Control and Estimation of the Boundary Heat Transfer Function in Stefan Problems

V. Barbu*
Institute of Mathematics
University of Iaşi,
RO-6600 Iaşi,
Romania

K. Kunisch†
Fachbereich Mathematik
Technische Universität Berlin
D-10623 Berlin,
Germany

W. Ring‡
Institut für Mathematik
Technische Universität Graz
A-8010 Graz,
Austria

April 1995

*Research partially supported by NSF Grant DMS-91-11794

†Research partially supported by Christian Doppler Laboratory for Parameter Identification and Inverse Problems

‡Research partially supported by Spezialforschungsbereich Optimierung und Kontrolle

Abstract

An approximation procedure for the identification of a nonlinear boundary heat transfer function in a one phase Stefan problem is presented. Alternatively the problem can be viewed as constructing a feedback control law for the control of the solidification surface in the Stefan problem. The analysis is based on Hilbert space methods and convex analysis techniques. Numerical results combining two regularization methods are presented.

Keywords: One phase Stefan problem, free moving boundary, inverse problem, feedback control, convex functions, regularization techniques.
AMS classification: 49A22, 32R30.

1 Introduction

Let Ω be a bounded domain in $\mathbb{R}^n$, $n \geq 1$ with a $C^{1,1}$ boundary , and let $\{\Omega_t : t \in [0,T]\}$ be a family of monotonically increasing (strict) subdomains of Ω with the property that , is contained in the boundary of Ω_t for all $t \in [0,T]$. To express $\{\Omega_t : t \in [0,T]\}$ analytically, the existence of a function $\sigma : \overline{\Omega}_T \to [0,T]$ is assumed with the properties that $\sigma \in C^2(\overline{\Omega}_T \backslash \Omega_0)$, $|\nabla \sigma(x)| \neq 0$ for all $x \in \overline{\Omega}_T \backslash \Omega_0$, $\sigma(x) = 0$ on $\overline{\Omega}_0$ and such that

$$\Omega_t = \{x \in \Omega_T : \sigma(x) < t\}, \Omega_t \subset \Omega_{\tilde{t}} \text{ for } 0 \leq t \leq \tilde{t} \leq T,$$

see Figure 1 below. We set

$$\begin{aligned} Q &= \{(x,t) \in \Omega_T \times (0,T) : \sigma(x) < t < T\}, \\ \Sigma &= \;, \times (0,T), \\ \Sigma_0 &= \left\{(x,t) : (\overline{\Omega}_T \setminus \Omega_0) \times (0,T) : t = \sigma(x)\right\}. \end{aligned}$$

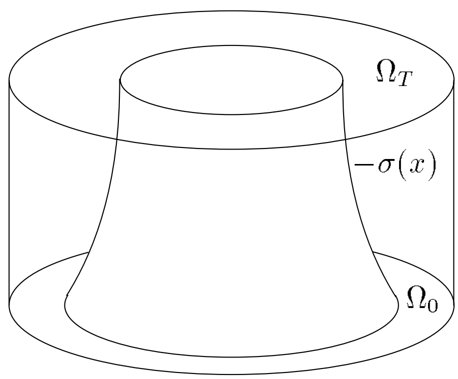

Figure 1: Space-time cylinder with moving boundary $\sigma(x)$

In this paper we consider the one phase Stefan problem

$$\left.\begin{aligned} y_t - \Delta y &= 0 \quad \text{in} \quad Q \\ y &= 0 \quad \text{in} \quad \Omega \times (0,T) \setminus Q \\ y &= 0, \nabla y \cdot \nabla \sigma = \rho \text{ in } \Sigma_0 \\ \tfrac{\partial y}{\partial \nu} + \beta(y) &= 0 \quad \text{in} \quad \Sigma \\ y(\cdot, 0) &= y_0 \text{ in } \quad \Omega_0 \\ y &< 0 \quad \text{in} \quad Q, \end{aligned}\right\} \tag{1.1}$$

where $y_0 \in H^1(\Omega_0)$, $y_0 < 0$ in Ω_0, $\beta : \mathbb{R} \to \mathbb{R}$ is a nonlinear mapping, $\rho > 0$ is the latent heat and $\frac{\partial}{\partial \nu} = \nabla \cdot \nu$ stands for the outer normal to , .

This problem describes the solidification of a volume of water occupying the domain Ω in the time interval $[0,T]$, having $,_t = \{x : t = \sigma(x)\}$ as the interface between solid and liquid regions. The state variable y stands for the temperature distribution, y_0 is the initial distribution and $\beta(y)$ describes the heat flux along the exterior boundary Σ. At time t, Ω_t is the solid (frozen) region, $\Omega \setminus \overline{\Omega}_t$ is the liquid (water) region and $\Sigma_0 = \bigcup\limits_{0 \leq t \leq T} ,_t$ describes the evolution of the free moving boundary $,_t$. The boundary condition

$$\tfrac{\partial y}{\partial \nu} + \beta(y) = 0 \text{ in } \Sigma \tag{1.2}$$

describes a possibly nonlinear boundary heat transfer law. In the linear case with $\beta(y) = \beta y$, $\beta > 0$, one refers to (1.2) as radiation condition. If β, together with ρ and y_0 are given, then the direct Stefan problem consists in determining the temperature distribution y together with the free boundary Σ_0 which is characterized by σ from (1.1).

Here we shall consider the following inverse problem: Given the function σ determine β from a class of admissible functions, such that $\Sigma_0 = \{(x,t) : t = \sigma(x) = \sigma(\beta; x)\}$ is the free boundary of the resulting one phase Stefan problem. This problem can be thought of in two different ways. First it constitutes the inverse problem of identifying the unknown boundary heat transfer coefficient from overspecified boundary data on Σ_0. Secondly it describes a feedback control problem for the one phase Stefan problem with boundary control:

$$\tfrac{\partial y}{\partial \nu} = v, \text{ in } \Sigma \tag{1.3}$$

and with the control u in the feedback form:

$$v = -\beta(y). \tag{1.4}$$

The objective is to steer the free boundary $\sigma = \sigma(\beta)$ to some a-priori desired solidification surface. We refer to [HN, HS] for results and references related to the control of Stefan problems.

The class of admissible heat transfer functions (or feedback control laws) is chosen to be

$$\begin{aligned} \mathcal{A} = \{\beta = \partial j :\ & \text{with } j : \mathbb{R} \to \mathbb{R} \text{ convex, continuous, } j(0) = 0,\ 0 \in \beta(0) \\ & \text{and } \alpha_0 + \omega_0 r^2 \leq j(r) \leq \alpha_1 + \omega_1 r^2 \text{ for all } r \in \mathbb{R}\}, \end{aligned}$$

where $0 < \omega_0 < \omega_1$, and $\alpha_0 < \alpha_1$ are constants and ∂j, mapping $\mathbb{R}$ into the set of all subsets of $\mathbb{R}$, is the subdifferential of j. Thus, β is a monotone graph and the boundary condition on Σ has to be replaced by

$$\tfrac{\partial y}{\partial \nu} \in -\beta(y). \tag{1.5}$$

The above inverse problem will be formulated as least squares problem:

$$\left.\begin{array}{c} \text{minimize} \displaystyle\int_{\Sigma_0} (\nabla y \cdot \nabla \sigma - \rho)^2 dx\, dt \\ \text{subject to } \beta \in \mathcal{A} \text{ and } y \in H^1(Q) \text{ satisfying} \end{array}\right\} \tag{1.6}$$

$$\left.\begin{array}{rcll} y_t - \Delta y & = & 0 & \text{in} \quad Q \\ y & = & 0 & \text{in} \quad \Sigma_0 \\ \frac{\partial y}{\partial \nu} + \beta(y) & \ni & 0 & \text{in} \quad \Sigma \\ y(\cdot, 0) & = & y_0 & \text{in} \quad \Omega. \end{array}\right\} \tag{1.7}$$

We note that the solution y of (1.7) is not effected by replacing j by j+constant. This motivates the constraint $j(0) = 0$ in the definition of $\mathcal{A}$. One of the main goals of this paper is the analysis of (1.6) by convex analysis techniques. In particular, the nonlinear boundary condition will be simplified by a substitution similar to operator splitting or the mixed finite element method. Numerically the solution of the inverse problem of identifying the heat transfer coefficient on one part of the boundary from measurements on other parts is related to the sideways heat equation [C] which is a notoriously illposed problem. The second goal of this paper is therefore the description of a numerical algorithm for the identification of the boundary heat transfer coefficient (or the feedback control law) β in (1.1), which proved to be successful on a series of test examples.

The plan of the paper is the following one. In Section 2 we shall study well-posedness of the closed loop system (1.7) in a Hilbert space framework. Section 3 is devoted to proving existence and convergence of suboptimal solutions to (1.6). An approximation process of similar kind was already used in [BK1, BK2] for the identification of nonlinear elliptic and parabolic boundary value problems. Roughly speaking the nonlinear boundary condition in (1.7) is decoupled via (1.3), (1.4), resulting in a parabolic optimal control problem on the non-cylindrical domain Q in the control variables v and j. In Section 4 a maximum principle type result for this problem is given. Numerical algorithms and tests are presented in Section 5.

We shall use standard notation for the spaces of square integrable functions and Sobolev spaces on Ω_t, Q and Σ. Given a lower semicontinuous function φ from a Hilbert space X to $\overline{\mathbb{R}} = (-\infty, \infty]$ we shall denote by $\partial\varphi$ its subdifferential, i.e.:

$$\partial\varphi(x) = \{\omega \in X : \varphi(x) \leq \varphi(u) + (\omega, x - u) \text{ for all } u \in X\},$$

and by

$$\begin{array}{l} \varphi^* : X \to \overline{\mathbb{R}} \text{ its conjugate function defined by,} \\ \varphi^*(p) = \sup \{(p, u) - \varphi(u) : u \in X\}, \text{ for } p \in X, \end{array}$$

where $(\cdot, \cdot)$ denotes the scalar product on X. We refer to [B1] for further results from convex analysis which will be used in this paper.

2 The nonlinear closed loop system

We shall study here the nonlinear system (1.7) which we repeat for convenience:

$$\left.\begin{array}{rcll} y_t - \Delta y & = & 0 & \text{in } Q \\ y & = & 0 & \text{in } \Sigma_0 \\ \frac{\partial y}{\partial \nu} + \beta(y) & \ni & 0 & \text{in } \Sigma \\ y(x,0) & = & y_0(x) & \text{in } \Omega_0. \end{array}\right\} \tag{2.1}$$

Here β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ such that

$$\text{dom } (\beta) = \mathbb{R} \text{ and } 0 \in \beta(0).$$

This implies that $\beta = \partial j$, where $j : \mathbb{R} \to \mathbb{R}$ is a convex and continuous function such that

$$j(0) = \inf \{ j(r) : r \in \mathbb{R} \},$$

see [B1], p. 71. We shall call y a (variational) solution of the boundary value problem (2.1), if it is an element of

$$V = \{ y \in H^1(Q) : y = 0 \ \text{ in } \Sigma_0 \},$$

such that

$$\begin{aligned} &\int_Q (-y(y-z)_t + \nabla y \cdot \nabla(y-z)) dx\, dt + \int_\Sigma (j(y) - j(z)) d\sigma_x dt \\ &+ \int_{\Omega_T} y(x,T)(y(x,T) - z(x,T)) dx - \int_{\Omega_0} y_0(x)(y(x,0) - z(x,0)) dx \le 0 \end{aligned} \tag{2.2}$$

for all $z \in V$. It is simple to check that every classical solution to (2.1) satisfies (2.2). Moreover we have the following result:

Proposition 2.1 *Assume that $\Delta\sigma \ge 0$ in $\Omega_T \setminus \Omega_0$, $\frac{\partial \sigma}{\partial n} \ge 0$ on $\partial\Omega_T \setminus$, , $y_0 \in H^1(\Omega_0)$, $y_0 = 0$ on $\partial\Omega_0 \setminus$, , $y_0 < 0$ a.e. in Ω_0 and that $j(y_0) \in L^1($, $)$. Then problem (2.1) has a unique solution $y \in V$ such that $\Delta y \in L^2(Q)$ and $\nabla y \in L^2(\Sigma_0)$. Moreover, $y \le 0$ a.e. in Q and the following estimate holds:*

$$\begin{aligned} &\int_Q (y_t^2 + |\nabla y|^2 + |\Delta y|^2) dx\, dt + \int_{\Omega_T \setminus \Omega_0} |\nabla y(x, \sigma(x))|^2 dx \\ &\quad + \int_{\Omega_T} (\tfrac{1}{2} y_t^2(x,T) + |\nabla y(x,T)|^2) dx + 2 \int_\Gamma j(y(x,T)) d\sigma_x + \int_\Sigma j(y) d\sigma_x dt \\ &\le \int_{\Omega_T} (\tfrac{1}{2} y_0^2(x) + |\nabla y_0(x)|^2) dx + 2 \int_\Gamma j(y_0(x)) d\sigma_x. \end{aligned} \tag{2.3}$$

Proof. First let us note that $x \to j(y(x,T)), x \in$, , and $(x,t) \to j(y(x,t))$, $(x,t) \in \Sigma$ are Lebesgue measurable and hence the corresponding integrals in (2.3) make sense as extended real values in $\overline{\mathbb{R}}$, [B1], p. 72.
We shall approximate (2.1) by the family of elliptic boundary value problems

$$\left.\begin{array}{rcll} \varepsilon y_{tt} + \Delta y - y_t & = & 0 & \text{in } Q \\ y & = & 0 & \text{in } \Sigma_0 \\ \frac{\partial y}{\partial \nu} + \beta(y) & \ni & 0 & \text{in } \Sigma \\ y_t(x,T) & = & 0 & \text{in } \Omega_T \\ y_t(x,0) & = & \frac{1}{\varepsilon}\left(y(x,0) - y_0(x)\right) & \text{in } \Omega_0, \end{array}\right\} \tag{2.4}$$

where $\varepsilon > 0$. By the general theory for elliptic variational inequalities [Br1, Li] it is known that (2.4) has a unique variational solution $y^\varepsilon \in V$ satisfying

$$\begin{aligned} &\int_Q [\varepsilon y_t^\varepsilon (y^\varepsilon - z)_t + \nabla y^\varepsilon \cdot \nabla (y^\varepsilon - z) + y_t^\varepsilon (y^\varepsilon - z)]\, dx\, dt \\ &+ \int_{\Omega_0} (y^\varepsilon(x,0) - y_0)(y^\varepsilon(x,0) - z(x,0))dx + \int_\Sigma (j(y^\varepsilon) - j(z)) d\sigma_x dt \le 0 \end{aligned} \tag{2.5}$$

for all $z \in V$. In fact, we may express (2.5) as

$$Ay^\varepsilon + \partial\psi(y^\varepsilon) \ni 0, \tag{2.6}$$

where $A : V \to V'$ is the continuous affine, monotone operator defined by

$$(Ay, z) = \int_Q (\varepsilon y_t z_t + \nabla y \cdot \nabla z + y_t z)\, dx\, dt + \int_{\Omega_0} (y(x,0) - y_0(x)) z(x,0) dx$$

for all $y, z \in V$, and $\psi : V \to \overline{\mathbb{R}}$ is the convex, lower semicontinuous function

$$\psi(y) = \int_\Gamma j(y) d\sigma_x \text{ for all } y \in V.$$

Since A is positive definite and ψ is a lower semicontinuous, proper convex function, we infer that (2.6) has a unique solution $y^\varepsilon \in V$, see e.g. [B1], Section 3.

