SIAM J. CONTROL OPTIM.
Vol. 39, No. 1, pp. 113–132

ADAPTIVE FINITE ELEMENT METHODS FOR OPTIMAL CONTROL OF PARTIAL DIFFERENTIAL EQUATIONS: BASIC CONCEPT*

ROLAND BECKER†, HARTMUT KAPP†, AND ROLF RANNACHER†

Abstract. A new approach to error control and mesh adaptivity is described for the discretization of optimal control problems governed by elliptic partial differential equations. The Lagrangian formalism yields the first-order necessary optimality condition in form of an indefinite boundary value problem which is approximated by an adaptive Galerkin finite element method. The mesh design in the resulting reduced models is controlled by residual-based a posteriori error estimates. These are derived by duality arguments employing the cost functional of the optimization problem for controlling the discretization error. In this case, the computed state and costate variables can be used as sensitivity factors multiplying the local cell-residuals in the error estimators. This results in a generic and simple algorithm for mesh adaptation within the optimization process. This method is developed and tested for simple boundary control problems in semiconductor models.

Key words. optimal control problem, finite elements, a posteriori error estimates, mesh adaptation, model reduction

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

PII. S0363012999351097

1. Introduction.

In this article, we develop an adaptive Galerkin finite element method for optimal control problems governed by elliptic partial differential equations. The main goal is the derivation of a posteriori error estimates as basis for guiding the mesh adaptation and for controlling the error in this model reduction. The problems considered have the form

$$J(u,q) \;\to\; \min!, \qquad A(u) = f + B(q), \tag{1.1}$$

where A is an elliptic differential operator for the state variable u, B an impact operator for the control variable q, and J is a cost functional. As prototypical examples, we will consider problems of boundary control in semiconductor models. Our approach utilizes the classical Lagrangian framework for reformulating the optimal control problem (1.1) as a boundary value problem for stationary points of the associated first-order necessary optimality condition. Introducing the Lagrangian functional

$$\mathcal{L}(u,\lambda,q) := J(u,q) + \langle \lambda, A(u,q) - B(q) - f \rangle, \tag{1.2}$$

with the costate variable λ (Lagrangian multiplier), the solutions of (1.1) are among the stationary points of $\mathcal{L}$, determined by the system of equations

$$\nabla \mathcal{L}(u,\lambda,q) = 0. \tag{1.3}$$

*Received by the editors January 27, 1999; accepted for publication (in revised form) May 3, 2000; published electronically August 17, 2000. This work was supported by the Deutsche Forschungsgemeinschaft (DFG), SFB 359 "Reaktive Strömungen, Diffusion und Transport," Universität Heidelberg.

http://www.siam.org/journals/sicon/39-1/35109.html

†Institut für Angewandte Mathematik, INF 294, Universität Heidelberg, D-69120 Heidelberg, Germany (roland.becker@iwr.uni-heidelberg.de, hartmut.kapp@iwr.uni-heidelberg.de, rolf.rannacher@iwr.uni-heidelberg.de).

We use a standard finite element method for discretizing this saddle-point problem which results in finite dimensional problems

$$\nabla \mathcal{L}(u_h, \lambda_h, q_h) = 0, \tag{1.4}$$

for the "discrete" states u_h, controls q_h, and costates λ_h. As long as the discretization procedure uses a pure Galerkin approach the discrete problem actually corresponds to a formulation of the original minimization problem on the discrete state space. Since discretization in partial differential equations is expensive, at least for praxis-relevant models, the question of how this "model reduction" affects the quality of the optimization result is crucial for a cost-efficient computation. The need for a posteriori error control is therefore evident.

The discretization of the state equation generally leads to approximate solutions $\{u_h, q_h\}$ which are *not admissible* in the strict sense for the original constrained minimization problem. Let S denote the solution operator which associates the state variable $u = u(q)$ to a given control function. The *optimal* control minimizes the functional $j(q) := J(S(q), q)$ for all controls. Then, discretization of the state equation also changes the functional. Denoting by S_h the discrete solution operator, the discrete optimal control q_h solves

$$j_h(q_h) := J(S_h(q_h), q_h) \;\rightarrow\; \min!. \tag{1.5}$$

If we want to perform numerical computation with controlled accuracy, we have to substitute the notion of an "admissible solution" by an error estimate for the state equation. Of course, the distance between the numerical and the exact solution should be measured with respect to the specific needs of the optimization problem, i.e., its effect on the functional to be minimized. This asks for a sensitivity analysis for the optimization problem with respect to perturbations in the state equation, particularly perturbations resulting from discretization. In this sense, our a posteriori error estimation aims to control the error due to replacing the infinite dimensional problem (1.1) by its finite dimensional analogue. The crucial question is now which quality measure is appropriate for controlling the discretization error. In general, forcing this error to be small uniformly in the whole computational domain, as is often required in ODE models, is not feasible for partial differential equations. Therefore, we need to develop control of the discretization error in accordance with the sensitivity properties of the optimization problem.

Our approach to this problem uses the general method developed in [3] and [4] for error control in the Galerkin finite element discretization of differential equations of the general form $A(u) = f$. Employing the linearized dual problem

$$A'(u_h)^* z = F(\cdot) \tag{1.6}$$

for an arbitrary output functional F, an a posteriori error estimate

$$|F(u) - F(u_h)| \le \eta_\omega(u_h) := \sum_{T \in \mathbb{T}_h} \rho_T(u_h)\, \omega_T(z) \tag{1.7}$$

can be derived. On a computational mesh $\mathbb{T}_h = \{T\}$ consisting of cells T, the local consistency errors, expressed in terms of residuals $\rho_T(u_h)$, are multiplied by weights $\omega_T(z)$ involving the "dual solution" z. These weights describe the dependence of the error on variations of the local residuals, i.e., on the local mesh size. In general the

estimate (1.7) has to be approximated by numerically solving the dual problem (1.6). This results in a feedback process for generating successively more and more accurate error bounds and solution-adapted meshes. In applying this approach to saddle-point problems arising from optimal control problems, it seems natural to base the error control on the given cost functional, i.e., to choose $F := J$. In this particular case the corresponding approximate dual solution can be expressed in terms of the computed solution $\{u_h, \lambda_h, q_h\}$. Hence, the evaluation of the corresponding a posteriori error estimate

$$|J(u,q) - J(u_h,q_h)| \le \eta_\omega(u_h,\lambda_h,q_h) \tag{1.8}$$

does not require much extra work and a posteriori error estimation is almost for free. This leads to a generic and simple strategy for mesh adaptation in discretizing optimal control problems.

It may be seen as a drawback that in this approach the accuracy in the discretization of the state equation is only controlled with respect to its effect on the cost functional. This can lead to discrete models which approximate the original optimization problem with minimal cost but the obtained discrete states and controls are "admissible" only in a very weak sense, possibly insufficient for the particular application. If satisfaction of the state equation is desired in a stronger sense, we can combine our method with traditional "energy-error control" leading to an a posteriori error estimate of the form

$$|J(u,q) - J(u_h,q_h)| + \beta \|A(u_h) - f - B(q_h)\|_* \ \le \ \eta_{\omega,E}(u_h,\lambda_h,q_h), \tag{1.9}$$

where $\|\cdot\|_*$ denotes the dual of the natural "energy norm" corresponding to the operator $A'(u)$, and β is a tuning factor. For a discussion of adaptive finite element methods using residual-based a posteriori error estimates, we refer to the survey papers [1], [14], [8], and [12].

First, we develop our approach within a general setting in order to abstract from inessential technicalities. Then, all steps are made concrete for a linear model problem of boundary control. Despite its simplicity this problem represents the main structure of optimal control and is chosen in order to clarify the idea underlying the proposed procedure. Some numerical results illustrate the main features of the adaptive algorithm particularly in comparison to more conventional methods based on global error control for the state equation. At the end, we extend our method to problems with nonlinear state equations with an example of boundary control in semiconductor models.

2. A linear model situation. We consider an abstract setting for optimal control: Let Q, V, and H be Hilbert spaces for the control variable $q \in Q$, the state variable $u \in V$, and given observations $c_0 \in H$. The inner product and norm of H are $(\cdot,\cdot)$ and $\|\cdot\|$, respectively. The state equation is given in the form

$$a(u,\varphi) = (f,\varphi) + b(q,\varphi) \quad \forall \varphi \in V, \tag{2.1}$$

where the bilinear form $a(\cdot,\cdot)$ represents a linear elliptic operator and the bilinear form $b(\cdot,\cdot)$ expresses the action of the control. The goal is to minimize the cost functional

$$J(u,q) = \tfrac{1}{2}\|cu - c_0\|^2 + \tfrac{1}{2}n(q,q), \tag{2.2}$$

where $c : V \to H$ is a linear bounded observation operator. For simplicity, we assume that $a(\cdot,\cdot)$ and $n(\cdot,\cdot)$ induce norms on the spaces V and Q denoted by $\|\cdot\|_a$ and $\|\cdot\|_n$, respectively. This guarantees the existence of a unique solution of the optimal control problem and the classical regularity theory for elliptic equations applies (see, e.g., [11]).

