SECOND-ORDER SUFFICIENT OPTIMALITY CONDITIONS FOR THE OPTIMAL CONTROL OF NAVIER-STOKES EQUATIONS

FREDI TRÖLTZSCH AND DANIEL WACHSMUTH[1]

Abstract. In this paper sufficient optimality conditions are established for optimal control of both steady-state and evolution Navier-Stokes equations. The second-order condition requires coercivity of the Lagrange function on a suitable subspace together with first-order necessary conditions. It ensures local optimality of a reference function in a L^s-neighborhood, whereby the underlying analysis allows to use weaker norms than L^∞.

Key words. Optimal control, Navier-Stokes equations, control constraints, second-order optimality conditions, first-order necessary conditions

AMS subject classifications. Primary 49K20, Secondary 49K27

1. Introduction. In this paper, we discuss second-order sufficient optimality conditions for optimal control problems governed by steady-state and instationary Navier-Stokes equations. These conditions form a central issue for different mathematical questions of optimal control theory. If second-order sufficient conditions hold true at a given control satisfying the first-order necessary conditions, then this control is locally optimal, it is unique as a local solution, and it is stable with respect to certain perturbations of given data. Moreover, the convergence of numerical approximations (say by finite elements) can be proven, and numerical algorithms such as SQP methods can be shown to locally converge.

Consequently, second-order conditions have been important assumptions in many papers on optimal control theory of ordinary differential equations, and it became important for partial differential equations as well. We only mention the case of elliptic equations studied by Casas, Unger, and Tröltzsch [10], Casas and Mateos [8], the discussion of pointwise state-constraints in Casas, Unger, and Tröltzsch [11], Raymond and Tröltzsch [22], or the convergence analysis of SQP methods in Arada, Raymond and Tröltzsch [3] and Tröltzsch [26]. The papers mentioned above are concerned with semilinear elliptic and parabolic equations with nonlinearities given by Nemytski operators. Therefore, the associated state functions have to be continuous to make these operators twice continuously differentiable.

The situation is, in some sense, easier for the Navier-Stokes equations. The nonlinearity $(y\cdot\nabla)\,y$ appearing in these equations is of quadratic type, and the associated Taylor expansion terminates after the second-order term with zero remainder. This property has been addressed by Hinze [20] for the optimal control of instationary Navier-Stokes equations. It simplifies the application of second-order conditions, since spaces of L^2-type for the control and $W(0,T)$-type for the state function are appropriate.

In [23], it was shown for the case of steady-state Navier-Stokes equations that second-order conditions are sufficient for Lipschitz stability of optimal solutions with respect

*This work was supported by SFB 557 "Control of complex turbulent shear flows", Technical University of Berlin.

[1]Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Str. des 17. Juni 136, Germany.

to perturbations. However, second-order conditions were applied in a quite strong form without showing their sufficiency for local optimality.

Here, the issue of second-order sufficiency is studied more detailed. We present the conditions in a fairly weak form that invokes also first-order sufficient conditions. More precisely, by using strongly active control constraints we shrink the subspace where the second derivative of the Lagrange function must be positive definite. Moreover, we carefully study the norms underlying the neighborhood, where local optimality can be assured, which enables us to prove local optimality in an L^s-neighborhood of the reference control with $s < \infty$. We discuss the steady-state and instationary Navier-Stokes equations in one paper, since the arguments are very similar for both cases.

As concerns strongly active constraints, we follow an approach by Dontchev, Hager, Poore and Yang [14] that has been successfully applied in other papers on second-order conditions as well. By this technique, a certain gap between second-order necessary and second-order sufficient conditions appears. This gap seems to be natural for problems in infinite-dimensional spaces. In a paper by Bonnans and Zidani [5], the gap was tightened under the assumption that the second-order derivative of the Lagrangian defines a Legendre form. Casas and Mateos [9] extended the applicability of this concept by an assumption of positivity on the second derivative of the Hamiltonian with respect to the control. Using these techniques, we also resolve the problem of the two-norm discrepancy: an appropriate formulation of the sufficient optimality condition implies L^2-quadratic growth of the objective in a L^2-neighborhood of the reference control.

Our arguments are influenced by various papers, where first-order necessary optimality conditions and numerical methods for optimal control of instationary Navier-Stokes equations are presented. We only mention Abergel and Temam [1], Casas [7], Gunzburger [17], Gunzburger and Manservisi [19], Fattorini and Sritharan [16], Hinze [20], Hinze and Kunisch [21], Sritharan [24] and the reference cited therein. We partially repeat some known arguments for proving first-order necessary conditions only for convenience.

2. The optimal control problems.

2.1. Control of the steady-state Navier-Stokes equations. In the first part of the paper, we consider the optimal control problem to minimize

$$J(u,y) = \int_\Omega |y(x) - y_d(x)|^2 dx + \frac{\gamma}{2} \int_\Omega |u(x)|^2 dx \tag{2.1}$$

subject to the steady-state Navier-Stokes equations,

$$\begin{aligned} -\nu\Delta y + (y\cdot\nabla)y + \nabla p &= u \qquad \text{in } \Omega, \\ \operatorname{div} y &= 0 \qquad \text{in } \Omega, \\ y &= 0 \qquad \text{on } \Gamma, \end{aligned} \tag{2.2}$$

and the box constraints

$$u_a(x) \le u(x) \le u_b(x)$$

to be fulfilled a.e. on Ω. In this setting Ω is an open bounded Lipschitz domain in $\mathbb{R}^n$ with boundary Γ. In the steady-state case, we will restrict the space dimension n to $2 \le n \le 4$. In this case, $H^1_0(\Omega)$ is continuously imbedded in $L^4(\Omega)$.

To complete the problem setting, we require the desired function y_d to be an element of $L^2(\Omega)^n$. The parameters γ and ν are assumed to be positive constants. In the box constraints on u two functions $u_a, u_b \in L^s(\Omega)^n$ are given, satisfying $u_{a,i}(x) \le u_{b,i}(x)$ for all $i = 1 \dots n$ and almost all $x \in \Omega$. The exponent s will be precised later. We set

$$U_{ad} = \{u \in L^s(\Omega)^n : \ u_{a,i}(x) \le u_i(x) \le u_{b,i}(x), \ i = 1 \dots n, \ \text{a.e. on } \Omega\}.$$

Up to now we did not explain, in which sense the state equations (2.2) has to be solved. The state y associated with u is defined as a weak solution of (2.2) in the next section.

2.2. The instationary case. In the second part, we consider the optimal control problem to minimize

$$\begin{aligned} J(u,y) = \frac{1}{2}\int_\Omega |y(x,T) - y_T(x)|^2 dx + \frac{1}{2}\int_0^T\int_\Omega |y(x,t) - y_Q(x,t)|^2 dxdt \\ +\frac{\gamma}{2}\int_0^T\int_\Omega |u(x,t)|^2 dxdt \end{aligned} \tag{2.3}$$

subject to the instationary Navier-Stokes equations.

$$\begin{aligned} y_t - \nu\Delta y + (y\cdot\nabla)y + \nabla p = u \quad & \text{in } Q, \\ \operatorname{div} y = 0 \quad & \text{in } Q, \\ y(0) = y_0 \quad & \text{in } \Omega, \end{aligned} \tag{2.4}$$

and the control constraints $u \in U_{ad}$ with control set re-defined below, where $Q = (0,T) \times \Omega$. Here, functions $y_T \in L^2(\Omega)^n$, $y_Q \in L^2(Q)^n$, and $y_0 \in H \subset L^2(\Omega)^n$ are given. The parameters γ and ν are adopted from the last section. Let two functions $u_a, u_b \in L^s(Q)^n$ be given such that $u_{a,i}(x,t) \le u_{b,i}(x,t)$ holds almost everywhere on Q and for all $i = 1, \dots, n$. The set of admissible controls is now defined by

$$U_{ad} = \{u \in L^s(Q)^n : \ u_{a,i}(x,t) \le u_i(x,t) \le u_{b,i}(x,t) \ \text{a.e. on } Q\}.$$

Again, the exponent s will be specified later.

3. Optimality conditions for the steady-state problem. In this section, we provide basic results on the state equation and first-order necessary optimality conditions. These results are more or less known from the literature. However, they are mostly presented in a different form and not directly applicable for our purposes. Therefore, we recall them for convenience.

3.1. The state equation. First, we define a solenoidal space that is frequently used in the literature,

$$V := \{v \in H_0^1(\Omega)^n : \operatorname{div} v = 0\}.$$

This space is a Hilbert space endowed with the standard scalar product of H_0^1,

$$(y,v)_V = \sum_{i=1}^n (\nabla y_i, \nabla v_i)_{L^2(\Omega)}.$$

The associated norm is denoted by $|\cdot|_V$. Further on, we will denote the pairing between V' and V as $\langle f, v\rangle$, where $f \in V'$ and $v \in V$. To simplify the notation, we define for $u \in L^q(\Omega)^n$

$$|u|_q := |u|_{L^q(\Omega)^n}.$$

The pairing between $L^q(\Omega)^n$ and $L^{q'}(\Omega)^n$ is denoted by $(\cdot,\cdot)_{q,q'}$, $1/q + 1/q' = 1$. For $q = q' = 2$ we get the usual scalar product of $L^2(\Omega)^n$, and we write $(\cdot,\cdot)_2 := (\cdot,\cdot)_{2,2}$. In the following, we will make use of the well-known interpolation inequality, cf. Brezis [6].

Lemma 3.1. *Let $1 \le q \le 2$ be given. Define s by $s = q/(2-q)$ for $q < 2$, or $s = \infty$ for $q = 2$, respectively. Further, let $D \subset \mathbb{R}^m$ be a bounded and measurable set. Then it holds for all $u \in L^s(D)$*

$$|u|^2_{L^q(D)} \le |u|_{L^1(D)} |u|_{L^s(D)}.$$

Additionally, we need the following well-known lemma of imbeddings of L^p-spaces, cf. Adams [2].

Lemma 3.2. *Let $D \subset \mathbb{R}^m$ be a bounded and measurable set with* $\mathrm{vol}(D) := \int_D 1dx < \infty$*, and let $1 \le p \le q \le \infty$ be given. Then for all $u \in L^q(D)$ it holds*

$$|u|_{L^p(D)} \le (\mathrm{vol}(D))^{1/p-1/q} |u|_{L^q(D)}.$$

Let us introduce for convenience a trilinear form $b : V \times V \times V \mapsto \mathbb{R}$ by

$$b(u, v, w) = ((u \cdot \nabla)v, w)_2 = \int_\Omega \sum_{i,j=1}^n u_i \frac{\partial v_j}{\partial x_i}\, w_j\, dx.$$

The following result was proven in [25].

Lemma 3.3. *For all $u, v, w \in V$ it holds*

$$b(u, v, w) = -b(u, w, v).$$

There is a positive constant C_n depending on the dimension n but not on u, v, w and Ω, such that

$$|b(u, v, w)| \le C_n\, |u|_4 |v|_V |w|_4 \tag{3.1}$$

holds for all $u, v, w \in V$.

As a simple conclusion of the previous lemma, we get $b(u, v, v) = 0$ for all $u, v \in V$. The estimate (3.1) expresses the continuity of b. We refer to [12, 20] for further estimates of b.

Furthermore, we introduce for $p \le 2n/(n-2)$ by N_p the norm of the imbedding of $H^1_0(\Omega)^n$ in $L^p(\Omega)^n$, i.e. $|y|_p \le N_p |y|_{H^1_0(\Omega)^n}$. For $2 \le n \le 4$, the imbedding of $H^1_0(\Omega)^n$ in $L^4(\Omega)^n$ is continuous. This fact will be frequently used. Moreover, we conclude from (3.1)

$$|b(u, v, w)| \le C_n N_4^2 |u|_V |v|_V |w|_V \quad \forall u, v, w \in V.$$

To obtain optimal regularity properties of the control-to-state mapping, we select real numbers q, q', s satisfying the following assumption

(A1) *The numbers q, q', $s \geq 1$ satisfy the following conditions:*
(i) The imbedding of $H_0^1(\Omega)$ in $L^{q'}(\Omega)$ is continuous.
(ii) The exponents q and q' are conjugate exponents, i.e. $1/q+1/q'=1$.
(iii) For all $u \in L^s(\Omega)^n$ it holds
$$|u|_q^2 \leq |u|_1 |u|_s.$$

Notice that condition (iii) implies $q \leq 2$. Here we have in mind two different situations. At first, $q = q' = 2$ and $s = \infty$ meet this assumption. Then the second-order sufficient condition of section 3.3 yields local optimality of the reference control in a $L^s = L^\infty$-neighbourhood. This means more or less that jumps of the optimal control have to be known a-priorily. To overcome this difficulty, we employ a second configuration, namely $q' = 4$, $q = 4/3$, $s = 2$, confer Lemma 3.1. Here we are able to work with a L^2-neighbourhood of the reference control.

