www.oeaw.ac.at

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

A-posteriori verification of optimality conditions for control problems with finite-dimensional control space

S. Akindeinde, D. Wachsmuth

RICAM-Report 2011-23

www.ricam.oeaw.ac.at

A-POSTERIORI VERIFICATION OF OPTIMALITY CONDITIONS FOR CONTROL PROBLEMS WITH FINITE-DIMENSIONAL CONTROL SPACE *

SAHEED AKINDEINDE AND DANIEL WACHSMUTH†

Abstract. In this paper we investigate non-convex optimal control problems. We are concerned with a-posteriori verification of sufficient optimality conditions. If the proposed verification method confirms the fulfillment of the sufficient condition then a-posteriori error estimates can be computed. A special ingredient of our method is an error analysis for the Hessian of the underlying optimization problem. We derive conditions under which positive definiteness of the Hessian of the discrete problem implies positive definiteness of the Hessian of the continuous problem. The article is complemented with numerical experiments.

Key words. non-convex optimal control problems, sufficient optimality conditions, a-posteriori error estimates

AMS subject classifications. 49M25, 49K20, 65N15, 65N25

1. Introduction. We study optimal control problems of the following type: Minimize the functional J given by

$$J(y, u) = g(y) + j(u) \tag{P}$$

over all $(y, u) \in Y \times U$ that satisfy the non-linear elliptic partial differential equation

$$E(y, u) = 0$$

and the control constraints

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

Here, Y is a Banach space, $U = \mathbb{R}^n$. The set $U_{ad} \subset U$ is a non-empty, convex and closed set given by

$$U_{ad} = \{u \in U : \ u_a \le u \le u_b\},$$

where the inequalities are to be understood component-wise. Examples that are covered by this framework include parameter identification and optimization problems with finitely many parameters, as for instance least-square problems as given by e.g.

$$J(y, u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} \tag{1.1}$$

over all $(y, u) \in H^1_0(\Omega) \times \mathbb{R}^n$ that satisfy the elliptic equation

$$\begin{aligned} -\Delta y + d(u; y) &= g \quad \text{in } \Omega, \\ y &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.2}$$

Here, $d(u; y)$ denotes a non-linear function of y parametrized by parameters $u \in \mathbb{R}^n$. The parameters have to be recovered by fitting the state y to the measured state y_d.

Another application is the optimization of material parameters by minimizing (1.1) subject to

$$\begin{aligned} -\text{div}(a\nabla y) &= g \quad \text{in } \Omega, \\ y &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.3}$$

where the coefficient function a is given by $a = \sum_{i=1}^n \chi_i u_i$ with $\chi_i = \chi_{\Omega_i}$ being characteristic functions of subsets $\Omega_i \in \Omega$. Both problems are complemented by the constraint $u \in U_{ad}$.

We are interested in the numerical solution and the solution accuracy for such type of problems. Given a numerical solution u_h of a discretization of (P), we are asking, under which conditions u_h is near a local solution $\bar{u}$ of (P). This is a non-trivial question, since the optimization problem (P) is non-convex due to the nonlinearity of the elliptic equation. Hence, all results on discretization errors are subject to a second-order sufficient optimality condition (SSC). This statement applies to both a-priori error estimates

*This work was partially supported by FWF grant P21564-N18.
†Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A–4040 Linz, Austria, saheed.akindeinde@oeaw.ac.at, daniel.wachsmuth@oeaw.ac.at

[2, 6] as well as a-posteriori error estimates [3, 4, 18]. The aim of this paper is to develop a method that allows to verify the fulfillment of the sufficient condition a-posteriori. As a side result, we obtain reliable a-posteriori error estimates.

Let us comment, why it is difficult to verify the second-order sufficient optimality condition. One obstacle is that the sufficient condition is required at a solution $\bar{u}$ of the original undiscretized problem (P). Even if $\bar{u}$ would be known, it would still be tedious to check that the SSC holds, as it requires the exact solution of linearized partial differential equations. Without knowledge of $\bar{u}$ the check for SSC appears to be of the same difficulty as the check whether (P) is convex.

Earlier work on verification of sufficient optimality conditions can be found for instance in Rösch and Wachsmuth [14, 15]. There the optimal control of semi-linear elliptic equations was studied. The principal idea was to verify the fulfillment of a infinite-dimensional second-order condition at the (known) discrete solution u_h, and by a careful analysis confirm that this property carries over to the unknown solution $\bar{u}$ of (P). This infinite-dimensional second-order condition at u_h was relatively easy to check due to the special structure of the Lagrangian associated to the considered problem, which makes it impossible to generalize the results to e.g. (1.3). Numerical studies on second-order sufficient conditions can be found in the work of Mittelmann [10, 11]. There the second-order sufficient optimality condition was checked for the *discrete* problem. The fulfillment of the discrete SSC is a strong indication for the fulfillment of the SSC for the continuous problem but not sufficient. In the present article we will combine both strategies: first check the discrete SSC as in [10, 11], then develop conditions, under which the discrete SSC implies a continuous second-order condition, and finally with an analysis as in [14, 15] conclude that this second-order condition carries over to the unknown solution $\bar{u}$ of (P).

As we want to develop an a-posteriori error analysis, let us comment on available a-posteriori error estimators in the literature. A-posteriori error estimates for non-linear control and identification problems can be found for instance in [3, 8, 9, 18]. Both dual-weighted residual type [4] and residual type error estimators are available. However, they depend on two crucial *a-priori* assumptions: the first is that a second-order sufficient condition has to hold at the solution of the continuous problem. With this assumption, error estimates of the type

$$\|\bar{u} - u_h\|_U \leq c\,\eta + \mathcal{R}$$

can be derived, where η is a computable error indicator and $\mathcal{R}$ is a second-order remainder term. Here, the second a-priori assumption comes into play: one has to assume that $\mathcal{R}$ is small enough, in order to guarantee that mesh refinement solely based on η is meaningful. A different approach with respect to mesh refinement was followed in [20]. There the residuals in the first-order necessary optimality condition were used to derive an adaptive procedure. However, smallness of residuals does not imply smallness of errors without any further assumption. Here again, SSC as well as smallness of remainder terms is essential to draw this conclusion. In this article, we will present conditions that allow reliable a-posteriori error estimates that *verify* these two conditions *a-posteriori* and that do not require them a-priori.

The goal of the paper is thus two-fold: first, we aim at verifying the sufficient condition a-posteriori. This allows to derive a-posteriori error estimators in a second step, see our main results Theorems 3.21 and 3.22 below.

The rest of this section is devoted to define all the ingredients of the abstract problem (P), its optimality conditions and discretization.

1.1. The abstract framework. Let us start with making the assumptions on the abstract problem (P) precise.

Assumption 1.

1. *The mapping $E : Y \times U \to Y^*$ is twice Fréchet-differentiable. Furthermore, we assume that the mapping E is strongly monotone with respect to the first variable, i.e. there is a constant $\delta > 0$ such that*

$$\langle E(y_1, u) - E(y_2, u),\, y_1 - y_2\rangle_{Y^*,Y} \geq \delta \|y_1 - y_2\|_Y^2$$

   for all $u \in U_{ad}, y_1, y_2 \in Y$.
2. *The functions $g : Y \to \mathbb{R}$ and $j : U \to \mathbb{R}$ are twice Fréchet-differentiable.*

The assumptions on E are met for instance for semilinear elliptic equations with monotone non-linearities, e.g. it is fulfilled for (1.2) if $d(u; \cdot)$ is monotonically increasing for all admissible u, and the assumption is fulfilled for (1.3) if the coefficients u are strictly positive. Moreover, the assumptions on g and j are met by the prototypical functional given in (1.1).

Proposition 1.1. *Under Assumption* (1)*, for each admissible control* $u \in U$ *the state equation* $E(y,u) = 0$ *is uniquely solvable. Furthermore, the partial derivative of operator* E *with respect to* y*, denoted by* E_y*, is continuously invertible, i.e.* $E_y(y_0,u_0)^{-1} \in \mathcal{L}(Y^*,Y)$ *for all* $(y_0,u_0) \in Y \times U_{ad}$*, and it holds* $\|E_y^{-1}(y_0,u_0)\|_{\mathcal{L}(Y^*,Y)} \le \delta^{-1}$ *for all* $(y_0,u_0) \in Y \times U_{ad}$*.*

Proof. The result is standard and the proof can be found for instance in [19]. ∎

Since E maps to the dual of Y, we can consider the weak formulation of the state equation

$$\langle E(y,u),\, v\rangle_{Y^*,Y} = 0 \qquad \forall v \in Y.$$

The solution mapping $u \to y$ that assigns to every control u a state y is denoted by S, i.e. $y = S(u)$. Its Fréchet derivative $S'(u)v$ for $v \in U$ is the unique solution of the linearized equation

$$E_y(S(u),u) \cdot S'(u)v + E_u(S(u),u) \cdot v = 0.$$

We define the Lagrange functional for the abstract problem

$$\mathcal{L}(u,y,p) = g(y) + j(u) - \langle E(y,u),\, p\rangle_{Y^*,Y}.$$

Let $(\bar{y},\bar{u})$ be locally optimal for (P). Then the first-order necessary optimality conditions can be expressed as $\mathcal{L}_y'(\bar{y},\bar{u},\bar{p}) = 0$ and $\mathcal{L}_u(\bar{y},\bar{u},\bar{p})(u-\bar{u}) \ge 0$ for all $u \in U_{ad}$, which is equivalent to

$$\begin{aligned} E_y(\bar{u},\bar{y})^*\bar{p} &= g'(\bar{y}) \\ \langle j'(\bar{u}) - E_u(\bar{u},\bar{y})^*\bar{p},\, u-\bar{u}\rangle_{U^*,U} &\ge 0 \quad \forall u \in U_{ad}. \end{aligned}$$

Since the problem (P) is in general non-convex, the fulfillment of these necessary conditions does not imply optimality. In order to guarantee this, one needs additional sufficient optimality conditions. One particular (and strong) instance is given by: There exists $\alpha > 0$ such that

$$\mathcal{L}''(\bar{u},\bar{y},\bar{p})[(z,v)^2] \ge \alpha\|v\|_U^2 \tag{1.4}$$

holds for all $v = u - \bar{u}$, $u \in U_{ad}$, and z solves the linearized equation $E_y(\bar{u},\bar{y})z + E_u(\bar{u},\bar{y})v = 0$. Later we will work with a weaker sufficient condition, where the subspace on which $\mathcal{L}''$ is required to be positive is shrunk taking strongly active inequality constraints into account.

Let us explain, why this condition is difficult to check numerically even in the case when $(\bar{u},\bar{y},\bar{p})$ are given. The main difficulty here is that the function z appearing in (1.4) is given as solution of a partial differential equation, which cannot be solved explicitely. Any discretization of this equation introduces another error that has to be analyzed.

Remark 1.2. *We can relax the differentiability requirements of Assumption 1 as follows: let* Y_∞ *be a Banach space with a continuous embedding in* Y*. Then it is sufficient to require Fréchet-differentiability of* E *and* g *with respect to* y *in the stronger topology of* Y_∞ *as long as the derivatives of* E *and* g *with respect to* y *satisfy the Lipschitz estimates in Assumption 4 below with respect to the weaker topology* Y*. This would allow for instance to choose* $Y = H_0^1(\Omega)$ *and* $Y_\infty = L^\infty(\Omega) \cap Y$ *or* $Y_\infty = H^2(\Omega) \cap Y$*. See also the comments after Assumption 4 and in Section 4.3 below.*

1.2. Discretization. In order to solve (P) the problem has to be discretized. Let Y_h be a finite-dimensional subspace of Y. Then a discretization of the state equation can be obtained in the following way: A function $y_h \in Y_h$ is a solution of the discretized equation for given $u \in U_{ad}$ if and only if

$$\langle E(y_h,u),\, \phi_h\rangle_{Y^*,Y} = 0 \qquad \forall \phi_h \in Y_h.$$

The discrete optimization problem is then given by: Minimize the functional $J(y_h,u_h)$ over all $(y_h,u_h) \in Y_h \times U_{ad}$, where y_h solves the discrete equation.

Let $(\bar{y}_h,\bar{u}_h)$ be a local solution of the discrete problem. Then it fulfills the discrete first-order necessary optimality condition, which is given as: there exists a uniquely determined discrete adjoint state $\bar{p}_h \in Y_h$ such that it holds

$$\begin{aligned} \langle E_y(\bar{y}_h,\bar{u}_h)^*\bar{p}_h,\, \phi_h\rangle_{Y^*,Y} &= \langle g'(\bar{y}_h),\, \phi_h\rangle_{Y^*,Y} \quad \forall \phi_h \in Y_h \\ \langle j'(\bar{u}_h) - E_u(\bar{y}_h,\bar{u}_h)^*\bar{p}_h, u-\bar{u}_h\rangle_{U^*,U} &\ge 0 \;\forall u \in U_{ad}. \end{aligned} \tag{1.5}$$

Throughout this work, we will assume that errors in discretizing the optimality system are controllable in the following sense.

Assumption 2. *Let (u_h, y_h, p_h) be approximations of the discrete optimal control and the corresponding state and adjoint. There are positive constants r_y, r_p, r_u such that the following holds*

$$\|E(y_h, u_h)\|_{Y^*} \leq r_y, \tag{1.6}$$
$$\|g'(y_h) - E_y(y_h, u_h)^* p_h\|_{Y^*} \leq r_p, \tag{1.7}$$
$$\langle j'(u_h) - E_u(y_h, u_h)^* p_h, u - u_h\rangle \geq -r_u \|u - u_h\|_U \ \forall u \in U_{ad}. \tag{1.8}$$

If $(\bar{y}_h, \bar{u}_h, \bar{p}_h)$ fulfills the first-order necessary optimality system (1.5) of the discrete problem then $r_u = 0$. The residuals r_y and r_p cannot be expected to vanish as they reflect the discretization error of the partial differential equation. We report on the computation of these residuals in Section 4.1.

As already mentioned, without any further assumption, smallness of the residuals in (1.6)–(1.8) does not imply smallness of the error $\|u - u_h\|_U$ in the control. In order to establish such a bound, it is essential to check that a second-order sufficient optimality condition is satisfied.

Here it is important to recognize that sufficient optimality conditions for the *discrete* problem alone are still not enough. The sufficient optimality condition for the discrete problem is given by: There exists $\alpha_h > 0$ such that

$$\mathcal{L}''(\bar{u}_h, \bar{y}_h, \bar{p}_h)[(z_h, v)^2] \geq \alpha_h \|v\|_U^2$$

holds for all $v = u - \bar{u}$, $u \in U_{ad}$, and z_h solves the linearized discrete equation

$$\langle E_y(\bar{u}_h, \bar{y}_h) z_h + E_u(\bar{u}_h, \bar{y}_h) v,\ \phi_h\rangle = 0 \quad \forall \phi_h \in Y_h.$$

This condition is equivalent to positive definiteness of a certain computable matrix, see Section 3.5. Hence, this condition can be checked computationally, see e.g. [10, 11]. Under conditions to be worked out in the sequel, the fulfillment of this discrete SSC implies the fulfillment of the continuous SSC. These conditions are fulfilled if α_h is relatively large compared to the residuals r_y, r_p, r_u, see Section 3.5 below.

If we can verify that the SSC (1.4) holds then we have a reliable error bound for the error in the optimal controls and states

$$\|\bar{y} - y_h\|_Y + \|\bar{u} - u_h\|_U \leq C\,(r_y + r_p + r_u)\,.$$

This allows to devise an adaptive refinement scheme, which refines elements with relatively large local error components in r_y, r_p.

We will first derive such an error representation for the reduced problem in Section 2, where the unknown y is eliminated in terms of $y = S(u)$. These results are then applied to the problem (P), the required error and Lipschitz estimates are carried out in Section 3. The main result on relation between discrete and continuous second-order conditions can be found in Section 3.5, Theorem 3.21. Finally, we state the a-posteriori error estimate and verification of optimality in Section 3.6, Theorem 3.22.

Notational convention. Throughout the paper we will use the following convention when naming constants: M_f and c_f will denote global bounds and Lipschitz constants for a function f; ϵ_x and r_x will denote errors estimates and residuals of a quantity x. Moreover, x^h will denote auxiliary quantities that have certain similarities to a discrete quantity x_h but do not need to be explicitely known.

2. Verification of optimality for reduced functional. Let us introduce the reduced objective functional $f : U \to \mathbb{R}$ by

$$f(u) = g(S(u)) + j(u). \tag{2.1}$$

Since J and S are twice Fréchet-differentiable, the reduced functional f inherits this property as well. This allows us to write the original abstract minimization problem in the control-reduced form:

$$\min_{u \in U_{ad}} f(u). \tag{2.2}$$

For the control-reduced problem (2.2), the first order optimality condition to be fulfilled by every optimal solution candidate $\bar{u}$ states

$$f'(\bar{u})(u - \bar{u}) \geq 0 \qquad \forall u \in U_{ad}.$$

The corresponding second order sufficient optimality condition to be fulfilled by a locally optimal solution $\bar{u}$ is given by the existence of $\alpha > 0$ such that

$$f''(\bar{u})[v, v] \geq \alpha \|v\|_U^2 \qquad \forall v = u - \bar{u} :\ u \in U_{ad},\ f'(\bar{u})v = 0. \tag{2.3}$$

Let us define the active set A as

$$A(u) = \{i \in \{1, ..., n\} : |f'(u)_i| > 0\}$$

and the corresponding inactive set as $I = \{1, ..., n\} \setminus A$. Here the notion of active set comes from the fact that for the solution $\bar{u}$ it holds $\bar{u}_i \in \{u_{a,i},\ u_{b,i}\}$ for $i \in A(\bar{u})$. That is, the inequality constraints are active for these components. Then the second order condition (2.3) can be written equivalently as: there exists $\alpha, \sigma > 0$ such that

$$|f'(\bar{u})| \geq \sigma \text{ on } A \tag{2.4}$$

and

$$f''(\bar{u})[v, v] \geq \alpha \|v\|_U^2 \qquad \forall v \in U : v = u - \bar{u},\ u \in U_{ad},\ v = 0 \text{ on } A. \tag{2.5}$$

Here, the notion "$v = 0$ on A" means $v_i = 0$ for all $i \in A$. One of the tasks in this paper is to verify conditions (2.4) and (2.5) numerically for the reduced problem (2.1).