Introducing a Lagrangian parameter $\lambda \in V$ and the Lagrangian function $\mathcal{L}(u,q,\lambda)$,

$$\mathcal{L}(u,q,\lambda) := J(u,q) + a(u,\lambda) - b(q,\lambda) - (f,\lambda),$$

the first-order necessary conditions (Euler–Lagrange equations) of the optimization problem,

$$\nabla\mathcal{L}(u,q,\lambda)(v,\mu,r) = 0 \quad \forall\{v,\mu,r\} \in V \times V \times Q \tag{2.3}$$

have the explicit form

$$\begin{aligned} a(v,\lambda) + (cu - c_0, cv) &= 0 \quad \forall v \in V, \\ a(u,\mu) - b(q,\mu) &= (f,\mu) \quad \forall \mu \in V, \\ -b(r,\lambda) + n(q,r) &= 0 \quad \forall r \in Q. \end{aligned} \tag{2.4}$$

This system has the usual saddle-point structure

$$\begin{aligned} (cu,cv) + a(v,\lambda) &= (c_0,cv) \quad \forall v \in V, \\ a(u,\mu) - b(q,\mu) &= (f,\mu) \quad \forall \mu \in V, \\ -b(r,\lambda) + n(q,r) &= 0 \quad \forall r \in Q. \end{aligned} \tag{2.5}$$

Introducing operators A, B, C, N which represent the corresponding bilinear forms, system (2.5) can also be written in matrix form as

$$\begin{bmatrix} C & A^T & 0 \\ A & 0 & -B \\ 0 & -B^T & N \end{bmatrix} \begin{bmatrix} u \\ \lambda \\ q \end{bmatrix} = \begin{bmatrix} c_0 \\ f \\ 0 \end{bmatrix}. \tag{2.6}$$

Below, we will consider the following realization of the foregoing abstract setting which represents the case of an elliptic linear state equation subjected to boundary control. Let $\Omega \subset \mathbb{R}^2$ be an open bounded domain with Lipschitz boundary $\partial\Omega$ which is decomposed into a homogeneous Neumann part Γ_N and a control part Γ_C on which the control acts ($\partial\Omega = \Gamma_C \cup \Gamma_N$),

$$\begin{aligned} -\Delta u + u &= f \quad \text{in } \Omega, \\ \partial_n u = 0 \ \text{ on } \Gamma_N, &\quad \partial_n u = q \ \text{ on } \Gamma_C. \end{aligned} \tag{2.7}$$

The observations are given on a part Γ_O of the boundary and the associated cost functional is

$$J(u,q) = \tfrac{1}{2}\|u - c_0\|^2_{\Gamma_O} + \tfrac{\alpha}{2}\|q\|^2_{\Gamma_C} \tag{2.8}$$

with a regularization parameter $\alpha > 0$. In this case the natural choice for the function spaces is $V = H^1(\Omega)$, the first-order Sobolev Hilbert-space over Ω, and $H = L^2(\Gamma_O)$, $Q = L^2(\Gamma_C)$, the usual Lebesgue Hilbert-spaces over Γ_C and Γ_O, respectively. The bilinear forms $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ and $n(\cdot,\cdot)$ are given by

$$a(u,v) = (\nabla u, \nabla v)_\Omega + (u,v)_\Omega, \quad b(q,v) = (q,v)_{\Gamma_C}, \quad n(q,r) = \alpha(q,r)_{\Gamma_C},$$

where $(\cdot,\cdot)_\Sigma$ denotes the L^2-inner product on the set Σ. The operator c in the cost functional is the trace operator, $cu = u_{|\Gamma_O}$. Then, the necessary optimality condition $\nabla\mathcal{L}(u,\lambda,q) = 0$ reads as follows:

$$\begin{aligned}(u,v)_{\Gamma_O} - (c_0,v)_{\Gamma_O} + (\nabla v,\nabla\lambda)_\Omega + (v,\lambda)_\Omega &= 0 \quad \forall v\in V,\\ (\nabla u,\nabla\mu)_\Omega + (u,\mu)_\Omega - (f,\mu)_\Omega - (q,\mu)_{\Gamma_C} &= 0 \quad \forall \mu\in V,\\ \alpha(q,r)_{\Gamma_C} - (\lambda,r)_{\Gamma_C} &= 0 \quad \forall r\in Q.\end{aligned} \tag{2.9}$$

3. A priori error estimate. For simplicity of notation, we introduce the product space $X = V\times V\times Q$, with elements of the form $x = \{u,\lambda,q\}$, which is equipped with the natural norm

$$\|x\|_X := \left(\|u\|_a^2 + \|\lambda\|_a^2 + \|q\|_n^2\right)^{1/2}.$$

Furthermore, we define a bilinear form $A(\cdot,\cdot)$ on X by

$$A(x,y) = A(\{u,\lambda,q\},\{v,\mu,r\}) := (cu,c\mu) + a(u,v) - b(q,v) + a(\mu,\lambda) - b(r,\lambda) + n(q,r).$$

Using this notation, system (2.5) can be written in compact form as

$$A(x,y) = F(y) \quad \forall y\in X, \tag{3.1}$$

with the linear functional $F(\cdot)$ defined by

$$F(y) = F(\{v,\mu,r\}) := (c_0,c\mu) + (f,v).$$

In order to simplify the analysis, we impose the following conditions:

$$|A(x,y)| \le c_A\|x\|_X\|y\|_X, \tag{3.2}$$

$$|b(r,v)| \le c_b\|r\|_n\|v\|_a. \tag{3.3}$$

The second condition, which relies on the presence of the regularization term $n(\cdot,\cdot)$ (requiring that $\alpha > 0$ in the above example), is rather strong. It can be substituted by an "inf-sup" condition for $b(\cdot,\cdot)$ under which the regularization could be omitted; see Remark 3.1, below. The bilinear form $A(\cdot,\cdot)$ satisfies the following stability condition.

PROPOSITION 3.1. *Under the assumptions* (3.2) *and* (3.3)*, there exists a constant* $\gamma > 0$ *such that*

$$\inf_{x\in X}\left\{\sup_{y\in X}\frac{A(x,y)}{\|x\|_X\|y\|_X}\right\} \ge \gamma. \tag{3.4}$$

Proof. For any fixed $x = \{u,\lambda,q\}$, we choose the test triple $y = \{v,\mu,r\} := \{u,\lambda,q\}$ to obtain

$$\begin{aligned}A(x,y) &= \|cu\|^2 + \|u\|_a^2 + \|\lambda\|_a^2 + \|q\|_n^2 - b(q,\lambda) - b(q,u)\\ &\ge \|cu\|^2 + \|u\|_a^2 + \|\lambda\|_a^2 + \|q\|_n^2 - \tfrac14\|q\|_n^2 - \tfrac34\|\lambda\|_a^2 - \tfrac14\|q\|_n^2 - \tfrac34\|u\|_a^2\\ &\ge \|cu\|^2 + \tfrac14\|u\|_a^2 + \tfrac14\|\lambda\|_a^2 + \tfrac12\|q\|_n^2.\end{aligned}$$

We conclude the asserted estimate by noting that $\|y\| = \|x\|$. ∎

We consider the discretization of the variational equation (3.1) by a standard Galerkin method using trial spaces $X_h := V_h\times V_h\times Q_h \subset X$. For each $x\in X$, there

shall exist an interpolation $i_h x \in X_h$, such that $\|x - i_h x\|_X \to 0$ $(h \to 0)$. Then, approximations $x_h \in X_h$ are determined by

$$A(x_h, y_h) = F(y_h) \quad \forall y_h \in X_h. \tag{3.5}$$

This discretization is automatically stable since a discrete analogue of (3.4) is fulfilled by the same argument as used above,

$$\inf_{x_h \in X_h} \left\{ \sup_{y_h \in X_h} \frac{A(x_h, y_h)}{\|x_h\|_X \|y_h\|_X} \right\} \geq \gamma > 0. \tag{3.6}$$

Combining (3.5) and (3.1), we get the Galerkin orthogonality

$$A(x - x_h, y_h) = 0, \quad y_h \in X_h. \tag{3.7}$$

This leads us to the following abstract a priori error estimate.

