SIAM J. CONTROL OPTIM.
Vol. 44, No. 5, pp. 1844–1863

A PRIORI ERROR ESTIMATES FOR THE FINITE ELEMENT DISCRETIZATION OF ELLIPTIC PARAMETER IDENTIFICATION PROBLEMS WITH POINTWISE MEASUREMENTS*

R. RANNACHER† AND B. VEXLER‡

Abstract. We develop an a priori error analysis for the finite element Galerkin discretization of parameter identification problems. The state equation is given by an elliptic partial differential equation of second order with a finite number of unknown parameters, which are estimated using pointwise measurements of the state variable.

Key words. parameter identification, finite elements, pointwise measurements, L^∞-error estimates

AMS subject classifications. 65K10, 65N30, 65N21, 49M25, 49K20

DOI. 10.1137/040611100

1. Introduction.

We consider parameter identification problems governed by an elliptic partial differential equation of second order. The finitely many unknown parameters are estimated using the measurements of point values of the *state variable*. Let $\Omega \subset \mathbb{R}^2$ be a convex polygonal domain; $L^2(\Omega)$ the corresponding Lebesgue space with inner product and norm denoted by $(\cdot,\cdot)$ and $\|\cdot\|_2$, respectively; and $H^m(\Omega)$ the Sobolev space of order $m \in \mathbb{N}$. With this notation, we set

$$V := \{v \in H^1(\Omega) \cap C(\overline{\Omega}) \,|\, v = 0 \text{ on } \partial\Omega\}.$$

The state variable $u \in V$ is determined by an elliptic partial differential equation (the *state equation*)

$$\begin{aligned} -\nabla \cdot (A(q)\nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

for given Hölder continuous $f \in C^\alpha(\overline{\Omega})$, $\alpha \in (0,1)$. Here, $Q \subset \mathbb{R}^{n_p}$ denotes the open admissible set of parameters $q \in \mathbb{R}^{n_p}$, for which $A(q) = (A_{ij}(q))$ is a symmetric and positive definite 2×2 matrix with twice continuously differentiable entries $A_{ij} : Q \to C^{1+\alpha}(\overline{\Omega})$. The above conditions guarantee that, for any admissible value of the parameter q, the corresponding solution u of the state equation (1.1) is in $H^2(\Omega)$ (see, e.g., Grisvard [12]). At the corner points of $\partial\Omega$, the second derivatives of the solution may become singular. However, u has Hölder continuous second derivatives, $u \in C^{2+\alpha}(\overline{\Omega}_d)$, for each subdomain $\Omega_d \subset \Omega$ with distance $d > 0$ to the corner points.

The usual weak formulation of (1.1) is

$$a(q)(u,\phi) = (f,\phi) \quad \forall \phi \in V, \tag{1.2}$$

---

*Received by the editors July 6, 2004; accepted for publication (in revised form) March 1, 2005; published electronically December 6, 2005. This work has been supported by the German Research Foundation (DFG) through the *Sonderforschungsbereich* 359 "Reactive Flows, Diffusion and Transport," and the *Graduiertenkolleg* "Complex Processes: Modeling, Simulation and Optimization" at the Interdisciplinary Center of Scientific Computing (IWR), University of Heidelberg.

http://www.siam.org/journals/sicon/44-5/61110.html

†Institute of Applied Mathematics, University of Heidelberg, Im Neuenheimer Feld 294, D-69120 Heidelberg, Germany (rannacher@iwr.uni-heidelberg.de).

‡Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenberger Strasse 69, A-4040 Linz, Austria (boris.vexler@oeaw.ac.at).

where the bilinear form $a(q)(\cdot,\cdot)$ is defined by

$$a(q)(u,\phi) := (A(q)\nabla u, \nabla\phi). \tag{1.3}$$

Further, the observation operator $C(\cdot)$ describing the mapping of the state variable u to the space of measurements $Z = \mathbb{R}^{n_m}$ is given by

$$C_i(v) = v(\xi_i), \quad i = 1, 2, \dots, n_m, \tag{1.4}$$

where $\{\xi_i\} \subset \Omega$ is a finite set of measurement points. We assume that $n_m \geq n_p$. The Euclidean product and norm on Q and Z are denoted by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$, respectively, and the same notation is also used for the corresponding natural norms of matrices.

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

$$\text{Minimize} \quad J(u) := \frac{1}{2}\|C(u) - \hat{C}\|^2 \tag{1.5}$$

under the constraint (1.2). Throughout, we assume the existence of a solution $(u, q) \in V \times Q$ of the problem (1.2), (1.5). For an analysis of the existence of solutions for parameter identification problems, see, e.g., Banks and Kunisch [2], Kravaris and Seinfeld [16], and Litvinov [17].

The state equation is discretized by a conforming finite element Galerkin method defined on a family $\{\mathcal{T}_h\}_{h>0}$ of shape regular quasi-uniform meshes $\mathcal{T}_h = \{K\}$ consisting of closed *cells* K which are either triangles or quadrilaterals. The straight parts which make up the boundary ∂K of a cell K are called *faces*. The mesh parameter h is defined as a cellwise constant function by setting $h_{|K} = h_K$, and h_K is the diameter of K. Usually we use the symbol h also for the maximal cell size, i.e.,

$$h = \max_{K\in\mathcal{T}_h} h_K. \tag{1.6}$$

For convenience, we assume that $0 < h < 1$. On the mesh $\mathcal{T}_h$ we define finite element spaces $V_h \subset V$ consisting of linear or bilinear shape functions; see, e.g., Brenner and Scott [5] or Johnson [14]. The corresponding discrete state $u_h \in V_h$ and parameter $q_h \in Q$ are determined by

$$\text{Minimize} \quad J(u_h), \tag{1.7}$$

under the constraint

$$a(q_h)(u_h, \phi_h) = (f, \phi_h) \quad \forall \phi_h \in V_h. \tag{1.8}$$

Since Q is finite dimensional, the parameter q_h is determined in the same space Q.

The main purpose of this paper is to analyze the behavior of the error in parameters $\|q - q_h\|$ for h tending to zero. There are a number of publications in which a priori error estimates are derived for optimal control problems governed by partial differential equations; see, e.g., Falk [8], Arada, Casas, and Tröltzsch [1], Deckelnic and Hinze [6], and Gunzburger and Hou [13]. However, there are only few published results on this topic in the context of parameter identification problems; see Falk [9], Neittaanmäki and Tai [18], and Kärkkäinen [15].

In Vexler [26], an a priori error analysis for the case of V-stable observation operators $C_i(\cdot)$ is developed and optimal-order convergence is shown,

$$\|q - q_h\| = \mathcal{O}(h^2), \tag{1.9}$$

essentially under the assumption that $|C_i(u) - C_i(u_h)| \leq ch^2\|u\|_{H^2}$. However, the generalization of this result to pointwise observations is not straightforward. For this case, we prove in this paper that, under certain regularity conditions,

$$\|q - q_h\| = \mathcal{O}(h^2|\log(h)|^2). \tag{1.10}$$

The proof uses the technique for estimating discrete Green functions developed in Frehse and Rannacher [10]. A complementary result of a posteriori error analysis for parameter identification problems is given in Becker and Vexler [4].

To the authors' knowledge, this is the first a priori error analysis for parameter identification problems with pointwise observations. The consideration of pointwise observations in determining discrete parameters seems very natural in view of practical measurement techniques; see [3] for applications in reactive flow analysis.

The paper is organized as follows. In the next section, we describe an algorithm for solving problem (1.7), (1.8). In section 3, we present a paradigm for a priori error analysis for discretization of a class of optimization problems. Thereafter, in section 4, we derive the announced error estimate using an L^∞-stability result, which is proven in section 5. In section 6, we present a numerical example confirming the asymptotic sharpness of our error estimate. Possible extensions are addressed in the last section.