If one takes $z = y^\varepsilon - (y^\varepsilon)^+$ in (2.5) one obtains

$$\begin{aligned} &\varepsilon |((y^\varepsilon)^+)_t|^2_{L^2(Q)} + |\nabla (y^\varepsilon)^+|^2_{L^2(Q)} + \tfrac{1}{2}|(y^\varepsilon)^+(\cdot,T)|^2_{L^2(\Omega_T)} + \tfrac{1}{2}|(y^\varepsilon)^+(\cdot,0)|^2_{L^2(\Omega_0)} \\ &\quad - \int_{\Omega_0} y_0\, (y^\varepsilon)^+(x,0)\, dx + \int_\Sigma (j(y^\varepsilon) - j(y^\varepsilon - (y^\varepsilon)^+)) d\sigma_x dt \le 0. \end{aligned}$$

Since $y_0 < 0$ a.e. on Ω_0 and $0 \in \partial j(0)$ it follows that $y^\varepsilon \le 0$ a.e. on Q for every $\varepsilon > 0$. Next, for $z = 0$ in (2.5) we obtain

$$
\begin{aligned}
&\varepsilon|y_t^\varepsilon|^2_{L^2(Q)} + |\nabla y^\varepsilon|^2_{L^2(Q)} + \tfrac{1}{2}|y^\varepsilon(\cdot,T)|^2_{L^2(\Omega_T)} + \int_\Sigma j(y^\varepsilon)d\sigma_x dt \\
&\qquad \le \tfrac{1}{2}|y_0|^2_{L^2(\Omega_0)} + \int_\Sigma j(0)d\sigma_x dt = \tfrac{1}{2}|y_0|^2_{L^2(\Omega_0)}.
\end{aligned} \tag{2.7}
$$

To obtain the following a-priori estimates we assume that β is continuously differentiable. This requirement will be eliminated at the end of the proof. In the sequel, without loss of generality, we may view y^ϵ as a smooth solution of (2.4). Indeed by interior and boundary regularity for nonlinear elliptic boundary value problems (see e.g. [Br1]) we know that y is $C^2(Q \setminus Q_\delta)$, where Q_δ represents a neighborhood of the corners of Q of order δ. To make the following calculations rigorous one has to replace Q by $Q \setminus Q_\delta$ and let δ tend to zero. Taking the inner product in $L^2(Q)$ of the first equation in (2.4) with y_t^ε we find

$$
\int_Q (y_t^\varepsilon)^2\, dx\, dt = \int_Q \Delta y^\varepsilon y_t^\varepsilon\, dx\, dt - \tfrac{\varepsilon}{2}\int_{\Omega_0} (y_t^\varepsilon)^2(x,0)dx - \tfrac{\varepsilon}{2}\int_{\Omega_T\setminus\Omega_0} y_t^\varepsilon(x,\sigma(x))^2 dx.
$$

This yields

$$
|y_t^\varepsilon|^2_{L^2(Q)} + \varepsilon|y_t^\varepsilon(\cdot,0)|^2_{L^2(\Omega_0)} \le |\Delta y^\varepsilon|^2_{L^2(Q)}. \tag{2.8}
$$

Finally, we shall multiply (2.4) by Δy^ε and integrate on Q. For that purpose we prepare some useful identities. Note that we defined $\sigma(x) = 0$ on Ω_0. We find

$$
\begin{aligned}
\int_{\Omega_T} \operatorname{div} \int_{\sigma(s)}^T y_t^\varepsilon \nabla y^\varepsilon dt dx \;&=\; -\int_{\Omega_T\setminus\Omega_0} y_t^\varepsilon(x,\sigma(x))\nabla y^\varepsilon(x,\sigma(x))\cdot\nabla\sigma(x)dx \\
&\quad + \int_{\Omega_T}\int_{\sigma(x)}^T y_t^\varepsilon \Delta y^\varepsilon dt dx + \tfrac{1}{2}\int_{\Omega_T}\int_{\sigma(x)}^T \tfrac{\partial}{\partial t}|\nabla y^\varepsilon|^2 dt dx.
\end{aligned} \tag{2.9}
$$

Moreover we have

$$
\nabla y^\varepsilon(x,\sigma(x)) + y_t^\varepsilon \nabla\sigma(x) = 0 \quad \text{a.e. on } \Omega_T \setminus \Omega_0. \tag{2.10}
$$

Combining (2.9), (2.10) and using the divergence theorem we obtain

$$
\begin{aligned}
\int_Q y_t^\varepsilon \Delta y^\varepsilon \, dx \, dt = & \int_{\Omega_T} \operatorname{div} \int_{\sigma(x)}^{T} y_t^\varepsilon \nabla y^\varepsilon dt dx \\
& + \int_{\Omega_T \setminus \Omega_0} y_t^\varepsilon(x,\sigma(x)) \nabla y^\varepsilon(x,\sigma(x)) \nabla \sigma(x) dx \\
& - \tfrac{1}{2} \int_{\Omega_T} |\nabla y^\varepsilon(x,T)|^2 dx + \tfrac{1}{2} \int_{\Omega_T \setminus \Omega_0} |\nabla y^\varepsilon(x,\sigma(x))|^2 dx \\
& + \tfrac{1}{2} \int_{\Omega_0} |\nabla y^\varepsilon(x,0)|^2 dx \\
= & \int_\Gamma \int_0^T y_t^\varepsilon \nabla y^\varepsilon \cdot \nu \, dt d\sigma_x - \int_{\Omega_T \setminus \Omega_0} |\nabla y^\varepsilon(x,\sigma(x))|^2 dx \\
& - \tfrac{1}{2} \int_{\Omega_T} |\nabla y^\varepsilon(x,T)|^2 dx + \tfrac{1}{2} \int_{\Omega_T \setminus \Omega_0} |\nabla y^\varepsilon(x,\sigma(x))|^2 dx \\
& + \tfrac{1}{2} \int_{\Omega_0} |\nabla y^\varepsilon(x,0)|^2 dx \\
= & \int_\Gamma [j(y^\varepsilon(x,0) - j(y^\varepsilon(x,T))] d\sigma_x - \tfrac{1}{2} \int_{\Omega_T \setminus \Omega_0} |\nabla y^\varepsilon(x,\sigma(x))|^2 dx \\
& - \tfrac{1}{2} \int_{\Omega_T} |\nabla y^\varepsilon(x,T)|^2 dx + \tfrac{1}{2} \int_{\Omega_0} |\nabla y^\varepsilon(x,0)|^2 dx,
\end{aligned}
\tag{2.11}
$$

where in the last step we used the boundary condition on Σ.
Similarily we find for a.e. $x \in \Omega_T$

$$
\begin{aligned}
\int_{\sigma(x)}^{T} y_{tt}^\varepsilon \Delta y^\varepsilon dt \quad = \quad & -y_t^\varepsilon(x,\sigma(x)) \Delta y^\varepsilon(x,\sigma(x)) - \operatorname{div} \int_{\sigma(x)}^{T} y_t^\varepsilon \nabla y_t^\varepsilon dt \\
& + \int_{\sigma(x)}^{T} |\nabla y_t^\varepsilon|^2 dt - y_t^\varepsilon(x,\sigma(x)) \nabla y_t^\varepsilon(x,\sigma(x)) \cdot \nabla \sigma(x),
\end{aligned}
$$

and upon integrating this equality on Ω_T

$$
\begin{aligned}
\int_Q y_{tt}^\varepsilon \Delta y^\varepsilon dx \, dt \quad = \quad & - \int_{\Omega_T \setminus \Omega_0} y_t^\varepsilon(x,\sigma(x)) \operatorname{div}_x(\nabla y^\varepsilon(x,\sigma(x)) dx \\
& - \int_{\Omega_0} y_t^\varepsilon(x,0) \Delta y^\varepsilon(x,0) dx + \int_Q |\nabla y_t^\varepsilon|^2 \, dx \, dt - \int_\Sigma y_t^\varepsilon \tfrac{\partial}{\partial t} \tfrac{\partial y^\varepsilon}{\partial \nu} d\sigma_x \, dt.
\end{aligned}
$$

Due to the boundary condition on Σ and the initial condition for y_t^ε we obtain

$$\begin{aligned}\int_Q y_{tt}^\varepsilon \Delta y^\varepsilon\,dx\,dt &= -\int_{\Omega_T\setminus\Omega_0} y_t^\varepsilon(x,\sigma(x))\,\mathrm{div}_x(\nabla y^\varepsilon(x,\sigma(x))dx \\ &\quad +\tfrac{1}{\varepsilon}\int_{\Omega_0}(y_0(x)-y^\varepsilon(x,0))\Delta y^\varepsilon(x,0)dx \\ &\quad +\int_Q |\nabla y_t^\varepsilon|^2\,dx\,dt+\int_\Sigma y_t^\varepsilon \tfrac{\partial}{\partial t}\beta(y^\varepsilon)d\sigma_x\,dt,\end{aligned} \tag{2.12}$$

where we use the temporary assumption that β is continuously differentiable. Combining (2.4), (2.11) and (2.12) implies that

$$\begin{aligned}&\int_Q |\Delta y^\varepsilon|^2\,dx\,dt = \int_Q y_t^\varepsilon \Delta y^\varepsilon\,dx\,dt - \varepsilon\int_Q y_{tt}^\varepsilon\Delta y^\varepsilon\,dx\,dt \\ &\le -\tfrac12\int_{\Omega_T}|\nabla y^\varepsilon(x,T)|^2dx+\tfrac12\int_{\Omega_0}|\nabla y^\varepsilon(x,0)|^2dx-\tfrac12\int_{\Omega_T\setminus\Omega_0}|\nabla y^\varepsilon(x,\sigma(x)|^2dx \\ &-\int_\Gamma j(y^\varepsilon(x,T))\,d\sigma_x+\int_\Gamma j(y^\varepsilon(x,0))\,d\sigma_x-\varepsilon\int_Q|\nabla y_t^\varepsilon|^2\,dx\,dt-\varepsilon\int_\Sigma y_t^\varepsilon\tfrac{\partial}{\partial t}\beta(y^\varepsilon)d\sigma_x dt \\ &-\int_{\Omega_0}\nabla y^\varepsilon(x,0)\nabla(y^\varepsilon(x,0)-y_0(x))dx+\int_\Gamma \beta(y^\varepsilon(x,0))(y_0(x)-y^\varepsilon(x,0))\,d\sigma_x \\ &+\varepsilon\int_{\Omega_T\setminus\Omega_0} y_t^\varepsilon(x,\sigma(x))\ \mathrm{div}\ (\nabla y^\varepsilon(x,\sigma(s))dx.\end{aligned} \tag{2.13}$$

Here we use the assumption that $y_0|_{\partial\Omega_0\setminus\Gamma}=0$. Let us note that

$$\begin{aligned}&-y_t^\varepsilon\tfrac{\partial}{\partial t}\beta(y^\varepsilon)\le 0 \text{ a.e. on } \Sigma,\\ &j(y^\varepsilon(x,0))+\beta(y^\varepsilon(x,0))(y_0-y^\varepsilon(x,0))\le j(y_0(x)) \text{ a.e. on },\ ,\\ &\tfrac12|\nabla y^\varepsilon(x,0)|^2-\nabla y^\varepsilon(x,0)\nabla(y^\varepsilon(x,0)-y_0(x))\\ &\qquad\le -\tfrac12|\nabla(y^\varepsilon(x,0)-y_0(x))|^2+\tfrac12|\nabla y_0(x)|^2 \text{ a.e. on } \Omega_0,\end{aligned}$$

and by (2.10)

$$\begin{aligned}&-\int_{\Omega_T\setminus\Omega_0} y_t^\varepsilon(x,\sigma(x))\ \mathrm{div}\ (\nabla y^\varepsilon(x,\sigma(x))dx=\int_{\Omega_T\setminus\Omega_0} y_t^\varepsilon(x,\sigma(x))\ \mathrm{div}\ (y_t^\varepsilon\nabla\sigma)dx\\ &=\int_{\Omega_T\setminus\Omega_0}(y_t^\varepsilon)^2(x,\sigma(x))\Delta\sigma dx+\tfrac12\int_{\Omega_T\setminus\Omega_0}\nabla(y_t^\varepsilon)^2\nabla\sigma dx\\ &=\tfrac12\int_{\Omega_T\setminus\Omega_0}(y_t^\varepsilon)^2(x,\sigma(x))\Delta\sigma dx+\tfrac12\int_{\partial(\Omega_T\setminus\Omega_0)}(y_t^\varepsilon)^2\nabla\sigma\cdot n\,ds\ge 0,\end{aligned}$$

where n denotes the outer normal to $\Omega_T\setminus\Omega_0$. In the last estimate we used $y_t^\varepsilon(x,0)=\frac{1}{\varepsilon}(y^\varepsilon(x,0)-y_0(x))=0$ a.e. on $\partial\Omega_0\setminus$, and the assumption that

