

## ADAPTIVE FINITE ELEMENT METHODS FOR PARAMETER ESTIMATION PROBLEMS IN LINEAR ELASTICITY

TAO FENG, MÅRTEN GULLIKSSON, AND WENBIN LIU

**Abstract.** In this paper, the Lamé coefficients in the linear elasticity problem are estimated by using the measurements of displacement. Some a posteriori error estimators for the approximation error of the parameters are derived, and then adaptive finite element schemes are developed for the discretization of the parameter estimation problem, based on the error estimators. The Gauss-Newton method is employed to solve the discretized nonlinear least-squares problem. Some numerical results are presented.

**Key Words.** parameter estimation, finite element approximation, adaptive finite element methods, a posteriori error estimates, linear elasticity.

### 1. Introduction

In this paper, we consider a parameter identification problem in the linear elasticity problem

$$(1) \quad \begin{aligned} -\mu\Delta\mathbf{u} - (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) &= \mathbf{f} && \text{in } \Omega, \\ \mathbf{u} &= \mathbf{u}_D && \text{on } \Gamma_D \end{aligned}$$

where  $\Omega$  is a polygonal domain in two dimensional space with the Lipschitz-continuous boundary, and the boundary  $\Gamma_D$  is positive  $d\gamma$ -measurable. As usual,  $\mathbf{u}$  denotes the displacement, and  $\mathbf{f}$  and  $\mathbf{u}_D$  represent the body force and the boundary displacement, respectively. Let  $U = \{\mathbf{u} \in (H_0^1(\Omega))^2\}$  and  $\mathbf{u} \in U + \mathbf{u}_D = Y$ , where the space  $U$  is assumed to have the product norm

$$\mathbf{u} = (u_1, u_2) \rightarrow \|\mathbf{u}\|_{1,\Omega} = \left( \sum_{i=1}^2 \|u_i\|_{1,\Omega}^2 \right)^{\frac{1}{2}}.$$

We define the strain tensor  $(\epsilon_{ij}(\mathbf{u}))$  as

$$\epsilon_{ij}(\mathbf{u}) = \epsilon_{ji}(\mathbf{u}) = \frac{1}{2}(\partial_j u_i + \partial_i u_j), \quad 1 \leq i, j \leq 2,$$

and the stress tensor  $(\sigma_{ij})$  is then given by Hooke's law for isotropic bodies

$$\sigma_{ij}(\mathbf{u}) = \sigma_{ji}(\mathbf{u}) = \lambda \left( \sum_{k=1}^2 \epsilon_{kk}(\mathbf{u}) \right) \delta_{ij} + 2\mu \epsilon_{ij}(\mathbf{u}), \quad 1 \leq i, j \leq 2,$$

where  $\delta_{ij}$  is Kronecker's symbol. The Lamé coefficients  $\lambda$  and  $\mu$  are given by

$$\mu = \frac{E}{2(1 + \nu)}, \quad \lambda = \frac{E\nu}{(1 + \nu)(1 - 2\nu)}$$

with Poisson's ratio  $\nu$  and Young's modulus  $E$ . It is well known that  $\lambda \geq \lambda' > 0$  and  $\mu \geq \mu' > 0$ .

---

Received by the editors November 28, 2007 and, in revised form, December 22, 2007.  
 2000 *Mathematics Subject Classification.* 65N30, 49J20, 74B05, 65M32.

In our parameter estimation problem, we aim to recover the constants  $\lambda$  and  $\mu$  by using the known measurements of displacement  $\mathbf{u}$ . To this end, the well-known output least-squares formulation is used, i.e., we solve

$$(2) \quad \min_{\mathbf{m}} \frac{1}{2} \|Q\mathbf{u}(\mathbf{m}) - \mathbf{z}\|_Z^2,$$

where  $\mathbf{u}$  is the solution of the linear elastic equation (1) and  $\mathbf{m} = (m_1, m_2)^T = (\lambda, \mu)^T$ . The vector  $\mathbf{z} \in Z$  is a given set of measurements and the observation space  $Z$  is supposed to be a Hilbert space. Furthermore, we set  $Q : Y \rightarrow Z$  as a linear bounded observation operator.

Usually, the parameter estimation problem is ill-posed or ill-conditioned; see [16]. Some regularization terms are added to the cost function (2) such that

$$(3) \quad \min_{\mathbf{m}} \left\{ \frac{1}{2} \|Q\mathbf{u}(\mathbf{m}) - \mathbf{z}\|_Z^2 + \frac{\beta}{2} \|\mathbf{m} - \mathbf{m}_{ref}\|^2 \right\},$$

where the penalty parameter  $\beta$  is assumed to be a very small positive number and  $\mathbf{m}_{ref}$  is a reference model. The regularization term  $\beta/2 \cdot \|\mathbf{m} - \mathbf{m}_{ref}\|^2$  improves the conditioning of the inverse problem. Here  $\|\cdot\|$  denotes the  $l^2$  norm of the vector. A good regularization parameter  $\beta$  should yield a fair balance between the perturbation error and the regularization error. Assume that the data  $\mathbf{z}$  contains noise with known standard deviation  $\mathbf{e}$ , then the regularization parameter should be chosen such that

$$\|Q\mathbf{u}(\mathbf{m}) - \mathbf{z}\|_Z = \|\mathbf{e}\|_Z;$$

see [23]. To solve the problems without known deviation, methods such as L-curve criterion, generalized cross-validation and the quasi-optimality criterion can be used for the regularization parameter selection; for more details, see [15, 23, 29].

To solve the parameter estimation problem, one must approximate the infinite-dimensional problem by introducing discretizations for the state space  $Y$  such as a finite element or difference approach. It is clear that the efficiency of our numerical methods will be influenced by the discretization scheme. In recent years, adaptive finite element method has been extensively and successfully investigated; see [1]. By using the adaptive finite element method, a numerical solution with a prescribed tolerance can be obtained with a minimal amount of work. This ensures a higher density of nodes in a certain area of the given domain, where the solution is more difficult to approximate. Although adaptive finite element approximation is widely used in the numerical simulation, it is not yet fully utilized in the parameter estimation problem. Very recently, some a posteriori error estimators have been derived for the parameter estimation problem [4, 8, 19]. In this paper, an adaptive finite element method for our parameter estimation problems is developed. Our emphasis here is to derive some a posteriori error estimators which control the error in the unknown parameters, instead of the cost function [4, 7]. Moreover, these error estimators are used to guide our mesh refinement.

We note that some efficient a posteriori error estimators have been derived by using the adjoint equation approach [4, 8]. In these error estimators, the local residuals of the solution are multiplied by weights which measure the dependence of the error on the local residuals. The weights are obtained by approximately solving an adjoint problem. However, the exact solution itself is included in these error estimators, which must be approximated by techniques such as higher order interpolation. Furthermore, since  $\lambda \geq \lambda' > 0$  and  $\mu \geq \mu' > 0$ , we get inequality constraints (see the optimal conditions (8) in section 2). In general, it is not clear how to apply the adjoint approach to this inequality constraint minimization problem. Thus, our corresponding error estimators are based on the approach developed by Kunisch, Liu and Yan [19]. In our error estimator, the weights are absorbed to

a simple constant  $C$  and the error estimator is the sum of the state and adjoint errors. Although some information is lost, it seems to be sufficient for guiding efficient mesh refinements.

The Levenberg-Marquardt method is employed to solve the optimization problem (2) on a fixed mesh; for details, see [13, 24, 27]. The algorithmic framework is given as follows:

1. Solve problem (2) on an initial coarse grid.
2. Refine the mesh gradually based on our error estimators. Then solve a sequence of the problem (2) on finer grids, using the solution of the most recently solved problem as the first iterate for the solution process of (2) on the next finer level.

This method is indeed a regularization and the penalty parameter  $\beta$  is updated adaptively [30].

The outline of this article is as follows. The parameter identification problem and its finite element discretization are described in section 2. In section 3, an a posteriori error estimator is developed for our parameter identification problem. Moreover, an a posteriori error estimator concerning the Neumann boundary condition is derived as an extension. Numerical experiments are presented in section 4.