PROPOSITION 3.2. *For the Galerkin approximation in $X_h \subset X$, there holds*

$$\begin{aligned} \|u - u_h\|_a + \|\lambda - \lambda_h\|_a + \|q - q_h\|_n \\ \leq c \Big\{ \inf_{\mu_h \in V_h} \|u - \mu_h\|_a + \inf_{v_h \in V_h} \|\lambda - v_h\|_a + \inf_{r_h \in Q_h} \|q - r_h\|_n \Big\}. \end{aligned} \tag{3.8}$$

Proof. The stability estimate (3.6) implies that

$$\gamma \|i_h x - x_h\| \leq \sup_{y_h \in X_h} \frac{A(i_h x - x_h, y_h)}{\|y_h\|_X} = \sup_{y_h \in X_h} \frac{A(i_h x - x, y_h)}{\|y_h\|_X} \leq c_A \|i_h x - x\|_X .$$

Here, we have used the Galerkin relation (3.7) and the continuity estimate (3.2). ∎

REMARK 3.1. *Of course, more precise error estimates could be given exploiting the structure of the underlying problem. For instance, it may be possible to equip the space Q with a different norm than the one induced by $n(\cdot, \cdot)$, in order to get robustness with respect to the regularization. This requires us to replace* (3.3) *by the following weaker inf-sup condition*

$$\inf_{q \in Q} \left\{ \sup_{v \in V} \frac{b(q, v)}{\|v\|_a} \right\} \geq \kappa > 0.$$

We note that for the model example with boundary control and boundary observations given above the conditions (3.2) *and* (3.3) *are satisfied.*

4. A posteriori error estimate. In this section, we derive an a posteriori error estimate for the model control problem. As discussed above, the error estimator to be derived should control the value of the cost functional. First, we carry out the analysis in the abstract functional analytic setting. Recalling the definition of the Lagrange functional $\mathcal{L}$,

$$\mathcal{L}(x) = \mathcal{L}(u, \lambda, q) = J(u, q) + a(u, \lambda) - (f, \lambda) - b(q, \lambda),$$

the continuous and discrete optimal control solutions $x = \{u, \lambda, q\} \in X$ and $x_h = \{u_h, \lambda_h, q_h\} \in X_h$ satisfy

$$\nabla \mathcal{L}(x)(\varphi) = 0, \quad \varphi \in X, \qquad \nabla \mathcal{L}(x_h)(\varphi_h) = 0, \quad \varphi_h \in X_h. \tag{4.1}$$

This implies the Galerkin orthogonality relation

$$\nabla^2\mathcal{L}(x - x_h, \varphi_h) = 0, \quad \varphi_h \in X_h, \tag{4.2}$$

for which it is essential that $\mathcal{L}$ is quadratic. Since the solutions u and u_h satisfy the corresponding state equations, the cost functional and the Lagrangian functional are related by

$$J(u, q) - J(u_h, q_h) = \mathcal{L}(x) - \mathcal{L}(x_h). \tag{4.3}$$

Further, there holds

$$\mathcal{L}(x) - \mathcal{L}(x_h) = \nabla\mathcal{L}(x)(x - x_h) - \tfrac{1}{2}\nabla^2\mathcal{L}(x - x_h, x - x_h). \tag{4.4}$$

The first term on the right-hand side vanishes since x is a stationary point of $\mathcal{L}$. Using the Galerkin orthogonality relation (4.2) in the second term, we obtain for arbitrary $\varphi_h \in X_h$ that

$$\nabla^2\mathcal{L}(x - x_h, x - x_h) = \nabla^2\mathcal{L}(x - x_h, x - x_h - \varphi_h) = -\nabla\mathcal{L}(x_h)(x - x_h - \varphi_h). \tag{4.5}$$

Therefore, choosing $\varphi_h = i_h x - x_h$, we get the following error representation.

PROPOSITION 4.1. *For the abstract model problem with linear state equation and quadratic cost functional, the following error identity holds:*

$$J(u, q) - J(u_h, q_h) = \tfrac{1}{2}\nabla\mathcal{L}(x_h)(x - i_h x). \tag{4.6}$$

In order to convert the abstract error identity (4.6) into a form which can be evaluated, we need to be more specific about the setting of the underlying problem and its discretization. As an example, we demonstrate this for the linear Neumann control problem described by (2.7) and (2.8). Here, the Galerkin finite element discretization of the saddle-point problem (2.9) uses subspaces $V_h \subset V = H^1(\Omega)$ and $Q_h \subset Q = L^2(\Gamma_C)$ of piecewise polynomial functions defined on regular decompositions $\mathbb{T}_h = \{T\}$ of the domain Ω into cells T (triangles or quadrilaterals); for a detailed description of such a setting see, e.g., Brenner and Scott [5]. Here, we use quadrilateral meshes where on each cell T the local shape functions are constructed by mapping bilinear functions defined on a reference square to the cell T. This ansatz is referred to as the "isoparametric" bilinear finite element. We assume that the space Q_h of discrete controls is given by the traces along Γ_C of the finite element functions of V_h. This is not necessary for our results but simplifies the notation. In order to avoid the technicalities caused by curved boundaries, we suppose the domain Ω to be polygonal. We use the notation $h_T := diam(T)$ and $h_\Gamma := diam(\Gamma)$ for the width of a cell $T \in \mathbb{T}_h$ and a corresponding cell edge $\Gamma \subset \partial T$. In order to ease local mesh refinement and coarsening, "hanging nodes" are allowed, but at most one per cell edge; see Figure 4.1. The degree of freedom at such a hanging node is eliminated by interpolation in order to keep the discretization "conforming" (see, e.g., [6] and [3]).

First, we turn (4.6) into a residual based a posteriori estimate as follows.

PROPOSITION 4.2. *For error control with respect to the cost functional J, there holds the weighted a posteriori error estimate*

$$|J(u, q) - J(u_h, q_h)| \;\le\; \eta_\omega(u_h, \lambda_h, q_h) = \sum_{T\in\mathbb{T}_h} \eta_T(u_h, \lambda_h, q_h), \tag{4.7}$$

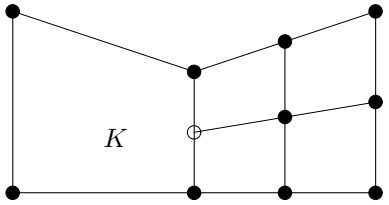

FIG. 4.1. *Quadrilateral mesh patch with a "hanging node."*

with the local error indicators

$$\eta_T(u_h,\lambda_h,q_h) := \rho_T^{(u)}\,\omega_T^{(\lambda)} + \rho_{\partial T}^{(u)}\,\omega_{\partial T}^{(\lambda)} + \rho_T^{(\lambda)}\,\omega_T^{(u)} + \rho_{\partial T}^{(\lambda)}\,\omega_{\partial T}^{(u)} + \rho_{\partial T}^{(q)}\,\omega_{\partial T}^{(q)}$$

and the cellwise residuals and weights

$$\begin{aligned}
\rho_T^{(u)} &:= \|R_h^{(u)}\|_T, & \omega_T^{(\lambda)} &:= \|\lambda - i_h\lambda\|_T,\\
\rho_{\partial T}^{(u)} &:= \|r_h^{(u)}\|_{\partial T}, & \omega_{\partial T}^{(\lambda)} &:= \|\lambda - i_h\lambda\|_{\partial T},\\
\rho_T^{(\lambda)} &:= \|R_h^{(\lambda)}\|_T, & \omega_T^{(u)} &:= \|u - i_h u\|_T,\\
\rho_{\partial T}^{(\lambda)} &:= \|r_h^{(\lambda)}\|_{\partial T}, & \omega_{\partial T}^{(u)} &:= \|u - i_h u\|_{\partial T},\\
\rho_{\partial T}^{(q)} &:= \|r_h^{(q)}\|_{\partial T\cap\Gamma_C}, & \omega_{\partial T}^{(q)} &:= \|q - j_h q\|_{\partial T\cap\Gamma_C}.
\end{aligned}$$

The "cell residuals" $R_h^{(u)}$, $R_h^{(\lambda)}$ and the "edge residuals" $r_h^{(u)}$, $r_h^{(\lambda)}$, $r_h^{(q)}$ are on cells T and cell edges Γ defined by

$$R_{h|T}^{(u)} := -\Delta u_h + u_h - f, \quad R_{h|T}^{(\lambda)} := -\Delta\lambda_h + \lambda_h, \qquad r_{h|\Gamma}^{(q)} := \alpha q_h - \lambda_h \ \text{ if } \Gamma\subset\Gamma_C,$$

$$r_{h|\Gamma}^{(u)} := \begin{cases} \frac{1}{2}h_\Gamma^{-1/2}[\partial_n\varphi_h] & \text{if } \Gamma\subset\partial T\setminus\partial\Omega,\\ h_\Gamma^{-1/2}\partial_n u_h & \text{if } \Gamma\subset\partial\Omega\setminus\Gamma_C,\\ h_\Gamma^{-1/2}(\partial_n u_h - q_h) & \text{if } \Gamma\subset\Gamma_C, \end{cases} \qquad r_{h|\Gamma}^{(\lambda)} := \begin{cases} \frac{1}{2}h_\Gamma^{-1/2}[\partial_n\varphi_h] & \text{if } \Gamma\subset\partial T\setminus\partial\Omega,\\ h_\Gamma^{-1/2}\partial_n\lambda_h & \text{if } \Gamma\subset\partial\Omega\setminus\Gamma_O,\\ h_\Gamma^{-1/2}(c_0 - u_h + \partial_n\lambda_h) & \text{if } \Gamma\subset\Gamma_O. \end{cases}$$

Here, $[\partial_n\varphi_h]$ denotes the jump of the normal derivative of φ_h across the interelement edges Γ, the boundary components Γ_C, Γ_O are the control and observation boundary, respectively, and i_h, j_h denote some local interpolation operators into the finite element spaces.

