



HAL
open science

Mesh Refinement and Numerical Sensitivity Analysis for Parameter Calibration of Partial Differential Equations

Roland Becker, Boris Vexler

► **To cite this version:**

Roland Becker, Boris Vexler. Mesh Refinement and Numerical Sensitivity Analysis for Parameter Calibration of Partial Differential Equations. *Journal of Computational Physics*, 2005, 206 (1), pp.95-110. hal-00158887

HAL Id: hal-00158887

<https://hal.science/hal-00158887v1>

Submitted on 1 Jul 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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*Institut für Angewandte Mathematik, Im Neuenheimer Feld 294, 69120
Heidelberg, Germany*

Abstract

We consider the calibration of parameters in physical models described by partial differential equations. This task is formulated in standard way as a constrained optimization problem with a cost functional of least squares type. The unknown parameters are calibrated using information obtained by 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: First, 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. 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 apply similar techniques as before to address the question of sensitivity. For this, 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.

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

Key words: Parameter estimation, adaptive mesh refinement, sensitivity analysis

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 only implicitly depend 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 a minimization problem, one has the following two main ingredients: First, one needs measurements, and second, one has to discretize the state equation 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.

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 weak form:

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

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 argument 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 $C : V \rightarrow Z$, which maps the *state variable* u to the Hilbert space of measurements Z . We denote by $(\cdot, \cdot)_Z$ 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 $J : Q \times V \rightarrow \mathbb{R}$:

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

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

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

and a regularization term involving prescribed $\alpha \geq 0$ 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 \mathcal{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:

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

under the constraint

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

The quantity of interest is described by a user-specified interest functional $I : Q \times V \rightarrow \mathbb{R}$. 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 $\hat{i} : Z \rightarrow \mathbb{R}$, which maps given measurements \bar{C} to the value of the quantity of interest for the solution of the corresponding problem (1,2). The aim of our sensitivity analysis is the computation of *relative condition numbers* κ_i describing the propagation of relative errors.

The outline of the paper is as follows: In the next section, 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 the last section we give the proofs for the propositions formulated in the paper.

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

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

In addition we suppose that the parameter space of the discrete problem is not reduced, $Q_h = Q$. Further, we do not incorporate constraints on the parameters. Our considerations are done such that the main results carry over to these generalizations.

The new contributions of this article 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 is an active area of research, see in the context of parabolic partial differential equations e.g. Malanowski [13], Tröltzsch [?] or Griesse [9]. Concerning a posteriori error estimation we generalize previous work; in [1] we have chosen as the interest functional the cost functional itself, and in [6] we have considered interest functionals depending on the parameters along with variants of the optimization algorithm. Here, we allow the interest functionals to depending on the parameters and on the state variable are allowed.

2 Optimization algorithm

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

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

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

Using this solution operator S we define the reduced observation operator $c : Q_0 \rightarrow Z$ by:

$$c(q) := C(S(q)) \quad (7)$$

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

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

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

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

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:*

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

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:

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

where e_j denotes the j -th vector of the standard orthonormal basis of Q .

Proof: The proof is given in Section 6.

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

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

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

$$M := - \sum_{i=1}^{n_Z} c''_i(q) R_i^{LS}. \quad (12)$$

Here, R_i^{LS} denotes the i -th component of the least-squares residual $R^{LS}(u)$ 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:*

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

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:

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

Proof: The proof is given in Section 6.

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:

$$H_k \delta q = G_k^* r_k, \quad (14)$$

where

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

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. For convergence theory of these methods see, e.g., Dennis & Schnabel [8] or Nocedal & Wright [14]. In our practical realization we use trust-region techniques for improving global convergence, see e.g. Conn, Gould & Toint [?].

This optimization algorithm, described above on the continuous level, is carried out for the discretized problem (4,5). To this end, we introduce the discrete solution operator $S_h : Q_0 \rightarrow Z$ 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.

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

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 our current mesh to a new one guided by the error estimator, which allows us to reduce the error with respect to the quantity of interest in an efficient way. This procedure is iterated until the value of the error estimator is below a given tolerance, see [?] for a detailed description of this algorithm.

