OPTIMALITY CONDITIONS FOR STATE-CONSTRAINED PDE CONTROL PROBLEMS WITH TIME-DEPENDENT CONTROLS

J.C. DE LOS REYES[‡]   P. MERINO[‡]   J. REHBERG[*]   F. TRÖLTZSCH[†]

ABSTRACT. The paper deals with optimal control problems for semilinear elliptic and parabolic PDEs subject to pointwise state constraints. The main issue is that the controls are taken from a restricted control space. In the parabolic case, they are $\mathbb{R}^m$ -vector-valued functions of the time, while the are vectors of $\mathbb{R}^m$ in elliptic problems. Under natural assumptions, first- and second-order sufficient optimality conditions are derived. The main result is the extension of second-order sufficient conditions to semilinear parabolic equations in domains of arbitrary dimension. In the elliptic case, the problems can be handled by known results of semi-infinite optimization. Here, different examples are discussed that exhibit different forms of active sets and where second-order sufficient conditions are satisfied at the optimal solution.

1. INTRODUCTION

This paper is a further contribution to second-order optimality conditions in the optimal control of partial differential equations. A large number of papers has been devoted to this issue so far and conditions of this type are used as an essential assumption in publications on numerical methods.

Sufficient conditions were investigated first for problems with *control constraints*. Their main focus was on second-order optimality conditions that are close to the associated necessary ones, for instance, by Bonnans [5], Casas, Unger and Tröltzsch [12], Goldberg and Tröltzsch [18]; see also the examples for the control of PDEs in the recent monograph by Bonnans and Shapiro [6]. The situation changes, if pointwise state constraints are given. Here, the theory is essentially more difficult as the Lagrange multipliers associated with the state constraints are Borel measures. Therefore, the associated theory is less complete than that for control constraints. Although considerable progress has been made in this issue, cf. Casas, Tröltzsch and Unger [13], Raymond and Tröltzsch [29], Casas and Mateos [10], Casas, De los Reyes and Tröltzsch [9], there remain open questions, if pointwise state constraints are formulated in the whole domain: In the elliptic case of boundary or distributed control with pointwise state-constraints, the conditions are sufficiently general only for spatial dimension 2 or 3, respectively, [13, 9]. For parabolic problems, only distributed controls in one-dimensional domains can be handled in full generality, [29].

The difficulties mentioned above are intrinsic for problems with pointwise state constraints and it seems that they cannot be entirely avoided. However, a review on the application of optimal control problems for partial differential equations shows that the situation is often easier in practice: In many cases, the control function depends only on finitely many parameters that may also depend on time in parabolic problems.

For instance, in all applications the group of the fourth author has been engaged so far, the controls are finite-dimensional in this sense. This finite-dimensionality seems to be characteristic for real applications of control theory. This concerns

the cooling of steel profiles by water discussed in [16], [31], [32], some problems of flow control, cf. Henning and King [22], the sublimation process for the production of SiC single crystals, cf. [26], and local hyperthermia in tumor therapy, [15]. In all of these problems, the controls are represented by finitely many real values, namely, the intensities of finitely many spray nozzles in cooling steel, amplitude and frequency of controlled suction or blowing in flow control, frequency and power of the electrical current in sublimation crystal growth, and the energy of finitely many microwave antennas in local hyperthermia.

In some cases, these finitely many real values depend on time. Moreover, in all the applications mentioned above, pointwise state constraints are very important. Problems of this type are the main issue of our paper. We address the following points:

First, we consider semilinear parabolic equations with distributed controls of the type $f(x,t) = \sum_{i=1}^{k} e_i(x)u_i(t)$, where the functions e_i are bounded and the control functions u_i are taken from a set of admissible controls. Thanks to the boundedness of the fixed functions e_i, we are able to extend a result of [9] for one-dimensional parabolic problems to domains Ω of arbitrary dimension. Although our regularity results can be extended to controls $f \in L^2(0,T;L^\infty(\Omega))$, we need the restriction to the finite-dimensional ansatz above to deal with second-order sufficient optimality conditions, cf. Remark 2.

Moreover, we consider different problems with finite-dimensional controls $u \in \mathbb{R}^m$. If the control is a vector of $\mathbb{R}^m$, pointwise state constraints generate a *semi-infinite optimization* problem. The associated constraint set is infinite by its nature. The Lagrange multipliers for the state-constraints remain to be Borel measures. Therefore, this class of problems is sufficiently interesting for the numerical analysis. This setting belongs to the class of semi-infinite optimization problems. Here, we are able to invoke the known theory of first- and second-order optimality conditions. In the case of partial differential equations, this theory needs special adaptations that are briefly sketched. Moreover, we present different examples of state-constrained control problems, where second-order sufficient conditions are satisfied at the optimal solution. These problems can be used to test numerical methods and show how diverse the active set may look like.

2. Semilinear parabolic problems

2.1. **Problem statement.** We consider the following distributed optimal control problem with time-dependent control functions,

$$
\textbf{(P1)} \begin{cases}
\min_{u \in \mathcal{U}_{ad}} J(u) = \int_Q L(x,t,y(x,t),u(t))\,dxdt + \int_\Sigma \ell(x,t,y(x,t))\,dS(x)dt \\
\qquad\qquad + \int_\Omega r(x,y(x,T))\,dx \\[4pt]
\text{subject to} \\
\begin{aligned}
y_t + A\,y(x,t) + d(x,t,y(x,t)) &= \sum_{i=1}^m e_i(x)\,u_i(t) &&\text{in}\quad Q, \\
\partial_\nu y(x,t) &= 0 &&\text{on}\quad \Sigma, \\
y(x,0) - y_0(x) &= 0 &&\text{in}\quad \Omega,
\end{aligned} \\[4pt]
\quad g(x,t,y(x,t)) \leqslant 0 \quad \text{for all } (x,t) \in K \subset \bar{Q},
\end{cases}
\tag{2.1}
$$

where $u = (u_1, \ldots, u_m)^\top$ and $\mathcal{U}_{ad}$ is defined by

$$
\mathcal{U}_{ad} = \{u \in L^\infty(0,T;\mathbb{R}^m) \,:\, u_a(t) \leq u(t) \leq u_b(t) \ \text{a.e. } t \in [0,T]\}.
$$

In this setting, Ω is a subset of $\mathbb{R}^n$ ($n \geqslant 1$) with boundary Γ, we have set $Q = \Omega \times (0, T)$, $\Sigma = \Gamma \times (0, T)$, A is a second-order elliptic differential operator, K is a non-empty compact subset of $\bar{Q}$, and $T > 0$ is a fixed real number. Moreover, functions $L : Q \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, $d : Q \times \mathbb{R} \to \mathbb{R}$, $\ell : \Sigma \times \mathbb{R} \to \mathbb{R}$, $r : \Omega \times \mathbb{R} \to \mathbb{R}$, $g : K \times \mathbb{R} \to \mathbb{R}$, an initial state $y_0 \in C(\bar{\Omega})$, fixed functions $e_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, m$, and fixed bounds $u_a, u_b \in L^\infty(0, T; \mathbb{R}^m)$ are given such that $u_a(t) \leq u_b(t)$ holds a.e. on $(0, T)$ in the componentwise sense. The symbol ∂_ν denotes the derivative in the direction of the outward unit normal ν at Γ.

2.2. **Main assumptions and well-posedness of the state equation.** In the parabolic case, we rely on the following assumptions:

(A.1) The set $\Omega \subset \mathbb{R}^n$ is an open and bounded Lipschitz domain in the sense of Nečas [28]. The differential operator A is defined by

$$Ay(x) = -\sum_{i,j=1}^{n} \partial_{x_j}(a_{ij}(x)\partial_{x_i} y(x))$$

with coefficients $a_{ij} \in L^\infty(\Omega)$ satisfying

$$\lambda_A |\xi|^2 \leq \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \quad \forall \xi \in \mathbb{R}^n, \ \forall x \in \Omega$$

for some $\lambda_A > 0$.

(A.2) (Carathèodory type assumption) For each fixed pair $(x, t) \in Q = \Omega \times (0, T)$ or $\Sigma = \Gamma \times (0, T)$, respectively, the functions $d = d(x, t, y)$ and $\ell = \ell(x, t, y)$ are twice partially differentiable with respect to y. For all fixed $y \in \mathbb{R}$, they are Lebesgue measurable with respect to $(x, t) \in Q$, or $x \in \Sigma$ respectively.

Analogously, for each fixed pair $(x, t) \in Q$, $L = L(x, t, y, u)$ is twice partially differentiable with respect to $(y, u) \in \mathbb{R}^{m+1}$. For all fixed $(y, u) \in \mathbb{R}^{m+1}$, L is Lebesgue measurable with respect to $(x, t) \in Q$.

The function $g = g(x, t, y)$ is supposed to be twice continuously differentiable with respect to y on $K \times \mathbb{R}$, i.e. g, $\dfrac{\partial g}{\partial y}$, and $\dfrac{\partial^2 g}{\partial y^2}$ are continuous on $K \times \mathbb{R}$.

(A.3) (Monotonicity) For almost all $(x, t) \in Q$ or $(x, t) \in \Sigma$, respectively, and all $y \in \mathbb{R}$ it holds that

$$\frac{\partial d}{\partial y}(x, t, y) \geq 0.$$

(A.4) (Boundedness and Lipschitz properties) The functions e_i, $i = 1, \ldots, m$, belong to $L^\infty(\Omega)$. There is a constant C_0 and, for all $M > 0$, a constant $C_L(M)$ such that the estimates

$$
\begin{aligned}
|d(x, t, 0)| + |\frac{\partial d}{\partial y}(x, t, 0)| + |\frac{\partial^2 d}{\partial y^2}(x, t, 0)| &\leqslant C_0 \\
|\frac{\partial^2 d}{\partial y^2}(x, t, y_1) - \frac{\partial^2 d}{\partial y^2}(x, t, y_2)| &\leqslant C_L(M)|y_1 - y_2|
\end{aligned}
$$

hold for almost all $(x, t) \in Q$ and all $|y_i| \leqslant M$, $i = 1, 2$. The functions ℓ, and g are assumed to satisfy these boundedness and Lipschitz properties on Σ and Q,

respectively. The function r is assumed to obey these assumptions with $x \in \Omega$ substituted for $(x,t) \in Q$. Analogously,

$$|L(x,t,0,0)| + |L'(x,t,0,0)| + |L''(x,t,0,0)| \leqslant C_0$$

$$|L''(x,t,y_1,u_1) - L''(x,t,y_2,u_2)| \leqslant C_L(M)\,(|y_1 - y_2| + |u_1 - u_2|)$$

hold for almost all $(x,t) \in Q$ and all $|y_i| \leqslant M$, $|u_i| \leqslant M$, $i = 1,2$. Here, L' and L'' denote the gradient and the Hessian matrix of L with respect to $(y,u) \in \mathbb{R}^{m+1}$.

2.3. **Well-posedness of the state equation.**

2.3.1. *State equation and existence of optimal controls.* In this subsection, we show that the control-to-state mapping $G : u \mapsto y$ is well defined for all admissible controls u. Moreover, we show certain continuity properties of G. Later, we will also discuss the differentiability of G. In all what follows, we denote the state function y associated with u by y_u, i.e. $y_u = G(u)$.

Theorem 1. *Suppose that the assumptions* **(A.1)** – **(A.4)** *are satisfied. Then, for every* $u \in L^q(0,T;\mathbb{R}^m)$ *with* $q > n/2 + 1$, *the state equation* (2.1) *has a unique solution* $y_u \in C(\bar{Q}) \cap W(0,T)$. *If* $u_k \rightharpoonup u$ *weakly in* $L^{\tilde{q}}(0,T;\mathbb{R}^m)$ *with* $\tilde{q} = \max\{q,2\}$, *then* $y_{u_k} \to y_u$ *strongly in* $C(\bar{Q})$.

Proof. The functions e_i are bounded, hence the right-hand side

$$v(x,t) = \sum_{i=1}^m e_i(x)\,u_i(t)$$

of equation (2.1) belongs to $L^q(Q)$ with $q > \frac{n}{2} + 1$. Therefore, existence, uniqueness and continuity of the solution $y \in W(0,T)$ of (2.1) follow from Casas [8]. It remains to show that $u_k \rightharpoonup u$ weakly in $L^{\tilde{q}}(0,T,\mathbb{R}^m)$ implies $y_{u_k} \to y_u$ strongly in $C(\bar{Q})$. To this aim, we first re-write the parabolic equation as

$$
\begin{array}{rcllr}
\dfrac{dy_k}{dt} + A\,y_k & = & v_k := \sum_{i=1}^m e_i(x)\,u_{k,i}(t) - d(x,t,y_k) & \text{in} & Q, \\[2mm]
\partial_\nu y_k & = & 0 & \text{on} & \Sigma, \\[1mm]
y_k(x,0) & = & y_0(x) & \text{in} & \Omega,
\end{array}
\tag{2.2}
$$

where we introduced $y_k := y_{u_k}$. The sequence of right-hand sides (v_k) is bounded in $L^{\tilde{q}}(Q)$. Therefore, the boundedness of the sequence y_k in $C(\bar{Q})$ is a standard conclusion, [8], so that $(d(x,t,y_k))$ is bounded in $L^\infty(Q)$. In view of this, the sequence (v_k) is bounded in $L^2(0,T;L^p(\Omega))$, where p is taken sufficiently large to meet the assumptions of Theorem 3. Therefore, (v_k) contains a sub-sequence (v_{k_l}) with $v_{k_l} \rightharpoonup v$ weakly in $L^2(0,T;L^p(\Omega))$, $l \to \infty$,

W.l.o.g we can assume $y_0 = 0$, since the solution $\hat{y}$ of $\hat{y}_t + A\hat{y} = 0$, $\partial_\nu \hat{y} = 0$ and $\hat{y}(x,0) = y_0(x)$ can be subtracted from the sequence y_k. This fixed function does not influence the convergence properties.

