R Becker, H Kapp and R Rannacher

Adaptive Finite Element Methods for Optimization Problems

Abstract We present a new approach to error control and mesh adaptivity in the numerical solution of optimal control problems governed by elliptic differential equations. The indefinite boundary value problems obtained by the Lagrangian formalism are discretized by the Galerkin finite element method. The mesh adaptation is driven by residual-based a posteriori error estimates which are derived by global duality arguments. This general approach facilitates control of the error with respect to any quantity of physical interest. In discretizing an optimization problem it seems natural to control the error with respect to the given cost functional. In this way, the computed solution can directly be used in weighting the cell residuals in the a posteriori error estimate. This approach has the features of model reduction as used in optimal control of complex dynamical systems. For illustration, we present some results of test computations for simple model problems in optimal control of super–conductivity.

1 Introduction

In this note we describe a new adaptive finite element method for optimal control problems governed by elliptic partial differential equations. The control problem is treated by the classical Lagrange formalism yielding the Euler-Lagrange equations as first-order optimality conditions. This indefinite system of partial differential equations for the state variable u, the Lagrange multiplier λ and the control variable q is discretized by a standard finite element method. In this approach the set of admissible solutions is also discretized and we generally obtain inadmissible states. 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 an economical computation. The need for adaptive error control is therefore evident.

For a posteriori error estimation, we apply the general optimal-control approach developed in [3] and [4] for Galerkin finite element discretizations of partial differential equations. In this method a posteriori estimates for the error with respect to arbitrary functional output are obtained via duality arguments. In these estimates local cell residuals of the computed solution are weighted by certain derivatives of the dual solution which describe the dependence of the error on variations of the local mesh size. In general theses weights have to be approximated by numerically solving the corresponding dual problem. This results in a feed-back process by which successively more and more accurate error bounds and economical meshes are generated.

In applying this approach to optimization problems it seems natural to base the error control on the given cost functional. In this particular case the corresponding dual solution is given in terms of the state variable, Lagrangian multiplier and control variable. Hence, the evaluation of the a posteriori error estimates does not require extra work for solving a dual problem and a posteriori error estimation is almost for free. This observation leads to a particularly simple and efficient strategy for mesh adaptation in the finite element discretization of optimal control problems. We provide the theoretical background for this approach and discuss some computational tests.

2 An optimization model problem

We consider optimal control problems of the form

$$J(u,q) \;\to\; \min!, \qquad A(u,q) = f, \tag{2.1}$$

where A is a partial differential operator relating a state variable u and a control variable q, and $J(\cdot,\cdot)$ is a cost functional. As prototypical model case, we consider a problem of boundary control in super–conductivity ("Ginzburg-Landau model"; see Du, Gunzburger and Peterson [6], and Ito and Kunisch [9]). The state equations are

$$\begin{aligned} -\Delta u + s(u) &= f \quad \text{on } \Omega, \\ \partial_n u = q \;\text{ on } \Gamma_C, &\quad \partial_n u = 0 \;\text{ on } \partial\Omega \setminus \Gamma_C, \end{aligned} \tag{2.2}$$

defined on an open bounded domain with boundary $\partial\Omega$. The control q acts on the boundary component Γ_C while the observations $u_{|\Gamma_O}$ are taken on a component Γ_O; see Figure 2.1. The cost functional is defined by

$$J(u,q) = \tfrac{1}{2}\|u - u_O\|^2_{\Gamma_O} + \tfrac{1}{2}\alpha\|q\|^2_{\Gamma_C}. \tag{2.3}$$

with a prescribed boundary function u_O and a regularization parameter $\alpha \geq 0$. The zero-order term may be linear $s(u) = u$ or nonlinear of the form $s(u) = u^3 - u$.

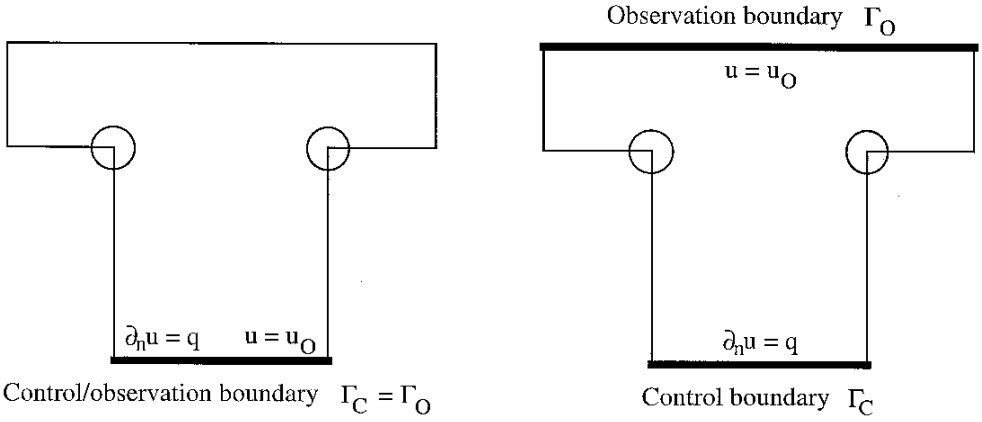

Figure 2.1: Configuration of the boundary control model problem on a T-domain (Ginzburg-Landau model): Configuration 1 (left), Configuration 2 (right).

We employ the Lagrangian formalism with the Langrange function

$$L(u,\lambda,q) = J(u,q) - (-\Delta u + s(u) - f, \lambda)_\Omega + (\partial_n u - q, \lambda)_{\Gamma_C}, \tag{2.4}$$

involving the adjoint variable λ. The corresponding Euler-Lagrange equations read expressed in weak form as follows:

$$(u,\psi)_{\Gamma_O} + (\nabla\psi, \nabla\lambda)_\Omega + (s'(u)\psi, \lambda)_\Omega = (u_O, \psi)_{\Gamma_O} \quad \forall \psi \in H, \tag{2.5}$$

$$(\nabla u, \nabla\pi)_\Omega + (s(u), \pi)_\Omega - (q, \pi)_{\Gamma_C} = (f, \pi)_\Omega \quad \forall \pi \in H, \tag{2.6}$$

$$-(\chi, \lambda)_{\Gamma_C} + \alpha(q, \chi)_{\Gamma_C} = 0 \quad \forall \chi \in Q. \tag{2.7}$$

Here, the natural function space for the state variable u and the adjoint variable λ is the Sobolev space $H := H^1(\Omega)$, while the control q is determined in the Lebesgue space $Q := L^2(\Gamma_C)$. The stationary points of this saddle-point problem yield possible extreme points of the optimization problem. The system (2.5) - (2.7) can be written in matrix form as follows:

$$\begin{bmatrix} C & A^T + N'(u) & 0 \\ A + \widetilde{N}'(u) & 0 & -B \\ 0 & -B^T & \alpha S \end{bmatrix} \begin{bmatrix} u \\ \lambda \\ q \end{bmatrix} = \begin{bmatrix} u_O \\ f \\ 0 \end{bmatrix}, \tag{2.8}$$

where $\widetilde{N}'(u) := N(0) + \int_0^1 N'(tu)\,dt$, and the operators C, A, B, S, $N(\cdot)$, and $N'(\cdot)$ have their obvious meaning. For solving this saddle–point problem, we use a Newton iteration which is defined on the continuous level in the function space $H \times H \times Q$. The Newton steps are discretized by an adaptive Galerkin finite element method using the conforming bilinear Q_1 element for all unknowns u, λ and q. The mesh adaptation is driven by residual-based a posteriori error estimates for the nonlinear problem. The linearized algebraic problems are solved by a GMRES method with multigrid preconditioning; for algorithmic details, we refer to [1], [2], and [10].