Of course, since the control space is finite-dimensional, the requirement (2.3) is equivalent to assuming $f''(\bar{u})[v, v] > 0$ for the mentioned test functions $v \neq 0$. Similarly, the requirement $|f'(\bar{u})| \geq \sigma$ can be replaced by $|f'(\bar{u})| > 0$. However, we will need later a quantification of these bounds. So we opted to present the sufficient optimality condition in this way.

Let us define the following notation that will be useful in this section. Let $A \subset \{1 \dots n\}$, $I = \{1 \dots n\} \setminus A$ be given. Then the restriction of a vector v to A is given by

$$(v|_A)_i := \begin{cases} v_i & \text{if } i \in A, \\ 0 & \text{if } i \notin A. \end{cases}$$

Additionally, we can split the norm on U as

$$\|u\|_U^2 = \sum_{i \in A} |u_i|^2 + \sum_{i \in I} |u_i|^2 = \left\|u|_A\right\|_U^2 + \left\|u|_I\right\|_U^2 =: \|u\|_A^2 + \|u\|_I^2.$$

Now let us suppose we computed an approximate solution u_h of the reduced problem (2.2). We want to verify that this approximation is close to a local solution of the reduced problem. In order to prove this we assume that u_h fulfills the following:

Assumption 3. *There is a subset $A \subset \{1 \dots n\}$ with $I = \{1 \dots n\} \setminus A$, and positive constants ϵ, α, σ such that the following hold*

$$f'(u_h)(u - u_h) \geq \sigma \|u - u_h\|_A - \epsilon \|u - u_h\|_I \qquad \forall u \in U_{ad}, \tag{2.6}$$

$$f''(u_h)[v, v] \geq \alpha \|v\|_U^2 \qquad \forall v \in U : v_i = 0 \text{ on } A. \tag{2.7}$$

Additionally there are positive constants $c_{f'}$, $c_{f''}$, and R such that it holds

$$\|f'(u) - f'(u_h)\|_{U^*} \leq c_{f'} \|u - u_h\|_U, \tag{2.8}$$

$$\|f''(u) - f''(u_h)\|_{(U \times U)^*} \leq c_{f''} \|u - u_h\|_U, \tag{2.9}$$

$$\|f''(u)\|_{(U \times U)^*} \leq M_{f''} \tag{2.10}$$

for all $u \in U_{ad}$ with $\|u - u_h\| \leq R$.

Some comments are in order. Inequality (2.6) means that the derivative $f'(u_h)$ has the right sign on the active set A, while $f'(u_h)$ is bounded by ϵ on the inactive set. We expect for computations that ϵ tends to zero with decreasing mesh size of the discretization, while σ should be bounded away from zero. Assumption (2.7) is exactly the second-order requirement (2.5) of the sufficient optimality condition for the reduced problem. The second part of Assumption 3 simply names the Lipschitz constants of f. An assumption similar to Assumption (3) was used in [15] without the notion of a set A, i.e. there $A = \emptyset$, $\sigma = 0$ was used.

Let us remark that Assumption 3 is fulfilled for the solution $\bar{u}$, if $\bar{u}$ satisfies the sufficient condition (2.4)–(2.5). Here it has to be noted that (2.4) implies $f'(\bar{u})(u - \bar{u}) \geq \sigma \|u - \bar{u}\|_{l^1(A)}$. Due to the finite-dimensional setting, the l^1-norm dominates the l^2-norm, which gives (2.6) with $\epsilon = 0$. In order to transfer the results to the infinite-dimensional case, in particular to $U = L^2(\Omega)$, one would have to work with two different norms of $L^1(\Omega)$ and $L^2(\Omega)$ type.

Let $u \in U_{ad}$ be an arbitrary feasible point. In the following, we want to analyze $f(u) - f(u_h)$ in terms of $\|u-u_h\|_A$ and $\|u-u_h\|_I$. To this end, let us introduce an auxiliary admissible control $\tilde{u}$ defined by

$$\tilde{u} = \begin{cases} u_h & \text{on the active set } A \\ u & \text{on the inactive set } I, \end{cases}$$

Furthermore, we define the abbreviations

$$r_I := \|u - u_h\|_I, \quad r_A := \|u - u_h\|_A.$$

Then we have $\|u - \tilde{u}\|_U = \|u - u_h\|_A$ and $\|\tilde{u} - u_h\|_U = \|u - u_h\|_I$. With the aid of Assumption 3 we will estimate the difference $f(u) - f(u_h)$. First we make use of the new control variable $\tilde{u}$ to split the difference

$$f(u) - f(u_h) = \big(f(u) - f(\tilde{u})\big) + \big(f(\tilde{u}) - f(u_h)\big). \tag{2.11}$$

Applying Taylor's expansion up to the second order on the first addend of (2.11), we have

$$\begin{aligned} f(u) - f(\tilde{u}) &= f'(\tilde{u})(u-\tilde{u}) + \int_0^1\int_0^s f''(\tilde{u} + t(u-\tilde{u}))[(u-\tilde{u})^2]\,\mathrm{d}t\mathrm{d}s \\ &= f'(u_h)(u-\tilde{u}) + (f'(\tilde{u}) - f'(u_h))(u-\tilde{u}) \\ &\quad + \int_0^1\int_0^s f''(\tilde{u} + t(u-\tilde{u}))[(u-\tilde{u})^2]\,\mathrm{d}t\mathrm{d}s, \end{aligned}$$

which, due to (2.6), (2.8) and (2.10), implies

$$f(u) - f(\tilde{u}) \geq \sigma r_A - c_{f'} r_A r_I - \frac{M_{f''}}{2} r_A^2.$$

Employing (2.6) and (2.7), the second addend of (2.11) is estimated as

$$\begin{aligned} f(\tilde{u}) - f(u_h) &\geq f'(u_h)(\tilde{u} - u_h) + \frac{1}{2} f''(u_h)(\tilde{u} - u_h)^2 \\ &\quad + \int_0^1\int_0^s (f''(u_h + t(\tilde{u} - u_h)) - f''(u_h))[(u-\tilde{u})^2]\,\mathrm{d}t\mathrm{d}s \\ &\geq -\epsilon r_I + \frac{\alpha}{2} r_I^2 - \frac{c_{f''}}{6} r_I^3. \end{aligned}$$

Altogether, we arrived at

$$f(u) - f(u_h) \geq \sigma r_A - c_{f'} r_A r_I - \frac{M_{f''}}{2} r_A^2 - \epsilon r_I + \frac{\alpha}{2} r_I^2 - \frac{c_{f''}}{6} r_I^3. \tag{2.12}$$

We can now prove a first result for the reduced problem: Under assumptions on the constants in Assumption 3 we obtain the existence of a local solution $\bar{u}$ near u_h.

THEOREM 2.1. *Let Assumption 3 be satisfied. If there exist $r_I, r_A > 0$ with $r_I^2 + r_A^2 < R^2$ such that*

$$\min(\sigma r_A, \frac{\alpha}{2} r_I^2) - c_{f'} r_A r_I - \frac{M_{f''}}{2} r_A^2 - \epsilon r_I - \frac{c_{f''}}{6} r_I^3 > 0 \tag{2.13}$$

then there exists a local solution $\bar{u}$ to the control-reduced problem (2.2) *satisfying*

$$\|\bar{u} - u_h\|_A < r_A, \quad \|\bar{u} - u_h\|_I < r_I.$$

If moreover with $r_+ := \sqrt{r_I^2 + r_A^2}$

$$\sigma - c_{f'} r_+ > 0, \quad \alpha - c_{f''} r_I > 0 \tag{2.14}$$

holds, then $\bar{u} = u_h$ on A and the second-order sufficient optimality conditions (2.4)–(2.5) *are fulfilled at $\bar{u}$.*

Proof. Let us define

$$B := \{u \in U : \ \|u - u_h\|_I \leq r_I, \ \|u - u_h\|_A \leq r_A\}.$$

Then B is bounded, closed, and non-empty. Hence, by the Weierstraß-Theorem, we have that the minimization problem

$$\min_{u \in U_{ad} \cap B} f(u)$$

admits a solution $\bar{u}$. Let us show that both of the constraints are not active on $\bar{u}$. Let us define

$$\rho_I := \|u_h - \bar{u}\|_I, \quad \rho_A := \|u_h - \bar{u}\|_A.$$

At first, we assume that $\|\bar{u} - u_h\|_A = r_A$ holds. Then according to (2.12) and (2.13) we have

$$\begin{aligned} f(\bar{u}) - f(u_h) &\geq \sigma r_A - c_{f'} r_A \rho_I - \frac{M_{f''}}{2} r_A^2 - \epsilon \rho_I + \frac{\alpha}{2}\rho_I^2 - \frac{c_{f''}}{6}\rho_I^3 \\ &\geq \sigma r_A - c_{f'} r_A r_I - \frac{M_{f''}}{2} r_A^2 - \epsilon r_I - \frac{c_{f''}}{6} r_I^3 > 0, \end{aligned}$$

which yields a contradiction, since by optimality of $\bar{u}$ we have $f(\bar{u}) - f(u_h) \leq 0$.

Second, suppose that it holds $\|\bar{u} - u_h\|_I = r_I$. Similarly as above, we get

$$f(\bar{u}) - f(u_h) \geq \frac{\alpha}{2} r_I^2 - c_{f'} r_A r_I - \frac{M_{f''}}{2} r_A^2 - \epsilon r_I - \frac{c_{f''}}{6} r_I^3 > 0,$$

which gives a contradiction as well. This proves that $\bar{u}$ lies in the interior of B, making it a local solution of the original problem. It remains to show that $\bar{u}$ satisfies SSC.

Let us take $u \in U_{ad}$ and define

$$\tilde{u} = \begin{cases} u_h & \text{on } I, \\ u & \text{on } A. \end{cases}$$

Then we can estimate due to (2.8) and (2.14)

$$f'(\bar{u})(\tilde{u}-\bar{u})|_A \geq f'(\bar{u})(\tilde{u}-u_h)|_A = (f'(\bar{u})-f'(u_h))(\tilde{u}-u_h)|_A + f'(u_h)(\tilde{u}-u_h)|_A \geq (-c_{f'} r_+ + \sigma)\|u-u_h\|_A.$$

Hence, $|f'(\bar{u})| > 0$ on A. Moreover, by optimality of $\bar{u}$

$$0 \leq f'(\bar{u})(u_h - \bar{u})|_A \leq (f'(\bar{u}) - f'(u_h))(u_h - \bar{u})|_A + f'(u_h)(u_h - \bar{u})|_A \leq (c_{f'} r_+ - \sigma)\|u_h - \bar{u}\|_A \leq 0,$$

which proves $\bar{u} = u_h$ on the active set A. Hence $\rho_A = 0$ and $\rho_I = r_+$ holds.

Similarly by (2.9) and (2.14), we obtain

$$f''(\bar{u})[(v,v)] \geq (\alpha - c_{f''}\rho_I)\|v\|_U^2 \geq (\alpha - c_{f''} r_I)\|v\|_U^2$$

implying the fulfillment of the positivity condition (2.4) and the coercivity condition (2.5) at the unknown solution $\bar{u}$. □

Under a condition slightly different from Theorem 2.1, a similar result can be obtained. Moreover, the conditions are more accessible than the ones from Theorem 2.1.

THEOREM 2.2. *Let Assumption 3 be satisfied. Let us suppose that there exist $r_I, r_A > 0$ such that with $r_+ := \sqrt{r_I^2 + r_A^2} < R$ it holds*

$$-\epsilon + \alpha r_I - \frac{c_{f''}}{2} r_I^2 > 0, \tag{2.15}$$

$$\sigma - c_{f'} r_+ > 0. \tag{2.16}$$

Then there exists a local optimal control $\bar{u}$ to the original problem (2.2) with

$$\|\bar{u} - u_h\|_U < r_+.$$

Moreover $\bar{u} = u_h$ on A and the second order sufficient optimality condition (2.4)–(2.5) holds at $\bar{u}$.

Proof. Let us define $B := \{u \in U : \ \|u - u_h\|_I \leq r_I, \ \|u - u_h\|_A \leq r_A\}$. As argued in the proof of Theorem 2.1, we have that the minimization problem $\min_{u \in U_{ad} \cap B} f(u)$ admits a solution $\bar{u}$. Again, it remains to prove that $\bar{u}$ is a local solution of the original problem. To this end, we will show that $\bar{u}$ lies not on the boundary of the ball B. We write

$$f'(\bar{u})(u_h - \bar{u}) = f'(\bar{u})(u_h - \bar{u})|_I + f'(\bar{u})(u_h - \bar{u})|_A,$$

and we will estimate the derivative of f on the active and inactive set separately. Let us define

$$\rho_+ := \|u_h - \bar{u}\|, \quad \rho_I := \|u_h - \bar{u}\|_I, \quad \rho_A := \|u_h - \bar{u}\|_A,$$

which implies $\rho_+ \leq r_+, \rho_I \leq r_I$, and $\rho_A \leq r_A$. For the active set, we obtain

$$\begin{aligned} f'(\bar{u})(u_h - \bar{u})|_A &= f'(u_h)(u_h - \bar{u})|_A + (f'(\bar{u}) - f'(u_h))(u_h - \bar{u})|_A \\ &\leq \rho_A(-\sigma + c_{f'}\rho_+). \end{aligned}$$

The contribution on the inactive set can be estimated as

$$\begin{aligned} f'(\bar{u})(u_h - \bar{u})|_I &= f'(u_h)(u_h - \bar{u})|_I + f''(u_h)[(u_h - \bar{u})|_I, \bar{u} - u_h] \\ &\quad + \int_0^1 (f''(u_h + s(\bar{u} - u_h)) - f''(u_h))[(u_h - \bar{u})|_I, \bar{u} - u_h]ds \\ &\leq \epsilon\rho_I - \alpha\rho_I^2 + M_{f''}\rho_I\rho_A + \frac{1}{2}c_{f''}\rho_I\rho_+^2 \\ &\leq \rho_I(\epsilon - \alpha\rho_I + M_{f''}\rho_A + \frac{1}{2}c_{f''}\rho_+^2) \end{aligned}$$

Furthermore, by the necessary optimality conditions, we have

$$f'(\bar{u})(u_h - \bar{u}) \geq 0.$$

First, let us assume that the constraint $\|u - u_h\|_A \leq r_A$ is active at $\bar{u}$, i.e. $\|\bar{u} - u_h\|_A = r_A > 0$. Then we obtain

$$0 \leq f'(\bar{u})(u_h - \bar{u})|_A \leq \rho_A(-\sigma + c_{f'}\rho_+) \leq r_A(-\sigma + c_{f'}r_+) < 0$$

by assumption (2.16), which is a contradiction. Moreover, this estimate implies $\rho_A = 0$, which means that $\bar{u} = u_h$ on the active set. Hence $\rho_+ = \rho_I$.

Second, let us assume that the the constraint $\|u - u_h\|_I \leq r_I$ is active at $\bar{u}$, i.e. $\|\bar{u} - u_h\|_I = r_I > 0$. Then it holds

$$0 \leq f'(\bar{u})(u_h - \bar{u})|_I \leq \rho_I(\epsilon - \alpha r_I + \frac{1}{2}c_{f''}r_I^2) < 0$$

by (2.15), which proves $\|\bar{u} - u_h\|_I < \rho_I$ by a similar reasoning as on the active set.

This implies that $\bar{u}$ is an interior point of B, and hence a local solution of the original problem. The inequality (2.15) implies the convexity condition (2.4), which can be proven as in [15, Theorem 2.5]. ◻

Let us now prove a much more explicit error bound.

Corollary 2.3. *Let the assumptions of Theorem 2.2 be satisfied. Then it holds*

$$\|\bar{u} - u_h\|_U \leq \frac{2\epsilon}{\alpha}.$$

Moreover, $f''(\bar{u})[v, v] \geq (\alpha - \|\bar{u} - u_h\|_U c_{f''}) \|v\|_U^2 > 0$ *for all* $v \in U$ *with* $v = 0$ *on* A.

Here, one can see clearly that our results imply $\|\bar{u} - u_h\|_U \to 0$ provided $\epsilon \to 0$ and SSC holds at $\bar{u}$. Surprisingly, this estimate does not involve any other of the constants present in Assumption 3 as e.g. $c_{f''}$.

Proof. By assumption, the polynomial $-\epsilon + \alpha r_I - \frac{c_{f''}}{2}r_I^2$ in (2.15) has a positive root. Since the polynomial is negative at $r_I = 0$, this implies that all roots are positive. The smallest root can be computed as

$$\tilde{r}_I = \frac{\alpha - \sqrt{\alpha^2 - 2\epsilon c_{f''}}}{c_{f''}},$$

which implies $\alpha^2 - 2\epsilon c_{f''} > 0$. If $\alpha^2 - 2\epsilon c_{f''} < 0$ then the polynomial would be strictly negative, which is a contradiction to the assumption (2.15). By elementary calculations, we find

$$\tilde{r}_I = \frac{2\epsilon}{\alpha + \sqrt{\alpha^2 - 2\epsilon c_{f''}}} \leq \frac{2\epsilon}{\alpha}.$$

The claim on the second derivative follows immediately. ◻

Let us close this section with the following observation, which gives a sufficient condition for the assumptions of Theorem 2.2. Moreover, these conditions are easier to check and independent of $M_{f''}$. In addition, they highlight the fact that if $\alpha, \sigma, c_{f'}, c_{f''}$ stay bounded, the assumption of Theorem 2.2 is satisfied if the discretization error ϵ goes to zero, which is guaranteed at least for uniform mesh refinement.