Proof. From the abstract error identity (4.6), we obtain that

$$\begin{aligned}
J(u,q) - J(u_h,q_h) &= \tfrac{1}{2}\nabla\mathcal{L}(x_h)(x - i_h x)\\
&= \tfrac{1}{2}(c_0 - u_h, u - i_h u)_{\Gamma_O} + \tfrac{1}{2}a(u - i_h u, \lambda_h)\\
&\quad + \tfrac{1}{2}a(u_h, \lambda - i_h\lambda) - \tfrac{1}{2}(f, \lambda - i_h\lambda) - \tfrac{1}{2}b(q_h, \lambda - i_h\lambda)\\
&\quad + \tfrac{1}{2}n(q_h, q - j_h q) - \tfrac{1}{2}b(q - j_h q, \lambda_h)\\
&=: I_\lambda + I_u + I_q.
\end{aligned}$$

Recalling the definition of the bilinear forms and integrating cellwise by parts, the

first term I_λ is rewritten as follows:

$$\begin{aligned}
2I_\lambda &= (c_0 - u_h, u - i_h u)_{\Gamma_O} + a(u - i_h u, \lambda_h) \\
&= (c_0 - u_h, u - i_h u)_{\Gamma_O} + (\nabla(u - i_h u), \nabla\lambda_h) + (u - i_h u, \lambda_h) \\
&= (c_0 - u_h + \partial_n\lambda_h, u - i_h u)_{\Gamma_O} + (u - i_h u, \partial_n\lambda_h)_{\partial\Omega\setminus\Gamma_O} \\
&\quad + \sum_{T\in\mathbb{T}_h} \Big\{ (u - i_h u, -\Delta\lambda_h + \lambda_h)_T + (u - i_h u, \partial_n\lambda)_{\partial T\setminus\partial\Omega} \Big\} \\
&= (c_0 - u_h + \partial_n\lambda_h, u - i_h u)_{\Gamma_O} + (u - i_h u, \partial_n\lambda_h)_{\partial\Omega\setminus\Gamma_O} \\
&\quad + \sum_{T\in\mathbb{T}_h} \Big\{ (u - i_h u, -\Delta\lambda_h + \lambda_h)_T + \tfrac{1}{2}(u - i_h u, [\partial_n\lambda])_{\partial T\setminus\partial\Omega} \Big\}.
\end{aligned}$$

Hence using the definition of the residuals $r_h^{(\lambda)}$ and $R_h^{(\lambda)}$, we find

$$\begin{aligned}
2I_\lambda &= (u - i_h u, r_h^{(\lambda)})_{\partial\Omega} + \sum_{T\in\mathbb{T}_h} \Big\{ (u - i_h u, R_h^{(\lambda)})_T + (u - i_h u, r_h^{(\lambda)})_{\partial T\setminus\partial\Omega} \Big\} \\
&= \sum_{T\in\mathbb{T}_h} \Big\{ (u - i_h u, R_h^{(\lambda)})_T + (u - i_h u, r_h^{(\lambda)})_{\partial T} \Big\},
\end{aligned}$$

and, consequently by the Cauchy–Schwarz inequality,

$$2|I_\lambda| \le \sum_{T\in\mathbb{T}_h} \Big\{ \|u - i_h u\|_T \, \|R_h^{(\lambda)}\|_T + \|u - i_h u\|_{\partial T} \, \|r_h^{(\lambda)}\|_{\partial T} \Big\}.$$

In the same way, we get for the other terms I_λ and I_q:

$$\begin{aligned}
2I_u &= a(u_h, \lambda - i_h\lambda) - (f, \lambda - i_h\lambda) - b(q_h, \lambda - i_h\lambda) \\
&= (\nabla u_h, \nabla(\lambda - i_h\lambda)) + (u_h, \lambda - i_h\lambda) - (f, \lambda - i_h\lambda) - (q_h, \lambda - i_h\lambda)_{\Gamma_C} \\
&= \sum_{T\in\mathbb{T}_h} \Big\{ (-\Delta u_h + u_h - f, \lambda - i_h\lambda)_T + \tfrac{1}{2}([\partial_n u_h], \lambda - i_h\lambda)_{\partial T\setminus\partial\Omega} \Big\} \\
&\quad + (\partial_n u_h, \lambda - i_h\lambda)_{\partial\Omega\setminus\Gamma_C} + (\partial_n u_h - q_h, \lambda - i_h\lambda)_{\Gamma_C} \\
&= \sum_{T\in\mathbb{T}_h} \Big\{ (R_h^{(u)}, \lambda - i_h\lambda)_T + (r_h^{(u)}, \lambda - i_h\lambda)_{\partial T\setminus\partial\Omega} \Big\} \\
&\quad + (r_h^{(u)}, \lambda - i_h\lambda)_{\partial\Omega\setminus\Gamma_C} + (r_h^{(u)}, \lambda - i_h\lambda)_{\Gamma_C} \\
&= \sum_{T\in\mathbb{T}_h} \Big\{ (R_h^{(u)}, \lambda - i_h\lambda)_T + (r_h^{(u)}, \lambda - i_h\lambda)_{\partial T} \Big\},
\end{aligned}$$

$$\begin{aligned}
2I_q &= n(q_h, q - j_h q) - b(q - j_h q, \lambda_h) = (\alpha q_h - \lambda_h, q - j_h q)_{\Gamma_C} \\
&= \sum_{\Gamma\subset\Gamma_C} (r_h^{(q)}, q - j_h q)_\Gamma,
\end{aligned}$$

and, consequently,

$$\begin{aligned}
2|I_u| &\le \sum_{T\in\mathbb{T}_h} \Big\{ \|R_h^{(u)}\|_T \, \|\lambda - i_h\lambda\|_T + \|r_h^{(u)}\|_{\partial T} \, \|\lambda - i_h\lambda\|\|_{\partial T} \Big\}, \\
2|I_q| &\le \sum_{\Gamma\subset\Gamma_C} \|r_h^{(q)}\|_\Gamma \, \|q - j_h q\|_\Gamma .
\end{aligned}$$

Collecting these estimates implies the asserted result. ∎

REMARK 4.1. *We note that in the a posteriori error estimate* (4.7)*, the residual of the state equation is weighted by terms involving the adjoint variable* λ *from the original equation* (2.5)*. This has a natural interpretation as it is well known from sensitivity analysis that the adjoint variable is a measure for the influence of perturbations on the cost functional. Since discretization can be interpreted as a special kind of perturbation, the appearance of* λ *in the estimator is not surprising. The special form of the weights involving the interpolation* $i_h z$ *is a characteristic feature of the Galerkin discretization.*

REMARK 4.2. *The a posteriori error estimate* (4.7) *is easier to understand if one recalls the model situation of approximating the boundary value problem*

$$-\Delta u = f \ \ in\, \Omega, \quad u = 0 \ \ on\, \partial\Omega,$$

by a Galerkin finite element method. The natural variational formulation of this problem is equivalent to an unconstraint optimization problem, namely, the minimization of the "energy functional" $J(u) := \frac{1}{2}\|\nabla u\|_\Omega^2 - (f,u)_\Omega$ *over the solution space* $V := H_0^1(\Omega)$*. In this context, in view of the identity*

$$J(u) - J(u_h) = -\tfrac{1}{2}\|\nabla e\|_\Omega^2,$$

error control with respect to the energy functional is equivalent to control of the error in the energy norm, $\|\nabla e\|_\Omega$*. It is well known that the latter can be achieved without referring to an additional dual problem since in this case the corresponding dual solution coincides with the error* e *itself; see* [3] *for a discussion of energy-error control in the context of "duality techniques."*

Evaluation of error estimators. The a posteriori error estimate (4.7) still involves the *continuous* solutions $\{u, \lambda, q\}$. As proposed in [4], we use the computed solutions $\{u_h, \lambda_h, q_h\}$ for approximating the weights $\omega_{\partial T}^{(\cdot)}$ and $\omega_T^{(\cdot)}$. To this end, we recall the well-known local approximation properties of finite elements, e.g.,

$$\|u - i_h u\|_T + h_\Gamma^{1/2}\|u - i_h u\|_\Gamma \le c_I h_T \,\|\nabla^2 u\|_T, \tag{4.8}$$

where $\Gamma \subset \partial T$, and i_h is the generic operator of cellwise nodal interpolation into V_h, with interpolation constant usually in the range $c_I \sim 0.1-1$; for details of the interpolation theory for finite elements, we refer to [5]. Analogous estimates hold for the terms involving q and λ. Then, the derivatives are approximated by suitable difference quotients, e.g.,

$$\|\nabla^2 u\|_T \approx \|\nabla_h^2 u_h\|_T. \tag{4.9}$$

For a more detailed discussion of this evaluation of weights and some of its alternatives, we refer to [4]. We emphasize that the proposed procedure for evaluating the a posteriori error estimate (4.7) uses only information in terms of the already computed solution $\{u_h, \lambda_h, q_h\}$.

