TECHNISCHE UNIVERSITÄT CHEMNITZ

Second Order Sufficient Optimality Conditions for some State–Constrained Control Problems of Semilinear Elliptic Equations

E. Casas, F. Tröltzsch, A. Unger

Preprint 97-19

Fakultät für Mathematik

SECOND ORDER SUFFICIENT OPTIMALITY CONDITIONS FOR SOME STATE–CONSTRAINED CONTROL PROBLEMS OF SEMILINEAR ELLIPTIC EQUATIONS*

EDUARDO CASAS†, FREDI TRÖLTZSCH‡, AND ANDREAS UNGER ‡

Abstract. In this paper we deal with a class of optimal control problems governed by elliptic equations with nonlinear boundary condition. The case of a boundary control is studied. We consider pointwise constraints on the control and certain equality and set–constraints on the state. Second order sufficient conditions for local optimality of controls are derived.

Key words. Boundary control, semilinear elliptic equations, sufficient optimality conditions, state constraints

AMS subject classifications. 49K20, 35J25

1. Introduction. In contrast to the optimal control of linear systems with convex objective, where first order necessary optimality conditions are already sufficient for optimality, higher order optimality conditions have to be used for nonlinear systems to verify optimality. Second order sufficient optimality conditions (SSC) are very useful to show certain important properties of optimal control problems, for instance local uniqueness of optimal controls and their stability with respect to perturbations of the problem. Moreover, they serve as important assumptions to guarantee the convergence of numerical methods in optimal control. We refer to the general expositions by Maurer and Zowe [14] and Maurer [13] concerned with different aspects of (SSC), to Alt [2], where the approximation of programming problems in Banach spaces is discussed, and to Alt [3], [4], containing a convergence analysis for Lagrange–Newton methods in Banach spaces.

Meanwhile, an extensive number of publications appeared discussing several aspects of (SSC) for control problems governed by ordinary differential equations. In this respect, the well known *two-norm discrepancy* led to new difficulties and interesting solutions. We refer for instance to Ioffe [12] and Maurer [13].

First considerations of (SSC) for control problems governed by partial differential equations were published by Goldberg and Tröltzsch [10], [11] for the boundary control of parabolic equations with nonlinear boundary conditions.

In the paper by Casas, Tröltzsch and Unger [9] the authors have extended these ideas to elliptic boundary control problems with pointwise constraints on the control. Moreover, the gap between second order necessary and sufficient optimality conditions was tightened taking into account first order sufficient optimality conditions (as introduced by Maurer and Zowe [14]). It should be mentioned that in this case already four norms have to be used (L^∞–norm for differentiation, L^2–norm to formulate (SSC), L^1–norm for the first order sufficient optimality condition, and a certain L^p–norm to obtain optimal regularity results).

* This research was partially supported by European Union, under HCM Project number ERBCHRXCT940471, and by Deutsche Forschungsgemeinschaft, under Project number Tr 302/1-2. The first author was also supported by Dirección General de Investigación Científica y Técnica (Spain)

† Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, 39071 Santander, Spain.

‡ Fakultät für Mathematik, Technische Universität Chemnitz-Zwickau, D-09107 Chemnitz, Germany

F.Bonnans [5] shows for a particular class of semilinear elliptic control problems with constraints on the control that a very weak form of second order sufficient conditions can be used to verify local optimality: If the second order derivative of the Lagrange function is a Legendre form, then it suffices to have its positivity in all critical directions.

In our paper, we extend the results of [9] allowing for additional constraints on the state. In this way, we continue the investigations of Casas and Tröltzsch [8] on second order necessary conditions. We rely also on general ideas of Maurer and Zowe [14] and refine them by means of a detailed splitting technique.

Incidentally, we aimed to establish second order sufficient optimality conditions for boundary control problems governed by semilinear elliptic equations in domains of arbitrary dimension with pointwise constraints on the control and the state. However, it turned out that *pointwise* state–constraints lead to essential and somehow surprising difficulties. To establish second order sufficient optimality conditions for problems with pointwise state–constraints given on the whole domain, we had to restrict ourselves to a 2–dimensional domain and controls appearing linearly in the boundary condition. These obstacles seem to indicate some limits for the "traditional" type of (SSC) for control problems governed by PDEs.

If pointwise state–constraints are imposed on compact subsets of the domain and the other quantities are sufficiently smooth, then arbitrary dimensions can be treated without restrictions on the nonlinearities. (In this case the adjoint state belongs to $L^\infty(\Gamma)$.) Moreover, we are able to get rid of the assumption of linearity of the boundary condition with respect to the control by introducing some extended form of second order optimality conditions.

2. The optimal control problem. We consider the problem to *minimize* the functional

$$F_0(y,u) = \int_\Omega f(x,y(x))\,dx + \int_\Gamma g(x,y(x),u(x))\,dS(x) \tag{2.1}$$

subject to the *equation of state*

$$\begin{cases} -\Delta y(x) + y(x) = 0 & \text{in } \Omega \\ \partial_\nu y(x) = b(x,y(x),u(x)) & \text{on } \Gamma, \end{cases} \tag{2.2}$$

to the constraints on the *state* y

$$F_i(y) = 0, \quad i = 1,\dots,m, \tag{2.3}$$

$$E(y) \in K, \tag{2.4}$$

and to the constraints on the *control* u

$$u_a(x) \le u(x) \le u_b(x) \quad \text{a. e. on } \Gamma. \tag{2.5}$$

In this setting, $\Omega \subset \mathbb{R}^n$ is a bounded domain (i. e. simply connected and open) with a Lipschitz boundary Γ according to the definition by Nečas [16]. By $f : \Omega \times \mathbb{R} \to \mathbb{R}$ and $g, b : \Gamma \times \mathbb{R}^2 \to \mathbb{R}$ sufficiently smooth functions are given, ∂_ν is the derivative in the direction of the unit outward normal ν on Γ. The functionals $F_i : C(\overline{\Omega}) \to \mathbb{R}$, $i = 1,\dots,m$, are supposed to be twice continuously Fréchet differentiable, i. e. being of class C^2. By E we denote a mapping of class C^2 from $C(\overline{\Omega})$ into a real Banach space Z. $K \subset Z$ is a non-empty convex closed set. u_a, $u_b : \Gamma \to \mathbb{R}$ are functions of $L^\infty(\Gamma)$ satisfying $u_a(x) \le u_b(x)$ on Γ.

The *control* u is looked upon the control space $\mathcal{U} = L^\infty(\Gamma)$, while the *state* y is defined as *weak solution* of (2.2) in the state space $C(\overline{\Omega}) \cap H^1(\Omega) = Y$, i. e.

$$\int_\Omega (\nabla y \nabla v + yv)\, dx = \int_\Gamma b(\cdot, y, u) v\, dS \quad \forall v \in H^1(\Omega). \tag{2.6}$$

We endow Y with the norm $\|y\|_Y = \|y\|_{C(\overline{\Omega})} + \|y\|_{H^1(\Omega)}$. The following assumptions are imposed on the given quantities:

(A1) For each fixed $x \in \Omega$ or Γ, respectively, the functions $f = f(x,y)$, $g = g(x,y,u)$, and $b = b(x,y,u)$ are of class C^2 with respect to (y,u). For (y,u) fixed, they are Lebesgue measurable with respect to $x \in \Omega$ or $x \in \Gamma$, respectively.

In the paper, partial derivatives are indicated in the usual way by subscripts. For instance, b_{yu} stands for $\partial^2 b/\partial y \partial u$. By $b'(x,y,u)$ and $b''(x,y,u)$ we denote the gradient and the Hessian matrix of b with respect to (y,u):

$$b'(x,y,u) = \begin{pmatrix} b_y(x,y,u) \\ b_u(x,y,u) \end{pmatrix}, \quad b''(x,y,u) = \begin{pmatrix} b_{yy}(x,y,u) & b_{yu}(x,y,u) \\ b_{uy}(x,y,u) & b_{uu}(x,y,u) \end{pmatrix},$$

$|b'|$ and $|b''|$ are defined by adding the absolute values of all entries.

In the next assumption **(A2)**, fixed parameters $p > n-1$ and s, r are used, which depend on n. For the possible (maximal) choices of s and r we refer to the discussion of regularity in (3.13). Roughly speaking, we have for the linearized system (2.2) that $y|_\Gamma \in L^s(\Gamma)$ and $y \in L^r(\Omega)$, if $u \in L^2(\Gamma)$. s' and r' are defined as conjugate numbers, for instance, $1/s' + 1/s = 1$.

(A2) For all $M > 0$ there are constants $C_M > 0$ and functions $\Psi_f^M \in L^{(r/2)'}(\Omega)$, $\Psi_g^{M,1} \in L^{(s/2)'}(\Gamma)$, $\Psi_g^{M,2} \in L^{2(s/2)'}(\Gamma)$, $\Psi_g^{M,3} \in L^\infty(\Gamma)$, and a continuous, monotone increasing function $\eta \in C(\mathbb{R}^+ \cup \{0\})$ with $\eta(0) = 0$ such that:

(i)

$$b_y(x,y,u) \le 0 \qquad \text{a. e. } x \in \Gamma, \forall (y,u) \in \mathbb{R}^2, \tag{2.7}$$

$b(\cdot, 0) \in L^p(\Gamma)$, for a $p > n-1$,
$|b'(x,y,u)| + |b''(x,y,u)| \le C_M$,
$|b''(x,y_1,u_1) - b''(x,y_2,u_2)| \le C_M\, \eta(|y_1 - y_2| + |u_1 - u_2|)$
for almost all $x \in \Gamma$ and all $|y|, |u|, |y_i|, |u_i| \le M$, $i = 1,2$.

(ii) $f(\cdot,0) \in L^1(\Omega)$, $f_y(\cdot,0) \in L^{r'}(\Omega)$, $f_{yy}(\cdot,0) \in L^{(r/2)'}(\Omega)$
$|f_{yy}(x,y_1) - f_{yy}(x,y_2)| \le \Psi_f^M(x)\, \eta(|y_1 - y_2|)$
for all $x \in \Omega$, $|y_i| \le M$, $i = 1,2$.

(iii) $g(\cdot,0) \in L^1(\Gamma)$, $g_y(\cdot,0) \in L^{s'}(\Gamma)$, $g_u(\cdot,0) \in L^2(\Gamma)$,
$g_{yy}(\cdot,0) \in L^{(s/2)'}(\Gamma)$, $g_{yu}(\cdot,0) \in L^{2(s/2)'}(\Gamma)$, $g_{uu}(\cdot,0) \in L^\infty(\Gamma)$
(here, $\cdot$ stands for x, 0 for $0 \in \mathbb{R}^2$)
$|g_{yy}(x,y_1,u_1) - g_{yy}(x,y_2,u_2)| \le \Psi_g^{M,1}(x)\eta(|y_1 - y_2| + |u_1 - u_2|)$
$|g_{yu}(x,y_1,u_1) - g_{yu}(x,y_2,u_2)| \le \Psi_g^{M,2}(x)\eta(|y_1 - y_2| + |u_1 - u_2|)$
$|g_{uu}(x,y_1,u_1) - g_{uu}(x,y_2,u_2)| \le \Psi_g^{M,3}(x)\eta(|y_1 - y_2| + |u_1 - u_2|)$
for almost all $x \in \Gamma$ and all $|y_i| \le M$, $|u_i| \le M$.

