# A POSTERIORI FINITE ELEMENT BOUNDS FOR SENSITIVITY DERIVATIVES OF PARTIAL–DIFFERENTIAL–EQUATION OUTPUTS *

ROBERT MICHAEL LEWIS †, ANTHONY T. PATERA ‡, AND JAUME PERAIRE §

**Abstract.** We present a Neumann-subproblem *a posteriori* finite element procedure for the efficient and accurate calculation of rigorous, "constant–free" upper and lower bounds for sensitivity derivatives of functionals of the solutions of partial differential equations. The design motivation for sensitivity derivative error control is discussed; the *a posteriori* finite element procedure is described; the asymptotic bounding properties and computational complexity of the method are summarized; and illustrative numerical results are presented.

**Key words.** *A posteriori* finite element bounds, sensitivity calculations, sensitivity equations

**Subject classification.** Applied and Numerical Mathematics

**1. Introduction.** We consider here an engineering system or component characterized by a design variable (or vector) $\beta$. We assume that the behavior of this system can be adequately represented by the solution of an appropriate partial differential equation, or perhaps set of partial differential equations. A typical engineering "forward" analysis is thus initiated by the specification of the design variable $\beta$; the partial differential equation then yields a field variable (or set of field variables) $u(\cdot;\beta)$; finally, on the basis of these intermediate field variables, the output(s) $s(\beta)$ can be evaluated. Here "output" denotes the engineering quantity of interest, that is, the metric relevant to the performance of the system. For our purposes here, we assume that this output is a *linear* functional of the field variable, $s(\beta) = \ell(u(\cdot;\beta))$.

Of ultimate interest, of course, is not the forward problem, but the "design problem." In brief, the design problem articulates the engineering objectives and constraints as a function of the outputs, and then seeks the best value of the design variable $\beta$ with respect to the selected criteria. Successful solution of the design problem requires repeated appeal to the forward problem in order to calculate (i) the output $s(\beta)$, and (ii) the sensitivity derivative of the output with respect to the design variable, $s' \equiv ds/d\beta$. The sensitivity derivatives can be important both in informal and formal design optimization contexts: in the former, $s'$ provides promising new search directions as well as an indication of design "robustness"; in the latter, $s'$ provides the gradients required by (rapidly convergent) quasi–Newton methods [11, 12].

For most problems of engineering interest, the underlying partial differential equations are far too complex to admit analytical solution, and a numerical approximation must thus be introduced; we shall consider here finite element methods. The requirements on the numerical approximation are twofold: the approxima-

*Revision: 2.0 Date: 1998/07/01 19:06:16. Submitted to Finite Elements in Design. Address all correspondence to Michael Lewis, ICASE, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia, 23681-0001, buckaroo@icase.edu.

†ICASE, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia, 23681-0001, buckaroo@icase.edu. This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-97046 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23681-0001.

‡Department of Mechanical Engineering, Masschusetts Institute of Technology. This work was supported by NASA under Grant NAG1-1978, by the AFOSR under Grant F49620-94-1-0121, and by DARPA and ONR under Grant N00014-91-J-1889.

§Department of Aeronautics and Astronautics, Masschusetts Institute of Technology. This work was supported by NASA under Grant NAG1-1978, by the AFOSR under Grant F49620-94-1-0121, and by DARPA and ONR under Grant N00014-91-J-1889.

tion must be sufficiently coarse so as to permit repeated appeal within the design context; the approximation must be sufficiently fine so that the numerical prediction of the desired outputs and associated sensitivity derivatives is representative of the true performance of the system.

*A posteriori* error control offers great promise in reconciling these often conflicting requirements. *A posteriori* analysis [3, 29] is composed of two critical ingredients: an estimation procedure which inexpensively assesses the error in a particular numerical solution; and an adaptive refinement procedure which exploits this error information to optimally improve the numerical solution. There are two objectives of *a posteriori* error control: to eliminate numerical uncertainty — arguably the single largest impediment to widespread adoption of simulation–based design; and to improve the numerical efficiency of the forward *and* optimization problem — thus permitting much more extensive design exploration.

In fact, greater certainty is a prerequisite for greater efficiency: we may consider a less expensive (or even *the* least expensive) discretization only if the associated error can be quantified, and hence constrained and controlled; we are no longer compelled to choose *either* certainty or efficiency — both can be achieved. Simultaneous control of approximation accuracy and approximation cost is particularly critical in the development of robust and effective design optimization procedures: for example, the accuracy of the sensitivity derivatives $s'$ strongly affects both the convergence, and the convergence rate, of (say) trust region [9, 19] and line–search [11, 20] quasi–Newton techniques [28].

In all earlier *a posteriori* error analysis techniques, either — in implicit approaches — the measure of the error is *not related to the actual engineering outputs of interest* (e.g., [15, 4, 2]), or — in explicit approaches — the error estimates for the engineering outputs of interest involve *numerous undetermined or uncertain constants or functions* (e.g.,[5, 6, 27]); in both cases, quantitative confirmation — and hence both certainty and efficiency — is seriously compromised, and the relevance to engineering design greatly reduced. In [21, 23, 25, 18] we propose a new class of *a posteriori* procedures that provide a critical new "enabling technology": the ability to obtain inexpensive, sharp, rigorous, and quantitative ("constant–free") *bounds* for the numerical error in the *engineering outputs of interest*. Our method is thus directly relevant to the design process, and should lay the foundation for systemic application of *a posteriori* error control within the engineering context.

Although our method provides a critical new capability, we are nevertheless indebted to earlier *a posteriori* implicit (Neumann subproblem) techniques [15, 4, 2] for several important conceptual and mathematical ingredients — in particular duality theory and flux "hybridization." The former, though not strictly necessary — and even sometimes restrictive — provides a derivational mechanism without which the requisite equations are very difficult to motivate; the latter — technically quite subtle — is the crucial component in ensuring computational efficiency. Our framework may thus be viewed as a generalization of earlier implicit techniques [2]; equivalently, these earlier techniques can be interpreted as special cases of our new formulation. We present in [25] a more detailed comparison between our approach and earlier implicit (and explicit) *a posteriori* error control proposals.

Our initial formulation [23, 25, 24] focused on symmetric coercive problems (e.g., Poisson and Linear Elasticity), and nonsymmetric coercive problems (e.g., Convection–Diffusion), as well as certain constrained problems (the Stokes equations, central to hydrodynamics [22]). However, we have recently developed a more general formulation [26, 17, 16] that greatly expands the class of equations and outputs that may be treated; in particular, noncoercive problems (e.g., the Helmholtz equation), nonlinear problems (e.g., the Burgers equation, the eigenvalue problem), and nonlinear outputs can now be addressed. The approach is relevant both to Galerkin techniques [17] and stabilized Galerkin methods [16].

In this paper we extend our framework to the case in which we are interested not only in error bounds and adaptive procedures for the output $s$, but also in error bounds and adaptive procedures for the sensitivity derivative $s'$. As already indicated, effective control of the accuracy of sensitivity derivatives is crucial in ensuring both the quality and efficiency of engineering optimization procedures [9, 10]. It is thus surprising that there is relatively little work on *a posteriori* error estimation relevant to sensitivity derivatives [7], in particular since there is considerable debate over the best way — discrete or continuous [28, 13, 8, 14] — to calculate the sensitivity functions and adjoints required to compute $s'$. In this work we aim to at least partially address the paucity of *a posteriori* results for this important class of problems.

In Section 2 we describe our model problem. For our purposes here we consider a simple one–dimensional coercive differential equation: noncoercive, semi–linear, and multi-dimensional problems [17] require virtually no modifications to the framework, and will be addressed in future publications. Furthermore, in this first paper we focus exclusively on *a posteriori* error *estimators* for $s'$; once these estimators are obtained, the adaptive procedures described in [25, 16] can be directly applied. We next introduce in Section 3 the relevant finite–element spaces and approximations. In Section 4 we describe the *a posteriori* procedure; Section 5 then discusses the computational complexity of this procedure and the asymptotic bounding properties of the resulting lower and upper estimators. Finally, in Section 6, we present a series of illustrative numerical results.

**2. Problem Statement.** As our model problem we shall consider a second–order inhomogeneous Dirichlet problem for $u(x;\beta)$,

$$-\nu(\beta)u_{xx} + \rho(\beta)u_x + \alpha(\beta)u = f \text{ in } \Omega,$$

$$u(0) = 0, u(1) = 1,$$

where $f$ is the prescribed inhomogeneity and $\Omega = ]0,1[$ is the domain. Here $\nu$, $\rho$, and $\alpha$ are respectively strictly positive, general, and non–negative real quantities which are independent of $x$ but may depend on the design parameter $\beta$. (Extension of the framework to permit these coefficients to vary with $x$ is straightforward, as will be clear from the exposition; examples will be presented in future publications.) Although we consider here the Dirichlet problem, Neumann problems can also be readily treated.