The use of the L^s-neighbourhoods with $s < \infty$ is possible since the control u appears linearly in the equation and quadratically in the objective. Moreover, control and state are separated in the objective funtional.

Definition 3.4 (Weak solution). *Let $u \in L^q(\Omega)^n$ be given. A function $y \in V$ is called weak solution of* (2.2) *if it satisfies the variational equation*
$$\nu\,(y,v)_V + b(y,y,v) = (u,v)_{q,q'} \quad \forall v \in V. \tag{3.2}$$

Observe that $v \in V$ implies $v \in L^{q'}(\Omega)$. This fact permits us to work with controls that are less regular than $L^2(\Omega)^n$. Moreover, we recall $(-\nu\Delta y, v)_2 = \nu\,(y,v)_V$. It is known that (3.2) admits a unique solution y if the norm of the inhomogeneity u is sufficiently small or the coefficient ν is sufficiently large:

Theorem 3.5 (Existence and uniqueness of weak solutions). *For given $f \in V'$ the equation*
$$\nu\,(y,v)_V + b(y,y,v) = \langle f, v\rangle \quad \forall v \in V. \tag{3.3}$$
admits at least one solution $y \in V$. If the smallness condition
$$\nu^2 > C_n N_4^2 |f|_{V'}$$
is satisfied, then this solution is unique.

This is proven for instance in [25, Theorems II.1.2, II.1.3].
If the functional f in (3.3) is generated by a L^q-function u,
$$\langle f, v\rangle = (u,v)_{q,q'},$$
then we have to impose some restrictions on the $L^q(\Omega)$-norm of u. Let the $L^q(\Omega)$-norm of the admissible controls be bounded by $\mathcal{M}_q$, i.e.
$$\mathcal{M}_q = \sup_{u \in U_{ad}} |u|_q.$$
Then the following condition ensures existence and uniqueness of $y = y(u)$:

(A2) *The set of admissible controls U_{ad} is bounded in $L^q(\Omega)^n$. The bound $\mathcal{M}_q$ satifies, together with the viscosity parameter ν,*
$$\frac{2C_n N_{q'} N_4^2}{\nu^2}\mathcal{M}_q \leq 1. \tag{3.4}$$

In the sections dealing with the steady-state case we assume that these two Assumptions (A1) and (A2) are satisfied.

LEMMA 3.6. *For all $u \in U_{ad}$, the variational equality* (3.2) *admits a unique solution $y \in V$. If y_1, $y_2 \in V$ are weak solutions of* (3.2) *corresponding to u_1, $u_2 \in U_{ad}$, then*

$$|y_i|_V \le \frac{N_{q'}}{\nu}|u_i|_q \le \frac{N_{q'}}{\nu}\mathcal{M}_q \quad \text{and} \quad |y_1 - y_2|_V \le \frac{2N_{q'}}{\nu}|u_1 - u_2|_q,$$

i.e. the solution mapping $u \mapsto y$ is Lipschitz on U_{ad}.

Proof. Existence and uniqueness of solutions follow by Theorem 3.5 in view of assumption (3.4). Testing (3.2) with $v = y$ yields

$$\nu|y|_V^2 + b(y,y,y) = (u,y)_{q,q'} \le |u|_q|y|_{q'} \le \frac{\nu}{2}|y|_V^2 + \frac{N_{q'}^2}{2\nu}|u|_q^2$$

by the Young inequality. Since $b(y,y,y) = 0$ for all $y \in V$, the first estimate follows immediately. The second is obtained in the following way: We test the variational equalities for y_1 and y_2 by $y_1 - y_2 =: z$ and substract them to get

$$\nu|z|_V^2 + b(y_1,y_1,z) - b(y_2,y_2,z) = (u_1 - u_2, z)_{q,q'}.$$

Since $b(y_1,z,z) = 0$ because of $z, y_1 \in V$, we can write

$$b(y_1,y_1,z) - b(y_2,y_2,z) = b(y_1,y_2,z) - b(y_1,z,z) - b(y_2,y_2,z) = b(z,y_2,z).$$

Then we obtain in view of (3.4)

$$|b(y_1,y_1,z) - b(y_2,y_2,z)| = |b(z,y_2,z)| \le C_n|z|_4^2|y_2|_V \le C_n N_4^2|z|_V^2\frac{N_{q'}}{\nu}\mathcal{M}_q \le \frac{\nu}{2}|z|_V^2. \tag{3.5}$$

Finally, using Young's inequality again, we arrive at

$$|z|_V^2 \le \frac{4N_{q'}^2}{\nu^2}|u_1 - u_2|_q^2,$$

and the claim is proven. ∎

To derive first-order necessary optimality conditions, we also need estimates of solutions of linearized equations.

COROLLARY 3.7. *Let $\bar y \in V$ be the state corresponding to a control $\bar u \in U_{ad}$. Then for every $f \in V'$ there exists a unique solution $y \in V$ of the linearized equation*

$$\nu\,(y,v)_V + b(y,\bar y,v) + b(\bar y,y,v) = \langle f,v\rangle \quad \forall v \in V. \tag{3.6}$$

It holds

$$|y|_V \le \frac{2}{\nu}|f|_{V'}. \tag{3.7}$$

Proof. Existence can be argued as in the proof of [25, Theorem II.1.2]. Here it is necessary that $\bar y$ is the state associated to some control $\bar u \in U_{ad}$. In this case, we have some smallness property of $\bar y$ which ensures the solvability.

The a-priori estimate (3.7) can be shown along the lines of the previous proof. We multiply (3.6) by y to obtain

$$\nu|y|_V^2 = \langle f, y\rangle - b(y, \bar{y}, y).$$

The nonlinear term is treated as in the previous Lemma, confer (3.5). Therefore, it holds $|b(y,\bar{y},y)| \leq \frac{\nu}{2}|y|_V^2$. Here we used that $\bar{y}$ is the state associated with an admissible control, hence $|\bar{y}|_V \leq \frac{N_{q'}}{\nu}\mathcal{M}_q$ holds by Lemma 3.6. We end up with

$$\nu|y|_V^2 \leq \frac{1}{\nu}|f|_{V'}^2 + \frac{\nu}{4}|y|_V^2 + \frac{\nu}{2}|y|_V^2,$$

which gives the claim immediately. ∎

3.2. First order necessary optimality conditions. So far, we provided results concerning the properties of the state equation. Now, we concentrate on the aspects of optimization. We denote by $G(u) = y$ the solution operator $u \mapsto y$ of the steady-state Navier-Stokes equations (3.2).

LEMMA 3.8. *The solution operator $G : L^q(\Omega)^n \mapsto V$ is Fréchet-differentiable. In particular, G is Fréchet-differentiable from $L^2(\Omega)^n$ to V. The derivative $G'(u)$ is given by $G'(\bar{u})h = z$, where z is a weak solution of*

$$\nu\,(z, v)_V + b(z, \bar{y}, v) + b(\bar{y}, z, v) = (h, v)_2 \quad \forall v \in V, \tag{3.8}$$

with $\bar{u} \in U_{ad}$, $\bar{y} = G(\bar{u})$, $h \in L^2(\Omega)^n$.

Proof. Let $\bar{u}$, $h \in L^q(\Omega)^n$ be given. Denote by $\bar{y}$ the state associated with $\bar{u}$ and by y_h the one associated wiht h, hence $\bar{y} = G(\bar{u})$ and $y_h = G(\bar{u} + h)$, and the following variational equalities hold:

$$\begin{aligned}
\nu\,(\bar{y}, v)_V + b(\bar{y}, \bar{y}, v) &= (\bar{u}, v)_{q,q'}, \\
\nu\,(y_h, v)_V + b(y_h, y_h, v) &= (\bar{u} + h, v)_{q,q'} \qquad \forall v \in V.
\end{aligned}$$

Since

$$\begin{aligned}
b(\bar{y}, \bar{y}, v) - b(y_h, y_h, v) &= b(\bar{y} - y_h, \bar{y}, v) + b(y_h, \bar{y} - y_h, v) \\
&= b(\bar{y} - y_h, \bar{y}, v) + b(\bar{y}, \bar{y} - y_h, v) - b(\bar{y} - y_h, \bar{y} - y_h, v),
\end{aligned}$$

the difference $d := \bar{y} - y_h$ solves

$$\nu\,(d, v)_V + b(d, \bar{y}, v) + b(\bar{y}, d, v) = (h, v)_{q,q'} + b(\bar{y} - y_h, \bar{y} - y_h, v) \qquad \forall v \in V.$$

Next we split this difference into functions z and r, $d = z + r$, that solve the two linear equations

$$\begin{aligned}
\nu\,(z, v)_V + b(z, \bar{y}, v) + b(\bar{y}, z, v) &= (h, v)_{q,q'} \\
\nu\,(r, v)_V + b(r, \bar{y}, v) + b(\bar{y}, r, v) &= b(\bar{y} - y_h, \bar{y} - y_h, v) \qquad \forall v \in V.
\end{aligned}$$

Existence and uniqueness of z and r follow by Corollary 3.7. Let us denote the solution operator of these linear equations by $A(\bar{y})$, then $z = A(\bar{y})h$. Clearly, this operator is linear. Its boundedness is a consequence of Corollary 3.7. We arrive at

$$\bar{y} - y_h - z = G(\bar{u}) - G(\bar{u} + h) - A(\bar{u})h = r.$$

To prove Fréchet-differentiability of G, we have to estimate the norm of r. By subsequent application of Corollary 3.7, Lemma 3.3, and Lemma 3.6 we obtain

$$|r|_V \leq \frac{2}{\nu}|b(\bar{y}-y_h,\bar{y}-y_h,\cdot)|_{V'} \leq \frac{2}{\nu}C_n N_4^2|\bar{y}-y_h|_V^2 \leq \frac{2}{\nu}C_n N_4^2\left(\frac{2N_{q'}}{\nu}\right)^2|h|_q^2.$$

Then it follows $|r|_V/|h|_q \to 0$ as $|h|_q \to 0$. In this way, the Fréchet-differentiability of G is proven, and we can identify $G'(\bar{u}) := A(\bar{u})$. Since $q \leq 2$ by Assumption (A1), G is Fréchet-differentiable from $L^2(\Omega)^n$ to V. ∎

Before discussing the second-order sufficient optimality condition, we derive for convenience the standard first-order necessary optimality condition.

Definition 3.9 (Locally optimal control). *A control $\bar{u} \in U_{ad}$ is called locally optimal in $L^2(\Omega)^n$, if there exists a constant $\rho > 0$ such that*

$$J(\bar{y},\bar{u}) \leq J(y_h,u_h)$$

holds for all $u_h \in U_{ad}$ with $|\bar{u}-u_h|_2 \leq \rho$. Here, $\bar{y}$ and y_h denote the states associated to $\bar{u}$ and u_h, respectively.

Theorem 3.10 (First-order necessary condition). *Let $\bar{u}$ be a locally optimal control for (2.1) with associated state $\bar{y} = y(\bar{u})$. Then there exists a unique solution $\bar{\lambda} \in V$ of the adjoint equation*

$$\nu\,(\bar{\lambda},v)_V + b(\bar{y},v,\bar{\lambda}) + b(v,\bar{y},\bar{\lambda}) = (\bar{y}-y_d,v)_2 \quad \forall v \in V. \tag{3.9}$$

Moreover, the variational inequality

$$(\gamma\bar{u}+\bar{\lambda},u-\bar{u})_2 \geq 0 \quad \forall u \in U_{ad} \tag{3.10}$$

is satisfied.