Thanks to Theorem 3, the mapping $v_k \mapsto y_k$ defined by (2.2) is linear and continuous from $L^2(0,T;L^p(\Omega))$ to the Hölder space $C^\alpha(Q)$ for some $\alpha > 0$. By the compactness of the injection $C^\alpha(Q) \hookrightarrow C(\bar{Q})$, (y_{k_l}) converges uniformly to some $y \in C(\bar{Q})$. This, implies $(d(x,t,y_{k_l})) \to (d(x,t,y))$ in $L^\infty(Q) \hookrightarrow L^2(0,T;L^p(\Omega))$ so that the right-hand side converges weakly to $v = \sum_{i=1}^m e_i\,u_i - d(\cdot,y)$. Therefore, y is associated with this function v and solves the semilinear equation with control u. By uniqueness, it must hold $y = y_u$, hence the same result is obtained for any subsequence of (v_k), and a standard argument yields $y_k \to y$ in $C(\bar{Q})$ as $k \to \infty$. $\square$

It is obvious that only those parts of the assumption **(A.4)** are needed in the theorem that are related to d and d_y.

Theorem 2. *Assume that the assumptions* **(A.1)** – **(A.4)** *are fulfilled, the function* $L = L(x, t, y, u)$ *is convex with respect to* $u \in \mathbb{R}^m$ *and the set of feasible controls is nonempty. Then the control problem* (P1) *has at least one solution.*

This theorem is a standard consequence of Theorem 1 and the lower semicontinuity of J that needs the convexity of L with respect to u.

2.3.2. *Hölder regularity for linear parabolic equations.* The results of this subsection are needed for the extension of second-order sufficient conditions to the case of n-dimensional domains, if the controls depend only on time. To derive second-order sufficient conditions, linearized equations are considered for L^2-controls, and the regularity of associated states is derived in Theorem 3, which has already been used in the proof of Theorem 1. The theorem is proved by recent results on maximal parabolic regularity. Throughout this section, Ω is allowed to be a Lipschitz domain in the sense of Grisvard [21]. This is more general than domains with Lipschitz boundary in the sense of Nečas [28].

We consider the linear parabolic problem

$$
\begin{aligned}
\frac{dz}{dt} + A\,z + c_0 z &= v &&\text{in} && Q, \\
\partial_\nu z &= 0 &&\text{in} && \Sigma, \\
z(x, 0) &= 0 &&\text{in} && \Omega,
\end{aligned}
\tag{2.3}
$$

where $c_0 \in L^\infty(Q)$ is given fixed. Later, c_0 stands for the partial derivative $\partial d / \partial y$ taken at the optimal state. To deal with equation (2.3), we provide some basic facts on maximal parabolic regularity.

We denote by $\mathcal{A} \in L^\infty(\Omega; \mathbb{R}^{n \times n})$ the matrix function $\mathcal{A}(x) = (a_{ij}(x))$, with a_{ij}, $i, j \in \{1, \ldots, n\}$ defined in **(A.1)**. Associated with our differential operator A, the linear and continuous operator $-\nabla \cdot \mathcal{A}\nabla : H^1(\Omega) \to H^1(\Omega)'$ is defined by

$$
< -\nabla \cdot \mathcal{A}\nabla v, w > := \int_\Omega \mathcal{A}\,\nabla v \cdot \nabla w\, dx \qquad \forall\, v, w \in H^1(\Omega).
\tag{2.4}
$$

The restriction of this operator to the spaces $L^p(\Omega)$ with $p \geq 2$ is denoted by A_p. Its domain is given by

$$
D(A_p) = \{ y \in H^1(\Omega) \,:\, A\,y \in L^p(\Omega) \}
$$

equipped with the associated graph norm. It is known that A_p incorporates a (homogeneous) Neumann boundary condition, see [17] Ch. II.2 or [14] Ch. 1.2.

Let us next recall the concept of maximal regularity: Let X be a Banach space and A denote a closed operator with dense domain $\mathcal{D} \subset X$ equipped with the graph norm. Moreover, let $S = (T_0, T) \subset \mathbb{R}$ be a bounded interval. Suppose $r \in (1, \infty)$. Then A is said to satisfy maximal parabolic $L^r(S; X)$-regularity iff for any $f \in L^r(S; X)$ there is a unique function $w \in W^{1,r}(S; X) \cap L^r(S; \mathcal{D})$ that satisfies

$$
\frac{\partial w}{\partial t} + Aw = f, \quad w(T_0) = 0.
\tag{2.5}
$$

By $W^{1,r}(S; X)$, we denote the set of those elements from $L^r(S; X)$ whose distributional derivative also belongs to $L^r(S; X)$. If X, Y are Banach spaces which form an interpolation couple, then we denote by $[X, Y]_\theta$ the corresponding complex interpolation space and by $(X, Y)_{\theta, r}$ the real interpolation space, see [30]. Below we well employ the following fact: There is a continuous injection

$$
\mathcal{E} : W^{1,r}(S; X) \cap L^r(S; \mathcal{D}) \hookrightarrow C(\bar{S}; (X, \mathcal{D})_{1-\frac{1}{r}, r}),
\tag{2.6}
$$

see [4] Ch. III Thm. 4.10.2, cf. also [30] Ch. 1.8. Moreover, for any $\eta \in (0, 1 - \frac{1}{r})$ there is a continuous embedding $(X, \mathcal{D})_{1-\frac{1}{r}, r} \hookrightarrow [X, \mathcal{D}]_\eta$ and, consequently, a

continuous embedding

$$C(\bar{S}; (X, \mathcal{D})_{1-\frac{1}{r}, r}) \hookrightarrow C(\bar{S}; [X, \mathcal{D}]_\eta). \tag{2.7}$$

Next, we need the following

Lemma 1. *Assume* $\tau \in (0, 1 - \frac{1}{r})$. *Then there is an index* $\kappa > 0$ *such that* $W^{1,r}(S; X) \cap L^r(S; \mathcal{D})$ *continuously embeds into* $C^\kappa(S; [X, \mathcal{D}]_\tau)$.

Proof. First of all, the estimate

$$\|w(t) - w(t_0)\|_X = \| \int_{t_0}^t w'(s) \, ds \|_X \leq \left(\int_{t_0}^t \|w'(s)\|_X^r \, ds \right)^{\frac{1}{r}} \left(\int_{t_0}^t ds \right)^{\frac{r-1}{r}} \leq$$

$$\leq \|w\|_{W^{1,r}(S;X)} \, |t - t_0|^{\frac{r-1}{r}}$$

gives us a (continuous) embedding from $W^{1,r}(S; X)$ into $C^\delta(S; X)$, where $\delta = \frac{r-1}{r}$. Let η be a number from $(\tau, 1 - \frac{1}{r})$. Then, putting $\lambda = \frac{\tau}{\eta}$ we obtain by the reiteration theorem for complex interpolation, see [30] Ch. 1.9.3,

$$\frac{\|w(t) - w(s)\|_{[X,D]_\tau}}{|t - s|^{\delta(1-\lambda)}} \leq c \, \frac{\|w(t) - w(s)\|_{[X,[X,D]_\eta]_\lambda}}{|t - s|^{\delta(1-\lambda)}} \leq$$

$$\leq c \, \frac{\|w(t) - w(s)\|_X^{1-\lambda}}{|t - s|^{\delta(1-\lambda)}} \|w(t) - w(s)\|_{[X,D]_\eta}^\lambda \leq$$

$$\leq c_1 \left(\frac{\|w(t) - w(s)\|_X}{|t - s|^\delta} \right)^{1-\lambda} \left(2\|w\|_{C(\bar{S};[X,D]_\eta)} \right)^\lambda \leq c_2 \|w\|_{W^{1,r}(S;X) \cap L^r(S;\mathcal{D})}.$$

$\square$

Corollary 1. *Assume that* A *satisfies maximal parabolic* $L^r(S; X)$-*regularity and let* $\tau \in (0, 1 - \frac{1}{r})$ *be given. Then the mapping that assigns to every* $f \in L^r(S; X)$ *the solution of* (2.5) *is continuous from* $L^r(S; X)$ *into* $C^\kappa(S; [X, \mathcal{D}]_\tau)$ *for a certain* $\kappa > 0$.

Proof. The mapping

$$\frac{\partial}{\partial t} + A : W^{1,r}(S; X) \cap L^r(S; \mathcal{D}) \cap \{w \,|\, w(T_0) = 0\} \to L^r(S; X)$$

is continuous and bijective. Hence, the inverse is also continuous by the open mapping theorem. Combining this with the preceding lemma gives the assertion.

$\square$

Now we are able to show our main result on parabolic regularity: Hölder regularity can be obtained for functions which exhibit only to L^2-regularity in time, if their spatial regularity is L^p and p is sufficiently large.

Theorem 3. *If* f *belongs to* $L^r(S; L^p(\Omega))$ *with* $r > 1$ *and sufficiently large* p, *then the solution* w *of*

$$\frac{\partial w}{\partial t} + A_p w = f, \quad w(0) = 0 \tag{2.8}$$

is from a space $C^\kappa(S; C^\beta(\Omega))$. *Moreover, the mapping* $f \mapsto w$ *is continuous from* $L^r(S; L^p(\Omega))$ *to* $C^\kappa(S; C^\beta(\Omega))$.

Proof. We apply the general results on maximal parabolic regularity to our operator A_p. It is known that A_p enjoys maximal parabolic $L^r(S; L^p(\Omega))$-regularity for every $p \in (1, \infty)$ and $r > 1$. We refer to [19], Thm. 7.4. Moreover, the following interpolation result is known: If $\theta \in (0, 1)$ and

$$\beta := \theta\alpha - (1 - \theta)\frac{n}{p} > 0, \tag{2.9}$$

then

$$[L^p(\Omega), C^\alpha(\Omega)]_\theta \hookrightarrow C^\beta(\Omega). \tag{2.10}$$

The result is shown in [19], see the proof of Thm. 7.1. In general, the domain of A_p is difficult to determine. However, the following embedding result is helpful: For $p > \frac{n}{2}$, there exists $\alpha > 0$ such that the continuous embedding

$$\mathrm{dom}(A_p) \hookrightarrow C^\alpha(\Omega) \tag{2.11}$$

holds true. This result is proved in [20]. Keeping $\alpha > 0$ fixed we can increase p so that also (2.9) is satisfied. Clearly, (2.10) and (2.11) remain true. Taking now into account Corollary 1 for $X = L^p(\Omega)$, then the assertion follows from (2.10) and (2.11). $\qquad\square$

Corollary 2. *If $r > 1$ and p is sufficiently large, then for all $v \in L^r(S; L^p(\Omega))$, the weak solution of (2.3) belongs to $C^\alpha(\bar{Q})$ with some $\alpha > 0$. The mapping $v \mapsto z$ is continuous from $L^r(0, T; L^p(\Omega))$ to $C^\alpha(\bar{Q})$.*

Proof. The operator $\frac{\partial}{\partial t} + A_p$ is a topological isomorphism between $L^r(S; D(A_p)) \cap W^{1,r}(S; L^p(\Omega)) \cap \{z : z(T_0) = 0\}$ and $L^r(S; L^p(\Omega))$, and is therefore a Fredholm operator of index zero. Obviously, the multiplication operator induced by c_0 is bounded on $L^2(S; L^p(\Omega))$; hence, the domain of $\frac{\partial}{\partial t} + A_p + c_0$ equals the domain of $\frac{\partial}{\partial t} + A_p$ – which, due to Theorem 3, compactly embeds into $L^r(S; L^p(\Omega))$. Thus, the multiplication operator induced by c_0 is relatively compact with respect to $\frac{\partial}{\partial t} + A_p$. By a well known perturbation theorem, $\frac{\partial}{\partial t} + A_p + c_0$ then also must be a Fredholm operator and is also of index zero (see Kato [23], Ch. IV.5.3). Let us show that it is injective: Let z be a solution of $\frac{\partial z}{\partial t} + A_p z + c_0 z = 0$ or, equivalently, $\frac{\partial z}{\partial t} + A_p z = -c_0 z$. As a solution of this parabolic equation with right hand side $-c_0 z$, z has then the representation $z(t) = -\int_0^t e^{(s-t)A_p} c_0(s, \cdot) z(s) \ ds$. But the semigroup e^{-tA_p} is contractive on $L^p(\Omega)$ (see [19] Thm. 4.11) and c_0 is essentially bounded; thus Gronwall's lemma yields $z \equiv 0$. Therefore, $\frac{\partial}{\partial t} + A_p + c_0$ is injective. Because it is Fredholm and, additionally, of index zero, it is also surjective. Consequently, the inverse operator maps $L^r(S; L^p(\Omega))$ continuously into the domain of $\frac{\partial}{\partial t} + A_p$, and this is continuously embedded into a space $C^\alpha(\bar{Q})$. $\qquad\square$

Notice that in particular $L^2(0, T; L^\infty(\Omega))$ is mapped continuously into a space $C^\alpha(\bar{Q})$. This is what we need for the discussion of second-order sufficient conditions. The theorem refers to the PDE (2.3) with homogeneous initial condition. It is obvious that the result extends to inhomogeneous Hölder continuous initial data.

Remark 1. *The Hölder regularity might also be deduced from Ladyzhenskaya et al. [25] under the assumption that Γ is a Lipschitz boundary in the sense of Nečas [28], cf. also [8]. Under this assumption, the continuity of the state function for our restricted class of controls was also shown in [29]. The use of maximal parabolic regularity permits to extend these results to Lipschitz domains in the sense of Grisvard [21]. Moreover, this approach is not restricted to Neumann conditions. It also allows for mixed boundary conditions and, of course, for Dirichlet conditions. Last, but not least, our approach essentially shortens and unifies the associated proofs.*

2.4. **Necessary optimality conditions.** The control-to-state mapping $G(u) = y_u$, $G : L^\infty(0, T; \mathbb{R}^m) \to C(\bar{Q}) \cap W(0, T)$, and the reduced objective functional J are of class C^2 from $L^\infty(0, T; \mathbb{R}^m)$ to their image spaces, provided that the assumptions **(A.1)-(A.4)** are satisfied. This follows by the arguments of [8].

