

## New trends in optimization in electromagnetics

**Abstract.** This paper reviews recent advances in optimisation of electromagnetic problems. CAD assisted optimal design often necessitates repetitive usage of numerically intensive field computation where cost-effective approaches are required. Modern algorithms increasingly rely on surrogate modelling, kriging-assisted methods, pareto-optimality and design sensitivity.

**Streszczenie.** Nowoczesne metody projektowania oparte na wspomaganium komputerowym z wykorzystaniem numerycznego obliczenia pól elektromagnetycznych wymagają skutecznych, a przede wszystkim szybkich i efektywnych metod optymalizacyjnych. Ostatnie lata przyniosły eksplozję nowych podejść i algorytmów. Artykuł jest próbą uogólnienia efektów ostatnich badań – w tym również dorobku autora – ze szczególnym uwzględnieniem najnowszych osiągnięć i spodziewanych nowych rozwiązań. (**Nowe tendencje optymalizacji w elektrodynamice technicznej.**)

**Keywords:** optimization, kriging, surrogate modelling, pareto-optimality, design sensitivity, electromagnetics, finite elements.

**Słowa kluczowe:** optymalizacja, elektromagnetyzm, elektrodynamika techniczna, komputerowe wspomaganie projektowania.

### Introduction

Electromagnetic solutions may be categorised as *direct problems*, in which the design is specified and the effects are investigated, and *inverse problems*, in which a design is sought producing a desired effect. If the effect is measured, we have an *identification* problem, if the effect is assumed, a *synthesis* problem is posed, to which a solution may not exist. With the advent of CAD and availability of powerful software for electromagnetic field analysis, design offices increasingly rely on powerful commercial packages offering high accuracy of solutions [1]. Static, quasi-static, transient, 2D or 3D, non-linear, anisotropic field systems can be solved, which are often connected to the driving circuitry and/or coupled to the associated mechanical or thermal fields. The optimal design problem expresses the desired effects as objectives to be minimized or maximized, thus posing a synthesis problem. In general, this can only be achieved by solving the direct problem multiple times, which is inherently more time-consuming than the direct problem.

Optimal electromagnetic design has recently become an area of very vigorous research involving mathematics, numerical analysis, software development and engineering design. A peculiarity of electromagnetic design, as argued above, is that the solutions are 'expensive' and thus 'cost-effective algorithms' have to be used. This paper is an attempt to provide a brief review of the most recent developments and highlights some new approaches.

### 'No free lunch' theorem

The 'no free lunch' (NFL) theorem [2] prohibits the existence of an algorithm which would outperform all other optimization algorithms, when averaged over all possible optimization problems. It argues that, averaged over all types of problems, every algorithm performs the same. However, design engineers are only interested in a subset of problems, thus – consistent with the NFL theorem – it is possible to identify a set of algorithms which outperform others over a particular domain of interest [3].

The spectrum of methods for solving optimization problems is vast. They may be categorized as single- or multi-objective, global or local, deterministic or stochastic, greedy or cost-effective, etc. Electromagnetic design does not fall neatly into one particular category, the objective function landscapes may be simple or complex, a single function evaluation may take anything from a few seconds to several days, problems are constrained or unconstrained, etc. Problems with computationally expensive objective functions need particular attention as they present the most common challenge in modern electromagnetic design; such problems are explored in this review.

Several methods exist for achieving cost-effectiveness in multi-objective optimization, including small population

genetic algorithms, hybrid algorithms, reduction of design variables and fitness inheritance. In this paper we focus on surrogate modelling and kriging-assisted methods [4, 5].

### Single-objective optimization

A single-objective optimization problem (SOOP) may be formulated as: minimize  $f(x)$  subject to  $g_j(x) \geq 0$  ( $j=1,2,\dots,J$ );  $h_k(x)=0$  ( $k=1,2,\dots,K$ );  $x_{iL} \leq x_i \leq x_{iU}$  ( $i=1,2,\dots,M$ ). The problem is unconstrained if  $J=K=0$  and each variable has bounds (some may vary discretely [6]). With the  $J$  inequality and  $K$  equality constraints, a *feasible region*  $S$  exists as a subset of the *decision variable space*, giving rise to the *feasible objective space*  $Z$  as a subset of the *objective space*. The aim of single-objective optimisation is to locate the design vector(s) in  $S$  which give the global minimum value  $f_{min}$ .

There are features of an objective function which make locating a global minimum a particular difficulty: degree of modality, size of basins of attraction of local minima, size of improving regions, degree of randomness in the positions of the minima and dimension of the search space [7]. The last one, known as the 'curse of dimensionality', is best approached through reduction of design variables.

Performance criteria used for comparison of algorithms include: best function value found, CPU time, number of function evaluations, accuracy, success rate and stopping criteria [8]. Some less obvious, but equally important, refer, for example, to the number of necessary tuning parameters, such as a mutation rate in genetic algorithms or cooling schemes in simulated annealing.

