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

How to check numerically the sufficient optimality conditions for infinite-dimensional optimization problems

A. RÃűsch, D. Wachsmuth

RICAM-Report 2008-23

How to check numerically the sufficient optimality conditions for infinite-dimensional optimization problems

Daniel Wachsmuth and Arnd Rösch

Abstract. We consider general non-convex optimal control problems. Many results for such problems rely on second-order sufficient optimality conditions. We propose a method to verify whether the second-order sufficient optimality conditions hold in a neighborhood of a numerical solution. This method is then applied to abstract optimal control problems. Finally, we consider an optimal control problem subject to a semi-linear elliptic equation that appears to have multiple local minima.

Mathematics Subject Classification (2000). Primary 49J20; Secondary 49K20, 65N15.

Keywords. optimal control, sufficient optimality conditions, verification of optimality conditions, problems with multiple minima.

1. Introduction

Let us consider the following model problem. Let U be a Hilbert space. Denote by U_{ad} the space of admissible controls, where U_{ad} is a closed, convex and non-empty subset of U. In addition, let $f : U \to \mathbb{R}$ be a twice continuously Fréchet-differentiable function. Then we are considering the problem

$$\min_{u \in U_{ad}} f(u). \tag{1.1}$$

The first-order necessary optimality condition for (1.1) reads as follows. Let $\bar{u}$ be a local solution of (1.1). Then the variational inequality

$$f'(\bar{u})(u - \bar{u}) \geq 0 \quad \forall u \in U_{ad} \tag{1.2}$$

is satisfied. An equivalent characterization is given by the inclusion

$$-f'(\bar{u}) \in N_{U_{ad}}(\bar{u}),$$

where $N_{U_{ad}}(\bar{u})$ denotes the normal cone of U_{ad} at $\bar{u}$. We will not consider second-order necessary conditions here, instead we refer to [3, 4].

A strong second-order sufficient optimality condition is satisfied at $\bar{u}$ if the condition (1.2) as well as the coercivity property

$$f''(\bar{u})[v,v] \geq \alpha\|v\|_U^2 \quad \forall v \in U \tag{1.3}$$

hold for some $\alpha > 0$. This condition ensures that $\bar{u}$ is locally optimal, and moreover, the quadratic growth condition holds: there are constants $r, \delta > 0$ such that

$$f(u) \geq f(\bar{u}) + \delta\|u - \bar{u}\|_U^2 \quad \forall u \in U_{ad} : \|u - \bar{u}\|_U \leq r.$$

Using the well-known concept of strongly active sets, see e.g. [3, 4, 8], the subspace, where f'' has to be positive definite, can be confined.

If the second-order sufficient conditions hold at the local minimum $\bar{u}$ one can prove several properties of the original optimization problem. At first, such a local solution is stable with respect to perturbations. That is, a small perturbation of the optimization problem leads only to a small perturbation in the solution. This stability is a major ingredient for convergence results, since one can interprete approximated problems as perturbations of the original one. This allows to prove local fast convergence of optimization methods (SQP, semi-smooth Newton) as well as convergence rates for finite-element discretizations of optimal control problems.

The importance of sufficient optimality conditions makes it desirable to verify whether these conditions are satisfied for a given problem. However, in condition (1.3) coercivity is assumed for the *unknown* solution. For finite-dimensional problems, one can compute eigenvalues of the Hessian matrix at some approximation of the solution since it is possible to compute this Hessian exactly. In infinite-dimensional problems, the computation of the second derivative is also prone to discretization errors. Hence, it is difficult to check whether the condition is fulfilled. This was the starting point for our investigations. We will propose a different condition, which is in fact a condition at a given approximation $\bar{u}_h$. Since only known quantities are involved, there is a chance to check this condition. For the details, we refer to Section 2. We have to admit that we can only deal with problems without two-norm discrepancy. The two-norm discrepancy occurs, whenever the the ingredients of the problem are differentiable with respect to a smaller space (say L^∞) and stronger norm, and coercivity of the second derivative only holds with respect to weaker norms (say L^2), see for instance [5].

The numerical solution of optimization problems in function spaces is often done by discretization. Let U_h be a finite-dimensional subspace of U with basis $\phi_h^1 \dots \phi_h^{N_h}$. Then an example discrete problem, which hopefully can be solved on a computer, reads as

$$\min_{u_h \in U_{ad} \cap U_h} f(u_h)$$

Given a discrete solution $\bar{u}_h$, one can introduce the discrete Hessian matrix associated with the discrete problem by

$$H = (h_{ij})_{i,j=1}^{N_h}, \quad h_{i,j} = f''(\bar{u}_h)[\phi_h^i, \phi_h^j].$$

Then one can compute the eigenvalues of H and check positive definiteness of H. If H is positive definite at $\bar{u}_h$ then $\bar{u}_h$ is a local minimum of the *discrete* problem. However, it may happen that $\bar{u}_h$ is not even close to a local solution of the original problem. Hence, the information in H is almost worthless in this case. We will present an example in Section 3, where exactly this situation occurs.

In Section 4, we will extend our approach to optimal control problems with partial differential equation. Here, we have in mind the following optimization problem

$$\min g(y) + j(u)$$

subject to

$$\begin{aligned} Ay + d(y) &= Bu, \\ u &\in U_{ad}. \end{aligned}$$

Here, a large class of semilinear elliptic state equation are covered by the analysis. in particular, steady-state Navier-Stokes equation are included. However, the differentiability requirements and the coercivity assumption are formulated with respect to the same spaces and norms. That is, problems with two-norm discrepancy are not covered.

In the article [12], the authors already suggested conditions for the numerical verification of optimality conditions. However, the analysis relied heavily on H^2-regularity of the solutions. We will overcome this restriction using a different approach for the treatment of the discretization errors.

The plan of the article is as follows. The verifiable condition is developed in Section 2. In Section 3, we introduce an example that shows that the computation of eigenvalues of the discrete Hessian cannot be taken as substitute for the condition of Section 2. The analysis concerned with optimal control problems for a semilinear elliptic equation is done in Section 4. We end the article with a report about an optimal control problem that admits two local solutions, see Section 5.

2. Coercivity condition for nonlinear programming

Let $u_h \in U_{ad} \cap U_h$ be an arbitrary, admissible point. Ideally, u_h would be the solution of a discretized problem or an approximation of it as the outcome of some iterative method. But we will not rely on this property, which is a major improvement over [12].

Now, let us present the coercivity condition. At first, we assume that we can find bounds of certain characteristics of f' and f''.

Assumption 2.1. There are constants ϵ, α, M, R such that the following three inequalities hold:

$$f'(u_h)(u-u_h) \geq -\epsilon\|u-u_h\|_U \qquad \forall u \in U_{ad}, \tag{2.1}$$

$$f''(u_h)[v,v] \geq \alpha\|v\|_U^2 \qquad \forall v \in U, \tag{2.2}$$

$$|(f''(u) - f''(u_h))[v_1,v_2]| \leq M\|u-u_h\|_U\|v_1\|_U\|v_2\|_U \qquad \forall u \in U_{ad},\ \|u-u_h\|_U \leq R,\ v_1, v_2 \in U. \tag{2.3}$$

Let us comment on the three inequalities involved in the assumption. The first one (2.1) measures in some sense the residuum in the variational inequality (1.2). The second inequality is a coercivity assumption on f'' at u_h. The essential difference to (1.3) is that the point, where we have to check for coercivity of f'', is known.

Moreover, these conditions are analogous to the pre-requisites of convergence theorems of Newton's method: smallness of initial residual, bounded invertibility, and local Lipschitz estimates. See also the comments below.

