Mathematical Models and Methods in Applied Sciences

IDENTIFICATION OF TEMPERATURE DEPENDENT PARAMETERS IN RADIATIVE HEAT TRANSFER

OLIVER TSE

Fachbereich Mathematik, Technische Universität Kaiserslautern, Erwin Schrdinger Straße
Kaiserslautern, 67663, Germany
tse@mathematik.uni-kl.de

RENÉ PINNAU

Fachbereich Mathematik, Technische Universität Kaiserslautern, Erwin Schrdinger Straße
Kaiserslautern, 67663, Germany
pinnau@mathematik.uni-kl.de

NORBERT SIEDOW

Transport Processes, Fraunhofer ITWM, Fraunhofer-Platz 1
Kaiserslautern, 67663, Germany
norbert.siedow@itwm.fraunhofer.de

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

Laser-induced thermotherapy (LITT) is an established minimally invasive percutaneous technique of tumor ablation. Nevertheless, there is a need to predict the effect of laser applications and optimizing irradiation planning in LITT. Optical attributes (absorption, scattering) change due to thermal denaturation. The work presents the possibility to identify these temperature dependent parameters from given temperature measurements via an optimal control problem. The solvability of the optimal control problem is analyzed and results of successful implementations are shown.

Keywords: Radiative heat transfer; SP_n-approximation; optimal control; parameter identification.

AMS Subject Classification: 22E46, 53C35, 57S20

1. Introduction

Laser Interstitial Thermo Therapy (LITT) is a well established minimally invasive method for cancer treatment, especially for irresectable liver tumors.[6]

An applicator device consisting of an optical laser fiber surrounded by water cooling is placed into the tumor tissue. The absorbed fraction of the laser light leads to a rise of the tissue temperature. For temperatures above $60°C$ coagulation starts due to protein denaturation leading to the destruction of tumor tissue. The optimal and safe clinical implementation of this technique depends critically on the precise knowledge of light distribution within the laser-treated tissue and its variation during thermal tissue denaturation.

The cancer treatment is guided by magnetic resonance imaging (MRI). Based on temperature-sensitive magnetic resonance parameters such as proton resonance frequency it is feasible to monitor the tissue temperature during the cancer treatment.[6] On the other hand, mathematical simulation may be used to predict the effects of the interstitial laser treatment and to optimize the irradiation planning in LITT. For that the knowledge about optical properties, like absorption or scattering, and their variations due to thermal denaturation, is indispensable. Combining both MR thermometry and mathematical simulation is a promising procedure to identify temperature depended tissue parameters and to optimize the cancer treatment.

For the mathematical modeling of radiative heat transfer in biological tissue the heat transfer equation has to be coupled with the radiative transfer equation. Because of the high dimensionality of the latter problem, the simpler SP_1-approximation is used instead of the full radiative transfer equation. A justification to this simplification for radiative transfer in biological tissues can be found in Ref. 1.

1.1. *Mathematical Problem Description*

Let $I \subset \mathbb{R}$ be a bounded time interval and $\Omega \subset \mathbb{R}^n$ be a bounded domain. Consider the SP_1-approximation to the radiative heat transfer equations given by the system

$$-\nabla \cdot \left(\frac{1}{3\beta(d,T)} \nabla \rho \right) + \mu(d,T)\rho = 0, \tag{1.1a}$$

$$c_p \partial_t T - \nabla \cdot (\kappa \nabla T) + b(T - T_b) - \mu(d,T)\rho = 0, \quad \text{in } Q := I \times \Omega, \tag{1.1b}$$

with boundary conditions

$$\frac{1}{3\beta(d,T)} \partial_n \rho + \gamma \rho = \gamma \rho_\partial, \tag{1.1c}$$

$$\kappa\, \partial_n T + \alpha T = \alpha T_\partial, \quad \text{on } \Sigma := I \times \partial\Omega, \tag{1.1d}$$

supplemented with an initial condition

$$T(0,x) = T_0(x) \quad \text{for all } x \in \Omega, \tag{1.1e}$$

where $\rho_\partial[\text{Wmm}^{-2}]$ and $T_\partial[\text{K}]$ denote the incident radiation and temperature at the boundary respectively, $T_b[\text{K}]$ the blood temperature, $\beta[\text{mm}^{-1}]$, $\mu[\text{mm}^{-1}]$, γ are optical parameters with β and μ depending on the temperature dependent rate constant d and temperature $T[\text{K}]$, and $c_p[\text{Jmm}^{-3}\text{K}^{-1}]$, $\kappa[\text{Wmm}^{-1}\text{K}^{-1}]$, $b[\text{Wmm}^{-3}\text{K}^{-1}]$, $\alpha[\text{Wmm}^{-2}]$ are thermal parameters. In general, the rate constant d models the denaturation of optical parameters due to temperature and may vary between different tissues. Throughout Sec. 2-4 we will for simplicity assume that $c_p = 1$.

The task at hand is to identify the rate constant d for given temperature measurements T_m[K] and common rate constant d_c. We consider the parameter identification problem as an optimal control problem, where we minimize a given cost functional J with the rate constant d being the control and the temperature T being the state, i.e

$$\min J(d,T) \quad \text{w.r.t.} \quad (d,T,\rho) \quad \text{subject to system (1.1).} \tag{1.2}$$

In this paper we provide an analysis for this approach. In Sec. 2 we study the state system, show the unique solvability of the state system and derive a priori estimates, which we will require in the following sections. We further show the unique solvability of the linearized state system along with its adjoint equations in Sec. 3. We then prove the existence of an optimal control d and derive regularity results for the control to state map in Sec. 4, which is essential for the introduction of the reduced cost functional. Sec. 5 will be devoted to examples and numerical implementations. Concluding remarks are given in Sec. 6.

1.2. *Notation*

For a domain $\Omega \subset \mathbb{R}^n$ with Lipschitz-boundary $\partial\Omega$, we denote the Lebesgue spaces with $L_p(\Omega)$ and the Sobolev spaces with $W_p^k(\Omega)$ ($k \in \mathbb{N}$, $p \in [1,\infty]$) and its norm by $\|\cdot\|_{L_p(\Omega)}$ and $\|\cdot\|_{W_p^k(\Omega)}$, respectively. We denote by p' the dual for p, i.e. $1/p'+1/p = 1$ such that $L_p{}^* \cong L_{p'}$. In the special case $p = 2$ we use $H^k(\Omega)$ to denote $W_2^k(\Omega)$. Further, let $\mathcal{D}(\Omega) = \mathcal{C}_0^\infty(\Omega)$ be the set of test functions and $H_0^k(\Omega)$ be the closure of $\mathcal{D}(\Omega)$ with respect to the $H^k(\Omega)$-norm. Its dual space $H_0^k(\Omega)^*$ is denoted by $H^{-k}(\Omega)$. The duality pairing of a Banach space X with its dual X^* is given by $\langle\cdot,\cdot\rangle_{X^*,X}$; if the spaces involved are clear, we simply write $\langle\cdot,\cdot\rangle$. For a Hilbert space H, its inner product is denoted by $(\cdot,\cdot)_H$; if H is clear we simply write $(\cdot,\cdot)$. We also denote $(\cdot,\cdot)_\partial$ to be the scalar product on the Hilbert space H_∂ of functions on the boundary $\partial\Omega$.

Moreover, for a bounded interval I and Banach space B, we define the Lebesgue-Bochner space $L_p(I;B)$ with $p \in [1,\infty]$ consisting of all measurable functions $f\colon I \to B$ for which the norm

$$\|f\|_{L_p(I;B)} = \Big(\int_I \|f(t)\|_B^p \,dt\Big)^{\frac{1}{p}}, \quad p \in [1,\infty),$$
$$\|f\|_{L_\infty(I;B)} = \sup_{t\in I} \|f(t)\|_B, \quad p = \infty$$

is finite. Further, we define the Sobolev-Bochner space $W_p^k(I;B)$ with $m \in \mathbb{N}$ and $p \in [1,\infty]$ consisting of all weakly absolutely continuous functions $f\colon I \to B$ such that f is m-times weakly differentiable, and $\partial_t^k f \in L_P(I;B)$ for all $k \le m$ (for details see Ref. 15). For $m = 1$, we just write $\dot f = \partial_t f$.

For notational convenience we denote

$$Q = I \times \Omega, \quad \Sigma = I \times \partial\Omega,$$
$$V_{p,r} = L_r(I; W_p^1(\Omega)), \quad \mathcal{W} = V_{2,2} \cap W_2^1(I; H^{-1}(\Omega)),$$
$$X_{p,r} = V_{p,r} \times \mathcal{W}, \quad Z = V_{2,2} \times V_{2,2} \times L_2(\Omega),$$

Note that for a bounded domain Ω, we have that the embedding $X_{p,r} \hookrightarrow X_{q,s}$ is continuous and dense for all $1 \le q \le p$ and $1 \le s \le r$.

Throughout this paper we will use the notations

$$\bar{u} = ess\sup_{x\in\Omega} u(x) < \infty \quad \text{and} \quad \underline{u} = ess\inf_{x\in\Omega} u(x) > -\infty,$$

when either exists. Unless otherwise stated, $\kappa, b \in L_{\infty,>0}(Q)$ and $\gamma, \alpha \in L_{\infty,\ge 0}(\Sigma)$, where

$$L_{\infty,>0\,(\ge 0)}(D) = \{u \in L_\infty(D) \,|\, \underline{u} > 0 \,(\underline{u} \ge 0)\}$$

for $D = Q, \Sigma$. We make the following assumption

(**A1**) Let $\Omega \subset \mathbb{R}^n$, $n = 2, 3$ be a bounded domain with $\mathcal{C}^{0,1}$-boundary $\partial\Omega$ and $I = (0, t_*)$, $t_* < \infty$.

2. The State System

2.1. *Nonlinearity*

We begin by discussing the nonlinearities in the system by means of Nemytskij operators. Known facts regarding Nemytskij operators and their properties can be found in Sec. 4.2 of Ref. 7, Sec. 5.2 of Ref. 18 and Ref. 8. We refer to Ref. 12 for an extensive study on nonlinear operators.

Theorem 2.1. *Assume (**A1**). Further, let $\mathcal{U} \subset \mathcal{C}_b^1(\mathbb{R})$ and $\mathcal{K} \subset L_\infty(Q)$ be open subsets. We define the operator $\varphi\colon \mathcal{U} \times \mathcal{K} \to L_\infty(Q)$, as follows:*

$$\varphi(d, u) = \varphi_0 + \varphi_1 \left(\int_0^{\cdot} d(u)(\tau)\, d\tau \right),$$

where $\varphi_0 \in L_\infty(\Omega)$ and $\varphi_1 \in \mathcal{C}_{b,loc}^1(\mathbb{R})$. Then, the operator φ is well-defined and continuously Fréchet differentiable with

$$D\varphi(d, u)\,(v_d, v_u) = \varphi_1' \left(\int_0^{\cdot} d(u)(s)\, ds \right) \int_0^{\cdot} \left(v_d(u) + d'(u)\, v_u\right)(\tau)\, d\tau,$$

for (d, u), $(v_d, v_u) \in \mathcal{U} \times \mathcal{K}$.

Proof. Let $(d, u) \in \mathcal{U} \times \mathcal{K}$. Since $d \in \mathcal{C}_b^1(\mathbb{R})$ we have $d(u) \in L_\infty(Q)$ for all $u \in \mathcal{K}$ by Theorem 1 of Ref. 8 and thus

$$\int_0^{\cdot} d(u)(\tau)\, d\tau \in W_\infty^1(I; L_\infty(\Omega))$$

by the definition of Bochner-Sobolev spaces. Note that the embedding $W^1_\infty(I; L_\infty(\Omega)) \hookrightarrow L_\infty(Q)$ is continuous. Using the same arguments as above we conclude the first assertion. The Fréchet differentiability follows by applying the chain rule. □

Example 2.1. Let $\mathcal{U} = H^2(\mathbb{R})$ and $\mathcal{K} = L_\infty(Q)$. Define $\varphi\colon \mathcal{U} \times \mathcal{K} \to L_\infty(Q)$ by

$$\varphi(d, u) = b - (b - a) \exp\left(- \int_0^{\cdot} d(u)(\tau)\, d\tau \right),$$

with constants a, $b > 0$.