COROLLARY 2.4. *Let Assumption 3 be satisfied. If*

$$\alpha^2 - 2c_{f''}\epsilon > 0, \tag{2.17}$$

$$\alpha\sigma - 2c_{f'}\epsilon > 0 \tag{2.18}$$

holds then the assumptions of Theorem 2.2 are satisfied.

Proof. As argued in the proof of the previous Corollary 2.3, condition (2.17) is sufficient for (2.15). Moreover, there the bound $r_I \leq 2\epsilon/\alpha$ was proven. Then we obtain

$$\sigma - c_{f'}r_+ \geq \sigma - c_{f'}(\frac{2\epsilon}{\alpha} + r_A) = \alpha^{-1}(\alpha\sigma - 2c_{f'}\epsilon) - c_{f'}r_A,$$

which shows that due to (2.18) we can choose $r_A > 0$ to satisfy (2.16). □

3. Application to the abstract problem.

Here, we will transform Assumption 3 in the previous section to assumptions on the solution (y_h, u_h, p_h) of the abstract problem (P). First we derive the Fréchet derivatives of the reduced functional f involving the abstract PDE operator $E(y, u)$.

Let us recall the definition of the reduced functional (2.1): $f(u) = g(S(u)) + j(u)$. Then the first derivative of the reduced functional is obtained as

$$f'(u)v = g'(S(u))z + j'(u) \cdot v \qquad \forall v \in U, \tag{3.1}$$

where $z = S'(u)v$. With p being the solution of $E_y(S(u), u)^*p = g'(S(u))$, an obvious computation gives the equality $g'(S(u))z = -\langle E_u(S(u), u)p,\, v\rangle_{U^*,U}$, which we use in (3.1) to obtain

$$f'(u)v = \langle -E_u(S(u), u)^*p + j'(u), v\rangle_{U^*,U}. \tag{3.2}$$

Similarly for $v_1, v_2 \in U$, taking the derivative of (3.2) yields

$$f''(u)[v_1,\, v_2] = g''(S(u))[S'(u)v_1,\, S'(u)v_2] + g'(S(u))S''(u)[v_1,\, v_2] + j''(u)[v_1,\, v_2]. \tag{3.3}$$

Since $S'(u)v = z$ solves

$$E_y(S(u), u)z + E_u(S(u), u)v = 0$$

it follows that $S''(u)[v_1, v_2] = \zeta$ is a solution of

$$E_y(S(u), u)\zeta + E''(S(u), u)[(v_1, S'(u)v_1),\, (v_2, S'(u)v_2)] = 0.$$

Again it can be shown that

$$g'(S(u))\zeta = -\langle E''(S(u), u)[(v_1, S'(u)v_1),\, (v_2, S'(u)v_2)],\, p\rangle_{Y^*,Y}. \tag{3.4}$$

Now using (3.4) in (3.3) yields

$$f''(u)[v_1,\, v_2] = g''(S(u))[z_1,\, z_2] + j''(u)[v_1,\, v_2] - \langle E''(S(u), u)[(v_1, z_1),\, (v_2, z_2)],\, p\rangle_{Y^*,Y}$$

with $z_i = S'(u)v_i$, $i = 1, 2$.

By Assumption 1, the functions E, j, g are twice Fréchet-differentiable with Lipschitz continuous second derivatives. In the sequel, we will need the associated Lipschitz constants. In order to get a compact as possible notation, we will introduce a short-hand notation of bounds of bilinear forms. If $G : X \times X \to Z$ is a bounded bilinear form, then

$$\|G\|_{\mathcal{B}(X,Z)} := \|G\|_{\mathcal{L}(X,\mathcal{L}(X,Z))} = \sup_{\|x_1\|_X = \|x_2\|_X = 1} \|G(x_1, x_2)\|_Z$$

is the associated bound of the bilinear form.

Let us fix the Lipschitz constants and bounds of derivatives with the following assumption.

Assumption 4. *Let R be a positive constant. We assume there are positive constants c_E, c_{E_y}, c_{E_u}, $c_{E''}$, $c_{g'}$, $c_{g''}$, $c_{j''}$ depending on R such that the estimates*

$$\|E(y,u) - E(y^h,u_h)\|_{Y^*} \le c_E(\|y^h - y\|_Y + \|u - u_h\|_U), \tag{3.5a}$$
$$\|E_y(y,u) - E_y(y^h,u_h)\|_{\mathcal{L}(Y,Y^*)} \le c_{E_y}(\|y^h - y\|_Y + \|u - u_h\|_U), \tag{3.5b}$$
$$\|E_u(y,u) - E_u(y^h,u_h)\|_{\mathcal{L}(U,Y^*)} \le c_{E_u}(\|y^h - y\|_Y + \|u - u_h\|_U), \tag{3.5c}$$
$$\|E''(y,u) - E''(y^h,u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \le c_{E''}(\|y^h - y\|_Y + \|u - u_h\|_U), \tag{3.5d}$$
$$\|g'(y) - g'(y^h)\|_{Y^*} \le c_{g'}\|y - y^h\|_Y, \tag{3.5e}$$
$$\|g''(y) - g''(y^h)\|_{(Y\times Y)^*} \le c_{g''}\|y - y^h\|_Y, \tag{3.5f}$$
$$\|j'(u) - j'(u_h)\|_{(U\times U)^*} \le c_{j'}\|u - u_h\|_U, \tag{3.5g}$$
$$\|j''(u) - j''(u_h)\|_{(U\times U)^*} \le c_{j''}\|u - u_h\|_U \tag{3.5h}$$

hold for all $u \in U_{ad}$, $\|u - u_h\|_U \le R$, $y = S(u)$, and $y^h = S(u_h)$ or $y^h = y_h$.

As already indicated in Section 1.1, we can relax the differentiability requirements for E and g, i.e. it is sufficient to have E and g to be Fréchet-differentiable with respect to a stronger space $Y_\infty \hookrightarrow Y$. Here, we have in mind to take $Y = H^1_0(\Omega)$ and $Y_\infty = H^1_0(\Omega) \cap L^\infty(\Omega)$. However, we still need the Lipschitz continuity w.r.t. y in the (weaker) space Y. Otherwise it would be necessary to have computable errors of y and p in Y_∞, which seems to be impossible for e.g. $Y_\infty = H^1_0(\Omega) \cap L^\infty(\Omega)$.

Let us recall the statement of Assumption 2:

$$\|E(y_h,u_h)\|_{Y^*} \le r_y,$$
$$\|g'(y_h) - E_y(y_h,u_h)^* p_h\|_{Y^*} \le r_p,$$
$$\langle j'(u_h) - E_u(y_h,u_h)^* p_h, u - u_h\rangle \ge -r_u\|u - u_h\|_U \;\forall u \in U_{ad}.$$

In the remainder of this section, we will express the constants in Assumption 3 by means of the residuals of Assumption 2 and the constants of Assumption 4.

3.1. Error estimates for state and adjoint equation, estimates for auxiliary functions. Let y^h and p^h be auxiliary variables that solve

$$E(y^h,u_h) = 0, \tag{3.6}$$
$$E_y(y^h,u_h)^* p^h = g'(y^h). \tag{3.7}$$

respectively. The following estimates hold for the introduced state and adjoint variables.

Lemma 3.1. *Let y^h, p^h be given by (3.6) and (3.7) respectively. Then it holds*

$$\|y^h - y_h\|_Y \le \epsilon_y, \tag{3.8}$$
$$\|p^h - p_h\|_Y \le \epsilon_p \tag{3.9}$$

with $\epsilon_y = \delta^{-1} r_y$ and $\epsilon_p = \delta^{-1}\left(c_{g'}\epsilon_y + r_p + c_{E_y}\epsilon_y\|p_h\|_Y\right)$.

Proof. Since y^h is a solution of the nonlinear equation, we have $E(y^h,u_h) - E(y_h,u_h) = -E(y_h,u_h)$. Testing this equation with $y^h - y_h$ and using the strong monotonicity of E we obtain

$$\delta\|y^h - y_h\|_Y^2 \le \|E(y_h,u_h)\|_{Y^*}\;\|y^h - y_h\|_Y.$$

The result then follows by using the residual estimate (1.6).

For estimating the adjoint state, observe that $p^h - p_h$ fulfills

$$E_y(y^h,u_h)^*(p^h - p_h) = g'(y^h) - g'(y_h) + (g'(y_h) - E_y(y_h,u_h)^* p_h) + \left(E_y(y_h,u_h)^* - E_y(y^h,u_h)^*\right) p_h.$$

Hence by Lipschitz properties (3.5e) and (3.5b) of g' and E_y, respectively, and the residual estimates (1.7) and (3.8), we obtain

$$\begin{aligned}
\|p^h - p_h\|_Y &\le \delta^{-1}\left(\|g'(y^h) - g'(y_h)\|_{Y^*} + \|g'(y_h) - E_y(y_h,u_h)^* p_h\|_{Y^*}\right) \\
&\quad + \delta^{-1}\|E_y(y_h,u_h) - E_y(y^h,u_h)\|_{\mathcal{L}(Y,Y^*)}\|p_h\|_Y \\
&\le \delta^{-1}\left(c_{g'}\|y_h - y^h\|_Y + r_p + c_{E_y}\|y^h - y_h\|_Y\|p_h\|_Y\right) \\
&\le \delta^{-1}\left(c_{g'}\epsilon_y + r_p + c_{E_y}\epsilon_y\|p_h\|_Y\right).
\end{aligned}$$

□

As one can see in the estimates above, it holds $\|y^h - y_h\|_Y \to 0$ and $\|p^h - p_h\|_Y \to 0$ if $r_y, r_p \to 0$. In addition to these results above, we derive bounds for the norms of y^h and p^h, which will turn out useful in the sequel.

COROLLARY 3.2. *Let y^h, p^h be as defined in* (3.6) (3.7) *respectively. Then it holds*

$$\|y^h\|_Y \leq M_y, \tag{3.10}$$

$$\|p^h\|_Y \leq M_p \tag{3.11}$$

with $M_y = \epsilon_y + \|y_h\|_Y$ and $M_p = \epsilon_p + \|p_h\|_Y$.

Proof. The claim is an easy consequence of Lemma 3.1 and the triangle inequality. □

3.2. Lipschitz estimate of f', computation of $c_{f'}$. Let $u \in U_{ad}$ be given with $\|u - u_h\|_U \leq R$. We define the associated state y and adjoint state p through

$$E(y, u) = 0 \tag{3.12}$$

$$E_y(y, u)^* p = g'(y). \tag{3.13}$$

In order to obtain the Lipschitz estimates for f' and f'' we have to estimate the differences $y - y^h$ and $p - p^h$.

LEMMA 3.3. *Let $u \in U_{ad}$ be given and y, p be the associated state and adjoint state solving* (3.12) *and* (3.13) *respectively. Then it holds*

$$\|y - y^h\|_Y \leq c_y \|u - u_h\|_U, \tag{3.14}$$

$$\|p - p^h\|_Y \leq c_p \|u - u_h\|_U \tag{3.15}$$

with

$$c_y = \delta^{-1} c_E,$$
$$c_p = \delta^{-1} \left(c_{g'} c_y + M_p c_{E_y} (c_y + 1) \right).$$

Proof. The functions y and y^h are the solutions of $E(y, u) = 0$ and $E(y^h, u_h) = 0$ respectively. Therefore we can write

$$\langle E(y, u) - E(y^h, u), y - y^h \rangle_{Y^*,Y} = \langle E(y^h, u_h) - E(y^h, u), y - y^h \rangle_{Y^*,Y}$$

By the monotonicity assumption on E we obtain using (3.5a)

$$\begin{aligned} \delta \|y - y^h\|_Y^2 &\leq \|E(y^h, u_h) - E(y^h, u)\|_{Y^*} \|y - y^h\|_Y \\ &\leq c_E \|u - u_h\|_U \|y - y^h\|_Y \end{aligned}$$

which gives the first estimate.

For the second estimate, recall that p and p^h are the solutions to the equations $E_y(y, u)^* p = g'(y)$ and $E_y(y^h, u_h)^* p^h = g'(y^h)$, respectively. Hence, the difference $p - p^h$ fulfills

$$E_y(y, u)^* (p - p^h) = g'(y) - g'(y^h) + \left(E_y(y^h, u_h)^* - E_y(y, u)^* \right) p^h$$

Using the Lipschitz estimates (3.5e),(3.5b) together with (3.11) and (3.14) we obtain the a-priori estimate

$$\begin{aligned} \|p - p^h\|_Y &\leq \delta^{-1} \left(\|g'(y) - g'(y^h)\|_{Y^*} + \|E_y(y^h, u_h) - E_y(y^h, u)\|_{\mathcal{L}(Y,Y^*)} \|p^h\|_Y \right) \\ &\leq \delta^{-1} \left(c_{g'} \|y - y^h\|_Y + c_{E_y} \left(\|y - y^h\|_Y + \|u - u_h\|_U \right) M_p \right) \\ &\leq \delta^{-1} \left(c_{g'} c_y \|u - u_h\|_U + c_{E_y} (c_y + 1) \|u - u_h\|_U M_p \right). \end{aligned}$$

□

LEMMA 3.4. *Let $u \in U_{ad}$ be given such that $\|u - u_h\|_U \leq R$ and set $v = u - u_h$. Then the first derivative f' of the reduced functional satisfies the Lipschitz estimate*

$$\|f'(u) - f'(u_h)\|_{U^*} \leq c_{f'} \|u - u_h\|_U$$

with

$$c_{f'} = c_{E_u} (c_y + 1)(R c_p + M_p) + c_p \left(\epsilon_y c_{E_u} + \|E_u(y_h, u_h)\|_{\mathcal{L}(U,Y^*)} \right) + c_{j'}.$$

Proof. Let p^h be as in (3.7). By previous computation $f'(u_h)v = \langle -E_u(y^h,u_h)^* p^h + j'(u_h), v\rangle_{U^*,U}$. Using the Lipschitz estimate (3.5g) we have

$$\begin{aligned}\|f'(u) - f'(u_h)\|_{U^*} &\le \|E_u(y^h,u_h)^*p^h - E_u(y,u)^*p + j'(u) - j'(u_h)\|_{U^*} \\ &\le \|E_u(y^h,u_h)^*p^h - E_u(y,u)^*p\|_{U^*}\|v\|_U + \|j'(u) - j'(u_h)\|_{U^*} \\ &\le \|E_u(y^h,u_h)^*p^h - E_u(y,u)^*p\|_{U^*}\|v\|_U + c_{j'}\|u-u_h\|_U.\end{aligned} \tag{3.16}$$

By adopting the splitting

$$\begin{aligned}E_u(y^h,u_h)^*p^h - E_u(y,u)^*p &= E_u(y,u)^*(p^h-p) + \left(E_u(y^h,u_h)^* - E_u(y,u)^*\right)p^h \\ &= \left(E_u(y,u) - E_u(y^h,u_h) + E_u(y^h,u_h) - E_u(y_h,u_h) + E_u(y_h,u_h)\right)^*(p^h-p) \\ &\quad + \left(E_u(y^h,u_h)^* - E_u(y,u)^*\right)p^h\end{aligned}$$

we estimate

$$\begin{aligned}\|E_u(y^h,u_h)^*p^h - E_u(y,u)^*p\|_{U^*} &\le \|E_u(y,u) - E_u(y^h,u_h)\|_{\mathcal{L}(U,Y^*)}\|p^h-p\|_Y \\ &\quad + \left(\|E_u(y^h,u_h) - E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)} + \|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)\|p^h-p\|_Y \\ &\quad + \|E_u(y^h,u_h)^* - E_u(y,u)^*\|_{\mathcal{L}(U,Y^*)}\|p^h\|_Y.\end{aligned}$$

Thanks to property (3.5c) of E_u, (3.14) and the residual estimate (3.8), it holds

$$\|E_u(y,u) - E_u(y^h,u_h)\|_{\mathcal{L}(U,Y^*)} \le c_{E_u}(\|y-y^h\|_Y + \|u-u_h\|_U) \le c_{E_u}(c_y+1)R \tag{3.17}$$

and

$$\|E_u(y^h,u_h) - E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)} \le c_{E_u}\|y^h-y_h\|_Y \le c_{E_u}\epsilon_y. \tag{3.18}$$

Hence by (3.15) and (3.11)

$$\begin{aligned}\|E_u(y^h,u_h)^*p^h - E_u(y,u)^*p\|_{U^*} &\le \left(c_{E_u}(c_y+1)R + c_{E_u}\epsilon_y + \|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)c_p\|u-u_h\|_U \\ &\quad + c_{E_u}(c_y+1)M_p\|u-u_h\|_U.\end{aligned} \tag{3.19}$$

Finally substituting (3.19) in (3.16) yields the desired estimate. ☐

3.3. Estimates for $f'(u_h)$, computation of ϵ and σ. After having computed the Lipschitz constant of f', we will now derive bounds for the constants ϵ and σ appearing in the first-order part (2.6) of Assumption 3.

At first, we will estimate the difference between $f'(u_h)$ and the gradient of the discrete problem defined by

$$f_h'(u_h) := -E_u(y_h,u_h)^*p_h + j'(u_h). \tag{3.20}$$

Here, we have the following.

