A Posteriori Finite Element Bounds for Linear–Functional Outputs of Elliptic Partial Differential Equations*

Marius Paraschivoiu† Jaime Peraire‡
Anthony T. Patera †

February 26, 1997

Abstract

We present a domain decomposition finite element technique for efficiently generating lower and upper bounds to outputs which are linear functionals of the solutions to symmetric or nonsymmetric second–order elliptic linear partial differential equations in two space dimensions. The method is based upon the construction of an augmented Lagrangian, in which the objective is a quadratic "energy" reformulation of the desired output, and the constraints are the finite element equilibrium equations and intersubdomain continuity requirements. The bounds on the output for a suitably fine "truth–mesh" discretization are then derived by appealing to a dual max min relaxation evaluated for optimally chosen adjoint and hybrid–flux candidate Lagrange multipliers generated by a K–element coarser "working–mesh" approximation. Independent of the form of the original partial differential equation, the computation on the truth mesh is reduced to K decoupled *subdomain–local, symmetric* Neumann problems. The technique

*Presented at the "Symposium on Advances in Computational Mechanics," a conference in honor of Professor J.T. Oden's 60th birthday, Austin, TX, January 13–15, 1997; submitted to Comp. Meth. Appl. Mech. Engrg.; Revision 1. Address all correspondence to Professor Anthony T. Patera, Room 3-266, M.I.T., 77 Massachusetts Avenue, Cambridge, MA, 02139, USA, `patera@mit.edu`.

†Department of Mechanical Engineering, M.I.T.

‡Department of Aeronautics and Astronautics, M.I.T.

is illustrated for the convection–diffusion and linear elasticity equations.

1 Introduction

Engineering design requires the prediction of system characteristics, or outputs, s — such as deformation, stress, drag and lift forces, flowrate, and heat transfer — that are functionals of field variables $u(x)$ representing displacement, velocity, or temperature; we shall assume here that the outputs are *linear* functionals of these field variables, $s = \ell(u)$. The field variables $u(x)$ are, in turn, solutions of partial differential equations that typically defy exact analysis. Engineering design efforts are therefore based on numerical (here finite element) approximations to the field variables $u(x)$ and associated outputs $s = \ell(u)$, $u_\delta(x)$ and $s_\delta = \ell(u_\delta)$, respectively, where δ represents the mesh spacing, and hence solution accuracy. The quantity of primary importance is typically *not* the full field solution, $u(x)$, but rather the *outputs* $s = \ell(u)$ which represent the relevant engineering design objectives and constraints.

It is often the case, implicitly or explicitly, that engineering design is based upon a (here two–level) hierarchy of numerical approximations. The first discretization, characterized by an element size $\delta = H$, is a "working" approximation which is relatively inexpensive, but which generates solutions $u_H(x)$ and associated outputs $s_H = \ell(u_H)$ that are deemed sufficiently accurate for the purposes of "preliminary" analysis. The second discretization, characterized by element size $\delta = h(< H)$, is a "truth" mesh which produces a solution $u_h(x)$ and associated output $s_h = \ell(u_h)$ for which $|s_h - s|$ is assumed negligibly small. (The latter is, in effect, a weak form of saturation hypothesis.) The h–discretization serves to verify the predictions of the H–discretization, either prior to design, as in a validated–surrogate framework [25], during design, as in trust–region optimization techniques [4], or after design, as final confirmation of anticipated performance. In any guise, validation or confirmation is a necessary precursor to the acceptance of numerical results as "admissible evidence" in the engineering design process, and must therefore provide *predictive* estimates for the *particular* outputs $s_h = \ell(u_h)$ of interest.