Error control for the state variable. The estimate (4.7) provides control of the error with respect to the cost functional which is the quantity of primary interest in the optimization problem. But this does not include control of the error in the state equation. Though the corresponding residuals $\rho_T^{(u)}$ and $\rho_{\partial T}^{(u)}$ are present in the error estimator, they are weighted according to their effect on the cost functional. In

case that the discrete state u_h is required to be admissible in a stronger sense, the error estimate (4.7) can be extended to also include control of the error in satisfying the state equation measured in the natural energy norm

$$\|u\|_E := \left(\|u\|^2 + \|\nabla u\|^2\right)^{1/2}.$$

The standard a posteriori error analysis for the boundary value problem

$$-\Delta u + u = f \ \text{ in } \Gamma_C, \quad \partial_n u = 0 \ \text{ on } \Gamma_N, \quad \partial_n u = q \ \text{ on } \Gamma_C, \tag{4.10}$$

for frozen control q, yields the following bound for the "energy error":

$$\|u - u_h\|_E^2 \ \le \ \eta_E(u_h) := c \sum_{T \in \mathbb{T}_h} \left\{ \rho_T^{(u)2} + \rho_{\partial T}^{(u)2} \right\}, \tag{4.11}$$

with the residuals as defined in Proposition 4.2. For the derivation of this estimate see, for example, [14], [4], and the literature cited therein. We see that all terms of $\eta_E(u_h)$ appear also in $\eta_\omega(u_h, \lambda_h, q_h)$, but are weighted in terms of the adjoint variable. The effect of this modification will be illustrated below by a numerical test. In practice, we may use a combination of our weighted error estimator and the energy-error estimator $\eta_{\omega,E}(u_h, \lambda_h, q_h) := \eta_\omega(u_h, \lambda_h, q_h) + \beta\eta_E(u_h)$, with a suitable weighting factor $\beta \ge 0$.

Strategies for mesh adaptation. Several strategies are possible for mesh adaptation on the basis of a posteriori error estimators η as developed above. Usually, a certain tolerance TOL for the error in the quantity $J(u, q)$ and an upper bound for the complexity of the discrete model, i.e. the maximum number of mesh cells $N_{\max}$, are given. It is assumed that an "optimal" mesh-size distribution is achieved if the local error indicators η_T are equilibrated over the mesh $\mathbb{T}_h$. This suggests use of the "error-balancing strategy"; i.e., we cycle through the mesh and try to equilibrate the local error indicators η_T according to $\eta_T \approx TOL/N$. This process requires iteration with respect to the number of mesh cells N and eventually results in $\eta \approx TOL$. However, this strategy may lead to very slow mesh refinement and is very delicate to use. More robust is the "fixed-fraction strategy" in which we order the cells according to the size of η_T and refine a certain threshold $X\%$ of cells with largest η_T (or those cells which contribute to a certain percentage of the error estimator η). A certain fraction $Y\%$ of cells with small η_T may be coarsened. By this strategy, one can achieve a prescribed rate of increase of N (or keep it constant as may be desirable in nonstationary computations). In the test computations described below the second version of the "Fixed-Fraction Strategy" has been used with threshold 30%.

5. Numerical results—linear case. We present a linear model problem as described in (2.7), where Ω is a T-shaped domain with maximum side length one; see Figure 5.1 (left). In this example the control acts along the lower boundary Γ_C, whereas the observation is taken along the upper boundary Γ_O. The cost functional is chosen as

$$J(u, q) := \tfrac{1}{2}\|u - c_0\|_{\Gamma_O}^2 + \tfrac{\alpha}{2}\|q\|_{\Gamma_C}^2,$$

with $c_0 \equiv 1.0$ and $\alpha = 1.0$. In this case, the regularization term $\frac{\alpha}{2}\|q\|_{\Gamma_C}^2$ may be viewed as part of the cost functional with its own physical meaning. We perform computations on a series of locally refined meshes. On each mesh, the system of the first-order necessary condition (2.5) is discretized by the Galerkin finite element

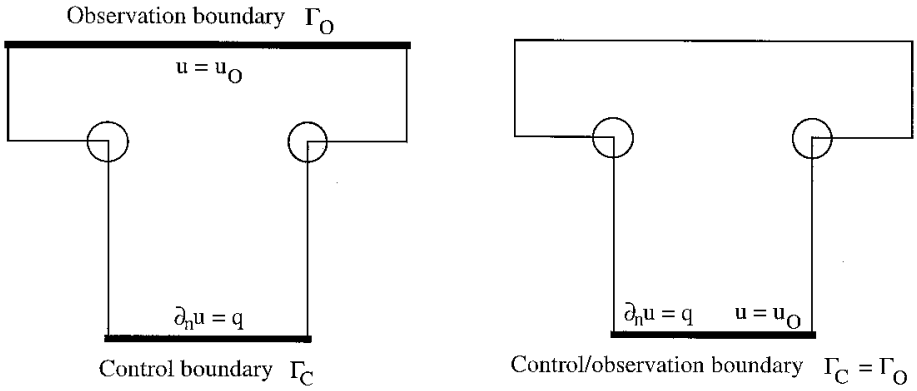

FIG. 5.1. *Configuration of the boundary control model problem on a T-domain (Ginzburg–Landau model): configuration* 1 *(left), configuration* 2 *(right).*

TABLE 5.1
Linear test (configuration 1*): Efficiency of the weighted error estimator.*

N	320	1376	4616	11816	23624	48716
E_h	$1.0e-3$	$3.5e-4$	$3.2e-5$	$1.6e-5$	$6.4e-6$	$2.8e-6$
I_{eff}	1.1	0.7	0.7	1.0	0.8	0.7

method described above. The resulting discrete saddle-point problems are solved iteratively by a GMRES method with multigrid preconditioning. The adaptive mesh refinement is based on an a posteriori error estimator. The weights in the error estimator (4.7) are evaluated using the strategy indicated in (4.8) and (4.9), with an interpolation constant set to $C_I = 0.1$. The mesh refinement uses the "fixed-fraction strategy" described above.

Table 5.1 shows the quality of the error estimator (4.7) for quantitative error control. The *efficiency index* is defined by $I_{eff} := E_h/\eta_h$, where $E_h := |J(u,q) - J(u_h,q_h)|$ is the error in the cost functional and $\eta_h := \eta(u_h,q_h)$ the value of the error estimator used. The reference value is obtained on a mesh with more than $200,000$ cells. We compare the weighted error estimator with a simple ad hoc approach based on the standard energy-error estimator (4.11) for the state equation. Figure 5.2 shows the computed "optimal" states over the meshes generated by the two different error estimators.

The two meshes are quite different: The energy-error estimator overemphasizes the steep gradients near the control boundary and it leaves the mesh too coarse along the observation boundary. The more selective *weighted* error estimator concentrates the mesh cells where they are needed for the optimization process. The quantitative effects on the mesh efficiency of these two different refinement criteria is shown in Figure 5.3 (E_h versus N in log/log-scale).

Finally, we check how well the approximation $\{u_h, \lambda_h, q_h\}$ obtained by the weighted error estimator (4.7) actually satisfies the state equation; for this the *global* energy-error estimator (4.11) is taken as quality measure. Table 5.2 shows a comparison of the two sequences of meshes generated by the weighted error estimator $\eta_\omega = \eta_\omega(u_h,\lambda_h,q_h)$ ("ω-meshes") and the energy-error estimator $\eta_E = \eta_E(u_h)$ ("E-meshes"). The first and second columns contain the values of η_ω and η_E on ω-meshes, while the third and fourth columns contain the values of η_ω and η_E on E-meshes.

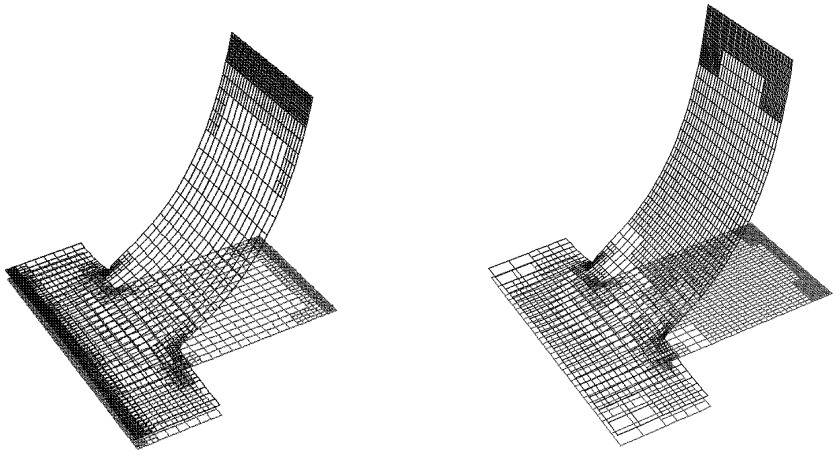

FIG. 5.2. *Linear test: Comparison of discrete solutions obtained by the weighted error estimator (left,* $N \sim 1600$ *cells) and the energy-error estimator (right,* $N \sim 1700$ *cells).*