LEMMA 3.5. *Let $f_h'(u_h)$ be defined by (3.20). Then it holds*

$$\|f'(u_h) - f_h'(u_h)\|_{U^*} \le \epsilon_{f'} \tag{3.21}$$

with

$$\epsilon_{f'} := c_{E_u}\epsilon_y M_p + \epsilon_p\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}.$$

Proof. Let $v \in U$. We estimate the difference

$$\begin{aligned}\|f'(u_h) - f_h'(u_h)\|_{U^*} &= \|E_u(y_h,u_h)^*p_h - E_u(y^h,u_h)^*p^h\|_{U^*} \\ &\le \|E_u(y_h,u_h) - E_u(y^h,u_h)\|_{\mathcal{L}(U,Y^*)}\|p^h\|_Y + \|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\|p_h-p^h\|_Y.\end{aligned}$$

Using (3.18), and the bounds of p^h and $p_h - p^h$ in (3.11) and (3.9), respectively, we have the estimate

$$\|E_u(y_h,u_h)^*p_h - E_u(y^h,u_h)^*p^h\|_{U^*} \le c_{E_u}\epsilon_y M_p + \epsilon_p\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}, \tag{3.22}$$

which proves (3.21). ∎

The estimate for the positivity constant σ can now be computed easily thanks to the result of the previous Lemma 3.5.

LEMMA 3.6. *For all admissible control $u \in U_{ad}$ with $\|u - u_h\|_U < R$ the following inequality holds on the active set of u_h*

$$f'(u_h)(u - u_h)|_A \geq \sigma \|u - u_h\|_A$$

where

$$\sigma := \sigma_h - \epsilon_{f'},$$
$$\sigma_h := \min_{i \in A} |f_h'(u_h)_i|.$$

Proof. We write

$$f'(u_h)(u-u_h)|_A = (f'(u_h) - f_h'(u_h))(u-u_h)|_A + f_h'(u_h)(u-u_h)|_A \geq (\sigma_h - \|f'(u_h) - f_h'(u_h)\|)\|u-u_h\|_A,$$

which yields the result upon applying the estimate (3.21) provided by Lemma 3.5. ∎

Furthermore, we have the following estimate for the first derivative of f on the inactive set.

LEMMA 3.7. *For all admissible control $u \in U_{ad}$ with $\|u - u_h\|_U < R$ the following inequality holds on the inactive set of u_h*

$$f'(u_h)(u - u_h)|_I \geq -\epsilon \|u - u_h\|_I$$

where

$$\epsilon := c_{E_u} \epsilon_y M_p + \epsilon_p \|E_u(y_h, u_h)\|_{\mathcal{L}(U, Y^*)} + r_u.$$

Proof. Applying the residual estimate (1.8) it holds

$$\begin{aligned} f'(u_h)(u - u_h)|_I &= \langle -E_u(y^h, u_h)^* p^h + j'(u_h), u - u_h \rangle_{U^*, U} \\ &= \langle E_u(y_h, u_h)^* p_h - E_u(y^h, u_h)^* p^h, u - u_h \rangle_{U^*, U} + \langle j'(u_h) - E_u(y_h, u_h)^* p_h, u - u_h \rangle_{U^*, U} \\ &\geq -\|E_u(y_h, u_h)^* p_h - E_u(y^h, u_h)^* p^h\|_{U^*} \|u - u_h\|_U - r_u \|u - u_h\|_U. \end{aligned}$$

The estimate for the leading term is given by (3.22), which gives

$$f'(u_h)(u - u_h)|_I \geq -\left(c_{E_u} \epsilon_y M_p + \epsilon_p \|E_u(y_h, u_h)\|_{\mathcal{L}(U, Y^*)} + r_u \right) \|u - u_h\|_U.$$

∎

3.4. Estimates of f'', computation of $c_{f''}$ and $M_{f''}$. Let us now turn to the relevant estimate for the second derivative f'', which was derived above as

$$f''(u)[v_1, v_2] = g''(S(u))[z_1, z_2] + j''(u)[v_1, v_2] - \langle E''(S(u), u)[(v_1, z_1), (v_2, z_2)], p \rangle_{Y^*, Y}$$

with p solving (3.13). Obviously, any change in u will not only change the argument of g'', j'', and E'', but also it will change the point, where the linearization of the solution operator S is made. This necessitates analysis of $S'(u) - S'(u_h)$ in order to be able to estimate $f''(u) - f''(u_h)$.

Let $v_i \in U$, $i = 1, 2$ be given. In the sequel, unless otherwise stated, z_i, z_i^h are defined as the solutions of

$$E_y(y, u) z_i + E_u(y, u) v_i = 0, \tag{3.23}$$
$$E_y(y^h, u_h) z_i^h + E_u(y^h, u_h) v_i = 0, \tag{3.24}$$

respectively. That is, we have $z_i = S'(u) v_i$ and $z_i^h = S'(u_h) v_i$. To be able to find the Lipschitz estimate for the second derivative f'', we derive the following useful estimates.

LEMMA 3.8. *Let $v_i \in U$ be given. Let z_i, z_i^h be defined by (3.23) and (3.24) respectively. Then for $u \in U_{ad}$, $\|u - u_h\|_U \leq R$ it holds*

$$\|z_i\|_Y \leq M_z \|v_i\|_U, \tag{3.25}$$
$$\|z_i^h\|_Y \leq M_{z^h} \|v_i\|_U, \tag{3.26}$$
$$\|z_i - z_i^h\|_Y \leq c_z \|u - u_h\|_U \|v_i\|_U. \tag{3.27}$$

The bounds are derived in the course of the proof.

Proof. First, testing (3.24) by z_i^h one finds the a priori estimate

$$\begin{aligned}\|z_i^h\|_Y &\le \delta^{-1}\|E_u(y^h,u_h)\|_{\mathcal{L}(U,Y^*)}\|v_i\|_U \\ &\le \delta^{-1}\left(\|E_u(y^h,u_h)-E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}+\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)\|v_i\|_U.\end{aligned}$$

Due to (3.18), we obtain

$$\|z_i^h\|_Y \le \delta^{-1}\left(c_{E_u}\epsilon_y+\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)\|v_i\|_U,$$

which implies $M_{z^h}=\delta^{-1}\left(c_{E_u}\epsilon_y+\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)$. Similarly by using the already obtained estimates (3.17) and (3.18), we obtain from (3.23)

$$\begin{aligned}\|z_i\|_Y &\le \delta^{-1}\|E_u(y,u)\|_{\mathcal{L}(U,Y^*)}\|v_i\|_U \\ &\le \delta^{-1}\left(\|E_u(y,u)-E_u(y^h,u_h)\|_{\mathcal{L}(U,Y^*)}+\|E_u(y^h,u_h)-E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right. \\ &\qquad \left.+\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)\|v_i\|_U \\ &\le \delta^{-1}\left(c_{E_u}R(c_y+1)+c_{E_u}\epsilon_y+\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)\|v_i\|_U,\end{aligned}$$

which gives $M_z=\delta^{-1}\left(c_{E_u}R(c_y+1)+c_{E_u}\epsilon_y+\|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right)$. To estimate $z_i-z_i^h$, observe that $z_i-z_i^h$ fulfills

$$E_y(y,u)(z_i-z_i^h)=\left(E_u(y^h,u_h)-E_u(y,u)\right)v_i+\left(E_y(y^h,u_h)-E_y(y,u)\right)z_i^h.$$

This implies the a priori estimate

$$\|z_i-z_i^h\|_Y \le \delta^{-1}\|E_u(y^h,u_h)-E_u(y,u)\|_{\mathcal{L}(U,Y^*)}\|v_i\|_U+\|E_y(y^h,u_h)-E_y(y,u)\|_{\mathcal{L}(Y,Y^*)}\|z_i^h\|_Y.$$

Employing (3.5c) and (3.5b) in estimating the first and second addends respectively, and using the estimates (3.14), (3.26) gives

$$\begin{aligned}\|z_i-z_i^h\|_Y &\le \delta^{-1}c_{E_u}\left(\|y^h-y\|_Y+\|u_h-u\|_U\right)\|v_i\|_U+c_{E_y}\left(\|y^h-y\|_Y+\|u_h-u\|_U\right)M_{z^h}\|v_i\|_U \\ &\le \delta^{-1}c_{E_u}(c_y+1)\|u_h-u\|_U\|v_i\|_U+c_{E_y}(c_y+1)M_{z^h}\|u_h-u\|_U\|v_i\|_U,\end{aligned}$$

which yields the last estimate (3.27) with $c_z=\delta^{-1}(c_y+1)(c_{E_u}+c_{E_y}M_{z^h})$. ∎

Now we are ready to do the first step in estimating $f''(u)-f''(u_h)$.

LEMMA 3.9. *Let $u\in U_{ad}$ and $v_i\in U$, $i=1,2$, be given. Let z_i, z_i^h, $i=1,2$ be defined as in the previous lemma. Then it holds*

$$|g''(y^h)[z_1^h,z_2^h]-g''(y)[z_1,z_2]|\le C_{g''}\|u-u_h\|_U\|v_1\|_U\|v_2\|_U$$

with

$$C_{g''}=c_{g''}c_yM_z^2+c_z\left(c_{g''}\epsilon_y+\|g''(y_h)\|_{(Y\times Y)^*}\right)(M_{z^h}+M_z).$$

Proof. We split

$$\begin{aligned}&g''(y^h)[z_1^h,z_2^h]-g''(y)[z_1,z_2] \\ &\qquad=\left(g''(y^h)-g''(y)\right)[z_1,z_2]+\left(g''(y^h)-g''(y_h)+g''(y_h)\right)\left([z_1^h,z_2^h]-[z_1,z_2]\right) \\ &\qquad=\left(g''(y^h)-g''(y)\right)[z_1,z_2]+\left(g''(y^h)-g''(y_h)+g''(y_h)\right)\left([z_1^h-z_1,z_2^h]+[z_1,z_2^h-z_2]\right),\end{aligned}$$

so that we can estimate

$$\begin{aligned}&|g''(y^h)[z_1^h,z_2^h]-g''(y)[z_1,z_2]| \\ &\quad\le \|g''(y^h)-g''(y)\|_{(Y\times Y)^*}\|z_1\|_Y\|z_2\|_Y \\ &\qquad+\left(\|g''(y^h)-g''(y_h)\|_{(Y\times Y)^*}+\|g''(y_h)\|_{(Y\times Y)^*}\right)\left(\|z_1^h-z_1\|_Y\|z_2^h\|_Y+\|z_2^h-z_2\|_Y\|z_1\|_Y\right).\end{aligned}$$

Now upon applying the estimates (3.25)-(3.27) in Lemma 3.8, and the Lipschitz estimate (3.5f) of g'', we obtain

$$\begin{aligned}|g''(y^h)[z_1^h, z_2^h] - g''(y)[z_1, z_2]| &\le c_{g''}\|y^h - y\|_Y M_z^2 \|v_1\|_U\|v_2\|_U \\ &\quad + \left(c_{g''}\|y^h - y_h\|_Y + \|g''(y_h)\|_{(Y\times Y)^*}\right)\left(c_z M_z^h + c_z M_z\right)\|u - u_h\|_U\|v_1\|_U\|v_2\|_U.\end{aligned}$$

Employing the estimates (3.14) and (3.8) for $\|y^h - y\|_Y$ and $\|y^h - y_h\|_Y$ we obtain finally

$$\begin{aligned}&|g''(y^h)[z_1^h, z_2^h] - g''(y)[z_1, z_2]| \\ &\qquad \le \left(c_{g''}c_y M_z^2 + c_z\left(c_{g''}\epsilon_y + \|g''(y_h)\|_{(Y\times Y)^*}\right)(M_{z^h} + M_z)\right)\|u - u_h\|_U\|v_1\|_U\|v_2\|_U,\end{aligned}$$

which is the desired result. ☐

In order to simplify the exposition of the Lipschitz estimates of the part of f'' that involves E'', we present the next lemma, where we have in mind to choose $G = E''(y, u)$.

Lemma 3.10. *Let G be a bounded bilinear form on the space $U\times Y$, i.e., $G : (U\times Y)\times(U\times Y) \mapsto Y^*$. Let $v_i \in U$ and for $i = 1, 2$ let z_i, z_i^h be defined by (3.23) and (3.24) respectively. Then for the pairs $d^h = [(v_1, z_1^h), (v_2, z_2^h)]$ and $d = [(v_1, z_1), (v_2, z_2)]$ of directions it holds*

$$\|G(d)\|_{Y^*} \le M_d\,\|G\|_{\mathcal{B}(U\times Y, Y^*)}\|v_1\|_U\|v_2\|_U, \tag{3.28}$$

$$\|G(d^h) - G(d)\|_{Y^*} \le c_d\,\|G\|_{\mathcal{B}(U\times Y, Y^*)}\|u - u_h\|_U\|v_1\|_U\|v_2\|_U \tag{3.29}$$

where

$$\begin{aligned}M_d &= (1 + M_z)^2, \\ c_d &= c_z\left(2 + M_z + M_{z^h}\right).\end{aligned}$$

Proof. It holds

$$\begin{aligned}\|G(d)\|_{Y^*} &\le \|G\|_{\mathcal{B}(U\times Y, Y^*)}(\|v_1\|_U + \|z_1\|_Y)(\|v_2\|_U + \|z_2\|_Y) \\ &\le \|G\|_{\mathcal{B}(U\times Y, Y^*)}\left(\|v_1\|_U\|v_2\|_U + \|v_1\|_U\|z_2\|_Y + \|z_1\|_Y\|v_2\|_U + \|z_1\|_Y\|z_2\|_Y\right) \\ &\le \|G\|_{\mathcal{B}(U\times Y, Y^*)}\left(1 + 2M_z + M_z^2\right)\|v_1\|_U\|v_2\|_U.\end{aligned}$$

For the second estimate, we have due to the bilinearity of G

$$\begin{aligned}G(d^h) - G(d) &= G[(v_1, z_1^h), (v_2, z_2^h)] - G[(v_1, z_1), (v_2, z_2)] \\ &= G[(v_1, 0), (v_2, 0)] + G[(v_1, 0), (0, z_2^h)] + G[(0, z_1^h), (v_2, 0)] + G[(0, z_1^h), (0, z_2^h)] \\ &\quad - G[(v_1, 0), (v_2, 0)] - G[(v_1, 0), (0, z_2)] - G[(0, z_1), (v_2, 0)] - G[(0, z_1), (0, z_2)] \\ &= G[(v_1, 0), (0, z_2^h - z_2)] + G[(0, z_1^h - z_1), (v_2, 0)] \\ &\quad + G[(0, z_1^h - z_1), (0, z_2^h)] + G[(0, z_1), (0, z_2^h - z_2)].\end{aligned}$$

Hence, we can estimate

$$\begin{aligned}\|G(d^h) - G(d)\|_{Y^*} &\le \|G\|_{\mathcal{B}(U\times Y, Y^*)}\big(\|v_1\|_U\|z_2^h - z_2\|_Y + \|v_2\|_U\|z_1^h - z_1\|_Y \\ &\quad + \|z_2^h - z_2\|_Y\|z_1\|_Y + \|z_1^h - z_1\|_Y\|z_2^h\|_Y\big) \\ &\le \|G\|_{\mathcal{B}(U\times Y, Y^*)}c_z\left(2 + M_z + M_{z^h}\right)\|u - u_h\|_U\|v_1\|_U\|v_2\|_U,\end{aligned}$$

where we used the estimates (3.25)–(3.27) provided by Lemma 3.8. ☐

Lemma 3.11. *Let $v_i \in U$, $i = 1, 2$, be given. Let z_i, z_i^h, $i = 1, 2$ be defined as in Lemma 3.8, i.e. z_i and z_i^h solve* (3.23) *and* (3.24), *respectively. For $i = 1, 2$ let z_i, z_i^h be defined by (3.23) and (3.24) respectively. Let $u \in U_{ad}$, $\|u - u_h\|_U \le R$ with associated adjoint state p that solves* (3.13) *be given. Then it holds*

$$\left|\left\langle E''(y^h, u_h)[(v_1, z_1^h)(v_2, z_2^h)],\, p^h\right\rangle_{Y^*,Y} - \langle E''(y, u)[(v_1, z_1)(v_2, z_2)],\, p\rangle_{Y^*,Y}\right| \le C_{E''}\|u - u_h\|_U\|v_1\|_U\|v_2\|_U$$

with

$$C_{E''} := \left(c_{E''}\epsilon_y + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y, Y^*)}\right)(c_d M_p + c_p M_d) + M_d c_{E''}(c_y + 1)(M_p + Rc_p).$$

Proof. Let $d^h = [(v_1, z_1^h), (v_2, z_2^h)]$ and $d = [(v_1, z_1), (v_2, z_2)]$. We write

$$\begin{aligned}\big\langle E''(y^h, u_h)(d^h),\, p^h\big\rangle_{Y^*,Y} &- \langle E''(y,u)(d),\, p\rangle_{Y^*,Y} \\ &= \big\langle E''(y,u)(d),\, p^h - p\big\rangle_{Y^*,Y} + \big\langle E''(y^h, u_h)(d^h) - E''(y,u)(d),\, p^h\big\rangle_{Y^*,Y} \end{aligned} \tag{3.30}$$

and

$$E''(y^h, u_h)(d^h) - E''(y,u)(d) = E''(y^h, u_h)(d^h) - E''(y^h, u_h)(d) \;+\; E''(y^h, u_h)(d) - E''(y,u)(d). \tag{3.31}$$

The first term on the right-hand side of (3.30) is estimated using the estimate (3.28) and (3.15) as

$$\begin{aligned}\langle E''(y,u)d, p^h - p\rangle_{Y^*,Y} &\le \|E''(y,u)d\|_{Y^*}\|p^h - p\|_Y \\ &\le M_d\|E''(y,u)\|_{\mathcal{B}(U\times Y,Y^*)}\|p^h - p\|_Y\|v_1\|_U\|v_2\|_U \\ &\le M_d c_p\|E''(y,u)\|_{\mathcal{B}(U\times Y,Y^*)}\|u - u_h\|_U\|v_1\|_U\|v_2\|_U.\end{aligned}$$

Applying the Lipschitz property (3.5d) of E'' we obtain

$$\begin{aligned}\|E''(y,u)\|_{\mathcal{B}(U\times Y,Y^*)} &\le \|E''(y,u) - E''(y^h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} + \|E''(y^h, u_h) - E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \\ &\quad + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \\ &\le c_{E''}\left(\|y - y^h\|_Y + \|u - u_h\|_U + \|y^h - y_h\|_Y\right) + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \\ &\le c_{E''}((c_y + 1)R + \epsilon_y) + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)},\end{aligned} \tag{3.32}$$

where we used the error estimates (3.14) and (3.8). Hence, we get

$$\langle E''(y,u)d, p^h - p\rangle_{Y^*,Y} \le M_d c_p \left(c_{E''}((c_y + 1)R + \epsilon_y) + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\right)\|u - u_h\|_U\|v_1\|_U\|v_2\|_U. \tag{3.33}$$

To estimate the first addend on the right-hand side of (3.31), we employ Lemma 3.10

$$\|E''(y^h, u_h)(d^h) - E''(y^h, u_h)(d)\|_{Y^*} \le c_d\|E''(y^h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\|u - u_h\|_U\|v_1\|_U\|v_2\|_U. \tag{3.34}$$

Applying again the Lipschitz property (3.5d) of E'' and the residual estimate (3.8) we have the estimate

$$\begin{aligned}\|E''(y^h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} &\le \|E''(y^h, u_h) - E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \\ &\le c_{E''}\|y^h - y_h\|_Y + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \\ &\le c_{E''}\epsilon_y + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}.\end{aligned} \tag{3.35}$$

Now substituting the estimate (3.35) in (3.34) we obtain

$$\|E''(y^h, u_h)(d^h) - E''(y^h, u_h)(d)\|_{Y^*} \le c_d\left(c_{E''}\epsilon_y + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\right)\|u - u_h\|_U\|v_1\|_U\|v_2\|_U. \tag{3.36}$$

The second addend in (3.31) is estimated as in (3.32) and applying (3.28)

$$\|E''(y^h, u_h)(d) - E''(y,u)(d)\|_{Y^*} \le c_{E''}(c_y + 1)M_d\|u - u_h\|_U\|v_1\|_U\|v_2\|_U. \tag{3.37}$$

Putting (3.33), (3.36), and (3.37) together the claim follows after simple factorization, with the bound $\|p^h\|_Y \le M_p$, cf. (3.11). ◻

We have now obtained all the necessary ingredients to compute the Lipschitz constant of f''. The estimate is given in the following lemma.