$\Delta\sigma \geq 0$ a.e. on $\Omega_T \setminus \Omega_0$ and $\frac{\partial\sigma}{\partial n} \geq 0$ on $\partial\Omega_T \setminus \Gamma$. Using these estimates in (2.13) one obtains

$$\begin{aligned}
&\int_Q |\Delta y^\varepsilon|^2\, dx\, dt + \tfrac{1}{2}\int_{\Omega_T} |\nabla y^\varepsilon(x,T)|^2 dx + \tfrac{1}{2}\int_{\Omega_T\setminus\Omega_0} |\nabla y^\varepsilon(x,\sigma(x))|^2 dx \\
&\quad + \int_\Gamma j(y^\varepsilon(x,T))\, d\sigma_x + \varepsilon \int_Q |\nabla y_t^\varepsilon|^2 dx\, dt + \tfrac{1}{2}\int_{\Omega_0} |\nabla(y^\varepsilon(x,0) - y_0(x))|^2 dx \\
&\leq \tfrac{1}{2}\int_{\Omega_0} |\nabla y_0|^2 dx + \int_\Gamma j(y_0(x))\, d\sigma_x.
\end{aligned} \tag{2.14}$$

Combining (2.7), (2.8) and (2.14) we find

$$\begin{aligned}
&\int_Q \left((y_t^\varepsilon)^2 + |\nabla y^\varepsilon|^2 + |\Delta y^\varepsilon|^2\right) dx dt + \int_{\Omega_T\setminus\Omega_0} |\nabla y^\varepsilon(x,\sigma(x))|^2 dx \\
&\quad + \int_{\Omega_0} |\nabla(y^\varepsilon(\cdot,0) - y_0)|^2 dx + \int_{\Omega_T} \left(\tfrac{1}{2} y^\varepsilon(x,T)^2 + |\nabla y^\varepsilon(x,T)|^2\right) dx \\
&\quad + 2\int_\Gamma j\left(y^\varepsilon(x,T)\right)\, d\sigma_x + \int_\Sigma j(y^\varepsilon) d\sigma_x\, dt \\
&\leq \int_{\Omega_0} \left(\tfrac{1}{2} y_0^2 + |\nabla y_0|^2\right) dx + 2\int_\Gamma j(y_0)\, d\sigma_x.
\end{aligned} \tag{2.15}$$

By estimate (2.15) it follows that there exists $y \in V$ with y_t and $\Delta y \in L^2(Q)$ such that on a subsequence of y^ε we have

$$\begin{array}{rcllll}
y^\varepsilon & \to & y & \text{weakly in} & H^1(Q), & \text{strongly in} \quad L^2(Q), \\
y_t^\varepsilon & \to & y_t & \text{weakly in} & L^2(Q), & \\
\Delta y^\varepsilon & \to & \Delta y & \text{weakly in} & L^2(Q), & \\
\varepsilon y_t^\varepsilon & \to & 0 & \text{strongly in} & L^2(Q), & \\
y^\varepsilon(x,0) & \to & y(x,0) & \text{strongly in} & L^2(\Omega_0), & \\
y^\varepsilon & \to & y & \text{strongly in} & L^2(\Sigma), & \\
\nabla y^\varepsilon & \to & \nabla y & \text{weakly in} & L^2(\Sigma_0), & \text{as} \quad \epsilon \to 0^+.
\end{array} \tag{2.16}$$

In particular we infer that the trace of ∇y on Σ_0 is well-defined in the sense of $\nabla y \in L^2(\Sigma_0)$. Letting ε tend to zero in (2.5) we find that y is a solution to (2.1) which satisfies the estimate (2.3). Since $y_\varepsilon \leq 0$ a.e. in Q for all $\varepsilon > 0$ we deduce that $y \leq 0$ a.e. in Q.

To eliminate the regularity requirement on β that we made before (2.8) one approximates β by a family of continuously differentiable, monotonically increasing functions $\beta^\lambda : \mathbb{R} \to \mathbb{R}$ with $\beta^\lambda(0) = 0$, which satisfy in addition

(i) $\liminf_{\lambda\to 0^+} j^\lambda(r_\lambda) \geq j(r)$ whenever $r_\lambda \to r$ in $\mathbb{R}$,

(ii) $\lim_{\lambda\to 0^+} j^\lambda(r) = j(r)$ for all $r \in \mathbb{R}$,

(iii) $|\beta^\lambda(r) - \beta_\lambda(r)| \le C$ for a constant C independent of $\lambda > 0, r \in \mathbb{R}$,

where

$$j^\lambda(r) = \int_0^r \beta^\lambda(s)ds \quad \text{and} \quad \beta_\lambda = \lambda^{-1}\left(1 - (1+\lambda\beta)^{-1}\right).$$

A specific choice for such a family of functions is given by

$$\beta^\lambda(r) = \int_{-\infty}^{\infty} \left(\beta_\lambda(r - \varepsilon^2 s) - \beta_\lambda(-\varepsilon^2 s)\right) ds.$$

For details we refer to [B1], pp. 157, 171, 322. Repeating the above arguments with β replaced by β^λ one obtains a double indexed family of variational solutions $y^{\varepsilon,\lambda}$ to (2.4) with β replaced by β^λ satisfying

$$\begin{aligned} &\int_Q \left[\varepsilon y_t^{\varepsilon,\lambda}(y^{\varepsilon,\lambda} - z)_t + \nabla y^{\varepsilon,\lambda}\cdot\nabla(y^{\varepsilon,\lambda} - z) + y_t^{\varepsilon,\lambda}(y^{\varepsilon,\lambda} - z)\right] dxdt \\ &\quad + \int_{\Omega_0} \left(y^{\varepsilon,\lambda}(x,0) - y_0\right)\left(y^{\varepsilon,\lambda}(x,0) - z(x,0)\right) dx \\ &\quad + \int_\Sigma \left(j^\lambda(y^{\varepsilon,\lambda}) - j^\lambda(z)\right) d\sigma_x dt \le 0. \end{aligned} \tag{2.17}$$

Moreover the following a priori estimate holds:

$$\begin{aligned} &\int_Q \left((y_t^{\varepsilon,\lambda})^2 + |\nabla y^{\varepsilon,\lambda}|^2 + |\Delta y^{\varepsilon,\lambda}|^2\right) dxdt + \int_{\Omega_T\setminus\Omega_0} |\nabla y^{\varepsilon,\delta}(x,\sigma(x))|^2 dx \\ &+ \int_{\Omega_0} |\nabla\left(y^{\varepsilon,\lambda}(\cdot,0) - y_0\right)|^2 dx + \int_{\Omega_T} \left(\tfrac{1}{2}(y_t^{\varepsilon,\lambda})^2(x,T) + |\nabla y^{\varepsilon,\lambda}(x,T)|^2\right) dx \\ &+2\int_\Gamma j^\lambda\left(y^{\varepsilon,\lambda}(x,T)\right) d\sigma_x + \int_\Sigma j^\lambda(y^{\varepsilon,\lambda}) d\sigma_x\, dt \\ &\le \int_{\Omega_0} \left(\tfrac{1}{2}y_0^2(x) + |\nabla y_0(x)|^2\right) dx + 2\int_\Gamma j^\lambda(y_0(x))\, d\sigma_x + \int_\Sigma j^\lambda(0) d\sigma_x\, dt. \end{aligned} \tag{2.18}$$

Now let us fix $\varepsilon > 0$. Due to (2.18) there exists $y^\varepsilon \in H^1(Q)$ with y_t^ε and $\Delta y^\varepsilon \in L^2(Q)$ and a subsequence of $\{y^{\varepsilon,\lambda}\}_{\lambda>0}$ that converges as $\lambda \to 0^+$ to y^ε in the sense of (2.16) with $\varepsilon \to 0^+$ replaced by $\lambda \to 0^+$. Moreover there exist functions $\xi_1 \in L^2(,\,)$ and $\xi_2 \in L^2(\Sigma)$ such that

$$\lim_{\lambda\to 0} j^\lambda\left(y^{\varepsilon,\lambda}(\cdot,T)\right) \;\to\; \xi_1 \quad \text{weakly in} \quad L^2(,\,) \tag{2.19}$$

$$\lim_{\lambda\to 0} j^\lambda(y^{\varepsilon,\lambda}) \;\to\; \xi_2 \quad \text{weakly in} \quad L^2(\Sigma). \tag{2.20}$$

Due to (iii) we have

$$|j^\lambda(r)| \leq |j_\lambda(r)| + C|r| \quad \text{for all} \quad r \in \mathbb{R}, \lambda > 0,$$

and hence using Fatou's lemma for the term $\int_\Sigma j^\lambda(y^{\varepsilon,\lambda}) d\sigma_x dt$ and Lebesgue's bounded convergence theorem for $\int_\Sigma j^\lambda(z) d\sigma_x dt$ we can pass to the limit as $\lambda \to 0^+$ in (2.17) and (2.18). In this way we obtain (2.5) and (2.15) for every $\varepsilon > 0$. Passing to the limit with respect to $\varepsilon \to 0$ we obtain existence of a solution to (2.1) which satisfies the estimate (2.3) for every β satisfying the condition specified at the beginning of this section.

The uniqueness of y is immediate by the variational formulation (2.2). This completes the proof.

3 A least squares approach to identify the boundary heat transfer function

We shall consider the optimization problem

$$\left.\begin{array}{l} \text{minimize} \displaystyle\int_{\Sigma_0} (\nabla y \cdot \nabla \sigma - \rho)^2 d\sigma_x dt \\ \text{subject to } \beta \in \mathcal{A} \text{ and to } y \in H^1(Q) \text{ a solution to (2.1).} \end{array}\right\} \qquad (P)$$

Throughout this section we assume that y_0 and σ satisfy the assumptions of Proposition 2.1. Then (P) is well-defined as a consequence of (2.3).

Theorem 3.1 *Problem (P) has at least one solution $(y^*, \beta^*) \in H^1(Q) \times \mathcal{A}$.*

Proof. Let d denote the infimum in problem (P) and let $\{(y_n, \beta_n)\} \subset H^1(Q) \times \mathcal{A}$ be a minimizing sequence satisfying

$$d \leq |\nabla y_n \cdot \nabla \sigma - \rho|^2_{L^2(\Sigma_0)} \leq d + \tfrac{1}{n} \tag{3.1}$$

with y_n the variational solution to

$$\left.\begin{array}{rcll} (y_n)_t - \Delta y_n & = & 0 & \text{in } Q \\ y_n & = & 0 & \text{in } \Sigma_0 \\ \frac{\partial y_n}{\partial \nu} + \beta_n(y_n) & \ni & 0 & \text{in } \Sigma \\ y_n(x,0) & = & y_0(x) & \text{in } \Omega_0. \end{array}\right\} \tag{3.2}$$

Due to (2.3) and the properties of elements in $\mathcal{A}$ there exists a constant C such that

$$|y_n|_{H^1(Q)} + |\Delta y_n|_{L^2(Q)} + |\nabla y_n|_{L^2(\Sigma_0)} + |j_n(y_n)|_{L^2(\Sigma)} \leq C, \tag{3.3}$$

for all n.

It follows that there exists $y \in H^1(Q)$ with $\Delta y \in L^2(Q)$ such that for a subsequence

$$y_n \to y \quad \text{weakly in} \quad H^1(Q) \quad \text{and} \quad \nabla y_n \to \nabla y \quad \text{weakly in} \quad L^2(\Sigma_0).$$

Taking lim inf in (3.1) we infer that $d = |\nabla y \cdot \nabla\sigma - \rho|^2_{L^2(\Sigma_0)}$. It remains to pass to the limit in

$$\int_Q (-y_n(y_n - z)_t + \nabla y_n \cdot \nabla(y_n - z))\, dx\, dt + \int_\Sigma (j_n(y_n) - j_n(z))\, d\sigma_x\, dt$$
$$+ \int_{\Omega_T} y_n(x,T)\,(y_n(x,T) - z(x,T))\, dx - \int_{\Omega_0} y_0(x)\,(y_n(x,0) - z(x,0))\, dx \leq 0$$

for all $z \in V$. This is simple for all terms except for the second. To take the lim inf on that term one uses the Arzela-Ascoli theorem to conclude convergence of $\{j_n\}$ in $C(I)$ for every compact $I \subset \mathbb{R}$. By Fatou's lemma and Lebesgue's bounded convergence theorem (using $j_n \in \mathcal{K}$) we obtain

$$\int_\Sigma j(y) d\sigma_x dt + \int_\Sigma j(z) d\sigma_x dt \leq \liminf \int_\Sigma j_n(y_n) d\sigma_x dt + \liminf \int_\Sigma j_n(z) d\sigma_x dt,$$

and the desired result follows.

Next we approximate problem (P) by the following family of optimal control problems in the control variable v:

$$\left.\begin{array}{l} \text{minimize} \left\{ \int_{\Sigma_0} (\nabla y \cdot \nabla\sigma - \rho)^2 d\sigma_x dt + \frac{1}{\lambda} \int_\Sigma (j(y) + j^*(-v) + vy)\ dx\, dt \right\} \\ \text{subject to } y \in H^1(Q), j \in \mathcal{K}, v \in L^2(\Sigma) \\ \text{and } y \text{ a variational solution to} \end{array}\right\} \quad (P_\lambda)$$

$$\left.\begin{array}{rcll} y_t - \Delta y & = & 0 & \text{in } Q \\ y & = & 0 & \text{in } \Sigma_0 \\ \frac{\partial y}{\partial \nu} & = & v & \text{in } \Sigma \\ y(x,0) & = & y_0(x) & \text{in } \Omega_0, \end{array}\right\} \quad (3.4)$$

where $\lambda > 0$, and

$$\begin{array}{rcl} \mathcal{K} & = & \{ j : \mathbb{R} \to \mathbb{R} \quad \text{is convex, continuous, } j(0) = 0,\ 0 \in \partial j(0), \\ & & \alpha_0 + \omega_0 r^2 \leq j(r) \leq \alpha_1 + \omega_1 r^2, \quad \text{for all} \quad r \in \mathbb{R}\}. \end{array} \quad (3.5)$$

Note that $\mathcal{A} = \{\beta = \partial j : j \in \mathcal{K}\}$.