Our point of departure shall be the weak formulation: Find $u \in X^D$ such that

$$a(u,v) = \langle f, v\rangle,\ \forall v \in X, \tag{2.1}$$

where $X = H_0^1(\Omega)$ and $X^D = H_D^1(\Omega)$. We recall that $H_0^1(\Omega)$ is the space of functions $v$ such that (i) $v$ and $v_x$ are in $L^2(\Omega)$ (that is, square integrable [1]), and (ii) $v(0) = v(1) = 0$; similarly, $H_D^1(\Omega)$ is the set of functions $v$ such that (i) $v$ and $v_x$ are in $L^2(\Omega)$, and (ii) $v(0) = 0$, $v(1) = 1$. The bilinear form $a(w,v)$ (or more precisely $a(w,v;\beta)$) is given by

$$a(w,v) = \int_0^1 \nu(\beta)w_x v_x + \rho(\beta)w_x v + \alpha(\beta)wv\,dx. \tag{2.2}$$

Finally, $f \in H^{-1}(\Omega)$ (the dual of $H^1(\Omega)$), and $\langle\ ,\ \rangle$ thus denotes the duality pairing.

Our interest is not directly in $u(x;\beta)$, but rather in the output $s(\beta) = \ell(u(x;\beta))$ and the sensitivity derivative $s'(\beta)$. The latter may be computed as

$$s'(\beta) = \ell(U(x;\beta)), \tag{2.3}$$

where $U \equiv u'(x;\beta) \equiv \frac{du(x;\beta)}{d\beta} \in X$, the sensitivity function, satisfies

$$a(U, V) = \langle f', V\rangle - a'(u, V),\ \forall V \in X, \tag{2.4}$$

in which

$$a'(w, v) = \int_0^1 \nu'(\beta) w_x v_x + \rho'(\beta) w_x v + \alpha'(\beta) w v \, dx; \tag{2.5}$$

here $'$ denotes, as always, $\frac{d}{d\beta}$. We could also readily permit the Dirichlet data to depend on $\beta$; in this case $U$ in (2.4) would satisfy inhomogeneous Dirichlet conditions. Finally, we note that the sensitivity function is only one of two approaches to the computation of sensitivity derivatives; the adjoint formulation will be considered in future publications.

For the formulation presented here the computation of the sensitivity derivatives thus requires, first, the solution of (2.1) for $u$, then the solution of (2.4) for $U$, and finally the evaluation of $s'(\beta)$ from (2.3): (2.1) and (2.4) are coupled but "triangular." However, if the only dependence on $\beta$ is through $f$ (or, in fact, the Dirichlet data) — that is, $\nu' = \rho' = \alpha' = 0$, and hence $a' = 0$ — then the system is no longer coupled: we can compute $U$ and hence $s'$ independently of $u$ and $s$. In this instance — which arises in many design contexts, including certain open–loop control problems and shape optimization formulations — we may directly apply our earlier bound procedures [23, 25, 17] to (2.4) and (2.3); as we shall see, this "direct" approach is, in fact, a special case of the more general framework developed in this paper.

Finally, we close this section by slightly generalizing our problem statement (2.1), (2.4), (2.3). In particular, we introduce a real positive parameter $\sigma$, and replace (2.1), (2.4), and (2.3) with

$$a(u, v) = \langle f, v\rangle,\ \forall v \in X, \tag{2.6}$$

$$a(U, V) = \sigma(\langle f', V\rangle - a'(u, V)),\ \forall V \in X, \tag{2.7}$$

and

$$s'(\beta) = \frac{1}{\sigma}\ell(U), \tag{2.8}$$

respectively. It is clear from the linearity of our output functional that (2.1), (2.4), (2.3) and (2.6), (2.7), (2.8) do indeed yield the same sensitivity derivative; this simply reflects the invariance of the sensitivity derivative to rescalings of the design parameter. However, (2.6), (2.7), (2.8) permits us to control the magnitude of $a'$ relative to the magnitude of $a$ in the coupled system (2.6), (2.7); this will be advantageous in the context of our error bound procedure.

It will prove very convenient in what follows to write (2.6), (2.7) more succinctly in terms of a product space. In particular, we define $Y^D = X^D \otimes X$, and $Y = X \times X$, and introduce the bilinear form

$$\mathcal{E}((w, W), (v, V)) = a(w, v) + a(W, V) + \sigma a'(w, V), \tag{2.9}$$

for $a$ and $a'$ as defined in (2.2) and (2.5), respectively. It is then simple to see that $(u, U) \in Y^D$ satisfies

$$\mathcal{E}((u, U), (v, V)) = \langle f, v\rangle + \sigma\langle f', V\rangle,\ \forall (v, V) \in Y, \tag{2.10}$$

since variations in $v$ and $V$ recreate (2.6) and (2.7), respectively. (Note that $\mathcal{E}$ in (2.9) can be further generalized: the $v$ and $V$ contributions need not be assigned the same weight.)

**3. Finite Element Approximation.** We shall consider here only uniform meshes, although in practice (adaptive) non–uniform meshes are preferred. We first introduce a uniform triangulation of $\Omega$, $\mathcal{T}_\delta$, comprising $K_\delta$ elements $T_\delta$ of uniform length $\delta = K_\delta^{-1}$; the corresponding $N_\delta = K_\delta + 1$ nodes of $\mathcal{T}_H$ are denoted by $x_{\delta\, i} = (i-1)\delta, i = 1, \ldots, N_\delta$. We then define our linear finite element approximations spaces as

$$X_\delta = \{v|_{T_\delta} \in I\!P_1(T_\delta),\ \forall T_\delta \in \mathcal{T}_\delta\} \cap X, \tag{3.1}$$

$$X_\delta^D = \{v|_{T_\delta} \in I\!P_1(T_\delta),\ \forall T_\delta \in \mathcal{T}_\delta\} \cap X^D, \tag{3.2}$$

where $I\!P_1(T_\delta)$ refers to the space of linear polynomials over $T_\delta$. Finally, we define our approximation product spaces as $Y_\delta^D = X_\delta^D \otimes X_\delta$ and $Y_\delta = X_\delta \otimes X_\delta$.

Our finite element approximation $(u_\delta, U_\delta) \in Y_\delta^D, s'_\delta \in I\!R$ to $(u, U) \in Y^D, s' \in I\!R$ of (2.6), (2.7), (2.8) is then given by

$$\mathcal{E}((u_\delta, U_\delta), (v, V)) = \langle f, v\rangle + \sigma\langle f', V\rangle,\ \forall (v, V) \in Y_\delta, \tag{3.3}$$

and

$$s'_\delta = \frac{1}{\sigma}\ell(U_\delta). \tag{3.4}$$

Summarizing the *a priori* theory, it is readily demonstrated that the problems (2.10) for $(u, U)$ and (3.3) for $(u_\delta, U_\delta)$ are well-posed, yielding unique solutions. It can be further shown that, given sufficient regularity, $\|u - u_\delta\|_1$ and $\|U - U_\delta\|_1$ vanish as $O(\delta)$, where $\|\cdot\|_1$ denotes the $H^1$ norm; in addition, application of Aubin–Nitsche theory to our product space formulation confirms that, as expected, $\|u - u_\delta\|_0$ and $\|U - U_\delta\|_0$ vanish as $O(\delta^2)$, where $\|\cdot\|_0$ denotes the $L^2$ norm. Finally, Aubin–Nitsche theory also readily reveals that both $|s - s_\delta|$ *and* $|s' - s'_\delta|$ vanish as $O(H^2)$ (here $s_\delta = \ell(u_\delta)$).

For the purposes of our bound formulation we shall be interested in two particular finite element spaces: a coarse, or design, approximation space, $X_H \equiv X_{\delta=H}$; and a fine, or "truth," approximation, $X_h \equiv X_{\delta=h}$. We require that $H/h$ is integral such that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$. The corresponding coarse and fine finite element approximations — solutions of (3.3), (3.4) — will be denoted $(u_H, U_H) \in Y_H^D, s'_H \in I\!R$ and $(u_h, U_h) \in Y_h^D, s'_h \in I\!R$, respectively. In brief, the coarse approximation is the approximation that we *must* use due to computational constraints; the fine approximation is the approximation that we would *like* to use — that is, the approximation on which we can (and will) safely assume that $u_h, U_h, s_h$, and $s'_h$ are all arbitrarily close to the corresponding exact quantities $u, U, s$, and $s'$. For future reference we define $e = u_h - u_H$ and $E = U_h - U_H$; it follows from our *a priori* results that, for $h$ sufficiently small compared to $H$, $\|e\|_1$ and $\|E\|_1$ vanish as $O(H)$, while $\|e\|_0, \|E\|_0$, and $|s'_h - s'_H|$ vanish as $O(H^2)$.

