Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 206 (2005) 95–110
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations

Roland Becker [a,*], Boris Vexler [b]

[a] Laboratoire de Mathématiques Appliquées, Université de Pau et des Pays de l'Adour BP 1155, 64013 PAU Cedex, France
[b] Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenberger Straße 69, 4040 Linz, Austria

Received 26 July 2004; received in revised form 24 November 2004; accepted 6 December 2004
Available online 17 March 2005

Abstract

We consider the calibration of parameters in physical models described by partial differential equations. This task is formulated as a constrained optimization problem with a cost functional of least squares type using information obtained from measurements. An important issue in the numerical solution of this type of problem is the control of the errors introduced, first, by discretization of the equations describing the physical model, and second, by measurement errors or other perturbations.

Our strategy is as follows: we suppose that the user defines an interest functional I, which might depend on both the state variable and the parameters and which represents the goal of the computation. First, we propose an a posteriori error estimator which measures the error with respect to this functional. This error estimator is used in an adaptive algorithm to construct economic meshes by local mesh refinement. The proposed estimator requires the solution of an auxiliary linear equation. Second, we address the question of sensitivity. Applying similar techniques as before, we derive quantities which describe the influence of small changes in the measurements on the value of the interest functional. These numbers, which we call relative condition numbers, give additional information on the problem under consideration. They can be computed by means of the solution of the auxiliary problem determined before.

Finally, we demonstrate our approach at hand of a parameter calibration problem for a model flow problem.

Keywords: Parameter estimation; Adaptive mesh refinement; Sensitivity analysis

* Corresponding author. Tel.: +559 40 7508; fax: +559 40 7555.

0021-9991/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2004.12.018

1. Introduction

A physical model described by a system of partial differential equations often involves unknown parameters, which cannot be measured directly, or whose measurement would require too much effort. This situation appears for example in the modeling of material properties or reaction velocities, or in the formulation of boundary conditions. In such situations, the estimation of unknown parameters is indispensable for successful simulation and optimization of the corresponding physical processes. The information required for parameter identification is usually obtained by observations of measurable quantities, like forces, fluxes, point values of pressure, velocity or concentration.

We distinguish two classes of such problems: parameter identification and model calibration problems. If the determination of the values of some unknown parameters is the primary goal of the computation, the problem is called *parameter identification problem*. If one is primarily interested in the computation of different physical quantities (*quantity of interest*) such as drag or lift coefficients, which depend only implicitly on the unknown parameters, we call this problem a *model calibration problem*. This distinction is important for the evaluation of the quality of a simulation. For example, one may think of the case, where the quantity of interest is not very sensitive with respect to some of the unknown parameters. Then, there is probably no need to estimate this parameter with high accuracy. In this paper, we wish to give a rigorous formulation of this idea.

For the formulation and numerical solution of such problems, one has the following two main ingredients: First, one needs measurements, and second, one has to discretize the physical model in order to obtain a finite-dimensional system. Both procedures introduce errors: On the one hand, we have measurement errors and on the other hand discretization errors. Both types of errors lead to inexact computation of the quantity of interest. The aim of this paper is to analyze the dependency of the computed quantity of interest on both, the discretization and the measurement errors. We first derive an a posteriori error estimator, which aims to control the error in the quantity of interest due to discretization. This error estimator is used in an adaptive mesh refinement algorithm, producing economical meshes with respect to the quantity of interest. Next, using similar techniques, we describe the computation of sensitivities of the quantity of interest with respect to the measurements, allowing to analyze the influence of the measurement errors on the quantity of interest. This is important for estimating the quality of the computed approximation and should be helpful for designing new experiments/measurements. We also show that the computation of these sensitivities requires nearly no additional numerical effort, if an adaptive mesh refinement algorithm based on our a posteriori error estimator is used.

In this paper, we consider the problem of parameter calibration formulated as follows: The state variable u, which represents the vector of all physical unknowns, is determined in an appropriate Hilbert space V by a partial differential equation (*state equation*) written in weak form:

$$a(q,u)(\phi) = f(\phi) \quad \forall \phi \in V. \tag{1}$$

Here q denotes the unknown parameters in a Hilbert space Q. The function $a(\cdot,\cdot)(\cdot)$ is defined on the Hilbert space $Q \times V \times V$ and is linear with respect to arguments in the second pair of parenthesis. The partial derivatives of the form $a(\cdot,\cdot)(\cdot)$ are denoted by $a'_u(\cdot,\cdot)(\cdot,\cdot)$, $a'_q(\cdot,\cdot)(\cdot,\cdot)$, etc.

Further, we have an observation operator $C : V \to Z$, which maps the *state variable* u to the Hilbert space of measurements Z. The form a and the observation C are assumed to be three times continuously differentiable. We denote by $\langle\cdot,\cdot\rangle_Z$ the scalar product of Z and by $\|\cdot\|_Z$ the corresponding norm. Similar notation is used for the scalar product and norm in the space Q.

The values of the parameters are estimated from a given set of measurements $\bar{C} \in Z$ using a least squares approach, such that we obtain the constrained minimization problem with cost functional $J : Q \times V \to \mathbb{R}$:

$$\text{Minimize} \quad J(q,u) := \frac{\alpha}{2}\|q-\bar{q}\|_Q^2 + \frac{1}{2}\|C(u)-\bar{C}\|_Z^2 \tag{2}$$

under the constraint (1). The cost functional defined in (2) is the sum of the squared norm of the so-called *least squares residual* defined by

$$R^{\mathrm{LS}}(u) := \bar{C} - C(u), \tag{3}$$

and a regularization term involving prescribed $\alpha \geqslant 0$ and $\bar{q} \in Q$.

The state equation is discretized by the Galerkin method based on a finite-dimensional space $V_h \subset V$. This space is constructed by finite element functions on a mesh $\mathscr{T}_h$. See Section 5 for an example in the context of a two-dimensional flow problem.

The discretized optimization problem for the discrete state $u_h \in V_h$ and parameter $q_h \in Q$ is formulated as follows:

$$\text{Minimize} \quad J(q_h,u_h) \tag{4}$$

under the constraint

$$a(q_h,u_h)(\phi_h) = f(\phi_h) \quad \forall \phi_h \in V_h. \tag{5}$$

The quantity of interest is described by a user-specified interest functional $I : Q \times V \to \mathbb{R}$. The proposed a posteriori error estimator controls the error

$$I(q,u) - I(q_h,u_h)$$

and is used in an adaptive algorithm for successive improvement of the accuracy by an appropriate local mesh refinement, see Section 3 for details.