REMARK 2.1. *Note that the estimates of (i)–(iii) imply boundedness and Lipschitz properties also for b, f, g, b', f', g' in several L-spaces. We omit them, because they follow by applying the mean value theorem.*

(A3) **(i)** Let us define for $y \in C(\overline{\Omega})$ and a certain measurable compact subset $A \subset \overline{\Omega}$ the norm

$$\|y\|_2 = \|y\|_{C(A)} + \|y\|_{L^r(\Omega)} + \|y\|_{L^s(\Gamma)}.$$

A stands for a subset, where we know $y \in C(A)$ for Neumann boundary data of $L^2(\Gamma)$. In the case $n = 2$, we may take $A = \overline{\Omega}$, for $n > 2$ we need $A \subset \Omega$. We put $\|y\|_{C(A)} = 0$, if $A = \emptyset$.
We require for a fixed reference state $\bar{y} \in C(\overline{\Omega})$:

$$|F_i'(\bar{y})y| \le C_F\|y\|_2 \quad \forall y \in C(\overline{\Omega})$$
$$|F_i''(\bar{y})[y_1, y_2]| \le C_F\|y_1\|_2\|y_2\|_2 \quad \forall y_1, y_2 \in C(\overline{\Omega}),$$

where $C_F > 0$, and

$$|F_i'(y_1)y - F_i'(y_2)y| \le C_M\|y_1 - y_2\|_2\|y\|_2$$
$$|(F_i''(y_1) - F_i''(y_2))[y, v]| \le C_M\, \eta(\|y_1 - y_2\|_{C(\overline{\Omega})})\|y\|_2\|v\|_2$$

for all y_j with $\|y_j\|_{C(\overline{\Omega})} \le M$, $j = 1, 2$, y and v from $C(\overline{\Omega})$ and $i = 1, \ldots, m$.

(ii) Analogous assumptions are imposed on $E : C(\overline{\Omega}) \to Z$, where $\|\cdot\|_Z$ is to be substituted for $|\cdot|$. For instance,

$$\|E'(\bar{y})y\|_Z \le C_E\|y\|_2 \quad \forall y \in C(\overline{\Omega})$$

is supposed.

Let us explain some characteristic examples for the possible choice of objective functionals and constraints:

Objective:

The functions f and g can be taken as

$$f(x, y) = f_0(x) + f_y(x)y + f_{yy}(x)y^2$$
$$g(x, y, u) = g_0(x) + g_y(x)y + g_u(x)u + g_{yy}(x)y^2 + g_{yu}(x)yu + g_{uu}(x)u^2,$$

where $f_0 \in L^1(\Omega)$, $f_y \in L^{r'}(\Omega)$, $f_{yy} \in L^{(r/2)'}(\Omega)$, $g_0 \in L^1(\Gamma)$, $g_y \in L^{s'}(\Gamma)$, $g_u \in L^2(\Gamma)$, $g_{yy} \in L^{(s/2)'}(\Gamma)$, $g_{yu} \in L^{2(s/2)'}(\Gamma)$, and $g_{uu} \in L^\infty(\Gamma)$ are given functions.

State–constraints:

A characteristic set of state–constraints might have the form

$$F_j(y) = \int_\Omega \omega_j(x, y(x))\, dx + \int_\Gamma \gamma_j(x, y(x))\, dS(x) + \int_A \pi_j(x, y(x))\, d\mu_j(x) = 0$$
$$j = 1, \ldots, m$$
$$E_1(y)(x) = e(x, y(x)) \in C \quad \forall x \in A$$
$$E_2(y)(x) = \int_\Omega k_1(x,\xi)\omega_0(\xi, y(\xi))\, dx + \int_\Gamma k_2(x,\xi)\gamma_0(\xi, y(\xi))\, dS(\xi)$$
$$+ \int_A k_3(x,\xi)\pi_0(\xi, y(\xi))\, d\mu_j(\xi) - c(x) \le 0 \quad \forall x \in D$$

with sufficiently smooth functions ω_j, γ_j, π_j, $j = 0, \ldots, m$, continuous functions c, k_i, $i = 1, 2, 3$, a convex closed subset $C \subset \mathbb{R}^l$, a closed set $D \subset \mathbb{R}^k$, and a sufficiently smooth mapping $e : A \times \mathbb{R} \to \mathbb{R}^l$.

It should be mentioned that the choice $A = \overline{\Omega}$ cannot be allowed for $n \geq 3$. However, we may approximate associated measures by functions of L-spaces. This motivates the choice of L^r- and L^s-spaces used above.

3. The state equation, first order necessary optimality conditions. It can be shown that the equation (2.2) admits for each $u \in \mathcal{U}^{ad}$ a unique weak solution $y = y(u) \in Y$, where $\mathcal{U}^{ad} = \{u \in L^\infty(\Gamma) \mid u_a(x) \leq u(x) \leq u_b(x)$ a. e. on $\Gamma\}$. Moreover, there is a constant M such that

$$\|y(u)\|_Y \leq M \quad \forall u \in \mathcal{U}^{ad}, \tag{3.1}$$

in particular $\|y\|_{C(\overline{\Omega})} \leq M$. In Casas and Tröltzsch [8] it was shown that the mapping $u \mapsto y(u)$ is of class C^2 from $L^\infty(\Gamma)$ into Y. Furthermore, there is a constant C_2 such that the Lipschitz property

$$\|y(u_1) - y(u_2)\|_2 \leq C_2 \|u_1 - u_2\|_{L^2(\Gamma)}$$

holds for all u_1, $u_2 \in \mathcal{U}^{ad}$ ($\|\cdot\|_2$ was defined in **(A3)**). For fixed $u \in \mathcal{U}^{ad}$ we have $b(\cdot, y, u) \in L^p(\Gamma)$, hence the weak solution $y \in Y$ of (2.2) belongs to the space

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

which is known to be continuously embedded into $Y = C(\overline{\Omega}) \cap H^1(\Omega)$ for every $q > n/2$ and every $p > n - 1$.

In all what follows we assume that a *reference pair* $(\overline{y}, \overline{u}) \in Y \times \mathcal{U}^{ad}$ is given, satisfying together with an associated *adjoint state* $\overline{\varphi} \in W^{1,\sigma}(\Omega)$, $\forall \sigma < n/(n-1)$, and *Lagrange multipliers*

$$\overline{\lambda} = (\overline{\lambda}_1, \ldots, \overline{\lambda}_m)^T \in \mathbb{R}^m, \; \overline{z}^* \in Z^*$$

the associated standard *first order necessary optimality condition.*

Let us state them for convenience. They can be proved following Casas [7], the general considerations in Casas and Bonnans [6] and in Zowe and Kurcyusz [21]. We just assume them. The first order optimality system to be satisfied by $(\overline{y}, \overline{u})$ consists of the state equations (2.2), the constraint $\overline{u} \in \mathcal{U}^{ad}$, the *adjoint equation*

$$-\Delta \overline{\varphi} + \overline{\varphi} = f_y(\cdot, \overline{y}) + \sum_{i=1}^m \overline{\lambda}_i F_i'(\overline{y})|_\Omega + (E'\overline{y})^* \overline{z}^*|_\Omega \qquad \text{in } \Omega \tag{3.2}$$

$$\partial_\nu \overline{\varphi} = b_y(\cdot, \overline{y}, \overline{u}) \overline{\varphi} + g_y(\cdot, \overline{y}, \overline{u}) + \sum_{i=1}^m \overline{\lambda}_i F_i'(\overline{y})|_\Gamma + (E'\overline{y})^* \overline{z}^*|_\Gamma \quad \text{on } \Gamma \tag{3.3}$$

for the adjoint state $\overline{\varphi}$, the *complementary slackness condition*

$$\langle \overline{z}^*, \kappa - E(\overline{y}) \rangle \leq 0 \quad \forall \kappa \in K, \tag{3.4}$$

and the *variational inequality*

$$\int_\Gamma (g_u(x, \overline{y}(x), \overline{u}(x)) + \overline{\varphi}(x) b_u(x, \overline{y}(x), \overline{u}(x))(u(x) - \overline{u}(x))\, dS(x) \geq 0 \tag{3.5}$$

for all $u \in \mathcal{U}^{ad}$. We have $F_i'(\bar y) \in C(\bar\Omega)^*$, $i = 1, \dots, m$, and $E'(\bar y)^* \bar z^* \in C(\bar\Omega)^*$, hence these quantities can be identified with real Borel measures on $\bar\Omega$. Owing to Casas [7], the problem

$$\begin{cases} -\Delta\varphi + \varphi = \mu_\Omega & \text{in } \Omega \\ \partial_\nu\varphi + \beta\varphi = \mu_\Gamma & \text{on } \Gamma \end{cases} \tag{3.6}$$

admits for given real Borel measures μ_Ω and μ_Γ concentrated on Ω and Γ, respectively, a unique solution $\varphi \in W^{1,\sigma}(\Omega)$ for all $\sigma < n/(n-1)$, if $\beta \in L^\infty(\Gamma)$ is nonnegative. This justifies to write

$$\bar\varphi = \varphi_0 + \sum_{i=1}^m \lambda_i \varphi_i + \varphi_E,$$

where φ_0, φ_i, and φ_E solve (3.6) for $\mu_\Omega = f_y$, $F_i'(\bar y)|_\Omega$, $E'(\bar y)^*\bar z^*|_\Omega$ and $\mu_\Gamma = g_y$, $F_i'(\bar y)|_\Gamma$, $E'(\bar y)^*\bar z^*|_\Gamma$, respectively. We have at least φ_0, φ_i and φ_E in $W^{1,\sigma}(\Omega)$. Moreover, $\bar\varphi$ satisfies the formula of integration by parts

$$\int_\Omega (-\Delta y + y)\varphi\, dx + \int_\Gamma (\partial_\nu y + \beta y)\varphi\, dS(x) = \int_\Omega y\, d\mu_\Omega + \int_\Gamma y\, d\mu_\Gamma \tag{3.7}$$

for all $y \in Y_{q,p}$, where $q > n/2$, $p > n-1$. Therefore, it is easy to check that the optimality conditions can be expressed by means of the *Lagrange function*

$$\begin{aligned} \mathcal{L}(y, u, \varphi, \lambda, z^*) = F_0(y, u) &- \int_\Omega (-\Delta y + y)\,\varphi\, dx - \int_\Gamma (\partial_\nu y - b(\cdot, y, u))\varphi\, dS \\ &+ \sum_{j=1}^m \lambda_j F_j(y) + \langle z^*, E(y)\rangle, \end{aligned} \tag{3.8}$$

$\mathcal{L} : Y_{q,p} \times \mathcal{U} \times W^{1,\sigma}(\Omega) \times \mathbb{R}^m \times Z^* \to \mathbb{R}$. The regularity of y and φ fit together, as $\varphi \in W^{1,\sigma}(\Omega)$ for all $\sigma < n/(n-1)$ ensures $\varphi \in L^s(\Omega)$ for all $s < n/(n-2)$ (cf. Nečas [16], Thm. 3.4, p. 69) and $\varphi|_\Gamma \in L^r(\Gamma)$ for all $r < 1 + 1/(n-2)$ ([16], Thm. 4.2, p.84). Hence this definition makes sense. In (3.8), $\langle\cdot,\cdot\rangle$ denotes the duality pairing between Z and its dual space Z^*. It is obvious that $\mathcal{L}$ is of class C^2 with respect to (y, u) for fixed φ, λ, and z^*.