Our goal will be to obtain *bounds* for $s'_h$ — the sensitivity derivative on the fine mesh — at considerably less cost than *direct computation* of $s'_h = \ell(U_h)$; indeed, the cost to compute the bounds will typically be only slightly greater than the cost to compute the *coarse* sensitivity derivative, $s'_H$. The coarse approximation will provide the "guesses" — in fact, Lagrange multipliers — which will ensure that these bounds are sufficiently accurate; in particular, given our *a priori* results, we require that the bounds approach $s'_h$ as $O(H^2)$. We will, indeed, achieve this desired (optimal) rate of convergence.

**4. Bound Formulation.**

**4.1. Preliminaries: Spaces and Forms.** We first introduce coarse and fine broken spaces,

$$\hat{X}_H = \{v|_{T_H} \in \mathbb{P}_1(T_H),\ \forall T_H \in \mathcal{T}_H\} \tag{4.1}$$

$$\hat{X}_h = \{v|_{T_h} \in \mathbb{P}_1(T_h),\ \forall T_h \in \mathcal{T}_h \,|\, v|_{T_H} \in C^0(T_H),\ \forall T_H \in \mathcal{T}_H\}, \tag{4.2}$$

where $C^0(T_H)$ is the space of continuous functions over $T_H$; note that the fine broken space is continuous *within* each $T_H$. We can thus form the associated broken product spaces $\hat{Y}_H = \hat{X}_H \otimes \hat{X}_H$ and $\hat{Y}_h = \hat{X}_h \otimes \hat{X}_h$. We next introduce the hybrid flux space $\mathcal{Q} \equiv \mathbb{R}^{N_H}$, the members $q_i$ of which are defined at the $x_{H\,i}, i = 1, \ldots, N_H$, the nodes of the coarse approximation; the corresponding product space is given by $Z = \mathcal{Q} \otimes \mathcal{Q}$.

We then define the bilinear form $b$: $\hat{Y}_h(\text{or } \hat{Y}_H) \times Z \to \mathbb{R}$ as

$$b((v, V), (q, Q)) = \sum_{n=1}^{N_H} [v]_n q_n + \sum_{n=1}^{N_H} [V]_n Q_n, \tag{4.3}$$

where $[w]$ denotes the jump at the interfaces: for any $1 < i < N_H$, $[w]_n = w(x_{H\,n}^+) - w(x_{H\,n}^-)$, where $w(x_{H\,n}^+)$ (respectively, $w(x_{H\,n}^-)$) refers to the limit as $x \to x_{H\,n}$ from the right (respectively, left); at the two endpoints, we define $[w]_1 = w(0)$ and $[w]_{N_H} = -w(1)$. The bilinear form imposes continuity (and homogeneous Dirichlet boundary conditions) on members of the broken spaces in the sense that

$$Y_H = \{(v, V) \in \hat{Y}_H \,|\, b((v, V), (q, Q)) = 0,\ \forall (q, Q) \in Z\}, \tag{4.4}$$

$$Y_h = \{(v, V) \in \hat{Y}_h \,|\, b((v, V), (q, Q)) = 0,\ \forall (q, Q) \in Z\}. \tag{4.5}$$

Extension of these spaces and bilinear forms to the multidimensional case is discussed in detail in [23, 25, 18].

We shall also require several forms related to $\mathcal{E}$ of (2.9). First, we define in standard fashion the symmetric part of $\mathcal{E}$,

$$\mathcal{E}^s((w, W), (v, V)) = a^s(w, v) + a^s(W, V) + \frac{\sigma}{2}(a'(w, V) + a'(v, W)), \tag{4.6}$$

where $a^s$ is the symmetric part of $a$ — for our boundary conditions simply $a$ with the convective term $(\rho(\beta) w_x v)$ omitted. We then further decompose $\mathcal{E}^s = \mathcal{E}^s_Y + \mathcal{E}^s_M$; we shall consider two such decompositions, Alternative A and Alternative B. In Alternative A we choose

$$\mathcal{E}^s_Y((w, W), (v, V)) = a^s(w, v) + a^s(W, V) + \frac{\sigma}{2} \int_0^1 \nu'(w_x V_x + v_x W_x) dx, \tag{4.7}$$

$$\mathcal{E}^s_M((w, W), (v, V)) = \frac{\sigma}{2}\Big(\int_0^1 \rho'(w_x V + v_x W) dx + \int_0^1 \alpha'(wV + vW) dx\Big). \tag{4.8}$$

In Alternative B we choose for $\mathcal{E}^s_Y$ and $\mathcal{E}^s_M$

$$\begin{aligned}\mathcal{E}^s_Y((w, W), (v, V)) = {} & a^s(w, v) + a^s(W, V) + \frac{\sigma}{2} \int_0^1 \nu'(w_x V_x + v_x W_x) dx \\ & + \frac{\sigma}{2} \int_0^1 \rho'(w_x V + v_x W) dx + \frac{\sigma}{2} |\rho'| \int_0^1 WV\, dx,\end{aligned} \tag{4.9}$$

and

$$\mathcal{E}^s_M((w, W), (v, V)) = \frac{\sigma}{2} \int_0^1 \alpha'(wV + vW) dx - \frac{\sigma}{2} |\rho'| \int_0^1 WV\, dx, \tag{4.10}$$

respectively.

**4.2. Estimator Procedure.** The estimator procedure comprises five steps [23, 18, 17], which we now describe. Note that we unfold some of the product forms into somewhat more transparent notation in the next section when we discuss computational complexity in greater detail.

1. Compute $(u_H, U_H) \in Y_H^D$ from

$$\mathcal{E}((u_H, U_H), (v, V)) = \langle f, v\rangle + \sigma\langle f', V\rangle,\ \forall (v, V) \in Y_H, \tag{4.11}$$

and define the associated residual

$$\mathcal{R}_H^u((v, V)) \equiv \langle f, v\rangle + \sigma\langle f', V\rangle - \mathcal{E}((u_H, U_H), (v, V)), \tag{4.12}$$

for any $(v, V)$ in $\hat{Y}_h$.

2. Compute the adjoint $(\psi_H, \Psi_H) \in Y_H$ from

$$\mathcal{E}((v, V), (\psi_H, \Psi_H)) = -\frac{1}{\sigma}\ell(V),\ \forall (v, V) \in Y_H, \tag{4.13}$$

and define the associated residual

$$\mathcal{R}_H^\psi((v, V)) = -\frac{1}{\sigma}\ell(V) - \mathcal{E}((v, V), (\psi_H, \Psi_H)), \tag{4.14}$$

for any $(v, V)$ in $\hat{Y}_h$.

3. Find the hybrid fluxes $(p_H^u, P_H^U) \in Z$ and $(p_H^\psi, P_H^\Psi) \in Z$ such that

$$b((v, V), (p_H^u, P_H^U)) = \mathcal{R}_H^u((v, V)),\ \forall (v, V) \in \hat{Y}_H, \tag{4.15}$$

$$b((v, V), (p_H^\psi, P_H^\Psi)) = \mathcal{R}_H^\psi((v, V)),\ \forall (v, V) \in \hat{Y}_H. \tag{4.16}$$

These equations have a unique solution thanks to equilibrium (4.11), (4.12); in higher space dimensions the matter is considerably more subtle [23, 18], though now well understood thanks to the important contributions in [15, 4, 2].

4. Find the reconstructed errors $(\hat{e}^u, \hat{E}^U) \in \hat{Y}_h$, $(\hat{e}^\psi, \hat{E}^\Psi) \in \hat{Y}_h$ such that

$$2\mathcal{E}_Y^s((\hat{e}^u, \hat{E}^U), (v, V)) = \mathcal{R}_H^u((v, V)) - b((v, V), (p_H^u, P_H^U)),\ \forall (v, V) \in \hat{Y}_h, \tag{4.17}$$

$$2\mathcal{E}_Y^s((\hat{e}^\psi, \hat{E}^\Psi), (v, V)) = \mathcal{R}_H^\psi((v, V)) - b((v, V), (p_H^\psi, P_H^\Psi)),\ \forall (v, V) \in \hat{Y}_h. \tag{4.18}$$

The well–posedness of this problem for our different choices of $\mathcal{E}_Y^s$ will be discussed shortly.