In order to analyze the dependency on measurements, we will introduce a functional $\hat{i} : Z \to \mathbb{R}$, which maps given measurements $\bar{C}$ to the value of the quantity of interest for the solution of the corresponding problem (1) and (2). The aim of our sensitivity analysis is the computation of *relative condition numbers* κ_i describing the propagation of relative errors from measurements.

The outline of the paper is as follows: In Section 2, we describe a typical optimization loop for the solution of the problem under consideration. In Section 3, we derive our a posteriori error estimator. Section 4 is devoted to sensitivity analysis. Thereafter, in Section 5, we illustrate our approach at hand of a flow problem. In Appendix A, we give the proofs for the propositions formulated in the paper.

For simplicity and clarity of presentation, we make the following assumptions. We suppose that both the control space Q and the measurement space Z are finite dimensional,

$$\dim Q = n_Q, \quad \dim Z = n_Z, \quad n_Z \geqslant n_Q.$$

In addition we suppose that the parameter space of the discrete problem is not reduced, i.e., $Q_h = Q$. The generalization of techniques to the case of infinite dimensional control space Q with $Q_h \subset Q$ is straightforward. Further, we do not incorporate inequality constraints on the parameters. The generalization to this case is the subject of forthcoming work.

The new contributions of this paper are the combination of local mesh refinement with sensitivity analysis and a generalization of a posteriori error estimators established before. Sensitivity analysis of parameter-dependent optimization problems is an active area of research, see in the context of parabolic partial differential equations, e.g., [12], or [9]. Concerning a posteriori error estimation, we generalize previous work; in [2] we have chosen as the interest functional the cost functional itself, and in [6] we have considered interest functionals depending on the parameters only, along with variants of the optimization algorithm. Here, we allow the interest functional to depend both on the parameters and on the state variable.

2. Optimization algorithm

In this section, we reformulate the problem under consideration as an unconstrained optimization problem and shortly discuss optimization algorithms for its solution.

Throughout the paper we assume that problem (1) and (2) admits a (locally) unique solution. Moreover, we assume the existence of a twice continuously differentiable solution operator $S : Q_0 \to V$ in a neighborhood $Q_0 \subset Q$ of the solution to this problem. For all $q \in Q_0$, we have:

$$a(q, S(q))(\phi) = (f, \phi) \quad \forall \phi \in V. \tag{6}$$

The existence of the solution operator S may be ensured by virtue of the implicit function theorem under the usual regularity assumption on $a'_u(q,u)$ in Q_0. Using this solution operator S, we define the reduced observation operator $c : Q_0 \to Z$ by

$$c(q) := C(S(q)) \tag{7}$$

in order to reformulate the problem under consideration as an unconstrained optimization problem with the reduced cost functional $j : Q \to \mathbb{R}$:

$$\text{Minimize} \quad j(q) := \frac{1}{2}\|c(q) - \bar{C}\|_Z^2 + \frac{\alpha}{2}\|q - \bar{q}\|_Q^2, \quad q \in Q. \tag{8}$$

Denoting by $G = c'(q)$, the Jacobian matrix of the reduced observation operator c, and by G^* the adjoint operator of G, the first-order necessary condition for (8) reads:

$$G^* c(q) + \alpha q = G^* \bar{C} + \alpha \bar{q}. \tag{9}$$

In the following proposition, we give a representation of the Jacobian G.

Proposition 1. *Let the reduced observation operator c be defined as in (7). Then its partial derivatives can be computed as follows*:

$$\frac{\partial c_i}{\partial q_j}(q) = G_{ij} = C_i'(u)(w_j), \quad i = 1, \ldots, n_Z, \;\; j = 1, \ldots, n_Q,$$

with $u = S(q)$, C_i and c_i denote the components of the observation and the reduced observation operators, respectively. The tangent solution $w_j \in V$ is determined by

$$a'_u(q,u)(w_j, \phi) = -a'_q(q,u)(e_j, \phi) \quad \forall \phi \in V, \tag{10}$$

where e_j denotes the jth vector of the an orthonormal basis of Q.

Proof. The proof is given in Appendix A. □

In the sequel, we will also need the second derivative of the reduced cost functional. We have

$$\nabla^2 j(q) = \alpha I + G^* G + M, \tag{11}$$

where the matrix $M \in \mathbb{R}^{n_Q \times n_Q}$ is defined by

$$M := -\sum_{i=1}^{n_Z} c_i''(q) R_i^{\mathrm{LS}}. \tag{12}$$

Here, $R_i^{\mathrm{LS}} \in \mathbb{R}$ denotes the ith component of the least-squares residual $R^{\mathrm{LS}}(u)$ with $u = S(q)$.

We collect the necessary information for computation of M in the next proposition.

Proposition 2. *The entries M_{jk} of the matrix M defined in (12) can be computed by*

$$M_{jk} = -a''_{uu}(q,u)(w_j,w_k,z) - a''_{uq}(q,u)(w_k,e_j,z) - a''_{qq}(q,u)(e_j,e_k,z) - a''_{uq}(q,u)(w_j,e_k,z) - \langle C''(u)(w_j,w_k), R^{\mathrm{LS}}(u)\rangle_Z,$$

where $u = S(q)$. Further, $w_j \in V$ is defined in (10) and $z \in V$ is the solution of the following adjoint equation:

$$a'_u(q,u)(\phi,z) = -\langle R^{\mathrm{LS}}(u), C'(u)(\phi)\rangle_Z \quad \forall \phi \in V. \tag{13}$$

Proof. The proof is given in Appendix A. □

The unconstrained optimization problem (8) is solved iteratively. Starting with an initial guess q^0, the next parameter is obtained by $q^{k+1} = q^k + \delta q$, where the update δq is the solution of the problem:

$$H_k \delta q = G_k^* r_k + \alpha(\bar{q} - q^k), \tag{14}$$

where

$$r_k := \overline{C} - c(q^k), \quad G_k := c'(q^k),$$

and H_k is an approximation of the Hessian $\nabla^2 j(q^k)$ of the reduced cost functional j. The choice of the matrix $H_k \in \mathbb{R}^{n_Q \times n_Q}$ leads to different variants of the optimization algorithm. Typical possibilities are $H_k = G_k^* G_k$ leading to the Gauss–Newton algorithm and $H_k = \nabla^2 j(q^k)$, which corresponds to the Newton method. In oder to improve the convergence behavior, one uses in addition step-length rules or trust-region techniques. For different algorithms and convergence theory see, e.g., [8,13,1].

Remark 1. Alternatively to building up the Hessian matrix as described above, one may compute only matrix–vector products to be used within a iterative procedure for solving system (14) (e.g., the conjugate gradient method). For the computation of the matrix–vector product of the Hessian matrix and a given vector, the solution of one tangent and one dual equation is needed.