Proof. The objective functional can be written as

$$\phi(u) = J(\tilde{G}(u),u) = \frac{1}{2}|\tilde{G}(u)-y_d|_2^2 + \frac{1}{2}|u|_2^2,$$

where $\tilde{G} : L^2(\Omega)^n \mapsto V$ stands for the solution operator G restricted to $L^2(\Omega)^n$. By Lemma 3.8, $\tilde{G}$ is also Frećhet-differentiable. The standard necessary condition for $\bar{u}$ to be a local optimum of $\phi(u)$ is $\phi'(u)(u-\bar{u}) \geq 0$ for all $u \in U_{ad}$, i.e.

$$\phi'(u)(u-\bar{u}) = (\tilde{G}(\bar{u})-y_d,\,\tilde{G}'(\bar{u})(u-\bar{u}))_2 + \gamma\,(\bar{u},u-\bar{u})_2 \geq 0 \quad \forall u \in U_{ad}. \tag{3.11}$$

We set $z := \tilde{G}'(\bar{u})(u-\bar{u})$, then z satisfies the linear equation (3.8). Let $\bar{\lambda}$ be the solution of (3.9). Its existence can be reasoned like in Corollary 3.7. Testing (3.8) by $\bar{\lambda}$, we get

$$\nu\,(z,\bar{\lambda})_V + b(z,\bar{y},\bar{\lambda}) + b(\bar{y},z,\bar{\lambda}) = (u-\bar{u},\bar{\lambda})_2. \tag{3.12}$$

Testing (3.9) by z yields

$$\nu\,(\bar{\lambda},z)_V + b(\bar{y},z,\bar{\lambda}) + b(z,\bar{y},\bar{\lambda}) = (\bar{y}-y_d,z)_2. \tag{3.13}$$

The left-hand sides in (3.12) and (3.13) are equal, so the right-hand sides are equal as well,

$$(u-\bar{u},\bar{\lambda})_2 = (\bar{y}-y_d,z)_2 = (\bar{y}-y_d,\,\tilde{G}'(\bar{u})(u-\bar{u}))_2.$$

Therefore, we obtain $\bar\lambda = \tilde G'(\bar u)^*(\bar y - y_d) = \tilde G'(\bar u)^*(\tilde G(\bar u) - y_d)$. The variational inequality now reads,

$$(\gamma\bar u + \bar\lambda, u - \bar u)_2 \geq 0 \quad \forall u \in U_{ad},$$

hence, the claim is proven. ∎

The solution $\bar\lambda$ of the adjoint equation (3.9) is said to be the adjoint state associated with $\bar y$. It can be easily verified that $\bar\lambda$ is a weak solution of the *adjoint partial differential equation*

$$\begin{aligned} -\nu\Delta\lambda - (\bar y\cdot\nabla)\lambda + (\nabla\bar y)^T\lambda + \nabla\mu &= \bar y - y_d && \text{on } \Omega,\\ \operatorname{div}\lambda &= 0 && \text{on } \Omega,\\ \lambda &= 0 && \text{on } \Gamma. \end{aligned}$$

The function μ might be interpreted as the adjoint pressure.

COROLLARY 3.11. *The adjoint state $\bar\lambda$, given by* (3.9)*, satisfies*

$$|\bar\lambda|_V \leq \frac{2}{\nu} N_2 |\bar y - y_d|_2.$$

Proof. Testing (3.9) by $\bar\lambda$, we get

$$\nu|\bar\lambda|_V^2 \leq N_2|\bar y - y_d|_2|\bar\lambda|_V + |b(\bar\lambda,\bar y,\bar\lambda)|.$$

The nonlinear term is estimated as in (3.5), which yields

$$|b(\bar\lambda,\bar y,\bar\lambda)| \leq C_n|\bar\lambda|_4^2|\bar y|_V \leq C_n N_4^2|\bar\lambda|_V^2\frac{N_{q'}}{\nu}\mathcal{M}_q \leq \frac{\nu}{2}|\bar\lambda|_V^2,$$

by (3.4). Now the claim follows by Young's inequality. ∎

To simplify notations we denote the pair (y,u) by v. It is called admissible, if u belongs to U_{ad} and y is the weak solution of (2.2) associated with u.

Let us introduce the Lagrange function $\mathcal{L} : V \times L^2(\Omega)^n \times V \mapsto \mathbb{R}$ for the optimal control problem as follows:

$$\mathcal{L}(y,u,\lambda) = J(u,y) - \nu\,(y,\lambda)_V - b(y,y,\lambda) + (u,\lambda)_2.$$

This function is twice Fréchet-differentiable with respect to u and y. The reader can readily verify that the necessary conditions can be expressed equivalently by

$$\mathcal{L}_y(\bar y,\bar u,\bar\lambda)\,y = 0 \quad \forall y \in V, \tag{3.14}$$

and

$$\mathcal{L}_u(\bar y,\bar u,\bar\lambda)(u-\bar u) \geq 0 \quad \forall u \in U_{ad}. \tag{3.15}$$

Here, $\mathcal{L}_y$, $\mathcal{L}_u$ denote the partial Fréchet-derivative of $\mathcal{L}$ with respect to y and u. The Fréchet-differentiability of $\mathcal{L}$ is shown in the next Lemma.

LEMMA 3.12. *The Lagrangian $\mathcal{L}$ is twice Fréchet-differentiable with respect to $v = (y,u)$ from $V\times L^2(\Omega)^n$ to $\mathbb{R}$. The second-order derivative at $\bar v = (\bar y,\bar u)$ fulfills together with the associated adjoint state $\bar\lambda$*

$$\mathcal{L}_{vv}(\bar v,\bar\lambda)[(z_1,h_1),(z_2,h_2)] = \mathcal{L}_{uu}(\bar v,\bar\lambda)[h_1,h_2] + \mathcal{L}_{yy}(\bar v,\bar\lambda)[z_1,z_2]$$

and

$$|\mathcal{L}_{yy}(\bar{v},\bar{\lambda})[z_1,z_2]| \le c_{\mathcal{L}}|z_1|_V|z_2|_V$$

for all $(z_i,h_i) \in V \times L^2(\Omega)^n$ *with some constant* $c_{\mathcal{L}} > 0$ *that does not depend on* $\bar{v},\bar{\lambda},z_1,z_2$.

Proof. The first-order derivatives of $\mathcal{L}$ with respect to y and u are

$$\begin{aligned}\mathcal{L}_y(\bar{v},\bar{\lambda})z &= (z,\bar{y}-y_d)_2 - \nu\,(z,\bar{\lambda})_V - b(z,\bar{y},\bar{\lambda}) - b(\bar{y},z,\bar{\lambda})\\ \mathcal{L}_u(\bar{v},\bar{\lambda})h &= \gamma\,(h,\bar{u})_2 + (u,\bar{\lambda})_2.\end{aligned}$$

The mappings $\bar{y} \mapsto \mathcal{L}_y(\bar{v},\bar{\lambda})$ and $\bar{u} \mapsto \mathcal{L}_u(\bar{v},\bar{\lambda})$ are affine linear. Their linear parts are bounded, hence continuous. Therefore, both mappings are Fréchet-differentiable. This shows that $\mathcal{L}$ is twice Fréchet-differentiable as well. The second-order derivative of $\mathcal{L}$ with respect to v is

$$\begin{aligned}\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[(z_1,h_1)(z_2,h_2)] &= \mathcal{L}_{uu}(\bar{v},\bar{\lambda})[h_1,h_2] + \mathcal{L}_{yy}(\bar{v},\bar{\lambda})[z_1,z_2]\\ &= \gamma\,(h_1,h_2)_2 + (z_1,z_2)_2 - b(z_1,z_2,\bar{\lambda}) - b(z_2,z_1,\bar{\lambda}),\end{aligned}$$

since mixed derivatives do not appear. Then we can estimate

$$\begin{aligned}|\mathcal{L}_{yy}(\bar{v},\bar{\lambda})[z_1,z_2]| &\le |z_1|_2|z_2|_2 + |b(z_1,z_2,\bar{\lambda})| + |b(z_2,z_1,\bar{\lambda})|\\ &\le |z_1|_2|z_2|_2 + 2C_n|z_1|_4|z_2|_4|\bar{\lambda}|_V\\ &\le c_{\mathcal{L}}|z_1|_V|z_2|_V.\end{aligned}$$

Here we used the estimates of b in Lemma 3.3 and the boundedness of the adjoint state, see Corollary 3.11. ∎

The Lagrangian has only non-zero derivatives up to order two. Derivatives of higher order vanish. Therefore, it holds

$$\mathcal{L}(y+z,u+h) = \mathcal{L}(y,u) + \mathcal{L}_y(y,u)z + \mathcal{L}_u(y,u)h + \frac{1}{2}\mathcal{L}_{yy}(y,u)[z,z] + \frac{1}{2}\mathcal{L}_{uu}(y,u)[h,h].$$

A remainder term does not appear. To shorten notations, we abbreviate $[v,v]$ by $[v]^2$, i.e.

$$\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[(z,h)]^2 := \mathcal{L}_{vv}(\bar{v},\bar{\lambda})[(z,h),(z,h)].$$

3.3. Second-order sufficient optimality condition. In the following, $\bar{v} = (\bar{y},\bar{u})$ is a fixed admissible reference pair. We suppose that the first-order necessary optimality conditions are fulfilled at $\bar{v}$.

Definition 3.13 (Strongly active sets). *For fixed* $\varepsilon > 0$ *and all* $i = 1,\ldots,n$ *we define sets* $\Omega_{\varepsilon,i}$ *by*

$$\Omega_{\varepsilon,i} = \{x \in \Omega :\ |\gamma\bar{u}_i(x) + \bar{\lambda}_i(x)| > \varepsilon\} \tag{3.16}$$

Here, $v_i(x)$ *denotes the value of the i-th component of a vector function* $v \in V$ *at* $x \in \Omega$. *Since* $\bar{u}$ *and* $\bar{\lambda}$ *are measurable functions, the sets* $\Omega_{\varepsilon,i}$ *are measurable, too. Moreover, for* $u \in L^p(\Omega)^n$ *and* $1 \le p < \infty$ *we define the* L^p*-norm with respect to the set of positivity by*

$$|u|_{L^p,\Omega_\varepsilon} := \left(\sum_{i=1}^n |u_i|^p_{L^p(\Omega_{\varepsilon,i})}\right)^{1/p}.$$

Notice that the variational inequality (3.10) uniquely determines $\bar{u}_i$ on $\Omega_{\varepsilon,i}$. If $\gamma\bar{u}_i(x) + \bar{\lambda}_i(x) \geq \varepsilon$ then $\bar{u}_i(x) = u_a(x)$ must hold. On the other hand, it follows $\bar{u}_i(x) = u_b(x)$, if $\gamma\bar{u}_i(x) + \bar{\lambda}_i(x) \leq -\varepsilon$ is satisfied.

COROLLARY 3.14. *It holds*

$$\sum_{i=1}^{n} \int_{\Omega_{\varepsilon,i}} (\gamma\bar{u}_i(x) + \bar{\lambda}_i(x))(u_i(x) - \bar{u}_i(x))dx \geq \varepsilon\,|u - \bar{u}|_{L^1,\Omega_\varepsilon}$$

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

Proof. From the variational inequality (3.10) we conclude the pointwise condition

$$(\gamma\,\bar{u}_i(x) + \bar{\lambda}_i(x))(u(x) - \bar{u}(x)) \geq 0$$

for almost all $x \in \Omega,\ i = 1, \ldots, n$. Therefore,

$$\begin{aligned}\int_{\Omega_{\varepsilon,i}} (\gamma\,\bar{u}_i(x) + \bar{\lambda}_i(x))(u_i(x) - \bar{u}_i(x))\,dx &= \int_{\Omega_{\varepsilon,i}} |\gamma\,\bar{u}_i(x) + \bar{\lambda}_i(x)|\ |(u_i(x) - \bar{u}_i(x)|\,dx \\ &\geq \varepsilon \int_{\Omega_{\varepsilon,i}} |(u_i(x) - \bar{u}_i(x)|\,dx \\ &= \varepsilon\,|u - \bar{u}|_{L^1(\Omega_{\varepsilon,i})}\end{aligned}$$

is satisfied. The claim follows by summing up this expression over $i = 1, \ldots, n$. ☐

We shall assume that the optimal pair $\bar{v} = (\bar{y}, \bar{u})$ and the associated adjoint state $\bar{\lambda}$ satisfy the following coercivity assumption on $\mathcal{L}''(\bar{v}, \bar{\lambda})$, henceforth called second-order sufficient optimality condition:

(SSC)
There exist $\varepsilon > 0$ and $\delta > 0$ such that

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z, h)]^2 \geq \delta\,|h|_q^2 \tag{3.17}$$

holds for all pairs $(z, h) \in V \times L^2(\Omega)^n$ with

$$h = u - \bar{u},\ u \in U_{ad},\ h_i = 0 \text{ on } \Omega_{\varepsilon,i} \text{ for } i = 1, \ldots, n,$$

and $z \in V$ being the weak solution of the linearized equation

$$\nu\,(z, w)_V + b(\bar{y}, z, w) + b(z, \bar{y}, w) = (h, w)_2 \quad \forall w \in V. \tag{3.18}$$

The parameter q is chosen according to Assumption (A1).

