Math. Program., Ser. B 101: 339–364 (2004)

Digital Object Identifier (DOI) 10.1007/s10107-004-0528-5

Giampaolo Liuzzi · Stefano Lucidi · Veronica Piccialli · Antonello Sotgiu

A magnetic resonance device designed via global optimization techniques*

Received: February 28, 2002 / Accepted: April 2, 2004
Published online: 21 July 2004 – © Springer-Verlag 2004

Abstract. In this paper we are concerned with the design of a small low-cost, low-field multipolar magnet for Magnetic Resonance Imaging with a high field uniformity. By introducing appropriate variables, the considered design problem is converted into a global optimization one. This latter problem is solved by means of a new derivative free global optimization method which is a distributed multi-start type algorithm controlled by means of a simulated annealing criterion. In particular, the proposed method employs, as local search engine, a derivative free procedure. Under reasonable assumptions, we prove that this local algorithm is attracted by global minimum points. Additionally, we show that the simulated annealing strategy is able to produce a suitable starting point in a finite number of steps with probability one.

Key words. Magnetic Resonance Imaging – Global optimization – Simulated annealing – Derivative free methods

1. Introduction

Magnetic Resonance Imaging (MRI) or Magnetic Resonance Tomography (MRT) [23, 16, 18, 3] is a powerful diagnosis instrument especially because of its non-invasive type, its high resolution and its ability to scan even soft tissues. Patients undergoing an MRI exam lie flat on a scanning table and are placed in a magnetic field created by a huge magnet.

MRI is strongly affected by the uniformity of the magnetic field generated by this magnet, in that, the more uniform the magnetic field is, the higher the resolution of the images produced during the exam is. Another important issue is connected to the dimension of the uniformity region. Indeed, it delimits the region that can be actually scanned by the MR apparatus.

The present technology of medical application of MRI has recently evolved toward very high magnetic field units (2-3 Tesla) based on super-conducting magnets with a cost of a few million dollars. This reflects on the cost of an MRI analysis. The result is that many fields of medicine cannot benefit from the use of these apparatus that are too

G. Liuzzi, S. Lucidi, V. Piccialli: Università di Roma "La Sapienza", Dipartimento di Informatica e Sistemistica "Antonio Ruberti", via Buonarroti, 12, 00185 Roma, Italy.
e-mail: {liuzzi, lucidi, piccialli}@dis.uniroma1.it

A. Sotgiu: INFM, Università de L'Aquila, Dipartimento di Tecnologie Biomediche, Coppito, 67100 L'Aq-uila, Italy. e-mail: sotgiu@univaq.it

* This work was supported by CNR/MIUR Research Program "*Metodi e sistemi di supporto alle decisioni*", Rome, Italy.

Mathematics Subject Classification (1991): 65K05, 62K05, 90C56

expensive for pathologies which are considered marginal in comparison, for example, to cardiac or cerebral diseases.

For this reason new kind of equipments are being devised with the purpose of lowering the cost of MRI analysis. They require new technology developments, particularly in the magnet design. The main purpose is to build economical MRI systems focused to dedicated applications. The more relevant ones are low-field MRI systems for orthopedic applications.

Naturally, low field apparatus are inherently less sensitive than high field ones. However, this drawback can be overcome by a high uniformity of the magnetic field which results in high resolution images.

Low field usually requires a lower magnet volume and weight, which imply a lower cost.

To date, the few small MR units produced, present the following drawbacks

- a low field uniformity (over 50 ppm);
- a high cost essentially due to the lack of market competition.

In this frame the development of dedicated equipment based on low field magnets is strategic for increasing the applications of MRI and their use in environments different from radiological sites.

One possible technique to build a low field dedicated magnet is that of using permanent magnets surrounded by an iron yoke to amplify the magnetic field. The presence of this iron yoke requires a careful design to obtain the required properties. It, in fact, introduces non linearities in the dependence of the field from the characteristics of the used magnetic material. For this reason the design procedure heavily relays on optimization procedures along with field computation programs.

To conclude, our main effort in this paper is aimed toward the design of a small low cost, low field (0.08 T) MR apparatus with the following innovative properties:

1. a big uniformity region with respect to the dimension of the structure;
2. a high level of field uniformity.

The above design problem can be converted into an optimization problem with the following properties. First of all, the objective function must represent the uniformity of the magnetic field within the target region and hence it depends on the relationship between the project variables and the magnetic field behavior. This implies that the objective function is nonlinear and non convex which does not allow us to rule out the presence of multiple local minimizers. Moreover, the behavior of the magnetic field is not known analytically, but it can only be computed by means of some simulation code. For this reason, the objective function can be considered as a black box with unavailable derivatives.

As regards the project variables, besides their usual boundedness, they must satisfy some design specifications. In particular, this turns out in nonlinear and non convex constraints involving the model geometry and its feasibility. Anyway, the particular structure of the problem allows us to tackle them without using sophisticated mathematical instruments.

In order to solve problems with the above features, we define a new derivative-free global optimization method. The idea behind the proposed approach consists in

performing several distributed derivative-free local minimizations starting from randomly generated feasible points satisfying a simulated annealing acceptability criterion.

In Section 2 we describe model of the magnet and we define the corresponding optimization problem. In Section 3 we introduce the new global optimization algorithm along with its properties and we report some numerical results on a set of test problems to show its efficiency. Finally, Section 4 is devoted to the results on the optimal design problem.

We conclude this section by introducing some notation used throughout the text. Given a vector or a scalar, say a, a subscript a_i indicates the i-th element of a set. Whereas, a superscript a^k identify the k-th iterate of a generic algorithm. We write a^+ to denote $\max\{0, a\}$. Moreover, for the sake of simplicity, given a m-tuple $(a_1, \ldots, a_m)$ we shall denote by $(a_1, \ldots, a_m)^k$ the m-tuple $(a_1^k, \ldots, a_m^k)$. Finally, we denote by $\mathcal{B}(x^*; \epsilon)$ the set $\{x \in \mathbb{R}^n : \|x - x^*\| \leq \epsilon\}$.

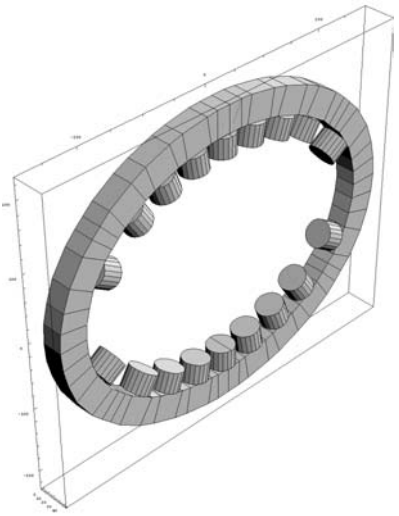
2. Model description and problem definition

In this section we describe an abstract model of the multipolar magnet. Then, after introducing and discussing some simplifying assumptions on the magnet, we define the variables which describe the model. Finally, we state the optimal design problem in terms of a global minimization constrained program. We remark that, since the practical realization of a physical model of the magnet is very expensive, this very task can be done only after the achievement of sound results which are precisely the purpose of the present work.

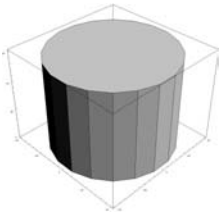
We consider a model (M) of the multipolar magnet which has the shape of a cylinder with an elliptical base. The elliptical shape of the magnet basis has been chosen to maximize the possibility of use in different situations (such as for the peripheral skeleton and paediatrics).

The magnet (M) is composed of twelve elliptical iron rings with eighteen small permanent magnets screwed on each of them. The arbitrary choice of the number of rings and magnets which compose the model, is due to the following reasons. As concerns the rings, we have to conciliate the need of limiting the border effect, which requires many rings, with the need of obtaining a low-cost structure, which, on the opposite, would drive us to use as few rings as possible. As for the small magnets, their number is directly connected with the requirement that the field intensity should be approximately 0.08 Tesla.

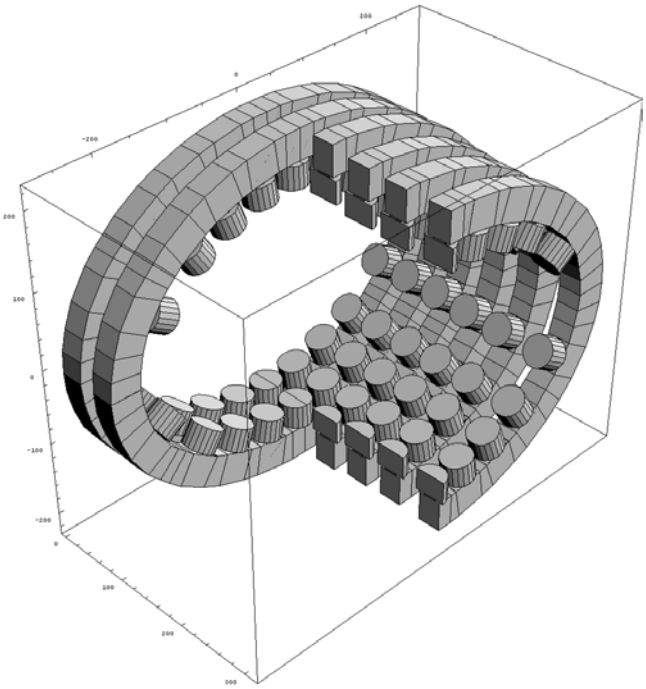
Moreover, for reasons which regard the future practical realization of the model, we limit ourselves to consider every ring equal to each other both in terms of dimensions of the semi-axes and in terms of the small magnets positioning. As a matter of fact, this restrictive assumption, though resulting in a less flexible structure, would ease the industrial building process. For the same reason, every small permanent magnet is assumed to be equal to each other. Therefore, every ring has $236mm \times 180mm$ long semi-axes and is 30mm wide and 40mm thick (like the one depicted in Figure 1(a)) and every small magnet has a cylindrical shape, with 42mm diameter and 30mm height (as shown in figure 1(b)), and a residual magnetization $M_r = 1.35$ Tesla.