## 2. The parameter estimation problem and its finite element approximation

Let  $\mathbf{f} = (f_1, f_2) \in (L^2(\Omega))^2$  and  $\mathbf{m} \in P = (R^+)^2$ , where  $R^+$  denotes the positive real number space. The weak formulation of (1) is given by the following: Find  $\mathbf{u} \in Y$  such that

$$(4) \quad a(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) \quad \forall \mathbf{v} \in U,$$

where

$$(5) \quad \begin{aligned} a(\mathbf{u}, \mathbf{v}) &= \int_{\Omega} \sum_{i,j=1}^2 \sigma_{ij}(\mathbf{u}) \epsilon_{ij}(\mathbf{v}) dx \\ &= \int_{\Omega} \{ \lambda \nabla \cdot \mathbf{u} \nabla \cdot \mathbf{v} + 2\mu \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{u}) \epsilon_{ij}(\mathbf{v}) \} dx \quad \forall \mathbf{v} \in U \end{aligned}$$

and

$$(6) \quad (\mathbf{f}, \mathbf{v}) = \int_{\Omega} \sum_{i=1}^2 f_i v_i \quad \forall \mathbf{v} \in U.$$

We define the composite function

$$(7) \quad s(\mathbf{m}) = \frac{1}{2} \| Q\mathbf{u}(\mathbf{m}) - \mathbf{z} \|_Z^2$$

and assume that  $s(\mathbf{m})$  is uniformly convex near the solution, at least when  $\mathbf{z}$  is attainable. Let us introduce the adjoint  $Q^*$  of  $Q$ . Then the optimal conditions for our optimization problem (2) (see, e.g., [22]) are to find  $(\mathbf{u}, \mathbf{p}, \mathbf{m}) \in Y \times U \times P$  such that

$$(8) \quad \begin{aligned} a(\mathbf{u}, \mathbf{v}) &= (\mathbf{f}, \mathbf{v}) & \forall \mathbf{v} \in U, \\ a(\mathbf{q}, \mathbf{p}) &= (Q^*(Q\mathbf{u} - \mathbf{z}), \mathbf{q}) & \forall \mathbf{q} \in U, \\ (s'(\mathbf{m}), \mathbf{m} - \mathbf{w}) &\leq 0 & \forall \mathbf{w} \in P, \end{aligned}$$

where

$$(s'(\mathbf{m}), \mathbf{m} - \mathbf{w}) = (-\tilde{\mathbf{p}}\tilde{\mathbf{u}}, \mathbf{m} - \mathbf{w}) = - \int_{\Omega} \sum_{i=1}^2 \tilde{p}_i \tilde{u}_i (m_i - w_i),$$

and we denote

$$(9) \quad \tilde{\mathbf{p}}\tilde{\mathbf{u}} = (\tilde{p}_1\tilde{u}_1, \tilde{p}_2\tilde{u}_2) = (\nabla \cdot \mathbf{u} \nabla \cdot \mathbf{p}, 2 \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{u})\epsilon_{ij}(\mathbf{p})).$$

Consider the finite element method for (8). Here, we are only interested in n-simplex elements and conforming finite elements. Let  $\Omega^h$  be a polygonal approximation of the problem (8). Let  $T^h$  be a partitioning of  $\Omega^h$  into disjoint regular n-simplex  $\tau$ , so that  $\bar{\Omega}^h = \cup_{\tau \in T^h} \bar{\tau}$ . Each element has at most one face on  $\partial\Omega^h$ ;  $\bar{\tau}$  and  $\bar{\tau}'$  have at most either one common vertex or a whole edge  $l$  if  $\bar{\tau}$  and  $\bar{\tau}' \in T^h$ . The maximum diameter of  $\tau$  is denoted by  $h_\tau$  and the mesh parameter  $h$  is defined as a cell-wise constant function by setting  $h|_\tau = h_\tau$ . We also require that  $E_i \in \partial\Omega^h \Rightarrow E_i \in \partial\Omega$  where  $\{E_i\}$  is the vertex set associated with the triangulation  $T^h$ . For simplicity, we assume that  $\Omega^h = \Omega$ , i.e.,  $\Omega$  is a convex polygon.

Associated with  $T^h$  is a finite element subspace  $W^h$  of  $C(\bar{\Omega}^h)$ , such that  $\varkappa|_\tau$  are polynomials of  $k$ -th order ( $k \geq 1$ ) for  $\varkappa \in W^h$  and  $\tau \in T^h$ . Let  $V^h = W^h \cap U$ , then it is easy to see that  $V^h \subset U$ .

The corresponding Galerkin solution  $(\mathbf{u}_h, \mathbf{p}_h, \mathbf{m}_h) \in Y^h \times V^h \times P$  of (8) is then given by

$$(10) \quad \begin{aligned} a_h(\mathbf{u}_h, \mathbf{v}_h) &= (\mathbf{f}, \mathbf{v}_h) & \forall \mathbf{v}_h \in V^h, \\ a_h(\mathbf{q}_h, \mathbf{p}_h) &= (Q^*(Q\mathbf{u}_h - \mathbf{z}), \mathbf{q}_h) & \forall \mathbf{q}_h \in V^h, \\ (s'_h(\mathbf{m}_h), \mathbf{m}_h - \mathbf{w}_h) &\leq 0 & \forall \mathbf{w}_h \in P, \end{aligned}$$

where  $(s'_h(\mathbf{m}_h), \mathbf{w}_h) = (-\tilde{\mathbf{p}}_h\tilde{\mathbf{u}}_h, \mathbf{w}_h)$ , and

$$a_h(\mathbf{u}_h, \mathbf{v}_h) = \int_{\Omega} \{ \lambda_h \operatorname{div} \mathbf{u}_h \operatorname{div} \mathbf{v}_h + 2\mu_h \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{u}_h)\epsilon_{ij}(\mathbf{v}_h) \} dx \quad \forall \mathbf{v}_h \in V^h.$$

In order to implement the local mesh refinement, we keep the information about the whole hierarchy of grids starting with the macro triangulation up to the actual one [17, 21]. Every element in this triangulation can be infinitely uniformly refined so that we can obtain an infinite hierarchy tree. Thus, every adaptive mesh is only a section of this hierarchy tree.

### 3. A posteriori error estimates

In this section, an a posteriori upper error estimator for our parameter estimation problem (2) is derived. Thereafter, we consider the case with Neumann boundary conditions and present the corresponding a posteriori error estimator.

We define the norm of  $\mathbf{t}$  as

$$|\mathbf{t}| = \left( \sum_{i,j=1}^2 |\epsilon_{ij}(\mathbf{t})|_{0,\Omega}^2 \right)^{\frac{1}{2}},$$

which is equivalent to the product norm [9]. A quasi-interpolation operator  $I_h$  is defined with the interpolation error estimates

$$(11) \quad \begin{aligned} \|\mathbf{v} - I_h \mathbf{v}\|_\tau &\leq ch_\tau \|\mathbf{v}\|_{1;\tilde{w}_\tau}, \\ \|\mathbf{v} - I_h \mathbf{v}\|_l &\leq ch_l^{\frac{1}{2}} \|\mathbf{v}\|_{1;\tilde{w}_l}, \end{aligned}$$

where  $\tilde{w}_\tau$  denotes the union of all elements sharing at least one point with  $\tau$  and  $\tilde{w}_l$  denotes the union of all elements sharing at least one point with  $l$ , and  $h_l$  is the

maximum diameter of the face  $l$  [28]. Further, we denote

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial \mathbf{n}} &= \left( \sum_{j=1}^2 \sigma_{1j}(\mathbf{u}) n_j, \sum_{j=1}^2 \sigma_{2j}(\mathbf{u}) n_j \right)^T, \\ \frac{\partial \mathbf{p}}{\partial \mathbf{n}} &= \left( \sum_{j=1}^2 \sigma_{1j}(\mathbf{p}) n_j, \sum_{j=1}^2 \sigma_{2j}(\mathbf{p}) n_j \right)^T\end{aligned}$$

and

$$(12) \quad \begin{aligned}\bar{\mathbf{p}} &= \left( \nabla \cdot \mathbf{p}, \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{p}) \right)^T = \left( \epsilon_{11}(\mathbf{p}) + \epsilon_{22}(\mathbf{p}), \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{p}) \right)^T, \\ \bar{\mathbf{u}} &= \left( \nabla \cdot \mathbf{u}, \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{u}) \right)^T = \left( \epsilon_{11}(\mathbf{u}) + \epsilon_{22}(\mathbf{u}), \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{u}) \right)^T.\end{aligned}$$