REMARK 3.15. *Notice that the definition of h implies $h(x) \geq 0$, where $\bar{u}(x) = u_a(x)$, and $h(x) \leq 0$, where $\bar{u}(x) = u_b(x)$. The condition $\varepsilon > 0$ can not be relaxed to $\varepsilon = 0$, see the counterexample in [15].*

Next we will prove that (SSC), together with the first-order necessary conditions, is sufficient for local optimality of $(\bar{y}, \bar{u})$.

THEOREM 3.16. *Let $\bar{v} = (\bar{y}, \bar{u})$ be admissible for the optimal control problem and suppose that $\bar{v}$ fulfills the first-order necessary optimality condition with associated adjoint state $\bar{\lambda}$. Assume further that (SSC) is satisfied at $\bar{v}$. Then there exist $\alpha > 0$ and $\rho > 0$ such that*

$$J(v) \geq J(\bar{v}) + \alpha\,|u - \bar{u}|_q^2$$

holds for all admissible pairs $v = (y, u)$ with $|u - \bar{u}|_s \le \rho$. The exponents s and q are chosen such that the Assumptions (A1) and (A2) are met.

Proof. Throughout the proof, c is used as a generic constant. Suppose that $\bar{v}$ fulfills the assumptions of the theorem. Let (y, u) be another admissible pair. We have

$$J(\bar{v}) = \mathcal{L}(\bar{v}, \bar{\lambda}) \quad \text{and} \quad J(v) = \mathcal{L}(v, \bar{\lambda}),$$

since $\bar{v}$ and v are admissible. Taylor-expansion of the Lagrange-function yields

$$\mathcal{L}(v, \bar{\lambda}) = \mathcal{L}(\bar{v}, \bar{\lambda}) + \mathcal{L}_y(\bar{v}, \bar{\lambda})(y - \bar{y}) + \mathcal{L}_u(\bar{v}, \bar{\lambda})(u - \bar{u}) + \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - \bar{v}, v - \bar{v}].$$

Notice that there is no remainder term due to the quadratic nature of the nonlinearities. Moreover, the necessary conditions (3.14), (3.15) are satisfied at $\bar{v}$ with adjoint state $\bar{\lambda}$. Therefore, the second term vanishes. The third term is nonnegative due to the variational inequality (3.15). However, we get even more by Corollary 3.14,

$$\begin{aligned}\mathcal{L}_u(\bar{v}, \bar{\lambda})(u - \bar{u}) &= \int_\Omega (\gamma \bar{u} + \bar{\lambda})(u - \bar{u})\, dx \\ &\ge \sum_{i=1}^n \int_{\Omega_{\varepsilon,i}} (\gamma \bar{u}_i + \bar{\lambda}_i)(u_i - \bar{u}_i)\, dx \ge \varepsilon\, |u - \bar{u}|_{L^1, \Omega_\varepsilon},\end{aligned}$$

confer the Definition 3.13 of $|\cdot|_{L^1,\Omega_\varepsilon}$. So we arrive at

$$\begin{aligned}J(v) &= J(\bar{v}) + \mathcal{L}_y(\bar{v}, \bar{\lambda})(y - \bar{y}) + \mathcal{L}_u(\bar{v}, \bar{\lambda})(u - \bar{u}) + \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - \bar{v}]^2 \\ &\ge J(\bar{v}) + \varepsilon\, |u - \bar{u}|_{L^1,\Omega_\varepsilon} + \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - \bar{v}]^2.\end{aligned} \tag{3.19}$$

Next, we investigate the second derivative of $\mathcal{L}$. Here, we invoke assumption (3.17) on the coercitivity of $\mathcal{L}_{vv}$ on a certain subspace. To do so, we introduce a new admissible control $\tilde{u} \in L^s(\Omega)^n$ by

$$\tilde{u}_i(x) = \begin{cases} \bar{u}_i(x) & \text{on } \Omega_{\varepsilon,i} \\ u_i(x) & \text{on } \Omega \setminus \Omega_{\varepsilon,i} \end{cases} \quad \text{for } i = 1, \ldots, n. \tag{3.20}$$

Then, we have $u - \bar{u} = (u - \tilde{u}) + (\tilde{u} - \bar{u})$, where $(u - \tilde{u})_i = 0$ on $\Omega \setminus \Omega_{\varepsilon,i}$ and $(\tilde{u} - \bar{u})_i = 0$ on $\Omega_{\varepsilon,i}$, so that $h := \tilde{u} - \bar{u}$ fits in the assumptions of (SSC). The difference $z := y - \bar{y}$ solves the equation

$$\nu\,(z, w)_V + b(z, \bar{y}, w) + b(\bar{y}, z, w) = (u - \bar{u}, w)_{q,q'} - b(y - \bar{y}, y - \bar{y}, w) \quad \forall w \in V.$$

We split $z = y - \bar{y}$ into $y_h + y_r$, where y_r and y_h solve the equations

$$\nu\,(y_h, w)_V + b(y_h, \bar{y}, w) + b(\bar{y}, y_h, w) = (h, w)_{q,q'} \quad \forall w \in V \tag{3.21}$$

and

$$\nu\,(y_r, w)_V + b(y_r, \bar{y}, w) + b(\bar{y}, y_r, w) = (u - \tilde{u}, w)_{q,q'} - b(y - \bar{y}, y - \bar{y}, w) \quad \forall w \in V. \tag{3.22}$$

Notice that (3.21) is linear and that (y_h, h) belongs to the subspace where (SSC) applies. The norm of these auxiliary states has to be estimated. Using Lemma 3.6 we obtain

$$|y_h|_V \le c\,|h|_q \le c\,(|\tilde{u}-u|_q + |u-\bar{u}|_q)\,. \tag{3.23}$$

To estimate $|y_r|_V$, we have to investigate the V'-norm of the right-hand side in (3.22), which defines a linear continuous functional on V. By Lemma 3.3 we find

$$|b(y-\bar{y}, y-\bar{y}, \cdot)|_{V'} \le C_n N_4^2 |y-\bar{y}|_V^2.$$

Now we apply Corollary 3.7 and get

$$|y_r|_V \le \frac{2}{\nu}\left(C_n N_4^2 |y-\bar{y}|_V^2 + N_{q'}|u-\tilde{u}|_q\right) \le c\left(|u-\bar{u}|_q^2 + |u-\tilde{u}|_q\right). \tag{3.24}$$

Denote the pair (y_h, h) by v_h. This pair fits in the assumptions of the theorem.We continue the investigation of the Lagrangian by

$$\begin{aligned}\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[v-\bar{v}]^2 &= \mathcal{L}_{uu}(\bar{v},\bar{\lambda})[u-\tilde{u}+h]^2 + \mathcal{L}_{yy}(\bar{v},\bar{\lambda})[y_h+y_r]^2\\ &= \mathcal{L}_{vv}(\bar{v},\bar{\lambda})[v_h]^2 + 2\mathcal{L}_{uu}(\bar{v},\bar{\lambda})[u-\tilde{u},h]\\ &\quad + \mathcal{L}_{uu}(\bar{v},\bar{\lambda})[u-\tilde{u}]^2 + 2\mathcal{L}_{yy}(\bar{v},\bar{\lambda})[y_r,y_h] + \mathcal{L}_{yy}(\bar{v},\bar{\lambda})[y_r]^2.\end{aligned} \tag{3.25}$$

(SSC) applies to the first term $\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[v_h]^2$. The second-order derivative with respect to u satisfies

$$2\mathcal{L}_{uu}(\bar{v},\bar{\lambda})[u-\tilde{u},h] + \mathcal{L}_{uu}(\bar{v},\bar{\lambda})[u-\tilde{u}]^2 = 2\gamma\,(u-\tilde{u},h)_2 + \gamma\,|u-\tilde{u}|_2^2.$$

By definition of $\tilde{u}$ we know that $(u-\tilde{u})_i$ vanishes on $\Omega\setminus\Omega_{\varepsilon,i}$ whereas $(h)_i$ vanishes on $\Omega_{\varepsilon,i}$. So their scalar product is zero. Therefore, it holds

$$2\mathcal{L}_{uu}(\bar{v},\bar{\lambda})[u-\tilde{u},h] + \mathcal{L}_{uu}(\bar{v},\bar{\lambda})[u-\tilde{u}]^2 = \gamma\,|u-\bar{u}|_{L^2,\Omega_\varepsilon}^2 \ge 0. \tag{3.26}$$

The remaining terms in (3.25) are treated by Lemma 3.12 and the estimates (3.23), (3.24),

$$\begin{aligned}&\left|2\mathcal{L}_{yy}(\bar{v},\bar{\lambda})[y_r,y_h] + \mathcal{L}_{yy}(\bar{v},\bar{\lambda})[y_r,y_r]\right| \le c\,\left(|y_h|_V|y_r|_V + |y_r|_V^2\right)\\ &\qquad \le c\left\{\left(|\tilde{u}-u|_q + |u-\bar{u}|_q\right)\left(|u-\bar{u}|_q^2 + |u-\tilde{u}|_q\right) + |u-\bar{u}|_q^4 + |u-\tilde{u}|_q^2\right\}.\end{aligned} \tag{3.27}$$

Now we can proceed with the investigation of $\mathcal{L}_{vv}$ in (3.25). Invoking (3.26) and (3.27), we obtain from (3.25)

$$\begin{aligned}\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[v-\bar{v}]^2 \ge \delta\,|h|_q^2 - c\,\big\{&\left(|\tilde{u}-u|_q + |u-\bar{u}|_q\right)\left(|u-\bar{u}|_q^2 + |u-\tilde{u}|_q\right)\\ &+ |u-\bar{u}|_q^4 + |u-\tilde{u}|_q^2\big\}.\end{aligned} \tag{3.28}$$

Our next aim is to eliminate h such that only terms containing $u-\bar{u}$ and $\tilde{u}-u$ appear. To achieve this goal, we first notice that

$$|u-\bar{u}|_q^2 = |u-\tilde{u}+h|_q^2 \le 2\left(|u-\tilde{u}|_q^2 + |h|_q^2\right),$$

which immediately gives

$$|h|_q^2 \ge \frac{1}{2}|u-\bar{u}|_q^2 - |u-\tilde{u}|_q^2.$$

Applying Young's inequality several times to separate the powers of $|u-\bar{u}|_q$ and $|\tilde{u}-u|_q$ we get from (3.28)

$$\begin{aligned}
\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[v-\bar{v}]^2 &\geq \frac{\delta}{2}|u-\bar{u}|_q^2 - c\big\{|u-\bar{u}|_q^4 + |u-\bar{u}|_q^3 + |u-\bar{u}|_q^2|\tilde{u}-u|_q \\
&\quad + |u-\bar{u}|_q|\tilde{u}-u|_q + |\tilde{u}-u|_q^2\big\} \\
&\geq \frac{\delta}{4}|u-\bar{u}|_q^2 - c\big\{|u-\bar{u}|_q^4 + |u-\bar{u}|_q^3 + |\tilde{u}-u|_q^2\big\} \\
&\geq |u-\bar{u}|_q^2\left(\frac{\delta}{4} - c\big\{|u-\bar{u}|_q^2 + |u-\bar{u}|_q\big\}\right) - c|\tilde{u}-u|_q^2,
\end{aligned}$$

