ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Computational Physics 187 (2003) 68–94

JOURNAL OF COMPUTATIONAL PHYSICS

www.elsevier.com/locate/jcp

Induction heating processes optimization a general optimal control approach

Y. Favennec *, V. Labbé, F. Bay

Centre de Mise en Forme des Matériaux, Ecole des Mines de Paris, B.P. 207, 06904 Sophia Antipolis Cédex, France

Received 5 November 2001; received in revised form 5 January 2003; accepted 31 January 2003

Abstract

A general automatic optimization procedure coupled to a finite element induction heating process simulation has been developed. The mathematical model and the numerical methods are presented along with results validating the model. The first part of this paper presents the direct induction heating mathematical model, the related main numerical choices and especially the ultra-weak coupling procedure. The general optimization problem is then presented with the full detailed transposition of the ultra-weak coupling procedure to the adjoint problem. Numerical results provided at the end prove the efficiency and robustness of the adjoint model in optimizing induction heating processes.

Keywords: Optimization; Inverse analysis; Control; Ajoint theory; Finite elements; Induction heating; Multiphysics coupling; Ultra-weak coupling; Semi-analytical derivation

1. Introduction

Induction heating processes have become increasingly used in these last years in industry. The main advantages of using these processes when compared to any other heating process (gas furnace...) are, among others, their fast heating rate, good reproducibility and low energy consumption [1]. The induction heating process basically consists in transmitting by electromagnetic means, energy from a coil through which an alternative current is circulating. Induced currents in the conductive part due to the well-known Foucault law then heat the workpiece thanks to the Joule effect. Induction heating processes are mainly used either at low frequencies (around 50 Hz), usually in order to reach a temperature distribution as uniform as possible within the material before any forming process, or at much higher frequencies (10^4–10^6 Hz) in order to heat very locally near the surface, usually for heat treatments [2].

Most induction heating processes are set up using engineering experience and a trial-and-error procedure in order to achieve the corresponding goal (grain size control, uniform prescribed temperature, hardness

* Corresponding author. Tel.: +33-4-93-95-75-86; fax: +33-4-92-38-97-52.
E-mail address: Yann.Favennec@ensmp.fr (Y. Favennec).

0021-9991/03/$ - see front matter
doi:10.1016/S0021-9991(03)00081-0

map, etc.). Induction heating process simulation, which couples electromagnetic and heat transfer equations, can be of great help for a more in depth understanding of occurring physical phenomena. So far, various numerical models have been developed coupling electromagnetism and heat transfer. Most models involve the well-known finite element approach [3–5] or mixed finite element and boundary element approaches [6–8]. Even though mixed methods are interesting due to their inherent ability to take into account open domains and inductor displacements, the global finite element approach has been preferred since it involves sparse matrices (leading to reductions in terms of CPU time and memory requirements) and is more suited for parallel computing. Most authors use the harmonic approximation, assuming that all electromagnetic fields are sine waves when the input current is a sine wave. This approximation, valid when considering linear magnetic materials, can yield to large errors when dealing with highly ferromagnetic materials [3,9]. That is the reason why the time-dependent formulation has been preferred. Time-dependent integration being very time consuming when using a traditional weak coupling between all problems, the ultra-weak strategy has been developed. The complete model developed in our laboratory, well described in [10], is extended here.

The use of a direct induction heating process simulation code can be a first step in optimizing the global process. Nevertheless, in order to save again on time computation and improve accuracy, general numerical models have to be developed in order to automatically optimize the process with respect to any industrial goal as long as they deal with temperature evolutions anywhere within the part to be heated. Very few optimization models (but [11–13]) applied on coupled magneto-thermal problems can be found in the literature. Models presented in [11,12] use the harmonic approximation and consider that electromagnetic and heat conduction problems are uncoupled. The same approach can also be found in the control of ultrasound surgery [14,15]. In [13], the optimization procedure is based on a zero order method leading to a too high computational cost. To go much further in induction optimization, this paper presents the optimization of induction heating processes when dealing with any material (with nonlinear physical properties) and using a gradient-type method based on the completely coupled formulation between both time-dependent nonlinear electromagnetic and heat transfer problems. Particular attention is given in the transposition of the ultra-weak coupling from the direct to the adjoint problem. Another major difference between [11] and this paper resides in the fact that the model presented here optimizes transient trajectories while [11] discusses only the final steady-state problem.

The paper is organized as follows. In Section 2, we review the direct induction heating modeling. The continuous equations as well as the space finite element discretization and the time integration procedure are presented. In Section 3, we formulate the optimization problem and derive the solution. In Section 4, we present results from non-trivial simulations. Numerical tests performed on several different objectives show the very good applicability of the developed optimization algorithm. Section 5 is dedicated to conclusions and extensions of the proposed approach.

2. The direct induction heating model

The mathematical model developed for induction heating normally involves three main physical phenomena related to electromagnetism, heat transfer and solid mechanics [10]. As far as this paper is concerned, we shall introduce only electromagnetism and heat transfer since objectives to be reached in the optimization procedure deal only with temperature evolutions through the control of electromagnetic process parameters. We have chosen to carry out a complete finite elements approach for both electromagnetic and thermal computations. For each model, proper continuous equations are written with proper applied boundary conditions. Space and time discretizations are then explained. A validation of the direct model is then performed.

2.1. The electromagnetic model

The global system of equations modeling electromagnetic wave propagation is based on the four Maxwell equations:

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{1}$$

$$\vec{\nabla} \cdot \varepsilon \vec{E} = 0, \tag{2}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{3}$$

$$\vec{\nabla} \times \vec{H} = \vec{j} + \frac{\partial \varepsilon \vec{E}}{\partial t}, \tag{4}$$

where $\vec{B}$ is the magnetic field, $\vec{H}$ is the magnetic field, ε is the dielectric constant, $\vec{E}$ is the electric field, $\vec{j}$ is the electric current density associated with free charges, $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ and $\times$ denotes the vector product. This system of equations is coupled with relations associated to material properties:

$$\mu(T, |\vec{H}|) = \frac{\partial \vec{B}}{\partial \vec{H}}, \tag{5}$$

$$\vec{j} = \sigma(T)\vec{E}, \tag{6}$$

where μ is the differential magnetic permeability, σ is the electrical conductivity and T is the temperature. The electromagnetic resolution consists in calculating fields, $\vec{E} = \vec{E}(r,t)$, $\vec{B} = \vec{B}(r,t)$, $\vec{H} = \vec{H}(r,t)$ and $\vec{j} = \vec{j}(r,t)$ at any location r and time t that satisfy all relations (1)–(6). The standard procedure consists in writing a single second-order wave propagation-like equation. When solving the electromagnetic problem using the electrical field for instance, the procedure is the following. By dividing (3) by the differential magnetic permeability (5) and taking its rotational on both sides, we first get

$$\vec{\nabla} \times \left(\frac{1}{\mu} \vec{\nabla} \times \vec{E} \right) = -\vec{\nabla} \times \frac{\partial \vec{H}}{\partial t} = -\frac{\partial}{\partial t} \left(\vec{\nabla} \times \vec{H} \right). \tag{7}$$

Substituting (4) in (7) then gives

$$\frac{\partial}{\partial t} \left(\vec{j} + \frac{\partial \varepsilon \vec{E}}{\partial t} \right) + \vec{\nabla} \times \left(\frac{1}{\mu} \vec{\nabla} \times \vec{E} \right) = 0. \tag{8}$$

The total current density $\vec{j}$ being the sum of the induced currents $\sigma\vec{E}$ and the imposed one $\vec{J}_s$, (8) becomes, introducing the Ohm law (6)

$$\varepsilon \frac{\partial^2 \vec{E}}{\partial t^2} + \sigma \frac{\partial \vec{E}}{\partial t} + \vec{\nabla} \times \left(\frac{1}{\mu} \vec{\nabla} \times \vec{E} \right) = -\frac{\partial \vec{J}_s}{\partial t}. \tag{9}$$

With the electrical field as the unknown, the electromagnetic problem consists then in calculating $\vec{E} = \vec{E}(r,t)$ satisfying (9) along with the null divergence condition (2). In the same manner, the electromagnetic problem consists in, when using the magnetic field $\vec{H} = \vec{H}(r,t)$ as the unknown, calculating $\vec{H} = \vec{H}(r,t)$ satisfying (10) along with its null divergence condition (1)

$$\varepsilon\frac{\partial^2\vec{H}}{\partial t^2}+\sigma\frac{\partial\vec{H}}{\partial t}+\vec{\nabla}\times\left(\frac{1}{\mu}\vec{\nabla}\times\vec{H}\right)=-\frac{\partial\vec{J}_s}{\partial t}. \tag{10}$$

In axis-symmetrical configurations, using a cylindrical coordinate system $[\vec{e}_r,\vec{e}_\theta,\vec{e}_z]$, electromagnetic fields do not depend on the angular coordinate ($\partial\cdot/\partial\theta=0$) and it is shown from (6) that when the current density circulates in the ortho-radial direction ($\vec{J}_S=J_S(r,z)\vec{e}_\theta$), the electrical field $\vec{E}=E_\theta(r,z)\vec{e}_\theta$ is then reduced to a single non-zero component in the ortho-radial direction with the null divergence condition implicitly taken into account [16,17]. Developing (9) in the cylindrical coordinate and taking into account of above remarks, the problem is then to find $E=E_\theta(r,z)$ satisfying

$$\varepsilon\frac{\partial^2E}{\partial t^2}+\sigma\frac{\partial E}{\partial t}-\vec{\nabla}\cdot\left(\frac{1}{\mu}\vec{\nabla}E\right)+\frac{1}{\mu}\frac{E}{r^2}-\frac{\partial}{\partial r}\left(\frac{1}{\mu}\right)\frac{E}{r}=-\frac{\partial J_S}{\partial t}, \tag{11}$$

where the last term in the left-hand side of (11) is usually neglected [5] even though it is numerically shown [17] that it affects consequently the results when dealing with magnetic materials. The input current density being most of the time a sine wave, this writes $J_S=J_0\sin(2\pi ft)$, where J_0 is the amplitude and f is the frequency. Next, a very standard approximation consists in neglecting the displacement currents (last term in the right-hand side of (4)). The general domain of validity when using this so-called magneto-quasi-static approximation can be found in any electromagnetism handbooks [18], and is given explicitly for induction heaters in [10]. Eventually, the electromagnetic problem consists in calculating $E=E_\theta(r,t)$ satisfying (12)

$$\sigma\frac{\partial E}{\partial t}-\vec{\nabla}\cdot\left(\frac{1}{\mu}\vec{\nabla}E\right)+\frac{1}{\mu}\frac{E}{r^2}-\frac{\partial}{\partial r}\left(\frac{1}{\mu}\right)\frac{E}{r}+2\pi fJ_0\sin(2\pi ft)=0. \tag{12}$$

It should be noted that Eq. (12) is a diffusion-like equation instead of the wave propagation-like equation (11). Two physical parameters, namely the electrical conductivity σ and the magnetic permeability μ, are of concern in (12). Both depend on temperature. For nonlinear ferromagnetic materials, the magnetic permeability also depends on the magnetic field strength. General formulations for linear and nonlinear properties are presented in [19]. Since most magnetic materials present nonlinearities with respect to the magnetic field [20], it is usually convenient to have a proper linearization transfer formulation from nonlinear to the linear formulation. This is also presented in [19,21]. It is shown in [22] (resp. [10]) how the nonlinearity is handled when using the magnetic field (resp. electric field) as the state variable. It is well known that all electromagnetic fields strengths decrease as the inverse of the source current distance to reach zero at infinity. When using the standard finite element method, it is necessary to take into account a closed domain using an artificial border. In axis-symmetrical cases, a null Dirichlet boundary condition on the symmetry axis Γ_0^e is prescribed to the electrical field and, to avoid artificial reflections or external borders $\Gamma^e-\Gamma_0^e$, an absorbing-type Robin-like condition is prescribed [4,12,23] (see Fig. 1).