We define, for $v \in L^\infty(0, T; \mathbb{R}^m)$, the function z_v as the unique solution to

$$\begin{cases} \dfrac{dz_v}{dt} + Az_v + \dfrac{\partial d}{\partial y}(x, t, y_u)z_v & = & \sum_{i=1}^m e_i(x)v_i(t) & \text{in } Q, \\ \partial_\nu z_v & = & 0 & \text{in } \Sigma, \\ y(x, 0) & = & 0 & \text{in } \Omega. \end{cases} \tag{2.12}$$

Then $G'(u)$, $G : L^\infty(0, T; \mathbb{R}^m) \to C(\bar{Q}) \cap W(0, T)$ is given by $G'(u)v = z_v$. Moreover, for v_1, $v_2 \in L^\infty(0, T; \mathbb{R}^m)$, we introduce $z_{v_i} = G'(u)v_i$, $i = 1, 2$, and obtain $G''(u)v_1v_2 = z_{v_1v_2}$, where $z_{v_1v_2}$ is the solution to

$$\begin{cases} \dfrac{dz_{v_1v_2}}{dt} + Az_{v_1v_2} + \dfrac{\partial d}{\partial y}(x, t, y_u)z_{v_1v_2} + \dfrac{\partial^2 d}{\partial y^2}(x, t, y_u)z_{v_1}z_{v_2} & = & 0 & \text{in } Q, \\ \partial_\nu z_{v_1v_2} & = & 0 & \text{in } \Sigma, \\ z_{v_1v_2}(x, 0) & = & 0 & \text{in } \Omega. \end{cases} \tag{2.13}$$

The adjoint state $\varphi_{0u} \in W(0, T)$ associated with u and J is introduced as the unique solution to

$$\begin{cases} -\dfrac{d\varphi}{dt} + A^*\varphi + \dfrac{\partial d}{\partial y}(x, t, y_u)\varphi & = & \dfrac{\partial L}{\partial y}(x, t, y_u, u) & \text{in } Q, \\ \partial_\nu \varphi & = & \dfrac{\partial \ell}{\partial y}(x, t, y_u) & \text{in } \Sigma, \\ \varphi(x, T) & = & \dfrac{\partial r}{\partial y}(x, y_u(x, T)) & \text{in } \Omega, \end{cases} \tag{2.14}$$

where A^* is the formally adjoint operator to A. Standard computations show

$$J'(u)v = \int_0^T \sum_{i=1}^m \left\{ \int_\Omega \left(\frac{\partial L}{\partial u_i}(x, t, y_u, u) + \varphi_{0u}e_i(x) \right) dx \right\} v_i(t)\, dt, \tag{2.15}$$

$$\begin{aligned} J''(u)v_1v_2 & = \int_Q \left[\frac{\partial^2 L}{\partial y^2}(x, t, y_u, u)z_{v_1}z_{v_2} + \frac{\partial^2 L}{\partial y \partial u}(x, t, y_u, u) \cdot (z_{v_1}v_2 + z_{v_2}v_1) \right. \\ & \qquad \left. + v_1^\top \frac{\partial^2 L}{\partial u^2}(x, t, y_u, u)v_2 - \varphi_{0u}\frac{\partial^2 d}{\partial y^2}(x, t, y_u)z_{v_1}z_{v_2} \right] dx\,dt, \\ & \quad + \int_\Sigma \frac{\partial^2 \ell}{\partial y^2}(x, t, y_u)z_{v_1}z_{v_2}\, dS\, dt \\ & \quad + \int_\Omega \frac{\partial^2 r}{\partial y^2}(x, y_u(x, T))z_{v_1}(x, T)z_{v_2}(x, T)\, dx. \end{aligned} \tag{2.16}$$

As in (2.16), we will use a dot to denote the inner product of $\mathbb{R}^m$. Notice that the z_{v_i} and φ_{0u} depend on (x, t), while the v_i depend on t only.

We require the following *linearized Slater condition*: There exists a function $u_0 \in L^\infty(0, T; \mathbb{R}^m)$ with $u_a(t) \leq u_0(t) \leq u_b(t)$ for a.e. $t \in (0, T)$ such that

$$g(x, t, \bar{y}(x, t)) + \frac{\partial g}{\partial y}(x, t, \bar{y}(x, t))z_{u_0 - \bar{u}}(x, t) < 0 \quad \forall (x, t) \in K. \tag{2.17}$$

Inserting $t = 0$, this imposes a condition on the initial state y_0. Therefore, to satisfy (2.17), we have to assume that

$$g(x, 0, y_0(x)) < 0 \quad \forall x \in \bar{\Omega} \text{ with } (x, 0) \in K. \tag{2.18}$$

Defining the Hamiltonian H by

$$H(x, t, y, u, \varphi) = L(x, t, y, u) + \varphi \left[\sum_{i=1}^m e_i(x)\, u_i - d(x, t, y) \right],$$

the first-order necessary conditions admit the following form:

Theorem 4. *Let $\bar{u}$ be a local solution of* (P1). *Suppose that the assumptions* (**A.1**) *–* (**A.4**) *hold and assume the Slater condition* (2.17) *with some $u_0 \in L^\infty(0,T;\mathbb{R}^m)$, $u_a(t) \le u_0(t) \le u_b(t)$ for a.e. $t \in (0,T)$. Then there exists a measure $\bar{\mu} \in M(K)$ and a function $\bar{\varphi} \in L^s(0,T;W^{1,\sigma}(\Omega))$ for all $s, \sigma \in [1,2)$ with $(2/s)+(n/\sigma) > n+1$ such that*

$$\begin{cases} -\dfrac{d\bar{\varphi}}{dt} + A^*\bar{\varphi} + \dfrac{\partial d}{\partial y}(x,t,\bar{y})\,\bar{\varphi} &= \dfrac{\partial L}{\partial y}(x,t,\bar{y},\bar{u}) + \dfrac{\partial g}{\partial y}(x,t,\bar{y})\bar{\mu}_{|Q}, \\[2mm] \partial_\nu \bar{\varphi}(x,t) &= \dfrac{\partial \ell}{\partial y}(x,t,y_u(x,t)) + \dfrac{\partial g}{\partial y}(x,t,\bar{y}(x,t))\bar{\mu}_{|\Sigma}, \\[2mm] \bar{\varphi}(x,T) &= \dfrac{\partial r}{\partial y}(x,\bar{y}(x,T)) + \dfrac{\partial g}{\partial y}(x,T,\bar{y}(x,T))\bar{\mu}_{|\Omega\times\{T\}} \end{cases}$$
(2.19)

for a.a. $x \in \Omega$, $t \in (0,T)$, where $\bar{\mu}_{|Q}$, $\bar{\mu}_{|\Sigma}$, and $\bar{\mu}_{|\Omega\times\{T\}}$ denote the restrictions of $\bar{\mu}$ to Q, Σ, and $\Omega \times \{T\}$, respectively,

$$\int_K (z(x,t) - g(x,t,\bar{y}(x,t))d\bar{\mu}(x,t) \le 0 \quad \forall z \in C(K) \quad \text{with } z(x,t) \le 0 \,\forall (x,t) \in K,$$
(2.20)

and, for almost all $t \in (0,T)$,

$$\int_\Omega \frac{\partial H}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(t),\bar{\varphi}(x,t))\,dx \cdot (u - \bar{u}(t)) \ge 0 \;\forall\, u \in [u_a(t), u_b(t)].$$
(2.21)

This theorem follows from Casas [8] using the admissible set

$$\tilde{U}_{ad} = \{v \mid v = \sum_{i=1}^m e_i(\cdot)u_i(\cdot), \; u_{a,i}(t) \le u_i(t) \le u_{b,i}(t) \text{ a.e. on } (0,T)\}$$

and writing the associated variational inequality in terms of u. The inequality (2.20) implies the well-known *complementary slackness condition*

$$\int_K g(x,t,\bar{y}(x,t))d\bar{\mu}(x,t) = 0.$$
(2.22)

Notice that $\partial H/\partial u$ is a $m-$vector function. Since we need the integrated form of H and its derivatives at the optimal point, we introduce the vector function

$$\mathbf{H}_u(t) = \int_\Omega \frac{\partial H}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(t),\bar{\varphi}(x,t))\,dx$$
(2.23)

and the (m,m)-matrix valued function

$$\mathbf{H}_{uu}(t) = \int_\Omega \frac{\partial^2 H}{\partial u^2}(x,t,\bar{y}(x,t),\bar{u}(t),\bar{\varphi}(x,t))\,dx$$
(2.24)

with entries

$$\mathbf{H}_{u_i u_j}(t) := (\bar{H}_{uu})_{ij}(t), \quad i,j \in \{1,\dots,m\}.$$

The (reduced) Lagrange function is defined in a standard way by

$$\begin{aligned} \mathcal{L}(u,\mu) &= \int_Q L(x,t,y_u(x,t),u(t))\,dxdt + \int_\Sigma \ell(x,t,y_u(x,t))\,dSdt \\ &+ \int_\Omega r(x,y_u(x,T))\,dx + \int_K g(x,t,y_u(x,t))\,d\mu(x,t). \end{aligned}$$

For later use, we establish the second-order derivative of $\mathcal{L}$ with increments $v_i \in L^\infty(0,T;\mathbb{R}^m)$, $i = 1,2$. The expressions below contain functions $z_{v_i}(x,t)$ and $v_i(t)$,

but we suppress their arguments for short.

$$
\begin{aligned}
\frac{\partial^2 \mathcal{L}}{\partial u^2}(u,\mu)v_1 v_2 = & \int_Q \left[\frac{\partial^2 L}{\partial y^2}(x,t,y_u,u)z_{v_1}z_{v_2} \right. \\
& + \frac{\partial^2 L}{\partial y \partial u}(x,t,y_u,u) \cdot (z_{v_1}v_2 + z_{v_2}v_1) + v_1^\top \frac{\partial^2 L}{\partial u^2}(x,t,y_u,u)v_2 \\
& \left. -\varphi_u \frac{\partial^2 d}{\partial y^2}(x,t,y_u)z_{v_1}z_{v_2} \right] dxdt + \int_\Omega \frac{\partial^2 r}{\partial y^2}(x,y_u(x,T))z_{v_1}(x,T)z_{v_2}(x,T)\,dx \\
& + \int_0^T \frac{\partial^2 \ell}{\partial y^2}(x,t,y_u)z_{v_1}z_{v_2}\,dt + \int_K \frac{\partial^2 g}{\partial y^2}(x,t,y_u)z_{v_1}z_{v_2}\,d\bar{\mu}(x,t),
\end{aligned}
$$
(2.25)

where φ_u is the solution of (2.19) with u taken for $\bar{u}$, y_u for $\bar{y}$, and μ for $\bar{\mu}$, respectively.

2.5. **Second-order sufficient optimality conditions.** The regularity results of the preceding section at hand, we are able to discuss second-order sufficient conditions for arbitrary dimensions of Ω. To establish them and to show their sufficiency, we follow the recent paper [9] by Casas et al. . The presentation here is similar to the one in [9], but there occur some differences due to the appearance of vector-valued control functions. We do not assume that the reader is familiar with the results of [9]. Therefore, we sketch the main steps of the analysis. For an easier comparison with the arguments presented in [9], we adopt also part of the notation used there.

First, we define the cone of critical directions associated with $\bar{u}$ by

$$C_{\bar{u}} = \{h \in L^2(0,T;\mathbb{R}^m) : h \text{ satisfies } (2.26), (2.27) \text{ and } (2.28) \text{ below}\},$$

$$
\forall i \in \{1,\dots,m\}, \quad h_i(t) = \begin{cases} \geq 0 & \text{if } \bar{u}_i(t) = u_{a,i}(t), \\ \leq 0 & \text{if } \bar{u}_i(t) = u_{b,i}(t), \\ = 0 & \text{if } \mathbf{H}_{u,i}(t) \neq 0, \end{cases}
$$
(2.26)

$$
\frac{\partial g}{\partial y}(x,t,\bar{y}(x,t))z_h(x,t) \leq 0 \quad \text{if} \quad g(x,t,\bar{y}(x,t)) = 0,
$$
(2.27)

$$
\int_K \frac{\partial g}{\partial y}(x,t,\bar{y}(x,t))z_h(x,t)\,d\bar{\mu}(x,t) = 0.
$$
(2.28)

The vector function $\mathbf{H}_u$ was defined in (2.23).

Moreover, we define, for fixed $\tau > 0$ and all $i \in \{1,\dots,m\}$, the sets of "sufficiently active control constraints"

$$
E_i^\tau = \{t \in [0,T] : |\mathbf{H}_{u,i}(t)| \geq \tau\}.
$$
(2.29)

In the next theorem, we write $\text{diag}(\chi_{E_i^\tau}(t))$ for the matrix $\text{diag}(\chi_{E_1^\tau}(t),\dots,\chi_{E_m^\tau}(t))$.

The second-order sufficient optimality conditions for $\bar{u}$ are stated in the following result:

Theorem 5. *Let $\bar{u}$ be a feasible control of problem* (P1) *that satisfies, together with the associated state $\bar{y}$ and $(\bar{\varphi}, \bar{\mu}) \in L^s(0,T;W^{1,\sigma}(\Omega)) \times M(K)$ for all s, $\sigma \in [1,2)$ with $(2/s) + (n/\sigma) > n + 1$, the first-order conditions* (2.19)-(2.21). *Assume in addition that there exist constants $\omega > 0, \alpha_0 > 0$, and $\tau > 0$ such that for all $\alpha > \alpha_0$*

$$
d^\top \Big(\mathbf{H}_{uu}(t) + \alpha\, diag(\chi_{E_i^\tau}(t)) \Big) d \quad \geq \quad \omega|d|^2 \ a.e. \ t \in [0,T], \ \forall d \in \mathbb{R}^m, \quad (2.30)
$$

$$
\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu})h^2 \quad > \quad 0 \quad \forall h \in C_{\bar{u}} \setminus \{0\}. \quad (2.31)
$$

Then there exist $\varepsilon > 0$ and $\delta > 0$ such that, for every admissible control u of problem (P1)*, the following inequality holds*

$$J(\bar{u}) + \frac{\delta}{2}\|u - \bar{u}\|_{L^2(0,T;\mathbb{R}^m)}^2 \leq J(u) \quad if \quad \|u - \bar{u}\|_{L^\infty(0,T;\mathbb{R}^m)} < \varepsilon. \qquad (2.32)$$

Proof. The proof is by contradiction. It follows the one presented in Casas et al. [9] for an elliptic control problem. Nevertheless, we sketch the main steps, since there are some essential changes due to the different nature of our problem.