In this paper we present a method for exploiting the H–discretization to *inexpensively* construct precise bounds for s_h,

$$(s_h)_{LB}(H) \leq s_h \leq (s_h)_{UB}(H), \tag{1}$$

thereby greatly reducing the cost of "truth" validation. In fact, the requisite calculations on the truth mesh are sufficiently inexpensive that h may be chosen rather conservatively, thus minimizing the effect of our saturation hypothesis. From the bounds (1) we can also derive a predictor for s_h,

$$(s_h)_{pre}(H) = \frac{1}{2}((s_h)_{LB}(H) + (s_h)_{UB}(H)), \tag{2}$$

which will satisfy

$$|(s_h)_{pre}(H) - s_h| \leq \Delta(H), \tag{3}$$

where

$$\Delta(H) = \frac{1}{2}|(s_h)_{UB}(H) - (s_h)_{LB}(H)|. \tag{4}$$

Note that, in some cases, the H–discretization may serve only as an intermediary to "bootstrap" the h–discretization bounds.

Our approach to the construction of the bounds (1) is closely related to Neumann subdomain *a posteriori* error estimation procedures for finite element discretizations of partial differential equations [3]. However, in contrast to earlier subdomain proposals, which measure the error in the energy, $\mathcal{H}^1$, or L^p norms [6, 15, 8, 2, 16], our technique predicts the error directly in the linear–functional outputs of engineering interest. Our methods rely on quadratic–linear duality theory, as in [2], and we develop our local Neumann problems on the truth mesh through domain decomposition and "hybridization" of the intersubdomain constraints [11], as in [15, 8, 2]; however, our particular augmented Lagrangian is "new" as regards both objectives and constraints, and the resulting estimation procedure is efficient, sharp, and general.

We emphasize that our focus is not on the error $s_h - s_H = \ell(u_h) - \ell(u_H) = \ell(u_h - u_H)$, but rather on bounds for $s_h = \ell(u_h)$; and not on adaptive mesh refinement, but rather on output validation and confirmation. It follows that the precise and quantitative calculation of all constants is imperative, even at the expense of some additional computational effort. Our approach is thus complementary to the "Aubin–Nitsche" [13] linear–functional residual estimators recently proposed by Becker and Rannacher [9, 10, 19]. In particular, these residual indicators — based on *a priori* stability and approximation constants and presumed–exact H–mesh adjoint functions — admit rapid evaluation for adaptive refinement, but not necessarily precise or rigorous quantification for final design confirmation. Our subdomain approach, the residual indicators of [9], and certain reconstruction techniques [7] are, to our knowledge, the only *a posteriori* error estimation approaches that directly address linear–functional norms.

In this paper we consider the extension of our technique, introduced in [18], from one space dimension to two space dimensions. The critical new ingredient is a procedure by which to generate appropriate hybrid fluxes in higher dimensions. Our approach to this "equilibration" problem is similar to, but not identical to, earlier proposals [15, 8, 2]; in particular, we must take into account the additional complexity of our Lagrangian in formulating the subdomain Neumann problems. We note that we shall present here the method in a variational context which renders transparent certain theoretical issues; however, the technique also admits a more general, purely algebraic perspective which, in turn, permits direct application to a broad variety of (non–variational) nodal discretization schemes. The reader interested in this alternative algebraic treatment is referred to our earlier paper [18].

The outline of the remainder of the paper is as follows. In Section 2 we introduce our model problems, the two–dimensional convection–diffusion and elasticity equations. In Section 3 we develop the finite element notions which we will need in the remainder of the paper. In Section 4 we describe the sequence of H–mesh and h–mesh problems that constitute the bound procedure. In Section 5 we prove that the quantities computed in Section 4 do, indeed, provide lower and upper bounds for s_h. Finally, in Section 6 we present and discuss our numerical results.

2 Model Problems

We shall formulate three model problems: the Poisson problem, the convection–diffusion problem, and the linear elasticity problem; we shall present numerical examples only for the latter two. These three problems provide examples of scalar equations (Poisson and convection–diffusion), of "vector" equations (elasticity), of symmetric problems (Poisson and elasticity), and of nonsymmetric problems (convection–diffusion). Other applications in progress include nonlinear [24] and hyperbolic [23] systems, time–dependent equations [22], and constrained problems (incompressible Stokes [17]).