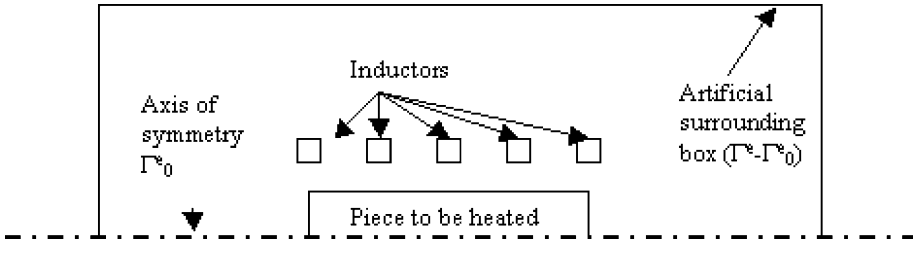

Fig. 1. Schematic axis-symmetrical representation of the global domain of study $\Omega^e=\Omega_{\text{part}}\cup\Omega_{\text{inductors}}\cup\Omega_{\text{air}}$ and of the border Γ^e.

2.2. The heat transfer model

Eddy currents derived from the electromagnetic model induce the heat dissipated within the workpiece due to Joule effects. Temperature evolution within the workpiece is governed by the classical heat transfer equation

$$\rho C\frac{\partial T}{\partial t}-\operatorname{div}(k\vec{\nabla}T)=\sigma E^2, \tag{13}$$

where ρ is the material density, C is the specific heat, k is the thermal conductivity, and σE^2 is the heat source term due to eddy currents [2]. The specific heat and thermal conductivity are also temperature dependent. Different kinds of boundary conditions for temperature or its normal derivative can be prescribed at interfaces, listing convection and radiation between the part and the air on Γ_0^t (14) and prescribed heat flux on Γ_1^t (15) [24]. We can also consider a prescribed temperature on Γ_2^t (16):

$$-k\vec{\nabla}T\cdot\vec{n}=h(T-T_{\text{ext}})+\varepsilon_{\text{emi}}\sigma_{\text{Ste}}(T^4-T_{\text{ext}}^4), \tag{14}$$

$$-k\vec{\nabla}T\cdot\vec{n}=\Phi_{\text{prescribed}}, \tag{15}$$

$$T=T_{\text{prescribed}}, \tag{16}$$

where $\vec{n}$ is the outward unit normal vector, h is the convection coefficient, ε_{emi} is the material emissivity, σ_{Ste} is the Stephan constant and T_{ext} is the room temperature. Although the heat conduction equation (13) is valid anywhere in the domain $\Omega^t\equiv\Omega^e$, one may choose to solve it only on the workpiece to be heated ($\Omega^t=\Omega_{\text{part}}$) – thus enabling easy enforcement of special boundary conditions such as convection and radiation for instance. Locations of boundaries Γ_0^t, Γ_1^t and Γ_2^t are given as an example in Section 4.

2.3. Integral formulations and space discretization

In order to establish weak formulations of (11) and (13), we multiply them, respectively, by test functions ψ and φ belonging to the functional space V and W that respect, respectively (17) and (18):

$$V=\left\{\psi\in H^1(\Omega^e),\frac{\psi}{r}\in L^2(\Omega^e),\psi=0\text{ on }\Gamma_O^e,\frac{\partial\psi}{\partial\theta}=0\right\}, \tag{17}$$

$$W=\left\{\varphi\in H^1(\Omega^t),\varphi=0\text{ on }\Gamma_2^t,\frac{\partial\varphi}{\partial\theta}=0\right\}, \tag{18}$$

with $H^1(\Omega)=\{\psi,\varphi\in L^2(\Omega),\nabla\psi,\nabla\varphi\in L^2(\Omega)\}$, and integrate them on the whole considered domains. After using the Green theorem, one gets the following Cauchy–Dirichlet–Neumann problems from which existence and uniqueness of solutions can be proved [25]

$$\int_\Omega\sigma\frac{\partial E}{\partial t}\psi+\int_\Omega\frac{1}{\mu}\vec{\nabla}E\cdot\vec{\nabla}\psi+\int_\Omega\frac{1}{\mu r^2}E\cdot\psi+\int_\Omega\frac{1}{\mu r}\frac{\partial}{\partial r}(E\psi)+\int_\Omega 2\pi f J_0\cos(2\pi ft)\psi=0\quad\forall\psi\in V, \tag{19}$$

$$\int_\Omega\rho C\frac{\partial T}{\partial t}\varphi+\int_\Omega k\nabla T\cdot\nabla\varphi+\int_{\Gamma_0^t}hT\cdot\varphi-\int_\Omega\sigma E^2\varphi-\int_{\Gamma_1^t}\Phi_{\text{prescribed}}\cdot\varphi-\int_{\Gamma_0^t}hT_{\text{ext}}\cdot\varphi=0\quad\forall\varphi\in W, \tag{20}$$

where the radiation boundary condition (14) has been linearized with respect to temperature pointing out that $\varepsilon_{\text{emi}}\sigma_{\text{Ste}}(T^4 - T_{\text{ext}}^4) \cong \varepsilon_{\text{emi}}\sigma_{\text{Ste}}(T^{*2} - T_{\text{ext}}^2)\cdot(T^* + T_{\text{ext}})(T - T_{\text{ext}})$, where T^* is an approximation of temperature (usually temperature of the previous time step). The generalized convection coefficient h involved in (20) and afterwards thus encompasses the classical convection coefficient as well as the linearized radiative condition [24].

Weak formulations (19) and (20) have now to be discretized in space. Functional spaces V and W are classically approached by, respectively, space discretized V^h and W^h, tests functions ψ and φ are approached by ψ^h and φ^h and unknowns E and T by E^h and T^h. Choosing a proper basis of the discretized space provided by the shape functions N_i associated to each node i of the quadratic triangular mesh, one gets the following linear systems, where E and T stand for the discretized versions E^h and T^h from now onwards:

$$\{R^e\} = [C^e]\left\{\frac{\partial E}{\partial t}(t)\right\} + [K^e]\{E(t)\} - \{B^e\} = \{0\}, \tag{21}$$

where

$$\begin{aligned}
[C^e]_{ij} &= \sum_{elt=1}^{nb.elts} \int_{elt} \sigma N_j N_i \,\mathrm{d}r,\\
[K^e]_{ij} &= \sum_{elt=1}^{nb.elts} \left[\int_{elt} \frac{1}{\mu}\nabla N_j\cdot\nabla N_i \,\mathrm{d}r + \int_{elt}\frac{1}{\mu r^2} N_j N_i \,\mathrm{d}r + \int_{elt}\frac{1}{\mu r}\cdot\frac{\partial}{\partial r}(N_j N_i)\,\mathrm{d}r\right],\\
\{B^e\}_i &= \sum_{elt=1}^{nb.elts}\left[\int_{elt} -2\pi f J_0\cos(2\pi f t)\cdot N_i \,\mathrm{d}r\right],
\end{aligned} \tag{22}$$

and

$$\{R^t\} = [C^t]\left\{\frac{\partial T}{\partial t}(t)\right\} + [K^t]\{T(t)\} - \{B^t\} = \{0\}, \tag{23}$$

where

$$\begin{aligned}
[C^t]_{ij} &= \sum_{elt=1}^{nb.elts} \int_{elt} \rho C N_j N_i \,\mathrm{d}r,\\
[K^t]_{ij} &= \sum_{elt=1}^{nb.elts}\left[\int_{elt} k\nabla N_j\cdot\nabla N_i \,\mathrm{d}r + \int_{\partial elt\cap\Gamma_0^t} h N_j N_i \,\mathrm{d}r\right],\\
\{B^t\}_i &= \sum_{elt=1}^{nb.elts}\left[\int_{elt}\sigma E^2\cdot N_i \,\mathrm{d}r + \int_{\partial elt\cap\Gamma_1^t}\Phi_{\text{prescribed}} N_j \,\mathrm{d}r + \int_{\partial elt\cap\Gamma_0^t} h T_{\text{ext}} N_j \,\mathrm{d}r\right].
\end{aligned} \tag{24}$$

2.4. Time discretization and coupling strategy

The space semi-discretized electromagnetic and thermal equations ((21), (22)) and ((23), (24)) must then be integrated in time. For accuracy reasons, we have selected second-order time step finite difference schemes detailed in [24,26] and recalled here below. Note that integration schemes are alike for both electromagnetic and thermal computations. Hereafter, variable X represents any of the main state variables that are the electrical field E or temperature T, and superscript X stand for either e or t with no misunderstanding. As a first stage, one defines a time t^* as

$$t^* = \alpha_1(t-\delta t) + \alpha_2 t + \alpha_3(t+\delta t), \tag{25}$$

with

$$\alpha_1 + \alpha_2 + \alpha_3 = 0. \tag{26}$$

The field X^* at time t^* and its time derivative are approximated by

$$X^* = \alpha_1 X^{t-\delta t} + \alpha_2 X^t + \alpha_3 X^{t+\delta t}, \tag{27}$$

$$\frac{\partial X^*}{\partial t} = \gamma \frac{X^{t+\delta t} - X^t}{\delta t} + (\gamma - 1)\frac{X^{t-\delta t} - X^t}{\delta t}. \tag{28}$$

System (29) is solved at time t^* for X^* with:

$$\begin{aligned} &\left(\frac{\gamma}{\alpha_3 \delta t}[C^X]^* + [K^X]^*\right)\{X^*\} = \{B^X\}^* + c_1[C^X]^*\{X^t\} + c_0[C^X]^*\{X^{t-\delta t}\}, \\ &c_1 = \frac{\gamma}{\delta t} + \frac{\gamma - 1}{\delta t} + \frac{\gamma \alpha_2}{\alpha_3 \delta t}, \\ &c_2 = \frac{\gamma \alpha_1}{\alpha_3 \delta t} - \frac{\gamma - 1}{\delta t}. \end{aligned} \tag{29}$$

Matrices are then linearized and do only depend on their values at times t and $t - \delta t$ [27]

$$[C]^* = (\alpha_1 - \alpha_3)[C]_{t-\delta t} + (\alpha_2 + 2\alpha_3)[C]_{t-\delta t}. \tag{30}$$

The corresponding field at time $t + \delta t$ is then computed

$$\{X\}^{t+\delta t} = \frac{1}{\alpha_3}(\{X\}^* - \alpha_1\{X\}^{t-\delta t} - \alpha_2\{X\}^t). \tag{31}$$

The choice of coefficients α_1, α_2, α_3 and γ greatly influences convergence [26]. Thus, several distinct schemes have been tested. All numerical results presented in [17] are briefly recalled here. Comparative studies have been performed on the test case (geometry, process and physical parameters) given by [5]. Fig. 2 presents the electrical field evolution on one location of the workpiece surface. It is shown that these second-order schemes which tend towards explicit schemes present high oscillations while the Dupont scheme is much more stable. As a consequence, this latter scheme converges much easier and faster. Although the time integration scheme choice is fundamental for the electromagnetic computation efficiency, this choice is less restrictive for heat transfer computations [17]. Therefore, the Dupont scheme ($\alpha_1 = 1/4$, $\alpha_2 = 0$, $\alpha_3 = 3/4$ and $\gamma = 1$) has been chosen for both computations.