Clearly $\exp \in \mathcal{C}^1_{b,\mathrm{loc}}(\mathbb{R})$. Due to standard embedding theorems, $H^2(\mathbb{R}) \hookrightarrow \mathcal{C}^1_b(\mathbb{R})$. Thus, φ is well-defined and continuously Fréchet differentiable on $\mathcal{U} \times \mathcal{K}$ by Theorem 2.1 with

$$\partial_1\varphi(d, u)\, v_d = (b - a) \exp\left(- \int_0^{\cdot} d(u)(s)\, ds \right) \int_0^{\cdot} v_d(u)(\tau)\, d\tau,$$
$$\partial_2\varphi(d, u)\, v_u = (b - a) \exp\left(- \int_0^{\cdot} d(u)(s)\, ds \right) \int_0^{\cdot} (d'(u)\, v_u)\,(\tau)\, d\tau,$$

for (d, u), $(v_d, v_u) \in \mathcal{U} \times \mathcal{K}$.

Remark 2.1. Most of our effort is intended to solve problems with φ as defined in the example above. Observe that in the case of non-negative d, i.e., $d \in \mathcal{U} = \{d \in H^2(\mathbb{R}) \mid d \geq 0\}$,

$$\varphi(\mathcal{U} \times \mathcal{K})(t, x) \in [\min\{a, b\}, \max\{a, b\}] \quad \text{for a.e. } (t, x) \in Q,$$

which shows that $\varphi(\mathcal{U} \times \mathcal{K})$ is uniformly bounded in $L_{\infty,>0}(Q)$.

We make the following assumption on β and μ:

(**A2**) β and μ are of type φ as defined in Theorem 2.1 and are uniformly bounded in $L_{\infty,>0}(Q)$ for all $(d, u) \in \mathcal{U} \times \mathcal{K}$.

2.2. *Radiation Equation*

Let $d \in \mathcal{U}$ be fixed throughout this section. Next, we deal with the radiation equation

$$-\nabla \cdot \left(\frac{1}{3\beta_d(T)} \nabla\rho \right) + \mu_d(T)\rho = 0, \qquad \text{in } Q \tag{2.1a}$$

with boundary condition

$$n \cdot \frac{1}{3\beta_d(T)}\, \nabla\rho + \gamma\rho = \gamma\rho_\partial, \qquad \text{on } \Sigma, \tag{2.1b}$$

where $\beta_d(T) = \beta(d, T)$ and $\mu_d(T) = \mu(d, T)$ are as in (**A2**).

For given $T \in \mathcal{K}$, we consider the weak formulation of (2.1) given by

$$F_1(\rho, T) = f_1 \qquad \text{in } {V_{2,r'}}^*, \tag{2.2}$$

where $F_1(\cdot, T) : V_{2,r} \to {V_{2,r'}}^*$ is induced by the bilinear form

$$\langle F_1(\rho, T), v\rangle = (\frac{1}{3\beta_d(T)}\nabla\rho, \nabla v) + (\mu_d(T)\rho, v) + (\gamma\rho, v)_\partial,$$

with right hand side

$$\langle f_1, v\rangle = (\gamma\rho_\partial, v)_\partial \quad \text{for all } v \in V_{2,r'}.$$

From standard elliptic theory we directly get the following result.

Lemma 2.1. *For an arbitrary but fixed $T \in \mathcal{K}$ there exists a unique solution $\rho \in V_{2,r}$ of (2.2) with*

$$\|\rho\|_{V_{2,r}} \le \frac{1}{\mathcal{H}_\rho}\,\overline{\gamma}\|\rho_\partial\|_{L_r(I;L_2(\partial\Omega))},$$

where $\mathcal{H}_\rho = \min\{1/(3\overline{\beta}_d), \underline{\mu}, \underline{\gamma}\}$.

Remark 2.2. Notice that, due to the uniform boundedness of $\overline{\beta}_d$ and $\overline{\mu}_d$, they do not depend on T, which then implies the uniform boundedness for ρ in $V_{2,r}$ with respect to T.

We further recall results obtained in Ref. 17 and especially refer to Theorem 3 of Ref. 17, which states as a corollary, the following: For $f_1 \in L_r(I; W^1_{p'}(\Omega)^*)$ with $p \ge n$ and sufficiently smooth boundary $\partial\Omega$, the solution $\rho \in V_{2,r}$ for (2.2) enjoys $V_{p,r}$-regularity, i.e. $\rho \in V_{p,r}$ with $n \le p \le p_0$ for some $p_0 < \infty$ depending only on $\underline{\beta}$, $\overline{\beta}$ and Ω.

2.3. *Heat Equation*

Let $w \in L_\infty(Q)$ and $\rho \in L_r(Q)$ for some $r \ge 2$. Now consider the system

$$\partial_t T - \nabla\cdot(\kappa\nabla T) + b\,T = b\,T_b + \mu_d(w)\rho, \qquad \text{in } Q \tag{2.3a}$$

with boundary condition

$$\kappa\,\partial_n T + \alpha T = \alpha T_\partial, \qquad \text{on } \Sigma, \tag{2.3b}$$

and initial condition $T(0, x) = T_0(x)$ for a.e. $x \in \Omega$.

Similarly, the weak formulation of (2.3) can be written as

$$\dot{T} + F_2(T) = f_2(w, \rho) \qquad \text{in } {V_{2,2}}^*, \tag{2.4}$$

with $T(0) = T_0$ where $F_2 : V_{2,2} \to {V_{2,2}}^*$ is induced by the bilinear form

$$\langle F_2(T), v\rangle = (\kappa\nabla T, \nabla v) + (b\,T, v) + (\alpha T, v)_\partial,$$

with right hand side

$$\langle f_2, v\rangle = (b\,T_b + \mu_d(w)\rho, v) + (\alpha T_\partial, v)_\partial \quad \text{for all } v \in V_{2,2}.$$

Lemma 2.2. *Assume (**A1**-**A2**) and let $p \ge n$ and $r > 4$. Then for $\rho, T_b \in L_r(I; L_p(\Omega))$, $T_\partial \in L_r(I; L_p(\partial\Omega))$ and $T_0 \in L_\infty(\Omega)$, there exists a unique solution*

$T \in \mathcal{W} \cap L_\infty(Q)$ for (2.4). Moreover, there exists a constant $c_\infty > 0$, independent of ρ, T_b, T_∂, T_0, such that

$$\|T\|_{\mathcal{W}} + \|T\|_{L_\infty(Q)} \leq \\ c_\infty \left(\|T_0\|_{L_\infty(\Omega)} + \|\rho\|_{L_r(I;L_p(\Omega))} + \|T_b\|_{L_r(I;L_p(\Omega))} + \|T_\partial\|_{L_r(I;L_p(\partial\Omega))}\right). \quad (2.5)$$

Proof. From the standard theory for linear parabolic equations,[14] we obtain a unique solution $T \in \mathcal{W}$ to problem (2.4) for $f_2 \in \mathcal{W}^*$ and $T_0 \in L_2(\Omega)$. Consider the weak formulation

$$-(u_1, \partial_t v) - (\kappa \nabla u_1, \nabla v) + (b\, u_1, v) + (\alpha\, u_1, v)_\partial = (T_0, v),$$

for all $v \in W_2^1(I; H^1(\Omega))$ with $v(T) = 0$. Similarly we obtain a solution $u_1 \in \mathcal{W}$ and further $u_1 \in L_\infty(Q)$ by maximum principle.[14] The difference between (2.4) and the above equation yields

$$-(u_2, \partial_t v) - (\kappa \nabla u_2, \nabla v) + (b\, u_2, v) + (\alpha\, u_2, v)_\partial = (b\, T_b + \mu_d(w)\rho, v) + (\alpha\, T_\partial, v)_\partial,$$

for all $v \in W_2^1(I; H^1(\Omega))$ with $v(T) = 0$, where $u_2 = T - u_1$. By introducing Sobolev-Morrey spaces and applying methods discussed by Griepentrog in Ref 10 and Ref 9, we obtain with the prescribed right hand sides a solution $u_2 \in \mathcal{C}(\overline{I}; \mathcal{C}^{0,\alpha}(\overline{\Omega})) \cap \mathcal{C}^{0,\frac{\alpha}{2}}(\overline{I}; \mathcal{C}(\overline{\Omega}))$ for some $\alpha = \alpha(p, r) > 0$. In particular, $u_2 \in L_\infty(Q)$ and thus $T = u_1 + u_2 \in L_\infty$ as desired. The asserted estimate is then obtained as a result of the triangle inequality and of the estimates for u_1 and u_2 respectively. □

Remark 2.3. Observe that the constant c_∞ given in Lemma 2.2 does not depend on $w \in L_\infty(Q)$ in any way due to (**A2**), which infers the uniform boundedness of T with respect to w.

2.4. *State Vectors*

Now we are ready to prove the existence and uniqueness for the radiative heat transfer problem (1.1). We begin by writing the system in its weak formulation given by

$$E_d(\rho, T) = 0 \qquad \text{in } Z^*, \quad (2.6)$$

where $E_d\colon X \to Z^*$ is a continuous map induced by

$$\langle E_{d,1}(\rho, T), v_1\rangle = (\frac{1}{3\beta_d(T)} \nabla\rho, \nabla v_1) + (\mu_d(T)\rho, v_1) + (\gamma(\rho - \rho_\partial), v_1)_\partial,$$
$$\langle E_{d,2}(\rho, T), v_2\rangle = \langle \dot{T}, v_2\rangle + (\kappa \nabla T, \nabla v_2) + (b(T - T_b) - \mu_d(T)\rho, v_2) \\ + (\alpha(T - T_\partial), v_2)_\partial,$$
$$\langle E_{d,3}(\rho, T), v_3\rangle = \langle T(0) - T_0, v_3\rangle,$$

for all $v = (v_1, v_2, v_3)^T \in Z$.

Theorem 2.2. *Assume (**A1**-**A2**) and let $p \geq n$ and $r > 4$. Then for $\rho_\partial \in L_r(I; L_2(\partial\Omega))$, $T_b \in L_r(I; L_p(\Omega))$, $T_\partial \in L_r(I; L_p(\partial\Omega))$ and $T_0 \in L_\infty(\Omega)$, there exists $(\rho, T) \in V_{2,r} \times \mathcal{K}$ fulfilling (2.6), where $\mathcal{K} := \mathcal{W} \cap L_\infty(Q)$.*

Proof. The outline of the proof is as follows: We start by freezing the nonlinearity and consider an auxiliary problem. We then define, with the help of the auxiliary problem, a compact fixed point mapping and later show uniform boundedness for the fixed points of the map. We then make use of the Leray-Schauder theorem (cf. Theorem 11.6 of Ref. 4) to conclude the theorem.

Let $w \in L_2(Q)$ and $\sigma \in [0, 1]$ be given. Consider the auxiliary problem: Find $(\rho, T) \in V_{2,r} \times \mathcal{W}$ such that

$$F_1(\rho, [w]_k) = \sigma f_1 \quad \text{in } {V_{2,r'}}^* \tag{2.7a}$$

$$\dot{T} + F_2(T) = \sigma f_2([w]_k, \rho) \quad \text{in } {V_{2,2}}^* \tag{2.7b}$$

with $T(0) = \sigma T_0$ in $L_\infty(\Omega)$ is fulfilled. Here, $[\cdot]_k \colon L_2(Q) \to L_2(Q)$ denotes the cut-function

$$[w]_k = \begin{cases} k, & w > k \\ w, & -k \leq w \leq k \\ -k, & w < -k \end{cases},$$

for any $k > 0$.

Note that in the auxiliary problem the two equations decouple. For a given $w \in L_2(Q)$, we have a unique solution $\rho \in V_{2,r}$ of the first equation in (2.7) due to Lemma 2.1. Inserting this into the second one gives the existence of a unique $T \in \mathcal{W}$ as discussed in Lemma 2.2.

Since solution operators are continuous and chains of continuous operators are continuous, this introduces a continuous fixed point mapping

$$H \colon L_2(Q) \times [0, 1] \to L_2(Q), \qquad (w, \sigma) \mapsto T,$$

which is well-defined and compact since $\mathcal{W} \hookrightarrow L_2(Q)$ is compact due to Aubin's Lemma. Also, $H(w, 0) = 0$ for all $w \in L_2(Q)$. All that is left to show is the uniform boundedness for fixed points.