Let us now take another admissible point $u \in U_{ad}$. With the help of Assumption 2.1, we can estimate the difference between $f(u)$ and $f(u_h)$ using Taylor expansion as

$$\begin{aligned} f(u) - f(u_h) &\geq f'(u_h)(u-u_h) + \frac{1}{2}f''(u_h)(u-u_h)^2 \\ &\quad + \int_0^1\int_0^s \left(f''(u_h + t(u-u_h)) - f''(u_h)\right)(u-u_h)^2 \mathrm{d}t\mathrm{d}s \\ &\geq -\epsilon\|u-u_h\|_U + \frac{\alpha}{2}\|u-u_h\|_U^2 - \frac{M}{6}\|u-u_h\|_U^3. \end{aligned} \tag{2.4}$$

In addition to Assumption 2.1, we need a further qualification, which relates the constants appearing there to each other.

Assumption 2.2. There exists a real number r_+ with $R > r_+ > 0$ such that

$$-\epsilon r_+ + \frac{\alpha}{2}r_+^2 - \frac{M}{6}r_+^3 > 0, \tag{2.5}$$

$$\alpha - Mr_+ > 0 \tag{2.6}$$

is satisfied.

The Assumptions 2.1 and 2.2 allow us to prove the main result of this section.

Theorem 2.3. *Let Assumptions* 2.1 *and* 2.2 *be satisfied. Then there exists a solution* $\bar{u}$ *of the original problem* (1.1) *with*

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

Furthermore, the second-order sufficient optimality condition (1.3) *is satisfied at* $\bar{u}$.

Proof. Let us consider the optimization problem (1.1) but restricted to the closed ball centered at u_h with radius r_+,

$$\min_{u\in\bar{B}(u_h,r_+)\cap U_{ad}} f(u).$$

Due to (2.6), the function f is convex on $\bar{B}(u_h, r_+)$. Hence, the auxiliary problem admits a global minimum $\bar{u}$ with $\|\bar{u}-u_h\|_H \le r_+$. Moreover, the second derivative of f is positive definite at $\bar{u}$ by (2.2) and (2.6).

By (2.5), $\bar{u}$ cannot lie on the boundary of $B(u_h, r_+)$, since the value of f is there larger than in u_h. That is, $\bar{u}$ is also the global minimum of f over the intersection of U_{ad} with the open ball at u_h with radius r_+. Hence, $\bar{u}$ is a local solution of the original problem (1.1). □

The consequences of this result are threefold: at first we obtain existence of a solution of the original problem in the specified neighborhood. Secondly, we can estimate the distance to the solution. And third, we can prove that this yet unknown solution fulfills the second-order optimality condition.

The inequality (2.5) is an assumption on the objective functional. We can replace it by an assumption on the first derivative f', and can prove an result analogous to Theorem 2.3.

Assumption 2.4. There exists a real number $\tilde{r}_+$ with $R > \tilde{r}_+ > 0$ such that

$$-\epsilon + \alpha\tilde{r}_+ - \frac{M}{2}\tilde{r}_+^2 > 0 \tag{2.7}$$

is satisfied.

Theorem 2.5. *Let the assumptions* 2.1 *and* 2.4 *be satisfied. Then there exists a solution* $\bar{u}$ *of the original problem* (1.1) *with*

$$\|\bar{u} - u_h\|_H < \tilde{r}_+.$$

Furthermore, the second-order sufficient optimality condition (1.3) *is satisfied.*

Proof. At first, we have to show that Assumption 2.4 implies the convexity of f is a neighborhood of u_h. Let us define a polynomial p by $p(r) = -\epsilon + \alpha\tilde{r} - \frac{M}{2}\tilde{r}^2$. We already know $p(0) < 0$ and $p(\tilde{r}_+) > 0$. Hence, there is a $\tilde{r}_0 \in (0, \tilde{r}_+)$ such that $p(\tilde{r}_0) = 0$. The root $\tilde{r}_0$ is given by $\tilde{r}_1 = \frac{\alpha}{M}\left(1 - \sqrt{1 - \frac{2M\epsilon}{\alpha^2}}\right)$. Moreover, it holds $\alpha - M\tilde{r}_0 = \alpha\sqrt{1 - \frac{2M\epsilon}{\alpha^2}} > 0$. Hence, there is a $\tilde{r}_1 \in (\tilde{r}_1, \tilde{r}_+)$ such that (2.6) and (2.7) are satisfied for $\tilde{r}_1$. This implies the convexity of f on the ball centered at u_h with radius $\tilde{r}_1$.

As in the proof of the previous Theorem 2.3, we obtain then the existence of a global solution $\bar{u}$ of the problem

$$\min_{u\in\bar{B}(u_h,\tilde{r}_1)\cap U_{ad}} f(u).$$

It remains to show that $\bar u$ is not on the boundary of $\bar B(u_h,\tilde r_1)$. Let us take an arbitray $u \in U_{ad}$ with $\|u-u_h\|_U = \tilde r_1$. Using (2.7), we obtain

$$\begin{aligned} f'(u)(u_h-u) &= f'(u_h)(u_h-u) + f''(u_h)(u_h-u,u-u_h) \\ &\quad + \int_0^1 \left(f''(u_h+t(u-u_h)) - f''(u_h)\right)(u_h-u,u-u_h)\mathrm{d}t \\ &\le \left(\epsilon - \alpha\|u-u_h\|_U + \frac{M}{2}\|u-u_h\|_U^2\right)\|u-u_h\|_U < 0. \end{aligned}$$

Hence, the necessary optimality conditions of (1.2) are not fulfilled for any control on the boundary of $\bar B(u_h,\tilde r_1)$. Thus, the solution $\bar u$ satisfies $\|\bar u - u_h\|_U < \tilde r_1 \le \tilde r_+$. Furthermore, it is a local solution of the original problem (1.1). □

A close inspection of the proof reveals that we can show an improved error estimate:

Corollary 2.6. *Let the assumptions of the previous theorem be fulfilled. Then it holds*

$$\|\bar u - u_h\|_U \le \frac{\alpha}{M}\left(1 - \sqrt{1 - \frac{2M\epsilon}{\alpha^2}}\right).$$

Proof. If Assumption 2.4 is satisfied with some $\tilde r_+$, then it will be satisfied for all r between $\frac{\alpha}{M}\left(1-\sqrt{1-\frac{2M\epsilon}{\alpha^2}}\right)$, which is the first root of the polynomial in (2.7), and $\tilde r_+$. Then Theorem 2.5 yields the claim. □

The Theorems 2.3 and 2.5 state that the yet unknown solution $\bar u$ satisfies the second-order sufficient optimality condition. This implies that it is possible apply deeper results, which rely on this conditions. For instance, we can apply results for the fast local convergence of optimization methods. That is, if the initial guess is close enough to the solution then the iterates will converge with a high convergence rate towards the solution $\bar u$.

Let us show exemplarily the fast convergence of Newton's method for generalized equations in the sense of [1, 7] if started at u_h. The key idea here is to write the variational inequality (1.2) as the inclusion

$$-f'(\bar u) \in N_{U_{ad}}(u).$$

Then the generalized Newton method solves for u_{k+1} the problem

$$-\left(f''(u_k)(u-u_k) + f'(u_k)\right) \in N_{U_{ad}}(u), \tag{2.8}$$

which is the first-order necessary optimality condition of

$$\min_{u\in U_{ad}} \frac{1}{2} f''(u_k)(u-u_k)^2 + f'(u_k)(u-u_k).$$

That is, only the objective function is linearized but not the constraint $u \in U_{ad}$. It turns out, that the conditions of Theorem 2.5 are sufficient to ensure local quadratic convergence of the simple iteration (2.8) for the initial choice $u_0 = u_h$.