The optimization algorithm described above on the continuous level, is carried out for the discretized problem (4) and (5). To this end, we introduce the discrete solution operator $S_h : Q_0 \to V_h$ and the discrete reduced observation operator $c_h(q_h) = C(S_h(q_h))$. Similar to the continuous case, the problem is reformulated as an unconstrained optimization problem, i.e.,

$$\text{Minimize} \quad j_h(q_h) = J(q_h, S_h(q_h)).$$

The derivatives of j_h are computed similar to Propositions 1 and 2.

3. A posteriori error estimation

In this section, we derive our a posteriori error estimator for the error with respect to the quantity of interest. Our aim is to prove the following error representation:

$$I(q,u) - I(q_h,u_h) = \eta + R,$$

where η is the a posteriori error estimator, which can in principle be evaluated, and R is a remainder term due to linearization. This error estimator is used within the following adaptive algorithm for error control and mesh refinement: We start on a coarse mesh, solve the discretized optimization problem and evaluate the error estimator. Thereafter, we refine the current mesh using local information obtained from the error estimator, reducing the error with respect to the quantity of interest. This procedure is iterated until the value of the error estimator is below a given tolerance, see [6] for a detailed description of this algorithm.

We define the Lagrange functional $\mathscr{L}$ by

$$\mathscr{L}(q,u,z) = J(q,u) + (f,z) - a(q,u)(z). \tag{15}$$

The first-order necessary condition for problem (1) and (2) is given by the stationarity of the Lagrangian $\mathscr{L}$. Setting for abbreviation $\xi = (q,u,z)$ it reads:

$$\mathscr{L}'(\xi)(\delta\xi) = 0 \quad \forall \delta\xi \in Q \times V \times V. \tag{16}$$

For the discretized problem, we have a similar first order necessary condition for $\xi_h = (q_h,u_h,z_h)$:

$$\mathscr{L}'(\xi_h)(\delta\xi_h) = 0 \quad \forall \delta\xi_h \in Q \times V_h \times V_h. \tag{17}$$

For the quantity of interest I, we introduce an additional functional $\mathscr{M}$:

$$\mathscr{M}(\xi,\chi) = I(q,u) + \mathscr{L}'(\xi)(\chi),$$

with $\chi = (p,v,y) \in Q \times V \times V$. Let now $x = (\xi,\chi)$ be a stationary point of $\mathscr{M}$. Then there holds:

$$I(q,u) = \mathscr{M}(x). \tag{18}$$

This equality, which also holds in similar way on the discrete level, is the starting point for our a posteriori error analysis.

As in [5,6], we obtain the following error representation:

Proposition 3. *Let $x = (\xi,\chi) \in X = (Q \times V \times V)^2$ be a stationary point of $\mathscr{M}$, i.e.,*

$$\mathscr{M}'(x)(\delta x) = 0 \qquad \forall \delta x \in X. \tag{19}$$

Further let $X_h = (Q \times V_h \times V_h)^2 \subset X$ be a subspace and $x_h = (\xi_h,\chi_h) \in X_h$ be the corresponding Galerkin solution satisfying

$$\mathscr{M}'(x_h)(\delta x_h) = 0 \qquad \forall \delta x_h \in X_h. \tag{20}$$

Then, there holds the following error representation:

$$I(q,u) - I(q_h,u_h) = \tfrac{1}{2}\mathscr{M}'(x_h)(x - \tilde{x}_h) + R, \tag{21}$$

where $\tilde{x}_h \in X_h$ is arbitrary and the remainder term R is given by

$$R = \frac{1}{2}\int_0^1 \mathscr{M}'''(x_h + se)(e,e,e)s(s-1)\,\mathrm{d}s, \tag{22}$$

with $e = x - x_h$.

Proof. The proof is given in Appendix A. □

For application of this result, the solution $x_h = (\xi_h, \chi_h)$ of (20) is needed. We note, that $\xi_h = (q_h, u_h, z_h)$ is the already computed solution of the first order optimality condition (17). It remains to compute the set of auxiliary variables $\chi_h = (p_h, v_h, y_h)$. At first glance, it seems as this might lead to huge additional computations. However, this can be avoided exploiting the special structure of the optimality system. In the next proposition, we describe the computation of $\chi = (p, v, y)$ on the continuous level (for clarity of notation). The corresponding discrete set χ_h is obtained in the same way.

Proposition 4. *Let $\xi = (q, u, z)$ be a stationary point of $\mathscr{L}$ defined in (15). Let moreover, $\{w_j\}_{1\leqslant j\leqslant n_Q}$ be the set of tangent solutions (10) and $H = \nabla^2 j(q)$ be the reduced Hessian. Then the auxiliary solution $\chi = (v, y, p)$ is given by*

$$Hp = g, \tag{23}$$

where the components of g are:

$$g_j = -I_q'(q,u)(e_j) - I_u'(q,u)(w_j),$$

$$v = \sum_{j=1}^{n_Q} w_j p_j, \tag{24}$$

and $y \in V$ *is determined by*

$$\begin{aligned} a'_u(q,u)(\phi,y) = {} & \langle C'(u)(v), C'(u)(\phi)\rangle_Z - \langle C''(u)(\phi,v), R^{\mathrm{LS}}(u)\rangle_Z - a''_{uu}(q,u)(\phi,v,z) \\ & - a''_{uq}(q,u)(\phi,p,z) + I'_u(q,u)(\phi) \quad \forall \phi \in V. \end{aligned} \tag{25}$$

Proof. The proof is given in Appendix A. □

From Propositions 3 and 4, we obtain the following result:

Theorem 1. *Let* $\xi = (q, u, z)$ *be a stationary point of the Lagrangian* $\mathscr{L}$ *defined in (15) and* $\xi_h = (q_h, u_h, z_h)$ *be the corresponding discrete solution. Let moreover,* $\chi = (v, p, y)$ *be defined as in Proposition* 4 *and* $\chi_h = (v_h, p_h, y_h)$ *be the corresponding discrete set of auxiliary variables. Then there holds*:

$$I(q,u) - I(q_h,u_h) = \tfrac{1}{2}\{\rho_u(\xi_h)(y - \tilde{y}_h) + \rho_z(\xi_h)(v - \tilde{v}_h)\} \tag{26}$$

$$+ \tfrac{1}{2}\{\rho_v(x_h)(z - \tilde{z}_h) + \rho_y(x_h)(u - \tilde{u}_h)\} + R, \tag{27}$$

where $\tilde{y}_h, \tilde{v}_h, \tilde{z}_h, \tilde{u}_h \in V_h$ *are arbitrary. The residual functionals are given by*

$$\begin{aligned} \rho_u(\xi_h)(\phi) &:= f(\phi) - a(q_h,u_h)(\phi), \\ \rho_z(\xi_h)(\phi) &:= -\langle C'(u_h)(\phi), R^{\mathrm{LS}}(u_h)\rangle_Z - a'_u(q_h,u_h)(\phi,z_h), \\ \rho_v(x_h)(\phi) &:= -a'_q(q_h,u_h)(p_h,\phi) - a'_u(q_h,u_h)(v_h,\phi), \\ \rho_y(x_h)(\phi) &:= I'(q_h,u_h)(\phi) + \langle C'(u_h)(v_h), C'(u_h)(\phi)\rangle_Z - \langle C''(u_h)(\phi,v_h), R^{\mathrm{LS}}(u_h)\rangle_Z \\ &\quad - a''_{uu}(q_h,u_h)(\phi,v_h,z_h) - a''_{uq}(q_h,u_h)(\phi,p_h,z_h) - a'_u(q_h,u_h)(\phi,y_h) \end{aligned} \tag{28}$$

and R *is a cubic remainder term due to linearization, see Proposition* 3.