2. Optimization algorithm. In this section, we reformulate the problem under consideration as an unconstrained optimization problem and describe a solution algorithm for it. Since the coefficient matrix $A(q)$ is assumed to be positive definite for parameters $q \in Q \subset \mathbb{R}^{n_p}$, the relation

$$a(q)(S(q), \phi) = (f, \phi) \quad \forall \phi \in H_0^1(\Omega) \tag{2.1}$$

defines an operator $S : Q \to H_0^1(\Omega)$. By the assumptions on the data of the problem and the Sobolev embedding theorem,

$$S : Q \to H_0^1(\Omega) \cap H^2(\Omega) \subset V. \tag{2.2}$$

The solution operator S can be shown to possess first and second derivatives which are continuous with respect to the norm of V; see Theorem 2.1 below. We recall that the existence of a solution $q \in Q$ of problem (1.5) is assumed. Let $Q_0 \subset Q$ be an open bounded set containing the optimal parameter q on which the coefficient matrix $A(q)$ is uniformly positive definite; i.e., there exists $\gamma \in \mathbb{R}_+$ such that

$$p^* A(q) p \geq \gamma \|p\|^2 \quad \forall p \in Q, \quad \forall q \in Q_0, \tag{2.3}$$

uniformly with respect to $x \in \overline{\Omega}$. We introduce the *reduced observation operator* $c : Q_0 \to Z$ by

$$c(q) := C(S(q)). \tag{2.4}$$

This allows us to reformulate the problem under consideration as an unconstrained optimization problem with the reduced cost functional $j : Q_0 \to \mathbb{R}$:

$$\text{Minimize} \quad j(q) := \frac{1}{2}\|c(q) - \hat{C}\|^2, \quad q \in Q_0. \tag{2.5}$$

Denoting by $G = c'(q) \in \mathbb{R}^{n_p \times n_m}$ the Jacobian matrix of the reduced observation operator $c(\cdot)$, the first-order necessary optimality condition $j'(q) = 0$ for (2.5) reads

$$G^*(c(q) - \hat{C}) = 0, \tag{2.6}$$

where G^* denotes the transpose of G. The positive semidefiniteness of the Hessian matrix $H := \nabla^2 j(q)$ is the second-order necessary optimality condition. A solution q of problem (2.5) is called *stable* if the sufficient optimality condition holds, i.e., if the Hessian H is positive definite. Throughout, we will assume the solution q to be stable. The stability of the solution is given, for instance, if the value of the cost functional $\|C(u) - \hat{C}\|$ is small enough and the matrix G has full rank n_p; see, e.g., [26] for details.

Since by assumption the matrix coefficient $A(\cdot)$ is twice continuously differentiable, there holds

$$\sup_{\xi \in Q_0} |||A(\xi)|||_{1,\infty} + \sup_{\xi \in Q_0} |||A'_{q_j}(\xi)|||_{1,\infty} + \sup_{\xi \in Q_0} |||A''_{q_j q_k}(\xi)|||_{1,\infty} < \infty, \tag{2.7}$$

where $|||B|||_{1,\infty} := \max_{i,j=1,2} \|B_{ij}\|_{1,\infty}$ for a matrix function $B = (B_{ij}) \in C^1(\overline{\Omega})^{2\times 2}$.

In the following propositions, we give representations of the Jacobian G of $c(\cdot)$, the Hessian H of $j(\cdot)$, and the Hessian of $c_i(\cdot)$.

THEOREM 2.1. *Let the reduced observation operator $c(\cdot)$ and the reduced functional $j(\cdot)$ be defined as in* (2.4) *and* (2.5)*, respectively.*

(i) *The elements of the Jacobian of $c(\cdot)$ at some $q \in Q_0$ are given by*

$$G_{ij} = \frac{\partial c_i}{\partial q_j}(q) = C_i(w_j), \quad i = 1, \dots, n_m,\ j = 1, \dots, n_p, \tag{2.8}$$

where $w_j \in V$ are the solutions of the problems

$$a(q)(w_j, \phi) = -(A'_{q_j}(q)\nabla u, \nabla\phi) \quad \forall \phi \in V, \tag{2.9}$$

with $u = S(q)$. The functions $w_j \in V$ depend continuously on $q \in Q$.

(ii) *The Hessian of $j(\cdot)$ can be expressed by*

$$H = G^* G + M, \tag{2.10}$$

where the matrix $M \in \mathbb{R}^{n_p \times n_p}$ is given by

$$M = \sum_{i=1}^{n_m} c_i''(q)(c_i(q) - \hat{C}_i). \tag{2.11}$$

The Hessian of $c_i(q)$ is given by

$$\frac{\partial^2}{\partial q_j\,\partial q_k} c_i(q) = C_i(v_{jk}), \tag{2.12}$$

where the $v_{jk} \in V$ are the solutions of the problems

$$\begin{aligned} a(q)(v_{jk}, \phi) = &-(A'_{q_j}(q)\nabla w_k, \nabla\phi) - (A'_{q_k}(q)\nabla w_j, \nabla\phi) \\ &- (A''_{q_j q_k}(q)\nabla u, \nabla\phi) \quad \forall \phi \in V, \end{aligned} \tag{2.13}$$

with w_j as defined in (2.9)*. The functions $v_{jk} \in V$ depend continuously on $q \in Q$.*

Proof. The derivation of the derivatives of $c(\cdot)$ uses the chain rule,

$$\frac{\partial c_i}{\partial q_j}(q) = \frac{\partial}{\partial q_j} C_i(S(q)) = S'_{q_j}(q)(\xi_i) =: w_j(\xi_i),$$

for $i = 1, \dots, n_m$, $j = 1, \dots, n_p$, where the functions w_j are determined by the relations (2.9). This is seen by considering the limit of difference quotients

$$\begin{aligned} 0 &= \lim_{t\to 0} \frac{1}{t}(a(q+tq_j)(S(q+tq_j),\phi) - (f,\phi) - a(q)(S(q),\phi) + (f,\phi)) \\ &= a(q)(S'_{q_j}(q),\phi) + a'_{q_j}(q)(S(q),\phi) = a(q)(w_j,\phi) + a'_{q_j}(q)(u,\phi). \end{aligned}$$

Analogously, we obtain

$$\frac{\partial^2 c_i}{\partial q_j \partial q_k}(q) = \frac{\partial^2}{\partial q_j \partial q_k} C_i(S(q)) = S''_{q_j q_k}(q)(\xi_i) =: v_{jk}(\xi_i),$$

where the functions v_{jk} are determined by the relations (2.13). To see this, we consider the limit of the difference quotient

$$\begin{aligned} 0 = \lim_{t\to 0} \frac{1}{t} &(a(q+tq_k)(S'_{q_j}(q+tq_k),\phi) + a'_{q_j}(q+tq_k)(S(q+tq_k),\phi) \\ &- a(q)(S'_{q_j}(q),\phi) - a'_{q_j}(q)(S(q),\phi)) \\ = a(q)&(S''_{q_j q_k}(q),\phi) + a'_{q_k}(q)(S'_{q_j}(q),\phi) + a'_{q_j}(q)(S'_{q_k}(q),\phi) + a''_{q_k q_j}(q)(S(q),\phi). \end{aligned}$$

In Lemma 2.2 below, we will show that all the functions u, w_j, v_{jk} are in $V \cap H^2(\Omega)$. This is due to the fact that they are determined by second-order elliptic boundary value problems on a convex domain, with smooth coefficients and right-hand sides in $L^2(\Omega)$ which depend continuously on the parameter $q \in Q$. This implies that also their solutions depend continuously on $q \in Q$ with respect to the norm of the solution space $H_0^1(\Omega) \cap H^2(\Omega) \subset V$. Hence, the solution operator $S : Q \to V$ is twice continuously differentiable. This completes the proof. ∎

In practice the Hessian H of $j(\cdot)$ is computed using the representation

$$M_{jk} = -(A'_{q_j}(q)\nabla w_k, \nabla z) - (A'_{q_k}(q)\nabla w_j, \nabla z) - (A''_{q_j q_k}(q)\nabla u, \nabla z), \tag{2.14}$$

with the function z determined by the dual equation

$$(A(q)\nabla\phi, \nabla z) = \langle C(u) - \bar{C}, C(\phi)\rangle. \tag{2.15}$$

For later purposes, we provide some a priori bounds for the solutions of the boundary value problems introduced in Theorem 2.1, which follow by standard results of elliptic regularity theory.

LEMMA 2.2. *For the solutions of the elliptic boundary value problems* (2.9) *and* (2.13) *there hold the global L^2 a priori estimates*

$$\|u\|_{2,2} + \|w_j\|_{2,2} + \|v_{jk}\|_{2,2} \le c, \tag{2.16}$$

where c is a generic constant depending only on the data of the problem. Further, for each subdomain $\Omega_d \subset \Omega$ with distance $d > 0$ to the corner points, there hold the L^∞ a priori estimates

$$\|u\|_{C^{2+\alpha}(\overline{\Omega}_d)} + \|w_j\|_{C^{2+\alpha}(\overline{\Omega}_d)} + \|v_{jk}\|_{C^{2+\alpha}(\overline{\Omega}_d)} \le c_d \tag{2.17}$$

with a generic constant $c_d \approx d^{-1}$.