Our approach relies on the techniques for functional-oriented a posteriori error estimation described in Becker & Rannacher [5]. However, the generalized framework developed here requires some substantial additional work. The same remark is true with respect to other approaches to functional-oriented a posteriori error estimation as presented in Oden & Prudhomme [15] or Machiels, Patera & Peraire [12].

We introduce the standard Lagrange \mathcal{L} functional by:

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

The first-order necessary conditions for the problem (1,2) are given by the stationarity of the Lagrangian \mathcal{L} . Denoting $\xi = (q, u, z)$ this condition reads:

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

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

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

For the quantity of interest I , we introduce an additional Lagrangian \mathcal{M} :

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

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

$$I(q, u) = M(x). \quad (18)$$

This equality, which holds similar also on the discrete level, is the starting point for our a posteriori error analysis. Moreover, the same equality will be used for the sensitivity analysis in Section 4.

As in [5?], 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 \mathcal{M} , i.e.*

$$\mathcal{M}'(x)(\delta x) = 0 \quad \forall \delta x \in X. \quad (19)$$

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

$$\mathcal{M}'(x_h)(\delta x_h) = 0 \quad \forall \delta x_h \in X_h. \quad (20)$$

Then, there holds the following error representation

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

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

$$R = \frac{1}{2} \int_0^1 \mathcal{M}'''(x_h + se)(e, e, e)s(s-1)ds, \quad (22)$$

with $e = x - x_h$.

Proof: The proof is given in Section 6.

For the application of this result the solution $x_h = (\xi_h, \chi_h)$ is needed. We note, that $\xi_h = (q_h, u_h, z_h)$ is a solution of the first order optimality condition (17). It remains to compute the set of auxiliary variables $\chi_h = (v_h, p_h, y_h)$. At the first glance, it seems as if it leads to a big coupled system to be solved. However, this can be avoided using the information available from the last step of the optimization loop. In the next proposition we describe the computation of $\chi = (v, p, y)$ 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 \mathcal{L} (15). Let moreover, $\{w_j\}_{1 \leq j \leq n_Q}$ be the set of tangent solutions (10) and $H = \nabla^2 j(q)$ be the reduced Hessian, both computed in the last step of the optimization algorithm. Then the auxiliary solution $\chi = (v, y, p)$ is given by:*

$$Hp = g, \quad (23)$$

where the components of g are:

$$g_j = -I'_q(q, u)(e_j) - I'_u(q, u)(w_j),$$

$$v = \sum_{j=1}^{n_Q} w_j p_j, \quad (24)$$

and $y \in V$ is determined by:

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

Proof: The proof is given in Section 6.

From Propositions 3 and 4 we obtain the following result:

Theorem 1 *Let $\xi = (q, u, z)$ be a stationary point of the Lagrangian \mathcal{L} (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:*

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

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

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

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

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

Proof: The proof is given in Section 6.

Remark 1 *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.*

The main computational cost for the a posteriori error estimator described above is the solution of one auxiliary equation (25). This is cheap, even in

comparison with only one step of the optimization loop, which includes solution of the state (nonlinear) and of the several (linear) tangent equations.

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 $\pi : Z_0 \rightarrow Q$, which maps a given measurement vector in Z_0 to a stable solution of the corresponding problem.*

Proof: The result relies on the implicit function theorem.

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* $\hat{i} : Z_0 \rightarrow \mathbb{R}$ as a function of measurements:

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

For derivation of the sensitivities we again make use of equality (18). This leads to the following result:

Theorem 2 *Let q be a stable solution of Problem 8 for the measurement vector \bar{C} . Moreover let $x = (q, u, z, v, p, y)$ be a stationary point of the Lagrangian M . Then, for a small perturbation $\delta\bar{C}$ of measurements the following error*

propagation holds:

$$\frac{\delta \hat{i}}{\hat{i}(q)} = \sum_{l=1}^{n_Z} \kappa_l \frac{\delta \bar{C}_l}{\bar{C}_l} + O(\|\delta \bar{C}\|_Z^2),$$

where

$$\delta \hat{i} = \hat{i}(\bar{C} + \delta \bar{C}) - \hat{i}(\bar{C})$$

and the relative condition numbers κ_l are given by:

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

Proof: The proof is given in Section 6.