According to (3.7), the optimality system can be written in the form (2.6), $u \in \mathcal{U}^{ad}$, and

$$\mathcal{L}_y(\bar y, \bar u, \bar\varphi, \bar\lambda, \bar z^*)y = 0 \quad \forall y \in Y \tag{3.9}$$

$$\mathcal{L}_u(\bar y, \bar u, \bar\varphi, \bar\lambda, \bar z^*)(u - \bar u) \geq 0 \quad \forall u \in \mathcal{U}^{ad} \tag{3.10}$$

$$\langle \bar z^*, \kappa - E(\bar y)\rangle \leq 0 \quad \forall \kappa \in K, \tag{3.11}$$

which is more convenient for our later computations. In order to simplify our notation, derivatives taken at $(\bar y, \bar u, \bar\varphi, \bar\lambda, \bar z^*)$ will be indicated by a bar. For instance, $\bar{\mathcal{L}}_y y$, $\bar{\mathcal{L}}_u(u - \bar u)$ would stand for the derivatives in (3.9) and (3.10), respectively. $\bar{\mathcal{L}}_{yy}[y_1, y_2]$ denotes the second derivative of $\mathcal{L}$ in the directions y_1, y_2 taken at $(\bar y, \bar u, \bar\varphi, \bar\lambda, \bar z^*)$. Moreover, $\bar{\mathcal{L}}_{ww}[w_1, w_2]$ is the second order derivative of $\mathcal{L}$ in the directions $w_1 = (y_1, u_1)$, $w_2 = (y_2, u_2)$. If $w_1 = w_2 = w$, we write for short $\bar{\mathcal{L}}_{ww}[w, w] = \bar{\mathcal{L}}_{ww}[w]^2$.

Next we state some useful results on certain linearized versions of the state equation. Regard first the linear system

$$\text{(3.12)} \qquad \begin{cases} -\Delta y + y = f & \text{in } \Omega \\ \partial_\nu y + \beta y = g & \text{on } \Gamma, \end{cases}$$

where $\beta \in L^\infty(\Gamma)$ is nonnegative. Owing to Casas [7], this system admits for each pair $(f, g) \in L^1(\Omega) \times L^1(\Gamma)$ a unique solution $y \in W^{1,\sigma}(\Omega)$ with $\sigma < n/(n-1)$. (Note that a function of L^1 can be considered as a Borel measure.) On the other hand, we have that the solution y of (3.12) belongs to $H^1(\Omega) \cap C(\overline{\Omega})$, if $(f, g) \in L^q(\Omega) \times L^p(\Gamma)$. This regularity result is well known for domains with C^1–boundary. Nevertheless, it holds also true for domains with Lipschitz boundary in the sense of Nečas [16] (cf. Stampacchia [17] and Murthy and Stampacchia [15]). In view of these results we see that the mapping $T : (f, g) \mapsto (y, y|_\Gamma)$ is a mapping from $L^1(\Omega) \times L^1(\Gamma)$ into $L^s(\Omega) \times L^t(\Gamma)$ with $s < n/(n-2)$ and $t < (n-1)/(n-2)$ by embedding theorems for $W^{1,\sigma}(\Omega)$ (cf. again Nečas [16]) and from $L^q(\Omega) \times L^p(\Gamma)$ into $L^\infty(\Omega) \times L^\infty(\Gamma)$. In both cases, the mapping is linear and continuous. Interpolation theory applies to show the following results for T viewed as a mapping defined on $L^2(\Omega) \times L^2(\Gamma)$:

$$\text{(3.13)} \quad y \in \begin{cases} C(\overline{\Omega}), & n = 2 \\ L^r(\Omega)\ \forall r < \infty, & n = 3 \\ L^r(\Omega)\ \forall r < \dfrac{2n}{n-3} & n \geq 4 \end{cases} \qquad y|_\Gamma \in \begin{cases} C(\Gamma) & n = 2 \\ L^s(\Gamma)\ \forall s < \infty & n = 3 \\ L^s(\Gamma)\ \forall s < \dfrac{2(n-1)}{n-3} & n \geq 4. \end{cases}$$

4. Regularity condition and linearization theorem. Recall that we consider a fixed reference pair $(\bar{y}, \bar{u})$ satisfying together with $(\bar{\varphi}, \bar{\lambda}, \bar{z}^*)$ the first order necessary conditions (3.9) - (3.11).

The *linearized cone* of $\mathcal{U}^{ad}$ at $\bar{u}$ is the set $\mathcal{C}(\bar{u}) = \{v \in L^\infty(\Gamma) \mid v = \varrho(u - \bar{u}),\ \varrho \geq 0,\ u \in \mathcal{U}^{ad}\}$. Let $F = F(y)$ denote the mapping $y \mapsto (F_1(y), \ldots, F_m(y))^T$ from Y to $\mathbb{R}^m$. For convenience, we introduce the set of all feasible pairs

$$\mathcal{M} = \{w = (y, u) \in Y \times \mathcal{U}^{ad} \mid y = G(u) \text{ and } y \text{ satisfies the state–constraints}\}$$

(note that G is the nonlinear control-state-mapping). Following Maurer and Zowe [14], the *linearized cone* $L(\mathcal{M}, \bar{w})$ at $\bar{w} = (\bar{y}, \bar{u})$ is defined by

$$L(\mathcal{M}, \bar{w}) = \{w = (y, u) \mid u \in \mathcal{C}(\bar{u}) \text{ and } (y, u) \text{ satisfies (4.1) –(4.3)}\},$$

where

$$\text{(4.1)} \qquad \begin{cases} -\Delta y + y &= 0 & \text{in } \Omega \\ \partial_\nu y &= b_y(\cdot, \bar{y}, \bar{u}) y + b_u(\cdot, \bar{y}, \bar{u}) u & \text{on } \Gamma \end{cases}$$

$$\text{(4.2)} \qquad F'(\bar{y}) y = 0$$

$$\text{(4.3)} \qquad E'(\bar{y}) y \in K(E(\bar{y})).$$

Here, $K(E(\bar{y})) = \{z \in Z \mid z = \varrho(\kappa - E(\bar{y})),\ \varrho \geq 0,\ \kappa \in K\}$ is the conical hull of $K - E(\bar{y})$.

REMARK 4.1. *The following choice of* $E : Y \to Z$ *is of particular interest:* $Z = \mathbb{R}^k$, $E(y) = (E_1(y), \ldots, E_k(y))^T$, $K = (\mathbb{R}^k)^-$. *Then (4.3) reduces to*

$$E_i'(\bar{y}) y \leq 0$$

for all active $i \in \{1, \ldots, k\}$, *i. e. for all* i, *where* $E_i(\bar{y}) = 0$.

The following *regularity assumption* **(R)** is essential for our analysis: For convenience we join the two constraints to one general constraint. To do so we put $\boldsymbol{Z} = \mathbb{R}^m \times Z$, $\boldsymbol{K} = \{0\} \times K$, define $T : Y \to \boldsymbol{Z}$ by $T(y) = (F(y), E(y))$ and $\boldsymbol{K}(T(\bar{y})) = \{0\} \times K(E(\bar{y}))$. Our regularity condition goes back to Zowe and Kurcyusz [21] and is

$$\textbf{(R)} \qquad T'(\bar{y})G'(\bar{u})\mathcal{C}(\bar{u}) - \boldsymbol{K}(T(\bar{y})) = \boldsymbol{Z}.$$

This condition is sufficient for the existence of a (non-degenerate) Lagrange multiplier associated to the state-constraint $E(y) \in K$ (cf. [21]). We should underline that **(R)** does not rely on the condition $\operatorname{int} K \neq \emptyset$. In the Appendix 7.1 we shall give some sufficient conditions for **(R)** to hold (which, however, require $\operatorname{int} K \neq \emptyset$). For $Z = \mathbb{R}^k$, $K = (\mathbb{R}^k)^-$, the condition **(R)** is equivalent to the well-known *Mangasarian–Fromowitz condition.*

THEOREM 4.2. *Suppose that* **(R)** *is satisfied. Then for all pairs* $(\hat{y}, \hat{u}) \in \mathcal{M}$ *there is a pair* $(y, u) \in L(\mathcal{M}, \overline{w})$ *such that the difference* $r = (r^y, r^u) = (\hat{y}, \hat{u}) - (\bar{y}, \bar{u}) - (y, u)$ *fulfils the following estimates:*

$$\|r\|_{Y \times L^\infty(\Gamma)} \le C_{L,p} \|\hat{u} - \bar{u}\|_{L^\infty(\Gamma)} \|\hat{u} - \bar{u}\|_{L^p(\Gamma)} \quad \forall p > n - 1 \tag{4.4}$$

$$\|r\| \le C_{L,2} \|\hat{u} - \bar{u}\|_{L^\infty(\Gamma)} \|\hat{u} - \bar{u}\|_{L^2(\Gamma)}, \tag{4.5}$$

where $\|r\| = \|r^y\|_2 + \|r^u\|_{L^2(\Gamma)}$. *If* $b(x, y, u) = b_1(x, y) + b_2(x)u$, *then*

$$\|r\|_{Y \times L^\infty(\Gamma)} \le C_{L,p} \|\hat{u} - \bar{u}\|^2_{L^p(\Gamma)} \quad \forall p > n - 1. \tag{4.6}$$

This theorem is proved in Appendix 7.2.

We conclude this section with some useful estimates for $\mathcal{L}''$ and certain remainder terms. First, we derive the expression for

$$\overline{\mathcal{L}}''[(y_1, u_1), (y_2, u_2)] = \mathcal{L}''(\bar{y}, \bar{u}, \overline{\varphi}, \overline{\lambda}, \overline{z}^*)[(y_1, u_1), (y_2, u_2)],$$

where $\mathcal{L}''$ denotes the second order derivative of $\mathcal{L}$ with respect to (y, u). We have

$$\begin{aligned} \overline{\mathcal{L}}''[(y_1, u_1), (y_2, u_2)] = & \int_\Omega f_{yy}(\cdot, \bar{y}) y_1 y_2 \, dx + \int_\Gamma (y_1, u_1) g''(\cdot, \bar{y}, \bar{u})(y_2, u_2)^T \, dS \\ & + \int_\Gamma \overline{\varphi} \cdot (y_1, u_1) b''(\cdot, \bar{y}, \bar{u})(y_2, u_2)^T \, dS \\ & + \sum_{i=1}^m \overline{\lambda}_i F_i''(\bar{y})[y_1, y_2] + \langle \overline{z}^*, E''(\bar{y})[y_1, y_2] \rangle. \end{aligned} \tag{4.7}$$

It is the term connected with $\overline{\varphi}$, which causes troubles, more precisely,

$$I = \int_\Gamma \overline{\varphi} \left(b_{yy}(\cdot, \bar{y}, \bar{u}) y_1 y_2 + b_{yu}(\cdot, \bar{y}, \bar{u})(y_1 u_2 + y_2 u_1) + b_{uu}(\cdot, \bar{y}, \bar{u}) u_1 u_2 \right) dS. \tag{4.8}$$

We shall need an estimate of I with respect to the norm $\|y\|_2 + \|u\|_{L^2(\Gamma)}$ (cf. (4.17)). This would require at least $\overline{\varphi} \in L^2(\Gamma)$ in the second item and $\overline{\varphi} \in L^\infty(\Gamma)$ in the third one. Without additional assumption only $\overline{\varphi} \in L^r(\Gamma)$ with $r < (n-1)/(n-2)$ follows from $\overline{\varphi} \in W^{1,\sigma}(\Omega)$, cf. Nečas [16], p. 84. This means in particular $\overline{\varphi} \in L^r(\Gamma)$ for all $r < \infty$, if $n = 2$, $\overline{\varphi} \in L^r(\Gamma)$, for $r < 2$, if $n = 3$. Therefore the following additional assumption is crucial for our analysis:

(A4) Let one of the following statements be true:

(i) $\overline{\varphi} \in L^\infty(\Gamma)$.

(ii) $b_{uu}(x,y,u) = 0$ on $\Gamma \times \mathbb{R}^2$ and, if $n \geq 3$, then $\overline{\varphi} \in L^r(\Gamma)$ for some $r > n-1$.

(iii) $b_{uu}(x,y,u) = b_{yu}(x,y,u) = 0$ on $\Gamma \times \mathbb{R}^2$ and, if $n \geq 4$, then $\overline{\varphi} \in L^r(\Gamma)$ for some $r > (n-1)/2$.