### Multi-objective optimization

In a multi-objective optimization problem (MOOP) we seek a minimum of  $f_m(x)$  ( $m=1,2,\dots,M$ ) subject to similar constraints as for SOOP. The aim is to try to simultaneously minimize the  $M$  different objectives  $f_m$ , giving rise to an *ideal objective vector* and an *utopian objective vector*, something which in general is not possible, leading to the necessity of defining exactly what constitutes a solution to the MOOP. The  $M$  objectives are almost always in competition making the concept of the ideal objective useless. *Nash Equilibrium* [9] is one possible answer, but the definition usually adopted is that of *Pareto-Optimality* [10].

### Pareto-optimality

Central to this idea is the notion of dominance and a non-dominated set. For any two solutions  $\mathbf{x}_1$  and  $\mathbf{x}_2 \in S$ ,  $\mathbf{x}_1$  is said to dominate  $\mathbf{x}_2$  if  $f_i(\mathbf{x}_1) \leq f_i(\mathbf{x}_2)$  for all  $i=1,2,\dots,M$  and  $f_i(\mathbf{x}_1) < f_i(\mathbf{x}_2)$  for at least one  $i \in \{1,2,\dots,M\}$ , so it is *better* in at least one of the objectives and *no worse* in all the others. We say that  $\mathbf{x}_1$  is non-dominated by  $\mathbf{x}_2$ . Among a set of solutions  $P \in S$ , the *non-dominated set*  $P'$  are those that are

not dominated by any member of the set  $P$ . The non-dominated set of the entire feasible space  $S$  is the *global Pareto-optimal set*, and its image in the feasible objective space is the *global Pareto-optimal front* (POF). Local sets and fronts can also be defined, as well as weak and proper Pareto-optimality [10]. A simple demonstration of these concepts is provided in Fig. 1 for two objectives.

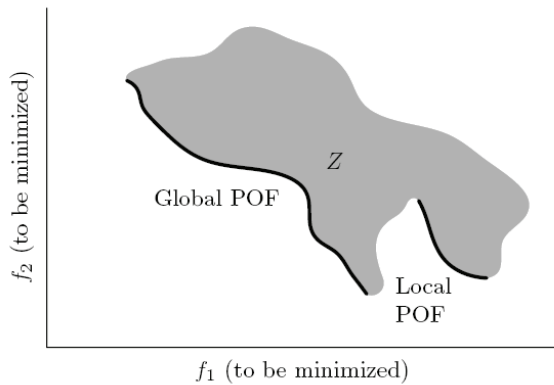


Fig. 1. Global and Local Pareto-optimal Fronts

In the absence of other criteria, all POF solutions are equally important and as many should be found as possible. The simultaneous two goals are to find a set of solutions as close as possible to the Pareto-optimal front and a set of solutions as diverse as possible. The former is simply stating that the solutions found are to be as close to the true optimal solutions as possible. The latter is specific to multi-objective optimization and is important because a diverse set of solutions assures us that no single objective is being favoured. Achieving a good balance between convergence to and diversity along the Pareto-optimal front is important to all multi-objective optimization algorithms.

The difficulties with locating the Pareto-optimal front are usually due to multi-modality, isolated optimum, convexity, discontinuity and non-uniformity [10]. Algorithms are verified and compared using special test functions with parameters allowing their difficulty to be fine-tuned [11]. At the same time new heuristic algorithms are developed inspired by all aspects of nature, like particle swarms [12] or artificial immune systems [13].

### Surrogate modelling

The simplest to visualise and construct are polynomial models; however, they have several inherent drawbacks: low-order polynomials are unable to model complex functions, whilst high-order ones often lead to ill-conditioned matrices; they can only be constructed after a certain number of observations, furthermore, it is only for this minimum number of observations that they are (necessarily) interpolating. In order for a surrogate model to be interpolating, it is necessary to use some additional *basis functions*, each centred around one of the  $n$  sampled points. Then the prediction made by the model may be written as

$$(1) \quad \hat{y}(x) = \sum_{k=1}^m \beta_k f_k(\mathbf{x}) + \sum_{j=1}^n b_j \phi(\mathbf{x} - \mathbf{x}_j)$$

where values of  $\beta_k$  are determined by least squares fitting, and there are several choices for  $\phi$ , the most popular being:  $\phi(r)=r$  (linear),  $r^3$  (cubic),  $(r^2+\gamma^2)^{1/2}$  (multiquadric),  $r^2 \log r$

(thin plate spline) and  $\exp(-\sum_{l=1}^d \theta_l |r_l|^{p_l})$ , where  $\theta_l \geq 0$ ,

$p \in (0,2]$  (kriging). The first four belong to a class known as *radial basis functions*, and various comparisons may be found in [14, 15]. The surrogate model which stands out, due to its solid statistical foundations, is kriging.