Both equations are solved in an incremental way. The weak coupling is the usual way for integrating in time both coupled equations. The general procedure is the following. At a given time step $t^n = n \cdot \delta t$, $n = 1, \ldots, N$, with $N \cdot \delta t = t_f$, one first calculates the electromagnetic field $E^n = E(r, t^n)$ at a fixed known temperature field $T^{n-1} = T(r, t^{n-1})$, then, as the second step, one calculates the temperature $T^n = T(r, t^n)$ using the pre-calculated fixed electromagnetic field E^n. These two sub-systems have to be solved N times in order to cover the whole time range $[t_0, t_f]$. The global direct problem resides thus in solving, where u^n are the controls:

$$\begin{cases} R^e(E^n, u^n) = 0 \\ R^t(E^n, T^n) = 0 \end{cases} \quad \forall n = 1, N. \tag{32}$$

Even though weak couplings are generally interesting due to their inherent high integration accuracy, the total number of integrations can become extremely high when considering high frequency. Indeed, the

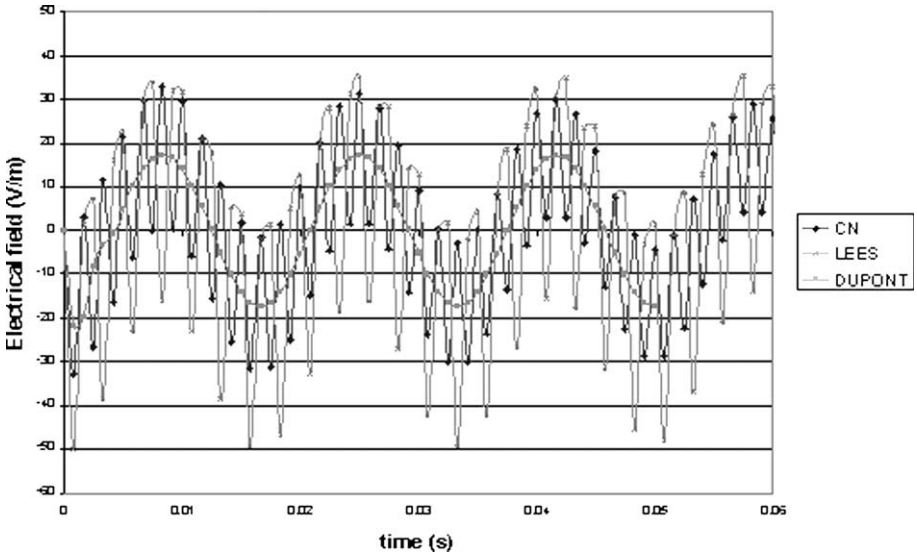

Fig. 2. Temporal evolution of the electrical field on the workpiece surface for several time integration schemes (Crank–Nicholson, Lees, Dupont).

characteristic time step scale for the thermal problem is of the order of magnitude of the second while the electromagnetic one is of the order of a fraction of an electromagnetic period and thus can reach 10^{-6}–10^{-9} s. In principle, both equations should be integrated using the same time step. One should thus consider the smaller of both. So, when discretizing electromagnetic periods by$N$ steps (usually 64), one has to compute $N * f$ times both equations of (32) per second of heating. This such high number of integrations leads to practical unfeasibility. A preferred ultra-weak coupling is performed.

Induction heating processes lasting several seconds in practice, temperature evolutions within a range related to an electromagnetic period τ or even several periods are very small and thus evolutions of all physical parameters are also negligible. As a consequence, for the small considered time interval, a periodic input current density $J_S(t)$ and thus a periodic loading vector $B^e(t)$ implies a periodic response $E(t)$. One then chooses a thermal discretization $[t_0, t_f] = \cup[t^{k-1}, t^k], k = 1, \ldots, K$ such that within each interval, evolutions of all physical parameters are small. Doing so, computation of the electromagnetic problem can be performed for each step k from time t^{k-1} to $t^{k-1} + \tau$ with a time step being still a fraction of the period τ. The response is then extrapolated during the rest of the intervals, namely within $[t^{k-1} + \tau, t^k]$. The heat transfer problem is then integrated with the chosen time step $|t^k - t^{k-1}|$ with a time averaged signal over the electromagnetic period $\tilde{E}^2(t^k)$ in order to compute $B^t(t)$. Formalizing, the ultra-weak coupling leads in solving both coupled problems by (33) as schematically shown in Fig. 3. It is to be noted that the ultra-weak coupling procedure has previously been validated in [10]

$$\forall k = 1, K \begin{cases} R^e(E^n, u^n) = 0 \quad \forall n = 1, N, \\ R^t(\tilde{E}^{2^k}, T^k) = 0. \end{cases} \tag{33}$$

2.5. The moving inductors case

Since the finite element discretization is used, there are mainly two ways to simulate moving inductors. The first one is the remeshing use as it is done in forming processes simulation [28,29]. Remeshing

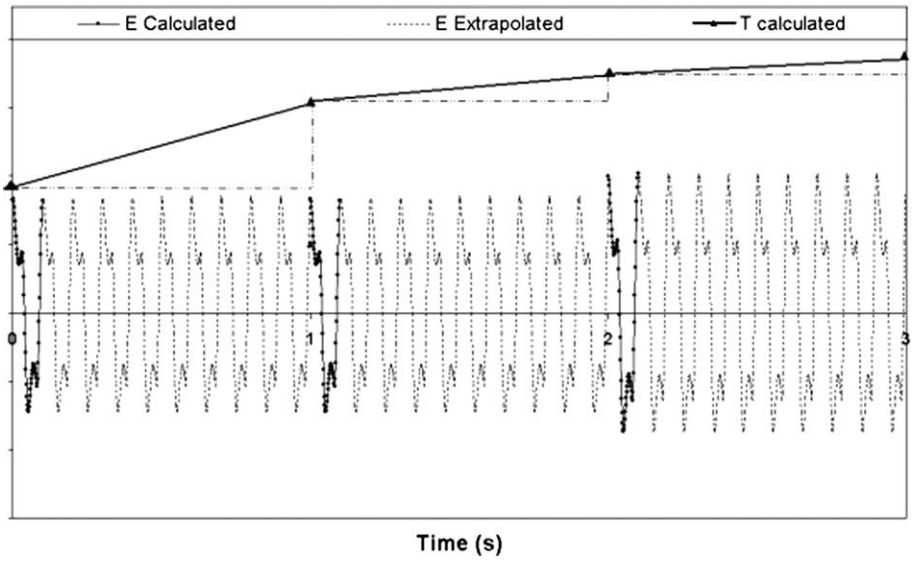

Fig. 3. The general ultra-weak coupling principle. At thermal intervals $k = 1, 2, 3$ the electromagnetic field is computed with its characteristic time step $\delta t = \tau/64$ during only one full period τ. Matrices C^e, K^e and the vector B^e for electromagnetic computation are assembled using the current step temperature $T^k = T(r, t^k)$. The electromagnetic field is then extrapolated until the next thermal step.

processes being very time consuming, we have preferred the following procedure where the area where the inductor is moving through is initially defined and meshed separately (see Fig. 16). The electromagnetic properties of this area are moved back and forth from air properties to inductor material properties. The inductors are moved virtually through an almost "continuous" change of physical

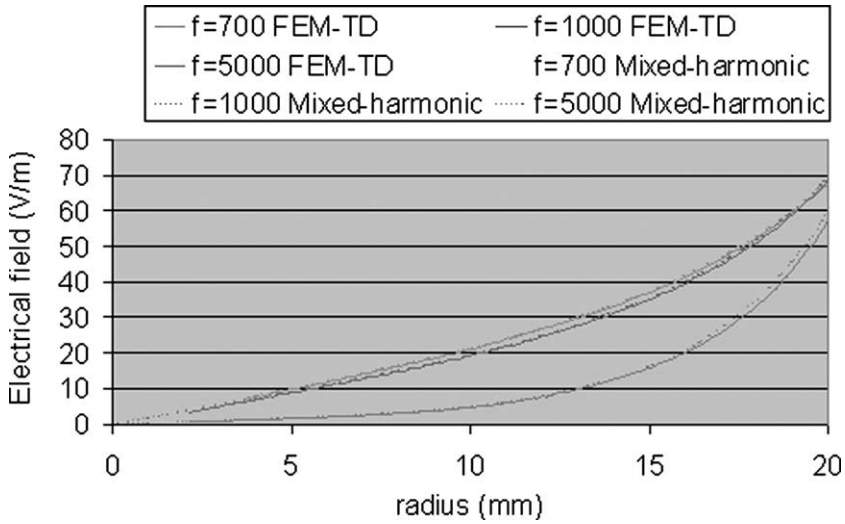

Fig. 4. Comparison of our finite element time-dependent model with the harmonic mixed boundary element–finite element model. Both give the same evolution for the electrical field as a function of the radius.

properties and location of electromagnetic source terms. The major advantage of using this method is that it enables an accurate simulation of the process without any mesh distortions, and thus avoids remeshing problems.

2.6. Conclusion

The direct modeling of induction heating processes presented upwards has been first confronted to an analytical solution given by [5], then to the numerical code developed at the laboratory FS-LNMS in Slovenia which is based on a mixed finite element–boundary element method [30,31]. Confrontation tests performed on the geometry and linear magnetic parameters given by [5] are presented in Fig. 4. Next, the comparison between the calculated electrical field evolution in the inductor surrounding a ferromagnetic cylindrical piece and the corresponding measured current intensity [32] (see Fig. 5) shows the robustness of the developed numerical model.

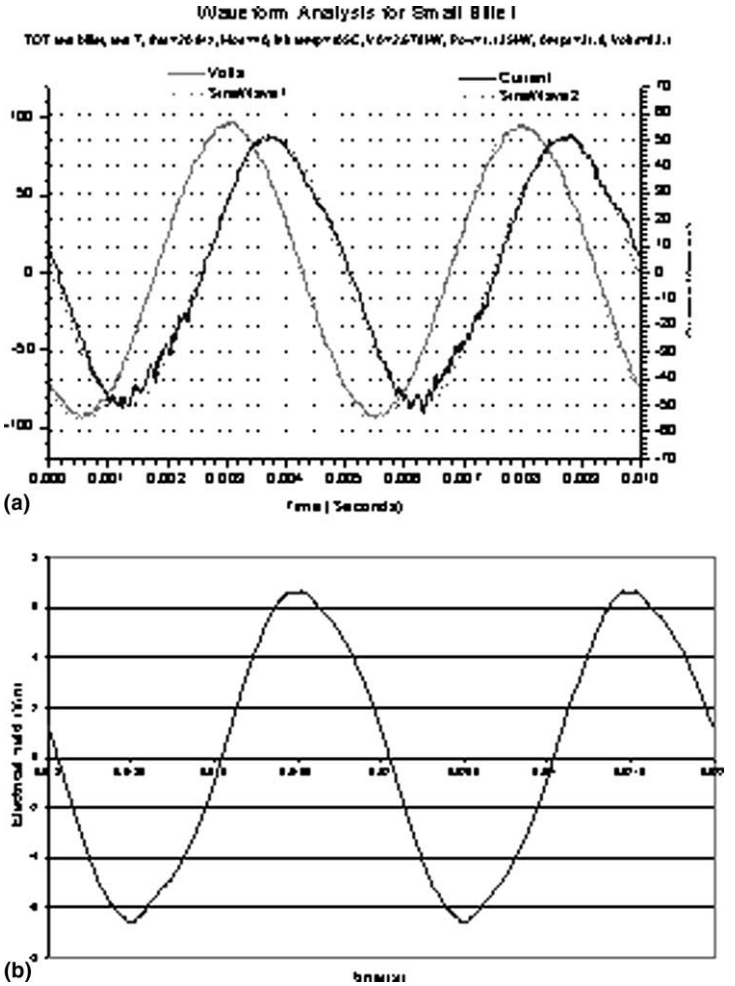

Fig. 5. Temporal evolution: (a) evolution with respect to time of the experimental intensity in the coil; (b) evolution with respect to time of the numerically computed electrical field in the part.