In what follows, we shall assume that indices, say i, that appear singly take on values $i = 1, \ldots, d$, where $d = 1$ for the scalar case, and $d = 2$ for the vector case; we shall also assume that any repeated indices are summed over $i = 1, 2$, *independent* of d. We shall denote by $\mathcal{W}^m(\mathcal{D})$ the spaces $(\mathcal{H}^m(\mathcal{D}))^d$, where $\mathcal{D}$ is a domain in $\mathbf{R}^2$, and $\mathcal{H}^m(\mathcal{D})$ is the usual Sobolev space of functions with square integrable derivatives of order m [1]. Finally, for any function $v \in \mathcal{W}^m(\mathcal{D})$, we introduce the norm

$$\|v\|_{m,\mathcal{D}} = \Big(\sum_{i=1}^{d} \|v_i\|^2_{\mathcal{H}^m(\mathcal{D})} \Big)^{1/2} . \tag{5}$$

For most of this paper we shall need only the case $m = 1$.

2.1 Governing Equations

Convection–Diffusion. We consider the convection diffusion problem for the scalar ($d = 1$) field variable, or "temperature," $u(x)$,

$$-(\nu u_{,j})_{,j} + U_j u_{,j} = f \quad \text{in}\, \Omega, \tag{6}$$

with boundary conditions

$$u = g_D \quad \text{on } \Gamma_D \ , \tag{7}$$

$$\nu u_{,i} \hat{n}_i = g_N \quad \text{on } \Gamma_N \ , \tag{8}$$

where $x = (x_1, x_2)$, with corresponding unit vectors $\hat{x}_1, \hat{x}_2$; Ω is the domain; $v_{i,j}$ refers to $\partial v_i / \partial x_j$; ν is the positive viscosity; $U = (U_1, U_2)$ is the prescribed continuous, divergence–free velocity; f is the volumetric data; Γ_D and Γ_N are the Dirichlet and Neumann portions of the domain boundary, $\overline{\Gamma}_D \cup \Gamma_N = \partial\Omega$; g_D and g_N are the Dirichlet and Neumann data; and $\hat{n}$ is the unit outward normal on $\partial\Omega$. We require that $U_i \hat{n}_i \geq 0$ on Γ_N, and that Γ_D be non–empty. Note that for $U = 0$ we recover the Poisson problem.

The variational form of (6)–(8) is: Given $f \in \mathcal{W}^{-1}(\Omega), g_D \in \mathcal{H}^{1/2}(\Gamma_D)$, and $g_N \in \mathcal{H}^{-1/2}(\Gamma_N)$, find $u \in \mathcal{W}^1_D(\Omega)$ such that

$$\int_\Omega \nu v_{,j} u_{,j} + v U_j u_{,j} \, dA = \int_\Omega v f dA + \int_{\Gamma_N} v g_N \, ds, \quad \forall v \in \mathcal{W}^1_0(\Omega) \ , \tag{9}$$

where dA is a differential area element, and ds is a differential line element. Here

$$\mathcal{W}^1_D(\Omega) = \{v \in \mathcal{W}^1(\Omega) \,|\, v|_{\Gamma_D} = g_D\} \;, \tag{10}$$

and

$$\mathcal{W}^1_0(\Omega) = \{v \in \mathcal{W}^1(\Omega) \,|\, v|_{\Gamma_D} = 0\} \;. \tag{11}$$

Since $U_i\hat{n}_i \geq 0$ on the Neumann boundary Γ_N, and Γ_D is non–empty, our bilinear form is coercive and continuous, and existence and uniqueness directly follow from the Lax–Milgram theorem.

We indicate here the *particular* problem for which we will obtain numerical results. We consider the square domain $\Omega =]0,2[\times]0,2[$, the four sides of which are denoted $\Gamma^j, j = 1,\ldots,4$, as shown in Figure 1. We take $f = 0$; $\Gamma_D = \Gamma^1 \cup \Gamma^2$, with $g_D|_{\Gamma^1} = x_1/2, g_D|_{\Gamma^2} = 1 - x_2/2$; and $\Gamma_N = \Gamma^3 \cup \Gamma^4$, with $g_N|_{\Gamma^3} = 1 - x_1/2$, $g_N|_{\Gamma^4} = x_2/2$. The velocity is prescribed as $U = (-1, 1)$, which thus produces a boundary layer on the Neumann sides of the domain. □