**Lemma 3.1.** *Let  $(\mathbf{u}, \mathbf{p}, \mathbf{m})$  and  $(\mathbf{u}_h, \mathbf{p}_h, \mathbf{m}_h)$  be the solution of (8) and (10), respectively. Then*

$$(13) \quad \|\mathbf{m} - \mathbf{m}_h\|^2 \leq C(\|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2 + \|\bar{\mathbf{p}}(\mathbf{m}_h) - \bar{\mathbf{p}}_h\|_{0,\Omega}^2).$$

*Proof.* Since  $s(\cdot)$  is uniformly convex near the solution  $\mathbf{m}$ , we have

$$(14) \quad (s'(\mathbf{m}_h), \mathbf{w}) = (-\tilde{\mathbf{p}}(\mathbf{m}_h) \tilde{\mathbf{u}}(\mathbf{m}_h), \mathbf{w}),$$

where  $(\mathbf{p}(\mathbf{m}_h), \mathbf{u}(\mathbf{m}_h))$  are the solutions of the following auxiliary equations

$$(15) \quad \begin{aligned}a_h(\mathbf{u}(\mathbf{m}_h), \mathbf{v}) &= (\mathbf{f}, \mathbf{v}) & \forall \mathbf{v} \in U, \\ a_h(\mathbf{q}, \mathbf{p}(\mathbf{m}_h)) &= (Q^*(Q\mathbf{u}(\mathbf{m}_h) - \mathbf{z}), \mathbf{q}) & \forall \mathbf{q} \in U.\end{aligned}$$

It follows from the local convexity of  $s(\cdot)$  that if  $\mathbf{m}_h$  is in the neighborhood of  $\mathbf{m}$ , then

$$(16) \quad c \|\mathbf{m} - \mathbf{m}_h\|^2 \leq (s'(\mathbf{m}), \mathbf{m} - \mathbf{m}_h) - (s'(\mathbf{m}_h), \mathbf{m} - \mathbf{m}_h),$$

where  $s'(\mathbf{m}_h)$  is defined in (14). By using (8), (10) and (16),

$$(17) \quad \begin{aligned}c \|\mathbf{m} - \mathbf{m}_h\|^2 &\leq -(s'(\mathbf{m}_h), \mathbf{m} - \mathbf{m}_h) \\ &\leq (s'_h(\mathbf{m}_h) - s'(\mathbf{m}_h), \mathbf{m} - \mathbf{m}_h).\end{aligned}$$

Since  $\lambda, \mu$  are constants, it can be shown that

$$\|\mathbf{p}(\mathbf{m}_h)\|_{1,\Omega} \leq C, \quad \|\mathbf{u}_h\|_{1,\Omega} \leq C$$

when  $h$  is small, see [20] for more details. Noting that  $\mathbf{m}$  belongs to a finite dimensional space, we have

$$(18) \quad \begin{aligned}&(s'_h(\mathbf{m}_h) - s'(\mathbf{m}_h), \mathbf{m} - \mathbf{m}_h) \\ &= (\tilde{\mathbf{p}}(\mathbf{m}_h) \tilde{\mathbf{u}}(\mathbf{m}_h), \mathbf{m} - \mathbf{m}_h) - (\tilde{\mathbf{p}}_h \tilde{\mathbf{u}}_h, \mathbf{m} - \mathbf{m}_h) \\ &= (\tilde{\mathbf{p}}(\mathbf{m}_h) \tilde{\mathbf{u}}(\mathbf{m}_h) - \tilde{\mathbf{p}}(\mathbf{m}_h) \tilde{\mathbf{u}}_h, \mathbf{m} - \mathbf{m}_h) \\ &\quad - (\tilde{\mathbf{p}}_h \tilde{\mathbf{u}}_h - \tilde{\mathbf{p}}(\mathbf{m}_h) \tilde{\mathbf{u}}_h, \mathbf{m} - \mathbf{m}_h) \\ &\leq C \|\mathbf{m} - \mathbf{m}_h\| (\|\tilde{\mathbf{p}}(\mathbf{m}_h) - \tilde{\mathbf{p}}_h\|_{0,\Omega} + \|\tilde{\mathbf{u}}(\mathbf{m}_h) - \tilde{\mathbf{u}}_h\|_{0,\Omega}).\end{aligned}$$

Therefore, Lemma 3.1 is a direct consequence of (17) and (18).  $\square$

In order to estimate  $\|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2$  and  $\|\bar{\mathbf{p}}(\mathbf{m}_h) - \bar{\mathbf{p}}_h\|_{0,\Omega}^2$ , we need the following lemma.

**Lemma 3.2.** *Let  $(\mathbf{p}(\mathbf{m}_h), \mathbf{u}(\mathbf{m}_h)), (\mathbf{p}_h, \mathbf{u}_h)$  be the solution of (15) and (10), respectively. Then*

$$(19) \quad \begin{aligned} \|\bar{\mathbf{p}}(\mathbf{m}_h) - \bar{\mathbf{p}}_h\|_{0,\Omega}^2 &\leq C \left\{ \|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2 + \sum_l h_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right]^2 \right. \\ &\quad \left. + \sum_\tau h_\tau^2 \int_\tau (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z}))^2 \right\}, \\ \|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2 &\leq C \left\{ \sum_l h_l \int_l \left[ \frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} \right]^2 + \sum_\tau h_\tau^2 \int_\tau (\mu_h \Delta \mathbf{u}_h \right. \\ &\quad \left. + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{u}_h) + \mathbf{f})^2 \right\}, \end{aligned}$$

where  $[\mathbf{r}]_l$  represents the jump of  $\mathbf{r}$  on the edge  $l$ . For convenience of presentation,  $[\mathbf{r}]_l = 0$  when  $l \subset \partial\Omega$ .

*Proof.* Let  $\mathbf{e}^p = \mathbf{p}(\mathbf{m}_h) - \mathbf{p}_h$ ,  $\mathbf{e}^p \in (H_0^1(\Omega))^2$ , then by using the positive definite property of the elasticity matrix and Korn's inequality (see [10]), also combined with (10), (11), (12) and (15), we can obtain

$$\begin{aligned} \|\bar{\mathbf{p}}(\mathbf{m}_h) - \bar{\mathbf{p}}_h\|_{0,\Omega}^2 &= |\epsilon_{11}(\mathbf{e}^p) + \epsilon_{22}(\mathbf{e}^p)|_{0,\Omega}^2 + \left| \sum_{i,j=1}^2 \epsilon_{ij}(\mathbf{e}^p) \right|_{0,\Omega}^2 \\ &\leq C \sum_{i,j=1}^2 |\epsilon_{ij}(\mathbf{e}^p)|_{0,\Omega}^2 = C |\mathbf{e}^p|^2 \leq C a_h(\mathbf{e}^p, \mathbf{e}^p) \\ &= C \left\{ (Q^*(Q\mathbf{u}(\mathbf{m}_h) - \mathbf{z}), \mathbf{e}^p) - a_h((\mathbf{e}^p - I_h \mathbf{e}^p), \mathbf{p}_h) - a_h(I_h \mathbf{e}^p, \mathbf{p}_h) \right\} \\ &= C \left\{ (Q^*(Q\mathbf{u}(\mathbf{m}_h) - Q\mathbf{u}_h), \mathbf{e}^p) - a_h((\mathbf{e}^p - I_h \mathbf{e}^p), \mathbf{p}_h) \right. \\ &\quad \left. - (Q^*(Q\mathbf{u}_h - \mathbf{z}), I_h \mathbf{e}^p - \mathbf{e}^p) \right\} \\ &\leq C \left\{ \|\mathbf{u}(\mathbf{m}_h) - \mathbf{u}_h\|_{0,\Omega} \|\mathbf{e}^p\|_{0,\Omega} + \sum_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right] (I_h \mathbf{e}^p - \mathbf{e}^p) \right. \\ &\quad \left. + \sum_\tau \int_\tau (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z})) (\mathbf{e}^p - I_h \mathbf{e}^p) \right\} \\ &\leq C \left\{ \|\mathbf{u}(\mathbf{m}_h) - \mathbf{u}_h\|_{0,\Omega} \|\mathbf{e}^p\|_{0,\Omega} + \left( \sum_l h_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right]^2 \right)^{\frac{1}{2}} \|\mathbf{e}^p\|_{1,\Omega} + \right. \\ &\quad \left. \left( \sum_\tau h_\tau^2 \int_\tau (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z}))^2 \right)^{\frac{1}{2}} \|\mathbf{e}^p\|_{1,\Omega} \right\}. \end{aligned}$$

