**Control and Cybernetics**

vol. **48** (2019) No. 1

# Optimal sparse boundary control for a semilinear parabolic equation with mixed control-state constraints* †

by

**Eduardo Casas**[1] **and Fredi Tröltzsch**[2]

[1]Departamento de Matemática Aplicada y Ciencias de la Computación,
E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria,
39005 Santander, Spain
eduardo.casas@unican.es
[2]Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin,
Germany
troeltzsch@math.tu-berlin.de.

*Dedicated to Günter Leugering on the occasion of His 65th birthday*

**Abstract:** A problem of sparse optimal boundary control for a semilinear parabolic partial differential equation is considered, where pointwise bounds on the control and mixed pointwise control-state constraints are given. A standard quadratic objective functional is to be minimized that includes a Tikhonov regularization term and the $L^1$-norm of the control accounting for the sparsity. Applying a recent linearization theorem, we derive first-order necessary optimality conditions in terms of a variational inequality under linearized mixed control state constraints. Based on this preliminary result, a Lagrange multiplier rule with bounded and measurable multipliers is derived and sparsity results on the optimal control are demonstrated.

**Keywords:** semilinear parabolic equation, optimal control, sparse boundary control, mixed control-state constraints

## 1. Introduction

In this paper, we investigate sparse optimal controls for a class of nonlinear parabolic boundary control problems with mixed pointwise control-state constraints. Let us motivate this class of problems by the following simplified example:

*Submitted: January 2019; Accepted: April 2019

†The first author was partially supported by the Spanish Ministerio de Economía y Competitividad under project MTM2017-83185-P. The second author was supported by the collaborative research center SFB 910, TU Berlin, project B6.

In a bounded domain $\Omega \subset \mathbb{R}^N$, $N \in \mathbb{N}$, with Lipschitz boundary $\Gamma$, consider the optimal control problem

$$\min J(y,u) := \int_0^T \int_\Omega \frac{1}{2}|y - y_Q|^2 \, dxdt + \int_0^T \int_\Gamma \Big(\frac{\nu}{2}|u|^2 + \kappa\,|u|\Big)\, d\sigma dt \quad (1.1)$$

subject to the parabolic initial-boundary value problem

$$\begin{array}{rcll} \partial_t y - \Delta y & = & 0 & \text{in } \Omega \times (0,T) \\ \partial_n y + y^3 & = & u & \text{in } \Gamma \times (0,T) \\ y(x,0) & = & 0 & \text{in } \Omega \end{array} \quad (1.2)$$

and to the mixed pointwise control-state constraints

$$u_a \le u(x,t) \quad (1.3)$$
$$u(x,t) - y(x,t) \le u_d \quad (1.4)$$

to be fulfilled for a.a. $(x,t) \in \Gamma \times (0,T)$. Here, $y_Q$ is a desired state function, while $T > 0$, $u_a$, $u_d$, $\nu > 0$, and $\kappa > 0$ are fixed real numbers. Throughout the paper, we will write $Q := \Omega \times (0,T)$ and $\Sigma := \Gamma \times (0,T)$. By $u \in L^p(\Sigma)$, $p > N+1$, we denote the control function and $y$ is the associated state; $\partial_n y$ denotes the outward normal derivative defined a.e. on $\Gamma$, and $\sigma$ is the Lebesgue surface measure induced on $\Gamma$.

If the state $y(x,t)$ is interpreted as the temperature of a point $x \in \Omega$ at time $t$ and $u$ is a controllable outside temperature, the constraints have the following meaning: The control constraint (1.3) restricts the control temperature from below, $u_a$ is the lowest possible temperature. The constraint (1.4) restricts the speed of heating: It requires that the difference between the boundary temperature $y$ and the outside control temperature $u$ be not too large. This restriction avoids too fast heating. To avoid too sudden cooling, the lower bound $u_a + y \le u$ might be posed instead of (1.4). This case can be considered analogously.

Our theory also works for two-sided control constraints $u_a \le u(x,t) \le u_b$ as in Casas and Tröltzsch (2018a), where some results are even easier to prove. However, the complexity of the presentation is higher. Therefore, we confine ourselves to the one-sided pure control constraint (1.3).

Problems with mixed control state constraints for partial differential equations were investigated first in Tröltzsch (1979) for boundary control of the linear heat equation. With $\kappa = 0$, the theory was discussed later for nonlinear equations in some papers by A. Rösch and the second author, see exemplarily Rösch and Tröltzsch (2007) and the references therein. Recently, in Casas and Tröltzsch (2018a) we extended the theory of problems with mixed control-state constraint to sparse control, i.e. to the case $\kappa > 0$. There, the linear heat equation was taken as state equation. By linearity of the equation, the discussion of first-order necessary conditions was simpler than for the problems posed here, because the feasible set was convex.

The main novelty of this new contribution is the nonlinearity of the equation in the context of sparsity and mixed control state constraints. By the nonlinearity, the feasible set is no longer convex and special techniques of analysis are needed. In particular, the proof of comparison theorems for certain solutions of parabolic equations is much more demanding.

## 2. The state equation and comparison theorems

We shall consider a more general version of the state equation (1.2), namely

$$\begin{aligned} \partial_t y - \Delta y + d(x,t,y) &= 0 && \text{in } Q \\ \partial_n y + b(x,t,y) &= u && \text{in } \Sigma \\ y(x,0) &= y_0(x) && \text{in } \Omega. \end{aligned} \tag{2.1}$$

For the state equation, we rely on the following assumptions that are adoptet in the whole paper without explicitly mentioning them any more.

Assumption 1
- *The initial state $y_0$ belongs to $L^\infty(\Omega)$.*
- *Throughout the paper, we fix real numbers $q > \frac{N}{2} + 1$ and $p > N + 1$.*
- *The functions $d : Q\times\mathbb{R} \to \mathbb{R}$ and $b : \Sigma\times\mathbb{R} \to \mathbb{R}$ are measurable with respect to $(x,t)$ and of class $C^1$ with respect to $y$. Moreover, $d(\cdot,\cdot,0) \in L^q(Q)$ and $b(\cdot,\cdot,0) \in L^p(\Sigma)$.*
- *With some constant $M_0 > 0$, it holds that*

$$\begin{aligned} \left|\frac{\partial d}{\partial y}(x,t,0)\right| &\le M_0 \quad \text{for a.a. } (x,t) \in Q, \\ \left|\frac{\partial b}{\partial y}(x,t,0)\right| &\le M_0 \quad \text{for a.a. } (x,t) \in \Sigma. \end{aligned} \tag{2.2}$$

- *There is some real number $M_1$ such that*

$$\begin{aligned} \frac{\partial d}{\partial y}(x,t,y) &\ge M_1 \quad \text{for a.a. } (x,t) \in Q, \\ \frac{\partial b}{\partial y}(x,t,y) &\ge M_1 \quad \text{for a.a. } (x,t) \in \Sigma. \end{aligned} \tag{2.3}$$

- *The functions $\frac{\partial d}{\partial y}$ and $\frac{\partial b}{\partial y}$ are locally Lipschitz with respect to $y$: This means that, for all $M > 0$, there is some constant $L_M$ such that*

$$\begin{aligned} \left|\frac{\partial d}{\partial y}(x,t,y_1) - \frac{\partial d}{\partial y}(x,t,y_2)\right| &\le L_M\,|y_1 - y_2| \ \text{ a.e. in } Q \\ \left|\frac{\partial b}{\partial y}(x,t,y_1) - \frac{\partial b}{\partial y}(x,t,y_2)\right| &\le L_M\,|y_1 - y_2| \ \text{ a.e. in } \Sigma \end{aligned} \tag{2.4}$$

*holds for all* $y_i \in \mathbb{R}$ *with* $|y_i| \le M$, $i = 1, 2$.
*Let us observe that* (2.2) *and* (2.4) *imply that for every* $|y| \le M$

$$\left|\frac{\partial d}{\partial y}(x,t,y)\right| \le M_0 + L_M M \quad \text{for a.a. } (x,t) \in Q.$$

*Moreover, from the mean value theorem we deduce*

$$|d(x,t,y_2) - d(x,t,y_1)| \le (M_0 + L_M M)|y_2 - y_1| \quad \text{for a.a. } (x,t) \in Q$$

*for all* $y_i \in \mathbb{R}$ *with* $|y_i| \le M$. *Analogous inequalities hold for* $b$.

Notice that (2.3) does not imply monotonicity of $d$ and $b$ with respect to $y$.

Hereafter, we will follow the standard notation

$$W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) : \partial_t y \in L^2(0,T;H^1(\Omega)^*)\}.$$

Theorem 2.1 *For every* $u \in L^p(\Sigma)$ *with* $p > N+1$, *the state equation* (2.1) *has a unique solution* $y_u \in W(0,T) \cap L^\infty(Q)$. *Moreover, there exist constants* $K_2$ *and* $K_\infty$ *such that*

$$\|y_u\|_{L^\infty(Q)} \le K_\infty\Big(\|y_0\|_{L^\infty(\Omega)} + \|d(\cdot,\cdot,0)\|_{L^q(Q)} + \|b(\cdot,\cdot,0)\|_{L^p(\Sigma)} + \|u\|_{L^p(\Sigma)}\Big), \tag{2.5}$$

$$\|y_u\|_{L^\infty(0,T;L^2(\Omega))} + \|y_u\|_{L^2(0,T;H^1(\Omega))} \le K_2\Big(\|y_0\|_{L^2(\Omega)} + \|d(\cdot,\cdot,0)\|_{L^2(Q)} + \|b(\cdot,\cdot,0)\|_{L^2(\Sigma)} + \|u\|_{L^2(\Sigma)}\Big). \tag{2.6}$$

*Finally, if* $\{u_k\}_{k=1}^\infty \subset L^p(\Sigma)$ *is a sequence converging weakly to* $u$ *in* $L^p(\Sigma)$, *then* $y_{u_k} \to y_u$ *strongly in* $L^2(0,T;H^1(\Omega)) \cap L^\infty(Q)$, $y_{u_k|\Sigma} \to y_{u|\Sigma}$ *strongly in* $L^\infty(\Sigma)$, *and* $y_{u_k}(T) \to y_u(T)$ *strongly in* $L^\infty(\Omega)$, $k \to \infty$.

Proof Though the proof follows standard lines, some complications arise from the fact that $b$ and $d$ are possibly not monotone increasing with respect to $y$. Since we do not know a precise reference for such a general partial differential equation, we comment the main steps of the proof for the convenience of the reader.

Without loss of generality, we can assume that the constant $M_1$ in (2.3) is strictly negative. According to Grisvard (1985), Theorem 1.5.1.10, there exists a constant $K$ only depending on $\Omega$ such that

$$\int_\Gamma |y(x)|^2\, d\sigma \le K\Big[\varepsilon \int_\Omega |\nabla y(x)|^2\, dx + \frac{1}{\varepsilon}\int_\Omega |y(x)|^2\, dx\Big] \tag{2.7}$$

holds for all $y \in H^1(\Omega)$ and all $\varepsilon \in (0,1)$; we select

$$0 < \varepsilon < \min\left\{1, \frac{1}{2K|M_1|}\right\}.$$

Now, associated with $\varepsilon$, we take a real number $\lambda$ satisfying

$$\lambda \geq \frac{1}{2} + |M_1|\Big(\frac{K}{\varepsilon} + 1\Big).$$

By the change of variables $y_\lambda(x,t) = \mathrm{e}^{-\lambda t} y(x,t)$ in (2.1), we obtain

$$\begin{array}{rcll} \partial_t y_\lambda - \Delta y_\lambda + d_\lambda(x,t,y_\lambda) & = & 0 & \text{in } Q \\ \partial_n y_\lambda + b_\lambda(x,t,y_\lambda) & = & \mathrm{e}^{-\lambda t} u & \text{in } \Sigma \\ y_\lambda(x,0) & = & y_0(x) & \text{in } \Omega, \end{array} \tag{2.8}$$

where

$$d_\lambda(x,t,y) = \mathrm{e}^{-\lambda t} d(x,t,\mathrm{e}^{\lambda t} y) + \lambda y \quad \text{and} \quad b_\lambda(x,t,y) = \mathrm{e}^{-\lambda t} b(x,t,\mathrm{e}^{\lambda t} y).$$

Then we have

$$\frac{\partial d_\lambda}{\partial y}(x,t,y) = \frac{\partial d}{\partial y}(x,t,\mathrm{e}^{\lambda t} y) + \lambda \geq M_1 + \lambda \geq \frac{1}{2} + \frac{|M_1|K}{\varepsilon}$$

and

$$\frac{\partial b_\lambda}{\partial y}(x,t,y) = \frac{\partial b}{\partial y}(x,t,\mathrm{e}^{\lambda t} y) \geq M_1.$$

Hence, with the mean value theorem and the above inequalities, the choice of $\lambda$ implies the following inequalities

$$\begin{aligned} &\int_\Omega [d_\lambda(x,t,y_\lambda) - d_\lambda(x,t,0)] y_\lambda \, \mathrm{d}x \geq \Big(\frac{1}{2} + \frac{|M_1|K}{\varepsilon}\Big) \int_\Omega |y_\lambda(x)|^2 \, dx, \\ &\int_\Gamma [b_\lambda(x,t,y_\lambda) - b_\lambda(x,t,0)] y_\lambda \, d\sigma \geq M_1 \int_\Gamma |y_\lambda(x)|^2 \, d\sigma. \end{aligned}$$

Using (2.7), we infer

$$\begin{aligned} &\int_\Gamma [b_\lambda(x,t,y_\lambda) - b_\lambda(x,t,0)] y_\lambda \, d\sigma \\ &\qquad \geq M_1 K \Big[\varepsilon \int_\Omega |\nabla y_\lambda(x)|^2 \, dx + \frac{1}{\varepsilon} \int_\Omega |y_\lambda(x)|^2 \, dx\Big] \\ &\qquad \geq -\frac{1}{2} \int_\Omega |\nabla y_\lambda(x)|^2 \, dx + \frac{M_1 K}{\varepsilon} \int_\Omega |y_\lambda(x)|^2 \, dx. \end{aligned}$$

Observe that $M_1$ is assumed to be negative.