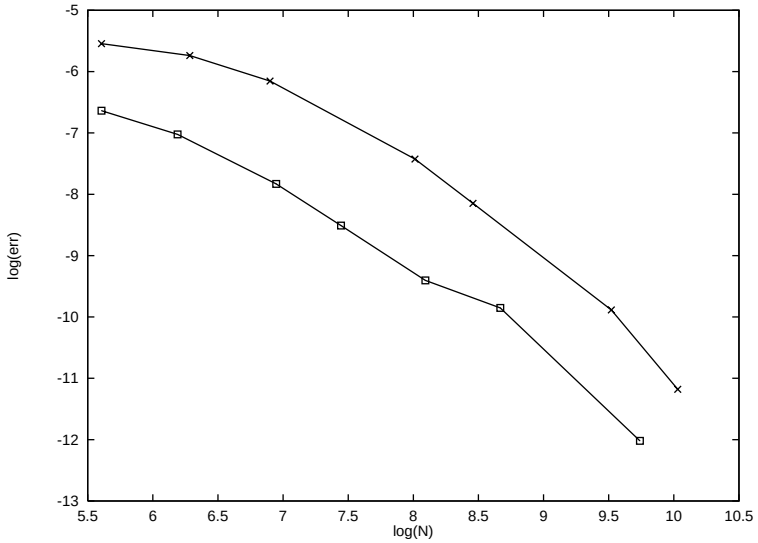

FIG. 5.3. *Linear test* (*configuration* 1): *Comparison of the efficiency of the meshes generated by the weighted error estimator (symbol* $\square$ *) and the energy-error estimator (symbol* $\times$ *) in* log / log *scale.*

We see that the energy-norm error bound η_E for the state equation on the ω-meshes is slightly larger than on the E-meshes. This is not surprising since the ω-meshes are not so much refined in the regions where the state variable has a steep gradient. The cells are rather concentrated along the control and observation boundaries which seems to be more effective for the optimization process. Indeed, the approximate solution $\{u_h, \lambda_h, q_h\}$ obtained by the weighted error estimator η_ω achieves a much smaller value (factor ~ 0.1) of the cost functional. However, for other data, e.g., $c_0 = \cos(2x)$ and $\alpha = 0.0001$, the discrepancy between the two kinds of meshes with respect to the satisfaction of the state equation may be more significant. In those cases it would be advisable to use the combined error estimator $\eta_{\omega,E}$ described

TABLE 5.2
Linear test (configuration 1): Values of the two error estimators η_ω and η_E obtained on "ω-meshes" and on "E-meshes."

$N \approx$	η_ω on ω-meshes	η_E on ω-meshes	η_ω on E-meshes	η_E on E-meshes
140	0.0040205	0.0193270	0.0043245	0.0162589
300	0.0022030	0.0157156	0.0026536	0.0112183
750	0.0008330	0.0092718	0.0020437	0.0074801
3700	0.0001660	0.0049598	0.0004870	0.0034197
11000	0.0000532	0.0026208	0.0002199	0.0019036
21000	0.0000317	0.0020740	0.0001189	0.0014285
28000	0.0000239	0.0016294	0.0001088	0.0012403
48000	0.0000108	0.0013373	0.0000722	0.0009399
145000	0.0000037	0.0006950	0.0000328	0.0005466

above, if stronger "admissibility" of the discrete state u_h is required.

6. The nonlinear case. Now, we consider a nonlinear analogue of the abstract linear model problem (2.1), (2.2),

$$a(u)(\varphi) = (f,\varphi) + b(q,\varphi) \quad \forall \varphi \in V, \tag{6.1}$$

where $a(\cdot)(\cdot)$ is a semilinear form on the Hilbert space V. The cost functional $J(\cdot,\cdot)$ is the same as in the linear case. The corresponding Lagrange functional

$$\mathcal{L}(u,q,\lambda) := J(u,q) + a(u)(\lambda) - b(q,\lambda) - (f,\lambda),$$

leads to the first-order necessary optimality condition

$$\begin{aligned} a'(u)(v,\lambda) + (cu - c_0, cv) &= 0 \quad \forall v \in V,\\ a(u)(\mu) - b(q,\mu) &= (f,\mu) \quad \forall \mu \in V,\\ -b(r,\lambda) + n(q,r) &= 0 \quad \forall r \in Q, \end{aligned} \tag{6.2}$$

where $a'(u)(\cdot,\cdot)$ denotes the tangent form of $a(\cdot)(\cdot)$ at u. As in the linear case, the discrete approximations $\{u_h, \lambda_h, q_h\}$ are determined as solutions of the saddle-point problem

$$\begin{aligned} a'(u_h)(v_h,\lambda_h) + (cu_h - c_0, cv) &= 0 \quad \forall v_h \in V_h,\\ a(u_h)(\mu_h) - b(q_h,\mu_h) &= (f,\mu_h) \quad \forall \mu_h \in V_h,\\ -b(r_h,\lambda_h) + n(q_h,r_h) &= 0 \quad \forall r_h \in Q_h, \end{aligned} \tag{6.3}$$

where the discretization is the same as in the linear case. We will use again the notation $x = \{u,\lambda,q\}$ and $x_h = \{u_h,\lambda_h,q_h\}$ for points in the spaces $X := V \times V \times Q$ and $X_h := V_h \times V_h \times Q_h$, respectively.

The a posteriori error estimation in the case of a nonlinear state equation follows the same pattern as in the linear case. First, we state an abstract result.

PROPOSITION 6.1. *For the Galerkin finite element approximation* (6.3) *of the abstract model problem* (6.2) *with nonlinear state equation and quadratic cost functional there holds*

$$J(u,q) - J(u_h,q_h) = \tfrac{1}{2}\nabla\mathcal{L}(x_h)(x - i_h x) + R(x,x_h)\,, \tag{6.4}$$

where the remainder term $R(x,x_h)$ can be estimated by

$$|R(x,x_h)| \;\le\; \sup_{\hat{x}\in[x_h,x]} |\nabla^3\mathcal{L}(\hat{x})(x-x_h,x-x_h,x-x_h)|. \tag{6.5}$$

Proof. The Galerkin orthogonality relation now reads

$$\nabla^2\mathcal{L}(\overline{xx_h})(x-x_h,\varphi_h) = \nabla\mathcal{L}(x)(\varphi_h) - \nabla\mathcal{L}(x_h)(\varphi_h) = 0, \quad \varphi_h \in X_h, \tag{6.6}$$

with the abbreviating notation

$$\mathcal{L}(\overline{xx_h}) := \int_0^1 \mathcal{L}(x + t(x_h - x))\, dt.$$

Since the solutions u and u_h satisfy the corresponding state equations, there holds again

$$J(u,q) - J(u_h,q_h) = \mathcal{L}(x) - \mathcal{L}(x_h).$$

By Taylor expansion, there holds

$$\begin{aligned} rcl\mathcal{L}(x) - \mathcal{L}(x_h) = \nabla\mathcal{L}(x)(x-x_h) - \tfrac{1}{2}\nabla^2\mathcal{L}(x)(x-x_h, x-x_h) \\ + \tfrac{1}{6}\nabla^3\mathcal{L}(\tilde{x})(x-x_h, x-x_h, x-x_h), \end{aligned}$$

where $\tilde{x}$ lies between x and x_h. Since x is a stationary point of $\mathcal{L}$, the first term on the right vanishes. In order to relate the second term to the Galerkin relation (6.6), we use again Taylor expansion,

$$\nabla^2\mathcal{L}(x)(x-x_h, x-x_h) = \nabla^2\mathcal{L}(\overline{xx_h})(x-x_h, x-x_h) + \nabla^3\mathcal{L}(\hat{x})(x-x_h, x-x_h, x-x_h),$$

where $\hat{x}$ is another point between x and x_h. In view of the identity

$$\nabla^2\mathcal{L}(\overline{xx_h})(x-x_h,\cdot) = \nabla\mathcal{L}(x)(\cdot) - \nabla\mathcal{L}(x_h)(\cdot) = -\nabla\mathcal{L}(x_h)(\cdot),$$

and the Galerkin relation (6.6), we conclude that

$$\begin{aligned} \mathcal{L}(x) - \mathcal{L}(x_h) &= -\tfrac{1}{2}\nabla^2\mathcal{L}(\overline{xx_h})(x-x_h, x-x_h) + R(x,x_h) \\ &= -\tfrac{1}{2}\nabla^2\mathcal{L}(\overline{xx_h})(x-x_h, x-x_h-\varphi_h) + R(x,x_h) \\ &= \tfrac{1}{2}\nabla\mathcal{L}(x_h)(x-x_h-\varphi_h) + R(x,x_h), \end{aligned}$$

with an arbitrary $\varphi_h \in X_h$, and the remainder term

$$R(x,x_h) = \nabla^3\mathcal{L}(\hat{x})(x-x_h, x-x_h, x-x_h) + \tfrac{1}{6}\nabla^3\mathcal{L}(\tilde{x})(x-x_h, x-x_h, x-x_h).$$

Taking here $\varphi_h = i_h x - x_h$ eventually results in

$$\mathcal{L}(x) - \mathcal{L}(x_h) = \tfrac{1}{2}\nabla\mathcal{L}(x_h)(x - i_h x) + R(x,x_h),$$

which completes the proof. ∎

We note that if the cost functional $J(\cdot)$ is quadratic and the control form $b(\cdot,\cdot)$ bilinear, then the only nonzero terms in $\nabla^3\mathcal{L}$ are

$$\frac{\partial^3\mathcal{L}}{\partial\lambda\partial^2 u}(x) = a''(u)(\cdot,\cdot,\cdot), \qquad \frac{\partial^3\mathcal{L}}{\partial^3 u}(x) = a'''(u)(\cdot,\cdot,\cdot,\lambda).$$

Further, if additionally the state equation is linear, then the remainder term $R(x,x_h)$ vanishes.

