On the switching behavior of sparse optimal controls for the one-dimensional heat equation

Fredi Tröltzsch and Daniel Wachsmuth

December 4, 2017

Fredi Tröltzsch *

Institut für Mathematik
Technische Universität Berlin
D-10623 Berlin, Germany

Daniel Wachsmuth

Institut für Mathematik
Universität Würzburg
D-97974 Würzburg, Germany

Dedicated to Prof. Dr. Eduardo Casas on the occasion of his 60th birthday

Abstract

An optimal boundary control problem for the one-dimensional heat equation is considered. The objective functional includes a standard quadratic terminal observation, a Tikhonov regularization term with regularization parameter ν, and the L^1-norm of the control that accounts for sparsity. The switching structure of the optimal control is discussed for $\nu \geq 0$. Under natural assumptions, it is shown that the set of switching points of the optimal control is countable with the final time as only possible accumulation point. The convergence of switching points is investigated for $\nu \searrow 0$.

1 Introduction

In this paper, we investigate the switching behavior of optimal controls for the following sparse optimal control problem with terminal observation in $\Omega = (0,1)$:

$$\min J(y,u) := \frac{1}{2} \int_\Omega |y(x,T) - y_\Omega(x)|^2 \, \mathrm{d}x + \frac{\nu}{2} \int_0^T |u(t)|^2 \, \mathrm{d}t + \mu \int_0^T |u(t)| \, \mathrm{d}t \tag{1.1}$$

subject to the parabolic initial-boundary value problem

$$\begin{aligned}
y_t - y_{xx} &= 0 && \text{in } (0,1) \times (0,T) \\
y_x(0,t) &= 0 && \text{in } (0,T) \\
y_x(1,t) + \alpha\, y(1,t) &= u(t) && \text{in } (0,T) \\
y(x,0) &= 0 && \text{in } (0,1)
\end{aligned} \tag{1.2}$$

and to the pointwise control constraints

$$a \le u(t) \le b, \quad \text{a.e. in } (0,T). \tag{1.3}$$

Bang-bang and switching properties for the solutions of optimal boundary control problems were extensively discussed from the 1960ties on. If $\nu = \mu = 0$, then it is well known that the optimal control is of bang-bang type provided that y_Ω is not attained by the optimal state. This result was discussed in several papers for linear parabolic equations, see [8, 13, 24], cf. also the monographies [15, 10, 18, 29]. For the case of the maximum norm as objective functional, the bang-bang principle was proved in [35, 12, 14]. The finite bang-bang-principle for supremum-norm problems was first proved in [16] and later in a different way in [11]. Bang-bang principles for nonlinear parabolic equations were discussed in [22, 25].

For $\nu > 0$ but $\mu = 0$, the switching behavior of optimal controls was investigated in [7, 28]. In particular, the convergence of switching points for $\nu \to 0$ was addressed. Numerical examples and numerical methods exploiting the switching structure were presented in [6, 19, 23]. Bang-bang properties for time-optimal parabolic boundary control problems were studied, e.g., in [8, 9, 20, 33, 17, 34]. This list of references on bang-bang principles and switching properties is by no means exhaustive. We also refer to the references of these papers.

The main novelty of our paper is the discussion of the switching structure for sparse optimal controls of parabolic boundary control problems (i.e., for the case $\mu > 0$). To our best knowledge, the switching

properties of sparse optimal boundary controls for parabolic problems were not yet discussed in the literature. In particular, this refers to the convergence of switching points for the limit $\nu \to 0$. In addition to proving convergence of switching points, we obtain also convergence rates with respect to ν for the approximation of switching points for $\nu \searrow 0$.

However, the general bang-bang structure of optimal sparse controls has already been investigated in a sequence of papers on semilinear elliptic control problems. We refer to [2]. Our paper was inspired by these general results.

Assumption 1 (Data)**.** In this setting, real numbers $T > 0$, $\nu \geq 0$, $\mu > 0$, $\alpha > 0$, and $a < 0 < b$ are fixed. The sign restrictions on a and b are needed only for some of the structural properties of the optimal control, neither for the existence of optimal controls nor for the necessary optimality conditions. The parameter μ is the so called sparse parameter.

Moreover, we fix a desired final state function $y_\Omega \in C[0,1]$. We require $y_\Omega \in C[0,1]$ in order to have continuity of the adjoint state up to the boundary. (For the optimality conditions, $y_\Omega \in L^2(0,1)$ would suffice.)

2 Well-posedness of the problem and necessary optimality conditions

2.1 Well-posedness of the problem

The linear initial-boundary value problem (1.1) is well posed. For each function $u \in L^2(0,T)$, there exists a unique solution $y \in W(0,T) \cap C(\bar{Q})$, where $Q := (0,1) \times (0,T)$ denotes the open space-time cylinder and $W(0,T) = \{y \in L^2(0,T; H^1(0,1)) : y' \in L^2(0,T; H^1(0,1)^*)\}$.

Let us denote the state y associated with u by y_u. The control-to-state mapping $u \mapsto y_u$ is linear and continuous from $L^2(0,T)$ to $W(0,T) \cap C(\bar{Q})$, cf. [1] or [29].

Let us introduce the functionals $j : L^1(0,T) \to \mathbb{R}$,

$$j(u) := \int_0^T |u(t)| \, \mathrm{d}t = \|u\|_{L^1(0,T)}$$

and $f\nu$,

$$f_\nu(u) := \frac{1}{2} \int_\Omega |y_u(x,T) - y_\Omega(x)|^2 \, \mathrm{d}x + \frac{\nu}{2} \int_0^T |u(t)|^2 \, \mathrm{d}t.$$

Then the reduced objective functional F is given by

$$F_\nu(u) := J(y_u, u) = f_\nu(u) + \mu\, j(u).$$

Introducing the set of admissible controls by

$$U_{ad} := \{u \in L^2(0,T) : a \le u(t) \le b, \ \text{ for a.a. } t \in (0,T)\},$$

we can re-write the optimal control problem (1.1)–(1.3) in the short form

$$\min_{u \in U_{ad}} F_\nu(u). \tag{P_ν}$$

The functional F_ν is continuous and convex, hence weakly lower semi-continuous. Moreover, the set U_{ad} is weakly compact and non-empty. Therefore, there exists at least one optimal control of the problem (P_ν), which will be denoted by u_ν to indicate the correspondence to the Tikhonov parameter ν. By $y_\nu := y_{u_\nu}$ we denote the optimal state associated with u_ν. In the case $\nu = 0$ we will drop the index 0 and write $\bar{u} := u_0$ and $\bar{y} := y_0$.

If $\nu > 0$, then F_ν is strictly convex and hence in this case the optimal control is unique. Under a natural assumption, we show later that this uniqueness also holds for $\nu = 0$. In addition, in the case $\nu = 0$ the optimal state is uniquely determined due to the strict convexity of f_ν with respect to y_u.

2.2 Fourier expansion for (1.2), Green's function

For convenience of the reader, we recall some known facts on the representation of the weak solution by a Green's function G. We consider the inhomogeneous initial-boundary value problem

$$\begin{aligned}
y_t(x,t) - y_{xx}(x,t) &= f(x,t) & \text{in } Q \\
y_x(0,t) &= 0 & \text{in } (0,T) \\
y_x(1,t) + \alpha\, y(1,t) &= u(t) & \text{in } (0,T) \\
y(x,0) &= y_0(x) & \text{in } (0,1),
\end{aligned} \tag{2.1}$$

where $f \in L^2(Q)$, $y_0 \in L^2(0,1)$, and $u \in L^2(0,T)$ are given. Below, we briefly address the case $\alpha = 0$, but later we will concentrate on positive α. The weak solution of (2.1) can be represented by a Green's function

4

$G = G(x, \xi, t)$ as

$$
\begin{aligned}
y(x, t) &= \int_0^1 G(x, \xi, t)\, y_0(\xi)\, \mathrm{d}\xi + \int_0^t \int_0^1 G(x, \xi, t - s)\, f(\xi, s)\, \mathrm{d}\xi\, \mathrm{d}s \\
&\quad + \int_0^t G(x, 1, t - s)\, u(s)\, \mathrm{d}s,
\end{aligned}
$$

$$(2.2)$$

where G is given by the following Fourier expansions:

$$
G(x, \xi, t) =
\begin{cases}
1 + 2 \displaystyle\sum_{n=1}^{\infty} \cos(n\pi x) \cos(n\pi\xi) \exp(-n^2\pi^2 t) & \text{for } \alpha = 0 \\
\displaystyle\sum_{n=1}^{\infty} \frac{1}{N_n} \cos(\rho_n x) \cos(\rho_n \xi) \exp(-\rho_n^2 t) & \text{for } \alpha > 0.
\end{cases}
$$

$$(2.3)$$

Here, (ρ_n) is the monotone increasing sequence of non-negative solutions to the equation

$$
\rho \tan \rho = \alpha
$$

and

$$
N_n = \int_0^1 \cos^2(\rho_n x)\, \mathrm{d}x = \frac{1}{2} + \frac{\sin(2\rho_n)}{4\rho_n} = \frac{1}{2} + \frac{\sin^2(\rho_n)}{2\alpha}
$$

are normalizing constants. The numbers $n\pi$ and ρ_n are the eigenvalues of the differential operator $\partial^2/\partial x^2$ subject to the homogeneous boundary conditions in (2.1) for $\alpha = 0$ and $\alpha > 0$, respectively. For the eigenvalues, we know that $\rho_n \sim (n-1)\pi$, $n \to \infty$. The functions $x \mapsto \cos(n\pi x)$ and $x \mapsto \cos(\rho_n x)$ are associated eigenfunctions, respectively. After normalization by the factors N_n, they form a complete orthonormal system in $L^2(0, 1)$; cf. [30].