Suppose that $\bar{u}$ does not satisfy the quadratic growth condition (2.32). Then there exists a sequence $\{u_k\}_{k=1}^\infty \subset L^\infty(0,T;\mathbb{R}^m)$ of feasible controls for (P1) such that $u_k \to \bar{u}$ in $L^\infty(0,T;\mathbb{R}^m)$ and

$$J(\bar{u}) + \frac{1}{k}\|u_k - \bar{u}\|_{L^2(0,T;\mathbb{R}^m)}^2 > J(u_k) \quad \forall k. \qquad (2.33)$$

Define $\rho_k = \|u_k - \bar{u}\|_{L^2(0,T;\mathbb{R}^m)}$ and

$$h_k = \frac{1}{\rho_k}\|u_k - \bar{u}\|_{L^2(0,T;\mathbb{R}^m)}.$$

Since $\|h_k\|_{L^2(0,T;\mathbb{R}^m)} = 1$, a weakly converging subsequence can be extracted. W.l.o.g. we can assume that $h_k \rightharpoonup h$ weakly in $L^2(0,T;\mathbb{R}^m)$, $k \to \infty$. Now, the proof is split into several steps.

Step 1: It is shown that

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = 0. \qquad (2.34)$$

The arguments are the same as in [9]. Moreover, they are analogous to the classical proof for finite-dimensional problems. Therefore, we omit them.

Step 2: $h \in C_{\bar{u}}$. We have to confirm (2.26)–(2.28). It is easy to verify that h satisfies the sign conditions of (2.26). To see that $h_i(t)$ vanishes, where $\mathbf{H}_{u,i}(t) \neq 0$, we notice that (2.21) implies after a standard discussion that, for all $i \in \{1, \ldots, m\}$

$$|\mathbf{H}_{u,i}(t)h_i(t)| = \mathbf{H}_{u,i}(t)h_i(t) \geq 0 \quad \text{for a.a. } t \in [0,T]. \qquad (2.35)$$

Therefore,

$$\int_0^T \sum_{i=1}^m |\mathbf{H}_{u,i}(t)\, h_i(t)|\, dt = \int_0^T \mathbf{H}_u(t) \cdot h(t)\, dt = \frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = 0$$

must hold in view of (2.34). This implies $h_i(t) = 0$ if $\mathbf{H}_{u,i}(t) \neq 0$, hence (2.26) is verified.

The proof of (2.27) is fairly standard and follows from

$$z_h = G'(\bar{u})h = \lim_{k \to \infty} \frac{y_{\bar{u}+\rho_k h_k} - \bar{y}}{\rho_k} \quad \text{in } C(\bar{Q}),$$

and from $g(x,t,y_{\bar{u}+\rho_k h_k}(x,t)) = g(x,t,y_k(x,t)) \leq 0$ for every $(x,t) \in K$, since u_k is feasible. Notice that $g(x,t,\bar{y}(x,t)) = 0$ holds for all (x,t) considered in (2.27).

It remains to show (2.28). We get from the complementary slackness condition (2.22) that

$$\begin{aligned} \int_K \frac{\partial g}{\partial y}(\cdot,\bar{y})z_h\, d\bar{\mu} &= \lim_{k \to \infty} \frac{1}{\rho_k}\int_K \left[g(\cdot, y_{\bar{u}+\rho_k h_k}) - g(\cdot,\bar{y})\right] d\bar{\mu} = \\ &= \lim_{k \to \infty} \frac{1}{\rho_k}\int_K g(\cdot, y_{u_k})\, d\bar{\mu} \leq 0, \qquad (2.36) \end{aligned}$$

since $\bar{\mu} \geq 0$ and $g(x,t,y_{u_k}(x,t)) \leq 0$ on K. On the other hand, (2.33) yields

$$J'(\bar{u})h = \lim_{k \to \infty} \frac{J(\bar{u} + \rho_k h_k) - J(\bar{u})}{\rho_k} = \lim_{k \to \infty} \frac{J(u_k) - J(\bar{u})}{\rho_k} \leq \lim_{k \to \infty} \frac{\rho_k}{k} = 0. \quad (2.37)$$

Now, (2.34), (2.36), (2.37) and the simple identity

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h = J'(\bar{u})h + \int_K \frac{\partial g}{\partial y}(x, t, \bar{y}(x, t))z_h(x, t)\, d\bar{\mu}(x, t)$$

imply that

$$J'(\bar{u})h = \int_K \frac{\partial g}{\partial y}(x, t, \bar{y}(x, t))z_h(x, t)\, d\bar{\mu}(x, t) = 0,$$

hence (2.28) holds and we have shown $h \in C_{\bar{u}}$.

Step 3: $h = 0$. In view of (2.31), it suffices to show

$$h^{\top}\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu})h \le 0. \tag{2.38}$$

For this purpose, we perform a second-order Taylor expansion of the Lagrangian,

$$\mathcal{L}(u_k, \bar{\mu}) = \mathcal{L}(\bar{u}, \bar{\mu}) + \rho_k \frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h_k + \frac{\rho_k^2}{2}\frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k, \bar{\mu})h_k^2, \tag{2.39}$$

where w_k is an intermediate point between $\bar{u}$ and u_k. Re-writing this equation, we get

$$\rho_k \frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h_k + \frac{\rho_k^2}{2}\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu})h_k^2$$

$$= \mathcal{L}(u_k, \bar{\mu}) - \mathcal{L}(\bar{u}, \bar{\mu}) + \frac{\rho_k^2}{2}\left[\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) - \frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k, \bar{\mu})\right]h_k^2. \tag{2.40}$$

Moreover, we mention that (2.33) can be written in terms of $\mathcal{L}$ as

$$\mathcal{L}(u_k, \bar{\mu}) - \mathcal{L}(\bar{u}, \bar{\mu}) \le \frac{\rho_k^2}{k}. \tag{2.41}$$

Taking into account the expression (2.25) of the second derivative of the Lagrangian, the assumptions **(A.1)**-**(A.4)**, Theorem 1, (2.12) and the fact that $u_k \to \bar{u}$ in $L^{\infty}(0, T; \mathbb{R}^m)$ and $\|h_k\|_{L^2(0,T;\mathbb{R}^m)} = 1$, we get that

$$\left|\left[\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) - \frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k, \bar{\mu})\right]h_k^2\right| \le \left\|\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) - \frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k, \bar{\mu})\right\|_{B(L^2(0,T;\mathbb{R}^m))} \|h_k\|^2_{L^2(0,T;\mathbb{R}^m)}$$

$$= \left\|\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) - \frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k, \bar{\mu})\right\|_{B(L^2(0,T;\mathbb{R}^m))} \to 0 \quad \text{when} \quad k \to \infty, \tag{2.42}$$

where $B(L^2(0, T; \mathbb{R}^m))$ is the space of quadratic forms in $L^2(0, T; \mathbb{R}^m)$.

From (2.35) and the definition of h_k we know that $\mathbf{H}_{u,i}(t)\, h_{k,i}(t) \ge 0$ for a.a. $t \in [0, T]$ and all $i \in \{1, \ldots, m\}$, therefore

$$\frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h_k = \int_0^T \mathbf{H}_u(t) \cdot h_k(t)\, dt \ge \sum_{i=1}^m \int_{E_i^{\tau}} |\mathbf{H}_{u,i}(t)||h_{k,i}(t)|\, dt$$

$$\ge \sum_{i=1}^m \tau \int_{E_i^{\tau}} |h_{k,i}(t)|\, dt. \tag{2.43}$$

For any $\varepsilon > 0$, there exists k_{ε} such that

$$\|\rho_k h_k\|_{L^{\infty}(0,T;\mathbb{R}^m)} = \|\bar{u} - u_k\|_{L^{\infty}(0,T;\mathbb{R}^m)} < \varepsilon \quad \forall k \ge k_{\varepsilon},$$

therefore

$$\frac{\rho_k^2 h_{k,i}^2(t)}{\varepsilon} \le \rho_k|h_{k,i}(t)| \quad \forall k \ge k_{\varepsilon}, \quad \text{a.e. } t \in [0, T].$$

From this inequality and (2.43) it follows that

$$\rho_k \frac{\partial \mathcal{L}}{\partial u}(\bar{u}, \bar{\mu})h_k \ge \rho_k \tau \sum_{i=1}^m \int_{E_i^{\tau}} |h_{k,i}(t)|\, dt \ge \frac{\rho_k^2 \tau}{\varepsilon}\sum_{i=1}^m \int_{E_i^{\tau}} h_{k,i}^2(t)\, dt. \tag{2.44}$$

Collecting (2.40)-(2.42) and (2.44) and dividing by $\rho_k^2/2$ we obtain for any $k \geq k_\varepsilon$

$$\frac{2\tau}{\varepsilon} \sum_{i=1}^{m} \int_{E_i^\tau} h_{k,i}^2(t)\, dt + \frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) h_k^2 \leq \frac{2}{k} + \left\| \frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) - \frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k, \bar{\mu}) \right\|_{B(L^2(0,T;\mathbb{R}^m))}.$$
(2.45)

Let us consider the left-hand side of this inequality. First of all we notice that from (2.25) and

$$\mathbf{H}_{uu}(t) = \int_\Omega \frac{\partial^2 L}{\partial u^2}(x, t, \bar{y}(x,t), \bar{u}(t))\, dx \qquad (2.46)$$

it follows that

$$\begin{aligned}
\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) h_k^2 &= \int_0^T \left[h_k(t)^\top \mathbf{H}_{uu}(t) h_k(t) + 2 \int_\Omega \bar{L}_{uy}(x,t) z_{h_k}(x,t)\, dx \cdot h_k(t) \right] dt \\
&\quad + \int_Q \bar{L}_{yy}(x,t) z_{h_k}^2(x,t)\, dxdt + \int_\Sigma \bar{\ell}_{yy}(x,t) z_{h_k}^2(x,t)\, dsdt \\
&\quad + \int_\Omega \bar{r}_{yy}(x) z_{h_k}^2(x,T)\, dx + \int_K \frac{\partial^2 g}{\partial y^2}(x,t,\bar{y}(x,t)) z_{h_k}^2(x,t)\, d\bar{\mu}(x,t),
\end{aligned}$$

where $\mathbf{H}_{uu}(t)$ was defined in (2.24) and $\bar{L}_{uy}$, $\bar{L}_{yy}$, $\bar{\ell}_{yy}$, and $\bar{r}_{yy}$ are defined by

$$\bar{L}_{uy}(x,t) = \frac{\partial^2 L}{\partial u \partial y}(x,t,\bar{y}(x,t),\bar{u}(t)), \quad \bar{L}_{yy}(x,t) = \frac{\partial^2 L}{\partial y^2}(x,t,\bar{y}(x,t),\bar{u}(t)),$$

$$\bar{\ell}_{yy}(x,t) = \frac{\partial^2 \ell}{\partial y^2}(x,t,\bar{y}(x,t)), \quad \bar{r}_{yy}(x) = \frac{\partial^2 r}{\partial y^2}(x,\bar{y}(x,T)).$$

The first integral of $\frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) h_k^2$ needs special care. We notice that

$$\frac{2\tau}{\varepsilon} \sum_{i=1}^{m} \int_{E_i^\tau} h_{k,i}^2(t)\, dt = \int_0^T h_k(t)^\top \left(\frac{2\tau}{\varepsilon} \operatorname{diag}(\chi_{E_i^\tau}(t)) \right) h_k(t) dt. \qquad (2.47)$$

Thanks to assumption (2.30), it holds that

$$d^\top \left(\mathbf{H}_{uu}(t) + \frac{2\tau}{\varepsilon} \operatorname{diag}(\chi_{E_i^\tau}(t)) \right) d \geq \omega\, |d|^2 \quad \forall\, 0 < \varepsilon < \varepsilon_0, \qquad (2.48)$$

if ε_0 is taken sufficiently small. Therefore, the matrix function $\mathbf{H}_{uu}(t) + \frac{2\tau}{\varepsilon} \operatorname{diag}(\chi_{E_i^\tau}(t))$ is uniformly positive definite, and hence we infer that

$$\begin{aligned}
\liminf_{k\to\infty} &\int_0^T h_k^\top(t) \left(\mathbf{H}_{uu}(t) + \frac{2\tau}{\varepsilon} \operatorname{diag}(\chi_{E_i^\tau})(t)) \right) h_k(t)\, dt \\
&\geq \int_0^T h^\top(t) \left(\mathbf{H}_{uu}(t) + \frac{2\tau}{\varepsilon} \operatorname{diag}\chi_{E_i^\tau}(t)) \right) h(t)\, dt.
\end{aligned} \qquad (2.49)$$

Finally, taking into account that, by Corollary 2 with $c_0 = \frac{\partial d}{\partial y}(\cdot, \bar{y})$, $z_{h_k} \to z_h$ strongly in $C(\bar{Q})$, we deduce from (2.45)-(2.49) and (2.42)

$$\begin{aligned}
\int_0^T &h^\top(t) \left(\mathbf{H}_{uu}(t) + \frac{2\tau}{\varepsilon} \operatorname{diag}\chi_{E_i^\tau}(t)) \right) h(t)\, dt + \int_Q 2\, \bar{L}_{uy}(x,t) z_h(x,t) \cdot h(t)\, dxdt \\
&+ \int_Q \bar{L}_{yy}(x,t) z_h^2(x,t)\, dxdt + \int_\Sigma \bar{\ell}_{yy}(x,t) z_h^2(x,t)\, dsdt \\
&+ \int_\Omega \bar{r}_{yy}(x) z_h^2(x,T)\, dx + \int_K \frac{\partial^2 g}{\partial y^2}(x,t,\bar{y}(x,t)) z_h^2(x,t)\, d\bar{\mu}(x,t) \leq 0.
\end{aligned}$$

This expression can be written as

$$\frac{2\tau}{\varepsilon} \sum_{i=1}^{m} \int_{E_i^\tau} h_i^2(t)\, dt + \frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \bar{\mu}) h^2 \leq 0,$$

which along with (2.31) and the fact that $h \in C_{\bar{u}}$ implies that $h = 0$.