Therefore, by combining the estimates for $d_\lambda$ and $b_\lambda$ we get

$$\begin{aligned}&\int_\Omega [d_\lambda(x,t,y_\lambda) - d_\lambda(x,t,0)]y_\lambda\,dx + \int_\Gamma [b_\lambda(x,t,y_\lambda) - b_\lambda(x,t,0)]y_\lambda\,d\sigma\\ &\geq -\frac{1}{2}\int_\Omega |\nabla y_\lambda(x)|^2\,dx + \frac{1}{2}\int_\Omega |y_\lambda(x)|^2\,dx.\end{aligned}$$

From this inequality, multiplying (2.8) by $y_\lambda$ and integrating in $\Omega$ yields an a posteriori a priori estimate for almost all $t \in (0,T)$,

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\int_\Omega |y_\lambda(t)|^2\,dx + \frac{1}{2}\int_\Omega |\nabla y_\lambda(t)|^2\,dx + \frac{1}{2}\int_\Omega |y_\lambda(t)|^2\,dx\\ &\leq \int_\Omega d_\lambda(x,t,0)y_\lambda(t)\,dx + \int_\Gamma [b_\lambda(x,t,0) + \mathrm{e}^{-\lambda t}u(x,t)]y_\lambda(t)\,d\sigma.\end{aligned} \tag{2.9}$$

The estimate (2.6) is a straightforward consequence of this inequality and the definition of $y_\lambda$. Let us sketch the proof of existence of a solution. To this end, for every integer $k \geq 1$ we consider the functions

$$d_\lambda^k(x,t,y) = d_\lambda(x,t,\mathbb{P}_{[-k,+k]}(y)) \quad \text{and} \quad b_\lambda^k(x,t,y) = b_\lambda(x,t,\mathbb{P}_{[-k,+k]}(y)),$$

where $\mathbb{P}_{[a,b]}:\mathbb{R} \longrightarrow [a,b]$ is the projection operator: $\mathbb{P}_{[a,b]}(s) = \max\{a, \min\{s,b\}\}$.

For every $k$ and every $(z,w) \in L^2(Q) \times L^2(\Sigma)$, we consider the linear boundary value problem

$$\begin{array}{rcll}\partial_t y - \Delta y + d_\lambda^k(x,t,z) &=& 0 & \text{in } Q\\ \partial_n y_\lambda + b_\lambda^k(x,t,w) &=& \mathrm{e}^{-\lambda t}u & \text{in } \Sigma\\ y(x,0) &=& y_0(x) & \text{in } \Omega.\end{array} \tag{2.10}$$

In view of

$$|d_\lambda^k(x,t,z)| \leq M_0 + (L_k + \lambda)k \quad \text{and} \quad |b_\lambda^k(x,t,w)| \leq M_0 + L_k k,$$

we deduce the existence and uniqueness of a solution $y \in W(0,T)$ of (2.10) satisfying

$$\|y\|_{W(0,T)} \leq C\Big(\|y_0\|_{L^2(\Omega)} + \|u\|_{L^2(\Sigma)} + M_0 + (L_k + \lambda)k\Big).$$

Then, using the compactness of the embedding $W(0,T) \subset L^2(Q)$ and of the trace mapping $W(0,T) \ni y \mapsto y_{|\Sigma} \in L^2(\Sigma)$ (see, for instance, Temam, 1979, Ch. III, Theorem 2.1), we deduce with Schauder's fixed point theorem, applied to the mapping

$$L^2(Q) \times L^2(\Sigma) \ni (z,w) \mapsto (y, y_{|\Sigma}) \in L^2(Q) \times L^2(\Sigma),$$

the existence of a fixed point $(y_\lambda^k, y_{\lambda|\Sigma}^k)$, i.e. a solution in $W(0,T)$ of

$$\begin{array}{rcll}\partial_t y - \Delta y + d_\lambda^k(x,t,y) &=& 0 & \text{in } Q\\ \partial_n y + b_\lambda^k(x,t,y) &=& \mathrm{e}^{-\lambda t}u & \text{in } \Sigma\\ y(x,0) &=& y_0(x) & \text{in } \Omega.\end{array}$$

Now, using that $\lambda > 0$ and arguing as we did to get the estimate (2.9), we can proceed similarly to Ladyzhenskaya, Solonnikov and Ural'tseva (1988), III-§7, to deduce the boundedness

$$\|y^k_\lambda\|_{L^\infty(Q)} \leq C'\Big(\|y_0\|_{L^\infty(\Omega)} + \|d(\cdot,\cdot,0)\|_{L^q(Q)} + \|b(\cdot,\cdot,0)\|_{L^p(\Sigma)} + \|u\|_{L^p(\Sigma)}\Big)$$

with $C'$ independent of $k$. Then, for every $k$ large enough, we have that $d^k_\lambda(x,t,y^k_\lambda) = d_\lambda(x,t,y^k_\lambda)$ and $b^k_\lambda(x,t,y^k_\lambda) = b_\lambda(x,t,y^k_\lambda)$ and, hence, $y^k_\lambda$ is solution of (2.8). Undoing the change of variable $y = \mathrm{e}^{\lambda t} y^k_\lambda$, we see that $y$ is solution of (2.1) and (2.5)-(2.6) hold. The uniqueness follows in a standard way from Gronwall's Lemma.

To conclude the proof, we assume that $u_k \rightharpoonup u$ in $L^p(\Sigma)$ and prove the strong convergence of $\{y_{u_k}\}$ to $y_u$. First, we observe that (2.5) and (2.6) imply that $\{y_{u_k}\}_k$ is bounded in $L^2(0,T;H^1(\Omega)) \cap L^\infty(Q)$. By subtracting the equations satisfied by $y_{u_k}$ and $y_u$ and using the mean value theorem, we get for $z_k = y_{u_k} - y_u$

$$\begin{array}{rcll} \partial_t z_k - \Delta z_k + \dfrac{\partial d}{\partial y}(x,t,\hat{y}^k) z_k &=& 0 & \text{in } Q \\ \partial_n z_k + \dfrac{\partial b}{\partial y}(x,t,\tilde{y}^k) z_k &=& u_k - u & \text{in } \Sigma \\ z_k(x,0) &=& 0 & \text{in } \Omega, \end{array} \tag{2.11}$$

where $\hat{y}^k = y_u + \theta_k(y_{u_k} - y_u)$ and $\tilde{y}^k = y_u + \rho_k(y_{u_k} - y_u)$ for some measurable functions $\theta_k : Q \to [0,1]$ and $\rho_k : \Sigma \to [0,1]$. From the above equation we infer that $\{z_k\}_k$ is bounded in the space of Hölder functions $C^{0,\mu}(\bar{Q})$ for some $\mu \in (0,1)$ as well as in $W(0,T)$; see Ladyzhenskaya, Solonnikov and Ural'tseva (1988), Ch. III, Theorem 10.1. Arguing as in (2.9), and undoing the change of variables, we get after integration in $[0,T]$

$$\int_Q |\nabla z_k|^2 \, dxdt + \int_Q |z_k|^2 \, dxdt \leq C \int_\Sigma (u_k - u) z_k \, d\sigma dt.$$

This inequality, along with the compactness of the embedding $C^{0,\mu}(\bar{Q}) \subset C(\bar{Q})$, implies the strong convergence $z_k \to 0$ in $C(\bar{Q}) \cap L^2(0,T;H^1(\Omega))$, hence the strong convergences stated in the theorem follow. □

We will heavily rely on certain comparison theorems for the solution of semilinear equations. There exist many associated results in literature, often proven for classical solutions; we refer exemplarily to Pao (1992). However, we were not able to find a rigorously proven result for weak solutions of the class of nonlinear parabolic equations below. This is not surprising in view of the technicalities needed to prove Theorem 2.1.

Therefore, we prove the following result, although it can be expected from the available literature for classical parabolic equations.

THEOREM 2.2 (COMPARISON THEOREM) *Let $u$, $v \in L^p(\Sigma)$ be given and let $y$ and $z$ be the weak solutions to*

$$\begin{aligned} \partial_t y - \Delta y + d(x,t,y) &= 0 && \text{in } Q \\ \partial_n y + b(x,t,y) - y &= u && \text{in } \Sigma \\ y(x,0) &= y_0(x) && \text{in } \Omega, \end{aligned} \tag{2.12}$$

$$\begin{aligned} \partial_t z - \Delta z + d(x,t,z) &= 0 && \text{in } Q \\ \partial_n z + b(x,t,z) - z &= v && \text{in } \Sigma \\ z(x,0) &= y_0(x) && \text{in } \Omega. \end{aligned} \tag{2.13}$$

*If $u(x,t) \le v(x,t)$ holds a.e. in $\Sigma$, then $y(x,t) \le z(x,t)$ is satisfied a.e. in $Q$ and $y_{|\Sigma}(x,t) \le z_{|\Sigma}(x,t)$ holds a.e. in $\Sigma$.*

PROOF Let us set $w = y - z$. Then, by subtracting the equations (2.12) and (2.13), we get

$$\begin{aligned} \partial_t w - \Delta w + \frac{\partial d}{\partial y}(x,t,\hat{w})w &= 0 && \text{in } Q \\ \partial_n w + \Big(\frac{\partial b}{\partial y}(x,t,\tilde{w}) - 1\Big)w &= u - v && \text{in } \Sigma \\ w(x,0) &= 0 && \text{in } \Omega, \end{aligned} \tag{2.14}$$

where $\hat{w} = y + \theta(z - y)$ and $\tilde{w} = y + \rho(z - y)$ for some measurable functions $\theta : Q \to [0,1]$ and $\rho : \Sigma \to [0,1]$. This linear system has a unique solution $w \in W(0,T) \cap L^\infty(Q)$, see Raymond and Zidani (1998, 1999). Now we select

$$0 < \varepsilon < \frac{1}{2K(|M_1| + 1)} \quad \text{and} \quad \lambda \ge \frac{1}{2} + |M_1| + \frac{K(|M_1| + 1)}{\varepsilon},$$

where $K$ satisfies (2.7) and $M_1$ is given by (2.3). By setting $w_\lambda(x,t) = \mathrm{e}^{-\lambda t} w(x,t)$, we obtain from (2.14)

$$\begin{aligned} \partial_t w_\lambda - \Delta w_\lambda + \Big(\frac{\partial d}{\partial y}(x,t,\hat{w}) + \lambda\Big)w_\lambda &= 0 && \text{in } Q \\ \partial_n w_\lambda + \Big(\frac{\partial b}{\partial y}(x,t,\tilde{w}) - 1\Big)w_\lambda &= \mathrm{e}^{-\lambda t}(u - v) && \text{in } \Sigma \\ w_\lambda(x,0) &= 0 && \text{in } \Omega. \end{aligned} \tag{2.15}$$

Multiplying this equation by $w_\lambda^+ = \max\{w_\lambda, 0\}$ and arguing as in the proof of (2.9) leads to

$$\frac{1}{2}\int_\Omega |w_\lambda^+(T)|^2\,dx + \frac{1}{2}\int_Q |\nabla w_\lambda^+|^2\,dxdt + \frac{1}{2}\int_Q |w_\lambda^+|^2\,dxdt \le$$
$$\int_\Sigma \mathrm{e}^{-\lambda t}(u - v)w_\lambda^+\,d\sigma dt \le 0,$$

which proves that $w_\lambda^+ = 0$. Since the solution of (2.15) is continuous in $\bar{Q}$, we infer that $w^+(x,t) = \mathrm{e}^{\lambda t} w_\lambda^+(x,t) \le 0$ for every $(x,t) \in \bar{Q}$. This proves the theorem. □

REMARK 1 *In the above proof we have used the fact that $w^+ \in W(0,T)$ for every $w \in W(0,T)$. This can be proven as follows. First, we recall that $z^+ \in H^1(\Omega)$ and $\|\nabla z^+\|_{H^1(\Omega)} \le \|\nabla z\|_{H^1(\Omega)}$ for every $z \in H^1(\Omega)$. As an immediate consequence we have that $w^+ \in L^2(0,T;H^1(\Omega))$ for every $w \in L^2(0,T;H^1(\Omega))$. Let us prove that $\frac{d}{dt}w^+ \in L^2(0,T;H^1(\Omega)^*)$. We take a sequence $\{w_k\}_k \subset C^\infty(\bar{Q}) \subset H^1(Q)$ such that $w_k \to w$ in $W(0,T)$. We observe that $\frac{d}{dt}w_k^+ \subset L^2(Q)$ and*

$$\left\|\frac{d}{dt}w_k^+\right\|_{L^2(0,T;H^1(\Omega)^*)} = \left(\int_0^T \left\|\frac{d}{dt}w_k^+(t)\right\|^2_{H^1(\Omega)^*} dt\right)^{1/2}$$
$$\le \left(\int_0^T \left\|\frac{d}{dt}w_k(t)\right\|^2_{H^1(\Omega)^*} dt\right)^{1/2} = \left\|\frac{d}{dt}w_k\right\|_{L^2(0,T;H^1(\Omega)^*)} \le C$$

*for some constant $C$ and for every $k$. Taking a subsequence, we can assume that $\{\frac{d}{dt}w_k^+\}_k$ converges weakly in $L^2(0,T;H^1(\Omega)^*)$ to some $g$. This means that*

$$\int_0^T \int_\Omega \frac{dw_k^+}{dt} z\,dx\,\phi(t)dt \to \int_0^T \langle g, z\rangle_{H^1(\Omega)^*,H^1(\Omega)}\phi(t)dt \tag{2.16}$$

*for all $z \in H^1(\Omega)$, $\phi \in D(0,T)$.*

*On the other hand, we have*

$$\int_0^T \int_\Omega \frac{dw_k^+}{dt} z\,dx\,\phi(t)dt = \int_0^T \frac{d}{dt}\int_\Omega w_k^+(t)\,z\,dx\,\phi(t)\,dt \tag{2.17}$$
$$= -\int_0^T \int_\Omega w_k^+(t)\,z\,dx\,\phi'(t)\,dt \to -\int_0^T \int_\Omega w^+(t)\,z\,dx\,\phi'(t)\,dt$$
$$= \left\langle \frac{d}{dt}\int_\Omega w^+\,z\,dx, \phi\right\rangle_{D'(0,T),D(0,T)}. \tag{2.18}$$

*From (2.16) and (2.18), we infer that*

$$\frac{d}{dt}\int_\Omega w^+\,z\,dx = \langle g, z\rangle_{H^1(\Omega)^*,H^1(\Omega)} \quad \forall z \in H^1(\Omega).$$

*From Temam (1979), Chpt. 3, Lemma 1.1, we find that*

$$\frac{d}{dt}w^+ = g \in L^2(0,T;H^1(\Omega)^*),$$

*which concludes the proof.*

The next theorem is a contribution to *inverse isotony*, also called *monotony in the sense of Collatz*, that was discussed extensively in the 1950-70ties. An operator $\mathcal{T}$ mapping a partially ordered vector space $X$ into another partially ordered vector space $Y$ is called *inverse isotone*, if the inequality $\mathcal{T}(u) \leq \mathcal{T}(v)$ in $Y$ implies $u \leq v$ in $X$. In this case, the problem $\mathcal{T}(x) = g$ is called a *problem of monotone type.* For the associated concept, we mention Collatz (1952, 1964), Glashoff and Werner (1979); or Redheffer and Walter (1979). Inverse isotony plays an important role in the inclusion of numerical solutions to differential equations by guaranteed upper and lower bounds, a topic of *interval mathematics*, see Ortega and Rheinboldt (1967) or Rheinholdt (1969).