Notice that, for our case $f = 0$ and $y_0 = 0$, the term $y(x, T)$ in the objective functional has the series representation

$$
y(x, T) = \sum_{n=1}^{\infty} \cos(\rho_n x) \frac{\cos(\rho_n)}{N_n} \int_0^T e^{-\rho_n^2 (T - s)} u(s)\, ds. \qquad (2.4)
$$

3 Necessary optimality conditions

3.1 The variational inequality

It is well known that the derivative of the differentiable functional f_ν can be represented in the form

$$
f_\nu'(u) v = \int_0^T (\varphi_u(1, t) + \nu\, u(t))\, v(t)\, \mathrm{d}t,
$$

5

where φ_u is the adjoint state associated with u. It is the unique weak solution to the adjoint equation

$$\begin{array}{rcll}
-\varphi_t - \varphi_{xx} & = & 0 & \text{in } Q \\
\varphi_x(0,t) & = & 0 & \text{in } (0,T) \\
\varphi_x(1,t) + \alpha\,\varphi(1,t) & = & 0 & \text{in } (0,T) \\
\varphi(x,T) & = & y_u(x,T) - y_\Omega(x) & \text{in } (0,1).
\end{array} \tag{3.1}$$

Notice that y_Ω is assumed to be continuous. Moreover, the function $x \mapsto y_u(x,T)$ is also continuous, because $u \in L^\infty(0,T)$. Therefore, we have $\varphi_u \in C(\bar{Q})$ and the continuity of the function $t \mapsto \varphi_u(1,t)$ on $[0,T]$.

Theorem 1 (Necessary optimality condition). *Let $\nu \geq 0$ be given, and let $u_\nu \in U_{ad}$ be optimal for the problem (P_ν). Then there exists a function $\lambda_\nu \in \partial j(u_\nu)$, such that the variational inequality*

$$\int_0^T (\varphi_\nu(1,t) + \nu u_\nu(t) + \mu\,\lambda_\nu(t))(u(t) - u_\nu(t))\,\mathrm{d}t \geq 0 \quad \forall u \in U_{ad}. \tag{3.2}$$

is satisfied with the adjoint state $\varphi_\nu := \varphi_{u_\nu}$.

This result is completely standard for the case $\mu = 0$, where F_ν is smooth, see [18]. If $\mu > 0$, the associated tools from subdifferential calculus can be found, for instance in [3, Thm. 3.1]. It is fairly obvious how these methods can be transferred to our problem. Therefore, we omit the proof. An equivalent formulation is obtained by replacing the subdifferential in (3.2) by directional derivatives, which is

$$\int_0^T (\varphi_\nu(1,t) + \nu u_\nu(t))(u(t) - u_\nu(t))\,\mathrm{d}t + \mu j'(u_\nu;\,u - u_\nu) \geq 0 \quad \forall u \in U_{ad}. \tag{3.3}$$

By a standard argument, we find that this variational inequality of integral type implies the following pointwise inequality:

$$(\varphi_\nu(1,t) + \nu u_\nu(t))(v - u_\nu(t)) + \mu j'(u_\nu(t); v - u_\nu(t)) \geq 0 \quad \forall a \leq v \leq b \tag{3.4}$$

for a.a. $t \in (0,T)$. Here, we denote by $j'(u,h)$ the directional derivative of the real function $j(u) := |u|$ at $u \in \mathbb{R}$ in the direction $h \in \mathbb{R}$.

3.2 The case $\nu = 0$, bang-bang-bang properties

In the case $\nu = 0$, a detailed discussion of the variational inequality (3.2) leads to nice structural properties of the optimal control $\bar{u}$. We

recall that for $\nu = 0$ the optimal control and the optimal state are denoted by $\bar{u}$ and $\bar{y}$, respectively.

Associated with $\bar{u}$, we introduce the following measurable sets

$$
\begin{aligned}
E_+ &= \{t \in (0,T) : \bar{u}(t) > 0\}, \\
E_0 &= \{t \in (0,T) : \bar{u}(t) = 0\}, \\
E_- &= \{t \in (0,T) : \bar{u}(t) < 0\}.
\end{aligned}
$$

The pointwise discussion of the variational inequality (3.2) yields the following result.

Lemma 1. *If $\nu = 0$, then for almost all $t \in (0,T)$, the following implications hold true:*

$$
t \in \begin{cases}
E_+ &\Rightarrow \bar{\varphi}(1,t) + \mu \leq 0, \\
E_- &\Rightarrow \bar{\varphi}(1,t) - \mu \geq 0, \\
E_0 &\Rightarrow \bar{\varphi}(1,t) + \mu\,\bar{\lambda}(t) = 0.
\end{cases}
\tag{3.5}
$$

Proof. For a.a. $t \in E_+$, we have $\bar{u}(t) > 0$, hence $\bar{\lambda}(t) = 1$. Therefore (3.2) almost everywhere implies

$$
\bar{\varphi}(1,t) + \mu\,\bar{\lambda}(t) = \bar{\varphi}(1,t) + \mu \leq 0.
$$

On E_-, the discussion is analogous. In E_0, we have a.e. $a < \bar{u}(t) < b$, hence the reduced gradient must vanish here, $\bar{\varphi}(1,t) + \mu\,\bar{\lambda}(t) = 0$ a.e. on E_0. $\qquad\square$

This discussion showed how the function $t \mapsto \bar{\varphi}(1,t) + \mu\,\bar{\lambda}(t)$ depends on the sign of $\bar{u}$. Another investigation will reveal the switching structure of $\bar{u}$ related to the function $t \mapsto \bar{\varphi}(1,t) + \mu\,\bar{\lambda}(t)$. To this end let us define the open sets

$$
\begin{aligned}
\Phi_+ &= \{t \in (0,T) : \bar{\varphi}(1,t) - \mu > 0\}, \\
\Phi_0 &= \{t \in (0,T) : |\bar{\varphi}(1,t)| < \mu\}, \\
\Phi_- &= \{t \in (0,T) : \bar{\varphi}(1,t) + \mu < 0\}.
\end{aligned}
$$

Lemma 2. *If $\nu = 0$, then for almost all $t \in (0,T)$, the following implications hold true:*

$$
t \in \begin{cases}
\Phi_+ &\Rightarrow \bar{u}(t) = a, \\
\Phi_- &\Rightarrow \bar{u}(t) = b, \\
\Phi_0 &\Rightarrow \bar{u}(t) = 0.
\end{cases}
\tag{3.6}
$$

Proof. We first prove the claim for Φ_+. The equation $\bar{u}(t) = a$ is obtained as follows: We have $|\bar{\lambda}(t)| \leq 1$, hence

$$0 < \bar{\varphi}(1,t) - \mu \leq \bar{\varphi}(1,t) + \mu\,\bar{\lambda}(t)$$

is satisfied a.e. in Φ_+. Now the variational inequality (3.2) implies $\bar{u}(t) = a$ a.e. in Φ_+. The continuity of the function $t \mapsto \bar{\varphi}(1,t)$ yields that Φ_+ is an open set. The proof for Φ_- is analogous. The statement of Lemma 1 shows that $\Phi_0 \cap (E_+ \cup E_-)$ has measure zero. Hence it holds $\bar{u} = 0$ almost everywhere on Φ_0. $\qquad\square$

3.3 Switching points of $\bar{u}$

For $\nu = 0$, the switching behavior of the optimal control depends on the solutions of the two equations

$$\bar{\varphi}(1,t) + \mu = 0 \quad \text{and} \quad \bar{\varphi}(1,t) - \mu = 0, \quad t \in [0,T].$$

To estimate their number, we need the following result:

Lemma 3. *The function $t \mapsto \bar{\varphi}(1,t)$ is continuous in $[0,T]$. It can be extended to a holomorphic function in the complex half plane $\{z \in \mathbb{C} : \Re(z) < T\}$.*

Proof. The continuity of the function $t \mapsto \bar{\varphi}(1,t)$ follows from $\bar{y} \in C(\bar{Q})$ and $y_\Omega \in C[0,1]$. Therefore, the terminal data for $\bar{\varphi}(T)$ are continuous in $[0,1]$, hence $\varphi \in C(\bar{Q})$ and the function $t \mapsto \bar{\varphi}(1,t)$ is continuous. We refer to [1, 21], and [29, Thm. 5.5].