LEMMA 3.12. *Let $u \in U_{ad}$ with $\|u - u_h\|_U \le R$ and $v_i \in U$, $i = 1, 2$, be given. Then the estimate*

$$|(f''(u) - f''(u_h))[v_1, v_2]| \le c_{f''}\|u - u_h\|_U\|v_1\|_U\|v_2\|_U$$

holds with $c_{f''} = c_{j''} + C_{g''} + C_{E''}$.

Proof. The second derivative f'' is given by

$$f''(u_h)[v_1, v_2] = j''(u_h)[v_1, v_2] + g''(y^h)[z_1^h, z_2^h] - \langle E''(y^h, u_h)[(v_1, z_1^h), (v_2, z_2^h)], p^h\rangle_{Y^*,Y}.$$

Then we obtain

$$\begin{aligned}|\left(f''(u)-f''(u_h)\right)[v_1,v_2]| \le \|j''(u)-j''(u_h)\|_{(U\times U)^*}\|v_1\|_U\|v_2\|_U + \left|g''(y)[z_1,z_2]-g''(y^h)[z_1^h,z_2^h]\right| \\ + \langle E''(y^h,u_h)[(v_1,z_1^h),(v_2,z_2^h)],p^h\rangle_{Y^*,Y} - \langle E''(y,u)[(v_1,z_1),(v_2,z_2)],p\rangle_{Y^*,Y}.\end{aligned}$$

With the help of Lipschitz estimate (3.5h) for j'', Lemma 3.9 and Lemma 3.11 one finds

$$|\left(f''(u)-f''(u_h)\right)[v_1,v_2]| \le (c_{j''}+C_{g''}+C_{E''})\|u-u_h\|_U\|v_1\|_U\|v_2\|_U,$$

which completes the proof. ∎

Using similar arguments and estimates, we derive now a uniform bound of f''.

LEMMA 3.13. *Let $u \in U_{ad}$ with $\|u-u_h\|_U \le R$ be given. Then there is a positive constant $M_{f''}$ such that*

$$\|f''(u)\|_{(U\times U)^*} \le M_{f''}$$

holds with

$$\begin{aligned}M_{f''} := c_{j''}R + \|j''(u_h)\|_{(U\times U)^*} + c_{g''}c_yR + c_{g''}\epsilon_y + \|g''(y_h)\|_{(Y\times Y)^*} \\ + M_d\left(c_{E''}\left((c_y+1)R+\epsilon_y\right) + \|E''(y_h,u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\right)(c_pR+M_p).\end{aligned}$$

Proof. Let us take $v_i \in U$ with $\|v_i\|_U = 1, i = 1,2$. Using (3.28) we have

$$\begin{aligned}|f''(u)[v_1,v_2]| &\le \|j''(u)\|_{(U\times U)^*} + \|g''(y)\|_{(Y\times Y)^*} + |\langle E''(y,u)[(v_1,z_1),(v_2,z_2)],p\rangle_{Y^*,Y}| \\ &\le \|j''(u)\|_{(U\times U)^*} + \|g''(y)\|_{(Y\times Y)^*} + M_d\|E''(y,u)\|_{\mathcal{B}(U\times Y,Y^*)}\|p\|_Y.\end{aligned} \tag{3.38}$$

We estimate each of the norms separately as follows. First, making use of the Lipschitz property (3.5h) it holds

$$\begin{aligned}\|j''(u)\|_{(U\times U)^*} &\le \|j''(u)-j''(u_h)\|_{(U\times U)^*} + \|j''(u_h)\|_{(U\times U)^*} \\ &\le c_{j''}R + \|j''(u_h)\|_{(U\times U)^*},\end{aligned} \tag{3.39}$$

where we have estimated the norm $\|u-u_h\|_U$ by its upper bound R. Secondly, applying property (3.5f) of g'', estimate (3.14) and the residual estimate (3.8) we obtain

$$\begin{aligned}\|g''(y)\|_{(Y\times Y)^*} &\le \|g''(y)-g''(y^h)\|_{(Y\times Y)^*} + \|g''(y^h)-g''(y_h)\|_{(Y\times Y)^*} + \|g''(y_h)\|_{(Y\times Y)^*} \\ &\le c_{g''}c_yR + c_{g''}\epsilon_y + \|g''(y_h)\|_{(Y\times Y)^*}.\end{aligned} \tag{3.40}$$

With the aid of estimates (3.15) and (3.11), the norm of the adjoint state p is estimated as

$$\|p\|_Y \le \|p-p^h\|_Y + \|p^h\|_Y \le c_pR + M_p.$$

Finally putting the already obtained estimate (3.32) of $\|E''(y,u)\|_{\mathcal{B}(U\times Y,Y^*)}$, (3.39) and (3.40) in (3.38) we obtain

$$\begin{aligned}|f''(u)[v_1,v_2]| \le c_{j''}R + \|j''(u_h)\|_{(U\times U)^*} + c_{g''}c_yR + c_{g''}\epsilon_y + \|g''(y_h)\|_{(Y\times Y)^*} \\ + M_d\left(c_{E''}\left((c_y+1)R+\epsilon_y\right) + \|E''(y_h,u_h)\|_{\mathcal{L}((U\times Y)^2,Y^*)}\right)(c_pR+\epsilon_p+\|p_h\|_Y).\end{aligned}$$

∎

3.5. Computation of coercivity constant α. Let us now describe how to determine the lower bound α in (2.5). The particular challenge is to find a computable estimate. Due to the finite-dimensional control space, the inequality

$$f''(u)[v,v] \ge \alpha\|v\|_U^2 \quad \forall v \in U \tag{3.41}$$

is equivalent to the inequality $\lambda_i \ge \alpha$ for all eigenvalues λ_i of all possible matrix realizations of f''. Let us choose

$$\{v_1 \dots v_n\} \text{ to be the canonical basis of } U = \mathbb{R}^n.$$

This choice also fits to the inequality constraints in U_{ad}: as they are posed component-wise, they are equivalent to inequality bounds on the coordinates of control vectors with respect to the chosen basis. With the basis $\{v_1 \dots v_n\}$ fixed, the inequality (3.41) is equivalent to the statement that all eigenvalues of the symmetric matrix F, $F_{ij} = f''(u)[v_i, v_j]$, $i, j = 1 \dots n$, are greater than α.

Let us recall the structure of f''. Let $u \in U_{ad}$ be given with the associated state y and adjoint p. Then we have

$$f''(u)[v_i, v_j] = j''(u)[v_i, v_j] + g''(y)[z_i, z_j] - \langle E''(y, u)[(v_i, z_i), (v_j, z_j)], p\rangle_{Y^*,Y} \tag{3.42}$$

where the functions z_i are the solutions of the linearized problem

$$E_y(y, u)z_i + E_u(y, u)v_i = 0.$$

As a consequence of representation (3.42), we have that the constant α in the inequality (2.7), which reads

$$f''(u_h)[v, v] \geq \alpha \|v\|_U^2 \quad \forall v \in U: \; v_i = 0 \; i \in A,$$

is equal to the smallest eigenvalue of the matrix L^h given by

$$L^h := \Big(j''(u_h)[v_i, v_j] + g''(y^h)[z_i^h, z_j^h] - \langle E''(y^h, u_h)[(v_i, z_i^h), (v_j, z_j^h)], p^h\rangle_{Y^*,Y}\Big)_{i,j\in I} \tag{3.43}$$

where the functions z_i^h solve

$$E_y(y^h, u_h)z_i^h + E_u(y^h, u_h)v_i = 0. \tag{3.44}$$

However, since y^h as well as p^h and z_i^h are solutions of the partial differential equations, the entries of L^h are not computable. We will overcome this difficulty by making use of the computable matrix

$$L_h := \Big(j''(u_h)[v_i, v_j] + g''(y_h)[z_{i,h}, z_{j,h}] - \langle E''(y_h, u_h)[(v_i, z_{i,h}), (v_j, z_{j,h})], p_h\rangle_{Y^*,Y}\Big)_{i,j\in I}, \tag{3.45}$$

where the functions $z_{i,h} \in Y_h$ are solutions of discrete linearized equations

$$\langle E_y(y_h, u_h)z_{i,h} + E_u(y_h, u_h)v_i, \psi_h\rangle_{Y^*,Y} = 0 \quad \forall \psi_h \in Y_h. \tag{3.46}$$

Let us emphasize that all the involved functions are discrete quantities and as such can be computed, which makes L_h computable as well. By construction, L^h and L_h are the matrix representation of $f''(u_h)$ and $f_h''(u_h)$, where f_h denotes the cost functional of the discrete problem.

We define α_h to be the minimum eigenvalue of matrix L_h. We will later on assume that $\alpha_h > 0$. This is a verifiable assumption since the entries and the eigenvalues of matrix L_h are computable. Moreover, if u_h is a solution of the discrete problem then $\alpha_h \geq 0$ holds by second-order necessary optimality conditions.

To be able to estimate the error between L^h and L_h, we make the following assumption on the discrete functions $z_{i,h}$ and the residuals of (3.46).

Assumption 5. *Let (y_h, u_h, p_h) be as in Assumption 2. Let $\{v_1 \dots v_n\}$ be the canonical basis of $\mathbb{R}^n$. Let $z_{i,h}$ be approximations of solutions of* (3.46). *We suppose that the upper bounds on the associated residuals are available as*

$$\|E_y(y_h, u_h)z_{i,h} + E_u(y_h, u_h)v_i\|_{Y^*} \leq r_{z,i} \quad \forall i = 1 \dots n.$$

Analogously to Lemma 3.1, we have the following result on the error between $z_{i,h}$ and z_i^h.

Lemma 3.14. *Let Assumption 5 be satisfied. Let z_i^h be given as solution of* (3.44). *Then it holds*

$$\|z_i^h - z_{i,h}\|_Y \leq \epsilon_{z,i} \tag{3.47}$$

with $\epsilon_{z,i} = \delta^{-1}((c_{E_u} + c_{E_y}\|z_{i,h}\|_Y)\epsilon_y + r_{z,i})$, $i = 1 \dots n$.

Proof. The difference $z_i^h - z_{i,h}$ satisfies

$$\begin{aligned} E_y(y^h, u_h)(z_i^h - z_{i,h}) &= -E_u(y^h, u_h)v_i - E_y(y^h, u_h)z_{i,h} \\ &= -(E_u(y^h, u_h) - E_u(y_h, u_h))v_i + (E_y(y^h, u_h) - E_y(y_h, u_h))z_{i,h} \\ &\quad + E_y(y_h, u_h)z_{i,h} + E_u(y_h, u_h))v_i. \end{aligned}$$

Using the Lipschitz estimates of E_u and E_y and the result of Lemma 3.1, we obtain

$$\|z_i^h - z_{i,h}\|_Y \le \delta^{-1}((c_{E_u} + c_{E_y}\|z_{i,h}\|_Y)\epsilon_y + r_{z,i}).$$

□

COROLLARY 3.15. *Let Assumption 5 be satisfied. Then it holds*

$$\|z_i^h\|_Y \le M_{z_i^h} \tag{3.48}$$

with $M_{z_i^h} := \epsilon_{z,i} + \|z_{i,h}\|_Y$, $i = 1 \dots n$.

The following lemmas, analogous to Lemma 3.9 and Lemma 3.10, are required in the subsequent computation.

LEMMA 3.16. *Let* z_i^h *be defined by* (3.24) *and* $z_{i,h}$ *by* (3.46). *Then the following inequality holds*

$$\left|g''(y^h)[z_i^h, z_j^h] - g''(y_h)[z_{i,h}, z_{j,h}]\right| \le \epsilon_{g''_{i,j}}$$

for $i, j = 1 \dots n$, *where*

$$\epsilon_{g''_{i,j}} := c_{g''} M_{z_i^h} M_{z_j^h} \epsilon_y + \|g''(y_h)\|_{(Y\times Y)^*} \left(M_{z_j^h}\epsilon_{z,i} + \|z_{i,h}\|_Y \epsilon_{z,j}\right).$$

Proof. We can write

$$\begin{aligned} g''(y^h)[z_i^h, z_j^h] &- g''(y_h)[z_{i,h}, z_{j,h}] \\ &= (g''(y^h) - g''(y_h))[z_i^h, z_j^h] + g''(y_h)([z_i^h, z_j^h] - [z_{i,h}, z_{j,h}]) \\ &= (g''(y^h) - g''(y_h))[z_i^h, z_j^h] + g''(y_h)\left([z_i^h - z_{i,h}, z_j^h] + [z_{i,h}, z_j^h - z_{j,h}]\right). \end{aligned} \tag{3.49}$$

We estimate the first addend of (3.49) using the Lipschitz estimate (3.5f) of g'', (3.48) and (3.8) to obtain

$$\begin{aligned} \left|(g''(y^h) - g''(y_h))[z_i^h, z_j^h]\right| &\le \|g''(y^h) - g''(y_h)\|_{(Y\times Y)^*}\|z_i^h\|_Y\|z_j^h\|_Y \\ &\le c_{g''}\|y^h - y_h\|_Y M_{z_i^h} M_{z_j^h} \\ &\le c_{g''} M_{z_i^h} M_{z_j^h}\epsilon_y. \end{aligned} \tag{3.50}$$

The second addend is likewise estimated using (3.47) and (3.48) as

$$\begin{aligned} \left|g''(y_h)\left([z_i^h - z_{i,h}, z_j^h] + [z_{i,h}, z_j^h - z_{j,h}]\right)\right| &\le \|g''(y_h)\|_{(Y\times Y)^*}\|z_i^h - z_{i,h}\|_Y\|z_j^h\|_Y + \|z_{i,h}\|_Y\|z_j^h - z_{j,h}\|_Y \\ &\le \|g''(y_h)\|_{(Y\times Y)^*}\left(M_{z_j^h}\epsilon_{z,i} + \|z_{i,h}\|_Y\epsilon_{z,j}\right). \end{aligned} \tag{3.51}$$

Now putting (3.50) and (3.51) in (3.49) yields the claim. □

LEMMA 3.17. *Let* G *be a bounded bilinear form on the space* $U \times Y$, *i.e.* $G : (U\times Y)\times(U\times Y) \mapsto Y^*$. *Let* $d_{i,j}^h := [(v_i, z_i^h), (v_j, z_j^h)]$ *and* $d_{i,j,h} := [(v_i, z_{i,h}), (v_j, z_{j,h})]$, $i, j = 1 \dots n$. *Then it holds*

$$\|G(d_{i,j,h})\|_{Y^*} \le M_{d_{i,j,h}}\|G\|_{\mathcal{B}(U\times Y, Y^*)}, \tag{3.52}$$

$$\|G(d_{i,j}^h)\|_{Y^*} \le M_{d_{i,j}}\|G\|_{\mathcal{B}(U\times Y, Y^*)}, \tag{3.53}$$

$$\|G(d_{i,j}^h) - G(d_{i,j,h})\|_{Y^*} \le \epsilon_{d_{i,j}}\|G\|_{\mathcal{B}(U\times Y, Y^*)} \tag{3.54}$$

for all $i, j = 1 \dots n$, *where the constants are given by*

$$\begin{aligned} M_{d_{i,j,h}} &:= (1 + \|z_{i,h}\|_Y)(1 + \|z_{j,h}\|_Y), \\ M_{d_{i,j}} &:= (1 + M_{z_i^h})(1 + M_{z_j^h}\|_Y), \\ \epsilon_{d_{i,j}} &:= \epsilon_{z,i}(1 + \|z_{j,h}\|_Y) + \epsilon_{z,j}(1 + M_{z_i^h}). \end{aligned}$$

Proof. For the first two estimates, we follow the steps of the proof of (3.28) in Lemma 3.10. Since $\|v_i\|_U = 1$, we have

$$\begin{aligned} \|G(d_{i,j,h})\|_{Y^*} &= \|G[(v_i, z_{i,h}), (v_j, z_{j,h})]\|_{Y^*} \\ &\le \|G\|_{\mathcal{B}(U\times Y, Y^*)}(\|v_i\|_U + \|z_{i,h}\|_Y)(\|v_j\|_U + \|z_{j,h}\|_Y) \\ &\le \|G\|_{\mathcal{B}(U\times Y, Y^*)}(1 + \|z_{i,h}\|_Y)(1 + \|z_{j,h}\|_Y). \end{aligned}$$

Similarly using (3.48) we obtain (3.53). For the last estimate, the proof is analogous to that of (3.29). Applying the estimates (3.47) and (3.48) we obtain

$$\begin{aligned}\|G(d^h_{i,j}) - G(d_{i,j,h})\|_{Y^*} &= \|G[(v_i, z^h_i), (v_j, z^h_j)] - G[(v_i, z_{i,h}), (v_j, z_{j,h})]\|_{Y^*}\\ &\le \|G\|_{\mathcal{B}(U\times Y,Y^*)}\big(\|v_i\|_U\|z^h_j - z_{j,h}\|_Y + \|v_j\|_U\|z^h_i - z_{i,h}\|_Y\\ &\qquad + \|z^h_j - z_{j,h}\|_Y\|z^h_i\|_Y + \|z^h_i - z_{i,h}\|_Y\|z_{j,h}\|_Y\big)\\ &\le \|G\|_{\mathcal{B}(U\times Y,Y^*)}\left(\epsilon_{z,i}(1 + \|z_{j,h}\|_Y) + \epsilon_{z,j}(1 + M_{z^h_i})\right).\end{aligned}$$

∎

Please note that the estimates $\epsilon_{g''_{i,j}}$ and $\epsilon_{d_{i,j}}$ are symmetric, e.g. it holds $\epsilon_{g''_{i,j}} = \epsilon_{g''_{j,i}}$, which follows from the structure of the bound $M_{z^h_i}$ given by (3.48). This is a nice coincidence as these error estimates are error bounds for symmetric perturbations of symmetric matrices.