Associated results were formulated for classical solutions. We prove the result for weak solutions and for the particular type of equations that we need for our analysis and for nonsmooth control functions as it is standard in control theory. To our best knowledge, this is not available from literature.

We recall that $y_u$ denotes the solution of (2.1) associated with $u$.

Theorem 2.3 (Inverse isotony) *For every function $g \in L^p(\Sigma)$, the equation*

$$v(x,t) = g(x,t) + y_v(x,t), \quad (x,t) \in \Sigma, \tag{2.19}$$

*has a unique solution $v \in L^p(\Sigma)$. If $u \in L^p(\Sigma)$ satisfies the inequality*

$$u(x,t) \leq g(x,t) + y_u(x,t), \quad (x,t) \in \Sigma, \tag{2.20}$$

*then $u(x,t) \leq v(x,t)$ holds for a.a. $(x,t) \in \Sigma$.*

Proof a) *Existence and uniqueness of the solution $v$ to* (2.19).

We consider the system

$$\begin{aligned} \partial_t y - \Delta y + d(x,t,y) &= 0 && \text{in } Q \\ \partial_n y + b(x,t,y) - y &= g && \text{in } \Sigma \\ y(x,0) &= y_0(x) && \text{in } \Omega. \end{aligned}$$

Thanks to Theorem 2.1, this system has a unique solution $y \in W(0,T) \cap L^\infty(Q)$. Notice that the function $\tilde{b}(x,t,y) := b(x,t,y) - y$ satisfies Assumption 1. Now we set $v := g + y_{|\Sigma}$. Then we have $v \in L^p(\Sigma)$. Shifting $y$ to the right hand side of the boundary condition above, we see that $y$ solves the state equation with control $v$ on the right-hand side. Therefore, it holds that $y = y_v$ and we see, in turn, that $v = g + y_v$ a.e. in $\Sigma$, hence $v$ solves the equation (2.19).

The uniqueness of $v$ is obtained as follows: If $v_i$, $i = 1, 2$, are solutions of (2.19), then we have

$$v_1 - v_2 = y_1 - y_2, \tag{2.21}$$

where $y_i = y_{v_i}$, $i = 1, 2$. Subtracting the differential equations satisfied by $y_1$ and $y_2$ and taking $w = y_1 - y_2$, we find

$$\begin{aligned} \partial_t w - \Delta w + \frac{\partial d}{\partial y}(x,t,\hat{w})w &= 0 \quad \text{in } Q \\ \partial_n w + \frac{\partial b}{\partial y}(x,t,\tilde{w})w &= 0 \quad \text{in } \Sigma \\ w(x,0) &= 0 \quad \text{in } \Omega. \end{aligned}$$

Using again Raymond and Zidani (1998, 1999) we infer that $w = 0$. This shows $y_1 = y_2$ and implies, in turn, $v_1 = v_2$ by equation (2.21). Another way of proving the uniqueness was communicated to us by one of the referees: The uniqueness directly follows from the inverse isotony shown in part b) below. We thank the unknown referee for this nice idea.

b) *Inverse isotony.* Assume now that $u \in L^p(\Sigma)$ satisfies the inequality (2.20). We have to show that $u \le v$ holds a.e. in $\Sigma$.

The inequality for $u$ means that $u = g - e + y_u$ with some $e \in L^p(\Sigma)$ that is a.e. non-negative. By inserting $u = g - e + y_u$ and $v = g + y_v$ as control functions in the state equation (2.1) for $y_u$ and $y_v$, respectively, we get the boundary conditions

$$\begin{aligned} &\partial_n y_u + b(x,t,y_u) - y_u = g - e \\ &\partial_n y_v + b(x,t,y_v) - y_v = g. \end{aligned}$$

Since $g - e \le g$, the Comparison Theorem 2.2 yields

$$y_u(x,t) \le y_v(x,t) \text{ for a.a. } (x,t) \in \Sigma.$$

Therefore, with (2.20) we infer

$$u \le g + y_u \le g + y_v = v \text{ a.e. in } \Sigma. \qquad \square$$

Let us consider the control-to-state mapping $G : L^p(\Sigma) \longrightarrow W(0,T) \cap L^\infty(Q)$, associating to every control $u$ the corresponding state $y_u = G(u)$ solving (2.1). According to Theorem 2.1, this mapping is well defined. Let us denote by $G_\Sigma : L^p(\Sigma) \longrightarrow L^p(\Sigma)$ the mapping given by

$$G_\Sigma(u) = y_{u|\Sigma}.$$

The term "inverse isotony" in the previous Theorem 2.3 refers to the following observation: The equation $v = g + y_v$ can be written as $v - G_\Sigma(v) = g$. Therefore, the result of the theorem can be formulated as

$$u - G_\Sigma(u) \le v - G_\Sigma(v) \implies u \le v$$

in $L^p(\Sigma)$, i.e. $I - G_\Sigma$ is inverse isotone and the equation $v - G_\Sigma(v) = g$ is a problem of monotone type.

Theorem 2.4 *The mapping $G$ is of class $C^1$. The derivative of $G$ at $u \in L^p(\Sigma)$ in the direction $v \in L^p(\Sigma)$ is given by $G'(u)v = z_v$, where $z_v \in W(0,T) \cap L^\infty(Q)$ is the unique solution to*

$$\begin{array}{rcll} \partial_t z - \Delta z + \dfrac{\partial d}{\partial y}(x,t,y_u)\, z &=& 0 & \text{in } Q \\ \partial_n z + \dfrac{\partial b}{\partial y}(x,t,y_u)\, z &=& v & \text{in } \Sigma \\ z(x,0) &=& 0 & \text{in } \Omega. \end{array} \tag{2.22}$$

Proof Let us define the function space

$$Y = \{y \in W(0,T) \cap L^\infty(Q) : \partial_t y - \Delta y \in L^q(Q) \ \text{ and } \ \partial_n y \in L^p(\Sigma)\}.$$

Let us recall that for any element $y \in L^2(0,T; H^1(\Omega))$ such that $\partial_t y - \Delta y \in L^2(Q)$ there exists the normal derivative $\partial_n y \in H_{00}^{1/2}(\Sigma)^*$; see, for instance, Dautray and Lions (2000), pp. 525-526. Moreover, the embedding $L^2(\Sigma) \subset H_{00}^{1/2}(\Sigma)^*$ holds. In $Y$ we consider the norm

$$\|y\|_Y = \|y\|_{W(0,T)} + \|y\|_{L^\infty(Q)} + \|\partial_t y - \Delta y\|_{L^q(Q)} + \|\partial_n y\|_{L^p(\Sigma)}.$$

Now, we consider the mapping

$$\mathcal{F} : Y \times L^p(\Sigma) \longrightarrow L^q(Q) \times L^p(\Sigma) \times L^\infty(\Omega),$$

$$\mathcal{F}(y,u) = (\partial_t y - \Delta y + d(x,t,y), \partial_n y + b(x,t,y) - u, y(0) - y_0).$$

It is easy to check that $\mathcal{F}$ is well defined and it is of class $C^1$ and

$$\frac{\partial \mathcal{F}}{\partial y}(y,u)z = \Big(\partial_t z - \Delta z + \frac{\partial d}{\partial y}(x,t,y)z, \partial_n z + \frac{\partial b}{\partial y}(x,t,y)z, z(0)\Big).$$

For every $(f, v, z_0) \in L^q(Q) \times L^p(\Sigma) \times L^\infty(\Omega)$ the problem

$$\begin{array}{rcll} \partial_t z - \Delta z + \dfrac{\partial d}{\partial y}(x,t,y)\, z &=& f & \text{in } Q \\ \partial_n z + \dfrac{\partial b}{\partial y}(x,t,y)\, z &=& v & \text{in } \Sigma \\ z(x,0) &=& z_0 & \text{in } \Omega \end{array}$$

has a unique solution $z \in Y$. Hence, $\frac{\partial \mathcal{F}}{\partial y}(y,u) : Y \longrightarrow L^q(Q) \times L^p(\Sigma) \times L^\infty(\Omega)$ is an isomorphism. This follows as a particular case from Theorem 2.1; see also Raymond and Zidani (1998, 1999). Therefore, since $\mathcal{F}(G(u),u) = 0$ for every $u \in L^p(\Sigma)$, we can apply the implicit function theorem to the equation $\mathcal{F}(y,u) = (0,0,0)$ to deduce that $G$ is of class $C^1$ and that $z_v = G'(u)v$ is the solution of (2.22). $\square$

Corollary 1 (Nonnegativity of $G'_\Sigma$) *For all $u \in L^\infty(\Sigma)$ and all directions $v \in L^\infty(\Sigma)$ that are a.e. nonnegative, also $G'_\Sigma(u)v$ is a.e. nonnegative, i.e.*

$$v \geq 0 \implies G'_\Sigma(u)v \geq 0.$$

Proof The result is obtained by applying the Comparison Theorem 2.2 to

$$\begin{aligned} \partial_t y - \Delta y + \frac{\partial d}{\partial y}(x,t,y_u(x,t))y &= 0 \quad \text{in } Q \\ \partial_n y + \Big[\frac{\partial b}{\partial y}(x,t,y_u(x,t)) + 1\Big] y - y &= 0 \quad \text{in } \Sigma \\ y(x,0) &= 0 \quad \text{in } \Omega, \end{aligned}$$

$$\begin{aligned} \partial_t z - \Delta z + \frac{\partial d}{\partial y}(x,t,y_u(x,t))z &= 0 \quad \text{in } Q \\ \partial_n z + \Big[\frac{\partial b}{\partial y}(x,t,y_u(x,t)) + 1\Big] z - z &= v \quad \text{in } \Sigma \\ z(x,0) &= 0 \quad \text{in } \Omega. \end{aligned}$$

Then, obviously, the assumptions of Theorem 2.2 are satisfied. Therefore, since $v \geq 0$, we find by applying the comparison theorem that $z = G'_\Sigma(u)v$ is nonnegative. □

## 3. Optimal control problem and its solvability

We shall investigate the following optimal control problem that is an extended version of the introductory example (1.1)-(1.4):

$$\text{Min} \quad J(u) := \frac{\nu_Q}{2}\int_Q |y_u(x,t) - y_Q(x,t)|^2\, dxdt + \frac{\nu_T}{2}\int_\Omega |y_u(x,T) - y_T(x)|^2 dx + \int_\Sigma \Big(\frac{\nu}{2}|u(x,t)|^2 + \kappa\,|u(x,t)|\Big)\, d\sigma dt \tag{3.1}$$

subject to the mixed pointwise control-state constraints

$$u_a \leq u(x,t) \leq u_d + y_u(x,t) \tag{3.2}$$

to be satisfied a.e. in $\Sigma$.

The next assumption is needed throughout the further paper. It has to be assumed in all further results and will not be explicitly mentioned.

Assumption 2 (Optimal control data) *In the problem above, the following quantities are given: Real numbers $\nu_Q \geq 0$, $\nu_T \geq 0$, a Tikhonov parameter $\nu > 0$, the sparse parameter $\kappa > 0$, a desired state $y_Q \in L^\infty(Q)$, a desired final state $y_T \in L^\infty(\Omega)$, and real numbers $u_a < 0$, $u_d > 0$ as bounds for the constraints* (3.2)*. We also assume that $\nu_Q + \nu_T > 0$.*

In all what follows, the Assumptions 1 and 2 are tacitly assumed.