Proof. The proof is given in Appendix A. □

Remark 2. For the computation of the discrete set of auxiliary variables $\chi_h = (p_h, v_h, y_h)$, we use the representations from Proposition 4, which are immediately translated to the discrete level.

In practice, one would solve the discrete analogon of Eq. (23) with the Hessian matrix computed in the last step of the optimization loop. However, this may introduce an additional remainder term (linearization error) depending on the iteration error.

Remark 3. For practical evaluation of the error estimator, terms like $u - \tilde{u}_h$ have to be approximated. Since $\tilde{u}_h$ is arbitrary, this term corresponds to a local interpolation error. In our numerical example, we use interpolation of the computed bilinear finite element solution y_h on the space of biquadratic finite elements on patches of cells, see [4] for details of this procedure. However, other reasonable procedures are available, see, e.g., [7].

If the Hessian matrix from the last step is used, as discussed in Remark 2, then the main computational cost for the a posteriori error estimator described above is the solution of one auxiliary Eq. (25). However, since one step of the optimization loop requires solution of the state (nonlinear) and of several (linear) tangent equations, the additional work for the a posteriori error estimator is relatively low.

4. Sensitivity analysis

In this section, we develop first-order sensitivity analysis for the quantity of interest. Our aim is to investigate the influence of perturbations in the measurements $\bar{C}$ on the quantity of interest $I(q, u)$. This will be done using similar techniques as in Section 3. Moreover, we will show, that the quantities computed for a posteriori error estimation may be directly used for our sensitivity analysis.

Let $q \in Q$ be a solution of the problem (8). Then, the Hessian matrix $\nabla^2 j(q)$ of the reduced cost functional is positive semidefinite due to the second-order necessary optimality condition. Throughout we assume the Hessian $\nabla^2 j(q)$ to be (strictly) positive definite, which corresponds to the standard second order sufficient optimality condition. Such a solution is called stable.

For addressing the question of the influence of the perturbations in the measurements on the solution of the problem, we have to ensure the existence of the solution for the perturbed problem. This is done in the following proposition.

Proposition 5. *Let q be a stable solution of the problem (8) for the measurement vector $\bar{C}$. Then there exists a neighborhood $Z_0 \subset Z$ of $\bar{C}$ and a continuously differentiable function π: $Z_0 \to Q$, which maps a given measurement vector in Z_0 to a stable solution of the corresponding problem.*

Proof. The proof is given in Appendix A. □

Without loss of generality, we assume, that $\pi(Z_0) \subset Q_0$. For a given measurement vector $\bar{C}$, the state variable at the optimum is given by $S(\pi(\bar{C}))$, where S is the solution operator defined in (6). This allows us to introduce the *reduced quantity of interest* $\widehat{i} : Z_0 \to \mathbb{R}$ as a function of measurements:

$$\widehat{i}(\bar{C}) = I(S(\pi(\bar{C})), \pi(\bar{C})).$$

Next we define the relative condition numbers κ_l describing the amplification of relative errors, uniquely determined by the following formula:

$$\frac{\widehat{i}(\bar{C} + \delta\bar{C}) - \widehat{i}(\bar{C})}{\widehat{i}(q)} = \sum_{l=1}^{n_Z} \kappa_l \frac{\delta\bar{C}_l}{\bar{C}_l} + \mathrm{O}(\|\delta\bar{C}\|_Z^2). \tag{29}$$

Here, we have assumed, that $\widehat{i}(q) \neq 0$ and $\bar{C}_l \neq 0$, $l = 1, 2, \ldots, n_Z$. Otherwise, this formula can be given by means of absolute perturbations.

In the following theorem, we propose an efficient way for the computation of κ_l. Again, we make use of the fundamental equality (18).

Theorem 2. *Let q be a stable solution of Problem (8) for the measurement vector $\bar{C}$. Moreover let $x = (q, u, z, p, v, y)$ be a stationary point of the Lagrangian $\mathscr{M}$. Then, for a small perturbation $\delta\bar{C}$ (29) holds and the relative condition numbers κ_l are given by*

$$\kappa_l = -C_l'(u)(v)\frac{\bar{C}_l}{I(q,u)}, \quad l = 1, 2, \ldots, n_Z.$$

Proof. The proof is given in Appendix A. □

Due to the above theorem, it turns out that the computation of the relative condition numbers κ_l is based on the same auxiliary solution (p, v, y) as the a posteriori error estimation in the previous section. Therefore, the κ_l can be computed with little additional computational effort.

Remark 4. The relative condition numbers κ_l allow also the following representation:

$$\kappa_l = -(Gp)_l \frac{\bar{C}_l}{I(q,u)}, \quad l = 1, 2, \ldots, n_Z,$$

where G is the Jacobian matrix of the reduced observation operator at the solution q.

The relative condition numbers κ_l describe the relative importance of the individual measurements for the determination of the quantity of interest. They may be used to assess the accuracy of the quantity of interest due to measurement errors. Moreover, these numbers give information for the design of new experiments.

Remark 5. On the discrete level similar considerations can be done, which leads to discrete relative condition numbers $\kappa_{h,l}$ given by

$$\kappa_{h,l} = -C_l'(u_h)(v_h)\frac{\bar{C}_l}{I(q_h,u_h)}, \quad l = 1,2,\ldots,n_Z.$$

Remark 6. The extension of this concept to a general parameter-dependent optimal control problem is straightforward. Let problem (1) and (2) depend on a perturbation parameter σ. Then, due to the fact that x is a stationary point of $\mathscr{M}$ there holds:

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\hat{i} = \frac{\partial}{\partial\sigma}\mathscr{M}(x).$$

This gives a possibility to compute the corresponding relative condition numbers.