(a)



(b)



(c)

Fig. 1.

Given all the above assumptions on model (M), we would need 227 variables to completely describe its geometry, that is, eleven displacement variables of the rings with respect to a fixed ring plus eighteen position variables of the magnets on every ring. Even though with all these variables the model would certainly have many degrees of freedom, a balance is to be found among a sufficient elasticity of the model (more variables) and the tractability of the underlying global optimization problem (less variables). In order to try and reduce the number of variables we have made some simplifying assumptions on model (M).

First of all, we assume that two magnets are fixed on every ring and directed along the short semi-axis. This seems reasonable if we seek for a uniform magnetic field directed along their axis. Moreover, we assume that only four magnets can be freely positioned on a quarter of ring and get the remaining twelve positions by reflection on the semi-axes of the ring. Furthermore, we assume that the whole structure is symmetric with respect to a plane parallel to the rings and which divides the model into two halves of six rings (see figure 1(c)).

Under the above assumptions, we come up with a model which is fine from a constructive point of view and which can be described by just 10 variables. However, practical experiences showed that this model has too few degrees of freedom. Thus, we relax the assumption that rings are equal to each other and we assume that only the first two rings are equal. So, we introduce four more variables, that is, the offsets of the four outermost rings with respect to the two innermost ones (see figure 2(b)). We chose these variables because in this way we have a better control on the border effect.

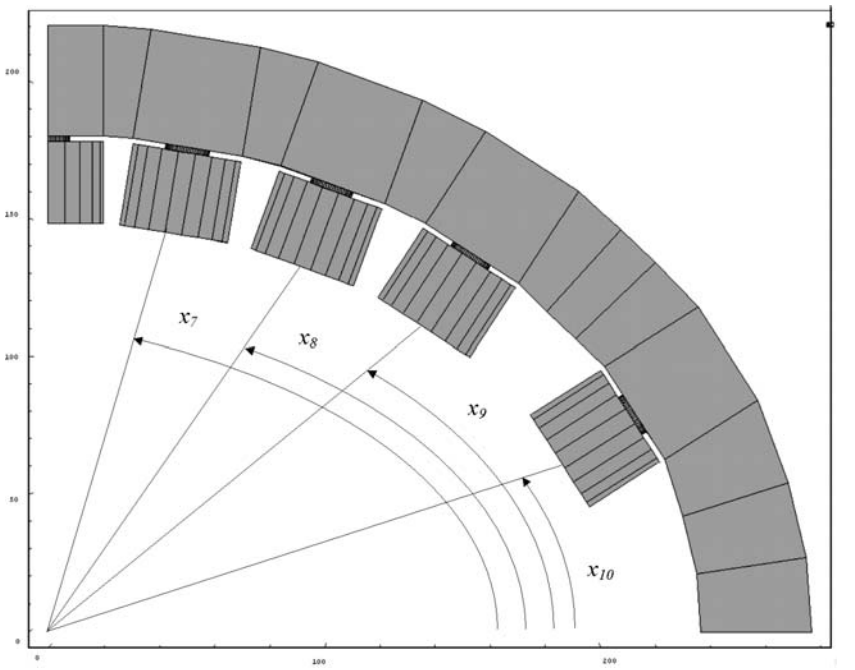
As already said, the key feature of every MR apparatus is the uniformity of the magnetic field within a specified region in space. In our case, this so-called uniformity region is a cylinder (12 cm long) having an elliptical base with major and minor semi-axes respectively 6 and 5 centimeters long, which is quite a large region with respect to the dimension of the multipolar magnet itself.

To summarize, we consider as optimization variables

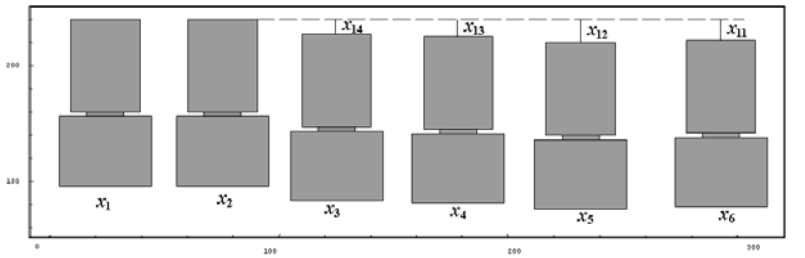
1. the positions $(x_1, \ldots, x_6)$ of the rings along the cylinder axis (see figure 2(b));
2. the angular positions (see figure 2(a)) of each row of small permanent magnets $(x_7, \ldots, x_{10})$, being a row constituted by magnets occupying the same position but on different rings (as shown by figure 1(c));
3. the offsets $(x_{11}, \ldots, x_{14})$ of the four outermost rings with respect to the two innermost ones (see figure 2(b)).

In spite of its simple structure and due to the high nonlinearities introduced by the iron yoke, the field generated by model (M) can not be analytically computed. For this reason, the multipolar magnet model, corresponding to a feasible choice of the above mentioned variables, is computer rendered within a field simulation program which predicts the behavior of the magnetic field by means of a sophisticated finite element analysis. In particular, we employ RADIA (see [10, 4]) a reliable finite element simulator released by the *European Synchrotron Radiation Facility* (ESRF) at Grenoble (France). The simulation process carried out by this code can take up to a minute of CPU time on a Pentium IV processor based computer.

Let XYZ be a system of Cartesian coordinates with origin at the center of the structure, X axis parallel to the cylinder axis and, Z and Y axes directed, respectively, along



(a) Angles



(b) Displacements and offsets

Fig. 2. Optimization variables

the shortest and longest semi-axes of the cylinder base. As already said, we want the magnetic field **B** generated by the multipolar magnet to be as uniform as possible within the target region and directed along the Z axis.

In order to measure the uniformity of the magnetic field within the target region, we just sample the three components of **B** on a grid of N_p points uniformly distributed inside the cylindrical region of interest. Let $B_X^{(i)}(x)$, $B_Y^{(i)}(x)$ and $B_Z^{(i)}(x)$ be the three components of the magnetic field measured at the i-th point of the grid, the objective function will be

$$U(x) = \frac{\sum_{i=1}^{N_p} \left((B_Z^{(i)}(x) - \bar{B}_Z(x))^2 + (B_X^{(i)}(x))^2 + (B_Y^{(i)}(x))^2 \right)}{\bar{B}_Z(x)^2},$$

where $x = (x_1 \dots x_{14})^\top$ and

$$\bar{B}_Z(x) = \frac{\sum_{i=1}^{N_p} B_Z^{(i)}(x)}{N_p}.$$

As regards the constraints, they can be divided into four groups.
The first one consists of the following box constraints on the variables:

$$\begin{aligned} 0 \le x_i \le 300, &\quad i = 1, \dots, 6 \\ 0 \le x_i \le 90, &\quad i = 7, \dots, 10 \\ -10 \le x_i \le 20, &\quad i = 11, \dots, 14. \end{aligned}$$

The second group ensures rings to be separated one from another, namely the following inequalities must be satisfied:

$$\begin{aligned} g_1(x) &= -x_1 + r < 0 \\ g_i(x) &= x_{i-1} - x_i + 2r < 0, \quad i = 2, \dots, 6, \end{aligned} \tag{1}$$

where $r = 21$ is the magnet radius. The third group of constraints ensures the ordering of the angular positions of the magnets:

$$g_i(x) = x_{i+1} - x_i < 0, \quad i = 7, \dots, 9. \tag{2}$$

Finally there is a group of constraints which are meant to avoid magnets overlapping:

$$g_i(x) < 0, \quad i = 10, \dots, 49. \tag{3}$$

These constraints are continuously differentiable but non convex. Since their expression is technical and does not enrich the paper, they are reported in Appendix.
We point out that the constraints $g(x) < 0$ can not be relaxed into $g(x) \le 0$ since the simulator outcome is not reliable whenever $g_i(x) \ge 0$ for $i = 1, \dots, 49$. On the other hand, the physics of the magnetic apparatus suggests that a good uniformity of the magnetic field can not be obtained with points x such that $g(x) = 0$. For this reason, it seems reasonable to assume that it is likely to find a feasible point $\tilde{x}$ such that $g(\tilde{x}) < 0$ and $U(\tilde{x}) \le U(x)$ for every x such that $g(x) = 0$.

Summarizing, our optimal design problem belongs to the following class of optimization problems[1]

$$\begin{aligned} &\min\ U(x) \\ &g(x) < 0 \\ &l \leq x \leq u, \end{aligned} \tag{M}$$

with the properties listed below:

(i) the objective function is of the black-box type in the sense that apart from its continuity in the feasible set, no information on its structure is available and its computation is expensive;
(ii) the objective function is not defined outside of the feasible set and the constraints are continuous but non convex; on the other hand it is "relatively" easy to generate feasible points and to remain in the feasible set;
(iii) it is possible to find a feasible point $\tilde{x}$ such that the set $\{x \in \mathbb{R}^n : g(x) < 0, l \leq x \leq u, U(x) \leq U(\tilde{x})\}$ is compact;
(iv) the objective function does not satisfy any convexity assumptions and hence the problem can have local minimizers. Indeed, a local minimization algorithm applied to this problem starting from different initial points produces different solutions (see the results reported in Section 4).

Feature (i) eliminates the possibility of using those methods that exploit derivative knowledge and forces us to employ methods which are not too expensive with respect to the number of function evaluations. Feature (ii), on the one hand, requires that during the minimization process only feasible points are explored and, on the other hand, implies that the feasibility can be maintained without using sophisticated techniques. Feature (iii) guarantees that Problem (M) admits a solution. Finally, feature (iv) qualifies (M) as a global optimization problem so that a global optimization method is to be used.

