

SPARSE OPTIMAL CONTROL FOR THE HEAT EQUATION WITH MIXED CONTROL-STATE CONSTRAINTS

EDUARDO CASAS

Departamento de Matemática Aplicada y Ciencias de la Computación
E.T.S.I. Industriales y de Telecomunicación
Universidad de Cantabria
39005 Santander, Spain

FREDI TRÖLTZSCH

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

(Communicated by Enrique Zuazua)

Dedicated to Prof. Dr. Frédéric Bonnans on the occasion of his 60th birthday

ABSTRACT. A problem of sparse optimal control for the heat equation is considered, where pointwise bounds on the control and mixed pointwise control-state constraints are given. A standard quadratic tracking type functional is to be minimized that includes a Tikhonov regularization term and the L^1 -norm of the control accounting for the sparsity. Special emphasis is laid on existence and regularity of Lagrange multipliers for the mixed control-state constraints. To this aim, a duality theorem for linear programming problems in Hilbert spaces is proved and applied to the given optimal control problem.

1. Introduction. In a bounded domain $\Omega \subset \mathbb{R}^N$ with Lipschitz boundary Γ , we investigate the following problem of optimal sparse control:

$$\min J(y, u) := \int_0^T \int_{\Omega} \left(\frac{1}{2} |y - y_Q|^2 + \frac{\nu}{2} |u|^2 + \kappa |u| \right) dx dt \quad (1.1)$$

subject to the parabolic initial-boundary value problem

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

2010 *Mathematics Subject Classification.* Primary: 49K20, 49N10; Secondary: 90C05, 90C46.

Key words and phrases. Heat equation, optimal control, sparse control, mixed control-state constraints.

The first author was partially supported by Spanish Ministerio de Economía, Industria y Competitividad under projects MTM2014-57531-P and MTM2017-83185-P. The second author was supported by the Collaborative Research Center SFB 910, TU Berlin, project B6.

* Corresponding author: Fredi Tröltzsch.

and to the mixed pointwise control-state constraints

$$u_a \leq u(x, t) \leq u_b, \quad (1.3)$$

$$u(x, t) \leq u_d + y(x, t) \quad (1.4)$$

to be fulfilled for a.a. $(x, t) \in Q := \Omega \times (0, T)$.

In this problem, a desired state function $y_Q \in L^q(Q)$ with some $q > N/2 + 1$, $q \geq 2$ if $N = 1$, and a fixed final time $T > 0$ are given. We denote by ∂_n the outward normal derivative on Γ and set $\Sigma := \Gamma \times (0, T)$.

Remark 1.1. The integrability index $q > N/2 + 1$ of y_Q is required for proving the boundedness of the adjoint states $\bar{\varphi}$ and $\bar{\psi}$ defined in (3.2) and (3.10), respectively. This is needed whenever the existence of Lagrange multipliers $\mu_i \in L^\infty(Q)$, $i = 1, 2$, associated with the two upper constraints is claimed. In particular, part (ii) of Theorem 3.8 on sparsity relies on bounded Lagrange multipliers. If the statement “ $(\bar{\mu}_1, \bar{\mu}_2) \in (L^\infty(Q))^2$ ” is replaced by “ $(\bar{\mu}_1, \bar{\mu}_2) \in (L^2(Q))^2$ ”, then associated results remain true for $q = 2$.

Moreover, we fix real constants $\nu > 0$ (Tikhonov parameter), $\kappa > 0$ (sparse parameter) and real bounds $u_a < 0$, $u_b > 0$, $u_d > 0$. The bounds might be functions as well, but we keep them constant for simplicity. We will consider in parallel the case $u_b = \infty$, where the upper bound for the control is missing.

We assume homogeneous initial data to simplify the presentation. The extension to a non-homogeneous initial condition requires only some obvious modifications. The weak solution $y \in W(0, T) := \{y \in L^2(0, T; H^1(\Omega)) : \partial_t y \in L^2(0, T; H^1(\Omega)')\}$ of (1.2) is the state of our control system, while the function $u \in L^2(Q)$ is the control.

The main novelty of our paper is the discussion of sparse optimal controls for a problem that includes pointwise mixed control-state constraints. While the issue of sparsity has been discussed quite extensively for problems with pointwise control constraints, the extension to mixed control-state constraints seems to be new. It requires special emphasis on existence and regularity of Lagrange multipliers associated with the mixed control-state constraints. To the best knowledge of the authors, this case has not yet been discussed in literature. The associated main result is Theorem 3.8 on sparsity of optimal controls.

To prove this theorem, we resolve two main difficulties: First, we show the existence of a Lagrange multiplier for inequality (1.4) that belongs to $L^\infty(Q)$. We cannot rely on a Slater type assumption since this would eventually lead to a Lagrange multiplier in $L^\infty(Q)^*$; notice that the cone of non-negative measurable functions has a non-empty interior only in $L^\infty(Q)$. As a key tool for overcoming this obstacle, we apply the duality theory for linear programming problems in Banach spaces, cf. Krabs [6] and Tröltzsch [11]. Since the papers [6, 11] are published in German, the associated duality theorem is recalled and improved in the Appendix. The proof of the duality theorem is presented for Hilbert spaces.

Another difficulty is to show the boundedness of the Lagrange multiplier, uniformly with respect to the sparse parameter κ . This property is stated in Theorem 4.7. The result on existence and uniform boundedness of the Lagrange multiplier associated with the mixed control-state constraint (1.4) has some auxiliary character. Nevertheless, it is a second main achievement of our paper that is the core of Theorem 3.4.

The technique of our paper can be extended to some other types of mixed control-state constraints, where the existence of bounded Lagrange multiplier functions can be shown by our duality method. For instance, the two-sided constraints $u_a \leq u \leq u_d + y$ or $u_d + y \leq u \leq u_b$ can be handled completely analogous by obvious modifications. In the latter case, the new control $v = -u$ can be introduced to transform the lower mixed control-state constraint to an upper one.

For the two-sided pure mixed constraints $u_c + y \leq u \leq u_d + y$, the existence of bounded Lagrange multipliers can be proved by introducing the function $v := u - y$ as a new control. Then these constraints are transformed to pure control constraints, where associated Lagrange multipliers can be easily constructed. Sparsity results can then be discussed after re-transforming the problem to $u = v + y$. The same holds true, if an additional bound on the control u is added. Then we have one control constraint and two-sided mixed control-state constraints. Here, the transformation $v = u - y$ leads to two-sided box constraints for v and one mixed control-state constraint. Then the existence of bounded Lagrange multipliers can be shown again as in this paper. The discussion of sparsity will need some modifications.

However, our method does not work for the full set of four-sided constraints

$$u_a \leq u \leq u_b \quad \text{and} \quad u_c + y \leq u \leq u_d + y.$$

Here, we were not able to prove the existence of associated Lagrange multipliers by duality. This difficulty is known since long time. There is a detour via a Slater assumption in $L^\infty(Q)$, Lagrange multipliers in $L^\infty(Q)^*$, and certain structural assumptions on active sets as in [10]. Unfortunately, the structural assumption cannot in general be verified in advance and has to be required.

2. Well-posedness of the optimal control problem.

2.1. The reduced optimization problem. It is well known that the state equation (1.2) is uniquely solvable: For each control $u \in L^2(Q)$ there exists a unique weak solution $y \in W(0, T)$ that we denote by y_u . The mapping $u \mapsto y_u$ is linear and continuous, cf. Lions [8]. We consider this mapping with range in $L^2(Q)$ and denote it by S , i.e. $S : L^2(Q) \rightarrow L^2(Q)$, $Su := y_u$.

Throughout the paper, for functions $u \in L^2(Q)$, we write $u \geq 0$ if and only if $u(x, t) \geq 0$ holds for almost all $(x, t) \in Q$.

In this way, the optimal control problem (1.1)-(1.4) can be re-formulated as

$$\min_{u \in C} f(u) := \frac{1}{2} \|Su - y_Q\|_{L^2(Q)}^2 + \frac{\nu}{2} \|u\|_{L^2(Q)}^2 + \kappa \|u\|_{L^1(Q)} \quad (\text{P})$$

where the *feasible set* $C \subset L^2(Q)$ is defined by

$$C := \{u \in L^2(Q) : u_a \leq u \leq u_b, \quad u \leq u_d + Su\}.$$

The next results are basic for our theory and will be frequently used in this paper.