3. The optimization problem

3.1. Formulation of the optimization problem

Direct modeling of induction heating processes enables the determination of temperature $T(r,t;ut)$ at any location $r \in \Omega$ and at any time t for a given set of process parameters u. The process parameters are the frequency f and the amplitude J_0 of the input source current density $J_S = J_0 \sin(2\pi ft)$. The inductor velocity v with respect to the workpiece to be heated can also be a process parameter. All process parameters are stationary in time, so $u = \{(f, J_0, v) \in \Re^+ \times \Re^+ \times \Re^+\}$. We recall that induction heating processes are mainly used either for pre-heating prior to forming operations in order to reach a temperature as uniform as possible within the part, or for heat treatment applications in which case temperature has to follow a precise path in space and time [1]. The optimization problem consists in finding controls u such that calculated temperature $T(r,t;u)$ are as close as possible to optimal ones denoted $T^{\text{opt}}(r,t)$. The goal is then to minimize, in the mean squared sense, the discrepancy between calculated and optimal temperature. The discrepancy $\hat{J}$, often called cost or objective function, is defined in the general continuous case as

$$\hat{J}(u) = \frac{1}{2}\int_{t_0}^{t_f}\int_{\Omega_{\text{part}}} (T(r,t;u) - T^{\text{opt}}(r,t))^2 \,\mathrm{d}r\,\mathrm{d}t + \frac{1}{2}\int_{\Omega_{\text{part}}} (T(r,t_f;u) - T^{\text{opt}}(r,t_f))^2 \,\mathrm{d}r, \tag{34}$$

where both terms in the right-hand side of (34) have a different meaning. The first term used to display the objective function time integration has been used for instance by [33] for an inverse heat conduction problem, by [34] for minimizing the energy spent during a forming process, and also by [35] for a shape optimization problem of a waveguide antenna. In our case, this first term will be used when heat treatment applications are dealt with. The second term, based on the final temperature field, has been used for instance by [11] in shape optimization in induction heating, and also by [35] when dealing with the control of phase volume fractions at the end of a laser hardening process. This term is used in our case for dealing with pre-heating applications. Temperature $T(r,t;u)$ being calculated numerically, the continuous objective function given in (34) has to be written in its space and time discretized form. Moreover, the objective function may be integrated only at predefined times $(t^k, k = 1,\ldots,K, 0 \leqslant t^k \leqslant t_f)$ and on time varying space domain of interest $\Omega^{\text{opt}}(t^k) \subset \Omega_{\text{part}}$. Thus, the discretized objective function $J(u)$ writes

$$J(u) = \frac{1}{2}\sum_{k=1}^{K}\sum_{elt=1}^{nb.elt}\sum_{\substack{int=1 \\ \text{if } int\in\Omega^{\text{opt}}(t^k)}}^{nb.int} (T(r_{\text{int}}, t^k; u) - T^{\text{opt}}(r_{\text{int}}, t^k))^2 \omega_{\text{int}}\xi_{\text{int}}, \tag{35}$$

where ω_{int} is the weight associated to the integration point *int*, and ξ_{int} is the binary weight associated to time t^k. The latter weight equals 0 or 1 depending on whether or not the functional J is integrated at this time. Eventually, the optimization problem consists in finding controls u that minimize the functional $J(u)$ under the constraints $R^e(E,u) = 0$ and $R^t(E,T) = 0$, formally

$$\text{Find } \overline{u} \text{ such that } J(\overline{u}) = \min_{\substack{R^e(E,u)=0, \\ R^t(E,T)=0}} J(u). \tag{36}$$

3.2. The Lagrangian method

A large number of methods can be found in literature to solve such optimization problem. One can roughly separate them in two kinds: zero order methods and gradient-type methods. Zero order methods

do not use any information concerning the objective function first- or higher-order derivatives and thus usually converge to the global minimum. The Pareto optimal theory has for instance been applied to solve a shape optimization problem of an inductor [13]. The Pareto method has proved to be efficient but needed on the other hand an extremely large number of direct model runs [13]. Gradient-type methods, based upon a local differentiation study of the objective function, are more interesting since they usually converge much more quickly [36] but, on the other hand, the direct model has to be differentiated. The objective function being calculated through the resolution of a discretized direct problem, it is highly preferable to differentiate the discretized direct model rather than the continuous one [37,38]. The direct differentiation of the discretized induction heating model has been performed by [12] in the case of a steady-state electromagnetic model coupled to a transient heat transfer model and by [19] in the case of a coupled transient model. Very good comparative studies between the direct differentiation and the use of the adjoint theory (Lagrangian) show that, when dealing with full coupled nonlinear problems, the latter method becomes interesting when the number of functionals for which design sensitivities are needed is less than the number of design parameters [39–41]. Next, it is well known [42] that Lagrangian methods are well suited when there is no explicit dependence of state variables – involved in the objective function – and the controls. That are the main reasons why the adjoint method has been preferred. It can be introduced through various approaches. We shall introduce it here using the Lagrangian approach. The classical general method when dealing with a single system of equations can be found in various books such as [43,44] and is developed here for both considered coupled systems. We define the Lagrangian of the problem (36) as

$$L(u,E,T,\lambda,\mu) = J(u) + \langle R^e(E,u),\lambda\rangle_{\Omega^e\times[t_0,t_f]} + \langle R^t(E,T),\mu\rangle_{\Omega^t\times[t_0,t_f]}, \tag{37}$$

where variables λ and μ are adjoint variables and where scalar products are defined as

$$\langle u,v\rangle_{\Omega\times[t_0,t_f]} = \int_{t_0}^{t_f}\int_{\Omega} uv \,\mathrm{d}r\,\mathrm{d}t. \tag{38}$$

Domains of study Ω^e and Ω^t (resp. for the electromagnetic and the heat transfer models) being included in the domain Ω, the Lagrangian quantity can writes

$$L(u,E,T,\lambda,\mu) = J(u) + \langle R^e(E,u),\lambda\rangle_{\Omega\times[t_0,t_f]} + \langle R^t(E,T),\mu\rangle_{\Omega\times[t_0,t_f]}. \tag{39}$$

We shall now prove that a necessary condition for u to be solution of (36) is that there exists a set (E,T,λ,μ) such that (u,E,T,λ,μ) is a saddle point of L. Indeed, the necessary condition writes

$$J'(u) = 0. \tag{40}$$

Let us show that this condition is equivalent to

$$\exists(E,T,\lambda,\mu), \frac{\partial L}{\partial u}(\cdot) = \frac{\partial L}{\partial E}(\cdot) = \frac{\partial L}{\partial T}(\cdot) = \frac{\partial L}{\partial \lambda}(\cdot) = \frac{\partial L}{\partial \mu}(\cdot) = 0. \tag{41}$$

Let first E and T verify, respectively, $(\partial L/\partial\lambda)(\cdot) = 0$ and $(\partial L/\partial\mu)(\cdot) = 0$, i.e.,

$$R^e(E,u) = 0, \quad R^t(E,T) = 0 \quad \forall t \in [t_0,t_f]. \tag{42}$$

Lagrangian derivation writes

$$\frac{\partial L}{\partial u}(u,E,T,\lambda,\mu) = \frac{\partial J}{\partial u} + \left\langle \lambda, \frac{\partial R^e}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]}. \tag{43}$$

Differentiation of (32) gives

$$\begin{aligned}\frac{\partial R^e}{\partial u} &= -\frac{\partial R^e}{\partial E}\frac{\partial E}{\partial u},\\ \frac{\partial R^t}{\partial u} &= -\frac{\partial R^t}{\partial E}\frac{\partial E}{\partial u} - \frac{\partial R^t}{\partial T}\frac{\partial T}{\partial u}.\end{aligned} \tag{44}$$

So, (43) becomes

$$\frac{\partial L}{\partial u} = \frac{\partial J}{\partial u} - \left\langle \lambda, \frac{\partial R^e}{\partial E}\frac{\partial E}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} - \left\langle \mu, \frac{\partial R^t}{\partial E}\frac{\partial E}{\partial u} + \frac{\partial R^t}{\partial T}\frac{\partial T}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]}. \tag{45}$$

Let now adjoint variables λ and μ verify, respectively

$$\begin{aligned}&\left\langle \lambda, \frac{\partial R^e}{\partial E}\frac{\partial E}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \mu, \frac{\partial R^t}{\partial E}\frac{\partial E}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} = 0,\\ &\left\langle \frac{\partial J}{\partial T}, \frac{\partial T}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \mu, \frac{\partial R^t}{\partial T}\frac{\partial T}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} = 0,\end{aligned} \tag{46}$$

which is equivalent for adjoint variables λ and μ to verify, respectively, $(\partial L/\partial E)(\cdot) = 0$ and $(\partial L/\partial T)(\cdot) = 0$. Eq. (45) thus becomes

$$\frac{\partial L}{\partial u}(\cdot) = \frac{\partial J}{\partial u} + \left\langle \frac{\partial J}{\partial T}, \frac{\partial T}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} = J'(u). \tag{47}$$

3.3. The adjoint problem

We recall here that the objective function gradient (43) is valid only when adjoint variables λ and μ are such that scalar expressions in (46) are satisfied within the whole time range $[t_0 t_f]$. The aim of the Lagrangian procedure is to get on one side of all scalar products all derivations of state variables with respect to the controls in order to avoid computing them. To do so, here is given the method to employ. One first expand both equations in (46) using (21)–(24)

$$\begin{aligned}&\left\langle \lambda, C^e\frac{\partial}{\partial t}\frac{\partial E}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \lambda, K^e\frac{\partial E}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \mu, \frac{\partial B^t}{\partial E}\frac{\partial E}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} = 0,\\ &\left\langle \frac{\partial J}{\partial T}, \frac{\partial T}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \mu, C^t\frac{\partial}{\partial t}\frac{\partial T}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \mu, K^t\frac{\partial T}{\partial u} \right\rangle_{\Omega\times[t_0,t_f]} = 0,\end{aligned} \tag{48}$$

where all involved operators have been detailed in Section 2.3 – except for the last term in the first equation of (48) which is given explicitly by

$$\left[\frac{\partial B^t}{\partial E}\right]_{ij} = \frac{\partial\{B^t\}_i}{\partial E_j} = \begin{cases} \sum_{elt=1}^{nb.elt} \int_{elt} 2\sigma E N_i \,\mathrm{d}r & \text{if } i = j\\ 0 & \text{else.}\end{cases} \tag{49}$$

In order to separate differentiations of state variables with respect to control parameters, the time integration is performed on (48), using the fact that operators C^e, K^e, C^t and K^t are self-adjoint (symmetric matrices) while the time derivative operator $\partial\cdot/\partial t$ is antisymmetric. System (48) thus becomes

$$
\begin{aligned}
&-\left\langle C^e\frac{\partial\lambda}{\partial t},\frac{\partial E}{\partial u}\right\rangle_{\Omega\times[t_0,t_f]}+\left\langle C^e\lambda(t_f),\frac{\partial E}{\partial u}(t_f)\right\rangle_{\Omega}-\left\langle C^e\lambda(t_0),\frac{\partial E}{\partial u}(t_0)\right\rangle_{\Omega}\\
&\quad+\left\langle K^e\lambda,\frac{\partial E}{\partial u}\right\rangle_{\Omega\times[t_0,t_f]}+\left\langle\frac{\partial B^t}{\partial E}\mu,\frac{\partial E}{\partial u}\right\rangle_{\Omega\times[t_0,t_f]}=0,\\
&-\left\langle C^t\frac{\partial\mu}{\partial t},\frac{\partial T}{\partial u}\right\rangle_{\Omega\times[t_0,t_f]}+\left\langle C^t\mu(t_f),\frac{\partial T}{\partial u}(t_f)\right\rangle_{\Omega}-\left\langle C^t\mu(t_0),\frac{\partial T}{\partial u}(t_0)\right\rangle_{\Omega}\\
&\quad+\left\langle K^t\mu,\frac{\partial T}{\partial u}\right\rangle_{\Omega\times[t_0,t_f]}+\left\langle\frac{\partial J}{\partial T},\frac{\partial T}{\partial u}\right\rangle_{\Omega\times[t_0,t_f]}=0.
\end{aligned}
\tag{50}
$$

The adjoint problem consists then in calculating $\lambda=\lambda(r,t),\mu=\mu(r,t)$ that satisfy (51) with initial conditions defined at time t_f given in (52):

$$
\begin{cases}
-C^e\frac{\partial\lambda}{\partial t}+K^e\lambda+\frac{\partial B^t}{\partial E}\mu=0\\
-C^t\frac{\partial\mu}{\partial t}+K^t\mu+\frac{\partial J}{\partial T}=0
\end{cases}
\quad\forall t\in[t_0,t_f],
\tag{51}
$$

$$
\begin{aligned}
&C^e\lambda(t_f)=0,\\
&C^t\mu(t_f)+\frac{\partial J}{\partial T}(t_f)=0.
\end{aligned}
\tag{52}
$$