First, we prove the existence of an optimal control. To this aim, we show the following result on boundedness of the feasible set:

Lemma 1 *There is some constant $M_F > 0$ such that*

$$\max\{\|u\|_{L^\infty(\Sigma)}, \|y_u\|_{L^\infty(Q)}, \|y_{u|\Sigma}\|_{L^\infty(\Sigma)}\} \le M_F \tag{3.3}$$

*holds for all feasible controls $u$, i.e. for all $u$ satisfying the constraints* (3.2), *and for their associated states $y_u$.*

Proof By Theorem 2.3 we know that there exists a unique solution $v$ to the equation $v = u_d + y_v$. Thanks to $p > N+1$, the state $y_v$ belongs to $L^\infty(Q)$ and $y_{v|\Sigma}$ to $L^\infty(\Sigma)$. Consequently, we have $v \in L^\infty(\Sigma)$; set $c_0 = \|v\|_{L^\infty(\Sigma)}$. Then, (3.2) and Theorem 2.3 on inverse isotony imply that $u(x,t) \le v(x,t) \le c_0$ holds for almost all $(x,t) \in \Sigma$. Moreover, we have $u \ge u_a$ and hence $\|u\|_{L^\infty(\Sigma)} \le \max\{|u_a|, c_0\}$. By inserting this in (2.5), we find another constant $c_1$ such that $\|y_u\|_{L^\infty(Q)} \le c_1$ and, as a consequence, also $\|y_{u|\Sigma}\|_{L^\infty(\Sigma)} \le c_1$. With the choice of $M_F := \max\{|u_a|, c_0, c_1\}$, the estimate (3.3) is fulfilled. □

Let us introduce for convenience the functionals

$$f(u) = \frac{\nu_Q}{2}\int_Q |y_u - y_Q|^2\,dxdt + \frac{\nu_T}{2}\int_\Omega |y_u(x,T) - y_T(x)|^2\,dx + \frac{\nu}{2}\int_\Sigma |u(x,t)|^2\,d\sigma dt, \tag{3.4}$$

$$j(u) = \int_\Sigma |u(x,t)|\,d\sigma dt,$$

and the *feasible set*

$$F = \{u \in L^\infty(\Sigma) : u_a \le u \le u_d + y_u\}.$$

Then we have $J(u) = f(u) + \kappa\, j(u)$ and the optimal control problem can be written in the short form as

$$\min\{f(u) + \kappa\, j(u) : u \in F\}. \tag{3.5}$$

We say that a control $\bar u \in F$ is *optimal* if

$$f(\bar u) + \kappa\, j(\bar u) \le f(u) + \kappa\, j(u) \quad \forall u \in F,$$

and *locally optimal* if this inequality holds for all $u \in F \cap B_\varepsilon(\bar u)$, where $B_\varepsilon(\bar u)$ is the open ball of $L^p(\Sigma)$ with radius $\varepsilon > 0$ centered at $\bar u$.

Theorem 3.1 (Existence of an optimal control) *Assume that there exists a control $u \in L^\infty(\Sigma)$ that satisfies the constraints* (3.2). *Then there exists at least one optimal control.*

Proof According to the assumptions of the theorem, the feasible set $F$ is non-empty. Therefore, we can select a minimizing sequence $\{u_k\}_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty}(f(u_k)+\kappa j(u_k)) = \inf\{f(u)+\kappa j(u) : u \in F\} =: inf \geq 0.$$

Thanks to Lemma 1, $F$ is bounded in $L^\infty(\Sigma)$, hence we can extract a subsequence of $\{u_k\}_k$ that converges weakly in $L^p(\Sigma)$ to some element $\bar{u}$; let us denote this subsequence by $\{u_k\}_k$, again. From the convergences stated in Theorem 2.1 we infer that $f(\bar{u}) \leq \liminf_{k\to\infty} f(u_k)$. Moreover, the convexity and strong continuity of $j$ imply that $j(\bar{u}) \leq \liminf_{k\to\infty} j(u_k)$. Hence, we get that $J(\bar{u}) \leq \liminf_{k\to\infty} J(u_k) = inf$. To conclude that $\bar{u}$ is an optimal control, we have to prove that $\bar{u} \in F$. This is a consequence of the fact that $F$ is sequentially weakly closed in $L^p(\Sigma)$. Indeed, first we observe that the sets

$$\mathcal{U}_1 = \{v \in L^\infty(\Sigma) : u_a \leq v\} \quad \text{and} \quad \mathcal{U}_2 = \{v \in L^\infty(\Sigma) : v \leq u_d\}$$

are convex and closed in $L^p(\Sigma)$, therefore weakly closed, as well. Since $\{u_k\}_k \subset \mathcal{U}_1$ and $\{u_k - y_{u_k}\}_k \subset \mathcal{U}_2$ and, due to Theorem 2.1, $u_k - y_{u_k} \rightharpoonup \bar{u} - y_{\bar{u}}$ in $L^p(\Sigma)$, we infer that $\bar{u} \in \mathcal{U}_1$ and $\bar{u} - y_{\bar{u}} \in \mathcal{U}_2$. This means $u_a \leq \bar{u} - y_{\bar{u}} \leq u_d$ and hence $\bar{u} \in F$. □

## 4. Necessary optimality conditions and sparsity

### 4.1. A convergence result

Let us start with a simple but important convergence result for $\kappa \to \infty$.

Assumption 3 *The control $u = 0$ strictly satisfies the mixed control-state constraint of the optimal control problem, i.e. there is some $\delta > 0$ such that*

$$\delta \leq u_d + G_\Sigma(0)(x,t) \quad \text{for a.a. } (x,t) \in \Sigma.$$

Since $u_a < 0$ holds, the above assumption implies that $u = 0$ is a feasible control. From now on, we will write

$$y^0 = G(0).$$

Lemma 2 *Let Assumption 3 be satisfied and let, for $\kappa > 0$, $\{\bar{u}_\kappa\}_\kappa$ denote a family of optimal controls of* (3.1)-(3.2)*, corresponding to the sparse parameters $\kappa$. Let $\bar{y}_\kappa = G(\bar{u}_\kappa)$ be the associated states. Then we have*

$$\lim_{\kappa\to\infty} \|\bar{u}_\kappa\|_{L^s(\Sigma)} = 0 \quad \forall s \in [1,\infty)$$

*and $\bar{y}_\kappa \to y^0$ strongly in $L^2(0,T;H^1(\Omega)) \cap L^\infty(Q)$, $\bar{y}_\kappa(\cdot,T) \to y^0(\cdot,T)$ strongly in $L^\infty(\Omega)$, and $\bar{y}_{\kappa|\Sigma} \to {y^0}_{|\Sigma}$ strongly in $L^\infty(\Sigma)$. Moreover, there is* *some $\kappa_0 > 0$ such that, for a.a. $(x,t) \in \Sigma$,*

$$u_d + \bar{y}_\kappa(x,t) \geq \tfrac{\delta}{2} \quad \forall \kappa \geq \kappa_0.$$

Proof Thanks to Assumption 3, the control $u = 0$ is feasible for all $\kappa$, hence

$$\begin{aligned} J(\bar{u}_\kappa) &= \frac{\nu_Q}{2}\|\bar{y}_\kappa - y_Q\|^2_{L^2(Q)} + \frac{\nu_T}{2}\|\bar{y}_\kappa(T) - y_T\|^2_{L^2(\Omega)} \\ &\quad + \frac{\nu}{2}\|\bar{u}_\kappa\|^2_{L^2(\Sigma)} + \kappa\|\bar{u}_\kappa\|_{L^1(\Sigma)} \\ &\leq J(0) = \frac{\nu_Q}{2}\|y^0 - y_Q\|^2_{L^2(Q)} + \frac{\nu_T}{2}\|y^0(T) - y_T\|^2_{L^2(\Omega)}. \end{aligned}$$

This immediately yields

$$\|\bar{u}_\kappa\|_{L^1(Q)} \leq \frac{1}{2\kappa}\Big(\nu_Q\|y^0 - y_Q\|^2_{L^2(Q)} + \nu_T\|y^0(T) - y_T\|^2_{L^2(\Omega)}\Big) \to 0, \quad \kappa \to \infty.$$

Moreover, we know from Lemma 1 that the set of feasible controls is bounded in $L^\infty(\Sigma)$. Therefore, the $L^1$-convergence above implies also

$$\|\bar{u}_\kappa\|_{L^s(\Sigma)} \to 0, \quad \kappa \to \infty \quad \forall s \in [1,\infty).$$

The convergences $\bar{y}_\kappa \to y^0$ and the associated ones for the final and boundary values of $\bar{y}_\kappa$ follow from Theorem 2.1. Finally, it is enough to notice that Assumption 3 can be written down as

$$\delta \leq u_d + y^0(x,t) \quad \text{for a.a. } (x,t) \in \Sigma,$$

to deduce the existence of $\kappa_0 > 0$ such that $u_d + \bar{y}_\kappa(x,t) \geq \frac{\delta}{2}$ holds for all $\kappa \geq \kappa_0$ and for a.a. $(x,t) \in \Sigma$. □

The proof reveals that Assumption 3 can be slightly relaxed: For the first part of the statement, we only need $0 \leq u_d + G_\Sigma(0)(x,t)$ for a.a. $(x,t) \in \Sigma$.

### 4.2. Linearization and preliminary optimality conditions

To prove our first-order necessary optimality conditions, we start with a linearization theorem. This complements a theorem on necessary optimality conditions in our related paper on sparsity properties for nonlinear problems with pointwise state constraints, Casas and Tröltzsch (2018b). There, we proved the first order necessary conditions, including regular Borel measures as Lagrange multipliers and prepared this result by an abstract Lagrange multiplier rule relying on a linearized Slater condition. Such a result was not known before, because the $L_1$-norm in the objective functional leads to a non-differentiable objective functional.

Here, we cannot apply this abstract result, because in our case of mixed control-state constraints we want to prove the existence of Lagrange multipliers for the constraint (1.4) that are functions of $L^\infty(\Sigma)$ rather than of $L^\infty(\Sigma)^*$. To this aim, we first perform a linearization of the problem in $L^\infty(\Sigma)$. Next, we change the control space to $L^2(\Sigma)$ without changing the feasible set. Finally we show the existence of Lagrange multipliers by special techniques that are tailored to problems with mixed control-state constraints.

We begin with an abstract result of linearization for differentiable constraints, but for an objective functional that is composed of a nonlinear differentiable and a convex function.

Let $U$ and $Y$ be two normed vector spaces, $K \subset U$ and $C \subset Y$ two convex sets, and let $H : U \longrightarrow Y$, $f : U \longrightarrow \mathbb{R}$, and $g : U \longrightarrow (-\infty, +\infty]$, be given mappings. Consider the optimization problem

$$\min\{f(u) + g(u) : u \in K \text{ and } H(u) \in C\}. \tag{4.1}$$

Theorem 4.1 (Linearization) *Let $\bar{u}$ be a local solution of* (4.1). *Assume that $f$ and $H$ are Gâteaux differentiable at $\bar{u}$, $g$ is convex and $int\, C \neq \emptyset$. If the linearized Slater condition*

$$\exists u_0 \in K : H(\bar{u}) + H'(\bar{u})(u_0 - \bar{u}) \in int\, C \tag{4.2}$$

*is satisfied, then*

$$f'(\bar{u})(u - \bar{u}) + g(u) - g(\bar{u}) \geq 0 \tag{4.3}$$

*holds for all $u \in K$ that satisfy*

$$H(\bar{u}) + H'(\bar{u})(u - \bar{u}) \in C. \tag{4.4}$$

Proof Let us define

$$U_0 = \{u \in K : H(\bar{u}) + H'(\bar{u})(u - \bar{u}) \in int\, C\}.$$

The linearized Slater assumption (4.2) implies that $u_0 \in U_0$, hence $U_0 \neq \emptyset$. Let us take an arbitrary element $u \in U_0$. Since

$$\lim_{\rho \to 0} \Big( H(\bar{u}) + \frac{H(\bar{u} + \rho(u - \bar{u})) - H(\bar{u})}{\rho} \Big) = H(\bar{u}) + H'(\bar{u})(u - \bar{u}) \in int\, C,$$

there exists $\rho_u \in (0, 1)$ such that

$$y_\rho = H(\bar{u}) + \frac{H(\bar{u} + \rho(u - \bar{u})) - H(\bar{u})}{\rho} \in int\, C \quad \forall \rho \in (0, \rho_u).$$

For every $\rho \in (0, \rho_u)$, we have

$$H(\bar{u} + \rho(u - \bar{u})) = \rho y_\rho + (1 - \rho) H(\bar{u}) \in C.$$

The local optimality of $\bar u$ along with the convexity of $g$ implies that for every $\rho$ sufficiently small we have

$$\begin{aligned} 0 &\le \frac{f(\bar u+\rho(u-\bar u))-f(\bar u)}{\rho}+\frac{g(\bar u+\rho(u-\bar u))-g(\bar u)}{\rho} \\ &\le \frac{f(\bar u+\rho(u-\bar u))-f(\bar u)}{\rho}+g(u)-g(\bar u). \end{aligned}$$

Passing to the limit as $\rho\to 0$ in the above inequality, we obtain

$$f'(\bar u)(u-\bar u)+g(u)-g(\bar u)\ge 0 \quad \forall u\in U_0. \tag{4.5}$$

Now, we take $u\in K$ such that $H(\bar u)+H'(\bar u)(u-\bar u)\in C$ but $u\notin U_0$. From (4.2) it follows that $u_\rho=u+\rho(u_0-u)\in U_0$ for every $\rho\in(0,1)$. Therefore, the convexity of $g$ and (4.5) implies

$$f'(\bar u)(u_\rho-\bar u)+g(u)-g(\bar u)+\rho(g(u_0)-g(u))\ge f'(\bar u)(u_\rho-\bar u)+g(u_\rho)-g(\bar u)\ge 0$$

for every $\rho\in(0,1)$. Upon passing to the limit $\rho\to 0$ in the left hand side of the above chain of inequalities, (4.3) follows. ☐

We apply this theorem to our optimal control problem for $U=Y=L^\infty(\Sigma)$ with

$$\begin{aligned} &g(u)=\kappa j(u),\\ &H(u)=u-u_d-G_\Sigma(u),\\ &K=\{u\in L^\infty(\Sigma): u\ge u_a \ \text{ a.e. in } \Sigma\},\\ &C=(L^\infty(\Sigma))^-=\{u\in L^\infty(\Sigma): u\le 0 \ \text{ a.e. in } \Sigma\}, \end{aligned}$$

where $G_\Sigma(u)=y_{u|\Sigma}$. Notice that $C$ has a nonempty interior in $L^\infty(\Sigma)$.