Theorem 2.7. *Let the assumptions of Theorem* 2.5 *be satisfied. Set* $u_0 := u_h$. *Then the sequence of iterates generated by the procedure* (2.8) *converges quadratically to* $\bar{u}$.

Proof. Let us assume first, that the equation (2.8) is solvable for some k. The iterate u_{k+1} satisfies the variational inequality

$$(f''(u_k)(u_{k+1} - u_k) + f'(u_k),\, u - u_{k+1}) \geq 0 \quad \forall u \in U_{ad}. \tag{2.9}$$

Setting $k = 0$, $u_0 = u = u_h$, and using (2.1), we obtain the following estimate of the initial step

$$\|u_1 - u_0\|_U \leq \frac{\epsilon}{\alpha}.$$

Let us denote by α_k the smallest eigenvalue of $f''(u_k)$. Then it holds $\alpha_0 = \alpha$ and $\alpha_{k+1} \geq \alpha_k - M\|u_k - u_{k+1}\|_U$. Setting $u := u_k$ in (2.9), we find applying (2.9) for u_k

$$\begin{aligned} f''(u_k)(u_k - u_{k+1})^2 &\leq f'(u_k)(u_k - u_{k+1}) \\ &= (f'(u_{k-1}) + f''(u_{k-1})(u_k - u_{k-1}),\, u_k - u_{k+1}) \\ &\quad + \int_0^1 \left(f''(u_{k-1} + t(u_k - u_{k-1})) - f''(u_{k-1})\right)(u_k - u_{k-1},\, u_k - u_{k+1})\mathrm{d}t. \end{aligned}$$

Setting $k-1$ for k in (2.9), we find the optimality relation for u_k, which implies that the first part of the right-hand side is non-positive. Applying Assumption 2.1, we obtain

$$f''(u_k)(u_k - u_{k+1})^2 \leq \frac{M}{2}\|u_k - u_{k-1}\|_U^2 \|u_k - u_{k+1}\|_U,$$

which gives

$$\|u_k - u_{k+1}\|_U \leq \frac{M}{2\alpha_k}\|u_k - u_{k-1}\|_U^2 \leq \frac{M}{2(\alpha_{k-1} - M\|u_k - u_{k-1}\|_U)}\|u_k - u_{k-1}\|_U^2.$$

Now, we can proceed as in Ortega's proof of the Newton-Kantorovich theorem [10], see also [6]. The technique applied there delivers (a) existence of solutions of (2.9) for all k, and (b) quadratic convergence. Moreover, the convergence region of Newton's method given by [10] is the ball at u_h with the radius given by Corollary 2.6. ☐

The similarities to the convergence proof of Newton-Kantorovich type is obvious. That is, the assumptions above can be interpreted as assumptions in the context of Newtons' method and vice-versa. This observation allows also to apply heuristic techniques to estimate the constants appearing in Assumption 2.1 during the procedure of Newton's method. For a detailed explanation of these techniques we refer to the monograph of Deuflhard [6].

3. On coercivity of f'' and positive definiteness of the discrete Hessian

The computation of the constants appearing in the Assumption 2.1 above is a difficult task, it is especially hard to find a lower bound for the smallest eigenvalue α of f''. Here, it would be advantageous if one could compute α as an eigenvalue of the discrete problem, which fulfills

$$f''(\tilde{u}_h)(v,v) \geq \alpha \|v\|_U^2 \quad \forall v \in U_h$$

for a finite-dimensional subspace $U_h \subset U$. This method is widely employed in numerical experiments to indicate optimality of computed solution. However, despite being attractive from a computational point of view, this method is not save in general and may lead to wrong conclusions.

We will now construct an example with numerical solution $\tilde{u}_h$ that has the following properties:

- $f'(\tilde{u}_h) = 0$,
- $f''(\tilde{u}_h)(v_h, v_h) \geq \alpha \|v_h\|_U^2 \quad \forall v_h \in U_h$ with $\alpha > 0$,
- $\tilde{u}_h$ is not close to a local minimum of the original problem.

That means in particular, that all eigenvalues of the discrete Hessian matrix are positive. Hence, $\tilde{u}_h$ is a local minimum of the discretized problem. Unfortunately, it appears that $\tilde{u}_h$ is not even in the neighborhood of a local minimum of the original problem. Thus, the positive definiteness of the discrete Hessian is misleading.

Minimizing a fourth-order polynomial

We will consider now a special objective function. Let be given $u_1 \neq u_2$ from the Hilbert space U. Then we want to minimize

$$f(u) = \frac{1}{2}\|u - u_1\|_U^2 \|u - u_2\|_U^2. \tag{3.1}$$

Of course, both u_1 and u_2 are global minima of this problem. Now, let us have a look on the derivatives of f. The first derivative is given by

$$f'(u) = (u - u_1)\|u - u_2\|_U^2 + (u - u_2)\|u - u_1\|_U^2. \tag{3.2}$$

And it turns out that $\tilde{u} := \frac{1}{2}(u_1 + u_2)$ is a stationary point. If U is one-dimensional then $\tilde{u}$ is a local maximum of f. For higher dimensional U, $\tilde{u}$ is actually a saddle point as we will see. Hence, let us assume in the sequel that the dimension of U is greater than one.

The second derivative of f is given as bilinear form by

$$\begin{aligned} f''(u)(v_1, v_2) = &\left(\|u - u_1\|_U^2 + \|u - u_2\|_U^2\right)(v_1, v_2) \\ &+ 2(u - u_2, v_1)(u - u_1, v_2) + 2(u - u_1, v_1)(u - u_2, v_2). \end{aligned} \tag{3.3}$$

Formally, one can decompose f'' into $D + 2VV^T$, where D is a positive multiple of the identity and VV^T is a two-rank perturbation. This simplifies the computation of eigenvalues of f''. Let us set $u = \tilde{u}$ and $v_1 = v_2 = v$ in (3.3). We obtain

$$f''(\tilde{u})(v,v) = \frac{1}{2}\|u_1 - u_2\|_U^2 \|v\|_U^2 - (v, u_1 - u_2)^2. \tag{3.4}$$

Let us decompose the space U as the direct sum: $\operatorname{span}\{u_1 - u_2\} \oplus \{u_1 - u_2\}^\perp$. Then we can write $v = v_1 + v_2$ with $(v_2, u_1 - u_2) = 0$, which gives

$$f''(\tilde{u})(v,v) = \frac{1}{2}\|u_1 - u_2\|_U^2 (\|v_1\|_U^2 + \|v_2\|_U^2) - (v_1, u_1 - u_2)^2. \tag{3.5}$$

For $v_1 \in \operatorname{span}\{u_1 - u_2\}$ it holds $(v_1, u_1 - u_2)^2 = \|u_1 - u_2\|_U^2 \|v_1\|_U^2$, which implies

$$f''(\tilde{u})(v,v) = \frac{1}{2}\|u_1 - u_2\|_U^2 (\|v_2\|_U^2 - \|v_1\|_U^2).$$

Thus, for the direction of negative curvature $v = s(u_1 - u_2)$ we have

$$f''(\tilde{u})(v,v) = -\frac{1}{2}\|u_1 - u_2\|_U^2 \|v\|_U^2.$$

With similar arguments, one finds the inequality

$$f''(u)(v,v) \geq \lambda_1(u)\|v\|_U^2$$

with

$$\lambda_1(u) = \|u - u_1\|_U^2 + \|u - u_2\|_U^2 - 2\|u - u_1\|_U \|u - u_2\|_U + 2(u - u_1, u - u_2). \tag{3.6}$$