Thus,

$$\begin{aligned} \|\bar{\mathbf{p}}(\mathbf{m}_h) - \bar{\mathbf{p}}_h\|_{0,\Omega}^2 &\leq C \left\{ \|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2 + \sum_l h_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right]^2 \right. \\ &\quad \left. + \sum_\tau h_\tau^2 \int_\tau (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z}))^2 \right\}. \end{aligned}$$

Similarly, it can be proved that

$$\begin{aligned} \|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2 &\leq C \left\{ \sum_l h_l \int_l \left[ \frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} \right]^2 + \sum_\tau h_\tau^2 \int_\tau (\mu_h \Delta \mathbf{u}_h \right. \\ &\quad \left. + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{u}_h) + \mathbf{f})^2 \right\}. \end{aligned}$$

This completes the proof of the lemma.  $\square$

Now, we give the a posteriori error estimator in the following theorem.

**Theorem 3.3.** *Let  $(\mathbf{u}, \mathbf{p}, \mathbf{m})$  and  $(\mathbf{u}_h, \mathbf{p}_h, \mathbf{m}_h)$  be the solution of (8) and (10), respectively. Then*

$$(20) \quad \|\mathbf{m}_h - \mathbf{m}\|^2 + |\mathbf{u}_h - \mathbf{u}|^2 + |\mathbf{p}_h - \mathbf{p}|^2 \leq C(\eta_1^2 + \eta_2^2),$$

where

$$\eta_1^2 = \sum_l h_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right]^2 + \sum_\tau h_\tau^2 \int_\tau (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z}))^2$$

and

$$\eta_2^2 = \sum_l h_l \int_l \left[ \frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} \right]^2 + \sum_\tau h_\tau^2 \int_\tau (\mu_h \Delta \mathbf{u}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{u}_h) + \mathbf{f})^2,$$

with  $[\mathbf{r}]_l$  representing the jump of  $\mathbf{r}$  on the edge  $l$ .

*Proof.* Following Lemma 3.1, Lemma 3.2 and Korn's inequality, we get

$$(21) \quad \|\mathbf{m}_h - \mathbf{m}\|^2 + |\mathbf{u}(\mathbf{m}_h) - \mathbf{u}_h|^2 + |\mathbf{p}(\mathbf{m}_h) - \mathbf{p}_h|^2 \leq C(\eta_1^2 + \eta_2^2).$$

According to (8) and (15), it can be shown that

$$(22) \quad \begin{aligned} |\mathbf{p}(\mathbf{m}_h) - \mathbf{p}| &\leq C(\|\mathbf{m}_h - \mathbf{m}\| + \|\mathbf{u}(\mathbf{m}_h) - \mathbf{u}\|_{0,\Omega}), \\ |\mathbf{u}(\mathbf{m}_h) - \mathbf{u}| &\leq C\|\mathbf{m}_h - \mathbf{m}\|. \end{aligned}$$

Thus, we have

$$(23) \quad \begin{aligned} |\mathbf{p}_h - \mathbf{p}| &\leq |\mathbf{p}(\mathbf{m}_h) - \mathbf{p}| + |\mathbf{p}(\mathbf{m}_h) - \mathbf{p}_h|, \\ |\mathbf{u}_h - \mathbf{u}| &\leq |\mathbf{u}(\mathbf{m}_h) - \mathbf{u}| + |\mathbf{u}(\mathbf{m}_h) - \mathbf{u}_h|, \end{aligned}$$

and (20) follows from (21)-(23).  $\square$

As an extension, we consider the optimization problem (2) governed by the linear elasticity problem including Neumann boundary condition on  $\Gamma_N$ , i.e., we solve

$$(24) \quad \min_{\mathbf{m}} \frac{1}{2} \|Q\mathbf{u}(\mathbf{m}) - \mathbf{z}\|_Z^2$$

subject to

$$a(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + (\mathbf{g}, \mathbf{v}), \quad \forall \mathbf{v} \in U,$$

where

$$(\mathbf{g}, \mathbf{v}) = \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{v} = \int_{\Gamma_N} \sum_{i=1}^2 g_i v_i, \quad \forall \mathbf{v} \in U,$$

and  $\Gamma = \Gamma_D + \Gamma_N$ . The Dirichlet part  $\Gamma_D$  of the boundary must have a positive one-dimensional Lebesgue measure in order to guarantee the unique solvability of the state problem. We set  $\mathbf{g} = (g_1, g_2) \in (L^2(\Gamma_N))^2$ , and define  $Y = U + u_D$  and  $U = \{\mathbf{u} = (u_1, u_2) \in (H^1(\Omega))^2, u_i = 0 \text{ on } \Gamma_D, 1 \leq i \leq 2\}$ . Assume that the composite function  $s(\mathbf{m})$  is locally uniformly convex, the optimal conditions for (24) are to find  $(\mathbf{u}, \mathbf{p}, \mathbf{m}) \in Y \times U \times P$  such that

$$(25) \quad \begin{aligned} a(\mathbf{u}, \mathbf{v}) &= (\mathbf{f}, \mathbf{v}) + (\mathbf{g}, \mathbf{v}) && \forall \mathbf{v} \in U, \\ a(\mathbf{q}, \mathbf{p}) &= (Q^*(Q\mathbf{u} - \mathbf{z}), \mathbf{q}) && \forall \mathbf{q} \in U, \\ (-\tilde{\mathbf{p}}\mathbf{u}, \mathbf{m} - \mathbf{w}) &\leq 0 && \forall \mathbf{w} \in P. \end{aligned}$$

The finite element approximation of (25) is to find  $(\mathbf{u}_h, \mathbf{p}_h, \mathbf{m}_h) \in Y^h \times V^h \times P$  such that

$$(26) \quad \begin{aligned} a_h(\mathbf{u}_h, \mathbf{v}_h) &= (\mathbf{f}, \mathbf{v}_h) + (\mathbf{g}, \mathbf{v}_h) && \forall \mathbf{v}_h \in V^h, \\ a_h(\mathbf{q}_h, \mathbf{p}_h) &= (Q^*(Q\mathbf{u}_h - \mathbf{z}), \mathbf{q}_h) && \forall \mathbf{q}_h \in V^h, \\ (-\tilde{\mathbf{p}}_h \tilde{\mathbf{u}}_h, \mathbf{m}_h - \mathbf{w}_h) &\leq 0 && \forall \mathbf{w}_h \in P. \end{aligned}$$

By employing similar techniques used in the proof of Theorem 3.3, the following theorem can be derived.