Now let $T \in L_2(Q)$ be a fixed point of $H(\cdot, \sigma)$. Since $\rho \in V_{2,r} \hookrightarrow L_r(I; L_p(\Omega))$, the requirements of Lemma 2.2 are fulfilled and we have $T \in \mathcal{K}$ for all $\sigma \in [0, 1]$ with estimate (2.5) being independent of σ. We recall Remark 2.3 stating that T is uniformly bounded with respect to $[w]_k$. Thus we may increase k until $[T]_k = T$ without effecting the estimate above yielding

$$\|T\|_{L_2(Q)} \leq \|T\|_{\mathcal{W}} + \|T\|_{L_\infty(Q)} \leq M(\rho_\partial, T_0, T_b, T_\partial) < \infty.$$

Applying the Leray-Schauder theorem concludes the proof of existence for $T \in \mathcal{K}$ and hence also for $\rho \in V_{2,r}$, i.e., $(\rho, T) \in V_{2,r} \times \mathcal{K}$. □

Theorem 2.3. *Assume (**A1**-**A2**). The solution $(\rho, T) \in V_{p,r} \times \mathcal{K}$, $p > 2$, of (2.6) is unique.*

Proof. Let $(\rho_1, T_1), (\rho_2, T_2) \in V_{p,r} \times \mathcal{K}$ be two solutions of (2.6). Then the difference $(\hat\rho, \hat T) = (\rho_1 - \rho_2, T_1 - T_2) \in V_{p,r} \times \mathcal{K}$ solves

$$\hat E_d(\hat\rho, \hat T) = 0 \quad \text{in } Z^* \tag{2.8}$$

with initial condition $\hat T_0 = 0$ in $L_\infty(\Omega)$, where $\hat E$ is given by

$$\langle \hat E_{d,1}(\hat\rho, \hat T), v_1\rangle = (\frac{1}{3\beta_d(T_1)}\nabla\hat\rho + \left(\frac{1}{3\beta_d(T_1)} - \frac{1}{3\beta_d(T_2)}\right)\nabla\rho_2, \nabla v_1) \\ + (\mu_d(T_1)\hat\rho + (\mu_d(T_1) - \mu_d(T_2))\,\rho_2, v_1) + (\gamma\hat\rho, v_1)_\partial, \tag{2.9a}$$

$$\langle \hat E_{d,2}(\hat\rho, \hat T), v_2\rangle = \langle \dot{\hat T}, v_2\rangle + (\kappa\nabla\hat T, \nabla v_2) + (b\hat T - \mu_d(T_1)\hat\rho, v_2) \\ - ((\mu_d(T_1) - \mu_d(T_2))\,\rho_2, v_2) + (\alpha\hat T, v_2)_\partial, \tag{2.9b}$$

$$\langle \hat E_{d,3}(\hat\rho, \hat T), v_3\rangle = \langle \hat T_0, v_3\rangle. \tag{2.9c}$$

Testing (2.9a) with $\hat\rho(t)$ and applying Hölder's and Young's inequalities, we get

$$\|\hat\rho(t)\|^2_{H^1(\Omega)} \le c_1\Big(\|\nabla\rho_2(t)\|^2_{L_p(\Omega)}\|\frac{1}{3\beta_d(T_2)}(t) - \frac{1}{3\beta_d(T_1)}(t)\|^2_{L_q(\Omega)} \\ + \|\rho_2(t)\|^2_{L_p(\Omega)}\|\mu_d(T_2)(t) - \mu_d(T_1)(t)\|^2_{L_q(\Omega)}\Big), \tag{2.10}$$

for a.e $t \in I$ with constant $c_1 > 0$ and $q = 2p/(p-2)$. Similarly we test (2.9b) with $\hat T(t)$ and apply Hölder's and Young's inequalities yielding for a.e. $t \in I$

$$\partial_t\|\hat T(t)\|^2_{L_2(\Omega)} + c_2\|\hat T(t)\|^2_{H^1(\Omega)} \le \\ c_3\Big(\|\hat\rho(t)\|^2_{L_2(\Omega)} + \|\rho_2(t)\|^2_{L_p(\Omega)}\|\mu_d(T_2)(t) - \mu_d(T_1)(t)\|^2_{L_q(\Omega)}\Big),$$

with constants c_2, $c_3 > 0$. Due to the continuous embedding $W^1_p(\Omega) \hookrightarrow L_2(\Omega)$ and inequality (2.10) we further obtain

$$\partial_t\|\hat T(t)\|^2_{L_2(\Omega)} \le c_4\|\rho_2(t)\|^2_{W^1_p(\Omega)}\Big(\|\frac{1}{3\beta_d(T_2)}(t) - \frac{1}{3\beta_d(T_1)}(t)\|^2_{L_q(\Omega)} \\ + \|\mu_d(T_2)(t) - \mu_d(T_1)(t)\|^2_{L_q(\Omega)}\Big),$$

with constant $c_4 > 0$.

Since T_1, $T_2 \in L_\infty(Q)$, there exists an $M > 0$ with $\max\{\overline{T}_1, \overline{T}_2\} \le M$. Furthermore, $(\cdot)^{-1}\colon \mathbb{R}_{>0} \to \mathbb{R}$, β_d and μ_d are locally Lipschitz-continuous. So we have constants $L_1(M)$, $L_2(M) > 0$ such that

$$\|\frac{1}{3\beta_d(T_2)}(t) - \frac{1}{3\beta_d(T_1)}(t)\|_{L_q(\Omega)} \le L_1(M)\|\hat T(t)\|_{L_q(\Omega)},$$
$$\|\mu_d(T_2)(t) - \mu_d(T_1)(t)\|_{L_q(\Omega)} \le L_2(M)\|\hat T(t)\|_{L_q(\Omega)},$$

thus implying

$$\partial_t\|\hat T(t)\|^2_{L_2(\Omega)} \le c_4\, L(M)\|\rho_2(t)\|^2_{W^1_p(\Omega)}\|\hat T(t)\|^2_{L_q(\Omega)},$$

for a.e. $t \in I$ with $L(M) = \max\{L_1(M), L_2(M)\}$. For the right hand side, we further have, due to interpolation inequalities, the bound

$$\|\hat{T}(t)\|_{L_q(\Omega)} \leq \|\hat{T}(t)\|^{1-\theta}_{L_{q^*}(\Omega)} \|\hat{T}(t)\|^{\theta}_{L_2(\Omega)} \leq |\Omega|^{\frac{1-\theta}{q^*}} \|\hat{T}(t)\|^{1-\theta}_{L_\infty(\Omega)} \|\hat{T}(t)\|^{\theta}_{L_2(\Omega)},$$

for all $q^* > q$ and $\theta \in (0,1)$ with $1/q = \theta/2 + (1-\theta)/q^*$. Altogether we get

$$\partial_t \|\hat{T}(t)\|^2_{L_2(\Omega)} \leq c_5 \|\rho_2(t)\|^2_{W^1_p(\Omega)} \|\hat{T}(t)\|^{2\theta}_{L_2(\Omega)},$$

for some $c_5 > 0$, which is equivalent to a nonlinear integral inequality of Gronwall-Bellman-Bihari type,[11] given by

$$\|\hat{T}(t)\|^2_{L_2(\Omega)} \leq \|\hat{T}_0\|^2_{L_2(\Omega)} + c_5 \int_0^t \|\rho_2(\tau)\|^2_{W^1_p(\Omega)} \Phi\Big(\|\hat{T}(\tau)\|^2_{L_2(\Omega)}\Big)\, d\tau,$$

with $\Phi(x) = x^\theta$ and $\|\hat{T}_0\|^2_{L_2(\Omega)} = 0$. Applying Theorem 3.2 of Ref. 11 to the above inequality, we obtain $\|\hat{T}(t)\|^2_{L_2(\Omega)} = 0$ for a.e. $t \in I$ and hence $\hat{T} = 0$ as well as $\hat{\rho} = 0$ a.e. in Q, which concludes the assertion. □

Remark 2.4. Observe that in Theorem 2.3, we required that $\rho(t) \in W^1_p(\Omega)$ with $p > 2$. This may be obtained by providing sufficiently smooth data and sufficiently smooth boundary $\partial\Omega$ as discussed in Remark 2.2.

We conclude this section by making the following assumption:

(**A3**) Let $p \geq n$ and $r > 4$. We assume $\rho_\partial \in L_r(I; L_p(\partial\Omega))$, $T_b \in L_r(I; L_p(\Omega))$, $T_\partial \in L_r(I; L_p(\partial\Omega))$ and $T_0 \in L_\infty(\Omega)$. We further assume that $p_0 \geq 3$ in Remark 2.2.

Theorem 2.4. *Under assumptions (**A1-A3**), we obtain a unique state $y = (\rho, T) \in \mathcal{X}$ for any given $d \in \mathcal{U}$, where $\mathcal{X} := V_{p_0,r} \times \mathcal{K}$, fulfilling the estimate*

$$\|y\|_{\mathcal{X}} \leq c_\chi \left(\|T_0\|_{L_\infty(\Omega)} + \|\rho_\partial\|_{L_r(I;L_p(\partial\Omega))} + \|T_b\|_{L_r(I;L_p(\Omega))} + \|T_\partial\|_{L_r(I;L_p(\partial\Omega))} \right).$$

3. The Linearized Equation and its Adjoint

3.1. *Linear State Vectors*

As in Sec. 2 we let $d \in \mathcal{U}$ be fixed but arbitrary throughout this section. Due to the continuous F-differentiability of β_d and μ_d on $\mathcal{K}$ we can consider the linearization of the nonlinear SP_1-system (2.6), given by

$$DE_d(y)[v] = g \quad \text{in } Z^*, \tag{3.1}$$

for $y, v \in \mathcal{X}$, where $DE_d\colon \mathcal{X} \to \mathcal{L}(\mathcal{X}; Z^*)$ is continuous and $g = (g_\rho, g_T, g_0) \in Z^*$. Due to density argument of the embedding $\mathcal{X} \hookrightarrow X_{2,2}$, we may extend the derivative at each state $y = (y_\rho, y_T) \in \mathcal{X}$ to a linear operator $A_y \in \mathcal{L}(X_{2,2}; Z^*)$, given by

$$\begin{aligned}\langle A_{y,1}\, v, w_1\rangle &= (\frac{1}{3\beta_d(y_T)}\nabla v_\rho, \nabla w_1) + (\mu_d(y_T)\, v_\rho, w_1) + (\gamma v_\rho, w_1)_\partial \\ &\qquad - (\frac{1}{3\beta_d^2(y_T)}\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho, \nabla w_1) + (\partial_2\mu_d(y_T)[v_T]\, y_\rho, w_1), \quad (3.2a)\\ \langle A_{y,2}\, v, w_2\rangle &= \langle \dot v_T, w_2\rangle + (\kappa\nabla v_T, \nabla w_2) + (b\, v_T - \partial_2\mu_d(y_T)[v_T]\, y_\rho, w_2) \\ &\qquad + (\alpha v_T, w_2)_\partial - (\mu_d(y_T)\, v_\rho, w_2), \quad (3.2b)\\ \langle A_{y,3}\, v, w_3\rangle &= (v_T(0), w_3), \quad (3.2c)\end{aligned}$$

for all $v = (v_\rho, v_T) \in X_{2,2}$, $w \in Z$. Note that we identified $\partial_2\beta_d(y_T)$ and $\partial_2\mu_d(y_T)$ in $\mathcal{L}(\mathcal{K}; L_2(Q))$ with their extensions in $\mathcal{L}(L_2(Q))$ respectively, which are well defined since $\mathcal{K}$ is dense in $L_2(Q)$.

Theorem 3.1. *Assume (**A1-A3**). Let $y = (y_\rho, y_T) \in \mathcal{X}$ and $g = (g_\rho, g_T, g_0) \in Z^*$. Then the problem: Find $v = (v_\rho, v_T) \in X_{2,2}$ such that*

$$A_y\, v = g \quad \text{in } Z^*, \tag{3.3}$$

where $A_y\colon X_{2,2} \to Z^*$ as defined in (3.2), has a unique solution.

Moreover, $A_y \in \mathcal{L}(X_{2,2}; Z^*)$ is a homeomorphism.