5. Compute the lower ($s'_-$) and upper ($s'_+$) estimators as $s'_\pm = \overline{s}' \pm \Delta'$, where

$$\overline{s}' = s'_H - 2\mathcal{E}_Y^s((\hat{e}^u, \hat{E}^U), (\hat{e}^\psi, \hat{E}^\Psi)), \tag{4.19}$$

$$\Delta' = 2\sqrt{\mathcal{E}_Y^s((\hat{e}^u, \hat{E}^U), (\hat{e}^u, \hat{E}^U))\, \mathcal{E}_Y^s((\hat{e}^\psi, \hat{E}^\Psi), (\hat{e}^\psi, \hat{E}^\Psi))}. \tag{4.20}$$

The difference between the upper and lower estimators — the bound *gap*, $\Delta'$ — can be decomposed into elemental contributions; the latter then serve as local indicators for adaptive refinement [25, 16].

Note that, for any given $\sigma$, these estimators have already been optimized over the scaling parameter $\kappa$ of [25, 18, 17].

We now address the well–posedness of Step 4; to begin, we consider Alternative A, with $\alpha = 0$. We first note from (4.7) that (4.17), (4.18) is singular in this case: $\mathcal{E}^s_Y((\ ,\ ),(v,V))$ vanishes for $(v,V)|_{T_H} = (c_1^{T_H}, c_2^{T_H})$, $\forall T_H \in \mathcal{T}_H$, where the $c_1^{T_H}, c_2^{T_H}$ are real constants. However, thanks to the definition of the hybrid fluxes, (4.15), (4.16), the problem is solvable; furthermore, it is clear from (4.19), (4.20) that the final bounds in Step 5 do not depend on the nullspace component chosen — though for theoretical purposes we should select the reconstructed errors to be of zero mean over each $T_H$ [18]. We next note that, for $0 < \sigma \le 2\nu/|\nu'|$, $\mathcal{E}^s_Y$ is positive semi–definite over $\hat{Y}_h$, since

$$\begin{aligned} \mathcal{E}^s_Y((v,V),(v,V)) &= a^s(v,v) + a^s(V,V) + \sigma\nu' \int_0^1 v_x V_x dx \\ &\ge \nu \int_0^1 (v_x^2 + V_x^2) dx - \frac{\sigma}{2}|\nu'| \int_0^1 (v_x^2 + V_x^2) dx; \end{aligned} \tag{4.21}$$

we have used here the standard inequality

$$|ab| \le \frac{1}{2}(\varepsilon a^2 + \frac{b^2}{\varepsilon}) \tag{4.22}$$

for $a$ and $b$ real numbers and $\varepsilon = 1$. It further follows from (4.21) that, first, (4.17), (4.18) is well–posed and stable (coercive) over the nullspace–excised quotient space, and second, the argument of the radical in Step 5 is always non–negative. In short, the entire bound procedure is well–posed. The case in which $\alpha \neq 0$ is, in fact, simpler: the problems are no longer singular, however the advantageous zero–mean property still obtains [18].

Alternative B leads to a rather similar analysis. We presume that $\rho' \neq 0$, as otherwise Alternative A and Alternative B are identical; and we first consider the more difficult case, $\alpha = 0$. We find from (4.9) that (4.17), (4.18) is again singular, however the nullspace is now different: $\mathcal{E}^s_Y((\ ,\ ),(v,V))$ vanishes for $(v,V)|_{T_H} = (c_1^{T_H}, 0)$, $\forall T_H \in \mathcal{T}_H$, where the $c_1^{T_H}$ are real constants. However, due to the definition of the hybrid fluxes, (4.15), (4.16), the problem is solvable; furthermore, it is clear from (4.19), (4.20) that the final bounds in Step 5 do not depend on the nullspace component chosen — for theoretical reasons we select $\hat{e}^u, \hat{e}^\psi$ to be of zero mean. (It can also be argued that the elemental mean of $\hat{E}^U, \hat{E}^\Psi$, though not zero, will be suitably small.) We next note that, for $0 < \sigma < 2\nu/(|\nu'| + |\rho'|)$, $\mathcal{E}^s_Y$ is positive semi–definite over $\hat{Y}_h$, since

$$\begin{aligned} \mathcal{E}^s_Y((v,V),(v,V)) \ge \nu \int_0^1 (v_x^2 + V_x^2) dx - \frac{\sigma}{2}|\nu'| \int_0^1 (v_x^2 + V_x^2) dx \\ -\frac{\sigma}{2}|\rho'| \int_0^1 (v_x^2 + V^2) dx + \frac{\sigma}{2}|\rho'| \int_0^1 V^2 dx, \end{aligned} \tag{4.23}$$

where we have again evoked (4.22) with $\varepsilon = 1$. It follows, as for Alternative A, that first, (4.19), (4.20) is well–posed and stable (coercive) over the nullspace–excised quotient space, and second, the argument of the radical in Step 5 is always non–negative. The case in which $\alpha \neq 0$ poses no problem: the problems are no longer singular, however the zero–mean property for $\hat{e}^u, \hat{e}^\psi$ still obtains.

**5. Estimator Properties.** In order to be of interest, the estimators must enjoy certain properties: (i) the estimators $s'_-, s'_+$ must correspond to bounds under appropriate hypotheses, and converge to the true result $s'_h$ sufficiently quickly; and (ii) the estimators must be inexpensive to compute relative to $\ell(U_h)$. We consider the former in Section 5.1, and the latter in Section 5.2. Note that, as regards computational complexity, we effectively anticipate multi–dimensional problems.

**5.1. Bounding Properties.** We consider only the lower estimator; similar results obtain for the upper estimator. To begin, we note that it follows directly from the general formulation of [17, 16] that, for both Alternative A and Alternative B,

$$s'_- = s'_h - \kappa^* E^s_Y((e-\hat{e}, E-\hat{E}), (e-\hat{e}, E-\hat{E})) - \kappa^* E^s_M((e,E),(e,E)), \tag{5.1}$$

where

$$(\hat{e}, \hat{E}) = (\hat{e}^u + \frac{1}{\kappa^*}\hat{e}^\psi, \hat{E}^U + \frac{1}{\kappa^*}\hat{E}^\Psi), \tag{5.2}$$

and

$$\kappa^* = \sqrt{E^s_Y((\hat{e}^\psi, \hat{E}^\Psi),(\hat{e}^\psi, \hat{E}^\Psi))/E^s_Y((\hat{e}^u, \hat{E}^U),(\hat{e}^u, \hat{E}^U))} \tag{5.3}$$

is the optimal scaling parameter of [25, 17]. We assume that $\sigma$ is chosen such that $E^s_Y$ is positive–definite.

Considering first the second — negative–definite — term of (5.1), it is clear from our *a priori* results for the $H^1$ error that this term should vanish no faster than $O(H^2)$. In fact, based on arguments similar to those developed in [17], it can be shown that, given sufficient regularity, this negative–definite term will indeed vanish quadratically. We now turn to the third — indefinite – term. For Alternative A we obtain from (4.8) and the inequality (4.22) that

$$|E^s_M((e,E),(e,E))| \le \frac{\sigma}{2}(|\rho'|\int_0^1 (\varepsilon e_x^2 + \frac{E^2}{\varepsilon})dx + |\alpha'|\int_0^1 (e^2+E^2)dx). \tag{5.4}$$

From our *a priori* results for the $H^1$ and $L^2$ error it follows that, for the choice $\varepsilon = H$, $|E^s_M((e,E),(e,E))|$ will vanish at least as fast as $O(H^3)$. For Alternative B, (4.10), we obtain in a similar fashion that

$$|E^s_M((e,E),(e,E))| \le \frac{\sigma}{2}(|\alpha'|\int_0^1 (e^2+E^2)dx + |\rho'|\int_0^1 E^2 dx), \tag{5.5}$$

which, from our *a priori* results for the $L^2$ error, should vanish as $O(H^4)$.

We can thus draw two conclusions. First, for $H$ sufficiently small, the indefinite term in (5.1) will be subdominant (in fact, vanishingly small) relative to the principal (negative–definite) contribution, and thus $s'_-$ will approach $s'_h$ from below — we obtain an *asymptotic lower bound*. Similar arguments demonstrate that $s'_+$ approaches $s'_h$ from above, and thus $s'_+$ is an *asymptotic upper bound*. For the case in which only $\nu'$ and $f'$ are non–zero, we obtain rigorous lower and upper bounds for all $H$; in practice, we find that, even for $\rho'$ and $\alpha'$ non–zero, it is very difficult *not* to obtain bounds — consistent with earlier applications of our bound procedure to noncoercive and nonlinear problems [26, 17, 16]. Second, we conclude that the estimators should converge to $s'_h$ at the optimal rate — $O(H^2)$ for our linear finite element approximation.