**Theorem 3.4.** *Let  $(\mathbf{u}, \mathbf{p}, \mathbf{m})$  and  $(\mathbf{u}_h, \mathbf{p}_h, \mathbf{m}_h)$  be the solution of (25) and (26), respectively. Then*

$$(27) \quad \|\mathbf{m}_h - \mathbf{m}\|^2 + |\mathbf{u}_h - \mathbf{u}|^2 + |\mathbf{p}_h - \mathbf{p}|^2 \leq C(\eta_1^2 + \eta_2^2),$$

where

$$\begin{aligned} \eta_1^2 &= \sum_{l \notin \Gamma_N} h_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right]^2 + \sum_{l \in \Gamma_N} h_l \int_l \left( \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right)^2 \\ &\quad + \sum_{\tau} h_{\tau}^2 \int_{\tau} (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z}))^2 \end{aligned}$$

and

$$\begin{aligned} \eta_2^2 &= \sum_{l \notin \Gamma_N} h_l \int_l \left[ \frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} \right]^2 + \sum_{l \in \Gamma_N} h_l \int_l \left( \mathbf{g} - \frac{\partial \mathbf{u}_h}{\partial \mathbf{n}} \right)^2 \\ &\quad + \sum_{\tau} h_{\tau}^2 \int_{\tau} (\mu_h \Delta \mathbf{u}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{u}_h) + \mathbf{f})^2, \end{aligned}$$

with  $[\mathbf{r}]_l$  representing the jump of  $\mathbf{r}$  on the edge  $l$ . We set  $[\mathbf{r}]_l = 0$  when  $l \subset \Gamma_D$ .

*Proof.* Let  $\mathbf{e}^p = \mathbf{p}(\mathbf{m}_h) - \mathbf{p}_h$ ,  $\mathbf{e}^p \in U$ , where  $\mathbf{p}(\mathbf{m}_h)$  is the solution to the equation

$$(28) \quad \begin{aligned} a_h(\mathbf{u}(\mathbf{m}_h), \mathbf{v}) &= (\mathbf{f}, \mathbf{v}) + (\mathbf{g}, \mathbf{v}) & \forall \mathbf{v} \in U, \\ a_h(\mathbf{q}, \mathbf{p}(\mathbf{m}_h)) &= (Q^*(Q\mathbf{u}(\mathbf{m}_h) - \mathbf{z}), \mathbf{q}) & \forall \mathbf{q} \in U. \end{aligned}$$

By using the positive definite property of the elasticity matrix and Korn's inequality, combined with (12), (11), (25), (26) and (28), we have

$$\begin{aligned} &\|\bar{\mathbf{p}}(\mathbf{m}_h) - \bar{\mathbf{p}}_h\|_{0,\Omega}^2 \leq C a_h(\mathbf{e}^p, \mathbf{e}^p) \\ &= C \{ (Q^*(Q\mathbf{u}(\mathbf{m}_h) - Q\mathbf{u}_h), \mathbf{e}^p) - a_h((\mathbf{e}^p - I_h \mathbf{e}^p), \mathbf{p}_h) \\ &\quad - (Q^*(Q\mathbf{u}_h - \mathbf{z}), I_h \mathbf{e}^p - \mathbf{e}^p) \} \\ &\leq C \left\{ \|\mathbf{u}(\mathbf{m}_h) - \mathbf{u}_h\|_{0,\Omega} \|\mathbf{e}^p\|_{0,\Omega} \right. \\ &\quad + \sum_{l \notin \Gamma_N} \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right] (I_h \mathbf{e}^p - \mathbf{e}^p) + \sum_{l \in \Gamma_N} \int_l \left( \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right) (I_h \mathbf{e}^p - \mathbf{e}^p) \\ &\quad \left. + \sum_{\tau} \int_{\tau} (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z})) (\mathbf{e}^p - I_h \mathbf{e}^p) \right\} \\ &\leq C \left\{ \|\mathbf{u}(\mathbf{m}_h) - \mathbf{u}_h\|_{0,\Omega} \|\mathbf{e}^p\|_{0,\Omega} \right. \\ &\quad + \left( \sum_{l \notin \Gamma_N} h_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right]^2 + \sum_{l \in \Gamma_N} h_l \int_l \left( \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right)^2 \right)^{\frac{1}{2}} \|\mathbf{e}^p\|_{1,\Omega} \\ &\quad \left. + \left( \sum_{\tau} h_{\tau}^2 \int_{\tau} (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z}))^2 \right)^{\frac{1}{2}} \|\mathbf{e}^p\|_{1,\Omega} \right\}. \end{aligned}$$

Thus,

$$\begin{aligned} \|\bar{\mathbf{p}}(\mathbf{m}_h) - \bar{\mathbf{p}}_h\|_{0,\Omega}^2 &\leq C \left\{ \|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2 + \sum_{l \notin \Gamma_N} h_l \int_l \left[ \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right]^2 + \sum_{l \in \Gamma_N} h_l \int_l \left( \frac{\partial \mathbf{p}_h}{\partial \mathbf{n}} \right)^2 \right. \\ &\quad \left. + \sum_{\tau} h_{\tau}^2 \int_{\tau} (\mu_h \Delta \mathbf{p}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{p}_h) + Q^*(Q\mathbf{u}_h - \mathbf{z}))^2 \right\}. \end{aligned}$$



Similarly, we obtain

$$\begin{aligned} \|\bar{\mathbf{u}}(\mathbf{m}_h) - \bar{\mathbf{u}}_h\|_{0,\Omega}^2 &\leq C \left\{ \sum_{\tau} h_{\tau}^2 \int_{\tau} (\mu_h \Delta \mathbf{u}_h + (\lambda_h + \mu_h) \nabla(\nabla \cdot \mathbf{u}_h) + \mathbf{f})^2 \right. \\ &\quad \left. + \sum_{l \notin \Gamma_N} h_l \int_l \left[ \frac{\partial \mathbf{u}_h}{\partial n} \right]^2 + \sum_{l \in \Gamma_N} h_l \int_l \left( \mathbf{g} - \frac{\partial \mathbf{u}_h}{\partial n} \right)^2 \right\}. \end{aligned}$$

(27) then can be proved with the technique used in Theorem 3.3 and Lemma 3.1.  $\square$

Since we assume that  $Q : Y \rightarrow Z$  is a linear bounded observation operator, the adjoint operator  $Q^* : Z \rightarrow Y$  is a linear bounded operator too [18]. However, in some applications, only  $n_m$  measurements at some particular points  $\mathbf{x}_i$ ,  $i = 1, 2, \dots, n_m$ , are available. If  $\mathbf{u} \in U = (H^1(\Omega))^2$ , the solution  $\mathbf{u}$  may not be continuous when the geometrical dimension  $d$  is equal to or higher than two. This implies that the solution in the measurement points may not be defined and that  $Q$  is not a linear bounded observation operator any more. In order to circumvent this issue, one may work with the average value of the quantity in a small neighborhood  $\varpi_{\epsilon}(\mathbf{x}_i)$  of the measurement points, or use mollification; see [1, 25]. It is customary to choose the mollifiers  $k_{\epsilon}$  of the form

$$k_{\epsilon}(\mathbf{x} - \mathbf{x}_i) = \begin{cases} C \exp[-\epsilon^2/(\epsilon^2 - (\mathbf{x} - \mathbf{x}_i)^2)] & \text{if } |\mathbf{x} - \mathbf{x}_i| < \epsilon, \\ 0 & \text{if } |\mathbf{x} - \mathbf{x}_i| \geq \epsilon, \end{cases}$$

where the constant  $C$ , which depends on  $d, \epsilon$  and  $\mathbf{x}_i$ , is selected to satisfy

$$\int_{\Omega} k_{\epsilon}(\mathbf{x} - \mathbf{x}_i) d\mathbf{x} = 1.$$

Hence, we can reformulate the optimization problem as

$$\min_{\mathbf{m}} \frac{1}{2} \sum_i \int_{\Omega} k_{\epsilon}(\mathbf{x} - \mathbf{x}_i) (\mathbf{u}(\mathbf{m}) - \mathbf{z}_i)^2 d\mathbf{x},$$

where  $\mathbf{z}_i$  denotes the observation data at the point  $\mathbf{x}_i$ . Our approach, developed in section 3, is then well-suited to the point-wise case; see [12] for more details.