The fact that the function $t \mapsto \bar{\varphi}(1,t)$ can be extended to a holomorphic function follows from its Fourier expansion. By the transformation of time $\tau := T - t$, we find from (2.2) and (2.3) that

$$
\begin{aligned}
\bar{\varphi}(1,t) &= \int_0^1 G(1,x,T-t)(\bar{y}(x,T) - y_\Omega(x))\,\mathrm{d}x \\
&= \sum_{n=1}^{\infty} \frac{\cos(\rho_n)}{N_n} e^{-\rho_n^2(T-t)} \int_0^1 d(x)\cos(\rho_n x)\,\mathrm{d}x,
\end{aligned}
\tag{3.7}
$$

where

$$d(x) := \bar{y}(x,T) - y_\Omega(x).$$

The eigenvalues ρ_n behave asymptotically like $(n-1)\pi$, $n \to \infty$. Therefore, the factor $e^{-\rho_n^2(T-t)}$ converges very fast to zero as $n \to \infty$ provided that $t < T$. For $t \leq T - \varepsilon$, $\varepsilon > 0$ fixed, the convergence of the series

(3.7) and of all of its derivatives w.r. to t is uniform in $-\infty < t \leq T - \varepsilon$. The same holds true for the complex extension

$$\bar{\varphi}(1, z) := \sum_{n=1}^{\infty} \frac{\cos(\rho_n)}{N_n} e^{-\rho_n^2(T-z)} \int_0^1 d(x) \cos(\rho_n x) \, \mathrm{d}x,$$

if $\Re(z) \leq T - \varepsilon$. Therefore, the series defines a holomorphic function in the half plane $\{z \in \mathbb{C} : \Re(z) < T\}$ and $\bar{\varphi}(1, t)$ is obtained as its real part. $\qquad\square$

Let us re-write the expansion of $\bar{\varphi}(1, t)$ in the shorter form

$$\varphi(1, t) = \sum_{n=1}^{\infty} \frac{\cos(\rho_n)}{N_n} e^{-\rho_n^2(T-t)} d_n, \tag{3.8}$$

where the numbers

$$d_n := \int_0^1 d(\xi) \cos(\rho_n \xi) \, \mathrm{d}\xi$$

correspond to the Fourier coefficients of d (the exact Fourier coefficients are given by $d_n/\sqrt{N_n}$). The decisive result for the switching behavior of $\bar{u}$ is the following:

Lemma 4. *Let $0 < \varepsilon < T$ be given and assume $\|\bar{y}(\cdot, T) - y_\Omega(\cdot)\|_{L^2(0,1)} > 0$. Then the equations*

$$\bar{\varphi}(1, t) + \mu = 0 \quad \text{and} \quad \bar{\varphi}(1, t) - \mu = 0 \tag{3.9}$$

have at most finitely many solutions in $[0, T - \varepsilon]$.

Therefore, in $[0, T]$ these equations have at most countably many solutions that may accumulate only at $t = T$.

Proof. Let us consider only the first equation,

$$\bar{\varphi}(1, t) + \mu = 0, \qquad t \in [0, T - \varepsilon].$$

Assume to the contrary that it has infinitely many solutions in $[0, T - \varepsilon]$. Then they must have an accumulation point $\bar{t} \in [0, T - \varepsilon]$. By the identity theorem for holomorphic functions, we deduce

$$\bar{\varphi}(1, t) + \mu = 0 \quad \forall t \in (-\infty, T - \varepsilon].$$

9

Differentiating this equation, we obtain $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\varphi}(1,t) = 0$, hence

$$\sum_{n=1}^{\infty} \frac{\cos(\rho_n)}{N_n} \rho_n^2 e^{-\rho_n^2(T-t)} d_n = 0 \quad \text{in } (-\infty, T-\varepsilon]. \tag{3.10}$$

We multiply (3.10) by $e^{\rho_1^2(T-t)}$ and get

$$\frac{\cos(\rho_1)}{N_1} \rho_1^2 d_1 + \sum_{n=2}^{\infty} \frac{\cos(\rho_n)}{N_n} \rho_n^2 e^{-(\rho_n^2 - \rho_1^2)(T-t)} d_n = 0 \quad \text{in } (-\infty, T-\varepsilon].$$

Now, we pass to the limit $t \to -\infty$. Since $\rho_n^2 > \rho_1^2$ holds for $n \geq 2$ and the series is uniformly convergent, it follows $\exp\left(-(\rho_n^2 - \rho_1^2)(T-t)\right) \to 0$, hence

$$\frac{\cos(\rho_1)}{N_1} \rho_1^2 d_1 = 0, \tag{3.11}$$

and hence $d_1 = 0$. Notice that $\cos(\rho_n) \neq 0$ holds for all $n \in \mathbb{N}$. Therefore, the first item in (3.10) is zero. Multiplying (3.10) by $e^{\rho_2^2(T-t)}$ and passing to the limit $t \to \infty$, we find $d_2 = 0$. Repeating this method infinitely many times, it follows $d_n = 0$ for all $n \in \mathbb{N}$.

The system of functions $\{\cos(\rho_n \cdot) : n \in \mathbb{N}\}$ is complete in $L^2(0,1)$, hence $d = 0$ must hold in the sense of $L^2(0,1)$. This contradicts the assumption that $d = \bar{y}(\cdot, T) - y_\Omega \neq 0$. $\qquad\square$

Theorem 2. *Suppose $\|\bar{y}(\cdot, T) - y_\Omega(\cdot)\|_{L^2(0,1)} > 0$. For each $\varepsilon \in (0,T)$, the sets $\Phi_+ \cap (0, T-\varepsilon)$, $\Phi_0 \cap (0, T-\varepsilon)$, and $\Phi_- \cap (0, T-\varepsilon)$ are the union of finitely many open intervals. Consequently, Φ_+, Φ_0, and Φ_- are the union of at most countably many open intervals (the components of Φ_+, Φ_0, and Φ_-) which can accumulate only at $t = T$.*

Proof. The continuity of the function $t \mapsto \bar{\varphi}(1,t)$ yields that Φ_+ is an open set. All components of Φ_+ are bounded by two zeros of the function $t \mapsto \bar{\varphi}(1,t) - \mu$ in $[0, T-\varepsilon]$ or by the numbers 0, $T-\varepsilon$. By Lemma 4, the number of solutions to the equation $\bar{\varphi}(1,t) - \mu = 0$ in $(0, T-\varepsilon)$ is finite, hence the number of components of Φ_+ in $(0, T-\varepsilon)$ is finite, too. The statement on the accumulation of components in $[0,T]$ is an obvious consequence. The claim for Φ_0 and Φ_- can be proven analogously. $\qquad\square$

In addition, it follows that the complement of $\Phi_+ \cup \Phi_0 \cup \Phi_+$, which is the set of solutions of (3.9), is countable. Hence, the switching conditions of Lemma 2 uniquely define $\bar{u}$ almost everywhere on $(0,T)$. This implies that $\bar{u}$ almost everywhere attains values from the discrete set

10

$\{a, 0, b\}$. Moreover, $\bar{u}$ is piecewise constant on $[0, T - \epsilon)$ for all $\epsilon > 0$ with discontinuities only located at the solutions of (3.9). These points will be called switching points in the sequel.

Definition 1. *All points $t \in (0, T)$, where one of the two functions $t \mapsto \bar{\varphi}(1, t) - \mu$ and $t \mapsto \bar{\varphi}(1, t) + \mu$ changes the sign, are said to be switching points of $\bar{u}$.*

Theorem 3 (Bang-Bang-Bang Principle). *Let $\nu = 0$ and assume $\|\bar{y}(\cdot, T) - y_\Omega\|_{L^2(0,1)} > 0$. Then the following switching properties hold true:*

(i) *For each $0 < \varepsilon < T$, the number of switching points of $\bar{u}$ in $[0, T - \varepsilon]$ is finite. Therefore, the set of switching points of $\bar{u}$ in $[0, T]$ is at most countable and switching points can only accumulate at $t = T$.*

Between two subsequent switching points, the optimal control $\bar{u}$ is identically constant and equal to one of the values b, a or 0.

(ii) *If $|\bar{\varphi}(1, T)| \neq \mu$, then the number of switching points of $\bar{u}$ is finite and there is a sufficiently small $\delta > 0$ such that, for a.a. $t \in (T - \delta, T]$,*

$$\bar{u}(t) = \begin{cases} b & if \quad \bar{\varphi}(1, T) < -\mu \\ 0 & if \quad -\mu < \bar{\varphi}(1, T) < \mu \\ a & if \quad \bar{\varphi}(1, T) > \mu \end{cases} \qquad (3.12)$$

is fulfilled.

(iii) *If $\bar{\varphi}(1, T) = \mu$, then, in a certain neighborhood $(T - \delta, T]$, the optimal control can switch at most countably many times between a and 0. In the case $\bar{\varphi}(1, T) = -\mu$ it can switch at most countably many times between b and 0 in $(T - \delta, T]$.*

(iv) *Switching-over of $\bar{u}$ between a and b cannot happen.*

Proof. (i) Switching points can only be boundary points of the sets Φ_+, Φ_-, and Φ_0. Therefore, they must solve one of the two equations (3.9). By Lemma 4, the number of their solutions is at most countable and can accumulate only at $t = T$. Between switching points, $\bar{u}$ can only attain the values b, a, and 0, cf. Lemma 2. This proves (i).

(ii) If $|\bar{\varphi}(1, T)| \neq \mu$, then we are in one of the cases $\bar{\varphi}(1, t) > \mu$, $-\mu < \bar{\varphi}(1, t) < \mu$, or $\bar{\varphi}(1, t) < -\mu$ at $t = T$. In either case, by continuity of the function $t \mapsto \bar{\varphi}(1, t)$, these inequalities remain valid for all t in a sufficiently small interval $(T - \delta, T]$.

Now we apply Theorem 1: In the first case, we have $t \in \Phi_+$ and hence $\bar{u} = a$ in $(T - \delta, T]$. In the second, we have $t \in \Phi_0$, hence $\bar{u}(t) = 0$ in $(T - \delta, T]$, and in the third we obtain analogously that $\bar{u}(t) = b$ in $(T - \delta, T]$.