Lemma 3.18. *Let $d^h_{i,j}$ and $d_{i,j,h}$ be as defined in Lemma 3.17. Then it holds*

$$\left|\langle E''(y_h, u_h)(d_{i,j,h}),\, p_h\rangle_{Y^*,Y} - \langle E''(y^h, u_h)(d^h_{i,j}),\, p^h\rangle_{Y^*,Y}\right| \le \epsilon_{E''_{i,j}}$$

with

$$\epsilon_{E''_{i,j}} := M_{d_{i,j}} c_{E''}\epsilon_y M_p + \|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\left(\epsilon_{d_{i,j}} M_p + M_{d_{i,j,h}}\epsilon_p\right).$$

Proof. It holds

$$\begin{aligned}&\langle E''(y_h, u_h)(d_{i,j,h}),\, p_h\rangle_{Y^*,Y} - \langle E''(y^h, u_h)(d^h_{i,j}),\, p^h\rangle_{Y^*,Y}\\ &\quad= \langle E''(y_h, u_h)(d_{i,j,h}) - E''(y^h, u_h)(d^h_{i,j}),\, p^h\rangle_{Y^*,Y} + \langle E''(y_h, u_h)(d_{i,j,h}),\, p_h - p^h\rangle_{Y^*,Y}\\ &\quad\le \|E''(y_h, u_h)d_{i,j,h} - E''(y^h, u_h)(d^h_{i,j})\|_{Y^*}\|p^h\|_Y + \|E''(y_h, u_h)(d_{i,j,h})\|_{Y^*}\|p_h - p^h\|_Y.\end{aligned} \tag{3.55}$$

We employ a similar splitting as in (3.31) to obtain

$$\begin{aligned}E''(y_h, u_h)(d_{i,j,h}) - E''(y^h, u_h)(d^h_{i,j}) &= \left(E''(y_h, u_h) - E''(y^h, u_h)\right)(d^h_{i,j})\\ &\quad + E''(y_h, u_h)(d_{i,j,h}) - E''(y_h, u_h)(d^h_{i,j}).\end{aligned}$$

Hence, by applying the estimates (3.53), (3.54), (3.5d) and (3.8) we obtain

$$\begin{aligned}&\|E''(y_h, u_h)(d_{i,j,h}) - E''(y^h, u_h)(d^h_{i,j})\|_{Y^*}\\ &\quad\le \|\left(E''(y_h, u_h) - E''(y^h, u_h)\right)(d^h_{i,j})\|_{Y^*} + \|E''(y_h, u_h)(d_{i,j,h}) - E''(y_h, u_h)(d^h_{i,j})\|_{Y^*}\\ &\quad\le M_{d_{i,j}}\|E''(y_h, u_h) - E''(y^h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} + \epsilon_{d_{i,j}}\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\\ &\quad\le M_{d_{i,j}} c_{E''}\|y_h - y^h\|_Y + \epsilon_{d_{i,j}}\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\\ &\quad\le M_{d_{i,j}} c_{E''}\epsilon_y + \epsilon_{d_{i,j}}\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}.\end{aligned} \tag{3.56}$$

Due to (3.52), it holds

$$\|E''(y_h, u_h)(d_{i,j,h})\|_{Y^*} \le M_{d_{i,j,h}}\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}. \tag{3.57}$$

Altogether, substituting (3.56) and (3.57) in (3.55) we arrive at

$$\begin{aligned}&\left|\langle E''(y_h, u_h)(d_{i,j,h}),\, p_h\rangle_{Y^*,Y} - \langle E''(y^h, u_h)(d^h_{i,j}),\, p^h\rangle_{Y^*,Y}\right|\\ &\qquad\le M_p\left(M_{d_{i,j}} c_{E''}\epsilon_y + \epsilon_{d_{i,j}}\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)}\right) + \epsilon_p M_{d_{i,j,h}}\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)},\end{aligned}$$

where we applied (3.11) and (3.9) to estimate the norms $\|p^h\|_Y$ and $\|p^h - p_h\|_Y$, respectively. ∎

Now we are in the position to prove the following bounds for the entries of the error matrix $L^h - L_h$.

Lemma 3.19. *Let the matrices L^h and L_h be given by* (3.43) *and* (3.45)*, respectively. Then it holds*

$$|L^h_{i,j} - L_{h_{i,j}}| \le \mathcal{E}_{i,j} := \epsilon_{g''_{i,j}} + \epsilon_{E''_{i,j}}, \quad i, j \in I.$$

Proof. By the definitions (3.43) and (3.45), the elements of the error matrix e_{ij} fulfill

$$\begin{aligned}
e_{ij} &= j''(u_h)(v_i,v_j) + g''(y^h)(z_i^h,z_j^h) - \langle E''(y^h,u_h)[(v_i,z_i^h),(v_j,z_j^h)],p^h\rangle_{Y^*,Y} \\
&\quad - j''(u_h)(v_i,v_j) - g''(y_h)(z_{i,h},z_{j,h}) + \langle E''(y_h,u_h)[(v_i,z_{i,h}),(v_j,z_{j,h})],p_h\rangle_{Y^*,Y} \\
&\le \left|g''(y^h)(z_i^h,z_j^h) - g''(y_h)(z_{i,h},z_{j,h})\right| \\
&\quad + \left|\langle E''(y_h,u_h)[(v_i,z_{i,h}),(v_j,z_{j,h})],\,p_h\rangle_{Y^*,Y} - \langle E''(y^h,u_h)[(v_i,z_i^h),(v_j,z_j^h)],\,p^h\rangle_{Y^*,Y}\right|.
\end{aligned}$$

Applying the results of Lemma 3.16 and Lemma 3.18 completes the proof. ☐

We finalize the computation of the coercivity constant by recalling the following result from matrix perturbation theory.

THEOREM 3.20. *Let the matrix A be perturbed by a symmetric matrix $\mathcal{E}$, and denote by $\lambda_k(A)$ the k-th eigenvalue of A. If A and $A+\mathcal{E}$ are $n\times n$ symmetric matrices, then*

$$|\lambda_k(A+\mathcal{E}) - \lambda_k(A)| \le \|\mathcal{E}\|_2$$

for $k = 1\ldots n$.

Proof. The simple proof can be found in [7]. ☐

THEOREM 3.21. *Let α, α_h be the minimum eigenvalues of matrices L^h and L_h respectively. Then it holds*

$$\alpha \ge \alpha_h - \|\mathcal{E}\|_2,$$

where the error matrix $\mathcal{E} = (\mathcal{E}_{i,j})$ is given in Lemma 3.19.

Proof. The claim follows from the previous Theorem 3.20, as L^h, L_h, as well as $\mathcal{E}$ are symmetric matrices. Moreover, $\mathcal{E}_{i,j}$ is an upper bound of $|L^h_{i,j} - L_{h_{i,j}}|$, which implies $\|L^h - L_h\|_2 \le \|\mathcal{E}\|_2$. ☐

If $\bar u$ is a solution of the optimization problem satisfying the SSC, then the bound α_h will eventually become positive. If mesh refinement is done is such a way that the residuals r_y, r_p, r_u vanish and $u_h \to u$, then the error $\|\mathcal{E}\|_2$ will tend to zero as well.

3.6. Main result. Let us summarize the results obtained in this section so far. The goal of all these work was to derive bounds to apply the results of Section 2. Let us recall that these results were given in terms of quantities ϵ, α, σ, and $c_{f'}, c_{f''}, M_{f''}$, cf. Assumption 3. All these constants were derived in the previous subsections. It remains to collect them and to present the main result, which is an error estimate for the error in the solution. Moreover, it allows to verify the fulfillment of the second-order sufficient condition a-priori.

THEOREM 3.22. *Let ϵ, α, σ, and $c_{f'}, c_{f''}, M_{f''}$ be computed according to the results in this section. Let us suppose that these constants satisfy the assumptions of Corollary 2.4. Then there exists a local solution $\bar u$ of* (P) *that satisfies the error bound*

$$\|\bar u - u_h\|_U \le \frac{2\epsilon}{\alpha}.$$

The solutions $\bar u$ fulfills the second-order sufficient condition given by (2.4)–(2.5). *Moreover, we have the a-posteriori error representation in terms of the residuals in Assumption 2*

$$\|\bar u - u_h\|_U \le \frac{2}{\alpha}\left(r_u + \omega_y\, r_y + \omega_p\, r_p\right)$$

with weights given by

$$\begin{aligned}
\omega_y &= c_{E_u}\delta^{-1}\left(\delta^{-2}\left(c_{g'} + c_{E_y}\|p_h\|_Y\right) r_y + \|p_h\|_Y\right) + \|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\delta^{-1}\left(c_{g'} + c_{E_y}\|p_h\|_Y\right), \\
\omega_p &= \delta^{-1}\left(c_{E_u} r_y + \|E_u(y_h,u_h)\|_{\mathcal{L}(U,Y^*)}\right).
\end{aligned}$$

Proof. The claim follows from Corollaries 2.3 and 2.4 as well as from the representation of ϵ in terms of r_u, r_y, r_p derived in Lemma 3.7. ☐

4. Application to parameter optimization problems. In this section, we apply the developed abstract framework to the parameter optimization problems (1.2) and (1.3), which were introduced in the first section. First, we fix the following settings, which are common to both problems.

Throughout this section, Ω is a two or three dimensional bounded Lipschitz domain with boundary $\partial\Omega$. The state space is $Y = H^1_0(\Omega)$, its dual $Y^* = H^{-1}(\Omega)$ and the control space is $U = \mathbb{R}^n$. The functions y_d, u_a, u_b, the regularization parameter $\kappa > 0$ and the source term g are all given in appropriate spaces. Furthermore, the functionals g, j in (1.1) are given by

$$g(y) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)}, \qquad j(u) := \frac{\kappa}{2}\|u\|^2_{\mathbb{R}^n}.$$

At first, let us argue that the cost functional, in particular the functions g and j met the requirements of Assumption 4. Indeed, we find that (3.5e)–(3.5h) are satisfied with

$$c_{g'} = I_2^2,\ c_{j'} = \kappa,\ c_{g''} = 0,\ c_{j''} = 0.$$

Here, we denoted by I_2 the norm of the embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$. Moreover, it holds $\|g''(y)\|_{(Y\times Y)^*} \leq I_2^2$ and $\|j''(u)\|_{(U\times U)^*} = \kappa$ for all $u \in U$, $y \in Y$, uniformly.

We will argue in the sequel that the resulting optimization problems subject to the nonlinear elliptic equations (1.2) or (1.3) fulfill Assumption 1. Let us first describe the employed discretization procedure. Afterwards, we will report on how the remaining estimates in Assumption 4 are computed for each problem.

4.1. Discretization and computation of residuals. We used standard finite element techniques to discretize the problem. The domain is divided into triangles. The finite element space Y_h is the classical spaces of piecewise quadratic and continuous elements (P2).

The critical part is the computation of the residuals. Here, 'computation' refers to the fact that we need constant-free error estimates, i.e. we have to determine r_y satisfying $\|E(y_h, u_h\|_{Y^*} \leq r_y$, no extra constants involved. That means, we cannot use standard residual-type a-posteriori error estimates. Nevertheless, there are quite a few options available, as for instance the so-called hypercircle method, see e.g. [5], estimates based on local $H(\mathrm{div})$-error representations [17], equilibrated residuals [1], or functional error estimates [13].

We used a related technique, as described in [12]. Let $\sigma \in H(\mathrm{div})$ be given, i.e. $\sigma \in L^2(\Omega)^d$ with $\mathrm{div}(\sigma) \in L^2(\Omega)$. Then we can estimate the residual in the equation $-\Delta y + d(y) = f$ at a discrete function y_h as

$$\begin{aligned}\|-\Delta y_h + d(y_h) - f\|_{H^{-1}} &\leq \|-\Delta y_h - \mathrm{div}(\sigma)\|_{H^{-1}} + \|\,\mathrm{div}(\sigma) + d(y_h) - f\|_{H^{-1}} \\ &\leq \|\nabla y_h + \sigma\|_{L^2} + I_2\|\,\mathrm{div}(\sigma) + d(y_h) - f\|_{L^2}.\end{aligned}$$

In our computations, we used the Raviart-Thomas elements RT_1 to discretize the space $H(\mathrm{div})$. In a post-processing step, we computed σ_h as minimizer of

$$\|\nabla y_h + \sigma\|^2_{L^2} + I_2^2\|\,\mathrm{div}(\sigma) + d(y_h) - f\|^2_{L^2}.$$

Then the residual was computed as

$$\|-\Delta y_h + d(y_h) - f\|_{H^{-1}} \leq \|\nabla y_h + \sigma_h\|_{L^2} + I_2\|\,\mathrm{div}(\sigma_h) + d(y_h) - f\|_{L^2}.$$

We applied this technique to compute bounds of the residuals for the state and adjoint equations as well as for the linearized equations appearing in the eigenvalue problem associated to f''.

4.2. Identification of coefficient in the main part of elliptic equation. Let us verify the assumptions for the optimization problems involving parameters in the differential operator. To this end, let disjoint, measurable sets $\Omega_i \subset \Omega$ be given, $i = 1 \ldots n$, with $\Omega = \bigcup_{i=1}^n \Omega_i$. In order to make the resulting differential operator coercive the lower bound on the coefficients is a positive number, $u_{a,i} = \tau > 0$. Upper bounds are taken into account as well.

We set $Y = H^1_0(\Omega)$ with norm $\|y\|^2_Y := \|\nabla y\|^2_{L^2(\Omega)} + \|y\|^2_{L^2(\Omega)}$. The mapping E is now defined as

$$\langle E(y, u),\, v\rangle = \sum_{i=1}^n u_i \int_{\Omega_i} (\nabla y \cdot \nabla v - gv)\ \mathrm{d}x.$$

Here, $g \in L^2(\Omega)$ is a given data function. With this definition, we have that the differentiability requirements of Assumption 1 are met. The following lemma states that also the strong monotonicity as well as Lipschitz continuity conditions of Assumptions 1 and 4 are fulfilled.

LEMMA 4.1. *Let $u \in U_{ad}$ and $y \in Y$ be given. Then it holds*

$$E_y(y, u) \leq \delta^{-1}$$

with $\delta = \tau$, which is the lower bound on the parameters. Moreover, we have that the inequalities (3.5a)–(3.5d) *of Assumption 1 hold with*

$$c_E = \max(\|y^h\|_Y,\, \|u_h\|_U + R), \quad c_{E_y} = c_{E_u} = 1, \quad c_{E''} = 0.$$

In addition, the inequalities $\|E_u(y_h, u_h)\|_{\mathcal{L}(U,Y^*)} \le \|y_h\|_Y$ and $\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \le 1$ are satisfied.