**5.2. Computational Complexity.** There are two issues that must be addressed: first, how does the computational cost of $s'_-$ (say) compare to that of $s'_h = \ell(U_h)$?; and second, how does the complexity scale when we permit several outputs, $s^m \equiv \ell^m(u), m = 1, \dots, M$, and several design parameters (or "inputs"), $\beta_j, j = 1, \dots, J$? The former is discussed in great detail in [23, 25, 18, 16]. The essential point is that the work on the fine mesh — Step 4 — is reduced to computations over the *broken* space $\hat{Y}_h$: it follows from the superlinearity of most solution strategies (at least in higher space dimensions) [16] that computation of $(\hat{e}^u, \hat{E}^U)$ and $(\hat{e}^\psi, \hat{E}^\Psi)$ is *much less expensive* than computation of $u_h, U_h$ — indeed, often no more expensive than computation of $u_H, U_H$. There are several other factors that further significantly decrease the cost of $(\hat{e}^u, \hat{E}^U)$ and $(\hat{e}^\psi, \hat{E}^\Psi)$ relative to $u_h, U_h$: the equations for $(\hat{e}^u, \hat{E}^U)$ and $(\hat{e}^\psi, \hat{E}^\Psi)$ will be symmetric, linear, and positive definite even when the original operator enjoys none of these properties; and the equations for $(\hat{e}^u, \hat{E}^U)$ and $(\hat{e}^\psi, \hat{E}^\Psi)$ admit obvious — communication–free, completely concurrent — medium–grained parallelism.

Turning now to the multiple-output, multiple-input question, our interest is in computing bounds for the $MJ$ quantities $(s_h)_j^m$, the derivatives of the $s_h^m = \ell^m(u_h)$ with respect to the $\beta_j$. We now address each step of the bound procedure. We shall assume that the usual nodal bases are evoked to transform the weak statements into appropriate linear algebraic systems, $\underline{A}\,\underline{x} = \underline{y}$; $\underline{A}$ and $\underline{y}$ shall be denoted the "matrix" and "right–hand side," respectively.

1. We first compute $u_H \in X_H^D$,

$$a(u_H, v) = \langle f, v\rangle,\ \forall v \in X_H, \tag{5.6}$$

and then $(U_H)_j \in X_H, j = 1, \ldots, J$,

$$a((U_H)_j, V) = \sigma(\langle f_j, V\rangle - a_j(u_H, V)),\ \forall V \in X_H, \tag{5.7}$$

where $a_j(w, v)$ (respectively $f_j$) refers to the derivative of the bilinear form $a$ (respectively $f$) with respect to $\beta_j$. We must thus solve one system in (5.6), and $J$ systems in (5.7). The $J + 1$ systems in (5.6), (5.7) share a common matrix; only the right–hand side varies.

2. We first compute $\Psi_H^m \in X_H, m = 1, \ldots, M$,

$$a(V, \Psi_H^m) = -\frac{1}{\sigma}\ell^m(V),\ \forall V \in X_H, \tag{5.8}$$

and then $(\psi_H)_j^m \in X_H, j = 1, \ldots, J, m = 1, \ldots, M$,

$$a(v, (\psi_H)_j^m) = -a_j(v, \Psi_H^m),\ \forall v \in X_H. \tag{5.9}$$

We must thus solve $M$ systems in (5.8), and $MJ$ systems in (5.9). All $M + MJ$ systems share a common matrix — in fact the transpose of the matrix associated with (5.6), (5.7); only the right–hand side varies.

3. We now compute the hybrid fluxes $(p_H^u, P_H^U)_j \in Z, j = 1, \ldots, J$, and $(p_H^\psi, P_H^\Psi)_j^m \in Z, j = 1, \ldots, J, m = 1, \ldots, M$, from

$$b((v, V), (p_H^u, P_H^U)_j) = (\mathcal{R}_H^u)_j((v, V)),\ \forall (v, V) \in \hat{Y}_H, \tag{5.10}$$

$$b((v, V), (p_H^\psi, P_H^\Psi)_j^m) = (\mathcal{R}_H^\psi)_j^m((v, V)),\ \forall (v, V) \in \hat{Y}_H, \tag{5.11}$$

where

$$(\mathcal{R}_H^u)_j((v, V)) = \langle f, v\rangle + \sigma\langle f_j, V\rangle - \mathcal{E}_j((u_H, (U_H)_j), (v, V)),$$

$$(\mathcal{R}_H^\psi)_j^m((v, V)) = -\frac{1}{\sigma}\ell^m(V) - \mathcal{E}_j((v, V), ((\psi_H)_j^m, \Psi_H^m)),$$

and $\mathcal{E}_j$ is $\mathcal{E}$ of (2.9) with $a'$ replaced by $a_j$. The systems (5.10) and (5.11) correspond to small, local problems — one problem (in fact four problems given the four hybrid fluxes) for each node of the coarse mesh. In the case of $M$ outputs and $J$ inputs we will now require $2(J+MJ)$ hybrid fluxes; in all $2(J+MJ)$ cases the particular matrices associated with each node are identical.

4. We find the reconstructed errors $(\hat{e}^u, \hat{E}^U)_j \in \hat{Y}_h, j = 1,\ldots,J$, and $(\hat{e}^\psi, \hat{E}^\Psi)_j^m \in \hat{Y}_h, j = 1,\ldots,J, m = 1,\ldots,M$, such that

$$2\mathcal{E}^s_{Y\,j}((\hat{e}^u, \hat{E}^U)_j, (v,V)) = (\mathcal{R}^u_H)_j((v,V)) - b((v,V),(p^u_H, P^U_H)_j), \ \forall (v,V) \in \hat{Y}_h, \tag{5.12}$$

$$2\mathcal{E}^s_{Y\,j}((\hat{e}^\psi, \hat{E}^\Psi)^m_j, (v,V)) = (\mathcal{R}^\psi_H)^m_j((v,V)) - b((v,V),(p^\psi_H, P^\Psi_H)^m_j), \ \forall (v,V) \in \hat{Y}_h, \tag{5.13}$$

where $\mathcal{E}^s_{Y\,j}$ corresponds to $\mathcal{E}^s_Y$ of (4.7) or (4.9) with $a'$ replaced by $a_j$. We must solve $J$ systems in (5.12) and $MJ$ systems in (5.13). The former correspond to $J$ different matrices; the latter to the same set of $J$ different matrices, each with $M$ different right–hand sides.

5. Finally, we compute the lower $((s_-)^m_j)$ and upper $((s_+)^m_j)$ estimators as $(s_\pm)^m_j = (\overline{s})^m_j + (\Delta)^m_j, j = 1,\ldots,J, m = 1,\ldots,M$, where

$$(\overline{s})^m_j = (s_H)^m_j - 2\mathcal{E}^s_{Y\,j}((\hat{e}^u, \hat{E}^U)_j, (\hat{e}^\psi, \hat{E}^\Psi)^m_j), \tag{5.14}$$

$$(\Delta)^m_j = 2\sqrt{\mathcal{E}^s_{Y\,j}((\hat{e}^u, \hat{E}^U)_j, (\hat{e}^u, \hat{E}^U)_j)\, \mathcal{E}^s_{Y\,j}((\hat{e}^\psi, \hat{E}^\Psi)^m_j, (\hat{e}^\psi, \hat{E}^\Psi)^m_j)}. \tag{5.15}$$

We now require $MJ$ inner products; in any event, this calculation does not contribute significantly to the computational cost.

We close with a brief summary. In the case of direct solution methods, in particular $LU$ (e.g, skyline or banded) solution procedures, the penalty for additional outputs and inputs is relatively small: on the coarse mesh, no additional $LU$ decompositions are required, only additional (much less expensive) forward/back solves; on the broken fine mesh, the penalty is somewhat more significant, in particular as regards multiple inputs — we must now perform $J$ $LU$ decompositions. (The stronger dependence on number of inputs is perhaps to be expected given that our formulation is based on the sensitivity function.) In the case of iterative solution methods it is more difficult to amortize additional right–hand sides, and thus the situation is less encouraging. Finally, we remark that even with the adjoint formulation the bounds require additional solves for multiple inputs; in essence, we obtain a bound on each individual sensitivity derivative, and thus the usual economies of scale do not apply.

