

SQP Methods for Parameter Identification Problems arising in Hyperthermia

T. Gänzler, Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), Takustraße
 7, D- 14195 Berlin, Germany

S. Volkwein, Institut für Mathematik und Wissenschaftliches Rechnen,
 Karl-Franzens-Universität Graz, Heinrichstraße 36, A-8010 Graz, Austria

M. Weiser, Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), Takustraße
 7, D- 14195 Berlin, Germany

(December 2004)

The paper introduces an identification problem arising in modern regional hyperthermia, a cancer therapy aiming at heating the tumor by microwave radiation. The task is to identify the highly individual perfusion, which affects the resulting temperature distribution, from MR measurements. The identification problem is formulated as an optimization problem. Existence of a solution and optimality conditions are analyzed. Different regularizations and problem variants are considered. For the numerical solution, a standard SQP method is used. Sufficient conditions for the convergence of the method are derived. Finally, numerical examples on artificial as well as clinical data are presented.

AMS MSC 2000: 65N21,92C55,49M15

Keywords: Parameter identification, optimality conditions, SQP methods

1 Introduction

Regional hyperthermia [3] is a cancer therapy that aims at heating the tumor above a therapeutic temperature of about 41°C in order to make it more susceptible to an accompanying radio or chemo therapy. Energy is injected into the human body by microwave radiation originating from a high frequency applicator. The resulting temperature is described by the elliptic bio heat transfer equation [15]

$$\begin{aligned} -\operatorname{div}(\kappa \nabla u) + w(u - u_a) &= f \quad \text{in } \Omega, \\ \gamma \frac{\partial u}{\partial n} + u &= g \quad \text{on } \Gamma = \partial \Omega, \end{aligned} \tag{1}$$

where the specific absorption rate (SAR) enters as the right hand side f . The Helmholtz term models the cooling of tissue by perfusion by arterial blood with a constant systemic temperature of $u_a = 37^\circ\text{C}$. The body Ω is a bounded open domain in \mathbf{R}^3 with boundary Γ . Its different tissue types occupy finitely many subdomains Ω_i , $i = 1, \dots, n_\Omega$, satisfying

$$\bar{\Omega} = \bigcup_{i=1}^{n_\Omega} \bar{\Omega}_i, \quad \Omega_i \cap \Omega_j = \emptyset \quad \text{for } i \neq j. \tag{2}$$

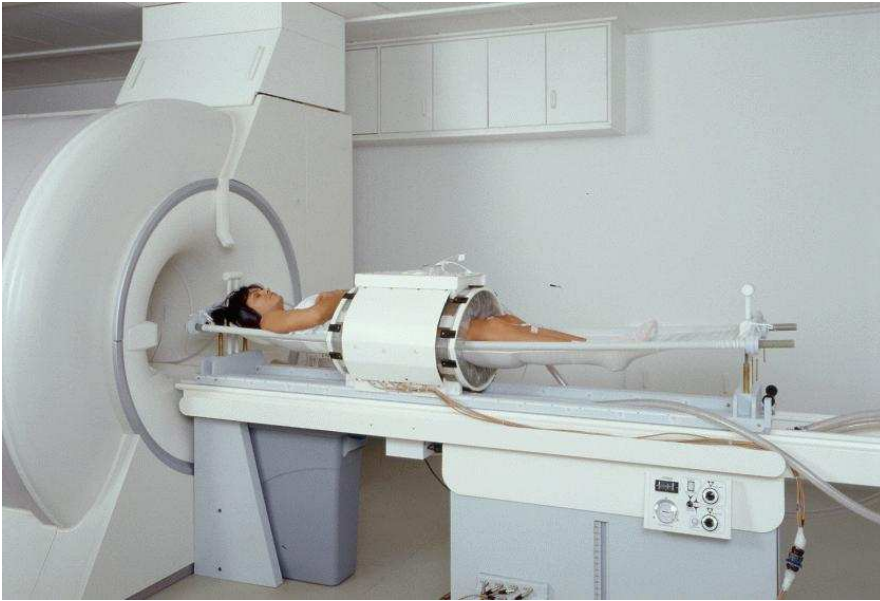


Figure 1. Combined MR hyperthermia applicator (Charité, Berlin Buch).

We assume the temperature diffusion coefficient $\kappa(\xi) \geq \underline{\kappa} > 0$ to be piecewise constant and the perfusion w to be piecewise continuous on the subdomains. The right-hand side f belongs to $L^2(\Omega)$ and $g \in L^2(\Gamma)$ is satisfied for the boundary data.

In therapy planning, the antenna parameters for the applicator are determined for each individual patient in order to achieve a therapeutically optimal temperature distribution. SAR values are obtained by solving the time-harmonic Maxwell equations, and the temperature distribution is predicted by (1) using a reference perfusion $w = \bar{w}(\xi, u(\xi))$ (see [4]). Due to the self-regulating capability of the tissue, this reference value depends on the temperature u .

Unfortunately, the actual perfusion value varies significantly between different patients, and even between different therapy sessions for the same patient. For a reliable optimization of the therapy parameters, the identification of the actual perfusion value is necessary.

Modern combined magnetic resonance (MR) hyperthermia applicators (see Fig. 1) now provide for the first time online distributed 3D measurements of temperature and perfusion. To the current best knowledge, the data m provided by MR measurements can be modeled by

$$m \approx a(u - u_{\text{bas}}) + b(w - w_{\text{bas}}) \quad \text{in } \Omega_m \subset \Omega, \tag{3}$$

where the piecewise constant and nonnegative sensitivities of m with respect to u and w , respectively, satisfy $a \approx b$ in muscle and $a \ll b$ in fat. $u_{\text{bas}}, w_{\text{bas}} \in L^2(\Omega)$ are known offsets. Due to measurement time constraints and artefacts in the data caused by motion of the patient and inhomogeneities of the magnetic field, however, the data are subject to noise and moreover cover only a subdomain Ω_m of the human body. In favor of a simpler presentation we extend the measurement parameters a, b , and m from Ω_m to Ω with value zero.

In this article we consider the identification of a nonnegative perfusion $w = w(\xi)$ from (1) and (3). We impose a Tichonov regularization for w . Since the perfusion can be assumed to depend continuously on the temperature in each subdomain Ω_i , but is typically discontinuous at tissue boundaries, we employ a piecewise H^1 -norm for regularization.

Parameter identification problems for elliptic PDEs were studied by several authors. For example, in the context of SQP methods in Hilbert space we refer to [12, 14] for reduced SQP methods, for the augmented Lagrangian algorithm to [7, 8, 11, 13] and to [9, 16] for augmented Lagrange-SQP methods.

The paper is organized as follows. In §2, we formulate the identification problem as an optimal

control problem and prove existence of a solution. The SQP algorithm is formulated in §3. § 4 is devoted to numerical examples. Finally, we draw some conclusions in §5.