Proof. Let $u, u_h \in U_{ad}$, $y^h \in Y$ be given with $\|u - u_h\| \le R$. Let $y = S(u)$ be the solution of $E(y, u) = 0$. At first, let us determine the Lipschitz constant of E:

$$\begin{aligned}\|E(y,u) - E(y^h,u_h)\|_{Y^*} &\le \sum_{i=1}^{n} \|(u_i - u_{h,i})\nabla y^h\|_{L^2(\Omega_i)} + \|u_i(\nabla y - \nabla y^h)\|_{L^2(\Omega_i)} \\ &\le \|u - u_h\|_U \|y^h\|_Y + (\|u_h\| + R)\|y - y^h\|_Y,\end{aligned}$$

so that (3.5a) holds with $c_E := \max(\|y^h\|_Y,\, \|u_h\|_U + R)$. Let us take $z \in Y$. Since $E_y(y, u)z = -\mathrm{div}(u\nabla z)$, it follows that

$$\|(E_y(y,u) - E_y(y^h,u_h))z\|_{Y^*} \le \sum_{i=1}^{n} \|(u_i - u_{h,i})\nabla z\|_{L^2(\Omega_i)} \le \|u - u_h\|_U \|z\|_Y,$$

which implies $c_{E_y} = 1$. A similar computation gives with $v \in U$

$$\|(E_u(y,u) - E_u(y^h,u_h))v\|_{Y^*} \le \sum_{i=1}^{n} \|v_i(\nabla y - \nabla y^h)\|_{L^2(\Omega_i)} \le \|y - y^h\|_Y \|v\|_U,$$

implying $c_{E_u} = 1$. With a similar estimate we immediately obtain $\|E_u(y_h, u_h)\|_{\mathcal{L}(U,Y^*)} \le \|y_h\|_Y$. Since E is bilinear with respect to (u, y), the second derivative $E''(y, u)$ is independent of (y, u), hence $c_{E''} = 0$. More precisely $\langle E_{yu}(y_h, u_h)[z, w],\ v\rangle_{Y^*,Y} = \sum_{i=1}^{n} w_i \int_{\Omega_i} \nabla z \nabla v$. This implies

$$\|E_{yu}(y_h,u_h)[z,w]\|_{Y^*} \le \sum_{i=1}^{n} \|w_i \nabla z\|_{L^2(\Omega_i)} \le \|w\|_U \|z\|_Y.$$

Noting that the second derivatives E_{yy}, E_{uu} vanish, we obtain $\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y,Y^*)} \le 1$. ◻

4.3. Parameter identification problem. Let us consider the elliptic problem (1.2). We will make the special choice

$$d(u; y) = \sum_{i=1}^{n} u_i d_i(y),$$

where $d_i : \mathbb{R} \to \mathbb{R}$ are assumed to be twice continuously differentiable, $i = 1 \ldots n$, with the second derivatives being Lipschitz continuous on intervals $[-M, M]$ for all $M \ge 0$. Furthermore, we assume d_i to be monotonically increasing. Here we have in mind to work with $d_i(y) = y|y|^{i-2}$.

As a result, we define the nonlinear operator E as

$$E(y,u) := -\Delta y + \sum_{i=1}^{n} u_i d_i(y) - g, \tag{4.1}$$

where $g \in L^2(\Omega)$ is a given function. In order to make the resulting operator monotonic with respect to y we impose positivity requirements on u, i.e. we set

$$U_{ad} = \{u \in \mathbb{R}^n :\ u_i \ge 0 \quad \forall i = 1 \ldots n\}.$$

Due to the choice of functions d_i, the operator E is Fréchet-differentiable from $H^1(\Omega)\cap L^\infty(\Omega) \to H^{-1}(\Omega)$. That is we have to work with the framework $Y = H_0^1(\Omega)$, $Y_\infty = L^\infty(\Omega) \cap Y$. We will use the following norm in Y: $\|y\|_Y^2 := \|\nabla y\|_{L^2(\Omega)}^2 + \|y\|_{L^2(\Omega)}^2$.

4.3.1. Lipschitz estimates. At first, let us consider the Nemyzki (superposition) operators induced by the functions d_i. For simplicity, we will denote them by d_i, too.

LEMMA 4.2. *The Nemyzki operators d_i are twice Fréchet-differentiable from $L^\infty(\Omega)$ to $L^\infty(\Omega)$. Moreover, we have*

$$\|d_i(y) - d_i(y^h)\|_{L^p(\Omega)} \le \Phi_i(\max(\|y\|_{L^\infty(\Omega)},\, \|y^h\|_{L^\infty(\Omega)}))\|y - y^h\|_{L^p(\Omega)},$$
$$\|d_i'(y) - d_i'(y^h)\|_{L^p(\Omega)} \le \Phi_i'(\max(\|y\|_{L^\infty(\Omega)},\, \|y^h\|_{L^\infty(\Omega)}))\|y - y^h\|_{L^p(\Omega)},$$
$$\|d_i''(y) - d_i''(y^h)\|_{L^p(\Omega)} \le \Phi_i''(\max(\|y\|_{L^\infty(\Omega)},\, \|y^h\|_{L^\infty(\Omega)}))\|y - y^h\|_{L^p(\Omega)}$$

for all $y, y^h \in L^\infty(\Omega)$, $p \in [1, +\infty]$. Here we used the functions

$$\Phi_i(M) := \max_{x\in[-M,M]} |d_i'(x)|, \Phi_i'(M) := \max_{x\in[-M,M]} |d_i''(x)|,$$

and $\Phi_i''(M)$ denotes the Lipschitz modulus of d_i'' on the interval $[-M, M]$, i.e. $|d_i''(x_1) - d_i''(x_2)| \le \Phi_i''(M)|x_1 - x_2|$ for $x_1, x_2 \in [-M, M]$.

Proof. Due to the assumptions on the functions d_i the Nemyzki operators d_i are twice continuously Fréchet differentiable. By the mean value theorem, we have $|d_i(x_1)-d_i(x_2)| \le \Phi_i(\max(|x_1|, |x_2|))|x_1-x_2|$ for all $x_1, x_2 \in \mathbb{R}$. Hence,

$$\|d_i(y) - d_i(y^h)\|_{L^p(\Omega)} \le \Phi_i(\max(\|y\|_{L^\infty(\Omega)},\, \|y^h\|_{L^\infty(\Omega)}))\|y - y^h\|_{L^p(\Omega)}.$$

With analogous arguments we obtain the estimates for the derivatives of d_i. ∎

In order to prove Lipschitz estimates of E, we need Lipschitz estimates and global bounds for solutions of the equation $E(y, u) = 0$. Let us denote norm of the embedding $H_0^1(\Omega) \hookrightarrow L^p(\Omega)$ by I_p if such a continuous embedding exists.

LEMMA 4.3. *Let $u \in U_{ad}$, $y \in Y_\infty$ be given. Then it holds*

$$\|E_y^{-1}(y, u)\|_{\mathcal{L}(Y^*,Y)} \le \delta$$

with $\delta = 1 - I_2^2$.

Proof. This follows from the fact that $d_y(u; y) \in L^\infty(\Omega)$ as well as u is non-negative and from the simple inequality $\|\nabla y\|^2_{L^2(\Omega)} = \|y\|^2_Y - \|y\|^2_{L^2(\Omega)} \ge (1 - I_2^2)\|y\|^2_Y$. ∎

LEMMA 4.4. *Let $E : Y \times \mathbb{R}^n \to Y^*$ be given as in (4.1). Let $u \in U_{ad}$ be given. Then it holds for $y = S(u)$ being the solution of $E(y, u) = 0$*

$$\|y\|_{L^\infty(\Omega)} \le M_{L^\infty}\|g\|_{L^2(\Omega)} + M_{u,L^\infty}\|u\|_U$$

with

$$M_{L^\infty} = 4\frac{I_6^2}{1 - I_2^2}|\Omega|^{1/6}, \qquad M_{u,L^\infty} = M_{L^\infty}|\Omega|^{1/2}(\sum_{i=1}^n |d_i(0)|^2)^{1/2}.$$

Proof. Due to Stampacchia [16], we have the estimate

$$\|y\|_{L^\infty(\Omega)} \le M_{L^\infty}\left(\|g\|_{L^2(\Omega)} + \sum_{i=1}^n |u_i| \cdot \|d_i(0)\|_{L^2(\Omega)}\right)$$

with $M_{L^\infty} = 4\frac{I_6^2}{1-I_2^2}|\Omega|^{1/6}$ computed in [14]. ∎

LEMMA 4.5. *Let $u, u_h \in U_{ad}$, $\|u - u_h\|_U \le R$, $y = S(u) \in Y_\infty$ and $y^h \in Y_\infty$ be given. Then it holds*

$$\|E(y, u) - E(y^h, u_h)\|_{Y^*} \le c_E(\|y^h - y\|_Y + \|u - u_h\|_U)$$

with

$$c_E := \max\left(1 + I_2^2\|\Phi(u_h, y^h, R)\|_U(\|u_h\|_U + R),\, I_2\|d(y^h)\|_U\right),$$

where we used the abbreviations

$$\|d(y^h)\|_U := \left(\sum_{i=1}^n \|d_i(y^h)\|^2_{L^2(\Omega)}\right)^{1/2}$$

$$\|\Phi(u_h, y^h, R)\|_U := \left(\sum_{i=1}^n \Phi_i\left(\max(\|y^h\|_{L^\infty(\Omega)},\, M_{L^\infty} + M_{u,L^\infty}(R + \|u_h\|_U))\right)\right)^{1/2}.$$

This implies inequality (3.5a).

Proof. The claim follows from the splitting

$$u_i d_i(y) - u_{h,i} d_i(y^h) = (u_i - u_{h,i}) d_i(y^h) + u_i(d_i(y) - d_i(y^h))$$

and applying Lemma 4.2. E.g. we have using the embedding $Y \hookrightarrow L^2(\Omega)$

$$\|u_i(d_i(y) - d_i(y^h))\|_{Y^*} \le I_2^2 |u_i| \cdot \Phi_i(\max(\|y\|_{L^\infty(\Omega)}, \|y^h\|_{L^\infty(\Omega)}))\|y - y^h\|_Y.$$

Applying the L^∞-bound given by Lemma 4.4 finishes the proof. ☐

LEMMA 4.6. *Let $u, u_h \in U_{ad}$, $\|u - u_h\|_U \le R$, $y = S(u) \in Y_\infty$ and $y^h \in Y_\infty$ be given. Then the inequality* (3.5b), *i.e.*

$$\|E_y(y,u) - E_y(y^h, u_h)\|_{\mathcal{L}(Y,Y^*)} \le c_{E_y}(\|y^h - y\|_Y + \|u - u_h\|_U),$$

is fulfilled with

$$c_{E_y} := \max\left(I_3^3 \|\Phi'(u_h, y^h, R)\|_U(\|u_h\|_U + R),\, I_3^2 \|d'(y^h)\|_U\right),$$

where we used the abbreviations

$$\|d'(y^h)\|_U := \left(\sum_{i=1}^n \|d_i'(y^h)\|_{L^3(\Omega)}^2\right)^{1/2}$$

$$\|\Phi'(u_h, y^h, R)\|_U := \left(\sum_{i=1}^n \Phi_i'\left(\max(\|y^h\|_{L^\infty(\Omega)}, M_{L^\infty} + M_{u,L^\infty}(R + \|u_h\|_U))\right)\right)^{1/2}.$$

Proof. The proof is analogous to the proof of the previous Lemma 4.5. ☐

LEMMA 4.7. *Let $u, u_h \in U_{ad}$, $\|u - u_h\|_U \le R$, $y = S(u) \in Y_\infty$ and $y^h \in Y_\infty$ be given. Then the inequalities*

$$\|E_u(y,u) - E_u(y^h, u_h)\|_{\mathcal{L}(U,Y^*)} \le c_{E_u}\|y^h - y\|_Y,$$
$$\|E_{yu}(y,u) - E_{yu}(y^h, u_h)\|_{\mathcal{L}(Y,\mathcal{L}(U,Y^*))} \le c_{E_{yu}}\|y^h - y\|_Y$$

are fulfilled with

$$c_{E_u} := I_2^2 \|\Phi(u_h, y^h, R)\|_U(\|u_h\|_U + R),$$
$$c_{E_{yu}} := I_3^3 \|\Phi'(u_h, y^h, R)\|_U(\|u_h\|_U + R),$$

where we used the notation of Lemmata 4.5 and 4.6. This gives (3.5b).

Proof. Let $w \in U$ be given. By construction, we have $E_u(y,u)w = \sum_{i=1}^n w_i d_i(y)$, which obviously implies $(E_u(y,u) - E_u(y^h, u_h))w = \sum_{i=1}^n w_i(d_i(y) - d_i(y^h))$. Using Lemma 4.2, we obtain

$$\|(E_u(y,u) - E_u(y^h, u_h))w\|_{Y^*} \le \|w\|_U \|\Phi_i(\max(\|y\|_{L^\infty(\Omega)}, \|y^h\|_{L^\infty(\Omega)}))\|y - y^h\|_Y.$$

The claim follows with the same argumentation as in the proof of Lemma 4.5. By analogous considerations we obtain the Lipschitz estimate for E_{yu}. ☐

LEMMA 4.8. *Let $u, u_h \in U_{ad}$, $\|u - u_h\|_U \le R$, $y = S(u) \in Y_\infty$ and $y^h \in Y_\infty$ be given. Then the inequality*

$$\|E_{yy}(y,u) - E_{yy}(y^h, u_h)\|_{\mathcal{B}(Y,Y^*)} \le c_{E_{yy}}(\|y^h - y\|_Y + \|u - u_h\|_U),$$

is fulfilled with

$$c_{E_{yy}} := \max\left(I_4^4 \|\Phi''(u_h, y^h, R)\|_U(\|u_h\|_U + R),\, I_4^3 \|d''(y^h)\|_U\right),$$

where we used the abbreviations

$$\|d''(y^h)\|_U := \left(\sum_{i=1}^n \|d_i''(y^h)\|_{L^4(\Omega)}^2\right)^{1/2}$$

$$\|\Phi''(u_h, y^h, R)\|_U := \left(\sum_{i=1}^n \Phi_i''\left(\max(\|y^h\|_{L^\infty(\Omega)}, M_{L^\infty} + M_{u,L^\infty}(R + \|u_h\|_U))\right)\right)^{1/2}.$$

Proof. The proof is analogous to the proof of Lemma 4.5. □

COROLLARY 4.9. *Let $u, u_h \in U_{ad}$, $\|u - u_h\|_U \le R$, $y = S(u) \in Y_\infty$ and $y^h \in Y_\infty$ be given. Then the inequality* (3.5d), *i.e.*

$$\|E''(y, u) - E''(y^h, u_h)\|_{\mathcal{B}(U\times Y, Y^*)} \le c_{E''}(\|y^h - y\|_Y + \|u - u_h\|_U)$$

is satisfied with

$$c_{E''} = c_{E_{yy}} + c_{E_{yu}}$$

where $c_{E_{yu}}$ and $c_{E_{yy}}$ are given by Lemmata 4.7 and 4.8, respectively.

Proof. The constant $c_{E''}$ can be determined as the spectral norm of the matrix $\begin{pmatrix} c_{E_{yy}} & c_{E_{yu}} \\ c_{E_{yu}} & 0 \end{pmatrix}$. The largest eigenvalue of this matrix is given by $\frac{1}{2}\left(c_{E_{yy}} + \sqrt{c_{E_{yy}}^2 + 4c_{E_{yu}}^2}\right) \le c_{E_{yy}} + c_{E_{yu}}$, which gives the bound $c_{E''}$. □

Let us close these considerations with stating the Lipschitz constants for the choice $d_1 = 1$, $d_i = y|y|^{i-2}$ for $i \ge 2$. Special care has to be taken as $d_3 = y|y|$ is not twice continuously differentiable. If we restrict all the considerations to positive values of y, then the previous results still hold.

COROLLARY 4.10. *Let the functions d_i be given by $d_1 = 1$, $d_i = y|y|^{i-2}$ for $i = 2 \dots n$. Then we have*

$$\begin{aligned}\Phi_i(M) &= (i-1)M^{i-2}\\ \Phi_i'(M) &= \max(0,\ (i-1)(i-2)M^{i-3})\\ \Phi_i''(M) &= \max(0,\ (i-1)(i-2)(i-3)M^{i-4}) \quad i \ne 3.\end{aligned}$$

If $y, y^h \in Y_\infty$ are non-negative then the claims of Lemma 4.8 and Corollary 4.9 are true with $\Phi_i''(M) = \max(0,\ (i-1)(i-2)(i-3)M^{i-4})$ for $i = 1 \dots n$.

LEMMA 4.11. *It holds*

$$\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y, Y^*)} \le \sum_{i=1}^{n} \left(I_2^2 \|d_i'(y_h)\|_{L^\infty} + I_2 I_4^2 u_{h,i} \|d_i''(y_h)\|_{L^\infty}\right).$$

Proof. Let $(w, z) \in U \times Y$. The following estimates hold

$$\begin{aligned}\|E_{yu}(y_h, u_h)[w, z]\|_{Y^*} &\le \sum_{i=1}^{n} \|w_i d_i'(y_h) z\|_{Y^*} \le I_2^2 \sum_{i=1}^{n} \|d_i'(y_h)\|_{L^\infty} \|w\|_U \|z\|_Y,\\ \|E_{yy}(y_h, u_h)[z, z]\|_{Y^*} &\le \sum_{i=1}^{n} \|u_{h,i}\, d_i''(y_h)[z, z]\|_{Y^*} \le I_2 I_4^2 \sum_{i=1}^{n} u_{h,i} \|d_i''(y_h)\|_{L^\infty} \|z\|_Y^2.\end{aligned}$$

Since the second derivative of E with respect to u vanishes, the claim then follows from the inequality $\|E''(y_h, u_h)\|_{\mathcal{B}(U\times Y, Y^*)} \le \|E_{yy}(y_h, u_h)\|_{\mathcal{L}(Y\times Y, Y^*)} + \|E_{yu}(y_h, u_h)\|_{\mathcal{L}(U\times Y, Y^*)}$. □

4.4. Numerical results. Let us report about the outcome of our numerical experiments. The first example is concerned with the optimization of coefficients in the main part of the operator, see e.g. section 4.2.

Let us comment briefly on the computation of the safety radius R, which appears severally in the previous sections. An adaptive procedure was employed. Starting with an initial guess $R = 0$, the control error bound r_+ was computed according to Corollary 2.3. If $r_+ > R$ the safety radius R was updated as $R := \theta r_+$ with $\theta = 1.01$. The computation of r_+ was then repeated until the condition $r_+ \le R$ is fulfilled.