5. Numerical example

In this section, we discuss numerical results for a model flow problem. We start with a description of our model configuration. Thereafter, we shortly describe the finite element discretization of the Navier–Stokes equations used in the presented computations. We show some numerical results concerning mesh refinement and numerical sensitivity analysis.

5.1. Configuration of the model problem

A typical difficulty in CFD is the prescription of in- and outflow boundary conditions. We consider a systems of pipes Ω, see Fig. 1, with a flow described by the Navier–Stokes equations, where the inflow and outflow boundary conditions are unknown. The circular hole in the lower branch represents the cross-section of a cylinder. The aim of the computation is the accurate prediction of the drag-coefficient of this cylinder.

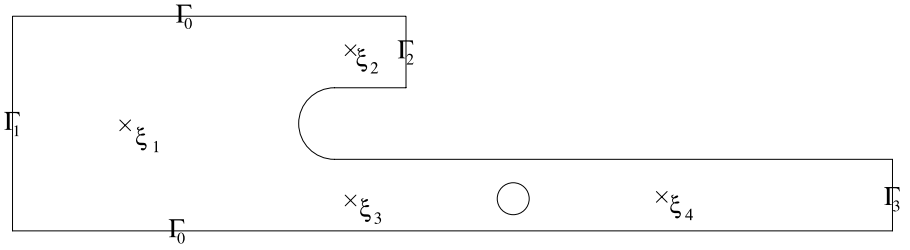

Fig. 1. Configuration of the system of pipes with measurement points marked crosses.

In order to embed this problem in our general setting, we parametrize the unknown boundary condition and obtain the following system of parameter-dependent state equations for the pressure p and velocity v:

$$\begin{aligned} &-\nu\Delta v + v\cdot\nabla v + \nabla p = f \quad \text{in } \Omega,\\ &\nabla\cdot v = 0 \quad \text{in } \Omega,\\ &v = 0 \quad \text{on } \Gamma_0,\\ &\nu\frac{\partial v}{\partial n} - pn = q_1 n \quad \text{on } \Gamma_1,\\ &\nu\frac{\partial v}{\partial n} - pn = q_2 n \quad \text{on } \Gamma_2,\\ &\nu\frac{\partial v}{\partial n} - pn = 0 \quad \text{on } \Gamma_3. \end{aligned} \tag{30}$$

The unknown parameter q in the boundary conditions is searched for in the parameter space $Q = \mathbb{R}^2$ and n denotes the outward unit normal vector to the boundary. This parameterization can be interpreted as follows: The parameters q_1 and q_2 describe the pressure difference between Γ_1 and Γ_3, and between Γ_2 and Γ_3 respectively, cf. Heywood et al. [10].

The solution of the state equation for the exact parameters $q = (0.03, 0.029)$ is shown in Fig. 2.

We assume the measurements $\bar{C} \in Z = \mathbb{R}^4$ to be given by point values of the velocity at four different point marked by crosses in Fig. 1, i.e., the components of the observation operator are given by

$$C_i(u) = v(\xi_i), \qquad i = 1,\dots,4. \tag{31}$$

The values of $\bar{C}_l$, which are used for the identification of the parameter q, are taken from the solution of the state equation for exact parameters, computed on a very fine mesh.

However, in this application, the values of the parameters do not describe the quantity of physical interest. These parameters are only used in order to deal with the problem of incorporating boundary conditions. The quantity we wish to compute is the drag-coefficient on the cylinder Γ_A.

Remark 7. The described problem does not fulfill the assumption that the observation operator C is bounded on the Hilbert space V, if we use for V the standard Sobolev spaces for velocities and pressure, $V = H^1(\Omega)^2 \times L^2(\Omega)$ (with standard modifications to incorporate Dirichlet data). Indeed, the RHS of the adjoint equation (13) is a weighted sum of Dirac measures, and therefore, its solution has the same singularities as the fundamental solution of the Stokes operator [11]. One possibility to obtain a well-posed formulation uses the Banach spaces $V = W^{1,p}(\Omega)^2 \times L^p(\Omega)$ for u and $V' = W^{1,p'}(\Omega)^2 \times L^{p'}(\Omega)$ for the

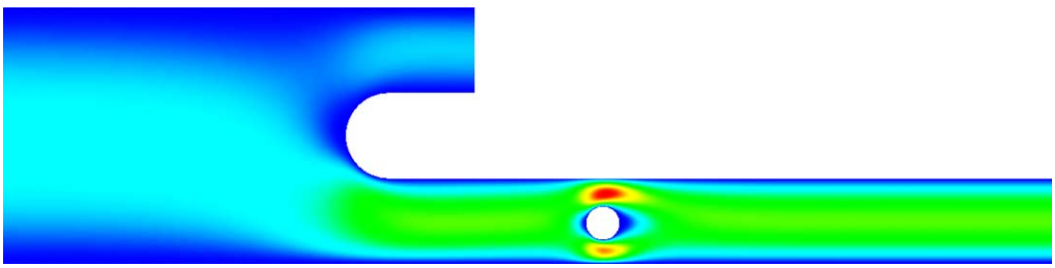

Fig. 2. Solution of the state equation (horizontal velocity) for the exact parameters $q = (0.03, 0.029)$.

Lagrange multiplier z, with $p > d$ and $\frac{1}{p} + \frac{1}{p'} = 1$. The finite element formulation remains unchanged. The developed theory for a posterior error estimation can be easily modified accordingly. The a priori error analysis of an elliptic parameter identification problems with pointwise measurements can be found in [14].

This quantity of physical interest is given by the functional I:

$$I(u) = c_0 \int_{\Gamma_A} n \cdot \sigma \cdot d \, \mathrm{d}s, \tag{32}$$

where $d = (1,0)$ is a chosen direction, c_0 is a given constant, and σ denotes the stress tensor given as usual by

$$\sigma = \frac{\nu}{2}(\nabla v + (\nabla v)^{\mathrm{T}}) - pI.$$

5.2. Discretization of the Navier–Stokes equations

The starting point for any finite element discretization of the Navier–Stokes equation (30) is the standard variational formulation. The space of test function for the velocities is

