Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2008, Article ID 369125, 15 pages
doi:10.1155/2008/369125

Research Article

Adjoint Variable Method for the Study of Combined Active and Passive Magnetic Shielding

Peter Sergeant,[1,2] Ivan Cimrák,[3] Valdemar Melicher,[3] Luc Dupré,[1] and Roger Van Keer[3]

[1] *Department of Electrical Energy, Systems and Automation, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium*
[2] *Department of Electrotechnology, Faculty of Applied Engineering Sciences, University College Ghent, Schoonmeersstraat 52, 9000 Gent, Belgium*
[3] *NfaM^2 Research Group, Department of Mathematical Analysis, Ghent University, Galglaan 2, 9000 Gent, Belgium*

Correspondence should be addressed to Peter Sergeant, peter.sergeant@ugent.be

Received 18 October 2007; Accepted 28 January 2008

Recommended by Alois Steindl

For shielding applications that cannot sufficiently be shielded by only a passive shield, it is useful to combine a passive and an active shield. Indeed, the latter does the "finetuning" of the field reduction that is mainly caused by the passive shield. The design requires the optimization of the geometry of the passive shield, the position of all coils of the active shield, and the real and imaginary components of the currents (when working in the frequency domain). As there are many variables, the computational effort for the optimization becomes huge. An optimization using genetic algorithms is compared with a classical gradient optimization and with a design sensitivity approach that uses an adjoint system. Several types of active and/or passive shields with constraints are designed. For each type, the optimization was carried out by all three techniques in order to compare them concerning CPU time and accuracy.

1. Introduction

As the computational cost has decreased significantly during the last years, more and more commercial software packages offer standard algorithms for optimization in order to solve design problems. Many researchers in industry and in academia now use commercial finite-element software in combination with commercial optimization routines, such as genetic algorithms. This approach is usually successful but very time consuming. For example, the design of a combined active and passive shield in [1] required between 8 and 14 days of CPU time.

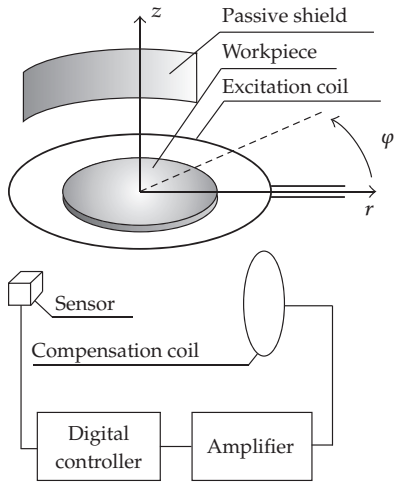

Figure 1: The axisymmetric induction heater application with shields discussed in [1].

In spite of the increased computational power of computers, the conventional combination of genetic algorithm and finite elements remains a computational challenge.

This paper is devoted to designing a passive and active shield for the same axisymmetric induction heater discussed in [1] and shown in Figure 1 in much less time. Therefore, not only another algorithm is explored (adjoint variable method, explained in Section 3), but also the optimization problem itself is changed in order to take advantage of the adjoint system. The optimization technique is similar to [2, 3] and allows the designer to use commercial software. The numerical model is quasistatic (and not static like in [2, 3]), which results in complex equations in the frequency domain (instead of real equations), and is based on Maxwell's equations. Both currents and geometrical shapes are optimized. Moreover, the optimization is multiobjective, and all of the five contributions to the cost are modelled in the same adjoint system. After validating the adjoint variable approach in Section 4, the method is compared in Section 5 with a classical gradient algorithm and with a genetic algorithm regarding the computation time and the efficiency of the shield.

2. Magnetic shielding problem

A magnetic shield is developed to reduce the magnetic stray field of an axisymmetric induction heating device for the heat treatment of aluminum discs. After adding the shield, the shielded device should comply with the reference levels of ICNIRP [4] or the European Community [5]. Next to the field reduction in a predefined "target area," several other constraints are involved with a magnetic shielding problem: limits on the resistive heating in the shields and on the influence of the heating process, development costs, and geometrical constraints to guarantee the accessibility of the shielded device. In the literature, many papers are published concerning passive shields [6], studying the effect of dimensions, permeability (including nonlinearity, hysteresis), and conductivity of the metal sheets on the shielding effectiveness. A number of papers were written about active shielding [7–9]: field reduction caused by a counter field that is produced by controlling the currents in a number of compensation coils.

Passive shields can be very efficient especially at frequencies above 1 kHz where induced currents become effective, especially if the passive shields are closed, that is, if they completely enclose the source like Faraday's cage. Active shields can be more attractive at lower frequencies especially if the compensation coils are close to the source, because the generator of the compensation current can have a lower voltage and current rating. However, a lot of devices cannot easily be shielded by only a passive or only an active shield because of geometrical constraints. It is shown in [1] that a combination of both types of shielding may improve the performance significantly. Here, the passive shield should be rather close to the source. By flux shunting or by induced currents, the passive shield causes a significant field reduction in the target area that is usually not homogeneous: a high field reduction is achieved in parts of the target region, but still high fields occur in other parts. Here, active shields can additionally reduce the field in these regions.

The equations to solve are Maxwell's equations in the frequency domain:

$$\begin{aligned} \nabla \times \mathbf{E} &= -\mathrm{j}\omega \mathbf{B}, \\ \nabla \times \mathbf{H} &= \mathbf{J}_\mathrm{e} + \mathbf{J}_\mathrm{i} = \mathbf{J}_\mathrm{e} + \sigma \mathbf{E}, \end{aligned} \tag{2.1}$$

where the total current density consists of the external current density $\mathbf{J}_\mathrm{e}$ and the induced current density $\mathbf{J}_\mathrm{i}$ due to the conductivity σ and the electric field $\mathbf{E}$. The system of (2.1) is solved together with the constitutive law $\mathbf{B} = \mu\mathbf{H}$ and with the following boundary conditions (see Figure 2):

$$\begin{aligned} \mathbf{E} \times \mathbf{n} &= 0 \quad \text{on } \Gamma_1, \\ \nabla \times \mathbf{E} \times \mathbf{n} &= 0 \quad \text{on } \Gamma_2, \end{aligned} \tag{2.2}$$

expressing the fact that the electric field should be perpendicular to Γ_1 and that the magnetic field should be perpendicular to Γ_2. Introducing the vector potential so that $\mathbf{B} = \nabla \times \mathbf{A}$ leads to

$$\begin{aligned} \nabla \times \mathbf{E} &= -\mathrm{j}\omega(\nabla \times \mathbf{A}), \\ \nabla \times \left[\frac{1}{\mu}(\nabla \times \mathbf{A})\right] &= \mathbf{J}_\mathrm{e} + \sigma \mathbf{E}. \end{aligned} \tag{2.3}$$