In practice, measurement data may not match exactly due to measurement errors. The statistical quality of the estimated parameter is proportional to the least square residual  $Q\mathbf{u}(\mathbf{m}) - \mathbf{z}$ , for more details, see [5]. This means that it is in general not efficient to reduce the error in the parameter by mesh refinement when the discretization error is smaller than the statistical error.

#### 4. Numerical examples

In this section, we carry out some numerical experiments to demonstrate the efficiency of our error estimators obtained in Section 3. Throughout, the finite element method is defined on a triangular mesh and piecewise quadratic polynomials are used in our finite element space. Furthermore, the same mesh will be used for the state and adjoint variables. Thus,  $\eta_1^2 + \eta_2^2$  will be the indicator of the mesh refinement. Moreover, we use the h-method and the general idea is to refine the mesh such that the error indicators are equally distributed over the computational meshes. To this end, an equidistribution strategy for mesh refinement is used in our algorithm [11]. Assume that an a posteriori error estimator  $\eta$  has the form  $\eta^2 = \sum_{e_i} \eta_{e_i}^2$ , where  $e_i$  is a finite element. In the process of mesh refinement, a tolerance  $tol$  and a parameter  $\theta$  are defined ( $tol$  could be the average value of the error estimator), and the element  $e_i$  will be refined if  $\eta_{e_i}^2 > \theta tol$ .

In order to solve the optimization problem, we employ the standard Gauss-Newton algorithm (see, e.g., [14]) with a trust region technique. It is well known

TABLE 1. The errors in  $\lambda$  and  $\mu$  for the uniform mesh and adaptive mesh.

Discretization	DOF	$ \lambda - \lambda_h $	$ \mu - \mu_h $
Uniform mesh	24929	9.108e-04	6.932e-04
Adaptive mesh	2956	8.90e-05	2.61e-05

that there are many types of trust-region methods. In our paper, the Levenberg-Marquardt method is utilized; see [13, 27]. All our computations are carried out with AFEPack, a generic C++ adaptive finite element library [21].

**4.1. Example 1.** In the first example, we consider the planar linear elasticity problem in an L-shaped domain (see Figure 1). The parameters  $(\lambda, \mu) = (2, 1)$  are

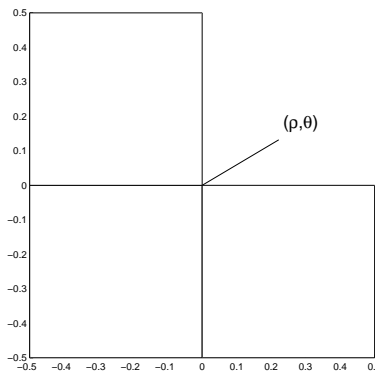


FIGURE 1. L-shaped domain.

estimated using the exact displacement given as follows

$$\mathbf{u} = \mathbf{u}^* + \frac{1}{\lambda + \mu} \phi + \frac{1}{\mu} \varphi,$$

where

$$\mathbf{u}^*(\rho, \theta) = \rho^\gamma \begin{pmatrix} \cos \gamma \theta & \sin \gamma \theta \\ -\sin \gamma \theta & \cos \gamma \theta \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

$$\phi(x_1, x_2) = \begin{pmatrix} -\cos \pi x_1 \cos \pi x_2 \\ \sin \pi x_1 \sin \pi x_2 \end{pmatrix}, \varphi(x_1, x_2) = \begin{pmatrix} \sin \pi x_1 \cos \pi x_2 \\ -\cos \pi x_1 \sin \pi x_2 \end{pmatrix},$$

with  $\alpha = \beta = \gamma = \frac{1}{2}$ , and  $\mathbf{f}$  is given by

$$\mathbf{f} = 2\pi^2 \begin{pmatrix} \sin \pi x_1 \cos \pi x_2 \\ -\cos \pi x_1 \sin \pi x_2 \end{pmatrix} + 2\pi^2 \frac{2\mu + \lambda}{\mu + \lambda} \begin{pmatrix} -\cos \pi x_1 \cos \pi x_2 \\ \sin \pi x_1 \sin \pi x_2 \end{pmatrix}.$$

On the boundary  $\Gamma_D$  of the domain, we set  $\mathbf{u}_D = \mathbf{u}$ . The polar coordinate  $(\rho, \theta)$  is shown in Figure 1.

We let the initial values of parameters to be  $\mathbf{m}^0 = (5, 0.6)^T$ . A typical mesh resulting from the application of our error estimator is shown in the left plot of Figure 2. It is obvious that the refinement is mainly around the origin. The reason is that the stress is singular at the origin. A comparison of the accuracy achieved on the uniform mesh and our adaptive mesh is made and the result is shown in Table 1, where *DOF* represents the degrees of freedom of the mesh. It turns out that the adaptive finite element method using our error estimator produces a mesh with obvious savings in degrees of freedom for a prescribed accuracy level.

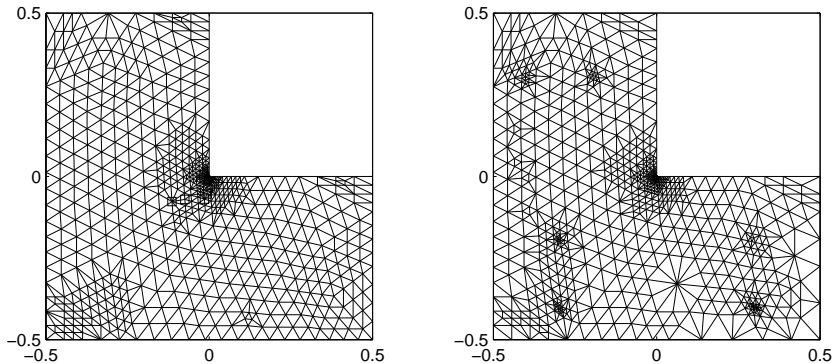


FIGURE 2. Left: Typical meshes produced by our error estimator with distributed measurement. Right: Typical meshes produced by our error estimator with point-wise measurement.

In some applications, measurements can be obtained only at distinct points. Assume that the measurements are given by values of the state variable at six different points:

$$(x_1, y_1) = (-0.2, 0.3), (x_2, y_2) = (-0.4, 0.3), (x_3, y_3) = (-0.2, -0.3),$$

$$(x_4, y_4) = (-0.4, -0.3), (x_5, y_5) = (0.3, -0.2), (x_6, y_6) = (0.3, -0.4).$$

The mesh resulting from the application of our error estimator is shown in the right-hand plot of Figure 2. During the test, we add random noise such that the observation  $\mathbf{z}$  is computed from

$$\mathbf{z}(x_i, y_i) = \mathbf{u}(x_i, y_i) * (1 + \delta\kappa)$$

for a random number  $\kappa$  in  $[-0.5, 0.5]$  with  $\delta$  as the noise level. The errors in  $\|\mathbf{m} - \mathbf{m}_h\|$  with different noise levels are shown in Figure 3. It indicates that when the noise is large, the errors in the parameters are large as well. Moreover, with a high level of noise, the errors in the parameters cannot be reduced efficiently by the mesh refinement. For instance: in case  $\delta = 10\%$ , further mesh refinement does not produce much more accuracy. However, when the noise level is low, the errors in the parameters can be reduced considerably by the mesh refinement.

Finally, we investigate the quantitative behavior of the estimator by computing the effectivity index of the error estimator, which is defined by

$$I_{eff} = \|\mathbf{m} - \mathbf{m}_h\|^2 / (\eta_1^2 + \eta_2^2).$$

The result is shown in Table 2 and it indicates that our error estimator is sufficient for guiding efficient mesh refinement.

TABLE 2. The effectivity index of the error estimator.