In terms of $H(u)=u-u_d-G_\Sigma(u)$, the condition (4.2) takes the form

$$u_0-u_d-G_\Sigma(\bar u)-G'_\Sigma(\bar u)(u_0-\bar u)\in int\, C.$$

For the optimal control problem (3.1)-(3.2), the linearized Slater condition admits the following form:

Assumption 4 (Linearized Slater condition) *Let $\bar u\in L^\infty(\Sigma)$ be fixed. Assume the existence of $u_0\in L^\infty(\Sigma)$ and $\delta>0$ such that*

$$u_0(x,t)\ge u_a \quad \text{and} \quad (u_0-u_d-G_\Sigma(\bar u)-G'_\Sigma(\bar u)(u_0-\bar u))(x,t)\le-\delta \tag{4.6}$$

*hold for a.a. $(x,t)\in\Sigma$.*

How strong is this assumption? It turns out that it follows from Assumption 3, provided that $\kappa$ is large enough.

Corollary 2 *If Assumption 3 is satisfied, then there exists $\kappa_1 \geq \kappa_0$ such that the linearized Slater condition is satisfied with $u_0 = 0$ for all $\kappa \geq \kappa_1$, where $\kappa_0$ was defined in Lemma 2.*

Proof By inserting the function $u_0 = 0$ in the linearized Slater condition (4.6), taken at $\bar{u}_\kappa$, we estimate the term $u_d + G_\Sigma(\bar{u}_\kappa) + G'_\Sigma(\bar{u}_\kappa)(-\bar{u}_\kappa)$. Using Lemma 2, we infer the existence of $\kappa_1 \geq \kappa_0$ such that

$$\|G_\Sigma(\bar{u}_\kappa) - G_\Sigma(0)\|_{L^\infty(\Sigma)} = \|\bar{y}_\kappa - y^0\|_{L^\infty(\Sigma)} \leq \delta/4$$

and

$$\|G'_\Sigma(\bar{u}_\kappa)\bar{u}_\kappa\|_{L^\infty(\Sigma)} \leq c\,\|\bar{u}_\kappa\|_{L^p(\Sigma)} \leq \delta/4,$$

provided that $\kappa \geq \kappa_1$. In the last estimate, we used the continuity of the mapping $u \mapsto G'_\Sigma(u)$ from $L^p(\Sigma)$ to $\mathcal{L}(L^p(\Sigma), L^\infty(\Sigma))$. Invoking Assumption 3, we obtain for all $\kappa \geq \kappa_1$ that

$$\begin{aligned} u_d + G_\Sigma(\bar{u}_\kappa) + G'_\Sigma(\bar{u}_\kappa)(-\bar{u}_\kappa) &= u_d + G_\Sigma(0) + (G_\Sigma(\bar{u}_\kappa) - G_\Sigma(0)) - G'_\Sigma(\bar{u}_\kappa)\bar{u}_\kappa \\ &\geq \delta - \frac{1}{4}\delta - \frac{1}{4}\delta = \frac{1}{2}\delta. \end{aligned}$$

Hence, the linearized Slater condition is satisfied with $\delta/2$. □

We have $G'_\Sigma(\bar{u})(u_0 - \bar{u}) = z_{u_0 - \bar{u}|\Sigma}$, where $z_{u_0-\bar{u}}$ is the unique solution to the linearized equation (2.22) with $v = u_0 - \bar{u}$; moreover, it holds that $G(\bar{u}) = y_{\bar{u}}$.

Corollary 3 (Linearization of the control problem) *Let $\bar{u}$ be a local solution of the control problem with associated state $\bar{y} = y_{\bar{u}}$ and assume that the linearized Slater condition* (4.6) *holds. Let $u \in L^\infty(\Sigma)$ obey the inequalities*

$$u_a \leq u(x,t) \leq u_d + \bar{y}(x,t) + z_{u-\bar{u}}(x,t) \tag{4.7}$$

*a.e. in $\Sigma$, where $z_{u-\bar{u}} = G'(\bar{u})(u - \bar{u})$ is the unique solution to the linearized equation* (2.22) *with $y_u = \bar{y}$ and $v = u - \bar{u}$ taken in the right-hand side. Then the inequality*

$$f'(\bar{u})(u - \bar{u}) + \kappa j(u) - \kappa j(\bar{u}) \geq 0 \tag{4.8}$$

*is satisfied for the functionals $f$ and $j$, defined in* (3.4).

The corollary follows straightforward from the general Theorem 4.1 after inserting the mappings related to the optimal control problem (3.1)-(3.2).

Let us denote the set of controls satisfying the linearized constraints by

$$L = \{u \in L^\infty(\Sigma) : u_a \leq u(x,t) \leq u_d + \bar{y}(x,t) + z_{u-\bar{u}}(x,t) \text{ a.e. in } \Sigma\}.$$

Thanks to the variational inequality (4.8), the control $\bar{u}$ solves the convex optimal control problem

$$\min\{f'(\bar{u})u + \kappa j(u) : u \in L\}. \tag{4.9}$$

Therefore, by subdifferential calculus, the following result is obtained as immediate conclusion:

Corollary 4 *Under the assumptions of Corollary 3, there is some $\bar{\lambda} \in \partial j(\bar{u})$ such that the following variational inequality is satisfied:*

$$f'(\bar{u})(u - \bar{u}) + \kappa \int_\Sigma \bar{\lambda}(x,t)(u(x,t) - \bar{u}(x,t))\, d\sigma dt \geq 0 \quad \forall u \in L. \tag{4.10}$$

We recall that a function $\bar{\lambda}$ belongs to the subdifferential $\partial j(\bar{u})$ if and only if

$$\bar{\lambda}(x,t) = \begin{cases} 1 & \text{if} \quad \bar{u}(x,t) > 0 \\ \in [-1,1] & \text{if} \quad \bar{u}(x,t) = 0 \\ -1 & \text{if} \quad \bar{u}(x,t) < 0. \end{cases}$$

For linearization, we considered $u$ and $\bar{u}$ as elements of $L^\infty(\Sigma)$, because the convex cone $C = (L^\infty(\Sigma))^-$ has a non-empty interior. Now, we formally extend the convex set $L$ to the space $L^2(\Sigma)$. It turns out that this set does not change by the transfer to $L^2(\Sigma)$. This can be confirmed by inverse isotony. Let us first prove the following linearized version of Theorem 2.3:

Corollary 5 *Let $u \in L^2(\Sigma)$ satisfy the inequality*

$$u(x,t) \leq u_d + \bar{y}(x,t) + z_{u-\bar{u}}(x,t) \quad \text{a.e. in } \Sigma.$$

*Then $u(x,t) \leq v(x,t)$ holds a.e. in $\Sigma$, where $v \in L^2(\Sigma)$ is the unique solution of*

$$v(x,t) = u_d + \bar{y}(x,t) + z_{v-\bar{u}}(x,t).$$

*Moreover, $v$ belongs to $L^\infty(\Sigma)$.*

Proof Analogously to Theorem 2.3, we obtain that any $u$ that obeys the assumption of the Corollary satisfies $u(x,t) \leq v(x,t)$, where $v$ is the unique solution to the equation

$$v(x,t) = u_d + \bar{y}(x,t) + z_{v-\bar{u}}(x,t) \tag{4.11}$$

and $z_{v-\bar{u}}$ is the solution to the linearized equation (2.22) with $v := v - \bar{u}$ as control. The solution $v$ of (4.11) can be constructed as in the proof of Theorem 2.3, part a). Indeed, we take $z \in W(0,T) \cap L^\infty(Q)$ as the solution of

$$\begin{aligned} \partial_t z - \Delta z + \frac{\partial d}{\partial y}(x,t,\bar{y})\, z &= 0 && \text{in } Q \\ \partial_n z + \frac{\partial b}{\partial y}(x,t,\bar{y})\, z - z &= u_d + \bar{y} - z_{\bar{u}} && \text{in } \Sigma \\ z(x,0) &= 0 && \text{in } \Omega \end{aligned} \tag{4.12}$$

and set $v(x,t) = u_d + \bar{y}(x,t) - z_{\bar{u}} + z(x,t)$ for $(x,t) \in \Sigma$. Then, the identity $z = z_v$ holds and $v(x,t) = u_d + \bar{y}(x,t) + z_{v-\bar{u}}(x,t)$ for $(x,t) \in \Sigma$. □

The solution $v$ is bounded and measurable. In view of this, if $u \in L$, then also $u$ is bounded, because we have $u(x,t) \in [u_a, v(x,t)]$ for a.a. $(x,t) \in \Sigma$. This shows that all $u \in L^2(\Sigma)$ that obey the inequalities (4.7) automatically belong to $L^\infty(\Sigma)$. Analogously to Lemma 1, there exists a constant $M_L > 0$, such that

$$\|u\|_{L^\infty(\Sigma)} \le M_L \qquad \forall u \in L.$$

Therefore, we have

$$L = \{u \in L^2(\Sigma) : u_a \le u(x,t) \le u_d + y_{\bar{u}}(x,t) + z_{u-\bar{u}}(x,t) \text{ a.e. in } \Sigma\}.$$

Notice that $G'(\bar{u})$ was initially defined in $L^p(\Sigma)$ via the linearized partial differential equation (2.22). By the solution properties of this equation, the mapping $v \mapsto z_v|_\Sigma$ is also linear and continuous in $L^2(\Sigma)$. Therefore, $G'(\bar{u})$ can be continuously extended to $L^2(\Sigma)$. Let us introduce the notation $S = G'_\Sigma(\bar{u})$ for this extended operator, i.e. $S : L^2(\Sigma) \to L^2(\Sigma)$. Then we have

$$Sv := G'_\Sigma(\bar{u})v = z_{v|\Sigma} \quad \forall v \in L^2(\Sigma).$$

Moreover, $L$ can be written down as follows

$$L = \{u \in L^2(\Sigma) : u_a \le u(x,t) \le u_d + y_{\bar{u}}(x,t) + S(u-\bar{u})(x,t) \text{ a.e. in } \Sigma\}.$$

Clearly, the nonnegativity of the operator $G'_\Sigma(\bar{u})$ extends from $L^\infty(\Sigma)$ to $L^2(\Sigma)$. Therefore, also $S$ is nonnegative, i.e.

$$v \ge 0 \implies Sv \ge 0 \quad \forall v \in L^2(\Sigma).$$

Obviously, this nonnegativity is also true for the adjoint operator $S^* : L^2(\Sigma) \to L^2(\Sigma)$.

REMARK 2 *Let us mention another property of $S$ that is used in the Appendix: The operator $I - S$ is bijective in $L^2(\Sigma)$. This is deduced analogously to the proof of Corollorary 5: For any $g \in L^2(\Sigma)$, the unique solution of $v - Sv = g$ is given by $v = g + z_{|\Sigma}$, where $z$ is the solution to the linear parabolic equation* (4.12) *with $g$ substituted for $u_d + \bar{y} - z_{\bar{u}}$ in the right-hand side. The continuity of $I - S$ and the open mapping theorem imply that $I - S$ is an isomorphism in $L^2(\Sigma)$.*

It is well known that the term $f'(\bar{u})(u - \bar{u})$ can be expressed in a more explicit way on using an adjoint state. We define the adjoint state $\bar{\varphi}$ as the unique solution to the *adjoint equation*

$$\begin{aligned} -\partial_t\bar{\varphi} - \Delta\bar{\varphi} + \frac{\partial d}{\partial y}(x,t,\bar{y})\,\bar{\varphi} &= \nu_Q(\bar{y} - y_Q) && \text{in } Q \\ \partial_n\bar{\varphi} + \frac{\partial b}{\partial y}(x,t,\bar{y})\,\bar{\varphi} &= 0 && \text{in } \Sigma \\ \bar{\varphi}(x,T) &= \nu_T(\bar{y}(x,T) - y_T(x)) && \text{in } \Omega. \end{aligned} \tag{4.13}$$

Then we have

$$f'(\bar{u})(u-\bar{u}) = \int_\Sigma (\bar{\varphi}(x,t) + \nu\bar{u}(x,t))(u(x,t) - \bar{u}(x,t))\, d\sigma dt. \tag{4.14}$$

Summarizing our findings, we obtain the following first-order condition:

Theorem 4.2 (Preliminary necessary optimality condition) *Let $\bar{u}$ be a local solution to the optimal control problem* (3.1)-(3.2) *that satisfies the linearized Slater condition* (4.6) *of Assumption 4. Then, there are an element $\bar{\lambda} \in \partial j(\bar{u})$ and a unique adjoint state $\bar{\varphi} \in W(0,T) \cap L^\infty(Q)$ such that*

$$\int_\Sigma (\bar{\varphi}(x,t) + \nu\bar{u}(x,t) + \kappa\bar{\lambda}(x,t))(u(x,t) - \bar{u}(x,t))\, d\sigma dt \geq 0 \quad \forall u \in L \tag{4.15}$$

*holds, where $\bar{\varphi}$ is the unique solution to* (4.13).