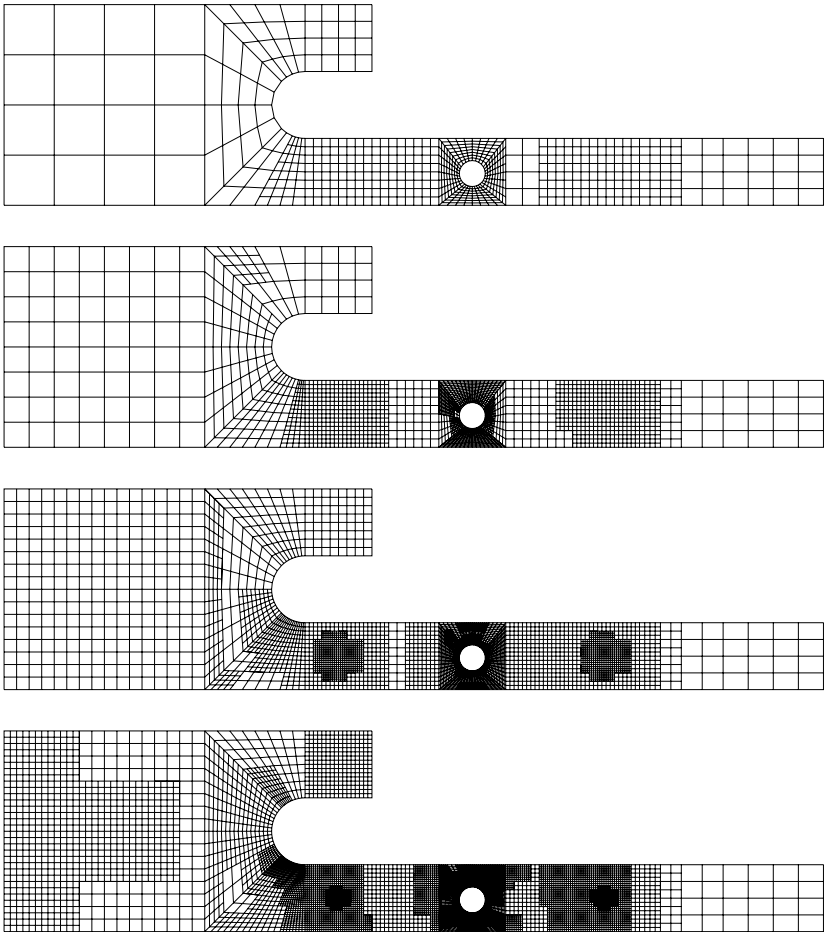

Fig. 3. Meshes generated by the adaptive algorithm with 746, 2170, 5094 and 11,068 nodes.

$$H = \{\psi \in H^1(\Omega)^2 | \psi = 0 \text{ on } \Gamma_0\}.$$

We set $V = H \times L^2(\Omega)$ and the form $a : Q \times V \times V \to \mathbb{R}$ is defined by

$$a(q,u)(\phi) = \nu(\nabla v, \nabla\psi) + (v \cdot \nabla v, \psi) - (p, \nabla \cdot \psi) + (\nabla \cdot v, \xi) - (q_1 n, \psi)_{\Gamma_1} - (q_2 n, \psi)_{\Gamma_2}, \tag{33}$$

where $\phi = (\psi, \xi) \in V$ denotes the test functions for velocity and pressure. The corresponding weak formulation of the state equation reads: Find $u = (v, p) \in V$ such that

$$a(q,u)(\phi) = 0 \quad \forall \phi \in V. \tag{34}$$

The state equation (34) is discretized using conforming finite elements on shape-regular quadrilateral meshes $\mathcal{T}_h$. However, in order to ease local mesh refinement we allow a cell to have nodes, which lie on midpoints of faces of neighboring cells. But at most one such *hanging node* is permitted for each face. We use isoparametric bilinear finite elements for both pressure and velocities. We add further terms to the semilinear form a (33) in order to obtain a stable formulation with respect to both the pressure-velocity coupling and convection dominated flow. The discretization is described in detail in [2].

5.3. Computational results

The optimization problem is solved by the Gauss–Newton method with the initial guess $q_0 = (0, 0)$, which corresponds to the state variable $u = 0$. The resulting nonlinear state equations are solved by Newton method and the solution of the linear subproblems are computed using a multigrid algorithm on locally refined meshes, see Becker and Braack [3]. With these ingredients, the total numerical cost for solution on a given mesh behaves like $O(N)$, where N is the number of nodes. All computations are done on the basis of the package *RoDoBo* for treating optimization problems governed by partial differential equations and the finite element toolkit *Gascoigne3D*.

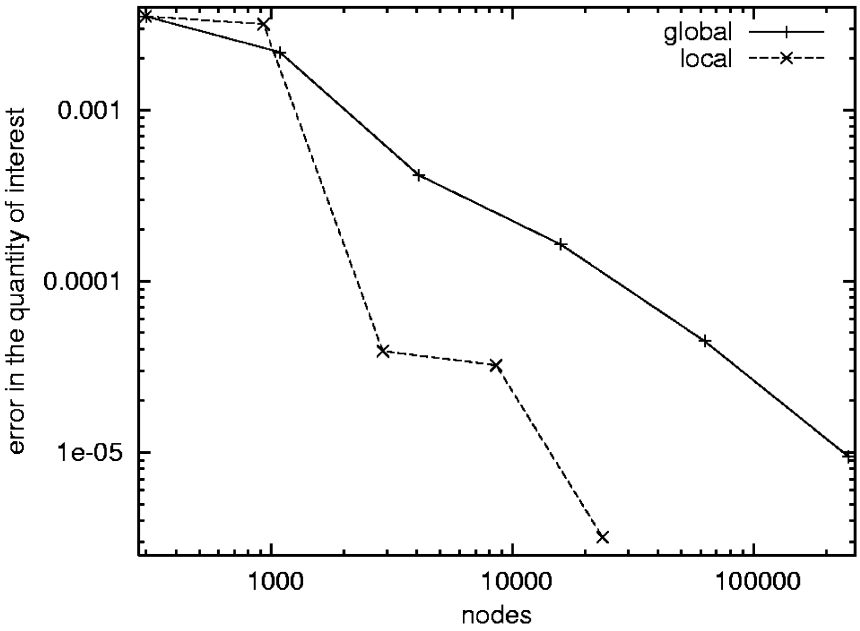

Fig. 4. Error in the quantity of interest I vs. number of mesh points for uniform mesh refinement and local refinement resulting from our a posteriori error estimator.