### Kriging

First introduced in the 1960s, kriging has formed the basis of geostatistics, but has also found applications in fields of research remote from geology. A version for predicting computer experiments with deterministic output, known as Design and Analysis of Computer Experiments (DACE), was developed in the late 1980s [16].

The second term in (1) may be treated as a functional departure from the polynomial terms. The theory of kriging models this functional departure as a stochastic process, or more specifically, as a Gaussian process characterized by its mean and its covariance function. In the DACE models this process is assumed to have mean zero and covariance  $\sigma^2 R(\mathbf{x}^i, \mathbf{x}^j)$  where  $\sigma^2$  is the variance of the process and  $R$  is a correlation function. Many different correlation functions

exist, the one used in DACE is  $\prod_{k=1}^n \exp(-\theta_k |x_k^i - x_k^j|^{p_k})$

where  $\theta_k$  determines how fast the correlation between design vectors drops away and  $p_k$  determines the smoothness of the function in the  $k^{th}$  coordinate direction. It should be noted that, as  $R(\mathbf{x}, \mathbf{x})=1$ , the DACE predictor is an interpolator. Finally, a  $n \times n$  correlation matrix  $\mathbf{R}$  may be defined which represents the correlation between each pair of evaluated design vectors and whose  $i$ - $j^{th}$  entry is  $R(x^i, x^j)$ . Several possibilities then exist, one leads to the so-called 'concentrated log-likelihood function'

$$(2) \quad -\frac{n}{2} \log(\hat{\sigma}^2) - \frac{1}{2} \log(|\mathbf{R}|)$$

which is maximized over  $\theta$  and  $\mathbf{p}$ . An expression for the mean squared error in the prediction may then found and having an estimate of such error can be very beneficial in deciding where to evaluate in the design variable space.

### Kriging-assisted SOOPs

Jones' taxonomy [4] uses two criteria: the kind of surrogate model used and the method of selecting search points. The former is further subdivided into those which interpolate the observed points and those which do not; the latter into two-stage and one-stage varieties. In two-stage algorithms first the surrogate model is fitted to the observed points and then a utility function is constructed to determine the next search point. In one-stage methods a design vector is determined which would yield the most credible response surface. Almost all existing algorithms are two-stage; however one-stage algorithms have been successfully constructed using both kriging [4] and radial basis function surrogate models [17].

An alternative taxonomy for kriging assisted methods classifies the methods according to the number of design vectors to be evaluated in each iteration and how 'tunable' each method is to the balance between exploration and exploitation. Non-tunable utility functions have no parameters to be defined in order to select a point to evaluate. One approach, which has been used by practitioners for many years, interprets the surrogate model as an accurate representation of the true function and evaluates the minimum of the surface. This often leads to a false minimum – a point which is a minimum of the surrogate model but not of the true function. Moreover, if the false minimum is actually an observed design vector, that is, one which has been used to construct the kriging

model, then the algorithm will choose to evaluate a design vector which has already been evaluated. Using the same solution twice in the construction of a surrogate model leads to the inversion of a singular matrix and the approach fails. If the minimum of the kriging model is close to an observed point, the resulting change in the kriging model during reconstruction in the subsequent iteration may be insignificant, thus many iterations are wasted sampling around a false optimum. This 'strawman approach' (as dubbed by Jones in [4]) does not take into account how unexplored a region of design variable space is when deciding where to evaluate: it simply exploits the most promising region of design variable space.

The Efficient Global Optimisation (EGO) algorithm [18] uses the concept of *expected improvement*, which may be viewed as a fixed compromise between exploration and exploitation. It is currently acknowledged as one of the best performing methods for SOOP.

Non-target based tunable utility functions allow the balance between exploration and exploitation to be controlled by a parameter, which is not an estimate of the global minimum. Two simple utility functions exist which are of this type (both based on the expected improvement utility function): the Generalized Expected Improvement (GEI), and the Weighted Expected Improvement (WEI) [19].

Target based tunable utility functions are characterized by allowing the balance between exploration and exploitation to be controlled by a parameter which is an estimate of the global minimum, while the concept of improvement is also used. Two methods have been recognised as most promising: the probability of improvement with multiple targets, and the one-stage credibility of hypotheses method (with either single or multiple targets each iteration) [4].

### Kriging-assisted MOOPs

Multi-objective methods using surrogate models may be divided into scalarizing and non-scalarizing. Scalarizing methods combine the multiple objectives of the MOOP using some function, and then use one of the methods for single-objective optimization. By varying the parameters which control how the multiple objectives are combined, an approximation to the Pareto-optimal front can be built up. The main purpose of a scalarizing function is to combine the multiple objectives of a MOOP in such a way that the contours of the resulting function are able to capture every point on the Pareto-optimal front. The most popular methods here include:  $\varepsilon$ -constraint (which considers only one of the objectives for minimization, whilst treating the other objectives as constraints to be satisfied), weighting (each objective is simply associated with a weighting coefficient), weighted and augmented weighted  $L_p$ , and weighted and modified weighted Tchebycheff metric (using different metrics to define the distance of a solution from the Utopian point) [10]. After transforming a MOOP to a SOOP, any method may then be used to solve the resulting SOOP. By varying the weights used in the scalarizing method, approximations of the Pareto-optimal front can then be built up. This gives rise to a huge number of possible cost-effective MOOP algorithms. Surprisingly few (other than the simplest, such as the weighting method combined with the 'strawman' approach) have been pursued in the literature, with two notable exceptions: EGO [18] and ParEGO [20].