Lemma 2.1. *If $u \in L^2(Q)$ is almost everywhere non-negative, then also Su has this property, i.e.*

$$u \geq 0 \Rightarrow Su \geq 0.$$

Proof. This is a well-known conclusion from the comparison principle for the linear heat equation with homogeneous initial and boundary data: If $u \geq 0$ holds, then the solution y of (1.2) is a.e. nonnegative, too. This implies the claim. The comparison

principle can be proven by testing the variational formulation of the heat equation with the positive part y^+ of the weak solution y . \square

Lemma 2.2 (Inverse non-negativity). *For all $g \in L^2(Q)$, the equation*

$$u - Su = g \tag{2.1}$$

has a unique solution $u \in L^2(Q)$. If $g \geq 0$ holds, then also $u \geq 0$ is satisfied. Further, if $g \in L^\infty(Q)$, then u belongs to $L^\infty(Q)$ as well.

Proof. The equation $u - Su = g$ is equivalent to the statement $u - y = g$, where y is the solution of (1.2). Subtracting y from both sides of the PDE, we find

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

First of all, given $g \in L^2(Q)$, this equation has a unique weak solution y . From $g \geq 0$ and the comparison principle for parabolic equations, we obtain $y \geq 0$.

Adding now y to both sides, the differential equation reads

$$\partial_t y - \Delta y = g + y.$$

Setting $u = g + y$, we have $y = Su$ and $u - Su = g$, hence the existence of a solution u is shown. The uniqueness follows from the fact that $y = 0$ is the unique solution to $g = 0$. Now we see $u = y + g \geq 0$, since both g and y are nonnegative. Finally, if $g \in L^\infty(Q)$, then we get from [7, §III.7] that $y \in L^\infty(Q)$ holds as well and, hence, $u = y + g \in L^\infty(Q)$. \square

The next simple result will be applied several times in our paper.

Corollary 2.3. *If $u, v \in L^2(Q)$ satisfy the relations*

$$\begin{aligned} u &\leq g + Su, \\ v &= g + Sv \end{aligned}$$

with some $g \in L^2(Q)$, then $u \leq v$ holds.

Proof. We have $u = g - e + Su$ with some a.e. nonnegative $e \in L^2(Q)$. Subtracting this from the equation for v yields

$$v - u = e + S(v - u).$$

Lemma 2.2 implies $v - u \geq 0$, hence $u \leq v$ is fulfilled. \square

Next, we prove the existence of an optimal solution to (P).

Lemma 2.4. *The optimization problem (P), and hence the optimal control problem (1.1)-(1.4), has a unique optimal control \bar{u} .*

Proof. In view of the assumptions $u_a < 0$, $u_b > 0$, and $u_d > 0$, the set C is not empty, since $u = 0$ satisfies all constraints. If both u_a and u_b are real numbers, then the result is obvious, since the set $\{u \in L^2(Q) : u_a \leq u \leq u_b\}$ is weakly compact. The additional pointwise control-state constraint does not change this, i.e. C is weakly compact. Notice that C is convex and closed, hence also weakly closed. Uniqueness follows by strict convexity of the functional f .

Let us therefore concentrate on the case $u_b = \infty$, where the set of admissible controls is defined by the constraints

$$u_a \leq u \leq u_d + Su.$$

From Corollary 2.3 we know that $u \leq v$ holds a.e. in Q , where $v \in W(0, T) \cap L^\infty(Q)$ is the solution to the equation $v = u_d + Sv$. Therefore, the feasible set C is bounded also in this case. Again, we can invoke weak compactness to get the existence of an optimal solution \bar{u} . \square

Remark 2.5. An inspection of the proof shows that the set C is bounded in $L^\infty(Q)$.

3. Optimality conditions and sparsity of optimal controls.

3.1. A variational inequality. In this section, we derive first order optimality conditions for the optimal control \bar{u} of (P). The main result is the “two-phase minimum principle” of Theorem 3.6.

We define the following particular functionals $I : L^2(Q) \rightarrow \mathbb{R}$ and $j : L^1(Q) \rightarrow \mathbb{R}$ that are included in the reduced objective functional f of (P):

$$\begin{aligned} I(u) &= \frac{1}{2} \|Su - y_Q\|_{L^2(Q)}^2 + \frac{\nu}{2} \|u\|_{L^2(Q)}^2, \\ j(u) &= \|u\|_{L^1(Q)}. \end{aligned}$$

In this way, we have $f = I + \kappa j$, hence (P) can be re-written as

$$\min_{u \in C} f(u) := I(u) + \kappa j(u). \quad (\text{P})$$

The functional f is the sum of the convex differentiable functional I and the convex directionally differentiable functional κj .

In all what follows, by $(\cdot, \cdot)_H$ the scalar product of a real Hilbert space H is denoted.

Since C is convex, we obtain the following necessary and sufficient optimality condition for the optimal solution \bar{u} :

Lemma 3.1. *A control $\bar{u} \in C$ is optimal for (P) if and only if an element $\bar{\lambda}$ of the subdifferential $\partial j(\bar{u})$ exists such that the variational inequality*

$$(S^*(S\bar{u} - y_Q) + \nu \bar{u} + \kappa \bar{\lambda}, u - \bar{u})_{L^2(Q)} \geq 0 \quad \forall u \in C \quad (3.1)$$

is satisfied.

The proof is standard; notice that we have $I'(\bar{u}) = S^*(S\bar{u} - y_Q) + \nu \bar{u}$. Moreover, we mention that $\partial j(u) \subset L^\infty(Q)$ holds for all $u \in L^1(Q)$ and that $\lambda \in \partial j(u)$ if and only if

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

holds for a.a. $(x, t) \in Q$. It is well known that the Hilbert space adjoint operator $S^* : L^2(Q) \rightarrow L^2(Q)$ can be expressed in terms of an adjoint state. We have $S^*(S\bar{u} - y_Q) = \bar{\varphi}$, where the adjoint state $\bar{\varphi}$ is the unique solution of the adjoint equation

$$\begin{aligned} -\partial_t \varphi - \Delta \varphi &= \bar{y} - y_Q && \text{in } Q \\ \partial_n \varphi &= 0 && \text{in } \Sigma \\ \varphi(x, T) &= 0 && \text{in } \Omega. \end{aligned} \quad (3.2)$$

By the adjoint state $\bar{\varphi}$, the variational inequality (3.1) can be re-formulated as

$$\iint_Q (\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda})(u - \bar{u}) \, dx dt \geq 0 \quad \forall u \in C. \quad (3.3)$$

In other words, \bar{u} solves the linear optimization problem

$$\begin{aligned} & \min \iint_Q (\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda}) u \, dx dt \\ & \text{subject to} \\ & u - Su \leq u_d, \\ & u_a \leq u \leq u_b. \end{aligned} \tag{3.4}$$

The included mixed control-state constraint $u - Su \leq u_d$ is posed in $L^2(Q)$, where the interior of the cone of nonnegative functions is empty. Therefore, it is not an easy task to show the existence of associated Lagrange multipliers. In the case of mixed control state constraints, there are special techniques to overcome this difficulty, we refer exemplarily to Röscher and Tröltzsch [10] and the references therein.

Here, we will employ the duality theory of continuous linear programming problems to establishing the next result. A similar technique was published in [12], but the proof was not presented in all details.

The further analysis depends on the following assumption.

Assumption 3.2. *The control u defined by $u(x, t) = u_a \, \forall (x, t) \in Q$ belongs to the feasible set C , i.e.*

$$u_a \leq u_d + y_a(x, t) \quad \forall (x, t) \in Q,$$

where y_a is the solution of (1.2) associated with $u = u_a$.

This assumption is quite natural. In the context of optimal heating of Ω by a heat source u , the state y stands for the temperature. The mixed control state constraint excludes a too sudden heating, since the difference $u - y$ cannot exceed the bound u_d . In this sense, Assumption 3.2 means that the minimal heat source is not too large related to the associated temperature bounds.

For discussing the sparsity of the optimal control, we need the following convergence result:

Lemma 3.3. *Let, for $\kappa \geq 0$, \bar{u}_κ denote the optimal control of (P) corresponding to the sparse parameter $\kappa \geq 0$ and let $\bar{y}_\kappa = S\bar{u}_\kappa$ be the associated state. Then we have*

$$\lim_{\kappa \rightarrow \infty} \|\bar{u}_\kappa\|_{L^p(Q)} = 0 \quad \forall p \in [1, \infty)$$

and

$$\lim_{\kappa \rightarrow \infty} \|\bar{y}_\kappa\|_{C(\bar{Q})} = 0.$$

Proof. The control $u = 0$ is feasible for (P), hence

$$f(\bar{u}_\kappa) = \frac{1}{2} \|\bar{y}_\kappa - y_Q\|_{L^2(Q)}^2 + \frac{\nu}{2} \|\bar{u}_\kappa\|_{L^2(Q)}^2 + \kappa \|\bar{u}_\kappa\|_{L^1(Q)} \leq f(0) = \frac{1}{2} \|y_Q\|_{L^2(Q)}^2.$$

This immediately yields

$$\|\bar{u}_\kappa\|_{L^1(Q)} \leq \frac{1}{2\kappa} \|y_Q\|_{L^2(Q)}^2 \rightarrow 0, \quad \kappa \rightarrow \infty.$$

Moreover, we know from Remark 2.5 that the set of feasible controls is bounded in $L^\infty(Q)$. Therefore, the L^1 -convergence above implies also

$$\|\bar{u}_\kappa\|_{L^p(Q)} \rightarrow 0, \quad \kappa \rightarrow \infty \quad \forall p \in [1, \infty).$$

Selecting an arbitrary $p > \frac{N}{2} + 1$, the continuity of the mapping $u \mapsto y_u$ from $L^p(Q)$ to $C(\bar{Q})$, [7, §III-7], immediately implies that $\lim_{\kappa \rightarrow \infty} \|\bar{y}_\kappa\|_{C(\bar{Q})} = 0$. \square

In all what follows, to simplify the presentation, we suppress the dependence of the optimal quantities on κ .