**6. Numerical Results.** We now present our numerical results. In all the examples we shall set $\alpha = 0$, $\rho = 10$, and $\nu = 1$; for $f = 0$ the solution is a boundary layer of thickness $O(\nu/\rho = 0.1)$ near the right–hand boundary, $x = 1$. In the first test case we look at variations in $\nu$; in the second test case we look at variations in $\rho$; and in the third test case we look at variations in $f$. (In the interest of brevity we do not present our results for variations in $\alpha$; no surprises are encountered.) We shall consider only Alternative A for the $\mathcal{E}^s_Y - \mathcal{E}^s_M$ splitting, as rather extensive tests indicate that the bounds generated by Alternative A and Alternative B are effectively indistinguishable. In all cases we take $h = .001$ for the fine mesh, and $h \le H \le H_{\max}$ for the coarse mesh, where $H_{\max} = .025$. We shall present our results in the form of $s'_H/s'_h$,

$\overline{s}'/s'_h$, $s'_-/s'_h$, and $s'_+/s'_h$ as a function of $H$. These ratios are represented on the plots as follows.

$$\text{Legend:}\quad s^H_{LB}/s^h = \diamond,\quad s^H_{UB}/s^h = \square,\quad s^H/s^h = \times,\quad s^H_M/s^h = \Delta.$$

In each of the three test cases we shall consider three different output functionals $\ell(v)$: the mean over the domain,

$$\ell^1(v) = \int_0^1 v\,dx\,, \tag{6.1}$$

which we shall denote the "mean" output; the value at a particular point in the domain (within the boundary layer), $x = 0.95$,

$$\ell^2(v) = v(.95)\,, \tag{6.2}$$

which we shall call the "point evaluation" output; and the flux at $x = 1$,

$$\ell^3(v) = a(v, x) - \langle f, x\rangle, \tag{6.3}$$

which we shall denote the "flux" output. As regards the flux output, it is readily shown by integration by parts that, for $u$ sufficiently smooth, $\ell^3(u) = \nu u_x(1)$; the advantage of $\ell^3$ of (6.3) over the more obvious representation $\ell(v) = \nu v_x(1)$ is that the former is a bounded functional while the latter is not. Boundedness (or equivalently, continuity) of the output functional (i) ensures that the adjoint problem is well–posed, and (ii) greatly improves the convergence rate of the bounds. Note that the other two outputs are also bounded: the mean output is bounded in all space dimensions; the point evaluation output is bounded only in one space dimension.

Proceeding with the numerical tests, our first test case is $\alpha = 0$, $\rho = 10$, $\nu = \beta$, $f = 0$, and $\beta = 1$; we take $\sigma = 1$, which ensures coercivity. We present in Figures 6.1, 6.2, and 6.3 on page 15 the bound results for the mean, point evaluation, and flux outputs, respectively. As expected, we obtain bounds; the bounds converge as $O(H^2)$; and the bounds are reasonably accurate even on the coarsest meshes. For the mean and point evaluation outputs $s'_H$ and $\overline{s}'$ are considerably more accurate than the bounds $s'_-$, $s'_+$; however for the point evaluation output, the effectiveness of the estimator is, in fact, quite good. In any event, we emphasize that the bounds provide *certainty* that can not be extracted from either $s'_H$ or $\overline{s}'$.

In our second test case we choose $\alpha = 0$, $\rho = \beta$, $\nu = 1$, $f = 0$, and $\beta = 10$; we again take $\sigma = 1$. We present in Figures 6.4, 6.5, and 6.6 on page 16 the bound results for the mean, point evaluation, and flux outputs, respectively. The results are very similar to those for our first test case — perhaps not surprising since this second test case is, in effect, a rescaling of the first test case. Note, however, that the numerical treatment of the two test cases is quite different: in the first case bounds are guaranteed for all $H$; in the second case bounds are assured only for $H$ sufficiently small. The numerical results of Figures 6.4, 6.5, and 6.6 demonstrate that, in fact, bounds are obtained even on the coarsest mesh — that is, the "$L^2 - H^1$ assumption" is robust.

In our third test case we choose $\alpha = 0$, $\rho = 10$, $\nu = 1$, $f = \sin(2\pi\beta x)$, and $\beta = 5$. We present in Figures 6.7, 6.8, and 6.9 on page 17 the bounds results for the mean, point evaluation, and flux outputs, respectively, for $\sigma = 1$. We present in Figures 6.10, 6.11, and 6.12 on page 18 the bound results for the mean, point evaluation, and flux outputs, respectively, but now for $\sigma$ very large ($\sigma = 100$) — effectively infinite (note there is no coercivity limit on $\sigma$ since $\nu' = 0$ for this test case). Comparison of Figures 6.7, 6.8, and 6.9 with Figures 6.10, 6.11, and 6.12 clearly demonstrates the important role that scaling–parameter optimization can play in achieving sharp bounds. In this particular example $\sigma \to \infty$ is easily identified as optimal: in the limit $\sigma \to \infty$ our general formulation reduces to the "direct" procedure to which we alluded in Section 2. The development of effective scaling–parameter optimization procedures for more general situations is a topic for future work.

In closing, we simply remark that sensitivity derivatives arise with increasing frequency in many different design and optimization contexts. Both certainty and efficiency are currently compromised by the lack of quantitative, relevant error estimation procedures; the techniques presented and illustrated in this paper represent an important first step in addressing this problem.

**Acknowledgments.** The authors would like to thank Dr. Luc Machiels of M.I.T. and Dr. Marius Parischivoiu of University of Toronto for helpful discussions concerning the formulation of the sensitivity derivative bounds and the implementation of the numerical tests.

REFERENCES

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] M. Ainsworth and J.T. Oden, *A unified approach to* a posteriori *error estimation using element residual methods*, Numer. Math., **65** (1993), pp. 23-50.

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

[4] R.E. Bank and A. Weiser, *Some* a posteriori *error estimators for elliptic partial differential equations*, Math. Comp., **44**:170 (1985), pp. 283–301.

[5] R. Becker and R. Rannacher, *Weighted* a posteriori *error control in finite element methods*, IWR Preprint 96–1 (SFB 359), Heidelberg, 1996.

[6] R. Becker and R. Rannacher, *A feedback approach to error control in finite element methods: basic analysis and examples,* IWR Preprint 96–52 (SFB 359), Heidelberg, 1996.

[7] J. Borggaard and D. Pelletier, *Computing design sensitivities using an adaptive finite element method*, AIAA Paper 96–1938, 27th AIAA Fluid Dynamics Conference, New Orleans, 1996.

[8] J. Burns, private communication.

[9] R.G. Carter, *On the global convergece of trust region algorithms using inexact gradient information*, SIAM Journal on Numerical Analysis, **28** (1991), pp. 251–265.

[10] R.G. Carter, *Numerical experience with a class of algorithms for nonlinear optimization using inexact function and gradient information*, SIAM Journal on Scientific Computing, **14** (1993), pp. 368–388.

[11] J.E. Dennis, Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[12] P.E. Gill, W. Murray, and M.H. Wright, *Practical Optimization*, Academic Press, London, 1981.

[13] M.D. Gunzburger, *Introduction to mathematical aspects of flow control and optimization*, in Inverse Design and Optimization, eds. R.A. Van den Braembussche and M. Manna, von Karman Institute for Fluid Dynamics, Rhode Saint Genèses, Belgium, 1997.

[14] A. Jameson, *Aerodynamic design via control theory*, Journal of Scientific Computing, **3** (1988), pp. 223–260.

[15] P. Ladeveze and D. Leguillon, *Error estimation procedures in the finite element method and applications*, SIAM J. Numer. Anal., **20** (1983), pp. 485–509.

[16] L. Machiels, A.T. Patera, J. Peraire, and Y. Maday, *A general framework for finite element* a posteriori

*error control: application to linear and nonlinear convection–dominated problems*, in Proc. ICFD Conference on Numerical Methods for Fluid Dynamics, Oxford, 1998.

[17] Y. Maday, A.T. Patera, and J. Peraire, *A general formulation for* a posteriori *bounds for output functionals of partial differential equations; application to the eigenvalue problem*, C.R. Acad. Sci. Paris A, to appear.

[18] Y. Maday and A.T. Patera, *Numerical analysis of* a posteriori *finite element bounds for linear–functional* outputs, Mathematical Models and Methods in Applied Science, submitted.

[19] J.J. Moré, *Recent developments in algorithms and software for trust region methods*, in Mathematical Programming: The State of the Art, Bonn, 1982, eds. A. Bachem, M. Grötschel, and B. Korte, Springer–Verlag, Berlin, 1983, pp. 258–287.