Taking into account of the objective function definition (35), the derivation of the objective function with respect to temperature writes at node i and time t^k

$$
\left\{\frac{\partial J}{\partial T}(t^k)\right\}_i=\begin{cases}\sum_{elt=1}^{nb.elt}\sum_{\text{if } int\in\Omega^{\text{opt}}(t^k)}^{nb.int}{}_{int=1}(T(r_{int},t^k;u)-T^{\text{opt}}(r_{\text{int}},t^k))\omega_{\text{int}}N_i & \text{if } \xi_k=1,\\ 0 & \text{else,}\end{cases}
\tag{53}
$$

where base functions N_i are previously defined in Section 2.3. Next, the loading component involved in the first equation of (51) writes explicitly at node i

$$
\left\{\frac{\partial B^t}{\partial E}\mu\right\}_i=\sum_{elt=1}^{nb.elt}\int_{elt}2\sigma E\mu N_i\,\mathrm{d}r.
\tag{54}
$$

Due to the sign change on the time operator, both equations involved in (51) have to be integrated backward in time in order to be well posed. Defining a new time variable $\tau=t_f-t$, both adjoint equations have to be solved forward from $\tau=0$ to t_f [33] with thermal discretization that follows $\tau_n=t_f-n\cdot\delta t$. Moreover, in order for the adjoint problem to be perfectly coherent with its direct model, the ultra-weak coupling strategy related to the direct model (33) as well as the time integration scheme (Section 2.4) are transposed to the adjoint. Condensing linear adjoint equations as in (33), we get to solve (55). When more features are integrated within the direct model, the transposition to the adjoint problem must follow. Transposition for more complicated direct models are presented in [45]

$$
\forall k=1,K\begin{cases}R^{\mu}(T^{K-k},\mu^{K-k})=0\\ R^{\lambda}\left(E^{N-n},T^{K-k},\mu^{K-k},\lambda^{N-n}\right)=0\end{cases}\quad\forall n=1,N.
\tag{55}
$$

3.4. The objective function gradient

The objective function gradient being given in (43) and (47), one needs to calculate the scalar product involved in (43). The calculation of the adjoint variable λ is done through the adjoint problem resolution

detailed in Section 3.3. The derivation of the electromagnetic system of equations with respect to controls remains. When controls are continuous, exact analytical derivations are preferred whereas when controls are not sufficiently continuous, one may rather use semi-analytical derivations. Frequency f and current density J_0 being continuous control parameters, one can derive analytically the electromagnetic residual R^e with respect to both controls. This gives explicitly, at node i and time t^k, using the same assembling procedure as for (22):

$$\left\{\frac{\partial R^e}{\partial J_0}(t^k)\right\}_i = \left\{\frac{\partial B^e}{\partial J_0}(t^k)\right\}_i = \sum_{elt=1}^{nb.elts} \int_{elt} -2\pi f \cos(2\pi f t^k) N_i \,\mathrm{d}s, \tag{56}$$

$$\left\{\frac{\partial R^e}{\partial f}(t^k)\right\}_i = \sum_{elt=1}^{nb.elts} \int_{elt} (-2\pi J_0 \cos(2\pi f t^k) + 4\pi^2 J_0 f t^k \sin(2\pi f t^k)) N_i \,\mathrm{d}s. \tag{57}$$

As previously explained in Section 2.5, the inductors are moved virtually and incrementally through a continuous change of physical properties and location of electromagnetic source terms. Thus, due to the space discretization, the electromagnetic residual is no longer continuous with respect to the inductor velocity. One way to handle this difficulty when differentiating the electromagnetic problem R^e with respect to the velocity v is to use semi-analytical derivations which is a compromise between analytical and local finite difference approaches. This approach has been used in various other areas for instance by [46,47]. Let δ be a dimensionless perturbation parameter applied on the velocity v. Then, depending on the current value v

$$\left\{\frac{\partial R^e}{\partial v}\right\} = \begin{cases} \frac{R^e(E,f,J_0,(1+\delta)v) - R^e(E,f,J_0,v)}{\delta \cdot v} & \text{if } v \neq 0, \\ \frac{R^e(E,f,J_0,\delta) - R^e(E,f,J_0,0)}{\delta} & \text{else.} \end{cases} \tag{58}$$

Since electromagnetic equilibrium is reached whatever t, previous relation writes

$$\left\{\frac{\partial R^e}{\partial v}\right\} = \begin{cases} \frac{R^e(E,f,J_0,(1+\delta)v)}{\delta \cdot v} & \text{if } v \neq 0, \\ \frac{R^e(E,f,J_0,\delta)}{\delta} & \text{else.} \end{cases} \tag{59}$$

This vector is assembled using the same assembling procedure as for the direct electromagnetic computation and also at the same times. A proper choice for the perturbation coefficient δ is not trivial. In general, the lowest the perturbation coefficient, the more accurate the calculated residual gradient is. Nevertheless, due to round-off calculations and especially due to the meshing which is obviously not continuous in space, one has to choose a perturbation coefficient δ such that space area differences between perturbed and non-perturbed locations of moving inductors share at least several integration points within the global mesh. Eventually, if the control vector $u = {}^t(f, J_0, v)$ is considered, the objective function gradient writes (60)

$$\nabla_u J = {}^t\left(\left\langle \lambda, \frac{\partial R^e}{\partial f}\right\rangle_{\Omega\times[t_0,t_f]}, \left\langle \lambda, \frac{\partial R^e}{\partial J_0}\right\rangle_{\Omega\times[t_0,t_f]}, \left\langle \lambda, \frac{\partial R^e}{\partial v}\right\rangle_{\Omega\times[t_0,t_f]}\right), \tag{60}$$

with the adjoint variable λ given by the adjoint problem resolution (55), and right-hand side vectors calculated, respectively, by (56), (57) and (59). Since the adjoint variable λ is calculated only on one electromagnetic period per thermal interval, one introduces the scalar χ defined as the ratio between the thermal time step $|t^k - t^{k-1}|$ and the electromagnetic period τ. The objective function gradient thus writes

$$\nabla_u J = \chi \cdot \sum_{k=1}^{K} {}^t\left(\left\langle \lambda, \frac{\partial R^e}{\partial f} \right\rangle_{\Omega\times[t^{k-1},t^{k-1}+\tau]}, \left\langle \lambda, \frac{\partial R^e}{\partial J_0} \right\rangle_{\Omega\times[t^{k-1},t^{k-1}+\tau]}, \left\langle \lambda, \frac{\partial R^e}{\partial v} \right\rangle_{\Omega\times[t^{k-1},t^{k-1}+\tau]} \right). \tag{61}$$

3.5. The global optimization problem resolution

The determination of the objective function gradient enables to use gradient-type methods. Given an initial set of controls u^0, one builds a series defined by

$$u^k = u^{k-1} + \alpha^k d^k, \tag{62}$$

where d^k is the direction of descent and α^k is the descent step size. At each iteration k the direction of descent d^k can be computed in various ways. The simplest method consists in choosing the so-called steepest descent [14] where

$$d^k = -\nabla_{u^k} J. \tag{63}$$

The main disadvantage of using the greatest slope algorithm comes from its low convergence as u^k gets close to the minimum [36,50,51]. In order to avoid erratic characteristics of steepest descent methods, an alternative consists in developing a conjugate gradient method applied on arbitrary functions. The best-known methods are the Fletcher and Reeves's [52] and the Polak–Ribiére's [53]. These methods are known to be very interesting, on one hand because they require storage of very little information (when compared to Newton or quasi-Newton methods [51]) and because their rate of convergence are very superior to that of ordinary gradient methods. The descent direction is calculated as follows:

$$\begin{aligned} d^k &= -\nabla_{u^k} J \quad \text{if } k = 1, \\ d^k &= -\nabla_{u^k} J + \beta^k d^{k-1} \quad \text{else}, \end{aligned} \tag{64}$$

where β^k is chosen such that d^k is conjugated with respect to d^{k-1}. The used Polak–Ribière method uses

$$\beta^k = \begin{cases} 0 & \text{if } k = 0, \\ \dfrac{\langle \nabla J(u^k), \nabla J(u^k) - \nabla J(u^{k-1}) \rangle}{\langle \nabla J(u^{k-1}), \nabla J(u^{k-1}) \rangle} & \text{else.} \end{cases} \tag{65}$$

Conjugate gradient methods are widely used in practice [33,54]. Eventually, the determination of an optimal descent step size is performed through the resolution of

$$\alpha^k = \underset{\alpha}{\mathrm{ArgMin}}\, j\left(u^{k-1} + \alpha d^k\right). \tag{66}$$

The resolution of problem (66) which is nothing more than the minimization of an application from $\Re$ into $\Re$ may be very time expensive since it needs, for the nonlinear model, several evaluations of J and thus several integrations of the direct model. Optimization codes libraries as presented in [48] for instance are widely used. The linear research algorithm which has been chosen relies on a dichotomial research coupled to a parabolic interpolation [19,49].

At each iteration k, the optimization procedure consists in, with a set of controls u^k,

1. integrating the objective function value (35) through the integration of the forward nonlinear direct model (32); store all state variables;
2. integrating the backward linear adjoint problem (55) with initial condition (52), all matrices being recomputed from stored state variables;
3. calculating the objective function gradient (60), the direction of descent (64);
4. solving the linear research algorithm (66) through several integrations of the nonlinear direct model.

This global optimization procedure is run until convergence criteria are reached. For the considered optimization problem, criteria are based on the objective function value, its evolution and its gradient

$$\begin{aligned} &J(u^k) \leqslant \chi_1, \\ &J(u^{k-1}) - J(u^k) \leqslant \chi_2, \\ &\left|\nabla_{u^k} J(u^k)\right| \leqslant \chi_3, \end{aligned} \tag{67}$$

where the choice for χ_i, $i = 1, \ldots, 3$, may be adjusted depending on whether a fine or a rough (for initialization for instance) optimization procedure is performed. In a practical point of view, it is not easy to choose χ_1 since it is to be chosen in accordance with the considered optimization problem. Thus, convergence criteria has been preferred [43] rather than on the functional value itself, this latter criterion being rather used when dealing with inverse analysis problems [55,56].

4. Numerical optimization results

Different process controls have been tested. For the sake of clarity, we have kept the same geometry for all cases. We present in Figs. 6 and 7 the used geometry [32] and the related mesh. The physical parameters of the magnetic EN3 steel billet are the following [32]. The relative magnetic permeability equals 90 at 0 K, with a temperature sensitivity of 6 (see formulation in [19]). The electrical conductivity equals 3.6 MS m^{-1} at 0 °C, 1 MS m^{-1} at 700 °C, and 0.7 MS m^{-1} at 1200 °C. The thermal conductivity equals 45 W m^{-1} K^{-1} at 0 °C, 42 W m^{-1} K^{-1} at 400 °C, 34 W m^{-1} K^{-1} at 700 °C, 27 W m^{-1} K^{-1} at 800 °C and 30 W m^{-1} K^{-1} at 1200 °C. The heat capacity equals 3.7 Mj m^{-3} K^{-1} at 0 °C, 6 Mj m^{-3} K^{-1} at 700 °C, 13.5 Mj m^{-3} K^{-1} at 770 °C, 5.2 Mj m^{-3} K^{-1} at 800 °C and 5.7 Mj m^{-3} K^{-1} at 1200 °C. The heat transfer problem is computed only in the part to be heated ($\Omega^t = \Omega_{\text{part}}$). Since the part is insulated as shown in Fig. 6, a null flux is

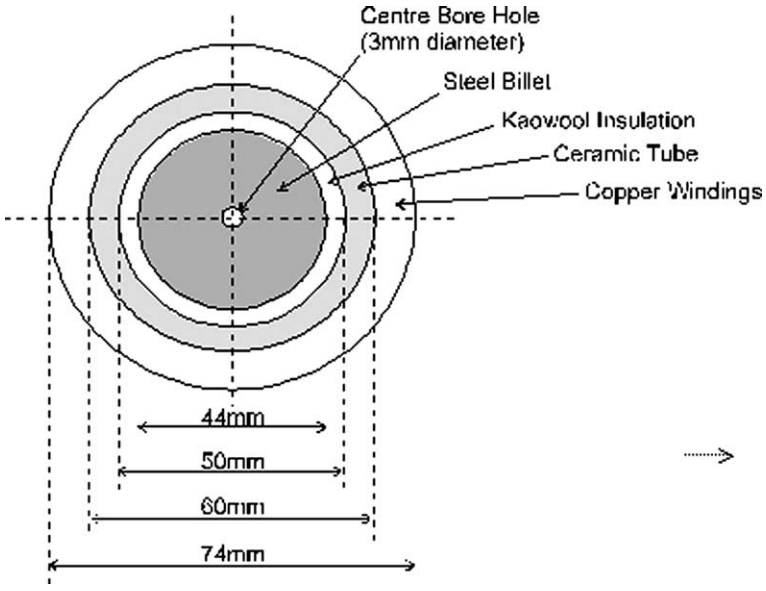

Fig. 6. Used geometry. Main characteristics are: billet height – 120 mm; number of layers for copper windings – 3; number of turns on inner layer – 55; number of turns on middle layer – 54; number of turns on outer layer – 55; diameter of copper wire – 2 mm; length of coil on former – 117 mm.

prescribed on the external surface where $r = 0.022$ m, and a free convection and radiation condition is prescribed on upper and lower surfaces where $z = 0.15$ m and $z = 0.27$ m. Electromagnetic parameters for the air and copper inductors are taken from [57].