Non-scalarizing methods consider each objective function individually, the simplest of which is just to evaluate the Pareto points predicted by the multiple surrogate models (equivalent to the strawman approach in SOOP); a good example is an optimisation of a switched reluctance motor reported in [21]. Many 'greedy' MOOP algorithms

exist which are non-scalarizing, in particular Multi-Objective Evolutionary Algorithms; however, non-scalarizing methods for cost-effective multi-objective optimization have only appeared fairly recently. In [5] and [22] the notion of 'equivalent' and 'dominating' designs is used to establish the *probability of improvement* utility function; unfortunately, this has a certain weakness of not favouring large improvements. Instead the *expected improvement* criteria can be employed for selecting design vectors – this leads to a more aggressive search and thus higher efficiency.

### Some practical issues

Surrogate models cannot be used to select every design vector to evaluate during the search: a certain minimum number of vectors need to be sampled before a kriging model can even be constructed. This initial set is called an *experimental design* and the theory behind selecting suitable points is known as Design of Experiments [23]. Classical experimental designs involved measurements, and had to account for features such as randomness and non-repeatability. Experimental designs for computer experiments do not have to take into account such features and are commonly referred to as modern experimental designs. The two most common modern experimental designs are the Latin Hypercube [18] and the Hammersley sequence [24].

Evaluating objective functions and constraints using CAD software – unlike validating algorithms using test functions – involves building a model, generating a mesh, creating a database and then running a solver. Any one of these stages is prone to failure, but this should not subsequently mean that the algorithm has failed. Instead mechanisms should be built into the algorithm to deal with such failed iterations. Although very important, this topic is almost entirely overlooked in the literature. Recently, however, methods were proposed in [25] to deal with missing data, which involved imputing values for failed iterations which consequently penalized that region of design variable space.

Constraints are a crucial part of any problem, as they dictate whether a solution is feasible or not; in many cases, finding a solution which is infeasible (no matter by how little), is as bad as finding no solution at all. Inequality constraints are usually handled using probability methods, penalty functions, expected violation technique or constrained utility function approach. Equality constraints are normally approximated by transforming the equality constraint into two inequality constraints or transforming the equality constraint into an objective.

### Novel algorithms

As already indicated, optimisation in electromagnetics appears to be having its 'second honeymoon' and as an active area of research could almost be considered as a 'born again' subject. A similar review of the state of the art published only three years ago [26] seems to be of historical value now in the context of the most recent developments. Several papers are in the pipeline and new ideas emerge faster than they can be published.

As an example, let us consider the credibility of hypothesis utility function which was recognized as a promising utility function for single-objective optimization [4]. Its most attractive feature is that it can perform well with deceiving experimental designs, but at a price that evaluating the credibility of a hypothesis itself becomes computationally expensive when the number of sampled points grows large. Expected improvement and its generalizations, on the other hand, remain computationally cheap to evaluate throughout the optimization process. Thus it should make sense to switch to using the expected

improvement utility function when the computational cost of using the credibility of hypothesis utility function becomes large, which is the main idea behind an algorithm which has only very recently appeared in press [27].

The proposed algorithm for locating the global minimum in  $d$ -dimensional design variable space consists of three steps: initialization, one-stage experimental design, and two-stage optimization search. The only purpose of the initialization is to sample enough points to allow a nontrivial kriging model to be constructed. The space-filling Hammersley Sequence experimental design, of size  $4d$ , is used to select the points. The experimental design size of  $4d$  is much smaller than is normally used ( $10d$  is suggested in [18] for example), as the philosophy of this algorithm is to use information about objective function space to search for the minimum at the earliest possible opportunity. The one-stage experimental design stage that follows is to strategically choose where to sample next. Let the minimum objective function value of the  $4d$  sampled points be  $f_{\min}$ , and the maximum  $f_{\max}$ . A design vector  $\mathbf{x}^*$  is then hypothesized to exist in design variable space which has an objective function value given by  $f^* = s_{\min} - \alpha(f_{\max} - f_{\min})$  where  $s(\mathbf{x})$  is the *root mean squared error* in the prediction, and the value of  $\alpha$  is varied using a cyclic scheme. The design vector chosen for evaluation at each iteration is the one which maximizes the credibility of the hypothesis of it having objective function value  $f^*$ . This step is repeated for  $6d$  iterations until  $10d$  points in total have been evaluated. In the final two-stage optimisation search, a kriging model is constructed using the  $10d$  sampled points. The *generalized expected improvement* (GEI) utility function is then used (with a cooling scheme) to select points for evaluation. Finally the *weighted expected improvement* (WEI) utility function is used to select points for evaluation: a cyclic scheme is used for varying the weighting parameter. Using WEI to finish the search allows more exploitative iterations to be used than is possible with the GEI utility function.