Step 4: $h_k \to 0$ *strongly in* $L^2(0,T;\mathbb{R}^m)$. We have already proved that $h_k \rightharpoonup 0$ weakly in $L^2(0,T;\mathbb{R}^m)$, therefore $z_{h_k} \to 0$ strongly in $C(\bar{Q})$ by Corollary 2. In view of (2.45) and $\|h_k\|_{L^2(0,T;\mathbb{R}^m)} = 1$ we get

$$
\begin{aligned}
0 \;<\; \omega \;=\; & \omega \limsup_{k\to\infty} \int_0^T |h_k(t)|^2\, dt \\
\leq\; & \limsup_{k\to\infty} \int_0^T h_k^\top(t)\Big(\mathbf{H}_{uu}(t) + \frac{2\tau}{\varepsilon}\,\mathrm{diag}(\chi_{E_i^\tau}(t))\Big)h_k(t)\, dt \\
\leq\; & \limsup_{k\to\infty}\left\{ \frac{2}{k} + \left\| \frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u},\bar{\mu}) - \frac{\partial^2 \mathcal{L}}{\partial u^2}(w_k,\bar{\mu}) \right\|_{B(L^2(0,T;\mathbb{R}^m))} \right. \\
& - \int_Q \Big[2\,\bar{L}_{uy}(x,t)z_h(x,t)\cdot h(t) - \bar{L}_{yy}(x,t)z_h^2(x,t) \Big]\,dxdt \\
& - \int_\Sigma \bar{\ell}_{yy}(x,t)z_h^2(x,t)\,dsdt - \int_\Omega \bar{r}_{yy}(x)z_h^2(x,T)\,dx \\
& - \int_K \frac{\partial^2 g}{\partial y^2}(x,t,\bar{y}(x,t))z_h^2(x,t)\,d\bar{\mu}(x,t) = 0.
\end{aligned}
$$

Thus we have got a contradiction so that our hypotheses (2.33) cannot be true. □

Remark 2. *A study of the proof reveals, why the special form of time-dependent controls is important: The sequence (h_k) converges in $L^2(0,T;\mathbb{R}^m)$. This space is associated with the domain $(0,T)$, where the controls can vary (notice that the functions $e_i \in L^\infty(\Omega)$ are fixed). Controls of $L^2(0,T;\mathbb{R}^m)$ are mapped continuously to continuous state functions. This fact is essential for the theory to work. For more general control functions $f \in L^2(0,T;L^\infty(\Omega))$, the associated domain of variability is $\Omega\times(0,T)$. Here, the space $L^2(\Omega\times(0,T))$ would underly the proof. However, state functions associated with controls of $L^2(\Omega\times(0,T))$ are not in general continuous so that the proof does not go through.*

Let us briefly comment on the consequences of the assumption (2.30) for the matrix function $\mathbf{H}_{uu}$. To this aim, we consider a certain index set $I \subset \{1,\ldots,m\}$ and assume that

$$
E = \bigcap_{i\in I} E_i^\tau \tag{2.50}
$$

is non-empty. To avoid a re-numbering, let us assume that $I = \{1,\ldots,k\}$. We write $\mathbf{H}_{uu} + \alpha\,\mathrm{diag}(\chi_E(t))$ as a block matrix in the form

$$
\mathbf{H}_{uu}(t) + \alpha\,\mathrm{diag}(\chi_E(t)) = \left[\begin{array}{cc} A(t) + \alpha\,Id & B(t) \\ B^\top(t) & C(t) \end{array} \right].
$$

Then (2.30) is satisfied on the set E, if $C(t)$ is uniformly positive definite there: There must exist $\beta > 0$ such that

$$
\xi^\top C(t)\xi \geq \beta\,|\xi|^2 \;\;\forall \xi \in \mathbb{R}^{m-k},
$$

but no condition on $A(t)$ is needed. This is seen as follows: We split $d \in \mathbb{R}^m$ in $d_1 \in \mathbb{R}^k$, containing the first k components and $d_2 \in \mathbb{R}^{m-k}$ with the remaining $m-k$ components. Then

$$
\begin{aligned}
d^\top\Big(\mathbf{H}_{uu}(t) + \alpha\,\mathrm{diag}(\chi_E(t))\Big)d \;=\; & d_1^\top A(t)d_1 + \alpha\,|d_1|^2 + 2d_1^\top B(t)d_2 + d_2^\top C(t)d_2 \\
\geq\; & \alpha\,|d_1|^2 - c_A|d_1|^2 - c_B|d_1|^2 - \varepsilon|d_2|^2 + \beta|d_2|^2 \\
\geq\; & \frac{\alpha}{2}|d_1|^2 + \frac{\beta}{2}|d_2|^2 \geq \omega|d|^2
\end{aligned}
$$

by the Young inequality, if α is sufficiently large, $\varepsilon > 0$ is taken sufficiently small, and $\omega = \min\{\alpha,\beta\}/2$.

In the particular case $k = m$, i.e. on $E = \bigcap_{i=1}^{m} E_i^\tau$, we do not need any condition on $\mathbf{H}_{uu}$, while $\mathbf{H}_{uu}$ must be uniformly positive definite on $[0, T] \setminus \bigcup_{i=1}^{m} E_i^\tau$, where no strong activity helps.

Certainly, there are many possible index sets I, and the discussion of all different cases is mainly a matter of combinatorics. We do not dwell upon this point. Instead, let us discuss the case $m = 2$, where $\mathbf{H}_{uu} + \alpha \operatorname{diag}(\chi_{E_1^\tau}, \chi_{E_2^\tau})$ is given as follows:

– In $E_1^\tau \cap E_2^\tau$, we have

$$\mathbf{H}_{uu} + \alpha \operatorname{diag}(\chi_{E_1^\tau}, \chi_{E_2^\tau}) = \left[\begin{array}{cc} H_{11}(t) + \alpha & H_{12}(t) \\ H_{12}(t) & H_{22}(t) + \alpha \end{array} \right].$$

This matrix is uniformly positive definite on $E_1^\tau \cap E_2^\tau$ for all sufficiently large α. Therefore, the matrix $\mathbf{H}_{uu}$ satisfies (2.30) on this subset without any further assumption on positive definiteness of $\mathbf{H}_{uu}$.

– In $E_1^\tau \setminus (E_1^\tau \cap E_2^\tau)$, it holds

$$\mathbf{H}_{uu} + \alpha \operatorname{diag}(\chi_{E_1^\tau}, \chi_{E_2^\tau}) = \left[\begin{array}{cc} H_{11}(t) + \alpha & H_{12}(t) \\ H_{12}(t) & H_{22}(t) \end{array} \right]$$

so that we need $H_{22}(t) \geq \beta > 0$ to make the matrix positive definite on $E_1^\tau \setminus (E_1^\tau \cap E_2^\tau)$ for sufficiently large α. An analogous condition must be imposed on $H_{11}(t)$ on $E_2^\tau \setminus (E_1^\tau \cap E_2^\tau)$.

– In $[0, T] \setminus (E_1^\tau \cup E_2^\tau)$, the matrix $\mathbf{H}_{uu}(t)$ must be uniformly positive definite to satisfy (2.30), since $\operatorname{diag}(\chi_{E_1^\tau}, \chi_{E_2^\tau})$ vanishes here.

Finally, we mention the case of a diagonal matrix $\mathbf{H}_{uu}(t) = \operatorname{diag}(H_{u_i u_i}(t))$. Here, (2.30) is satisfied, if and only if

$$\forall i \in \{1, \ldots, m\}: \quad H_{u_i u_i}(t) \geq \omega \quad \forall t \in [0, T] \setminus E_i^\tau.$$

This happens in the standard setting where u appears only in a Tikhonov regularization term, i.e.

$$L = L(x, t, y) + \lambda |u|^2.$$

3. Semilinear elliptic case

3.1. **Problem statement.** In this section, the following elliptic problem with pointwise state constraints is considered:

$$\textbf{(P2)} \begin{cases} \min_{u \in \mathcal{U}_{ad}} J(u) = \displaystyle\int_\Omega L(x, y(x), u) \, dx + \int_\Gamma \ell(x, y(x)) \, dS(x) \\[2mm] \text{subject to} \\[2mm] \begin{aligned} -\Delta y(x) + y(x) + d(x, y(x), u) &= 0 \quad \text{in} \quad \Omega \\ \partial_\nu y(x) + b(x, y(x)) &= 0 \quad \text{on} \quad \Gamma, \end{aligned} \\[2mm] g(x, y(x)) \leqslant 0 \quad \text{for all } x \in \bar{\Omega}, \end{cases} \tag{3.1}$$

where $\mathcal{U}_{ad} := \{u \in \mathbb{R}^m : u_a \leqslant u \leqslant u_b\}$. In this setting, $\Omega \subset \mathbb{R}^n$ $(n \geqslant 2)$ is a bounded domain with Lipschitz boundary Γ. Moreover, sufficiently smooth functions $L, d : \Omega \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, $g : \bar{\Omega} \times \mathbb{R} \to \mathbb{R}$, and $b, \ell : \Gamma \times \mathbb{R} \to \mathbb{R}$ are given; u_a and u_b with $u_a \leqslant u_b$ are vectors of $\mathbb{R}^m$. We adopt the notation from the preceding sections.

Similar to [10], [12], the following assumptions are imposed on L, ℓ, d, b:

(A.5) (Carathèodory type assumption) For each fixed $x \in \Omega$ or Γ respectively, the functions $L = L(x, y, u)$, $\ell = \ell(x, y)$, $d = d(x, y, u)$, $b = b(x, y)$, are of class C^2 with respect to (y, u). For all fixed (y, u) or fixed y, respectively, they are Lebesgue measurable with respect to the variable $x \in \Omega$, or $x \in \Gamma$ respectively. The function $g = g(x, y)$ is supposed to be twice continuously differentiable with respect to y on $\bar{\Omega} \times \mathbb{R}$.

(A.6) (Monotonicity) For almost all $x \in \Omega$, or $x \in \Gamma$, respectively, and any fixed $u \in \mathcal{U}_{ad}$ when applies, it holds

$$\frac{\partial d}{\partial y}(x, y, u) \geq 0, \quad \frac{\partial b}{\partial y}(x, y) \geq 0.$$

(A.7) (Boundedness and Lipschitz properties) There is a constant C_0 and, for all $M > 0$, an $C_L(M)$ such that the estimates

$$
\begin{aligned}
|d(x, 0, 0)| + |d'(x, 0, 0)| + |d''(x, 0, 0)| &\leqslant C_0 \\
|d''(x, y_1, u_1) - d''(x, y_2, u_2)| &\leqslant C_L(M)(|y_1 - y_2| + |u_1 - u_2|)
\end{aligned}
$$

hold for almost all $x \in \Omega$, all $u, u_i \in \mathcal{U}_{ad}$, and all $|y_i| \leqslant M$, $i = 1, 2$.
The functions L, φ, b, and g are assumed to satisfy **(A.7)** as well.

3.2. **First-order necessary conditions.** In this section, we consider optimality conditions for problem (P2). Since the controls are in $\mathbb{R}^m$ and infinitely many pointwise state constraints are given, this problem belongs to the class of semi-infinite mathematical programming problems.

Therefore, the first- and second-order optimality conditions might be deduced from the theory of semi-infinite programming. Nevertheless, the transfer of these results to the control of PDEs needs the handling of the associated partial differential equations and to discuss the differentiability properties of the underlying control-to-state mappings so that these facts are worth mentioning.

First, we derive some basic results for the control-to-state mapping that are standard for problems with control functions appearing linearly on the right hand side of the PDE. Since our setting includes controls appearing nonlinearly, we have to slightly update the arguments, and we state them for convenience of the reader.

Theorem 6. *Let $\Omega \subset \mathbb{R}^n$ be a bounded and Lipschitz domain. Suppose that the conditions* **(A.5)** *and* **(A.6)** *are satisfied. Then, for all $u \in \mathcal{U}_{ad}$, the partial differential equation*

$$
\begin{aligned}
-\Delta y(x) + y(x) + d(x, y(x), u) &= 0 \quad \text{in} \quad \Omega \\
\partial_\nu y(x) + b(x, y(x)) &= 0 \quad \text{on} \quad \Gamma,
\end{aligned}
\tag{3.2}
$$

has a unique solution $y_u \in H^1(\Omega) \cap C(\bar{\Omega})$ and the estimate

$$\|y_u\|_{H^1(\Omega)} + \|y_u\|_{C(\bar{\Omega})} \leqslant c_M \tag{3.3}$$

holds with a constant c_M that does not depend on u, if u belongs to $\mathcal{U}_{ad}$.

Proof. The result follows by setting

$$\tilde{d}(x, y) := d(x, y, u),$$

and applying regularity results for semilinear elliptic equations obtained in [3] or [7]. Notice that $\tilde{d}$ is a monotone function with respect to y. □

Remark 3. *Due to (3.3), the control-to-state mapping $G : \mathbb{R}^m \to H^1(\Omega) \cap C(\bar{\Omega})$ that assigns to each $u \in \mathcal{U}_{ad}$ the solution y_u of equation (3.2), satisfies the boundedness condition $\|G(u)\|_{H^1(\Omega) \cap C(\bar{\Omega})} \leqslant c_M$ for all $u \in \mathcal{U}_{ad}$.*

Remark 4. *Due to our assumptions, the weak solution of (3.2) lies in the space*

$$Y_{q,p} = \{y \in H^1(\Omega) : -\Delta y + y \in L^q(\Omega), \quad \partial_\nu y \in L^p(\Gamma)\}$$

for all $p, q \geq 1$. $Y_{q,p}$ *is known to be continuously embedded in* $H^1(\Omega) \cap C(\bar{\Omega})$ *for each* $q > n/2$ *and each* $p > n - 1$, *see* [3] *or* [13].