(iii), (iv) If $\bar{\varphi}(1, T) = \mu$ or $\bar{\varphi}(1, T) = -\mu$ is satisfied, we cannot exclude an accumulation of switching points at $t = T$. By the continuity of $t \mapsto \bar{\varphi}(1, t)$, these can only be a switchover between b and 0 or a and 0, respectively. □

Theorem 4. *If $\nu = 0$ and $\|\bar{y}(\cdot, T) - y_\Omega\|_{L^2(0,1)} > 0$, then the optimal control $\bar{u}$ is unique.*

Proof. Let two optimal controls $\bar{u}$ and $\bar{v}$ be given. Due to strict convexity of f_ν, the optimal state is unique, which gives $\bar{y} = y_{\bar{u}} = y_{\bar{v}}$. Thanks to Theorem 3, both controls must be of bang-bang-bang type: Almost everywhere and in open intervals, they admit only the values a, b or 0. Since the control problem is convex, every convex combination $\theta \bar{u} + (1 - \theta) \bar{v}$ is an optimal control and bang-bang-bang. This is only possible if $\bar{u} = \bar{v}$ holds almost everywhere. □

Definition 2. *Assume that $\|\bar{y}(\cdot, T) - y_\Omega\|_{L^2(0,1)} > 0$ holds for the solution of (P_0). Then the optimal control $\bar{u}$ is has at most countably many switching points. The switching points in $(0, T)$ solving $\bar{\varphi}(1, t) = \mu$ are denoted by t_j^μ and the ones solving $\bar{\varphi}(1, t) = -\mu$ are denoted by $t_j^{-\mu}$, $j \geq 1$. These switching points are ordered such that $t_j^\mu < t_{j+1}^\mu$ and $t_j^{-\mu} < t_{j+1}^{-\mu}$ holds for all $j \geq 1$.*

3.4 The case $\nu > 0$

Now we assume $\nu > 0$ and consider the problem (P_ν), i.e., the problem

$$\min_{u \in U_{ad}} \left\{ \int_\Omega |y_u(x, T) - y_\Omega(x)|^2 \, dx + \mu \int_0^T |u(t)| \, dt + \frac{\nu}{2} \int_0^T |u(t)|^2 \, dt \right\}. \tag{3.13}$$

We recall that y_u is defined as solution of the equation (1.2) associated to u.

Again, this problem has an optimal control u_ν with associated optimal state $y_\nu := y_{u_\nu}$. By strict convexity of the functional in (3.13), the optimal control is unique. The associated adjoint state is $\varphi_\nu := \varphi_{u_\nu}$. The necessary optimality condition is stated in Theorem 1.

By a detailed pointwise discussion of the variational inequality (3.2), the following result is deduced completely analogous to a result of [3, 27]

12

for a class of elliptic equations. We also refer to a later result for a parabolic problem in [4]. In the theorem, the projection function $\mathbb{P}_{[s_1,s_2]} : \mathbb{R} \to [s_1, s_2]$ is defined by

$$\mathbb{P}_{[s_1,s_2]}(s) = \max\{s_1, \min\{s, s_2\}\}.$$

Theorem 5. *For almost all $t \in [0, T]$, the following equations are fulfilled:*

$$u_\nu(t) \;=\; \mathbb{P}_{[a,b]}\left(-\frac{1}{\nu}(\varphi_\nu(1,t) + \mu\,\lambda_\nu(t))\right), \qquad (3.14)$$

$$u_\nu(t) \;=\; 0 \quad \textit{if and only if} \quad |\varphi_\nu(1,t)| \leq \mu, \qquad (3.15)$$

$$\lambda_\nu(t) \;=\; \mathbb{P}_{[-1,1]}\left(-\frac{1}{\mu}\,\varphi_\nu(1,t)\right). \qquad (3.16)$$

The relation (3.15) expresses the sparsity of the optimal control, while (3.16) extracts a single element out of the subdifferential of $j(u_\nu)$. We skip the proof, because it is completely analogous to the one in [3].

As a simple conclusion of Theorem 5 we get that, for $\nu > 0$, the functions λ_ν and u_ν are continuous on $[0, T]$: Indeed, the function $t \mapsto \varphi_\nu(1,t)$ is continuous, hence (3.16) yields the continuity of λ_ν. Inserting this in (3.14), we see the continuity of u_ν.

Let us determine the structure of u_ν. We might follow the presentation in [5], but for the convenience of the reader we prove the results again in our framework. Inserting (3.16) in (3.14), we find

$$u_\nu(t) = \mathbb{P}_{[a,b]}\left(-\frac{1}{\nu}\left(\varphi_\nu(1,t) + \mu\,\mathbb{P}_{[-1,1]}\left(-\frac{1}{\mu}\,\varphi_\nu(1,t)\right)\right)\right). \qquad (3.17)$$

Discussing this representation, we find the following result:

Theorem 6. *Assume $\nu > 0$. Then the implications*

$$\varphi_\nu(1,t) \in (-\infty, -\mu - \nu b) \;\Rightarrow\; u_\nu(t) = b \qquad (3.18)$$

$$\varphi_\nu(1,t) \in (-\mu - \nu b, -\mu) \;\Rightarrow\; u_\nu(t) = -\frac{1}{\nu}(\varphi_\nu(1,t) + \mu) \quad (3.19)$$

$$\varphi_\nu(1,t) \in (-\mu, \mu) \;\Rightarrow\; u_\nu(t) = 0 \qquad (3.20)$$

$$\varphi_\nu(1,t) \in (\mu, \mu - \nu a) \;\Rightarrow\; u_\nu(t) = -\frac{1}{\nu}(\varphi_\nu(1,t) - \mu) \quad (3.21)$$

$$\varphi_\nu(1,t) \in (\mu - \nu a, \infty) \;\Rightarrow\; u_\nu(t) = a \qquad (3.22)$$

hold almost everywhere in $[0, T]$.

Proof. (a) The implication (3.20) follows immediately from (3.15).

(b) Now we show (3.19). Here, the inclusion for φ_ν is equivalent to

$$1 < -\frac{1}{\mu}\varphi_\nu(1,t) < 1 + \frac{\nu b}{\mu}. \tag{3.23}$$

The left-hand side implies that $\mathbb{P}_{[-1,1]}(-\frac{1}{\mu}\varphi_\nu(1,t)) = 1$, hence (3.17) yields

$$u_\nu(t) = \mathbb{P}_{[a,b]}(-\frac{1}{\nu}(\varphi_\nu(1,t) + \mu)). \tag{3.24}$$

The last inequality of (3.23) is equivalent with $-\frac{1}{\nu}(\varphi_\nu(1,t) + \mu)) < b$, hence

$$u_\nu(t) = -\frac{1}{\nu}(\varphi_\nu(1,t) + \mu),$$

i.e., (3.19) is shown. (c) To prove (3.18), we mention that the inclusion $\varphi_\nu(1,t) \in (-\infty, -\mu - \nu b)$ is equivalent with

$$-\frac{1}{\mu}\varphi_\nu(1,t)) > 1 + \frac{\nu}{\mu}.$$

Again, we arrive at (3.24). However, the inclusion above also yields the inequality $-\frac{1}{\nu}(\varphi_\nu(1,t) + \mu) > b$, and hence from (3.24) we obtain the conclusion of (3.18).

The implications (3.21) and (3.22) can be confirmed in the same way. $\qquad\square$

This theorem reveals that the solutions of the four equations

$$\varphi_\nu(1,t) + \mu + \nu b = 0 \tag{3.25}$$
$$\varphi_\nu(1,t) + \mu = 0 \tag{3.26}$$
$$\varphi_\nu(1,t) - \mu = 0 \tag{3.27}$$
$$\varphi_\nu(1,t) - \mu + \nu a = 0 \tag{3.28}$$

determine the switching behavior of u_ν. In other words, u_ν can only switch in the zeros of the four functions standing in the left-hand side of (3.25)–(3.28).

Definition 3. *For $\nu > 0$, any $t \in (0,T)$, where one of the functions in the left-hand side of (3.25)–(3.28) changes its sign, is said to be a switching point of u_ν.*

Lemma 5. *Assume $\nu > 0$. If $\|y_\nu(\cdot,T) - y_\Omega\|_{L^2(0,1)} > 0$, then each of the equations (3.25)–(3.28) can have at most countably many solutions that can accumulate only at $t = T$. Therefore, u_ν can have at most countably many switching points that can accumulate only at $t = T$.*

Proof. The proof is almost identical with that of Lemma 3, since the adjoint state φ_ν solves the same adjoint equation as $\bar{\varphi}$, but with terminal value $y_\nu(\cdot, T) - y_\Omega$. Therefore, we have

$$
\begin{aligned}
\varphi_\nu(1, t) &= \int_0^1 G(1, \xi, T - t)(\bar{y}(\xi, T) - y_\Omega(\xi)) \, \mathrm{d}\xi \\
&= \sum_{n=1}^{\infty} \frac{\cos(\rho_n)}{N_n} e^{-\rho_n^2(T-t)} \int_0^1 d_\nu(\xi) \cos(\rho_n \xi) \, \mathrm{d}\xi,
\end{aligned}
$$

where $d_\nu = y_\nu(\cdot, T) - y_\Omega$. Now we proceed as in the proof of Lemma 3. $\qquad\square$

Definition 4. *The switching points of u_ν, i.e., the solutions in $(0, T)$ of (3.25), (3.26), (3.27), and (3.28), where the associated function changes its sign, are denoted by $t_j^{-\mu-\nu b}$, $t_j^{-\mu,\nu}$, $t_j^{\mu,\nu}$, and $t_j^{\mu-\nu a}$, respectively, for all $j \geq 1$ that may occur.*

Since the set of switching points is countable with possible accumulation only at $t = T$, we can assume that the sequences of switching points are ordered w.r. to j, namely $t_j^s < t_{j+1}^s$ holds for all $j \geq 1$ that appear and all upper indices $s \in \{-\mu - \nu b, (-\mu, \nu), (\mu, \nu), \mu - \nu a\}$.

4   Pass to the limit $\nu \to 0$

In this section, we discuss the convergence of controls u_ν for $\nu \searrow 0$. In addition, the convergence of switching points can be shown. First, we discuss the convergence of the sequences (u_ν), (y_ν) and (φ_ν) of optimal quantities for (P_ν).