3 Galerkin finite element discretization

The Galerkin finite element discretization of the saddle–point problem (2.5) - (2.7) uses subspaces $H_h \subset H$ and $Q_h \subset Q$ of piecewise polynomial functions defined on regular decompositions $\mathbb{T}_h = \cup_{T\in\mathbb{T}_h}\{T\}$ of the domain Ω into cells T (triangles or quadrilaterals); see Brenner and Scott [5]. Here, we think of bilinear shape functions. In order to avoid unnecessary complications due to 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 $T \in \mathbb{T}_h$ or a cell edge $\Gamma \subset \partial\Omega$. In order to ease local mesh refinement and coarsening, hanging nodes are allowed but at most one per edge; see Figure 3.1. The corresponding degrees of freedom are eliminated by interpolation in order to keep the discretization conforming. For simplicity, we assume that the space Q_h of discrete controls is given by the traces of the finite element functions of V_h. This is not necessary for our results but simplifies notation.

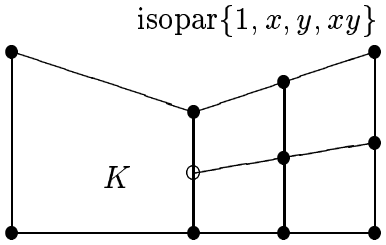

Figure 3.1: Quadrilateral mesh with 'hanging nodes"

The discrete problems seek $\{u_h, \lambda_h.q_h\} \in H_h \times H_h \times Q_h$, such that

$$(u_h, \psi_h)_{\Gamma_O} + (\nabla\psi_h, \nabla\lambda_h)_\Omega + (s'(u_h)\psi_h, \lambda_h)_\Omega = (u_O, \psi_h)_{\Gamma_O} \quad \forall\psi_h \in H_h, \tag{3.1}$$

$$(\nabla u_h, \nabla\pi_h)_\Omega + (s(u_h), \pi_h)_\Omega - (q_h, \pi_h)_{\Gamma_C} = (f, \pi_h)_\Omega \quad \forall\pi_h \in H_h, \tag{3.2}$$

$$-(\chi_h, \lambda_h)_{\Gamma_C} + \alpha(q_h, \chi_h)_{\Gamma_C} = 0 \quad \forall\chi_h \in Q_h. \tag{3.3}$$

For this approximation, we have the following theoretical a priori estimate for the errors $e_u := u - u_h$, $e_\lambda := \lambda - \lambda_h$, and $e_q := q - q_h$ (see Gunzburger and Hou [8], and also [2]):

$$\|\nabla e_u\|_\Omega + \|e_q\|_{\Gamma_C} + \|\nabla e_\lambda\|_\Omega \;\le\; c\,\big\{\|\nabla(u - \psi_h)\|_\Omega + \|\nabla(\lambda - \pi_h)\|_\Omega + \|q - \chi_h\|_{\Gamma_C}\big\},$$

for arbitrary approximations $\psi_h,\ \pi_h \in H_h$ and $\chi_h \in Q_h$.

4 A posteriori error estimation

4.1 A model problem

We give a brief introduction to a posteriori error estimation using duality techniques. For illustration, we consider the simple model problem

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega. \tag{4.1}$$

The variational formulation of this problem seeks $u \in H = H_0^1(\Omega)$ such that

$$(\nabla u, \nabla\phi)_\Omega = (f, \phi)_\Omega \quad \forall\phi \in H. \tag{4.2}$$

The discrete approximations $u_h \in H_h$ are determined by the Galerkin equations

$$(\nabla u_h, \nabla\phi_h)_\Omega = (f, \phi_h)_\Omega \quad \forall\phi_h \in H_h. \tag{4.3}$$

The essential feature of this approximation scheme is the "Galerkin orthogonality" of the error $e := u - u_h$,

$$(\nabla e, \nabla\phi_h)_\Omega = 0, \quad \phi_h \in H_h. \tag{4.4}$$

Next, we seek to derive a posteriori error estimates. Let $J(\cdot)$ be a *linear* "error functional" defined on H and $z \in H$ the solution of the corresponding dual problem

$$(\nabla\phi, \nabla z)_\Omega = J(\phi) \quad \forall \phi \in H. \tag{4.5}$$

Taking $\phi = e$ in (4.5) and using the Galerkin orthogonality, cell-wise integration by parts (observing that $z_{|\partial\Omega} = 0$) results in the error representation

$$\begin{aligned} J(e) &= (\nabla e, \nabla z)_\Omega = (\nabla e, \nabla(z - \phi_h))_\Omega \\ &= \sum_{T\in\mathbb{T}_h} \left\{(-\Delta u + \Delta u_h, z - \phi_h)_T - (\partial_n u_h, z - \phi_h)_{\partial T}\right\} \\ &= \sum_{T\in\mathbb{T}_h} \left\{(f + \Delta u_h, z - \phi_h)_T - \tfrac{1}{2}(n\cdot[\nabla u_h], z - \phi_h)_{\partial T\setminus\partial\Omega}\right\}, \end{aligned} \tag{4.6}$$

where $[\nabla u_h]$ denotes the jump of ∇u_h across the interelement boundary, and $\phi_h \in H_h$ is an arbitrary approximation. In this relation the factors $z - \phi_h$ may be viewed as weights expressing the sensitivity of the error quantity $J(e)$ with respect to changes of the equation residuals $(f + \Delta u_h)_{|T}$ and jump residuals $n\cdot[\nabla u_h]_{|\partial T}$. From the error identity (4.6), we infer the following a posteriori error estimate

$$|J(e)| \le \eta(u_h) := \sum_{T\in\mathbb{T}_h} h_T^2 \left\{\rho_T^{(u)}\omega_T^{(z)} + \rho_{\partial T}^{(u)}\omega_{\partial T}^{(z)}\right\}, \tag{4.7}$$

with the cell residuals and weights

$$\begin{aligned} \rho_T^{(u)} &:= h_T^{-1}\|f + \Delta u_h\|_T, & \rho_{\partial T}^{(u)} &:= \tfrac{1}{2}h_T^{-3/2}\|n\cdot[\nabla u_h]\|_{\partial T\setminus\partial\Omega}, \\ \omega_T^{(z)} &:= h_T^{-1}\|z - \phi_h\|_T, & \omega_{\partial T}^{(z)} &:= h_T^{-1/2}\|z - \phi_h\|_{\partial T\setminus\partial\Omega}. \end{aligned}$$

In view of the local approximation properties of finite elements, there holds

$$\omega_T^{(z)} + \omega_{\partial T}^{(z)} \le c_I h_T^2 \max_T |\nabla^2 z|\,. \tag{4.8}$$

We remark that in a finite difference discretization of the model problem (4.1) the corresponding "influence factors" behave like $\omega_T^{(z)} \approx \max_T |z|$.

In practice the weights $\omega_T^{(z)}$, $\omega_{\partial T}^{(z)}$ have to be determined computationally. Let $z_h \in H_h$ be the Galerkin approximation of z defined by

$$(\nabla\phi_h, \nabla z_h)_\Omega = J(\phi_h) \quad \forall \phi_h \in V_h. \tag{4.9}$$

In view of the estimate (4.8), we can approximate

$$\omega_T^{(z)} + \omega_{\partial T}^{(z)} \approx c_I h_T^2 \max_T |\nabla_h^2 z_h|, \tag{4.10}$$

where $\nabla_h z_h$ is a suitable difference quotient approximating $\nabla^2 z$. The interpolation constant is usually in the range $c_I \approx 0.1 - 1$ and can be determined by calibration.