The above theorem turns out, that the computation of the relative condition numbers κ_l is based on the same auxiliary solution (v, p, y) as the a posteriori error estimation in the previous section. Therefore, they can be computed without very few additional computational effort.

Remark 2 *The relative condition numbers κ_l allow also the following representation:*

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

where G is the jacobian matrix of the reduced observation operator computed in the last step of the optimization algorithm.

The relative condition numbers κ_l describe the relative importance of the measurements for the quantity of interest. They may be used for the assessment of the error in the quantity of interest due to measurement errors. Moreover, they give information, which should be helpful by design of new experiments. The application of this concepts for automatic experiment design is a subject of future work.

Remark 3 *On the discrete level similar considerations can be done, which leads to discrete relative condition numbers $\kappa_{h,l}$ given by*

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

Remark 4 *The extension of this concept to a general parameter-depended optimal control problem is straightforward. Let the problem depend on a parameter σ . Then, due to the fact that x is a stationary point of \mathcal{M} there holds:*

$$\frac{\partial}{\partial \sigma} \hat{i} = \frac{\partial}{\partial \sigma} \mathcal{M}(x).$$

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 discuss finite element discretization of Navier-Stokes equations in our context and present some computational results illustrating our approach.

5.1 Configuration of the model problem

A typical difficulty in a broad class of CFD problems is the modeling of boundary conditions. We consider a systems of pipes Ω , see Figure 1, with a flow described by Navier-Stokes equation, where the precise inflow and outflow boundary conditions are unknown. The circular hole in the lower branch represents the cross-section of the cylinder and the aim of the computation is a precise approximation of the drag-coefficient on this cylinder.

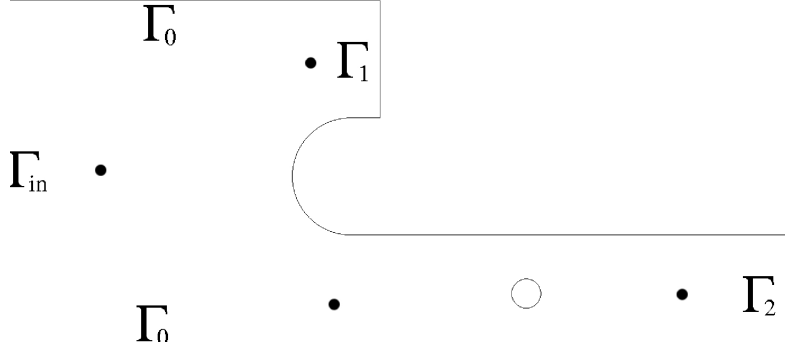


Fig. 1. Configuration of the system of pipes with measurement points marked by black circles

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

$$\begin{aligned}
 -\nu\Delta v + v \cdot \nabla v + \nabla p &= f & \text{in } \Omega, \\
 \nabla \cdot v &= 0 & \text{in } \Omega, \\
 v &= 0 & \text{on } \Gamma_0, \\
 \nu \frac{\partial v}{\partial n} - pn &= q_1 n & \text{on } \Gamma_{in}, \\
 \nu \frac{\partial v}{\partial n} - pn &= q_2 n & \text{on } \Gamma_1, \\
 \nu \frac{\partial v}{\partial n} - pn &= 0 & \text{on } \Gamma_2.
 \end{aligned} \tag{29}$$

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_{in} and Γ_2 , and between Γ_1 and Γ_2 respectively, cf. Heywood, Rannacher & Turek [11].

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 black circles in Figure 1. The solution of the state equation for the exact parameters $q = (0.03, 0.029)$ is shown in Figure 2. However, the values of the parameters do not describe the quantity of physical interest in this application. The quantity we wish to compute is the drag-coefficient on the cylinder with the boundary denoted by Γ_A .