(iv) $b''(\cdot,y,u) = 0$.

Let us briefly comment on the consequences of these assumptions:
(i) holds true, if $\overline{f}_y \in L^q(\Omega)$, $\overline{g}_y \in L^p(\Gamma)$, and the restrictions of F_i', $i = 1,\dots,m$, and $E'(\overline{y})^*\overline{z}^*$ to Ω and Γ belong to $L^q(\Omega)$, $L^p(\Gamma)$, too. Moreover, this holds for functionals F_i', $i = 1,\dots,m$, and $E'(\overline{y})^*\overline{z}^*$ of $C(\overline{\Omega})^*$, where the associated real Borel measures are concentrated on the set $A \subset \Omega$.
(ii) requires linearity of b with respect to u, i. e. $b(x,y,u) = b_0(x,y) + b_1(x,y)u$.
(iii) means that $b(x,y,u) = b_1(x,y) + b_2(x)u$, while
(iv) is only true for an affine–linear boundary condition (however, still for a nonlinear functional F_0).

As a consequence of **(A3)** and **(A4)**, *pointwise state-constraints* on the whole set $\overline{\Omega}$ can only be handled by the standard part of our theory, if u appears linearly in the boundary condition and $n = 2$. In the considerations below we denote by r_i^T the remainder terms of ith order of the Taylor expansion of a mapping T. In this way, the following first and second order expansions of $b(x,y,u)$ will be used for triplets (x,y,u) and $(x,\overline{y},\overline{u}) \in \mathbb{R}^3$:

$$b(x,y,u) - b(x,\overline{y},\overline{u}) = b'(x,\overline{y},\overline{u})(y-\overline{y},u-\overline{u}) + r_1^b, \tag{4.9}$$

where

$$r_1^b = (b_y^\vartheta - \overline{b}_y)(y-\overline{y}) + (b_u^\vartheta - \overline{b}_u)(u-\overline{u}), \tag{4.10}$$

and $b_y^\vartheta, b_u^\vartheta, \overline{b}_y, \overline{b}_u$ denote b_y, b_u taken at $(x, \overline{y}+\vartheta(y-\overline{y}), \overline{u}+\vartheta(u-\overline{u}))$ and $(x,\overline{y},\overline{u})$, respectively, with some $\vartheta \in (0,1)$. Analogously,

$$\begin{aligned} b(x,y,u) - b(x,\overline{y},\overline{u}) &= b'(x,\overline{y},\overline{u})(y-\overline{y},u-\overline{u}) \\ &\quad + \frac{1}{2}(y-\overline{y},u-\overline{u})b''(x,\overline{y},\overline{u})\begin{pmatrix} y-\overline{y} \\ u-\overline{u}\end{pmatrix} + r_2^b, \end{aligned} \tag{4.11}$$

where

$$r_2^b = \frac{1}{2}(y-\overline{y},u-\overline{u})[b'',^\vartheta - \overline{b}''](y-\overline{y},u-\overline{u})^T \tag{4.12}$$

and $b'',^\vartheta$, $\overline{b}''$ denote the Hessian matrix of b with respect to (y,u) taken at the same triplets as above. According to our assumptions on b' and b'', the estimates

$$|r_1^b| \leq C_M(|y-\overline{y}|^2 + |u-\overline{u}|^2) \tag{4.13}$$

$$|r_2^b| \leq C_M\,\eta(|y-\overline{y}| + |u-\overline{u}|)(|y-\overline{y}|^2 + |u-\overline{u}|^2) \tag{4.14}$$

are valid for all $|y|, |\overline{y}|, |u|, |\overline{u}| \leq M$. We continue with the discussion of the remainders $r_1^{\mathcal{L}}$ and $r_2^{\mathcal{L}}$. It holds

$$\begin{aligned} &\mathcal{L}(y,u,\overline{\varphi},\overline{\lambda},\overline{z}^*) - \mathcal{L}(\overline{y},\overline{u},\overline{\varphi},\overline{\lambda},\overline{z}^*) \\ &= \overline{\mathcal{L}}_y(y-\overline{y}) + \overline{\mathcal{L}}_u(u-\overline{u}) + r_1^{\mathcal{L}} \\ &= \overline{\mathcal{L}}_y(y-\overline{y}) + \overline{\mathcal{L}}_u(u-\overline{u}) + \frac{1}{2}\left(\overline{\mathcal{L}}_{yy}[y-\overline{y}]^2 + 2\overline{\mathcal{L}}_{yu}[y-\overline{y},u-\overline{u}] + \overline{\mathcal{L}}_{uu}[u-\overline{u}]^2\right) + r_2^{\mathcal{L}}, \end{aligned}$$

where $\overline{\mathcal{L}}$ indicates that $\mathcal{L}$ and its derivatives are taken at $(\overline{y}, \overline{u}, \overline{\varphi}, \overline{\lambda}, \overline{z}^*)$. We have

$$r_1^{\mathcal{L}} = (\mathcal{L}_y^{\vartheta} - \overline{\mathcal{L}}_y)(y - \overline{y}) + (\mathcal{L}_u^{\vartheta} - \overline{\mathcal{L}}_u)(u - \overline{u})$$

$$r_2^{\mathcal{L}} = \frac{1}{2}((\mathcal{L}_{yy}^{\vartheta} - \overline{\mathcal{L}}_{yy})[y - \overline{y}]^2 + 2(\mathcal{L}_{yu}^{\vartheta} - \overline{\mathcal{L}}_{yu})[y - \overline{y}, u - \overline{u}] + (\mathcal{L}_{uu}^{\vartheta} - \overline{\mathcal{L}}_{uu})[u - \overline{u}]^2).$$

$\mathcal{L}^{\vartheta}$ indicates that $(\overline{y} + \vartheta(y - \overline{y}), \overline{u} + \vartheta(u - \overline{u}), \overline{\varphi}, \overline{\lambda}, \overline{z}^*)$ is inserted for $(y, u, \overline{\varphi}, \overline{\lambda}, \overline{z}^*)$ in $\mathcal{L}'$ and $\mathcal{L}''$, where $\vartheta \in (0, 1)$. Relying on the assumptions **(A1)**–**(A4)** we are able to verify

$$|r_1^{\mathcal{L}}| \le C_{\mathcal{L}}(\|y - \overline{y}\|_2^2 + \|u - \overline{u}\|_{L^2(\Gamma)}^2) \tag{4.15}$$

$$|r_2^{\mathcal{L}}| \le C_{\mathcal{L}}\, \eta(\|y - \overline{y}\|_{C(\overline{\Omega})} + \|u - \overline{u}\|_{L^\infty(\Gamma)}) \cdot (\|y - \overline{y}\|_2^2 + \|u - \overline{u}\|_{L^2(\Gamma)}^2) \tag{4.16}$$

and

$$|\overline{\mathcal{L}}''[(y_1, u_1), (y_2, u_2)]| \le C_{\mathcal{L}}(\|y_1\|_2 + \|u_1\|_{L^2(\Gamma)})(\|y_2\|_2 + \|u_2\|_{L^2(\Gamma)}) \tag{4.17}$$

with some $C_{\mathcal{L}} > 0$, which depends in particular on $\overline{\varphi}$. This analysis is performed in the Appendix 7.3.

5. Standard second order sufficient optimality condition. We aim to establish sufficient optimality conditions, which are close to the necessary ones derived in Casas and Tröltzsch [8]. Therefore, we take into account *first order sufficient optimality conditions*. We combine an approach going back to Zowe and Maurer [14] with a splitting technique, which was known for the optimal control of ordinary differential equations and has been extended to the case of elliptic equations without state–constraints by the authors in [9].

In [14], Maurer and Zowe introduce first order sufficient optimality conditions taking into account a general constraint $g(w) \le 0$. Aiming to apply this approach to our problem in its full generality, we observed that this type of first order sufficient optimality conditions considerably complicates the presentation and the assumptions. Therefore, we introduce in a first step the first order sufficient optimality condition only for the constraints on the control. Later, we deal in the same way with the state–constraints, too. Define for fixed $\tau > 0$ (arbitrarily small)

$$\Gamma_\tau = \{x \in \Gamma \mid |g_u(x, \overline{y}(x), \overline{u}(x)) + \overline{\varphi}(x) b_u(x, \overline{y}(x), \overline{u}(x))| \ge \tau\}.$$

Γ_τ is a subset of "strongly active" control constraints (cf. (3.5)). Moreover, we mention that

$$\langle \overline{z}^*, E'(\overline{y})y \rangle \le 0 \tag{5.1}$$

holds, if $(y, u) \in L(\mathcal{M}, \overline{w})$. This follows from $\langle \overline{z}^*, E'(\overline{y})y \rangle = \varrho \langle \overline{z}^*, \kappa - E(\overline{y}) \rangle \le 0$ according to (3.4). Let $\mathcal{P}_\tau : L^\infty(\Gamma) \to L^\infty(\Gamma)$ denote the projection operator $u \mapsto \chi_{\Gamma \setminus \Gamma_\tau} u = \mathcal{P}_\tau u$. In other words, $(\mathcal{P}_\tau u)(x) = u(x)$ on $\Gamma \setminus \Gamma_\tau$ and $(\mathcal{P}_\tau u)(x) = 0$ on Γ_τ. We start with the following "standard" *second order sufficient optimality condition.*

(SSC) There exist positive numbers τ and δ such that

$$\mathcal{L}''(\overline{y}, \overline{u}, \overline{\varphi}, \overline{\lambda}, \overline{z}^*)[w_2, w_2] \ge \delta \|u_2\|_{L^2(\Gamma)}^2 \tag{5.2}$$

holds for all $w_2 = (y_2, u_2)$ obtained in the following way: For every $w = (y, u) \in L(\mathcal{M}, \overline{w})$ we split up the control part u by $u_1 = (u - \mathcal{P}_\tau u)$ and $u_2 = \mathcal{P}_\tau u$. Finally, we denote by y_i the linearized state associated to u_i, i. e.

$$\begin{cases} -\Delta y_i + y_i = 0 & \text{in } \Omega \\ \partial_\nu y_i = b_y(\cdot, \overline{y}, \overline{u}) y_i + b_u(\cdot, \overline{y}, \overline{u}) u_i & \text{on } \Gamma. \end{cases} \tag{5.3}$$

According to this, we get the splitting $w = w_1 + w_2 = (y_1, u_1) + (y_2, u_2)$.