The next theorem is basic for the investigation of sparsity properties of the optimal control. Its proof relies on results of the duality theory of linear continuous programming problems that is presented later in the Section 4. However, we formulate the theorem right here in order to come faster to the theory of sparsity. The results of Section 4 do not rely on Theorem 3.4.

Theorem 3.4. (i) *Let Assumption 3.2 be satisfied, \bar{u} be optimal for the control problem (1.1)-(1.4), and \bar{y} be the associated state. Then a pair $(\bar{\mu}_1, \bar{\mu}_2)$ of Lagrange multipliers $\bar{\mu}_i \in L^\infty(Q)$, $i = 1, 2$, exists such that the variational inequality*

$$\iint_Q (\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} + \bar{\mu}_1 + \bar{\mu}_2 - S^* \bar{\mu}_2)(u - \bar{u}) \, dxdt \geq 0 \quad (3.5)$$

holds for all $u \in L^2(Q)$ with $u \geq u_a$. Moreover, the complementarity conditions

$$\begin{aligned} \bar{\mu}_1 &\geq 0, \quad \bar{u} - u_b \leq 0, \quad \iint_Q (\bar{u} - u_b) \bar{\mu}_1 \, dxdt = 0, \\ \bar{\mu}_2 &\geq 0, \quad \bar{u} - \bar{y} - u_d \leq 0, \quad \iint_Q (\bar{u} - \bar{y} - u_d) \bar{\mu}_2 \, dxdt = 0 \end{aligned} \quad (3.6)$$

are satisfied.

(ii) *There exist a constant $M > 0$ not depending on κ and a pair of multipliers $(\bar{\mu}_1, \bar{\mu}_2)$ satisfying all conditions above such that $\|\bar{\mu}_i\|_{L^\infty(Q)} \leq M$, $i = 1, 2$, holds.*

Proof. Part (i): The existence of Lagrange multipliers $(\bar{\mu}_1, \bar{\mu}_2) \in (L^\infty(Q))^2$ is ensured by Theorems 4.5 and 4.6 that are proven in Section 4 by duality theory. Once this existence is shown, the variational inequality (3.5) and the complementarity conditions (3.6) express well-known properties of Lagrange multipliers. To see this, associated with the linear programming problem (3.4), we define the Lagrangian function

$$\begin{aligned} \mathcal{L}(u, \mu_1, \mu_2) &= \iint_Q (\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda}) u \, dxdt \\ &\quad + \iint_Q (u - u_b) \mu_1 \, dxdt + \iint_Q (u - Su - u_d) \mu_2 \, dxdt. \end{aligned}$$

Then (3.5) is the standard condition

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu}_1, \bar{\mu}_2)(u - \bar{u}) \geq 0 \quad \forall u \geq u_a,$$

while (3.6) contains the associated complementarity conditions.

Part (ii) on uniform boundedness of $(\bar{\mu}_1, \bar{\mu}_2)$ is nothing more than statement (ii) of Theorem 4.7. \square

Remark 3.5. In Theorem 3.4, we introduced Lagrange multipliers only for the two upper constraints, while the lower bound $u \geq u_a$ was not “eliminated” by a multiplier. We need the lower bound in the pointwise minimum principle below, hence it is not useful to eliminate it. Nevertheless, we might define a Lagrange multiplier $\bar{\mu}_3 \geq 0$ by

$$\bar{\mu}_3 := \bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} + \bar{\mu}_1 + \bar{\mu}_2 - S^* \bar{\mu}_2.$$

A simple discussion of the variational inequality (3.5) reveals that this function $\bar{\mu}_3$ is nonnegative, indeed. Moreover, let us introduce the Lagrangian function

$$\tilde{\mathcal{L}}(u, \mu_3) = \iint_Q \{(\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} + \bar{\mu}_1 + \bar{\mu}_2 - S^* \bar{\mu}_2) u + \mu_3(u_a - u)\} \, dx dt$$

that is associated with the variational inequality (3.5) and the lower bound constraint, eliminated by μ_3 . It is easy to confirm that the derivative $\partial \tilde{\mathcal{L}} / \partial u$ vanishes at $(\bar{u}, \bar{\mu}_3)$. This is a direct consequence of the definition of $\bar{\mu}_3$.

In addition the complementarity condition

$$\iint_Q \bar{\mu}_3(u_a - \bar{u}) \, dx dt = 0$$

is satisfied. Therefore, $\bar{\mu}_3$ defined above is a Lagrange multiplier associated with the lower bound constraint.

The Lagrange multiplier rule of Theorem 3.4 is not yet easily applicable to the discussion of sparsity. To facilitate this discussion, we perform an intermediate step and derive a pointwise minimum principle as a useful tool for proving sparsity properties of the optimal control \bar{u} . This type of minimum principles was introduced in the framework of continuous linear programming by Grinold [4]. It was later invoked to proving a generalized bang-bang principle in [12].

Theorem 3.6 (Pointwise minimum principle). *Let Assumption 3.2 be satisfied. If \bar{u} is an optimal control and $(\bar{\mu}_1, \bar{\mu}_2) \in (L^2(Q))^2$ is an associated pair of Lagrange multipliers that exists according to Theorem 3.4, then for almost all $(x, t) \in Q$ the solution u of the problem*

$$\min (\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} - S^* \bar{\mu}_2)(x, t) u \quad (3.7)$$

subject to

$$u_a \leq u \leq \min\{u_b, u_d + \bar{y}(x, t)\} \quad (3.8)$$

is attained by $u = \bar{u}(x, t)$.

Proof. For convenience, we define

$$e(x, t) = (\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} - S^* \bar{\mu}_2)(x, t).$$

For given $(x, t) \in Q$, the minimum in (3.7) is attained by

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

The minimum is attained at any value on the set, where $e(x, t) = 0$.

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,

$$E_1 = \{(x, t) \in Q : e(x, t) > 0 \text{ but } \bar{u}(x, t) > u_a\},$$

$$E_2 = \{(x, t) \in Q : e(x, t) < 0 \text{ but } \bar{u}(x, t) < \min\{u_b, u_d + \bar{y}(x, t)\}\}.$$

In the points of E_1 , we have $\bar{u}(x, t) > u_a$. Here, the variational inequality (3.5) can only hold, if

$$(\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} + \bar{\mu}_1 + \bar{\mu}_2 - S^* \bar{\mu}_2)(x, t) \leq 0$$

is true a.e. in E_1 . From $\bar{\mu}_i \geq 0$, $i = 1, 2$, we get a.e. in E_1

$$(\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} - S^* \bar{\mu}_2)(x, t) = e(x, t) \leq 0$$

contradicting the definition of E_1 . Therefore, E_1 cannot have positive measure.

In a.a. points of E_2 , the functions $\bar{\mu}_1$ and $\bar{\mu}_2$ vanish, because $\bar{\mu}_1$ is a Lagrange multiplier for the inequality $u \leq u_b$, $\bar{\mu}_2$ is a multiplier for the inequality $u \leq u_d + Su$, and both inequalities are inactive in E_2 at \bar{u} ; notice that $\bar{y} = S\bar{u}$.

Here, the variational inequality (3.5) can only be true, if

$$(\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} + \bar{\mu}_1 + \bar{\mu}_2 - S^* \bar{\mu}_2)(x, t) \geq 0 \quad (3.9)$$

holds a.e. in Q . Since a.e. in E_2 both multipliers vanish, we get

$$(\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda} - S^* \bar{\mu}_2)(x, t) = e(x, t) \geq 0$$

contradicting the definition of E_2 . Therefore, also E_2 cannot have positive measure. This completes the proof. \square

Remark 3.7. The constraints (3.8) have the form $u_a \leq u \leq \min\{u_b, u_d + \bar{y}\}$ with fixed function \bar{y} , hence in this context they are pointwise control constraints. This is characteristic for a so-called two-phase maximum principle introduced by Grinold [4] for continuous linear programming problems, here formulated as minimum principle.