The algorithm was first tested on a "Humps" test function, taken from Matlab [28]:

$$(3) \quad f(x) = 6 - \frac{1}{(x-0.3)^2 + 0.01} - \frac{1}{(x-0.9)^2 + 0.04}$$

where  $x \in [-5, 5]$ . The number of evaluations taken to locate the minimum within 1% tolerance was 8, which was during step 2 of the algorithm. An interesting observation was that the kriging prediction when the optimum was found was not very accurate; however this should be of no concern. The purpose of optimization is simply to locate the optimal point, not to accurately predict the function being optimized as well. This is related to Vapnik's principle from machine learning [29]: "When solving a problem of interest, do not solve a more general problem as an intermediate step. Try to get the answer that you really need but not a more general one." During optimization the aim is to locate the minimum of an unknown function, while we are not concerned with the more general problem of approximating the unknown function as accurately as possible. If the kriging approximation is not very accurate, this is not an issue: the optimum point has been located regardless.

In this test case the optimum has been found at the 8th evaluation, and so the algorithm can terminate. However, in general it is not known if the optimum has been found, and so the algorithm will proceed to step 3. This happened in the next example: the optimal design of an electron gun.

The voltage on, and position of, the focus electrode of an electron gun was varied so as to focus the beam of electrons on the centre of the anode as much as possible.

Thus, denoting the voltage on the focus electrode by  $V$  and its perpendicular distance from the emitting surface by  $d$ , the objective function was:

$$(4) \quad f(V, d) = \int_{\text{anode}} J(r) r^2 dS$$

with  $V \in [0, 1000]$  and  $d \in [5, 10]$ , where  $r$  is the radial distance from the centre of the anode surface,  $J(r)$  is the current density at  $r$ , and the integral is taken over the surface of the anode. Each analysis was carried out using commercial software OPERA, with the space charge solver SCALA. A random search of 100 iterations was also carried out for comparison and the best solution obtained had a value of  $f=0.1493$ , whereas using the new algorithm and a stopping criterion of  $30d$  ( $=60$ ) evaluations in total, the final solution – found during the third step – had a better objective function value of  $f=0.0867$ . The algorithm was found to outperform a random search of 100 iterations in just 14 iterations. The configuration of the final design is shown in Fig. 2.

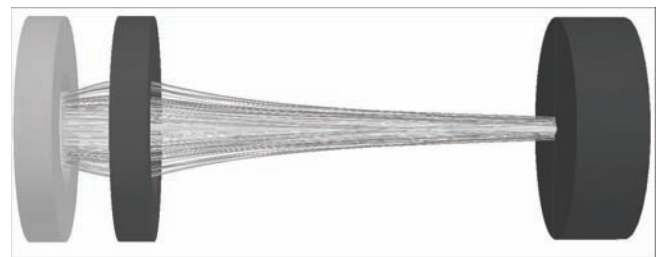


Fig. 2. Final configuration of the optimized electron gun [27]

For kriging-assisted multi-objective optimisation, a generalized ParEGO algorithm, utilizing *generalized expected improvement*, *weighted expected improvement* and *probability of improvement*, is currently under development and will be reported very soon. The dynamic weight vector used is not changed at each iteration; instead it changes in cycles, in order to build up an approximation of the Pareto-optimal front. Two further algorithms are also being derived, a scalarized one-stage algorithm and a non-scalarizing algorithm based on probability of specific improvement.

### Sensitivity analysis

Special consideration needs to be given to a class of methods based on sensitivity analysis. Such approaches offer the advantage of having computation times independent of the number of design variables, thus making them particularly useful in topology optimisation, where the parameterization enables all feasible shapes of electromagnetic devices to be explored. These algorithms are still at early stages of development and not versatile enough to be considered as competitive against surrogate modelling, but their performance is very impressive.

From the point of view of accuracy and time-efficiency in finding the optimum solution in design space, the design sensitivity analysis (DSA) appears to be very competitive compared with other optimization methods. Depending on the technique used to compute the derivative of an objective function, the DSA can be categorized as the discrete DSA (DDSA) or the continuum DSA (CDSA). The former obtains gradient information from direct differentiation of the discretized algebraic system matrix, whereas the latter uses an analytically derived sensitivity formula by exploiting the material derivative-adjoint variable (MDAV) method. From the practical point of view, as a general-purpose approach, the DDSA has a critical drawback as it requires some amendments to the software

source code to perform sensitivity calculation. On the other hand, the CDSA can be incorporated into existing EM software packages without the need to modify the source code. In fact the approach is independent of the actual method used and is adaptable to finite elements, boundary elements, finite differences, etc. For example, when dealing with the optimization of magnetostatic devices, the objectives can be classified into the following three categories (see Fig. 3(a)):

1. Control of the global quantities (i.e. energy, force or inductance) connected with the magnetic vector potential,  $\mathbf{A}$ , inside the region of interest  $\Omega_f$ .
2. Shaping of the local quantity distribution (i.e. magnetic flux density or field intensity) inside the region  $\Omega_g$ .
3. Adjustment of the surface field distribution (i.e. magnetic force density or tensile stress) on  $\gamma_h$ , part of the interface  $\gamma$ .