THEOREM 5.1. *Let the feasible pair $\overline{w} = (\overline{y}, \overline{u})$ satisfy the regularity condition* **(R)**, *the first order necessary optimality conditions* (3.9)-(3.11) *and the second order sufficient optimality condition* **(SSC)**. *Suppose further that the general assumptions* **(A1)**-**(A4)** *are satisfied. Then there are constants $\varrho > 0$ and $\delta' > 0$ such that*

$$F_0(\hat{y}, \hat{u}) \geq F_0(\overline{y}, \overline{u}) + \delta' \|\hat{u} - \overline{u}\|^2_{L^2(\Gamma)} \tag{5.4}$$

holds for all feasible pairs $\hat{w} = (\hat{y}, \hat{u})$ such that

$$\|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)} < \varrho. \tag{5.5}$$

Proof. Let $\hat{w} = (\hat{y}, \hat{u})$ be a given feasible pair. We use for convenience the notation $\bar{l} = (\overline{\varphi}, \overline{\lambda}, \overline{z}^*)$ for the triplet of Lagrange multipliers appearing in the first order necessary optimality conditions. Obviously,

$$F_0(\hat{w}) - F_0(\overline{w}) = \mathcal{L}(\hat{w}, \bar{l}) - \mathcal{L}(\overline{w}, \bar{l}) - \langle \overline{z}^*, E(\hat{y}) - E(\bar{y}) \rangle \tag{5.6}$$

follows from $F(\hat{w}) = F(\overline{w}) = 0$. It holds

$$-\langle \overline{z}^*, E(\hat{y}) - E(\bar{y})) \rangle \geq 0,$$

hence we may avoid this term, and a second order Taylor expansion yields

$$\begin{aligned} F_0(\hat{w}) - F_0(\overline{w}) &\geq \mathcal{L}(\overline{w}, \bar{l}) - \mathcal{L}(\overline{w}, \bar{l}) \\ &\geq \int_\Gamma l_u (\hat{u} - \overline{u})\, dS + \frac{1}{2} \mathcal{L}''(\overline{w}, \bar{l})[\hat{w} - \overline{w}]^2 + r_2^{\mathcal{L}}(\overline{w}, \hat{w} - \overline{w}) \end{aligned}$$

where $l_u(x) = g_u(x, \overline{y}(x), \overline{u}(x)) + \overline{\varphi}(x) b_u(x, \overline{y}(x), \overline{u}(x))$. Hence

$$F_0(\hat{w}) - F(\overline{w}) \geq \tau \int_{\Gamma_\tau} |\hat{u} - \overline{u}|\, dS + \frac{1}{2} \mathcal{L}''(\overline{w}, \bar{l})[\hat{w} - \overline{w}]^2 + r_2^{\mathcal{L}}(\overline{w}, \hat{w} - \overline{w}). \tag{5.7}$$

We introduce for convenience the bilinear form $B = \mathcal{L}''(\overline{w}, \bar{l})$ and approximate $\hat{w} - \overline{w}$ by $w = (y, u) \in L(\mathcal{M}, \overline{w})$, according to Theorem 4.2. In this way we get $r = (r^y, r^u)$ such that $\hat{w} - \overline{w} = w + r$ and

$$\|r\| \leq C_L \|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)} \|\hat{u} - \overline{u}\|_{L^2(\Gamma)}. \tag{5.8}$$

Then $B[\hat{w} - \overline{w}]^2 = B[w]^2 + 2B[r, w] + B[r]^2$. We have $w \in L(\mathcal{M}, \overline{w})$, hence **(SSC)** applies to $B[w]^2$. Splitting up $w = w_1 + w_2$ as in **(SSC)**,

$$\begin{aligned} B[w]^2 &= B[w_2]^2 + 2B[w_1, w_2] + B[w_1]^2 \\ &\geq \delta \|u_2\|^2_{L^2(\Gamma)} - 2C_{\mathcal{L}}(\|y_1\|_2 + \|u_1\|_{L^2(\Gamma)})(\|y_2\|_2 + \|u_2\|_{L^2(\Gamma)}) \\ &\quad - C_{\mathcal{L}}(\|y_1\|_2 + \|u_1\|_{L^2(\Gamma)})^2 \end{aligned}$$

by **(SSC)** and (4.17). Suppose $\varrho < 1$ and $\|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)} < \varrho$. In the following, c denotes a generic constant. By $\|y_i\|_2 \leq c \|u_i\|_{L^2(\Gamma)}$ and Young's inequality,

$$B[w]^2 \geq \delta \|u_2\|^2_{L^2(\Gamma)} - \frac{\delta}{2} \|u_2\|^2_{L^2(\Gamma)} - c \|u_1\|^2_{L^2(\Gamma)}$$

$$\geq \frac{\delta}{2}\int\limits_{\Gamma\setminus\Gamma_r} u^2\,dS - c\int\limits_{\Gamma_r} u^2\,dS$$

$$\geq \frac{\delta}{2}\int\limits_{\Gamma\setminus\Gamma_r} |\hat{u}-\overline{u}|^2\,dS - c\int\limits_{\Gamma\setminus\Gamma_r} |\hat{u}-\overline{u}|\,|r^u|\,dS - c\int\limits_{\Gamma_r} |\hat{u}-\overline{u}|^2\,dS$$

$$-c\int\limits_{\Gamma_r} |\hat{u}-\overline{u}|\,|r^u|\,dS - c\int\limits_{\Gamma_r} |r^u|^2\,dS.$$

The third integrand is estimated by $\|\hat{u}-\overline{u}\|_{L^\infty(\Gamma)}|\hat{u}-\overline{u}|$, in the other integrals (excepting the first) we insert the estimate (5.8). This leads to

$$B[w]^2 \geq \frac{\delta}{2}\int\limits_{\Gamma\setminus\Gamma_r} |\hat{u}-\overline{u}|^2\,dS - c\varrho\int\limits_{\Gamma_r} |\hat{u}-\overline{u}|\,dS - c\varrho\|\hat{u}-\overline{u}\|^2_{L^2(\Gamma)}. \tag{5.9}$$

The estimation of $B[r,w]$ and $B[r]^2$ is simpler.

$$\begin{aligned}|B[r,w]| &\leq c\|r\|\|u\|_{L^2(\Gamma)} = c\|r\|\|\hat{u}-\overline{u}+r^u\|_{L^2(\Gamma)}\\ &\leq c\varrho\|\hat{u}-\overline{u}\|^2_{L^2(\Gamma)}.\end{aligned}$$

The same estimate applies to $B[r]^2$. Altogether,

$$B[\hat{w}-\overline{w}]^2 \geq \frac{\delta}{2}\int\limits_{\Gamma\setminus\Gamma_r} |\hat{u}-\overline{u}|^2\,dS - c\varrho\int\limits_{\Gamma_r} |\hat{u}-\overline{u}|\,dS - c\varrho\|\hat{u}-\overline{u}\|^2_{L^2(\Gamma)} \tag{5.10}$$

is obtained. Inserting (5.10) in (5.7), we get

$$\begin{aligned}F_0(\hat{w}) - F_0(\overline{w}) &\geq (\tau - c\varrho)\int\limits_{\Gamma_r} |\hat{u}-\overline{u}|\,dS + \frac{\delta}{2}\int\limits_{\Gamma\setminus\Gamma_r} |\hat{u}-\overline{u}|^2\,dS - c\varrho\|\hat{u}-\overline{u}\|^2_{L^2(\Gamma)}\\ &\quad -|r_2^{\mathcal{L}}(\overline{w},\hat{w}-\overline{w})|\\ &\geq \frac{\tau}{2}\int\limits_{\Gamma_r} |\hat{u}-\overline{u}|\,dS + \frac{\delta}{2}\int\limits_{\Gamma\setminus\Gamma_r} |\hat{u}-\overline{u}|^2\,dS - c\varrho\|\hat{u}-\overline{u}\|^2_{L^2(\Gamma)}\\ &\quad -|r_2^{\mathcal{L}}(\overline{w},\hat{w}-\overline{w})|.\end{aligned}$$

Because of $\|\hat{u}-\overline{u}\|_{L^\infty(\Gamma)} \leq 1$, we have $|\hat{u}-\overline{u}| \geq |\hat{u}-\overline{u}|^2$ almost everywhere. Using this in the first integral, setting $\delta' = \min\{\tau/2, \delta/2\}$, and inserting the estimate (4.16) for $r_2^{\mathcal{L}}$, we arrive at

$$\begin{aligned}F_0(\hat{w}) - F_0(\overline{w}) &\geq \|\hat{u}-\overline{u}\|^2_{L^2(\Gamma)}(\delta' - c\varrho - \eta(c\|\hat{u}-\overline{u}\|_{L^\infty(\Gamma)}))\\ &\geq \frac{\delta'}{2}\|\hat{u}-\overline{u}\|^2_{L^2(\Gamma)}\end{aligned}$$

for sufficiently small $\varrho > 0$. □

The study of the paper [14] reveals that first order sufficient optimality conditions can be extended also to state–constraints. However, this leads to a quite involved construction and more restrictive assumptions. We have to suppose that the function

b is linear with respect to the control u and $n = 2$. The associated theorem is stated below. For fixed $\beta > 0$ and $\tau > 0$ we define the following subset of $L(\mathcal{M}, \overline{w})$:

$$L_{\beta,\tau}(\mathcal{M}, \overline{w}) = \{w = (y, u) \mid w \in L(\mathcal{M}, \overline{w}) \text{ and satisfies (5.11) below}\}.$$

The decisive inequality characterising $L_{\beta,\tau}$ is

$$\langle \overline{z}^*, E'(\overline{y})y\rangle \geq -\beta \int_{\Gamma\setminus\Gamma_\tau} |u(x)|\, dS(x). \tag{5.11}$$

$L_{\beta,\tau}(\mathcal{M}, \overline{w})$ is the subset of $L(\mathcal{M}, \overline{w})$, where first order sufficient optimality conditions are not very supported by the term $\langle \overline{z}^*, E(y)\rangle$. It is only this set, where we have to require second order conditions, namely

(SSC') There exist positive numbers β, τ, and δ such that

$$\mathcal{L}''(\overline{y}, \overline{u}, \overline{\varphi}, \overline{\lambda}, \overline{z}^*)[w_2, w_2] \geq \delta \|u_2\|^2_{L^2(\Gamma)} \tag{5.12}$$

holds for all $w_2 = (y_2, u_2)$ obtained in the same way as in **(SSC)** by elements w taken from the smaller set $L_{\beta,\tau}(\mathcal{M}, \overline{w})$.

Using this condition we formulate

THEOREM 5.2. *Let the feasible pair $\overline{w} = (\overline{y}, \overline{u})$ satisfy the regularity condition* **(R)**, *the first order necessary optimality conditions* (3.9)-(3.11) *and the second order sufficient optimality condition* **(SSC')**. *Suppose further that the general assumptions* **(A1)**-**(A4)** *are satisfied. Moreover, assume that $n = 2$ and $b(x, y, u) = b_1(x, y) + b_2(x)u$. Then there are constants $\varrho > 0$ and $\delta' > 0$ such that*

$$F_0(\hat{y}, \hat{u}) \geq F_0(\overline{y}, \overline{u}) + \delta' \|\hat{u} - \overline{u}\|^2_{L^2(\Gamma)} \tag{5.13}$$

holds for all feasible pairs $\hat{w} = (\hat{y}, \hat{u})$ such that

$$\|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)} < \varrho. \tag{5.14}$$

Proof. We start exactly in the same way we have shown Theorem 5.1 and arrive at

$$F_0(\hat{w}) - F_0(\overline{w}) = \mathcal{L}(\hat{w}, \bar{l}) - \mathcal{L}(\overline{w}, \bar{l}) - \langle \overline{z}^*, E(\hat{y}) - E(\overline{y})\rangle. \tag{5.15}$$

Once again, $\hat{w} - \overline{w} = w + r$. Now we distinct between two cases.

Case I: First order sufficiency yields (5.4):

In this case

$$-\langle \overline{z}^*, E'(\overline{y})y\rangle > \beta \int_{\Gamma\setminus\Gamma_\tau} |u(x)|\, dS(x), \tag{5.16}$$

i. e. $w = (y, u) \in L(\mathcal{M}, \overline{w}) \setminus L_{\beta,\tau}(\mathcal{M}, \overline{w})$. Here we handle (5.15) as follows

$$\begin{aligned} F_0(\hat{w}) - F_0(\overline{w}) &= \mathcal{L}'(\overline{w}, \bar{l})(\hat{w} - \overline{w}) + r_1^{\mathcal{L}}(\overline{w}, \hat{w} - \overline{w}) - \langle \overline{z}^*, E(\hat{y}) - E(\overline{y})\rangle \\ &= \mathcal{L}_y(\overline{w}, \bar{l})(\hat{y} - \overline{y}) + \mathcal{L}_u(\overline{w}, \bar{l})(\hat{u} - \overline{u}) - \langle \overline{z}^*, E'(\overline{y})(\hat{y} - \overline{y})\rangle \\ &\quad + r_1^{\mathcal{L}}(\overline{w}, \hat{w} - \overline{w}) - \langle \overline{z}^*, r_1^E(\overline{y}, \hat{y} - \overline{y})\rangle \\ &= 0 + \int_\Gamma l_u(x)(\hat{u}(x) - \overline{u}(x))\, dS(x) - \langle \overline{z}^*, E'(\overline{y})y\rangle \\ &\quad + r_1^{\mathcal{L}}(\overline{w}, \hat{w} - \overline{w}) - \langle \overline{z}^*, E'(\overline{y})r^y + r_1^E(\overline{y}, \hat{y} - \overline{y})\rangle, \end{aligned} \tag{5.17}$$

where $l_u(x) = g_u(x, \bar{y}(x), \bar{u}(x)) + \bar{\varphi}(x) b_u(x, \bar{y}(x), \bar{u}(x))$.