From the above discussion it emerges that the solution method should be able to find a global minimum point among the local ones without requiring any derivative knowledge and using a limited number of function evaluations. Most of the global derivative free methods usually require too many function evaluations to get a solution and are able to efficiently tackle problems with only a few variables. Therefore, they are not particularly well-suited to tackle problems with features (i)-(iv) (see also the introduction of [2] and the references therein). On the opposite, sufficiently good solutions of minimization problems showing some or all of the features (i)-(iv) have been obtained by using some algorithms belonging to the particular class of Controlled Random Search (CRS) methods [8, 2, 5, 24, 25, 12, 17]. All these algorithms, though intended for unconstrained problems, are easily extendible to handle constraints of type (ii). Such methods are usually very efficient in identifying the region where the global minimum point is located, but they do not have a similar ability in determining a good approximation of the global minimum point. Since the quality of the solution is of fundamental importance

[1] A description of Problem (M) in the AMPL [11] syntax can be found at the URL `http://www.dis.uniroma1.it/~liuzzi/papers/Model/MRmodel.zip`. Although the model is of public domain, to compute the objective function one should have a licensed version of Mathematica [27] along with the freeware RADIA simulator package.

to our application, we have tried to define a new algorithm which, hopefully, had the same global capacity of CRS methods but an higher accuracy. The description of this new algorithm is the topic of the next section.

3. Global optimization approach

We consider the following constrained problem

$$\min f(x) \qquad (4)$$
$$x \in \mathcal{F} \subseteq \mathbb{R}^n,$$

where $\mathcal{F} = \mathcal{D} \cap \mathcal{S}$ with $\mathcal{D} = \{x \in \mathbb{R}^n : l \leq x \leq u\}$ and $\mathcal{S}$ an open set. For a given $x^\circ \in \mathbb{R}^n$, we denote the level set of $f(x)$ by

$$\Omega(x^\circ) = \{x \in \mathcal{F} : f(x) \leq f(x^\circ)\}.$$

Throughout this section we shall assume that, even though no derivative knowledge is available for the objective function, $f(x)$ is twice continuously differentiable.
Given a feasible point $x \in \mathcal{F}$, we denote by

$$A(x) = \{d \in \mathbb{R}^n : \exists \bar{\alpha} > 0,\ \forall \alpha \in [0, \bar{\alpha}],\ x + \alpha d \in \mathcal{F}\}$$

the set of feasible directions in x.
We define a stationary point of Problem (4) as a point $x^* \in \mathcal{F}$ such that

$$\nabla f(x^*)^T d \geq 0,\ \forall\, d \in A(x^*).$$

The algorithm we propose is based on an idea which is not new among optimization practitioners, namely the repeated use of local minimization algorithms as local search engine of a global minimization. In fact, it is well-known that good approximations of minimum points can be obtained by using efficient local minimization algorithms. However, when a local algorithm starts far from a global minimum point, it can be trapped into a local minimum. In order to overcome this last difficulty, we have followed the approach proposed in [20] where:

- the starting points of the local searches are chosen at random according to a probability density function which is updated during the iterates of the algorithm so as to be concentrated around a global minimum point of the problem;
- the local searches are performed by an algorithm which has the property that it is attracted by any global minimum point.

As concerns the latter property, Proposition 1.12 in [1] shows that any gradient related method is attracted by every strong local minimum point. An analogous result holds for trust region algorithms (see e.g. [6]) which require derivative knowledge too. Unfortunately, in our context we can not use any derivative-based method. However, the following proposition shows that a similar result can be obtained without requiring any derivative information.

Proposition 1. *Let $f \in \mathcal{C}^2$ and $\{x^k\}$ be a sequence of feasible points generated by an iterative method $x^{k+1} = x^k + \alpha^k d^k$ satisfying*

$$f(x^{k+1}) \leq f(x^k) - \theta(\alpha^k)^2 \|d^k\|^2, \tag{5}$$

for all k, where $\theta > 0$. Then, for every global minimum x^* of $f(x)$ on $\mathcal{F}$ where $\nabla^2 f(x^*)$ is positive definite, there exists an open set $\mathcal{L}$ containing x^* such that, if $x^{\bar{k}} \in \mathcal{L}$ for some $\bar{k} \geq 0$

- *then $x^k \in \mathcal{L}$ for all $k \geq \bar{k}$;*
- *if any accumulation point of the sequence $\{x^k\}$ is a stationary point of Problem (4), then $\{x^k\} \to x^*$.*

Proof. Let x^* be a global minimum of $f(x)$ on $\mathcal{F}$ with $\nabla^2 f(x^*)$ positive definite. Then an $\bar{\epsilon} > 0$ exists such that $\mathcal{B}(x^*, \bar{\epsilon}) \subset \mathcal{S}$ and for all $x \in \mathcal{B}(x^*, \bar{\epsilon})$, the matrix $\nabla^2 f(x)$ is also positive definite. Denote by

$$\gamma = \min_{\substack{\|x - x^*\| \leq \bar{\epsilon} \\ \|z\| = 1}} z^T \nabla^2 f(x) z. \tag{6}$$

It obviously results $\gamma > 0$. Consider the open set

$$\mathcal{L} = \{x \in \mathcal{F} \mid \|x - x^*\| < \bar{\epsilon},\ f(x) < f(x^*) + \frac{\bar{\epsilon}^2 \gamma \theta}{2(2\theta + \gamma)}\} \tag{7}$$

which, by continuity, is not empty. We claim that, if $x^{\bar{k}} \in \mathcal{L}$ for some $\bar{k} \geq 0$, then $x^k \in \mathcal{L}$ for all $k \geq \bar{k}$ and $\{x^k\} \to x^*$.
Indeed, by using Taylor's theorem and the mean value theorem, we have

$$\begin{aligned} f(x^{\bar{k}}) - f(x^*) = {} & \nabla f(x^*)^T (x^{\bar{k}} - x^*) \\ & + \frac{1}{2}(x^{\bar{k}} - x^*)^T \nabla^2 f(x^* + \lambda^{\bar{k}}(x^{\bar{k}} - x^*))(x^{\bar{k}} - x^*), \end{aligned} \tag{8}$$

where $\lambda^{\bar{k}} \in (0, 1)$. Since, $\mathcal{D} \cap \mathcal{B}(x^*, \bar{\epsilon})$ is a convex set and $\mathcal{B}(x^*, \bar{\epsilon}) \subset \mathcal{S}$ by hypothesis, then $x^{\bar{k}} - x^*$ is a feasible direction for Problem (4). By the assumption that x^* is a stationary point of Problem (4), we have that

$$\nabla f(x^*)^T (x^{\bar{k}} - x^*) \geq 0. \tag{9}$$

Now, (9), (8) and (6) imply

$$\begin{aligned} f(x^{\bar{k}}) - f(x^*) &\geq \frac{1}{2}(x^{\bar{k}} - x^*)^T \nabla^2 f(x^* + \lambda^{\bar{k}}(x^{\bar{k}} - x^*))(x^{\bar{k}} - x^*) \\ &\geq \frac{1}{2}\gamma \|x^{\bar{k}} - x^*\|^2. \end{aligned} \tag{10}$$

Recalling the hypothesis $f(x^{k+1}) \leq f(x^k) - \theta(\alpha^k)^2 \|d^k\|^2$ we have

$$\begin{aligned} (\alpha^{\bar{k}})^2 \|d^{\bar{k}}\|^2 &\leq \frac{1}{\theta}(f(x^{\bar{k}}) - f(x^{\bar{k}+1})) \\ &\leq \frac{1}{\theta}(f(x^{\bar{k}}) - f(x^*)), \end{aligned} \tag{11}$$

where the second inequality follows from the hypothesis that x^* is a global minimum. By using (10) and (11), we have

$$\begin{aligned}\|x^{\bar{k}+1} - x^*\|^2 &= \|x^{\bar{k}} - x^* + \alpha^{\bar{k}} d^{\bar{k}}\|^2 \\ &\leq 2\|x^{\bar{k}} - x^*\|^2 + 2(\alpha^{\bar{k}})^2 \|d^{\bar{k}}\|^2 \\ &\leq \frac{4}{\gamma}(f(x^{\bar{k}}) - f(x^*)) + \frac{2}{\theta}(f(x^{\bar{k}}) - f(x^*)) \\ &= \frac{2(2\theta+\gamma)}{\gamma\theta}(f(x^{\bar{k}}) - f(x^*)).\end{aligned} \tag{12}$$

Since $x^{\bar{k}} \in \mathcal{L}$ we have

$$f(x^{\bar{k}}) - f(x^*) < \frac{\bar{\epsilon}^2 \gamma\theta}{2(2\theta+\gamma)},$$

which, combined with (12) yields

$$\|x^{\bar{k}+1} - x^*\| < \bar{\epsilon}. \tag{13}$$

Furthermore, by using again the hypothesis $f(x^{k+1}) \leq f(x^k) - \theta(\alpha^k)^2\|d^k\|^2$, we have that $f(x^{k+1}) \leq f(x^k)$ for all k and hence

$$f(x^{\bar{k}+1}) \leq f(x^{\bar{k}}) < f(x^*) + \frac{\bar{\epsilon}^2\gamma\theta}{2(2\theta+\gamma)} \tag{14}$$

From the above two inequalities, it follows that $x^{\bar{k}+1} \in \mathcal{B}(x^*, \bar{\epsilon})$ and similarly $x^k \in \mathcal{L}$ for all $k \geq \bar{k}$. In particular, since $\mathcal{L} \subset \mathcal{B}(x^*, \bar{\epsilon})$, then $x^k \in \mathcal{B}(x^*, \bar{\epsilon})$ for all $k \geq \bar{k}$. Since, by definition, $\mathcal{B}(x^*, \bar{\epsilon})$ is a compact set, the sequence $\{x^k\}$ will have at least one limit point which by assumption must be a stationary point of Problem (4). Now, since $f(x)$ is strictly convex within $\mathcal{B}(x^*, \bar{\epsilon})$, the only stationary point of Problem (4) within $\mathcal{B}(x^*, \bar{\epsilon})$ is the point x^*. Hence $x^k \to x^*$. □

Remark. Proposition 1 has been stated in the case $\mathcal{F} = \mathcal{D} \cap \mathcal{S}$ with $\mathcal{D} = \{x \in \mathbb{R}^n : l \leq x \leq u\}$. However, the same result holds even if $\mathcal{D}$ is a generic closed convex set.