DOF	$\ \mathbf{m} - \mathbf{m}_h\ ^2$	$\eta_1^2 + \eta_2^2$	$I_{eff}$
119	4.763e-3	9.660e-3	0.493
429	1.840e-3	4.524e-3	0.407
1550	8.642e-4	2.230e-3	0.388
2301	3.878e-4	1.110e-3	0.349
3042	1.665e-4	5.509e-4	0.302
3161	7.520e-5	2.768e-4	0.272
3348	3.142e-5	1.395e-4	0.225

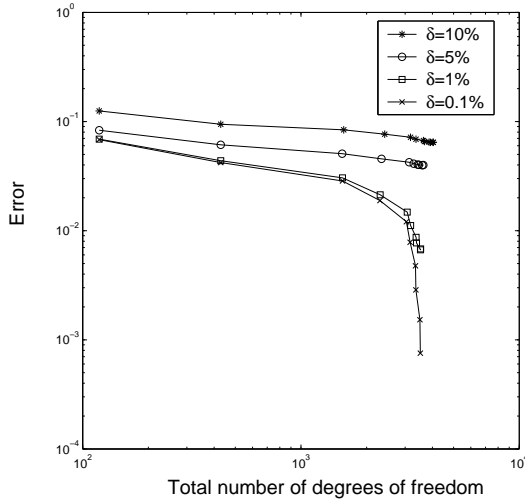


FIGURE 3. Error reduction in  $\|\mathbf{m} - \mathbf{m}_h\|$  for the different noise levels.

It is important to remark that, in the point-wise measurement case, care must be taken in selecting the size of  $\epsilon$ . In our finite element codes, the integration is carried out using classical Gauss quadrature rules and accuracy is directly correlated to the number of Gaussian points used in each element. Therefore, the size of the support of  $k_\epsilon(\mathbf{x} - \mathbf{x}_i)$ , equal to  $2\epsilon$ , should be selected with respect to the mesh size  $h$  of the element containing the point  $\mathbf{x}_i$ . Some suggestions can be found in [25], where some numerical experiments are carried out to show a possible good choice of the size of  $\epsilon$ .

A mesh that is efficient for our parameter estimation problem will depend not only on the forward problem, but also on the objective function. In general, a mesh efficient for the discretization scheme of the parameter estimation is quite different from a mesh suited for the forward problem. For instance, the mesh refinements will not be around the measurement points for the point-wise case by only considering a mesh suited for the solution of the forward problem. Of course, for this particular case, we know the probable refinement areas a priori so that it is possible to investigate a priori local mesh refinements, which is impossible in general.

Here we report investigations on a priori local mesh refinements and comparisons with our adaptive mesh refinement strategy. The main idea is to use a priori knowledge about the singularity of the forward solution and the measurement points to determine mesh densities of the elements near the corner and the measurement points [2, 3]. We first tried a priori indicators of power type defined by  $r > 0$

$$\eta'_{priori} = \sum_{\tau} \frac{1}{(\mathbf{x}_{\tau} - \mathbf{x}_i)^r},$$

where the points  $\mathbf{x}_i$  represent the measurement points and the re-entrant corner point and  $\mathbf{x}_{\tau}$  denotes the barycenter of the element  $\tau$ . It seemed that such indicators did not work well. Then we used the mollifiers  $k_\epsilon$  as the indicator, which is defined in section 3. By using only the coordinator information, an a priori indicator can be constructed as

$$\eta_{priori} = \sum_{\tau} \int_{\tau} k_\epsilon(\mathbf{x} - \mathbf{x}_i).$$

For convenience, we let  $\epsilon = \epsilon_{k+1} = \xi * \epsilon_k, 0 < \xi < 1, k = 0, 1, \dots$ , where  $k$  denotes the refinement step. In the process of mesh refinement, at each iteration, an average value  $\eta_{avg}$  is calculated and a parameter  $\theta$  is defined,  $0 < \theta < 1$ . Then the element  $\tau$  will be refined if  $\eta_{priori,\tau} > \theta\eta_{avg}$ , where  $\eta_{priori,\tau} = \int_{\tau} k_{\epsilon}(\mathbf{x} - \mathbf{x}_i)$ . In our experiments, we first put  $\xi = 0.8, \epsilon_0 = 0.05$ , the same as those defined in the objective function. The results were not satisfactory, however. Then we tried several different combinations of these parameters. For the case  $\xi = 0.8, \theta = 0.75$ , different choices of the size of  $\epsilon_0$  with the convergence results are shown in Figure 4.

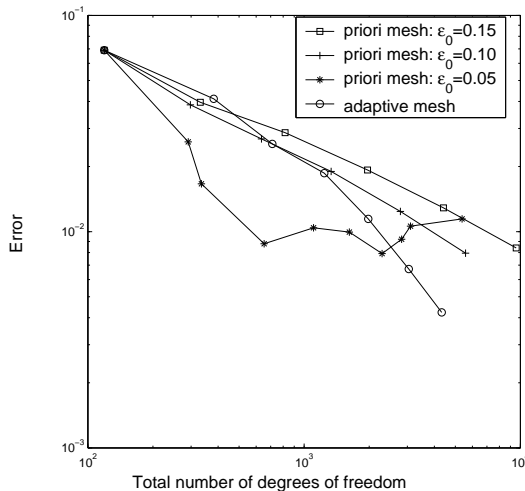


FIGURE 4. Error reduction in  $\|\mathbf{m} - \mathbf{m}_h\|$  for a priori mesh and adaptive mesh depending on the degrees of freedom.

As shown in Figure 4, it is clear that an appropriate mesh generated by a priori mesh refinement strategy can produce rather accurate solutions compared to our adaptive mesh, e.g., in the case where  $\epsilon_0 = 0.1$ , although it does not seem that they can match the adaptive method in terms of accuracy. Furthermore, it is much cheaper from a computational point of view. However, it is not straightforward to select the constants in the a priori mesh grading algorithms, such as  $\epsilon_0$ ,  $\theta$  and  $\xi$ ; i.e., it is not easy to decide a good density of the mesh around measurement points, re-entrant corner and other parts of the domain. As we can see from Figure 4, in the case  $\epsilon_0 = 0.05$ , an a priori mesh cannot produce high-accuracy solutions. Such problems are quite common in a priori mesh approaches; see [2]. In our adaptive mesh refinement algorithm, after we define the parameter  $\theta$ , the mesh is generated completely automatically with respect to the error estimators and a good balance between the refined and unrefined regions is obtained. For more details on how to choose the parameter  $\theta$ , please refer to [6]. We should mention that several different values of  $\theta$  have been tested for our adaptive algorithm, all with satisfactory results.

Moreover, although the a priori mesh refinement strategy may manage to work for this individual problem, our goal is to produce efficient meshes for general applications. The adaptive finite element method suits this purpose. In our next parameter estimation problem, the locations of measurement points are unknown and are in fact design variables. This will lead to the following formulation:

$$\min_{\mathbf{m}} \max_{\mathbf{x}_i} \frac{1}{2} \sum_i \int_{\Omega} k_{\epsilon}(\mathbf{x} - \mathbf{x}_i) (\mathbf{u}(\mathbf{m}, \mathbf{x}) - \mathbf{z}_i)^2 d\mathbf{x},$$

where  $\mathbf{z}_i$  is the measurement value of  $\mathbf{u}(\mathbf{m}_0, \mathbf{x}_i)$  at point  $\mathbf{x}_i$ , and we assume that  $\mathbf{m}_0$  is the exact value of  $\mathbf{m}$ . It is clear that we cannot design an a priori mesh refinement strategy for this problem. However, the general framework presented in this paper still works. Finally, our adaptive method is also suitable for the distributed parameter estimation problems.

**4.2. Example 2.** In the second example, we consider the linear elastic problem with the Neumann boundary conditions. As shown in Figure 5, a short cantilever plate of elastic isotropic material is analyzed [26]. The computational domain

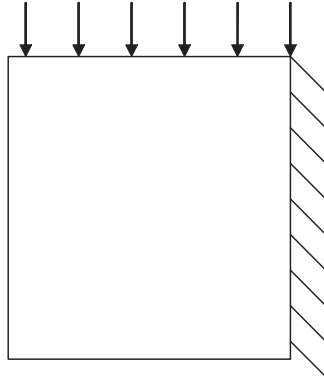


FIGURE 5. Short cantilever plate.