Let us denote by U_h a finite-dimensional subspace of U. The orthogonal projection from U onto U_h is denoted by Π_h. Let us recall the expression for $f''(\tilde{u})$, cf. 3.4,

$$f''(\tilde{u})(v,v) = \frac{1}{2}\|u_1 - u_2\|_U^2 \|v\|_U^2 - (v, u_1 - u_2)^2.$$

We will now derive conditions such that $f''(\tilde{u})[v_h, v_h] > 0$ is fulfilled for all $v_h \neq 0$ from the finite-dimensional space U_h. Let us consider for a moment directions v_h with $\|v_h\|_U = 1$. The supremum of $(v_h, u_1 - u_2)$ over all such v_h is attained at $v_h = \frac{\Pi_h(u_1 - u_2)}{\|\Pi_h(u_1 - u_2)\|_U}$, which implies for $\|v_h\|_U = 1$

$$\begin{aligned} f''(\tilde{u})(v_h, v_h) &= \frac{1}{2}\|u_1 - u_2\|_U^2 - (v_h, u_1 - u_2)^2 \\ &\geq \frac{1}{2}\|u_1 - u_2\|_U^2 - \|\Pi_h(u_1 - u_2)\|_U^2. \end{aligned}$$

Using $\|u_1 - u_2\|_U^2 = \|(I - \Pi_h)(u_1 - u_2)\|_U^2 + \|\Pi_h(u_1 - u_2)\|_U^2$, we find that $f''(\tilde{u})$ is positive definite on U_h if

$$\|(I - \Pi_h)(u_1 - u_2)\|_U \geq \|\Pi_h(u_1 - u_2)\|_U$$

holds. That is, if the L^2-norm of the projection $\Pi_h(u_1 - u_2)$ captures less than one half of the L^2-norm of $u_1 - u_2$, then the bilinear form $f''(\tilde{u})$ is positive definite on U_h despite being indefinite on whole U. Or in other words, if the discretization is too coarse to approximate the direction of negative curvature $\tilde{v} = u_1 - u_2$ the bilinear form is positive definite on U_h.

Let us demonstrate that such a situation may occur for a concrete optimization problem. Let us define U to be the set of square integrable functions on $I = (0, 1)$, $U := L^2(0, 1)$. Take an integer number N, set $h := 1/N$. The interval I is subdivided into N subintervals I_j of equal length h, $j = 1 \dots N$. The discrete subspace U_h is chosen as the space of piecewise constant function on the intervals I_j. Let us denote by Π_h the L^2-projector onto U_h, $\Pi_h : L^2(I) \to U_h$. The functions u_1, u_2 we will choose such that

1. $\tilde{u} = (u_1 + u_2)/2$ is in U_h for all h,
2. the direction of negative curvature $\tilde{v}$ is 'hard to approximate' with functions from U_h.

Let $\epsilon > 0$ be a small number. We define

$$u_1(x) = x^{-1/2+\epsilon}, \quad u_2(x) = -u_1(x).$$

Then obviously we have $u_1 + u_2 = 0 \in U_h$. Moreover, it holds $u_1 \in L^2(I)$ and $u_1 \notin H^1(I)$. The latter property is the reason, why u_1 can only be approximated with low convergence rates with respect to h.

Lemma 3.1. *For the above choice of U_h and u_1 it holds*

$$\|u_1 - \Pi_h u_1\|_U \geq g(\epsilon) h^\epsilon$$

with $g(\epsilon) = \frac{|2\epsilon-1|}{\sqrt{2\epsilon}|2\epsilon+1|}$.

Proof. Let us only consider the approximation of u_1 by a function w that is constant on the first subinterval $(0, h)$ and equal to u on $(h, 1)$,

$$w(x) = \begin{cases} w_0 & \text{if } x \in (0, h) \\ u_1(x) & \text{if } x \in [h, 1) \end{cases}$$

with w_0 in $\mathbb{R}$. It is clear that $\|u_1 - \Pi_h u_1\|_U \geq \|u_1 - w\|_U$ holds. A short computation yields that $w_0 = \frac{h^{-1/2+\epsilon}}{1/2+\epsilon}$ minimizes $\|u_1 - w\|_U^2$ over all choices of w_0, which in turn gives

$$\int_0^h (u(x) - w_0)^2 \mathrm{d}x = h^{2\epsilon} \frac{(\epsilon - 1/2)^2}{2\epsilon(\epsilon + 1/2)^2},$$

and the claim is proven. □

Let us remark, that the previous lemma not only gives an arbitrary small convergence rate for $\epsilon \to 0+$ but also states that the constant explodes for ϵ tending to zero.

In Table 1 we computed the "critical values" of h^0. If the mesh size h is larger than h^0, then one gets a wrong indication by the eigenvalues of the discrete Hessian: The smallest eigenvalue of the Hessian of the discretized problem is positive, but the computed solution is only a saddle point of the original problem. Consequently, the usual strategy to look at the smallest eigenvalue of the Hessian fails for this simple problem. The last line of Table 1 shows that this wrong indication can occur even for very small discretization parameters.

ϵ	h_0	
0.05	1/18	$= 0.056$
0.04	1/106	$= 0.0094$
0.03	1/1917	$= 5.2 \cdot 10^{-4}$
0.02	1/619660	$= 1.6 \cdot 10^{-6}$

TABLE 1. Critical mesh sizes

4. Application to an abstract optimal control problem

Let us consider now a more complicated optimization problem. We will introduce an additional constraint, which will mimic a partial differential equation. We investigate the minimization of the functional

$$J(y,u) := g(y) + j(u) \tag{4.1}$$

subject to

$$Ay + d(y) = Bu, \tag{4.2}$$

$$u \in U_{ad}. \tag{4.3}$$

Here, y denotes the state of the system, u the control. Let U_{ad} be a closed, convex, non-empty subset of the Hilbert space U.

Assumption 4.1. Let Y be a Banach space. Let $A : Y \to Y'$ and $B : U \to Y'$ be linear operators. Moreover, we assume A to be coercive, i.e. it holds $\langle Ay, y\rangle_{Y',Y} \geq \delta \|y\|_Y^2$ for some $\delta > 0$ and all $y \in Y$.

The functions d, g, j are twice Fréchet-differentiable as functions from Y to Y', Y to $\mathbb{R}$, and U to $\mathbb{R}$, respectively. Moreover, we assume for simplicity that d is monotone with $d(0) = 0$.

Thus, the state equation (4.2) has to hold in Y'. Under the assumptions above, this equation is uniquely solvable for each control u. Let us denote the solution mapping by S, i.e. $y = S(u)$ is the solution of (4.2). Since d is monotone, the linearized equation

$$Ay + d'(\tilde{y})y = Bu$$

is solvable, where we set $\tilde{y} = S(\tilde{u})$. In addition, there exists an upper bound of the norm of its solution operator $S'(\tilde{u})$,

$$\|S'(u)\|_{L(Y',Y)} = \|(A + d'(y))^{-1}\|_{L(Y',Y)} \leq c_A \quad \forall u \in U, y = S(u).$$

In view of this estimate, we can directly give a Lipschitz estimate for solutions of (4.2)

$$\|S(u_1) - S(u_2)\|_Y \leq c_A \|B\| \cdot \|u_1 - u_2\|_U. \tag{4.4}$$

Assumption 4.2. Let us take $R > 0$ and $u_h \in U_{ad}$. Then we assume that there exist constants $c_{g'}, c_{d'}, c_{d''}, c_{g''}, c_{j''}$ depending on R such that the local Lipschitz estimates

$$\begin{aligned}
\|d'(y) - d'(y_h)\|_{L(Y,Y')} &\leq c_{d'}\|y - y_h\|_Y \\
\|g'(y) - g'(y_h)\|_{Y'} &\leq c_{g'}\|y - y_h\|_Y \\
\|d''(y) - d''(y_h)\|_{L(Y\times Y,Y')} &\leq c_{d''}\|y - y_h\|_Y \\
\|g''(y) - g''(y_h)\|_{(Y\times Y)'} &\leq c_{g''}\|y - y_h\|_Y \\
\|j''(u) - j''(u_h)\|_{(U\times U)'} &\leq c_{j''}\|u - u_h\|_Y
\end{aligned}$$

hold for all $u \in U_{ad}$ with $\|u - u_h\|_U < R$ and $y = S(u)$.