Let us conclude this section by a slight reformulation of the variational inequality (3.5). We know that $\bar{\varphi} = S^*(\bar{y} - y_Q)$, hence we have $\bar{\varphi} - S^* \bar{\mu}_2 = S^*(\bar{y} - y_Q - \bar{\mu}_2) = \bar{\psi}$, where $\bar{\psi}$ is the unique weak solution to the adjoint equation

$$\begin{aligned} -\partial_t \psi - \Delta \psi &= \bar{y} - y_Q - \bar{\mu}_2 && \text{in } Q \\ \partial_n \psi &= 0 && \text{in } \Sigma \\ \psi(x, T) &= 0 && \text{in } \Omega. \end{aligned} \quad (3.10)$$

By $\bar{\psi}$ and the minimum principle (3.7)-(3.8), the variational inequality (3.5) admits the form

$$\iint_Q (\bar{\psi} + \nu \bar{u} + \kappa \bar{\lambda})(u - \bar{u}) \, dx dt \geq 0 \quad \forall u \in L^2(Q) : u_a \leq u \leq \min\{u_b, u_d + \bar{y}\}. \quad (3.11)$$

3.2. Sparsity of the optimal control. The variational inequality (3.11) is the main tool for proving sparsity properties of \bar{u} .

Theorem 3.8 (Sparsity). *(i) Let Assumption 3.2 be satisfied and let \bar{u} be optimal for the control problem (1.1)-(1.4). Then a pair of Lagrange multipliers $(\bar{\mu}_1, \bar{\mu}_2) \in (L^\infty(Q))^2$ exists such that the implications*

$$\begin{aligned} |\bar{\psi}(x, t)| \leq \kappa &\Rightarrow \bar{u}(x, t) = 0 \\ \bar{u}(x, t) = 0 &\Rightarrow \bar{\psi}(x, t) \leq \kappa \end{aligned} \quad (3.12)$$

are satisfied for a.a. $(x, t) \in Q$ with the adjoint state $\bar{\psi}$ solving (3.10). Moreover, for some constant $\kappa_0 > 0$ and for a.a. $(x, t) \in Q$ we have that

$$\bar{u}(x, t) = 0 \Leftrightarrow |\bar{\psi}(x, t)| \leq \kappa \quad \forall \kappa \geq \kappa_0. \quad (3.13)$$

(ii) There is a value $\kappa_1 > 0$, such that \bar{u} vanishes for all sparse parameters $\kappa \geq \kappa_1$.

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

$$\bar{\lambda}(x, t) = \mathbb{P}_{[-1, 1]} \left\{ -\frac{1}{\kappa} \bar{\psi}(x, t) \right\}, \quad (3.14)$$

where the projection $\mathbb{P}_{[-1, 1]} : \mathbb{R} \rightarrow [-1, 1]$ is defined by

$$\mathbb{P}_{[-1, 1]}(\alpha) = \max\{-1, \min\{1, \alpha\}\}.$$

Proof. The main ideas are inspired by the proof of sparsity for pointwise control constraints in [2]. However, some changes are needed to tackle mixed control-state constraints.

(i) First, we confirm the sparsity relations (3.12). We define the sets

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

Let us show $(x, t) \in E_0 \Rightarrow \bar{\psi}(x, t) \leq \kappa$. In E_0 , we have $u_a < \bar{u}(x, t) = 0 < u_b$ and $\bar{u}(x, t) \leq u_d + \bar{y}(x, t)$. The lower inequality is not active. From the variational inequality (3.11), we find

$$0 \geq \bar{\psi}(x, t) + \nu \bar{u}(x, t) + \kappa \bar{\lambda}(x, t) = \bar{\psi}(x, t) + \kappa \bar{\lambda}(x, t) \geq \bar{\psi}(x, t) - \kappa \quad (3.15)$$

a.e. in E_0 . Therefore, $\bar{\psi}(x, t) \leq \kappa$ must hold. This confirms the lower implication of (3.12). To show the upper one, assume conversely that $|\bar{\psi}(x, t)| \leq \kappa$. A standard result for solutions \bar{u} of the variational inequality (3.11) is the projection formula

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

For a.a. $(x, t) \in E_+$, 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 $|\bar{\psi}(x, t)| > \kappa$ in contrary to the assumption above. Hence the measure of E_+ must be zero. Analogously, E_- cannot have positive measure. 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 (3.12).

To show (3.13), we consider the points (x, t) , where $\bar{u}(x, t) = 0$ holds. We invoke Lemma 3.3 that implies $\|\bar{y}\|_{C(\bar{Q})} < u_d$ for all sufficiently large κ , say $\kappa \geq \kappa_1$. 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 (3.15), 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 (3.12), this proves (3.13).

(ii) According to Remark 2.5, the sets of all feasible controls u and associated states y_u are bounded in $L^\infty(Q)$. The same follows for the associated adjoint states φ solving equation (3.2). The adjoint state $\bar{\psi} = \bar{\varphi} - S^* \bar{\mu}_2$ depends on $\bar{\mu}_2$. However, by Theorem 3.4, part (ii), we can assume the Lagrange multiplier $\bar{\mu}_2$ to be bounded in $L^\infty(Q)$ by some $M > 0$, independently of κ . Therefore, we can assume

$$\|\bar{\psi}\|_{L^\infty(Q)} \leq M_1$$

with some constant $M_1 > 0$ not depending on κ . Notice that the assumption $q > N/2 + 1$ is invoked for this property. For all $\kappa \geq \kappa_0 = \max\{M_1, \kappa_1\}$, the relation (3.13) yields $\bar{u} = 0$.

(iii) The projection formula (3.14) is confirmed as follows: For a.a. $(x, t) \in E_+$, we have $\bar{\lambda}(x, t) = 1$. Moreover, (3.11) implies

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

and hence $\bar{\lambda}(x, t) = 1 \leq -\kappa^{-1}(\bar{\psi}(x, t) + \nu \bar{u}(x, t))$. Therefore, the projection formula is true a.e. in E_+ . In the same way, E_- can be treated.

In E_0 , all inequalities are inactive, hence $\bar{\psi}(x, t) + \bar{u}(x, t) + \kappa \bar{\lambda}(x, t) = 0$ follows from (3.11). Again, the projection formula (3.14) is an immediate conclusion. \square

Remark 3.9. In view of Lemma 3.3 and the fact that $u_d > 0$, the optimal state \bar{y} satisfies $u_d + \bar{y} > 0$ for all sufficiently large κ . For such values of κ , all inequalities of problem (P) are inactive in the points (x, t) where $\bar{u}(x, t) = 0$. Hence, for large κ , we were able to prove (3.13) in the same way as for pure pointwise control constraints, cf. [2].

4. Dual linear programming problems.

4.1. Dual problem and weak duality. Thanks to (3.4), the optimal control \bar{u} solves the linear continuous programming problem

$$\max_{u \in C} \iint_Q a(x, t) u(x, t) \, dx dt \quad (4.1)$$

where

$$a(x, t) = -(\bar{\varphi} + \nu \bar{u} + \kappa \bar{\lambda})(x, t).$$

By the definition of C , this problem is posed in the Hilbert space $L^2(Q)$.

Problem (4.1) does not yet have the standard form of continuous linear programming that we prefer. Instead of the lower bound $u \geq u_a$ we want to have a nonnegativity restriction. To this aim, we introduce a new (shifted) control

$$v = u - u_a$$

and transform problem (4.1) accordingly. Then \bar{u} is optimal for (P) if and only if $\bar{v} = \bar{u} - u_a$ is optimal for the problem

$$\begin{aligned} & \max \iint_Q a(x, t) v(x, t) \, dx dt \\ & \text{subject to the constraints} \\ & v(x, t) \leq v_b \\ & v(x, t) \leq b(x, t) + (Sv)(x, t) \\ & v(x, t) \geq 0 \end{aligned} \quad (\text{PP})$$

to be satisfied a.e. in Q ,

where $v_b := u_b - u_a > 0$ and the function b is defined by

$$b = u_d + Su_a - u_a.$$

If Assumption 3.2 is satisfied, then $b(x, t) \geq 0$ is fulfilled for a.a. $(x, t) \in Q$. We consider this problem in $L^2(Q)$, i.e. we assume $v \in L^2(Q)$ and also the constraints are viewed in $L^2(Q)$.

Problem (PP) is our *primal problem*. We know that $\bar{v} = \bar{u} - u_a$ solves this problem and want to prove the existence of bounded and measurable Lagrange multipliers associated with \bar{v} . For this purpose, we invoke the theory of linear programming problems in function spaces.