2 Optimality conditions

In this section we formulate the parameter identification problem as an optimal control problem and present necessary optimality conditions for an optimal solution. Throughout we suppose that $\Omega \subset \mathbf{R}^d$ with $d \in \{1, 2, 3\}$ satisfies a uniform interior cone condition (that is, there exists a fixed cone K_Ω such that each $\xi \in \Gamma$ is the vertex of a cone $K_\Omega(\xi) \subset \bar{\Omega}$ congruent to K_Ω).

2.1 The optimal control problem

In view of (2) let us introduce the inner product

$$\langle w, \tilde{w} \rangle_W = \sum_{i=1}^{n_\Omega} \int_{\Omega_i} (\alpha w \tilde{w} + \beta \nabla w \cdot \nabla \tilde{w}) \, d\xi \quad \text{for } w, \tilde{w} \in H^1(\Omega) \quad (4)$$

with scalar parameters $\alpha > 0, \beta \geq 0$ and the induced norm $\|w\|_W = \sqrt{\langle w, w \rangle_W}$. We define the Hilbert space W to be the closure of $H^1(\Omega)$ with respect to $\|\cdot\|_W$. For $\beta = 0$ we obtain $W = L^2(\Omega)$ and

$$W = \{w \in L^2(\Omega) : w|_{\Omega_i} \in H^1(\Omega_i) \text{ for } i = 1, \dots, n_\Omega\}$$

for $\beta > 0$. Throughout the remainder of the paper we suppose $\beta > 0$.

To write the elliptic differential equation (1) in a compact form we define the operator $\tilde{e} : H^1(\Omega) \times W \rightarrow H^1(\Omega)'$ by

$$\langle \tilde{e}(u, w), \phi \rangle_{(H^1)'} = \int_{\Omega} \kappa \nabla u \cdot \nabla \phi + (w(u - u_a) - f) \phi \, d\xi + \frac{1}{\gamma} \int_{\Gamma} \kappa(u - g) \phi \, ds$$

for all $\phi \in H^1(\Omega)$ and for $(u, w) \in H^1(\Omega) \times W$, where $\langle \cdot, \cdot \rangle_{(H^1)'}$ denotes the dual pairing associated with $H^1(\Omega)$ and its dual $H^1(\Omega)'$. Then we set

$$e = (-\Delta + I)^{-1} \tilde{e} : H^1(\Omega) \times W \rightarrow H^1(\Omega).$$

Here, $(-\Delta + I)^{-1} : H^1(\Omega)' \rightarrow H^1(\Omega)$ is the Neumann solution operator or Riesz isomorphism, i.e., for $h \in H^1(\Omega)'$, $v = (-\Delta + I)^{-1} h$ solves

$$\int_{\Omega} (\nabla v \cdot \nabla \phi + v \phi) \, d\xi = \langle h, \phi \rangle_{(H^1)'}$$
 for all $\phi \in H^1(\Omega)$.

Remark 1 We express the equality constraint in the Hilbert space $H^1(\Omega)$ to simplify the presentation throughout Sections 2 and 3. However in our numerical realization we deal with the operator $\tilde{e}(u, w)$ instead of $e(u, w)$. For a numerical comparison of both formulations we refer the reader to [16].

Since Ω satisfies a uniform interior cone condition, $H^1(\Omega)$ is continuously embedded into $L^6(\Omega)$ for $d \leq 3$, cf. [5, p. 151]. From the Hölder inequality we find that the integral

$$\int_{\Omega} w(u - u_a) \phi \, dx \quad \text{for } (u, w, \phi) \in H^1(\Omega) \times W \times H^1(\Omega)$$

and its Fréchet-derivatives with respect to the pair (u, w) are well-defined. Since the other terms in the definition of the operator \tilde{e} are linear, we conclude that the operator e is twice continuously Fréchet-differentiable and its second Fréchet-derivative is globally Lipschitz-continuous. The next theorem ensures existence of a unique solution to the state equations for arbitrary non-negative $w \in W$.

THEOREM 2.1 *Suppose that Ω is a bounded connected open set in \mathbf{R}^d with smooth boundary, i.e., Γ is a variety of dimension $d - 1$ of class C^∞ and Ω lies locally on one side of Γ . Then for every $w \in L^2(\Omega)$ with $w \geq 0$ a.e. in Ω there exists a unique solution $u = u(w) \in H^1(\Omega)$ of the equation $e(u, w) = 0$. In particular, the set*

$$E = \{(u, w) \in H^1(\Omega) \times W : e(u, w) = 0 \text{ in } H^1(\Omega)\}$$

is nonempty. Moreover, u satisfies the estimate

$$\|u\|_{H^1(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Gamma)} + \|w\|_{L^2(\Omega)})$$

for some constant $C > 0$.

Proof Since Ω is a bounded connected open set with smooth boundary, it is known that $H^1(\Omega)$ is a Hilbert space endowed with the inner product

$$\langle u, v \rangle_\Gamma = \int_\Omega \nabla u \cdot \nabla v \, dx + \int_\Gamma uv \, ds \quad \text{for } u, v \in H^1(\Omega)$$

and the induced norm $\|u\|_\Gamma = \sqrt{\langle u, u \rangle_\Gamma}$ for $u \in H^1(\Omega)$, and that $\|\cdot\|_\Gamma$ is an equivalent norm to $\|\cdot\|_{H^1(\Omega)}$ on $H^1(\Omega)$ (see [1, p. 133]). Utilizing Hölder's and Young's inequalities, $\kappa(x) \geq \kappa > 0$, $\gamma > 0$, the continuous embedding $H^1(\Omega) \hookrightarrow L^6(\Omega)$ and $u_a \in L^3(\Omega)$, the claim follows by the Lax-Milgram lemma. \square

Remark 2 If Ω is only an open Lipschitz domain, we have to require that $w \geq \underline{w}$ on Ω a.e. for a positive constant \underline{w} . The existence of a unique solution $u = u(w) \in H^1(\Omega)$ to the state equation $e(u, w) = 0$ can be proved by utilizing the Lax-Milgram lemma. This positivity requirement for the perfusion is realistic, since $w = 0$ only for necrotic tissue.

For brevity we introduce the Hilbert space

$$Z = H^1(\Omega) \times W$$

endowed with the Hilbert space product topology. Next we define the set of admissible parameters $W_{\text{ad}} \subset W$ by

$$W_{\text{ad}} = \{w \in W : w \geq 0 \text{ a.e. in } \Omega\}.$$

Moreover, we set $Z_{\text{ad}} = H^1(\Omega) \times W_{\text{ad}}$.