The above proposition states that, under reasonable assumptions, all the iterates are attracted by a strong global minimum point provided that one of them is sufficiently close to it. Note that usually a well posed physical problem has isolated global minima. On the other hand, if the problem has been carefully formulated, it is reasonable to assume that the hessian of the objective function is positive definite on the global minimum points. For this reason, the assumption on the positive definitiveness of $\nabla f(x^*)$ does not seem too restrictive and it is likely to be verified.

In the literature some derivative free algorithms satisfying property (5) have been proposed, see, for instance, [14], [21] and [13]. In [22] one of such methods has been defined to solve box constrained minimization problems. In this work we adapt the approach described in [22] to solve Problem (4). In particular, the proposed algorithm is based on the following procedure, which performs a sampling of the objective function along the coordinate axes.

Procedure DF(x, α)

Data. $\gamma > 0$, $\delta \in (0, 1)$, $\theta \in (0, 1)$, $d^i = e^i$ and $\tilde{\alpha}^i = \alpha$ for $i = 1, \ldots, n$.

1. *Initialization:* Set $i = 1$ and $x^i = x$.
2. *Direction choice:*
   2.1 Compute α^i_{max} s.t. $x^i + \alpha^i_{max} d^i \in \partial\mathcal{D}$ and set $\alpha = \min\{\tilde{\alpha}^i, \alpha^i_{max}\}$.
   If $\alpha > 0$, $f(x^i + \alpha d^i) \leq f(x^i) - \gamma(\alpha)^2$ and $x^i + \alpha d^i \in \mathcal{S}$,
   then go to Step 4.
   2.2 Compute α^i_{max} s.t. $x^i - \alpha^i_{max} d^i \in \partial\mathcal{D}$ and set $\alpha = \min\{\tilde{\alpha}^i, \alpha^i_{max}\}$.
   If $\alpha > 0$, $f(x^i - \alpha d^i) \leq f(x^i) - \gamma(\alpha)^2$ and $x^i - \alpha d^i \in \mathcal{S}$,
   then set $d^i = -d^i$ and go to Step 4.
3. *Direction failure:* Set $\bar{\alpha} = 0$, $\tilde{\alpha}^i = \theta\alpha$, and go to Step 5.
4. *Linesearch:*
   4.1 Let $\hat{\alpha} = \min\{\alpha^i_{max}, \frac{\alpha}{\delta}\}$.
   If $\alpha = \alpha^i_{max}$ or $f(x^i + \hat{\alpha} d^i) > f(x^i) - \gamma\hat{\alpha}^2$ or $x^i + \hat{\alpha} d^i \notin \mathcal{S}$,
   then set $\bar{\alpha} = \alpha$, $\tilde{\alpha}^i = \alpha$ and go to step 5.
   4.2 Set $\alpha = \hat{\alpha}$ and go to step 4.1.
5. *New point:* Set $x^{i+1} = x^i + \bar{\alpha} d^i$.
6. *Stopping criterion:* If $i = n$, then $(\tilde{x}, \tilde{\alpha}) = (x^{i+1}, \max_{i=1,\ldots,n}\{\tilde{\alpha}^i\})$, return $(\tilde{x}, \tilde{\alpha})$
   else set $i = i + 1$ and go to Step 2.

Then the local minimization algorithm is described below.

Procedure DFA$(x^\circ, \alpha_{\text{tol}})$

Data. $\alpha^\circ > 0$.

1. Set $j = 1$ and $x^j = x^\circ$.
2. Set $(x^{j+1}, \alpha^j) = DF(x^j, \alpha^{j-1})$.
3. If $\alpha^j > \alpha_{\text{tol}}$ then set $j := j + 1$ and go to step 2.
   else return (x^{j+1}, α^j).

Whenever $\alpha_{\text{tol}} = 0$, it is possible to show that Procedure DFA satisfies the following properties.

Proposition 2. *Assume that $f \in C^1$ and that the level set $\Omega(x^\circ)$ is compact. Let $\{x^j\}$ and $\{\alpha^j\}$ be the sequences produced by Procedure DFA$(x^\circ, 0)$, then:*

(i) for every $x^\circ \in \mathbb{R}^n$, every accumulation point of the sequence $\{x^j\}$ is a stationary point of Problem (4) and

$$\lim_{j\to\infty} \alpha^j = 0; \tag{15}$$

(ii) for every global minimum $x^\star$ where $\nabla^2 f(x^\star)$ is positive definite, then an open set $\mathcal{L}$ exists such that if $x^\circ \in \mathcal{L}$ then $x^j \in \mathcal{L}$ for every index j and $\lim_{j\to\infty} x^j = x^\star$.

Proof. As regards the proof of point (i) we refer to [21]. Point (ii) follows directly from Proposition 1 and the instructions of Procedure DF. □

Equation (15) of the above proposition guarantees that Procedure DFA terminates provided that $\alpha_{\text{tol}} > 0$. As for point (ii), it guarantees that if the starting point of a local minimization is sufficiently close to a global minimum point with positive definite Hessian, then the whole sequence converges to it. In order to ensure that this happens, in the following algorithm (DFSA) we make use of the simulated annealing acceptability criterion proposed in [20]. In particular, the basic idea is to let a local minimization start from points chosen at random according to a probability density function proportional to $e^{\frac{-(f(x)-f_{\min})^+}{T}}$, where T is a parameter called "annealing temperature" which is progressively decreased during the minimization process.

Algorithm DFSA

Data. $\alpha_\circ > 0$ and $\alpha_{\text{tol}} > 0$.

1. Compute an initial temperature T°, randomly choose a feasible point x°, set $f_{\min} = f(x^\circ)$ and $x_{\min} = x^\circ$. Set $k := 1$.
2. Randomly choose a feasible point $x^k \in \mathbb{R}^n$ and a scalar $z^k \in (0, 1)$.
3. If $z^k \leq e^{\frac{-(f(x^k)-f_{\min})^+}{T^k}}$ then go to Step 4,
   else compute T^{k+1} and go to Step 6.
4. Compute $(\tilde{x}^k, \alpha) = \text{DFA}(x^k, \alpha_{\text{tol}})$.
5. If $f(\tilde{x}^k) < f_{\min}$ then set $f_{\min} = f(\tilde{x}^k)$ and $x_{\min} = \tilde{x}^k$.
6. Set $k := k + 1$ and go to Step 2.

As concerns the updating of the temperature parameter T^k at Step 3, we refer to the updating rule reported in [20].

Recalling the analysis carried out in [20, 26] it is possible to prove the following result.

Proposition 3. *For every global minimum point $x^\star$ of $f(x)$ on $\mathcal{F}$ and for every $\epsilon > 0$, Step 3 of Algorithm DFSA accepts a point $x^k \in \mathcal{B}(x^*; \epsilon) \cap \mathcal{F}$ in a finite number of iterations with probability one.*

Proposition 3 guarantees that, for every global minimum point $x^\star$ of $f(x)$ on $\mathcal{F}$ with positive definite Hessian, a point $x^k \in \mathcal{L}$ (where $\mathcal{L}$ is the neighborhood of $x^\star$ defined in Proposition 1) is accepted in a finite number of steps with probability one. In turn, Proposition 1 guarantees that a local minimization starting from such a point is attracted by the global minimum.

In order to evaluate the efficiency of DFSA we have tested it on a set of well-known global optimization test problems (see [2] and the references therein). These test problems are essentially unconstrained in the sense that the bounds on the variables are inactive at the global minimum point. Our choice is motivated by the fact that we want to obtain a comparison as realistic as possible with a particular algorithm of the Controlled Random search class, namely ACRS[2] (see [5] for a complete description of ACRS).

[2] ACRS has been recently released on the NEOS server [15, 7, 9] under the Global Optimization category.

Algorithm ACRS iteratively improves a working set of p points initially chosen at random on the feasible set $\mathcal{F}$. At each iteration, the algorithm, selects at random $n+1$ points of the current working set privileging those points with a lower function value. Then the weighted centroid of the selected points is computed and the new trial point is obtained by doing a weighted reflection of the worst point (among the selected ones) on the centroid. The weights used to compute the centroid and to do the reflection adaptively change during the iterations to better exploit the information gathered on the objective function as the algorithm goes on. Finally, the working set is updated by replacing the point with the highest function value with the new point provided that it is feasible and has a function value lower than the maximum value over the working set. The algorithm stops when the difference between the maximum and the minimum function values over the working set, is below a prefixed tolerance *tol*.

Our choice of comparing DFSA with ACRS is mainly due to the fact that ACRS has already been employed to tackle the magnetic resonance optimal design problem (see [19]) as well as other optimal design problems of various types (see, e.g., [5, 8, 2]).

The results of Table 1 were obtained choosing $p = 25n$ and $tol = 10^{-6}$ for ACRS. As concernes DFSA, we set $\alpha_{\text{tol}} = 10^{-6}$ and chose to stop it whenever the test at Step 5 fails ten times consecutively. We note that *tol* and α_{tol} are intentionally equal in order to get comparable results.

Moreover, since DFSA and ACRS use randomly generated points, the reported results refer to the average behavior on 100 runs.

From Table 1, it clearly emerges the superiority of DFSA in terms of precision of the computed solution for problems with dimension ranging from 10 to 100, which is a feature we are interested in since we aim at obtaining a high uniformity level in the optimal design problem (which, as already said, has 14 variables). Moreover, we note that DFSA is worse than ACRS with respect to the number of objective function evaluations when $n \leq 10$. The two algorithms are comparable when $n = 15, 20$, while DFSA becomes more and more efficient when $n > 20$. This behavior is essentially imputable to the intimate difference between the two algorithms. As concerns DFSA, its multi-start nature along with a poor stopping criterion are responsible for the high computational burden of the algorithm. On the other hand, since every local minimization of algorithm DFSA is carried out using a derivative free linesearch strategy, DFSA is able to locate the global minimum point with a high precision. On the opposite, ACRS employs a simplex based local search which allows it to locate a rough estimate of the global minimum point very rapidly, especially for small problems, but at the expense of the accuracy.