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

This quantity of interest is given by the functional I :

$$I(u) = c_0 \int_{\Gamma_A} n \cdot \sigma \cdot d ds, \quad (30)$$

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

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

5.2 Discretization of the Navier-Stokes equations

The starting point for any finite element discretization of the Navier-Stokes equations (29) is a variational formulation. The space of test function for the velocity is

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

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

$$a(q, u)(\phi) = \nu(\nabla v, \nabla \psi) + (v \cdot \nabla v, \psi) - (p, \nabla \cdot \psi) + (\nabla \cdot v, \xi) - (q_1 n, \psi)_{\Gamma_{in}} - (q_2 n, \psi)_{\Gamma_1}, \quad (31)$$

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

$$a(q, u)(\phi) = 0 \quad \forall \phi \in V. \quad (32)$$

The state equation (32) is discretized using conforming finite elements on shape-regular quadrilateral meshes \mathcal{T}_h , see e.g. Ciarlet [7]. 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 finite elements of equal order for each component leading to the following finite element space:

$$V_h = \{\phi_h = (\psi_h, \xi_h) \in V \mid \phi_h|_K \in \tilde{Q}_1(K)^3, K \in \mathcal{T}_h\},$$

where $\tilde{Q}_1(K)$ consists of shape functions obtained via a bilinear transformation from the space of bilinear functions $Q_1(\hat{K})$ on the reference cell $\hat{K} = (0, 1)^2$. We add stabilization terms to the semilinear form a (31) in order to obtain a stable formulation. The discretization is described in [1].

$$s(\beta)(u_h, \phi_h) = \sum_{K \in \mathcal{T}_h} \left\{ \alpha_K \left(\nabla(\pi_h p_h), \nabla(\pi_h \xi_h) \right)_K + \delta_K \left(\beta \cdot \nabla(\pi_h v_h), \beta \cdot \nabla(\pi_h \psi_h) \right)_K \right\},$$

where the cell-wise coefficients δ_K and α_K are chosen:

$$\delta_K := \frac{\delta_0 h_K^2}{6\nu + h_K |\beta|_K}, \quad \alpha_K = \alpha_0 h_K^2 / \nu.$$

Here, h_K denotes the diameter of the cell K and $|\beta|_K$ is the cell-wise value of the velocity field β . The parameters δ_0 and α_0 are usually chosen between 0.2 and 1. The stabilized discrete formulation reads: Find $u_h = (v_h, p_h) \in \hat{V}_h$ such that:

$$a(q_h, u_h)(\phi_h) + s(v_h)(u_h, \phi_h) = 0 \quad \forall \phi_h \in V_h.$$

For the analysis of stability and convergence of the described approximation scheme see Becker & Braack [3]. Similar stabilization techniques for the convective term are analyzed in Guermond [10].

5.3 Computational results

The described 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 & Braack [2]. 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 *AdmiralSlonik*

for treating optimization problems governed by partial differential equations and the finite element toolkit *Gascoigne3D*.

Application of the a posteriori error estimator for the quantity of interest (30) leads to the sequence of locally refined meshes, see in Figure 3.