Proof. The variational equations defining u as well as w_j and v_{jk} can be rewritten in such a form that they represent second-order elliptic boundary value problems with smooth coefficients and right-hand sides which are bounded functionals on $L^2(\Omega)$ and $C^\alpha_{\mathrm{loc}}(\Omega)$, respectively, as follows:

$$a(q)(u,\phi) = (f,\phi) \quad \forall \phi \in V, \tag{2.18}$$

$$a(q)(w_j,\phi) = (\nabla\cdot A'_{q_j}(q)\nabla u,\phi) \quad \forall \phi \in V, \tag{2.19}$$

$$\begin{aligned} a(q)(v_{jk},\phi) &= (\nabla\cdot A'_{q_j}(q)\nabla w_k,\phi) + (\nabla\cdot A'_{q_k}(q)\nabla w_j,\phi) \\ &\quad + (\nabla\cdot A''_{q_jq_k}(q)\nabla u,\phi) \quad \forall \phi\in V. \end{aligned} \tag{2.20}$$

By assumption the coefficient functions $A'_{q_j}(q)$ and $A''_{q_jq_k}(q)$ are smooth. In view of the convexity of the polygonal domain Ω, the H^2-regularity estimates then follow by results from Grisvard [12]. A reference for the corresponding $C^{2+\alpha}$-estimates is Gilbarg and Trudinger [11]. In the first step, from (2.18), we get

$$\|u\|_{2,2} \le c\|f\|_2 \le c,$$
$$\|u\|_{C^{2+\alpha}(\overline{\Omega}_d)} \le c_d\big\{\|f\|_{C^\alpha(\overline{\Omega}_{d/2})} + \|u\|_{2,2}\big\} \le c_d.$$

Then, using this in (2.19), we conclude that

$$\|w_j\|_{2,2} \le c\|u\|_{2,2} \le c,$$
$$\|w_j\|_{C^{2+\alpha}(\overline{\Omega}_d)} \le c_d\big\{\|u\|_{C^{2+\alpha}(\overline{\Omega}_{d/2})} + \|u\|_{2,2}\big\} \le c_d.$$

Finally, this is used in (2.20) and allows us to conclude that

$$\|v_{jk}\|_{2,2} \le c\max\big\{\|w_j\|_{2,2}, \|w_k\|_{2,2}\big\} + c\|u\|_{2,2} \le c,$$
$$\|v_{jk}\|_{C^{2+\alpha}(\overline{\Omega}_d)} \le c_d\big\{\|w_j\|_{C^{2+\alpha}(\overline{\Omega}_{d/2})} + \|u\|_{C^{2+\alpha}(\overline{\Omega}_{d/2})} + \|w_j\|_{2,2}\big\} \le c_d.$$

This completes the proof. ◻

Similar to the continuous case, we introduce a discrete solution operator S_h: $Q_0 \to V_h$ by the equation

$$a(q_h)(S_h(q_h),\phi_h) = (f,\phi_h) \quad \forall \phi_h \in V_h,\ q_h \in Q_0. \tag{2.21}$$

As before, we turn the discrete problem (1.7), (1.8) into an unconstrained minimization problem,

$$\text{Minimize} \quad j_h(q_h) := \frac{1}{2}\|c_h(q_h) - \hat{C}\|^2, \quad q_h \in Q_0, \tag{2.22}$$

where the discrete reduced observation operator c_h is defined by

$$c_h(q_h) = C(S_h(q)). \tag{2.23}$$

Denoting the corresponding Jacobian by $G_h = c'_h(q_h)$, the necessary optimality condition $j'_h(q_h) = 0$ reads

$$G_h^*(c_h(q_h) - \hat{C}) = 0. \tag{2.24}$$

The derivatives of the discrete observation operator c_h can be computed in a way analogous to that in Theorem 2.1.

Problem (2.22) is solved iteratively starting with an initial guess q_h^0 and using the recursive setting $q_h^{k+1} = q_h^k + \delta q_h$. The update δq_h is obtained as the solution of the system of linear equations

$$H_k\,\delta q_h = G_h^*(\hat{C} - c_h(q_h^k)), \tag{2.25}$$

where $G_h = c_h'(q_h^k)$, and H_k is an appropriate symmetric approximation of the Hessian $\nabla^2 j_h(q_h^k)$. The most widely used choice of the matrix $H_k = G_h^* G_h$ leads us to the Gauß–Newton algorithm; see, e.g., Nocedal and Wright [21].

For one step of the Gauß–Newton algorithm the state equation and n_p tangent problems (2.9) have to be solved, which involves the same linear operator but with different right-hand sides. Due to the small dimension n_p of the parameter space Q, the solution of (2.25) is uncritical. For discussing other Newton-type methods and trust-region techniques for globalization of the convergence in this context, see [26].

3. A paradigm for a priori error analysis. In this section we present a general approach to the error analysis of a class of optimization problems such as are considered in this paper. The main result is stated in the following theorem. It is a variant of well-known perturbation theorems for differentiable mappings, which is particularly tailored to the present situation. However, it seems easier to include the elementary proof than to search for the precise reference.

THEOREM 3.1. *Let $F, F_h : \mathbb{R}^n \to \mathbb{R}^n$, for a discretization parameter $h \in \mathbb{R}_+$, be continuously differentiable operators, and $x \in \mathbb{R}^n$ be a solution of $F(x) = 0$. Let the following conditions be fulfilled:*

(i) *The derivative $F'(x)$ is positive definite; i.e., there is a constant $\gamma > 0$ such that*

$$p^* F'(x) p \geq \gamma \|p\|^2, \quad p \in \mathbb{R}^n. \tag{3.1}$$

(ii) *There is a neighborhood U of x and a positive number $L(h) \in \mathbb{R}_+$ such that*

$$\|F_h'(\xi) - F_h'(\eta)\| \leq L(h) \|\xi - \eta\| \quad \forall \xi, \eta \in U. \tag{3.2}$$

(iii) *With the h-dependent constant $L(h)$, there holds*

$$\lim_{h \to 0} L(h) \|F(x) - F_h(x)\| = 0. \tag{3.3}$$

(iv) *There holds*

$$\lim_{h \to 0} \|F'(x) - F_h'(x)\| = 0. \tag{3.4}$$

Then, for h small enough, there exists $x_h \in U$ such that $F_h(x_h) = 0$, and $F_h'(x_h)$ is positive definite uniformly in h. Further, there holds the a priori error estimate

$$\|x - x_h\| \leq \frac{2}{\gamma} \|F(x) - F_h(x)\|. \tag{3.5}$$

Proof. Due to condition (iv), we can choose a positive number $h_1 \in \mathbb{R}_+$ such that for $h \leq h_1$ there holds

$$\|F'(x) - F_h'(x)\| \leq \frac{1}{4} \gamma. \tag{3.6}$$

Moreover, for $\rho = \rho(h) = \frac{\gamma}{kL(h)}$, with some $k \geq 4$ sufficiently large, there holds

$$B_\rho(x) = \{\xi \in \mathbb{R}^n, \; \|x - \xi\| \leq \rho\} \subset U. \tag{3.7}$$

For this choice, we obtain that, for $h \leq h_1$, $F_h'(\cdot)$ is positive definite on $B_\rho(x)$:

$$\begin{aligned} p^* F_h'(\xi) p &= p^* F'(x) p + p^* (F_h'(x) - F'(x)) p + p^* (F_h'(\xi) - F_h'(x)) p \\ &\geq \gamma \|p\|^2 - \|F_h'(x) - F'(x)\| \, \|p\|^2 - \|F_h'(\xi) - F_h'(x)\| \, \|p\|^2 \\ &\geq \left(\gamma - \frac{1}{4} \gamma - L(h) \rho \right) \|p\|^2 \geq \frac{1}{2} \gamma \|p\|^2. \end{aligned}$$

In a similar way, we conclude that, for $h \le h_1$, $F_h'(\cdot)$ is also bounded on $B_\rho(x)$:

$$\|F_h'(\xi)\| \le \beta := \|F'(x)\| + \frac{1}{2}\gamma.$$

Next, we prove that there exists a unique $x_h \in B_\rho(x)$ with $F_h(x_h) = 0$. To this end, we define an operator $D_s : \mathbb{R}^n \to \mathbb{R}^n$, for $s \in \mathbb{R}_+$, by

$$D_s(\xi) = \xi - sF_h(\xi).$$

For a certain choice of s, we show that D_s is a contraction on $B_\rho(x)$, and we use the Banach fixed point theorem. For $\xi \in B_\rho(x)$, $h \le h_1$, and an arbitrary $p \in \mathbb{R}^n$, there holds

$$\begin{aligned}\|D_s'(\xi)p\|^2 = \|p - sF_h'(\xi)p\|^2 &= \|p\|^2 - 2sp^*F_h'(\xi)p + s^2\|F_h'(\xi)p\|^2 \\ &\le (1 - s\gamma + s^2\beta^2)\|p\|^2.\end{aligned}$$

For the choice $s = \gamma(2\beta^2)^{-1}$, we obtain

$$\|D_s'(\xi)p\|^2 \le \Big(1 - \frac{\gamma^2}{4\beta^2}\Big)\|p\|^2,$$

and consequently,

$$\|D_s'(\xi)\| \le \Big(1 - \frac{\gamma^2}{4\beta^2}\Big)^{1/2} < 1.$$

Moreover, for arbitrary $\xi \in B_\rho(x)$, there holds

$$\begin{aligned}\|x - D_s(\xi)\| &= \|D_s(x) - D_s(\xi) + sF_h(x)\| \\ &\le \|D_s(x) - D_s(\xi)\| + s\|F_h(x) - F(x)\| \\ &\le \|D_s'(\eta)\|\,\|x - \xi\| + s\|F_h(x) - F(x)\|\end{aligned}$$

for a certain $\eta \in B_\rho$. Hence, the above estimate implies

$$\begin{aligned}\|x - D_s(\xi)\| &\le \Big(1 - \frac{\gamma^2}{4\beta^2}\Big)^{1/2}\rho + s\,\|F_h(x) - F(x)\| \\ &= \rho\left\{\Big(1 - \frac{\gamma^2}{4\beta^2}\Big)^{1/2} + s\frac{k}{\gamma}L(h)\|F_h(x) - F(x)\|\right\}.\end{aligned}$$