Substitution of the electric field yields the equation to be solved by a finite-element model (FEM) [10]:

$$\nabla \times \left[\frac{1}{\mu}(\nabla \times \mathbf{A})\right] + \mathrm{j}\omega\sigma\mathbf{A} = \mathbf{J}_\mathrm{e} \tag{2.4}$$

which is written in weak formulation with test variable φ:

$$\left(\frac{1}{\mu(\mathbf{p})}\nabla \times \mathbf{A}(\mathbf{p}), \nabla \times \varphi\right) + \left(\mathrm{j}\omega\sigma(\mathbf{p})\mathbf{A}(\mathbf{p}), \varphi\right) = \left(\mathbf{J}_\mathrm{e}(\mathbf{p}), \varphi\right), \tag{2.5}$$

taking into account the boundary conditions. The variable $\mathbf{p}$ contains all parameters to be optimized: the position r_p and height h_p of the passive shield with thickness t_p in steel, and the (complex) currents $I_{\mathrm{r}i} + \mathrm{j}I_{\mathrm{i}i}$ and positions (r_i, z_i), $i = 1 \cdots n$ of the n active shield coils in the axisymmetric problem. Symbol $(\cdot, \cdot)$ denotes the standard scalar product defined by

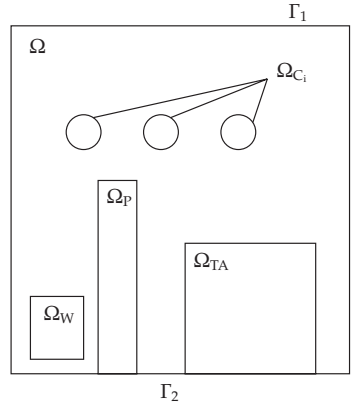

Figure 2: Domain Ω. Γ_2 is the bottom horizontal border of Ω and Γ_1 is the rest of the border of Ω; Ω_W is the workpiece; Ω_P is the passive shield; Ω_{TA} is the target area; Ω_{C_i} are the compensation coils of the active shield.

$(\mathbf{u}, \mathbf{v}) = \int_\Omega \mathbf{u}\cdot\mathbf{v}\,d\Omega$. Notice that μ and σ depend on the position and size of the passive shield. Therefore, they both depend on $\mathbf{p}$. The weak formulation has the advantage that μ and σ can be discontinuous quantities.

The shielding application that we study is axisymmetric so that the vector potential can be written as $\mathbf{A} = 0\,\mathbf{1}_r + A\,\mathbf{1}_\phi + 0\,\mathbf{1}_z$. Here, A is the azimuthal component of the vector potential of which the r- and z-components are zero. Maxwell's equations in the frequency domain (2.5) are in the axisymmetric case:

$$\left(\frac{1}{\mu(\mathbf{p})}\nabla A(\mathbf{p}), \nabla\varphi\right) + \left(\frac{1}{r\mu(\mathbf{p})}A(\mathbf{p}), \frac{\partial\varphi}{\partial r}\right) + (\mathrm{j}\omega\sigma(\mathbf{p})A(\mathbf{p}), \varphi) = (J_e(\mathbf{p}), \varphi). \tag{2.6}$$

On the boundary (see Figure 2) we prescribe

$$\begin{aligned} A &= 0 \quad \text{on } \Gamma_1, \\ \nabla A\cdot\mathbf{n} &= 0 \quad \text{on } \Gamma_2, \end{aligned} \tag{2.7}$$

where $\nabla A := (\partial A/\partial r, \partial A/\partial z)$.

The objective value is a function of the scalar potential A and of ∇A. First of all, it takes into account the magnetic field in the target area (F_B): the main objective is to minimize the magnetic induction $\mathbf{B} = \nabla\times\mathbf{A}$ in the reference region Ω_{TA}. This can be achieved by minimizing the cost functional

$$\begin{aligned} F(\mathbf{p}) &= \frac{1}{2}\int_{\Omega_{TA}} |\mathbf{B}|^2\,d\Omega \\ &= \frac{1}{2}\int_{\Omega_{TA}} |\nabla\times\mathbf{A}|^2\,d\Omega \\ &= \frac{1}{2}\int_{\Omega_{TA}} \left(|\nabla A|^2 + 2\frac{A}{r}\frac{\partial A}{\partial r} + \frac{A^2}{r^2}\right)d\Omega. \end{aligned} \tag{2.8}$$

In order to reduce the complexity of the cost function, we ignore the last terms and choose $(1/2)\int_{\Omega_{TA}} |\nabla A|^2\,d\Omega$ as cost value. Other objective values in the multiobjective optimization are

the dissipation in the passive shield (F_P), the dissipation in the active shield (F_A), the change of the heating of the workpiece by the adding of shields (F_W), and the volume of the passive shield as an investment cost (F_V):

$$F(\mathbf{p}) = w_1 F_\mathrm{B} + w_2 F_\mathrm{P} + w_3 F_\mathrm{A} + w_4 F_\mathrm{W} + w_5 F_\mathrm{V}, \text{ wherein} \tag{2.9}$$

$$F_\mathrm{B} = \frac{1}{2}\left\|\nabla A(\mathbf{p})\right\|^2_{\Omega_{\mathrm{TA}},} \tag{2.10}$$

$$F_\mathrm{P} = \frac{1}{2}\left\|\mathbf{J}_\mathrm{i}\cdot\mathbf{E}\right\|_{\Omega_\mathrm{P}} = \frac{1}{2}\sigma\omega^2\left\|A(\mathbf{p})\right\|^2_{\Omega_{\mathrm{P},}} \tag{2.11}$$