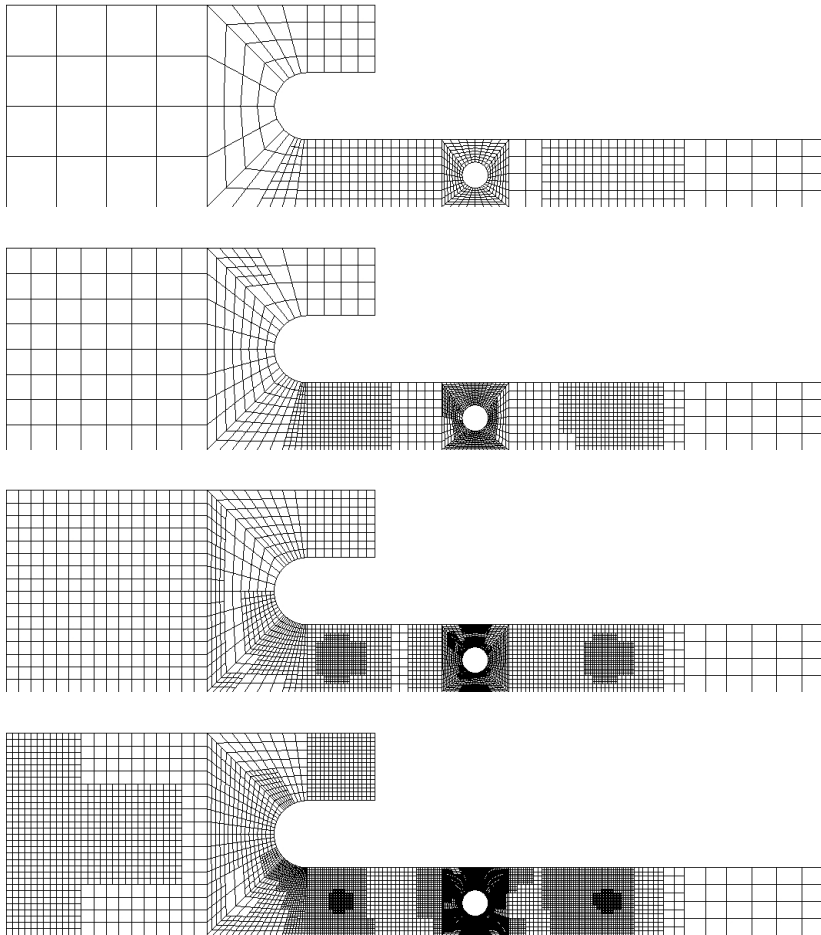


Fig. 3. Meshes generated by the adaptive algorithm with 746, 2 170, 5 094 and 11 068 nodes.

The comparison of the error in the quantity of interest (30) for uniform mesh refinement and the refinement produced by our a posteriori error estimator is done in Figure 4. 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 point measurements on sequence of uniformly refined meshes. In Table 1 the results are listed.

This turn out that for example 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 similar perturbation of the measurement \bar{C}_4 introduces an error in the quantity of interest of about 5.4%. In table 2 we investigate the quality of

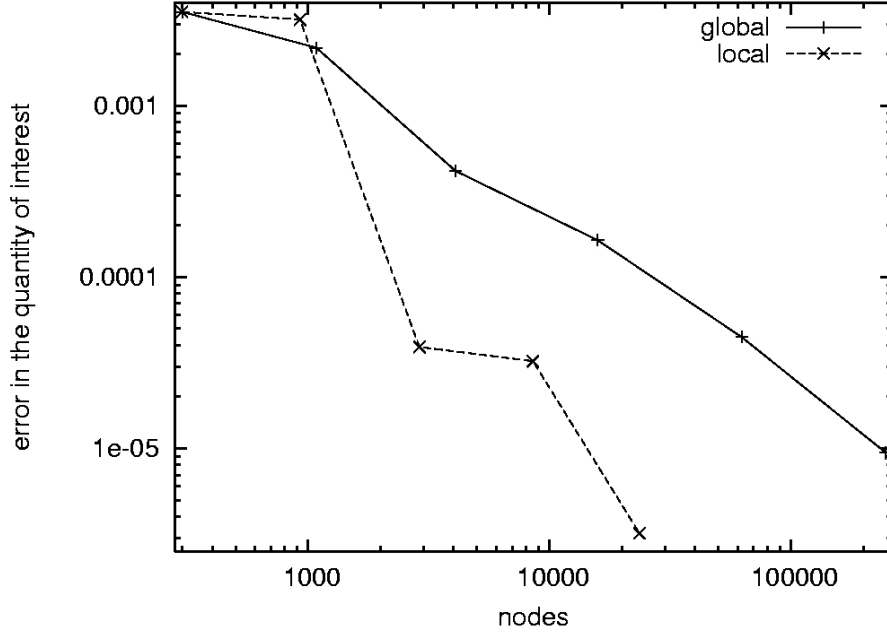


Fig. 4. Error in the quantity of interest E vs. number of mesh points for uniform mesh refinement and local refinement resulting from our a posteriori error estimator.