Theorem 7. *(i) For $\nu \searrow 0$, the sequence (u_ν) contains a strongly convergent subsequence in $L^2(0, 1)$, denoted w.l.o.g. by (u_ν) again, such that $u_\nu \to \hat{u}$ as $\nu \searrow 0$. The control $\hat{u}$ is optimal for (P_0). The associated subsequences (y_{u_ν}) and (φ_{u_ν}) converge uniformly in $\bar{Q}$ to $y_{\hat{u}}$ and $\varphi_{\hat{u}}$, respectively, as $\nu \searrow 0$.*

(ii) If the optimal control $\bar{u}$ of (P_0) is unique, then these convergence properties hold for the whole sequences (u_ν), (y_{u_ν}), and (φ_{u_ν}).

Proof. (i) Since U_{ad} is weakly compact, the existence of a weakly convergent subsequence (u_ν) with weak limit $\hat{u} \in U_{ad}$ is obvious. Let $\bar{u}$ be optimal for (P_0). Then we have

$$
F_\nu(\bar{u}) \geq F_\nu(u_\nu) \geq F_0(u_\nu) \qquad \forall \nu > 0. \tag{4.1}
$$

15

Passing to the limit, we obtain from (4.1)

$$F_0(\bar{u}) = \lim_{\nu \searrow 0} F_\nu(\bar{u}) \geq \liminf_{\nu \searrow 0} F_0(u_\nu) \geq F_0(\hat{u}),$$

where we used the weak lower semicontinuity of F_0 to get the last inequality. Therefore, $\hat{u}$ must also be optimal for (P$_0$).

From optimality of u_ν and $\hat{u}$ for (P$_\nu$) and (P$_0$), respectively, we get

$$F_\nu(u_\nu) \leq F_\nu(\hat{u}) = \frac{1}{2} \int_\Omega |y_{\hat{u}}(x,T) - y_\Omega(x)|^2 \, \mathrm{d}x + \mu j(\hat{u}) + \frac{\nu}{2} \int_0^T |\hat{u}(t)|^2 \, \mathrm{d}t$$

$$\leq \frac{1}{2} \int_\Omega |y_{u_\nu}(x,T) - y_\Omega(x)|^2 \, \mathrm{d}x + \mu j(u_\nu) + \frac{\nu}{2} \int_0^T |\hat{u}(t)|^2 \, \mathrm{d}t.$$

This implies

$$\frac{\nu}{2} \int_0^T |u_\nu(t)|^2 \, \mathrm{d}t \leq \frac{\nu}{2} \int_0^T |\hat{u}(t)|^2 \, \mathrm{d}t$$

and hence, dividing by $\nu/2$ we obtain

$$\int_0^T |\hat{u}(t)|^2 \, \mathrm{d}t \leq \liminf_{\nu \searrow 0} \int_0^T |u_\nu(t)|^2 \, \mathrm{d}t \leq \limsup_{\nu \searrow 0} \int_0^T |u_\nu(t)|^2 \, \mathrm{d}t$$

$$\leq \int_0^T |\hat{u}(t)|^2 \, \mathrm{d}t.$$

This implies convergence of norms and strong convergence $u_\nu \to \hat{u}$ in $L^2(0,T)$ for $\nu \searrow 0$. The strongly convergent subsequence (u_ν) in $L^2(0,1)$ is transformed to a uniformly convergent subsequence (y_{u_ν}), i.e., $y_{u_\nu} \to y_{\hat{u}}$ in $C(\bar{Q})$. Therefore, we also have $\varphi_{u_\nu} \to \varphi_{\hat{u}}$ in $C(\bar{Q})$, because the mapping associating the solution of the adjoint equation to the final datum is continuous from $C[0,1]$ to $C(\bar{Q})$ and we have $\varphi_{u_\nu}(T) = y_{u_\nu}(T) - y_\Omega$.

(ii) If the optimal control of (P$_0$) is unique, say $\bar{u}$, then all subsequences of u_ν contain a subsequence converging weakly to the same limit $\bar{u}$. Then the whole sequence (u_ν) converges to $\bar{u}$. This transfers to the sequences (y_{u_ν}) and (φ_{u_ν}). □

Lemma 6. *For each $k \in \mathbb{N}$ and $\varepsilon \in (0,T)$, there is a constant $c > 0$ such that*

$$\left\| \frac{\mathrm{d}^k}{\mathrm{d}t^k} \varphi_{u_\nu}(1,\cdot) - \frac{\mathrm{d}^k}{\mathrm{d}t^k} \varphi_{\hat{u}}(1,\cdot) \right\|_{C([0,T-\varepsilon])} \leq c \, \|y_{u_\nu}(\cdot,T) - y_{\hat{u}}(\cdot,T)\|_{L^2(\Omega)}$$

holds for all $\hat{u} \in U_{ad}$.

16

Proof. In $[0, T - \varepsilon]$, the formally differentiated Fourier series is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_{u_\nu}(1,t) = \sum_{n=1}^{\infty} \rho_n^2 \frac{\cos\rho_n}{N_n} \int_0^1 \cos(\rho_n x)(y_{u_\nu}(x,T) - y_\Omega(x))\,\mathrm{d}x\, e^{-\rho_n^2(T-t)}.$$

It is uniformly convergent, since the series

$$(M + \|y_\Omega\|_{C[0,1]}) \sum_{n=1}^{\infty} \frac{1}{N_n} \rho_n^2\, e^{-\rho_n^2 \varepsilon}$$

with $M = \sup_{u \in U_{ad}} \|y_u\|_{C(\bar{Q})}$ is a convergent majorant. Let $\hat{u} \in U_{ad}$ be given. Then $\frac{\mathrm{d}}{\mathrm{d}t}\varphi_{\hat{u}}(1,t)$ has an analogous series representation. Hence, we can estimate

$$\left| \frac{\mathrm{d}}{\mathrm{d}t}\varphi_{u_\nu}(1,t) - \frac{\mathrm{d}}{\mathrm{d}t}\varphi_{\hat{u}}(1,t) \right|$$

$$= \left| \sum_{n=1}^{\infty} \rho_n^2 \frac{\cos\rho_n}{N_n} \int_0^1 \cos(\rho_n x)(y_{u_\nu}(x,T) - y_{\hat{u}}(x))\,\mathrm{d}x\, e^{-\rho_n^2(T-t)} \right|$$

$$\leq \|y_{u_\nu}(\cdot, T) - y_{\hat{u}}(\cdot, T)\|_{L^2(\Omega)} \sum_{n=1}^{\infty} \rho_n^2 \frac{1}{\sqrt{N_n}} e^{-\rho_n^2(T-t)},$$

which proves the claim for $k = 1$. The proof can be completed by an induction argument with respect to k. $\qquad\square$

The norm $\|y_{u_\nu}(\cdot, T) - y_{\hat{u}}(\cdot, T)\|_{L^2(\Omega)}$ can be estimated with the help of the following result.

Lemma 7. *Let $\hat{u}$ be optimal for (P_0). Then it holds*

$$\|y_{u_\nu}(\cdot, T) - y_{\hat{u}}(\cdot, T)\|_{L^2(\Omega)}^2 + \nu \|u_\nu - \hat{u}\|_{L^2(0,T)}^2 \leq \nu(\hat{u}, \hat{u} - u_\nu)_{L^2(0,T)}.$$

Proof. This is an immediate consequence of the optimality conditions, see, e.g., [31, Lemma 2.5]. $\qquad\square$

Combining these two results, we obtain a convergence rate for $\nu \searrow 0$ for the adjoint states.

Lemma 8. *Let $\hat{u}$ be optimal for (P_0). Then for each $k \in \mathbb{N}$ and $\varepsilon \in (0, T)$, there is a constant $c > 0$ such that*

$$\left\| \frac{\mathrm{d}^k}{\mathrm{d}t^k}\varphi_{u_\nu}(1,\cdot) - \frac{\mathrm{d}^k}{\mathrm{d}t^k}\varphi_{\hat{u}}(1,\cdot) \right\|_{C([0,T-\varepsilon])} \leq c\,\nu^{1/2}$$

is satisfied for all $\nu > 0$.

Proof. This is consequence of the previous two lemmas and the bound-edness of U_{ad}. □

Now we are able to prove the convergence of switching points of optimal controls for $\nu \searrow 0$.

Theorem 8. *Assume that $\|\bar{y}(\cdot, T) - y_\Omega\|_{L^2(0,1)} > 0$ is fulfilled. Let $t_j^\mu \in (0, T)$ be such that $\bar{\varphi}(1, t_j^\mu) = \mu$, and let n be the smallest positive integer such that*

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} \bar{\varphi}(1, t_j^\mu) \neq 0. \tag{4.2}$$

Then there are $\nu_0 > 0$ and $\tau > 0$ with the following properties:

(i) For all $\nu \in (0, \nu_0)$, each of the equations (3.27) and (3.28) has at most n solutions in the interval $(t_j^\mu - \tau, t_j^\mu + \tau)$. The equations (3.25) and (3.26) do not have solutions in $(t_j^\mu - \tau, t_j^\mu + \tau)$.

(ii) Additionally, there is $c > 0$ such that

$$|t_j^\mu - t_\nu| \leq c \, \nu^{\frac{1}{2n}}, \tag{4.3}$$

$$|t_j^\mu - t_\nu| \leq c \, (\|y_{u_\nu}(\cdot, T) - y_{\hat{u}}(\cdot, T)\|_{L^2(\Omega)}^{\frac{1}{n}} + \nu^{\frac{1}{n}}), \tag{4.4}$$

for all $\nu < \nu_0$ and all $t_\nu \in (t_j^\mu - \tau, t_j^\mu + \tau)$ solving one of the equations (3.27)–(3.28).

(iii) If n is odd, in particular if t_j^μ is a switching point of $\bar{u}$, then there is $\nu_1 \in (0, \nu_0)$ such that for all $\nu \in (0, \nu_1)$ there exist switching points of u_ν in the interval $(t_j^\mu - \tau, t_j^\mu + \tau)$ that solve one of the equations (3.27)–(3.28).