Alternatively, we may construct from $z_h \in H_h$ a patchwise bi-quadratic interpolation $I_h^2 z_h$ and replace $z - \phi_h$ in the weights by $I_h^2 z_h - z_h$. This gives an approximation which is free of any interpolation constant. The quality of these approximations for the model problem has been analysed in [4].

Next, we apply the foregoing argument for deriving an energy-norm error estimate. This is intended to prepare for our analysis of the error in approximating optimal control problems. The solution $u \in H$ of (4.1) is characterized by the property

$$L(u) = \tfrac{1}{2}\|\nabla u\|_\Omega^2 - (f,u)_\Omega \;\rightarrow\; min! \quad \text{on } H. \tag{4.11}$$

Further, we note that

$$\begin{aligned} L(u) - L(u_h) &= \tfrac{1}{2}\|\nabla u\|_\Omega^2 - (f,u)_\Omega - \tfrac{1}{2}\|\nabla u_h\|_\Omega^2 + (f,u_h)_\Omega \\ &= -\tfrac{1}{2}\|\nabla u\|_\Omega^2 - \tfrac{1}{2}\|\nabla u_h\|_\Omega^2 + (\nabla u, \nabla u_h)_\Omega = -\tfrac{1}{2}\|\nabla e\|_\Omega^2. \end{aligned}$$

Hence, energy-error control means control of the error with respect to the "energy functional" $L(\cdot)$. This can be accomplished by using the linear error functional

$$J(\phi) = -\tfrac{1}{2}(\nabla e, \nabla \phi)_\Omega$$

in the corresponding dual problem (4.5). Obviously, the dual solution is then simply given by $z = -\frac{1}{2}e$. Accordingly, the general a posteriori error estimate (4.7) takes the particular form

$$\|\nabla e\|_\Omega^2 \le \sum_{T\in\mathbb{T}_h} h_T^2 \big\{ \rho_T^{(u)} \omega_T^{(u)} + \rho_{\partial T}^{(u)} \omega_{\partial T}^{(u)} \big\}, \tag{4.12}$$

with the weights $\omega_T^{(u)} = h_T^{-1}\|u - \phi_h\|_T$ and $\omega_{\partial T}^{(u)} = h_T^{-1/2}\|u - \phi_h\|_{\partial T \setminus \partial\Omega}$. Then, using the local approximation estimate

$$\inf_{\phi_h \in H_h} \Big(\sum_{T\in\mathbb{T}_h} \big\{ h_T^{-2}\|u - \phi_h\|_T^2 + h_T^{-1}\|u - \phi_h\|_{\partial T}^2 \big\} \Big)^{1/2} \le c_I \|\nabla e\|_\Omega, \tag{4.13}$$

we conclude from (4.12) that

$$\|\nabla e\|_\Omega^2 \le c_I \Big(\sum_{T\in\mathbb{T}_h} h_T^4 \big\{ \rho_T^{(u)2} + \rho_{\partial T}^{(u)2} \big\} \Big)^{1/2} \|\nabla e\|_\Omega.$$

This implies the standard residual–based energy-norm a posteriori error estimate (see, e.g. Verfürth [12]):

$$|L(u) - L(u_h)| = \tfrac{1}{2}\|\nabla e\|_\Omega^2 \le \tfrac{1}{2} c_I^2 \sum_{T\in\mathbb{T}_h} h_T^4 \big\{ \rho_T^{(u)2} + \rho_{\partial T}^{(u)2} \big\}. \tag{4.14}$$

Below, we will see that the peculiar relation $z = -\frac{1}{2}e$ for the dual solution corresponding to the "energy functional" $L(\cdot)$ follows from a general principle which can be used also for the discretization of the optimal control problem described above.

4.2 A general paradigm for a posteriori error estimation

Next, we consider an abstract situation. Let $L(u)$ be a twice differentiable functional on some Hilbert space V, e.g., the energy functional related to the Poisson problem or the Lagrangian functional defined for the Ginzburg-Landau model. For its first and second differentials at u, we use the notation $L'(u;\cdot)$ and $L''(u;\cdot,\cdot)$, respectively. Notice that $L''(u;\cdot,\cdot)$ is symmetric. We seek stationary points $u \in V$ of $L(\cdot)$,

$$L'(u;\phi) = 0 \quad \forall \phi \in V. \tag{4.15}$$

Corresponding approximations $u_h \in V_h$ are defined in finite dimensional subspaces $V_h \subset V$ by the Galerkin equations

$$L'(u_h;\phi_h) = 0 \quad \forall \phi_h \in V_h. \tag{4.16}$$

Let $J(\cdot)$ be a functional chosen for measuring the error $e = u - u_h$. Then,

$$J(u) - J(u_h) = \int_0^1 J'(u_h + te; e)\, dt, \tag{4.17}$$

$$L'(u;\cdot) - L'(u_h;\cdot) = \int_0^1 L''(u_h + te; e, \cdot)\, dt, \tag{4.18}$$

leads us to consider the "dual problem"

$$\int_0^1 L''(u_h + te; \phi, z)\, dt = \int_0^1 J'(u_h + te; \phi)\, dt \quad \forall \phi \in V, \tag{4.19}$$

which is assumed to have a solution $z \in V$. Then, taking $\phi = e$ in (4.19) and using the Galerkin equation (4.16) results in the error identity

$$J(u) - J(u_h) = L'(u; z) - L'(u_h; z) = -L'(u_h; z - \phi_h), \tag{4.20}$$

with arbitrary $\phi_h \in V_h$. In general, this error representation cannot be evaluated since the left-hand side as well as the right-hand side in the dual problem (4.19) depend on the unknown solution u. The simplest way of approximation is to replace u by u_h, which yields the perturbed dual problem

$$L''(u_h; \phi, \tilde{z}) = J'(u_h; \phi) \quad \forall \phi \in V. \tag{4.21}$$

Controlling the effect of this perturbation on the accuracy of the resulting error estimate may be a delicate task and depends strongly on the particular problem under consideration. Our own experiences with different types of applications (e.g., the Navier-Stokes equations and models in elasto-plasticity) indicate that this problem is not critical as long as the solution to be computed is stable. The crucial problem is the approximation of the perturbed dual solution by solving a discretized dual problem

$$L''(u_h; \phi_h, \tilde{z}_h) = J'(u_h; \phi_h) \quad \forall \phi_h \in V_h. \tag{4.22}$$

So far, the derivation of the error representation (4.20) did not use that the variational equation (4.15) stems from an "energy functional". In fact it can be used for much more general situations; see the surveys given in Eriksson, et al. [7], and in [11]. It seems natural to control the error $e = u - u_h$ with respect to the given "energy" functional $L(\cdot)$. Observing that $L'(u;\phi) = 0$, it follows by integration by parts that

$$\int_0^1 L'(u_h + te;\phi)\,dt = -\int_0^1 L''(u_h + te; e, \phi)t\,dt = -\int_0^1 L''(u_h + te;\phi, e)t\,dt.$$

Hence, in this case the dual problem (4.19) takes the special form

$$\int_0^1 L''(u_h + te;\phi, z)\,dt = -\int_0^1 L''(u_h + te;\phi, e)t\,dt \quad \forall\phi \in V. \tag{4.23}$$