Remark 3 If Ω is only an open Lipschitz domain, we have to define

$$W_{\text{ad}} = \{w \in W : w \geq \underline{w} \text{ a.e. in } \Omega\}$$

with some constant $\underline{w} > 0$ instead. Due to Remark 2-b) the existence of a unique solution u to (1) for every $w \in W_{\text{ad}}$ is then ensured.

For $z = (u, w) \in Z$ we introduce the cost functional $J : Z \rightarrow \mathbf{R}^+$ with $\mathbf{R}^+ = \{t \in \mathbf{R} : t \geq 0\}$ by

$$J(z) = \frac{1}{2} \|a(u - u_{\text{bas}}) + b(w - w_{\text{bas}}) - m\|_{L^2(\Omega)}^2 + \frac{1}{2} \|w - w_{\text{ref}}\|_W^2, \quad (5)$$

where $w_{\text{ref}} \in W$ is a given reference perfusion. Moreover, let $a, b \in L^\infty(\Omega)$ satisfy $a, b \geq 0$ in Ω a.e. and m belongs to $L^2(\Omega)$. It is obvious that J is twice continuously Fréchet-differentiable and its second Fréchet-derivative is Lipschitz-continuous. Note that the regularization term contains the scalar parameters α and β from (4), which have to be chosen in accordance to the measurement error.

Summarizing, the parameter identification problem (1)–(3) can be written as

$$\min J(z) \quad \text{subject to} \quad z \in Z_{\text{ad}} \text{ and } e(z) = 0. \quad (\text{P})$$

To prove existence of an optimal solution to (P) we shall make use of the following lemma.

LEMMA 2.2 *If $D \subset Z$ is weak sequentially closed and bounded, then the set $D \cap E$ is weak sequentially closed, where we have introduced the set E in Theorem 2.1.*

Proof If $D \cap E$ is empty, nothing has to be shown. Otherwise let $z = (u, w) \in H^1(\Omega) \times W$ and $\{z_n\}_{n \in \mathbf{N}}$, $z_n = (u_n, w_n)$, a sequence in $D \cap E$ such that $z_n \rightarrow z$ as n tends to infinity. From $z_n \in E$ for every n we infer that $e(u_n, w_n) = 0$ holds. Then we obtain

$$\lim_{n \rightarrow \infty} \int_{\Omega} \kappa \nabla(u_n - u) \cdot \nabla \phi \, d\xi = 0 \quad \text{for all } \phi \in H^1(\Omega). \quad (6)$$

Since Ω satisfies a uniform interior cone condition and $\beta > 0$, $H^1(\Omega)$ and W are compactly embedded into $L^4(\Omega)$ for $d \leq 3$, see [5, p. 161]. Hence, there exists a subsequence $\{z_{n_k}\}_{k \in \mathbf{N}}$, such that $z_{n_k} \rightarrow z$ in $L^4(\Omega) \times L^4(\Omega)$. Furthermore, the sequence $\{z_{n_k}\}_{k \in \mathbf{N}}$ is bounded in $L^2(\Omega) \times L^2(\Omega)$. Using Hölder's inequality we infer that

$$\begin{aligned} & \left| \int_{\Omega} (w_{n_k} u_{n_k} - wu) \phi \, d\xi \right| \\ &= \left| \int_{\Omega} (w_{n_k} u_{n_k} - w_{n_k} u + w_{n_k} u - wu) \phi \, d\xi \right| \\ &\leq (\|w_{n_k}\|_{L^2(\Omega)} \|u_{n_k} - u\|_{L^4(\Omega)} + \|w_{n_k} - w\|_{L^4(\Omega)} \|u\|_{L^4(\Omega)}) \|\phi\|_{L^4(\Omega)} \\ &\stackrel{k \rightarrow \infty}{\rightarrow} 0 \quad \text{for all } \phi \in H^1(\Omega). \end{aligned} \quad (7)$$

Analogously, we get

$$\left| \int_{\Omega} (w_{n_k} - w) u_a \phi \, d\xi \right| \leq \|w_{n_k} - w\|_{L^4(\Omega)} \|u_a\|_{L^3(\Omega)} \|\phi\|_{L^6(\Omega)} \stackrel{k \rightarrow \infty}{\rightarrow} 0 \quad (8)$$

for all $\phi \in H^1(\Omega)$. Recall that the trace operator $\tau_{\Gamma} : H^1(\Omega) \rightarrow H^{1/2}(\Gamma)$ is continuous and that $H^{1/2}(\Gamma)$ is compactly embedded into $L^2(\Gamma)$. Thus, we find

$$\left| \frac{1}{\gamma} \int_{\Gamma} \kappa(u_{n_k} - u) \phi \, ds \right| \leq \frac{1}{\gamma} \|\kappa\|_{L^{\infty}(\Omega)} \|u_{n_k} - u\|_{L^2(\Omega)} \|\phi\|_{L^2(\Gamma)} \stackrel{k \rightarrow \infty}{\rightarrow} 0 \quad (9)$$

for all $\phi \in H^1(\Omega)$. Combining (6)-(9) it follows that

$$\begin{aligned} \|e(z_{n_k}) - e(z)\|_{H^1(\Omega)} &= \|\tilde{e}(z_{n_k}) - \tilde{e}(z)\|_{H^1(\Omega)}, \\ &= \sup_{\|\phi\|_{H^1(\Omega)}=1} \left| \int_{\Omega} \kappa \nabla(u_{n_k} - u) \cdot \nabla \phi + (w_{n_k}(u_{n_k} - u_a) - w(u - u_a)) \phi \, d\xi \right. \\ &\quad \left. + \frac{1}{\gamma} \int_{\Gamma} \kappa(u_{n_k} - u) \phi \, ds \right| \stackrel{k \rightarrow \infty}{\rightarrow} 0. \end{aligned}$$

Hence, $e(z_n) = 0$ implies that $e(z) = 0$ and $z \in E$. Since D is weak sequentially closed, $z \in D$ holds, such that the claim follows from $z \in D \cap E$. \square

Utilizing Theorem 2.1 and Lemma 2.2, the existence of an optimal solution to (P) can be proved.

THEOREM 2.3 *Let Ω satisfy the hypothesis of Theorem 2.1. Then (P) possesses at least one optimal solution, $z^* = (u^*, w^*)$.*

Proof From (5) it follows that J is radially unbounded, i.e., the level sets

$$D_{\rho} = \{(u, w) \in Z_{\text{ad}} : J(u, w) \leq \rho\}$$

are closed, bounded, and convex, and hence weak sequentially closed (see [21, Proposition 38.2]).

By Lemma 2.2, $D_{\rho} \cap E$ is nonempty and weak sequentially closed for sufficiently large ρ .

Moreover, convexity and continuity of J imply that J is weak sequentially lower continuous on Z_{ad} (see [21, Proposition 37.7]). Therefore a solution of (P) exists (see [21, Theorem 38.A]). \square

The following result implies a standard constraint qualification condition.