Analogous results hold for solutions $t_i^{-\mu}$ of $\bar{\varphi}(1, t) = -\mu$.

Proof. Denote $\bar{t} := t_j^\mu$. Assume for the moment that $\Lambda := \frac{\mathrm{d}^n}{\mathrm{d}t^n} \bar{\varphi}(1, \bar{t}) > 0$ holds in (4.2). Take $\tau > 0$ such that $\bar{t} + \tau < T$ and

$$\bar{\varphi}(1, t) \geq \frac{\mu}{2}$$

as well as

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} \bar{\varphi}(1, t) \geq \frac{\Lambda}{2} > 0 \tag{4.5}$$

are satisfied for all $t \in (0, T)$ with $|t - \bar{t}| < \tau$. Thanks to Lemma 6, we have uniform convergence of $\frac{\mathrm{d}^n}{\mathrm{d}t^n} \varphi_\nu(1, \cdot)$ to $\frac{\mathrm{d}^n}{\mathrm{d}t^n} \bar{\varphi}(1, \cdot)$ in $[\bar{t} - \tau, \bar{t} + \tau]$. Therefore, there is $\nu_0 > 0$ such that

$$\varphi_\nu(1, t) \geq 0$$

18

and

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\varphi_\nu(1,t) \geq \frac{\Lambda}{4} > 0$$

hold for all $\nu \in (0,\nu_0)$ and $t \in (0,T)$ with $|t - \bar{t}| < \tau$. Hence, the equation $\varphi_\nu(1,t) = \mu$ can have at most n distinct solutions in $(\bar{t}-\tau, \bar{t}+\tau)$. Analogously, the equation $\varphi_\nu(1,t) = \mu - \nu a$ has at most n distinct solutions in $(\bar{t}-\tau, \bar{t}+\tau)$. In addition, by $\varphi_\nu(1,t) \geq 0$ in $(\bar{t}-\tau, \bar{t}+\tau)$, the equations $\varphi_\nu(1,t) = -\mu$ and $\varphi_\nu(1,t) = -\mu - \nu b$ do not have solutions in $(\bar{t} - \tau, \bar{t} + \tau)$.

Let now $t \in (0,T)$ be given with $|t - \bar{t}| < \tau$. Then, performing a Taylor expansion and invoking the assumption, we find

$$\bar{\varphi}(1,t) = \mu + \frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n}\bar{\varphi}(1,\xi)(t - \bar{t})^n \tag{4.6}$$

with some intermediate point ξ.

Let n be an odd integer. Setting $t = \bar{t} - \tau$ and $t = \bar{t} + \tau$ and taking (4.5) into account, yields

$$\bar{\varphi}(1, \bar{t} - \tau) - \mu \leq -\frac{\Lambda}{2n!}\tau^n,$$

$$\bar{\varphi}(1, \bar{t} + \tau) - \mu \geq +\frac{\Lambda}{2n!}\tau^n.$$

By uniform convergence of the adjoint state, there is $\nu_1 \in (0,\nu_0)$ such that

$$\|\bar{\varphi}(1,\cdot) - \varphi_\nu(1,\cdot)\|_{C([\bar{t}-\tau, \bar{t}+\tau])} \leq \frac{\Lambda}{4n!}\tau^n$$

for all $\nu < \nu_1$. This implies

$$\varphi_\nu(1, \bar{t} - \tau) - \mu \leq -\frac{\Lambda}{4n!}\tau^n < 0,$$

$$\varphi_\nu(1, \bar{t} + \tau) - \mu \geq +\frac{\Lambda}{4n!}\tau^n > 0.$$

By the intermediate value theorem, there is a solution to $\varphi_\nu(1,t)-\mu = 0$ in $(\bar{t} - \tau, \bar{t} + \tau)$. At least one of these solutions is indeed a switching point of u_ν, as $\varphi_\nu(1,t) - \mu$ changes sign in the interval $(\bar{t} - \tau, \bar{t} + \tau)$.

Analogously, we can show existence of solutions of $\varphi_\nu(1,t)-\mu+a\nu = 0$. Here, we obtain for $\nu < \nu_1$ making ν_1 smaller if necessary

$$\varphi_\nu(1, \bar{t} - \tau) - \mu + a\nu \leq a\nu - \frac{\Lambda}{4n!}\tau^n \leq -\frac{\Lambda}{8n!}\tau^n < 0,$$

$$\varphi_\nu(1, \bar{t} + \tau) - \mu + a\nu \geq a\nu + \frac{\Lambda}{4n!}\tau^n \geq \frac{\Lambda}{8n!}\tau^n > 0.$$

This shows existence of solutions of (3.28) close to $\bar{t}$ for small ν.

Let now $t \in (\bar{t} - \tau, \bar{t} + \tau)$ be a solution of (3.27). Then we get

$$|\bar{\varphi}(1,t) - \varphi_\nu(1,t)| = \left| \mu + \frac{1}{n!} \frac{d^n}{dt^n} \bar{\varphi}(1,\xi)(t - \bar{t})^n - \mu \right| \geq \frac{\Lambda}{2n!} |t - \bar{t}|^n.$$

Since $|\bar{\varphi}(1,t) - \varphi_\nu(1,t)| \leq c\nu^{1/2}$ by Lemma 8, we obtain $|t - \bar{t}| \leq c' \, \nu^{\frac{1}{2n}}$. If t is a solution of (3.28) in $(\bar{t} - \tau, \bar{t} + \tau)$, then we obtain

$$|\bar{\varphi}(1,t) - \varphi_\nu(1,t)| = \left| \mu + \frac{1}{n!} \frac{d^n}{dt^n} \bar{\varphi}(1,\xi)(t - \bar{t})^n - \mu + \nu a \right|$$

$$\geq \frac{\Lambda}{2n!} |t - \bar{t}|^n - \nu|a|.$$

This proves the estimate $|t - \bar{t}| \leq c' \, \nu^{\frac{1}{2n}}$ for all $\nu < \nu_0$. The inequality (4.4) can be proven by using Lemma 6 instead of Lemma 8 to estimate $|\bar{\varphi}(1,t) - \varphi_\nu(1,t)|$.

With obvious modifications, the result can be proven if $\Lambda < 0$ or $t = t_i^{-\mu}$ holds. $\qquad\square$

Remark 1. *(i) An integer n satisfying (4.2) must exist, since the mapping $t \mapsto \bar{\varphi}(1,t)$ is analytic in $(-\infty, T)$.*

(ii) For $n = 1$, we obtain the following particular case of Theorem 8: Assume that t_j^μ is a switching point of $\bar{u}$ such that (4.2) is satisfied with $n = 1$. Then, for all $0 < \nu < \nu_1$, exactly two switching points $t_j^{\mu,\nu}$ and $t_j^{\mu - \nu a}$ of u_ν exist that solve equation (3.27) and (3.28), respectively. We have $\lim_{\nu \searrow 0} t_j^{\mu,\nu} = \lim_{\nu \searrow 0} t_j^{\mu - \nu a} = t_j^\mu$. An analogous result holds for $t_j^{-\mu}$ with points $t_j^{-\mu,\nu}$ and $t_j^{-\mu - \nu b}$.

Corollary 1. *Under the same conditions as in the previous Theorem 8, for every point $\bar{t} \in (0, T)$ with $|\bar{\varphi}(1,t)| = \mu$ there exists $c > 0$ such that*

$$\int_{\bar{t} - \tau}^{\bar{t} + \tau} \bar{\varphi}(1,t)(u(t) - \bar{u}(t)) + \mu \jmath'(\bar{u}(t); u(t) - \bar{u}(t)) \, dt \geq c \|u - \bar{u}\|_{L^1(\bar{t} - \tau, \bar{t} + \tau)}^{n+1}$$

holds for all $u \in U_{ad}$, where n and τ are as in Theorem 8. Here we used again the notation $\jmath(u) := |u|$.

Proof. Let $\bar{t} := t_j^\mu$ be a switching point of $\bar{u}$. As in the proof of Theorem 8, cf., (4.5) and (4.6), there is $\tau > 0$ and $K = \frac{|\Lambda|}{2 \cdot n!} > 0$ such that for all $t \in (\bar{t} - \tau, \bar{t} + \tau)$ it holds $\bar{\varphi}(1,t) \geq \frac{\mu}{2}$ and

$$|\bar{\varphi}(1,t) - \mu| \geq K|t - \bar{t}|^n. \tag{4.7}$$

In addition, we have $\bar{u}(t) \in \{a, 0\}$ for almost all t in this interval. Define for $\epsilon > 0$

$$A_\epsilon := \{t \in (\bar{t} - \tau, \bar{t} + \tau) : |\bar{\varphi}(1,t) - \mu| \geq \epsilon\}, \quad I_\epsilon = (\bar{t} - \tau, \bar{t} + \tau) \setminus A_\epsilon.$$

Hence, due to (4.7), there is a constant $c' > 0$ such that it holds $|I_\epsilon| \leq c' \, \epsilon^{1/n}$ for all $\epsilon > 0$.