N	κ_1	κ_2	κ_3	κ_4
300	6.778e-2	-1.640e-2	5.032e-1	5.110e-1
1080	6.870e-2	-1.651e-2	5.268e-1	5.337e-1
4080	6.894e-2	-1.657e-2	5.336e-1	5.415e-1
15840	6.932e-2	-1.667e-2	5.353e-1	5.434e-1
62400	6.941e-2	-1.670e-2	5.358e-1	5.438e-1
247680	6.943e-2	-1.670e-2	5.359e-1	5.439e-1

Table 1

Relative condition numbers for point measurements on sequence of uniformly refined meshes.

the prediction of the relative error $\delta I/I$ in the quantity of interest for different error levels $\delta C_i/C_i$ in measurements.

6 Appendix

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

$\frac{\delta C_1}{C_1} \cdot 100\%$	$\frac{\delta I}{I} \cdot 100\%$	$\kappa_1 \frac{\delta C_1}{C_1} \cdot 100\%$
10 %	0.6904 %	0.6894 %
20 %	1.3824 %	1.3788 %
50 %	3.4682 %	3.4470 %
$\frac{\delta C_2}{C_2} \cdot 100\%$	$\frac{\delta I}{I} \cdot 100\%$	$\kappa_2 \frac{\delta C_2}{C_2} \cdot 100\%$
10 %	-0.1657 %	-0.1657 %
20 %	-0.3315 %	-0.3314 %
50 %	-0.8288 %	-0.8285 %
$\frac{\delta C_3}{C_3} \cdot 100\%$	$\frac{\delta I}{I} \cdot 100\%$	$\kappa_3 \frac{\delta C_3}{C_3} \cdot 100\%$
10 %	5.376 %	5.336 %
20 %	10.831 %	10.672 %
50 %	27.724 %	26.680 %
$\frac{\delta C_4}{C_4} \cdot 100\%$	$\frac{\delta I}{I} \cdot 100\%$	$\kappa_4 \frac{\delta C_4}{C_4} \cdot 100\%$
10 %	5.460 %	5.415 %
20 %	11.007 %	10.830 %
50 %	28.185 %	27.075 %

Table 2
 ??????????

Proof of Proposition 1:

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

$$a'_u(q, u)(\delta u, \phi) + a'_q(q, u)(\delta q, \phi) = 0 \quad \forall \phi \in V, \quad (33)$$

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). We obtain using the

Lagrange functional \mathcal{L} defined in (15):

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

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

$$j'(q)(\delta q) = \mathcal{L}'_u(q, u, z)(\delta u) + \mathcal{L}'_q(q, u, z)(\delta q), \quad (34)$$

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:

$$\mathcal{L}'_u(q, u, z)(\phi) = 0 \quad \forall \phi \in V. \quad (35)$$

We take derivative of (34) with respect to q in the direction τq and obtain:

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

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 (35) and moreover there holds:

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

due to (33). We complete the proof by setting $\delta q = e_j$, $\tau q = e_k$ and calculating the second derivatives of \mathcal{L} .

#

Proof of Proposition 3:

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

$$\mathcal{L}'(\xi)(\delta \xi) = 0 \quad \forall \delta \xi \in Q \times V \times V \quad (36)$$

and ξ_h is the corresponding Galerkin solution

$$\mathcal{L}'(\xi_h)(\delta \xi_h) = 0 \quad \forall \delta \xi_h \in V_h \times Q \times V_h. \quad (37)$$

Therefore, we obtain:

$$I(q, u) - I(q_h, u_h) = \mathcal{M}(x) - \mathcal{M}(x_h). \quad (38)$$

We rewrite the right hand side of (38) as follows:

$$\mathcal{M}(x) - \mathcal{M}(x_h) = \int_0^1 \mathcal{M}'(x_h + s e)(e) ds, \quad (39)$$

approximate the integral by the trapezoidal rule and obtain:

$$\mathcal{M}(x) - \mathcal{M}(x_h) = \frac{1}{2}\mathcal{M}'(x)(e) + \frac{1}{2}\mathcal{M}'(x_h)(e) + R, \quad (40)$$

where the remainder term R is given by

$$R = \frac{1}{2} \int_0^1 \mathcal{M}'''(x_h + se)(e, e, e)s(s-1)ds,$$

The term $\mathcal{M}'(x)(e)$ vanishes, and due to Galerkin orthogonality the term $\mathcal{M}'(x_h)(e)$ can be replaced by $\mathcal{M}'(x_h)(x - \tilde{x}_h)$ 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 \mathcal{L} . Therefore, there holds:

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

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

$$\begin{aligned} \mathcal{M}'_z(\xi, \chi)(\phi) &= \mathcal{L}''_{uz}(\xi)(v, \phi) + \mathcal{L}''_{qz}(\xi)(p, \phi) \\ &= \sum_{j=1}^{n_Q} \left(a'_u(q, u)(w_j, \phi) + a'_q(q, u)(e_j, \phi) \right). \end{aligned}$$

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

$$\mathcal{L}''_{zu}(\xi)(y, \phi) = -I'_u(q, u)(\phi) - \mathcal{L}''_{uu}(\xi)(v, \phi) - \mathcal{L}''_{qu}(\xi)(p, \phi) \quad \forall \phi \in V, \quad (41)$$

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

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

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

$$\begin{aligned} \mathcal{M}'_q(\xi, \chi)(e_j) &= \langle Hp, e_j \rangle_Q + I'_q(q, u)(e_j) \\ &\quad - \mathcal{L}''_{zq}(\xi)(y, e_j) - \mathcal{L}''_{uu}(\xi)(v, w_j) - \mathcal{L}''_{qu}(\xi)(p, w_j). \end{aligned}$$

Due to the definition of w_j and the equation (41) we obtain:

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

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

#

Proof of Theorem 1:

We apply Proposition 3 and Proposition 4. Due to the choice $Q_h = Q$ we may set $\tilde{x}_h = (\tilde{u}_h, q, \tilde{z}_h, \tilde{v}_h, p, \tilde{y}_h)$. This completes the proof.

#

Proof of Theorem 2:

There holds:

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

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

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

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 \mathcal{M} there holds:

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

This completes the proof.

#

References

- [1] R. Becker. Mesh adaptation for stationary flow control. *J. Math. Fluid Mech.*, 3(4):317–341, 2001.
- [2] R. Becker and M. Braack. Multigrid techniques for finite elements on locally refined meshes. *Numerical Linear Algebra with Applications*, 7:363–379, 2000. Special Issue.
- [3] R. Becker and M. Braack. A finite element pressure gradient stabilization for the Stokes equations based on local projections. *Calcolo*, 38(4):173–199, 2001.
- [4] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East-West J. Numer. Math.*, 4:237–264, 1996.
- [5] R. Becker and R. Rannacher. An optimal control approach to a-posteriori error estimation. In A. Iserles, editor, *Acta Numerica 2001*, pages 1–102. Cambridge University Press, 2001.
- [6] R. Becker and B. Vexler. A posteriori error estimation for finite element discretizations of parameter identification problems. *Numer. Math.*, 96(3):435–459, 2004.

- [7] P.G. Ciarlet. *Finite Element Methods for Elliptic Problems*. North-Holland, Amsterdam, 1978.
- [8] J.E. Dennis and 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. Technical report, Universit"at Bayreuth, 2003.
- [10] J.-L. Guermond. Stabilization of Galerkin approximations of transport equations by subgrid modeling. *Modél. Math. Anal. Numér.*, 33(6):1293–1316, 1999.
- [11] J.G. Heywood, R. Rannacher, and S. Turek. Artificial boundaries and flux and pressure conditions for the incompressible Navier-Stokes equations. *Int. J. Numer. Math. Fluids.*, 22:325–352, 1992.
- [12] L. Machiels, A.T. Patera, and J. Peraire. Output bound approximation for partial differential equations; application to the incompressible Navier-Stokes equations. In S. Biringen, editor, *Industrial and Environmental Applications of Direct and Large Eddy Numerical Simulation*. Springer, Berlin-Heidelberg-New York, 1998.
- [13] K. Malanowski. Sensitivity analysis for parametric optimal control of semilinear parabolic equations. *J. Convex Anal.*, 9(2):543–561, 2002.
- [14] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer Series in Operations Research. Springer New York, 1999.
- [15] J.T. Oden and S. Prudhomme. On goal-oriented error estimation for elliptic problems: Application to the control of pointwise errors. *Comput. Methods Appl. Mech. Engrg.*, 176:313–331, 1999.