(PP) has the structure of linear continuous programming problems that were discussed extensively in the 1970ties, cf. Grinold [3, 4] and Tyndall [13], related to so-called bottleneck control problems for ordinary differential equations. For an extension to the control of partial differential equations, we refer to Tröltzsch

[12]. The associated *dual problem*, also considered in $L^2(Q)$, can be established by Lagrange duality. For $u_b < \infty$, it is the following:

$$\begin{aligned} & \min \iint_Q (v_b \mu_1(x, t) + b(x, t) \mu_2(x, t)) \, dx dt \\ & \text{subject to the constraints} \\ & \mu_1(x, t) + \mu_2(x, t) \geq a(x, t) + (S^* \mu_2)(x, t) \\ & \mu_1(x, t) \geq 0 \\ & \mu_2(x, t) \geq 0 \\ & \text{to be fulfilled a.e. in } Q. \end{aligned} \tag{DP}$$

If $u_b = \infty$, then the dual problem admits a simpler form, namely

$$\begin{aligned} & \min \iint_Q b(x, t) \mu(x, t) \, dx dt \\ & \text{subject to} \\ & \mu(x, t) \geq a(x, t) + (S^* \mu)(x, t) \\ & \mu(x, t) \geq 0 \\ & \text{a.e. in } Q. \end{aligned} \tag{DP}$$

In what follows, we concentrate on the case $u_b < \infty$. As a rule of thumb, the theory for the case $u_b = \infty$ remains true with $\mu_1 = 0$ and $\mu_2 =: \mu$.

Let us state some important properties of this pair of dual linear problems. To shorten the arguments, functions v and $\mu = (\mu_1, \mu_2)$ are said to be *feasible* for (PP) or (DP), respectively, if v satisfies all restrictions of (DP) and μ the ones of (DP).

Lemma 4.1 (Weak duality). *For all $v \in L^2(Q)$ that are feasible for the primal problem (PP) and all pairs $(\mu_1, \mu_2) \in L^2(Q)^2$ being feasible for the dual problem (DP), we have*

$$\iint_Q a(x, t) v(x, t) \, dx dt \leq \iint_Q (v_b \mu_1(x, t) + b(x, t) \mu_2(x, t)) \, dx dt. \tag{4.2}$$

Proof. The statement follows from Lemma 5.3 of the Appendix that we prove for more general pairs of dual linear programming problems posed in Hilbert spaces. \square

4.2. Strong duality. Next, we discuss the solvability of the dual problem. The dual feasible set is never empty:

For (DP), the pair $(\mu_1, \mu_2) = (|a|, 0)$ satisfies all constraints. If $u_b = \infty$, then the dual problem is (\widehat{DP}) . In this case, the solution μ of the equation $\mu = |a| + S^* \mu$ is feasible by inverse non-negativity.

For the solvability, the following result is essential:

Lemma 4.2. *Let Assumption 3.2 be satisfied. Then, to any feasible pair (μ_1, μ_2) for (DP), another feasible pair $(\widehat{\mu}_1, \widehat{\mu}_2)$ exists such that*

$$\widehat{\mu}_1(x, t) + \widehat{\mu}_2(x, t) = [a(x, t) + S^*(\widehat{\mu}_2)(x, t)]_+ \quad \forall (x, t) \in Q \tag{4.3}$$

and the objective value of $(\widehat{\mu}_1, \widehat{\mu}_2)$ is not larger, i.e.

$$\iint_Q (v_b \widehat{\mu}_1 + b \widehat{\mu}_2) \, dx dt \leq \iint_Q (v_b \mu_1 + b \mu_2) \, dx dt. \tag{4.4}$$

Proof. Let $E \subset Q$ be the set of all points with

$$\mu_1(x, t) + \mu_2(x, t) > [a(x, t) + S^*(\mu_2)(x, t)]_+ \quad \forall (x, t) \in E. \quad (4.5)$$

If the Lebesgue measure $|E|$ of E is zero, then we can easily modify μ_1, μ_2 on E such that (4.3) is satisfied everywhere in Q and (4.4) holds with equality. Assume therefore $|E| > 0$. We construct a componentwise and pointwise monotone decreasing sequence of feasible pairs $(\mu_{1,k}, \mu_{2,k}) \in (L^2(Q))^2$ converging to a pair $(\hat{\mu}_1, \hat{\mu}_2)$ having the desired properties. To this aim, we define the first element of the sequence by

$$\mu_{i,1} := \mu_i, \quad i = 1, 2,$$

and set $k = 1$. Let $E_0 := \{(x, t) \in E : [a(x, t) + S^*(\mu_{2,k})(x, t)]_+ = 0\}$ and $E_+ = E \setminus E_0$. The next feasible pair $(\mu_{1,k+1}, \mu_{2,k+1})$ with not greater objective functional value is constructed as follows: For $i = 1, 2$ we set

$$\mu_{i,k+1}(x, t) = \begin{cases} \mu_{i,k}(x, t), & (x, t) \in Q \setminus E, \\ 0, & (x, t) \in E_0, \\ \alpha(x, t)\mu_{i,k}(x, t), & (x, t) \in E_+, \end{cases}$$

where

$$\alpha(x, t) = \frac{[a(x, t) + S^*(\mu_{2,k})(x, t)]_+}{\mu_{1,k}(x, t) + \mu_{2,k}(x, t)}, \quad (x, t) \in E_+.$$

Notice that $0 < \alpha(x, t) \leq 1$. We have that

$$0 \leq \mu_{i,k+1}(x, t) \leq \mu_{i,k}(x, t), \quad i = 1, 2, \quad \forall (x, t) \in Q$$

and

$$\mu_{1,k+1}(x, t) + \mu_{2,k+1}(x, t) = [a(x, t) + (S^*\mu_{2,k})(x, t)]_+.$$

With S , also the operator S^* is nonnegative. This follows from

$$(S^*v, w)_{L^2(Q)} = (v, Sw)_{L^2(Q)} \geq 0 \quad \forall v \geq 0, w \geq 0.$$

The last inequalities and $S^* \geq 0$ ensure

$$\begin{aligned} & \mu_{1,k}(x, t) + \mu_{2,k}(x, t) \\ & \geq \mu_{1,k+1}(x, t) + \mu_{2,k+1}(x, t) = [a(x, t) + (S^*\mu_{2,k})(x, t)]_+ \\ & \geq [a(x, t) + (S^*\mu_{2,k+1})(x, t)]_+, \end{aligned} \quad (4.6)$$

hence $(\mu_{i,k+1}, \mu_{2,k+1})$ is feasible as well. Next, we set $k := k + 1$ and determine the next iterate.

In this way, pointwise monotone decreasing sequences of functions $(\mu_{i,k})_{k=1}^\infty$, $i = 1, 2$, are constructed. All pairs $(\mu_{1,k}, \mu_{2,k})$ are feasible for (DP). The sequences are bounded by $0 \leq \mu_{i,k} \leq \mu_i$, $i = 1, 2$, and pointwise convergent to measurable limit functions $\hat{\mu}_i$, $i = 1, 2$. In the limit, the equation (4.3) must be satisfied. This follows from passing to the limit $k \rightarrow \infty$ in (4.6).

By Assumption 3.2, we have that $b = u_d + Su_a - u_a \geq 0$ and $v_b = u_b - u_a \geq 0$. Therefore, the smaller the multipliers are, the smaller is the associated objective value of (DP), hence (4.4) must hold. \square

Corollary 4.3. *All feasible pairs (μ_1, μ_2) for (DP) that obey the equation (4.3) satisfy the inequalities*

$$0 \leq \mu_1(x, t) \leq |a(x, t)| + (S^*\mu_2)(x, t), \quad (4.7)$$

$$0 \leq \mu_2(x, t) \leq |a(x, t)| + (S^*\mu_2)(x, t). \quad (4.8)$$

Proof. The second inequality follows from

$$[a(x, t) + (S^* \mu_2)(x, t)]_+ \leq |a(x, t) + (S^* \mu_2)(x, t)| \leq |a(x, t)| + (S^* \mu_2)(x, t),$$

since $S^* \mu_2 \geq 0$. Now the first inequality is a conclusion of (4.3) and the non-negativity of $\bar{\mu}_1$ and $\bar{\mu}_2$. \square

Lemma 4.4. *If Assumption 3.2 is fulfilled, then the set of all feasible pairs (μ_1, μ_2) for (DP) satisfying the equation (4.3) is bounded in $(L^\infty(Q))^2$.*

Proof. By Corollary 4.3, (4.8) and Corollary 2.3 applied to S^* instead of S , we get

$$0 \leq \mu_2 \leq \mu,$$

where $\mu \geq 0$ is the solution to the equation

$$\mu = |a| + S^* \mu$$

that corresponds to the inequality (4.8). Since a is bounded and measurable, the solution μ enjoys the same property. This can be shown by a bootstrapping argument. First, we know $\mu \in L^2(Q)$. By the smoothing property of parabolic solution operators, we have that $S^* \mu \in L^\infty(0, T; L^2(\Omega))$. Hence, $\mu = |a| + S^* \mu$ belongs to $L^\infty(0, T; L^2(\Omega))$. Finally, from [7, Theorem III-7.1], we infer that $S^* \mu \in L^\infty(Q)$ and $\mu = |a| + S^* \mu \in L^\infty(Q)$.