4.1. Test case 1

The first test case consists in reaching, on the surface, within the radius range (0.018 m; 0.022 m) (see Fig. 7), an average optimal temperature of 1500 °C after 10 s of heating through the control of the frequency f and the current density J_0. Initial guessed frequency f^0 is 500 Hz, and initial current density in the coil J_0^0 is 10^9A m^{-2}. The general objective function writes

$$J(u) = \int_{\Omega_{\mathrm{opt}}(t_f)} \left(T(\omega, t_f; u) - T^{\mathrm{opt}}(\omega, t_f)\right)^2 \mathrm{d}\omega, \tag{68}$$

where

$$u = \left\{(f, J_0) \in \Re^+ \times \Re^+\right\}; \quad t_f = 10 \text{ s}; \quad T^{\mathrm{opt}}(\omega, t_f) = 1773 \text{ K};$$

$$\Omega_{\mathrm{opt}}(t_f) = \{\omega(r,z), 0.018 \text{ m} \leqslant r \leqslant 0.02 \text{ m and } 0.15 \text{ m} \leqslant z \leqslant 0.27 \text{ m}\}.$$

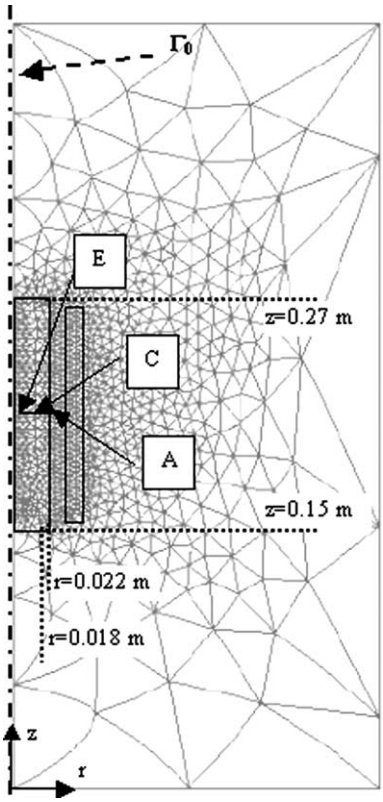

Fig. 7. Mesh used for all test cases and location points A, C and E. Non-presented points B and D are, respectively, in between points A and C, and C and E.

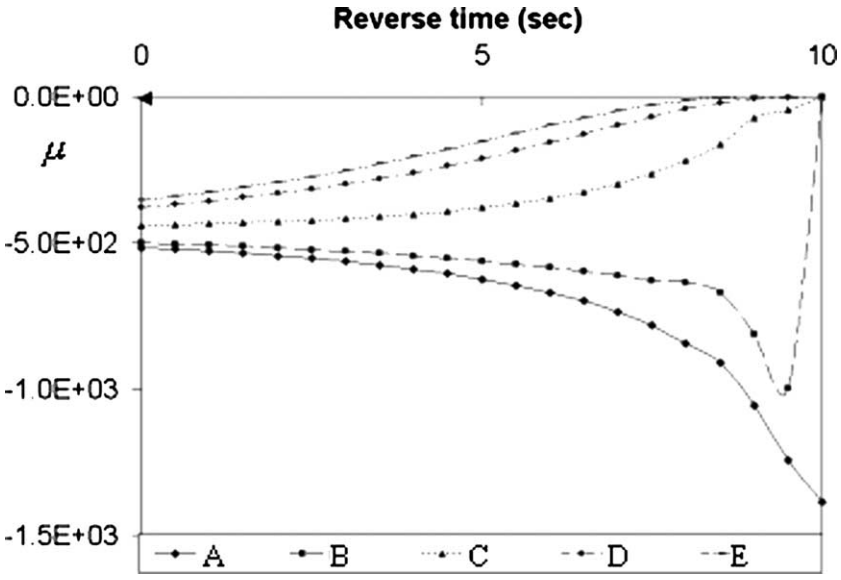

Fig. 8. Backward evolution of the adjoint variable $\mu(r,t)$ on locations A, B, C, D and E (see Fig. 7).

The direct electromagnetic and thermal equations (32) are first computed. The objective function (35) is then calculated along with calculation of the derivation of it with respect to temperature at final time (53). Linear parabolic adjoint equations (55) are then backward integrated. Figs. 8 and 9 present the backward evolutions of adjoint variables μ and λ. Fig. 10 presents evolutions of the calculated process parameters f and J_0 with respect to iterations k, and Fig. 11 presents the evolution of the decreasing objective function value $J(u^k)$ with k.

Iteration 0 is directly related to the first guessed parameters (f^0, J_0^0). The objective function $J(u^0)$ is calculated, as well as its gradient $\nabla_{u^0} J$. Iterations 1–5 are still related to the first minimization loop, where

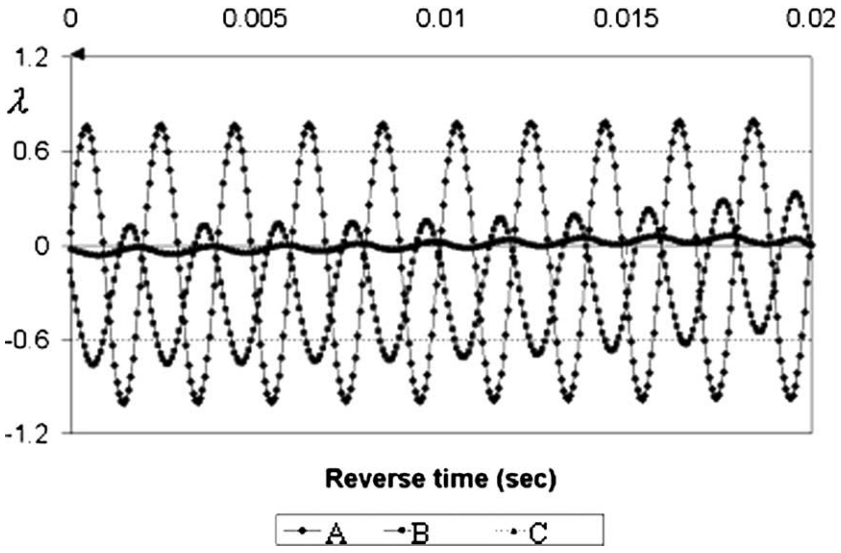

Fig. 9. Backward evolution of the adjoint variable $\lambda(r,t)$ on locations A, B and C.

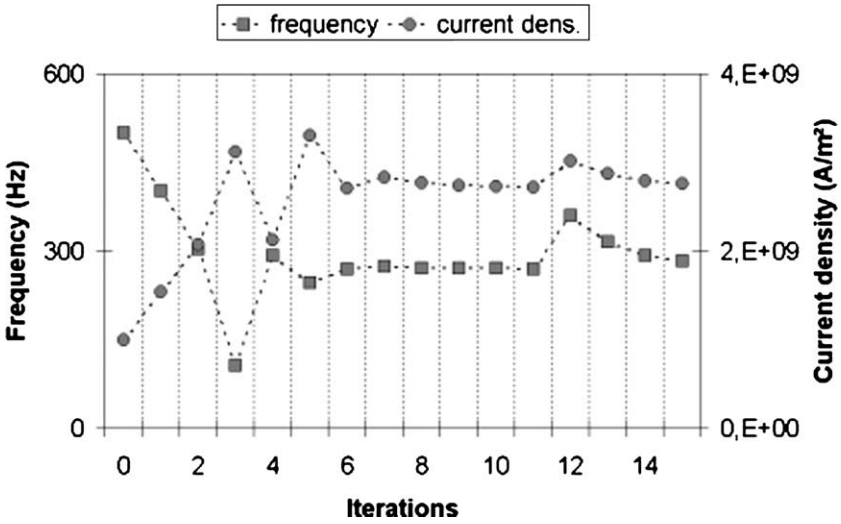

Fig. 10. Control values $u^k = \{(f^k, J_0^k) \in \Re^+ \times \Re^+\}$ evolution with respect to iterations k.

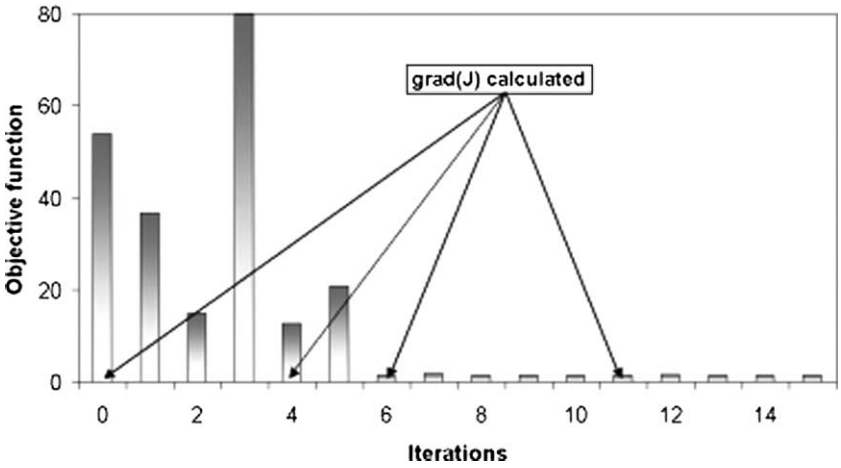

Fig. 11. Objective function value evolution $J(u^k)$ with respect to iterations k.

the line search algorithm (66) is running. Then, the gradient is calculated once again. Following iterations are related to the second and third external loops where objective function and its gradient are calculated followed again by the line search algorithm. It is to be seen that only 3 objective function gradient calculations and 16 runs of the direct model are needed to decrease the objective function by a factor greater than 40.

4.2. Test case 2

This second test consists in minimizing the objective function (69). Figs. 12 and 13 present evolutions of controls u^k and of the objective function value $J(u^k)$ with respect to iterations k. Eight iterations are needed to decrease the functional by a factor 20.

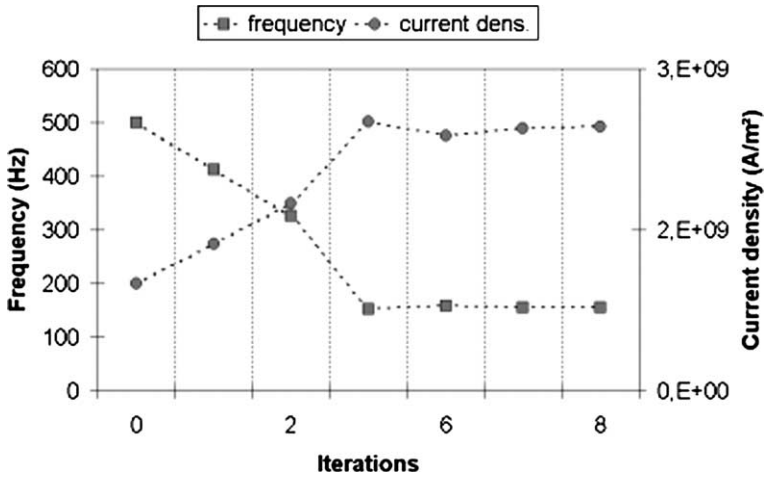

Fig. 12. Control values $u^k = \{(f^k, J_0^k) \in \Re^+ \times \Re^+\}$ evolution with respect to iterations k.