Due to condition (iii), there is a number $h_2 \in \mathbb{R}_+$ such that, for $h \le h_2$, there holds

$$L(h)\|F_h(x) - F(x)\| \le \frac{\gamma}{ks}\left\{1 - \Big(1 - \frac{\gamma^2}{4\beta^2}\Big)^{1/2}\right\}.$$

Hence, for $h < h_0 := \min\{h_1, h_2\}$,

$$\|x - D_s(\xi)\| \le \rho,$$

and consequently $D_s(\xi) \in B_\rho(x)$. For $h \le h_0$, by the Banach fixed point theorem, we obtain the existence of $x_h \in B_\rho(x)$ with $F_h(x_h) = 0$. By construction of $B_\rho(x)$, the derivative $F_h'(x_h)$ is positive definite with the h-independent constant $\frac{1}{2}\gamma$. This implies that, for a certain $\xi \in B_\rho(x)$,

$$(x - x_h)^*(F_h(x) - F_h(x_h)) = (x - x_h)^*F_h'(\xi)(x - x_h) \ge \frac{\gamma}{2}\|x - x_h\|^2.$$

Hence, using $F(x) = F_h(x_h) = 0$,

$$\|x - x_h\|^2 \leq \frac{2}{\gamma}(x - x_h)^*(F_h(x) - F_h(x_h)) = \frac{2}{\gamma}(x - x_h)^*(F_h(x) - F(x)) \\ \leq \frac{2}{\gamma}\|F_h(x) - F(x)\|\,\|x - x_h\|.$$

This completes the proof. □

4. A priori error estimation.

In this section we apply the paradigm presented in section 3 to the problem under consideration. We prove the following theorem.

THEOREM 4.1. *Let $q \in Q$ be a stable solution of (2.5). Then, for h small enough, there exists a stable solution $q_h \in Q$ of (2.22), and there holds the following a priori error estimate:*

$$\|q - q_h\| = \mathcal{O}(h^2|\log(h)|^2). \tag{4.1}$$

On the basis of the estimate (4.1), we can also derive optimal-order estimates for the error $u - u_h$ in the corresponding states. However, since this would be a simple exercise using the arguments developed below, and since the optimal states are of only minor practical interest in parameter estimation problems, we do not state these estimates.

The proof of Theorem 4.1 is given by checking the conditions from Theorem 3.1 for the operators

$$F(\xi) := \nabla j(\xi), \quad F_h(\xi) := \nabla j_h(\xi).$$

The constant in (4.1) turns out to depend in a reciprocal way on the distance

$$\delta := \min_{i=1,\dots,n_m} \operatorname{dist}(\xi_i, \Sigma)$$

of the set of measurement points ξ_i to the set Σ of corner points of $\partial\Omega$. Therefore, we will use generic constants c and c_δ, where c depends only on the domain Ω, the force f, and the characteristics of the mesh family $\{\mathcal{T}_h\}_h$, while c_δ may additionally depend on the distance δ like $c_\delta \approx \delta^{-1}$. Further, by $L^p(\Omega)$ and $W^{m,p}(\Omega)$, for $m \in \mathbb{N}$ and $1 \leq p \leq \infty$, we denote the standard Lebesgue and Sobolev spaces, respectively, and by $\|\cdot\|_p$ and $\|\cdot\|_{m,p}$ the corresponding norms. The restriction of such a norm to a subset $\Omega' \subset \Omega$ is indicated by $\|\cdot\|_{m,p;\Omega'}$.

By $i_h : C(\overline{\Omega}) \to V_h$ we denote the usual (linear) operator of nodal interpolation for which the following cellwise estimate is well known (see Brenner and Scott [5]):

$$h_K^{-2}\|v - i_h v\|_{p;K} + h_K^{-1}\|\nabla(v - i_h v)\|_{p;K} + \|\nabla^2 i_h v\|_{p;K} \leq c\|\nabla^2 v\|_{p;K}, \tag{4.2}$$

for $1 \leq p \leq \infty$, with constants c independent of h.

An important ingredient of the proof of Theorem 4.1 is the following L^∞-stability theorem. For $d > 0$, we define the subset $\Omega_d \subset \Omega$ by

$$\Omega_d := \{x \in \Omega,\ \operatorname{dist}(x, \Sigma) > d\}.$$

THEOREM 4.2 (stability theorem). *Let $q \in Q_0$, $\psi \in H_0^1(\Omega) \cap C(\bar{\Omega})$, and a matrix $B = B(x) \in W^{1,\infty}(\Omega)^{2\times 2}$ be given. Moreover, let $v_h \in V_h$ be a solution of*

$$a(q)(v_h, \phi_h) = (B\nabla\psi, \nabla\phi_h) \quad \forall \phi_h \in V_h. \tag{4.3}$$

Then, there hold the L^2*-stability estimate*

$$\|v_h\|_2 + h\|\nabla v_h\|_2 \le c\,|||B|||_{1,\infty}\big\{\|\psi\|_2 + h\|\nabla\psi\|_2\big\} \tag{4.4}$$

and the local L^∞*-stability estimate*

$$\|v_h\|_{\infty;\Omega_d} \le c_d\,|||B|||_{1,\infty}\big\{|\log(h)|\,\|\psi\|_{\infty;\Omega_{d/2}} + \|\psi\|_2 + h\|\nabla\psi\|_2\big\}, \tag{4.5}$$

with a constant $c_d \approx d^{-1}$.

The L^2 estimate (4.4) is a standard result from finite element analysis, while the L^∞ estimate (4.5) can be concluded by estimates of discrete Green functions such as these developed in Frehse and Rannacher [10] and Rannacher and Scott [24] (see also Brenner and Scott [5, Chapter 7]). The proof for this is given in section 5 below. A similar L^∞-stability result has been proven in Rannacher [22] in the time-dependent parabolic case. For the solution q of problem (2.5), we introduce $\bar{u}_h \in V_h$ determined by

$$a(q)(\bar{u}_h, \phi_h) = (f, \phi_h) \quad \forall \phi_h \in V_h. \tag{4.6}$$

Further, we define $w_{j,h} \in V_h$ and $v_{jk,h} \in V_h$, for $j,k = 1,2,\ldots,n_p$, as the solutions of the problems

$$a(q)(w_{j,h}, \phi_h) = -(A'_{q_j}(q)\nabla\bar{u}_h, \nabla\phi_h) \quad \forall \phi_h \in V_h \tag{4.7}$$

and

$$\begin{aligned} a(q)(v_{jk,h}, \phi_h) = &-(A'_{q_j}(q)\nabla w_{k,h}, \nabla\phi_h) - (A'_{q_k}(q)\nabla w_{j,h}, \nabla\phi_h) \\ &- (A''_{q_j q_k}(q)\nabla\bar{u}_h, \nabla\phi_h) \quad \forall \phi_h \in V_h, \end{aligned} \tag{4.8}$$

respectively. The next lemma provides necessary estimates for the errors $u - \bar{u}_h$, $w_j - w_{j,h}$, and $v_{jk} - v_{jk,h}$. We recall the notation $\delta := \min_{i=1,\ldots,n_m} \operatorname{dist}(\xi_i, \Sigma)$.

LEMMA 4.3. *Under the above assumptions the following estimates hold:*

$$\|C(u - \bar{u}_h)\| \le c_\delta h^2 |\log(h)|, \tag{4.9}$$

$$\|C(w_j - w_{j,h})\| \le c_\delta h^2 |\log(h)|^2, \quad j = 1,2,\ldots,n_p, \tag{4.10}$$

$$\|C(v_{jk} - v_{jk,h})\| \le c_\delta h^2 |\log(h)|^3, \quad j,k = 1,2,\ldots,n_p. \tag{4.11}$$

Proof. The proof uses the a priori bounds (2.16) and (2.17) provided in Lemma 2.2 for u, and the auxiliary functions w_j, v_{jk}, $j,k = 1,\ldots,n_p$, corresponding to arbitrary $q \in Q_0$.