Elasticity. We consider the linear elasticity generalized planar stress problem for the vector $(d = 2)$ displacement $u(x)$,

$$-\,\sigma_{ij,j}(u) = f_i \quad \text{in}\,\Omega, \tag{12}$$

with boundary conditions

$$u_i\hat{n}_i = g^n_D \quad \text{on } \Gamma^n_D \;, \tag{13}$$

$$\hat{n}_i\sigma_{ij}\hat{n}_j = g^n_N \quad \text{on } \Gamma^n_N \;, \tag{14}$$

$$u_i\hat{t}_i = g^t_D \quad \text{on } \Gamma^t_D \;, \tag{15}$$

$$\hat{t}_i\sigma_{ij}\hat{n}_j = g^t_N \quad \text{on } \Gamma^t_N \;. \tag{16}$$

Here

$$\sigma_{ij}(u) = \hat{\lambda}_2(u_{i,j} + u_{j,i}) + \hat{\lambda}_1 u_{k,k}\delta_{ij} \tag{17}$$

is the stress as a function of the displacement, in which δ_{ij} is the Kronecker–delta symbol, and $\hat{\lambda}_1, \hat{\lambda}_2$ are related to the Lamé coefficients: $\hat{\lambda}_1 = \nu_p E/(1 - \nu_p^2)$, $\hat{\lambda}_2 = E/2(1 + \nu_p)$, where E is Young's modulus, and ν_p Poisson's ratio. Also, f is the imposed force; $\hat{t}$ is the unit tangent vector on $\partial\Omega$ such that $\hat{t} \times \hat{n} = \hat{x}_1 \times \hat{x}_2$; Γ^n_D and Γ^n_N are the Dirichlet and Neumann portions of the domain boundary for the normal displacements and tractions, respectively, with corresponding data g_D^n and g_N^n; and Γ_D^t and Γ_N^t are the Dirichlet and Neumann portions of the domain boundary for the tangential displacements and tractions, respectively, with corresponding data g_D^t and g_N^t. Note that $\partial\Omega = \overline{\Gamma}_D^n \cup \Gamma_N^n = \overline{\Gamma}_D^t \cup \Gamma_N^t$. We assume that the boundary data is nonsingular at corners, and sufficiently restrictive that translation and rigid–body rotation are precluded.

The variational form of (12)–(17) can then be expressed as: Given $f \in \mathcal{W}^{-1}(\Omega), g_D^n \in \mathcal{H}^{1/2}(\Gamma_D^n), g_N^n \in \mathcal{H}^{-1/2}(\Gamma_N^n), g_D^t \in \mathcal{H}^{1/2}(\Gamma_D^t)$, and $g_N^t \in \mathcal{H}^{-1/2}(\Gamma_N^t)$, find $u \in \mathcal{W}_D^1(\Omega)$ such that

$$\int_\Omega v_{i,j}\sigma_{ij}(u)\, dA = \int_\Omega v_i f_i dA + \int_{\Gamma_N^n} v_i g_N^n \hat{n}_i\, ds + \int_{\Gamma_N^t} v_i g_N^t \hat{t}_i, \quad \forall v \in \mathcal{W}_0^1(\Omega)\ , \tag{18}$$

where

$$\mathcal{W}_D^1(\Omega) = \{v \in \mathcal{W}^1(\Omega)\,|\, v_i\hat{n}_i|_{\Gamma_D^n} = g_D^n, v_i\hat{t}_i|_{\Gamma_D^t} = g_D^t\}\ , \tag{19}$$

and

$$\mathcal{W}_0^1(\Omega) = \{v \in \mathcal{W}^1(\Omega)\,|\, v_i\hat{n}_i|_{\Gamma_D^n} = v_i\hat{t}_i|_{\Gamma_D^t} = 0\}\ . \tag{20}$$

Given our assumptions on the boundary data, the problem admits a unique solution.