$$F_\mathrm{A} = \frac{1}{2}\,\frac{2\pi r\rho}{S_\mathrm{A}}\sum_{i=1}^{n}\left(I^2_{\mathrm{r},i} + I^2_{\mathrm{i},i}\right), \tag{2.12}$$

$$F_\mathrm{W} = P_{\mathrm{W},0} - \frac{1}{2}\left\|\mathbf{J}_\mathrm{i}\cdot\mathbf{E}\right\|_{\Omega_\mathrm{W}} = P_{\mathrm{W},0} - \frac{1}{2}\sigma\omega^2\left\|A(\mathbf{p})\right\|^2_{\Omega_{\mathrm{W},}} \tag{2.13}$$

$$F_\mathrm{V} = 2\pi r_\mathrm{p} h_\mathrm{p} t_\mathrm{p}. \tag{2.14}$$

Herein, S_A is the cross-section of an active shield coil and ρ is the resistivity of the coil material. The constant $P_{\mathrm{W},0}$ is the power dissipated in the workpiece without shields present. The weighting factors w_k, $k = 1,\ldots,5$ should be chosen as explained in Section 4.

3. Adjoint system

In order to use any gradient-based optimization technique, one has to evaluate the derivative of $F(\mathbf{p})$ with respect to $\mathbf{p}$ in some direction $\boldsymbol{\zeta}$. For simplicity of the exposition, we use the symbol δ defined as

$$\delta u = \frac{\partial u}{\partial \mathbf{p}}\boldsymbol{\zeta} = \lim_{\epsilon\to 0}\frac{u(\mathbf{p}+\epsilon\boldsymbol{\zeta}) - u(\mathbf{p})}{\epsilon}. \tag{3.1}$$

Formal derivation of the cost functional gives us

$$\delta F = w_1\delta F_\mathrm{B} + w_2\delta F_\mathrm{P} + w_3\delta F_\mathrm{A} + w_4\delta F_\mathrm{W} + w_5\delta F_\mathrm{V}, \tag{3.2}$$

where particular derivatives read as

$$\begin{aligned}
\delta F_\mathrm{B} &= \frac{1}{2}\left[\left(\nabla\delta A^*, \nabla A(\mathbf{p})\right)_{\Omega_{\mathrm{TA}}} + \left(\nabla\delta A, \nabla A^*(\mathbf{p})\right)_{\Omega_{\mathrm{TA}}}\right],\\
\delta F_\mathrm{P} &= \frac{1}{2}\delta\sigma\omega^2\left\|A(\mathbf{p})\right\|^2_{\Omega_\mathrm{P}} + \frac{1}{2}\sigma\omega^2\left[\left(\delta A^*, A(\mathbf{p})\right)_{\Omega_\mathrm{P}} + \left(\delta A, A^*(\mathbf{p})\right)_{\Omega_\mathrm{P}}\right]\\
&\quad + \frac{1}{2}\sigma\omega^2\int_{\delta\Omega_\mathrm{P}} A(\mathbf{p})\cdot A^*(\mathbf{p})\,\mathrm{d}(\delta\Omega),\\
\delta F_\mathrm{A} &= \frac{2\pi r\rho}{S_\mathrm{A}}\sum_{i=1}^{n}\delta I_{\mathrm{r},i} I_{\mathrm{r},i} + \frac{2\pi r\rho}{S_\mathrm{A}}\sum_{i=1}^{n}\delta I_{\mathrm{i},i} I_{\mathrm{i},i} + \delta r\frac{\pi\rho}{S_\mathrm{A}}\sum_{i=1}^{n}\left(I^2_{\mathrm{r},i} + I^2_{\mathrm{i},i}\right),\\
\delta F_\mathrm{W} &= -\frac{1}{2}\delta\sigma\omega^2\left\|A(\mathbf{p})\right\|^2_{\Omega_\mathrm{W}} - \frac{1}{2}\sigma\omega^2\left[\left(\delta A^*, A(\mathbf{p})\right)_{\Omega_\mathrm{W}} + \left(\delta A, A^*(\mathbf{p})\right)_{\Omega_\mathrm{W}}\right],\\
\delta F_\mathrm{V} &= 2\pi r_\mathrm{p} t_\mathrm{p}\delta h_\mathrm{p}.
\end{aligned} \tag{3.3}$$

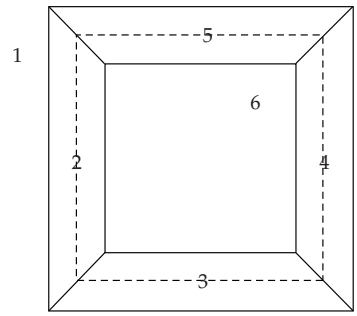

Figure 3: Layout of a domain of which the position or shape can change.

The symbol "$*$" denotes the complex conjugate. The extraboundary integral in δF_{P} is necessary because the domain Ω_{P}, on which F_{P} is evaluated, changes in size when the height of the passive shield changes. Here, we have to evaluate the topological derivative, which depends on concrete settings. In our case of a rectangular passive shield domain between r-coordinates (a, b) and z-coordinates $(0, h_{\mathrm{p}})$, where we change just one dimension, height h_{p}, we have to evaluate

$$\lim_{\epsilon\to 0} \frac{\left(\int_0^{h+\epsilon}\int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y - \int_0^h\int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y\right)}{\epsilon} = \int_a^b \lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_h^{h+\epsilon} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_a^b f(x,h)\,\mathrm{d}x. \tag{3.4}$$

We used here the mean value theorem. Thus in our case,

$$\int_{\delta\Omega_{\mathrm{P}}} A(\mathbf{p})\cdot A^*(\mathbf{p})\,\mathrm{d}\delta\Omega = \int_{\Gamma_{\mathrm{P}}} A(\mathbf{p})\cdot A^*(\mathbf{p})\,\mathrm{d}\delta\Omega, \tag{3.5}$$

where Γ_{P} is the upper boundary of Ω_{P}.

The derivative $\delta\sigma$ is used to handle the variation of domain size, considering a transition region. For $\delta\sigma$, one can use an explicit expression: instead of modelling one solid domain with high conductivity (dashed line in Figure 3), we model next to the domain with high conductivity a small transition domain (the domains 2-3-4-5 in solid line in Figure 3) that surrounds the domain with high σ. In this transition zone, the conductivity decreases linearly from σ to zero. The derivative $\delta\sigma$ differs from zero in this transition domain only.