We will apply this abstract result for a nonlinear problem of optimal control in the "Ginzburg–Landau model" of superconductivity in semiconductors; for references see Du, Gunzburger, and Peterson [7], Itô and Kunish [10], and also Tinkham [13]. It has the same structure as the linear model problem considered above,

$$\begin{aligned} -\Delta u + s(u) &= f \quad \text{in } \Omega, \\ \partial_n u = 0 \ \text{ on } \Gamma_N, &\quad \partial_n u = q \ \text{ on } \Gamma_C, \end{aligned} \tag{6.7}$$

with the nonlinearity $s(u) := u^3 - u$, and the quadratic cost functional

$$J(u,q) = \tfrac{1}{2}\|u - c_0\|^2_{\Gamma_O} + \tfrac{\alpha}{2}\|q\|^2_{\Gamma_C}.$$

The corresponding first-order necessary condition (6.2) uses the notation

$$a(u)(v) = (\nabla u, \nabla v)_\Omega + (s(u), v)_\Omega, \quad b(q,v) = (q,v)_{\Gamma_C}, \quad n(q,r) = \alpha(q,r)_{\Gamma_C},$$

and is approximated by the scheme (6.3). The well-posedness of this optimization problem, the existence of the adjoint variable λ, as well as a priori error estimates for its discretization have been discussed by Gunzburger and Hou [9]. From Proposition 6.1, we conclude the following a posteriori result.

PROPOSITION 6.2. *For error control with respect to the cost functional J, there holds the weighted a posteriori error estimate*

$$|J(u,q) - J(u_h,q_h)| \ \le \ \eta_\omega(u_h,\lambda_h,q_h) + R(\{u,\lambda,q\},\{u_h,\lambda_h,q_h\}), \tag{6.8}$$

where the local error indicators $\eta_T(u_h,\lambda_h,q_h)$ in the linearized error estimator

$$\eta_\omega(u_h,\lambda_h,q_h) := \sum_{T\in\mathbb{T}_h} \eta_T(u_h,\lambda_h,q_h) \tag{6.9}$$

are defined as in the linear case (Proposition 4.2), here with the "cell residuals"

$$\begin{aligned} R^{(u)}_{h|T} &:= -\Delta u_h + s(u_h) - f, \quad R^{(\lambda)}_{h|T} := -\Delta\lambda_h + s'(u_h)\lambda_h, \\ r^{(q)}_{h|\Gamma} &:= \alpha q_h - \lambda_h, \ \ \textit{if } \Gamma \subset \Gamma_C. \end{aligned} \tag{6.10}$$

For the remainder term, there holds the a priori estimate

$$\big|R(\{u,\lambda,q\},\{u_h,\lambda_h,q_h\})\big| \ \le \ 6\int_\Omega \Big\{ \max\{|u|,|u_h|\}|u-u_h|^3 + |u-u_h|^2|\lambda-\lambda_h| \Big\}\, dx. \tag{6.11}$$

As in the linear case, the weights are evaluated numerically using the approximations $\{u_h,\lambda_h,q_h\}$, but now the weighted error estimator contains an additional linearization error represented by the remainder R. Theory as well as practical experience show that, in the present case, this additional error is of higher order on well-adapted meshes and can therefore be neglected. In fact, assuming sufficient smoothness of the solution $\{u,\lambda,q\}$, there holds

$$\big|R(\{u,\lambda,q\},\{u_h,\lambda_h,q_h\})\big| \le c(u,u_h)\, h^6_{\max}, \tag{6.12}$$

with the maximum step size $h_{\max}$ of the mesh. The proof of this order-optimal estimate employing techniques from L^∞-error analysis of finite elements would be

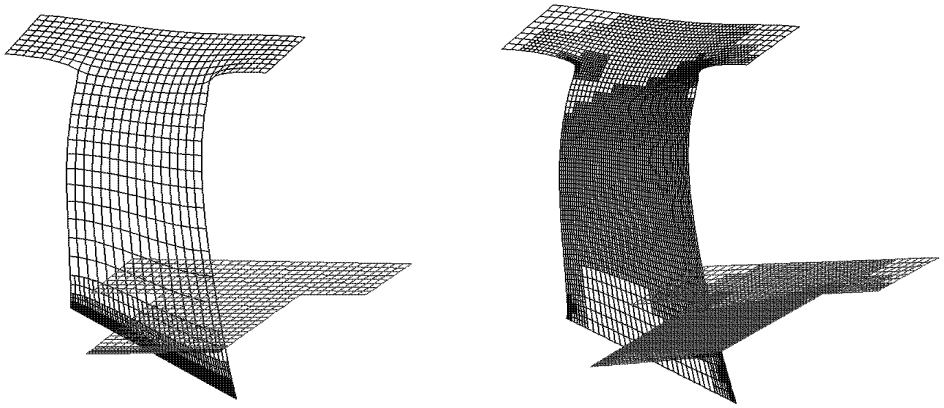

FIG. 7.1. *Nonlinear test (configuration 2, $\alpha=0$): Comparison of discrete solutions obtained by the weighted error estimator (left, $N \sim 5000$ cells) and the energy-error estimator (right, $N \sim 4800$ cells).*

TABLE 7.1
Nonlinear test (configuration 2, $\alpha=0$): Efficiency of the weighted error estimator in the case $\Gamma_C = \Gamma_O$.

N	596	1616	5084	8648	15512
E_h	2.56e-04	2.38e-04	8.22e-05	4.21e-05	3.99e-05
I_{eff}	0.34	0.81	0.46	0.29	0.43

rather lengthy and is therefore omitted. In view of this observation, we neglect the remainder term in the a posteriori error estimate (6.8) and base the mesh adaptation on its main part $\eta_\omega(u_h, \lambda_h, q_h)$.

The discrete problems (6.3) are solved by a quasi-Newton iteration which is derived from a corresponding scheme formulated on the continuous level. On each discrete level the Newton iteration is carried to the limit before the error estimator is applied for mesh refinement. The results of this process may significantly differ from those obtained if each Newton step is discretized separately, mixing iteration and discretization errors together; see the preceding publication [2] for the latter approach.

7. Numerical results—nonlinear case. We again compare the weighted error estimator with a simple ad hoc energy-error estimator of the form (4.11) using the modified cell residuals (6.2). For illustrating our approach, we consider two different choices for the boundaries of control and observation shown in Figure 5.1 as configuration 1 and configuration 2. The notation I_{eff}, E_h, and η_h is as defined above.

(i) *First test:* First, we consider configuration 2 in which the same boundary is taken for control and observation, $\Gamma_C = \Gamma_O$ (lower boundary of the T-shaped domain) and set $\alpha = 0$. In this configuration, we do not expect any need for strong mesh refinement "far away" from this boundary if we only want to deal with the optimization problem.

The observations are taken as $c_0(x) = \sin(0.19x)$. Table 7.1 shows the quality of the weighted error estimator for quantitative error control for this nonlinear test case. The reference value for the objective function $J(u, q)$ is computed on a refined mesh with about $130,000$ cells. Due to the special choice $\Gamma_C = \Gamma_O$, the adjoint variable λ equals zero almost everywhere away from Γ_C, i.e., the weighted error

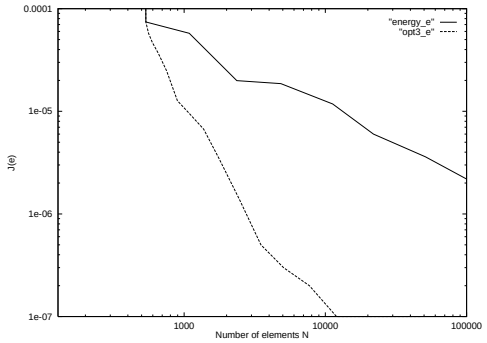

FIG. 7.2. *Nonlinear test (configuration* 2, $\alpha = 0$*): Comparison of efficiency of meshes generated by the two estimators* η_ω *(broken line) and* η_E *(solid line) in* log /*log scale.*

TABLE 7.2

Nonlinear test (configuration 1, $\alpha = 0.1$*): Efficiency of the weighted error estimator for computing a secondary stationary point.*

N	512	15368	27800	57632	197408
E_h	9.29e-05	8.14e-07	4.86e-07	2.31e-07	4.58e-08
I_{eff}	1.32	0.56	0.35	0.42	0.32