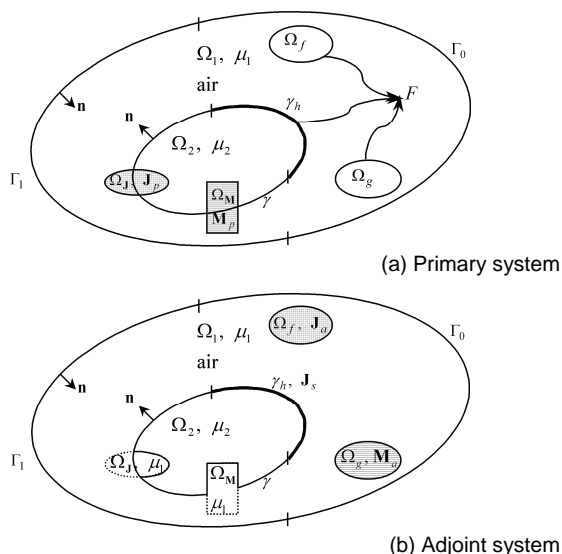


Fig. 3. Dual system of the CDSA [30]

The sensitivity formula can then be derived for both the *primary system* and the *adjoint system* (Fig. 3(b)) [30, 31]. By exploring the analogy between the two formulations the physical interpretation of various aspects of the adjoint system may be established. In particular, the geometric and material properties are found to be the same as those of the primary system and sources may be recognised as electric current or permanent magnet. In other words, the adjoint system – despite having been introduced as a mathematical derivation – is physically well based. Thus both the primary and the adjoint system can be solved using a standard EM package. Moreover, in cases where the objective function is associated with the system energy, the adjoint system does not actually need to be calculated [30], making the optimisation process even more straightforward.

To demonstrate the technique we will refer to a case in *topology optimisation* (TO). In [31] a very fast and efficient TO algorithm for optimizing source distribution in linear magnetostatic problems was put forward. A unified sensitivity formula has been derived as

$$(5) \quad \frac{d\bar{F}}{d\mathbf{p}} = \int_{\Omega} \left[ \left(-\frac{\partial v}{\partial \mathbf{p}}\right) \text{curl} \mathbf{A} \cdot \text{curl} \boldsymbol{\lambda} + \frac{\partial \mathbf{J}}{\partial \mathbf{p}} \cdot \boldsymbol{\lambda} + \frac{\partial \mathbf{M}}{\partial \mathbf{p}} \cdot \text{curl} \boldsymbol{\lambda} \right] d\Omega$$

where the three integrands are sensitivity coefficients with respect to system parameters  $\mu$ ,  $\mathbf{J}$  and  $\mathbf{M}$ , respectively,  $v$  is the reluctivity,  $\boldsymbol{\lambda}$  means the Lagrange multiplier vector interpreted as the adjoint variable and  $\mathbf{p}$  is the system

parameter vector. After solving the dual formulation, consisting of the primary and the adjoint systems, the design sensitivity is thus easily obtained. However, a distinction should be made between finding the Optimized Material Distribution (OMD) and the Optimized Source Distribution (OSD) in electromagnetic systems. In the case of OMD, the permeability is still present in the adjoint system, as well as in the primary one, and thus the fields of both systems are strongly distorted. In this case, the magnitude of the objective function that indicates how far the current design is from the optimum is not exactly reflected in the design sensitivity. Thus an abrupt change of the permeability is liable to result in OMD being trapped in local minima. It is therefore essential that the material is forced to vary gradually from a void to solid state.

On the other hand, the current density and permanent magnetization of OSD are replaced in the adjoint system by air regions. Thus the adjoint field distribution depends not on the relative value of distributed materials but on the magnitude and location of the objective function evaluated. This means that OSD itself, unlike OMD, does not have lots of local minima in the design space. Moreover, even if there are abrupt changes of material values assigned to design cells, OSD can easily converge to an optimum solution. These properties form the basis of a novel material updating algorithm for OSD, which allows each design cell to have only one state, that of a void or a solid [31].