THEOREM 2.4 *If Ω satisfies the hypothesis of Theorem 2.1, then for every $z = (u, w) \in Z_{\text{ad}}$ the operator $e_u(z)$ is bijective. Here $e_u(z)$ denotes the Fréchet-derivative of the operator e with respect to the variable u at the point z . In particular, the Fréchet-derivative $e'(z)$ is surjective.*

Proof The operator $e_u(z)$ is bijective if and only if for every $h \in H^1(\Omega)'$ there exists a unique $v \in H^1(\Omega)$ such that

$$\begin{aligned} -\operatorname{div}(\kappa \nabla v) + wv &= h && \text{in } \Omega, \\ \gamma \frac{\partial v}{\partial n} + v &= 0 && \text{on } \Gamma. \end{aligned} \tag{10}$$

The existence of a unique solution v to (10) follows by the same arguments as in the proof of Theorem 2.1. \square

2.2 First-order necessary optimality conditions

Let us introduce the Lagrangian functional $L : Z \times H^1(\Omega) \rightarrow \mathbf{R}$ associated with (P):

$$L(z, \lambda) = J(z) + \langle e(z), \lambda \rangle_{H^1(\Omega)}$$

Problem (P) is a nonconvex programming problem, so that different local minima might occur. Numerical methods will deliver a local minimum close to their starting point. Therefore, we do not restrict our investigations to global solutions of (P). We will assume that a fixed reference solution $z^* = (u^*, w^*) \in Z_{\text{ad}}$ is given, satisfying certain first- and second-order optimality conditions (ensuring local optimality of the solution). The active set at z^* is given by

$$\mathcal{A} = \{\xi \in \Omega : w^*(\xi) = 0 \text{ for } \xi \in \Omega\}.$$

The corresponding inactive set is $\mathcal{I} = \Omega \setminus \mathcal{A}$.

In the following theorem we present first-order necessary optimality conditions for (P).

THEOREM 2.5 *Suppose that $z^* = (u^*, w^*) \in Z_{\text{ad}}$ is a local solution to (P). Then there exist Lagrange multipliers $\lambda^* \in H^1(\Omega)$ and $\eta^* \in W$ such that*

$$-\operatorname{div}(\kappa \nabla \lambda^*) + w^* \lambda^* + a(a(u^* - u_{\text{bas}}) + b(w^* - w_{\text{bas}}) - m) = 0 \quad \text{in } \Omega, \tag{11a}$$

$$\gamma \frac{\partial \lambda^*}{\partial n} + \lambda^* = 0 \quad \text{on } \Gamma, \tag{11b}$$

$$e(u^*, w^*) = 0, \tag{11c}$$

$$\begin{aligned} & \mathcal{C}(b(a(u^* - u_{\text{bas}}) + b(w^* - w_{\text{bas}}) - m)) + \\ & \mathcal{C}(\lambda^*(u^* - u_a)) + a(w^* - w_{\text{ref}}) + \eta^* = 0 \quad \text{in } \mathcal{A}, \end{aligned} \tag{11d}$$

$$\eta^*|_{\mathcal{A}} \leq 0 \quad \text{and} \quad \eta^*|_{\mathcal{I}} = 0, \tag{11e}$$

where, for instance, $\eta^*|_{\mathcal{A}}$ denotes the restriction of the multiplier η^* on the subset \mathcal{A} of Ω . Furthermore, the linear bounded operator $\mathcal{C} : L^2(\Omega) \rightarrow W$ in (11d) is given as follows: For $w = Ch \in W$ we have $w(\xi) = w_i(\xi)$ for $\xi \in \Omega_i$ a.e., $i = 1, \dots, n_\Omega$, where

$$-\Delta w_i + w_i = h|_{\Omega_i} \quad \text{in } \Omega_i, \tag{12a}$$

$$-\frac{\partial w_i}{\partial n} = 0 \quad \text{on } \partial\Omega_i. \tag{12b}$$

Proof Due to Theorem 2.4 there exists a unique Lagrange multiplier $\lambda^* \in H^1(\Omega)$ such that

$$L_u(z^*, \lambda^*)u = 0 \quad \text{for all } u \in H^1(\Omega). \tag{13}$$

From (13) we infer (11a) and (11b). Condition (11c) denotes feasibility and is clearly satisfied. Due to the optimality of w^* the following variational inequality holds:

$$L_w(z^*, \lambda^*)(w - w^*) \geq 0 \quad \text{for all } w \in W_{\text{ad}}.$$

We set

$$\begin{aligned} \langle -\eta^*, w - w^* \rangle_W &= L_w(u^*, w^*, \lambda^*)(w - w^*) \\ &= \int_{\Omega} (b(a(u^* - u_{\text{bas}}) + b(w^* - w_{\text{bas}}) - m)(w - w^*) \, d\xi \\ &\quad + \int_{\Omega} \lambda^*(u^* - u_a)(w - w^*) \, d\xi + \alpha \langle w^* - w_{\text{ref}}, w - w^* \rangle_W \end{aligned} \quad (14)$$

for all $w \in W_{\text{ad}}$. Setting

$$r = b(a(u^* - u_{\text{bas}}) + b(w^* - w_{\text{bas}}) - m) + \lambda^*(u^* - u_{\text{bas}}) \in L^2(\Omega) \quad (15)$$

we infer from (12) that

$$\begin{aligned} \int_{\Omega} r(w - w^*) \, d\xi &= \sum_{i=1}^{n_{\Omega}} \int_{\Omega_i} r(w - w^*) \, d\xi = \sum_{i=1}^{n_{\Omega}} \langle Cr, w - w^* \rangle_{H^1(\Omega_i)} \\ &= \langle Cr, w - w^* \rangle_W \quad \text{for all } w \in W_{\text{ad}}. \end{aligned} \quad (16)$$

Combining, (14) and (16) we obtain (11d). For the proof of (11e) we refer the reader to [6]. \square
For second order conditions we refer to [18].

3 SQP Method

In the following we neglect the inequality constraint $w \geq 0$ and consider instead of (P) the optimal control problem

$$\min J(z) \quad \text{subject to} \quad e(z) = 0. \quad (17)$$

Next we introduce the standard (local) Lagrange-SQP method:

Algorithm 1 (Lagrange-SQP method)

- a) Choose $(z^0, \lambda^0) \in Z$ and set $n = 0$.
- b) Solve the following linear-quadratic minimization problem for δz^n

$$\begin{cases} \min L_z(z^n, \lambda^n) \delta z^n + \frac{1}{2} L_{zz}(z^n, \lambda^n)(\delta z^n, \delta z^n) \\ \text{subject to } e'(z^n) \delta z^n + e(z^n) = 0. \end{cases} \quad (\text{QP}^n)$$

- c) Set $z^{n+1} = z^n + \delta z^n$, $\lambda^{n+1} = \Lambda(z^n, \lambda^n)$ and go back to step b) provided a certain stopping criterium is not satisfied.