In order to improve the efficiency of algorithm DFSA retaining its accuracy, we can draw our inspiration from ACRS itself. Indeed, an important aspect of the latter algorithm is that its working set is iteratively updated by performing, at each iteration, a single step of a simplex based local search. Thus, no local search is carried out completely, as opposed to DFSA, and this allows for a considerable reduction of function evaluations. Moreover, the presence of such an array of points conveys a certain degree of nonmonotonicity thus improving the ability to find a global minimum point. Following these lines we have introduced a *working set* W of m triples $(x, f(x), \alpha)$ where, in particular, α is the current step of the derivative free algorithm at x. The working set W can be updated either by performing a single step of the derivative free algorithm (using Procedure DF) or by replacing the worst point, namely the one with highest objective function value,

Table 1. Numerical comparison between DFSA and ACRS algorithms.

		DFSA			ACRS		
PROBLEM	n	nf	F. min	F. aver.	nf	F. min	F. aver
S-H. Camel B.	2	2068	−1.0316	−1.0316	522	−1.0316	−1.0316
Treccani	2	970	0.000D+00	0.000D+00	485	0.178D-09	0.118D-07
Quartic	2	1697	−0.3524	−0.3524	524	−0.3524	−0.3524
Schubert	2	2246	−186.7309	−186.7309	2271	−186.7309	−186.7309
Schub. pen. 1	2	2098	−186.7309	−179.3470	1255	−186.7309	−186.1988
Schub. pen. 2	2	1979	−186.7309	−178.5920	1078	−186.7309	−186.3615
Shekel $m = 5$	4	4126	−10.1532	−10.1027	3704	−10.1532	−8.7748
Shekel $m = 7$	4	3896	−10.4029	−10.3502	2119	−10.4029	−10.3362
Shekel $m = 10$	4	3610	−10.5364	−10.4694	2191	−10.5364	−10.2402
Expon.	2	1026	−1.0000	−1.0000	379	−1.0000	−1.0000
Expon.	4	2052	−1.0000	−1.0000	1002	−1.0000	−1.0000
Cos-mix	2	895	−0.2000	−0.2000	471	−0.2000	−0.2000
Cos-mix	4	2156	−0.4000	−0.4000	1310	−0.4000	−0.4000
Hartman	3	3248	−3.8628	−3.8628	754	−3.8628	−3.8628
Hartman	6	7075	−3.3224	−3.3224	2217	−3.3224	−3.3224
5n loc-min	2	959	0.197D-13	0.497D-12	545	0.345D-10	0.106D-07
5n loc-min	5	2696	0.891D-14	0.198D-12	1587	0.443D-08	0.251D-07
5n loc-min	10	5474	0.548D-14	0.691D-13	3912	0.288D-07	0.268D-05
5n loc-min	15	8847	0.314D-14	0.603D-13	7935	0.136D-05	0.582D-04
5n loc-min	20	11285	0.248D-14	0.530D-13	12676	0.561D-04	0.331D-03
5n loc-min	30	17026	0.161D-14	0.304D-13	23262	0.428D-03	1.074E-03
5n loc-min	50	29261	0.826D-15	0.119D-13	53077	0.493D-03	1.371E-03
5n loc-min	100	58473	0.501D-15	0.467D-14	139575	3.463E-03	1.021E-02
10n loc-min	2	1101	0.282D-13	0.116D-10	599	0.795D-11	0.959D-08
10n loc-min	5	2991	0.226D-13	0.219D-11	1715	0.209D-08	0.252D-07
10n loc-min	10	6968	0.119D-13	0.460D-12	4769	0.282D-06	0.469D-04
10n loc-min	15	10214	0.644D-14	0.267D-12	9441	0.227D-04	0.844D-03
10n loc-min	20	13260	0.566D-14	0.134D-12	14245	0.287D-03	0.0031
10n loc-min	30	22682	0.252D-14	0.639D-13	25575	2.220E-03	1.063E-02
10n loc-min	50	39270	0.932D-15	0.622D-03	59789	6.952E-03	2.623E-02
10n loc-min	100	83315	0.884D-15	0.144D-13	155816	6.743E-02	2.523E-01
15n loc-min	2	1057	0.377D-14	0.808D-12	496	0.433D-10	0.951D-08
15n loc-min	5	2958	0.218D-14	0.221D-12	1487	0.338D-08	0.286D-07
15n loc-min	10	6866	0.309D-14	0.187D-12	3820	0.500D-07	0.337D-05
15n loc-min	15	10871	0.23D-14	0.728D-13	8566	0.881D-06	0.377D-03
15n loc-min	20	15725	0.142D-14	0.109D-03	11888	0.362D-04	0.397D-03
15n loc-min	30	24601	0.196D-14	0.493D-13	25050	1.613E-03	7.902E-03
15n loc-min	50	42792	0.149D-14	0.11D-03	63912	1.215E-02	4.635E-02
15n loc-min	100	91848	0.139D-14	0.248D-13	160518	2.852E-01	9.126E-01
Griewank	2	1284	0.000D+00	0.873D-02	781	0.117D-09	0.912D-03
Griewank	5	3565	0.000D+00	0.164D-01	2306	0.524D-08	0.291D-07
Griewank	10	5375	0.000D+00	0.000D+00	5882	0.269D-07	0.129D-06
Griewank	15	8049	0.00D+00	0.00D+00	11210	0.511D-06	0.646D-05
Griewank	20	11325	0.000D+00	0.000D+00	17400	0.2D-05	0.377D-04
Griewank	30	17767	0.00D+00	0.00D+00	31969	0.440D-04	0.158D-03
Griewank	50	29579	0.00D+00	0.00D+00	71379	0.201D-03	0.550D-03
Griewank	100	59263	0.00D+00	0.00D+00	199537	1.054E-03	2.144E-03

with a point produced by a partial minimization carried out by means of Procedure DFA starting from a point accepted by the simulated annealing criterion.

Algorithm DDFSA

1. Data: $\alpha^\circ > 0$.
2. Compute an initial temperature T°, set $k := 0$.
3. Generate the initial working set $W^k = \{(x_i, f(x_i), \alpha_i)^k,\ i = 1, 2, \ldots m\}$.
4. Set

$$(x_{\min}, f_{\min}, \alpha_{\min})^k := (\bar{x}, \bar{f}, \bar{\alpha}) \in W^k : \bar{f} = \min_{(x, f(x), \alpha) \in W^k} f(x),$$

$$(x_{\max}, f_{\max}, \alpha_{\max})^k := (\bar{x}, \bar{f}, \bar{\alpha}) \in W^k : \bar{f} = \max_{(x, f(x), \alpha) \in W^k} f(x),$$

$$\alpha^k_{\text{stop}} = \max_{(x, f(x), \alpha) \in W^k} \alpha.$$

5. Randomly choose a feasible point $x^k \in \mathbb{R}^n$ and a scalar $z^k \in (0, 1)$.
6. If $z^k > e^{\frac{-(f(x^k) - f_{\min})^+}{T^k}}$ compute T^{k+1} and go to step 9.
7. Compute $(\tilde{x}^k, \tilde{\alpha}^k) = \text{DFA}(x^k, \alpha^k_{\text{stop}})$.
8. If $f(\tilde{x}^k) < f^k_{\max}$ then set

$$W^{k+1} := (W^k \setminus \{(x_{\max}, f_{\max}, \alpha_{\max})^k\}) \cup \{(\tilde{x}^k, f(\tilde{x}^k), \tilde{\alpha}^k)\},$$

   $k := k + 1$ and go to step 4.
9. For every triple $(x_i, f(x_i), \alpha_i)^k \in W^k$, set $(\tilde{x}^k_i, \tilde{\alpha}^k_i) = DF(x^k_i, \alpha^k_i)$. Set

$$W^{k+1} := \bigcup_{i=1}^{m} \{\tilde{x}^k_i, f(\tilde{x}^k_i), \tilde{\alpha}^k_i)\},$$

   $k := k + 1$ and go to step 4.

As regards the initial working set generation at step 3, we have chosen the following strategy: every element of W° is obtained by applying procedure DF starting from a feasible point accepted by the simulated annealing criterion.

Algorithm DDFSA enjoys an analogous property to that stated by Proposition 3. In particular, for the sake of completeness, we report the following proposition.

Proposition 4. *For every global minimum point $x^\star$ of $f(x)$ on $\mathcal{F}$ and for every $\epsilon > 0$, Step 6 of Algorithm DDFSA accepts a point $x^k \in \mathcal{B}(x^*; \epsilon) \cap \mathcal{F}$ in a finite number of iterations with probability one.*

In order to exploit the above proposition, we need a further property, which actually guarantees the existence of a gap between the global minimum value and the lowest value of $f(x)$ on a local, non global, minimum point.

Proposition 5. *Assume that $\Omega(x^\circ)$ is a compact set and that every global minimum point $x^\star$ of $f(x)$ on $\mathcal{F}$ is such that $\nabla^2 f(x^\star)$ is positive definite, then $\delta > 0$ exists such that*

$$f(x^\star) + \delta < \inf_{x \in \tilde{S}(x^\circ)} f(x),$$

where $\tilde{S}(x^\circ)$ is the set of local-nonglobal minimum points of $f(x)$ in $\Omega(x^\circ)$.