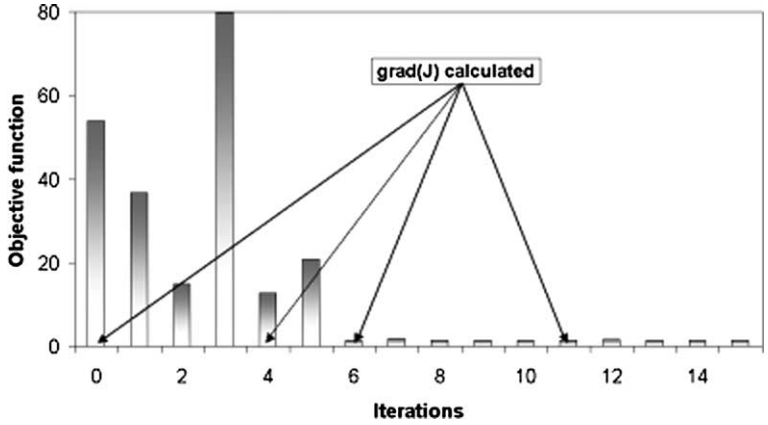

Fig. 13. Objective function value evolution $J(u^k)$ with respect to iterations k.

$$J(u) = \int_{\Omega^{\text{opt}}} \left(T(\omega, t_1; u) - T^{\text{opt}}(\omega, t_1)\right)^2 \mathrm{d}\omega + \int_{\Omega^{\text{opt}}} \left(T(\omega, t_2; u) - T^{\text{opt}}(\omega, t_2)\right)^2 \mathrm{d}\omega + \int_{\Omega^{\text{opt}}} \left(T(\omega, t_f; u) - T^{\text{opt}}(\omega, t_f)\right)^2 \mathrm{d}\omega, \tag{69}$$

where

$$u = \{(f, J_0) \in \Re^+ \times \Re^+\}; \quad t_1 = 2.5 \text{ s}; \quad t_2 = 3.5 \text{ s}; \quad t_f = 5 \text{ s};$$

$$T^{\text{opt}}(\omega, t_1) = 850 \text{ K}; \quad T^{\text{opt}}(\omega, t_2) = 1030 \text{ K}; \quad T^{\text{opt}}(\omega, t_f) = 1273 \text{ K};$$

$$\Omega^{\text{opt}} = \{\omega(r, z), 0.018 \text{ m} \leqslant r \leqslant 0.02 \text{ m and } 0.15 \text{ m} \leqslant z \leqslant 0.27 \text{ m}\}.$$

4.3. Test case 3

The sample is heated during 20 s and cooled down freely for 10 s. The present aim is to find optimal frequency and current density such that the final temperature is as uniform as possible and the closest as possible to 800 °C. The objective function to be minimizes writes (70). Figs. 14 and 15 present evolutions of controls u^k and of the objective function value $J(u^k)$ with respect to iterations k. Six iterations are needed to decrease drastically the functional by a factor greater than 100.

$$J(u) = \int_{\Omega^{\mathrm{opt}}} \left(T(\omega, t_f; u) - T^{\mathrm{opt}}(\omega, t_f)\right)^2 \mathrm{d}\omega, \tag{70}$$

where

$$u = \left\{(f, J_0) \in \Re^+ \times \Re^+, 0 \leqslant t \leqslant t_1\right\}; \quad t_1 = 20 \text{ s}; \quad t_f = 30 \text{ s};$$

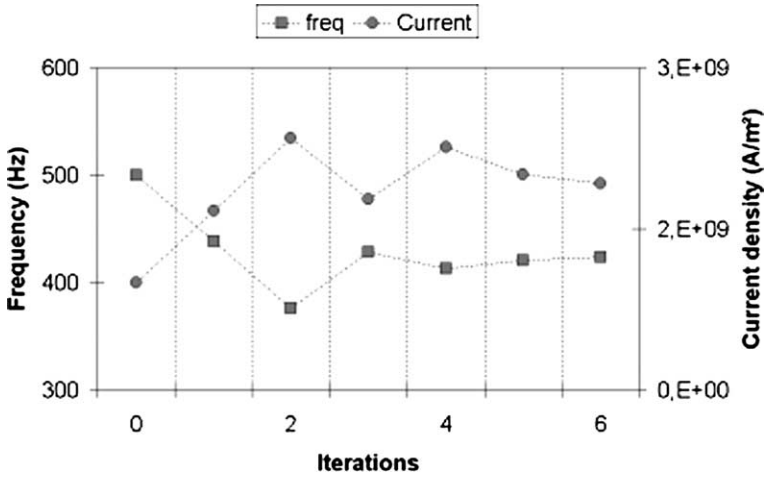

Fig. 14. Control values $u^k = \{(f^k, J_0^k) \in \Re^+ \times \Re^+\}$ evolution with respect to iterations k.

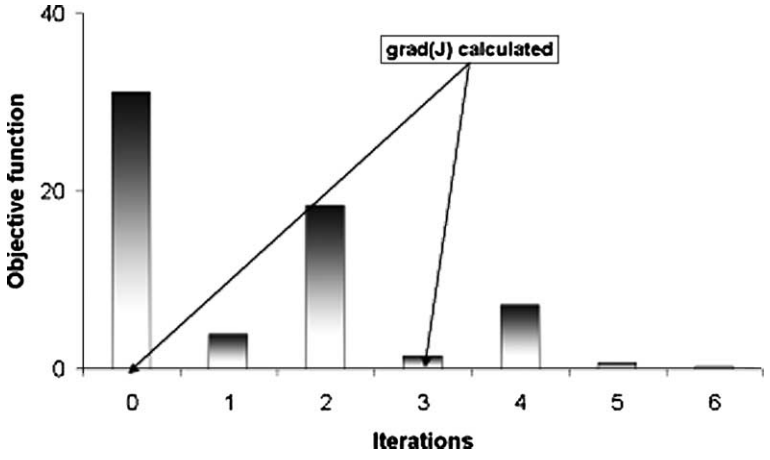

Fig. 15. Objective function value evolution $J(u^k)$ with respect to iterations k.

$$T^{\mathrm{opt}}(\omega, t_f) = 1073 \text{ K}; \quad \Omega^{\mathrm{opt}} = \Omega_{\mathrm{part}}.$$

4.4. Test case 4

This fourth test case uses the same global geometry as for all previous test cases except that the inductor is moving along the z-axis as shown in Fig. 16. The aim here is to find $u = \{(f, J_0, vit) \in \Re^+ \times \Re^+ \times \Re^+\}$ such that, after 10 s of heating, the surface between $z = 22.5$ and $z = 24.7$ cm (see Fig. 16) is as close as possible to 850 K. The objective function to be minimized in given by (71). Fig. 17 presents the evolutions, with respect to iterations, of frequency, input current and coil velocity, and Fig. 18 presents the evolution of the objective function value. Only eight full calculations are needed for decreasing the objective function by a factor of almost 100.

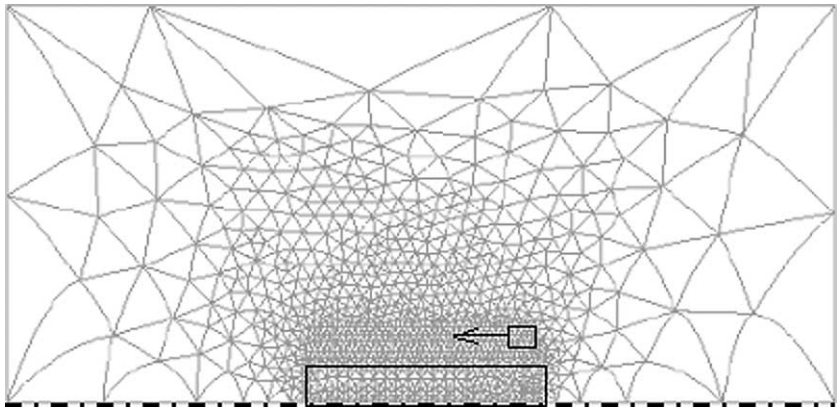

Fig. 16. Geometry used in test case 4. The inductor is initially moving at velocity $vit^0 = 10^{-2}$ m s^{-1}.

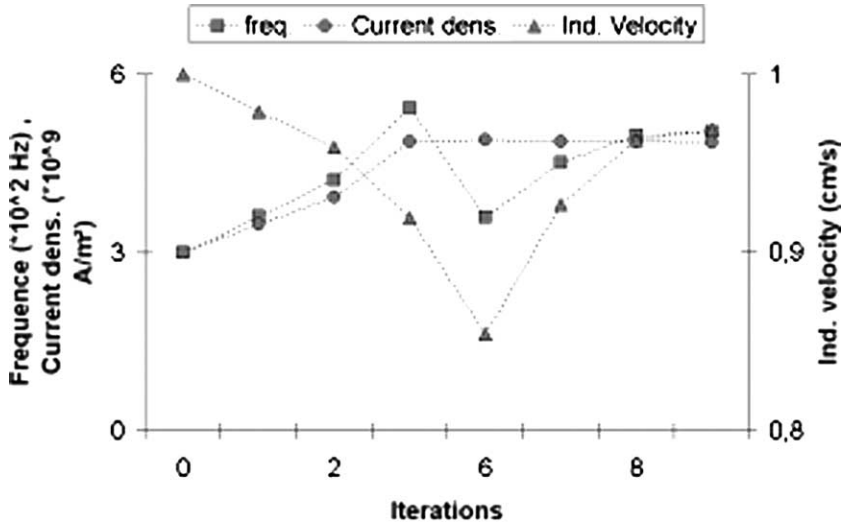

Fig. 17. Control values $u^k = \{(f^k, J_0^k, vit^k) \in \Re^+ \times \Re^+ \times \Re^+\}$ with respect to iterations k.

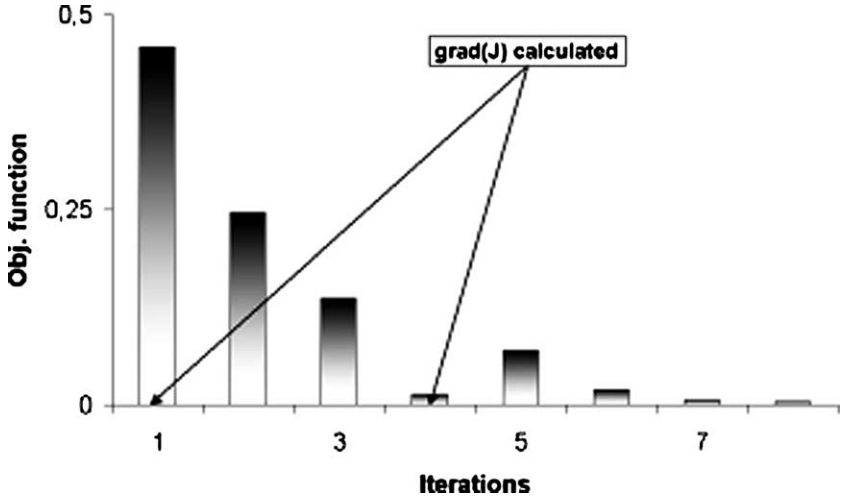

Fig. 18. Objective function value evolution $J(u^k)$ with respect to iterations k.

$$J(u) = \int_{\Omega^{\text{opt}}} \left(T(\omega, t_f; u) - T^{\text{opt}}(\omega, t_f)\right)^2 \mathrm{d}\omega, \tag{71}$$

where

$$u = \left\{(f, J_0, vit) \in \Re^+ \times \Re^+ \times \Re^+\right\}; \quad t_f = 15 \text{ s}; \quad T^{\text{opt}}(\omega, t_f) = 850 \text{ K};$$

$$\Omega^{\text{opt}} = \{\omega(r,z),\ 0.018 \text{ m} \leqslant r \leqslant 0.02 \text{ m and } 0.225 \text{ m} \leqslant z \leqslant 0.247 \text{ m}\}$$

5. Conclusion

A complete optimization model coupling electromagnetism and heat transfer phenomena has been presented. It is based on a gradient-type method related to an ultra-weak coupling between both non-linear time-dependent problems, and thus enables dealing with magnetic materials very accurately. Numerical optimization tests performed on various distinct objective functions have shown the efficiency and the robustness of the proposed method where only several iterations are needed to find optimality conditions.