Remark 1 The linear-quadratic minimization problem (QP^n) is well-defined provided the operator $L_{zz}(z^n, \lambda^n)$ is coercive on the null space $N(e'(z^n))$ of $e'(z^n)$ and $e'(z^n)$ is surjective. Due to Theorem 2.4 the operator $e'(z)$ is surjective for any $z \in Z_{\text{ad}}$. Moreover, the operator $L_{zz}(z^n, \lambda^n)$ is coercive on $N(e'(z^n))$ provided $\|a(u^* - u_0) + b(w^* - w_0) - m\|_{L^2(\Omega)}$ is small; see [18, Theorem 4.2].

Note that an open issue in Algorithm 1 is the Lagrange multiplier update formula described by the mapping $\Lambda : Z \times H^1(\Omega) \rightarrow H^1(\Omega)$. Different Lagrange multiplier updates possibly lead to different rates of convergence. Here, we concentrate on the Newton Lagrange multiplier update.

For brevity, let us introduce the operator

$$F(z, \lambda) = \begin{pmatrix} L_z(z, \lambda) \\ e(z) \end{pmatrix} \quad \text{for } (z, \lambda) \in Z.$$

Then the first-order necessary optimality conditions can be expressed as

$$F(z^*, \lambda^*) = 0. \quad (18)$$

To find z^* numerically we solve (18) by the Newton method. For that purpose we require the Fréchet-derivative ∇F of the operator F . Setting $M = \nabla F$ we have

$$M(z, \lambda) = \begin{pmatrix} L_{zz}(z, \lambda) & e'(z)^* \\ e'(z) & 0 \end{pmatrix} \quad \text{for } (z, \lambda) \in Z.$$

Now we formulate the Lagrange-SQP method with the Newton Lagrange multiplier update.

Algorithm 2 (Lagrange-Newton-SQP method)

- a) Choose $(z^0, \lambda^0) \in Z$ and put $n = 0$.
- b) Solve for $(\delta z^n, \delta \lambda^n)$ the linear system

$$M(z^n, \lambda^n) \begin{pmatrix} \delta z^n \\ \delta \lambda^n \end{pmatrix} = -F(z^n, \lambda^n). \quad (19)$$

- c) Set $(z^{n+1}, \lambda^{n+1}) = (z^n + \delta z^n, \lambda^n + \delta \lambda^n)$, $n = n + 1$ and go back to b) provided a certain stopping criterium is not satisfied.

Remark 2

- a) In the context of Algorithm 1, the Newton multiplier update Λ_N is given by $\Lambda_N(z^n, \lambda^n) = \lambda^n + \Delta \lambda^n$, where

$$\begin{aligned} \Delta \lambda^n &= (e'(z^n) L_{zz}^c(z^n, \lambda^n)^{-1} e'(z^n)^*)^{-1} \\ &\quad \cdot (e'(z^n) - e'(z^n) L_{zz}^c(z^n, \lambda^n)^{-1} L_z^c(z^n, \lambda^n)). \end{aligned}$$

If $z \in Z_{\text{ad}}$ holds and if $\|a(u^* - u_0) + b(w^* - w_0) - m\|_{L^2(\Omega)}$ is small, the operator Λ_N is well-defined; see Remark 1 and [10, Lemma 2.2.3].

- b) Notice that in each iteration of Algorithm 2 the saddle-point problem (19) needs to be solved.
- c) If $L_{zz}(z, \lambda)$ is coercive on $N(e'(z))$ and $e'(z)$ is surjective for all $(z, \lambda) \in Z$ belonging to a neighborhood of $(z^*, \lambda^*) \in Z$, the SQP method is locally equivalent with the Newton method applied to $L'(z, \lambda) = 0$. Notice, that

In the following we show that for sufficiently well-matched data certain local solutions are unique, and that the sequential quadratic programming approach converges quadratically.

THEOREM 3.1 *There exist constants $C_D > 0$ and $\rho > 0$ independent of the actual measured data m and the reference perfusion w_{ref} , such that the following holds:*

If

$$r = \|a(u(w_{\text{ref}}) - u_{\text{bas}}) + b(w_{\text{ref}} - w_{\text{bas}}) - m\|_{L^2} < C_D \quad (20)$$

for $w_{\text{ref}} \geq \underline{w} > 0$ and $u(w_{\text{ref}})$ satisfying (1), then there exists a local solution z^* of (17) in the ball $B(z^0, \rho r)$, which is unique in $B(z^0, C_D \rho)$. If the SQP method is started at $(u(w_{\text{ref}}), w_{\text{ref}}, 0)$, it converges quadratically towards z^* .

Proof We will use Kantorovich's Theorem to prove convergence of Newton's method applied to $F(x) = L'(z, \lambda)$ with $x = (z, \lambda)$. Thus we have to verify its assumptions: invertibility of $F'(x^0)$, a Lipschitz condition for F' , and a sufficiently small initial residual $F(x^0)$.

(i) *Invertibility of $F'(x^0)$.* $F'(x^0)$ is a saddle point operator with positive definite upper left block $L_{zz}(x^0)$, since for any $\delta z \in N(e'(z^0))$ we have

$$L_{zz}(x^0)[\delta z, \delta z] = \|a\delta u + b\delta w\|_{L^2}^2 + \alpha \|\delta w\|_W^2 \geq \alpha \|\delta w\|_W^2. \quad (21)$$

Due to Theorem 2.4 there is a constant C_S such that $\|\delta u\|_{H^1} \leq C_S \|\delta w\|_{L^2}$ and hence

$$L_{zz}(x^0)[\delta z, \delta z] \geq \frac{\|\delta z\|_Z^2}{\sqrt{1 + C_S^2}}.$$

Note that C_S is indeed independent of w_{ref} . Moreover, $e'(z^0)$ satisfies the usual inf-sup-condition again due to Theorem 2.4, such that the Brezzi theorem establishes a bound C_I of $F'(x^0)^{-1}$ in the

space $\mathcal{E}(Z)$, where \mathcal{L} denotes the common space of all linear and continuous operators from Z into itself.