Proof. For the two last terms in (3.2a) we have the following bounds

$$|(\frac{1}{3\beta_d^2(y_T)}\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho, \nabla w_1)| \le \frac{1}{3\underline{\beta_d}^2}\|\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho\|_{L_2(Q)}\|\nabla w_1\|_{L_2(Q)},$$

with

$$\begin{aligned}\|\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho\|_{L_2(Q)} &\le \|\nabla y_\rho\|_{L_2(I;L_3(\Omega))}\|\partial_2\beta_d(y_T)[v_T]\|_{L_\infty(I;L_6(\Omega))}\\ &\le c'_\beta\|y_\rho\|_{V_{3,2}}\|v_T\|_{L_2(I;L_6(\Omega))} \le c_\beta\|y_\rho\|_{V_{3,2}}\|v_T\|_{V_{2,2}},\end{aligned}$$

as given in Theorem 2.1. Similarly, we obtain

$$\begin{aligned}|(\partial_2\mu_d(y_T)[v_T]\, y_\rho, w_1)| &\le c'_\mu\|y_\rho\|_{L_2(I;L_3(\Omega))}\|v_T\|_{L_2(I;L_6(\Omega))}\|w_1\|_{L_2(Q)}\\ &\le c_\mu\|y_\rho\|_{V_{3,2}}\|v_T\|_{V_{2,2}}\|w_1\|_{L_2(Q)}.\end{aligned}$$

Suppose that $v_T \in V_{2,2}$ is given. Consider the problem: For $g_\rho \in {V_{2,2}}^*$, find $v_\rho \in V_{2,2}$ such that

$$a_\rho(v_\rho, w_1) = (\frac{1}{3\beta_d^2(y_T)}\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho, \nabla w_1) - (\partial_2\mu_d(y_T)[v_T]\, y_\rho, w_1) + \langle g_\rho, w_1\rangle,$$

where a_ρ is the continuous bilinear form given by

$$a_\rho(v_\rho, w_1) = (\frac{1}{3\beta_d(y_T)}\nabla v_\rho, \nabla w_1) + (\mu_d(y_T)v_\rho, w_1) + (\gamma v_\rho, w_1)_\partial,$$

which is clearly coercive in $V_{2,2}$ since

$$a_\rho(v_\rho, v_\rho) \ge \frac{1}{3\overline{\beta}_d}\|\nabla v_\rho\|^2_{L_2(Q)} + \underline{\mu_d}\|v_\rho\|^2_{L_2(Q)} + \underline{\gamma}\|v_\rho\|^2_{L_2(\Sigma)} \ge c_\rho\|v_\rho\|^2_{V_{2,2}},$$

with $c_\rho = \min\{(3\overline{\beta}_d)^{-1}, \underline{\mu_d}, \underline{\gamma}\}$. Thus by Lax-Milgram, we obtain a unique solution $v_\rho \in V_{2,2}$ with the bound

$$\|v_\rho\|_{V_{2,2}} \le \frac{1}{c_\rho}\Big((c_\beta + c_\mu)\|y_\rho\|_{V_{3,2}}\|v_T\|_{V_{2,2}} + \|g_\rho\|_{V_{2,2}^*}\Big). \tag{3.4}$$

Now define the bilinear form a_T as follows.

$$a_T(v_T, w_2) = (\kappa\nabla v_T, \nabla w_2) + (b\, v_T, w_2) + (\alpha\, v_T, w_2)_\partial \\ - (\partial_2\mu_d(y_T)[v_T]\, y_\rho + \mu_d(y_T)\, v_{\rho,1}, w_2),$$

where

$$v_{\rho,1} = v_{\rho,1}[\frac{1}{3\beta_d^2(y_T)}\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho - \partial_2\mu_d(y_T)[v_T]\, y_\rho] \quad \text{and} \quad v_{\rho,2} = v_{\rho,2}[g_\rho],$$

which is well-defined due to linearity. Clearly a_T is continuous on $V_{2,2} \times V_{2,2}$.

We claim that a_T is weakly coercive in $V_{2,2} \hookrightarrow L_2(Q)$, i.e. it fulfills a Gårding inequality.[13] Indeed, by applying Hölder's and Young's inequalities together with the bounds derived so far, we obtain for $\epsilon > 0$:

$$\begin{aligned} a_T(v_T, v_T) &\ge \underline{\kappa}\|\nabla v_T\|^2_{L_2(Q)} + \underline{b}\|v_T\|^2_{L_2(Q)} + \underline{\alpha}\|v_T\|^2_{L_2(\Sigma)} \\ &\qquad - c_\mu\|y_\rho\|_{V_{3,2}}\|v_T\|_{V_{2,2}}\|v_T\|_{L_2(Q)} - \overline{\mu}_d\|v_{\rho,1}\|_{L_2(Q)}\|v_T\|_{L_2(Q)} \\ &\ge \lambda_1\|v_T\|^2_{V_{2,2}} - \lambda_2\|v_T\|^2_{L_2(Q)} \end{aligned}$$

where

$$\lambda_1 = c_T - \frac{\epsilon}{2}\Big(c_\mu^2 + \frac{\overline{\mu}_d^2}{c_\rho^2}(c_\beta + c_\mu)^2\Big)\|y_\rho\|^2_{V_{3,2}} \quad \text{and} \quad \lambda_2 = \frac{1}{\epsilon} - \underline{b},$$

with $c_T = \min\{\underline{\kappa}, \underline{\alpha}\}$. With an appropriate $\epsilon > 0$ such that $\lambda_1 > 0$, we finally obtain,

$$a_T(v_T, v_T) + \lambda_2\|v_T\|^2_{L_2(Q)} \ge \lambda_1\|v_T\|^2_{V_{2,2}}, \tag{3.5}$$

which affirms our claim.

Now consider the auxiliary problem: Find $v_T \in \mathcal{W}$ such that

$$\langle \dot{v}_T, w\rangle + a_T(v_T, w) = \langle \mu_d(y_T)v_{\rho,2} + g_T, w\rangle \quad \text{for all } w \in V_{2,2}, \tag{3.6}$$

with initial condition $v_T(0) = g_0 \in L_2(\Omega)$.

Since $a_T\colon V_{2,2} \times V_{2,2} \to \mathbb{R}$ is continuous and weakly coercive in $V_{2,2} \hookrightarrow L_2(Q)$ as shown in (3.5), standard theory for linear parabolic equations gives us the existence and uniqueness of a solution $v_T \in \mathcal{W}$ fulfilling (3.6) (cf. Sec. 11.1 of Ref 13), with a constant $c_2(y_\rho) > 0$ depending on $y_\rho \in V_{p_0,r}$, the bound

$$\|v_T\|_{V_{2,2}} \le c_2(y_\rho)\left(\|v_0\|_{L_2(\Omega)} + \|g_\rho\|_{V_{2,2}^*} + \|g_T\|_{V_{2,2}^*}\right) = c_2(y_\rho)\|g\|_{Z^*},$$

which further yields for $v_\rho \in V_{2,2}$ its existence, uniqueness and the bound

$$\|v_\rho\|_{V_{2,2}} \le \frac{1}{c_\rho}\left(c_3(y_\rho)\|g\|_{Z^*} + \|g_\rho\|_{V_{2,2}^*}\right) \le c_4(y_\rho)\|g\|_{Z^*},$$

with constants $c_3(y_\rho)$, $c_4(y_\rho) > 0$, according to (3.4). Since $\dot{v}_T \in {V_{2,2}}^*$ fulfills (3.6), we have also the bound

$$\|\dot{v}_T\|_{{V_{2,2}}^*} \leq c_5(y_\rho)\|g\|_{Z^*},$$

with a constant $c_5(y_\rho) > 0$, which yields altogether the assertion. □

Lemma 3.1. *Assume (**A1-A3**). Let $y \in \mathcal{X}$ and $(g_\rho, g_T, g_0) \in Y$ be given, where*

$$Y := L_r(I; W^1_{p'}(\Omega)^*) \times L_r(I; L_p(\Omega)) \times L_\infty(\Omega),$$

for $p \geq n$ and $r > 4$. Then the unique solution $v = (v_\rho, v_T) \in X_{2,2}$ of (3.3) is in fact in $\mathcal{X}$.

Proof. We start with considering the auxiliary problem given in (3.6). Notice that $v_{\rho,2} \in V_{p_0,r}$ since $g_\rho \in L_r(I; W^1_{p'}(\Omega)^*)$. Define $\tilde{v}_T = v_T e^{-\lambda_2 t}$, with λ_2 as given in (3.5). Due to the linearity of a_T, (3.6) then becomes

$$\langle \dot{\tilde{v}}_T, w\rangle + a_{T,\lambda_2}(\tilde{v}_T, w) = \langle \big(\mu_d(y_T)\, v_{\rho,2} + g_T\big)e^{-\lambda_2 t}, w\rangle, \tag{3.7}$$

for all $w \in V_{2,2}$, where $a_{T,\lambda_2}\colon V_{2,2} \times V_{2,2} \to \mathbb{R}$ is the bilinear form

$$a_{T,\lambda_2}(w_1, w_2) = a_T(w_1, w_2) + \lambda_2(w_1, w_2) \quad \text{for all } (w_1, w_2) \in V_{2,2} \times V_{2,2}.$$

Following the arguments made in Lemma 2.2 for $\tilde{v}_T$ with $g_0 \in L_\infty(\Omega)$, we conclude that $\tilde{v}_T \in L_\infty(Q)$ and thus also $v_T = \tilde{v}_T e^{\lambda_2 t} \in L_\infty(Q)$.

Notice that since $v_T \in L_\infty(Q)$, the right hand side to the problem

$$a_\rho(v_{\rho,1}, w) = (\frac{1}{3\beta_d^2(y_T)}\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho, \nabla w) - (\partial_2\mu_d(y_T)[v_T]\, y_\rho, w),$$

for all $w \in V_{2,2}$ with $y_\rho \in V_{p_0,r}$ is indeed in $L_r(I; W^1_{p'}(\Omega)^*)$, thus implying that $v_{\rho,1} \in V_{p_0,r}$ and consequently $v_\rho = v_{\rho,1} + v_{\rho,2} \in V_{p_0,r}$. Altogether we have $(v_\rho, v_T) \in \mathcal{X}$ as claimed. □

3.2. *Adjoint State Vectors*

Next we study the adjoint operator.

Theorem 3.2. *Assume (**A1-A3**). Let $y = (y_\rho, y_T) \in \mathcal{X}$ and $h = (h_\rho, h_T) \in {X_{2,2}}^*$. Then the problem: Find $\xi = (\xi_\rho, \xi_T, \xi_0) \in Z$ such that*

$$A_y^*\,\xi = h \quad \text{in } {X_{2,2}}^*,$$

where $A_y^*\colon Z \to {X_{2,2}}^*$ is the adjoint operator to A_y, has a unique solution.

Furthermore, if $h \in {V_{2,2}}^* \times {V_{2,2}}^*$, then we have that $(\xi_\rho, \xi_T) \in X_{2,2}$, and ξ can be characterized as the variational solution of

$$-\nabla \cdot \left(\frac{1}{3\beta_d(y_T)}\nabla\xi_\rho\right) + \mu_d(\xi_\rho - \xi_T) = h_\rho, \tag{3.8a}$$

$$-\partial_t\xi_T - \nabla\cdot(\kappa\nabla\xi_T) + b\,\xi_T - \partial_2\beta_d(y_T)^*[\frac{1}{3\beta_d^2(y_T)}\nabla y_\rho \cdot \nabla\xi_\rho] + \partial_2\mu_d(y_T)^*[y_\rho\,\xi_\rho - y_\rho\,\xi_T] = h_T \quad \text{in } Q, \tag{3.8b}$$

with boundary conditions

$$\frac{1}{3\beta_d(y_T)}\,\partial_n\xi_\rho + \gamma\xi_\rho = 0, \tag{3.8c}$$

$$\kappa\,\partial_n\xi_T + \alpha\xi_T = 0 \quad \text{on } \Sigma, \tag{3.8d}$$

with initial and terminal conditions $\xi_T(0) = \xi_0$ *and* $\xi_T(t_*) = 0$ *in* $L_2(\Omega)$ *respectively.*

Proof. We start by giving a formal representation of the adjoint, i.e.

