrogate modelling (i.e., a one-off surrogate model construction in the optimization process), a sparsely connected backward propagation neural network is employed in [31]. The surrogate model in [31] is constructed only after the network parameters are adaptively tuned using hybrid real-binary particle swarm optimization algorithm [37] to promote global optimization capability, and a time-varying transfer function is set to lower the tendency of easily getting trapped into a local optimal and to improve convergence speed. It is then used to replace computationally expensive full-wave EM simulations in a standard multi-objective optimization framework using multi-objective evolutionary algorithms (MOEAs) to generate the PF.

To improve the convergence speed and to promote a better spread of design solutions for the multi-objective optimization of antennas via local search mechanisms, a non-dominated
and local search-assisted multi-objective optimization method is proposed in [35]. The method in [35] differs from the conventional multi-objective optimization methods via the adoption of a local search method to generate an improved population in the optimization process. To obtain the population for consecutive iterations, the traditional non-dominated sorting method [38] and the farthest-candidate method [38] are used. The population is updated by the local search for improved convergence speed by replacing current solutions with neighbouring solutions using the replacement strategy in [38]. As a result, boundary solutions (i.e., solutions with minimum and maximum fitness values) are prioritized for selection to ensure an even distribution of the Pareto-optimal solutions on the PF.

To lower the computational overhead and improve the convergence speed of the multi-objective optimization of antennas via surrogate modelling and variable-fidelity EM models, a robust methodology is presented in [34], [36]. The essential component of this method is the sequential domain patching of the design space. This is carried out by firstly generating extreme Pareto-optimal designs (two designs) realized by using an auxiliary low-fidelity model of the antenna conjunctively with a GP interpolation model in a standard multi-objective optimization framework using MOEA [39]. The patching process is then implemented as a stencil-based search targeted at linking the extreme Pareto-optimal designs through an iterative generation of subregions within the design space to have an initial approximation of the Pareto set. Since the initial pareto set is realized at the level of a low-fidelity model, the final Pareto set is generated by refining the selected coarse designs using the output space mapping (OSM) procedure [40]. The OSM correction mainly ensures that at the beginning of each iteration the fitness of the refined models correspond to the fitness of the second-order polynomial approximation of the low-fidelity models [41].

D. Process variation-aware or yield-driven optimization

A majority of optimization-driven antenna design methods do not account for the likely discrepancies that may exist between the nominal antenna structure (often the optimal design obtained after optimization) and the actual (fabricated) antenna structure. For a robust design of antennas and to ensure a full design closure, statistical analysis is required for the quantification of the fabricated antenna deviations from its nominal design values. This procedure is referred to as process variation-aware or yield-driven design and it is aimed at maximizing the probability that a fabricated prototype will meet the performance specifications within the range of assumed statistical deviations from its nominal design [16], [42].

To carry out statistical modelling for yield evaluation, the traditional Monte Carlo (MC) method is the most common and generic method [42]. However, the number of trials required to have an accurate MC yield estimate is often large [42] leading to a slow convergence speed. Due to the slow convergence speed of MC and the dimensionality of contemporary antennas, a very large number of samples (i.e., many hundreds to thousands of computationally expensive full-wave EM simulations and their responses) will be required for the yield analysis of antenna design problems [16]. This is computationally prohibitive; consequently, conventional statistical methods such as MC are not very popular for antenna design.

Though a number of techniques have been proposed for the expedited statistical analysis of EM models and microwave structures, these techniques have mostly been applied to the design of integrated circuits, multiconductors, microstrip and filters [43], [44]. With a focus on yield-driven design of antennas, a low-cost statistical analysis and yield optimization of antennas using auxiliary response surface approximation is proposed in [16]. In [16], a fast GP surrogate model is constructed within the vicinity of the nominal design (i.e., the optimal design obtained after an optimization run) and a low-cost MC analysis is carried out using the fast surrogate model. The yield estimate (i.e., the likelihood that the performance specifications stipulated for the antenna are satisfied within the range of assumed statistical deviations with respect to the fabrication and/or material tolerances) is then maximized using sequential approximate optimization by employing the local interpolation surrogate (statistically optimized) and rebuilt in a new domain.

As a way of lowering the computational overhead of yield-driven design of antennas, a performance-based nested surrogate modelling method using a two-level kriging is proposed in [45]. In [45], two varying surrogate models are considered - a first-level surrogate model which is used to identify the "promising region" of the parameter space by mapping the objective space into the geometry parameter space of the antenna to have the surrogate domain and the final (second level) surrogate model is built over the domain to connote the antenna responses. Since the dimension of design space is not very high, the mapping or surjective transformation adopted for parameter extraction makes the allocation of uniform training data points very straightforward in [45]. The same mapping subsequently allows for a suitable optimization of the surrogate which can then be directly used in a yield-driven optimization framework.

III. Conclusion

In this paper, present-day machine learning-assisted antenna optimization methods are reviewed. These methods have been discussed under the general forms or approaches to antenna optimization. Methods which are suitable for high-dimensional parameter spaces and multiple specifications without the requirement of initial designs and/or ad-hoc processes are recommended for use due to their generality, robustness and optimization capability to handle a variety of modern antenna design problem cases.

REFERENCES

et al.