TABLE 7.3

Nonlinear test (configuration 1, $\alpha = 1$*): Efficiency of the weighted error estimator for computing a secondary stationary point.*

N	512	8120	25544	42608	126284
E_h	2.08e-03	4.35e-05	9.26e-06	5.95e-06	8.94e-07
I_{eff}	0.52	0.73	0.88	1.21	0.98

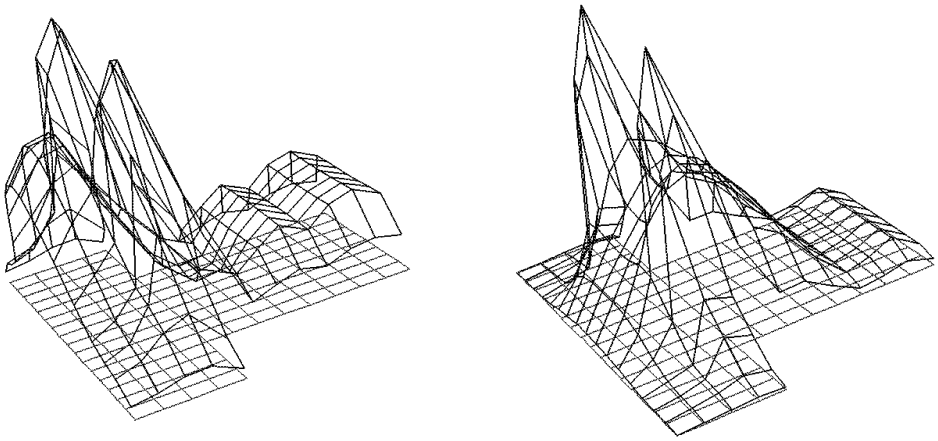

FIG. 7.3. *Nonlinear test (configuration* 1, $\alpha = 0.1$*): Distributions of local error indicators in the weighted error estimator* η_ω *(left) and the energy-error estimator* η_E *(right).*

estimator pays attention only to the neighborhood of the control boundary. The energy-error estimator mainly recognizes the singularities in the primal solution at the two reentrant corners (see Figure 7.1).

In Figure 7.2, we compare the efficiency of the meshes generated by the two

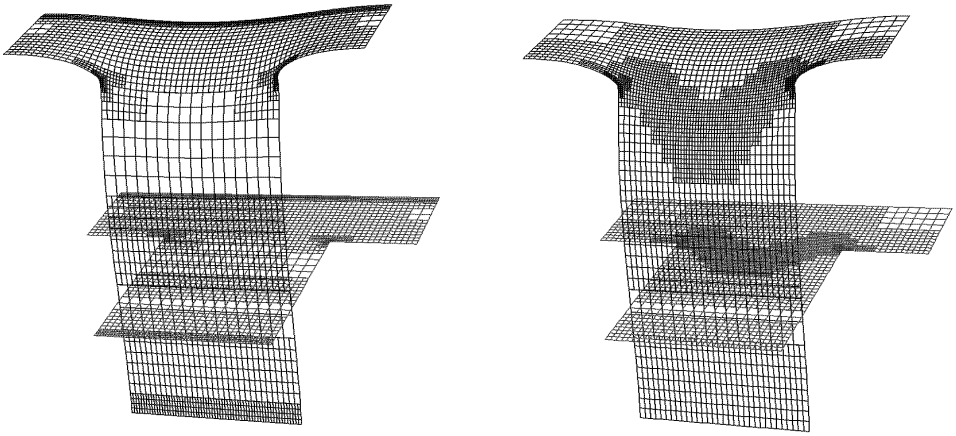

FIG. 7.4. *Nonlinear test (configuration* 1, $\alpha = 0.1$*): Comparison of discrete solutions obtained by the weighted error estimator* η_ω *(left,* $N \sim 3000$ *cells) and the energy-error estimator* η_E *(right,* $N \sim 3300$ *cells).*

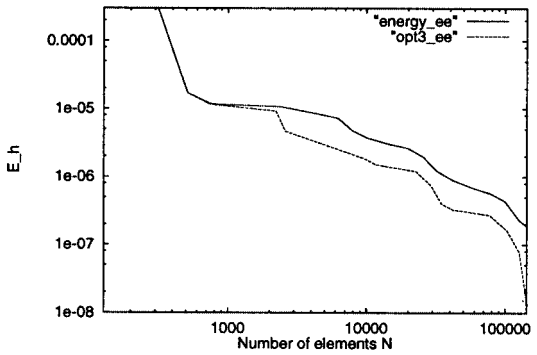

FIG. 7.5. *Nonlinear test (configuration* 1, $\alpha = 0.1$*): Comparison of efficiency of meshes generated by the two error estimators:* η_ω *(broken line) and* η_E *(solid line) in* log */log scale.*

estimators. We see that in this situation, the solution of the optimization problem is approximated with significantly less cells using the weighted error estimator which exploits the "extreme" feature $\Gamma_C = \Gamma_O$ of this problem.

(ii) *Second test:* Now, we consider configuration 1 in which the control and the observation are taken on opposite boundaries, $\Gamma_C \cap \Gamma_O = \emptyset$. In this case, we expect better results for the energy-error estimator than in configuration 1, because the information must pass from the control to the observation boundary and the corner singularities will have a stronger effect. Nevertheless, the weighted error estimator should perform better since it also considers the critical control and observation boundaries.

We take the observation as $c_0 \equiv 1$, as in the linear case, and set $\alpha = 0.1$. In this configuration, there exist several stationary points of $\mathcal{L}(u, q, \lambda)$, which can be obtained by varying the starting values for the Newton iteration. One trivial solution (actually the global minimum) is a constant equal to c_0, with $q \equiv 0$. In this case, we have $J(u, q) = 0$ (up to round-off error) and match the observations already

on a very coarse mesh. We do not show the results of this computation. The two other stationary points are symmetric to each other with respect to the center plane $\{x = 0\}$. Tables 7.2 and 7.3 show the quality of the weighted error estimator for quantitative error control for one of these local minima. The reference values for $J(u, q)$ are obtained on an adaptive mesh with about $550,000$ cells. The numerical results demonstrate the correct qualitative behavior of the weighted error estimator. For the choice $\alpha = 1$, we get slightly better results than for $\alpha = 0.1$ because of higher stability in the optimization problem.

Next, Figure 7.3 shows the distribution of the cell error indicators η_T in the weighted error estimator η_ω and in the energy-error estimator η_E for $\alpha = 0.1$. We clearly see the different ways in which these error estimators put their weight: η_ω observes the control and observation boundary which is critical for the optimization process while η_E emphasizes the corner singularities. Figure 7.4 shows the resulting meshes together with the computed discrete solutions. Finally, in Figure 7.5, we see the faster convergence toward the minimum of the objective functional using the weighted error estimator compared to the energy-error estimator.

REFERENCES

[1] M. Ainsworth and J.T. Oden, *A posteriori error estimation in finite element analysis*, Comput. Methods Appl. Mech. Engrg., 142 (1997), pp. 1–88.

[2] R. Becker and H. Kapp, *Optimization in PDE models with adaptive finite element discretization*, in Proceedings ENUMATH 97 (Heidelberg), H.G. Bock et al., eds., World Scientific, River Edge, NJ, 1998, pp. 147–155.

[3] R. Becker and R. Rannacher, *Weighted a posteriori error control in FE methods*, ENUMATH-95, Paris, 1995, in Proceedings ENUMATH-97, H.G. Bock et al., eds., World Scientific, River Edge, NJ, 1998, pp. 621–637.

[4] R. Becker and R. Rannacher, *A feed-back approach to error control in finite element methods: Basic analysis and examples*, East-West J. Numer. Math, 4 (1996), pp. 237–264.

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

[6] G.F. Carey and J.T. Oden, *Finite Elements, Computational Aspects*, Vol. III, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[7] Q. Du, M.D. Gunzburger, and J.S. Peterson, *Analysis and approximation of the Ginzburg–Landau model of superconductivity*, SIAM Rev., 34 (1992), pp. 54–81.

[8] K. Eriksson, D. Estep, P. Hanspo, and C. Johnson, *Introduction to adaptive methods for differential equations*, Acta Numerica 1995, A. Iserles, ed., Cambridge University Press, Cambridge, UK, 1995, pp. 105–158.

[9] 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.

[10] K. Itô and K. Kunisch, *Augmented Lagrangian-SQP methods for nonlinear optimal control problems of tracking type*, SIAM J. Control Optim., 34 (1996), pp. 874–891.

[11] J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.

[12] R. Rannacher, *Error control in finite element computations*, in Error Control and Adaptivity in Scientific Computing, H. Bulgak and C. Zenger, eds., NATO Sci. Ser. C Math. Phys. Sci., 536, Kluwer, Dordrecht, the Netherlands, 1999, pp. 247–278.

[13] M. Tinkham, *Introduction to Superconductivity*, McGraw-Hill, New York, 1975.

[14] R. Verfürth, *A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques*, Adv. Numer. Math., Wiley-Teubner, New York, Stuttgart, 1996.