[20] J. Nocedal, *Theory of algorithms for unconstrained optimization*, Acta Numerica, **1** (1992), pp. 199–242.

[21] M. Paraschivoiu and A.T. Patera, *A hierarchical duality approach to bounds for the outputs of partial differential equations,* Comp. Meth. Appl. Mech. Engrg., to appear.

[22] M. Paraschivoiu and A.T. Patera, A posteriori *bounds for linear–functional outputs of Crouzeix–Raviart finite element discretizations of the incompressible Stokes problem*, Int. J. Num. Methods Fluids, submitted.

[23] M. Paraschivoiu, J. Peraire, and A.T. Patera, A posteriori *finite element bounds for linear–functional outputs of elliptic partial differential equations*, Comp. Meth. Appl. Mech. Engrg., **150** (1997), pp. 289–312.

[24] M. Paraschivoiu, J. Peraire, Y. Maday, and A.T. Patera, *Fast bounds for outputs of partial differential equations*, in Proceedings Workshop on Optimal Design and Control, Washington, D.C., eds. J.A. Burns, E.M. Cliff, and B. Grossman, to appear.

[25] J. Peraire and A.T. Patera, *Bounds for linear–functional outputs of coercive partial differential equations: local indicators and adaptive refinement*, in Proceedings of the Workshop on New Advances in Adaptive Computational Methods in Mechanics, Cachan, France, eds. P. Ladeveze and J.T. Oden, Elsevier, 1997.

[26] J. Peraire and A.T. Patera, *Asymptotic* a posteriori *finite element bounds for the outputs of noncoercive problems: the Helmholtz and Burgers equations*, Comp. Method. in Appl. Mech. and Engrg., submitted.

[27] R. Rannacher and F.-T. Suttmeier, *A feedback approach to error control in finite element methods: application to linear elasticity*, Preprint 96–42 (SFB 359), Heidelberg, 1996.

[28] G.R. Shubin and P.D. Frank, *A comparison of the implicit gradient approach and the variational approach to aerodynamic design optimization*, Tech. Rep. AMS-TR-163, Boeing Computer Services, Applied Mathematics and Statistics Division, 1991.

[29] R. Verfürth, *A Review of* a Posteriori *Error Estimates and Adaptive Mesh–Refinement Techniques*, Wiley–Teubner, 1997.

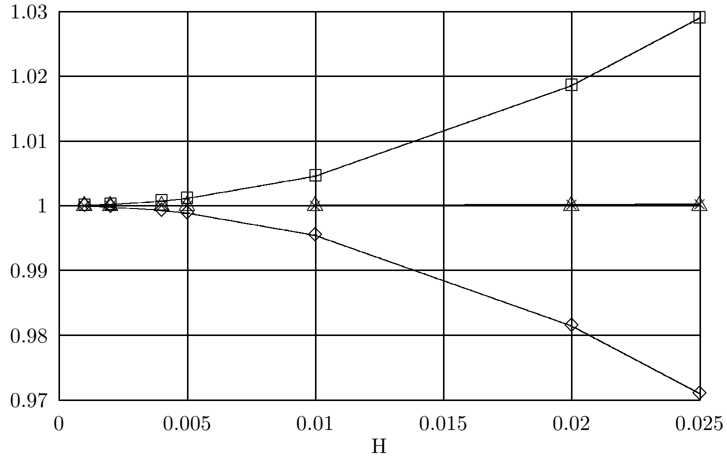


Fig. 6.1. *Results for the mean functional and $\nu = \beta$ at $\beta = 1$.*

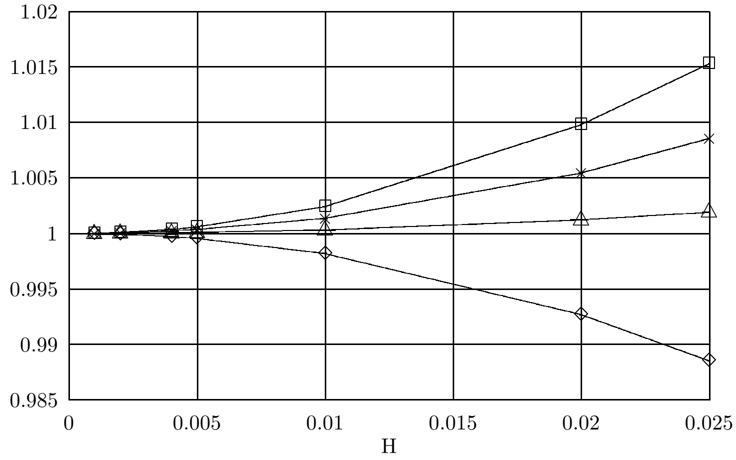


Fig. 6.2. *Results for the point evaluation functional and $\nu = \beta$ at $\beta = 1$.*

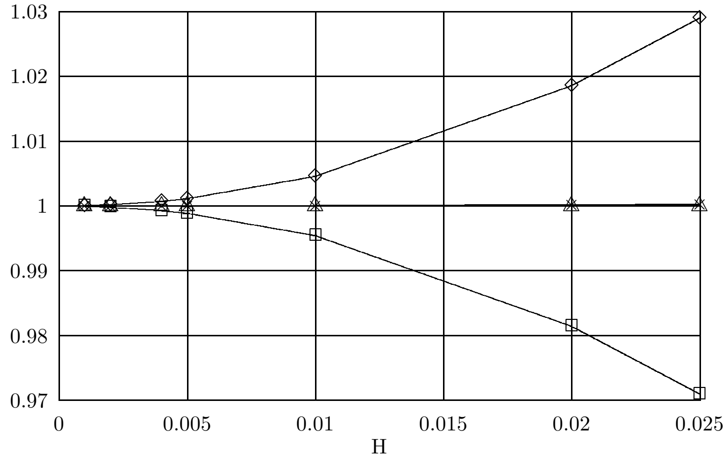


Fig. 6.3. *Results for the flux functional and $\nu = \beta$ at $\beta = 1$.*

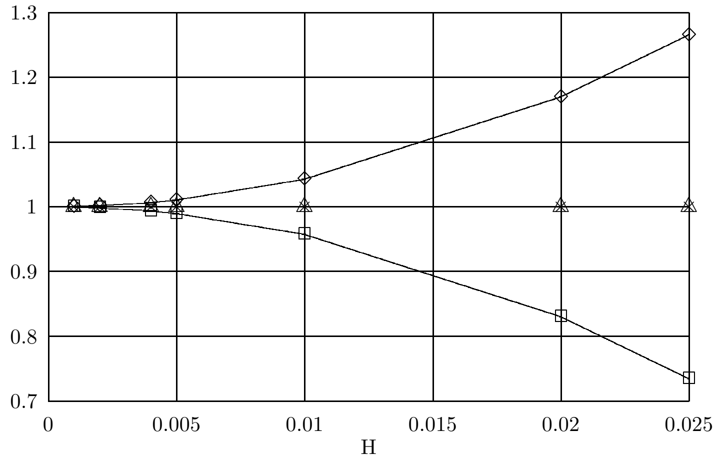


Fig. 6.4. *Results for the mean functional and $\rho = \beta$ at $\beta = 1$.*

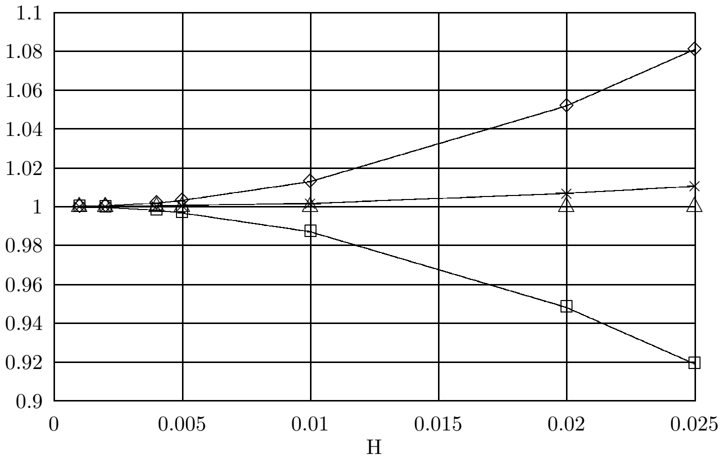


Fig. 6.5. *Results for the point evaluation functional and $\rho = \beta$ at $\beta = 1$.*