The problem class covered by Assumptions 4.1 and 4.2 is wide enough to cover distributed or boundary control problems for semiliner elliptic equations. Moreover, the case of the steady-state Navier-Stokes equations fits also in the assumption. However, we have to admit that optimal control problems with two-norm discrepancy are not included.

Let us define the reduced cost functional $\phi : U \to \mathbb{R}$ by

$$\phi(u) := g(S(u)) + j(u).$$

The conditions in Section 2, i.e. Assumptions 2.1, 2.2, and 2.4, have now to be interpreted as conditions on the reduced cost functional. The reduced functional of course inherits the structure of the optimal control problem (4.1)–(4.3). So we will express the conditions on ϕ in terms of the original problem.

Let $(\bar{u}, \bar{y})$ be an admissible pair for (4.1)–(4.3). If $\bar{u}$ is locally optimal, then there exists an adjoint state $\bar{p} \in Y$ such that it holds

$$A^*\bar{p} + d'(\bar{y})^*\bar{p} = g'(\bar{y})$$

and

$$(j'(\bar{u}) + B^*\bar{p},\, u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}.$$

Let now (u_h, y_h, p_h) be some triple consisting of approximations of a locally optimal control, state, and adjoint. Suppose u_h is an admissible control. Let us assume that we can control the residuals of the optimality system.

Assumption 4.3. There are positive constants $\epsilon_u, \epsilon_y, \epsilon_p$ such that it holds

$$(j'(u_h) + B^*p_h,\, u - u_h) \geq -\epsilon_u\|u - u_h\|_U \quad \forall u \in U_{ad}, \tag{4.5}$$

$$\|Ay_h + d(y_h) - Bu_h\|_{Y'} \leq \epsilon_y, \tag{4.6}$$

$$\|A^*p_h + d'(y_h)^*p_h - g'(y_h)\|_{Y'} \leq \epsilon_p. \tag{4.7}$$

This assumption corresponds to (2.1) in Assumption 2.1 of Section 2. We will now investigate the error in the variational inequality (2.1), i.e. we want to estimate ϵ in

$$\phi'(u_h)(u - u_h) \geq -\epsilon\|u - u_h\|_U \quad \forall u \in U_{ad}.$$

To characterize the derivative ϕ' in terms of the original problem let us introduce two auxiliary functions y^h and p^h as the solutions of

$$\begin{aligned} Ay^h + d(y^h) &= u_h, \\ A^*p^h + d'(y^h)^*p^h &= g'(y^h). \end{aligned} \tag{4.8}$$

We have the following error estimates for these states and adjoints:

Lemma 4.4. *Let y^h and p^h be given by (4.8). Then it holds*

$$\|y^h - y_h\|_Y \le c_A\epsilon_y, \tag{4.9}$$

$$\|p^h - p_h\|_Y \le c_A\left((c_{g'} + c_{d'}\|p_h\|_Y)\|y^h - y_h\|_Y + \epsilon_p\right). \tag{4.10}$$

Proof. The difference $p^h - p_h$ solves the equation

$$\begin{aligned} A^*(p^h - p_h) + d'(y^h)^*(p^h - p_h) &= g'(y^h) - g'(y_h) \\ &\quad - (A^*p^h + d'(y_h)^*p_h - g'(y_h)) + (d'(y_h)^* - d'(y^h)^*)p_h, \end{aligned}$$

which immediately gives (4.10) using the notations of Assumption 4.2. The difference $y^h - y_h$ can be treated similarly, and one obtains $\|y^h - y_h\|_Y \le c_A\epsilon_y$. □

Observe, that y^h and p^h can be written as $y^h = S(u_h)$ and $p^h = S'(u_h)^*g'(y^h)$. Hence, we can rewrite the first derivative ϕ' as

$$\begin{aligned} \phi'(u_h)(u - u_h) &= (j'(u_h) + B * S'(u_h)^*g'(y^h),\ u - u_h) \\ &= (j'(u_h) + B^*p^h, u - u_h). \end{aligned}$$

Lemma 4.5. *The following inequality is satisfied for all admissible controls $u \in U_{ad}$*

$$\phi'(u_h)(u - u_h) \ge -\epsilon\|u - u_h\|_U,$$

where ϵ is given by

$$\epsilon := \epsilon_u + \|B\| \cdot \|p^h - p_h\|_Y$$

Proof. The claim follows immediately from

$$\begin{aligned} \phi'(u_h)(u - u_h) &= (j'(u_h) + B^*p^h, u - u_h) = (j'(u_h) + B^*p_h + B^*(p^h - p_h), u - u_h) \\ &\ge -\left(\epsilon_u + \|B\| \cdot \|p^h - p_h\|_Y\right)\|u - u_h\|_U. \end{aligned}$$

□

At next, we need a coercivity condition on the second derivative of the Lagrangian involving known quantities only.

Assumption 4.6. There is a constant $\delta > 0$ such that

$$j''(u_h)(v, v) + g''(y_h)(z, z) + d''(y_h)(z, z)p_h \ge \delta\|v\|^2 \tag{4.11}$$

holds for all $v = u - u_h$, $u \in U_{ad}$ with z being the solution of the linearized equation.

$$Az + d'(y_h)z = Bv. \tag{4.12}$$

This condition is especially fulfilled for convex j and g and under the sign condition $d''(y_h)p_h > 0$. We will now derive an lower bound for the eigenvalues of ϕ'' analogously to (2.2).

Lemma 4.7. *Let $v = u - u_h$, $u \in U_{ad}$ be given. Then it holds*

$$\phi''(u_h)(v, v) \geq \alpha \|v\|_U^2$$

with α given by

$$\alpha = \delta - \left(c_{g''} + \|p^h - p_h\|_Y (c_{d''} + \|d''(y_h)\|) + c_{d''}\|p_h\|_Y\right) \|y^h - y_h\|_Y (c_A\|B\|)^2$$
$$-\left(\|g''(y_h)\|_{(Y\times Y)'} + \|p_h\|_Y \|d''(y_h)\|_{L(Y\times Y,Y')}\right) (2c_A\|B\|)(c_A c_{d'}\|y_h - y^h\|_Y c_A\|B\|).$$

Proof. We can write the second derivative of ϕ as

$$\phi''(u_h)(v, v) = j''(u_h)(v, v) + g''(y^h)(z^h, z^h) + d''(y^h)(z^h, z^h)p^h, \tag{4.13}$$

where z^h solves

$$Az^h + d'(y^h)z^h = Bv. \tag{4.14}$$

Let us denote by z the solution of (4.12). A-priori bounds of z and z^h can be calculated as above, and we obtain

$$\|z\|_Y, \|z^h\|_Y \leq c_A\|B\| \cdot \|v\|_U.$$

The difference $z^h - z$ solves $A(z^h - z) + d'(y^h)(z^h - z) = (d'(y_h) - d'(y^h))z$, hence it holds

$$\|z^h - z\|_Y \leq c_A c_{d'} \|y_h - y^h\|_Y c_A\|B\|\|v\|_U.$$

Let us introduce the abbreviations $s^h := g''(y^h) + d''(y^h)(\cdot, \cdot)p^h$, $s_h := g''(y_h) + d''(y_h)(\cdot, \cdot)p_h$; $s^h, s_h : Y \times Y \to \mathbb{R}$. Then we write

$$s^h(z^h, z^h) = s_h(z, z) + (s^h - s_h)(z^h, z^h) + s_h((z^h, z^h) - (z, z)).$$