So far, all control parameters have been taken constant in time. A direct extension of the proposed adjoint method consists in finding time-dependent optimal control parameters. From the algorithmic point of view, very little development is of concern since just the time integration definition domain has to be changed in the scalar product (43) and thus in (60). Nevertheless, this should lead to an infinity of optimality solutions especially when dealing with functionals defined at final time. Therefore, inequality constraints on process parameters will have to be added along with a supplementary term related to control parameters evolutions added to the objective function. This method has proved its efficiency in other domains as in the control of ultrasound heating [14].

In the same spirit, realistic thermo-electromagnetic processes simulations often need a very fine mesh on magnetic piece surfaces. To deal with such cases, the global optimization procedure has been parallelized through the single program multiple domains (SPMDs) method [58]. In order to save again on time

computation, the next step may be to use the ultra-weak variational formulation (UWVF) applied on the electromagnetic space discretization instead of the finite element approach [59,60]. Though this new method will drastically change the direct model space discretization, derivation of the direct model being performed on the discretized problem, the whole proposed adjoint method will remain unchanged.

Acknowledgements

The authors thank the European Community for its financial support provided for this work (Esprit Project No. 28283).

References

[1] E.J. Davies, Conduction and Induction Heating, Peregrinus, London, 1990.
[2] S. Zinn, S.L. Semiatin, Elements of induction heating – Design, control and applications, ASM International, Electronic Power Research Institute, 1988.
[3] C. Chaboudez, S. Clain, R. Glardon, D. Mari, J. Rappaz, M. Swierkosz, Numerical modeling in induction heating for axisymmetric geometries, IEEE Trans. Magn. 33 (1) (1997) 735–745.
[4] C. Marchand, A. Foggia, A 2-D finite element program for magnetic induction heating, IEEE Trans. Magn. 19 (6) (1983) 2647–2649.
[5] K.F. Wang, S. Chandrasekar, H.T.Y. Yang, Finite element simulation of induction heat treatment, J. Mater. Eng. Perform. 1 (1) (1992).
[6] T. Myoshi, M. Sumiya, H. Omori, Analysis of an induction heating system by the finite element method combined with the boundary integral equation, IEEE Trans. Magn. 23 (2) (1987) 1827–1832.
[7] N. El-Kaddah, R. Craen, W. Loue, A mathematical model of induction heating of thixoformable aluminium billets, 127th TMS Annual Meeting, San Antonio, February 15–19, 1998.
[8] S.J. Salon, J.M. Schneider, A hybrid finite element–boundary integral formulation of eddy current problem, IEEE Trans. Magn. MAG-18 (1982) 461–466.
[9] J. Rappaz, M. Swierkosz, Mathematical modeling and simulation of induction heating processes, Appl. Math. Comp. Sci. 6 (2) (1996) 207–221.
[10] F. Bay, V. Labbé, Y. Favennec, J.L. Chenot, A numerical model for induction heating processes coupling electromagnetism and thermomechanics, Int. J. Numer. Meth. Eng., to be published.
[11] O. Bodart, A.V. Boureau, R. Touzani, Numerical investigation of optimal control of induction heating processes, Appl. Math. Modeling 25 (2001) 697–712.
[12] T.H. Pham, S.R.H. Hooles, Unconstrained optimization of coupled magneto-thermal problems, IEEE Trans. Magn. 31 (3) (1995) 1988–1991.
[13] M. Battistetti, P. Di Barba, F. Dughiero, M. Farina, S. Lupi, A. Savini, Multiobjective design optimization of an inductor for surface heating: an innovative approach, in: 9th IEEE Conference on Electromagnetic Field Computation, Milwaukee, USA, June 2000.
[14] M. Malinen, T. Huttunen, J.P. Kaipio, Optimal control for the ultrasound induced heating of a tumor, in: 4th International Conference on Inverse Problems in Engineering, Rio de Janeiro, Brasil, 2002.
[15] P. VanBaren, E.S. Ebbini, Multi-point temperature control during hyperthermia treatments: theory and simulations, IEEE Trans. Bio. Eng. (1995).
[16] A. Bossavit, Méthodes numériques en électromagnétisme, Eyrolles, Collection EDF, 1991.
[17] V. Labbé, Modélisation numérique du chauffage par induction, Approche éléments finis et calcul parallèle, Ph.D. Thesis, Ecole des Mines de Paris, 2002.
[18] H. Gie, J.P. Sarmant, Electromagnétisme, vols. 1&2, Lavoisier, Coll. Des Techniques et Sciences, 1985.
[19] Y. Favennec, V. Labbé, F. Bay, Identification of magnetic parameters by inverse analysis coupled with finite element modeling, IEEE Trans. Magn. 38 (6) (2002).
[20] R.M. Bozorth, Ferromagnetism, Van Nostrand, New York, 1951.
[21] C. Guérin, Détermination des pertes par courant de foucault dans les cuves de transformateur, Ph.D. Thesis, INP Grenoble, France, 1994.
[22] A. Bossavit, J.C. Vérité, The Trifou code: solving the 3-D eddy currents problem by using as state variable, IEEE Trans. Magn. 19 (6) (1983) 2465–2470.

[23] C. Saltiel, M. Leon, Thermal and stress analysis during diffusion bonding of metal drum rotor disks using electromagnetic heating, Numer. Heat Transfer A (1996) 35–54.
[24] N. Soyris, L. Fourment, T. Coupez, J.P. Cescutti, G. Brachotte, J.L. Chenot, Forging of a connecting rod : 3D finite element calculation, Eng. Comput. 9 (1) (1992) 63–80.
[25] A. Quarteroni, A. Valli, in: Numerical Approximation of Partial Differential Equations, Springer Series in Computational Mathematics, Berlin, 1994.
[26] M.A. Hogge, A comparison of two and three level integration schemes for non-linear heat conduction, in: R.W. Lewis, K. Morgan, O.C. Zienkiewicz (Eds.), Numerical Methods in Heat Transfer, Wiley, New York, 1981, pp. 75–90.
[27] M. Zlamal, Finite element methods for nonlinear parabolic equations, RAIRO Numer. Anal. 11 (1) (1977).
[28] T. Coupez, J.L. Chenot, Larges deformations and automatic remeshing, in: D.J.R. Owen, et al. (Eds.), Computational Plasticity (COMPLASIII), Pineridge Press, 1992, pp. 1077–1088.
[29] T. Coupez, A mesh improvement method for 3D automatic remeshing, in: N.P. Weatherill, et al. (Eds.), Numerical Grid Generation in Computational Fluid Dynamics and related Fields, Pineridge Press, 1994, pp. 615–626.
[30] B. Stok, N. Mole, Coupling FE and BE approach in axisymmetric eddy current problems solution, in: Int. Conf. on Comput. Engn & Science – ICES'01, Puerto Vallarta, Mexico, 2001.
[31] N. Mole, B. Stok, Coupled FEM–BEM modelling in view of heat induction processes simulation, in: 8th Seminar of the International Federation for Heat Treatment and Surface Engineering – IFHTSE2001 Dubrovnik, Croatia, 2001.
[32] A. Dexter, Heatmaster Workpackage 3: Experimental validation, EA-Technology Report No. 5348, Project No. 42020, October 2000.
[33] Y. Jarny, M.N. Ozisik, J.P. Bardon, A general optimization method using adjoint equation for solving multidimensional inverse heat conduction, Int. J. Heat Mass Transfer 34 (11) (1991) 2911–2919.
[34] L. Fourment, D. Vieilledent, J.L. Chenot, Shape optimization of the axisymmetric preform tool in forging using a direct differentiation method, Int. J. Forming Processes 1 (4) (1998) 399–423.
[35] D. Hömberg, J. Sokolowski, Optimal control of laser hardening, Adv. Math. Sci. Appl., Gakkotosho, Tokyo 8 (1998) 911–928.
[36] P.E. Gill,, W. Murray, M.H. Wright, Practical Optimization, Academic Pres, London, 1992.
[37] F.X. Le Dimet, J. Blum, Assimilation de données pour les fluides géophysiques, Matapli, No. 67, 2002.
[38] J.R. Roche, Gradient of the discretized energy method and discretized, Appl. Math. Comput. Sci. (3) (1997) 545–565.
[39] P. Michaleris, D.A. Tortorelli, C. Vidal, Tangent operator and design sensitivity formulations for transient non-linear coupled problems with applications to elastoplasticity, Int. J. Numer. Meth. Eng. 37 (1994) 2471–2499.
[40] D.A. Tortorelli, P. Michaleris, Design sensitivity analysis: overview and review, Inverse Probl. Eng. 1 (1994) 71–105.
[41] A. Anju, M. Kawahara, Comparison of sensitivity equation and adjoint equation method for parameter identification problems, Int. J. Numer. Meth. Eng. 40 (1997) 1015–1024.
[42] K.A. Woodbury, Inverse Analysis Handbook, CRC Press, Boca Raton, FL, 2002.
[43] J. Céa, Optimisation, théorie et algorithmes, Dunod, 1971.
[44] B. Larrouturou, P.L. Lions, Méthodes mathématiques pour les sciences de l'ingénieur: optimisation et analyse numérique, Cours de l'Ecole Polytechnique, 1994.
[45] Y. Favennec, V. Labbé, F. Bay, An efficient coupling strategy for optimizing coupled periodical electromagnetic and heat transfer problems, Int. J. Numer. Meth. Eng. (submitted).
[46] E.S. Jochen, D.A. Tortorelli, Inverse heat conduction problem solutions via second order design sensitivities and Newton's method, Inverse Probl. Eng. 2 (1995) 227–262.
[47] B. Boyer, E. Massoni, Inverse analysis for identification of parameters during thermomechanical tests, Numiform 2001, Toyohashi University of Technology, Japan, June 2001.
[48] J.C. Gilbert, C. Lemarechal, Some numerical experiments with variable storage quasi-Newton algorithms, Math. Prog. B 25 (1989) 407–435.
[49] K. Schittkowski, Nonlinear Programming Codes. Informations, Test, Performance, Lecture Notes in Economic and Mathematical System, Springer, Berlin, 1980.
[50] H. Curry, The method of steepest descent for nonlinear minimization problems, Quart. Appl. Math. (2) (1944) 258–261.
[51] M. Minoux, Mathematical Programming, Theory and Applications, Wiley, Chichester, 1986.
[52] R. Fletcher, C.M. Reeves, Function minimization by conjugate gradients, Comput. J. (7) (1964) 149–154.
[53] E. Polak, Computational Methods in Optimization. A Unified Approach, Academic Press, New York, 1971.
[54] N. Zabaras, S. Kang, Control of the freezing interface motion in two-dimensional solidification processes using adjoint method, Int. J. Numer. Meth. Eng. (38) (1995) 63–80.
[55] J.V. Beck, B. lackwell, C.R.St. Clair, Inverse Heat Conduction, Wiley, New York, 1985.
[56] A.N. Tikhonov, V.Y. Arsenin, Solutions of Ill-posed Problems, Winston, Washington, DC, 1977.
[57] D. Pollock, Physical Properties of Materials for Engineers, second ed., CRC Press, Boca Raton, FL, 1993.
[58] V. Labbé, Y. Favennec, F. Bay, A parallel computing strategy of an optimization procedure designed for electro-thermal applications, J. Parallel Distrib. Comput. (submitted).

[59] O. Cessenat, B. Despres, Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem, SIAM J. Numer. Anal. (136) (1997) 68–82.
[60] T. Huttunen, P. Monk, J.P. Kaipio, Computational aspects of the ultra weak variational formulation, J. Comput. Phys. (accepted).