By Corollary 4.3, (4.7), we also have $\|\mu_1\|_{L^\infty(Q)} \leq \|\mu\|_{L^\infty(Q)}$ completing the proof. \square

Later, we shall prove that this boundedness is even uniform with respect to the sparse parameter κ .

Theorem 4.5 (Dual existence). *If Assumption 3.2 is fulfilled, then the dual problem (DP) has at least one optimal solution $(\bar{\mu}_1, \bar{\mu}_2)$ that belongs to $(L^\infty(Q))^2$.*

Proof. By Lemma 4.2, the search for an optimal solution of (DP) can be restricted to the set of all feasible pairs (μ_1, μ_2) that solve equation (4.3). Thanks to Lemma 4.4, this set is bounded in $(L^\infty(Q))^2$, hence also in $(L^2(Q))^2$. Therefore, the search for the minimum can be restricted to the intersection of the feasible set of (DP) with a sufficiently large closed ball of $(L^2(Q))^2$. This is a non-empty, convex and closed set, hence weakly compact. Now the claim follows immediately. \square

Theorem 4.6 (Strong duality). *Let Assumption 3.2 be fulfilled, let \bar{v} be optimal for (PP) and $(\bar{\mu}_1, \bar{\mu}_2) \in (L^2(Q))^2$ be optimal for (DP). Then*

$$\iint_Q a(x, t) \bar{v}(x, t) \, dx dt = \iint_Q (v_b \bar{\mu}_1(x, t) + b(x, t) \bar{\mu}_2(x, t)) \, dx dt.$$

We prove this theorem in the Appendix. Thanks to the theorem, the optimal solution \bar{v} of the primal problem (PP) is a Lagrange multiplier for the inequality $\mu_1 + \mu_2 \geq a + S^* \mu_2$ of (DP) while any pair of solutions $(\bar{\mu}_1, \bar{\mu}_2)$ of (DP) is a pair of Lagrange multipliers for the upper restrictions of (PP), i.e. $\bar{\mu}_1$ is a multiplier for the inequality $v \leq u_b - u_a$ while $\bar{\mu}_2$ is one for the mixed control-state constraint $v - Sv \leq b$.

Therefore, the following complementarity conditions are satisfied by \bar{v} and $(\bar{\mu}_1, \bar{\mu}_2)$:

$$\begin{aligned} \bar{v} &\geq 0, \quad \bar{\mu}_1 + \bar{\mu}_2 \geq a + S^* \bar{\mu}_2, \quad \iint_Q \bar{v} (\bar{\mu}_1 + \bar{\mu}_2 - a - S^* \bar{\mu}_2) \, dx dt = 0, \\ \bar{\mu}_1 &\geq 0, \quad \bar{v} \leq v_b, \quad \iint_Q \bar{\mu}_1 (\bar{v} - v_b) \, dx dt = 0, \\ \bar{\mu}_2 &\geq 0, \quad \bar{v} \leq b + S\bar{v}, \quad \iint_Q \bar{\mu}_2 (\bar{v} - b - S\bar{v}) \, dx dt = 0. \end{aligned} \quad (4.9)$$

4.3. Uniform boundedness of Lagrange multipliers with respect to κ . All Lagrange multipliers μ_i satisfying the equation (4.3) are bounded and measurable. However, their $L^\infty(Q)$ -norm might grow with the sparse parameter κ . We show that this case can be excluded.

Theorem 4.7 (Uniform boundedness of Lagrange multipliers). *(i) There is some constant $\kappa_0 > 0$ with the following property: For all $\kappa \geq \kappa_0$, we find an optimal solution $(\bar{\mu}_1, \bar{\mu}_2)$ of (DP) satisfying the inequalities*

$$\bar{\mu}_i(x, t) \leq |\bar{\varphi}(x, t)| + (S^* \bar{\mu}_2)(x, t) \text{ a.e. in } Q, \quad i = 1, 2. \quad (4.10)$$

(ii) A constant $M > 0$ not depending on κ exists, such that

$$\|\bar{\mu}_i\|_{L^\infty(Q)} \leq M, \quad i = 1, 2, \quad (4.11)$$

is satisfied for at least one optimal solution $(\bar{\mu}_1, \bar{\mu}_2)$ of (DP).

Proof. Part (i): In view of Lemma 4.2 and Corollary 4.3, there is a solution $(\bar{\mu}_1, \bar{\mu}_2)$ of (DP) that obeys

$$\bar{\mu}_i(x, t) \leq [a(x, t) + (S^* \bar{\mu}_2)(x, t)]_+ \text{ a.e. in } Q, \quad i = 1, 2. \quad (4.12)$$

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

Case 1. $a(x, t) + (S^* \bar{\mu}_2)(x, t) \leq 0$. Then $\bar{\mu}_i(x, t) = 0$, hence we can estimate

$$\bar{\mu}_i(x, t) = 0 \leq |\bar{\varphi}(x, t)| + (S^* \bar{\mu}_2)(x, t), \quad i = 1, 2.$$

Case 2. $a(x, t) + (S^* \bar{\mu}_2)(x, t) > 0$. Here, we get from (4.12)

$$\bar{\mu}_i(x, t) \leq a(x, t) + (S^* \bar{\mu}_2)(x, t), \quad i = 1, 2.$$

To verify the claim in Case 2, we show for the associated points (x, t) that

$$a(x, t) \leq |\bar{\varphi}(x, t)|.$$

We recall that $a(x, t) = -\bar{\varphi}(x, t) - \nu \bar{u}(x, t) - \kappa \bar{\lambda}(x, t)$. The further discussion depends on the sign of $\bar{u}(x, t)$.

Case 2a. $\bar{u}(x, t) \leq 0$.

Here both upper inequalities of (PP) are inactive, if κ is large enough. Let us show this:

Since $u_b > 0$ and $\bar{u}(x, t) \leq 0$, the upper control constraint $\bar{u} \leq u_b$ cannot be active in this case. Moreover, we know from Lemma 3.3 that $\|\bar{y}\|_{L^\infty(Q)} < u_d$ holds for all sufficiently large κ , say $\kappa \geq \kappa_0$. Then we have $u_d + \bar{y} = u_d + S\bar{u} > 0$ a.e. in Q and the inequality $\bar{u} \leq u_d + S\bar{u}$ is also inactive.

Therefore, $\bar{\mu}_1(x, t) = \bar{\mu}_2(x, t) = 0$ follows from the second and third complementary condition of (4.9) in Case 2a, provided that $\kappa \geq \kappa_0$. The first restriction of

(DP) implies $0 \geq a(x, t) + (S^* \bar{\mu}_2)(x, t)$ and hence we are not in Case 2. We found out that $\bar{u}(x, t) > 0$ must hold a.e. in Case 2.

Case 2b. If $\bar{u}(x, t) > 0$, then $\bar{\lambda}(x, t) = 1$, hence

$$a(x, t) = -\bar{\varphi}(x, t) - \nu \bar{u}(x, t) - \kappa < -\bar{\varphi}(x, t) \leq |\bar{\varphi}(x, t)|.$$

Part (ii): For all $\kappa \leq \kappa_0$, we have for at least one solution $(\bar{\mu}_1, \bar{\mu}_2)$ of (DP) that

$$\begin{aligned} |a(x, t)| &= |\bar{\varphi}(x, t) + \nu \bar{u}(x, t) + \kappa \bar{\lambda} - (S^* \bar{\mu}_2)(x, t)| \\ &\leq |\bar{\varphi}(x, t)| + \nu |\bar{u}(x, t)| + \kappa_0 + (S^* \bar{\mu}_2)(x, t) \\ &\leq c + (S^* \bar{\mu}_2)(x, t), \end{aligned}$$

because the set of all feasible controls u and hence also the set of all possible adjoint states φ are uniformly bounded in $L^\infty(Q)$ with respect to κ , cf. Remark 2.5 and the proof of Lemma 4.4. Here, again the assumption $q > N/2 + 1$ enters.

Now we invoke the inequality (4.8) and get

$$\bar{\mu}_2 \leq c + S^* \bar{\mu}_2.$$

Since $I - S^*$ is inverse non-negative, we obtain the estimate

$$\bar{\mu}_2 \leq z,$$

where $z \in L^\infty(Q)$ is the unique solution to $z = c + S^* z$. This proves part (ii), since $\|\bar{\mu}_2\|_{L^\infty(Q)} \leq \|z\|_{L^\infty(Q)} =: M$. The estimate for $\bar{\mu}_1$ is now obtained from (4.7) by

$$|\bar{\mu}_1| \leq |a| + S^* \bar{\mu}_2 \leq c + S^* \bar{\mu}_2 = z \quad \text{a.e. in } Q.$$

□

Now we are able to complete the proof of Theorem 3.4.