$$\begin{aligned}
&\langle v, A_y^*\,\xi\rangle = \\
&= (\nabla v_\rho, \frac{1}{3\beta_d(y_T)}\nabla\xi_\rho) + (v_\rho, \mu_d(y_T)(\xi_\rho - \xi_T)) + (v_\rho, \gamma\xi_\rho)_\partial \\
&\quad + (v_T(0), \xi_0) + \langle \dot v_T, \xi_T\rangle + (\nabla v_T, \kappa\nabla\xi_T) + (v_T, b\,\xi_T) + (v_T, \alpha\,\xi_T)_\partial \\
&\qquad - (v_T, \partial_2\beta_d(y_T)^*[\frac{1}{3\beta_d^2(y_T)}\nabla y_\rho\cdot\nabla\xi_\rho]) + (v_T, \partial_2\mu_d(y_T)^*[y_\rho\,\xi_\rho - y_\rho\,\xi_T]) \\
&= (\frac{1}{3\beta_d(y_T)}\nabla v_\rho - \frac{1}{3\beta_d^2(y_T)}\partial_2\beta_d(y_T)[v_T]\,\nabla y_\rho, \nabla\xi_\rho) + (\mu_d(y_T)v_\rho, \xi_\rho) \\
&\quad + (\partial_2\mu_d(y_T)[v_T]\,y_\rho, \xi_\rho) - (\partial_2\mu_d(y_T)[v_T]\,y_\rho, \xi_T) + (\gamma v_\rho, \xi_\rho)_\partial + (v_T(0), \xi_0) \\
&\qquad + \langle \dot v_T, \xi_T\rangle + (\kappa\nabla v_T, \nabla\xi_T) + (b\,v_T, \xi_T) - (\mu_d(y_T)\,v_\rho, \xi_T) + (\alpha\,v_T, \xi_T)_\partial \\
&= \langle A_y\,v, \xi\rangle.
\end{aligned}$$

Due standard results from functional analysis we obtain the continuous invertibility of the adjoint operator $A_y^* \in \mathcal{L}(Z; {X_{2,2}}^*)$, i.e. $A_y^{-*} \in \mathcal{L}({X_{2,2}}^*; Z)$. Moreover, we have the bound

$$\|\xi\|_Z \le \|A_y^{-*}\|_{\mathcal{L}({X_{2,2}}^*;Z)}\|h\|_{{X_{2,2}}^*}.$$

Now let $h \in {V_{2,2}}^* \times {V_{2,2}}^*$ and $\dot\xi_T$ denote the distributional time derivative of $\xi_T \in V_{2,2}$. Notice that the function

$$\begin{aligned}
t \mapsto B(t) := \Big(\nabla\cdot(\kappa\nabla\xi_T) - b\,\xi_T + \partial_2\beta_d(y_T)^*[\frac{1}{3\beta_d^2(y_T)}\nabla y_\rho\cdot\nabla\xi_\rho] \\
- \partial_2\mu_d(y_T)^*[y_\rho\,\xi_\rho - y_\rho\,\xi_T] + h_T\Big)(t)
\end{aligned}$$

is in ${V_{2,2}}^*$. Then

$$\int_I \langle -\dot\xi_T(t), v\rangle\,\varphi(t)\,dt = \int_I (B(t), v)\,\varphi'(t)\,dt, \quad \text{for all} \quad v \in V_{2,2}, \quad \varphi \in \mathcal{C}_0^\infty(I;\mathbb{R}),$$

which by definition implies that $\dot\xi_T \in {V_{2,2}}^*$. Due to the bound above, we obtain $\xi \in X_{2,2} \times L_2(\Omega)$. From the embedding $W \hookrightarrow \mathcal{C}(I; L_2(\Omega))$ we obtain the initial and terminal conditions $\xi_T(0) = \xi_0$ and $\xi_T(t_*) = 0$ in $L_2(\Omega)$ respectively. □

4. Existence of an Optimal Control

In this section we make the following assumption regarding the cost functional of the optimal control problem.

(**A4**) Let $\mathcal{U} = H^2(\mathbb{R})$ and $J\colon \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ denote a cost functional which is assumed to be twice continuously F-differentiable with locally Lipschitz continuous second derivatives. Further, let J be of separated type, i.e., $J(d, y) = J_1(y) + J_2(d)$ and radially unbounded with respect to d for every y, bounded from below and weakly lower semi-continuous.

Next, we want to give the precise mathematical statement of the optimal control problem (1.2). We define the control/state pair $(d, y = (y_\rho, y_T)) \in \mathcal{U} \times \mathcal{X}$ and the nonlinear operator $E : \mathcal{U} \times \mathcal{X} \to Z^*$ as in (2.6). Now let $J\colon \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ be a cost functional that fulfills assumption (**A4**), the minimization problem (1.2) can then be written as

$$\min J(d, y) \text{ over } (d, y) \in \mathcal{U} \times \mathcal{X} \text{ subject to } E(d, y) = 0 \text{ in } Z^*. \tag{4.1}$$

Example 4.1. Assume (**A1**). Let $\epsilon > 0$ be arbitrary and $\{\delta^\epsilon\}_{\epsilon>0}$ be a Dirac-sequence. We define for each i the sequence $\{\delta^\epsilon_{x_i}\}_{\epsilon>0}$ as follows

$$\delta^\epsilon_{x_i} * u = \int_\Omega u(x)\, \delta^\epsilon_{x_i}(x)\, dx = \int_\Omega u(x)\, \delta^\epsilon(x_i - x)\, dx, \tag{4.2}$$

for any $u \in H^1(\Omega)$. Now let $p \in (1, \infty)$ and consider the cost functional $J_\epsilon\colon \mathcal{U}\times\mathcal{X} \to \mathbb{R}$ given by

$$J_\epsilon(d, y) = \frac{1}{p} \sum_i \|\big(\delta^\epsilon_{x_i} * y_T(\cdot)\big) - T_{m,i}\|^p_{L_p(I)} + \frac{\lambda}{2}\|d - d_c\|^2_{\mathcal{U}}, \tag{4.3}$$

for finitely many given measurements $T_{m,i} \in L_p(I)$ at points $x_i \in \Omega$, common parameter $d_c \in \mathcal{U}$ and some $\lambda > 0$. Notice that $\lim_{\epsilon\to 0} \delta^\epsilon_{x_i} = \delta_{x_i}$ in $\mathcal{D}(\Omega)^*$, where δ_{x_i} is the Dirac-distribution on x_i given by $\delta_{x_i} * u = u(x_i)$ for $u \in H^1(\Omega)$. Due to the embedding $H^1(\Omega) \to \mathcal{C}(\bar{\Omega})$ for $n = 1$, $u(x_i)$ exists and hence $\delta_{x_i} \in H^1(\Omega)^*$. Since $\delta^\epsilon_{x_i}$ is also in $H^1(\Omega)^*$ for all $\epsilon > 0$, we have that $\lim_{\epsilon\to 0} \delta^\epsilon_{x_i} = \delta_{x_i}$ in $H^1(\Omega)^*$ and thus

$$\lim_{\epsilon\to 0} J_\epsilon(d, y) = \frac{1}{p} \sum_i \|y_T(\cdot)(x_i) - T_{m,i}\|^p_{L_p(I)} + \frac{\lambda}{2}\|d - d_c\|^2_{H^2(\mathbb{R})} =: J(d, y),$$

for $(d, y) \in \mathcal{U} \times \mathcal{X}$, which easily follows from the continuity of norms.

Due to the lack of an embedding theorem for $n = 2, 3$ respectively, this convergence fails. However, the membership of $\delta^\epsilon_{x_i}$ in $H^1(\Omega)^*$ for all $\epsilon > 0$ still holds and so we may make use of J_ϵ with arbitrarily small $\epsilon > 0$.

4.1. *Existence of Minimizer*

In this subsection we prove the existence of a minimizer. In general, uniqueness does not hold since the set of solutions for $E(d, y) = 0$ in Z^* may not be convex. The existence however can easily be shown.

Theorem 4.1. *Assume (**A1-A4**). Then there exists a $(d_*, y_*) \in \mathcal{U} \times \mathcal{X}$ solving the constraint minimization problem (4.1).*

Proof. Let $\{(d_k, y_k)\}_{k\in\mathbb{N}} \in \mathcal{U} \times \mathcal{X}$ be a minimizing sequence such that

$$j = \inf_{(d,y)\in\mathcal{U}\times\mathcal{X}} J(d, y) = \liminf_{k\in\mathbb{N}} J(d_k, y_k) \quad \text{and} \quad E(d_k, y_k) = 0 \text{ in } Z^*,$$

for all $k \in \mathbb{N}$, where $j > -\infty$ by definition of J. The radial unboundedness of J with respect to d implies that $\{d_k\}_{k\in\mathbb{N}}$ is bounded in $\mathcal{U}$. Since $\mathcal{U}$ is reflexive there exists a weakly convergent subsequence, denoted again by $\{d_k\}_{k\in\mathbb{N}}$ such that

$$d_k \rightharpoonup d_* \quad \text{in} \quad \mathcal{U}.$$

Since $\mathcal{U}$ is closed and convex, $d_* \in \mathcal{U}$. From (**A2**) and the uniform bounds with respect to d_k for the solutions of (2.6) obtained in Theorem 2.4, we conclude the boundedness of $\{y_k\}_{k\in\mathbb{N}}$ in $\mathcal{X}$. Similarly, we obtain a weakly convergent subsequence, denoted again by $\{y_k\}_{k\in\mathbb{N}}$ such that

$$y_k \rightharpoonup y_* \quad \text{in} \quad \mathcal{X}.$$

Due to the weak lower semicontinuity of J, we have

$$J(d_*, y_*) \le \liminf_{k\in\mathbb{N}} J(d_k, y_k) = j,$$

which directly implies $J(d_*, y_*) = j$.

We are left to show that (d_*, y_*) fulfills the constraints, i.e. (d_*, y_*) solves (2.6). Due to the standard compact embedding theorems for $H^2(\mathbb{R}) \hookrightarrow \mathcal{C}_b^1(\mathbb{R})$, we obtain a strongly convergent subsequence, denoted again by $\{d_k\}_{k\in\mathbb{N}}$ such that

$$d_k \to d_* \quad \text{in} \quad \mathcal{C}_b^1(\mathbb{R}).$$

Similarly, standard compact embedding theorems imply the strong convergence of a subsequence of $\{y_{T,k}\}_{k\in\mathbb{N}}$, denoted again by $\{y_{T,k}\}_{k\in\mathbb{N}}$ in $L_2(Q)$, i.e.,

$$y_{T,k} \to y_{T,*} \quad \text{in} \quad L_2(Q).$$

Since $d_* \in \mathcal{C}_b^1(\mathbb{R})$, we further have that $d_*\colon L_2(Q) \to L_2(Q)$ is continuous as a Nemytskij operator.[8] Thus, we have a strongly convergent sequence $\{d_*(y_{T,k})\}_{k\in\mathbb{N}}$ in $L_2(Q)$ and consequently a subsequence, denoted again by $\{d_*(y_{T,k})\}_{k\in\mathbb{N}}$ such that

$$d_*(y_{T,k}) \to d_*(y_{T,*}) \quad \text{a.e. in} \quad Q.$$

Due to its uniform boundedness in $L_\infty(Q)$ we have, by Lebesgue's dominated convergence theorem, that

$$d_*(y_{T,k}) \to d_*(y_{T,*}) \quad \text{in} \quad L_\infty(Q),$$

which yields together with Theorem 2.1

$$\begin{aligned}\beta(d_k, y_{T,k}) &\to \beta(d_*, y_{T,*}) \quad \text{in} \quad L_\infty(Q),\\ \mu(d_k, y_{T,k}) &\to \mu(d_*, y_{T,*}) \quad \text{in} \quad L_\infty(Q).\end{aligned}$$

From the continuity of the function $(\cdot)^{-1}\colon \mathbb{R}_{>0} \to \mathbb{R}$ and the uniform boundedness of $\beta(d, y_T)$ in $L_{\infty,>0}(Q)$ we may pass to the limit in (2.6), thus concluding the assertion. □

4.2. *Control-to-State Map and Derivatives*

Let $d \in \mathcal{U}$ be fixed but arbitrary. Suppose E is given by (2.6) and fulfills the requirements of Theorem 2.4, then we have the existence of a state $y \in \mathcal{X}$. This implicitly defines a control-to-state map $d \mapsto y(d)$. The main task in this section is to study and analyze this mapping.