If u is sufficiently close to $\bar{u}$, i.e. $|u-\bar{u}|_q \leq N_{s,q}\,|u-\bar{u}|_s \leq N_{s,q}\,\rho_1$, then the term in brackets is greater than $\delta/8$. Hence we arrive at

$$\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[v-\bar{v}]^2 \geq \frac{\delta}{8}\,|u-\bar{u}|_q^2 - c\,|\tilde{u}-u|_q^2.$$

Now we are able to complete the estimation of the objective functional. We continue (3.19) by

$$\begin{aligned}
J(v) &\geq J(\bar{v}) + \varepsilon\,|u-\bar{u}|_{L^1,\Omega_\varepsilon} + \frac{1}{2}\mathcal{L}_{vv}(\bar{v},\bar{\lambda})[v-\bar{v}]^2 \\
&\geq J(\bar{v}) + \varepsilon\,|u-\bar{u}|_{L^1,\Omega_\varepsilon} + \frac{\delta}{16}|u-\bar{u}|_q^2 - \frac{c}{2}|\tilde{u}-u|_q^2.
\end{aligned}$$

By definition, $\tilde{u}$ and u differ only on the sets $\Omega\setminus\Omega_{i,\varepsilon}$, while $\tilde{u}$ and $\bar{u}$ coincide on $\Omega\setminus\Omega_{i,\varepsilon}$, hence we conclude using Lemma 3.1,

$$|\tilde{u}-u|_q^2 \leq |\tilde{u}-u|_1|\tilde{u}-u|_s = |u-\bar{u}|_{L^1,\Omega_\varepsilon}|u-\bar{u}|_s \leq \rho_2|u-\bar{u}|_{L^1,\Omega_\varepsilon},$$

if the L^s-norm of the difference is sufficiently small, i.e. $|u-\bar{u}|_s \leq \rho_2$. Hence,

$$\begin{aligned}
J(v) &\geq J(\bar{v}) + \varepsilon\,|u-\bar{u}|_{L^1,\Omega_\varepsilon} + \frac{\delta}{16}\,|u-\bar{u}|_q^2 - \frac{c}{2}\,|\tilde{u}-u|_q^2 \\
&\geq J(\bar{v}) + \left(\varepsilon - \frac{c}{2}\rho_2\right)|u-\bar{u}|_{L^1,\Omega_\varepsilon} + \frac{\delta}{16}\,|u-\bar{u}|_q^2.
\end{aligned}$$

Choosing ρ_2 so small that $\varepsilon - \frac{c}{2}\rho_2 > 0$, we prove the claim with $\alpha = \delta/16$ and $\rho = \min(\rho_1,\rho_2)$. ∎

The next result is a immediate conclusion.

THEOREM 3.17. *Suppose that the assumptions of Theorem 3.16 hold true. Then $\bar{u}$ is a locally optimal control in the sense of $L^s(\Omega)^n$.*

REMARK 3.18. *The second-order sufficient optimality condition can be adapted to general objective functionals following [10]. However, then one obtains differentiability of the functional J with respect to control and state only in L^∞-type spaces. Consequently, one has to work with L^∞-neighborhoods of the reference control.*

3.4. An equivalent formulation of second-order sufficient optimality conditions. Here, we comment on other formulations of second-order suficient conditions known from literature, [4, 5, 9]. Let us consider the sufficient optimality condition (SSC) with parameters $q = q' = 2$. We assume in this section that the Assumptions (A1) and (A2) are satisfied. Let us recall (SSC) for $q = 2$ for convenience:

(SSC) *There exist $\varepsilon > 0$ and $\delta > 0$ such that*

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z,h)]^2 \geq \delta\, |h|_2^2 \tag{3.29}$$

holds for all pairs $(z,h) \in V \times L^2(\Omega)^n$ with

$$h = u - \bar{u},\ u \in U_{ad},\ h_i = 0 \text{ on } \Omega_{\varepsilon,i} \text{ for } i = 1, \dots, n, \tag{3.30}$$

and $z \in V$ being the weak solution of the homogeneous and linearized equation

$$\nu\, (z,w)_V + b(\bar{y}, z, w) + b(z, \bar{y}, w) = (h,w)_2 \quad \forall w \in V. \tag{3.31}$$

We prove that (SSC) is equivalent to another formulation, introduced first by Bonnans [4, 5]. The tangent cone on U_{ad} at $\bar{u}$, denoted by $T(\bar{u})$, is defined by

$$T(\bar{u}) = \left\{ h \in L^2(\Omega)^n \,\middle|\, h = \lim_{k\to\infty} \frac{u_k - \bar{u}}{t_k},\ u_k \in U_{ad},\ t_k \downarrow 0 \right\}.$$

$T(\bar{u})$ is convex, non-empty and closed in $L^2(\Omega)^n$, hence also weakly closed. By $T(\bar{u})$, we are able to formulate (SSC) in the following way:

(SSC$_0$) *It holds*

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z,h)]^2 > 0 \tag{3.32}$$

for all pairs $(z,h) \in V \times L^2(\Omega)^n$ with $h \not\equiv 0$, $h \in T(\bar{u})$, and

$$h_i = 0 \text{ on } \Omega_{0,i}$$

for all $i = 1, \dots, n$, where z is the solution of the associated linearized equation (3.31).

Notice that $\Omega_{0,i} = \{x \in \Omega : |\gamma \bar{u}_i(x) + \bar{\lambda}_i(x)| > 0\}$.

THEOREM 3.19. *The conditions (SSC) and (SSC$_0$) are equivalent.*

Proof. It is easy to see that (SSC) implies (SSC$_0$). Let $0 \neq h \in T(\bar{u})$ with $h_i = 0$ a.e. on $\Omega_{0,i}$. Since $\Omega_{\varepsilon,i} \subset \Omega_{0,i}$, it holds $h_i = 0$ on $\Omega_{\varepsilon,i}$. Further, there exists a sequence $h_k = (\bar{u} - u_k)/t_k$ converging to h in $L^2(\Omega)^n$. After extracting a subsequence if necessary, we find that $\bar{u}_i(x) - u_{k,i}(x) \to 0$ a.e. on $\Omega_{0,i}$. Hence, we can choose u_k such that $u_{k,i}(x) = \bar{u}_i(x)$ on $\Omega_{0,i}$. This implies $h_{k,i} = 0$ on $\Omega_{0,i}$, and h_k can be used as test function in (SSC). Let z, z_k be the associated solutions of the linearized equation and $v := (z,h)$, $v_k := (z_k, h_k)$. Then it holds

$$\begin{aligned}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v]^2 &= \mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - v_k + v_k]^2 \\ &= \mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v_k]^2 + 2\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - v_k, v_k] + L_{vv}(\bar{v}, \bar{\lambda})[v - v_k]^2.\end{aligned}$$

Using assumption (3.29), estimates of $\mathcal{L}_{vv}$ in Lemma 3.12, and Corollary 3.7, we obtain

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v]^2 \geq \delta |h_k|_2^2 - c_1 |h - h_k|_2 |h_k|_2 - c_2 |h - h|_2^2 \geq \frac{\delta}{2} |h_k|_2^2 - c|h - h_k|_2^2,$$

which gives in the limit $k \to \infty$

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v]^2 \geq \frac{\delta}{2}|h|_2^2 > 0,$$

and (SSC_0) is satisfied.

Let us prove the converse direction. Assume, that (SSC_0) holds true but not (SSC). Then for all $\varepsilon > 0$ and $\delta > 0$ there exists $h_{\delta,\varepsilon} \in L^2(\Omega)^n$ such that $(h_{\delta,\varepsilon})_i = 0$ on $\Omega_{\varepsilon,i}$, $h_{\delta,\varepsilon} = u_{\delta,\varepsilon} - \bar{u}$, $u_{\delta,\varepsilon} \in U_{ad}$, and

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z_{\delta,\varepsilon}, h_{\delta,\varepsilon})]^2 < \delta\,|h_{\delta,\varepsilon}|_2^2$$

is fulfilled with associated $z_{\delta,\varepsilon}$. Multiplying $h_{\delta,\varepsilon}$ by some positive constant, we can assume $|h_{\delta,\varepsilon}|_2 = 1$ and $h_{\delta,\varepsilon} \in T(\bar{u})$. Choosing $\delta_k = \varepsilon_k = 1/k$, $h_k := h_{\delta_k,\varepsilon_k}$, we find

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z_k, h_k)]^2 < \frac{1}{k},$$

where z_k is the weak solution of (3.31), hence

$$\limsup_k \mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z_k, h_k)]^2 \leq 0.$$

Since the set $\{h_k\}_{k=1}^\infty$ is bounded in $L^2(\Omega)^n$, there exists an element $\tilde{h} \in L^2(\Omega)^n$, such that, after extracting a subsequence if necessary, the h_k converge weakly in $L^2(\Omega)^n$ to $\tilde{h}$. The tangent cone $T(\bar{u})$ is weakly closed, therefore $\tilde{h} \in T(\bar{u})$.

Next, we want to show $h_{k,i}(x) \to 0$ a.e. pointwise on $\Omega_{0,i}$. Let $x_0 \in \Omega_{0,i}$ be given. Then it holds $|\gamma\bar{u}_i(x_0) + \bar{\lambda}_i(x_0)| = \tau' > 0$, which implies by definition $x_0 \in \Omega_{\tau,i}$ for all $0 \leq \tau \leq \tau'$. Hence, there exists an index $k_i(x_0)$ such that $x_0 \in \Omega_{\varepsilon_k,i} = \Omega_{1/k,i}$ for all $k > k_i(x_0)$. By construction of h_k we conclude $h_{k,i}(x_0) = 0$ for all $k > k_i(x_0)$. It follows $\tilde{h}(x) = 0$ on Ω_0 almost everywhere.

We decompose $\mathcal{L}_{vv}$ and use $|h_k|_2 = 1$ to get

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z_k, h_k)]^2 = \gamma|h_k|_2^2 + Q(z_k) = \gamma + Q(z_k), \tag{3.33}$$

with $Q(z) = |z|_2^2 - 2b(z, z, \bar{\lambda})$. The solution mapping $h \mapsto z$ associated with (3.31) is linear and continuous from $L^2(\Omega)^n$ to V. Thus, we obtain $z_k \rightharpoonup \tilde{z}$ in V and $z_k \to \tilde{z}$ in H, since V is compactly imbedded in H. A well-known result of Temam [25, Lemma II.1.5] yields $b(z_k, z_k, \bar{\lambda}) \to b(\tilde{z}, \tilde{z}, \bar{\lambda})$. We conclude $\lim_{k\to\infty} Q(z_k) = Q(\tilde{z})$. Passing to the limit in (3.33), we get

$$Q(\tilde{z}) \leq \limsup_k \mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z_k, h_k)]^2 - \gamma \leq -\gamma < 0,$$

which proves that $\tilde{h}$ cannot vanish, remember $Q(0) = 0$. Finally,

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(\tilde{z}, \tilde{h})]^2 = \gamma|\tilde{h}|_2^2 + Q(\tilde{z}) \leq \gamma - \gamma \leq 0$$

is obtained, which contradicts (SSC_0). $\square$

Another second-order sufficient optimality condition introduced by Casas and Mateos [9] involves the Hamiltonian of the optimal control problem. Due to the special form of our objective functional, this formulation is equivalent to (SSC_0).

Following the lines of Bonnans [4], we can prove that (SSC_0) even implies a L^2-growth condition in a L^2-neighborhood around $(\bar{y}, \bar{u})$.

THEOREM 3.20. *Let $\bar{v} = (\bar{y}, \bar{u})$ be admissible for the optimal control problem and suppose that $\bar{v}$ fulfills the first-order necessary optimality condition with associated adjoint state $\bar{\lambda}$. Assume further that (SSC_0) is satisfied at $\bar{v}$. Then there exist $\alpha > 0$ and $\rho > 0$ such that*

$$J(v) \geq J(\bar{v}) + \alpha \, |u - \bar{u}|_2^2 \tag{3.34}$$

holds for all admissible pairs $v = (y, u)$ with $|u - \bar{u}|_2 \leq \rho$.