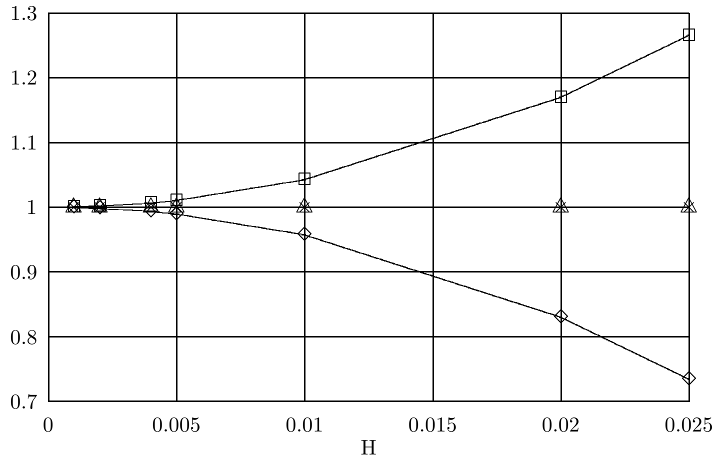


Fig. 6.6. *Results for the flux functional and $\rho = \beta$ at $\beta = 1$.*

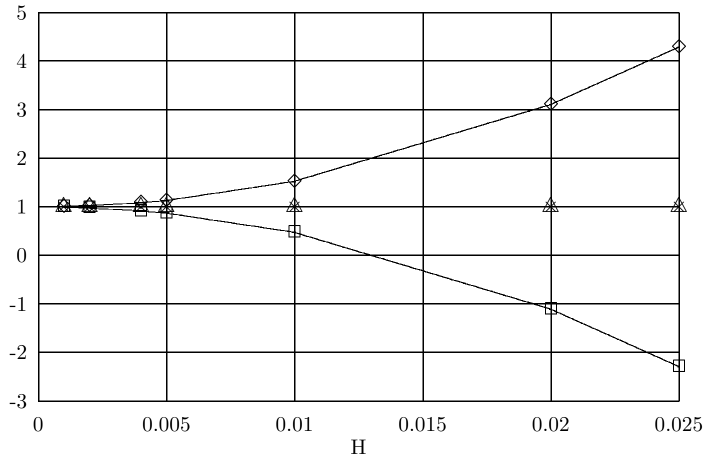


FIG. 6.7. *Results for the mean functional and* $f(x) = \sin(2\pi\beta x)$ *at* $\beta = 5$ *with* $\sigma = 1$.

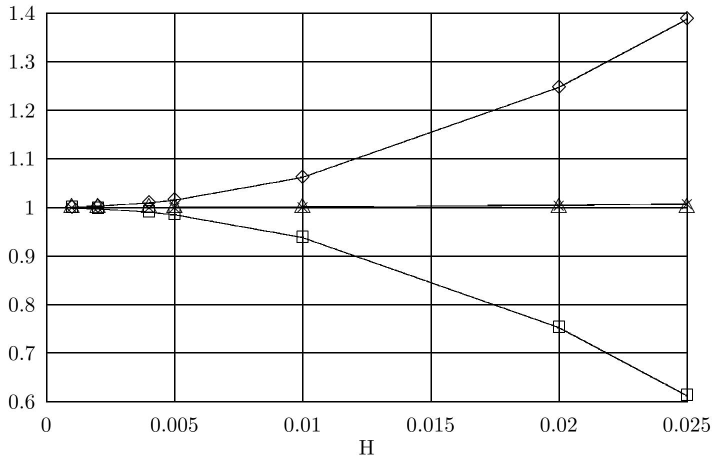


FIG. 6.8. *Results for the point evaluation functional and* $f(x) = \sin(2\pi\beta x)$ *at* $\beta = 5$ *with* $\sigma = 1$.

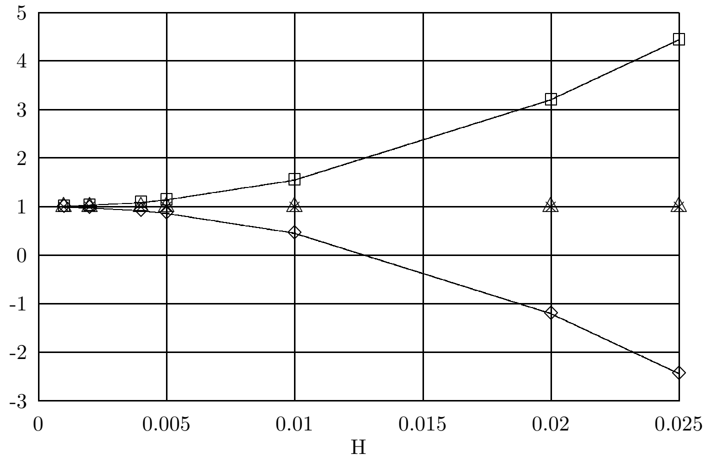


FIG. 6.9. *Results for the flux functional and* $f(x) = \sin(2\pi\beta x)$ *at* $\beta = 5$ *with* $\sigma = 1$.

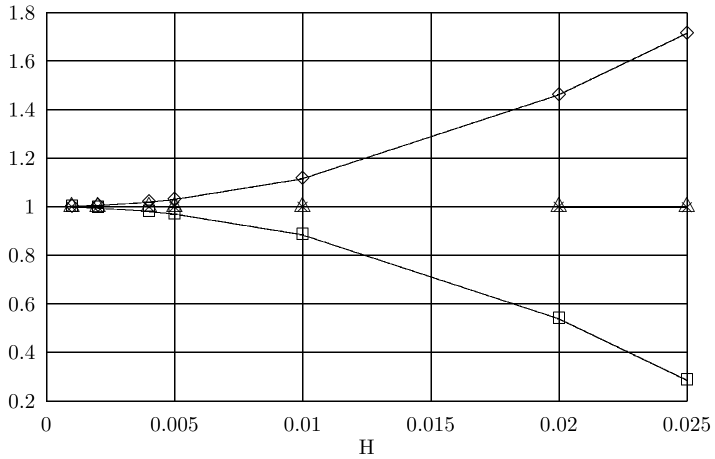


FIG. 6.10. *Results for the mean functional and* $f(x) = \sin(2\pi\beta x)$ *at* $\beta = 5$ *with* $\sigma = 100$.

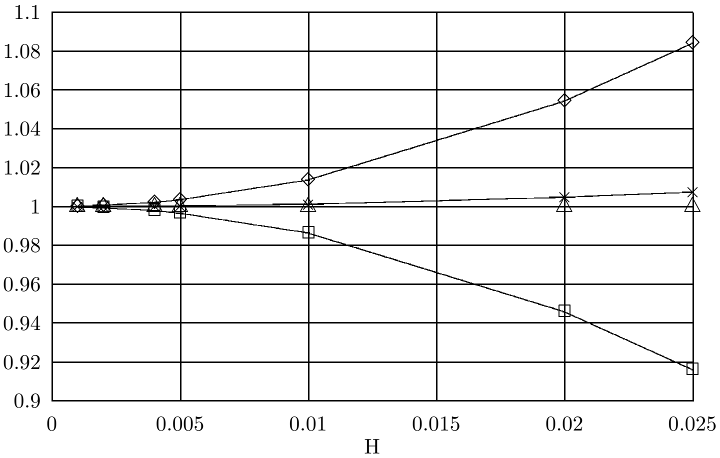


FIG. 6.11. *Results for the point evaluation functional and* $f(x) = \sin(2\pi\beta x)$ *at* $\beta = 5$ *with* $\sigma = 100$.

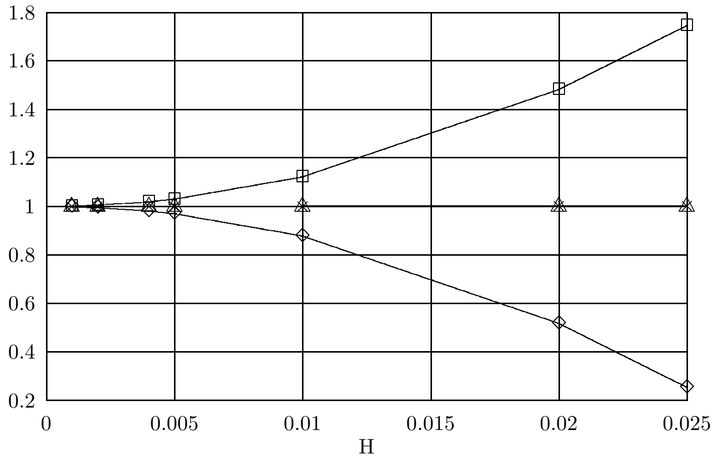


FIG. 6.12. *Results for the flux functional and* $f(x) = \sin(2\pi\beta x)$ *at* $\beta = 5$ *with* $\sigma = 100$.