(i) By definition, $\bar{u}_h$ is the Ritz projection of u corresponding to the energy form $a(q)(\cdot,\cdot)$, i.e.,

$$a(q)(\bar{u}_h, \phi_h) = a(q)(u, \phi_h) \quad \forall \phi_h \in V_h.$$

By the standard L^2-error estimate for finite elements, there holds

$$\|u - \bar{u}_h\|_2 + h\|\nabla(u - \bar{u}_h)\|_2 \le ch^2. \tag{4.12}$$

Further, applying the L^∞-stability estimate (4.5) of Theorem 4.2 for the equation

$$a(q)(i_h u - \bar{u}_h, \phi_h) = a(q)(i_h u - u, \phi_h) \quad \forall \phi_h \in V_h,$$

with the nodal interpolant $i_h u \in V_h$ of u, yields the estimate

$$\|i_h u - \bar{u}_h\|_{\infty;\Omega_\delta} \leq c_\delta \big\{ |\log(h)|\, \|i_h u - u\|_{\infty;\Omega_{\delta/2}} + \|i_h u - u\|_2 + h\|\nabla(i_h u - u)\|_2 \big\}.$$

From this, using the approximation properties (4.2) of i_h, we conclude the error estimate

$$\|u - \bar{u}_h\|_{\infty;\Omega_\delta} \leq \|u - i_h u\|_{\infty;\Omega_\delta} + \|i_h u - \bar{u}_h\|_{\infty;\Omega_\delta} \leq c_\delta h^2 |\log(h)|. \tag{4.13}$$

Here, the constant c_δ depends on the global H^2 norm and the local $W^{2,\infty}$ norm of the solution, which are both known to be bounded in view of the a priori bounds (2.16) and (2.17). Since $\xi_i \in \overline{\Omega}_\delta$, we obtain the estimate (4.9).

(ii) For proving (4.10), we introduce an additional discrete variable $\bar{w}_{j,h}$ determined by the equation

$$a(q)(\bar{w}_{j,h}, \phi_h) = -(A'_{q_j}(q)\nabla u, \nabla \phi_h) \quad \forall \phi_h \in V_h.$$

The error $e = w_j - w_{j,h}$ is split like $e = e_1 + e_2$, with $e_1 = w_j - \bar{w}_{j,h}$ and $e_2 = \bar{w}_{j,h} - w_{j,h}$. For the Ritz-projection error e_1, as before, there holds the L^2-error estimate

$$\|e_1\|_2 + h\|\nabla e_1\|_2 \leq ch^2 \|w_j\|_{2,2} \leq ch^2$$

and the pointwise error estimate

$$\|e_1\|_{\infty;\Omega_\delta} \leq c_\delta h^2 \big\{ |\log(h)|\, \|\nabla^2 w_j\|_{\infty;\Omega_{\delta/2}} + \|w_j\|_{2,2} \big\} \leq c_\delta h^2 |\log(h)|.$$

For $e_2 \in V_h$, we have

$$a(q)(e_2, \phi_h) = -(A'_{q_j}(q)\nabla(u - \bar{u}_h), \nabla \phi_h) \quad \forall \phi_h \in V_h.$$

Hence, the L^2-stability estimate (4.4) of Theorem 4.2 and the estimate (4.12) imply

$$\|e_2\|_2 + h\|\nabla e_2\|_2 \leq c\big\{ \|u - \bar{u}_h\|_2 + h\|\nabla(u - \bar{u}_h)\|_2 \big\} \leq ch^2.$$

This shows that, for $j = 1, \ldots, n_p$,

$$\|w_j - w_{j,h}\|_2 + h\|\nabla(w_j - w_{j,h})\|_2 \leq ch^2. \tag{4.14}$$

Further, the L^∞-stability estimate (4.5) of Theorem 4.2 yields

$$\|e_2\|_{\infty;\Omega_\delta} \leq c_\delta \big\{ |\log(h)|\, \|u - \bar{u}_h\|_{\infty;\Omega_{\delta/2}} + \|u - \bar{u}_h\|_2 + h\|\nabla(u - \bar{u}_h)\|_2 \big\},$$

which, by (4.12) and (4.13), implies $\|e_2\|_{\infty;\Omega_\delta} \leq c_\delta |\log(h)|^2 h^2$. We obtain

$$\|w_j - w_{j,h}\|_{\infty;\Omega_\delta} \leq c_\delta |\log(h)|^2 h^2, \quad j = 1, \ldots, n_p, \tag{4.15}$$

which implies the desired estimate (4.10).

(iii) The proof of (4.11) uses the same line of argument as before. Using the additional discrete variable $\bar{v}_{jk,h}$ determined by the equation

$$\begin{aligned} a(q)(\bar{v}_{jk,h}, \phi_h) = &-(A'_{q_j}(q)\nabla w_k, \nabla \phi_h) - (A'_{q_k}(q)\nabla w_j, \nabla \phi_h) \\ &- (A''_{q_j q_k}(q)\nabla u, \nabla \phi_h) \quad \forall \phi_h \in V_h, \end{aligned}$$

the error $e = v_{jk} - v_{jk,h}$ is split like $e = e_1 + e_2$, with $e_1 = v_{jk} - \bar{v}_{jk,h}$ and $e_2 = \bar{v}_{jk,h} - v_{jk,h}$. For the Ritz-projection error e_1, as before, we conclude the pointwise error estimate

$$\|e_1\|_{\infty;\Omega_\delta} \le c_\delta h^2\big\{|\log(h)|\,\|\nabla^2 v_{jk}\|_{\infty;\Omega_{\delta/2}} + \|v_{jk}\|_{2,2}\big\} \le c_\delta h^2|\log(h)|.$$

For $e_2 \in V_h$, we have

$$\begin{aligned} a(q)(e_2,\phi_h) = &-(A'_{q_j}(q)\nabla(w_k - w_{k,h}),\nabla\phi_h) - (A'_{q_k}(q)\nabla(w_j - w_{j,h}),\nabla\phi_h) \\ &-(A''_{q_j q_k}(q)\nabla(u - \bar{u}_h),\nabla\phi_h) \quad \forall\phi_h \in V_h, \end{aligned}$$

and therefore, again by the L^∞-stability estimate (4.5) of Theorem 4.2,

$$\begin{aligned} \|e_2\|_{\infty;\Omega_\delta} \le c_\delta|\log(h)| &\Big\{ \max_{j=1,\dots,n_p} \|w_j - w_{j,h}\|_{\infty;\Omega_{\delta/2}} + \|u - \bar{u}_h\|_{\infty;\Omega_{\delta/2}} \Big\} \\ &+ c \max_{j=1,\dots,n_p} \big\{\|w_j - w_{j,h}\|_2 + h\|\nabla(w_j - w_{j,h})\|_2\big\} \\ &+ c\big\{\|u - \bar{u}_h\|_2 + h\|\nabla(u - \bar{u}_h)\|_2\big\}. \end{aligned}$$

Then, by the foregoing error estimates, we obtain $\|e_2\|_{\infty;\Omega_\delta} \le c_\delta h^2|\log(h)|^3$, and consequently,

$$\|v_{jk} - v_{jk,h}\|_{\infty;\Omega_\delta} \le c_\delta|\log(h)|^3 h^2, \quad j,k = 1,\dots,n_p. \tag{4.16}$$

This eventually yields the desired estimated (4.11). □

A direct application of Lemma 4.3 leads to the following result.

LEMMA 4.4. *Under the above assumptions, there holds*

$$\Big|\frac{\partial}{\partial q_j}(j - j_h)(q)\Big| \le c_\delta h^2|\log(h)|^2, \quad j = 1,2,\dots,n_p, \tag{4.17}$$

$$\Big|\frac{\partial^2}{\partial q_j \partial q_k}(j - j_h)(q)\Big| \le c_\delta h^2|\log(h)|^3, \quad j,k = 1,2,\dots,n_p. \tag{4.18}$$

Proof. We have the representation

$$\begin{aligned} \frac{\partial}{\partial q_j}(j - j_h)(q) &= \langle C(u) - \hat{C}, C(w_j)\rangle - \langle C(\bar{u}_h) - \hat{C}, C(w_{j,h})\rangle \\ &= \langle C(u) - \hat{C}, C(w_j - w_{j,h})\rangle + \langle C(u - \bar{u}_h), C(w_{j,h})\rangle, \end{aligned}$$

from which we obtain

$$\Big|\frac{\partial}{\partial q_j}(j - j_h)(q)\Big| \le \|C(u) - \hat{C}\|\,\|C(w_j - w_{j,h})\| + \|C(u - \bar{u}_h)\|\,\|C(w_{j,h})\|.$$

By the a priori bounds (2.16) and (2.17) and the Sobolev embedding theorem, we see that

$$\|C(u)\| + \|C(w_j)\| + \|C(v_{jk})\| \le c. \tag{4.19}$$

Combining this with the error estimate (4.10) implies $\|C(w_{j,h})\| \le c$. Then, we can conclude the first estimate (4.17) from the error estimates of Lemma 4.3. To prove (4.18), we write

$$\begin{aligned}\frac{\partial^2}{\partial q_j q_k}(j-j_h)(q) &= \langle C(w_j), C(w_k)\rangle + \langle C(u)-\hat{C}, C(v_{jk})\rangle \\ &\quad - \langle C(w_{j,h}), C(w_{k,h})\rangle - \langle C(\bar{u}_h)-\hat{C}, C(v_{jk,h})\rangle \\ &= \langle C(w_j - w_{j,h}), C(w_k)\rangle + \langle C(w_{j,h}), C(w_k - w_{k,h})\rangle \\ &\quad + \langle C(u-\bar{u}_h), C(v_{jk})\rangle + \langle C(\bar{u}_h)-\hat{C}, C(v_{jk}-v_{jk,h})\rangle.\end{aligned}$$

Using as before the bounds (4.19) and the error estimates of Lemma 4.3 completes the proof. ◻

For the application of Theorem 3.1 it remains to check the Lipschitz condition (3.2). For two arbitrary parameter sets $\xi, \eta \in Q_0$, we set $u_\xi = S_h(\xi)$ and $u_\eta = S_h(\eta)$. Correspondingly, we define $w_{j,\xi}, w_{j,\eta} \in V_h$ and $v_{jk,\xi}, v_{jk,\eta} \in V_h$ similarly to $w_{j,h}$ and $v_{j,h}$ for $q=\xi$ and $q=\eta$, respectively.