If the functional $L(\cdot)$ is quadratic or in the general case by linearization $u \to u_h$, we obtain the perturbed dual problem

$$L''(u_h;\phi,\tilde{z}) = -\tfrac{1}{2}L''(u_h;\phi,e) \quad \forall\phi \in V, \tag{4.24}$$

with the solution $\tilde{z} = -\frac{1}{2}e$. The resulting a posteriori error estimate has the form

$$|L(u) - L(u_h)| \approx \inf_{\phi_h \in V_h} |L'(u_h;\tilde{z} - \phi_h)| = \inf_{\phi_h \in V_h} \tfrac{1}{2}|L'(u_h;\tilde{u} - \phi_h)|. \tag{4.25}$$

In the ideal case of a quadratic functional $L(\cdot)$ linearization is not required and this error bound becomes exact. Here, again the quantity

$$\tilde{z} - \phi_h = -\tfrac{1}{2}e - \phi_h = \tfrac{1}{2}(u - \psi_h)$$

has to be approximated as described above by using the computed solution $u_h \in H_h$.

We emphasize that in this particular case the evaluation of the a posteriori error estimate with respect to the "energy functional" does not require the explicit solution of the dual problem. This abstract reasoning can be taken as guide-line for systematically deriving a posteriori error estimates in concrete situations. In the following, we will carry this out for the optimal control problem described in Section 2.

4.3 A posteriori error analysis for the optimization problem

Now, we want to apply the formalism of the previous section to the particular situation of the boundary control problem described in Section 2. The functional of interest is the Lagrangian functional of the optimal control problem,

$$L(v) = J(u,q) + (\nabla u, \nabla\lambda)_\Omega + (s(u) - f, \lambda)_\Omega - (q,\lambda)_{\Gamma_C},$$

defined for triples $v = \{u, \lambda, q\}$ in the Hilbert space $V := H \times H \times Q$. We recall the equations for stationary points $v = \{u, \lambda, q\} \in V$ of $L(\cdot)$ in slightly rearranged form:

$$(\psi, u - u_O)_{\Gamma_O} + (\nabla\psi, \nabla\lambda)_\Omega + (\psi, s'(u)\lambda)_\Omega = 0 \quad \forall\psi \in H, \tag{4.26}$$

$$(\nabla u, \nabla\pi)_\Omega + (s(u) - f, \pi)_\Omega - (q,\pi)_{\Gamma_C} = 0 \quad \forall\pi \in H, \tag{4.27}$$

$$(\lambda - \alpha q, \chi)_{\Gamma_C} = 0 \quad \forall\chi \in Q. \tag{4.28}$$

The corresponding discrete approximations $v_h = \{u_h, \lambda_h\, q_h\}$ are determined in the finite element space $V_h = H_h \times H_h \times Q_h \subset V$ by

$$(\psi_h, u_h - u_O)_{\Gamma_O} + (\nabla\psi_h, \nabla\lambda_h)_\Omega + (\psi_h, s'(u_h)\lambda_h)_\Omega = 0 \quad \forall\psi_h \in H_h, \tag{4.29}$$

$$(\nabla u_h, \nabla\pi_h)_\Omega + (s(u_h) - f, \pi_h)_\Omega - (q_h, \pi_h)_{\Gamma_C} = 0 \quad \forall\pi_h \in H_h, \tag{4.30}$$

$$(\lambda_h - \alpha q_h, \chi_h)_{\Gamma_C} = 0 \quad \forall\chi_h \in Q_h. \tag{4.31}$$

Following the formalism of the previous section, we seek to estimate the error $e = \{e_u, e_\lambda, e_q\}$ with respect to the Lagrangian functional $L(\cdot)$. The corresponding linearized dual problem

$$L''(u_h; \phi, \tilde{z}) = -\tfrac{1}{2}L''(u_h; \phi, e) \quad \forall\phi \in V, \tag{4.32}$$

then has the solution $\tilde{z} = -\frac{1}{2}\{e_u, e_\lambda, e_q\}$. Hence, we do not have to built this dual problem nor to spend extra work for solving it. This leads us to the following result.

Theorem. *For the finite element discretization of the variational equation (2.5) - (2.7), there holds the a posteriori error relation*

$$\begin{aligned} |J(u,q) - J(u_h,q_h)| \approx &\sum_{\Gamma\subset\partial\Omega} h_\Gamma^2\big\{\rho_\Gamma^{(\lambda)}\omega_\Gamma^{(u)} + \rho_\Gamma^{(u)}\,\omega_\Gamma^{(\lambda)}\big\} + \sum_{\Gamma\subset\Gamma_C} h_\Gamma^2\rho_\Gamma^{(q)}\,\omega_\Gamma^{(q)} \\ &+ \sum_{T\in\mathbb{T}_h} h_T^2\big\{\rho_T^{(u)}\,\omega_T^{(\lambda)} + \rho_{\partial T}^{(u)}\,\omega_{\partial T}^{(\lambda)} + \rho_T^{(\lambda)}\,\omega_T^{(u)} + \rho_{\partial T}^{(\lambda)}\,\omega_{\partial T}^{(u)}\big\}, \end{aligned} \tag{4.33}$$

with the cell residuals and weights

$$\begin{aligned} &\rho_\Gamma^{(\lambda)} = h_\Gamma^{-3/2}\|u_h - u_O + \partial_n\lambda_h\|_\Gamma, \quad \Gamma\subset\Gamma_O, && \omega_\Gamma^{(u)} = h_\Gamma^{-1/2}\|u - \psi_h\|_\Gamma, \\ &\rho_\Gamma^{(\lambda)} = h_\Gamma^{-3/2}\|\partial_n\lambda_h\|_\Gamma, \quad \Gamma\subset\partial\Omega\backslash\Gamma_O, && \\ &\rho_\Gamma^{(u)} = h_\Gamma^{-3/2}\|\partial_n u_h - q_h\|_\Gamma, \quad \Gamma\in\Gamma_C, && \omega_\Gamma^{(\lambda)} = h_\Gamma^{-1/2}\|\lambda - \pi_h\|_\Gamma, \\ &\rho_\Gamma^{(u)} = h_\Gamma^{-3/2}\|\partial_n u_h\|_\Gamma, \quad \Gamma\subset\partial\Omega\backslash\Gamma_C, && \\ &\rho_\Gamma^{(q)} = h_\Gamma^{-3/2}\|\lambda_h - \alpha q_h\|_\Gamma, && \omega_\Gamma^{(q)} = h_\Gamma^{-1/2}\|q - \chi_h\|_\Gamma, \end{aligned}$$

$$\begin{aligned} &\rho_T^{(u)} = h_T^{-1}\|\Delta u_h - s(u_h) + f\|_T, && \omega_T^{(\lambda)} = h_T^{-1}\|\lambda - \pi_h\|_T, \\ &\rho_{\partial T}^{(u)} = \tfrac{1}{2}h_T^{-3/2}\|n\cdot[\nabla u_h]\|_{\partial T\backslash\partial\Omega}, && \omega_{\partial T}^{(\lambda)} = h_T^{-1/2}\|\lambda - \pi_h\|_{\partial T\backslash\partial\Omega}, \\ &\rho_T^{(\lambda)} = h_T^{-1}\|\Delta\lambda_h - s'(u_h)\lambda_h\|_T, && \omega_T^{(u)} = h_T^{-1}\|u - \psi_h\|_T, \\ &\rho_{\partial T}^{(\lambda)} = \tfrac{1}{2}h_T^{-3/2}\|n\cdot[\nabla\lambda_h]\|_{\partial T\backslash\partial\Omega}, && \omega_{\partial T}^{(u)} = h_T^{-1/2}\|u - \psi_h\|_{\partial T\backslash\partial\Omega}, \end{aligned}$$

where $\psi_h, \pi_h \in H_h$ and $\chi_h \in Q_h$ are arbitrary approximations. If the Lagrangian functional $L(\cdot)$ is quadratic this error relation yields a true upper bound.