$\Omega = (0, 1)^2$ . The boundary conditions on the top edge of the computational domain are set to  $\mathbf{g} \cdot \mathbf{n} = (0, -1.0)$ , the boundary conditions on the right are set to  $\mathbf{u} = \mathbf{0}$ , and the boundary conditions are set to homogeneous natural boundary conditions on the bottom and left edges. Young's modulus  $E = 1.0$  and Poisson's ratio  $\nu = 0.3$ .

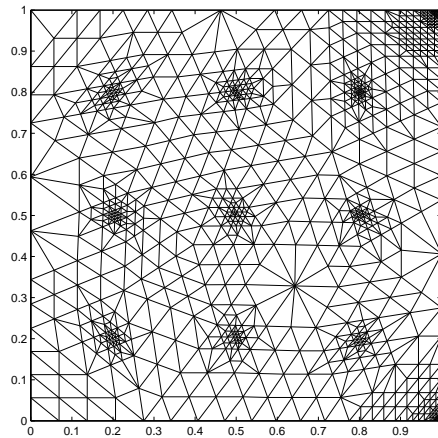


FIGURE 6. Typical meshes produced by our error estimator.

The exact solution for this problem is not known analytically, so we use a very fine mesh to solve the problem to get the "exact" solution at the measurement points

(0.2, 0.2), (0.2, 0.5), (0.2, 0.8), (0.5, 0.2), (0.5, 0.5), (0.5, 0.8), (0.8, 0.2), (0.8, 0.5), (0.8, 0.8). Typical meshes resulting from the application of our error estimator are shown in Figure 6. The mesh refinement procedures are not only around the top-right corner and the bottom-right corner, where the stress is relatively larger, but also around the measurement points.

A comparison of the accuracy in  $\mu$  achieved on the uniform mesh and the adaptive mesh is shown in Figure 7. It can be seen that our proposed algorithm is much more efficient. We also execute some numerical computations with different observation noise levels. In Table 3, the errors in  $\lambda$  and  $\mu$  with different noise levels are shown. Again, the result indicates that when the noise level is low, the errors in the parameters can be reduced considerably by mesh refinement.

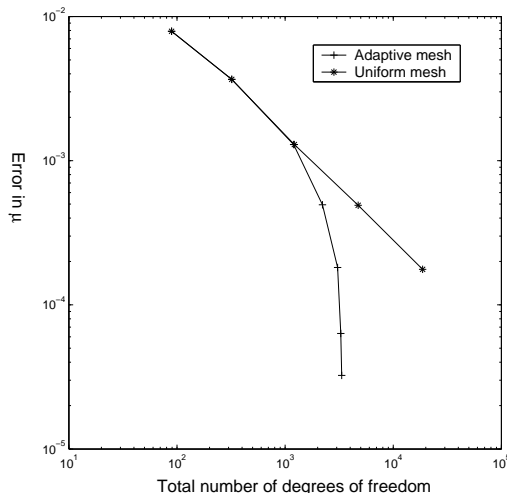


FIGURE 7. Error reduction in  $|\mu - \mu_h|$  for the uniform mesh and adaptive mesh depending on the degrees of freedom.

TABLE 3. Computation with noisy data.

$\delta$	DOF	$ \lambda - \lambda_h $	$ \mu - \mu_h $
10%	1172	0.0885	4.597e-3
5%	3089	0.0115	2.749e-3
1%	3010	1.277e-3	1.628e-4
0.1%	2910	7.220e-4	8.634e-5
0.0%	3336	3.066e-4	3.24e-05

## References

- [1] M. Ainsworth and J. T. Oden, A Posterior Error Estimation in Finite Element Analysis, John Wiley & Sons, New York, 2000.
- [2] T. Apel and F. Milde, Comparison of various mesh refinement strategies near edges, Comm. Numer. Methods Engrg, 12 (1996) 373-381.
- [3] I. Babuška, R. B. Kellogg and J. Pitkäranta, Direct and inverse error estimates for finite elements with mesh refinements, Numer. Math., 33 (1979) 447-471.
- [4] W. Bangerth, Adaptive Finite Element Methods for the Identification of Distributed Parameters in the Partial Differential Equations, Ph.D. thesis, University of Heidelberg, Institut für Angewandte Mathematik, 2002.
- [5] D. Bates and D. Watts, Nonlinear Regression Analysis and its Applications, John Wiley & Sons, New York, 1988.

- [6] 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 (1996) 237-264.
- [7] R. Becker, H. Kapp and R. Rannacher, *Adaptive finite element methods for optimization problems*, Proceedings of 18th Biennial Conference on Numerical Analysis (Ed. D. F. Griffiths and G. A. Watson), p. 21-42, Chapman & Hall/CRC, 2000.
- [8] R. Becker and B. Vexler, A posteriori error estimation for finite element discretization of parameter identification problems, *Numer. Math.*, 96 (2004) 435-459.
- [9] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [10] G. Duvaut and J. L. Lions, *Inequalities in Mechanics and Physics*, Springer-Verlag, Berlin, 1976.
- [11] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems I: A linear model problem, *SIAM J. Numer. Anal.*, 28 (1991) 43-77.
- [12] T. Feng, M. Gulliksson and W. B. Liu, Adaptive finite element methods for identification of elastic constants, *J. Sci. Comput.*, 26 (2006) 217-235.
- [13] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, 2000.
- [14] E. Haber, U. M. Ascher and D. Oldenburg, On optimization techniques for solving nonlinear inverse problems, *Inverse Problems*, 16 (2000) 1263-1280.
- [15] P. C. Hansen and D. P. O'leary, The use of L-curve in the regularization of discrete ill-posed problems, *SIAM J. Sci. Comput.*, 14 (1993) 1487-1503.
- [16] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
- [17] Y. Q. Huang, R. Li, W. B. Liu and N. N. Yan, Efficient discretization for finite element approximation of constrained optimal control problems, to appear.
- [18] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, 1991.
- [19] K. Kunisch, W. B. Liu and N. N. Yan, *A posteriori error estimates for a model parameter estimation problem*, Numerical mathematics and advanced applications, p. 723-730, Springer, Berlin, 2003.
- [20] O. A. Ladyzhenskaya and H. H. Ural'tseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York, 1968.
- [21] R. Li, on Multi-Mesh  $h$ -Adaptive Algorithm, *J. Sci. Comput.*, 24 (2005) 321-341. AFEPack is available online at <http://dsec.pku.edu.cn/~rli>.
- [22] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.
- [23] V. A. Morozov, *Methods for Solving Incorrectly Posed Problems*, Springer-Verlag, New York, 1984.
- [24] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research, Springer, New York, 1999.
- [25] S. Prudhomme and J. T. Oden, On goal-oriented error estimation for elliptic problems: Application to the control of pointwise errors, *Computer Methods in Applied Mechanics and Engineering*, 176 (1999) 313-331.
- [26] H. Steeb, A. Maute and E. Ramm, *Goal-oriented error estimation in solid mechanics*, Error-controlled adaptive finite elements in solid mechanics (Ed. E. Stein), p. 211-261, John Wiley & Sons, 2002.
- [27] N. Z. Sun, *Inverse Problems in Groundwater Modeling*, Kluwer Academic, Netherlands, 1994.
- [28] R. Verfürth, A Review of A Posteriori Error Estimation Techniques for Elasticity Problems, *Comput. Meth. Appl. Mech. Engrg.*, 176 (1999) 419-440.
- [29] G. Wahba, *Spline Models for Observation Data*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, SIAM, Philadelphia, 1990.
- [30] Y. F. Wang and Y. X. Yuan, Convergence and regularization of trust region methods for nonlinear ill-posed inverse problem, *Inverse Problem*, 21 (2005) 821-838

CIPR, University of Bergen, P.O. Box 7800, N-5020 Bergen, Norway  
*E-mail:* tao.feng@cipr.uib.no

Department of Engineering, Physics and Mathematics, Mid Sweden Univ., SE-851 70 Sundsvall, Sweden  
*E-mail:* marten.gulliksson@miun.se

Institute of Mathematics and Statistics, University of Kent, Canterbury, CT2 7NF, England  
*E-mail:* w.b.liu@kent.ac.uk