Owing to $n = 2$ and $b(x, y, u) = b_1(x, y) + b_2(x)u$, we are able to apply the strong estimate (4.6) for $p = 2$. That is

$$\|r\|_{Y \times L^\infty(\Gamma)} \leq C_{L,2} \|\hat{u} - \bar{u}\|^2_{L^2(\Gamma)}. \tag{5.18}$$

By Theorem 4.2, (5.18), (4.15), and **(A3), (ii)** we have

$$\max\{\|r^y\|_2, |r_1^{\mathcal{C}}|, \|r_1^E\|_Z\} \leq c(\|\hat{y} - \bar{y}\|_2^2 + \|\hat{u} - \bar{u}\|^2_{L^2(\Gamma)}).$$

Thus the Lipschitz property of $u \mapsto y(u) = G(u)$ from $L^2(\Gamma)$ into $C(\bar{\Omega})$ (note that $n = 2$) implies that the last three items of (5.17) can be estimated by $c\,\|\hat{u} - \bar{u}\|^2_{L^2(\Gamma)}$. To the second one we apply (5.16), while the first one is treated by Γ_τ:
We know that

$$l_u(x)(\hat{u}(x) - \bar{u}(x)) \geq 0 \quad \text{a. e. on } \Gamma,$$

hence

$$\int_\Gamma l_u(\hat{u} - \bar{u})\, dS \geq \int_{\Gamma_\tau} l_u(\hat{u} - \bar{u})\, dS = \int_{\Gamma_\tau} |l_u|\, |\hat{u} - \bar{u}|\, dS \geq \tau \int_{\Gamma_\tau} |\hat{u} - \bar{u}|\, dS.$$

Now (5.17) can be continued by

$$\begin{aligned} F_0(\hat{w}) - F_0(\bar{w}) &\geq \tau \int_{\Gamma_\tau} |\hat{u} - \bar{u}|\, dS + \beta \int_{\Gamma \setminus \Gamma_\tau} |u|\, dS - c\|\hat{u} - \bar{u}\|^2_{L^2(\Gamma)} \\ &\geq \tau \int_{\Gamma_\tau} |\hat{u} - \bar{u}|\, dS + \beta \int_{\Gamma \setminus \Gamma_\tau} |\hat{u} - \bar{u}|\, dS - c\|\hat{u} - \bar{u}\|^2_{L^2(\Gamma)} \end{aligned}$$

as $\|r^u\|_{L^\infty(\Gamma)} \leq c\|\hat{u} - \bar{u}\|^2_{L^2(\Gamma)}$. Proceeding with the estimation, we have

$$\begin{aligned} F_0(\hat{w}) - F_0(\bar{w}) &\geq \min\{\beta, \tau\} \|\hat{u} - \bar{u}\|_{L^1(\Gamma)} - c\varrho \|\hat{u} - \bar{u}\|_{L^1(\Gamma)} \\ &\geq \beta' \|\hat{u} - \bar{u}\|_{L^1(\Gamma)} \end{aligned}$$

with some $\beta' > 0$, provided that $\|\hat{u} - \bar{u}\|_{L^\infty(\Gamma)} \leq \varrho \leq \varrho_1$, where ϱ_1 is sufficiently small. Assume additionally that $\varrho_1 \leq 1$. Then $|\hat{u} - \bar{u}|^2 \leq |\hat{u} - \bar{u}|$ a. e., hence

$$F_0(\hat{w}) - F_0(\bar{w}) \geq \beta' \|\hat{u} - \bar{u}\|^2_{L^2(\Gamma)} \tag{5.19}$$

for $\|\hat{u} - \bar{u}\|_{L^\infty(\Gamma)} \leq \varrho_1$.

Case II: Partial use of first order sufficient optimality conditions

Here, we avoid the term $\langle \bar{z}^*, E(\hat{y}) - E(\bar{y}) \rangle$ and proceed word for word as in the proof of Theorem 5.1, using $L_{\beta,\tau}$ instead of L. □

REMARK 5.3. *In applications, it will be quite difficult to describe in an explicit way, which $(y, u) \in L(\mathcal{M}, \bar{w})$ fall into the different classes, where case I or case II applies. Therefore, this type of first order sufficient condition is only of limited value.*

Theorem 5.1 follows from Theorem 5.2 by setting $\beta = 0$, where we can avoid the restrictions $n = 2$ and $b(x, y, u) = b_1(x, y) + b_2(x)u$.

By definition, $C(\bar{u}) = \{\rho(u - \bar{u}) \mid u \in \mathcal{U}^{ad}, \rho \geq 0\}$. The closure of $C(\bar{u})$ in $L^2(\Gamma)$ is

$$\mathrm{cl}\, C(\bar{u}) = \{v \in L^2(\Gamma) \mid v(x) \leq 0, \text{ if } \bar{u}(x) = u_b(x),\ v(x) \geq 0, \text{ if } \bar{u}(x) = u_a(x)\}.$$

If we require **(SSC)** *for $\mathrm{cl}\, C(\bar{u})$ instead of $C(\bar{u})$, then Theorem 5.1 holds as well, since $\mathrm{cl}\, C(\bar{u}) \supset C(\bar{u})$. It can be shown by means of* **(R)** *and the generalized open mapping theorem that (SSC) formulated for $\mathrm{cl}\, C(\bar{u})$ is in fact equivalent to (SSC) for $C(\bar{u})$.*

6. Extended second order conditions. A study of the preceding sections reveals that **(SSC)** is sufficient for local optimality in any dimension of Ω and without any restriction to the form of the nonlinear function b, whenever we know $\overline{\varphi} \in L^\infty(\Gamma)$. This holds, if pointwise state-constraints are required only in compact subsets of Ω and the other quantities are sufficiently smooth. We can allow for pointwise state-constraints up to the boundary Γ for $n = 2$, if $b(x, y, u)$ is linear w.r. to u. An extension to $\overline{\varphi} \in L^r(\Gamma)$ was possible only under stronger assumptions on b. We shall briefly sketch in this section that some extended form of **(SSC)** can partially improve the results for $n \leq 3$.

For $\overline{\varphi} \notin L^\infty(\Gamma)$ it seems to be natural to introduce in $L^\infty(\Gamma)$ the norm

$$\|u\|_\varphi = \left(\int_\Gamma (1 + |\overline{\varphi}(x)|) u^2(x)\, dS(x) \right)^{1/2}.$$

If $\overline{\varphi} \in L^\infty(\Gamma)$, this norm is equivalent to $\|u\|_{L^2(\Gamma)}$. Note that $u \in L^\infty(\Gamma)$ and $y \in C(\overline{\Omega})$ holds in all parts of our paper. However, we apply also different L^2-norms in this spaces. In order to get rid of the restrictions imposed on b in **(A4)** we re-define the set of strongly active control constraints Γ_τ by

$$\text{(6.20)} \quad \Gamma_{\tau,\varphi} = \{x \in \Gamma \,|\, |g_u(x, \overline{y}(x), \overline{u}(x)) + \overline{\varphi}(x) b_u(x, \overline{y}(x), \overline{u}(x))| \geq \tau(1 + |\overline{\varphi}(x)|)\}.$$

Moreover, we require instead of (5.2) the condition

$$\text{(6.21)} \qquad \mathcal{L}''(\overline{y}, \overline{u}, \overline{\varphi}, \overline{\lambda}, \overline{z}^*)[w_2, w_2] \geq \delta \|u_2\|_\varphi^2.$$

If we proceed in this way, then the statement of Theorem 5.1 remains true without assumption **(A4)** *for* $n = 2, 3$.

This can be seen as follows: The sections 1–3 are not influenced by introducing $\|u\|_\varphi$, while in section 4 only the estimates (4.15)–(4.17) have to be changed. This is the decisive point. We are able to replace $\|\cdot\|_{L^2(\Gamma)}$ by $\|\cdot\|_\varphi$ there, as the basic inequalities (7.15)–(7.17) (Appendix 7.3.) can be slightly re-formulated: (7.15) is nothing more than

$$\text{(6.22)} \qquad \int_\Gamma |\overline{\varphi}| u^2 dS \leq \|u\|_\varphi^2,$$

(7.17) remains unchanged ($n = 2, 3$), and (7.16) is substituted by

$$\begin{aligned} \int_\Gamma |\overline{\varphi}|\, |y|\, |u|\, dS &= \int_\Gamma |\overline{\varphi}|^{1/2} |y|\, |\overline{\varphi}|^{1/2} |u|\, dS \leq \|u\|_\varphi \left(\int_\Gamma |\overline{\varphi}| y^2 dS \right)^{1/2} \\ \text{(6.23)} \qquad &\leq \|\overline{\varphi}\|_{L^{s/(s-2)}(\Gamma)}^{1/2} \|y\|_{L^s(\Gamma)} \|u\|_\varphi, \end{aligned}$$

where we have used (7.16) for sufficiently large s ($n = 2, 3$). Now a careful study of the proof of Theorem 5.1 shows that **(A4)** can be removed on using (6.22) and (6.23). Assuming (6.21), we arrive at the estimate (5.4) with $\|\hat{u} - \overline{u}\|_\varphi^2$ instead of $\|\hat{u} - \overline{u}\|_{L^2(\Gamma)}^2$. Then (5.4) follows from $\|u\|_\varphi \geq \|u\|_{L^2(\Gamma)}$. The same arguments apply to the first order sufficient conditions in Theorem 5.2 for $n = 2$, if we re-define $L(\mathcal{M}, \overline{w})$ by means of the inequality

$$\langle \overline{z}^*, E'(\overline{y})y\rangle \geq -\beta \int_{\Gamma\setminus\Gamma_r} (1+|\overline{\varphi}|)|u|\,dS \tag{6.24}$$

substituted for (5.11).