Proof. In the present case, there holds

$$\begin{aligned} L(v) - L(v_h) &= J(u,q) + (\nabla u, \nabla\lambda)_\Omega + (s(u) - f, \lambda)_\Omega - (q,\lambda)_{\Gamma_C} \\ &\quad - J(u_h,q_h) - (\nabla u_h, \nabla\lambda_h)_\Omega - (s(u_h) - f, \lambda_h)_\Omega + (q_h,\lambda_h)_{\Gamma_C} \\ &= J(u,q) - J(u_h,q_h), \end{aligned}$$

since $\{u,\lambda,q\}$ and $\{u_h,\lambda_h,q_h\}$ satisfy the equations (4.27) and (4.30), respectively. Hence, error control with respect to the Lagrangian functional $L(\cdot)$ and the cost functional $J(\cdot)$ is equivalent. Now, the general error identity (4.25) implies that

$$|J(u,q) - J(u_h,q_h)| \approx \inf_{\phi_h \in V_h} |L'(v_h; v - \phi_h)|, \tag{4.34}$$

where $v_h = \{u_h,\lambda_h,q_h\}$ and $v = \{u,\lambda,q\}$. Notice that this relation is an identity if the functional $J(\cdot)$ is quadratic. From (4.29) - (4.31), we see that

$$\begin{aligned} L'(v_h; v - \phi_h) &= (u_h - u_O, u - \psi_h)_{\Gamma_O} + (\nabla(u - \psi_h), \nabla\lambda_h)_\Omega + (u - \psi_h, s'(u_h)\lambda_h)_\Omega \\ &\quad + (\nabla u_h, \nabla(\lambda - \pi_h))_\Omega + (s(u_h) - f, \lambda - \pi_h)_\Omega - (q_h, \lambda - \pi_h)_{\Gamma_C} \\ &\quad + (\lambda_h - \alpha q_h, q - \chi_h)_{\Gamma_C}. \end{aligned}$$

Splitting the global integrals into the contributions from each single cell $T \in \mathbb{T}_h$ and each cell edge $\Gamma \subset \partial\Omega$, respectively, and integrating locally by parts yields

$$\begin{aligned} L'(v_h; v - \phi_h) &= \sum_{\Gamma\subset\Gamma_O} (u_h - u_O + \partial_n\lambda_h, u - \psi_h)_\Gamma + \sum_{\Gamma\subset\partial\Omega\setminus\Gamma_O} (\partial_n\lambda_h, u - \psi_h)_\Gamma \\ &\quad + \sum_{\Gamma\subset\Gamma_C} (\partial_n u_h - q_h, \lambda - \pi_h)_\Gamma + \sum_{\Gamma\subset\partial\Omega\setminus\Gamma_C} (\partial_n u_h, \lambda - \pi_h)_\Gamma \\ &\quad + \sum_{\Gamma\subset\Gamma_C} (\lambda_h - \alpha q_h, q - \chi_h)_\Gamma \\ &\quad + \sum_{T\in\mathbb{T}_h} \big\{(-\Delta u_h + s(u_h) - f, \lambda - \pi_h)_T + \tfrac{1}{2}(n\cdot[\nabla u_h], \lambda - \pi_h)_{\partial T\setminus\partial\Omega}\big\} \\ &\quad + \sum_{T\in\mathbb{T}_h} \big\{(u - \psi_h, -\Delta\lambda_h + s'(u_h)\lambda_h)_T + \tfrac{1}{2}(u - \psi_h, n\cdot[\nabla\lambda_h])_{\partial T\setminus\partial\Omega}\big\}. \end{aligned}$$

From this the asserted relation follows by applying the Hölder inequality. □

In the computational examples presented in the next section, we evaluate the weights by using the local interpolation estimate

$$h_T^{-1}\|\phi - I_h\phi\|_T + h_T^{-1/2}\|\phi - I_h\phi\|_{\partial T} \le c_I h_T^2 \max_T |\nabla^2\phi|. \tag{4.35}$$

The derivatives $\nabla^2 u$, $\nabla^2\lambda$ and $\partial_s^2 q$ are then approximated by corresponding difference quotients of the computed solution $\{u_h,\lambda_h,q_h\}$.

Below, we will compare the performance of the dual–weighted error estimator (4.33) with more traditional error indicators. Control of the error in the "energy norm" of the state equation alone leads to the a posteriori error indicator

$$\eta_E(u_h) := c_I \sum_{T\in\mathbb{T}_h} h_T^3\, \rho_{\partial T}^{(u)2} + c_I \sum_{\Gamma\subset\partial\Omega} h_\Gamma^3\, \rho_\Gamma^{(u)2}, \tag{4.36}$$

with the cell residuals $\rho_{\partial T}^{(u)}$ and $\rho_\Gamma^{(u)}$ as defined above. Further, incorporating error control for the adjoint equation results in

$$\eta_E(u_h,\lambda_h) := c_I \sum_{T\in\mathbb{T}_h} h_T^3\big\{\rho_{\partial T}^{(u)2} + \rho_{\partial T}^{(\lambda)2}\big\} + c_I \sum_{\Gamma\subset\partial\Omega} h_\Gamma^3\big\{\rho_\Gamma^{(u)2} + \rho_\Gamma^{(\lambda)2}\big\}. \tag{4.37}$$

Both ad-hoc criteria aim at satisfying the state equation and the adjoint equation uniformly with good accuracy. However, this concept seems questionable since it does not take into account the sensitivity of the cost functional with respect to the local perturbations introduced by discretization. Capturing these dependencies is the particular feature of our approach. In practice, one may use a simplified version of the error estimator (4.33) which incorporates only the essential jump-residual terms and the residuals along the observation and control boundary:

$$\eta_\omega(u_h,\lambda_h,q_h) := \sum_{T\in\mathbb{T}_h} h_T^3\, \rho_{\partial T}^{(u)}\rho_{\partial T}^{(\lambda)} + \sum_{\Gamma\subset\partial\Omega} h_\Gamma^2\big\{\rho_\Gamma^{(\lambda)}\omega_\Gamma^{(u)} + \rho_\Gamma^{(u)}\omega_\Gamma^{(\lambda)}\big\}. \tag{4.38}$$

The performance of these different error indicators is compared by an example in the next section.

Strategies for mesh adaptation

Finally, we briefly describe strategies for mesh adaptation on the basis of a posteriori error estimates of the kind

$$\eta := \sum_{T\in\mathbb{T}_h} \eta_T, \tag{4.39}$$

with certain cell indicators, e.g. $\eta_T := h_T^3 \rho(u_h)_{\partial T}\rho(\lambda_h)_{\partial T}$. We aim at achieving a prescribed tolerance TOL for the quantity $J(u)$ and the number of mesh cells N which measures the complexity of the computational model. Usually the admissible complexity is constrained by some maximum value $N_{\max}$.

i) Error balancing strategy: We cycle through the mesh and 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 results in $\eta \approx TOL$.

ii) Fixed fraction strategy: We order the cells according to the size of η_T and refine a certain percentage (say 30%) of cells with largest η_T, or those cells which contribute to a certain percentage of the value of $\eta(u_h)$. A certain part 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).