Table 1
Relative condition numbers for point measurements on sequence of locally refined meshes

N	κ_1	κ_2	κ_3	κ_4
300	6.778E − 2	−1.640E − 2	5.032E − 1	5.110E − 1
746	6.920E − 2	−1.582E − 2	5.264E − 1	5.329E − 1
2170	6.908E − 2	−1.638E − 2	5.322E − 1	5.404E − 1
11,068	6.931E − 2	−1.661E − 2	5.346E − 1	5.429E − 1

Table 2
Comparison of the computed and predicted relative error in the quantity of interest for different levels of measurement errors

$\frac{\delta\bar{C}_1}{\bar{C}_1}\cdot 100\%$	$\frac{\delta I}{I}\cdot 100\%$	$\kappa_1\frac{\delta\bar{C}_1}{\bar{C}_1}\cdot 100\%$
10	0.6904	0.6894
20	1.3824	1.3788
50	3.4682	3.4470
$\frac{\delta\bar{C}_2}{\bar{C}_2}\cdot 100\%$		$\kappa_2\frac{\delta\bar{C}_2}{\bar{C}_2}\cdot 100\%$
10	−0.1657	−0.1657
20	−0.3315	−0.3314
50	−0.8288	−0.8285
$\frac{\delta\bar{C}_3}{\bar{C}_3}\cdot 100\%$		$\kappa_3\frac{\delta\bar{C}_3}{\bar{C}_3}\cdot 100\%$
10	5.376	5.336
20	10.831	10.672
50	27.724	26.680
$\frac{\delta\bar{C}_4}{\bar{C}_4}\cdot 100\%$		$\kappa_4\frac{\delta\bar{C}_4}{\bar{C}_4}\cdot 100\%$
10	5.460	5.415
20	11.007	10.830
50	28.185	27.075

Application of the a posteriori error estimator for the quantity of interest (32) leads to the sequence of locally refined meshes shown Fig. 3. Obviously, there is strong local mesh refinement at the points of measurement and around the cylinder where the quantity of interest is defined. Note the different strength of refinement at the points ξ_i. They are explained by the automatic weighting built in the a posteriori error estimator.

Next we investigate the quality of this sequence of meshes generated by our a posteriori error estimator. In Fig. 4, we show the corresponding error in the quantity of interest (32) compared with the errors obtained by uniform mesh refinement. It turns out, that the refinement strategy based on the error estimator for the quantity of interest leads to very efficient meshes.

We compute the relative condition numbers κ_l for the four points measurement on the above sequence of locally refined meshes. The results are listed in Table 1. The measurements at ξ_3 and ξ_4 clearly have more influence on I than the others.

It turns out that a perturbation of the measurement $\bar{C}_1$ of 10% would lead only to a perturbation of about 0.69% in the quantity of interest. However, a 10% perturbation of the measurement $\bar{C}_4$ introduces an error in the quantity of interest of about 5.4%.

Finally, we investigate the quality of κ_l for the prediction in the relative change of I, i.e., we check the validity of (29). In Table 2, we show the relative error $\delta I/I$ in the quantity of interest for different error levels $\delta\bar{C}_l/\bar{C}_l$ in the measurements. In order to do this comparison, since $\delta I = \widehat{i}(\bar{C}+\delta\bar{C}) - \widehat{i}(\bar{C})$, the solution corresponding to $\bar{C}+\delta\bar{C}$ is computed in addition. The quality of the prediction based on κ_l is very satisfactory.

Appendix A. In this section, we give the proofs of the propositions and theorems formulated above.

Proof of Proposition 1. Let $u = S(q)$ be a solution of the state equation (1). Taking the derivative of Eq. (6) with respect to q in the direction δq, we obtain

$$a'_u(q,u)(\delta u,\phi) + a'_q(q,u)(\delta q,\phi) = 0 \quad \forall \phi \in V, \tag{A.1}$$

where $\delta u = S'(q)(\delta q)$. Moreover, there holds:

$$c'(q)(\delta q) = C'(u)(\delta u),$$

and we complete the proof by setting $\delta q = e_j$. □

Proof of Proposition 2. Let $u = S(q)$ be a solution of the state equation (1). Using the Lagrange functional $\mathscr{L}$ defined in (15) we obtain:

$$j(q) = J(q,u) = \mathscr{L}(q,u,z)$$

for arbitrary $z \in V$. Taking the derivative with respect to q in the direction δq, we obtain:

$$j'(q)(\delta q) = \mathscr{L}'_u(q,u,z)(\delta u) + \mathscr{L}'_q(q,u,z)(\delta q), \tag{A.2}$$

where $\delta u = S'(q)(\delta q)$. Let now $z \in V$ be a solution of the adjoint equation (13), which corresponds to the condition:

$$\mathscr{L}'_u(q,u,z)(\phi) = 0 \quad \forall \phi \in V. \tag{A.3}$$

We take the derivative of (A.2) with respect to q in the direction τq and obtain:

$$\begin{aligned} j''(q)(\delta q,\tau q) = {} & \mathscr{L}'_u(q,u,z)(\delta^2 u) + \mathscr{L}''_{uu}(q,u,z)(\delta u,\tau u) + \mathscr{L}''_{uq}(q,u,z)(\delta u,\tau q) + \mathscr{L}''_{uz}(q,u,z)(\delta u,\tau z) \\ & + \mathscr{L}''_{qu}(q,u,z)(\delta q,\tau u) + \mathscr{L}''_{qq}(q,u,z)(\delta q,\tau q) + \mathscr{L}''_{qz}(q,u,z)(\delta q,\tau z), \end{aligned}$$

where $\delta^2 u = S''(q)(\delta q,\tau q)$, $\tau u = S'(q)(\tau q)$ and $\tau z \in V$ is the derivatives of z with respect to q in the direction τq. The first term vanishes due to (A.3) and moreover there holds:

$$\mathscr{L}''_{uz}(q,u,z)(\delta u,\tau z) + \mathscr{L}''_{qz}(q,u,z)(\delta q,\tau z) = a'_u(q,u)(\delta u,\tau z) + a'_q(q,u)(\delta q,\tau z) = 0,$$

due to (A.1). We complete the proof by setting $\delta q = e_j, \tau q = e_k$ and calculating the second derivatives of $\mathscr{L}$. □

Proof of Proposition 3. We note, that $\xi = (q, u, z)$ is a stationary point of $\mathscr{L}$, i.e.

$$\mathscr{L}'(\xi)(\delta\xi) = 0 \qquad \forall \delta\xi \in Q \times V \times V \tag{A.4}$$

and ξ_h is the corresponding Galerkin solution

$$\mathscr{L}'(\xi_h)(\delta\xi_h) = 0 \qquad \forall \delta\xi_h \in Q \times V_h \times V_h. \tag{A.5}$$

Therefore, we obtain:

$$I(q,u) - I(q_h,u_h) = \mathscr{M}(x) - \mathscr{M}(x_h). \tag{A.6}$$

We rewrite the RHS of (A.6) as follows:

$$\mathscr{M}(x) - \mathscr{M}(x_h) = \int_0^1 \mathscr{M}'(x_h + se)(e)\,\mathrm{d}s, \tag{A.7}$$

approximate the integral by the trapezoidal rule and obtain:

$$\mathscr{M}(x) - \mathscr{M}(x_h) = \frac{1}{2}\mathscr{M}'(x)(e) + \frac{1}{2}\mathscr{M}'(x_h)(e) + R, \tag{A.8}$$

where the remainder term R is given by

$$R = \frac{1}{2}\int_0^1 \mathscr{M}'''(x_h + se)(e,e,e)s(s-1)\,\mathrm{d}s.$$

The term $\mathscr{M}'(x)(e)$ vanishes, and due to Galerkin orthogonality the term $\mathscr{M}'(x_h)(e)$ can be replaced by $\mathscr{M}'(x_h)(x - \tilde{x}_h)$ with $\tilde{x}_h \in X_h$ arbitrarily chosen. This completes the proof. □