Here, the first addend appears in the coercivity assumption (4.11). The second addend is estimated as

$$\|s^h - s_h\|_{(Y\times Y)'} \leq \Big(c_{g''} + \|p^h - p_h\|_Y \|d''(y_h)\|_{L(Y\times Y,Y')}$$
$$+ \left(\|p_h\|_Y + \|p^h - p_h\|_Y\right) c_{d''}\Big) \|y^h - y_h\|_Y.$$

For the third one we obtain

$$\|s_h((z^h, z^h) - (z, z))\|_{(Y\times Y)'}$$
$$\leq \left(\|g''(y_h)\|_{(Y\times Y)'} + \|p_h\|_Y \|d''(y_h)\|_{L(Y\times Y,Y')}\right) \|z^h + z\|_Y \|z^h - z\|_Y.$$

Putting everything together we finally find

$$\phi''(u_h)(v, v) \geq \alpha \|v\|_U^2$$

with α equal to

$$\delta - \left(c_{g''} + \|p^h - p_h\|_Y (c_{d''} + \|d''(y_h)\|_{L(Y\times Y,Y')}) + c_{d''}\|p_h\|_Y\right) \|y^h - y_h\|_Y (c_A\|B\|)^2$$
$$-\left(\|g''(y_h)\|_{(Y\times Y)'} + \|p_h\|_Y \|d''(y_h)\|_{L(Y\times Y,Y')}\right) (2c_A\|B\|)(c_A c_{d'}\|y_h - y^h\|_Y c_A\|B\|).$$

□

According to Lemma 4.4, that the negative terms in the estimate of α are of the order of ϵ_y, ϵ_p. That is, there is hope that α is positive for small residuals in the optimality system. That will be true in particular, if (u_h, y_h, p_h) solves a very fine discretized problem and a second-order sufficient optimality condition is fulfilled for the original problem (4.1)–(4.3).

Corollary 4.8. *If Assumption* 4.6 *holds with the linearized equation* (4.14) *instead of* (4.12)*, then the statement of Lemma* 4.7 *is valid with*

$$\alpha = \delta - \left(c_{g''} + \|p^h - p_h\|_Y (c_{d''} + \|d''(y_h)\|) + c_{d''}\|p_h\|_Y\right) \|y^h - y_h\|_Y (c_A\|B\|)^2.$$

Finally, we have to compute the Lipschitz constant of ϕ'' as equivalent to inequality (2.3) in Section 2.

Lemma 4.9. *There is a constant $M > 0$ such that it holds for all $u \in U_{ad}$ with $\|u - u_h\|_U < R$*

$$|(\phi''(u) - \phi''(u_h))(v_1, v_2)| \le M\|u - u_h\|_U \|v_1\|_U \|v_2\|_U \quad \forall v_1, v_2 \in U.$$

An upper bound of M is given in the course of the proof.

Proof. Let y and p be the solutions of the state and adjoint equations associated with u, i.e. they satisfy

$$\begin{aligned} Ay + d(y) &= u, \\ A^*p + d'(y)^*p &= g'(y). \end{aligned}$$

Then it holds

$$\begin{aligned}(\phi''(u) - \phi''(u_h))(v_1, v_2) = (j''(u) - j''(u_h))(v_1, v_2) + (g''(y) - g''(y^h))(z_1, z_2) \\ + (pd''(y) - p^h d''(y^h))(z_1, z_2),\end{aligned}$$

where the z_i, $i = 1, 2$, are the solutions of the linearized equations $Az_i + d'(y^h)z_i = Bv_i$. Using Lipschitz continuity of the solution mapping, cf. (4.4), we obtain

$$\begin{aligned} \|y - y^h\|_Y &\le c_A\|B\| \cdot \|u - u_h\| && \le c_A\|B\|R, \\ \|y - y_h\|_Y &\le c_A(\|B\| \cdot \|u - u_h\| + \epsilon_y) && \le c_A(\|B\|R + \epsilon_y). \end{aligned} \tag{4.15}$$

Similarly to (4.10), the difference of the adjoint states is estimated by

$$\begin{aligned} \|p - p^h\|_Y &\le c_A(c_{g'} + c_{d'}(\|p^h - p_h\|_Y + \|p_h\|_Y))\|y - y^h\|_Y \\ &\le c_A(c_{g'} + c_{d'}(\|p^h - p_h\|_Y + \|p_h\|_Y))c_A\|B\|R. \end{aligned} \tag{4.16}$$

Employing the splitting

$$\begin{aligned} pd''(y) - p^h d''(y^h) &= (p - p^h)d''(y) + p^h(d''(p) - d''(y^h)) \\ &= (p - p^h)\left(d''(y_h) + d''(y) - d''(y_h)\right) \\ &\quad + (p_h + p^h - p_h)(d''(y) - d''(y^h)) \end{aligned}$$

we can estimate

$$\begin{aligned}\|d''(y)(\cdot,\cdot)p - d''(y^h)(\cdot,\cdot)p^h\|_{(Y\times Y)'} &\le \|p-p^h\|_Y\|d''(y_h)\|_{L(Y\times Y,Y')}\\ &+ c_{d''}\left(\|p-p^h\|_Y\|y-y_h\|_Y + (\|p_h\|_Y + \|p^h-p_h\|_Y)\|y-y^h\|_Y\right)\end{aligned}$$

And the claim of the Lemma holds with

$$\begin{aligned}M \ge c_{j''} + (c_A\|B\|)^2\Big(&c_{g''}\|y-y^h\|_Y + \|p-p^h\|_Y\|d''(y_h)\|_{L(Y\times Y,Y')}\\ &+ c_{d''}\left(\|p-p^h\|_Y\|y-y_h\|_Y + (\|p_h\|_Y + \|p^h-p_h\|_Y)\|y-y^h\|_Y\right)\Big),\end{aligned}$$

where for $\|y-y^h\|_Y$, $\|y-y_h\|_Y$, $\|p-p^h\|_Y$ the corresponding upper bounds (4.15)–(4.16) has to be used. □

The Lemmata 4.5, 4.7, and 4.9 give the possibility to estimate the constants ϵ, α, M that are needed to proceed with the results of Section 2.

Theorem 4.10. *Let the constants given by Lemmata* 4.5, 4.7, *and* 4.9 *fulfill the Assumption* 2.4. *Then there exists a local solution* $\bar{u}$ *of the optimal control problem* (4.1)–(4.3), *which satisfies*

$$\|\bar{u}-u_h\|_U \le \alpha\sqrt{1-\frac{2M\epsilon}{\alpha^2}}.$$

Moreover, the second-order sufficient condition holds at $\bar{u}$.