For δA, however, one needs to set up corresponding PDE. For clarity, the dependence of $\mathbf{p}$ is not explicitly written any more in the following equations. Differentiation of (2.6) with respect to $\mathbf{p}$ results in

$$\begin{aligned}&\left((\mu)^{-1}\nabla\delta A, \nabla\varphi\right) + \left(\mathrm{j}\omega\sigma\delta A, \varphi\right) + \left((\mu r)^{-1}\delta A, \frac{\partial\varphi}{\partial r}\right)\\ &\quad = \left(\delta J_{\mathrm{e}}, \varphi\right) - \left(\delta\left(\mu^{-1}\right)\nabla A, \nabla\varphi\right) - \left(\mathrm{j}\omega\delta\sigma A, \varphi\right) - \left(r^{-1}\delta\left(\mu^{-1}\right)A, \frac{\partial\varphi}{\partial r}\right).\end{aligned} \tag{3.6}$$

We obtain a PDE for δA. This is, however, not so practical. For the implementation, one needs to know a partial derivative in one component of $\mathbf{p}$. Imagine $\mathbf{p} = (a, b, c)$ consists of three

components. So for computations we need to know three values $\partial A(\mathbf{p})/\partial a$, $\partial A(\mathbf{p})/\partial b$, and $\partial A(\mathbf{p})/\partial c$, that is, we need to compute δA for three different values of $\boldsymbol{\zeta}$, namely, $\boldsymbol{\zeta}_1 = (1,0,0)$, $\boldsymbol{\zeta}_2 = (0,1,0)$, and $\boldsymbol{\zeta}_3 = (0,0,1)$. This is, however, time-consuming since every computation of δA means solving one PDE.

To solve this problem, we apply the so-called adjoint variable method. This method was successfully used in the optimal design problem in micromagnetism involving the shape optimization of the ferromagnetic core in MRAM memories [11]. We set up a dual problem to (3.6). Find $\xi \in H_{0,\Gamma_1}(\Omega)$ such that for all $\varphi \in H_{0,\Gamma_1}(\Omega)$, the following equality holds:

$$\left((\mu)^{-1}\nabla\varphi, \nabla\xi\right) + \left(\mathrm{j}\omega\sigma\varphi, \xi\right) + \left((\mu r)^{-1}\varphi, \frac{\partial\xi}{\partial r}\right) = a(\varphi), \tag{3.7}$$

where the functional $a(\varphi)$ is composed of those terms from δF that contain δA, namely,

$$\begin{aligned} a(\varphi) = {} & \frac{w_1}{2}\left[\left(\nabla\varphi^*, \nabla A\right)_{\Omega_{\mathrm{TA}}} + \left(\nabla\varphi, \nabla A^*\right)_{\Omega_{\mathrm{TA}}}\right] + \frac{w_2}{2}\sigma\omega^2\left[\left(\varphi^*, A\right)_{\Omega_{\mathrm{P}}} + \left(\varphi, A^*\right)_{\Omega_{\mathrm{P}}}\right] \\ & - \frac{w_4}{2}\sigma\omega^2\left[\left(\varphi^*, A\right)_{\Omega_{\mathrm{W}}} + \left(\varphi, A^*\right)_{\Omega_{\mathrm{W}}}\right]. \end{aligned} \tag{3.8}$$

The functional $a(\varphi)$ serves as a pseudosource for the adjoint problem derived from the cost functional. Since the terms in the cost functional are nonzero only on the corresponding subdomains $\Omega_{\mathrm{TA}}, \Omega_{\mathrm{P}}$, and Ω_{W}, so is the pseudosource. Notice that the cost terms of the active shield dissipation F_{A} and the volume of the passive shield (F_{V}) in (2.9) do not result in a pseudosource in the dual problem. Although (3.7) and (3.8) represent only one adjoint system, its solution requires two finite-element evaluations. Indeed, the test functions φ are real-valued functions in the commercial software we use while they should be complex in order to solve (3.7) and (3.8). In practice, the adjoint system is solved a first time using only the real parts of the pseudosources. Afterwards, it is solved a second time using the imaginary parts of the pseudosources. This means that in total three finite-element calculations are needed to obtain the objective value and all gradients. Possibly, as the two adjoint calculations are independent from each other, they can be executed in parallel on appropriate computer architectures.

Knowing the solution ξ of the dual problem (3.7), one can settle the following result: by setting $\varphi = \delta A$ in (3.7) and by setting $\varphi = \xi$ in (3.6) we arrive at

$$a(\delta A) = \left(\delta J_{\mathrm{e}}, \xi\right) - \left(\delta\left(\mu^{-1}\right)\nabla A, \nabla\xi\right) - \left(\mathrm{j}\omega\delta\sigma A, \xi\right) - \left(r^{-1}\delta\left(\mu^{-1}\right)A, \frac{\partial\xi}{\partial r}\right). \tag{3.9}$$

When comparing $a(\delta A)$ obtained by setting $\varphi = \delta A$ in (3.8) with the derivative of the cost (3.3), it can be observed that all terms occurring in (3.8) also occur in (3.3). By adding the missing terms, we obtain the explicit expression for the derivative of the cost:

$$\begin{aligned} \delta F = {} & \left(\delta J_{\mathrm{e}}, \xi\right) - \left(\delta\left(\mu^{-1}\right)\nabla A, \nabla\xi\right) - \left(\mathrm{j}\omega\delta\sigma A, \xi\right) - \left(r^{-1}\delta\left(\mu^{-1}\right)A, \frac{\partial\xi}{\partial r}\right) \\ & + \frac{w_2}{2}\delta\sigma\omega^2\|A\|^2_{\Omega_{\mathrm{P}}} + \frac{w_2}{2}\sigma\omega^2\int_{\delta\Omega_{\mathrm{P}}} A\cdot A^*\,\mathrm{d}(\delta\Omega) \\ & + \frac{w_3 2\pi r\rho}{S_{\mathrm{A}}}\sum_{i=1}^{n}\left(\delta I_{\mathrm{r},i}\right)I_{\mathrm{r},i} + \frac{w_3 2\pi r\rho}{S_{\mathrm{A}}}\sum_{i=1}^{n}\left(\delta I_{\mathrm{i},i}\right)I_{\mathrm{i},i} + \delta r\frac{w_3\pi\rho}{S_{\mathrm{A}}}\sum_{i=1}^{n}\left(I^2_{\mathrm{r},i} + I^2_{\mathrm{i},i}\right) \\ & - \frac{w_4}{2}\delta\sigma\omega^2\|A\|^2_{\Omega_{\mathrm{W}}} + w_5 2\pi r_{\mathrm{p}}t_{\mathrm{p}}\delta h_{\mathrm{p}}. \end{aligned} \tag{3.10}$$