Example 1. The domain was chosen as $\Omega = (0, 1)^2$. The domain was split into four subdomains

$$\Omega_1 = (0, 0.5)^2,\ \Omega_2 = (0, 0.5) \times (0.5, 1),\ \Omega_3 = (0.5, 1) \times (0, 0.5),\ \Omega_4 = (0.5, 1)^2.$$

The problem data was given as

$$u_a = 0.1,\ u_b = +\infty,\ \kappa = 10^{-1},\ y_d(x_1, x_2) = (1 - x_1)^2(1 - x_2)x_1x_2^2.$$

We solved the discretized problem on a sequence of uniformly refined grids. The solution vector on the finest grid was computed to

$$\bar{u}_h = (0.8047,\ 0.8062,\ 0.8020,\ 0.8047).$$

As can be seen, the inequality constraints are not active at $\bar{u}_h$, hence $A = \emptyset$.

The results for the verification process are shown in Tables 4.1 and 4.2. As can be seen from the second column of Table 4.1, the discrete sufficient optimality condition is satisfied on all grids, as α_h is uniformly positive. However, $f''(u_h)$ is only positive definite for the grids with $h \leq 0.0177$ as the error $\|\mathcal{E}\|_2$ is larger than α_h on the coarser grids. But since $\|\mathcal{E}\|_2$ decays like h^2 the condition $\alpha = \alpha_h - \|\mathcal{E}\|_2$ eventually is satisfied. In this example, we see that as soon as the bound α is positive, the conditions of Theorem 3.22 are satisfied, and hence an error estimate $\|\bar{u} - u_h\|_U \leq r_+$ is available. This includes the statement that we are able to verify the existence of a local solution of (P) in the neighborhood of u_h. Moreover, the error bound r_+ decays with h^2, which is expected, since $r_+ \leq 2\alpha/\epsilon$ holds and ϵ tends to zero with the rate h^2, as can be seen Table 4.1.

h	α_h	$\|\mathcal{E}\|_2$	α	ϵ	r_+
0.0707	$1.4644 \cdot 10^{-1}$	$8.0093 \cdot 10^{-1}$	$-6.5449 \cdot 10^{-1}$	$3.1342 \cdot 10^{-2}$	–
0.0354	$1.4644 \cdot 10^{-1}$	$2.1288 \cdot 10^{-1}$	$-6.6443 \cdot 10^{-2}$	$8.3960 \cdot 10^{-3}$	–
0.0177	$1.4644 \cdot 10^{-1}$	$5.6723 \cdot 10^{-2}$	$8.9718 \cdot 10^{-2}$	$2.2376 \cdot 10^{-3}$	$4.9880 \cdot 10^{-2}$
0.0088	$1.4644 \cdot 10^{-1}$	$1.5054 \cdot 10^{-2}$	$1.3139 \cdot 10^{-1}$	$5.9291 \cdot 10^{-4}$	$9.0255 \cdot 10^{-3}$
0.0044	$1.4644 \cdot 10^{-1}$	$3.9754 \cdot 10^{-3}$	$1.4247 \cdot 10^{-1}$	$1.5627 \cdot 10^{-4}$	$2.1938 \cdot 10^{-3}$

TABLE 4.1
Example 1: verification results, α, ϵ, r_+

For convenience, we also report about the other quantities involved in the verification process, namely the Lipschitz constants of the reduced functional f. As can be seen in Table 4.2, the Lipschitz constants $c_{f'}$ and $c_{f''}$ as well as the bound $M_{f''}$ are bounded uniformly for all discretizations. They are monotonically decreasing due to their computation, which is the expected behavior in the light of the derivation in Section 3. Of course, all these constants are expected to be bounded away from zero.

h	$\epsilon_{f'}$	$c_{f'}$	$c_{f''}$	$M_{f''}$
0.0707	$3.1342 \cdot 10^{-2}$	8.3591	$1.7697 \cdot 10^{2}$	5.3334
0.0354	$8.3960 \cdot 10^{-3}$	8.1958	$1.7307 \cdot 10^{2}$	5.0768
0.0177	$2.2376 \cdot 10^{-3}$	8.1519	$1.7203 \cdot 10^{2}$	5.0081
0.0088	$5.9291 \cdot 10^{-4}$	8.1402	$1.7175 \cdot 10^{2}$	4.9897
0.0044	$1.5627 \cdot 10^{-4}$	8.1371	$1.7167 \cdot 10^{2}$	4.9848

TABLE 4.2
Example 1: verification results, Lipschitz constants

4.4.1. Example 2. Let us present the results of the computations for a problem similar to Example 1, but with changed parameters

$$u_a = 0.4572,\ u_b = +\infty,\ \kappa = 10^{-1},\ y_d(x_1, x_2) = \sin(15x_1x_2)e^{\frac{x_1}{2}+\frac{x_2}{3}}.$$

The discrete solution was computed to

$$\bar{u}_h = (0.4572,\ 0.7669,\ 0.7780,\ 0.8871).$$

As one can see, the inequality constraint $u_{a,1} \leq u_1$ is active, giving rise to the choice $A = \{1\}$ of the active set.

h	α_h	$\|\mathcal{E}\|_2$	α	ϵ	r_+
0.0707	$1.3108 \cdot 10^{-1}$	1.0209	$-8.8984 \cdot 10^{-1}$	$6.2166 \cdot 10^{-2}$	–
0.0354	$1.3107 \cdot 10^{-1}$	$2.7589 \cdot 10^{-1}$	$-1.4482 \cdot 10^{-1}$	$1.6510 \cdot 10^{-2}$	–
0.0177	$1.3107 \cdot 10^{-1}$	$7.8072 \cdot 10^{-2}$	$5.2997 \cdot 10^{-2}$	$4.4240 \cdot 10^{-3}$	$1.6695 \cdot 10^{-1}$
0.0088	$1.3107 \cdot 10^{-1}$	$2.4100 \cdot 10^{-2}$	$1.0697 \cdot 10^{-1}$	$1.2418 \cdot 10^{-3}$	$2.3217 \cdot 10^{-2}$
0.0044	$1.3107 \cdot 10^{-1}$	$8.5596 \cdot 10^{-3}$	$1.2251 \cdot 10^{-1}$	$3.8949 \cdot 10^{-4}$	$6.3585 \cdot 10^{-3}$

TABLE 4.3
Example 2: verification results, α, ϵ, r_+

Table 4.3 depicts the computed error bounds for different mesh sizes h. Similar to example 1, the discrete sufficient optimality condition $\alpha_h > 0$ is satisfied for all meshes, while the positive definiteness of $f''(u_h)$ can be proven only for fine meshes. As in example 1, we get the convergence of ϵ and r_+ like h^2.

h	σ_h	$\epsilon_{f'}$	σ	$\sigma - c_{f'} r_+$
0.0707	$1.1202 \cdot 10^{-2}$	$6.2166 \cdot 10^{-2}$	$-5.0964 \cdot 10^{-2}$	–
0.0354	$1.1206 \cdot 10^{-2}$	$1.6510 \cdot 10^{-2}$	$-5.3042 \cdot 10^{-3}$	–
0.0177	$1.1207 \cdot 10^{-2}$	$4.4240 \cdot 10^{-3}$	$6.7826 \cdot 10^{-3}$	-1.9550
0.0088	$1.1207 \cdot 10^{-2}$	$1.2418 \cdot 10^{-3}$	$9.9649 \cdot 10^{-3}$	$-2.6231 \cdot 10^{-1}$
0.0044	$1.1207 \cdot 10^{-2}$	$3.8949 \cdot 10^{-4}$	$1.0817 \cdot 10^{-2}$	$-6.3714 \cdot 10^{-2}$

TABLE 4.4
Example 2: verification results, strongly active constraints

Now, let us have a closer inspection of the results with respect to the strongly active inequality constraints, the associated numbers can be found in Table 4.4. As can be seen, the active constraints are strongly active for the discrete problem, i.e. $\sigma_h > 0$. Moreover, they become strongly active for the continuous problem too, as σ is positive for the fine meshes. However, we were not to be able to verify that the constraints are active at the solution of the continuous problem, too. This would require to find $\sigma - c_{f'} r_+ > 0$, which was not the case in our computations. We expect, that this condition will become true for even finer discretizations, since σ and $c_{f'}$ converge to some fixed positive value, while r_+ decays for uniform refinement. For the Lipschitz constants of f we observe a similar behavior like in Example 1, see Table 4.5.

h	$c_{f'}$	$c_{f''}$	$M_{f''}$
0.0707	$1.2160 \cdot 10^{+1}$	$2.7978 \cdot 10^{+2}$	$1.4084 \cdot 10^{+1}$
0.0354	$1.1836 \cdot 10^{+1}$	$2.7135 \cdot 10^{+2}$	$1.3347 \cdot 10^{+1}$
0.0177	$1.1750 \cdot 10^{+1}$	$2.6911 \cdot 10^{+2}$	$1.3152 \cdot 10^{+1}$
0.0088	$1.1728 \cdot 10^{+1}$	$2.6852 \cdot 10^{+2}$	$1.3100 \cdot 10^{+1}$
0.0044	$1.1721 \cdot 10^{+1}$	$2.6836 \cdot 10^{+2}$	$1.3086 \cdot 10^{+1}$

TABLE 4.5
Example 2: verification results, Lipschitz constants

Example 3. For the identification problem as analyzed in Section 4.3, the following choices were made

$$\Omega = (0,1)^2,\ u_a = -\infty,\ u_b = 0.5,\ \kappa = 10^{-2},\ y_d(x_1, x_2) = 0.5 \sin(2\pi x_1 x_2).$$

The discrete solution was computed to

$$\bar{u}_h = (0.5000, 0.2640, 0.1363, 0.0750),$$

which necessitates the choice $A = \{1\}$.

The computed bounds and constants can be found in Tables 4.6 and 4.7. As can be seen from Table 4.6, the verification assumptions were satisfied already on the coarsest mesh.

h	α_h	$\|\mathcal{E}\|_2$	α	ϵ	r_+
0.0707	$9.9631 \cdot 10^{-3}$	$4.7558 \cdot 10^{-3}$	$5.2073 \cdot 10^{-3}$	$5.3201 \cdot 10^{-4}$	$2.0433 \cdot 10^{-1}$
0.0354	$9.9631 \cdot 10^{-3}$	$1.2338 \cdot 10^{-3}$	$8.7293 \cdot 10^{-3}$	$1.3759 \cdot 10^{-4}$	$3.1523 \cdot 10^{-2}$
0.0177	$9.9631 \cdot 10^{-3}$	$3.2033 \cdot 10^{-4}$	$9.6427 \cdot 10^{-3}$	$3.5641 \cdot 10^{-5}$	$7.3924 \cdot 10^{-3}$
0.0088	$9.9631 \cdot 10^{-3}$	$8.3029 \cdot 10^{-5}$	$9.8800 \cdot 10^{-3}$	$9.2241 \cdot 10^{-6}$	$1.8672 \cdot 10^{-3}$
0.0044	$9.9631 \cdot 10^{-3}$	$2.1468 \cdot 10^{-5}$	$9.9416 \cdot 10^{-3}$	$2.3825 \cdot 10^{-6}$	$4.7931 \cdot 10^{-4}$

TABLE 4.6
Example 3: verification results, α, ϵ, r_+

Before discussing the observed behavior with respect to decreasing mesh-size, let us turn to the inspection of the results for one fixed discretization. For $h = 0.0177$, we obtained the fulfillment of Assumption 3 with $\sigma = 7.0463 \cdot 10^{-4}$, $\epsilon = 3.5641 \cdot 10^{-5}$, and $\alpha = 9.6427 \cdot 10^{-3}$. The corresponding values of $c_{f'}$ and $c_{f''}$ can be found in Table 4.7. By Theorem 3.22, there exists an optimal control $\bar{u}$ in the neighborhood of the discrete solution $\bar{u}_h$ with

$$\|\bar{u} - u_h\|_U \leq 7.3924 \cdot 10^{-3}.$$

The safety radius was adaptively computed to $R = 7.5 \cdot 10^{-3}$. Hence the assumption $r_+ \le R$ in Theorem 2.2 is fulfilled. Furthermore the condition (2.16) is satisfied with $\sigma - c_{f'} r_{+_1} = 2.5434 \cdot 10^{-1} > 0$, which implies $u_h = \bar{u}_h$ on the active set A. Additionally, we have that the second-order sufficient optimality condition is fulfilled with

$$f''(\bar{u})[v,v] \ge 3.5291 \cdot 10^{-4} \|v\|_U^2$$

for all $v \in U$ with $v|_A = 0$.

h	$c_{f'}$	$c_{f''}$	$\sigma - c_{f'} r_+$	$\alpha - c_{f''} r_+$
0.0707	$5.0265 \cdot 10^{-2}$	$2.7562 \cdot 10^{-1}$	$-1.0063 \cdot 10^{-2}$	$-5.1112 \cdot 10^{-2}$
0.0354	$4.8313 \cdot 10^{-2}$	$2.5972 \cdot 10^{-1}$	$-9.2030 \cdot 10^{-4}$	$5.4221 \cdot 10^{-4}$
0.0177	$4.7579 \cdot 10^{-2}$	$2.5434 \cdot 10^{-1}$	$3.5291 \cdot 10^{-4}$	$7.7625 \cdot 10^{-3}$
0.0088	$4.7278 \cdot 10^{-2}$	$2.5238 \cdot 10^{-1}$	$6.4277 \cdot 10^{-4}$	$9.4088 \cdot 10^{-3}$
0.0044	$4.7144 \cdot 10^{-2}$	$2.5159 \cdot 10^{-1}$	$7.1530 \cdot 10^{-4}$	$9.8210 \cdot 10^{-3}$

TABLE 4.7
Example 3: verification results, Lipschitz constants

Similar as in the previous examples, we observe a convergence rate $r_+ \sim h^2$. Moreover, for fine grids, we find that the sufficient second-order condition is satisfied at the *still unknown* local solution $\bar{u}$ of (P). First, the inequality constraints are strongly active at $\bar{u}$ on A, see the column '$\sigma - c_{f'} r_+$' in Table 4.7, which contains an estimate $|f'(\bar{u})_i| \ge \sigma - c_{f'} r_+$ for $i \in A$. And second, also $f''(\bar{u})$ is positive definite, as we have the lower bound $f''(\bar{u})[v,v] \ge (\alpha - c_{f''} r_+)\|v\|_U^2$ for $v = 0$ on A, where $\alpha - c_{f''} r_+$ can be found in Table 4.7 as well.

REFERENCES

[1] M. Ainsworth. A framework for obtaining guaranteed error bounds for finite element approximations. *J. Comput. Appl. Math.*, 234(9):2618–2632, 2010.
[2] N. Arada, E. Casas, and F. Tröltzsch. Error estimates for a semilinear elliptic control problem. *Computational Optimization and Application*, 23:201–229, 2002.
[3] R. Becker. Estimating the control error in discretized pde-constrained optimization. *Journal of Numerical Mathematics*, 14:163–185, 2006.
[4] R. Becker and R. Rannacher. An optimal control approach to a priori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.
[5] D. Braess. *Finite elements*. Cambridge University Press, Cambridge, third edition, 2007. Theory, fast solvers, and applications in elasticity theory.
[6] E. Casas, M. Mateos, and F. Tröltzsch. Error estimates for the numerical approximation of boundary semilinear elliptic control problems. *Comput. Optim. Appl.*, 31:193–219, 2005.
[7] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.
[8] A. Griesbaum, B. Kaltenbacher, and B. Vexler. Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization. *Inverse Problems*, 24(2):025025, 20, 2008.
[9] K. Kunisch, W. Liu, Y. Chang, N. Yan, and R. Li. Adaptive finite element approximation for a class of parameter estimation problems. *J. Comput. Math.*, 28(5):645–675, 2010.
[10] H. D. Mittelmann. Sufficient optimality for discretized parabolic and elliptic control problems. In *Fast solution of discretized optimization problems (Berlin, 2000)*, volume 138 of *Internat. Ser. Numer. Math.*, pages 184–196. Birkhäuser, Basel, 2001.
[11] H. D. Mittelmann. Verification of second-order sufficient optimality conditions for semilinear elliptic and parabolic control problems. *Comput. Optim. Appl.*, 20(1):93–110, 2001.
[12] M. Plum. Computer-assisted enclosure methods for elliptic differential equations. *Linear Algebra Appl.*, 324(1-3):147–187, 2001.
[13] S. Repin. *A posteriori estimates for partial differential equations*, volume 4 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH & Co. KG, Berlin, 2008.
[14] A. Rösch and D. Wachsmuth. Numerical verification of optimality conditions. *SIAM J. Control Optim.*, 47(5):2557–2581, 2008.
[15] A. Rösch and D. Wachsmuth. How to check numerically the sufficient optimality conditions for infinite-dimensional optimization problems. In *Control of Coupled Partial Differential Equations*, volume 158 of *Internat. Ser. Numer. Math.*, pages 297–317. Birkhäuser, Basel, 2009.
[16] G. Stampacchia. Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. *Ann. Inst. Fourier (Grenoble)*, 15(fasc. 1):189–258, 1965.
[17] R. Verfürth. A note on constant-free a posteriori error estimates. *SIAM J. Numer. Anal.*, 47(4):3180–3194, 2009.
[18] B. Vexler and W. Wollner. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM J. Control Optim.*, 47(1):509–534, 2008.
[19] E. Zeidler. *Nonlinear functional analysis and its applications. II/B*. Springer-Verlag, New York, 1990. Nonlinear monotone operators.
[20] J. C. Ziems and S. Ulbrich. Adaptive multilevel inexact SQP methods for PDE-constrained optimization. *SIAM J. Optim.*, 21(1):1–40, 2011.