Remark 5. *The operator* $G : \mathcal{U}_{ad} \to H^1(\Omega) \cap C(\bar{\Omega})$ *obeys the Lipschitz property*

$$\|y_1 - y_2\|_{H^1(\Omega)} + \|y_1 - y_2\|_{C(\bar{\Omega})} \leqslant C_L(M)|u_1 - u_2| \tag{3.4}$$

for all y_i *such that* $y_i = G(u_i)$, *with* $u_i \in \mathbb{R}^m$ *and* $|u_i| \leq M$ *for* $i = 1, 2$.

Remark 6. *Under Assumption* **(A.7)**, *the control-to-state mapping* $G : \mathbb{R}^m \to H^1(\Omega) \cap C(\bar{\Omega})$ *is twice continuously Fréchet-differentiable. For arbitrary elements* $\bar{u}$ *and* u *of* $\mathcal{U}_{ad}$, *and for* $\bar{y} = G(\bar{u})$, *the function* $y = G'(\bar{u})v$ *is the unique solution of the problem*

$$\begin{cases} -\Delta z_v + z_v + \dfrac{\partial d}{\partial y}(x, \bar{y}, \bar{u})z_v &= -\dfrac{\partial d}{\partial u}(x, \bar{y}, \bar{u})v & \text{in } \Omega \\[2mm] \partial_\nu z_v + \dfrac{\partial b}{\partial y}(x, \bar{y})z_v &= 0 & \text{on } \Gamma. \end{cases} \tag{3.5}$$

Moreover, y *satisfies the inequality*

$$\|z_v\|_{H^1(\Omega)} + \|z_v\|_{C(\bar{\Omega})} \leqslant c_\infty |v| \tag{3.6}$$

for some constant c_∞ *independent of* u.

The function $z_{v_1 v_2}$, defined by $z_{v_1 v_2} = G''(u)[v_1, v_2]$, is the unique solution of

$$\begin{cases} -\Delta z_{v_1 v_2} + z_{v_1 v_2} + \dfrac{\partial d}{\partial y}(x, y_u, u)z_{v_1 v_2} &= -(y_{u_1}, u_1^\top)\, d''(x, y_u, u)\,(y_{u_2}, u_2^\top)^\top \ \text{in } \Omega \\[2mm] \partial_\nu z_{v_1 v_2} + \dfrac{\partial b}{\partial y}(x, y_u)z_{v_1 v_2} &= -\dfrac{\partial^2 b}{\partial y^2}(x, y_u)y_{u_1}y_{u_2} & \text{on } \Gamma. \end{cases} \tag{3.7}$$

The proofs of these results are fairly standard. For control functions appearing linearly in the right hand side of (3.2), it is given in [11]. It can also be found in [33]. The adaptation to the vector case needed here is more or less straightforward. We omit the associated arguments.

We first note that the set feasible set of this problem is closed and bounded in $\mathbb{R}^m$ due to the continuity of g. Therefore, the reduced functional f defined above is continuous and compactness guarantees existence of an optimal control $\bar{u}$ provided that the feasible set is non-empty.

Notice that this existence result is not in general true for control functions appearing nonlinearly. In the sequel, $\bar{u}$ stands for an optimal control with state $\bar{y} = G(\bar{u})$. Later, in the context of second-order conditions, it is again a candidate for local optimality. Henceforth we assume the following linearized Slater condition:

(A.8) (Regularity Condition) There exists $u_0 \in \mathcal{U}_{ad}$ such that

$$g(x, \bar{y}(x)) + g_y(x, \bar{y}(x))z_{u_0 - \bar{u}}(x) < 0 \quad \text{for all } x \in \bar{\Omega}, \tag{3.8}$$

where $\bar{y}$ satisfies (3.2).

It is well known, see [24], that this condition guarantees the existence of a nonnegative Lagrange multiplier μ in the space $M(\bar{\Omega}) = C(\bar{\Omega})^*$ of regular Borel measures on $\bar{\Omega}$, and an associated adjoint state $\bar{\varphi} \in W^{1,\sigma}(\Omega)$ for all $1 \leq \sigma < n/(n-1)$, defined by the adjoint equation

$$\begin{cases} -\Delta\bar{\varphi} + \bar{\varphi} + \dfrac{\partial d}{\partial y}(\cdot, \bar{y}, \bar{u})\bar{\varphi} &= \dfrac{\partial L}{\partial y}(\cdot, \bar{y}, \bar{u}) + \bar{\mu}_\Omega \dfrac{\partial g}{\partial y}(\cdot, \bar{y}) & \text{in} \quad \Omega \\[2mm] \partial_\nu\bar{\varphi} + \dfrac{\partial b}{\partial y}(\cdot, \bar{y})\bar{\varphi} &= \dfrac{\partial \ell}{\partial y}(\cdot, \bar{y}) + \bar{\mu}_\Gamma \dfrac{\partial g}{\partial y}(\cdot, \bar{y}) & \text{on} \quad \Gamma, \end{cases} \tag{3.9}$$

where $\bar{\mu}_\Omega$ and $\bar{\mu}_\Gamma$ denote the restrictions of $\bar{\mu} \in M(\bar{\Omega})$ to Ω and Γ, respectively. Following [7] or [13], we know that equation (3.9) admits a unique solution $\bar{\varphi} \in W^{1,\sigma}(\Omega)$, for all $\sigma < n/(n-1)$. Moreover the variational inequality

$$\mathbf{H}_u^{\mathsf{T}}(u - \bar{u}) \geqslant 0, \quad \forall u \in \mathcal{U}_{ad}, \tag{3.10}$$

holds, where the vector $\mathbf{H}_u \in \mathbb{R}^m$ is defined by $\mathbf{H}_u = (\mathbf{H}_{u,i})_{i=1,\dots,m}$, with

$$\mathbf{H}_{u,i} := \int_\Omega \frac{\partial L}{\partial u_i}(x, \bar{y}(x), \bar{u}) - \bar{\varphi}(x)\,\frac{\partial d}{\partial u_i}(x, \bar{y}(x), \bar{u})\ dx,$$

and the complementarity condition

$$\int_{\bar{\Omega}} g(x, \bar{y}(x))d\bar{\mu}(x) = 0 \tag{3.11}$$

is satisfied.

For the ease of later computations, we introduce a different form of the *Lagrange function* $\mathcal{L} : Y_{q,p} \times \mathbb{R}^m \times W^{1,\sigma}(\Omega) \times M(\bar{\Omega}) \to \mathbb{R}$,

$$\mathcal{L}(y, u, \varphi, \mu) = J(y, u) - \int_\Omega (-\Delta y + y + d(x, y, u))\varphi\ dx$$

$$- \int_\Gamma (\partial_\nu y + b(x, y))\varphi\ dS + \langle \mu, g(\cdot, y)\rangle_\Omega, \tag{3.12}$$

which accounts also for the state-equation. Clearly, it holds that $\mathcal{L} = \mathcal{L}$, if y satisfies the state equation. Moreover, it is known that it holds for the partial derivatives of $\mathcal{L}$ with respect to y and u

$$\mathcal{L}_y(\bar{y}, \bar{u}, \varphi, \mu)y = 0 \qquad \forall y \in H^1(\Omega) \tag{3.13}$$

$$\mathcal{L}_u(\bar{y}, \bar{u}, \varphi, \mu)(u - \bar{u}) = \mathbf{H}_u^{\mathsf{T}}(u - \bar{u}) \geqslant 0 \quad \forall u \in \mathcal{U}_{ad}. \tag{3.14}$$

Therefore, the Lagrange function is an appropriate tool to express optimality conditions in a convenient way, in particular it is useful to verify second-order sufficient conditions by checking $\mathcal{L}''(\bar{y}, \bar{u}, \bar{\varphi}, \bar{\mu})[(z_h, h)]^2 > 0$ for all z_h satisfying the linearized equation (3.5). This second derivative of $\mathcal{L}$ with respect to (y, u) is expressed by

$$\mathcal{L}''(\bar{y}, \bar{u}, \varphi, \mu)[(y_1, u_1), (y_2, u_2)] =$$

$$\int_\Omega \frac{\partial^2 L}{\partial y^2}y_1 y_2 + \frac{\partial^2 L}{\partial u \partial y}y_2 \cdot u_1 + \frac{\partial^2 L}{\partial y \partial u}\cdot u_2 y_1 + u_1^{\mathsf{T}}\frac{\partial^2 L}{\partial u^2}u_2\ dx$$

$$+ \int_\Gamma \frac{\partial^2 \ell}{\partial y^2}y_1 y_2\ dS - \int_\Omega \Big(\frac{\partial^2 d}{\partial y^2}y_1 y_2 + \frac{\partial^2 d}{\partial u \partial y}y_2 \cdot u_1$$

$$+ \frac{\partial^2 d}{\partial y \partial u}\cdot u_2 y_1 + u_1^{\mathsf{T}}\frac{\partial^2 d}{\partial u^2}u_2\Big)\varphi\ dx - \int_\Gamma \frac{\partial^2 b}{\partial y^2}y_1 y_2 \varphi\ dS$$

$$+ \Big\langle \mu, \frac{\partial^2 g}{\partial y^2}(\cdot, \bar{y})y_1 y_2\Big\rangle_\Omega, \tag{3.15}$$

where the derivatives of L and d are taken at x, $\bar{y}$, and $\bar{u}$, respectively.

3.3. Second-Order Sufficient Optimality Conditions.
In this section, we discuss second-order sufficient conditions for problem (P2). Since our controls belong to $\mathbb{R}^m$, the low regularity of the adjoint state does not cause troubles in the estimations of $\mathcal{L}''$. Therefore, second-order sufficient conditions can be obtained for arbitrary dimension of the domain.

The proof of these conditions can either be performed analogous to Section 2.4. or derived by transferring the known second-order conditions from the theory of semi-infinite programming problems to our control problem.

Therefore, we state the second-order sufficient optimality conditions without proof. Let $d = (\mathbf{H}_u)$ denote the first-order derivative of the Lagrangian function

with respect to u introduced in (3.10). For convenience, we introduce the following sets: $\mathcal{A}_+ := \{i : d_i > 0\}$, $\mathcal{A}_- := \{i : d_i < 0\}$, and $\mathcal{A} := \mathcal{A}_+ \cup \mathcal{A}_-$. Let $\tau := \min\{|d_i| : i \in \mathcal{A}\} > 0$. We also define the critical cone associated with $\bar{u}$

$$C_{\bar{u}} = \{h \in \mathbb{R}^m : h_i = 0 \quad \forall i \in \mathcal{A}, \text{ and satisfies (3.16)-(3.18)}\}$$

$$h_i = \begin{cases} \geqslant 0 & if \quad \bar{u}_i = u_{a,i} \\ \leqslant 0 & if \quad \bar{u}_i = u_{b,i} \end{cases} \tag{3.16}$$

$$\frac{\partial g}{\partial y}(x, \bar{y}(x))z_h(x) \leqslant 0, \text{ if } g(x, \bar{y}(x)) = 0 \tag{3.17}$$

$$\int_{\bar{\Omega}} \frac{\partial g}{\partial y}(x, \bar{y}(x))z_h(x)d\mu = 0, \tag{3.18}$$

where z_h stands for $G'(\bar{u})h$.

Remark 7. *It is not difficult to see that this definition of the critical cone $C_{\bar{u}}$ is equivalent with the one used for semi-infinite programing in Bonnans and Shapiro [6]. We only include the restrictions on the control in the restriction function of the semi-infinite problem.*

Theorem 7. *Let $\bar{u}$ be a feasible control for problem $(P2)$ with associated state $\bar{y}$ satisfying the first-order necessary conditions formulated in Section 3.2. Assume that*

$$h^\top \frac{\partial^2 \mathcal{L}}{\partial u^2}(\bar{u}, \mu)h > 0, \qquad \forall h \in C_{\bar{u}} \backslash \{0\}. \tag{3.19}$$

Then there exist $\varepsilon > 0$ and $\delta > 0$ such that for every feasible control of $(P2)$

$$J(\bar{u}) + \frac{\delta}{2}|u - \bar{u}|^2 \leqslant J(u)$$

holds for all feasible controls with $|u - \bar{u}| < \varepsilon$.

Remark 8. *As mentioned above, this theorem might be proven as for our parabolic problem. The proof is even simpler, since our controls belong to the space $\mathbb{R}^m$, where the unit sphere is compact. Moreover, it follows from regularity results Casas [7] or Alibert and Raymond [3] that the control-to-state mapping $u \mapsto y$ is of class C^2 from $\mathbb{R}^m$ to $H^1(\Omega) \cap C(\bar{\Omega})$. However, the second-order conditions follow also from the theory of semi-infinite programming, cf. Bonnans and Shapiro [6], if the associated results are re-written in terms of partial differential equations.*

In the same way, parabolic problems for the type

$$\begin{cases} \min_{u \in \mathcal{U}_{ad}} J(y, u) = \int_Q L(x, t, y(x, t)) \, dxdt + \int_\Sigma \ell(x, t, y(x, t), u) \, dS(x)dt \\ \\ \text{subject to} \\ \\ \begin{aligned} y_t - \Delta y(x, t) + d(x, t, y(x, t)) &= 0 \quad \text{in} \quad Q, \\ \partial_\nu y(x, t) + b(x, t, y(x, t), u) &= 0 \quad \text{on} \quad \Sigma, \\ y(x, 0) - y_0(x) &= 0 \quad \text{on} \quad \Omega, \end{aligned} \\ \\ g(x, t, y(x, t)) \leqslant 0 \quad \text{for all } (x, t) \in \bar{Q} \end{cases} \tag{3.20}$$

can be dealt with, where $\mathcal{U}_{ad}$ is defined as before. Notice that here the control appears in the boundary condition, where second-order sufficient conditions have not yet been proven for control *functions* under the presence of pointwise state constraints. Here, the control-to-state mapping is from $\mathbb{R}^m$ to $C(\bar{Q})$, cf. Casas [8]. Again, the second-order conditions can also be derived from the conditions for semi-infinite programming in [6].

3.4. **Examples.** Let us finally illustrate the situation by some examples, which exhibit different types of active sets and some kinds of nonlinearities.