5. An optimal control problem with two local minima

Now, let us apply the technique described above to the following optimal control problem: Minimize

$$\frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{1}{2}\|u-u_1\|^2_{L^2(\Omega)}\|u-u_2\|^2_{L^2(\Omega)} \tag{5.1}$$

subject to the semilinear state equation

$$\begin{aligned} -\Delta y(x) + y(x)^3 &= u(x) \quad \text{in } \Omega\\ y(x) &= 0 \quad \text{on } \Gamma \end{aligned} \tag{5.2}$$

and the control constraints

$$u_a(x) \le u(x) \le u_b(x) \text{ a.e. on } \Omega. \tag{5.3}$$

Let $\Omega \subset \mathbb{R}^d$, $d = 2,3$, be a bounded domain with Lipschitz(?) boundary Γ. Furthermore, functions $y_d, u_a, u_b \in L^2(\Omega)$, $u_a(x) \le u_b(x)$ a.e. on Ω, are given.

At first, let us choose the function spaces. We set $Y := H^1_0(\Omega)$ with $\|y\|_Y := \|\nabla y\|_{L^2(\Omega)}$ and $U = L^2(\Omega)$, $\|u\|_U = \|u\|_{L^2(\Omega)}$. The right-hand side operator B is the embedding operator $L^2(\Omega) \to H^{-1}(\Omega)$. Its norm is bounded as $\|B\|_{L(U,Y')} \le I_2$. Using the notation of the previous section, we define

$$\begin{aligned} d(y) &:= y^3, \\ g(y) &:= \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)}, \\ j(u) &:= \frac{\nu}{2}\|u - u_1\|^2_{L^2(\Omega)}\|u - u_2\|^2_{L^2(\Omega)} \end{aligned}$$

Due to the embedding $H^1(\Omega) \to L^4(\Omega)$, the function d is differentiable from $Y = H^1_0(\Omega)$ to Y'. Let us denote upper bounds of the embedding constants $H^1_0(\Omega) \to L^p(\Omega)$ by I_p, $p < \infty$. They can be computed by eigenvalue estimates for the Laplacian. Furthermore, formulas for I_p are given in [11].

We computed (u_h, y_h, p_h) as the solution of the discretized optimal control problem. The discretization was carried out using $P2$-elements for the state and $P1$-elements for the control.

In Section 4, many constants have to be computed. Let us report, how we computed them for the particular example.

Solution estimates. By monotonicity of the semilinearity $d(y)$ we have,

$$\|S(u_1) - S(u_2)\|_Y \le I_2\|u_1 - u_2\|_U.$$

Since $d'(y)$ is non-negative, it holds,

$$\|(A + d'(y))^{-1}\|_{Y',Y} \le 1 =: c_A.$$

Lipschitz estimates. Let $u \in U$ be taken with $\|u - u_h\|_U \le R$. Then we have $\|y - y^h\|_Y := \|S(u) - S(u_h)\|_Y \le I_2R$. Some of the Lipschitz constants, will depend on R. After an easy computation, one finds

$$\begin{aligned} \|(d'(y) - d'(y_h))z\|_{Y'} &\le I_4^4(\|y_h\|_Y + I_2R)\|y - y_h\|_Y\|z\|_Y \\ &=: c_{d'}(R)\|y - y_h\|_Y\|z\|_Y \\ \|g'(y) - g'(y_h)\|_{Y'} &\le I_2^2\|y - y_h\|_Y \\ \|d''(y) - d''(y_h)\|_{L(Y\times Y,Y')} &\le I_4^4\|y - y_h\|_Y \\ \|g''(y) - g''(y_h)\| &= 0. \end{aligned}$$

The function j coincides with the function f analyzed in Section 3, its derivative was derived in (3.3). Let us now write

$$\begin{aligned} (j''(u) - j''(u_h))(v_1, v_2) &= \left(\|u - u_1\|^2_U + \|u - u_2\|^2_U - \|u_h - u_1\|^2_U - \|u_h - u_2\|^2_U\right)(v_1, v_2) \\ &\quad + 2(u - u_2, v_1)(u - u_1, v_2) + 2(u - u_1, v_1)(u - u_2, v_2) \\ &\quad - 2(u_h - u_2, v_1)(u_h - u_1, v_2) - 2(u_h - u_1, v_1)(u_h - u_2, v_2). \end{aligned}$$

Using the identity

$$\|u - u_1\|^2_U - \|u_h - u_1\|^2_U = (u - u_h, 2(u_h - u_1) + (u - u_h)),$$

we find for $\|u - u_h\|_U \le R$

$$\big|\|u - u_1\|_U^2 - \|u_h - u_1\|_U^2\big| \le (2\|u_h - u_1\|_U + R)\|u - u_h\|_U.$$

Analogously, we get

$$\begin{aligned}(u - u_2, v_i)&(u - u_1, v_j) - (u_h - u_2, v_i)(u_h - u_1, v_j)\\ &= \Big((u - u_2, v_i) - (u_h - u_2, v_i)\Big)(u_h - u_1, v_j) + (u - u_2, v_i)(u - u_h, v_j)\\ &= (u - u_h, v_i)(u_h - u_1, v_j) + (u - u_h + u_h - u_2, v_i)(u - u_h, v_j),\end{aligned}$$

which implies the estimate

$$\begin{aligned}|(u - u_2, v_i)&(u - u_1, v_j) - (u_h - u_2, v_i)(u_h - u_1, v_j)\|\\ &\le (\|u_h - u_1\|_U + \|u_h - u_2\|_U + R)\|u - u_h\|_U\|v_i\|_U\|v_j\|_U.\end{aligned}$$

Hence, we obtain the following value of $c_{j''}$:

$$c_{j''}(R) := 6(\|u_h - u_1\|_U + \|u_h - u_2\|_U + R).$$

Residual estimates. Now, let us explain how we obtained bounds for the residuals in Assumption 4.3. If one could compute a function q such that the inequality

$$(j'(u_h) + B^*p_h + q,\, u - u_h) \ge 0$$

holds for *all* admissible controls $u \in U_{ad}$, then the lower bound in (4.5) is realized by $\epsilon_u = \|q\|_U$. The computation of such a function q is described for instance in [9].

There are quite a few possibilities to estimate the residuals in the state and the adjoint equation. For instance, one can apply standard a-posteriori error estimators of residual type. We used another possibility, as described in [11].

Let $\sigma \in H(div)$ be given, i.e. $\sigma \in L^2(\Omega)$ with $\operatorname{div}(\sigma) \in L^2(\Omega)$. Then we can estimate

$$\begin{aligned}\| - \Delta y_h + d(y_h) - Bu_h\|_{H^{-1}} &\le \| - \Delta y_h - \operatorname{div}(\sigma)\|_{H^{-1}} + \|\operatorname{div}(\sigma) + d(y_h) - Bu_h\|_{H^{-1}}\\ &\le \|\nabla y_h - \sigma\|_{L^2} + I_2\|\operatorname{div}(\sigma) + d(y_h) - Bu_h\|_{L^2}.\end{aligned}$$

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

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

A similar technique was applied to compute the adjoint residual (4.7).

Coercivity check. The lower coercivity bound δ as in Assumption 4.6 was computed as $\delta = \lambda_1(u_h)$, where $\lambda_1(u)$ is defined by (3.6). Since it holds

$$g''(y_h)(z, z) + d''(y_h)(z, z)p_h = \int_\Omega (g''(y_h(x)) + d''(y_h(x))p_h(x))(z(x))^2\mathrm{d}x,$$

we checked the sign of $g''(y_h) + d''(y_h)p_h$. If the sign was positive, δ was chosen as above, and we could use the estimate given by Corollary 4.8. Otherwise, the computation were repeated on a finer mesh.