Proof. Let $S^\star$ be the set of global minimum points of $f(x)$ on $\mathcal{F}$, namely,

$$S^\star = \{x^\star \in \mathbb{R}^n : f(x^\star) \leq f(x), \forall\, x \in \mathbb{R}^n\}.$$

By assumption, for every $x^\star \in S^\star$, $f(x)$ is strictly convex within a neighborhood of $x^\star$, that is there exists a positive constant $\sigma_{x^\star}$ such that:

$$f(x) > f(x^*), \quad \|\nabla f(x)\| \neq 0, \qquad \forall x : \|x - x^\star\| < \sigma_{x^\star}, \qquad x \neq x^\star. \tag{16}$$

Now it is possible to define the following set

$$\hat{S}(x^\circ) = \bigcap_{x^\star \in S^\star} \left\{x \in \Omega(x^\circ) : \|x - x^\star\| \geq \sigma_{x^\star}\right\}.$$

The compactness assumption on the level set $\Omega(x^\circ)$ implies that also the set $\hat{S}(x^\circ)$ is compact and, hence, there exists $\hat{x} \in \hat{S}(x^\circ)$ such that:

$$f(\hat{x}) = \min_{x \in \hat{S}(x^\circ)} f(x).$$

By the definition of $\hat{S}(x^\circ)$, we can write

$$f(\hat{x}) = f(x^*) + 2\delta.$$

Since $S^\star \bigcap \hat{S}(x^\circ) = \emptyset$ and $\tilde{S}(x^\circ) \subseteq \hat{S}(x^\circ)$, we have for all $x^\star \in S^\star$ and for all $x \in \tilde{S}(x^\circ)$:

$$f(x^\star) + \delta < f(\hat{x}) \leq f(x),$$

which proves the proposition. □

The above proposition and the continuity of $f(x)$ guarantee that the set

$$\mathcal{A}(x^\star) = \left\{x \in \mathcal{F} : f(x) \leq f(x^\star) + \frac{\tilde{f} - f(x^\star)}{2}\right\},$$

where $\tilde{f} = \inf_{x \in \tilde{S}(x^\circ)} f(x)$, is not empty. Hence, under the assumptions of Proposition 5, an $\bar{\epsilon} > 0$ exists such that $\mathcal{B}(x^*; \bar{\epsilon}) \cap \mathcal{F} \subseteq \mathcal{L} \cap \mathcal{A}(x^\star)$(where $\mathcal{L}$ is the neighborhood of $x^\star$ defined in Proposition 1). Hence a point $x^k \in \mathcal{L} \cap \mathcal{A}(x^\star)$ is accepted and inserted into the working set W^k as x_i^k in a finite number of steps with probability one. Moreover, such a point x_i^k can be substituted, at Step 8 of Algorithm DDFSA, only by a point $\tilde{x}^k$ which belongs to $\mathcal{L} \cap \mathcal{A}(x^\star)$ too. Now, Proposition 1 implies that starting from such a point x_i^k the local minimization at step 9 produces a sequence of points which is attracted by the global minimum.

Table 2. Numerical results of DDFSA.

		DDFSA		
PROBLEM	n	nf	F. min	F. aver.
S-H. Camel B.	2	749	−1.0316	−1.0316
Treccani	2	999	0.106D-13	0.247D-12
Quartic	2	759	−0.3524	−0.3524
Schubert	2	1296	−186.7309	−182.9417
Schub. pen.1	2	853	−186.7309	−181.1954
Schub. pen.2	2	793	−186.7309	−183.4487
Shekel $m = 5$	4	1833	−10.1532	−10.1532
Shekel $m = 7$	4	2080	−10.4029	−10.2447
Shekel $m = 10$	4	2046	−10.5364	−10.5364
Expon.	2	431	−1.0000	−1.0000
Expon.	4	937	−1.0000	−1.0000
Cos-mix.	2	531	−0.2000	−0.2000
Cos-mix.	4	1160	−0.4000	−0.4000
Hartman	3	704	−3.8628	−3.8628
Hartman	6	1642	−3.3224	−3.3224
5n loc-min	2	615	0.329D-13	0.228D-12
5n loc-min	5	1398	0.114D-13	0.106D-12
5n loc-min	10	2447	0.462D-14	0.551D-13
5n loc-min	15	4350	0.345D-14	0.233D-13
5n loc-min	20	6623	0.242D-14	0.198D-13
5n loc-min	30	10537	0.143D-14	0.119D-13
5n loc-min	50	18643	0.105D-14	0.592D-14
5n loc-min	100	41161	0.513D-15	0.343D-14
10n loc-min	2	763	0.114D-12	0.730D-11
10n loc-min	5	1347	0.392D-13	0.182D-11
10n loc-min	10	2802	0.152D-13	0.479D-12
10n loc-min	15	5013	0.826D-14	0.415D-12
10n loc-min	20	7504	0.737D-14	0.0016
10n loc-min	30	12320	0.360D-14	0.176D-12
10n loc-min	50	21464	0.470D-14	0.742D-13
10n loc-min	100	47366	0.574D-15	0.374D-13
15n loc-min	2	676	0.234D-13	0.165D-12
15n loc-min	5	1460	0.568D-14	0.525D-12
15n loc-min	10	2656	0.241D-14	0.220D-03
15n loc-min	15	5053	0.513D-14	0.330D-03
15n loc-min	20	7800	0.159D-14	0.167D-12
15n loc-min	30	13098	0.191D-14	0.440D-03
15n loc-min	50	23072	0.347D-14	0.220D-03
15n loc-min	100	49998	0.897D-14	0.549D-03
Griewank	2	637	0.259D-14	0.0146
Griewank	5	1281	0.135D-13	0.886D-13
Griewank	10	3023	0.000D+00	0.898D-13
Griewank	15	5930	0.000D+00	0.761D-13
Griewank	20	9963	0.000D+00	0.603D-13
Griewank	30	15070	0.000D+00	0.649D-13
Griewank	50	28435	0.000D+00	0.621D-13
Griewank	100	62652	0.000D+00	0.709D-13

Table 2 reports the behavior of algorithm DDFSA on the same set of test problems used to compare DFSA and ACRS. The results are obtained by using $m = \min\{20, \max\{10, n\}\}$ as dimension of the working set W and stopping Algorithm DDFSA when $\alpha_{\text{stop}} < 10^{-6}$. The instructions of step 9 and Eq. (15) imply that this stopping condition is eventually satisfied.

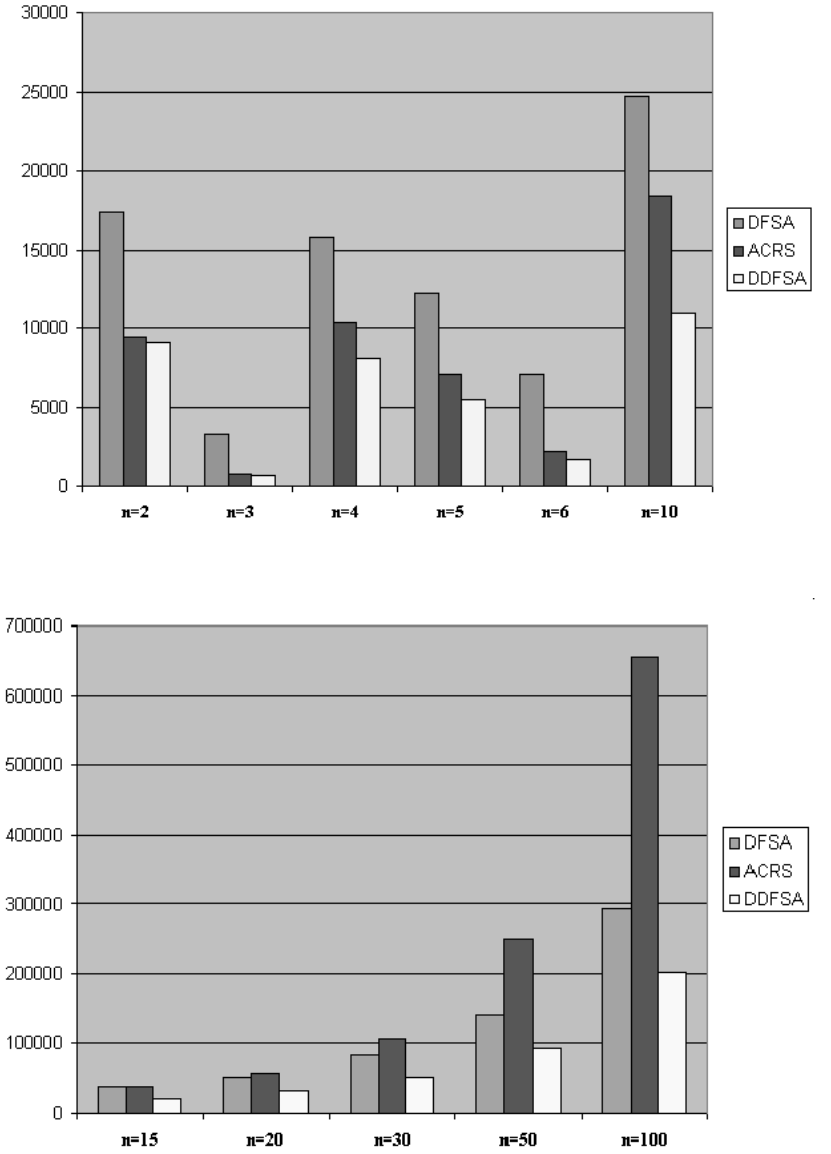


Fig. 3. Comparison among DFSA, ACRS and DDFSA in terms of function evaluations

By comparing tables 1 and 2, it emerges that the distributed algorithm DDFSA outperforms DFSA and ACRS in terms of function evaluations but retains a high accuracy of the solution as well. This superiority is clearly pointed out by Figure 3 which summarizes the behavior of the three algorithms on the test problems by reporting the total number of function evaluations needed to solve problems with the same dimension.