Example 1. *Here we study the optimal control problem*

$$\textbf{(E1)} \quad \begin{cases} \min_{u \in \mathbb{R}^4} J(y, u) = \dfrac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \dfrac{1}{2}|u - u_d|^2 \\[2mm] \text{subject to} \\[1mm] \begin{aligned} -\Delta y(x) + y(x) + \tfrac{1}{7}y(x)^3 &= \textstyle\sum_{i=1}^4 u_i e_i(x) & in & \quad \Omega \\ \partial_\nu y(x) &= 0 & on & \quad \Gamma, \end{aligned} \\[2mm] y(x) \geqslant b(x) \quad \text{for all } x \in \bar\Omega = [0,1] \times [0,1], \end{cases}$$

where

$$b(x) = \begin{cases} 2x_1 + 1, & x_1 < \tfrac{1}{2}, \\ 2 & x_1 \geqslant \tfrac{1}{2}, \end{cases} \quad \text{and } e_i = \begin{cases} 1, & x \in \Omega_i, \\ 0 & \text{otherwise}, \end{cases} \quad i = 1, 2, 3, 4,$$

and $\Omega_1 = (0, \tfrac{1}{2}] \times (0, \tfrac{1}{2}]$, $\Omega_2 = (0, \tfrac{1}{2}] \times (\tfrac{1}{2}, 1)$, $\Omega_3 = (\tfrac{1}{2}, 1) \times [\tfrac{1}{2}, 1)$, *and* $\Omega_4 = (\tfrac{1}{2}, 1) \times (0, \tfrac{1}{2})$. *We shall see that the optimal state is* $\bar y = 2$, *hence the active set for this problem is* $\overline{\Omega_3 \cup \Omega_4}$.

The first-order conditions for this problem are as follows: For the optimal solution $(\bar y, \bar u)$ *in* $H^2(\Omega) \times \mathbb{R}^4$ *of the optimal control problem (E1), there exist a Lagrange multiplier* $\bar\mu \in M(\bar\Omega)$ *and an adjoint state* $\bar\varphi \in W^{1,\sigma}(\Omega)$ *such that*

$$\begin{cases} \begin{aligned} -\Delta \bar y + \bar y + \tfrac{1}{7}\bar y^3 &= \textstyle\sum_{i=1}^4 \bar u_i e_i & in & \quad \Omega \\ \partial_\nu \bar y &= 0 & on & \quad \Gamma, \\ -\Delta \bar\varphi + \bar\varphi + \tfrac{12}{7}\bar\varphi &= \bar y - y_d - \bar\mu & in & \quad \Omega \\ \partial_\nu \bar\varphi &= 0 & on & \quad \Gamma, \end{aligned} \\[2mm] (\bar u - u_d) + \left[\displaystyle\int_\Omega \bar\varphi e_1 \, dx, \ldots, \int_\Omega \bar\varphi e_4 \, dx \right]^\top = 0, \\[3mm] \displaystyle\int_{\bar\Omega} (\bar y - b) d\mu = 0, \quad \bar y(x) \geqslant b(x) \quad \forall x \in \bar\Omega \\[2mm] and \ \bar\mu \geqslant 0. \end{cases} \qquad (3.21)$$

The following quantities satisfy the optimality system:

$$\bar y(x) = 2, \quad \bar u = \frac{22}{7}[1,1,1,1]^\top, \quad \bar\varphi(x_1, x_2) = \begin{cases} -x_1^2 + \dfrac{1}{2}, & x_1 < \tfrac{1}{2}, \\[2mm] \dfrac{1}{4}, & x_1 \geqslant \tfrac{1}{2}, \end{cases}$$

together with the Lagrange multiplier $\bar\mu = \delta_{x_1}(\tfrac{1}{2})$, *where* $\delta_{x_i}(z)$ *denotes the Dirac measure with respect to the variable* x_i, *concentrated at* $x_i = z$.

In fact, it is easy to see that $\bar y$ *and* $\bar u$ *fulfill the state equation in (3.21). Since* $\bar\varphi$ *does not depend on* x_2, *we find that* $\Delta\bar\varphi = \partial_{x_1}^2 \bar\varphi = \delta_{x_1}(\tfrac{1}{2}) - \psi(x_1)$, *where*

$$\psi(x_1) = \begin{cases} 2, & x_1 < \tfrac{1}{2}, \\ 0, & x_1 \geqslant \tfrac{1}{2}. \end{cases}$$

Therefore,

$$-\Delta\bar\varphi + \bar\varphi + \frac{12}{7}\bar\varphi = -\delta_{x_1}(\tfrac{1}{2}) + \begin{cases} 2 + (1 + \tfrac{12}{7})(\tfrac{1}{2} - x_1^2), & x_1 < \tfrac{1}{2}, \\[2mm] \tfrac{1}{4}(1 + \tfrac{12}{7}), & x_1 \geqslant \tfrac{1}{2}, \end{cases}$$

and a simple computation shows that this is equal to $\bar{y} - y_d - \bar{\mu}$, with

$$y_d = \begin{cases} (x_1^2 - \frac{1}{2})(1 + \frac{12}{7}) & x_1 < \frac{1}{2}, \\ 2 - \frac{1}{4}(1 + \frac{12}{7}) & x_1 \geqslant \frac{1}{2}, \end{cases} \qquad u_d = \bar{u} + \left[\frac{5}{48}, \frac{5}{48}, \frac{1}{16}, \frac{1}{16} \right]^\top.$$

To check the second order conditions notice that $\bar{\varphi} \leqslant \frac{1}{2}$, hence we have

$$\mathcal{L}''(\bar{y}, \bar{u}, \bar{\varphi}, \bar{\mu})[(z_h, h)]^2 = \|z_h\|_{L^2(\Omega)}^2 + |h|^2 - \int_\Omega \frac{12}{7} \bar{\varphi} z_h^2 \, dx$$

$$\geq |h|^2 + \frac{1}{7} \int_\Omega z_h^2 \, dx \geqslant |h|^2$$

for all $h \in \mathbb{R}^4$.

Example 2. *This example includes a semilinear elliptic equation with controls appearing nonlinearly. The active set has measure zero. The problem is the following*

(E2) $\begin{cases} \min_{u \in \mathbb{R}^2} J(y, u) = \dfrac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \dfrac{1}{2} |u - u_d|^2 \\[2mm] \text{subject to} \\[2mm] -\Delta y(x) + y(x) + \frac{1}{5} y(x)3 = \frac{1}{5}(u_1 e_1(x) + u_2 e_2(x))^2 \quad in \quad \Omega \qquad (3.22) \\[1mm] \qquad\qquad\qquad \partial_\nu y(x) = 0 \qquad\qquad\qquad\qquad on \quad \Gamma, \\[2mm] \quad y(x) \geqslant b(x) \quad for \ all \ x \in \bar{\Omega} = [0,1] \times [0,1], \end{cases}$

where

$$b(x) = \begin{cases} \frac{1}{2} - \frac{1}{4}(x_1 + x_2), & x_1 + x_2 \geqslant 1, \\ \frac{1}{4}(x_1 + x_2), & x_1 + x_2 < 1, \end{cases}$$

$$e_1 = \begin{cases} \frac{1}{8}, & x_1 + x_2 \geqslant 1, \\ 0 & otherwise, \end{cases} \qquad e_2 = \begin{cases} \frac{1}{4}, & x_1 + x_2 < 1, \\ 0 & otherwise. \end{cases}$$

The optimality system is as in the previous example, with the gradient equation

$$(\bar{u} - u_d) + \frac{2}{5} \begin{bmatrix} \int_\Omega (\bar{u}_1 e_1(x) + \bar{u}_2 e_2(x)) \bar{\varphi} e_1 \, dx \\ \int_\Omega (\bar{u}_1 e_1(x) + \bar{u}_2 e_2(x)) \bar{\varphi} e_2 \, dx \end{bmatrix}$$

$$= (\bar{u} - u_d) + \frac{2}{5} \begin{bmatrix} \int_\Omega \bar{u}_1 e_1(x)^2 \bar{\varphi} \, dx \\ \int_\Omega \bar{u}_2 e_2(x)^2 \bar{\varphi} \, dx \end{bmatrix} = 0. \qquad (3.23)$$

To satisfy the optimality system, we define the quantities

$$\bar{y}(x) = \frac{1}{4}, \quad \bar{u} = \left[9, \frac{9}{2} \right]^\top, \quad \bar{\varphi}(x) = \begin{cases} -(x_1 2 - 2x_1 + 2), & x_1 + x_2 \geqslant 1, \\ -(x_2 2 + 1), & x_1 + x_2 < 1, \end{cases}$$

note that $\bar{\varphi}$ satisfies $\partial_\nu \bar{\varphi} = 0$. Moreover, we define the Lagrange multiplier $\bar{\mu}$ by the positive measure

$$\bar{\mu} = 2(1 - x_1)\delta_{x_2}(1 - x_1) + 2x_2 \delta_{x_1}(1 - x_2).$$

where $\delta_{x_2}(1 - x_1)$ and $\delta_{x_1}(1 - x_2)$ are the Dirac measures with respect to x_2 and x_1 concentrated at $x_2 = 1 - x_1$ and $x_1 = 1 - x_2$, respectively. These quantities satisfy the state equation in (3.22) and the identity (3.23). The active set is the diagonal of the unit square: $x_2 = 1 - x_1$. Let us discuss the adjoint equation. It holds

$$\frac{\partial 2\bar{\varphi}}{\partial x_1 2}(x_1, x_2) = 2x_2 \delta_{x_1}(1 - x_2) + \begin{cases} -2, & x_1 + x_2 \geqslant 1, \\ 0, & x_1 + x_2 < 1, \end{cases},$$

and

$$\frac{\partial 2\bar{\varphi}}{\partial x_2 2}(x_1, x_2) = 2(1 - x_1)\delta_{x_2}(1 - x_1) + \begin{cases} 0, & x_1 + x_2 \geqslant 1, \\ -2, & x_1 + x_2 < 1. \end{cases}$$

Inserting the defined functions in the adjoint equation, we get

$$-\Delta\bar{\varphi} + \bar{\varphi} + \frac{3}{5}\bar{y}2\bar{\varphi} =$$

$$2(1 - x_1)\delta_{x_2}(1 - x_1) + 2x_2\delta_{x_1}(1 - x_2) + \begin{cases} 2 - \frac{83}{80}(x_1 2 - 2x_1 + 2), & x_1 + x_2 \geqslant 1, \\ 2 - \frac{83}{80}(x_2 2 + 1), & x_1 + x_2 < 1. \end{cases}$$

The right-hand side of the last equation must be equal to $\bar{y} - y_d - \bar{\mu}$, hence we define

$$y_d = \begin{cases} \frac{83}{80}(x_1 2 - 2x_1 + 2) - \frac{7}{4} & x_1 + x_2 \geqslant 1, \\ \frac{83}{80}(x_2 2 + 1) - \frac{7}{4} & x_1 + x_2 < 1. \end{cases}$$

From (3.23), we find $u_d = \bar{u} - \frac{1}{10}\left[\frac{21}{64}, \frac{27}{32}\right]^\top$. Finally, we confirm the second-order sufficient conditions. Since $0 \geqslant \bar{\varphi} \geqslant -2$, we have for all $h \in \mathbb{R}^2 \backslash \{0\}$ that

$$\mathcal{L}''(\bar{y}, \bar{u}, \bar{\varphi}, \bar{\mu})[(z_h, h)]^2 = |h|^2 + \int_\Omega z_h^2\left(1 - \frac{6}{5}\bar{y}\bar{\varphi}\right) dx + \frac{2}{5}\int_\Omega (h_1 e_1(x) + h_2 e_2(x))^2 \bar{\varphi}\, dx$$

$$\geqslant |h|^2 + \int_\Omega z_h^2\, dx - \frac{4}{5}\int_\Omega ((h_1 e_1(x))^2 + (h_2 e_2(x))^2)\, dx$$

$$\geqslant |h|^2 - \frac{4}{5}\left(\frac{1}{4}\right)^2 |\Omega||h|^2 \geqslant \frac{19}{20}|h|^2.$$

In the last estimate, we have used $e_i(x) \leq 1/4$ for $i = 1, 2$.