(ii) *Lipschitz condition for F'* . Since

$$(F'(x_1) - F'(x_2))[\delta x, \delta x] = 2 \int_{\Omega} ((\lambda_1 - \lambda_2)\delta u \delta w + (u_1 - u_2)\delta w + (w_1 - w_2)\delta u) d\xi,$$

F' satisfies the Lipschitz condition

$$\begin{aligned} \|F'(x_1) - F'(x_2)\| &\leq 2\|(\delta u, \delta w, \delta \lambda)\|^{-2} \int_{\Omega} ((\lambda_1 - \lambda_2)\delta u \delta w + (u_1 - u_2)\delta w + (w_1 - w_2)\delta u) d\xi \\ &\leq 2\|(\delta u, \delta w, \delta \lambda)\|^{-2} c(\Omega) \left(\|\lambda_1 - \lambda_2\|_{L^4(\Omega)} \|\delta u\|_{L^4(\Omega)} \|\delta w\|_{L^2(\Omega)} \right. \\ &\quad \left. + \|\delta \lambda\|_{L^4(\Omega)} \|u_1 - u_2\|_{L^4(\Omega)} \|\delta w\|_{L^2(\Omega)} + \|\delta \lambda\|_{L^4(\Omega)} \|\delta u\|_{L^4(\Omega)} \|w_1 - w_2\|_{L^2(\Omega)} \right) \\ &\leq C_L \|(u_1 - u_2, w_1 - w_2, \lambda_1 - \lambda_2)\| \\ &\leq C_L \|x_1 - x_2\| \end{aligned}$$

with Lipschitz constant C_L .

(iii) *Small initial residual*. The starting point x^0 leads to an initial residual

$$F(x^0) = \begin{bmatrix} a(a(u^0 - u_{\text{bas}}) + b(w^0 - w_{\text{bas}}) - m) \\ b(a(u^0 - u_{\text{bas}}) + b(w^0 - w_{\text{bas}}) - m) \\ 0 \end{bmatrix}$$

that leads to an initial Newton correction with

$$\|F'(x^0)^{-1} F(x^0)\| \leq 2C_I \max(a, b)r.$$

(iv) *Kantorovich theorem*. Combining the estimates (i) and (ii) yields the affine invariant Lipschitz condition

$$\|F'(x^0)^{-1}(F'(x_1) - F'(x_2))\| \leq \bar{\omega}_0 \|x_1 - x_2\|$$

with Lipschitz constant $\bar{\omega}_0 = C_I C_L$. Defining $C_D = (4\bar{\omega}_0 C_I \max(a, b))^{-1}$, the assumption $r < C_D$ and (iii) lead to

$$h_0 = \bar{\omega}_0 \|F'(x^0)^{-1} F(x^0)\| = 2\bar{\omega}_0 C_I \max(a, b)r < \frac{1}{2}.$$

Application of Kantorovich's theorem (cf. [2, Thm. 2.1]) yields the existence of a solution point $x^* \in B(x^0, \rho_-)$ with

$$\rho_- = \frac{1 - \sqrt{1 - 2h_0}}{\bar{\omega}_0} \leq 4C_I \max(a, b)r = \rho r$$

that satisfies the first order necessary condition $L'(z^*, \lambda^*) = F(x^*) = 0$. x^* is unique in $B(x^0, \rho_+)$ with

$$\rho_+ = \frac{1 + \sqrt{1 - 2h_0}}{\bar{\omega}_0} \geq \bar{\omega}_0^{-1} = C_D \rho.$$

Moreover, the Newton iteration converges quadratically to x^* .

(v) *Second order sufficient condition*. Let us introduce the path $x(\tau) = x^0 + \tau(x^* - x^0)$, $\tau \in [0, 1]$, with associated derivatives $F'(\tau) = F'(x(\tau))$. Since $x(\tau) \in B(x^0, r\rho)$, the inverses $F'(\tau)^{-1}$ are uniformly bounded for all τ due to the Banach perturbation lemma. Moreover, due to the Lipschitz continuity of F' , the mapping $\tau \rightarrow F'(\tau)^{-1}$ is continuous. We consider the function

$$s(\tau) = \inf_{\xi \in Z^*} \langle (\xi, 0), F'(\tau)^{-1} (\xi, 0) \rangle,$$

which is again continuous. Note that due to the injectivity of $F'(\tau)$

$$s(\tau) = \inf_{\zeta \in N(e'(z(\tau)))} \langle \zeta, L_{zz}(x(\tau))\zeta \rangle$$

holds. Now assume that $s(1) \leq 0$. Due to the continuity of s there is a $\bar{\tau} \in [0, 1]$ such that $s(\bar{\tau}) = 0$, which implies the existence of a minimizing sequence $\zeta_k \in N(e'(z(\bar{\tau})))$ with $\langle \zeta, L_{zz}(x(\bar{\tau}))\zeta \rangle \rightarrow 0$.

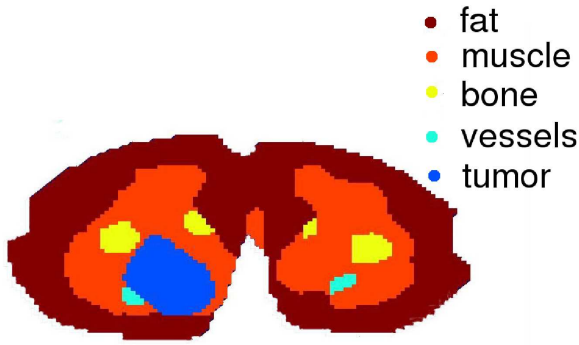


Figure 2. Materials and geometry of the cross section. Measurements are available on $\Omega_m = \{\text{fat, muscle, tumor}\}$. The area of the cross section is approximately 0.2×0.2 square meters.

Since $L_{zz}(x(\bar{\tau}))$ is symmetric, this contradicts the boundedness of $F'(\bar{\tau})^{-1}$. Thus, $s(1) > 0$ and hence $L_{zz}(x^*)$ is positive definite on $N(e'(z^*))$. Therefore, x^* is indeed a local solution of the minimization problem (17). \square

Remark 3 Theorem 3.1 does not guarantee the admissibility of the solution. With L_∞ -regularity results for elliptic problems with Robin boundary conditions, lower bounds on w^* can be established, including nonnegativity in the case of sufficiently well matched data.

Remark 4 The constants C_D and ρ established in Theorem 3.1 do typically increase with increasing α . If assumption (20) is not satisfied, choosing a larger regularization parameter may lead to a convergent method. Hence, performing a homotopy in α , starting with comparatively large values, can be expected to perform well.

4 Numerical examples

To begin with, we study the identification capability of the algorithm on a 2D cross section of a hyperthermia patient with artificial data. The cross section is chosen from the end of a treatment, such that the temperature is already relatively high. See Fig. 2 for the geometry and material distribution of the cross section. Choosing realistic perfusion profiles w and w_{bas} corresponding to hyperthermia and normal conditions, respectively, we obtain the related temperatures u and u_{bas} by solving (1). The measurements m are evaluated on a 128×128 voxel grid representing Ω_m according to (3), adding independent white noise with variance $\sigma = 0.5$. The noise level has been chosen to produce the same total variation of the signal as is observed in clinical measurements. The measurement parameters a and b (cmp. 3) are set equal to 1 on muscle and tumor, $a = 0.1, b = 1$ on fat and zero elsewhere.