Given $\mu \in L^2(\Sigma)$, by standard arguments we can easily check that $S^*\mu = \phi_\Sigma$, where $\phi \in W(0,T)$ is the unique solution of

$$\begin{aligned} -\partial_t\phi - \Delta\phi + \frac{\partial d}{\partial y}(x,t,\bar{y})\,\phi &= 0 \quad \text{in } Q \\ \partial_n\phi + \frac{\partial b}{\partial y}(x,t,\bar{y})\,\phi &= \mu \quad \text{in } \Sigma \\ \phi(x,T) &= 0 \quad \text{in } \Omega. \end{aligned} \tag{4.16}$$

### 4.3. Lagrange multiplier rule and minimum principle

Let us start this section by a short exposition on the notion of Lagrange multipliers, associated with the mixed control-state constraint. We can have two different views, but both lead to the same result. The first is the consideration of the original nonlinear and non-differentiable problem (3.5). The associated Lagrangian function that "eliminates" only the upper constraint by a Lagrange multiplier $\mu$, is

$$L(u,\mu) = f(u) + \kappa j(u) + \int_\Sigma (u - u_d - G_\Sigma(u))\,\mu\, d\sigma dt.$$

A function $\bar{\mu} \in L^2(\Sigma)$ is called a Lagrange multiplier associated with a local solution $\bar{u}$ of (3.5), if $\bar{\mu} \geq 0$, $\bar{u}$ obeys the necessary optimality conditions for the optimization problem $\min_{u \geq u_a} L(u,\bar{\mu})$, and the complementarity condition

$$\int_\Sigma (\bar{u} - u_d - G_\Sigma(\bar{u}))\,\bar{\mu}\, d\sigma dt = 0 \tag{4.17}$$

is satisfied. The expected necessary optimality conditions are

$$f'(\bar{u})(u-\bar{u}) + \int_\Sigma [\kappa\bar{\lambda}(u-\bar{u}) + ((u-\bar{u}) - G'_\Sigma(\bar{u})(u-\bar{u}))\bar{\mu}]\, d\sigma dt \geq 0 \quad \forall u \geq u_a,$$

where $\bar\lambda \in \partial j(\bar u)$. Upon inserting (4.14) and with $S = G'_\Sigma(\bar u)$, this amounts to

$$\int_\Sigma (\bar\varphi + \nu\bar u + \kappa\bar\lambda + \bar\mu - S^*\bar\mu)(u - \bar u)\, d\sigma dt \geq 0 \quad \forall u \geq u_a. \tag{4.18}$$

A second point of view regarding the Lagrange multiplier rule is the linearized optimal control problem: If Assumption 4 is fulfilled, then - owing to Theorem 4.3 - the locally optimal control $\bar u$ solves the linear optimization problem with mixed control-state constraints

$$\min \int_\Sigma (\bar\varphi(x,t) + \nu\bar u(x,t) + \kappa\bar\lambda(x,t))\, u(x,t)\, d\sigma dt \geq 0 \tag{4.19}$$

subject to $u \in L^2(\Sigma)$ and

$$u_a \leq u(x,t) \leq (u_d + y_{\bar u} + S(u - \bar u))(x,t) \tag{4.20}$$

for a.a. $(x,t) \in \Sigma$, where $\bar\varphi$ is the unique solution to (4.13). Now we introduce the Lagrangian function associated with this problem, where we include the linearized mixed control-state constraint by a multiplier $\mu \in L^2(\Sigma)$. This Lagrangian is

$$\mathcal{L}(u,\mu) = \int_\Sigma (\bar\varphi + \nu\bar u + \kappa\bar\lambda)\, u\, d\sigma dt + \int_\Sigma (u - u_d - y_{\bar u} - S(u - \bar u))\, \mu\, d\sigma dt.$$

The associated necessary optimality conditions consist – again with a Lagrange multiplier $\bar\mu$ – of

$$\frac{\partial \mathcal{L}}{\partial u}(\bar u, \bar\mu)(u - \bar u) \geq 0 \quad \forall u \geq u_a,$$

joint with the complementarity condition, associated with the linearized upper inequality constraint. A simple computation reveals that this amounts to the same conditions as above, namely to (4.17)-(4.18). This observation shows that any Lagrange multiplier $\bar\mu \in L^2(\Sigma)$ for the linearized problem, associated with $\bar u$, is also a multiplier for the original nonlinear and non-differentiable problem and vice versa. We are justified to concentrate on the linearized problem.

Assumption 5 *The lower bound $u_a$ is assumed to be feasible for the linearized problem, i.e. there holds*

$$u_a - u_d - y_{\bar u} - S(u_a - \bar u) \leq 0 \quad \text{a.e. in } \Sigma.$$

Remark 3 *(i) If $u_a$ satisfies Assumption 5, then it is close to being a Slater point.*

*Indeed, if the inequality holds in the stronger form*

$$u_a - u_d - y_{\bar u} - S(u_a - \bar u) \leq e - \delta$$

*with some $\delta > 0$, then $u_0 := u_a$ is a Slater point. Observe that Assumption 5 is satisfied for every $\kappa$ large enough if $u_a - u_d - y^\emptyset - Su_a \leq e - \delta$ for some $\delta > \emptyset$.*

*(ii) The linearized Slater Assumption 4 and Assumption 5 are independent conditions. None of them can be deduced from the other one. Due to Assumption 4, we can apply the linearization technique. Of course, this assumption also leads to the existence of a Lagrange multiplier $\bar\mu \in L^\infty(\Sigma)^*$. However, one of our main goals is to obtain a more regular multiplier, namely $\bar\mu \in L^\infty(\Sigma)$. For this purpose, we use Assumption 5.*

Theorem 4.3 (Existence of a Lagrange multiplier) *Let Assumption 5 be fulfilled and let $\bar u \in L^2(\Sigma)$ solve the linear optimization problem* (4.19)-(4.20). *Denote by $\bar y := y_{\bar u}$ the associated state. Then there exists a non-negative Lagrange multiplier $\bar\mu \in L^\infty(\Sigma)$ associated with the mixed control-state constraint. This is a function $\bar\mu$ such that*

$$\int_\Sigma (\bar\varphi + \nu\,\bar u + \kappa\,\bar\lambda + \bar\mu - S^*\bar\mu)(u - \bar u)\,d\sigma dt \geq 0 \quad \forall u \geq u_a \tag{4.21}$$

*holds and the following complementarity conditions are satisfied,*

$$\bar\mu \geq 0, \quad \bar u - \bar y - u_d \leq 0, \quad \int_\Sigma (\bar u - \bar y - u_d)\,\bar\mu\,d\sigma dt = 0. \tag{4.22}$$

In view of the similarities of the proof to an analogous one in Casas and Tröltzsch (2018a), we will sketch it in the Appendix. Notice that the Lagrange multiplier $\bar\mu$ is a Lagrange multiplier for the nonlinear optimal control problem (3.1)-(3.2), since it fulfills the conditions (4.17)-(4.18), characterizing a Lagrange multiplier.

Let us formulate the optimality conditions in a slightly different form that later will simplify the proof of sparsity properties of the optimal control.

Theorem 4.4 (Pointwise minimum principle) *Let the Assumptions 4 and 5 be satisfied. If $\bar u$ is a locally optimal control and $\bar\mu \in L^\infty(\Sigma)$ is an associated Lagrange multiplier that exists according to Theorem 4.3, then, for almost all $(x,t) \in \Sigma$, the value $u = \bar u(x,t)$ is a solution of the problem*

$$\min\,(\bar\varphi + \nu\,\bar u + \kappa\,\bar\lambda - S^*\bar\mu)(x,t)\,u \tag{4.23}$$

*subject to*

$$u_a \leq u \leq u_d + \bar y(x,t). \tag{4.24}$$

Proof The proof is similar to that of Theorem 2 in Casas and Tröltzsch (2018a). We rely on the variational inequality (4.21) that needs Assumption 4 for linearization and Assumption 5 for the existence of the Lagrange multiplier $\bar\mu$. For convenience, we define

$$e(x,t) = (\bar\varphi + \nu\,\bar u + \kappa\,\bar\lambda - S^*\bar\mu)(x,t).$$

For given $(x,t) \in \Sigma$, the minimum in (4.23) is attained by

$$u = \begin{cases} u_a, & \text{if} \quad e(x,t) > 0, \\ \in [u_a, u_d + \bar{y}(x,t)], & \text{if} \quad e(x,t) = 0, \\ u_d + \bar{y}(x,t), & \text{if} \quad e(x,t) < 0. \end{cases}$$

Assume that the result of the theorem is not true. Then one of the following two measurable sets $E_1$, $E_2$ must have positive measure,

$$\begin{aligned} E_1 &= \{(x,t) \in \Sigma : e(x,t) > 0 \text{ but } \bar{u}(x,t) > u_a\}, \\ E_2 &= \{(x,t) \in \Sigma : e(x,t) < 0 \text{ but } \bar{u}(x,t) < u_d + \bar{y}(x,t). \end{aligned}$$

In the points of $E_1$, we have $\bar{u}(x,t) > u_a$. Here, the variational inequality (4.21) can only hold, if

$$(\bar{\varphi} + \nu\,\bar{u} + \kappa\,\bar{\lambda} + \bar{\mu} - S^*\bar{\mu})(x,t) = 0 \text{ a.e. in } E_1,$$

hence $e(x,t) \leq 0$ follows from $\bar{\mu} \geq 0$, contradicting the definition of $E_1$. Therefore, $E_1$ cannot have positive measure.

To analyze the set $E_2$, we use that the variational inequality (4.21) implies

$$(\bar{\varphi} + \nu\,\bar{u} + \kappa\,\bar{\lambda} + \bar{\mu} - S^*\bar{\mu})(x,t) \geq 0 \tag{4.25}$$

a.e. in $\Sigma$. We also observe that in a.a. points of $E_2$ the multiplier $\bar{\mu}$ vanishes, because the upper (mixed control-state-) constraint is inactive. Hence, by $\bar{\mu} = 0$ in $E_2$, the inequality $e(x,t) = (\bar{\varphi} + \nu\,\bar{u} + \kappa\,\bar{\lambda} - S^*\bar{\mu})(x,t) \geq 0$ follows from (4.25), contradicting the definition of $E_2$. Therefore, also $E_2$ cannot have positive measure. This completes the proof. □

Remark 4 *The constraints* (4.24) *are* $u_a \leq u \leq u_d + \bar{y}$ *with fixed function* $\bar{y}$. *In this form, they are pointwise control constraints. This is characteristic for the so-called two-phase maximum principle, introduced by Grinold (1970) for continuous linear programming problems, here formulated as minimum principle.*

Let us conclude this section by a slight reformulation of the variational inequality (4.21). From (4.13) and (4.16) we infer that $\bar{\varphi}_{|\Sigma} - S^*\bar{\mu} = \bar{\psi}_{|\Sigma}$ , where $\bar{\psi}$ is the unique weak solution to the adjoint equation

$$\begin{aligned} -\partial_t \bar{\psi} - \Delta \bar{\psi} + \frac{\partial d}{\partial y}(x,t,\bar{y})\,\bar{\psi} &= \nu_Q(\bar{y} - y_Q) && \text{in } Q \\ \partial_n \bar{\psi} + \frac{\partial b}{\partial y}(x,t,\bar{y})\,\bar{\psi} &= -\bar{\mu} && \text{in } \Sigma \\ \bar{\psi}(x,T) &= \nu_T(\bar{y}(x,T) - y_T(x)) && \text{in } \Omega. \end{aligned} \tag{4.26}$$

By $\bar{\psi}$ and the minimum principle (4.23)-(4.24), the variational inequality (4.15) admits the final form

$$\int_\Sigma (\bar{\psi} + \nu\,\bar{u} + \kappa\,\bar{\lambda})(u - \bar{u})\,d\sigma dt \geq 0 \qquad \forall u \in L^2(\Sigma) : u_a \leq u \leq u_d + \bar{y}. \tag{4.27}$$

### 4.4. Sparsity of the optimal control

The variational inequality (4.27) is a useful tool for proving the following theorem on sparsity properties of $\bar{u}$ that is the main result of our paper:

Theorem 4.5 (Sparsity) *(i) Let the Assumptions 3, 4, and 5 be satisfied and let $\bar{u}$ be locally optimal for the control problem* (3.1)-(3.2). *Then, owing to Theorem 5.2 that is demonstrated in the Appendix independently of Theorem 4.5, a Lagrange multiplier $\bar{\mu} \in L^\infty(\Sigma)$, associated with the upper constraint of* (3.2) *exists. For any such Lagrange multiplier, the implications*

$$\begin{aligned} |\bar{\psi}(x,t)| \le \kappa \quad &\Rightarrow \quad \bar{u}(x,t) = 0 \\ \bar{u}(x,t) = 0 \quad &\Rightarrow \quad \bar{\psi}(x,t) \le \kappa \end{aligned} \tag{4.28}$$

*are satisfied for a.a.* $(x,t) \in \Sigma$ *with the adjoint state $\bar{\psi}$ solving* (4.26). *Moreover, if $\kappa \ge \kappa_0$ with $\kappa_0$ introduced in Lemma 2, we have that*

$$\bar{u}(x,t) = 0 \;\Leftrightarrow\; |\bar{\psi}(x,t)| \le \kappa \quad \text{for a.a. } (x,t) \in \Sigma. \tag{4.29}$$

*(ii) There is a value $\kappa_1 > 0$, such that $\bar{u} = 0$ holds for all sparse parameters $\kappa \ge \kappa_1$ such that Assumption 5 is satisfied.*

*(iii) The element $\bar{\lambda}$ of the subdifferential $\partial j(\bar{u})$ is given by*

$$\bar{\lambda}(x,t) = \mathbb{P}_{[-1,1]}\left\{-\frac{1}{\kappa}\bar{\psi}(x,t)\right\}. \tag{4.30}$$

Proof The main ideas are inspired by the proof of sparsity for pointwise control constraints from Casas, Herzog and Wachsmuth (2012). However, some changes are needed to tackle mixed control-state constraints.