Proof. Let us suppose that (SSC_0) is satisfied, whereas (3.34) does not hold. Then for all $\alpha > 0$ and $\rho > 0$ there exists $u_{\alpha,\rho} \in U_{ad}$ with $|u_{\alpha,\rho} - \bar{u}|_2 \leq \rho$ and

$$J(v_{\alpha,\rho}) < J(\bar{v}) + \alpha \, |u_{\alpha,\rho} - \bar{u}|_2^2, \tag{3.35}$$

where $v_{\alpha,\rho} = (u_{\alpha,\rho}, y_{\alpha,\rho})$ and $y_{\alpha,\rho}$ is the solution of (3.2) associated with $u_{\alpha,\rho}$. We choose $\alpha_k = \rho_k = 1/k$ and $u_k = u_{\alpha_k,\rho_k}$, $y_k = y_{\alpha_k,\rho_k}$.

By construction, it follows $u_k \to \bar{u}$ in $L^2(\Omega)^n$ as $k \to \infty$. Hence, we can write $u_k = \bar{u} + t_k h_k$, $|h_k|_2 = 1$ and $t_k \to 0$ as $k \to \infty$. Because the set of these h_k is bounded in $L^2(\Omega)^n$ we can extract a subsequence denoted again by (h_k) converging weakly to $\tilde{h} \in T(\bar{u}) \subset L^2(\Omega)^n$. In the following, let z_k be the solution of (3.31) associated with h_k.

Since $(\bar{u}, \bar{y})$ and (u_k, y_k) satisfy the state equation, it holds $\mathcal{L}(\bar{v}, \bar{\lambda}) = J(\bar{v})$ and $\mathcal{L}(v_k, \bar{\lambda}) = J(v_k)$. Then we obtain

$$J(v_k) = \mathcal{L}(v_k, \bar{\lambda}) = \mathcal{L}(\bar{v}, \bar{\lambda}) + t_k \mathcal{L}_u(\bar{v}, \bar{\lambda}) h_k + t_k \mathcal{L}_y(\bar{v}, \bar{\lambda}) z_k + t_k^2 \mathcal{L}_{vv}[(z_k, h_k)^2]. \tag{3.36}$$

The first-order necessary conditions (3.14), (3.15) are fulfilled, so we find $\mathcal{L}_y(\bar{v}, \bar{\lambda}) z_k = 0$ and $\mathcal{L}_u(\bar{v}, \bar{\lambda}) h_k \geq 0$. At first, we show $\tilde{h} = 0$ a.e. on Ω_0. We derive from (3.35) and (3.36)

$$0 \leq \mathcal{L}_u(\bar{v}, \bar{\lambda}) h_k = \frac{1}{t_k} \left(J(v_k) - J(\bar{v}) \right) - t_k \mathcal{L}_{vv}[(z_k, h_k)^2] < t_k \left\{ \frac{1}{k} - \mathcal{L}_{vv}[(z_k, h_k)^2] \right\}, \tag{3.37}$$

which gives $\mathcal{L}_u(\bar{v}, \bar{\lambda}) \tilde{h} = 0$ since $\mathcal{L}_{vv}[(z_k, h_k)^2]$ is bounded. The variational inequality

$$(\gamma \bar{u}_i(x) + \bar{\lambda}_i(x)) h_{k,i}(x) \geq 0$$

holds a.e. on Ω, $i = 1, \ldots, n$, so the weak limit $\tilde{h}_i(x)$ satisfies

$$(\gamma \bar{u}_i(x) + \bar{\lambda}_i(x)) \tilde{h}(x) \geq 0$$

as well. This, together with $\mathcal{L}_u(\bar{v}, \bar{\lambda}) \tilde{h} = 0$, yields $\tilde{h}(x) = 0$ on Ω_0, cf. the definition of Ω_0.

Finally, we show that (3.35) contradicts (SSC_0). Obviously (3.37) implies

$$\mathcal{L}_{vv}[(z_k, h_k)^2] < \frac{1}{k}.$$

Arguing as in the proof of the previous Theorem 3.19, we find that $\tilde{h}$ satisfies

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(\tilde{z}, \tilde{h})]^2 \leq 0,$$

with $h \not\equiv 0$. Since $\tilde{h}$ is admissible as test function in (SSC_0), this shows that the positivity assumption of (SSC_0) is violated. ∎

REMARK 3.21. *Observe that this theorem overcomes the two-norm discrepancy typically appearing in optimal control of semilinear equations. This is due to the very special form of the quadratic cost functional* (2.1)*, the linear appearance of the control u in the state equation, and the differentiability of the nonlinearity of the Navier-Stokes equations and the associated solution operator G in weaker than L^∞-norms.*

REMARK 3.22. *Casas and Mateos [9] require positivity of $\mathcal{L}_{vv}$ for increments vanishing on $\Omega \setminus \Omega_0$ together with uniform positivity of the second derivative of the Hamiltonian with respect to the control on $\Omega \setminus \Omega_\tau$ for some $\tau > 0$. The last property is fulfilled for our optimal control problem. The Hamiltonian is given by*

$$H(x, y, u, \lambda) = \frac{1}{2}|y - y_d(x)|^2 + \frac{\gamma}{2}|u|^2 + \lambda \cdot u.$$

Its second derivative with respect to u is

$$\frac{\partial^2 H}{\partial u^2}(x, y, u, \lambda) = \gamma,$$

which is uniform positive on Ω. Therefore, we are able to work with active sets $\Omega_{0,i}$ in (SSC_0).

In Dunn's counterexample [15], the second derivative of the Hamiltonian with respect to the control is nonnegative on $\Omega \setminus \Omega_0$ but indefinite on $\Omega \setminus \Omega_\tau$ for every $\tau > 0$. Hence, the use of the active set Ω_0 in (SSC) causes a contradiction.

4. The instationary case. In this secion, we consider in a very similar way the optimal control problem (2.3)–(2.4) for the instationary Navier-Stokes equations. The similarity of arguments will permit to shorten the presentation.

4.1. Notations and preliminary results. Here, we will restrict ourselve to the two-dimensional case, $n = 2$, since a satisfactory theory of the instationary Navier-Stokes equations is only available for this space dimension. In the two-dimensional case, a unique weak solution of (2.4) exists that depends continuously on the given data. First, we introduce some notations and provide some results that we need later on.

To begin with, we define the solenoidal space

$$H := \{v \in L^2(\Omega)^2 : \operatorname{div} v = 0\}.$$

Endowed with the usual L^2-scalar product, denoted by $(\cdot,\cdot)_H$, this space is a Hilbert space. The associated norm is denoted by $|\cdot|_H$. We shall work in the standard spaces of abstract functions from $[0, T]$ to a real Banach space X, $L^p(0, T; X)$ and $C([0, T]; X)$, endowed with their natural norms,

$$\|y\|_{L^p(X)} := \|y\|_{L^p(0,T;X)} = \left(\int_0^T |y(t)|_X^p \, dt \right)^{1/p},$$
$$\|y\|_{C([0,T];X)} := \max_{t \in [0,T]} |y(t)|_X,$$

$1 \le p < \infty$. To deal with the time derivative in (2.4), we introduce the following spaces of functions y whose time derivative y_t exists as abstract function,

$$W^\alpha(0,T;V) := \{y \in L^2(0,T;V) :\ y_t \in L^\alpha(0,T;V')\},$$

where $1 \le \alpha < \infty$. Moreover, we write for convenience

$$W(0,T) := W^2(0,T;V).$$

Endowed with the norm

$$\|y\|_{W^\alpha} := \|y\|_{W^\alpha(0,T;V)} = \|y\|_{L^2(V)} + \|y_t\|_{L^\alpha(V')},$$

these spaces are Banach spaces, respectively Hilbert spaces in the case of $W(0,T)$. In the sequel, we will use for $u \in L^p(Q)^2$ the notation

$$\|u\|_p := |u|_{L^p(Q)^2}.$$

In all what follows, $\|\cdot\|$ stands for norms of abstract functions, while $|\cdot|$ denotes norms of "stationary" spaces like H and V.

Corollary 4.1. *Let $v \in V$ and $y \in W^\alpha(0,T;V)$ be given. It holds*

$$|v|_4 \le 2^{1/4}|v|_H^{1/2}|v|_V^{1/2}. \tag{4.1}$$

If $\alpha > 1$ then y is, up to changes on sets of zero measure, equivalent to a function of $C([0,T],H)$, and there is a constant $c > 0$ such that

$$\|y\|_4 + \|y\|_{C([0,T],H)} \le c\,\|y\|_{W(0,T)}. \tag{4.2}$$

Proof. The first claim is proven in [25, Lemma III.3.3]. Note that $W^\alpha(0,T;V)$ for $\alpha > 1$ is continuously imbedded in $C([0,T],H)$, cf. [13, p. 483]. The L^4-claim follows from integrating (4.1) over $[0,T]$,

$$\begin{aligned}\|y\|_4^4 = \int_0^T |y(t)|_4^4 dt &\le 2\int_0^T |y(t)|_H^2|y(t)|_V^2 dt \le 2\,\|y\|_{C([0,T],H)}^2 \int_0^T |y(t)|_V^2 dt \\ &\le 2\,\|y\|_{C([0,T],H)}^2\|y\|_{L^2(V)}^2 \le c\,\|y\|_{W(0,T)}^4.\end{aligned}$$

□

In view of inequality (4.1), we can state another estimate of the trilinear form b. In the two-dimensional case it holds

$$|b(u,v,w)| \le \sqrt{2}|u|_H^{1/2}|u|_V^{1/2}|v|_V|w|_H^{1/2}|w|_V^{1/2} \tag{4.3}$$

for all $u,v,w \in V$. This follows directly from the estimate given in (3.1) and the previous corollary.

To specify the problem setting, we introduce a linear operator $A : L^2(0,T;V) \mapsto L^2(0,T;V')$ by

$$\int_0^T \langle (Ay)(t),\, v(t)\rangle_{V',V} dt := \int_0^T (y(t),v(t))_V dt,$$

and a nonlinear operator $B : L^2(0,T;V) \mapsto L^1(0,T;V')$ by

$$\int_0^T \big\langle \big(B(y)\big)(t),\, w(t)\big\rangle_{V',V}\, dt := \int_0^T b(y(t), y(t), w(t))dt,$$

where y, $v \in L^2(0,T;V)$ and $w \in L^\infty(0,T;V)$, respectively.

We need a bound on the admissible controls to establish a Lipschitz estimate of solutions of (2.4). Without loss of generality, we assume in the sequel that the set U_{ad} is bounded in $L^2(Q)^2$, i.e. there exists a constant $\mathcal{M} > 0$ such that

$$\sup_{u \in U_{ad}} \|u\|_2 \leq \mathcal{M}. \tag{4.4}$$

If this assumption is violated then we can introduce an artificial bound. For that purpose, let $\bar{u}$ be the optimal control and $\bar{J} = J(\bar{u})$ the corresponding value of the objective. Then $\bar{u}$ is also optimal for the same optimal control problem with changed set of admissible controls $\tilde{U}_{ad} = U_{ad} \cap \{u \in L^2(Q)^2 : \|u\|_2 \leq 2(\bar{J}+1)/\gamma\}$.

As in the stationary case, we want to derive a sufficient optimality condition that ensures local optimality of the reference control not only in $L^\infty(Q)^2$ but also in $L^s(Q)^2$, with some $s < \infty$. It remains to specify the exponent s.

(A3) *Let q, q', s be real numbers such that the following statements are true.*
- *(i)* $q' \leq 4$.
- *(ii) The exponents q and q' are conjugate exponents, i.e.* $\frac{1}{q} + \frac{1}{q'} = 1$.
- *(iii) For all $u \in L^s(Q)^2$ it holds*

$$\|u\|_q^2 \leq \|u\|_1 \|u\|_s.$$

Here we find that the two triplets $(q,q',s) = (4/3, 4, 2)$ and $(q,q',s) = (2,2,\infty)$ fulfill this assumptions as they did in the stationary case. The assumption (i) is needed to obtain by Lemma 3.2

$$\|y\|_{q'} \leq (\operatorname{vol} Q)^{1/q' - 1/4} \|y\|_4 \tag{4.5}$$

for all $y \in W(0,T)$. In the rest of this section we assume that Assumption (A3) is satisfied.