5 Numerical results

We perform computations for the Ginzburg-Landau model described in Section 2 on a series of locally refined meshes. On each mesh, the Euler-Lagrange equations are discretized by the Galerkin finite element method as described above using piecewise bilinear shape functions for both the state and adjoint variables u and λ, while the traces on Γ_C of the bilinear shape functions form the control space Q_h. Then, the resulting discrete systems are solved iteratively and new meshes are generated on the basis of a posteriori error estimators. In all cases, the weights are evaluated by using difference approximation as described in the previous section with interpolation constants set to $c_I = 0.1$. The mesh refinement follows the "fixed fraction" strategy. The *effectivity index*

$$I_{eff} := \frac{|J(u_{ref}, q_{ref}) - J(u_h, q_h)|}{\eta(u_h, q_h)},$$

is used as a measure for the quality of our a posteriori error estimator. The reference solutions are obtained on adaptively generated meshes with about $2 \cdot 10^5$ cells.

5.1 A "forward" test case

First, we want to illustrate the difference between traditional "energy error control" and our functional-oriented "dual-weighted error control" by considering the following linear primal test example:

$$\begin{aligned} -\Delta u + u &= 0 \quad \text{on } \Omega, \\ \partial_n u = q \ \text{ on } \Gamma_C, &\quad \partial_n u = 0 \ \text{ on } \partial\Omega \setminus \Gamma_C; \end{aligned} \tag{5.1}$$

see Configuration 2 in Figure 2.1. The boundary control is frozen as $q \equiv 0.0503455$ (taken from an optimization result). The corresponding discrete equations read

$$(\nabla u_h, \nabla \psi_h)_\Omega + (u_h, \psi_h)_\Omega = (q, \psi_h)_{\Gamma_C} \quad \forall \psi_h \in V_h. \tag{5.2}$$

We control the error $e = u - u_h$ with respect to the *quadratic* observation functional

$$J(u) = \tfrac{1}{2}\|u - u_O\|^2_{\Gamma_O}.$$

The corresponding dual solution $z \in V$ is obtained by solving the corresponding system (2.5) - (2.7) with frozen boundary function q and linearized right-hand side $J'(u_h;\psi)$. The resulting a posteriori error bound is

$$|J(u) - J(u_h)| \le \eta_\omega(u_h) := \sum_{T\in\mathbb{T}_h} h_T^2 \{\rho_T^{(u)}\omega_T^{(z)} + \rho_{\partial T}^{(u)}\omega_{\partial T}^{(z)}\} + \sum_{\Gamma\in\partial\Omega} h_\Gamma^2 \rho_\Gamma^{(u)}\omega_\Gamma^{(z)}, \tag{5.3}$$

with cell residuals and weights defined as above. The asymptotic correctness of this error estimator is demonstrated in Table 5.1.

Table 5.1: Effectivity index of the dual–weighted error estimator $\eta_\omega(u_h)$ applied to the linear primal model problem; $E(u_h) := |J(u) - J(u_h)|$.

N	1376	5840	22544	57104	84368
$E(u_h)$	1.64e-05	4.17e-06	1.01e-06	3.5e-07	2.49e-07
I_{eff}	0.81	0.91	0.92	0.95	0.88

We compare the dual–weighted error estimator with the traditional energy-norm error estimator which in this case reads as follows:

$$\|\nabla e\|_\Omega^2 \leq \eta_E(u_h) := c_I \sum_{T\in\mathbb{T}_h} h_T^4 \{\rho_T^{(u)2} + \rho_{\partial T}^{(u)2}\} + c_I \sum_{\Gamma\in\partial\Omega} h_\Gamma^4 \rho_\Gamma^{(u)2}, \tag{5.4}$$

with the notation as introduced above, Clearly, small $\|\nabla e\|_\Omega$ implies small $E(u_h)$, but not vice versa. Hence, mesh adaptation based on the energy-error estimator may result in overly refined meshes. This is clearly seen in Figure 5.1.

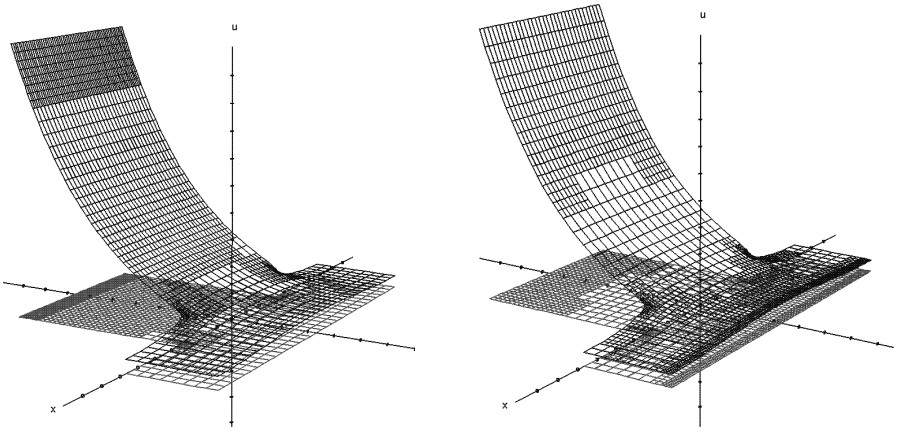

Figure 5.1: Meshes obtained by the error estimators $\eta_E(u_h)$ (left) and $\eta_\omega(u_h)$ (right) with $N \sim 5000$ cells in both cases; the graph of the solution is strongly scaled up.

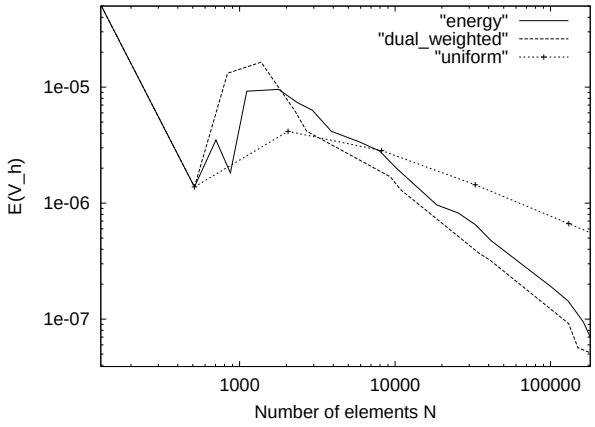

Figure 5.2: Efficiency of meshes generated by the error estimators $\eta_E(u_h)$ (solid line) and $\eta_\omega(u_h)$ (dashed line), and by uniform refinement (crosses) (log/log scale).

Obviously, the energy-error estimator puts too much emphasis on refining at the reentrant corners which is obviously less important for achieving good accuracy along the observation boundary Γ_O. In contrast to that, the dual–weighted error estimator provides a better balance between resolving the corner singularities and the neighborhood of Γ_O. This results in a higher mesh economy as shown by the corresponding error plots in Figure 5.2. This demonstrates the value of capturing the sensitivities inherent to the problem under consideration. This effect will become even more pronounced in solving the optimal control problem.

5.2 Optimal control test cases

Next, we consider the optimal control problem for the *nonlinear* Ginzburg-Landau model with $s(u) = u^3 - u$ and $f \equiv 0$. The solution approach is as described in the preceding sections.