4. Numerical results

As a first attempt, we have tried to solve the optimal design problem by using a local minimization algorithm. In particular, we have used procedure DFA setting $\alpha_{\text{tol}} = 10^{-6}$ and taking ten different starting points randomly chosen on the feasible set. We have obtained ten different local minimizers, where the worst local minimum point is

$$\begin{aligned} x_1 &= 27.91mm & x_4 &= 181.85mm \\ x_2 &= 87.20mm & x_5 &= 235.67mm \\ x_3 &= 138.85mm & x_6 &= 296.87mm \\ x_7 &= 20.53^\circ & x_9 &= 60.58^\circ \\ x_8 &= 38.62^\circ & x_{10} &= 83.57^\circ \\ x_{11} &= 20.00mm & x_{13} &= -10.00mm \\ x_{12} &= 20.00mm & x_{14} &= -10.00mm \end{aligned}$$

which has an objective function value of $3.32 \cdot 10^{-6}$ and corresponds to a uniformity of 102 ppm for the z component, 30 ppm for the x component and 55 ppm for the y component. The best local minimum is

$$\begin{aligned} x_1 &= 25.01mm & x_4 &= 174.64mm \\ x_2 &= 72.23mm & x_5 &= 228.75mm \\ x_3 &= 123.52mm & x_6 &= 298.08mm \\ x_7 &= 18.72^\circ & x_9 &= 54.40^\circ \\ x_8 &= 36.71^\circ & x_{10} &= 77.82^\circ \\ x_{11} &= -10.00mm & x_{13} &= 5.53mm \\ x_{12} &= 20.00mm & x_{14} &= 5.31mm \end{aligned}$$

which has an objective function value of $2.51 \cdot 10^{-7}$ and corresponds to a uniformity of 28 ppm for the z component and 12 ppm for the x and y components.

The above results clearly show the need of using a global strategy to get a design with the best uniformity. To this end we have employed algorithm DDFSA (where the values of the constants are the same as stated in Section 3). DDFSA, after 19408 function evaluations, produced a point with a function value of $1.095 \cdot 10^{-7}$ which corresponds to a uniformity of 17 ppm for the z component, 10 ppm for the x component and 9 ppm for the y component of the magnetic field within the uniformity region. Since the precision obtainable by the manufacturing process is 1/100 mm and 1/1000 deg., we report the solution point appropriately rounded:

$$\begin{aligned} x_1 &= 28.97mm & x_4 &= 197.80mm \\ x_2 &= 88.79mm & x_5 &= 243.98mm \\ x_3 &= 142.95mm & x_6 &= 299.97mm \\ x_7 &= 19.165^\circ & x_9 &= 56.069^\circ \\ x_8 &= 37.869^\circ & x_{10} &= 81.225^\circ \\ x_{11} &= 13.02mm & x_{13} &= 1.45mm \\ x_{12} &= -3.79mm & x_{14} &= -8.90mm. \end{aligned}$$

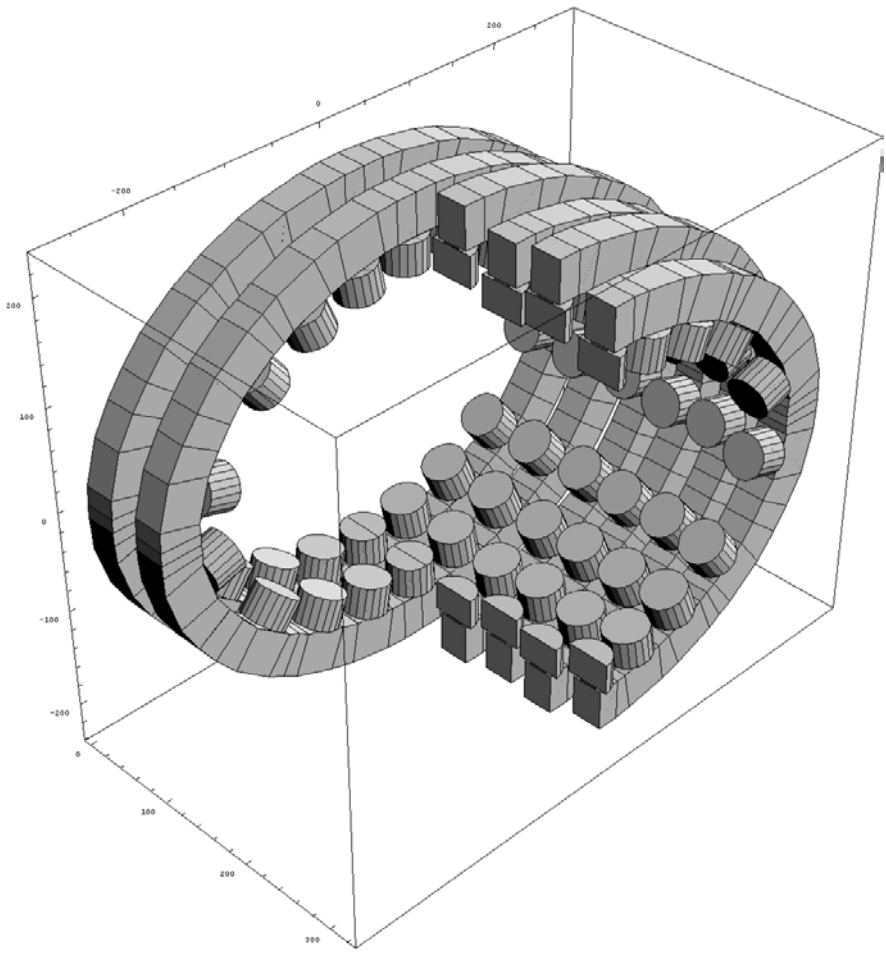


Fig. 4. Optimal design

The rounded point has an objective function value of $1.141 \cdot 10^{-7}$ which corresponds to a uniformity of 18 ppm for the z component, 11 ppm for the x component and 9 ppm for the y component of the magnetic field within the uniformity region.

The z component of the magnetic field has an average value of $\bar{B}_z = 743.5$ Gauss. In Figure 4 we report the optimal structure as determined by the optimization procedure.

The global optimization algorithm allowed us to obtain a satisfactory field uniformity. We recall that existing low-field small MR magnets present 50 ppm of field uniformity. To have an idea of the achieved improvement in the field uniformity level, we plot the field behavior on the YZ-plane. Figure 5 corresponds to the optimal structure detected by DDFSA, whereas Figure 6, corresponds to the best local minimum found by DFA, which, as already said, has a field uniformity of approximately 28 ppm.

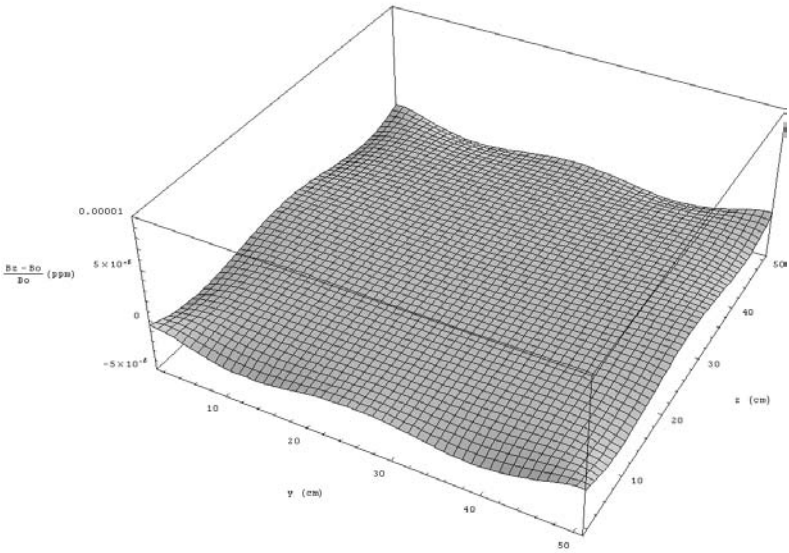


Fig. 5. Optimal configuration uniformity

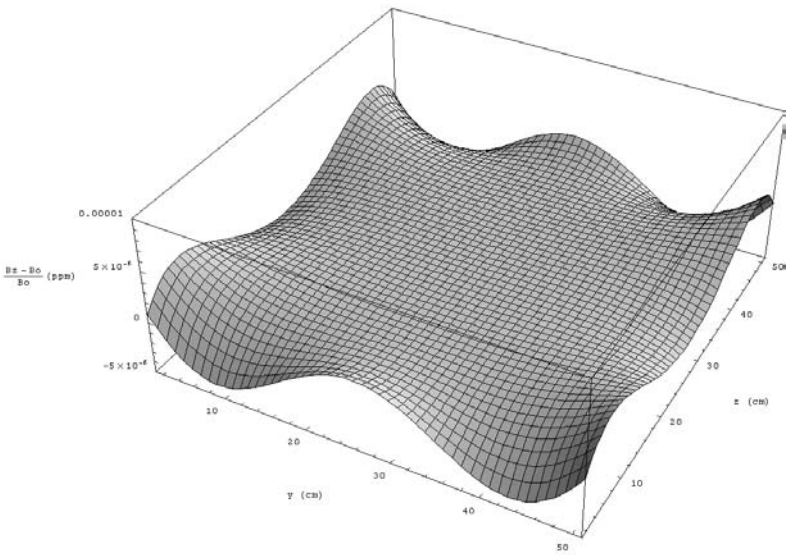


Fig. 6. 28 ppm configuration

In conclusion, we can say that Algorithm DDFSA performed well on this practical problem both in terms of number of function evaluations and in terms of the attained quality of the solution. However, we believe that further improvements are possible. In particular the following aspects of the proposed solution technique deserve a better understanding and consideration:

- a closer examination of different global searches and their combination with the derivative free local engine;
- an improvement of the local search phase aimed to decreasing the number of function evaluations;
- the definition of algorithms which can handle both continuous and integer variables. In the Magnetic Resonance problem, it would be interesting to consider as variables the number of magnets as well as the number of rings.

The above points are subject of continuing work.

Acknowledgements. The authors would like to thank two anonymous referee and the guest editor for their helpful comments, suggestions and corrections which greatly contributed in improving the paper.

5. Appendix

In this section, for the sake of completeness, we report the analytical expression of constraints $g_h(x)$, $h = 10, \ldots, 49$ of Problem (M).