1. $i = 0$ and estimate the initial value $\mathbf{p}_0$.
2. Compute $A(\mathbf{p}_i)$ from the direct problem (2.5) setting $\mathbf{p} = \mathbf{p}_i$.
3. Compute ξ from the dual problem (3.7) using $A(\mathbf{p}_i)$ instead of $A(\mathbf{p})$.
4. Using explicit expression (3.10) for the derivative, compute $\partial F(\mathbf{p}_i)/\partial \mathbf{p}$.
5. By a line search algorithm determine the optimal step α_i.
6. Compute $\mathbf{p}_{i+1} = \mathbf{p}_i - \alpha_i(\partial F(\mathbf{p}_i)/\partial \mathbf{p})$.
7. If convergence is reached (controlling e.g., the difference $F(\mathbf{p}_i) - F(\mathbf{p}_{i+1})$) then stop, otherwise set $i := i + 1$ and go to step 2 of the algorithm.

Algorithm 1: Gradient-adjoint algorithm.

With this explicit expression for the derivative of $F(\mathbf{p})$ with respect to $\mathbf{p}$, we are able to evaluate δF for any $\boldsymbol{\zeta}$. We just need to solve the PDE for ξ.

The actual minimization algorithm starts from some initial guess, for example, $\mathbf{p}_0 = 0$ or other value obtained, for example, from a genetic algorithm.

The gradient-adjoint optimization algorithm can in principle be seen as a classical gradient algorithm that tries to minimize a cost function by iteratively evaluating a finite-element model (step 2). There is, however, one major difference: the gradients are obtained by solving another (adjoint) finite-element model (step 3) and by using the solution in an expression for the gradient (step 4). In the classical algorithm, the gradients are determined by applying perturbations in the parameters.

4. Validation of the adjoint variable method

The studied application to design shields for is an axisymmetric induction heating device with an excitation coil of 0.2012 m radius carrying a current of 4000 A at 1 kHz. The workpiece of 10 mm height and 191 mm radius is made from aluminum.

We now add a passive shield of 0.65 mm thickness at radial position 0.3 m. Figure 4 shows the several cost terms from (2.9) using the weighting factors $w_1 = 4 \times 10^9$, $w_2 = 8 \times 10^{-3}$, $w_3 = 2 \times 10^{-4}$, $w_4 = 2 \times 10^{-2}$, and $w_5 = 10^4$ as a function of the passive shield height. The choice of the weighting factors depends on the relative importance that is assigned to each cost term. No active shield was considered in this situation ($F_A = 0$). It can be seen that the term F_B related to the magnetic field in the target area decreases strongly with increasing height of the passive shield, which is a purely conductive shield with $\mu_r = 1$ and $\sigma = 5.9 \times 10^6$ S/m. The losses caused by induced currents in this shield F_P increase with the height for rather low shields with $h_p < 0.1$ m, but remain almost constant for $h_p > 0.1$ m. The cost F_W increases with the height of the passive shield: as the shield is conductive, it reduces the flux produced by the excitation current. Consequently, the heating induced in the workpiece decreases, which causes the cost F_W.

Figure 5 illustrates the derivatives to the parameter h_p, calculated on the one hand by the adjoint method and on the other hand by the conventional approach: applying a perturbation of 2 mm in the height of the shield, re-evaluating the cost, and approximating the derivative by dividing the subtraction of both cost values by the perturbation step. The correspondence is good for this conductive shield. However, the finite-element mesh in the 1.25 mm wide and 2 mm high transition domain where the relevant terms of (3.10) are integrated had a maximal

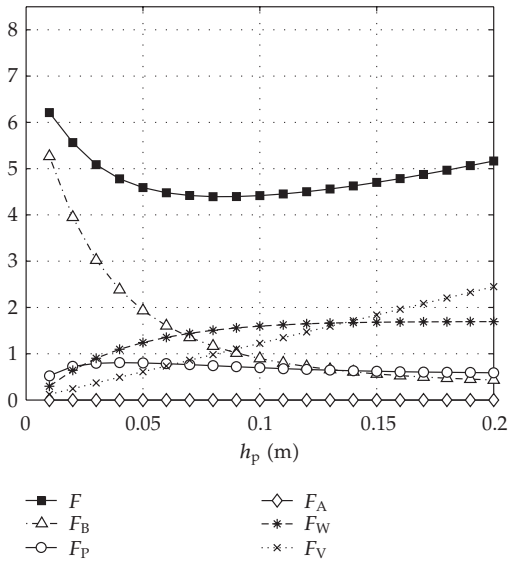

Figure 4: Several contributions to the objective value F as a function of the passive shield height (no active shield).

edge size of 0.1 mm: much finer than needed to obtain accurate gradients by the conventional method (0.25 mm).

In order to illustrate the limits of the adjoint method, we compared the conventional and adjoint gradients for several permeabilities in Figure 6. In the expression for the adjoint gradients (3.10), the term $(\delta(\mu^{-1})\nabla A, \nabla\xi)$ appears. For high permeability in the shield, the coefficient $\delta(\mu^{-1}) = -\mu^{-2}\mathrm{d}\mu/\mathrm{d}h_p$ is very small in the region of the transition zone near the shield because here the permeability is high and μ^{-2} is very small. In the region of the transition zone near the air, the relative permeability is close to 1 and the term is much larger. The mesh refinement for sufficient accuracy is here much more critical than for a conductive shield. As can be observed in Figure 6 for $\mu_r = 372$, the position at which the adjoint gradient crosses zero—important to find the correct solution in optimization problems—is about 30% too low.