Let now $u \in U_{ad}$ be given. Take $t \in A_\epsilon$ with $\bar{\varphi}(1,t) - \mu \geq \epsilon$. This implies $\bar{u}(t) = a < 0$ and

$$\bar{\varphi}(1,t)(u(t) - \bar{u}(t)) + \mu j'(\bar{u}(t); u(t) - \bar{u}(t)) = (\bar{\varphi}(1,t) - \mu)(u(t) - \bar{u}(t))$$
$$\geq \epsilon |u(t) - \bar{u}(t)|.$$

On the other hand, take $t \in A_\epsilon$ with $\bar{\varphi}(1,t) - \mu \leq -\epsilon$. As τ was chosen such that $\bar{\varphi}(1,t) \geq \frac{\mu}{2} > 0$ holds on $(\bar{t} - \tau, \bar{t} + \tau)$, it follows $\bar{u}(t) = 0$ and

$$\bar{\varphi}(1,t)(u(t) - \bar{u}(t)) + \mu j'(\bar{u}(t); u(t) - \bar{u}(t))$$
$$= \bar{\varphi}(1,t)(u(t) - \bar{u}(t)) + \mu |u(t) - \bar{u}(t)|$$
$$\geq (\mu - \bar{\varphi}(1,t))|u(t) - \bar{u}(t)|$$
$$\geq \epsilon |u(t) - \bar{u}(t)|.$$

Invoking (3.4), we conclude

$$\int_{\bar{t}-\tau}^{\bar{t}+\tau} \bar{\varphi}(1,t)(u(t) - \bar{u}(t)) + \mu j'(\bar{u}(t); u(t) - \bar{u}(t)) \, \mathrm{d}t$$

$$\geq \int_{A_\epsilon} \bar{\varphi}(1,t)(u(t) - \bar{u}(t)) + \mu j'(\bar{u}(t); u(t) - \bar{u}(t)) \, \mathrm{d}t$$
$$\geq \epsilon \|u - \bar{u}\|_{L^1(A_\epsilon)}$$
$$\geq \epsilon(\|u - \bar{u}\|_{L^1(\bar{t}-\tau,\bar{t}+\tau)} - \|u - \bar{u}\|_{L^1(I_\epsilon)})$$
$$\geq \epsilon(\|u - \bar{u}\|_{L^1(\bar{t}-\tau,\bar{t}+\tau)} - c'|b - a| \, \epsilon^{1/n}).$$

Setting
$$\epsilon := (2c'|b - a|)^{-n}\|u - \bar{u}\|^n_{L^1(\bar{t}-\tau,\bar{t}+\tau)}$$

yields

$$\int_{\bar{t}-\tau}^{\bar{t}+\tau} \bar{\varphi}(1,t)(u(t) - \bar{u}(t)) + \mu j'(\bar{u}(t); u(t) - \bar{u}(t)) \, \mathrm{d}t \geq \frac{\epsilon}{2}\|u - \bar{u}\|_{L^1(\bar{t}-\tau,\bar{t}+\tau)}$$
$$= c\|u - \bar{u}\|^{1+n}_{L^1(\bar{t}-\tau,\bar{t}+\tau)}$$

with $c = (2c'|b - a|)^{-n}/2$. $\qquad\square$

Theorem 9. *Assume that $\|\bar{y}(\cdot, T) - y_\Omega\|_{L^2(0,1)} > 0$ is fulfilled. Furthermore, we require that $|\bar{\varphi}(1,T)| \neq \mu$. Then there are constants $n \in \mathbb{N}$, $\nu_2 > 0$, and $c > 0$ such that it holds*

$$\|u_\nu - \bar{u}\|_{L^1(0,T)} \leq c\, \nu^{1/n}$$

for all $\nu \in (0, \nu_2)$.

Proof. The condition $|\bar{\varphi}(1,T)| \neq \mu$ implies that there are finitely many solutions of $|\bar{\varphi}(1,t)| = \mu$ in $(0,T)$. Let $t_1 \ldots t_m$ be the solutions of $|\bar{\varphi}(1,t)| = \mu$ in $(0,T)$. If $|\bar{\varphi}(1,0)| = \mu$, then we add the point $t = 0$ to this set. Denote by n_i and τ_i the constants given by Theorem 8 associated with the points t_i, $i = 1 \ldots m$. Note that Theorem 8 is also true if $t = 0$ is taken, with obvious modifications of the proof. Set $I_i := (t_i - \tau_i, t_i + \tau_i)$, $I = \bigcup_{i=1}^m I_i$, $J := (0,T) \setminus I$, $n := \max_{i=1\ldots m} n_i$.

By continuity, there is $\sigma > 0$ such that $\big||\bar{\varphi}(1,t)| - \mu\big| \geq \sigma$ for all $t \in J$. Due to uniform convergence $\varphi_\nu \to \bar{\varphi}$, there is $\nu_2 > 0$ such that it holds $\big||\varphi_\nu(1,t)| - \mu\big| \geq \sigma/2 > \nu(b - a)$ for all $t \in J$ and $\nu \in (0, \nu_2)$. Consequently u_ν and $\bar{u}$ coincide on J.

The intervals I_i were constructed in Theorem 8 such that $|\bar{\varphi}(1,t)| > \mu/2$ holds for all $t \in I_i$. Due to uniform convergence of the adjoint states, we get $|\varphi_\nu(1,t)| > \mu/4$ for all $t \in I$ and for all $\nu \in (0, \nu_2)$ by making ν_2 smaller if necessary. Hence, the signs of $\bar{\varphi}(1,t)$ and $\varphi_\nu(1,t)$ coincide on I, which implies $\bar{u}(t)u_\nu(t) \geq 0$ for all $t \in I$.

In the next step we will invoke Corollary 1. Let c_i be the constant given by Corollary 1 associated to the point t_i. Then we find

$$\int_0^T \bar{\varphi}(1,t)(u_\nu(t) - \bar{u}(t)) + \mu j'(\bar{u}(t);\, u_\nu(t) - \bar{u}(t))\, \mathrm{d}t$$

$$= \sum_{i=1}^m \int_{I_i} \bar{\varphi}(1,t)(u_\nu(t) - \bar{u}(t)) + \mu j'(\bar{u}(t);\, u_\nu(t) - \bar{u}(t))\, \mathrm{d}t$$

$$\geq \sum_{i=1}^m c_i \|u_\nu - \bar{u}\|_{L^1(I_i)}^{1+n_i}.$$

Since $|\bar{u}_\nu - u_\nu| \leq b - a$ and $n_i \leq n$, we obtain with

$$\tilde{c} := \min_{i=1\ldots m} c_i (2\tau_i(b - a))^{n_i - n}$$

$$\int_0^T \bar{\varphi}(1,t)(u_\nu(t) - \bar{u}(t)) + \mu j'(\bar{u}(t); \, u_\nu(t) - \bar{u}(t)) \, \mathrm{d}t$$

$$\geq \sum_{i=1}^m c_i \|u_\nu - \bar{u}\|_{L^1(I_i)}^{1+n_i}$$

$$= \sum_{i=1}^m c_i \|u_\nu - \bar{u}\|_{L^1(I_i)}^{n_i-n} \|u_\nu - \bar{u}\|_{L^1(I_i)}^{1+n} \qquad (4.8)$$

$$\geq \sum_{i=1}^m c_i (2\tau_i(b-a))^{n_i-n} \|u_\nu - \bar{u}\|_{L^1(I_i)}^{1+n}$$

$$\geq \tilde{c} \|u_\nu - \bar{u}\|_{L^1(I)}^{n+1}.$$

Testing the variational inequality (3.3) for u_ν with $\bar{u}$ and adding it to the inequality (4.8), we obtain

$$\int_0^T (\nu u_\nu(t) + \varphi_\nu(1,t) - \bar{\varphi}(1,t))(\bar{u}(t) - u_\nu(t)) \, \mathrm{d}t$$

$$+ \mu(j'(u_\nu; \, \bar{u} - u_\nu) + j'(\bar{u}; \, u_\nu - \bar{u})) \geq \tilde{c} \|u_\nu - \bar{u}\|_{L^1(I)}^{n+1}.$$

Since $\bar{u} = u_\nu$ on J and $\bar{u}(t)u_\nu(t) \geq 0$ on I it holds $j'(u_\nu; \, \bar{u} - u_\nu) + j'(\bar{u}; \, u_\nu - \bar{u}) = 0$. Due to the definition of the adjoint equation, it holds

$$\int_0^T (\varphi_\nu(1,t) - \bar{\varphi}(1,t))(\bar{u}(t) - u_\nu(t)) \, \mathrm{d}t = -\|y_{u_\nu}(\cdot, T) - y_{\bar{u}}(\cdot, T)\|_{L^2(\Omega)}^2.$$

Combining these facts, we find

$$\|y_{u_\nu}(\cdot, T) - y_{\bar{u}}(\cdot, T)\|_{L^2(\Omega)}^2 + \nu \|u_\nu - \bar{u}\|_{L^2(0,T)}^2 + \tilde{c} \|u_\nu - \bar{u}\|_{L^1(I)}^{n+1} \leq \nu(\bar{u}, \bar{u} - u_\nu)_{L^2(I)}.$$

See also [26, Lemma 1.3] and [32, Lemma 3.1] for similar results. By Young's inequality, we can estimate the right-hand side,

$$\nu(\bar{u}, \bar{u} - u_\nu)_{L^2(I)} = \nu(\bar{u}, \bar{u} - u_\nu)_{L^2(I)}$$

$$\leq \nu \,|\max\{b, |a|\}| \, \|u_\nu - \bar{u}\|_{L^1(I)}$$

$$\leq \frac{n}{n+1} \tilde{c}^{-1/n} (\nu|b-a|)^{\frac{n+1}{n}} + \frac{\tilde{c}}{n} \|u_\nu - \bar{u}\|_{L^1(I)}^{n+1}.$$

The last item can be absorbed by the left-hand side. This shows that the convergence rates

$$\|y_{u_\nu}(\cdot, T) - y_{\bar{u}}(\cdot, T)\|_{L^2(\Omega)} \leq c \, \nu^{\frac{1}{2} + \frac{1}{2n}},$$

$$\|u_\nu - \bar{u}\|_{L^2(0,T)} \leq c \, \nu^{\frac{1}{2n}},$$

$$\|u_\nu - \bar{u}\|_{L^1(0,T)} \leq c \, \nu^{\frac{1}{n}}$$

are satisfied for ν small enough with some constant $c > 0$ independent of ν. Here, we used again that $u_\nu = \bar{u}$ on J, which implies $\|u_\nu - \bar{u}\|_{L^1(I)} = \|u_\nu - \bar{u}\|_{L^1(0,T)}$. $\qquad\square$

Remark 2. *Let us point out an extension of the previous theorem: The assumption $|\bar{\varphi}(1, T)| \neq \mu$ can be omitted if we require instead that there are finitely many solutions of $|\bar{\varphi}(1, t)| = \mu$ in $(0, T)$, and that there exists $\tau > 0$, $n > 0$, and $c > 0$ such that it holds*

$$\left| \left\{ t \in (T - \tau, T) : \ \big||\bar{\varphi}(1, t)| - \mu\big| < \epsilon \right\} \right| \leq c \, \epsilon^{1/n} \quad \forall \epsilon > 0.$$

Then the conclusion of Corollary 1 is valid for the interval $(T - \tau, T)$, and the proof of Theorem 8 remains valid with minor modifications.