7. Appendix.

7.1. On the regularity condition. **(R)** is satisfied in the following particular cases: Let $\hat{Y} \subset H^1(\Omega)$ denote the set of all solutions of the state equation (4.1) linearized at $(\overline{y},\overline{u})$ associated to $u \in L^\infty(\Gamma)$, i. e. $\hat{Y} = G'(\overline{u})L^\infty(\Gamma)$.

a) $K = Z$ (no inequality constraints)
Then **(R)** means $F'(\overline{y})G'(\overline{u})\mathcal{C}(\overline{u}) = \mathbb{R}^m$. This is implied by the stronger condition that $F'(\overline{y})\hat{Y} = \mathbb{R}^m$ and that there is an $\tilde{u} \in \text{int}_{L^\infty(\Gamma)}\mathcal{U}^{ad}$ with $F'(\overline{y})\tilde{y} = 0$, where $\tilde{y}$ is the solution of the linearized state equation (4.1) associated to $\tilde{u} - \overline{u}$, i. e. $\tilde{y} = G'(\overline{u})(\tilde{u} - \overline{u})$. The proof follows from Tröltzsch [20], Lemma 1.2.2.

b) $F = 0$ (no equality constraints)
In this case **(R)** reads $E'(\overline{y})G'(\overline{u})\mathcal{C}(\overline{u}) - K(E(\overline{u})) = Z$. Sufficient for this is $E'(\overline{y})\hat{Y} - K(E(\overline{y})) = Z$ and the existence of an $\tilde{u} \in \text{int}_{L^\infty(\Gamma)}\mathcal{U}^{ad}$ such that for $\tilde{y} = G'(\overline{u})(\tilde{u} - \overline{u})$ it holds $E'(\overline{y})\tilde{y} \in K(E(\overline{y}))$ (confirm again [20], Lemma 1.2.2). Case a) follows from b).

c) General case
(R) is implied by $\text{int}_Z K \neq \emptyset$, $\text{int}_{L^\infty(\Gamma)}\mathcal{U}^{ad} \neq \emptyset$,

$$F'(\overline{y})\hat{Y} = \mathbb{R}^m, \tag{7.1}$$

and the existence of an $\tilde{u} \in \text{int}_{L^\infty(\Gamma)}\mathcal{U}^{ad}$ such that it holds for $\tilde{y} = G'(\overline{u})(\tilde{u} - \overline{u})$

$$E(\overline{y}) + E'(\overline{y})\tilde{y} \in \text{int}_Z K, \tag{7.2}$$

$$F'(\overline{y})\tilde{y} = 0. \tag{7.3}$$

To show that c) implies **(R)** we first mention the simple fact that $\tilde{z} \in \text{int}_Z K$ implies $\tilde{z} + z/\varrho \in K$ for arbitrary $z \in Z$, if ϱ is sufficiently large.We show that the system

$$F'(\overline{y})y = z_1 \tag{7.4}$$

$$E'(\overline{y})y - \varrho(k - E(\overline{y})) = z_2 \tag{7.5}$$

is solvable for all $z_1 \in \mathbb{R}^m$, $z_2 \in Z$ with some $y \in G'(\overline{u})\mathcal{C}(\overline{u})$, $k \in K$, and $\varrho \geq 0$: By (7.1) we find $u_1 \in L^\infty(\Gamma)$ such that

$$F'(\overline{y})y_1 = z_1,$$

where $y_1 = G'(\overline{u})u_1$. Now we add to y_1 a multiple of $\tilde{y}$. Then

$$F'(\overline{y})(y_1 + \varrho\tilde{y}) = F'(\overline{y})y_1 = z_1$$

by (7.3). Thus, (7.4) holds for $y = y_1 + \varrho\tilde{y}$. Moreover,

$$E(\overline{y}) + E'(\overline{y})\tilde{y} - \frac{1}{\varrho}(z_2 - E'(\overline{y})y_1) = k \in K$$

for sufficiently large ϱ by (7.2). This means

$$E'(\overline{y})(y_1 + \varrho\tilde{y}) - \varrho(k - E(\overline{y})) = z_2.$$

Therefore, (7.5) holds for $y = y_1 + \varrho\tilde{y}$. Furthermore, $u_1 + \varrho(\tilde{u} - \overline{u}) = \varrho(\tilde{u} + (1/\varrho)u_1 - \overline{u}) \in \mathcal{C}(\overline{u})$ for sufficiently large ϱ, as $\tilde{u} + (1/\varrho)u_1 \in \mathcal{U}^{ad}$ for ϱ large enough ($\tilde{u} \in \text{int}_{L^\infty(\Gamma)}\mathcal{U}^{ad}$). Thus $y \in G'(\overline{u})\mathcal{C}(\overline{u})$ what remained to be shown.

7.2. Proof of the linearization theorem. To show Theorem 4.2 we need the following auxiliary result:

LEMMA 7.1. *Let $\overline{u}, \hat{u} \in \mathcal{U}^{ad}$ be given with associated states $\overline{y}, \hat{y}$ defined by (2.2). Define $y \in Y$ as the solution of the linearized state equation*

$$\left\{\begin{array}{ll} -\Delta y + y = 0 & \text{in } \Omega \\ \partial_\nu y = b_y(\cdot, \overline{y}, \overline{u})y + b_u(\cdot, \overline{y}, \overline{u})(\hat{u} - \overline{u}) & \text{on } \Gamma. \end{array}\right. \tag{7.6}$$

Then there are constants C_p, C_2 such that

$$\|\hat{y} - \overline{y} - y\|_Y \leq C_p \|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)} \|\hat{u} - \overline{u}\|_{L^p(\Gamma)} \quad \forall p > n-1 \tag{7.7}$$

$$\|\hat{y} - \overline{y} - y\|_2 \leq C_2 \|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)} \|\hat{u} - \overline{u}\|_{L^2(\Gamma)}. \tag{7.8}$$

In the case, where $b_u(x, y, u)$ does not depend on y and u, it holds

$$\|\hat{y} - \overline{y} - y\|_Y \leq C_p \|\hat{u} - \overline{u}\|^2_{L^p(\Gamma)} \quad \forall p > n-1. \tag{7.9}$$

Proof. We use the first order expansion of b at $(x, \hat{y}, \hat{u})$ and $(x, \overline{y}, \overline{u})$ and obtain from (2.2), (7.6), and (4.9) that

$$\begin{array}{rl} -\Delta(\hat{y} - \overline{y} - y) + (\hat{y} - \overline{y} - y) = 0 & \text{in } \Omega \\ \partial_\nu(\hat{y} - \overline{y} - y) - b_y(\cdot, \overline{y}, \overline{u})(\hat{y} - \overline{y} - y) = r_1^b & \text{on } \Gamma, \end{array}$$

where

$$|r_1^b(x)| \leq C_M(|\hat{y}(x) - \overline{y}(x)|^2 + |\hat{u}(x) - \overline{u}(x)|^2)$$

and M depends on $\mathcal{U}^{ad}$ (note that the boundedness of $\mathcal{U}^{ad}$ implies a uniform bound on all admissible states). Therefore the discussion of (3.12) yields for $p > n-1$

$$\begin{aligned} \|\hat{y} - \overline{y} - y\|_Y &\leq c\|r_1^b\|_{L^p(\Gamma)} \\ &\leq c\left(\left(\int_\Gamma |\hat{y} - \overline{y}|^{2p} dS\right)^{\frac{1}{p}} + \left(\int_\Gamma |\hat{u} - \overline{u}|^{2p} dS\right)^{\frac{1}{p}}\right). \end{aligned}$$

If $p > n-1$, then the mapping $u \mapsto y = G(u)$ is Lipschitz from $L^p(\Gamma)$ to $C(\overline{\Omega})$. In the case $p = 2$ this holds in the norm $\|y\|_2$ for y. For $p > n-1$ we continue

$$\|\hat{y} - \overline{y} - y\|_Y \leq c\left(\|\hat{u} - \overline{u}\|^2_{L^p(\Gamma)} + \|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)}\|\hat{u} - \overline{u}\|_{L^p(\Gamma)}\right),$$

while $p = 2$ yields only

$$\|\hat{y} - \overline{y} - y\|_2 \leq c\|\hat{u} - \overline{u}\|_{L^\infty(\Gamma)}\|\hat{u} - \overline{u}\|_{L^2(\Gamma)},$$

thus (7.7) and (7.8) are shown. If b_u does not depend on (y,u), then $b_u(\cdot,\overline{y}+\vartheta(\hat{y}-\overline{y}),\overline{u}+\vartheta(\hat{u}-\overline{u}))=b_u(\cdot,\overline{y},\overline{u})$, hence

$$|r_1^b|=|(b_y^\vartheta-\bar{b}_y)(\hat{y}-\overline{y})|\le c|\hat{y}-\overline{y}|^2.$$

This yields

$$\|r_1^b\|_{L^p(\Gamma)}\le c\left(\|\hat{y}-\overline{y}\|_{C(\Gamma)}\|\hat{y}-\overline{y}\|_{L^p(\Gamma)}\right)\le c\|\hat{u}-\overline{u}\|^2_{L^p(\Gamma)},$$

i. e. (7.9). □

Proof of Theorem 4.2. We follow the main line of proof given by Maurer and Zowe [14] furnishing it with a detailed discussion of remainders. Define $v=\hat{u}-\overline{u}$ and let $\tilde{y}$ denote the solution of the linear system (4.1) associated to $u:=v$. We have $\tilde{y}=G'(\overline{u})v$, where $G:L^\infty(\Gamma)\to Y$ is the control–state mapping $u\mapsto y=G(u)$ for the nonlinear system (2.2). By Lemma 7.1,

$$\|\hat{y}-\overline{y}-\tilde{y}\|_Y\le e(v), \tag{7.10}$$

where $e(v)$ denotes the right hand side of the estimates (7.7) and (7.9), respectively, depending on the assumptions on b. Further,

$$T(\hat{y})-T(\overline{y})=T'(\overline{y})(\hat{y}-\overline{y})+r^T,$$

where

$$\|r^T\|_{\boldsymbol{Z}}\le c\|\hat{y}-\overline{y}\|_2^2 \tag{7.11}$$

according to the Lipschitz estimates in **(A3)** for F_i' and E'. Thus

$$T(\hat{y})-T(\overline{y})=T'(\overline{y})\tilde{y}+\tilde{r}+r^T, \tag{7.12}$$

where $\tilde{r}=T'(\overline{y})(\hat{y}-\overline{y}-\tilde{y})$ satisfies an estimate of the type (7.10). Put $\Phi(u)=T(G(u))$. Then $\Phi'(\overline{u})v=T'(\overline{y})G'(\overline{u})v=T'(\overline{y})\tilde{y}$, hence (7.12) reads

$$\Phi(\hat{u})-\Phi(\overline{u})=\Phi'(\overline{u})v+\tilde{r}+r^T.$$

(R) can be re–written as

$$\Phi'(\overline{u})\mathcal{C}(\overline{u})-\boldsymbol{K}(\Phi(\overline{u}))=\boldsymbol{Z}.$$

Now the *generalized open mapping theorem* by Zowe and Kurcyusz [21] applies: There is an $\alpha>0$ such that

$$\alpha[\Phi'(\overline{u})(\mathcal{C}(\overline{u})\cap B_{\mathcal{U}}(0))-(\boldsymbol{K}(\Phi(\overline{u}))\cap B_{\boldsymbol{Z}}(0))]\supset B_{\boldsymbol{Z}}(0),$$

where $B_{\mathcal{U}}(0)$ and $B_{\boldsymbol{Z}}(0)$ denote the closed unit balls of $L^\infty(\Gamma)$ and $\boldsymbol{Z}$ around 0, respectively. This implies

$$\alpha\Phi'(\overline{u})(\mathcal{C}(\overline{u})\cap B_{\mathcal{U}}(0))-\boldsymbol{K}(\Phi(\overline{u}))\supset B_{\boldsymbol{Z}}(0).$$