Theorem 4.2. *Assume (**A1**-**A3**). Then the mapping $d \mapsto y(d)$ is continuously F-differentiable as a mapping $\mathcal{U} \to \mathcal{X}$ and its derivative is given by*

$$y'(d) = -D_y E(d, y(d))^{-1} D_d E(d, y(d)). \tag{4.4}$$

Proof. The idea (see also Ref. 3, 16) is to split the nonlinear operator E into its linear part L acting on y, as well as its nonlinear part N and constant part f, i.e.,

$$E(d, y) = L\,y + N(d, y) - f,$$

where $L\colon X_{2,2} \to Z^*$, $N\colon \mathcal{U} \times \mathcal{X} \to Y$, as given in Lemma 3.1, and $f \in Z^*$ are defined by

$$\begin{aligned}\langle L\,y, w\rangle &= (\epsilon\nabla y_\rho, \nabla w_1) + (\epsilon\, y_\rho, w_1) + (\gamma\, y_\rho, w_1)_\partial + (y_T(0), w_3)\\ &\quad + \langle \dot{y}_T, w_2\rangle + (\kappa\nabla y_T, \nabla w_2) + (b\, y_T, w_2) + (\alpha\, y_T, w2)_\partial - (\epsilon\, y_\rho, w_2),\\ \langle N(d, y), w\rangle &= (\Big(\frac{1}{3\beta(d, y_T)} - \epsilon\Big)\nabla y_\rho, \nabla w_1) + (\big(\mu(d, y_T) - \epsilon\big)y_\rho, w_1 - w_2),\\ \langle f, w\rangle &= (\gamma\,\rho_\partial, w_1)_\partial + (b\,T_b, w_2) + (\alpha\,T_\partial, w_2)_\partial + (T_0, w_3),\end{aligned}$$

with $0 < \epsilon < \min\{(1/3\bar{\beta}), \underline{\mu}\}$.

By assumption (**A3**) and Theorem 2.4, we have $L^{-1}f \in \mathcal{X}$. Notice that Theorem 2.4 also holds true for elements from Y, i.e. $L^{-1}\colon Y \to \mathcal{X}$. Define the operator $R\colon \mathcal{U} \times \mathcal{X} \to \mathcal{X}$ by

$$R(d, y) = y + L^{-1}N(d, y) - L^{-1}f,$$

which is well-defined by the arguments above.

First, note that R is continuously F-differentiable. Indeed, $N\colon \mathcal{U} \times \mathcal{X} \to Y$ is continuously F-differentiable due to Theorem 2.1. Since the linear operator L^{-1} is also continuously F-differentiable, we may apply the chain rule to affirm our claim.

Next, we claim that $D_y R(d,y)\colon \mathcal{X} \to \mathcal{X}$ is invertible for all $(d,y) \in \mathcal{U} \times \mathcal{X}$, i.e., we have to show that for any $g \in \mathcal{X}$ there exists a unique $u \in \mathcal{X}$ such that

$$D_y R(d,y)\, u = u + L^{-1} D_y N(d,y)\, u = g \quad \text{in } \mathcal{X}.$$

By introducing $v = u - g$, we get

$$v + L^{-1} D_y N(d,y)\,(v+g) = 0 \quad \text{in } \mathcal{X},$$

which is equivalent to

$$L\, v + D_y N(d,y)\, v = -D_y N(d,y)\, g \quad \text{in } Z^*. \tag{4.5}$$

Notice that the left hand side corresponds to the linearized system A_y given in Sec. 3. Since the right hand side belongs to Y, Theorem 3.1 and Lemma 3.1 asserts the existence and uniqueness of a $v \in \mathcal{X}$ solving (4.5); thus also a unique $u = v + g \in \mathcal{X}$.

We then facilitate the implicit function theorem for R, which gives us the continuous F-differentiability of $d \mapsto y(d)$ and the equation

$$y'(d) = -D_y R(d, y(d))^{-1} D_d R(d, y(d)).$$

Since E is equivalent to R by the fact that $R = L^{-1}E$, the results obtained for R are valid for E. Due to linearity of L we finally obtain

$$\begin{aligned} y'(d) &= -D_y (L^{-1}E)(d, y(d))^{-1} D_d (L^{-1}E)(d, y(d)) \\ &= -D_y E(d, y(d))^{-1} D_d E(d, y(d)), \end{aligned}$$

which concludes the proof. □

4.3. *Reduced Optimal Control Problem*

Let $J\colon \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ be a cost functional fulfilling (**A4**). Due to the existence of an F-differentiable control-to-state map $d \mapsto y(d)$ given by Theorem 4.2, we may introduce the reduced optimal control problem, which reads as follows:

$$\min \hat{J}(d) \text{ over } d \in \mathcal{U} \text{ subject to } \hat{E}(d) = 0 \text{ in } Z^*, \tag{4.6}$$

where $\hat{J}(d) = J(d, y(d))$ and $\hat{E}(d) = E(d, y(d))$. Similarly, we set $\hat{\beta}(d) = \beta(d, \mathcal{P}_T[y](d))$ and $\hat{\mu}(d) = \mu(d, \mathcal{P}_T[y](d))$, where $\mathcal{P}_T$ is the canonical projection from $\mathcal{X}$ into $\mathcal{K}$.

Example 4.2. As an example, we consider the reduced optimal control for the cost functional (4.3) given by

$$\hat{J}_\epsilon(d) = \frac{1}{p} \sum_i \|\big(\delta^\epsilon_{x_i} * y_T(d)(\cdot)\big) - T_{m,i}\|^p_{L_p(Q)} + \frac{\lambda}{2}\|d - d_c\|^2_{\mathcal{U}}, \tag{4.7}$$

for any $\epsilon > 0$ and $p \in (0, \infty)$. By definition of Dirac-sequences we have

$$\int_\Omega \delta^\epsilon(x)\, dx = 1 \quad \text{for all } \epsilon > 0.$$

Thus, (4.7) can be rewritten as

$$\hat{J}_\epsilon(d) = \frac{1}{p}\sum_i \|\big(y_T(d) - T_{m,i}\big)\delta^\epsilon_{x_i}\|^p_{L_p(Q)} + \frac{\lambda}{2}\|d - d_c\|^2_{\mathcal{U}}, \tag{4.8}$$

where we used (4.2).

4.4. *The First-Order Optimality Condition*

Let $J\colon \mathcal{U}\times\mathcal{X}\to\mathbb{R}$ be a cost functional fulfilling (**A4**) and $\hat{J}$ its corresponding reduced cost functional as in (4.6). The necessary first-order optimality condition is given by

$$\hat{J}'(d) = 0.$$

Using the chain rule and applying (4.4) of Theorem 4.2 we obtain

$$\begin{aligned}\hat{J}'(d)[v_d] &= \langle D_y J(d, y(d)), y'(d)[v_d]\rangle_{\mathcal{X}^*,\mathcal{X}} + \langle D_d J(d, y(d)), v_d\rangle_{\mathcal{U}^*,\mathcal{U}}\\ &= \langle D_d E(d, y(d))^*[\xi] + D_d J(d, y(d)), v_d\rangle_{\mathcal{U}^*,\mathcal{U}},\end{aligned}$$

for all $v_d \in \mathcal{U}$, where we introduced the adjoint variable

$$\xi = -D_y E(d, y(d))^{-*} D_y J(d, y(d)) \quad \text{in } Z.$$

Since the above equality holds for all $v_d \in \mathcal{U}$, we have

$$\hat{J}'(d) = D_d\hat{E}(d)^*[\xi] + D_d J(d, y(d)) \quad \text{in } \mathcal{U}^*.$$

From the representation of the derivative $\hat{J}'$ and the adjoint variable $\xi \in Z$, we obtain the following theorem.

Theorem 4.3. *Let $J\colon \mathcal{U}\times\mathcal{X}\to\mathbb{R}$ be a cost functional fulfilling (**A4**) and $(d_*, y_*) \in \mathcal{U}\times\mathcal{X}$ be a solution of the constrained minimization problem (4.1). Then there exists a unique Lagrange multiplier $\xi_* \in Z$, which together with the optimal solution (d_*, y_*) satisfy the first-order optimality system*

$$\begin{aligned} E(d_*, y_*) &= 0 \quad \textit{in } Z^*,\\ D_y\hat{E}(d_*)^*[\xi_*] + D_y J(d_*, y_*) &= 0 \quad \textit{in } {X_{2,2}}^*,\\ D_d\hat{E}(d_*)^*[\xi_*] + D_d J(d_*, y_*) &= 0 \quad \textit{in } \mathcal{U}^*.\end{aligned}$$

Proof. Clearly $D_y\hat{E}(d_*) = A_{y_*}$. Since $D_y J(d_*, y_*) \in {X_{2,2}}^*$, by Theorem 3.2 we obtain a unique solution to the adjoint problem

$$A^*_{y_*}\xi = D_y J(d_*, y_*) \quad \text{in } {X_{2,2}}^*,$$

which is none other than the second equality; thus yielding the assertion. □

As an example, we consider the reduced cost functional $\hat{J}_\epsilon$ as given in (4.8) and give an explicit representation for its derivative $\hat{J}'_\epsilon$.

Theorem 4.4. *Let $p \in (0, \infty)$ and $\epsilon > 0$ be sufficiently small such that the support for each $\delta^\epsilon_{x_i}$ are disjoint. Then $\hat{J}_\epsilon$, as defined in (4.8) is F-differentiable with*

$$\hat{J}'_\epsilon(d) = \partial_1\hat{\beta}(d)^*[-\frac{1}{3\hat{\beta}^2(d)}\nabla y_\rho \cdot \nabla\xi_\rho] + \partial_1\hat{\mu}(d)^*[y_\rho(\xi_\rho - \xi_T)] + \lambda(d - d_c), \quad (4.9)$$

in $\mathcal{U}^*$ for all $d \in \mathcal{U}$, where $\xi = (\xi_\rho, \xi_T, \xi_0) \in Z$ is the solution to the adjoint problem

$$-A^*_{y(d)}\,\xi = h \quad \text{in } {X_{2,2}}^*,$$

with $h = (0, \big(y_T(d) - T_{m,i}\big)\delta^\epsilon_{\{x_i\}_i}) \in {X_{2,2}}^*$ and $\delta^\epsilon_{\{x_i\}_i}$ defined as in the proof.

Proof. The F-differentiability follows from the F-differentiability of norms and of the control-to-state map $d \mapsto y(d)$ as given in Theorem 4.2. We define $\delta^\epsilon_{\{x_i\}_i}$ simply as the sum of all $\delta^\epsilon_{x_i}$, i.e. $\delta^\epsilon_{\{x_i\}_i} = \sum_i \delta^\epsilon_{x_i}$. Since the support for each $\delta^\epsilon_{x_i}$ are disjoint by assumption, we have

$$\sum_i \big(y_T(d) - T_{m,i}\big)\delta^\epsilon_{x_i} = \big(y_T(d) - T_{m,i}\big)\delta^\epsilon_{\{x_i\}_i}.$$

Using (4.4) of Theorem 4.2 and the above equality we get by formal computations

$$\begin{aligned}\hat{J}'_\epsilon(d)[v_d] &= \langle (\big(y_T(d) - T_{m,i}\big)\delta^\epsilon_{\{x_i\}_i})^{p-1}, y'_T(d)[v_d]\rangle_{L_q(Q),\,L_p(Q)} + \lambda(d - d_c, v_d)_{\mathcal{U}} \\ &= \langle D_d\hat{E}(d)^*[\xi] + \lambda(d - d_c), v_d\rangle_{\mathcal{U}^*,\mathcal{U}},\end{aligned}$$

for all $v_d \in \mathcal{U}$, where $\xi = \xi(d) \in Z$ is the solution to the adjoint problem

$$-A^*_{y(d)}\,\xi = h \quad \text{in } {X_{2,2}}^*,$$

with $h = (0, (\big(y_T(d) - T_{m,i}\big)\delta^\epsilon_{\{x_i\}_i})^{p-1}) \in {X_{2,2}}^*$.