5 Extensions

The results of this paper can be easily extended to the following slightly more general situations:

(i) We considered problems with homogeneous initial condition $y(\cdot, 0) = 0$. All results remain true for the non-homogeneous initial condition $y(\cdot, 0) = y_0(\cdot)$ with $y_0 \in L^2(\Omega)$. To see this, we solve the heat equation with homogeneous boundary data and initial condition $y(\cdot, 0) = y_0$ and denote the solution by $\hat{y}$. Then $y(x, T) - y_\Omega = y_u(x, t) - (y_\Omega - \hat{y}(x, T))$ so that the results can be proven with $\hat{y}_\Omega := y_\Omega - \hat{y}(x, T)$.

(ii) For distributed controls of the form $f(x, t) = e(x)u(t)$ that act in the right-hand side of the heat equation with homogeneous boundary conditions, the solution y is given by the series representation

$$y(x, T) = \sum_{n=1}^{\infty} \frac{e_n}{N_n} \int_0^T e^{-\rho_n^2 (T-s)} u(s) \, ds, \tag{5.9}$$

where

$$e_n := \int_0^1 \cos(\rho_n \xi) e(\xi) \, d\xi.$$

In this way, the Fourier coefficients e_n replace the numbers $\cos \rho_n$ in (2.4). For proving the switching properties, in (3.11) we used the fact that $\cos \rho_n \neq 0$ holds for all $n \in \mathbb{N}$. Therefore, an easy inspection of the proofs shows that all results of the paper remain true for distributed controls of the form $f(x, t) = e(x)u(t)$ with fixed $e \in L^2(\Omega)$, if the condition

$$\int_0^1 \cos(\rho_n \xi) e(\xi) \, d\xi \neq 0 \quad \forall n \in \mathbb{N}$$

is fulfilled. In other words, the theory remains true for functions e where all Fourier coefficients with respect to the system $\cos(\rho_n x)$ are non-vanishing.

References

[1] E. Casas. Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control Optim.*, **35** (1997), 1297–1327.

[2] E. Casas. Second order analysis for bang-bang control problems of PDEs. *SIAM J. Control Optim.*, **50** (2012), 2355–2372.

[3] E. Casas, R. Herzog, and G. Wachsmuth. Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional. *SIAM J. Optim.*, **22** (2012), 795–820.

[4] E. Casas, C. Ryll, and F. Tröltzsch. Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems. *Comput. Methods Appl. Math.*, **13** (2013), 415–442.

[5] E. Casas, C. Ryll, and F. Tröltzsch. Second order and stability analysis for optimal sparse control of the FitzHugh-Nagumo equation. *SIAM J. Control Optim.*, **53** (2015), 2168–2202.

[6] V. Dhamo and F. Tröltzsch. Some aspects of reachability for parabolic boundary control problems with control constraints. *Comput. Optim. Appl.*, **50** (2011), 75–110.

[7] K. Eppler and F. Tröltzsch. On switching points of optimal controls for coercive parabolic boundary control problems. *Optimization*, **17** (1986), 93–101.

[8] H. O. Fattorini. Time-optimal control of solutions of operational differential equations. *SIAM J. on Control*, **2** (1964), 54–59.

[9] H. O. Fattorini. The time-optimal problem for boundary control of the heat equation. In *Calculus of variations and control theory (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975; dedicated to Laurence Chisholm Young on the occasion of his 70th birthday)*, Math. Res. Center, Univ. Wisconsin, Publ. **36**, 305–320, Academic Press, New York, 1976.

[10] H. O. Fattorini. *Infinite dimensional linear control systems*, volume 201 of *North-Holland Mathematics Studies*. Elsevier Science B.V., Amsterdam, 2005. The time optimal and norm optimal problems.

[11] K. Glashoff. Restricted approximation by strongly sign-regular kernels: the finite bang-bang principle. *J. Approx. Theory*, **29** (1980), 212–217.

[12] K. Glashoff and W. Krabs. Dualität und Bang-Bang-Prinzip bei einem parabolischen Rand-Kontrollproblem. In *Numerische Behandlung von Variations und Steuerungsproblemen (Tagung, Sonderforschungsber. 72 "Approximation und Optimierung", Inst. Angew. Math., Univ. Bonn, Bonn, 1974)*. Bonn. Math. Schriften, **77** (1975), 1–8.

[13] K. Glashoff and E. Sachs. On theoretical and numerical aspects of the bang-bang-principle. *Numer. Math.*, **29** (1977/78), 93–113.

[14] K. Glashoff and N. Weck. Boundary control of parabolic differential equations in arbitrary dimensions: supremum-norm problems. *SIAM J. Control Optimization*, **14** (1976), 662–681.

[15] W. A. Gruver and E. Sachs. *Algorithmic methods in optimal control*, volume 47 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass.-London, 1981.

[16] A. Karafiat. The problem of the number of switches in parabolic equations with control. *Ann. Polish Math.*, **XXXIV** (1977), 289–316.

[17] K. Kunisch and L. Wang. Bang-bang property of time optimal controls of semilinear parabolic equation. *Discrete Contin. Dyn. Syst.*, **36** (2016), 279–302.

[18] J.-L. Lions. *Optimal control of systems governed by partial differential equations.* Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170. Springer-Verlag, New York-Berlin, 1971.

[19] U. Mackenroth. Some remarks on the numerical solution of bang-bang type optimal control problems. *Numer. Funct. Anal. Optim.*, **5** (1982/83), 457–484.

[20] V. J. Mizel and T. I. Seidman. An abstract bang-bang principle and time-optimal boundary control of the heat equation. *SIAM J. Control Optim.*, **35** (1997), 1204–1216.

[21] J. P. Raymond and H. Zidani. Pontryagin's principle for state-constrained control problems governed by parabolic equations with unbounded controls. *SIAM J. Control Optim.*, **36** (1998), 1853–1879.

[22] E. Sachs. A parabolic control problem with a boundary condition of the Stefan-Boltzmann type. *Z. Angew. Math. Mech.*, **58** (1978), 443–449.

[23] K. Schittkowski. Numerical solution of a time-optimal parabolic boundary value control problem. *J. Optim. Theory Appl.*, **27** (1979), 271–290.

[24] E. J. P. Georg Schmidt. The "bang-bang" principle for the time-optimal problem in boundary control of the heat equation. *SIAM J. Control Optim.*, **18** (1980), 101–107.

[25] E. J. P. Georg Schmidt. Boundary control for the heat equation with nonlinear boundary condition. *J. Differential Equations*, **78** (1989), 89–121.

[26] M. Seydenschwanz. Convergence results for the discrete regularization of linear-quadratic control problems with bang-bang solutions. *Comput. Optim. Appl.*, **61** (2015), 731–760.

[27] G. Stadler. Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices. *Comput. Optim. Appl.*, **44** (2009), 159–181.

[28] F. Tröltzsch. Semidiscrete finite element approximation of parabolic boundary control problems—convergence of switching points. In *Optimal control of partial differential equations, II (Oberwolfach, 1986), Internat. Schriftenreihe Numer. Math.*, **78**, 219–232, Birkhäuser, Basel, 1987.

[29] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[30] A. N. Tychonov and A. A. Samarski. *Partial differential equations of mathematical physics. Vol. I.* Translated by S. Radding. Holden-Day, Inc., San Francisco, Calif.-London-Amsterdam, 1964.

[31] D. Wachsmuth and G. Wachsmuth. Regularization error estimates and discrepancy principle for optimal control problems with inequality constraints. *Control Cybernet.*, **40** (2011), 1125–1158.

[32] G. Wachsmuth and D. Wachsmuth. Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM Control Optim. Calc. Var.*, **17** (2011), 858–886.

[33] G. Wang and L. Wang. The bang-bang principle of time optimal controls for the heat equation with internal controls. *Systems Control Lett.*, **56** (2007), 709–713.

[34] L. Wang and Q. Yan. Bang-bang property of time optimal null controls for some semilinear heat equation. *SIAM J. Control Optim.*, **54** (2016), 2949–2964.

[35] N. Weck. Über Existenz, Eindeutigkeit und das "Bang-Bang-Prinzip" bei Kontrollproblemen aus der Wärmeleitung. In *Numerische Behandlung von Variations und Steuerungsproblemen (Tagungsband, Sonderforschungsber. 72 "Approximation und Optimierung", Inst. Angew. Math., Univ. Bonn, Bonn, 1974)*, Bonn. Math. Schriften, **77** (1975), 9–19.