(i) First, we confirm the sparsity relations (4.28). We define the sets

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

Let us show $(x,t) \in E_0 \Rightarrow \bar{\psi}(x,t) \le \kappa$. In $E_0$, we have $u_a < \bar{u}(x,t) \le u_d + \bar{y}(x,t)$. The lower inequality is not active. From the variational inequality (4.27) that needs the Assumptions 4 and 5, we find

$$0 \ge \bar{\psi}(x,t) + \nu\bar{u}(x,t) + \kappa\bar{\lambda}(x,t) = \bar{\psi}(x,t) + \kappa\bar{\lambda}(x,t) \ge \bar{\psi}(x,t) - \kappa \tag{4.31}$$

a.e. in $E_0$. Therefore, $\bar{\psi}(x,t) \le \kappa$ holds. This confirms the lower implication of (4.28). To show the upper one, assume conversely that $|\bar{\psi}(x,t)| \le \kappa$. A standard result for solutions $\bar{u}$ of the variational inequality (4.27) is the projection formula

$$\bar{u}(x,t) = \mathbb{P}_{[u_a\,,\,u_d + \bar{y}(x,t)]}\left\{-\nu^{-1}(\bar{\psi}(x,t) + \kappa\,\bar{\lambda}(x,t))\right\}.$$

In $E_+$, we have $\bar{\lambda}(x,t) = 1$, hence the projection formula implies

$$0 < -\frac{1}{\nu}(\bar{\psi}(x,t) + \kappa).$$

This yields $\bar{\psi}(x,t) < -\kappa$ and therefore $|\bar{\psi}(x,t)| > \kappa$, contrary to the assumption above. Analogously, the inequality $\bar{\psi}(x,t) > \kappa$ holds for a.a. $(x,t) \in E_-$. Therefore, $\bar{u}(x,t) = 0$ must be satisfied in a.a. points $(x,t)$ with $|\bar{\psi}(x,t)| \leq \kappa$; we have confirmed (4.28).

To show (4.29), we consider the points $(x,t)$, where $\bar{u}(x,t) = 0$ holds. We invoke Lemma 2 that implies $\frac{\delta}{2} \leq u_d + \bar{y}(x,t)$ for all $\kappa \geq \kappa_0$. In this case, for $\bar{u}(x,t) = 0$ also the upper inequality $\bar{u} \leq u_d + \bar{y}$ is inactive. Now, instead of (4.31), we obtain the equation

$$0 = \bar{\psi}(x,t) + \nu\bar{u}(x,t) + \kappa\bar{\lambda}(x,t) = \bar{\psi}(x,t) + \kappa\bar{\lambda}(x,t)$$

that yields $|\bar{\psi}(x,t)| \leq \kappa$. Along with the upper implication of (4.28), this proves (4.29).

(ii) According to Lemma 1, the sets of all feasible controls $u$ and associated states $y_u$ are bounded in $L^\infty(\Sigma)$. The same follows for the associated adjoint states $\varphi$, solving equation (4.13). The adjoint state $\bar{\psi}_{|\Sigma} = \bar{\varphi}_{|\Sigma} - S^*\bar{\mu}$ depends on $\bar{\mu}$. However, by Theorem 5.2, part (ii), we can assume the Lagrange multiplier $\bar{\mu}$ to be bounded in $L^\infty(\Sigma)$ by some $M_2 > 0$, independently of $\kappa$. Notice that in Theorem 5.2 we do not claim that all existing Lagrange multipliers are bounded. Therefore, we can assume

$$\|\bar{\psi}\|_{L^\infty(\Sigma)} \leq M_3$$

with some constant $M_3 > 0$ not depending on $\kappa$. Notice that the assumption $p > N + 1$ is invoked for this property. For all $\kappa \geq \kappa_1 = \max\{M_3, \kappa_0\}$, relation (4.29) yields $\bar{u} = 0$.

(iii) The projection formula (4.30) is standard, see the proof of Theorem 3 in Casas and Tröltzsch (2018a). □

Remark 5 (Use of the assumptions) *The general Assumptions 1 and 2 are needed for the whole theory of the optimal control problem. In view of Corollary 2, the linearized Slater Assumption 4 is not needed for all sufficiently large $\kappa$, if Assumption 3 holds. Moreover, it is not needed, if Assumption 5 is satisfied in strong form, see Remark 3. Therefore, it is sufficient to require Assumption 3 and Assumption 5 in strong form.*

## 5. Appendix

The results of this paper are based on a linearization method. The associated analysis occupies a major part of our presentation. After linearization, the further steps up to the final result on sparsity are similar to the ones in Casas and Troltzsch (2018a). However, they differ from the arguments in Casas and Troltzsch (2018a): Here, we do not include an upper bound $u \leq u_b$ on the control in addition to the other given constraints. This enables us to simplify the presentation. It might be time consuming for a reader to adapt the proofs of Casas and Troltzsch (2018a) to the problem posed here. Therefore, we briefly discuss the associated issues in this Appendix to provide the reader with the associated main ideas.

### 5.1. A pair of dual linear programming problems

The aim of this section is to prove the existence of a Lagrange multiplier, associated with the mixed control-state constraint that is a function and belongs to $L^\infty(\Sigma)$. The starting point is the variational inequality (4.15). In view of this inequality, the considered locally optimal control $\bar{u}$ is a solution to the following linear continuous optimization problem:

$$\min \int_\Sigma (\bar{\varphi}(x,t) + \nu\bar{u}(x,t) + \kappa\bar{\lambda}(x,t))\, u(x,t)\, d\sigma dt$$
$$u(x,t) \leq u_d + y_{\bar{u}}(x,t) + z_{u-\bar{u}}(x,t),$$
$$u(x,t) \geq u_a,$$
$$\text{for a.a. } (x,t) \in \Sigma,$$

where $z_{u-\bar{u}}$ is defined by the linearized equation (2.22).

To cover this problem by one of the standard forms of linear programming problems with nonnegativity constraints, we substitute

$$v := u - u_a$$

and recall that $z_v = Sv$. Then, the linearized mixed control-state constraint above reads

$$v \leq u_d - u_a + y_{\bar{u}} + S(u_a - \bar{u}) + Sv.$$

Now we set

$$b = u_d - u_a + y_{\bar{u}} + S(u_a - \bar{u}) \quad \text{and} \quad -a = \bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda},$$

and arrive at the following final form of our linear programming problem that we call the *primal problem* and denote it by (PP):

$$\max \int_\Sigma a(x,t)v(x,t)\, d\sigma dt$$
$$\text{subject to } v \in L^2(\Sigma) \text{ and} \tag{PP}$$
$$v \leq b + Sv$$
$$v \geq 0.$$

Here, all inequalities are to be understood pointwise a.e. in $\Sigma$. Notice that $a$ belongs to $L^\infty(\Sigma)$, since all of the functions defining $a$ are bounded. We know by our construction that $\bar v = \bar u - u_a$ is a solution of (PP). Moreover, the function $b$ defined above is nonnegative, if Assumption 5 is fulfilled. Indeed, this assumption is equivalent to $b \ge 0$; notice that $z_{u_a - \bar u} = S(u_a - \bar u)$.

### 5.2. Dual problem

Our next goal is to confirm the existence of a Lagrange multiplier, associated with the upper bound of (PP). We apply the duality theory of continuous linear programming problems as in Casas and Tröltzsch (2018a). That technique has the particular advantage that Lagrange multipliers are constructed in a fairly explicit way. This is useful to prove some special sparsity properties of locally optimal controls of the optimal control problem (3.1)-(3.2).

To this aim, we establish the dual problem to (PP); it is

$$\begin{aligned} &\min \int_\Sigma b(x,t)\,\mu(x,t)\,d\sigma dt \\ &\qquad \text{subject to } \mu \in L^2(\Sigma) \text{ and} \\ &\mu \ge a + S^*\mu \\ &\mu \ge 0. \end{aligned} \tag{DP}$$

This dual problem can be constructed as the Lagrangian dual to the conic linear programming (PP); we refer to Bonnans and Shapiro (2000), Section 2.5.6. Two questions have to be answered: The first is the equality of the maximum of (PP) with the infimum of its dual problem (DP), the so-called strong duality. The second is the solvability of (DP), i.e. if the infimum is attained as minimum. We begin with the strong duality, which is more difficult than the second question.

The duality theorem of Casas and Tröltzsch (2018a) was inspired by early results for linear programming problems of bottleneck type, see, e.g., Grinold (1970). The techniques for such problems initiated duality results for more general linear programming problems that rely on the so-called *boundedness condition*, introduced in Grinold (1970). We refer, for instance, to Krabs (1968). This condition fits very well the mixed control-state constraints as in our paper. It does not require that the cone of nonnegative functions be non-nempty, as Slater type conditions assume. Moreover, for our problem, it can be easily confirmed, see Lemma 3 below.

The linear programming problem (PP) satisfies a *boundedness condition* that ensures the equality of the primal optimal value with the associated dual one. We refer to our exposition in Casas and Tröltzsch (2018a). To confirm this condition, we define for given $d \in L^2(\Sigma)$ the set

$$P(d) = \{v \in L^2(\Sigma) : v \le d + Sv,\ v \ge 0\}.$$

This is the feasible set of (PP), associated with varying right-hand side $d$ in the upper bound of (PP).

Lemma 3 (Boundedness condition) *There exists $\eta > 0$ independent of $d$, such that*

$$\|v\|_{L^2(\Sigma)} \leq \eta\, \|d\|_{L^2(\Sigma)} \quad \forall v \in P(d).$$

Proof Owing to Theorem 2.3 on inverse isotony that also holds for the linearized equation, the following property is satisfied: If $v$ belongs to $P(d)$, then we have $v \leq w$, where

$$w = d + Sw.$$

The solution $w$ exists and belongs to $L^2(\Sigma)$. Using that $I - S$ is boundedly invertible (see Remark 2), it follows that

$$\|v\|_{L^2(\Sigma)} \leq \|w\|_{L^2(\Sigma)} = \|(I-S)^{-1}d\|_{L^2(\Sigma)} \leq \|(I-S)^{-1}\|_{\mathcal{L}(L^2(\Sigma))}\, \|d\|_{L^2(\Sigma)}.$$

Therefore, the desired estimate holds true with $\eta = \|(I-S)^{-1}\|_{\mathcal{L}(L^2(\Sigma))}$. □

Corollary 6 *The primal problem (PP) and the dual problem (DP) have the same optimal value, i.e.*

$$\max_{\substack{v \leq b + Sv \\ v \geq 0}} \int_\Sigma a\, v\, d\sigma dt \quad = \quad \inf_{\substack{\mu \geq a + S^*\mu \\ \mu \geq 0}} \int_\Sigma b\, \mu\, d\sigma dt.$$

Proof In view of Lemma 3, the boundedness condition of Assumption 2 in Casas and Tröltzsch (2018a) is satisfied. The claim follows from the general duality Theorem 7 in Casas and Tröltzsch (2018a). □

### 5.3. Solvability of the dual problem and Lagrange multipliers

To show the existence of Lagrange multipliers, again we follow the lines of Casas and Tröltzsch (2018a). We have to show that (DP) has an optimal solution. Then the infimum of the dual problem is a minimum and the solution of (DP) is a Lagrange multiplier for (PP).

Theorem 5.1 (Dual existence) *If Assumption 5 is satisfied, then the dual problem (DP) has at least one optimal solution $\bar\mu$ that belongs to $L^\infty(\Sigma)$ and satisfies the equation*

$$\mu(x,t) = \max\{0\,,\, a(x,t) + (S^*\mu)(x,t)\}. \tag{5.32}$$

Proof The proof is analogous to the steps that prove Theorem 4 in Casas and Tröltzsch (2018a). We refer to Lemma 7-8 of Casas and Tröltzsch (2018a).

a) *The feasible set of (DP) is not empty:*

We take the unique solution $\hat{\mu}$ of the equation

$$\hat{\mu} = |a| + S^* \hat{\mu}.$$

By inverse isotony, the non-negativity of the operator $S^*$ implies that $\hat{\mu} \geq 0$ holds. Obviously, $\hat{\mu}$ satisfies the mixed constraint of (DP), since we have $|a(x,t)| \geq a(x,t)$.

b) *Construction of a dual solution*

We begin with a feasible solution $\mu$ for the dual problem that does not satisfy the equation (5.32). Then, there is a set $E \subset \Sigma$ with positive measure, where

$$\mu(x,t) > \max\{0\,,\, a(x,t) + (S^*\mu)(x,t)\} \qquad \text{for a.a. } (x,t) \in E.$$

Let $E_+$ be the subset of $E$, where the function $a + S^*\mu$ is nonnegative and $E_-$ be the subset of $E$, where it is negative. We construct a new feasible solution $\tilde{\mu}$ that is smaller than $\mu$ on $E$.

For all $(x,t) \in E_+$, we have $\mu(x,t) > a(x,t) + (S^*\mu)(x,t)$. Here, we set

$$\tilde{\mu}(x,t) := a(x,t) + (S^*\mu)(x,t) < \mu(x,t).$$

In $E_-$, there holds $\mu(x,t) > 0$. Here, we fix $\tilde{\mu}(x,t) := 0$ and hence

$$\mu(x,t) > \tilde{\mu}(x,t) = 0 > a(x,t) + (S^*\mu)(x,t).$$

In $\Sigma \setminus E$ we define $\tilde{\mu}(x,t) := \mu(x,t)$.