There is still to show the explicit representation of $D_d\hat{E}(d)^*[\xi]$. Differentiating E with respect to d at the point $(d, y) \in \mathcal{U} \times \mathcal{X}$ gives

$$\begin{aligned}\langle D_dE(d, y)[v_d], \xi\rangle = (&-\frac{1}{3\beta^2(d, y_T)}\partial_1\beta(d, y_T)[v_d]\,\nabla y_\rho, \nabla\xi_\rho) \\ &+ (\partial_1\mu(d, y_T)[v_d]\,y_\rho, \xi_\rho - \xi_T), \quad (4.10)\end{aligned}$$

for $v_d \in \mathcal{U}$ and $\xi \in Z$, where

$$\partial_1\beta(d, y_T)[v_d] = \varphi'_{\beta,1}\left(\int_0^\cdot d(y_T)(s)\,ds\right)\int_0^\cdot v_d(y_T)(\tau)\,d\tau, \quad (4.11)$$

$$\partial_1\mu(d, y_T)[v_d] = \varphi'_{\mu,1}\left(\int_0^\cdot d(y_T)(s)\,ds\right)\int_0^\cdot v_d(y_T)(\tau)\,d\tau. \quad (4.12)$$

Since $\mathcal{U}$ is a separable Hilbert space, it admits a countable orthonormal basis and is therefore isometrically isomorphic to l^2, via the map

$$i_{\mathcal{U}}\colon l^2 \to \mathcal{U}; \quad \{v_k\}_k \mapsto \sum_k v_k e_k,$$

for any given countable orthonormal basis $\{e_k\}_k \subset \mathcal{U}$. Using this fact, we may rewrite (4.11) with $v_d = \sum_k v_{d,k} e_k$ as

$$\partial_1\beta(d, y_T)[v_d] = \varphi'_{\beta,1}\left(\int_0^{\cdot} d(y_T)(s)\, ds\right) \sum_k \int_0^{\cdot} v_{d,k} e_k(y_T)(\tau)\, d\tau.$$

By simple computations, a change of integrals with the above equation, and the isometric isomorphism $i_{\mathcal{U}}$, we obtain for the first part of (4.10)

$$\begin{aligned}
&(-\frac{1}{3\beta^2(d, y_T)}\partial_1\beta(d, y_T)[v_d]\, \nabla y_\rho, \nabla\xi_\rho) \\
&\quad = \sum_k v_{d,k}\beta_k^*[-\frac{1}{3\beta^2(d, y_T)}\nabla y_\rho \cdot \nabla\xi_\rho] \\
&\quad\quad = \langle\{\beta_k^*[-\frac{1}{3\beta^2(d, y_T)}\nabla y_\rho \cdot \nabla\xi_\rho]\}_k, \{v_{d,k}\}_k\rangle_{l^{2*}, l^2} \\
&\quad\quad\quad = \langle\sum_k \beta_k^*[-\frac{1}{3\beta^2(d, y_T)}\nabla y_\rho \cdot \nabla\xi_\rho] e_k, v_d\rangle_{\mathcal{U}^*, \mathcal{U}} \\
&\quad\quad\quad\quad = \langle\partial_1\beta(d, y_T)^*[-\frac{1}{3\beta^2(d, y_T)}\nabla y_\rho \cdot \nabla\xi_\rho], v_d\rangle_{\mathcal{U}^*, \mathcal{U}},
\end{aligned}$$

where

$$\beta_k^*[w] = \langle e_k(y_T), \int_{\cdot}^{t_*} \varphi'_{\beta,1}\left(\int_0^{\tau} d(y_T)(s)\, ds\right) w(\tau)\, d\tau\rangle_{L_\infty(Q),\, L_1(Q)},$$

for all $w \in L_1(Q)$ and $k \in \mathbb{N}$. This holds analogously for (4.12) with

$$\mu_k^*[w] = \langle e_k(y_T), \int_{\cdot}^{t_*} \varphi'_{\mu,1}\left(\int_0^{\tau} d(y_T)(s)\, ds\right) w(\tau)\, d\tau\rangle_{L_\infty(Q),\, L_1(Q)},$$

for all $w \in L_1(Q)$ and $k \in \mathbb{N}$. Altogether we obtain for (4.10)

$$\begin{aligned}
&\langle D_d\hat{E}(d)^*[\xi], v_d\,\rangle = \\
&\quad \langle\partial_1\hat{\beta}(d)^*[-\frac{1}{3\hat{\beta}^2(d)}\nabla y_\rho \cdot \nabla\xi_\rho] + \partial_1\hat{\mu}(d)^*[y_\rho\big(\xi_\rho - \xi_T\big)], v_d\rangle_{\mathcal{U}^*, \mathcal{U}},
\end{aligned}$$

for all $v_d \in \mathcal{U}$ and $\xi \in Z$ with $\partial_1\hat{\beta}(d)^*$ and $\partial_1\hat{\mu}(d)^*$ explicitly given by

$$\partial_1\hat{\beta}(d)^*[w] = \sum_k \beta_k^*[w] e_k \quad \text{and} \quad \partial_1\hat{\mu}(d)^*[w] = \sum_k \mu_k^*[w] e_k,$$

respectively for a given countable orthonormal basis $\{e_k\}_k \subset \mathcal{U}$. □

Remark 4.1. Note that the requirement for $\epsilon > 0$ to be sufficiently small was not necessary in the proof. It was only required to simplify the notations for computations.

5. Numerical Simulation and Optimization

In this section we present numerical results underlining the feasibility of our approach.

5.1. *Forward Simulation and Measurements Generation*

To produce measurements for the identification of the temperature dependent rate constant d, we consider an ex-vivo experiment,[2] in which a porcine liver is exposed to a 30mm×3mm (length×width) Nd:YAG laser fiber with water cooling kept at 298.15 K (25°C). The treatment is conducted with a constant power of 28 W over a period of 845 seconds (≈ 14 minutes). We assume that the porcine liver is homogeneous and has an initial temperature of $T_0 = 298.15$ K. This allows for a reduction of the problem (due to radial symmetry) into a 2-dimensional problem given by

$$-\nabla \cdot \left(\frac{1}{3\beta(d,T)} \nabla \rho \right) + \mu(d,T)\rho = 0, \qquad (5.1a)$$

$$c_p \partial_t T - \nabla \cdot (\kappa \nabla T) - \mu(d,T)\rho = 0, \qquad (5.1b)$$

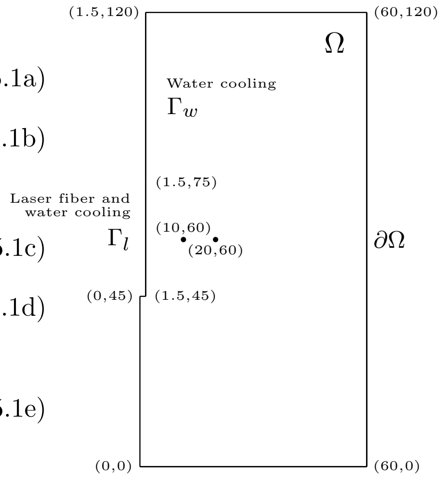
in Q, with boundary conditions

$$\frac{1}{3\beta(d,T)}\, \partial_n \rho + \frac{1}{2}(\rho - \rho_\partial) = 0, \qquad (5.1c)$$

$$\kappa\, \partial_n T + \alpha(T - T_\partial) = 0, \qquad (5.1d)$$

on Σ and initial condition

$$T(0,x) - T_0 = 0 \quad \text{for a.e. } x \in \Omega, \qquad (5.1e)$$



where thermal parameters c_p, κ are the product of density with specific heat capacity, and heat conductivity respectively, as given in Table 1. The functions ρ_∂, T_∂, α are defined as follows

$$\rho_\partial = \begin{cases} \frac{28}{\pi|\Gamma_l|} & \text{on } \Gamma_l \\ 0 & \text{otherwise,} \end{cases} \qquad T_\partial = \begin{cases} 298.15 & \text{on } \Gamma_w \cup \Gamma_l \\ 0 & \text{otherwise,} \end{cases} \qquad \alpha = \begin{cases} 1 \cdot 10^6 & \text{on } \Gamma_w \cup \Gamma_l \\ 0 & \text{otherwise.} \end{cases}$$

	Native	Coagulated
$\mu_a[\text{mm}^{-1}]$	$1.950 \cdot 10^{-2}$	$1.300 \cdot 10^{-2}$
$\mu_s[\text{mm}^{-1}]$	4.350	30.590
g	$9.310 \cdot 10^{-1}$	$9.165 \cdot 10^{-1}$
$c_p[\text{Jmm}^{-3}\text{K}^{-1}]$	$1.040 \cdot 10^{-6} \times 3.640 \cdot 10^3$	$1.040 \cdot 10^{-6} \times 3.640 \cdot 10^3$
$\kappa[\text{Wmm}^{-1}\text{K}^{-1}]$	$5.180 \cdot 10^{-4}$	$5.180 \cdot 10^{-4}$
$A\ [\text{s}^{-1}]$	$9.510 \cdot 10^{48}$	$9.510 \cdot 10^{48}$
$E_a[\text{Jmol}^{-1}]$	$3.304 \cdot 10^5$	$3.304 \cdot 10^5$

Table 1. Optical and Thermal Parameters for Measurements generation

Further, we define the temperature dependent optical parameters β and μ as follows

$$\beta(d,T) = \beta_c - (\beta_c - \beta_n)\exp\left(-\int_0^{\cdot} d(T)(\tau)\,d\tau\right),$$
$$\mu(d,T) = \mu_{a,c} - (\mu_{a,c} - \mu_{a,n})\exp\left(-\int_0^{\cdot} d(T)(\tau)\,d\tau\right),$$

with

$$\beta_n = \mu_{a,n} + (1-g)\mu_{s,n} \quad \text{and} \quad \beta_c = \mu_{a,c} + (1-g)\mu_{s,c},$$

where $\mu_{a,n}$, $\mu_{a,c}$, $\mu_{s,n}$, $\mu_{s,c}$, g are constants denoting the natural absorption coefficient, coagulated absorption coefficient, natural scattering coefficient, coagulated scattering coefficient and the anisotropy factor respectively, as given in Table 1. For simplicity, we consider an Ansatz for the temperature dependent rate constant d given by the Arrhenius equation

$$d(y_T) = Ae^{-E_a/Ry_T}, \tag{5.2}$$

where A is the frequency factor and E_a the activation energy, which are as given in Table 1, and $R[\text{Jmol}^{-1}\text{K}^{-1}]$ the universal gas constant.

The solution of (5.1) was done semi-implicitly with 2019 triangular linear elements and a time step of 13 seconds. Measurements for identification were taken at points $x_1 = (10,60)$[mm] and $x_2 = (20,60)$[mm].

5.2. *Optimization Algorithm*

Note that due to (5.2), the identification problem is reduced to identifying an optimal pair $u = (A, E_a) \in \mathcal{U} \subset \mathbb{R}^2$. Now consider the reduced cost functional

$$\hat{J}_\epsilon(A,E_a) = \frac{1}{4}\sum_{i=1}^{2} \|\big(\delta^\epsilon_{x_i} * (y_T \circ d)(A,E_a)(\cdot)\big) - T_{m,i}\|^4_{L_4(I)} + \frac{\lambda}{2}\|u-u_0\|^2_{\mathcal{U}}, \tag{5.3}$$

for $\epsilon < \min\{\frac{1}{2}\text{diam}(T_h)\,|\,T_h \in \mathcal{T}_h\}$, where $\mathcal{T}_h$ denotes the set of triangular elements.

The optimization was performed using a modified BFGS method for nonconvex minimization[5] with Armijo rule for the line search and stops as soon as the gradient norm of the reduced cost functional is less than 10^{-3}. The regularization parameter λ was set to 10^{-5}. An outline of the optimization algorithm is given as follows:

0. Choose initial point $u_0 = (A_0, E_{a,0})$, positive definite matrix B_0, and numerical constants $\sigma \in (0,1)$ and $\varrho \in (0,1)$. Set $k = 0$.
1. Solve for $\bar{u}_k$ the system
$$B_k\bar{u}_k + \nabla\hat{J}_\epsilon(u_k) = 0 \quad \text{in } \mathcal{U}. \tag{5.4}$$
2. Find the smallest non-negative integer j, say j_k, satisfying
$$\hat{J}_\epsilon(u_k + \varrho^j\bar{u}_k) \le \hat{J}_\epsilon(u_k) + \sigma\varrho^j\nabla\hat{J}_\epsilon(u_k)\cdot\bar{u}_k \tag{5.5}$$
and let $s_k = \varrho^{j_k}$.

3. Set $u_{k+1} = u_k + s_k \bar{u}_k$ for the next iterate.
4. Update B_{k+1} using the formula

$$B_{k+1} = B_k - \frac{B_k p_k p_k^T B_k}{p_k^T B_k p_k} + \frac{q_k q_k^T}{q_k^T p_k}, \tag{5.6}$$

where $p_k = u_{k+1} - u_k = s_k \bar{u}_k$ and

$$q_k = r_k + \tau_k \|\nabla \hat{J}_\epsilon(d_k)\| p_k,$$

with $r_k = \nabla \hat{J}_\epsilon(u_{k+1}) - \nabla \hat{J}_\epsilon(u_k)$ and $\tau_k = 1 + \max\left\{-\frac{r_k^T p_k}{\|p_k\|^2}, 0\right\}$.
5. $k = k + 1$ and go to 1. while $\|\bar{u}_k\| > \delta$ for some $\delta > 0$.