Test for Configuration 1

We begin with Configuration 1 shown in Figure 2.1. The cost functional is chosen with $u_O = \sin(0.19x)$ and $\alpha = 0$. This test case represents an extreme situation since here the observation $u_{|\Gamma_O}$ is evaluated right at the control boundary Γ_C, i.e., the information relevant for the optimization problem does not have to pass through the domain. Hence, the corner singularities do not much affect the optimization process and should therefore not induce mesh refinement. This is accomplished by our dual–weighted error estimator as shown by Table 5.2. In this case, the heuristic energy-error indicators (4.36) and (4.37) produce inefficient meshes as seen in Figures 5.4, 5.3 and 5.5. In fact, in (4.38) the boundary indicators are dominant over the interior indicators.

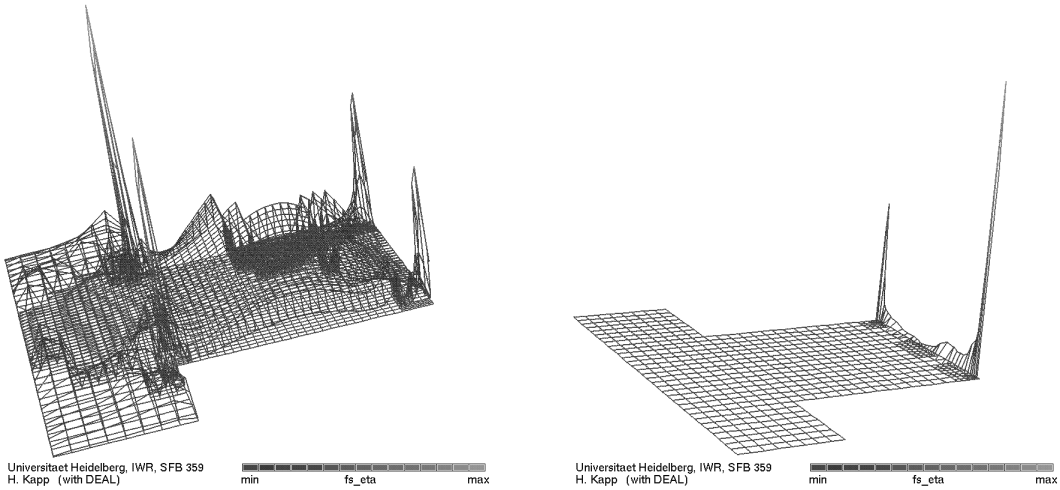

Figure 5.3: Configuration 1: Size of cell indicators η_T in the error estimators $\eta_E(u_h)$ (left) and $\eta_\omega(u_h, \lambda_h, q_h)$ (right).

Table 5.2: Configuration 1: Effectivity index of the dual-weighted error estimator $\eta_\omega(u_h, \lambda_h, q_h)$ (reference value $J(u,q) = 0.0119467...$).

N	596	1616	5084	8648	15512
$E(v_h)$	0.0002555	0.0002375	8.22e-05	4.21e-05	3.99e-05
I_{eff}	0.34	0.81	0.46	0.29	0.43

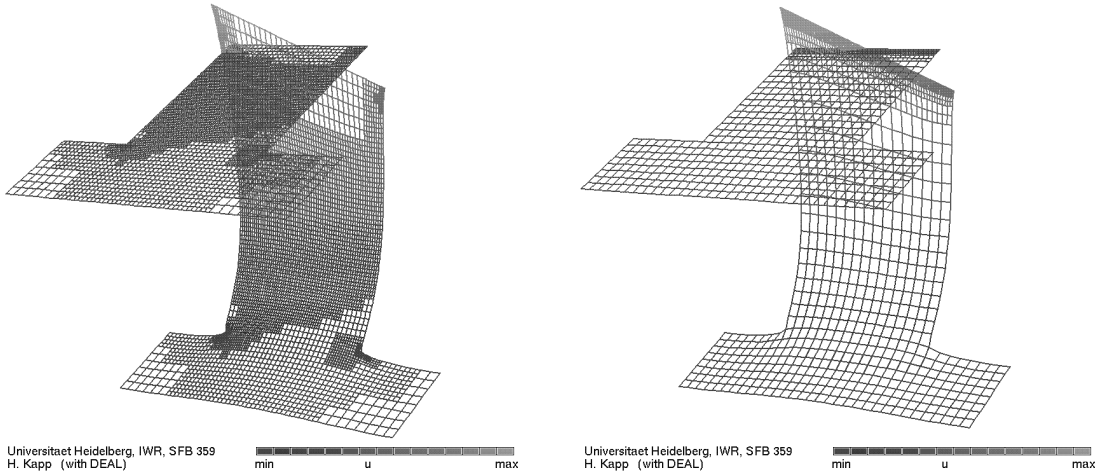

Figure 5.4: Configuration 1: Comparison of discrete solutions obtained by the error estimators $\eta_E(u_h)$ (right, $N \sim 4800$ cells) and $\eta_\omega(u_h, \lambda_h, q_h)$ (left, $N \sim 5000$ cells).

In Figure 5.5, we compare the efficiency of the meshes generated by the dual–weighted error estimator and the energy-error indicator. The first one yields significantly more economical meshes; the value 0.011948 of the cost function is obtained with only $N \sim 3500$ cells compared to the $N \sim 100000$ cells needed by the energy-norm indicator .

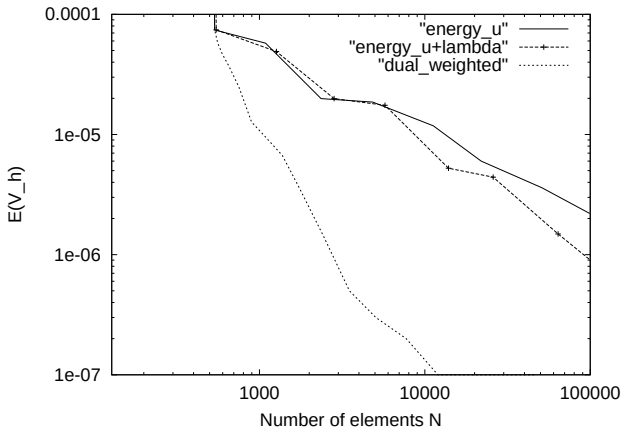

Figure 5.5: Configuratione 1: Comparison of efficiency of meshes generated by the error indicators $\eta_E(u_h)$ (solid line), $\eta_E(u_h, \lambda_h)$ (crosses), and $\eta_\omega(u_h, \lambda_h, q_h)$ (dashed line).

Test for Configuration 2

For the last test, we take the observations as $u_O \equiv 1$ and set the regularization parameter to $\alpha = 0.1$. Now, depending on the nonlinearity $s(u)$, there may be several stationary points of the Euler-Lagrange equations. By varying the starting values for the Newton iteration, we can approximate these solutions. A trivial solution corresponds to $u \equiv u_O$ and is actually the global minimum of the cost functional. The two other computed solutions correspond to a local minimum and a local maximum. The effectivity of the dual–weighted error estimator for computing the local minimum is shown in Table 5.3.

Table 5.3: Configuration 2: Effectivity index of the weighted error estimator $\eta_\omega(u_h, \lambda_h, q_h)$ for computing the local minimum (reference value $J(u,q) = 0.04888934625...$).