The introduction of the second term in the payoff of (P_λ) is suggested by the equivalence between

$$\frac{\partial y}{\partial \nu} + \beta(y) \ni 0 \quad \text{and} \quad j(y) + j^*\left(-\frac{\partial y}{\partial \nu}\right) = -y\frac{\partial y}{\partial \nu}$$

at all points $(x,t) \in \Sigma$ where these equations are well-defined. Moreover

$$j(y) + j^*\left(-\frac{\partial y}{\partial \nu}\right) + y\frac{\partial y}{\partial \nu} \geq 0 \tag{3.6}$$

whenever this expression is defined. Upon setting $v = \frac{\partial y}{\partial \nu}$ and integrating (3.6) on Σ we obtain

$$\int_\Sigma (j(y) + j^*(-v) + vy)\, d\sigma_x dt \geq 0,$$

with equality holding if and only if

$$j(y) + j^*(-v) + vy = 0 \text{ a.e. in } \Sigma. \tag{3.7}$$

Thus the second term in the payoff of (P_λ) is a penalty term realizing (3.7) and (P_λ) constitutes a splitting or mixed finite element method with respect to the variable $\frac{\partial y}{\partial \nu}$ for problem (P).

Before going further, some comments concerning (3.4) are in order. For $y_0 \in H^1(\Omega_0)$ with $y_0 = 0$ a.e. in $\partial\Omega_0 \setminus ,$, and if $v_t \in L^2(\Sigma)$, then (3.4) has a unique variational solution y satisfying

$$\int_Q (\nabla y \cdot \nabla z - y z_t)\, dx dt + \int_{\Omega_T} y(x,T) z(x,T) dx = \int_\Sigma vz\, d\sigma_x dt + \int_{\Omega_0} y_0(x) z(x,0)\, dx, \tag{3.8}$$

for all $\{z \in H^1(Q) : z = 0 \text{ in } \Sigma_0\}$ and

$$y = 0 \text{ a.e. in } \Sigma_0. \tag{3.9}$$

Moreover y satisfies $y \in H^1(Q)$ with $\Delta y \in L^2(Q)$ and $\nabla y \cdot \nabla \sigma \in L^2(\Sigma_0)$. The proof follows from the proof of Proposition 2.1, see also [B2]. It is simple to argue that

$$\int_Q |\nabla y|^2\, dx\, dt + \int_{\Omega_t} y^2(x,t) dx \leq \tilde{C}(\int_{\Omega_0} y_0^2 dx + \int_\Sigma v^2 d\sigma_x\, dt), \tag{3.10}$$

with $\tilde{C}$ independent of y_0, v and $t \in [0,T]$.

For general $v \in L^2(\Sigma)$ it follows by density arguments based on (3.10) that (3.4) admits an unique variational solution satisfying (3.8) and (3.10) as well. Since $y(t) \in H^1(\Omega_t)$ for a.e. $t \in [0,T]$ and $\int_0^T |y(\cdot,t)|^2_{H^1(\Omega_t)} dt < \infty$, the trace of $y(t)$ on $\partial\Omega_t$ is well-defined and belongs to $L^2(\partial\Omega_t)$ for a.e. $t \in (0,T)$, so that (3.9) makes sense. In the neighborhood of Σ_0 the solution y of (3.4) is in fact more regular. For that purpose we introduce the domain $\Omega' \subset \Omega_0$ (and hence $\Omega' \subset \Omega_t$ for all $t \in [0,T]$) with the property that $\partial\Omega'$ consists of two connected components, one of which coincides with , and the other lies in the interior of Ω_0. Then it is straightforward to argue, e.g. by a Galerkin procedure, that the solution of (3.4) satisfies in addition to (3.10)

$$\int_0^T |y(\cdot,t)|^2_{H^2(\Omega_t \setminus \Omega')} dt \leq \bar{C}(|y_0|^2_{H^1(\Omega)} + |v|^2_{L^2(\Sigma)}) \tag{3.11}$$

with $\bar{C}$ independent of $y_0 \in H^1(\Omega)$ and $v \in L^2(\Sigma)$. In particular the payoff in (P_λ) is well-defined for $v \in L^2(\Sigma)$.

We shall say that $(y_\lambda, j_\lambda, v_\lambda)$ converges in $(L^2(Q))_w \times \mathcal{K} \times (L^2(\Sigma))_w$ to (y, j, v), if $y_\lambda \to y$ weakly in $L^2(Q)$, $v_\lambda \to v$ weakly in $L^2(\Sigma)$ and $j_\lambda \to j$ uniformly on compact subsets of $\mathbb{R}$.

Theorem 3.2 *For every $\lambda > 0$ there exists at least one solution $(y_\lambda, j_\lambda, v_\lambda) \in L^2(Q) \times \mathcal{K} \times L^2(\Sigma)$ of (P_λ). Moreover $\{y_\lambda, j_\lambda, v_\lambda\}_{\lambda>0}$ contains a clusterpoint in $(L^2(Q))_w \times \mathcal{K} \times (L^2(\Sigma))_w$, the first two components of every such clusterpoint are a solution of (P) and $\liminf_{\lambda \to 0^+} P_\lambda = \inf P$.*

Proof. Let (y_n, j_n, v_n) be a minimizing sequence for (P_λ) satisfying

$$\begin{aligned} d_\lambda \le \int_{\Sigma_0} (\nabla y_n \cdot \nabla \sigma - \rho)^2 \, d\sigma_x dt + \tfrac{1}{\lambda} \int_{\Sigma} (j_n(y_n) + j_n^*(-v_n) + v_n y_n) \, d\sigma_x dt \\ \le d_\lambda + \tfrac{1}{n}, \end{aligned} \tag{3.12}$$

for every $n \in \mathbb{N}$, where y_n, v_n satisfy (3.4), $j_n \in \mathcal{K}$ and $d_\lambda = \inf(P_\lambda)$. Due to the properties of $\mathcal{K}$ using (3.8) to eliminate the term $\int_\Sigma v_n y_n d\sigma_x dt$ it follows that

$$|\nabla y|^2_{L^2(Q)} + |v_n|_{L^2(\Sigma)} + |y_n|_{L^2(\Sigma)} \le C,$$

for a constant C independent of n. From (3.11) it follows that $\{|y_n|_{H^2(\Omega_t \backslash \Omega')} : t \in [0,T]\}_{n=1}^\infty$ is bounded as well. Consequently there exist $y_\lambda \in L^2(Q)$ with $\nabla y_\lambda \in L^2(Q)$ and $\nabla y_\lambda \cdot \nabla \sigma \in L^2(\Sigma_0)$, and $v_\lambda \in L^2(\Sigma)$, such that on a subsequence, again denoted by $\{n\}$, we have

$$\begin{array}{rcll} y_n & \to & y & \text{weakly in } & L^2(Q) \\ \nabla y_n & \to & \nabla y & \text{weakly in } & L^2(Q) \\ \nabla y_n \cdot \nabla \sigma & \to & \nabla y \cdot \nabla \sigma & \text{strongly in } & L^2(\Sigma_0) \\ v_n & \to & v & \text{weakly in } & L^2(\Sigma). \end{array}$$

Since $j_n \in \mathcal{K}$ for all n there exists a constant C_1 independent of n such that

$$|\partial j_n^*(r)| + |\partial j_n(r)| \le C_1(|r|^2 + 1) \quad \text{for all} \quad r \in \mathbb{R}.$$

Consequently by the Arzela Ascoli theorem there exists $j_\lambda \in \mathcal{K}$ such that

$$j_n(r) \to j_\lambda(r) \quad \text{and} \quad j_n^*(r) \to j_\lambda^*(r) \quad \text{uniformly in } r \text{ on compact subsets of } \mathbb{R}. \tag{3.13}$$

It is simple to argue that (y_λ, v_λ) is a variational solution of (3.4), i.e. that (3.8) is satisfied.

We next argue that $\{\gamma_0 y_n\}_{n=1}^\infty$ is compact in $L^2(\Sigma)$. Here γ_0 denotes the trace of y on Σ. Let $H^1_m(\Omega') = \{\varphi \in H^1(\Omega') : \varphi | \partial\Omega' \backslash , = 0\}$. Using the fact that $\{v_n\}$ is bounded in $L^2(\Sigma)$ it is simple to argue that $\{y_n\}$ is bounded in $L^2((0,T);H^1(\Omega'))$ and $\{(y_n)_t\}$ is bounded in $L^2((0,T);H^1_m(\Omega')^*)$. For every $\varepsilon \in (0,1/2)$ one has the continuous injections

$$H^1(\Omega') \subset H^{1/2+\varepsilon}(\Omega') \subset H^1_m(\Omega')^*,$$

with the first one being also compact. By Aubin's compactness theorem it follows that $\{y_n\}$ is precompact in $L^2((0,T);H^{1/2+\varepsilon}(\Omega'))$ and hence $\{\gamma_0 y_n\}$ is precompact in $L^2(\Sigma)$. Hence for a subsequence, again denoted by the same symbol, we have

$$\gamma_0 y_n \to \gamma_0 y_\lambda \text{ strongly in } L^2(\Sigma). \tag{3.14}$$

Since $j_n \in \mathcal{K}$ it follows that $j_n(y_n) \geq \alpha_0$ a.e. and hence by Fatou's lemma and (3.13), (3.14) we find

$$\liminf_{n\to\infty} \int\limits_\Sigma j_n(y_n) d\sigma_x\, dt \geq \int\limits_\Sigma j_\lambda(y_\lambda) d\sigma_x\, dt. \tag{3.15}$$

Since $v_n \to v_\lambda$ only weakly in $L^2(\Sigma)$, taking the $\liminf$ in $\int_\Sigma j_n^*(-v_n)d\sigma_x dt$ is more delicate. Based on (3.13) one can argue as in the proof of Theorem 2.1 of [BK1] to obtain

$$\liminf_{n\to\infty} \int\limits_\Sigma j_n^*(-v_n) d\sigma_x\, dt \geq \int\limits_\Sigma j_\lambda^*(-v_\lambda) d\sigma_x\, dt. \tag{3.16}$$

We also have

$$\lim_{n\to\infty} \int\limits_\Sigma y_n v_n d\sigma_x\, dt = \int\limits_\Sigma y_\lambda v_\lambda d\sigma_x\, dt. \tag{3.17}$$

Combining (3.12) and (3.15) – (3.17) we find

$$d_\lambda = \int\limits_{\Sigma_0} (\nabla y_\lambda \nabla \sigma - \rho)^2 d\sigma_x\, dt + \tfrac{1}{\lambda} \int\limits_\Sigma (j_\lambda(y_\lambda) + j_\lambda^*(-v) + y_\lambda v_\lambda) d\sigma_x\, dt,$$

and hence $(y_\lambda, j_\lambda, v_\lambda)$ is a solution to (P_λ).

We turn to the asymptotic behavior of (P_λ) as $\lambda \to 0^+$. For every $\lambda > 0$ we have

$$\begin{aligned} &\int\limits_{\Sigma_0} (\nabla y_\lambda \cdot \nabla \sigma - \rho)^2 d\sigma_x\, dt + \tfrac{1}{\lambda} \int\limits_\Sigma \left(j_\lambda(y_\lambda) + j_\lambda^*(-v_\lambda) + y_\lambda v_\lambda \right) d\sigma_x\, dt \leq \\ &\qquad \leq \int\limits_{\Sigma_0} (\nabla y^0 \nabla \sigma - \rho)^2 d\sigma_x\, dt = \inf P, \end{aligned} \tag{3.18}$$

where (y^0, j^0, v^0) is any solution to (P).

Arguing as above, with $(y_\lambda, j_\lambda, v_\lambda)$ replacing (y_n, j_n, v_n) we find that $\{|y_\lambda|_{L^2(Q)}\}$, $\{|\nabla y_\lambda|_{L^2(Q)}\}$, $\{|y_\lambda|_{H^2(\Omega_t \setminus \Omega')} : t \in [0,T]\}$, and $\{|v_\lambda|_{L^2(\Sigma)}\}$ are bounded uniformly for $\lambda > 0$. Consequently there exist $y \in L^2(Q)$ with $\nabla y \in L^2(Q)$ and

$\nabla y \cdot \nabla\sigma \in L^2(\Sigma_0)$, and $v \in L^2(\Sigma)$ such that on a subsequence $\{\lambda_n\}$ of $\{\lambda\}_{\lambda>0}$, with $\lim \lambda_n = 0$ we have

$$\begin{array}{rcll} y_{\lambda_n} & \to & y & \text{weakly in} \quad L^2(Q) \\ \nabla y_{\lambda_n} & \to & \nabla y & \text{weakly in} \quad L^2(Q) \\ \nabla y_{\lambda_n} \cdot \nabla\sigma & \to & \nabla y \nabla\sigma & \text{strongly in} \quad L^2(\Sigma_0) \\ v_{\lambda_n} & \to & v & \text{weakly in} \quad L^2(\Sigma), \end{array}$$

as $n \to \infty$. Moreover there exists $j \in \mathcal{K}$ such that

$$j_{\lambda_n}(r) \to j(r) \text{ and } j^*_{\lambda_n}(r) \to j^*(r) \text{ uniformly in } r \text{ on compact subsets of } \mathbb{R}.$$

As above one argues that

$$\liminf_{n\to\infty} \int\limits_{\Sigma} \left(j_{\lambda_n}(y_{\lambda_n}) + j^*_{\lambda_n}(-v_n) + y_{\lambda_n} v_{\lambda_n} \right) d\sigma_x\, dt$$
$$ge \int\limits_{\Sigma} \left(j(y) + j^*(-v) + yv \right) d\sigma_x\, dt \geq 0,$$

and by (3.18)

$$j(y) + j^*(-v) + yv = 0 \text{ a.e. in } \Sigma.$$

It follows that

$$v \in \beta(y) \text{ a.e. in } \Sigma.$$

Clearly (y, v) is a variational solution of (3.4) and hence (y, j, v) is a variational solution of (2.1). Moreover from (3.18)

$$\int\limits_{\Sigma_0} (\nabla y \cdot \nabla\sigma - \rho)^2 d\sigma_x\, dt \leq \int\limits_{\Sigma_0} (\nabla y^0 \cdot \nabla\sigma - \rho)^2 d\sigma_x\, dt,$$

and thus (y, j, v) is a solution to (P). By (3.18) it follows that $\liminf_{\lambda\to 0^+} P_\lambda = \inf P$. This ends the proof.