Fig. 4 shows a quarter of an axi-symmetric model of a permanent magnet assembly for an MRI device. The design goal is to find an optimized distribution of shimming magnets over the pole piece surface to obtain homogenous field distribution in a spherical volume in the centre. The shimming magnets have known residual flux density and thickness and the domain is subdivided into 120 separate regions. The objective function is defined as

$$(6) \quad F = \sum_{i=1}^{45} (H_{zi} - H_{zo})^2, \quad \mathbf{M}(x, y) = \mathbf{M}_s(P)p^n$$

where  $H_{zi}$  is the  $z$ -component of the magnetic field intensity computed over the objective regions and  $H_{zo}$  is the desired value.  $\mathbf{M}_s(P)$  in each cell is forced to be  $\pm \mathbf{M}_s$  according to the sign of the accumulated design sensitivity.

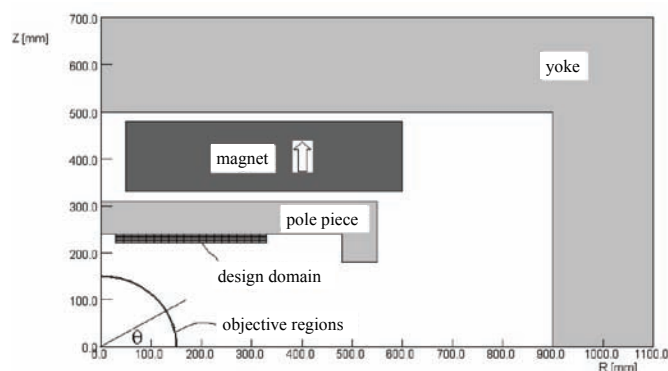


Fig. 4. A quarter model of a permanent-magnet type MRI [31]

Changes to the distribution of shimming magnets during the optimization process are shown in Fig. 5, while Fig. 6 compares the  $z$ -component of magnetic field over the surface of the central spherical volume before and after optimization. Despite the huge number of possibilities in terms of the shimming magnets distribution, the convergence was achieved in about 10 iterations, while the uniformity of the field has been improved by nearly an order of magnitude compared with the initial design [31].

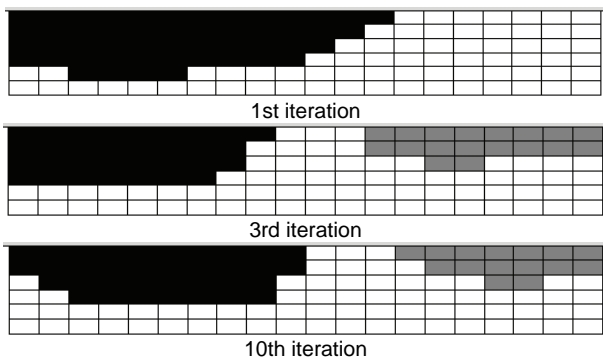


Fig. 5. Shimming magnet distribution (Black cell:  $+M_s$ , Grey:  $-M_s$ )

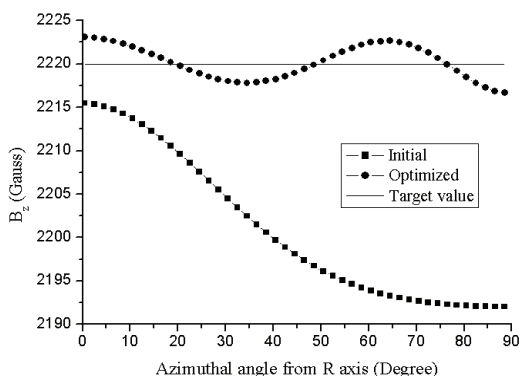


Fig. 6. Field distributions before and after optimization

### Concluding remarks

Software for electromagnetic field computation is now a mature tool commonly employed in design offices across all sections of industry, mainly for analysis and performance prediction (virtual prototyping). Optimal design calls for new cost-effective algorithms to reduce the number of necessary function evaluations and thus avoid excessive computation times. Optimisation in electromagnetics has recently become fashionable again and an area of very active research, as demonstrated by a vast number of new publications and even existence of dedicated conferences [32]. This review reflects on the state of the art and highlights new trends. Surrogate modelling, kriging assisted methods and design sensitivity seem to dominate the current way of thinking of researchers around the world.

### REFERENCES

- [1] Neittaanmaki P., Rudnicki M., Savini A., Inverse Problems and Optimal Design in Electricity and Magnetism, *Oxford University Press* 1996
- [2] Wolpert D. H., Macready W. G., No free lunch theorems for optimization, *IEEE Trans. Evol. Comp.*, 1 (1997), No. 1, 67–82
- [3] Baritompä W. P., Dur M., Hendrix E. M. T., Noakes L., Pullan W. J., Wood G. R., Matching stochastic algorithms to objective function landscapes, *Journal of Global Optimization*, 31 (2005), No. 4, 579–598
- [4] Jones D. R., A taxonomy of global optimization methods based on response surfaces, *Journal of Global Optimization*, 21 (2001), 345–374
- [5] Keane A. J., Statistical improvement criteria for use in multiobjective design optimization, *AIAA Journal*, 44 (2006), No. (4), 879–891
- [6] Di Barba P., Forghani B., Lowther D. A., Discrete-valued design optimisation of a multiple-coil inductor for uniform surface heating, *COMPEL*, 24 (2005), 271–280
- [7] Locatelli M., Wood G. R., Objective function features providing barriers to rapid global optimization. *Journal of Global Optimization*, 31 (2005), 549–565
- [8] Khompatraporn C., Pinter J. D., Zabinsky B., Comparative assessment of algorithms and software for global optimization, *Journal of Global Optimization*, 31 (2005), 613–633
- [9] Di Barba P., Nash equilibrium and pareto front for the optimal shape design in electromechanics, In *Proceedings of 6th*