N	512	15368	27800	57632	197408
$E(v_h)$	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

Next, Figure 5.6 shows the distribution of the local cell indicators η_T for the two error estimators; the corresponding meshes are seen in Figure 5.7. Obviously, the weighted error estimator induces a stronger refinement along the observation and control boundaries which seems more relevant for the optimization process than resolving the corner singularities. However, the total contribution of the interior cell indicators is dominant over that of the boundary indicators. This explains why the gain in efficiency (about 25%) of the dual-weighted error estimator over the energy-norm indicator is less significant here compared to the previous example.

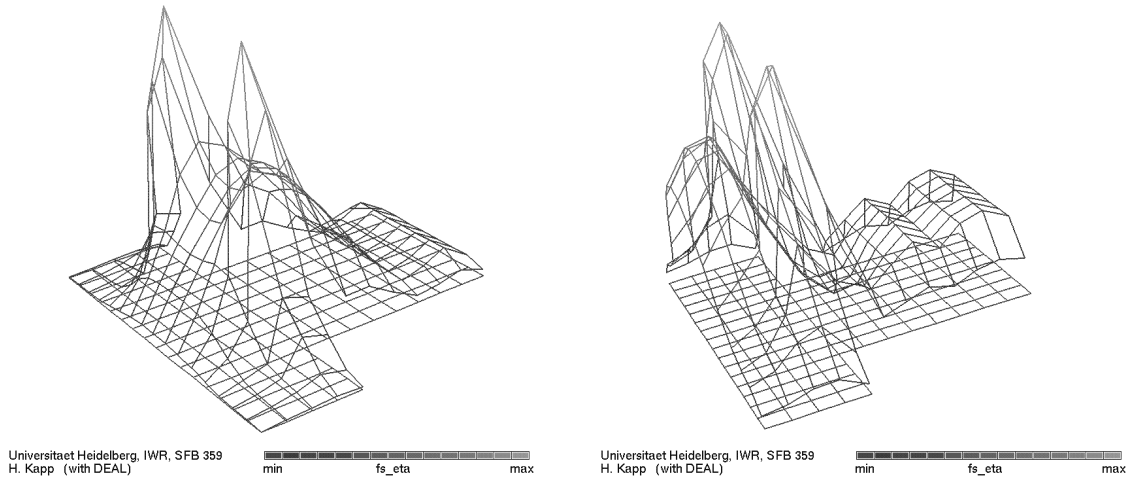

Figure 5.6: Configuration 2: Distributions of the cell indicators η_T in the error estimators $\eta_E(u_h)$ (left) and $\eta_\omega(u_h, \lambda_h, q_h)$ (right).

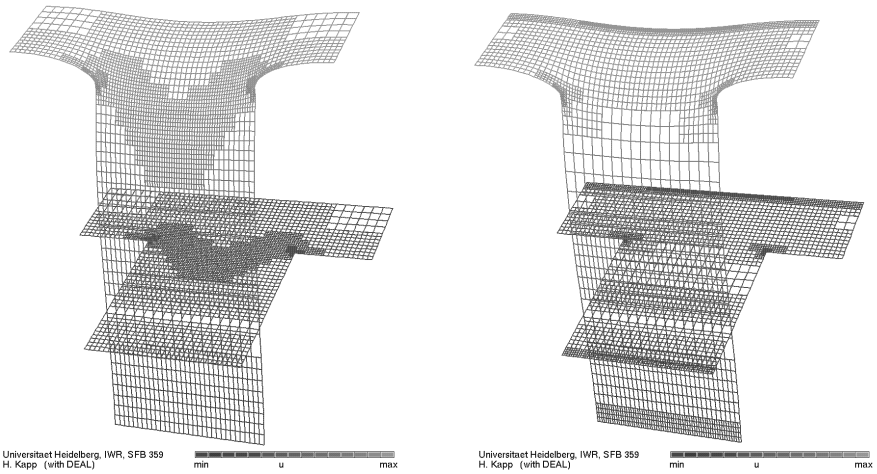

Figure 5.7: Configuration 2: Comparison of discrete solutions obtained by the error estimators $\eta_E(u_h)$ (right, $\sim$ 3300 cells) and $\eta_\omega(u_h, \lambda_h, q_h)$ (left, $\sim$ 3000 cells).

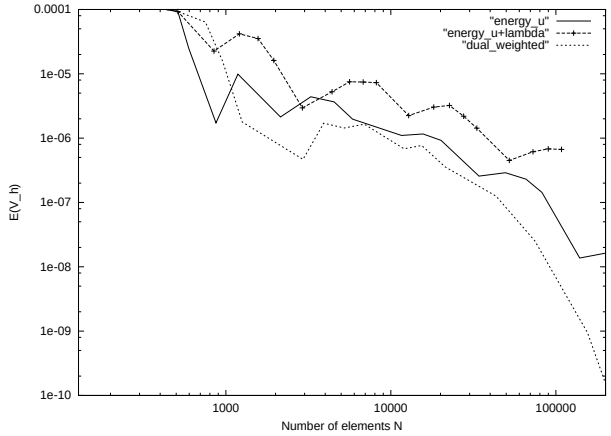

Figure 5.8: Configuration 2: Comparison of efficiency of meshes generated by the error indicators $\eta_E(u_h)$ (solid line), $\eta_E(u_h, \lambda_h)$ (crosses), and $\eta_\omega(u_h, \lambda_h, q_h)$ (dashed line).

References

[1] R. Becker and H. Kapp, *Optimization in PDE models with adaptive finite element discretization*, Proc. ENUMATH'97, Heidelberg, Sept.29 - Oct.3, 1997, World Scientific Publishers, Singapore, 1998.

[2] R. Becker, H. Kapp, and R. Rannacher, *Adaptive finite element methods for optimal control of partial differential equations: basic concepts*, Preprint 98-55, SFB 359, Univ. of Heidelberg, Nov. 1998, submitted for publication.

[3] R. Becker and R. Rannacher, *Weighted a posteriori error control in FE methods*, ENUMATH'95, Paris, Sept. 18-22, 1995, Proc. ENUMATH'97 (H. G. Bock, et al., eds.), pp. 621–637, World Scientific Publishers, Singapore, 1998.

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

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

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

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

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

[9] K. Ito and K. Kunisch, *Augmented Lagrangian-SQP methods for nonlinear optimal control problems of tracking type*, SIAM J. Contr. Opt. 34, 874-891 (1996).

[10] H. Kapp, *Adaptive Galerkin Finite Element Methods for Optimal Control of Partial Differential Equations*, Doctor Thesis, Institute of Applied Mathematics, University of Heidelberg, in preparation.

[11] R. Rannacher, *Error control in finite element computations*, Proc. NATO-Summer School "Error Control and Adaptivity in Scientific Computing" Antalya (Turkey), Aug. 1998 (H. Bulgak and C. Zenger, eds.), pp. 247–278, NATO Science Series, Series C, Vol. 536, Kluwer, Dordrecht, 1999.

[12] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley/Teubner, New York-Stuttgart, 1996.

Acknowledgement

The authors acknowlege support by the German Research Association (DFG) through the SFB 359 "Reactive Flows, Diffusion and Transport", University of Heidelberg.

Roland Becker
Hartmut Kapp
Rolf Rannacher

Institute of Applied Mathematics
University of Heidelberg
INF 293, D-69120 Heidelberg, Germany
roland.becker@iwr.uni-heidelberg.de
hartmut.kapp@iwr.uni-heidelberg.de
rannacher@iwr.uni-heidelberg.de
http://gaia.iwr.uni-heidelberg.de