In the final result of this section we return to the control theory interpretation of (P). We shall assert that the feedback control law

$$u \in -\partial j_\lambda(y) \text{ in } \Sigma$$

with j_λ the second component of the solution $(y_\lambda, j_\lambda, v_\lambda)$ of (P_λ) applied to the original problem (P) gives a suboptimal approximating solution for that problem.

Theorem 3.3 *Let $(y_\lambda, j_\lambda) \in L^2(Q) \times \mathcal{K}$ be given by Theorem 3.2 and let y^*_λ be the solution to (2.1) with $\beta = \partial j_\lambda$. Then*

$$y^*_\lambda - y_\lambda \to 0 \text{ and } \nabla(y^*_\lambda - y_\lambda) \to 0 \text{ strongly in } L^2(Q) \text{ as } \lambda \to 0^+. \tag{3.19}$$

Moreover, every subsequence $\{(y^*_{\lambda_n}, j_{\lambda_n})\}$ of $\{(y_\lambda, j_\lambda)\}_{\lambda>0}$ contains a cluster point (y^*, j^*) in $(L^2(Q))_w \times \mathcal{K}$, which is a solution to (P).

Proof. For every $\lambda > 0$ we have

$$\int_Q \left((y_\lambda^*)_t(y_\lambda^* - z) + \nabla y_\lambda^* \cdot \nabla(y_\lambda^* - z)\right)\, dx\, dt + \int_\Sigma \left(j_\lambda(y_\lambda^*) - j_\lambda(z)\right) d\sigma_x\, dt \leq 0$$

and

$$\int_Q \left((y_\lambda)_t w + \nabla y_\lambda \cdot \nabla w\right)\, dx\, dt - \int_\Sigma v_\lambda w\, d\sigma_x\, dt = 0$$

for all $z, w \in V$. Setting $z = y_\lambda, w = y_\lambda^* - y_\lambda$ and subtracting the above equality from the inequality we find

$$\begin{aligned} &\int_{\Omega_T} |y_\lambda^*(x,T) - y_\lambda(x,T)|^2 dx + \int_Q |\nabla(y_\lambda^* - y_\lambda)|^2\, dx\, dt \\ &\quad + \int_\Sigma \left[j_\lambda(y_\lambda^*) - j_\lambda(y_\lambda) + v_\lambda(y_\lambda^* - y_\lambda)\right] d\sigma_x\, dt \leq 0. \end{aligned} \tag{3.20}$$

We recall from (3.18) that

$$\int_\Sigma \left(j_\lambda(y_\lambda) + j_\lambda^*(-v_\lambda) + y_\lambda v_\lambda\right) d\sigma_x\, dt \leq C\lambda \quad \text{for all} \quad \lambda > 0$$

and hence

$$\begin{aligned} &-\int_\Sigma \left(j_\lambda(y_\lambda^*) - j_\lambda(y_\lambda) + v_\lambda(y_\lambda^* - y_\lambda)\right) d\sigma_x\, dt \\ &\qquad \leq C\lambda - \int_\Sigma \left(v_\lambda y_\lambda^* + j_\lambda(y_\lambda^*) + j_\lambda^*(-v_\lambda)\right) d\sigma_x\, dt \leq C\lambda. \end{aligned}$$

Inserting this estimate into (3.20) the validity of (3.19) follows. For any sequence $\{(y_{\lambda_n}^*, j_{\lambda_n})\}$ the sequence $\{(y_{\lambda_n}, j_{\lambda_n}, v_{\lambda_n})\}$ contains, due to Theorem 3.2, a convergent subsequence, the limit of which has the property that its first two components are a solution to (P).

4 Solving Problems (P_λ)

Besides the obvious fact that the problems (P_λ) are infinite dimensional and that numerical realization requires discretization, these problems represent some serious structural difficulties. In this section we shall address the problems related to the numerical treatment of the set $\mathcal{K}$, the elements of which are defined over an unbounded domain and are required to be strictly convex. Moreover we characterize the gradient of the cost with respect to (y, j, v). These considerations are independent of the spatial dimension. In the following section we describe a specific numerical realization in spatial dimension one. There we shall also take into consideration the inherent illposedness of the optimization problems: In fact, the cost functional in (P_λ) is not coercive with respect to v or j.

Throughout this section we shall assume that $y_0 \in H^1(\Omega_0) \cap C(\Omega_0)$, that $y_0 < 0$ in Ω_0 and that $y_0 = 0$ on $\partial\Omega_0 \setminus$, . Then for any $\beta \in \mathcal{A}$ the solution y to (2.1) satisfies

$$-a := \inf_{\Omega_0} y_0 \leq y(x,t) \leq 0 \quad \text{in} \quad Q \tag{4.1}$$

and

$$-a \leq y(x,t) < 0 \quad \text{on} \quad \Sigma. \tag{4.2}$$

This follows from the strong maximum principle for parabolic equations (see e.g. [12]), since the extrema of y in Q are not attained on Σ. Since in view of (2.1) only the values of j on the range of y contribute to the value of the cost functional in (P), it therefore suffices to restrict the domain of j to $[-a, 0]$. Concerning the actual problem formulation on a bounded domain two conflicting issues arise. On the one hand one would like to enlarge the domain for j beyond $[-a, 0]$ so that perturbed problems (e.g. (P_λ)) still have the property that their solutions are not effected by restricting the domain and on the other hand, enlarging the domain introduces some indeterminancy into the problem (with adverse numerical consequences), since the limit problem is not effected by the values of j on the complement of $[-a, 0]$. For the purpose of the present section we shall restrict the domain of j to $I = [-a-\varepsilon, \varepsilon]$ for some $\varepsilon \geq 0$.

Another practical issue consists in the numerical realization of the convexity assumption involved in the definition of $\mathcal{A}$. This will be accomplished by an additional regularity assumption for j. For computational purposes we shall therefore consider

$$\left.\begin{array}{l} \text{minimize} \displaystyle\int_{\Sigma_0} (\nabla y \cdot \nabla \sigma - \rho)^2 d\sigma_x dt \\ \text{on} (y, j) \in H^1(Q) \times \mathcal{K}^1, \text{ subject to (2.1) with } \beta = \partial j \end{array}\right\} \tag{P^1}$$

where

$$\mathcal{K}^1 = \{ j \in H^2(I) : j(0) = j'(0) = 0, \;\; 2\omega_0 \leq j''(r) \leq 2\omega_1, \text{ for a.e. } r \in I \}.$$

Accordingly (P_λ) is replaced by

$$\left.\begin{array}{l} \text{minimize} \displaystyle\int_{\Sigma_0} (\nabla y \cdot \nabla \sigma - \rho)^2 d\sigma_x dt + \frac{\eta}{2} |v|^2_{L^2(\Sigma)} + \\ \qquad \displaystyle\frac{1}{\lambda} \int_{\Sigma} (j(y) + j^*(-v) + vy) d\sigma_x dt \\ \text{on} (y, j, v) \in H^1(Q) \times \mathcal{K}^1 \times L^2(\Sigma) \text{ subject to (3.4).} \end{array}\right\} \tag{P^1_λ}$$

In (P^1_λ) it is understood that $j(y)=\infty$ for $y\notin I$ and $j^*:\mathbb{R}\to\mathbb{R}$ is defined by

$$j^*(p)=\sup\{py-j(y): y\in I\},\ \text{for } p\in\mathbb{R}.$$

The regularization term $\frac{\eta}{2}|v|^2_{L^2(\Sigma)}$ in (P^1_λ) has been added for numerical purposes.

Since $\mathcal{K}^1$ is a closed convex subset of

$$\begin{aligned}&\mathcal{K}_I=\{j:I\to\mathbb{R}, j \text{ convex and continuous}, j(0)=0,\ 0\in\partial j(0),\\ &\alpha_0+\omega_0r^2\le j(r)\le\alpha_1+\omega_1r^2, \quad \text{for all } r\in I\},\end{aligned}$$

and, as seen above, problem (P) can be restricted to $\mathcal{K}_I$, Theorems 3.1, 3.2, and 3.3 remain valid for (P^1) and (P^1_λ).

We next characterize the gradients of (P^1_λ) with respect to j and (y,v). At the same time we note that (P^1_λ) can be decomposed into two convex optimization problems. These are:

1) For fixed $j\in\mathcal{K}^1$ solve the optimal control problem

$$\begin{aligned}&\text{minimize} && \int_{\Sigma_0}(\nabla y\cdot\nabla\sigma-\rho)^2d\sigma_x dt+\tfrac{\mu}{2}|v|^2_{L^2(\Sigma)}\\ &&& +\tfrac{1}{\lambda}\int_{\Sigma}(j(y)+j^*(-v)+yv)d\sigma_x dt \\ &\text{on} && (y,v)\in H^1(Q)\times L^2(\Sigma) \text{ subject to (3.4)}.\end{aligned} \tag{4.3}$$

2) For fixed $(y,v)\in H^1(Q)\times L^2(\Sigma)$ solve the minimization problem

$$\begin{aligned}&\text{minimize} && \int_{\Sigma}(j(y)+j^*(-v))d\sigma_x dt\\ &\text{subject to} && j\in\mathcal{K}^1\end{aligned} \tag{4.4}$$

We turn to the characterization of the gradients of the cost functionals in (4.3) and (4.4). Problem (4.3) is a convex optimal control problem, in fact, using (3.4), the term $\int_\Sigma y\,v\,d\sigma_x\,dt$ can be replaced by

$$\frac{1}{2}\int_{\Omega_T}y(x,T)^2\,dx-\frac{1}{2}\int_{\Omega_0}y_0^2\,dx+\int_Q|\nabla y|^2d\sigma_x\,dt.$$

This problem has a unique solution that is characterized by

$$v=-\left(\mu+\tfrac{1}{\lambda}\partial j^*\right)^{-1}\left(\tfrac{1}{\lambda}y+p\right), \tag{4.5}$$

or equivalently

$$v=\begin{cases}\dfrac{1}{\mu\lambda}\Big[(1+\lambda\mu j')^{-1}-1\Big](y+\lambda p) & \text{a.e. in } \{(1+\lambda\mu\,\partial j)^{-1}(y+\lambda p)\in \operatorname{int} I\}\\ -\dfrac{1}{\mu\lambda}(y+\lambda p-\epsilon) & \text{a.e. in } \{(1+\lambda\mu\,\partial j)^{-1}(y+\lambda p)\ge\epsilon\}\\ -\dfrac{1}{\mu\lambda}(y+\lambda p+a+\epsilon) & \text{a.e. in } \{(1+\lambda\mu\,\partial j)^{-1}(y+\lambda p)\le -a-\epsilon\},\end{cases}$$

where p is the variational solution to

$$\begin{array}{rcll} p_t + \Delta p & = & 0 & \text{in } Q \\ \frac{\partial p}{\partial \nu} & = & \frac{1}{\lambda}(v + j'(y)) & \text{in } \Sigma \\ p & = & -2(\nabla y \cdot \nabla \sigma - \rho)|\nabla \sigma| & \text{in } \Sigma_0 \\ p(\cdot, T) & = & 0 & \text{in } \Omega_T. \end{array}$$

Let $\mathcal{J}$ denote the cost functional in (4.3). Then its gradient in direction $\delta v \in L^2(\Sigma)$ at $v \in L^2(\Sigma)$ is given by

$$\nabla \mathcal{J}(v)(\delta v) = \int_\Sigma (\mu v - p - \tfrac{1}{\lambda}\partial j^*(-v) - \tfrac{1}{\lambda} y)\delta v \, d\sigma_x dt.$$

We next turn to (4.4) which we rewrite as

$$\left.\begin{array}{ll} \text{minimize} & \int_\Sigma (j(y) + j^*(-v))\, d\sigma_x dt \\ \text{over} & \vartheta \in L^2(I) \text{ subject to} \\ & j''(r) = \vartheta(r) \ \text{ a.e. in } I \\ & j(0) \ = j'(0) = 0 \\ & 2\omega_0 \le \vartheta \le 2\omega_1. \end{array}\right\} \tag{4.6}$$

Solving the differential equation in (4.6) for j as a function of ϑ we find

$$j(r) = \int_0^r (r - s)\vartheta(s) ds \quad \text{for } r \in I$$

and

$$j^*(p) = \int_0^{\gamma_\vartheta(p)} s \, ds \quad \text{for } p \in \mathbb{R},$$

where

$$\gamma_\vartheta(p) = \begin{cases} \beta_\vartheta^{-1}(p) & \text{if} \quad p \in \beta_\vartheta(I) \\ \varepsilon & \text{if} \quad p \ge \beta_\vartheta(\varepsilon) \\ -a - \varepsilon & \text{if} \quad p \le \beta_\vartheta(-a - \varepsilon), \end{cases}$$

and

$$\beta_\vartheta(r) = j'(r) = \int_0^r \vartheta(r) ds.$$

Since $\vartheta \ge 2\omega_0 > 0$ invertibility of β_ϑ follows. The cost functional in (4.4) can be expressed as a function of ϑ in the following way:

$$\Phi(\vartheta) = \int_\Sigma \Big(\int_0^{y(x,t)} (y(x,t) - r)\vartheta(r) ds) + \int_0^{\gamma_\vartheta(-v(x,t))} r\vartheta(r) ds \Big) d\sigma_x dt,$$

where $(y, v) \in H^1(Q) \times L^2(\Sigma)$, with $y(x,t) \in I$ a.e., is fixed.