Figure 7 depicts the gradients as a function of the current in an active shield coil and as a function of the position of such a coil. For the amplitude of the current, the accuracy is comparable to the one of classical derivatives, even for a coarse mesh. The first of the three main reasons for the better accuracy compared to δh_p of the passive shield is that the shape is square in contrast to the shape of the passive shield, which is a very small and high rectangle. Secondly, the term δJ_e is constant and not quadratic like $\delta(\mu^{-1})$. Thirdly, the quantities A and ξ to integrate on the active shield coil transition domain are much more smooth than the quantities ∇A and $\nabla\xi$ to integrate on the passive shield transition domain. For the position, however, the difference is larger (up to 20% as can be seen in Figure 7(b)). The reason for this is that the calculation of the gradient requires the subtraction of integrals in two domains: the domains 2 and 4 in Figure 3 if the object is an active shield coil that moves in horizontal direction. The integrals have the same order of magnitude, causing the difference to be sensitive to numerical noise.

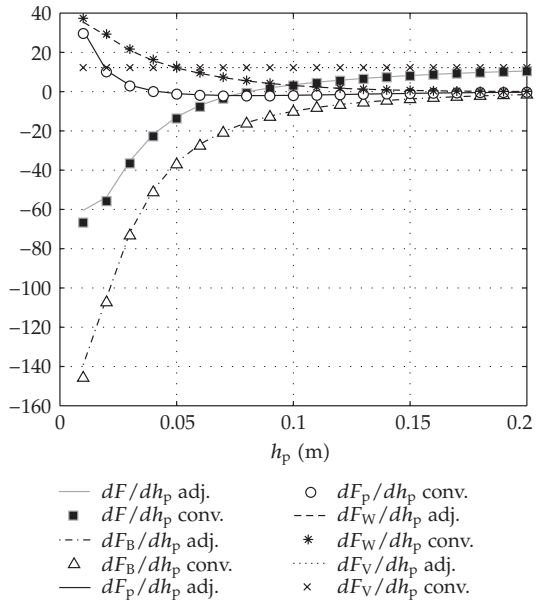

Figure 5: Conventional and adjoint gradient of all cost terms separately as a function of the height of the passive shield for $\mu_r = 1$ and $\sigma = 5.9 \times 10^6$ S/m (no active shield).

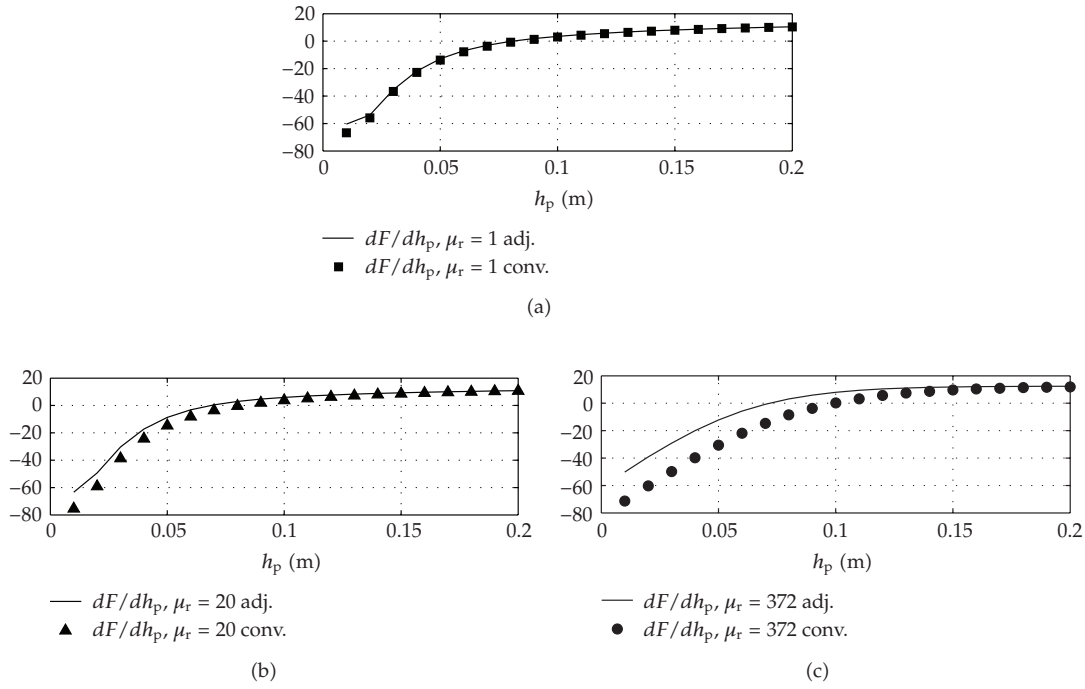

Figure 6: Total conventional and adjoint gradients as a function of the height of the passive shield for several μ_r and $\sigma = 5.9 \times 10^6$ S/m (no active shield).

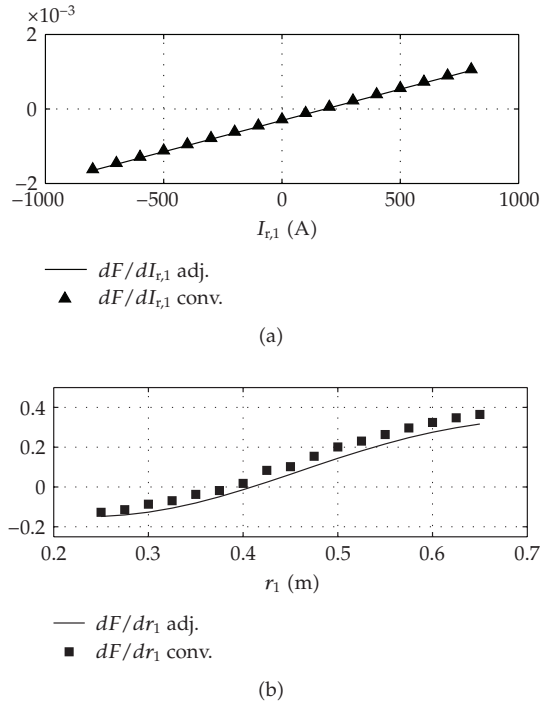

Figure 7: Total conventional and adjoint gradients as a function (a) of the current in a coil and (b) as a function of the position of a coil.