Let $(r, h) = (21, 30)$, $(\bar{a}, \bar{b}) = (204, 148)$ and $\mathcal{S} = \{(\bar{a}, \bar{b}), (\bar{a} - x_i, \bar{b} - x_i), i = 11, \ldots, 14\}$. For every pair $(a, b) \in \mathcal{S}$ and every $i = 7, \ldots, 10$, let

$$c_{Y,i} = t_i \cos x_i \qquad c_{Z,i} = t_i \sin x_i,$$

with $t_i = \dfrac{ab}{\sqrt{(b \cos x_i)^2 + (a \sin x_i)^2}}$, be the coordinates (in the first quadrant) of the intersection point between an ellipse with a and b semiaxes and a line with slope $\operatorname{tg} x_i$ (as shown in figure 7). By simple geometric reasoning, we can compute the following quantities:

$$\begin{aligned}
\alpha_i &= \operatorname{arctg}\left(\frac{b^2}{a^2 \operatorname{tg} x_i}\right) \\
c^{\circ}_{Y,i} &= c_{Y,i} + h \sin \alpha_i \qquad c^{\circ}_{Z,i} = c_{Z,i} + h \cos \alpha_i \\
v^{(1)}_{Y,i} &= c_{Y,i} + r \cos \alpha_i \qquad v^{(1)}_{Z,i} = c_{Z,i} - r \sin \alpha_i \\
v^{(4)}_{Y,i} &= c_{Y,i} - r \cos \alpha_i \qquad v^{(4)}_{Z,i} = c_{Z,i} + r \sin \alpha_i,
\end{aligned}$$

where, in particular, points $(v^{(j)}_{Y,i}, v^{(j)}_{Z,i})$ for $j = 1, 2, 3, 4$ are the corners of the orthogonal projection of the i-th magnet onto the YZ-plane, that is a plane parallel to the cylinder

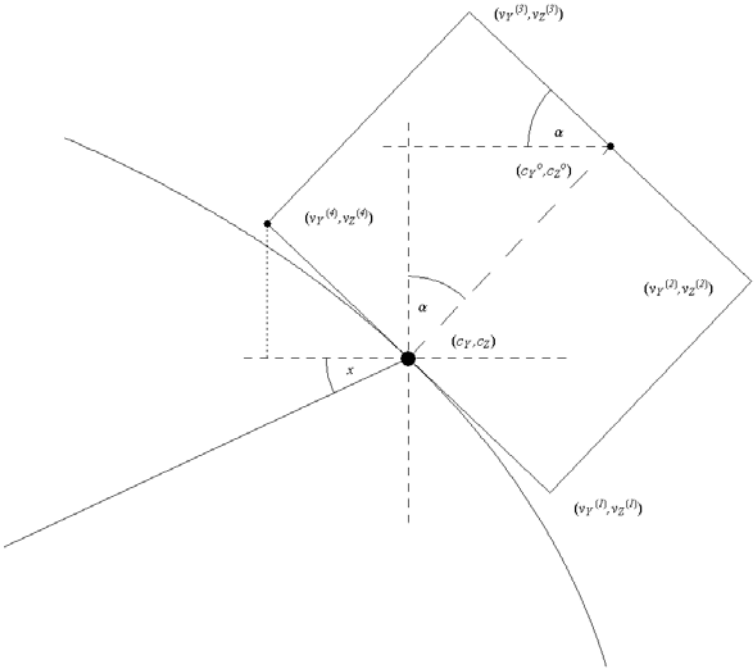


Fig. 7.

basis. These points are crucial since, as it is readily seen, two consecutive magnets do not overlap on each other if and only if their projections do not. Then, for every choice of $(a, b) \in \mathcal{S}$, we get the following non convex constraints:

$$v_{Z,i}^{(4)} - v_{Z,i-1}^{(1)} - (v_{Y,i}^{(4)} - v_{Y,i-1}^{(1)})\text{ctg}\alpha_{i-1} < 0, \tag{17}$$

$$\frac{\left[v_{Z,i}^{(4)} - v_{Z,i}^{(1)} - \text{ctg}\alpha_{i-1}(v_{Y,i}^{(4)} - v_{Y,i}^{(1)})\right]^2}{1 + \text{ctg}\alpha_{i-1}^2} - (0.5)^2 > 0, \tag{18}$$

for $i = 8, 9, 10$ and

$$v_{Y,7}^{(4)} - r > 0, \tag{19}$$

$$v_{Z,10}^{(1)} > 0. \tag{20}$$

In particular, constraints (17) guarantee that the i-th magnet lies completely on the right of the $(i-1)$-th magnet. This task is accomplished by assuring that point $(v_{Y,i}^{(4)}, v_{Z,i}^{(4)})$ lies below the line passing through point $(v_{Y,i-1}^{(1)}, v_{Z,i-1}^{(1)})$ with angular coefficient $\text{ctg}\alpha_{i-1}$. The second set of constraints, that is constraints (18), ensure that pairs of

consecutive magnets be strictly more than 0.5 millimeters apart from each other. Indeed, inequality (18) ensures that the distance between point $(v^{(4)}_{Y,i}, v^{(4)}_{Z,i})$ and the line passing through point $(v^{(1)}_{Y,i-1}, v^{(1)}_{Z,i-1})$ with angular coefficient $\mathrm{ctg}\alpha_{i-1}$ is greater than 0.5mm. Finally, constraints (19) and (20) guarantee that the first and the last magnet do not touch, respectively, the central one and the upper most magnet in the symmetric part of the ring.

In the end, we come up with a group of eight constraints (17)-(20) for every pair $(a, b) \in \mathcal{S}$, that is forty constraints which constitute $g_h(x)$, for all $h = 10, \ldots, 49$.

References

1. Bertsekas, D.P.: Constrained Optimization and Lagrange Multipliers Methods. Academic Press, New York, 1982
2. Brachetti, P., De Felice Ciccoli, M., Di Pillo, G., Lucidi, S.: A new version of the Price's algorithm for global optimization. J. Global Optim. **10**, 165–184 (1997)
3. Cho, Z., Jones, J.P., Singh, M.: Foundations of Medical Imaging. John Wiley and Sons, New York, 1993
4. Chubar, O., Elleaume, P., Chavanne, J.: A 3d magnetostatics computer code for insertion devices. J. Synchrotron Radiation **5**, 481–484 (1998)
5. Cirio, L., Lucidi, S., Parasiliti, F., Villani, M.: A global optimization approach for the synchronous motors design by finite element analysis. Int. J. Appl. Electromagnetics and Mechanics **16**, 13–27 (2002)
6. Conn, A.R., Gould, N.I.M., Toint, Ph.: Trust-region methods. MPS-SIAM Series on Optimization. SIAM publications, Philadelphia, PA, USA, 2000, p. 22
7. Czyzyk, J., Mesnier, M., Moré, J.: The neos server. IEEE J. Comput. Sci. Engin. **5**, 68–75 (1998)
8. Daidone, A., Parasiliti, F., Villani, M., Lucidi, S.: A new method for the design optimization of three-phase induction motors. IEEE Trans. Magnetics **34**, 2932–2935 (1998)
9. Dolan, E.: The neos server 4.0 administrative guide. Tech. Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, USA, 2001
10. Elleaume, P., Chubar, O., Chavanne, J.: Computing 3d magnetic field from insertion devices. In: Proceeding of the PAC97 Conference May, 1997, pp. 3509–3511
11. Fourer, R., Gay, D., Kernighan, B.: AMPL a modeling language for mathematical programming. Boyd & Fraser publishing company, Massachusetts, 1993
12. Garcia, I., Ortigosa, P.M., Casado, L.G., Herman, G.T., Matej, S.: Multidimensional optimization in image reconstruction from projections. In: Developments in Global Optimization, I. M. Bomze, T. Csendes, R. Horst, and P. Pardalos, (eds.), Kluwer, Dordrecht, 1997, pp. 289–300
13. García-Palomares, U.M., Rodríguez, J.F.: New Sequential and Parallel Derivative-Free Algorithms for Unconstrained Minimization. SIAM J. Optim. **13**, 79–96 (2002)
14. Grippo, L., Lampariello, F., Lucidi, S., Sciandrone, M.: Global convergence and stabilization of unconstrained minimization methodswithout derivatives. J. Optim. Theory Appl. **56**, 385–406 (1988)
15. Gropp, W., Moré, J.: Optimization environments and the neos server. In: Approximation Theory and Optimization, M. Buhmann and A. Iserles, (eds.), Cambridge University Press, 1997, pp. 167–182
16. Haacke, E.M., Brown, R.W., Thompson, M.R., Vankatesan, R.: Magnetic Resonance Imaging: Physical Principles and Sequence Design. John Wiley and Sons, New York, 1999
17. Hendrix, E., Ortigosa, P., Garcia, I.: On success rates for controlled random search. J. Global Optim. **21**, 239–263 (2001)
18. Liang, Z., Lauterbur, P.C.: Principles of Magnetic Resonance Imaging: A Signal Processing Approach. IEEE Press, 2000
19. Liuzzi, G., Lucidi, S., Placidi, G., Sotgiu, A.: A magnetic resonance device designed via global optimization techniques. TR 09-02, Department of Computer and System Sciences "A. Ruberti", University of Rome "La Sapienza", Rome, Italy, 2002
20. Lucidi, S., Piccioni, M.: Random tunneling by means of acceptance-rejection sampling for global optimization. J. Optim. Theory Appl. **62**, 255–279 (1989)
21. Lucidi, S., Sciandrone, M.: On the global convergence of derivative free methods for unconstrained optimization. SIAM J. Optim. **13**, 97–116 (2002)
22. Lucidi, S., Sciandrone, M.: A derivative-free algorithm for bound constrained optimization. Comput. Optim. Appl. **21** (2), 119–142 (2002)
23. Nishimura, D.G.: Magnetic Resonance Imaging. MRSRL Press
24. Price, W.L.: Global optimization algorithms for a CAD workstation. J. Optim. Theory Appl. **55**, 133–146 (1983)

25. Price, W.L., Woodhams, F.: Optimising accelerator for CAD workstations. IEEE Proc. **135**, 214–221 (1988)
26. Solis, F.J.-B., Wets, R.J.: Minimization by random search techniques. Math. Oper. Res. **6**, 19–30 (1981)
27. Wolfram Research Inc.: Mathematica, http://www.wolfram.com/, 2003