- International Conference on Computational Electromagnetics*, (2006), 123–124
- [10] Miettinen K. M., *Nonlinear Multiobjective Optimization*, *Kluwer Academic Publishers* 1999
- [11] Huband S., Hingston P., Barone L., While L., A review of multiobjective test problems and a scalable test problem toolkit, *IEEE Trans. Evol. Comp.*, 10 (2006), No 5, 477–506
- [12] Reddy M. J., Kumar D. N., An efficient multi-objective optimization algorithm based on swarm intelligence for engineering design, *Journal of Global Optimization*, 39 (2007), No. 1, 49–68
- [13] Canova A., Freschi F., Repetto M., Hybrid method coupling ais and zeroth order deterministic search, *COMPEL*, 24 (2005), No. 3, 784–795
- [14] Mullur A. A., Messac A., Metamodeling using extended radial basis functions: a comparative approach, *Engineering with Computers*, 21 (2006), 203–217
- [15] Chen V. C. P., Tsui K-W., Barton R. R., Meckesheimer M., A review on design, modeling and applications of computer experiments, *IIE Transactions*, 38 (2006), 273–291
- [16] Santner T. J., Williams B. J., Notz W. I., *The Design and Analysis of Computer Experiments*, *Springer*, 2003
- [17] Regis R. G., Shoemaker C. A., Improved strategies for radial basis function methods for global optimization, *Journal of Global Optimization*, 37 (2007), No. 1, 113–135
- [18] Jones D. R., Schonlau M., Welch W. J., Efficient global optimization of expensive black-box functions, *Journal of Global Optimization*, 13 (1998), 455–492
- [19] Sobester A., Leary S. J., Keane A. J., On the design of optimization strategies based on global response surface approximation models, *Journal of Global Optimization*, 33 (2005), 31–59
- [20] Knowles J., ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems, *IEEE Trans. Evol. Comp.*, 10 (2006), No. 1, 50–66
- [21] Lebensztajn L., et al, A multi-objective analysis of a special switched reluctance motor, *COMPEL*, 24 (2005), No. 3, 931–941
- [22] Emmerich M. T. M., Giannakoglou K. C., Naujoks B., Single- and multiobjective evolutionary optimization assisted by gaussian random field metamodels, *IEEE Trans. Evol. Comp.*, 10 (2006), No. 4, 421–439
- [23] Montgomery D., *Design and Analysis of Experiments*. *John Wiley and Sons*, 2001
- [24] Kalagnanam J. R., Diwekar U. M., An efficient sampling technique for off-line quality control, *Technometrics*, 39 (1997), No. 3, 308–318
- [25] Forrester A. I. J., Sobester A., Keane A. J., Optimization with missing data. *Proceedings of the Royal Society A*, 462 (2006), 935–945
- [26] Sykulski J. K., Reducing computational effort in field optimisation problems, *COMPEL*, 23 (2004), No. 1, 159–172
- [27] Hawe G. I., Sykulski J. K., A hybrid one-then-two stage algorithm for computationally expensive electromagnetic design optimization, *COMPEL*, 26 (2007), No. 2, 240–250
- [28] Matlab, High-Performance Computation and Visualization Software Reference Guide, *The MathWorks Inc.*, 2004
- [29] Vapnik, V., *Estimation of Dependences Based on Empirical Data*, 2nd edition, *Springer*, 2006
- [30] Kim, D-H., Ship, K. S., Sykulski J. K., Applying continuum design sensitivity analysis combined with standard EM software to shape optimisation in magnetostatic problems, *IEEE Trans. Magn.*, 40 (2004), No. 2, 1156–1159
- [31] Kim, D-H., Sykulski J. K., Lowther D. A., A novel scheme for material updating in source distribution optimization of magnetic devices using sensitivity analysis, *IEEE Trans. Magn.*, 41 (2005), No. 5, 1752–1755
- [32] Selected papers from the 9<sup>th</sup> Workshop on Optimisation and Inverse Problems in Electromagnetism, Sorrento 2006, *COMPEL*, 26 (2007), No. 2, 229–402

**Author:** Professor Jan Sykulski, FIEE, FInstP, FBCS, University of Southampton, Electrical Power Engineering, School of ECS, Highfield, Southampton SO17 1BJ, UK, E-mail: [jks@soton.ac.uk](mailto:jks@soton.ac.uk)