LEMMA 4.5. *For $\xi, \eta \in Q_0$, there hold*

$$\|C(u_\xi - u_\eta)\| \le c_\delta |\log(h)|\, \|\xi-\eta\|, \tag{4.20}$$

$$\|C(w_{j,\xi} - w_{j,\eta})\| \le c_\delta |\log(h)|^2\, \|\xi-\eta\|, \tag{4.21}$$

$$\|C(v_{jk,\xi} - v_{jk,\eta})\| \le c_\delta |\log(h)|^3\, \|\xi-\eta\|. \tag{4.22}$$

Proof. Due to the definition of u_ξ and u_η, we have

$$(A(\xi)\nabla(u_\xi - u_\eta), \nabla\phi_h) = -((A(\xi)-A(\eta))\nabla u_\eta, \nabla\phi_h) \quad \forall \phi_h \in V_h.$$

Using Theorem 4.2, with $d=\delta$, we obtain

$$\|C(u_\xi - u_\eta)\| \le c\, |||(A(\xi)-A(\eta)|||_{1,\infty} \left\{ |\log(h)|\, \|u_\eta\|_{\infty;\Omega_{\delta/2}} + \|u_\eta\|_2 + h\|\nabla u_\eta\|_2 \right\}.$$

Since u_η is the Ritz projection of an H^2 function, all its norms occurring on the right-hand side are bounded independent of h and $\eta \in Q_0$ by standard estimates from finite element analysis. This implies (4.20) since

$$|||A(\xi)-A(\eta)|||_{1,\infty} \le c\|\xi-\eta\|.$$

The estimates (4.21) and (4.22) are obtained in a similar way. ◻

LEMMA 4.6. *For $\xi, \eta \in Q_0$, there holds*

$$\left| \frac{\partial^2}{\partial q_j q_k} j_h(\xi) - \frac{\partial^2}{\partial q_j q_k} j_h(\eta) \right| \le L(h)\|\xi-\eta\|, \tag{4.23}$$

where $L(h) = c_\delta |\log(h)|^3$.

Proof. We have

$$\begin{aligned}\frac{\partial^2}{\partial q_j q_k} j_h(\xi) - \frac{\partial^2}{\partial q_j q_k} j_h(\eta) &= \langle C(w_{j,\xi}), C(w_{k,\xi})\rangle + \langle C(u_\xi)-\hat{C}, C(v_{jk,\xi})\rangle \\ &\quad - \langle C(w_{j,\eta}), C(w_{k,\eta})\rangle - \langle C(u_\eta)-\hat{C}, C(v_{jk,\eta})\rangle \\ &= \langle C(w_{j,\xi} - w_{j,\eta}), C(w_{k,\xi})\rangle + \langle C(w_{j,\eta}), C(w_{k,\xi} - w_{k,\eta})\rangle \\ &\quad + \langle C(u_\xi - u_\eta), C(v_{jk,\xi})\rangle + \langle C(u_\eta)-\hat{C}, C(v_{jk,\xi} - v_{jk,\eta})\rangle,\end{aligned}$$

and, consequently,

$$\begin{aligned}\Big|\frac{\partial^2}{\partial q_j q_k} j_h(\xi) - \frac{\partial^2}{\partial q_j q_k} j_h(\eta)\Big| &\le \|C(w_{j,\xi} - w_{j,\eta})\| \, \|C(w_{k,\xi})\| \\ &+ \|C(w_{j,\eta})\| \, \|C(w_{k,\xi} - w_{k,\eta})\| + \|C(u_\xi - u_\eta)\| \, \|C(v_{jk,\xi})\| \\ &+ \|C(u_\eta) - \hat{C}\| \, \|C(v_{jk,\xi} - v_{jk,\eta})\|.\end{aligned}$$

Now, the assertion follows by the estimates of Lemma 4.5 if we can bound the terms $C(u_\eta)$, $C(w_{j,\eta})$, and $C(v_{jk,\xi})$. This is achieved by using the bounds for $C(u)$, $C(w_j)$, and $C(v_{jk})$ in (4.19) together with the error estimates of Lemma 4.3. □

To complete the proof of Theorem 4.1 we check the conditions of Theorem 3.1. Condition (3.1) is fulfilled due to the stability of the solution q of the problem (2.5). Condition (3.2) is shown in Lemma 4.6. Condition (3.3) is obtained by Lemma 4.4 and Lemma 4.6 using $\lim_{h\to 0} h^2|\log(h)|^5 = 0$. Finally, condition (3.4) holds due to Lemma 4.4. Hence, the estimate (3.5) of Theorem 3.1 completes the proof.

5. Proof of Theorem 4.2. (i) We begin with the L^2-stability estimate. Taking $\phi_h := v_h$ in (4.3), we obtain

$$\|\nabla v_h\|_2 \le c|||B|||_{1,\infty}\|\nabla\psi\|_2. \tag{5.1}$$

To estimate $\|v_h\|_2$, we use the solution $z \in V \cap H^2(\Omega)$ of the auxiliary equation

$$a(q)(\phi, z) = (\phi, v_h)\|v_h\|_2^{-1} \quad \forall \phi \in V.$$

Taking $\phi := v_h$ as test function and integrating by parts, we have

$$\begin{aligned}\|v_h\|_2 &= a(q)(v_h, z) = a(q)(v_h, z - i_h z) + a(q)(v_h, i_h z) \\ &= a(q)(v_h, z - i_h z) + (B\nabla\psi, \nabla i_h z) \\ &= a(q)(v_h, z - i_h z) + (B\nabla\psi, \nabla(i_h z - z)) - (\psi, \nabla\cdot B^T\nabla z).\end{aligned}$$

Then, using the approximation properties (4.2) of the interpolant $i_h z \in V_h$, we conclude that

$$\begin{aligned}\|v_h\|_2 &\le c\|\nabla v_h\|_2\|\nabla(z - i_h z)\|_2 + |||B|||_{1,\infty}\|\nabla\psi\|_2\|\nabla(z - i_h z)\|_2 \\ &\quad + |||B|||_{1,\infty}\|\psi\|_2\|z\|_{2,2} \\ &\le c\big\{h\|\nabla v_h\| + |||B|||_{1,\infty} h\|\nabla\psi\|_2 + |||B|||_{1,\infty}\|\psi\|_2\big\}\,\|z\|_{2,2}.\end{aligned}$$

Hence, observing (5.1) and the bound $\|z\|_{2,2} \le c$, we obtain

$$\|v_h\|_2 \le c|||B|||_{1,\infty}\big\{\|\psi\|_2 + h\|\nabla\psi\|_2\big\}. \tag{5.2}$$

(ii) Next, we prove the L^∞-stability estimate. Let $a \in \Omega_\delta$ be an arbitrary point lying in a cell K. For any fixed h, there exists a cellwise polynomial function δ_h with $\mathrm{supp}(\delta_h) \subset K$ such that

$$(\phi_h, \delta_h) = \phi_h(a) \quad \forall \phi_h \in V_h.$$

The function δ_h plays the role of an approximate Dirac function. Correspondingly, we introduce a regularized Green function $g \in V \cap H^2(\Omega)$ by

$$(A(q)\nabla\phi, \nabla g) = (\delta_h, \phi) \quad \forall \phi \in V,$$

and the corresponding Ritz projection $g_h \in V_h$ by

$$(A(q)\nabla\phi_h, \nabla g_h) = (\delta_h, \phi_h) \quad \forall \phi_h \in V_h.$$

For functions which are only cellwise defined, we will use the "broken" norm $\|v\|'_p := \sum_{K\in\mathcal{T}_h} \|v\|_{p;T}$. The following three lemmas provide the key estimates for the proof of the theorem.