5. Comparison of optimization techniques for the shielding problem

The parameters to optimize are the height of the passive shield (between 10 and 200 mm), the horizontal position of each compensation coil (between 0.250 and 0.900 m and $r_{i+1} > r_i$+0.06 m), and the real and imaginary component of the compensation current in each coil (between −400 and +400 A, which is 10% of the excitation current). The radius of the passive shield (0.3 m) and the vertical position of the compensation coils (1.15 m) were not optimized. For all optimizations described below, genetic algorithms are compared with the conventional gradient method and with the gradient method that uses the adjoint system. The efficiency of several designed shields is also compared with the efficiency of the shields in [1] for the same induction heating application. Genetic algorithms are stochastic optimization routines that iteratively find a global optimum by applying selection, mutation, and recombination (cross-over) techniques [12] on a "population." The initial population consists of individuals whose parameters are randomly chosen within the boundaries.

The classical gradient algorithm finds the solution of the constrained nonlinear optimization problem by using sequential quadratic programming (SQP) [13]. The minimum is found using numerical approximations of the Jacobian and the Hessian. Each element of the Jacobian is determined by applying a small perturbation in one parameter and re-evaluating the cost value. In case of m parameters to optimize, this procedure requires $m + 1$ evaluations if the gradient is approximated by $[F(p + \epsilon) - F(p)]/\epsilon$ or $2m + 1$ if the gradient is approximated by

$[F(p+\epsilon) - F(p-\epsilon)]/(2\epsilon)$. For gradient algorithms, the starting value is important in order to have convergence. The height of the passive shield was chosen randomly to be 60 mm. The starting values for the active shield parameters were chosen taken into account the following issues: the coils should be positioned rather close to the axisymmetric axis, because the field lines enter the target area from that zone. The coil with the smallest radius should have the highest currents because the field intensity decreases with increasing r. The ratio between the real and imaginary part of the current was chosen to be equal to

$$\frac{\int_{\mathrm{TA}} \sqrt{\left[\mathrm{Re}(B_r)\right]^2 + \left[\mathrm{Re}(B_z)\right]^2}\,\mathrm{d}a}{\int_{\mathrm{TA}} \sqrt{\left[\mathrm{Im}(B_r)\right]^2 + \left[\mathrm{Im}(B_z)\right]^2}\,\mathrm{d}a} = \frac{6.163\,\mu T}{3.742\,\mu T} = 1.65 \tag{5.1}$$

in case of a 90-mm high passive shield with $\mu_r = 372$. The design sensitivity approach also solves a constrained nonlinear optimization problem without, however, the need to evaluate the model $m+1$ or $2m+1$ times in order to determine the gradients. The gradients are retrieved from the solution of the adjoint system, which requires two extra solutions (one for the real part of the source term and one for the imaginary part), however, without the need to generate a new geometry or mesh. Similar to the classical gradient algorithm, the starting value is important and the solution may be a local minimum.

The following optimizations are carried out each time using the three algorithms (see Table 1).

(1) Optimization of the height of the passive shield (1 parameter)

In optimization 1a, the passive shield has conductivity 5.9×10^6 S/m and permeability $\mu_r = 1$. All terms in the cost function and their derivatives can be seen in Figures 4 and 5, respectively, as well as the position of the optimum at 83.9 mm and its cost of 4.3924. Table 1 shows that the optimum is found by the conventional gradient method and by the genetic algorithm. The adjoint method finds a very good approximation (80.7 mm) of the optimum and a slightly higher objective value. This can be explained by the fact that the gradients of the two methods almost coincide in Figure 5. Evidently, for only one parameter to optimize, the adjoint method is slower than the conventional gradient method. The genetic algorithm is very slow.

In optimization 1b, the shield has the same conductivity 5.9×10^6 S/m and also a permeability of $\mu_r = 372$. In Figure 6, it is observed that in this case the accuracy of the adjoint method is much lower. This explains why the optimal height found by the adjoint method is 71.7 mm—this is where the gradient of the adjoint method crosses zero—instead of 100.1 mm. The difference in cost, however, is very small (see Table 1) so that the solution obtained by the adjoint method is still an acceptable solution. Notice that the shield with high permeability has a lower cost value than the purely conductive one. The average field in the target area is 7.76 μT versus 12.05 μT for the non-ferromagnetic shield: a reduction compared to the average induction in the unshielded case (28.9 μT) of 11.4 and 7.6 decibel, respectively. The optimal height of the non-ferromagnetic shield is lower than the optimal height of the ferromagnetic one, because it has more losses ($F_P = 0.73$ versus 0.51) and much more disturbance of the heating process ($F_W = 1.53$ versus 0.61).

Table 1: Comparison of optimization techniques concerning calculation time (CPU) and optimal cost value $F_{\min}$ for several shielding problems.

	Optimized parameters		Genetic		Gradient conv.		Gradient adjoint	
	Pas.	Act.	CPU	$F_{\min}$	CPU	$F_{\min}$	CPU	$F_{\min}$
1a	h_p ($\mu_r = 1$)	—	1h 24′ 29″	4.3924	3′ 15″	4.3924	8′ 10″	4.3939
1b	h_p ($\mu_r = 372$)	—	1h 00′ 39″	2.9562	6′ 28″	2.9562	9′ 24″	3.1376
2	—	I_{r1}, I_{i1}, r_1	2h 19′ 33″	2.9023	9′ 31″	2.9166	11′ 20″	2.9194
3a	h_p ($\mu_r = 372$)	I_{r1}, I_{i1}, r_1	2h 19′ 45″	2.9077	8′ 40″	2.9285	11′ 59″	3.0154
3b					8′ 13″	2.9100	8′ 15″	2.9681
4	—	I_{ri}, I_{ii}, r_i, $i = 1 \cdots 3$	2h 48′ 01″	2.8635	1h 16′ 54″	3.0523	16′ 13″	2.9832

(2) Optimization of an active shield consisting of only one coil

The complex current $I_{r1} + jI_{i1}$ and the horizontal position r_1 are optimized (3 parameters). A passive shield with height $h_p = 90\,\text{mm}$ is present but not optimized. In order to test the gradient algorithms, a "bad" starting value was chosen $[I_{r1}, I_{i1}, r_1] = [0, 0, 0.40]$. The genetic algorithm found a very good solution [120.9, −67.4, 0.503] while the gradient algorithms found [143.3, − 88.30, 0.415] (conventional) and [157.6, − 90.3, 0.400] (adjoint), both having almost the same cost and requiring almost the same CPU time. The correct position is not found by the gradient algorithms, but their solution has a cost that is only 0.5% higher than the one of the genetic algorithm.