By this construction, we have $\tilde{\mu}(x,t) \leq \mu(x,t)$, hence

$$\int_\Sigma b(x,t)\tilde{\mu}(x,t)\, dxdt \leq \int_\Sigma b(x,t)\mu(x,t)\, dxdt$$

follows from $b \geq 0$. Here, Assumption 5 enters. Moreover, since the operator $S^*$ is nonnegative, we find

$$\tilde{\mu}(x,t) \geq a(x,t) + (S^*\mu)(x,t) \geq a(x,t) + (S^*\tilde{\mu})(x,t) \quad \text{for a.a. } (x,t) \in \Sigma.$$

Therefore, $\tilde{\mu}$ is feasible for the dual problem and has an objective value not exceeding that of $\mu$. Furthermore, the following identity holds

$$\tilde{\mu}(x,t) = \max\{0\,,\, a(x,t) + (S^*\mu)(x,t)\} \qquad \text{for a.a. } (x,t) \in \Sigma.$$

Repeating this process with $\mu := \tilde{\mu}$ and continuing in this way, we construct a sequence $\{\mu_n\}$ that is pointwise non-increasing and converges to a function $\bar{\mu}$. Moreover, by construction, $\mu_{n+1}$ satisfies

$$\mu_{n+1}(x,t) = \max\{0\,,\, a(x,t) + (S^*\mu_n)(x,t)\} \qquad \text{for a.a. } (x,t) \in \Sigma.$$

Passing to the limit in this identity, we infer that $\bar{\mu}$ satisfies the equation (5.32). Moreover, it has an objective value not greater than that of the initial function $\mu$. Therefore, we can restrict the search for an optimum to solutions $w$ of (5.32).

c) *Existence of an optimal solution to (DP)*

The set of functions $\mu$ satisfying (5.32) is bounded in $L^\infty(\Sigma)$. Indeed, for all such functions, there holds

$$\mu \leq |a| + S^*\mu.$$

By inverse isotony, it follows that $\mu \leq w$, where $w \in L^\infty(\Sigma)$ is the unique solution to

$$w = |a| + S^* w. \tag{5.33}$$

We already pointed out that the function $a$ belongs to $L^\infty(\Sigma)$, hence the same holds true for all $\mu$ that obey (5.32). Therefore, the search for a minimum of (DP) can be restricted to a bounded subset of $L^\infty(\Sigma) \subset L^2(\Sigma)$ that is sequentially weakly compact in $L^2(\Sigma)$. The objective functional is linear and continuous. Hence, the existence of an optimal solution follows immediately. □

It turns out that any (optimal) solution of (DP) is a Lagrange multiplier for the mixed control state constraint of the optimal control problem, associated with the optimal solution $(\bar{y}, \bar{u})$ :

Lemma 4 *Any solution $\bar{\mu}$ of (DP) is a Lagrange multiplier for the mixed control state constraint of the optimal control problem.*

Proof As a solution to (DP), $\bar{\mu}$ is a Lagrange multiplier for (PP). This is a standard result of duality theory. We set up the Lagrangian function $\mathcal{L}(v, \mu)$ for (PP) and use the optimality condition

$$\partial_v \mathcal{L}(\bar{v}, \bar{\mu})(v - \bar{v}) \geq 0 \quad \forall v \geq 0.$$

Then we obtain

$$\int_\Sigma [(-a)(v - \bar{v}) + ((v - \bar{v}) - S(v - \bar{v}))\,\bar{\mu}]\, d\sigma dt \geq 0 \quad \forall v \geq 0.$$

Moreover, the complementarity condition

$$\int_\Sigma (\bar{v} - S\bar{v} - b)\bar{\mu}\, d\sigma dt = 0$$

is satisfied. After re-substituting $v = u - u_a$, $b = u_d - u_a + y_{\bar{u}} + z_{u_a - \bar{u}}$ and $z_{u_a - \bar{u}} = Sv$, $v - \bar{v} = u - \bar{u}$, $-a = \bar{\varphi} + \nu\bar{u} + \kappa\bar{\lambda}$, an easy computation shows the equivalence of the last two formulas with (4.21). Therefore, $\bar{\mu}$ is a Lagrange multiplier. □

### 5.4. Uniform boundedness of the multiplier $\mu$ with respect to $\kappa$

We know that the function $-a = \bar\varphi + \kappa\bar\lambda + \nu\bar u$ is bounded for each fixed $\kappa$. However, its $L^\infty(\Sigma)$-norm might tend to infinity as $\kappa \to \infty$. The aim of this subsection is to show that this cannot happen and $\|a\|_{L^\infty(\Sigma)}$ remains bounded, independently of $\kappa$. Then also the solution $w$ of (5.33) is bounded independently of $\kappa$ and the same property follows for at least one Lagrange multiplier $\mu$, satisfying $\mu \le w$.

This uniform boundedness is needed to show the property of sparsity that the optimal control $\bar u$ vanishes if $\kappa$ is sufficiently large. In this section, it is useful to indicate the dependence of the solutions to the control problems on $\kappa$ by an associated index. For this purpose, we shall write $\bar u_\kappa$, $\bar y_\kappa$, $a_\kappa$, $\bar\varphi_\kappa$, $\bar\mu_\kappa$ etc. for the optimal quantities of the problems associated to $\kappa$.

Theorem 5.2 (Uniform boundedness of Lagrange multipliers) *Let the Assumptions 3 and 5 be satisfied. Then,*

*(i) for all $\kappa \ge \kappa_0$ with $\kappa_0$ introduced in Lemma 2, we find an optimal solution $\bar\mu_\kappa$ of (DP), satisfying the inequality*

$$\bar\mu_\kappa(x,t) \le |\bar\varphi_\kappa(x,t)| + (S^*\bar\mu_\kappa)(x,t) \text{ a.e. in } \Sigma. \tag{5.34}$$

*(ii) a constant $\hat M > 0$ not depending on $\kappa \ge 0$ exists, such that*

$$\|\bar\mu_\kappa\|_{L^\infty(\Sigma)} \le \hat M \tag{5.35}$$

*is satisfied for at least one optimal solution $\bar\mu_\kappa$ of (DP).*

Proof Part (i): In view of Theorem 5.1, there is a solution $\bar\mu_\kappa$ of (DP) that obeys

$$\bar\mu_\kappa(x,t) = \max\{0\,,\, a_\kappa(x,t) + (S^*\bar\mu_\kappa)(x,t)\} \text{ a.e. in } \Sigma. \tag{5.36}$$

Now we distinguish between two cases for $(x,t)$:

*Case 1:* $a_\kappa(x,t) + (S^*\bar\mu_\kappa)(x,t) \le 0$. Then $\bar\mu_\kappa(x,t) = 0$ and we can estimate

$$\bar\mu_\kappa(x,t) = 0 \le |\bar\varphi_\kappa(x,t)| + (S^*\bar\mu_\kappa)(x,t);$$

notice that $S^*\bar\mu_\kappa \ge 0$ follows from $\bar\mu_\kappa \ge 0$.

*Case 2:* $a_\kappa(x,t) + (S^*\bar\mu_\kappa)(x,t) > 0$. Here, we get from (5.36)

$$\bar\mu_\kappa(x,t) = a_\kappa(x,t) + (S^*\bar\mu_\kappa)(x,t).$$

To verify the claim in Case 2, we show for the associated points $(x,t)$ that $a_\kappa(x,t) \le |\bar\varphi_\kappa(x,t)|$ holds. Then, (5.34) follows immediately.

We recall that $a_\kappa(x,t) = -\bar\varphi_\kappa(x,t) - \nu\,\bar u_\kappa(x,t) - \kappa\bar\lambda_\kappa(x,t)$. The further discussion depends on the sign of $\bar u_\kappa(x,t)$.

*Case 2a,* $\bar{u}_\kappa(x,t) \le 0$: Here, the inequality $\bar{v}_\kappa \le b_\kappa + S\bar{v}_\kappa$ of (PP) is inactive, if $\kappa$ is large enough. Let us show this: as above, we set $\bar{v}_\kappa = \bar{u}_\kappa - u_a$. By inserting this and the definition of $b$ in the inequality $v \le b + Sv$, we find

$$\bar{u}_\kappa - u_a \le u_d - u_a + y_{\bar{u}_\kappa} + S(u_a - \bar{u}_\kappa) + S(\bar{u}_\kappa - u_a).$$

Simplifying, we find $\bar{u}_\kappa \le u_d + \bar{y}_\kappa$. By Lemma 2, we conclude that

$$u_d + \bar{y}_\kappa(x,t) \ge \frac{\delta}{2}$$

a.e. in $\Sigma$ for all sufficiently large $\kappa$, say $\kappa \ge \kappa_0$. Therefore, the inequality $0 = \bar{u}_\kappa \le u_d + \bar{y}_\kappa$ is inactive in Case 2a. In view of this, $\bar{\mu}_\kappa(x,t) = 0$ follows from the complementary condition (4.22) for all $\kappa \ge \kappa_0$.

Then, however, the restriction of (DP) implies $\bar{\mu}_\kappa = 0 \ge a_\kappa + S^*\bar{\mu}_\kappa$ in the associated points $(x,t)$ and hence we are not in Case 2, contradicting our initial assumption. In this way, we found out that $\bar{u}_\kappa(x,t) > 0$ must hold a.e. in Case 2. This is the subject of the next case:

*Case 2b,* $\bar{u}_\kappa(x,t) > 0$: Then, $\bar{\lambda}_\kappa(x,t) = 1$ holds, hence we found the desired inequality

$$a_\kappa(x,t) = -\bar{\varphi}_\kappa(x,t) - \nu\bar{u}_\kappa(x,t) - \kappa < -\bar{\varphi}_\kappa(x,t) \le |\bar{\varphi}_\kappa(x,t)|.$$

Part (ii): As a consequence of Lemma 1, all possible adjoint states $\bar{\varphi}_\kappa$, associated to feasible controls, are uniformly bounded, hence there exists a number $M_1 > 0$ such that $\|\bar{\varphi}_\kappa\|_{L^\infty(\Sigma)} \le M_1 \quad \forall \kappa \ge 0$. Due to (5.34), we can assume

$$\bar{\mu}_\kappa(x,t) \le M_1 + (S^*\bar{\mu}_\kappa)(x,t) \text{ a.e. in } \Sigma \quad \forall \kappa \ge \kappa_0. \tag{5.37}$$

For all $0 \le \kappa \le \kappa_0$, we have for at least one solution $\bar{\mu}_\kappa$ of (DP) that

$$\begin{aligned}\bar{\mu}_\kappa(x,t) &\le |(a_\kappa + S^*\bar{\mu}_\kappa)(x,t)| = |-(\bar{\varphi}_\kappa + \nu\,\bar{u}_\kappa + \kappa\bar{\lambda})(x,t) + (S^*\bar{\mu}_\kappa)(x,t)| \\ &\le M_1 + \nu M_F + \kappa_0 + (S^*\bar{\mu}_\kappa)(x,t) = M_2 + (S^*\bar{\mu}_\kappa)(x,t).\end{aligned}$$

In view of (5.37), the inequality

$$\bar{\mu}_\kappa(x,t) \le M_2 + (S^*\bar{\mu}_\kappa)(x,t) \text{ a.e. in } \Sigma \quad \forall \kappa \ge 0$$

holds with $M_2 \ge M_1$. By inverse isotony, we obtain $0 \le \bar{\mu}_\kappa(x,t) \le w(x,t)$ for the selected multiplier $\bar{\mu}_\kappa$ that obeys (5.36), where $w \in L^\infty(\Sigma)$ is the unique solution to

$$w(x,t) = M_2 + (S^*w)(x,t).$$

Therefore, we obtain $\|\bar{\mu}\|_{L^\infty(\Sigma)} \le \hat{M} := \|w\|_{L^\infty(\Sigma)}$ that verifies (5.35). □

## References


Bonnans, F. and Shapiro, A. (2000), *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York.

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

Casas, E. and Tröltzsch, F. (2018a) Sparse optimal control for the heat equation with mixed control-state constraints. To appear in *Math. Control Rel. Fields*.

Casas, E. and Tröltzsch, F. (2018b) State-constrained semilinear elliptic optimization problems with unrestricted sparse controls. To appear in *Math. Control Rel. Fields*.

Collatz, L. (1952) Aufgaben monotoner Art. *Arch. Math.*, **3**, 366–376.

Collatz, L. (1964) *Funktionalanalysis und numerische Mathematik. Die Grundlehren der mathematischen Wissenschaften*, **120**, Springer-Verlag, Berlin.

Dautray, R. and Lions, J. (2000) *Mathematical Analysis and Numerical Methods for Science and Technology*, **5**. Springer-Verlag, Berlin-Heidelberg-New York.

Glashoff, K. and Werner, B. (1979) Inverse monotonicity of monotone $L$-operators with applications to quasilinear and free boundary value problems. *J. Math. Anal. Appl.*, **72** (1), 89–105.

Grinold, R.C. (1970) Symmetric duality for continuous linear programs. *SIAM J. Appl. Math.*, **18**, 84–97.

Grisvard, P. (1985) *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston-London-Melbourne.

Krabs, W. (1968) Lineare Optimierung in halbgeordneten Vektorräumen. *Numer. Math.*, **11**, 220–231.

Ladyzhenskaya, O., Solonnikov, V., and Ural'tseva, N. (1988) *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society, Providence.

Ortega, J.M. and Rheinboldt, W.C. (1967) Monotone iterations for nonlinear equations with application to Gauss-Seidel methods. *SIAM J. Numer. Anal.*, **4**, 171–190.

Pao, C.V. (1992) *Nonlinear Parabolic and Elliptic Equations*. Plenum Press, New York.

Raymond, J.P. and Zidani, H. (1998) Pontryagin's principle for state-constrained control problems governed by parabolic equations with unbounded controls. *SIAM J. Control Optim.*, **36** (6), 1853–1879.

Raymond, J.R. and Zidani, H. (1999) Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations. *Appl. Math. Optim.*, **39** (2), 143–177.

Redheffer, R. and Walter, W. (1979) Inequalities involving derivatives. *Pacific J. Math.*, **85** (1), 165–178.

Rheinboldt, W.C. (1969) *On M-functions and their application to nonlinear Gauss-Seidel iterations and to network flows.* BMwF-GMD-23, Gesellschaft für Mathematik und Datenverarbeitung, Bonn.

Rösch, A. and Tröltzsch, F. (2007) On regularity of solutions and Lagrange multipliers of optimal control problems for semilinear equations with mixed pointwise control-state constraints. *SIAM J. Control and Optimization*, **46** (3), 1098–1115.

Temam, R. (1979) *Navier-Stokes Equations.* North-Holland, Amsterdam.

Tröltzsch, F. (1979) A minimum principle and a generalized bang-bang principle for a distributed optimal control problem with constraints on control and state. *Z. Angew. Math. Mech.*, **59** (12), 737–739.