Proof of Theorem 3.4. Part (i): The existence of a pair of Lagrange multipliers $(\bar{\mu}_1, \bar{\mu}_2) \in (L^\infty(Q))^2$ is ensured by Theorems 4.5 and 4.6. The variational inequality (3.5) and the complementarity conditions (3.6) are standard properties of Lagrange multipliers (cf. the remarks after Theorem 3.4).

Part (ii) on uniform boundedness of $\bar{\mu}_2$ is nothing more than statement (ii) of Theorem 4.7. □

5. Appendix – A duality theorem.

5.1. Duality theorem in Hilbert spaces. The result of this section will be proved for general pairs of dual linear programming problems that are not necessarily related to the specific pair of dual problems defined in Section 4. However, this general result will be applied to them.

Let $\{U, \|\cdot\|_U\}$ and $\{V, \|\cdot\|_V\}$ be real Hilbert spaces with inner products $(\cdot, \cdot)_U$ and $(\cdot, \cdot)_V$, respectively, $A : U \rightarrow V$ a linear and continuous operator, and let $K_U \subset U$ and $K_V \subset V$ be nonempty, convex and closed cones.

In U and V we define partial orderings \geq_U and \geq_V by $u \geq_U 0$ iff $u \in K_U$ and $v \geq_V 0$ iff $v \in K_V$. The converse inequalities \leq_U and \leq_V are defined accordingly, e.g. $u \leq_U 0$ iff $-u \geq_U 0$. Moreover, we fix elements $a \in U$ and $c \in V$.

We consider the *primal linear programming problem*

$$\max (a, u)_U \quad \text{subject to} \quad Au \leq_V c, \quad u \geq_U 0. \quad (\mathcal{PP})$$

By the Riesz theorem, we identify the dual spaces U^* of U and V^* of V with U and V , respectively. To the cones K_U and K_V , we associate dual cones K_U^* and K_V^* by

$$K_U^* = \{w \in U : (w, u)_U \geq 0 \quad \forall u \in K_U\}.$$

The dual cone $K_V^* \subset V$ is defined analogously.

Remark 5.1. If $U = L^2(Q)$ and K_U is the cone of a.e. nonnegative functions of U , then we have $K_U^* = K_U$. This is the case of interest for the application to our control problem (P), where $K_U = K_V = L^2(Q)^+$.

The partial orderings induced in U and V by K_U^* and K_V^* are denoted by \geq_U^* resp. \geq_V^* . We have

$$(w, u)_U \geq 0 \quad \forall w \geq_U^* 0, \forall u \geq_U 0.$$

Let $A^* : V \rightarrow U$ be the Hilbert space adjoint operator to A defined by $(Au, v)_V = (u, A^*v)_U \quad \forall u \in U, v \in V$.

The *dual problem* is defined by

$$\min (c, v)_V \quad \text{subject to} \quad A^*v \geq_U^* a, \quad v \geq_V^* 0. \quad (\mathcal{DP})$$

If an element $u \in U$ (resp. $v \in V$) obeys the constraints of (\mathcal{PP}) (resp. (\mathcal{DP})), then it is called *feasible* for the associated problem.

Remark 5.2. The pair of linear dual problems (\mathcal{PP}) and (\mathcal{DP}) fits in the duality theory of convex optimization, a well elaborated part of nonlinear programming. Associated strong duality theorems are often based on a Slater condition as in Luenberger [9] or on the assumption of calmness that is discussed in Bonnans and Shapiro [1], Sect. 1.5. A Slater condition cannot be satisfied in the space selected for our given optimal control problem and calmness is difficult to verify. Therefore, we invoke another method. We prove a duality theorem that relies on a certain boundedness condition. It turns out that this condition is fulfilled in our application, if Assumption 3.2 is satisfied.

Lemma 5.3 (Weak duality). *For all u that are feasible for (\mathcal{PP}) and all v being feasible for (\mathcal{DP}) , we have*

$$(a, u)_U \leq (c, v)_V.$$

Proof. We obtain

$$(a, u)_U \leq (A^*v, u)_U = (v, Au)_V \leq (v, c)_V,$$

where we employed the inequalities $u \geq_U 0$, $a \leq_U^* A^*v$ to get the first inequality and $v \geq_V^* 0$, $Au \leq_V c$ for the second one. \square

Let us define for (varying) $d \in V$ and $e \in U$ the sets

$$\begin{aligned} P(d) &= \{u \in U : u \geq_U 0, \quad Au \leq_V d\}, \\ D(e) &= \{v \in V : v \geq_V^* 0, \quad A^*v \geq_U^* e\}. \end{aligned}$$

Then $P(c)$ is the primal feasible set and $D(a)$ is the dual one. As a consequence of the last lemma, the relation of weak duality

$$\sup_{u \in P(c)} (a, u)_U \leq \inf_{v \in D(a)} (c, v)_V$$

is fulfilled. In this section, we show that the associated equality holds, if the *boundedness condition* below is satisfied for the primal problem.

Assumption 5.4 (Boundedness condition). *There exist $\eta > 0$ independent of d and, for all $d \in V$, a closed set $K(d) \subset P(d)$ such that the following two conditions are satisfied:*

$$\begin{aligned} \|u\|_U &\leq \eta \|d\|_V \quad \forall u \in K(d) \\ \forall u \in P(d) \quad \exists \hat{u} \in K(d) : \quad (a, u)_U &\leq (a, \hat{u})_U. \end{aligned}$$

Theorem 5.5 (Strong duality). *If the feasible set $P(c)$ of the primal problem (PP) is nonempty and Assumption 5.4 is satisfied, then the primal problem has an optimal solution. Moreover, the strong duality relation*

$$\max_{u \in P(c)} (a, u)_U = \inf_{v \in D(a)} (c, v)_V \quad (5.1)$$

is fulfilled.

Proof. (i) *Solvability of the primal problem.* Let $s = \sup_{u \in P(c)} (a, u)_U$ be the primal supremum. Since $P(c)$ is non-empty, we have $s > -\infty$. Thanks to the boundedness condition, the search for a maximum in the primal problem can be restricted to the bounded set $K(c)$. Taking a sequence $(u_n)_n$ of feasible elements for (PP) with $\lim_{n \rightarrow \infty} (a, u_n)_U = s$, we can assume without limitation of generality $u_n \in K(c)$. By the boundedness of $K(c)$, a weakly convergent subsequence can be selected, w.l.o.g. $(u_n)_n$ itself. In this way, we can assume $u_n \rightharpoonup \bar{u}$, $n \rightarrow \infty$. It is easy to verify that \bar{u} is a feasible solution, hence it attains the primal supremum and is optimal.

(ii) *A convex closed cone.* We define the set

$$K = \{(d, \delta) \in V \times \mathbb{R} : \exists u \in P(d) \text{ with } (a, u)_U \geq \delta\}.$$

It is easy to verify that K is a convex cone. Moreover, the boundedness condition implies that K is closed: To this end, take a sequence of elements $(d_n, \delta_n)_n$ of elements of K converging to (d, δ) as $n \rightarrow \infty$. Then we find $u_n \in U$ such that

$$Au_n \leq_V d_n, \quad u_n \geq_U 0, \quad (a, u_n)_U \geq \delta_n \quad \forall n \in \mathbb{N}.$$

By the boundedness condition, there exists an $\eta > 0$ and $\hat{u}_n \in U$ such that

$$A\hat{u}_n \leq_V d_n, \quad \hat{u}_n \geq_U 0, \quad (a, \hat{u}_n)_U \geq (a, u_n)_U \geq \delta_n \quad \forall n \in \mathbb{N}$$

and

$$\|\hat{u}_n\|_U \leq \eta \|d_n\|_V \quad \forall n \in \mathbb{N}$$

hold. By the last inequality, the sequence $(\hat{u}_n)_n$ is bounded in U , hence we can select a weakly convergence subsequence, w.l.o.g. $\hat{u}_n \rightharpoonup u$, $n \rightarrow \infty$. Since the cones K_U and K_V are convex and closed, they are also weakly closed. This permits to show in turn

$$Au \leq_V d, \quad u \geq_U 0, \quad (a, u)_U \geq \delta,$$

hence $(d, \delta) \in K$ is proved.