Thus (4.4) can be expressed as

$$\begin{array}{l} \text{minimize } \Phi(\vartheta) \text{ over } \vartheta \in U \\ \text{where } U = \{\vartheta \in L^\infty(\Sigma) : 2\omega_0 \le \vartheta \le 2\omega_1 \text{ a.e. in } I\}. \end{array} \tag{4.7}$$

Arguing as in [BK1] we see that Φ is convex. Moreover it is Gateaux differentiable in a neighborhood of U and we find for the derivative at ϑ in direction w

$$\begin{aligned} \nabla\Phi(\vartheta)(w) \;=\; & \int_\Sigma \Big[\int_0^{y(x,t)} (y(x,t) - r)\, w(r) dr \\ & + \int_0^{\beta_\gamma^{-1}(-v(x,t))} \Big(r - \beta_\gamma^{-1}(-v(x,t))\Big)\, w(r) dr \Big] d\sigma_x dt, \end{aligned}$$

and the Riesz representator, denoted by μ, is given by

$$\begin{aligned} \mu \;=\; & -\int_\Sigma (y(x,t) - r)\, \chi_{[y(x,t),0]}(r) d\sigma_x dt \\ & + \int_\Sigma \Big(\beta_\gamma^{-1}(-v(x,t) - r\Big)\, \chi_{[\beta_\gamma^{-1}(-v(x,t)),0]}(r) d\sigma_x dt, \end{aligned}$$

where χ denotes the characteristic function of the indicated interval. We therefore deduce that the solutions ϑ to (4.7) satisfy

$$\begin{array}{l} 0 \in \mu + N_U(\vartheta), \\ \text{where } N_U(\vartheta) \text{ is the normal cone to } U \text{ at } \vartheta, \text{ defined by} \\ N_U(\vartheta) = \{\eta \in L^1(I) : \eta(r) = 0 \quad \text{if} \quad 2\omega_0 < \vartheta(r) < 2\omega_1;\; \eta(r) \ge 0 \quad \text{if} \\ \vartheta(r) = 2\omega_1;\; \eta(r) \le 0 \quad \text{if} \quad \vartheta(r) = 2\omega_0\}. \end{array} \tag{4.8}$$

From (4.8) we conclude that the optimal solution ϑ satisfies

$$\vartheta(r) \begin{cases} = 2\omega_0 & \text{if} \quad \mu(r) > 0 \\ = 2\omega_1 & \text{if} \quad \mu(r) < 0 \\ \in (2\omega_0, 2\omega_1) & \text{if} \quad \mu(r) = 0. \end{cases} \tag{4.9}$$

5 Numerical Experiments

We carried out numerical tests in spatial dimension $n = 1$ with the following specifications:

- $\Omega = (0, 1)$
- $\Omega_0 = (\frac{1}{2}, 1)$

- $T = \frac{1}{2}$; $\Omega_T = (l, 1)$ where $0 < l < \frac{1}{2}$
- $\sigma : [l, 1) \to [0, T]$ such that $\sigma = 0$ on $[\frac{1}{2}, 1)$, $\sigma > 0$ and strictly decreasing on $[l, \frac{1}{2})$, $\sigma(l) = T$ and $\sigma \in C^2([l, \frac{1}{2}])$ with $\sigma''(x) \geq 0$ for all $x \in [l, \frac{1}{2}]$,
- $Q = \{(x, t) \mid x \in (l, 1),\ t > \sigma(x)\}$,
- $\Omega_t = (\sigma^{-1}(t), 1)$ for all $t \in (0, T)$.

In dimension one it is more convenient to work with $\sigma^{-1} : [0, T] \to [l, \frac{1}{2}]$ rather than σ. We slightly modify the cost functional (P_λ) by multiplying the first integrand by $({\sigma^{-1}}'(t))^2 = 1/(\sigma'(\sigma^{-1}(t)))^2$ and consider the following problem:

$$\left.\begin{array}{l} \text{minimize } J(v, j) = \displaystyle\int_0^T \left| y_x(\sigma^{-1}(t), t) - \rho {\sigma^{-1}}'(t) \right|^2 dt \\ \displaystyle +\frac{1}{\lambda} \int_0^T \left[j(y(1, t)) + j^*(-v(t)) + v(t) y(1, t) \right] dt + \frac{\eta}{2} \int_0^T |v(t)|^2 \, dt \\ \text{over } (v, j) \in L^2(0, T) \times \mathcal{K}^1 \end{array}\right\} \quad (5.10)$$

where y is the solution of

$$\left.\begin{array}{rcl} y_t - y_{xx} & = & 0 \quad \text{in } Q \\ y(\sigma^{-1}(t), t) & = & 0 \quad \text{on } \Sigma_0 \\ \dfrac{\partial y}{\partial \nu}(1, t) & = & v(t) \quad \text{on } \Sigma \\ y(x, 0) & = & y_0(x) \quad \text{on } \Omega_0. \end{array}\right\} \quad (5.11)$$

Here we suppose that $y_0 \in C^1([\frac{1}{2}, 1])$; $y_0(x) < 0$ for all $x > \frac{1}{2}$ and $y_0(\frac{1}{2}) = 0$.

Next we characterize the 'Neumann to Neumann' mapping $v \mapsto y_x(\sigma^{-1}(t), t)$ and the 'Dirichlet to Neumann' mapping $v \mapsto y(1, t)$ explicitly als Volterra integral operators acting on v, and we thus eliminate the constraint (5.11) from the optimization problem (5.10). We preferred this procedure over simply calculating $y_x(\sigma^{-1}(t), t)$ and $y(1, t)$ in (5.11) by a finite element discretization, because in the latter case it turned out that the number of elements in spatial direction and the number of timesteps must be extremly large in order to obtain results of reasonable accuracy. Recall that the fundamental solution for the heat equation is given by

$$,\ (x, t; \xi, \tau) = \frac{1}{2\sqrt{\pi}} (t - \tau)^{-\frac{1}{2}} \exp\left(\frac{(x - \xi)^2}{4(t - \tau)} \right)$$

We set

$$G(x, t; \xi, \tau) = ,\ (x, t; \xi, \tau) - ,\ (2 - x, t; \xi, \tau)$$

and

$$N(x, t; \xi, \tau) = ,\ (x, t; \xi, \tau) + ,\ (2 - x, t; \xi, \tau)$$

and define the operators

$$L_1 f(t) \;=\; \int_0^t N_x(\sigma^{-1}(t), t; \sigma^{-1}(\tau), \tau)\, f(\tau)\, d\tau \tag{5.12}$$

$$L_2 f(t) \;=\; \int_0^t N_x(\sigma^{-1}(t), t; 1, \tau)\, f(\tau)\, d\tau \tag{5.13}$$

$$M_1 f(t) \;=\; \int_0^t N(1, t; 1, \tau)\, f(\tau)\, d\tau \tag{5.14}$$

$$M_2 f(t) \;=\; \int_0^t N(1, t; \sigma^{-1}(\tau), \tau)\, f(\tau)\, d\tau \tag{5.15}$$

and

$$G_1 f(t) \;=\; \int_{\frac{1}{2}}^1 G(\sigma^{-1}(t), t; \xi, 0)\, f(\xi)\, d\xi \tag{5.16}$$

$$G_2 f(t) \;=\; \int_{\frac{1}{2}}^1 N(1, t; \xi, 0)\, f(\xi)\, d\xi \tag{5.17}$$

for $t > 0$. Note that the operators (5.12) and (5.14) are of the form $\int_0^t k(t,\tau)(t-\tau)^{-\frac{1}{2}} f(\tau) d\tau$ and the operators (5.13) and (5.15) are of the form $\int_0^t k(t,\tau) f(\tau) d\tau$ with some function k continuous on $D = \{(t,\tau) : 0 \le t \le T; 0 \le \tau \le t\}$. Following [GLS], proof of Theorem 2.2 (i), p.64, and proof of Theorem 2.5, p.66, it can be proved that (5.12)–(5.15) define compact operators on $L^2(0,T)$. Moreover it is easy to see that G_1 and G_2 in (5.16) and (5.17) respectively are bounded from $L_2(\frac{1}{2}, 1)$ into $L_2(0,T)$. Therefore we can define compact operators $\mathcal{L}$ and $\mathcal{M}$ from $L_2(0,T)$ into itself by

$$\mathcal{L} f = 2(I + 2L_1)^{-1} L_2 f \tag{5.18}$$

$$\mathcal{M} f = (M_1 - M_2 \mathcal{L}) f. \tag{5.19}$$

Moreover we define

$$d_1 = 2(I + 2L_1)^{-1} G_1 \frac{dy_0}{dx} \;\in L^2(0,T), \tag{5.20}$$

$$d_2 = G_2 y_0 - M_2 d_1 \;\in L^2(0,T), \tag{5.21}$$

where I denotes the identity on $L^2(0,T)$. The existence of $(I + 2L_1)^{-1}$ follows from the fact that we can decompose the interval $[0,T]$ into finitely many subintervals such that the norm of the restriction of $2L_1$ to these subintervals is less than 1. The solution $f \in L^2(0,T)$ of

$$(I + 2L_1) f = g \in L^2(0,T)$$

is unique and can be expressed in the form

$$f(t) = g(t) - \int_0^t r(t,\tau)\, g(\tau)\, d\tau,$$

where $r(t,\tau)$ is the resolvent of $2L_1$. (c.f. [GLS], Theorem 3.6, p.234, and Corollary 3.14, p.238). It follows that

$$\mathcal{L}f(t) = 2\int_0^t \Big[k_2(t,\tau) - \int_\tau^t r(t,s)\,k_2(s,\tau)\,ds\Big] f(\tau)d\tau = \int_0^t \kappa(t,\tau)f(\tau)d\tau \quad (5.22)$$

is a Volterra operator of the first kind. Here k_2 denotes the kernel of the operator L_2.

We now address the problem of determining $y_x(\sigma^{-1}(t),t)$ and $y(1,t)$ in (5.11) from known boundary- and initial values.

Proposition 5.1 *Suppose* $y \in H^1(Q)$ *is a variational solution of (5.11) with* $v \in L^2(0,T)$. *Then*

$$y_x(\sigma^{-1}(t),t) = \mathcal{L}v(t) + d_1(t) \quad (5.23)$$

and

$$y(1,t) = \mathcal{M}v(t) + d_2(t) \quad (5.24)$$

a.e. on $[0,T]$.

Proof. It is known ([KMP], Theorem 2.4) that (5.23) and (5.24) hold for $v \in C([0,T])$. A density argument together with (3.11) implies the claim for all $v \in L^2(0,1)$.

Proposition 5.1 allows to write the cost functional in (5.10) as

$$\begin{aligned} J(v,j) = &\int_0^T \Big|\mathcal{L}v(t) - d(t)\Big|^2 dt + \frac{\eta}{2}\int_0^T |v(t)|^2\,dt \\ &+\frac{1}{\lambda}\int_0^T \Big[j\Big(\mathcal{M}v(t)+d_2(t)\Big) + j^*(-v(t)) + v(t)\Big(\mathcal{M}v(t)+d_2(t)\Big)\Big]\,dt \end{aligned} \quad (5.25)$$

where $d(t) = d_1 - \rho\,\sigma^{-1\prime}(t)$. For the minimization of (5.25) we used an iterative (SQP) method from the MATLAB optimization toolbox with analytically provided gradients. We chose I in $\mathcal{K}^1$ as $I = (-M, 0)$, where $-M < \min\{y_0(x) : x \in [\frac{1}{2}, 1]\}$. We provided for the fact that $y(1,t) \in I$ may not be satisfied during the iteration by extending j outside of I as a linear function with very steep slope and such that j is convex on all of $\mathbb{R}$. Moreover, just as in in Section 4, we replaced the independent variable j by $\vartheta = j''$, where j and ϑ are related by

$$j(y) = \int_0^y (y-r)\vartheta(r)\,dr \quad \text{for } y \in I, \quad (5.26)$$

and $\vartheta \in \mathcal{K}_2 = \{\vartheta \in L^2(I) : 2\omega_0 \le \vartheta \le 2\omega_1\}$. We next give the explicit form of the gradients. Due to the fact that we use the boundary element formulation for the solution y on Σ_0 and Σ the gradient of J with respect to v is simple, and the adjoint equation (compare (4.6)) is realized through the adjoints of the operators $\mathcal{L}$ and $\mathcal{M}$. We find for the derivative of J with respect to v in direction $w \in L^2(0,T)$:

$$\begin{aligned}\Big\langle \frac{\partial}{\partial v} J(v,j), w\Big\rangle_{L^2(0,T)} &= \langle 2\mathcal{L}^*(\mathcal{L}v - d) + \eta v, w\rangle_{L^2(0,T)} \\ &+ \frac{1}{\lambda}\langle \mathcal{M}^* j'(\mathcal{M}v + d_2) - j^{*\prime}(-v) + (\mathcal{M}+\mathcal{M}^*)v + d_2, w\rangle_{L^2(0,T)}\end{aligned} \tag{5.27}$$

To calculate the derivate of $\vartheta \to J(v, j(\vartheta))$ one proceeds as in the computation of the necessary optimality condition of subproblem 2 of Section 4. We find for the derivative in direction $\xi \in L^\infty(-M,0)$:

$$\begin{aligned}\Big\langle \frac{\partial}{\partial \vartheta} J(v,j(\vartheta)), \xi\Big\rangle_{L^2(-M,0)} \;=\; \frac{1}{\lambda}\int_0^T \Big[&\int_0^{y(1,t)} \Big(y(1,t) - r\Big)\,\xi(r)\,dr \\ &+ \int_0^{\gamma_\vartheta(-v(t))} \Big(r - \gamma_\vartheta(-v(t))\Big)\,\xi(r)\,dr\Big]\,dt,\end{aligned}$$

where

$$\gamma_\vartheta(p) = \begin{cases} 0 \quad \text{if } p > 0 \\ \beta^{-1}(p) \quad \text{if } p \in I \\ \gamma_\vartheta(-M) \quad \text{if } p < \beta(-M), \end{cases}$$

and

$$\beta(r) = \int_0^r \vartheta(s)\,ds \quad \text{for } r \in T,$$

and we assumed that $y(1,t) \in I$ for all $t \in [0,T]$. (If $y(1,t) \notin I$, then j is extended outside of I as explained above, and ξ is set 0 on the complement of I.) For the numerical realization all t-depending functions were discretized as sums of linear spline functions with respect to the grid $\{i\frac{T}{m}\}_{i=0}^m$. The Volterra operators that need to be discretized have either continuous kernels ((5.13) and (5.15)) or weakly singular kernels with singularity of the type $(t-\tau)^{-\frac{1}{2}}$ ((5.12) and (5.14)). In the first case the trapezoidal rule was used to evaluate the integrals and for the singular case we used first order Gauss integration with weight function $(t-\tau)^{-\frac{1}{2}}$ on each of the subintervals. The function ϑ is discretized as sum of elementary step functions

$$B_i^n(y) = \begin{cases} 1 \quad \text{for } -\dfrac{i}{n}M < y \le -\dfrac{i-1}{n}M \\ 0 \quad \text{otherwise,} \end{cases}$$

for $i = 1, \cdots, n$. If $\vartheta^n(y) = \Sigma_{i=1}^n \vartheta_i B_i^n(y)$ with $\vartheta_i \in \mathbb{R}$ and j^n is calculated from ϑ^n via (5.26) then $(j^n)'(0) = 0$. The condition $2\omega_0 \le j'' \le 2\omega_1$ is clearly equivalent to

$$2w_0 \le \vartheta_i \le 2w_1 \quad \text{for all } i = 1, \cdots, n,$$

and can easily be realized as a box constraint in the computations. In numerical experiments the constraint $\vartheta_i \le 2\omega_1$ plays no significant role. For the constraint

$2\omega_0 \le \vartheta_i$ we generally took $\omega_0 = 0$ or ω_0 very small. Especially for the nonattainable case this constraint is essential. Here we call the problem (P) attainable, if there exists $(\beta, y) \in \mathcal{A} \times H^1(Q)$ such that the value of the cost functional is zero if $\eta = 0$. Minimizing $J(v, j(\vartheta))$ involves solving the first-kind Volterra equation $\mathcal{L}v = d$ in a least squares sense. Since $\mathcal{L}$ is compact on $L^2(0,T)$ this is an illposed problem and requires regularization.