LEMMA 5.1. *The following global L^2 estimates hold:*

$$\|g\|_2 + |\log(h)|^{-1/2}\|\nabla g\|_2 + h\|\nabla^2 g\|_2 \le c, \tag{5.3}$$

$$h^{-1}\|g - g_h\|_2 + \|\nabla(g - g_h)\|_2 + h\|\nabla^2 g_h\|'_2 \le c. \tag{5.4}$$

Proof. The assertion follows by standard L^2 a priori and error estimates for g and $g - g_h$, respectively. We skip the details and refer to [10]. Note that $\|\nabla^2 g_h\|'_2$ vanishes for linear finite elements. In the case of bilinear elements, we estimate using the interpolant $i_h g$ as follows:

$$\|\nabla^2 g_h\|'_2 \le \|\nabla^2(g_h - i_h g)\|'_2 + \|\nabla^2(g - i_h g)\|'_2 + \|\nabla^2 g\|_2.$$

For the first term, we obtain, using an inverse inequality,

$$\begin{aligned}\|\nabla^2(g_h - i_h g)\|'_2 &\le ch^{-1}\|\nabla(g_h - i_h g)\|_2\\ &\le ch^{-1}\big\{\|\nabla(g - i_h g)\|_2 + \|\nabla(g - g_h)\|_2\big\}\end{aligned}$$

and obtain by the interpolation estimate (4.2), with $p = 2$, and by the other estimates derived before,

$$\|\nabla^2 g_h\|'_2 \le c\big\{h^{-1}\|\nabla(g - g_h)\|_2 + \|\nabla^2 g\|_2\big\} \le ch^{-1}.$$

This completes the proof. ∎

LEMMA 5.2. *For sufficiently small $h \ll \delta$, the following local L^2 estimate holds:*

$$\|\nabla(g - g_h)\|_{2;\Omega\setminus\Omega_{\delta/2}} + h\|\nabla^2 g\|_{2;\Omega\setminus\Omega_{\delta/2}} \le c_\delta h, \tag{5.5}$$

with a constant $c_\delta \approx \delta^{-1}$ but independent of h.

Proof. The assertion follows by standard local elliptic a priori estimates and by arguments from the local L^2 error analysis for finite elements, as provided in Nitsche and Schatz [20]:

$$\begin{aligned}\|\nabla^2 g\|_{2;\Omega\setminus\Omega_{\delta/2}} &\le c\|\Delta g\|_{2;\Omega\setminus\Omega_{3\delta/4}} + c_\delta\|g\|_2,\\ \|\nabla(g - g_h)\|_{2;\Omega\setminus\Omega_{\delta/2}} &\le c\|\nabla(g - i_h g)\|_{2;\Omega\setminus\Omega_{3\delta/4}} + c_\delta\|g - g_h\|_2,\end{aligned}$$

with constants $c_\delta \approx \delta^{-1}$. Now, the assertion follows by the interpolation estimate (4.2) and the other estimates already proven. ∎

LEMMA 5.3. *The following L^1 a priori and error estimates hold:*

$$\|\nabla g\|_1 + \|\nabla^2 g\|'_1 \le c|\log(h)|, \tag{5.6}$$

$$\|\nabla(g - g_h)\|_1 + h\|\nabla^2(g - g_h)\|'_1 \le ch|\log(h)|, \tag{5.7}$$

with a constant c independent of h and δ.

Proof. The proof can be found in [10]. ∎

For the point $a \in \Omega_d$, there holds

$$v_h(a) = (v_h, \delta_h) = (A(q)\nabla v_h, \nabla g_h) = (B\nabla\psi, \nabla g_h).$$

We employ a standard localization argument. Let $\omega \in C_0^\infty(\Omega)$ be a smooth function with the properties

$$0 \le \omega \le 1, \quad \omega_{|\Omega_{\delta/2}} \equiv 1, \quad \omega_{|\Omega\setminus\Omega_{\delta/4}} \equiv 0.$$

With this notation, we have

$$(B\nabla\psi, \nabla g_h) = (B\nabla(\omega\psi), \nabla g_h) + (B\nabla((1-\omega)\psi), \nabla g_h) =: \Sigma_1 + \Sigma_2.$$

First, we estimate the term Σ_1. By integration by parts and observing that $\psi_{|\partial\Omega} = 0$, we obtain

$$(B\nabla(\omega\psi), \nabla g_h) = \sum_{K\in\mathcal{T}_h} \big\{(\omega\psi, -\nabla\cdot(B\nabla g_h))_T + (\omega\psi, n\cdot B\nabla g_h)_{\partial K\setminus\partial\Omega}\big\},$$

where n is the outward unit normal vector to ∂K. Let $[\nabla g_h]$ denote the jump of the gradient across the interior faces $\Gamma \subset \partial K$. Using this notation, we obtain

$$\Sigma_1 \le \|\psi\|_{\infty;\Omega_{\delta/2}} \sum_{K\in\mathcal{T}_h} \left\{ \|\nabla\cdot(B\nabla g_h)\|_{1;K} + \frac{1}{2}\|n\cdot[B\nabla g_h]\|_{1;\partial K\setminus\partial\Omega} \right\}.$$

First, the estimates of Lemma 5.3 yield

$$\sum_{K\in\mathcal{T}_h} \|\nabla\cdot(B\nabla g_h)\|_{1;K} \le c\,|||B|||_{1,\infty}\left\{\|\nabla g_h\|_1 + \|\nabla^2 g_h\|_1'\right\} \le c\,|||B|||_{1,\infty}\,|\log(h)|.$$

Next, observing that $g \in H^2(\Omega)$ and therefore $[B\nabla g] = 0$, we obtain by a trace theorem

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} \|n\cdot[B\nabla g_h]\|_{1;K\setminus\partial\Omega} &= \sum_{K\in\mathcal{T}_h} \|n\cdot[B\nabla(g_h - g)]\|_{1;\partial K\setminus\partial\Omega} \\ &\le c\,|||B|||_{1,\infty} \sum_{K\in\mathcal{T}_h} \big\{h^{-1}\|\nabla(g-g_h)\|_{1;K} + \|\nabla^2(g-g_h)\|_{1;K}'\big\}.\end{aligned}$$

Hence, collecting the foregoing estimates,

$$\Sigma_1 \le c\,|||B|||_{1,\infty}\,\|\psi\|_{\infty;\Omega_{\delta/2}}\big\{|\log(h)| + h^{-1}\|\nabla(g-g_h)\|_1 + \|\nabla^2(g-g_h)\|_1'\big\}.$$

Again using the estimates of Lemma 5.3, we obtain

$$\Sigma_1 \le c\,|||B|||_{1,\infty}\,\|\psi\|_{\infty;\Omega_{\delta/2}}|\log(h)|. \tag{5.8}$$

For the term Σ_2, we estimate as follows:

$$\begin{aligned}\Sigma_2 &= (B\nabla((1-\omega)\psi), \nabla(g_h - g)) + (B\nabla((1-\omega)\psi), \nabla g) \\ &\le |||B|||_{1,\infty}\big\{\|\nabla\psi\|_2\,\|\nabla(g_h-g)\|_{2;\Omega\setminus\Omega_{\delta/2}} + c_\delta\|\psi\|_2\,\|\nabla(g_h-g)\|_2\big\} \\ &\quad + c\,|||B|||_{1,\infty}\|\psi\|_2\left\{\|\nabla^2 g\|_{2;\Omega\setminus\Omega_{\delta/2}} + \|\nabla g\|_{2;\Omega\setminus\Omega_{\delta/2}}\right\}.\end{aligned}$$

Then, by the L^2 estimates of Lemmas 5.1 and 5.2, it follows that

$$\Sigma_2 \leq c_\delta |||B|||_{1,\infty} \{ \|\psi\|_2 + h \|\nabla \psi\|_2 \}. \tag{5.9}$$

This completes the proof of the theorem.

6. Numerical results. In this section, we discuss a sample problem confirming the a priori error estimate of Theorem 4.1. The state equation is given by

$$\begin{aligned} -\nabla \cdot (A(q) \nabla u) &= 2 \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{6.1}$$

where Ω is the unit square. The matrix $A(q)$ is a function of the parameter $q = (q_1, q_2) \in Q = \mathbb{R}^2$, given by

$$A(q) = \begin{pmatrix} q_1^2 & q_1 q_2 \\ q_1 q_2 & \exp(q_2) \end{pmatrix}.$$

In this case the admissible set of parameters is $Q_0 = \{(q_1, q_2) \in Q : q_1 \neq 0, e^{q_2} > q_2^2\}$. The parameters are estimated from the measurements of the state variable at nine different points $\xi_i \in \Omega$; see Figure 6.1.

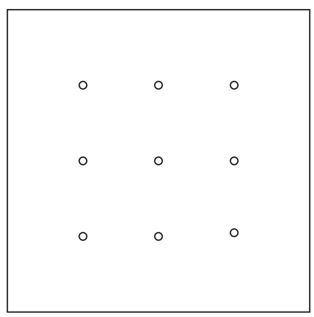

Fig. 6.1. *The computational domain with measurement points marked by circles.*

The vector of measurements $\hat{C}$ is given by

$$\hat{C}_i = C_i(S(\hat{q}))(1 + \varepsilon_i), \quad i = 1, \dots, 9,$$

where the reference parameter is $\hat{q} = (5, 6)$ and $\varepsilon = (\varepsilon_i)$ describes the data perturbation. We consider two cases:

(a) $\varepsilon \approx 0$, (b) $\varepsilon \approx (0.12, -0.26, 0.29, -0.37, -0.49, 0.13, -0.04, -0.45, 0.20)$.