Computation of r_+. As one can see above, some of the constants depend on the safety radius R. That implies that the constants ϵ, α, M depend on R as well. Let us report, how we computed the value $\tilde{r}_+$, cf. Assumption 2.4. By a bisection method, we computed an interval $[r_1, r_2]$ that contains the smallest positive root of the polynomial

$$p(r) = -\epsilon(r+\rho) + \alpha(r+\rho) - \frac{1}{2}M(r+\rho)r^2$$

with small $\rho = 10^{-5}$. Then $\tilde{r}_+$ was chosen as the right border of the interval, $\tilde{r}_+ = r_2$. That is, all the assumptions are fulfilled for $\tilde{r}_+$ and $R := \tilde{r}_+ + \rho$. If the bisection method was not able to find r such that $p(r)$ was positive, the whole computation was restarted on a finer mesh.

Data. The domain Ω was chosen as $\Omega = (0,3)^2 \setminus [1,2]^2$. Hence, Ω is not convex. This implies that the solution of the elliptic equation does not belong to $H^2(\Omega)$ in general. Thus, the theory as developed in [12] cannot be applied. Furthermore, we took

$$y_d(x_1, x_2) = 0.02 \cdot \sin(\pi x_1)\sin(\pi x_2)$$

and

$$u_1(x) = 0.1, \qquad u_2(x) = 0.4,$$
$$u_a(x) = 0.394, \qquad u_b(x) = 0.099.$$

Solution method and results. The mesh was chosen as a uniform triangulation of the domain with 25.600 triangles, which yields a mesh size of about $h = 0.035$. We solved the discretized optimal control problem by the SQP-method with semi-smooth Newton's method for the inner problems. As initial guesses we used $y_h^0 = 0$ and $p_h^0 =$ for state and adjoint. Starting the SQP-method at $u_h^0 = 0$ yields the solution depicted in Figure 1.

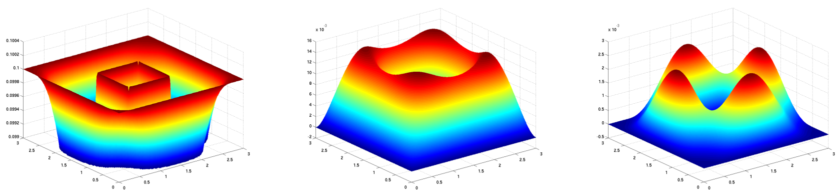

FIGURE 1. First solution: control $\bar{u}_h^1$, state $\bar{y}_h^1$, adjoint state $\bar{p}_h^1$

For a different starting point, we obtained a different solution. Choosing $u_0 = 0.5$ yields the solution triple shown in Figure 2.

The results of the numerical verification technique are as follows. The radius $\tilde{r}_+^1 = 5.773 \cdot 10^{-4}$ satisfies Assumption 2.4 and thus the requirements of Theorem 4.10. Hence, there exists a local solution $\bar{u}^1$ of (5.1)–(5.3) in the neighborhood

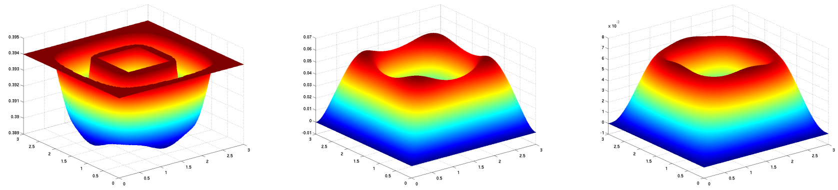

FIGURE 2. Second solution: control $\bar{u}_h^2$, state $\bar{y}_h^2$, adjoint state $\bar{p}_h^2$

of u_h^1 with the error estimate

$$\|\bar{u}^1 - \bar{u}_h^1\|_U \le 5.773 \cdot 10^{-4}.$$

Moreover, the second derivative of the reduced cost functional is positive definite and it holds

$$\phi''(\bar{u}^1)(v,v) \ge 0.7202\|v\|_U^2 \quad \forall v \in U.$$

Similarly, we found the radius $\tilde{r}_+^2 = 2.727 \cdot 10^{-3}$ for the second discrete solution, which gives the existence of a locally optimal control $\bar{u}^2$ with

$$\|\bar{u}^2 - \bar{u}_h^2\|_U \le 2.727 \cdot 10^{-3}$$

and

$$\phi''(\bar{u}^2)(v,v) \ge 0.5991\|v\|_U^2 \quad \forall v \in U.$$

Since $\|u_h^1 - u_h^2\|_U$ is much larger than $\tilde{r}_+^1 + \tilde{r}_+^2$, the controls $\bar{u}^1$ and $\bar{u}^2$ are clearly separated. Consequently, the optimal control problem (5.1)–(5.3) with the data as given above has at least to locally optimal controls.

Convergence rates. We computed solutions for different mesh sizes. The coarsest mesh was obtained by a uniform triangulation with 400 triangles. The meshes were then refined using a grading strategy [2] to cope with the re-entrant corners.

The convergence behaviour of the SQP-method did not change: depending on the initial guess the obtained solutions were either near $\bar{u}^1$ or $\bar{u}^2$. In Table 2 we listed the error bounds r_+^1 and r_+^2.

h	r_+^1	r_+^2
0.28284	$2.7311 \cdot 10^{-3}$	$1.2382 \cdot 10^{-2}$
0.18284	$1.2450 \cdot 10^{-3}$	$5.7057 \cdot 10^{-3}$
0.09913	$5.8972 \cdot 10^{-4}$	$2.7202 \cdot 10^{-3}$
0.05134	$2.8629 \cdot 10^{-4}$	$1.3268 \cdot 10^{-3}$
0.02609	$1.4096 \cdot 10^{-4}$	$6.6074 \cdot 10^{-4}$
0.01315	$6.9948 \cdot 10^{-5}$	$3.3381 \cdot 10^{-4}$

TABLE 2. Error bounds for different meshes

As one can see, the error bounds r_+^1 and r_+^2 decrease like h. For a uniform discretization of the non-convex domain Ω one would expect lower convergence rates. The optimal convergence rate is then recovered using mesh-grading in the vicinity of the re-entrant corners.

References

[1] W. Alt. The Lagrange-Newton method for infinite dimensional optimization problems. *Numerical Functional Analysis and Optimization*, 11:201–224, 1990.

[2] T. Apel, A. Sändig, and J. Whiteman. Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains. *Math. Methods Appl. Sci.*, 19(1):63–85, 1996.

[3] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems.* Springer, New York, 2000.

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

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

[6] P. Deuflhard. *Newton methods for nonlinear problems*, volume 35 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, 2004.

[7] A. L. Dontchev. Local analysis of a Newton-type method based on partial linearization. *AMS Lectures in Appl. Math.*, pages 295–306, 1996.

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

[9] K. Krumbiegel and A. Rösch. A new stopping criterion for iterative solvers for control constrained optimal control problems. *Archives of Control Sciences*, 18(1):17–42, 2008.

[10] J. M. Ortega. The Newton-Kantorovich theorem. *Amer. Math. Monthly*, 75:658–660, 1968.

[11] M. Plum. Computer-assisted enclosure methods for elliptic differential equations. *Linear Algebra Appl.*, 324(1-3):147–187, 2001.

[12] A. Rösch and D. Wachsmuth. Numerical verification of optimality conditions. *SIAM J. Control Optim.*, 2008. to appear.

Daniel Wachsmuth
Johann Radon Institute for Computational and Applied Mathematics (RICAM),
Austrian Academy of Sciences,
Altenbergerstrae 69,
A-4040 Linz,
Austria
e-mail: `daniel.wachsmuth@ricam.oeaw.ac.at`

Arnd Rösch
Universität Duisburg-Essen,
Fachbereich Mathematik,
Forsthausweg 2,
D-47057 Duisburg,
Germany
e-mail: arnd.roesch@uni-due.de