4.2. The state equation. We begin with the notation of weak solutions for the instationary Navier-Stokes equations (2.4)

DEFINITION 4.2 (Weak solution). *Let $f \in L^2(0,T;V')$ and $y_0 \in H$ be given. A function $y \in L^2(0,T;V)$ with $y_t \in L^1(0,T;V')$ is called weak solution of* (2.4) *if*

$$\begin{aligned} y_t + \nu A y + B(y) &= f, \\ y(0) &= y_0. \end{aligned} \tag{4.6}$$

Results concerning the solvability of (4.6) are standard, cf. [12, 25] for proofs and further details.

THEOREM 4.3 (Existence and uniqueness of solutions). *For every $f \in L^2(0,T;V')$ and $y_0 \in H$, the equation* (4.6) *has a unique solution $y \in W(0,T)$.*

Notice that the regularity $y \in W(0,T)$ is more than the regularity needed to define weak solutions. As in the stationary case, we want to work with the weakest norms of the control as possible. In the presence of a distributed control $u \in L^2(Q)^2$, the inhomogeneity f is formed by

$$\langle f(t), v(t)\rangle_{V',V} := (u(t), v(t))_{q,q'} \quad v \in L^2(0,T;V), t \in [0,T],$$

where q is an exponent less or equal 2. Next we will derive some useful estimates of weak solutions. Observe, that we need $u \in L^2(Q)^2$ to prove that the solutions are of class $C([0,T],H)$, but the estimates contain $L^q(Q)^2$-norms of u, which are weaker since $q \le 2$.

LEMMA 4.4. *For each $u \in L^2(Q)^2$ there exists a unique weak solution $y \in W(0,T)$ of* (2.4)*. It holds*

$$\|y\|_{L^2(V)} + \|y\|_{C([0,T],H)} \le c_B\left(|y_0|_H + \|u\|_q\right), \tag{4.7}$$

where $c_B = c_B(q)$ is independent of y_0 and u. If y_1, y_2 are two solutions of (2.4) *associated with control functions u_1, $u_2 \in U_{ad}$, repectively, then the Lipschitz estimate*

$$\|y_1 - y_2\|_{L^2(V)} + \|y_1 - y_2\|_{C([0,T],H)} \le c_L\ \|u_1 - u_2\|_q \tag{4.8}$$

is satisfied with some constant $c_L > 0$.

Proof. Existence and regularity follow from Theorem 4.3. Let y be the unique weak solution of (2.4) defined by (4.6). We test (4.6) by y. Then the nonlinear term vanishes due to $b(y(t), y(t), y(t)) = 0$ for almost all $t \in [0,T]$. We get the following differential equation:

$$\frac{1}{2}\frac{d}{dt}|y(t)|_H^2 + \nu|y(t)|_V^2 = (u(t), y(t))_{q,q'} \qquad \text{a.e. on } [0,T].$$

Integration from 0 to $t \in [0,T]$ yields

$$\frac{1}{2}|y(t)|_H^2 - \frac{1}{2}|y_0|_H^2 + \nu\int_0^t |y(s)|_V^2 ds = \int_0^t (u(s), y(s))_{q,q'} ds. \tag{4.9}$$

Using Hölders inequality, the inequalities (4.1), (4.5), and Young's inequality, we derive

$$\begin{aligned}\int_0^t (u,y)_{q,q'} ds &\le \int_0^t |u(s)|_q |y(s)|_{q'} ds \le \left(\int_0^t |u(s)|_q^q ds\right)^{1/q}\left(\int_0^t |y(s)|_{q'}^{q'} ds\right)^{1/q'} \\ &\le c_a\|u\|_q\left(\int_0^t |y(s)|_H^{q'/2}|y(s)|_V^{q'/2} ds\right)^{1/q'} \\ &\le c_a\|u\|_q\|y\|_{C([0,T],H)}^{1/2}\|y\|_{L^{q'/2}(0,t;V)}^{1/2},\end{aligned}$$

where $c_a = 2^{1/4}(\operatorname{vol} Q)^{\mu}$ and $\mu = 1/q' - 1/4$ are given by (4.1) and (4.5). Notice, that $q' \le 4$ implies $q'/2 \le 2$, hence we can apply Lemma 3.2 with respect to the time interval $[0,t]$ to proceed

$$\begin{aligned}&\le c_a T^{2\mu}\|u\|_q\|y\|_{C([0,T],H)}^{1/2}\|y\|_{L^2(0,t;V)}^{1/2} \\ &\le c_b\|u\|_q^2 + \frac{1}{4}\|y\|_{C([0,T],H)}^2 + \frac{\nu}{2}\|y\|_{L^2(0,t;V)}^2,\end{aligned} \tag{4.10}$$

where $c_b = \frac{1}{2}(\operatorname{vol} Q)^{2\mu} T^{4\mu} \nu^{-1/2}$. Putting (4.9) and (4.10) together, we find that

$$\frac{1}{2}|y(t)|_H^2 + \frac{\nu}{2}\|y\|_{L^2(0,t;V)}^2 \le \frac{1}{2}|y_0|_H^2 + c_b\|u\|_q^2 + \frac{1}{4}\|y\|_{C([0,T],H)}^2$$

holds for almost all $t \in [0,T]$. Here, the $C([0,T],H)$-norm of y appears on the right-hand side to bound $|y(t)|_H$. Since $y \in C([0,T],H)$ is given by Theorem 4.3, this inequality makes sense. Taking the maximum for $t \in [0,T]$ on the left-hand side we get

$$\frac{1}{4}\|y\|_{C([0,T],H)}^2 \le \frac{1}{2}|y_0|_H^2 + c_b\|u\|_q^2.$$

The $L^2(0,T;V)$-estimate of y follows immediately,

$$\|y\|_{L^2(V)}^2 \le \frac{2}{\nu}|y_0|_H^2 + \frac{4c_b}{\nu}\|u\|_q^2 \le \frac{2}{\nu}|y_0|_H^2 + \frac{2}{\nu^{3/2}}\mathcal{M}^2 =: \mathcal{K}. \tag{4.11}$$

In this way, we have derived a uniform bound on $\|y\|_{L^2(V)}$ for all states y associated with admissible controls. It remains to prove the Lipschitz-estimate. Let y_1, y_2 be two solutions of (2.4) associated with the control functions u_1, u_2. Denote by y and u the difference of them, $y = y_1 - y_2$ and $u = u_1 - u_2$. We substract the corresponding variational equalities, test with $v = y$, and integrate over $[0,t]$. This yields

$$\frac{1}{2}|y(t)|_H^2 + \nu\int_0^t |y(s)|_V^2\,ds = \int_0^t (u(s),y(s))_{q,q'}\,ds - \int_0^t b(y(s),y_2(s),y(s))ds, \tag{4.12}$$

since $y(0) = y_1(0) - y_2(0) \equiv 0$. For the treatment of the nonlinear terms we refer to equation (3.6). Analogously as above, we conclude

$$\int_0^t (u(s),y(s))_{q,q'}\,ds \le c_u\|u\|_q^2 + \frac{1}{\mathcal{N}}\|y\|_{C([0,T],H)}^2 + \frac{\nu}{2}\|y\|_{L^2(0,t;V)}^2,$$

with $c_u = \frac{1}{4}(\operatorname{vol} Q)^{2\mu} T^{4\mu} \mathcal{N}^{1/2}\nu^{-1/2}$ and a constant $\mathcal{N} > 0$ to be specified later. The nonlinear term is estimated by (4.3),

$$\begin{aligned}\left|\int_0^t b(y(s),y_2(s),y(s))ds\right| &\le \sqrt{2}\int_0^t |y(s)|_H|y(s)|_V|y_2(s)|_V\,ds \\ &\le \frac{\nu}{4}\int_0^t |y(s)|_V^2\,ds + \frac{2}{\nu}\int_0^t |y(s)|_H^2|y_2(s)|_V^2\,ds.\end{aligned}$$

Inserting these estimates in (4.12), we obtain

$$\frac{1}{2}|y(t)|_H^2 + \frac{\nu}{4}\|y\|_{L^2(0,t;V)}^2 \le c_u\|u\|_q^2 + \frac{1}{\mathcal{N}}\|y\|_{C([0,T],H)}^2 + \frac{2}{\nu}\int_0^t |y(s)|_H^2|y_2(s)|_V^2\,ds.$$

Since $y_2 \in L^2(0,T;V)$, the norm square $|y_2(\cdot)|_V^2$ is integrable and Gronwall's lemma applies to get

$$|y(t)|_H^2 \le \exp\left(\frac{4}{\nu}\|y_2\|_{L^2(V)}^2\right)\left(c_u\|u\|_q^2 + \frac{2}{\mathcal{N}}\|y\|_{C([0,T],H)}^2\right).$$

We choose $\mathcal{N} := 8\exp\left(\frac{4}{\nu}\mathcal{K}\right)$, where $\mathcal{K}$ is given by (4.11). The uniform bound derived also in equation (4.11) yields that the following inequality holds for all $t \in [0,T]$:

$$\frac{1}{2}|y(t)|_H^2 \le \frac{1}{4}\|y\|_{C([0,T],H)}^2 + c\|u\|_q^2.$$

With the same arguments as above, we conclude

$$\|y\|_{C([0,T],H)} + \|y\|_{L^2(V)} \le c\|u\|_q,$$

and the Lipschitz dependence of the states on the controls is proven. ∎

To establish optimality conditions, we will also need estimates of solutions of linearized equations. Therefore, we introduce the derivative $B'(y_l)$ of the nonlinear operator B which is given by

$$\int_0^T \langle B'(y_l(t))y(t), w(t)\rangle_{V',V} dt := \int_0^T \{b(y_l(t), y(t), w(t)) + b(y(t), y_l(t), w(t))\}\, dt.$$

In view of (4.3), it can be shown that $B'(y_l) \in \mathcal{L}(L^2(0,T;V),\, L^{4/3}(0,T;V'))$ for $y_l \in W(0,T)$.

LEMMA 4.5. *Let $y_l \in W(0,T)$ be the state associated with a control $u_l \in U_{ad}$. Then, for all $u \in L^2(Q)^2$, there exists a unique weak solution $y \in W(0,T)$ of the linearized equation*

$$\begin{aligned} y_t + \nu Ay + B'(y_l)y &= u, \\ y(0) &= 0. \end{aligned} \tag{4.13}$$

It satisfies the estimate

$$\|y\|_{C([0,T],H)} + \|y\|_{L^2(V)} \le c_l\|u\|_q. \tag{4.14}$$

Proof. For the proof of existence we refer to [19]. A similar result was proven in [7] for the three-dimensional case. The estimate (4.14) can be shown as in the previous lemma. The uniqueness of solutions is a consequence of the linearity of the equation and the continuity estimate (4.14). ∎

4.3. First order necessary optimality conditions. Now we return to our optimal control problem. Before stating the second-order sufficient optimality condition, we briefly recall the necessary conditions for local optimality. For the proofs and further discussion see [1, 7, 18, 20] and the references cited therein.

DEFINITION 4.6 (Locally optimal control). *A control $\bar u \in U_{ad}$ is said to be locally optimal in $L^2(Q)^2$, if there exists a constant $\rho > 0$ such that*

$$J(\bar y, \bar u) \le J(y_h, u_h)$$

holds for all $u_h \in U_{ad}$ with $\|\bar u - u_h\|_2 \le \rho$. Here, $\bar y$ and y_h denote the states associated with $\bar u$ and u_h, respectively.

In the following, we denote by $B'(\bar y)^*$ the formal adjoint of $B'(\bar y)$. For $\bar y \in W(0,T)$, it is a continuous linear operator from $L^2(0,T;V)$ to $L^{4/3}(0,T;V')$.

THEOREM 4.7 (Necessary condition). *Let $\bar u$ be a locally optimal control with associated state $\bar y = y(\bar u)$. Then there exists a unique solution $\bar\lambda \in W^{4/3}(0,T;V)$ of the adjoint equation*

$$\begin{aligned} -\bar\lambda_t + \nu A\bar\lambda + B'(\bar y)^*\bar\lambda &= \bar y - y_Q \\ \bar\lambda(T) &= \bar y(T) - y_T. \end{aligned} \tag{4.15}$$

Moreover, the variational inequality

$$(\gamma\bar{u} + \bar{\lambda}, u - \bar{u})_{L^2(Q)^2} \geq 0 \quad \forall u \in U_{ad} \tag{4.16}$$

is satisfied.