(3) Optimization of both the passive and the active shield with one coil

Next to the three parameters of case 2, also the height of the passive shield is optimized (4 parameters in total). In case 3a, the starting value was chosen according to the description above: $[h_p, I_{r1}, I_{i1}, r_1] = [0.060, 200, -121, 0.300]$, resulting in a slightly higher cost for the conventional gradient algorithm compared to the genetic algorithm that found the optimal value: [0.1038, 89.38, − 103.62, 0.342]. In case 3b, the starting values for the active shield are chosen equal to the optimum of case 2. This time, the gradient algorithm finds the global minimum. Both for 3a and 3b, the gradient method using the adjoint system ends up with a higher cost because the height of the passive shield is chosen too low for the same reason as in case 1b. The calculation time of both gradient algorithms has the same magnitude.

(4) Optimization of an active shield consisting of 3 coils

This optimization results in 9 optimization parameters (three times real part of the current, imaginary part of the current, and horizontal position). The starting positions were 0.3, 0.4, and 0.6 m, the starting values for the currents were $I_{r1,0} = [300, -100, 50]$ and $I_{i1,0} = -I_{r1,0}/1.65$. None of both gradient algorithms found the low cost value of the genetic algorithm (2.86) or the corresponding optimum [219.69, − 72.442, 89.033, 17.53, − 73.202, − 16.1, 0.274, 0.404, 0.891]. For 9 variables, the approach using the adjoint variable is much faster than the conventional algorithm. Moreover, it finds a lower cost value although the gradients do not deviate much from the conventional ones.

When comparing these results to the ones from [1], it is seen that the calculation time of the genetic algorithm is much faster than for the genetic algorithm in the cited paper. This is because the objective function in the latter paper contained one finite-element calculation for every compensation coil considered, as well as an inner genetic algorithm nested in the main objective function, and up to nine compensation coils. The resulting shield had a better efficiency than the ones presented here (with less compensation coils), but the CPU time could be reduced from several days to a few hours. The gradient techniques solve the problem in less than one hour, but usually, their solution has a slightly higher cost value.

6. Conclusion

The gradient algorithm using an adjoint system was compared to a gradient method and a genetic algorithm for the design of a passive and active shield. It is observed that the accuracy of the adjoint variable method is good regarding the height of a conductive, non-ferromagnetic shield and regarding the current in compensation coils. The accuracy is somewhat less concerning the position of active shield coils. For a shield with high permeability, the accuracy may be unacceptable. For one evaluation of the objective function, the CPU time of the adjoint approach is about three times longer (as it requires two extra finite-element calculations). The optimization using the adjoint variables is slower than the conventional gradient method in the case of less than three parameters to optimize, comparable in case of three or four parameters, and faster in the case of more than four parameters.

Acknowledgments

This work was supported by the FWO project G.0082.06, by the GOA project BOF 07/GOA/006 and the IAP project P6/21. P. Sergeant and I. Cimrak are supported by the Fund for Scientific Research—Flanders FWO (Belgium).

References

[1] P. Sergeant, L. Dupré, M. De Wulf, and J. Melkebeek, "Optimizing active and passive magnetic shields in induction heating by a genetic algorithm," *IEEE Transactions on Magnetics*, vol. 39, no. 6, pp. 3486–3496, 2003.

[2] D. H. Kim, K. S. Ship, and J. K. Sykulski, "Applying continuum design sensitivity analysis combined with standard EM software to shape optimization in magnetostatic problems," *IEEE Transactions on Magnetics*, vol. 40, no. 2, pp. 1156–1159, 2004.

[3] D. H. Kim, J. K. Sykulski, and D. A. Lowther, "A novel scheme for material updating in source distribution optimization of magnetic devices using sensitivity analysis," *IEEE Transactions on Magnetics*, vol. 41, no. 5, pp. 1752–1755, 2005.

[4] ICNIRP-Guidelines, "Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz)," *Health Physics*, vol. 74, no. 4, pp. 494–522, 1998.

[5] "Corrigendum to directive 2004/40/EC of the European parliament and of the council of 29 April 2004 on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (electromagnetic fields)," *Official Journal of the European Union*, vol. 47, no. L 184, pp. 1–9, 2004.

[6] R. B. Schulz, V. C. Plantz, and D. R. Brush, "Shielding theory and practice," *IEEE Transactions on Electromagnetic Compatibility*, vol. 30, no. 3, pp. 187–201, 1988.

[7] R. A. Walling, J. J. Paserba, and C. W. Burns, "Series-capacitor compensated shield scheme for enhanced mitigation of transmission line magnetic fields," *IEEE Transactions on Power Delivery*, vol. 8, no. 1, pp. 461–469, 1993.

[8] P. Pettersson, "Principles in transmission line magnetic field reduction," *IEEE Transactions on Power Delivery*, vol. 11, no. 3, pp. 1587–1593, 1996.
[9] M. L. Hiles, R. G. Olsen, K. C. Holte, D. R. Jensen, and K. L. Griffing, "Power frequency magnetic field management using a combination of active and passive shielding technology," *IEEE Transactions on Power Delivery*, vol. 13, no. 1, pp. 171–179, 1998.
[10] P. P. Silvester and R. L. Ferrari, *Finite Elements for Electrical Engineers*, Cambridge University Press, Cambridge, UK, 1990.
[11] I. Cimrák and V. Melicher, "Sensitivity analysis framework for micromagnetism with application to optimal shape design of MRAM memories," *Inverse Problems*, vol. 23, pp. 563–588, 2007.
[12] D. B. Fogel, "An introduction to simulated evolutionary optimization," *IEEE Transactions on Neural Networks*, vol. 5, no. 1, pp. 3–14, 1994.
[13] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, London, UK, 1981.