For that purpose we already included the Tikhonov regularization term $\frac{\eta}{2}|v|^2_{L^2}$ in the cost functional. Alternatively we used $\frac{\eta}{2}|v'|^2_{L^2}$ as a regularization term, but this did not change the numerical results significantly. In the case of noisy data, Tikhonov regularization, however, did not produce completely satisfactory results. We therefore combined the Tikhonov regularization terms with a method that is suitable specifically for Volterra problems. A frequently used method in this context, sometimes refered to as sequential regularization method, is due to Beck (c.f. [BBC], [La]). In every time-step $t_j \in \{t_0, t_1, \cdots, t_{m+1-(r-1)}\}$, with $r > 1$, the coefficient v_j is chosen such that a constant continuation of v with value v_j fits best the data for the next $r-1$ time-steps in the least square sense. The normal equation of this least squares problem has the form

$$\tilde{\mathcal{L}}^{m,r} v^{m,r} = \tilde{d}^{m,r},$$

where $\tilde{\mathcal{L}}^{m,r}$ is a well conditioned lower triangular matrix and $\tilde{d}^{m,r}$ can be derived from the discretization of the given data function d; for details we refer to [BBC], [La]. When using this method in the context of our problem we replace the first term in the discretized cost functional by

$$\int_0^T \left|\tilde{\mathcal{L}}^{m,r} v^{m,r} - \tilde{d}^{m,r}\right|^2 dt.$$

Note that Beck's method uses information from $r-1$ future time-steps and hence, if the original data vector has dimension $m+1$, then the vector $v^{m,r}$ has dimension $m+2-r$, and consequently v is only defined on the subinterval $[0, T^{m,r}]$, with $T^{m,r} = T(1 - \frac{r-1}{m})$, of $[0,T]$. In the first two experiments we considered cases where the prescribed boundary σ is attainable by a boundary heat transfer law of the form $\frac{\partial y}{\partial \nu} = -\beta(y)$ at $x = 1$, i.e. there exist 'true' functions $\overline{v}$ and $\overline{j}$, such that $J(\overline{v}, \overline{j}) = 0$, if $\eta = 0$ in the cost functional (5.25). In order to obtain data for such a problem we solved the forward *Stefan* problem

$$\left.\begin{array}{rcl} y_t &=& y_{xx} \quad \text{on } Q \\ y(\sigma^{-1}(t), t) &=& 0 \quad \text{on } (0,T) \\ y_x(1,t) &=& \overline{v}(t) \quad \text{on } (0,T) \\ y_x(\sigma^{-1}(t), t) &=& \rho\, \sigma^{-1\prime}(t) \quad \text{on } (0,T) \\ y(x,0) &=& y_0(x) \end{array}\right\} \qquad (5.28)$$

with some fixed, monotonically decreasing function $\overline{v}$ and *unknown* boundary σ. We specifically chose

$$y_0(x) = 4\Big(x - \frac{1}{2}\Big)\Big(x - \frac{3}{2}\Big), \qquad T = \frac{1}{2} \qquad \text{and} \quad \rho = \frac{3}{2},$$

and

$$\overline{v}(t) = 0.2\exp(-4t).$$

Let $(\sigma, \overline{y})$ denote the corresponding solution of (5.28). It can be seen that $\overline{y}(1,t)$ is monotonically increasing on $(0,T)$ and hence $\overline{v}$ and $\overline{y}(1,\cdot)$ are related via some monotonically increasing function $\overline{\beta} = \overline{j}' : [-1, \overline{y}(1,T)] \rightarrow [\overline{v}(T), 0.2]$ via

$$\overline{v}(t) + \overline{\beta}(\overline{y}(1,t)) = 0, \quad \text{for } t \in [0,T].$$

Since $\overline{y}_x(\sigma(t),t) - \rho{\sigma^{-1}}'(t) = 0$ for all t it follows that $J(\overline{v}, \overline{j}) = 0$, if one uses as free boundary the function σ which was calculated as solution of (5.28), and if $\eta = 0$. For this σ as input data we solved the regularized problem:

$$\begin{array}{l} \text{minimize } J(v, j(\vartheta)) \\ \text{subject to } \vartheta \geq 0, \end{array} \tag{5.29}$$

with the following set of parameters:

$$\eta = 10^{-7},\ \lambda = 10^4,\ n = 64\ (t\text{-discretization}),\ m = 48\ (\text{discretization for } j).$$

Let (v_{opt}, β_{opt}) indicate the calculated solution of (5.29). Figure 2 compares v_{opt} and $\overline{v}$, Figure 3 shows β_{opt} and $\overline{\beta} = \int_0^r \overline{\vartheta}(s)ds$. In Figure 4 we plotted $\beta_{opt}(y_{opt}(1,\cdot))$ and $-v_{opt}$ as functions of t. This plot shows how well the boundary condition $v(t) + \beta(y(1,t)) = 0$ is fulfilled for the calculated solution. Note that $y(1,t) = \mathcal{M}v(t) + d_2(t)$ occurs as a by-product in the calculation of the cost functional.

The plots in Figures 5–7 are analogous to those of Figures 2–4, but here we added some uniformly distributed noise to the input data $\sigma^{-1}(t.)$ The noise-level was chosen 0.25% with respect to the $|\cdot|_\infty$-norm. This may seem to be a small noise level, but the derivative of $\sigma^{-1}(t)$ occurs at a prominent place in the cost functional in the data vector d. To obtain d, we had to carry out a numerical differentiation resulting in an error on d of about 18% in the $|\cdot|_\infty$-norm. In this noisy data case, we used Beck's future regularization method as describe before with the following parameters:

$$\eta = 5 \cdot 10^{-4},\ \ \lambda = 1,\ \ n = 128,\ \ m = 32,\ \ r = 16\ (\text{number of future time-steps}).$$

By the use of future information, the solution v_{opt} is only defined on the interval $[0, \frac{1}{2} - \frac{r-1}{m}] = [0, 0.4414]$.

For the last two plots in Figures 8 and 9 we chose an 'arbitrary' function σ for the free boundary, which cannot be attained exactly by a boundary *control* of the given type. This fact is reflected in the fact that for the calculated optimal solution, the constraints on the monotonicity of β are active. The intervals of constant function values in Figure 8 are due to this fact. Figure 9 shows how well the prescribed $(\sigma^{-1})'$ can be approximated by the optimal pair (v_{opt}, β_{opt}).

For comparison, we also plotted the free boundary corresponding to the pair of starting values $(v_0(t), \beta_0(y)) = (0.2, 0.2y)$ at the beginning of the optimization. Here we used again ordinary Tikhonov regularization (no future regularization) with the following parameters:

$$\eta = 10^{-5}, \quad \lambda = 10^{-2}, \quad n = 64, \quad m = 48.$$

References

[B1] V. Barbu: Analysis and Control of Nonlinear Infinite Dimensional Systems, Academic Press, New York, 1993.

[B2] V. Barbu: The approximate solvability of the inverse one phase Stefan problem, Internat. Series Numer. Math., Vol. 99, 33–43, Birkhäuser Verlag, Basel 1991.

[BBC] J. V. Beck, B. Blackwell, and C. Clair: Inverse Heat Conduction, Wiley-Interscience, 1985.

[BDZ] V. Barbu, G. Da Prato, J. P. Zolésio: Feedback controllability of the free boundary of the one phase Stefan problem, Diff. Integral Equs. 4 (1991), 225–239.

[BK1] V. Barbu, K. Kunisch: Identification of nonlinear elliptic equations (to appear).

[BK2] V. Barbu, K. Kunisch: Identification of nonlinear parabolic equations (to appear).

[Br1] H. Brézis: Problèmes unilatéraux, J. Math. Pures Appl. 51 (1972), 1–64.

[Br2] H. Brézis: Analyse Fonctionnelle, Dunod, Paris, 1983.

[C] J. R. Cannon: The One-Dimensional Heat Equation, Addison-Weseley Publ. Comp., Menlo Park, 1984.

[DZ] G. Da Prato, J. P. Zolésio: An optimal control problem for a parabolic equation in noncylindrical domains, Systems & Control Letters 11 (1988), 73–77.

[GLS] Gripenberg, S. Londen, and Staffans: Volterra Integral and Functional Equations, Cambridge University Press, 1990.

[HN] K. H. Hoffmann, M. Niezgodka: Control of evolutionary free boundary problems, in Free Boundary Problems: Theory and Applications, pp. 439–450, K. H. Hoffmann and J. Sprekels, eds., Pitman Research Notes in Mathematics 186, Longman, London (1990).

[HS] K. H. Hoffmann, J. Sprekels: Real time control of the free boundary in a two phase Stefan problem, Numerical Functional Anal. Optimiz. 5 (1982), 47–76.

[KMP] K. Kunisch, K. Murphy, and G. Peichl: Estimation of the conductivity in the one-phase Stefan problem: Basic results, to appear in Bolletino Unione Mat. Italiana.

[La] P. K. Lamm: Future-sequential regularization methods for ill-posed Volterra equations, to appear in J. Math. Anal. and Appl.

[Li] J. L. Lions: Quelques Méthodes de Résolutions de Problèmes aux Limites Nonlinéaires, Dunod, Paris, 1969.

[PW] H. Proter, H. Weinberger: The Maximum Principle, Springer-Verlag, Berlin, 1983.

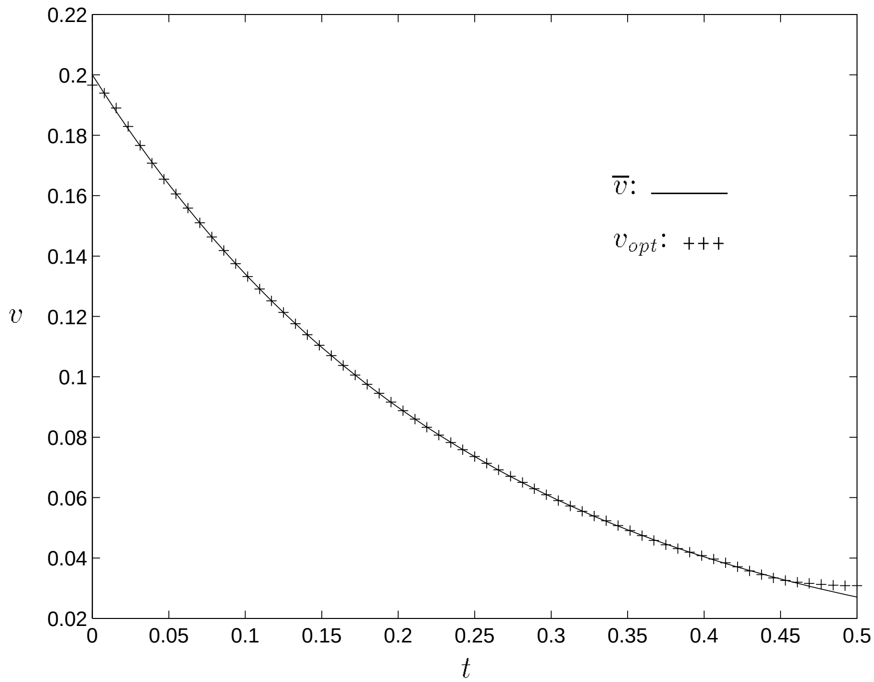

Figure 2: Comparison of true and calculated boundary heat transfer function $v(t) = \frac{\partial y}{\partial \nu}(1, t)$ (no data-noise)

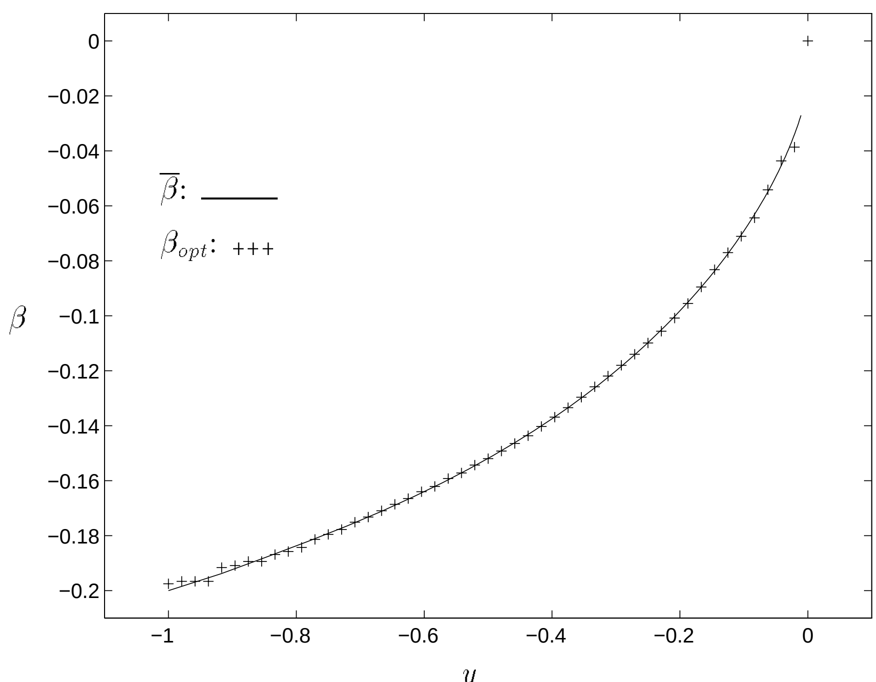

Figure 3: Comparison of true and calculated feedback law β (no data-noise)