Proof of Proposition 4. We note, that $\xi = (q,u,z)$ is a stationary point of $\mathscr{L}$. Therefore, there holds:

$$\mathscr{M}'_{\chi}(\xi,\chi)(\delta\chi) = \mathscr{L}'(\xi)(\delta\chi) = 0.$$

It remains to show that $\mathscr{M}'_{\xi}(\xi,\chi) = 0$. Due to the definition of v (24) there holds:

$$\mathscr{M}'_z(\xi,\chi)(\phi) = \mathscr{L}''_{uz}(\xi)(v,\phi) + \mathscr{L}''_{qz}(\xi)(p,\phi) = \sum\nolimits_{j=1}^{n_Q}\Big(a'_u(q,u)(w_j,\phi) + a'_q(q,u)(e_j,\phi)\Big).$$

This sum vanishes because of the definition of w_j (10). The equation for y (25) can be rewritten in the following form:

$$\mathscr{L}''_{zu}(\xi)(y,\phi) = -I'_u(q,u)(\phi) - \mathscr{L}''_{uu}(\xi)(v,\phi) - \mathscr{L}''_{qu}(\xi)(p,\phi) \quad \forall\phi \in V, \tag{A.9}$$

which is equivalent to $\mathscr{M}'_u(\xi,\chi) = 0$. Finally, we show, that the derivative $\mathscr{M}'_q(\xi,\chi)$ vanishes. There holds:

$$\mathscr{M}'_q(\xi,\chi)(\delta q) = I'_q(q,u)(\delta q) + \mathscr{L}''_{uq}(\xi)(v,\delta q) + \mathscr{L}''_{qq}(\xi)(p,\delta q) + \mathscr{L}''_{zq}(\xi)(y,\delta q).$$

Using the representation of second derivatives of $j(q)$ from Proposition 2 and setting $\delta q = e_j$ we have:

$$\mathscr{M}'_q(\xi,\chi)(e_j) = \langle Hp, e_j\rangle_Q + I'_q(q,u)(e_j) - \mathscr{L}''_{zq}(\xi)(y,e_j) - \mathscr{L}''_{uu}(\xi)(v,w_j) - \mathscr{L}''_{qu}(\xi)(p,w_j). \tag{A.10}$$

Due to the definition of w_j and the Eq. (A.9) we obtain:

$$\mathscr{M}'_q(\xi,\chi)(e_j) = \langle Hp, e_j\rangle_Q + I'_q(q,u)(e_j) + I'_u(q,u)(w_j).$$

We complete the proof using the definition of p (23). □

Proof of Proposition 5. The solution q of problem (8) is assumed to be stable. Therefore the symmetric matrix $\nabla^2 j(q)$ is positive definite and consequently invertible. Due to the finite dimension of Q the inverse of $\nabla^2 j(q)$ is bounded and the implicit function theorem can be applied to the optimality condition (9). This completes the proof. □

Proof of Theorem 1. We apply Propositions 3 and 4. Due to the choice $Q_h = Q$ we may set $\tilde{x}_h = (\tilde{u}_h, q, z_h, \tilde{v}_h, p, \tilde{y}_h)$. This completes the proof. □

Proof of Theorem 2. There holds:

$$\hat{i}(\bar{C}) = \mathscr{M}(x).$$

We take derivatives in respect to $\bar{C}$ and obtain:

$$\hat{i}'(\bar{C})(\delta\bar{C}) = \mathscr{M}'_{\bar{C}}(x)(\delta\bar{C}) + \mathscr{M}'_x(\delta x),$$

where δx is the derivative of x with respect to $\bar{C}$ in the direction $\delta\bar{C}$. Due to the fact that x is a stationary point of $\mathscr{M}$ there holds:

$$\hat{i}'(\bar{C})(\delta\bar{C}) = \mathscr{M}'_{\bar{C}}(x)(\delta\bar{C}) = -\langle C_I'(u)(v), \delta\bar{C}\rangle_Z.$$

This completes the proof. □

References

[1] N. Gould, A.R. Conn, Ph.L. Toint, Trust-region Methods, SIAM, 2000.
[2] R. Becker, Mesh adaptation for stationary flow control, J. Math. Fluid Mech. 3 (4) (2001) 317–341.
[3] R. Becker, M. Braack, Multigrid techniques for finite elements on locally refined meshes, Numer. Linear Algebra Appl. (2000).
[4] R. Becker, R. Rannacher, A feed-back approach to error control in finite element methods: Basic analysis and examples, East-West J. Numer. Math. 4 (1996) 237–264.
[5] R. Becker, R. Rannacher, An optimal control approach to a-posteriori error estimation, in: A. Iserles (Ed.), Acta Numerica 2001, Cambridege University Press, 2001, pp. 1–102.
[6] R. Becker, B. Vexler, A posteriori error estimation for finite element discretizations of parameter identification problems, Numer. Math. 96 (3) (2004) 435–459.
[7] M. Braack, T. Richter. Solutions of 3D Navier–Stokes benchmark problems with adaptive finite elements. Comput. Fluids (accepted).
[8] J.E. Dennis, R.B. Schnabel, Numerical methods for unconstrained optimization and nonlinear equations, Number 16 in Classics in Applied Mathematics. SIAM, Society for Industrial and Applied Mathematics, 1996.
[9] R. Griesse, Parametric sensitivity analysis in optimal control of a reaction-diffusion system—part I solution differentiability, Numer. Funct. Anal. Optim. 25 (1–2) (2004) 93–117.
[10] J.G. Heywood, R. Rannacher, S. Turek, Artificial boundaries and flux and pressure conditions for the incompressible Navier–Stokes equations, Int. J. Numer. Math. Fluids 22 (1992) 325–352.
[11] O.A. Ladyzhenskaya, The Mathematical Theory of Viscous Incompressible Flow, Gordon and Breach Science Publishers, 1969.
[12] K. Malanowski, Sensitivity analysis for parametric optimal control of semilinear parabolic equations, J. Convex Anal. 9 (2) (2002) 543–561.
[13] J. Nocedal, S.J. Wright, Numerical Optimization. Springer Series in Operations Research, Springer, New York, 1999.
[14] R. Rannacher, B. Vexler. A priori error estimates for the finite element discretization of elliptic parameter identification problems with pointwise measurements. SIAM J. Contr. Optim. (accepted).