Since the values of $C_i(S(\hat{q}))$ are not available analytically, they are computed approximately by solving state equation (6.1) on a very fine mesh with about 10^6 degrees of freedom. For case (a) the solution $q^{(a)}$ of the parameter identification problem matches the reference parameter $\hat{q}$, and hence the cost functional $J(u)$ almost vanishes in $q^{(a)}$. Case (b) is more realistic because of the "measurement errors" modeled by a randomly chosen ε. Moreover, in this case in contrast to case (a), the solution $q^{(b)}$ of the corresponding parameter identification problem and the reference parameter $\hat{q}$ differ.

The parameter identification problem is discretized using bilinear finite elements on uniformly refined meshes. The results are listed in Tables 6.1 and 6.2. For both cases the theoretically predicted orders of convergence are achieved.

Table 6.1

Case (a)*: the error and the order of convergence with respect to the components of q without data perturbation; $N \sim h^{-2}$ number of unknowns.*

N	$q_1^{(a)} - q_{1,h}^{(a)}$	$q_2^{(a)} - q_{2,h}^{(a)}$
81	5.955e-1	9.902e-4
289	1.407e-1	1.731e-4
1089	3.436e-2	4.343e-5
4225	8.509e-3	1.080e-5
16641	2.098e-3	2.668e-6
66049	4.993e-4	6.352e-7
Order	2.05	2.04

Table 6.2

Case (b)*: the error and the order of convergence with respect to the components of q with data perturbation, $N \sim h^{-2}$ number of unknowns.*

N	$q_1^{(b)} - q_{1,h}^{(b)}$	$q_2^{(b)} - q_{2,h}^{(b)}$
81	2.059e-0	1.874e-2
289	5.172e-1	2.999e-3
1089	1.467e-1	8.341e-4
4225	3.771e-2	2.111e-4
16641	9.832e-3	5.640e-5
66049	2.350e-3	1.348e-5
Order	2.01	1.98

7. Conclusions and extensions. In this paper we have derived an a priori error estimate for the finite element discretization of an elliptic discrete parameter identification problem with pointwise measurements. The crucial point in our argument is the stability estimate of Theorem 4.2. The result of Theorem 4.1 can be extended to situations in which such a stability estimate is available. We list some possible directions of generalization.

1. *More general meshes.* For simplicity, we have assumed a quasi-uniform mesh family $\{\mathcal{T}_h\}_h$. The analysis can be extended to locally refined meshes, provided that the ratio of $h_{\min}$ and $h = h_{\max}$ is polynomial, $h_{\min} \approx h^p$, with some $p \geq 1$. For such meshes the stability result of Theorem 4.2 holds true with $|\log(h_{\min})| \approx p|\log(h)|$. This will be shown in the forthcoming paper [23]. Related results for L^∞-error estimates can be extracted (with some work) from Schatz and Wahlbin [25].

2. *More general domains.* Our argument uses that the solution operator $S(\cdot)$ maps Q into $H_0^1(\Omega) \cap H^2(\Omega)$, which is guaranteed on smoothly bounded or convex domains. In the case of a domain with reentrant corners or edges this regularity property is lost. This lack of regularity of the solution can be compensated by an appropriate refinement of the mesh near the critical corner points or edges. The stability estimate of Theorem 4.2 also holds in this situation, in two as well as in three dimensions. This will be shown in a forthcoming paper.

3. *Higher-order approximation.* The result of Theorem 4.1 can be also extended to the case of higher-order finite elements, similar to the analysis of Nitsche [19]. In this case the logarithmic factor $|\log(h)|$ can be dropped in the stability Theorem 4.2.

4. *More general equations.* Theorem 4.1 can also be extended to more general elliptic equations or systems of the form

$$-\nabla \cdot (A(q)\nabla u + b_1(q)u) + b_2(q)u = f,$$

with parameter-dependent coefficients $A(q)$, $b_1(q)$, $b_2(q)$. Corresponding L^∞-error estimates for very general (nonlinear) elliptic systems have been derived in Dobrowolski and Rannacher [7].

REFERENCES

[1] N. Arada, E. Casas, and F. Tröltzsch, *Error estimates for the numerical approximation of a semilinear elliptic control problem*, J. Comput. Optim. Appl., 23 (2002), pp. 201–229.

[2] H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser Boston, Cambridge, MA, 1989.

[3] R. Becker, M. Braack, and B. Vexler, *Numerical parameter estimation for chemical models in multidimensional reactive flows*, Combust. Theory Modelling, 8 (2004), pp. 661–682.

[4] R. Becker and B. Vexler, *A posteriori error estimation for finite element discretization of parameter identification problems*, Numer. Math., 96 (2004), pp. 435–459.

[5] S. Brenner and R. L. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, Berlin, Heidelberg, New York, 1994.

[6] K. Deckelnick and M. Hinze, *Error estimates in space and time for tracking-type control of the instationary Stokes system*, in Proceedings of the 8th Conference on Control of Distributed Parameter Systems, Graz, Austria, 2001, Internat. Ser. Numer. Math., 143, Birkhäuser, Basel, Switzerland, 2002, pp. 87–104.

[7] M. Dobrowolski and R. Rannacher, *Finite element methods for nonlinear elliptic problems of second order*, Math. Nachr., 94 (1980), pp. 155–172.

[8] R. S. Falk, *Approximation of a class of optimal control problems with order of convergence estimates*, J. Math. Anal. Appl., 44 (1973), pp. 28–47.

[9] R. S. Falk, *Error estimates of the numerical identification of a variable coefficient*, Math. Comp., 40 (1983), pp. 537–546.

[10] J. Frehse and R. Rannacher, *Eine L^1-Fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente*, in Tagungsband Finite Elemente, Bonner Math. Schrifter 89, University of Bonn, Bonn, Germany, 1976, pp. 92–114.

[11] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd ed. Springer, Berlin-Heidelberg, 1983.

[12] P. Grisvard, *Singularities in Boundary Value Problems*, Masson, Paris, and Springer, Berlin, 1992.

[13] M. D. Gunzburger and L. S. Hou, *Finite-dimensional approximation of a class of constrained nonlinear optimal control problems*, SIAM J. Control Optim., 34 (1996), pp. 1001–1043.

[14] C. Johnson, *Numerical Solution of Partial Differential Equations by Finite Element Method*, Cambridge University Press, Cambridge, UK, 1987.

[15] T. Kärkkäinen, *Error estimates for distributed parameter identification in linear elliptic equations*, J. Math. Systems Estim. Control, 6 (1996), pp. 117–120.

[16] C. Kravaris and J. H. Seinfeld, *Identification of parameters in distributed parameter systems by regularization*, SIAM J. Control Optim., 23 (1985), pp. 217–241.

[17] V. G. Litvinov, *Optimization in Elliptic Problems with Applications to Mechanics of Deformable Bodies and Fluid Mechanics*, Oper. Theory Adv. Appl. 119, Birkhäuser, Basel, Switzerland, 2000.

[18] P. Neittaanmäki and X.-C. Tai, *Error estimates for numerical identification of distributed parameters*, J. Comput. Math., 10 (1992), pp. 66–78.

[19] J. Nitsche, *Über L^∞-Abschätzungen von Projektoren auf Finite Elemente*, in Tagungsband Finite Elemente, Bonner Math. Schrifter 89, University of Bonn, Bonn, Germany, 1976, pp. 13–30.

[20] J. Nitsche and A. Schatz, *Interior estimates for Ritz–Galerkin methods*, Math. Comput., 28 (1976), pp. 937–958.

[21] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Ser. Oper. Res., Springer, New York, 1999.

[22] R. Rannacher, *L^∞-stability estimates and asymptotic error expansion for parabolic finite element equations*, in Tagungsband Extrapolation and Defect Correction, Bonner Math. Schrifter 228, University of Bonn, Bonn, Germany, 1991, pp. 74–94.

[23] R. Rannacher, *On L^∞-Stability in the Finite Element Method on Meshes without Uniform Size and Shape Property*, Preprint SFB 359, Institute of Applied Mathematics, University of Heidelberg, Heidelberg, Germany, 2005.

[24] R. Rannacher and L. R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp., 38 (1982), pp. 437–445.

[25] A. H. Schatz and L. B. Wahlbin, *Maximum norm estimates in the finite element method on plane polygonal domains. Part* I, Math. Comp., 32 (1978), pp. 73–109.

[26] B. Vexler, *Adaptive Finite Element Methods for Parameter Identification Problems*, Ph.D. thesis, Institute of Applied Mathematics, University of Heidelberg, Heidelberg, Germany, 2004.