(iii) *Convex cones C_n to be separated from K .*

For all $n \in \mathbb{N}$, we have

$$(c, s + 1/n) \notin K.$$

Since K is closed, there are open balls $B_{r_n} \subset V \times \mathbb{R}$ of radius $r_n > 0$ centered at $(c, s + 1/n)$ such that

$$K \cap B_{r_n} = \emptyset \quad \forall n \in \mathbb{N}.$$

Define for all $n \in \mathbb{N}$

$$C_n = \bigcup_{\lambda > 0} \lambda B_{r_n}.$$

These sets have the following properties:

- Obviously, all C_n are cones.
- All intersections $C_n \cap K$ are empty: If $C_n \cap K$ would contain an element u , then there are $\lambda > 0$ and $\tilde{u} \in B_{r_n}$ such that $u = \lambda \tilde{u}$. Since K is a cone, also $u/\lambda = \tilde{u}$ belongs to K and also to B_{r_n} contradicting $K \cap B_{r_n} = \emptyset$.

– C_n is convex as it can be readily verified. Moreover, B_{r_n} is contained in C_n , hence the interior of C_n is not empty.

Therefore, all C_n are convex cones with nonempty interior.

(iv) *Separation of K and C_n .* By the Eidelheit separation theorem, see [5], there is a closed hyperplane separating the convex closed cone K and the convex set C_n with nonempty interior. Therefore, $z_n \in V$, $\xi_n \in \mathbb{R}$, and $\sigma_n \in \mathbb{R}$ exist such that

$$(d, z_n)_V + \delta \xi_n \leq \sigma_n < (w, z_n)_V + \alpha \xi_n \quad \forall (d, \delta) \in K, \forall (w, \alpha) \in C_n. \quad (5.2)$$

From this inequality, we draw several conclusions to finally arrive at strong duality.

– We have that $(d, \delta) \in K$ for all $d \geq_V 0$ and $\delta \leq 0$. Therefore, (5.2) yields $\sigma_n \geq 0$, $z_n \leq_V^* 0$, and $\xi_n \geq 0$.

– Since C_n is a cone, with (w, α) also $\lambda(w, \alpha)$ belongs to C_n for all positive λ . By $\lambda \searrow 0$, we obtain $\sigma_n = 0$ from (5.2).

– For all $u \geq_U 0$, the pair $(Au, (a, u)_U)$ belongs to K . Inserting this in (5.2), we deduce

$$(u, A^* z_n + \xi_n a)_U \leq 0 \quad \forall u \geq_U 0,$$

hence

$$A^* z_n + \xi_n a \leq_U^* 0. \quad (5.3)$$

– The numbers ξ_n cannot vanish. If $\xi_n = 0$ would hold, inserting $(c, (c, u)_U)$ in the left hand side of (5.2) and $(c, s + 1/n)$ in the right hand side, we arrive at the contradiction $(c, z_n)_V < (c, z_n)_V$. Therefore, $\xi_n > 0$ must hold.

– Define $v_n = -z_n/\xi_n$, then $A^* v_n \geq_U^* a$ follows from (5.3). Moreover, $z_n \leq_V^* 0$ and $\xi_n \geq 0$ yield $v_n \geq_V^* 0$, hence v_n is feasible for the dual problem (\mathcal{DP}) .

– Finally, dividing the right-hand side of (5.2) by ξ_n , we get

$$0 < (w, z_n/\xi_n)_V + \alpha.$$

Inserting the elements $w = c$ and $\alpha = s + 1/n$, we arrive in view of $v_n = -z_n/\xi_n$ at

$$(c, v_n)_V < s + 1/n \quad \forall n \in \mathbb{N}.$$

Passing to the limit $n \rightarrow \infty$, we see that the dual objective value $(c, v_n)_V$ can be taken arbitrarily close to the primal maximum s . In view of the weak duality result of Lemma 5.3, this shows (5.1). \square

5.2. Application to the optimal control problem (P). Finally, we prove Theorem 4.6. We consider the primal and dual problem defined in Section 4. To show Theorem 4.6 on strong duality, we can take two ways: The first is to view the dual problem (DP) as primal one and to verify the boundedness assumption 5.4 for (DP). This can be done invoking Lemma 4.2. The other way around is to verify the boundedness assumption for the primal problem. We prefer this variant and consider two different primal problems.

The first is related to the case, where $u_b < \infty$, i.e. where the pointwise control constraints have upper and lower bounds. Then the primal problem (PP) can be written as

$$\begin{aligned} \max & (a, v)_{L^2(Q)} \\ & v \leq v_b \\ & v \leq b + Sv \\ & v \geq 0, \end{aligned}$$

where v is taken from $U = L^2(Q)$. Now we set $V = U \times U$, $K_U = L^2(Q)^+$, $K_V = K_U \times K_U$, and

$$A = \begin{pmatrix} I_d \\ I_d - S \end{pmatrix}, \quad c = \begin{pmatrix} v_b \\ b \end{pmatrix}.$$

Then the primal problem (PP) is equivalent to

$$\max (a, u)_U, \quad Au \leq_V c, \quad u \geq_U 0.$$

Let us verify Assumption 5.4 of boundedness. We select a varying $d = (d_1, d_2) \in V$ and consider the set

$$P(d) = \{u \in L^2(Q) : u \geq 0, u \leq d_1, u \leq d_2 + Su\}.$$

For all $u \in P(d)$, we have

$$\|u\|_{L^2(Q)} \leq \|d_1\|_{L^2(Q)}.$$

Moreover, by inverse non-negativity, we find

$$\|u\|_{L^2(Q)} \leq \|z\|_{L^2(Q)}$$

where z is the unique solution to $z = d_2 + Sz$. Since $(I - S)$ is continuously invertible and non-negative, an $\alpha > 0$ exists such that $\|z\|_{L^2(Q)} \leq \alpha \|d_2\|_{L^2(Q)}$. Therefore,

$$\|u\|_{L^2(Q)} \leq \max\{1, \alpha\} (\|d_1\|_{L^2(Q)} + \|d_2\|_{L^2(Q)})$$

holds for all $u \in P(d)$. This property is even stronger than Assumption 5.4. Therefore, by Theorem 5.5 the strong duality relation is satisfied. This proves Theorem 4.6 for the first form of (PP).

In the second form of (PP), the upper control constraint is missing, i.e. $u_b = \infty$, hence $v_b = \infty$. Here, the primal problem takes the form

$$\begin{aligned} \max (a, v)_{L^2(Q)} \\ v \leq b + Sv \\ v \geq 0. \end{aligned}$$

We define $V = L^2(Q)$, $K_V = K_U$, and $A = I - S$. For $d \in L^2(Q)$, the set $P(d)$ is

$$P(d) = \{u \in L^2(Q) : u \geq 0, u \leq d + Su\}.$$

By inverse non-negativity, Corollary 2.3, we find

$$\|u\|_{L^2(Q)} \leq \alpha \|d\|_{L^2(Q)} \quad \forall u \in P(d).$$

Again, Assumption 5.4 is fulfilled and Theorem 4.6 is proved by Theorem 5.5.

REFERENCES

- [1] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer-Verlag, New York, 2000.
- [2] 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.
- [3] R. C. Grinold, *Continuous programming. I. Linear objectives*, *J. Math. Anal. Appl.*, **28** (1969), 32–51.
- [4] R. C. Grinold, *Symmetric duality for continuous linear programs*, *SIAM J. Appl. Math.*, **18** (1970), 84–97.
- [5] J. Jahn, *Vector Optimization. Theory, Applications, and Extensions*, Springer-Verlag, Berlin, 2004.
- [6] W. Krabs, *Zur Dualitätstheorie bei linearen Optimierungsproblemen in halbgeordneten Vektorräumen*, *Math. Z.*, **121** (1971), 320–328.

- [7] O. A. Ladyzhenskaya, V. A. Solonnikov and N. N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, Providence, R.I., 1968.
- [8] J.-L. Lions, *Contrôle Optimal de Systèmes Gouvernés par des Équations aux Dérivées Partielles*, Avant Propos de P. Lelong Dunod, Paris, Gauthier-Villars, Paris, 1968.
- [9] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, Inc., New York-London-Sydney, 1969.
- [10] A. Rösch and F. Tröltzsch, [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** (2007), 1098–1115.
- [11] F. Tröltzsch, [Existenz- und Dualitätsaussagen für lineare Optimierungsaufgaben in reflexiven Banach-Räumen](#), *Math. Operationsforschung und Statistik*, **6** (1975), 901–912.
- [12] F. Tröltzsch, [A minimum principle and a generalized bang-bang-principle for a distributed optimal control problem with constraints on the control and the state](#), *Z. Angew. Math. Mech.*, **59** (1979), 737–739.
- [13] W. F. Tyndall, [A duality theorem for a class of continuous linear programming problems](#), *J. Soc. Indust. Appl. Math.*, **13** (1965), 644–666.

Received February 2018; revised December 2018.

E-mail address: eduardo.casas@unican.es

E-mail address: troeltzsch@math.tu-berlin.de