Example 3. *This is a slight modification of an example in [27], where the optimal state is active in one single point.*

$$\textbf{(E3)} \quad \begin{cases} \min\limits_{u \in \mathbb{R}^3} J(y, u) = \dfrac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \dfrac{1}{2}|u - u_d|^2 \\[2mm] \text{subject to} \\[1mm] -\Delta y(x) + y(x) + y(x)^3 = \sum_{i=1}^3 u_i e_i(x) \quad in \quad \Omega \\ \qquad\qquad\qquad \partial_\nu y(x) = 0 \qquad\qquad\qquad on \quad \Gamma, \\[2mm] y(x) \geqslant 2 - |x|^2 \quad for\ all\ x \in \bar{\Omega} = \overline{B_1(0)}, \end{cases} \qquad (3.24)$$

In this case, we define Ω_1, Ω_2, Ω_3 as the subsets of $\bar{\Omega}$ determined by $(r, \theta) \in [0, 1] \times [0, \pi/2]$, $(r, \theta) \in (0, 1] \times (\pi/2, 3\pi/2)$, $(r, \theta) \in (0, 1] \times [3\pi/2, 2\pi)$, respectively. Furthermore we set

$$e_1(x) = \begin{cases} 1 & in\ \Omega_1, \\ 0 & elsewhere, \end{cases} \quad e_2(x) = \begin{cases} 2 & in\ \Omega_2, \\ 0 & elsewhere, \end{cases} \quad e_3(x) = \begin{cases} 10 & in\ \Omega_3, \\ 0 & elsewhere. \end{cases}$$

The associated optimality system is

$$
\begin{cases}
\begin{aligned}
-\Delta\bar{y} + \bar{y} + \bar{y}^3 &= \sum_{i=1}^3 u_i e_i(x) && in && \Omega \\
\partial_\nu y(x) &= 0 && on && \Gamma,
\end{aligned} \\[2mm]
\begin{aligned}
-\Delta\bar{\varphi} + \bar{\varphi} + 3\bar{y}^2\bar{\varphi} &= \bar{y} - y_d - \bar{\mu} && in && \Omega \\
\partial_\nu\bar{\varphi}(x) &= 0 && on && \Gamma,
\end{aligned} \\[2mm]
(\bar{u} - u_d) + \left[\int_\Omega \bar{\varphi} e_1(x)\, dx, \int_\Omega \bar{\varphi} e_2(x)\, dx, \int_\Omega \bar{\varphi} e_3(x)\, dx \right]^\top = 0 \\[2mm]
\int_{\bar{\Omega}} (\bar{y} - 2 + |x|^2) d\bar{\mu} = 0, \quad \bar{y}(x) \geqslant 2 - |x|^2 \quad \forall x \in \bar{\Omega} \\[2mm]
\bar{\mu} \geqslant 0.
\end{cases}
$$

It is not difficult to check that $\bar{y} \equiv 2$ and the optimal control $\bar{u} = [10 \quad 5 \quad 1]^\top$ satisfy the state equation. The active set consists of the single point $x = (0, 0)$. The Lagrange multiplier $\mu = \delta_0$ satisfies the complementarity condition, where δ_0 is the Dirac measure concentrated at the origin. Let us define as adjoint state $\bar{\varphi} = \frac{1}{2\pi} \log|x| - \frac{1}{4\pi}|x|^2$, then we have that $\partial_\nu\bar{\varphi} = 0$ at the boundary,

$$
-\Delta\bar{\varphi} + \bar{\varphi} + 3\bar{y}^2\bar{\varphi} = \frac{1}{\pi} - \delta_0 + 13\left(\frac{1}{2\pi} \log|x| - \frac{1}{4\pi}|x|^2\right),
$$

and it is easy to confirm that the right-hand side of the last identity is equal to $\bar{y} - y_d - \bar{\mu}$, if we define

$$
y_d = 2 - \frac{1}{\pi} - 13\left(\frac{1}{4\pi}|x|^2 - \frac{1}{2\pi}\log|x|\right), \quad u_d = \left[\frac{317}{32}, \frac{37}{8}, \frac{1}{16}\right]^\top.
$$

Therefore, the adjoint equation and the optimality system are satisfied. The second order sufficient conditions are fulfilled, too. We have

$$
\mathcal{L}''(\bar{y}, \bar{u}, \bar{\varphi}, \bar{\mu})[(z_h, h)]^2 = |h|^2 + \int_\Omega z_h^2(1 - 12\bar{\varphi})\, dx \geqslant |h|^2
$$

for all $h \in \mathbb{R}^2$, since $\bar{\varphi} \leqslant 0$.

Example 4. *In this example we consider a problem with some coefficients of the equation as controls.*

$$
(E4) \quad
\begin{cases}
\displaystyle \min_{u \in \mathbb{R}^2} J(y, u) = \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \frac{1}{2}|u - u_d|^2 \\[2mm]
\text{subject to} \\[2mm]
\begin{aligned}
-\Delta y(x) + (u_1 + \pi^2) y(x) + u_2 y(x)^3 &= f(x) && in && \Omega \\
\partial_\nu y(x) &= 0 && on && \Gamma,
\end{aligned} \\[4mm]
y(x) \leqslant b(x) := \begin{cases} (2 - 3x_1)^2 - \frac{1}{4} & x_1 \leqslant \frac{1}{2}, \\ 0 & x_1 \geqslant \frac{1}{2}, \end{cases} \quad \text{for all } x \in \bar{\Omega} = [0, 1] \times [0, 1], \\[4mm]
0 \leqslant u = [u_1, u_2]^\top.
\end{cases}
\qquad (3.25)
$$

We impose nonnegativity of the components of u to have a monotone operator in the state equation. The function f is given by

$$
f(x) = 3\pi^2 \cos(\pi x_1) + \frac{2}{3}\cos(\pi x_1)^3.
$$

The optimality system is

$$
\begin{cases}
\begin{aligned}
-\Delta\bar{y} + (\bar{u}_1 + \pi^2)\bar{y} + \bar{u}_2\bar{y}^3 &= f(x) &\text{in}\quad &\Omega \\
\partial_\nu y(x) &= 0 &\text{on}\quad &\Gamma,
\end{aligned} \\[2mm]
\begin{aligned}
-\Delta\bar{\varphi} + (\bar{u}_1 + \pi^2)\bar{\varphi} + 3\bar{u}_2\bar{\varphi}\bar{y}^2 &= \bar{y} - y_d + \bar{\mu} &\text{in}\quad &\Omega \\
\partial_\nu\bar{\varphi}(x) &= 0 &\text{on}\quad &\Gamma
\end{aligned} \\[2mm]
\left((\bar{u} - u_d) - \begin{bmatrix} \int_\Omega \bar{\varphi}\bar{y}\,dx \\ \int_\Omega \bar{\varphi}\bar{y}^3\,dx \end{bmatrix} \right)^\top (u - \bar{u}) \geqslant 0 \quad \forall 0 \leqslant u \in \mathbb{R}^2, \\[2mm]
\int_{\bar{\Omega}} (\bar{y}(x) - b(x))d\bar{\mu} = 0, \quad \bar{y}(x) \leqslant b(x) \quad \forall x \in \bar{\Omega} \\[2mm]
\text{with } \bar{\mu} \geqslant 0.
\end{cases}
$$

We define

$$
\bar{y}(x) = \cos(\pi x_1), \quad \bar{u} = \begin{bmatrix} \pi^2 \\ \frac{2}{3} \end{bmatrix}, \quad \bar{\varphi}(x) = \begin{cases} \dfrac{x_1^2}{2} - \dfrac{1}{8}, & x_1 \leqslant \frac{1}{2}, \\ 0, & x_1 > \frac{1}{2}, \end{cases}
$$

and the Lagrange multiplier

$$
\bar{\mu} = \frac{1}{2}\delta_{x_1}(1/2).
$$

Since $\bar{y}(x) = b(x)$ at $x_1 = 1/2$, the state is active in the set: $\{(\frac{1}{2}, x_2) : 0 \leqslant x_2 \leqslant 1\}$. Inserting $\bar{\varphi}$ in the adjoint equation, we obtain

$$
\begin{aligned}
\bar{y} - y_d + \bar{\mu} &= -\Delta\bar{\varphi} + \bar{\varphi}(\bar{u}_1 + \pi^2 + 3\bar{u}_2\bar{y}^2) \\
&= \frac{1}{2}\delta_{x_1}(1/2) + \begin{cases} -1 + 2(\pi^2 + \bar{y}^2)\left(\dfrac{x_1^2}{2} - \dfrac{1}{8}\right) & x_1 \leqslant \frac{1}{2}, \\ 0 & x_1 > \frac{1}{2}. \end{cases}
\end{aligned}
$$

Therefore, adjoint equation is satisfied with the choice

$$
y_d = \cos(\pi x_1) + \begin{cases} 1 - \left(x_1^2 - \dfrac{1}{4}\right)\left(\pi^2 + \cos(\pi x_1)^2\right), & x_1 \leqslant \frac{1}{2}, \\ 0 & x_1 > \frac{1}{2}. \end{cases}
$$

The variational inequality holds true with

$$
u_d = \left[\pi^2 + \frac{1}{\pi^3}, \; \frac{2}{3} + \frac{20}{27}\frac{1}{\pi^3}\right]^\top.
$$

To verify the second-order sufficient conditions, we note that $-\frac{1}{8} < \bar{\varphi} \leqslant 0$, and compute $\mathcal{L}''$. This leads to the expression

$$
\mathcal{L}''(\bar{y}, \bar{u}, \bar{\varphi}, \bar{\mu})[(z_h, h)]^2 = \int_\Omega \begin{bmatrix} z_h \\ h_1 \\ h_2 \end{bmatrix}^\top H \begin{bmatrix} z_h \\ h_1 \\ h_2 \end{bmatrix} dx > 0
$$

for all $h = [h_1, h_2]^\top \in \mathbb{R}^2 \backslash \{0\}$, with

$$
H = \begin{bmatrix} 1 - 4\bar{y}\bar{\varphi} & -\bar{\varphi} & -3\bar{y}^2\bar{\varphi} \\ -\bar{\varphi} & 1 & 0 \\ -3\bar{y}^2\bar{\varphi} & 0 & 1 \end{bmatrix}
$$

Positive definiteness of the matrix H holds for all $x \in [0, 1]$, because it is a symmetric diagonal dominant matrix with positive diagonal.

References

[1] S. Agmon. On the eigenfunctions and on the eigenvalues of general elliptic boundary value problems. *Comm. Pure Appl. Math.*, 15:119–147, 1962.

[2] S. Agmon, A. Douglis, L. Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I. *Comm. Pure Appl. Math.*, 12:623–727, 1959. II. *Comm. Pure Appl. Math.*, 17:35–92, 1964.

[3] J.-J. Alibert and J.-P. Raymond. Boundary control of semilinear elliptic equations with discontinuous leading coefficients and unbounded controls. *Numer. Funct. Anal. and Optimization*, 3&4:235–250, 1997.

[4] Amann, H. *Linear and quasilinear parabolic problems*. Basel-Boston-Berlin, Birkhäuser, 1995.

[5] F. Bonnans. Second-order analysis for control constrained optimal control problems of semilinear elliptic systems. *Appl. Math. and Optimization*, 38:303–325, 1998.

[6] F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York, 2000.

[7] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control and Optimization*, 4:1309–1322, 1986.

[8] E. Casas. Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control and Optimization*, 35:1297–1327, 1997.

[9] E. Casas, J.C. de los Reyes, and F. Tröltzsch. Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints. *submitted*, 2007.

[10] E. Casas and M. Mateos. Second order sufficient optimality conditions for semilinear elliptic control problems with finitely many state constraints. *SIAM J. Control and Optimization*, 40:1431–1454, 2002.

[11] E. Casas and F. Tröltzsch. Second order necessary and sufficient optimality conditions for optimization problems and applications to control theory. *SIAM J. Optimization*, 13:406–431, 2002.

[12] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for a nonlinear elliptic control problem. *Z. für Analysis und ihre Anwendungen (ZAA)*, 15:687–707, 1996.

[13] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations. *SIAM J. Control and Optimization*, 38(5):1369–1391, 2000.

[14] Ciarlet, P.G. *The finite element method for elliptic problems*, Studies in Mathematics and its Applications, North Holland, Amsterdam-New York-Oxford, 1979.

[15] P. Deuflhard, M. Seebass, D. Stalling, R. Beck, and H.-C. Hege. Hyperthermia treatment planning in clinical cancer therapy: Modelling, simulation, and visualization. In A. Sydow, editor, *Computational Physics, Chemistry and Biology*, pages 9–17. Wissenschaft und Technik-Verlag, 1997.

[16] K. Eppler and F. Tröltzsch. Fast optimization methods in the selective cooling of steel. In M. Grötschel, S. O. Krumke, and J. Rambau, editors, *Online optimization of large scale systems*, pages 185–204. Springer-Verlag, 2001.

[17] Gajewski, H., Gröger, K., Zacharias, K. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Akademie-Verlag, Berlin, 1974.

[18] H. Goldberg and F. Tröltzsch. Second order optimality conditions for a class of control problems governed by nonlinear integral equations with application to parabolic boundary control. *Optimization*, 20:687–698, 1989.

[19] J. A. Griepentrog, H. C. Kaiser, J. Rehberg. Heat kernel and resolvent properties for second order elliptic differential operators with general boundary conditions on L^p. *Adv. Math. Sci. Appl.*, 11:87–112, 2001.

[20] J. A.Griepentrog. Linear elliptic boundary value problems with non-smooth data: Campanato spaces of functionals. *Math. Nachr.*, 243: 19–42, 2002.

[21] Grisvard, P. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston 1985.

[22] L. Henning and R. King. Drag reduction by closed-loop control of a separated flow over a bluff body with a blunt trailing edge. 44th IEEE Conference on Decision and Control and European Control Conference ECC 2005, Seville, Spain. accepted.

[23] Kato, T. *Perturbation theory for linear operators.* Corr. printing of the 2nd ed., Grundlehren der mathematischen Wissenschaften. Springer, Berlin-Heidelberg-New York.

[24] S. Kurcyusz J. Zowe. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, (5):49–62, 1979.

[25] Ladyzhenskaya, O. A., Solonnikov, V. A, and Ural'ceva, N. N. *Linear and Quasilinear Equations of Parabolic Type.* American Math. Society, Providence, R.I. 1968.

[26] C. Meyer and P. Philip. Optimizing the temperature profile during sublimation growth of sic single crystals: Control of heating power, frequency and coil position. *Crystal Growth & Design*, 5:1145–1156, 2005.

[27] C. Meyer, U. Prüfert, and F. Tröltzsch. On two numerical methods for state-constrained elliptic control problems. Technical report, Institut für Mathematik, Technische Universität Berlin, 2005. Report 5-2005, submitted.

[28] Nečas, J. *Les méthodes directes en théorie des equations elliptiques*. Academia, Prague 1967.

[29] J.-P. Raymond and F. Tröltzsch. Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints. *Discrete and Continuous Dynamical Systems*, 6:431–450, 2000.

[30] Triebel, H. *Interpolation theory, Function Spaces, Differential Operators*. North Holland, Amsterdam-New York-Oxford, 1978.

[31] F. Tröltzsch, R. Lezius, R. Krengel, and H. Wehage. Mathematische Behandlung der optimalen Steuerung von Abkühlungsprozessen bei Profilstählen. In K.H. Hoffmann, W. Jäger, T. Lohmann, and H. Schunck, editors, *Mathematik– Schlüsseltechnologie für die Zukunft, Verbundprojekte zwischen Universität und Industrie*, pages 513–524. Springer-Verlag, 1997.

[32] F. Tröltzsch and A. Unger. Fast solution of optimal control problems in the selective cooling of steel. *ZAMM*, 81:447–456, 2001.

[33] Tröltzsch, F. *Optimale Steuerung partieller Differentialgleichungen – Theorie, Verfahren und Anwendungen*. Vieweg, 2005.

[‡]Department of Mathematics, EPN Quito, Ecuador

[⋆]Weierstrass-Institut für Angewandte Analysis und Stochastik (WIAS), Berlin, Germany

[†]Institut für Mathematik, TU Berlin, Germany