As a reference perfusion we choose w_{ref} as $w + \delta w_1 + \delta w_2$, with a geometry-conforming systematic offset $\delta w_1 = -0.2w$ continuous on each subdomain Ω_i , and an additional geometry-violating offset $\delta w_2 = -0.2w$ that is discontinuous inside a subdomain. The chosen perfusions are depicted in Fig. 3, with temperatures and measurements shown in Fig. 4.

As a starting point for the homotopy in the regularization parameter α we choose $w_{\text{ref}}, u(w_{\text{ref}})$, and $\lambda = 0$ with $\alpha = 100$, performing classical continuation up to $\alpha = 0.05$. L^2 errors are shown for different values of β in Fig. 5. We divided the L^2 errors by the square root of the areas Ω_m and $\Omega \setminus \Omega_m$ to obtain average errors on these areas. With decreasing regularization the temperature errors decrease until a certain point and increase for less regularization. The results with the least errors are obtained for $\alpha = 0.1, \beta = 0.001$ and shown in Fig. 6. The average error of the starting temperature $u(w_{\text{ref}})$ (see Fig. 4 Middle) is 0.46°C on Ω_m . Despite up to 40 percent offset and 50 percent white noise in the measuring data the error is reduced to 0.15°C by our method (Fig. 5).

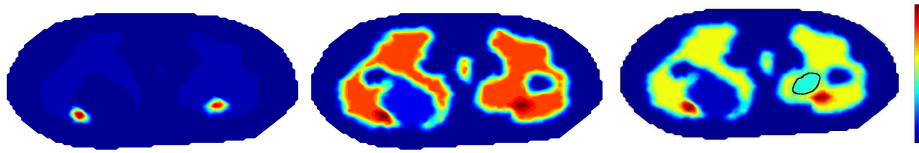


Figure 3. Perfusion data for the artificial setting. Left: normal perfusion w_{bas} . Middle: perfusion w under hyperthermia. Right: reference perfusion w_{ref} with the geometry violating offset δw_2 outlined in black.

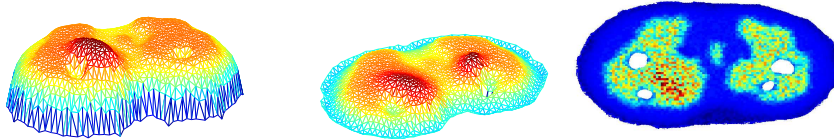


Figure 4. Temperature and measurement data for the artificial setting. Left: actual temperature $u(w)$. Middle: initial temperature error $u(w) - u(w_{\text{ref}})$. Right: noisy measurements on Ω_m .

As we can see the perfusion errors concentrate around $\Omega \setminus \Omega_m = \{ \text{vessels, bone} \}$. One reason for that behavior is due to the fact, that there is no measuring data available on $\Omega \setminus \Omega_m$. Also the perfusion in the vessels is more than 30 times higher than anywhere else, which leads to higher absolute errors on $\Omega \setminus \Omega_m$, see Fig. 5.

As a further example, we perform the identification for a clinical data set with unknown temperature and perfusion. The only data available for verification is a local measurement from a temperature probe, the position of which is indicated in the MR measurement data in Fig. 7, left. On the right the SAR-values are shown, which enter the right hand side of the bio heat transfer equation (1). Since the measuring data in the artificial example has the same total variation than the clinical measurement we expect regularization parameters α, β of the same order of magnitude to work well. For $\alpha = 5$ and $\beta = 0.001$ the identified temperature and perfusion are shown in Fig. 8. The temperature in the position of the local temperature measurement has been identified as $44,5^\circ\text{C}$ compared to 44°C measured directly. An area of low perfusion inside the tumor is identified, which indicates a region of necrotic tissue in the tumor's center.

5 Conclusion

Despite significant offsets and noise in the signal, the method is capable of identifying the temperature with an error reduction to 30 % of the initial error on Ω_m . The resulting perfusion is admissible except for a small area around the vessels. A rigorous treatment of the inequality constraints seems to be necessary, see e.g. [17, 19, 20]. Since the violation of the constraints is relatively small, however, a significant effect on the temperature field cannot be expected. More information about the measurement parameters a and b , i.e. about the dependence of the signal on perfusion and temperature, is needed to improve the clinical identifications. A homotopy in the regularization parameters α and β is crucial for the unconstrained SQP method to converge reliably. If SQP steps leave the admissible set W_{ad} , the BHTE (1) is not necessarily solvable, and divergence is expected and occasionally observed.

Acknowledgment

The authors would like to thank J. Gellermann for providing clinical measurement data.

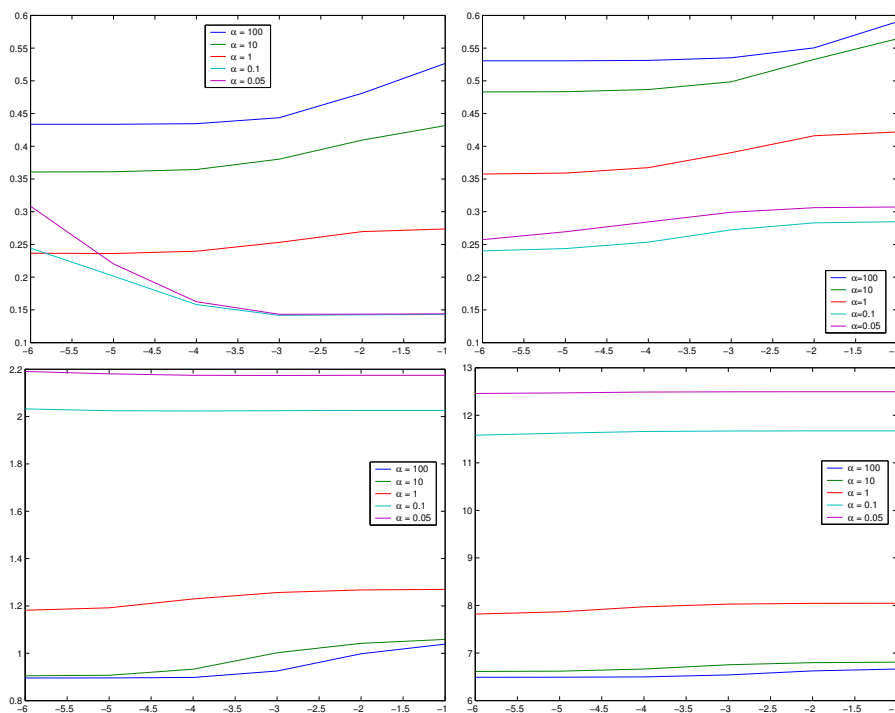


Figure 5. L2-Identification errors: Top left: L2-temperature error on Ω_m , Top right: L2-temperature error on $\Omega \setminus \Omega_m$, Bottom left: L2-perfusion error on Ω_m , Bottom right: L2-perfusion error on $\Omega \setminus \Omega_m$ for different values of $\log(\beta)$. Note that the scales in the two bottom plots are different.