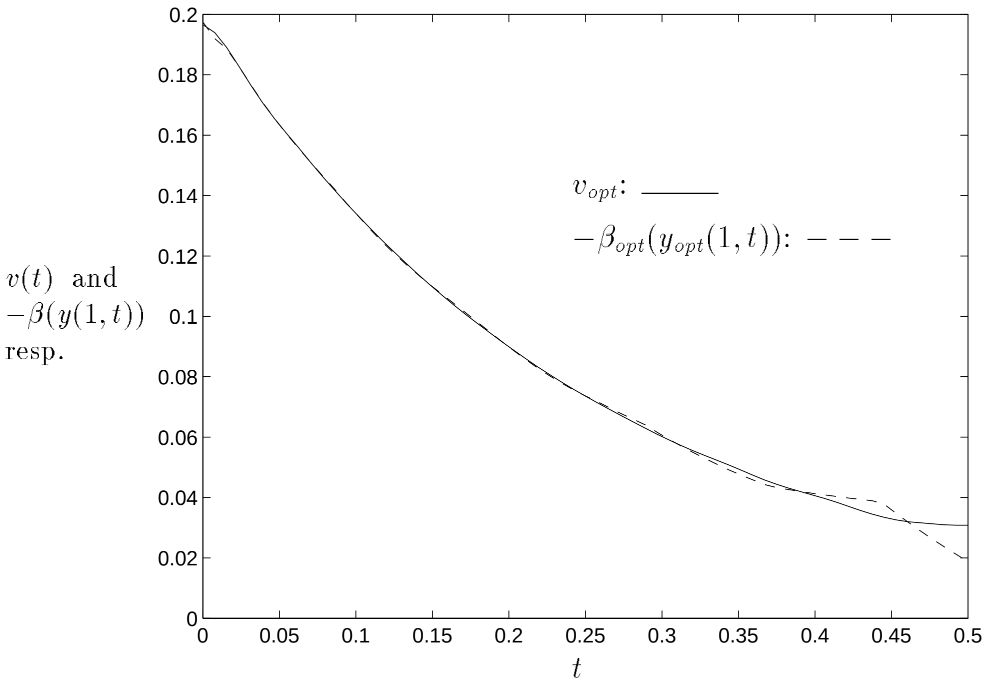

Figure 4: How well is the feedback law fulfilled for the calculated solution? (no data-noise)

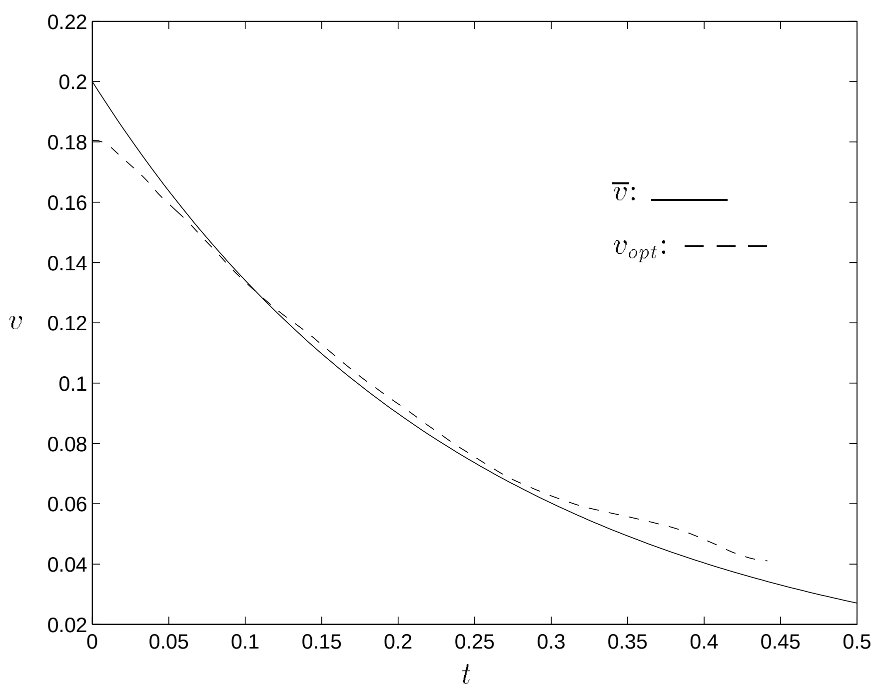

Figure 5: Comparison of true and calculated boundary heat transfer function $v(t) = \frac{\partial y}{\partial \nu}(1,t)$ (noisy data)

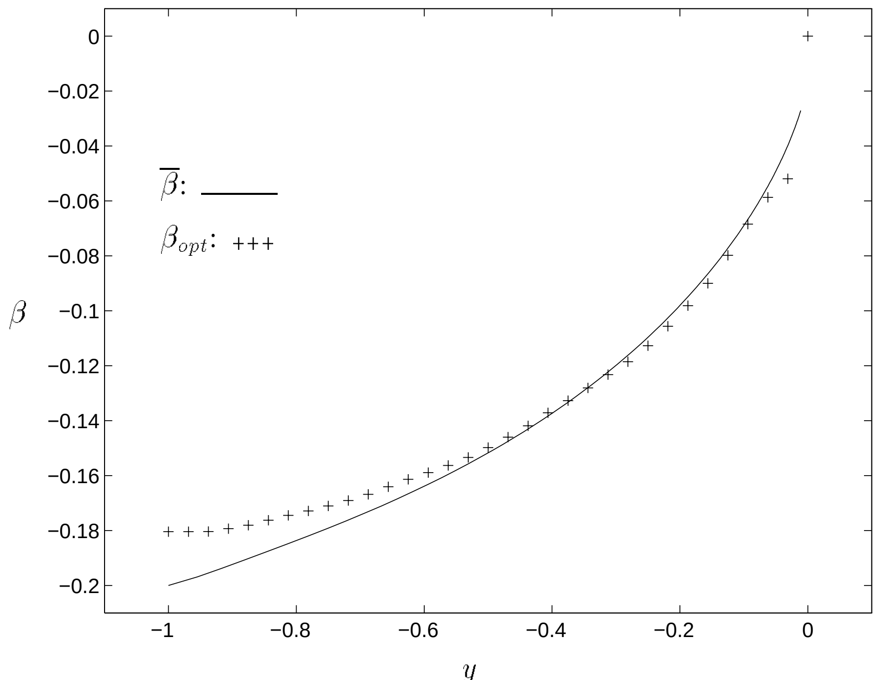

Figure 6: Comparison of true and calculated feedback law β (noisy data)

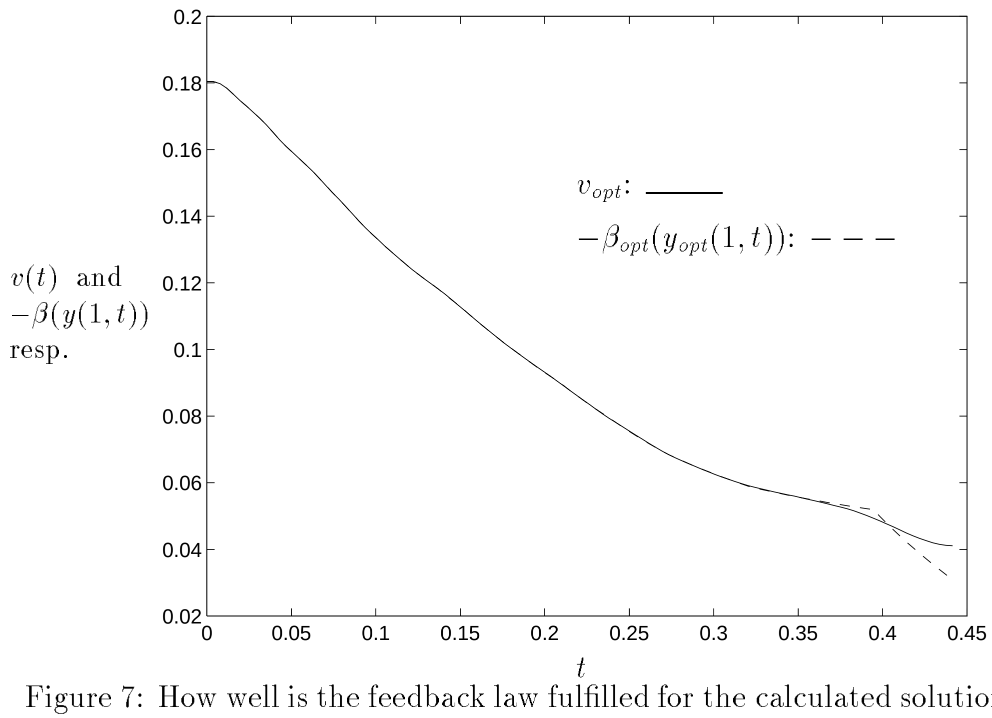

Figure 7: How well is the feedback law fulfilled for the calculated solution? (noisy data)

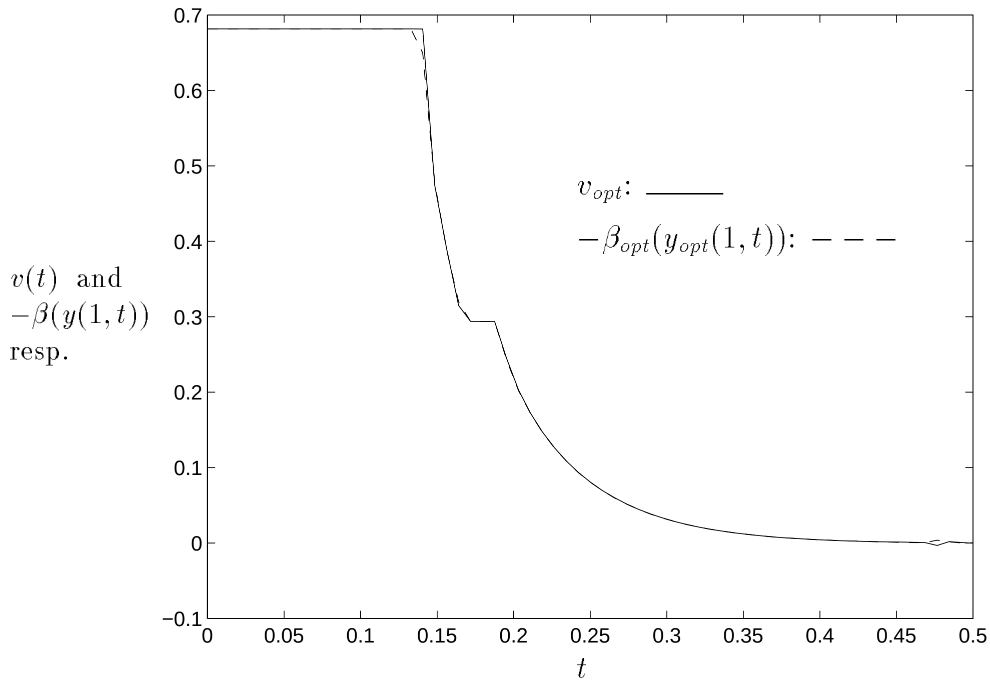

Figure 8: How well is the feedback law fulfilled for the calculated solution? (non-attainable data)

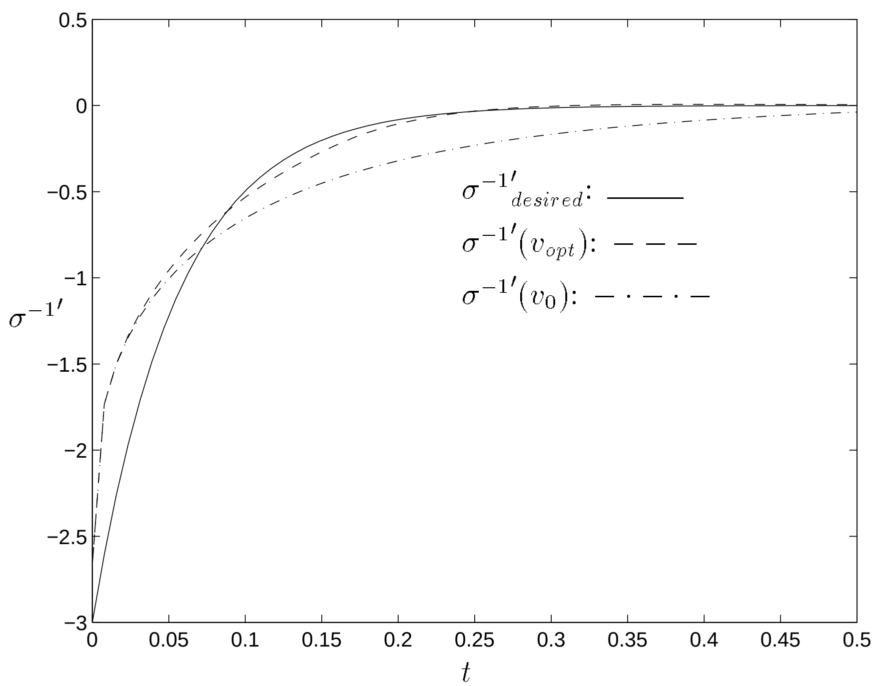

Figure 9: Comparison of desired, optimal, and initial values for $\sigma^{-1\prime}$ (non-attainable data)