Remark 5.1. Observe that an evaluation of the gradient $\nabla \hat{J}_\epsilon(u_k)$ in (5.4) and (5.6) involves the following steps

1-1. Solve for y_k the forward system

$$E(d(u_k), y_k) = 0 \quad \text{in } Z^*.$$

1-2. Solve for ξ_k the adjoint system

$$D_y E(d(u_k), y_k)^* \xi_k = -D_y J(u_k, y_k) \quad \text{in } {X_{2,2}}^*.$$

1-3. Compute $\hat{J}_\epsilon'(u_k) \in \mathcal{U}^*$ as in Theorem 4.4 and identify $\nabla \hat{J}_\epsilon(u_k) \in \mathcal{U}$ with $\hat{J}_\epsilon'(u_k)$ via Riesz identification,

while an evaluation of the reduced cost functional $\hat{J}_\epsilon(u_k)$ in (5.5) involves only the steps

2-1. Solve for y_k the forward system

$$E(d(u_k), y_k) = 0 \quad \text{in } Z^*.$$

2-2. Compute $\hat{J}_\epsilon(u_k)$ via (5.3).

Thus, by choosing appropriate numerical constants $\sigma \in (0, 1)$ and $\varrho \in (0, 1)$, it is possible to obtain sufficiently low complexity for the optimization problem.

The algorithm was initialized with $d_0 = (1.0 \cdot 10^{50}, 3.5 \cdot 10^5) \in \mathcal{U}$ where $\mathcal{U} = \mathbb{R}^2$. The initial state corresponding to d_0 can be seen in Fig. 1.

The optimization was done for both exact measurements and noisy measurements, as seen in Table 2. At first glance, one might think that the variations to the optimal solutions are high. These variations are, however, relatively low when scaled to the given problem. Furthermore, the optimized values are physical, i.e. within the predicted intervals $[1 \cdot 10^{40}, 1 \cdot 10^{100}]$ for A and $[3 \cdot 10^5, 6 \cdot 10^5]$ for E_a.

Figure 2 and 3 show results of the optimization procedure under noiseless and noisy measurement data respectively. Note that the results of their respective gradient norm and cost functional show fast convergence of the modified BFGS method in obtaining optimal parameters $(A_*, E_{a,*}) \in \mathcal{U}$ for both, with and without noise. One also notices the lack of convergence to zero in the cost functional in the presence of noise, which is as expected.

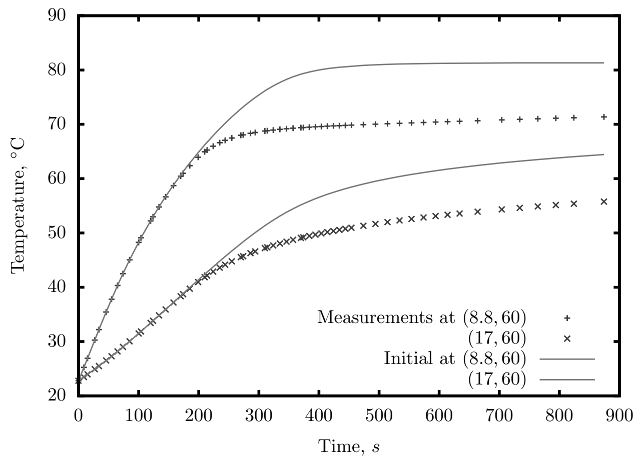


Fig. 1. Initial state

Measurement Noise	Optimized Value, d_{opt}	Optimal State
0%	$(5.554 \cdot 10^{51}, 3.474 \cdot 10^5)$	Fig. 2
5%	$(1.375 \cdot 10^{64}, 4.283 \cdot 10^5)$	Fig. 3

Table 2. Optimal Values

Acknowledgment

The authors acknowledge support from DFG via contract SI 1289/1-1 and PI 408/4 in the SPP 1253, and from the DAAD.

References

1. A. D. Klose and E. W. Larsen. Light Transport in Biological Tissue Based on the Simplified Spherical Harmonics Equations. *Journal of Computational Physics 220 (2006), 441-470*.
2. A. Roggan. Dosimetrie thermischer Laseranwendungen in der Medizin: Untersuchung der speziellen Gewebeeigenschaften un physikalisch-mathematische Modellbildung. Fortschritte in der Lasermedizin, Vol. 16. *Hüthig Jehle Rehm, Landsberg/Lech (1997)*.
3. C. Meyer and P. Philip and F. Tröltzsch. Optimal Control of a Semilinear PDE with Nonlocal Radiation Interface Conditions. *SIAM J. Control Optim., 45 (2006), 688-721*.
4. D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*, 1 ed. Springer-Verlag, Berlin, 1983.
5. D. H. Li and M. Fukushima. A modified BFGS method and its global convergence in nonconvex minimization. *Journal of Computational and Applied Mathematics 129 (2001) 15-35*.

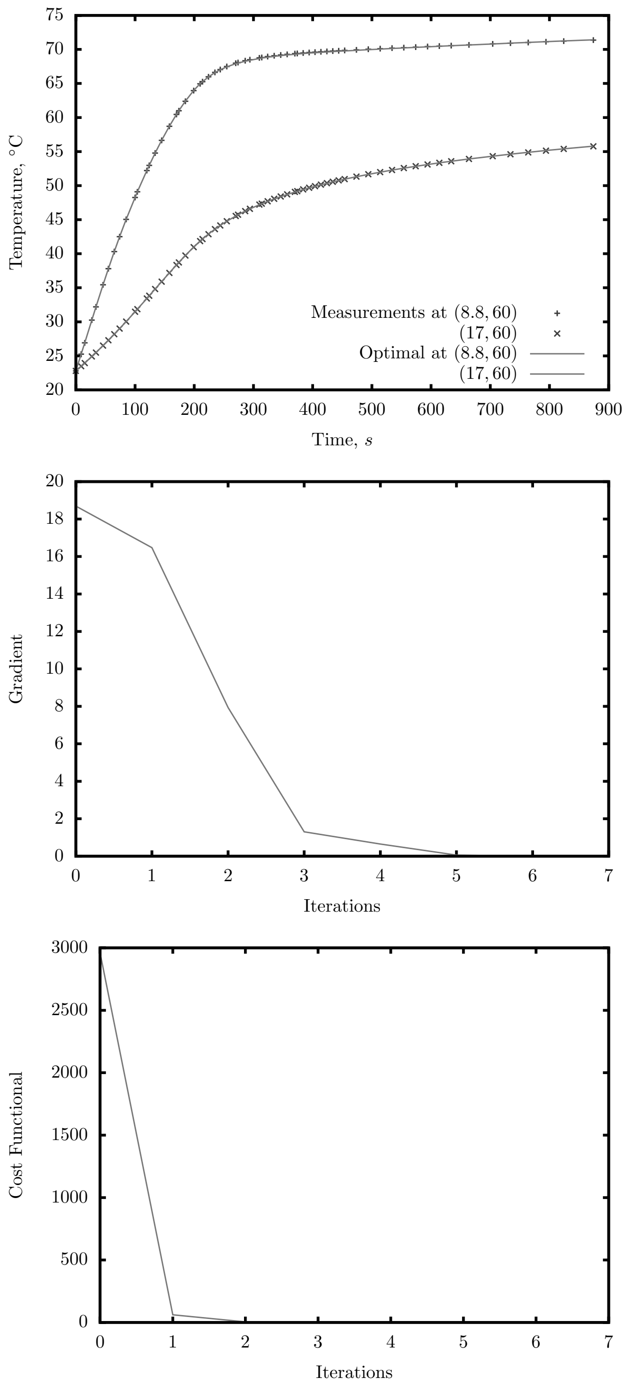


Fig. 2. Optimal temperature state, Gradient norm, Cost functional.

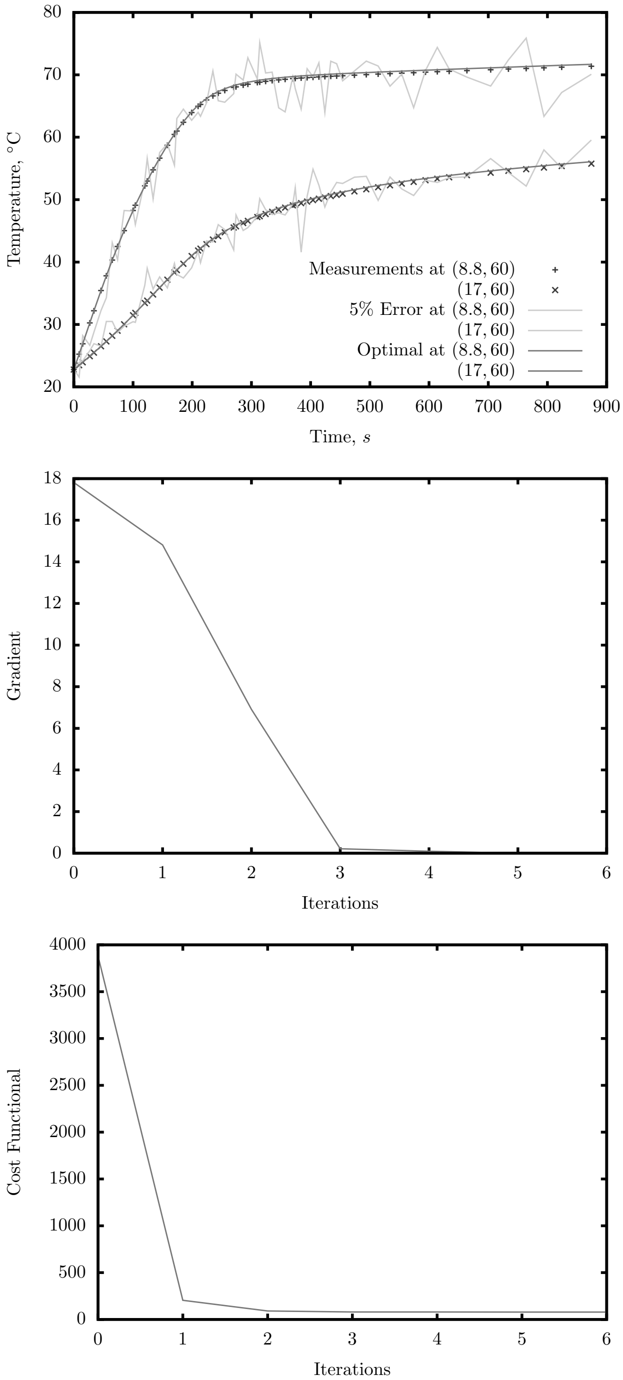


Fig. 3. Optimal temperature state with 5% noise, Gradient norm, Cost functional.

6. Editors: G. Müller and A. Roggan. *Laser-induced Interstitial Thermotherapy.* SPIE Optical Engineering Press, 1995.
7. F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen: Theorie, Verfahren und Anwendungen.* Vieweg, 2005.
8. H. Goldberg and W. Kampowsky and F. Tröltzsch. On Nemytskij Operators in L_p-Spaces of Abstract Functions. *Math. Nachr. 155 (1992), 127-140.*
9. J. A. Griepentrog. Maximal regularity for nonsmooth parabolic problems in Sobolev-Morrey spaces. *Preprint.*
10. J. A. Griepentrog. Sobolev-Morrey spaces associated with evolution equations. *Preprint.*
11. J. A. Oguntuase. On Integral Inequalities of Gronwall-Bellman-Bihari Type in Several Variables. *J. Ineq. Pure and Appl. Math. 1(2) Art. 20 (2000).*
12. J. Appell and P. P. Zabrejko. *Nonlinear Superposition Operators.* Cambridge Tracts in Mathematics, Vol. 95, Cambridge University Press, Cambridge, 1990.
13. M. Renardy and R. C. Rogers. *An Introduction to Partial Differential Equations.* Springer-Verlag New York, Inc., 2004.
14. O. A. Ladyženskaya and V. A. Solonnikov and N. N. Ural'ceva. *Linear and Quasi-linear Equations of Parabolic Type.* American Mathematical Society, Translations of Mathematical Monographs Vol. 23, 1968.
15. R. E. Showalter. *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations.* Mathematical surveys and Monographs, Volume 49, 1997.
16. R. Pinnau. Analysis of Optimal Boundary Control for Radiative Heat Transfer Modeled by the SP_1-System. Commun. Math. Sci. *Vol. 5, No. 4, pp. 951-969.*
17. T. Gallout and A. Monier. On the regularity of Solutions to Elliptic Equations. *Rendiconti di Matematica, Serie VII, Volume 19, Roma (1999), 471-488.*
18. T. Runst and W. Sickel. *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations.* De Gruyter Series in Nonlinear Analysis and Applications 3, de Gruyter, 1996.