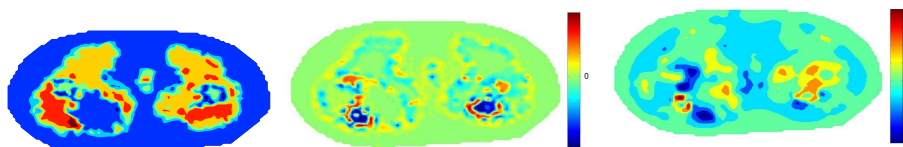


Figure 6. Errors for $\alpha = 0.1$, $\beta = 0.001$, Left: Identified perfusion, Middle: perfusion error $w_{id} - w$, Right: temperature error $u_{id} - u(w)$

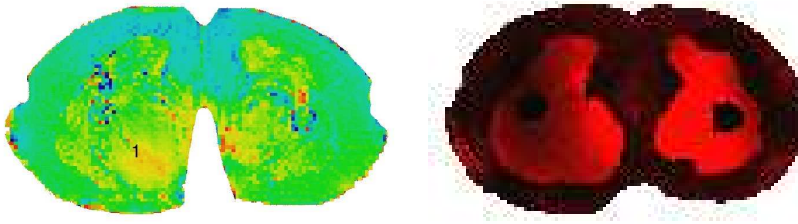


Figure 7. Left: Clinical measurement data. Invasive temperature measurement $\approx 44^\circ\text{C}$ in position 1. Right: Clinical SAR-values.

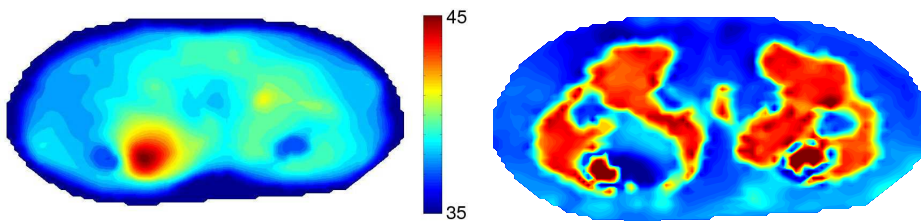


Figure 8. Identified temperature (left) and perfusion (right) for $\alpha = 5, \beta = 1e - 3$. Lower perfusion in necrotic tumor tissue is identified.

References

- [1] Dautray, R. and Lions, J.-L., 1988, *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 2. Springer.
- [2] Deuffhard, P., 2004, *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*, volume 35 of *Computational Mathematics*. Springer.
- [3] Deuffhard P., Seebaß, M., Stalling, D., Beck, R., and Hege, H.-C., 1997, Hyperthermia Treatment Planning in Clinical Cancer Therapy: Modelling, Simulation and Visualization. In A. Sydow, editor, *Proc. of the 15th IMACS World Congress 1997 on Scientific Computation: Modelling and Applied Mathematics*, volume 3, pages 9–17. Wissenschaft und Technik Verlag.
- [4] Deuffhard, P., Weiser, M. and Seebaß, M., 2000, A new nonlinear elliptic multilevel fem applied to regional hyperthermia. *Comput. Vis. Sci.*, 3(3):115–120.
- [5] Gilbarg, D. and Trudinger, N.S., *Elliptic Partial Differential Equations of Second Order*. Springer, 1977.
- [6] Hintermüller, M., A primal-dual active set algorithm for bilaterally control constrained optimal control problems. *Quarterly of Applied Mathematics*, to appear.
- [7] Ito, K., Krollner, M. and Kunisch, K., 1991, A numerical study of an augmented Lagrangian method for the estimation of parameters in elliptic systems. *SIAM J. Sci. Stat. Comput.*, 12(4):884–910.
- [8] Ito, K. and Kunisch, K., 1990, The augmented Lagrangian method for parameter estimation in elliptic systems. *SIAM J. Control Optim.*, 28:113–136.
- [9] Ito, K. and Kunisch, K., 1996, Augmented Lagrangian-SQP-methods in Hilbert spaces and application to control in the coefficient problems. *SIAM J. Optim.*, 6:96–125.
- [10] Kleis, D., 1997, *Augmented Lagrange SQP methods and application to the sterilization of prepackaged food*. PhD thesis, Fachbereich IV, University of Trier.
- [11] Kunisch, K. and Peichl, G., 1991, Estimation of a temporally and spatially varying diffusion coefficient in a parabolic system by an augmented Lagrangian technique. *Numer. Math.*, 59:473–509.
- [12] Kunisch, K. and Sachs, E.W., 1992, Reduced SQP methods for parameter identification problems. *SIAM J. Numer. Anal.*, 29:1793–1820.
- [13] Kunisch, K. and Tai, X.-C., 1997, Sequential and parallel splitting methods for bilinear control problems in Hilbert spaces. *SIAM J. Numer. Anal.*, 34:91–118.
- [14] Logashenko, D., Maar, B., Schulz, V. and Wittum, G., 2001, Parameter estimation and geometrical optimal design for Bingham measurement devices. Technical Report 642, WIAS, Berlin.
- [15] Pennes, H.H., 1948, Analysis of tissue and arterial blood temperatures in the resting human forearm. *J. Appl. Phys.*, 1:93–122.
- [16] Volkwein, S., 2002, Mesh-independence of Lagrange-SQP methods with Lipschitz-continuous Lagrange multiplier updates. *Optimization Methods and Software*, 17:77–111.
- [17] Volkwein, S., 2003 Lagrange-SQP techniques for the control constrained optimal boundary control problems for the Burgers equation. *Computational Optimization and Applications*, 26:253–284.
- [18] Volkwein, S. and Weiser, M., 2004, Optimality conditions for a constrained parameter identification problem in hyperthermia. Special Research Center F 003 *Optimization and Control*, Project Area Continuous Optimization and Control, Report No. 305.
- [19] Weiser, M., Gänzler, T., Schiela, A., 2004 A control reduced primal interior point method for PDE constrained optimization. ZIB-Report 04-38.
- [20] Weiser, M., Schiela, A., 2004 Function space interior point methods for PDE constrained optimization. *PAMM* 4 (1): 43-46.
- [21] Zeidler, E., 1985, *Nonlinear Functional Analysis and its Applications*, volume III. Springer, New York.