Proof. A proof can be found in [18, 19]. It can be carried out along the lines of the proof of Theorem 3.10 for the stationary case. First, one can show the Fréchet-differentiablity of the solution operator of the instationary equation. The adjoint system and the variational inequality is then derived by the method of transposition. The regularity of $\bar{\lambda}$ is proven in [21]. ☐

Next, we state an estimate of the norm of the adjoint state, see [20] for the details.

COROLLARY 4.8. *Let $\lambda \in W^{4/3}(0,T;V)$ be the weak solution of (4.15), where $y \in W(0,T)$ is a state associated with an admissible control $u \in U_{ad}$. Then it holds*

$$\|\lambda\|^2_{L^2(V)} \leq c\left(|y(T) - y_T|^2_{L^2(\Omega)} + \|y - y_Q\|_2^2\right).$$

Let us introduce the Lagrange function $\mathcal{L} : W(0,T) \times L^2(Q)^2 \times W^{4/3}(0,T;V)$ of the instationary optimal control problem by

$$\mathcal{L}(y,u,\lambda) = J(u,y) + \int_0^T \{\langle y_t, \lambda\rangle_{V',V} - \nu(y,\lambda)_V - b(y,y,\lambda) + (u,\lambda)\}\, dt.$$

One can easily verify that the necessary optimality conditions given in Theorem 4.7 are equivalent to

$$\mathcal{L}_u(\bar{y},\bar{u},\bar{\lambda})(u - \bar{u}) \geq 0 \quad \forall u \in U_{ad},$$

and

$$\mathcal{L}_y(\bar{y},\bar{u},\bar{\lambda})h = 0 \quad \forall h \in W(0,T) \text{ with } h(0) = 0.$$

As in the stationary case, it can be proven that $\mathcal{L}$ is twice Fréchet differentiable with respect to y and u, confer Lemma 3.12. Here we derive an estimate of the norm of $\mathcal{L}_{yy}$.

In the analysis of the second-order condition estimations of the time derivative y_t of a state $y \in W(0,T)$ are not needed. Therefore, we introduce a space $\tilde{W}$ by

$$\tilde{W} = L^2(0,T;V) \cap C([0,T],H).$$

equipped with the norm

$$\|y\|^2_{\tilde{W}} := \|y\|^2_{L^2(V)} + \|y\|^2_{C([0,T],H)}.$$

LEMMA 4.9. *The second derivative of the Lagrangian $\mathcal{L}$ at $y \in W(0,T)$ with associated adjoint state λ in the direction $z_1, z_2 \in W(0,T)$ satisfies the estimate*

$$|\mathcal{L}_{yy}(y,u,\lambda)[z_1,z_2]| \leq c_{\mathcal{L}}\, \|z_1\|_{\tilde{W}}\|z_2\|_{\tilde{W}}$$

for all $z_1, z_2 \in W(0,T)$.

Proof. The second derivative of $\mathcal{L}$ is given by

$$\mathcal{L}_{yy}(y,u,\lambda)[z_1,z_2] = (z_1(T), z_2(T))_2 + \int_0^T (z_1,z_2)_2 - b(z_1,z_2,\lambda) - b(z_2,z_1,\lambda)\, dt.$$

The nonlinear terms are estimated by (4.3),

$$\begin{aligned}\left|\int_0^T b(z_1,z_2,\lambda)\,dt\right| &\le c\int_0^T |z_1(t)|_H^{1/2}|z_1(t)|_V^{1/2}|z_2(t)|_H^{1/2}|z_2(t)|_V^{1/2}|\lambda(t)|_V\,dt\\ &\le c\,\|z_1\|_{C([0,T],H)}^{1/2}\|z_2\|_{C([0,T],H)}^{1/2}\int_0^T |z_1(t)|_V^{1/2}|z_2(t)|_V^{1/2}|\lambda(t)|_V\,dt\\ &\le c\,\|z_1\|_{C([0,T],H)}^{1/2}\|z_2\|_{C([0,T],H)}^{1/2}\|z_1\|_{L^2(V)}^{1/2}\|z_2\|_{L^2(V)}^{1/2}\|\lambda\|_{L^2(V)}\\ &\le c\,\|z_1\|_{\tilde W}\|z_2\|_{\tilde W}\|\lambda\|_{L^2(V)}.\end{aligned}$$

Corollary 4.8 together with Lemma 4.4 and the boundedness of the controls (4.4) yields a uniform bound on all adjoint states, $\|\lambda\|_{L^2(V)} \le C$, independently of y and u. Now the claim follows immediately. □

4.4. Second-order sufficient optimality condition. In what follows we fix $\bar v := (\bar y, \bar u)$ to be an admissible reference pair. We suppose that $\bar v$ satisfies the first-order necessary optimality conditions.

Definition 4.10 (Strongly active sets). *Let $\varepsilon > 0$ and $i \in \{1,2\}$ be given. Define sets $Q_{\varepsilon,i} \subseteq Q = \Omega \times [0,T]$ by*

$$Q_{\varepsilon,i} = \{(x,t) \in Q :\, |\gamma\bar u_i(x,t) + \bar\lambda_i(x,t)| > \varepsilon\}.$$

For $u \in L^p(Q)^2$ and $1 \le p < \infty$ we define the L^p-norm with respect to the sets of strongly active control constraints

$$\|u\|_{L^p,Q_\varepsilon} := \left(\sum_{i=1}^2 \|u_i\|_{L^p(Q_{\varepsilon,i})}^p\right)^{1/p}.$$

As in the previous section, we can show the following conclusion, cf. Corollary 3.14.

Corollary 4.11. *For all $u \in U_{ad}$ it holds*

$$\sum_{i=1}^2 \int_0^T \int_{Q_{\varepsilon,i}} (\gamma\bar u_i(x,t) + \bar\lambda_i(x,t))(u_i(x,t) - \bar u_i(x,t))dx\,dt \ge \varepsilon\,\|u - \bar u\|_1.$$

We assume that the reference pair $\bar v = (\bar y, \bar u)$ satisfies the following coercivity assumption on $\mathcal{L}''(\bar v, \bar\lambda)$, in the sequel called second-order sufficient condition:

(SSC) *There exist $\varepsilon > 0$ and $\delta > 0$ such that*

$$\mathcal{L}_{vv}(\bar v, \bar\lambda)[(z,h)]^2 \ge \delta\,\|h\|_q^2$$

holds for all pairs $(z,h) \in W(0,T) \times L^2(Q)^2$ with

$$h = u - \bar u,\ u \in U_{ad},\ h_i = 0 \text{ on } Q_{\varepsilon,i} \text{ for } i = 1,2,$$

and $z \in W(0,T)$ being the weak solution of the linearized equation

$$\begin{aligned} z_t + Az + B'(\bar y)z &= h\\ z(0) &= 0.\end{aligned}$$

Now, we collected all tools to prove that (SSC) is sufficient for local optimality of $(\bar{y}, \bar{u})$, provided the first-order necessary conditions are fulfilled. The proof in the instationary case follows exactly the lines of the proof in the stationary case, cf. Theorem 3.16. So we only state the associated results without proof.

THEOREM 4.12. *Let $\bar{v} = (\bar{y}, \bar{u})$ be admissible for the optimal control problem and suppose that $\bar{v}$ fulfills the first-order necessary optimality condition with associated adjoint state $\bar{\lambda}$. Assume further that (SSC) is satisfied at $\bar{v}$. Then there exist $\alpha > 0$ and $\rho > 0$ such that*

$$J(v) \geq J(\bar{v}) + \alpha \|u - \bar{u}\|_q^2$$

holds for all admissible pairs $v = (y, u)$ with $\|u - \bar{u}\|_s \leq \rho$, where the exponents s and q are chosen according to Assumption (A3).

REMARK 4.13. *As in the stationary case, cf. Section 3.4, one case establish an equivalent sufficient condition (SSC_0), which ensures together with first-order necessary optimality conditions local optimality of a reference control without any two-norm discrepancy.*

Acknowledgements. The authors are very grateful to E. Casas for his comments that led us to the remarks in Section 3.4.

REFERENCES

[1] F. Abergel and R. Temam. On some control problems in fluid mechanics. *Theoret. Comput. Fluid Dynam.*, 1:303–325, 1990.

[2] R. A. Adams. *Sobolev spaces.* Academic Press, San Diego, 1978.

[3] N. Arada, J.-P. Raymond, and F. Tröltzsch. On an augmented Lagrangian SQP method for a class of optimal control problems in Banach spaces. *Computational Optimization and Applications*, 22:369–398, 2002.

[4] J. F. Bonnans. Second-order analysis for control constrained optimal control problems of semilinear elliptic equations. *Appl. Math. Optim.*, 38:303–325, 1998.

[5] J. F. Bonnans and H. Zidani. Optimal control problems with partially polyhedric constraints. *SIAM J. Control and Optim.*, 37:1726–1741, 1999.

[6] H. Brezis. *Analyse fonctionelle.* Masson, Paris, 1983.

[7] E. Casas. An optimal control problem governed by the evolution Navier-Stokes equations. In S. S. Sritharan, editor, *Optimal control of viscous flows*, Frontiers in applied mathematics. SIAM, Philadelphia, 1993.

[8] E. Casas and M. Mateos. Uniform convergence of the FEM. Applications to state constrained control problems. *Comp. Appl. Math., To appear*, 2001.

[9] E. Casas and M. Mateos. Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints. *SIAM J. Control and Optimization*, 40(5):1431–1454, 2002.

[10] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for a nonlinear elliptic control problem. *J. for Analysis and its Applications*, 15:687–707, 1996.

[11] 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.

[12] P. Constantin and C. Foias. *Navier-Stokes equations.* The University of Chicago Press, Chicago, 1988.

[13] R. Dautray and J. L. Lions. *Evolution problems I*, volume 5 of *Mathematical analysis and numerical methods for science and technology.* Springer, Berlin, 1992.

[14] A. L. Dontchev, W. W. Hager, A. B. Poore, and B. Yang. Optimality, stability, and convergence in optimal control. *Appl. Math. Optim.*, 31:297–326, 1995.

[15] J. C. Dunn. On second-order sufficient conditions for structured nonlinear programs in infinite-dimensional function spaces. In A. Fiacco, editor, *Mathematical programming with data perturbations*, pages 83–107. Marcel Dekker, 1998.

[16] H. O. Fattorini and S. Sritharan. Necessary and sufficient for optimal controls in viscous flow problems. *Proc. Royal Soc. of Edinburgh*, 124:211–251, 1994.

[17] M. D. Gunzburger, editor. *Flow control*. Spinger, New York, 1995.

[18] M. D. Gunzburger and S. Manservisi. The velocity tracking problem for Navier-Stokes flows with bounded distributed controls. *SIAM J. Contr. Optim.*, 37:1913–1945, 1999.

[19] M. D. Gunzburger and S. Manservisi. Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed control. *SIAM J. Numer. Anal.*, 37:1481–1512, 2000.

[20] M. Hinze. *Optimal and instantaneous control of the instationary Navier-Stokes equations.* Habilitation, TU Berlin, 1999.

[21] M. Hinze and K. Kunisch. Second order methods for optimal control of time-dependent fluid flow. *Siam J. Contr. Optim.*, 40:925–946, 2001.

[22] 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.

[23] T. Roubíček and F. Tröltzsch. Lipschitz stability of optimal controls for the steady-state navier-stokes equations. Tech. Report 749-2002, Inst. of Math, TU Berlin, to appear, 2002.

[24] S. Sritharan. Dynamic programming of the Navier-Stokes equations. *Systems & Control Letters*, 16:299–307, 1991.

[25] R. Temam. *Navier-Stokes equations*. North Holland, Amsterdam, 1979.

[26] F. Tröltzsch. On the Lagrange-Newton-SQP method for the optimal control of semilinear parabolic equations. *SIAM J. Control and Optimization*, 38:294–312, 1999.