We choose $z=(\tilde{r}+r^T)/\|\tilde{r}+r^T\|_{\boldsymbol{Z}}$ and conclude that there exists an $h\in\mathcal{C}(\overline{u})$ such that $\|h\|_{L^\infty(\Gamma)}\le\alpha\|\tilde{r}+r^T\|_{\boldsymbol{Z}}$ and

$$\Phi'(\overline{u})h-k=\tilde{r}+r^T,$$

where $k \in \boldsymbol{K}(\Phi(\overline{u})) = \boldsymbol{K}(T(\overline{y}))$. Define $y_h = G'(\overline{u})h$. Then

$$\tilde{r} + r^T = T'(\overline{y})y_h - k. \tag{7.13}$$

Put $y = \tilde{y} + y_h$, $u = v + h$ and insert (7.13) in (7.12). Then

$$\begin{aligned} T'(\overline{y})y &= T(\hat{y}) - T(\overline{y}) + k \\ &\in \boldsymbol{K} - T(\overline{y}) + \boldsymbol{K}(T(\overline{y})) \\ &\subset \boldsymbol{K}(T(\overline{y})) + \boldsymbol{K}(T(\overline{y})) = \boldsymbol{K}(T(\overline{y})), \end{aligned}$$

since $T(\hat{y}) \in \boldsymbol{K}$ and $\boldsymbol{K}(T(\overline{y}))$ is a cone. Moreover, $r^u = \hat{u} - \overline{u} - u = -h$,

$$\|r^u\|_{L^\infty(\Gamma)} = \|\hat{u} - \overline{u} - u\|_{L^\infty(\Gamma)} = \|h\|_{L^\infty(\Gamma)} \le \alpha\|\tilde{r} + r^T\|_{\boldsymbol{Z}}$$

and $r^y = \hat{y} - \overline{y} - y$,

$$\begin{aligned} \|r^y\|_Y = \|\hat{y} - \overline{y} - y\|_Y &\le \|\hat{y} - \overline{y} - \tilde{y}\|_Y + \|y_h\|_Y \le e(v) + \|y_h\|_Y \\ &\le e(v) + c\|h\|_{L^\infty(\Gamma)} \le e(v) + c\|\tilde{r} + r^T\|_{\boldsymbol{Z}}. \end{aligned}$$

Thus

$$\begin{aligned} \|r\|_{Y\times L^\infty(\Gamma)} = \|r^y\|_Y + \|r^u\|_{L^\infty(\Gamma)} &\le e(v) + 2c\|\tilde{r} + r^T\|_Z \\ &\le e(v) + ce(v) + c\|\hat{y} - \overline{y}\|_2^2 \end{aligned} \tag{7.14}$$

by (7.10) and (7.11). Therefore,

$$\begin{aligned} \|r\|_{Y\times L^\infty(\Gamma)} &\le c(\|v\|_{L^\infty(\Gamma)}\|v\|_{L^p(\Gamma)} + \|v\|_{L^p(\Gamma)}^2) \\ &\le c\|v\|_{L^\infty(\Gamma)}\|v\|_{L^p(\Gamma)} \end{aligned}$$

in the case of (7.7) and

$$\|r\|_{Y\times L^\infty(\Gamma)} \le c\|v\|_{L^p(\Gamma)}^2$$

in the case that (7.9) holds. (4.5) is shown completely analogous. Here, $e(v)$ is defined by (7.8), $\|\cdot\|_Y$ is to be replaced by $\|\cdot\|_2$ and $\|\cdot\|_{L^\infty(\Gamma)}$ by $\|\cdot\|_{L^2(\Gamma)}$. We rely on the continuity of $T'(\overline{y})$ in the L^2-norm. (4.5) follows directly from (7.14). □

7.3. Estimates of the Lagrange function. In this subsection we derive the estimates (4.15)–(4.17) for $r_1^{\mathcal{L}}$, $r_2^{\mathcal{L}}$, and $\mathcal{L}''$. They depend mainly on the estimation of I defined in (4.8), which is performed by the discussion of the following integrals:

$$\int_\Gamma |\overline{\varphi}|\, u^2\, dS \le c\|u\|_{L^2(\Gamma)}^2 \tag{7.15}$$

provided that assumption **(A4)**, **(i)** is fulfilled, and

$$\begin{aligned} \int_\Gamma |\overline{\varphi}|\,|y|\,|u|\, dS &\le c\|\overline{\varphi}y\|_{L^2(\Gamma)}\|u\|_{L^2(\Gamma)} \le c\|\overline{\varphi}^2\|_{L^{(s/2)'}(\Gamma)}^{1/2}\|y^2\|_{L^{s/2}(\Gamma)}^{1/2}\|u\|_{L^2(\Gamma)} \\ &\le c\|\overline{\varphi}\|_{L^{2s/(s-2)}(\Gamma)}\|y\|_{L^s(\Gamma)}\|u\|_{L^2(\Gamma)}. \end{aligned} \tag{7.16}$$

These estimates are justified by **(A4)**, **(ii)**: For $n = 2$ we know $y \in C(\Gamma)$ and $\varphi \in L^r(\Gamma)$ $\forall r < \infty$. If $n \ge 3$, then $y \in L^s(\Gamma)$ for all $s < 2(n-1)/(n-3)$ (including

$s < \infty$ for $n = 3$). The function $2s/(s-2) = 2/(1-1/s)$ is monotone decreasing. Therefore, $s \uparrow 2(n-1)/(n-3)$ implies $2s/(s-2) \downarrow n-1$, so that $\overline{\varphi} \in L^r(\Gamma)$ for some $r > n-1$ is sufficient to justify (7.16) with a sufficiently large s. Finally,

$$\int_\Gamma |\overline{\varphi}| y^2 \, dS \le \|\overline{\varphi}\|_{L^{(s/2)'}(\Gamma)} \|y^2\|_{L^{s/2}(\Gamma)} = \|\overline{\varphi}\|_{L^{s/(s-2)}(\Gamma)} \|y\|^2_{L^s(\Gamma)} \tag{7.17}$$

is estimated by **(A4), (iii)**: In the case $n = 2$ we can take $s = \infty$, as $y \in C(\Gamma)$ and $\varphi \in L^1(\Gamma)$ is true without any additional assumption. For $n = 3$ we know $y \in L^s(\Gamma)$ for all $s < \infty$. If $s \uparrow \infty$, then $s/(s-2) \downarrow 1 < n/(n-1)$. Since $\varphi \in L^r(\Gamma)$ for all $r < n/(n-1)$, (7.17) is true for sufficiently large s. In the case $n \ge 4$ the same analysis as in the case $n \ge 3$ above leads to the additional assumption $\overline{\varphi} \in L^r(\Gamma)$ for some $r > \frac{n-1}{2}$. Now it is very easy to derive the estimates (4.15)–(4.17) for $\mathcal{L}''$, $r_1^{\mathcal{L}}$, and $r_2^{\mathcal{L}}$: For instance, I in (4.8) is handled by (7.15)–(7.17) and

$$\begin{aligned} |I| &\le \int_\Gamma |\overline{\varphi}| (|\overline{b}_{yy}| \, |y_1 y_2| + |\overline{b}_{yu}| (|y_1 u_2| + |y_2 u_1|) + |\overline{b}_{uu}| \, |u_1 u_2| \, dS \\ &\le c(\|y_1\|_2 + \|u_1\|_{L^2(\Gamma)})(\|y_2\|_2 + \|u_2\|_{L^2(\Gamma)}), \end{aligned}$$

as $\overline{b}_{yy}$, $\overline{b}_{yu}$, and $\overline{b}_{uu}$ belong to $L^\infty(\Gamma)$. The other parts of $\mathcal{L}''$ are discussed by means of **(A1)–(A3)**. This yields after easy computations (4.17). In the same way, the remainder terms are investigated. Here, the terms connected with I are the most difficult ones again. For instance, (7.15)–(7.17) applies to discuss

$$\begin{aligned} |r_2^I| = \int_\Gamma |\overline{\varphi}| \; &\{ |b_{yy}^\vartheta - \overline{b}_{yy}| \, |y - \overline{y}|^2 + 2|b_{yu}^\vartheta - \overline{b}_{yu}| \, |y - \overline{y}| \, |u - \overline{u}| \\ &+ |b_{uu}^\vartheta - \overline{b}_{uu}| \, |u - \overline{u}|^2 \} \; dS \\ \le c\eta(&\|y - \overline{y}\|_{C(\Gamma)} + \|u - \overline{u}\|_{L^\infty(\Gamma)})(\|y - \overline{y}\|_2^2 + \|u - \overline{u}\|^2_{L^2(\Gamma)}), \end{aligned}$$

which contributes to $r_2^{\mathcal{L}}$. The other terms of $r_2^{\mathcal{L}}$ are handled by the estimates for second order derivatives in **(A1)–(A3)** in a direct way. Simple computations of this type verify (4.15)–(4.16). We leave the details to the reader.

REFERENCES

[1] Adams, R. A.: Sobolev spaces. Academic Press, 1978.

[2] Alt, W.: On the approximation of infinite optimization problems with an application to optimal control problems. Appl. Math. Optimization 12 (1984), pp. 15–27.

[3] Alt, W.: The Lagrange-Newton method for infinite-dimensional optimization problems. Numer. Funct. Anal. and Optim. 11 (1990), pp. 201–224.

[4] Alt, W.: Discretization and mesh-independence of Newton's method for generalized equations, to appear.

[5] Bonnans, F.: Second order analysis for control constrained optimal control problems of semilinear elliptic systems. To appear in Appl. Math. Optimization.

[6] Bonnans, F. and Casas, F.: Controle de systemes elliptiques semilineaires comportant des contraintes sur l' etat. In: H. Brezis, J. L. Lions (Eds.), Nonlinear partial differential equations and their applications, College de France Seminar Vol. VIII; Pitman Research Notes in Mathematics Series, Vol. 166, pp. 69–86, Longman Scientific & Technical, Burnt Mill, Harlow, 1988.

[7] Casas, E.: Boundary control with pointwise state constraints. SIAM J. Control Optim. 31 (1993), pp. 993–1006.

[8] Casas, E. and F. Tröltzsch: Second order necessary optimality conditions for some state-constrained control problems of semilinear elliptic equations. To appear in Appl. Math. Optimization.

[9] Casas, E., Tröltzsch, F., and A. Unger: Second order sufficient optimality conditions for a nonlinear elliptic control problem. J. for Analysis and its Appl. 15 (1996), pp. 687–707.

[10] Goldberg, H. and Tröltzsch, F.: Second order optimality conditions for a class of control problems governed by non-linear integral equations with applications to parabolic boundary control. Optimization 20 (1989), pp. 687–698.

[11] Goldberg, H. and Tröltzsch, F.: Second order sufficient optimality conditions for a class of non-linear parabolic boundary control problems, SIAM J. Control Optim. 31 (1993), pp. 1007–1027.

[12] Ioffe, A.D.: Necessary and sufficient conditions for a local minimum, part 3: Second order conditions and augmented duality. SIAM J. Control Optimization 17 (1979), pp. 266–288.

[13] Maurer, H: First and second order sufficient optimality conditions in mathematical programming and optimal control. Math. Programming Study 14 (1981), pp. 163–177.

[14] Maurer, H. and Zowe, J.: First- and second-order conditions in infinite-dimensional programming problems. Math. Programming 16 (1979), pp. 98–110.

[15] Murthy, M. K. V. and Stampacchia, G.: A variational inequality with mixed boundary conditions. Israel J. Math. 13 (1972), pp. 188–224.

[16] Nečas, J.: Les Méthodes Directes en Théorie des Equations Elliptiques. Editeurs Academia, Prague, 1967.

[17] Stampacchia, G.: Problemi al contorno ellittici con dati discontinui dotati di soluzioni Hölderiane. Ann. Math. Pura Appl. 51 (1960), pp. 1–38.

[18] Stampacchia, G.: Equations Elliptiques du Second Ordre a Coefficients Discontinus. Les Presses de l'Universite de Montreal, Montreal, 1966.

[19] Triebel, H.: Interpolation theory, function spaces, differential operators. 2nd edition, J.A. Barth Verlag, Heidelberg, 1995.

[20] Tröltzsch, F.: Optimality conditions for parabolic control problems and applications, Teubner-Texte zur Mathematik, 62, B. G. Teubner Verlagsgesellschaft, Leipzig, 1984.

[21] Zowe, J. and Kurcyusz, S.: Regularity and stability for the mathematical programming problem in Banach spaces. Appl. Math. Optimization 5 (1979), pp. 49–62.