We indicate here the *particular* problem for which we will obtain numerical results. We consider the domain Ω shown in Figure 2, corresponding to one-fourth of a symmetrically loaded plate with two rectangular holes; the domain boundaries are denoted $\Gamma_j, j = 1,\ldots,6$. We take $f = 0$; $\Gamma_D^n = \Gamma^1 \cup \Gamma^2 \cup \Gamma^5$, with $g_D^n|_{\Gamma_D^n} = 0$; $\Gamma_N^n = \Gamma^3 \cup \Gamma^4 \cup \Gamma^6$, with $g_N^n|_{\Gamma^3} = 1$, and $g_N^n|_{\Gamma^4} = g_N^n|_{\Gamma^6} = 0$; $\Gamma_D^t = \emptyset$; and $\Gamma_N^t = \partial\Omega$, with $g_N^t = 0$. $\square$

2.2 Output Linear Functionals

We assume that our output s may be expressed as a linear functional

$$\ell : \mathcal{W}^1(\Omega) \to \mathbf{R}\ , \tag{21}$$

that is, $s = \ell(u)$. Examples of possible linear functionals include the average of u_i over a region of Ω, the average of u_i over a curve in Ω or a portion of Γ_N, and the integral of the flux or traction over a portion of Γ_D. We recall that, if $\ell \in (\mathcal{W}^1(\Omega))' = \mathcal{W}^{-1}(\Omega)$, then ℓ is a bounded (and hence continuous) functional: $\ell(v) \leq C\|v\|_{1,\Omega}, \forall v \in \mathcal{W}^1(\Omega)$. In general, if the required output *can* be expressed as a continuous functional, the convergence rate of the output, and associated bounds (1), will certainly benefit [18]; however, if u is sufficiently regular, our technique can also successfully treat linear functionals which are *not* bounded [17]. The class of outputs that should be admitted, how these outputs should be expressed, and how less regular outputs should be mollified, all remain open questions [9, 10]. We now indicate the particular linear functionals we shall investigate in this paper.

Convection–Diffusion. Here we shall look at the average temperature in the vicinity of the corner $(x_1, x_2) = (0, 2)$ at which we expect the maximum temperature in the system to occur. In particular, for the symmetric boundary strip Γ^O identified in Figure 1, we define

$$\ell(v) = \frac{1}{|\Gamma^O|} \int_{\Gamma^O} v \, ds \; , \tag{22}$$

where $|\Gamma^O|$ is the length of Γ^O. Note that, for Γ^O of non-zero measure, $\ell(v)$ is a continuous functional. Finally, from (21) we then compute our desired output as $s = \ell(u)$. $\Box$

Elasticity. Here we shall look at two outputs. We consider the scenario in which the boundary segment Γ^5 represents a "bonded" surface; we thus measure the normal force on this surface, denoted s^1, as an indication of possible failure. In addition, we evaluate the average normal displacement over a segment $\Gamma^O \subset \Gamma^3$, denoted s^2; for $\Gamma^O = \Gamma^3$, s^2 is proportional to the compliance.

The force linear functional we write as

$$\ell^1(v) = \int_\Omega \chi_{i,j} \sigma_{ij}(u) \, dA - \int \chi_i f_i \, ds - \int_{\Gamma_N^n} \chi_i g_N^n \hat{n}_i \, ds - \int_{\Gamma_N^t} \chi_i g_N^t \hat{t}_i \, ds \; , \tag{23}$$

where χ is any function in $\mathcal{W}^1(\Omega)$ such that $\chi_i \hat{n}_i = 1$ on Γ^5 and $\chi_i \hat{n}_i = 0$ on Γ^1 and Γ^2; our particular choice for χ will be discussed in Section 6. For u sufficiently regular, $\ell^1(u) = \int_{\Gamma^5} \hat{n}_i \sigma_{ij}(u) \hat{n}_j \, ds$; to demonstrate this, we simply integrate (23) by parts and appeal to the equilibrium statement (18). Note however, that (23) is bounded for all v in $\mathcal{W}^1(\Omega)$, whereas the more obvious alternative linear functional, $\int_{\Gamma^5} \hat{n}_i \sigma_{ij}(v) \hat{n}_j \, ds$, is *not* in $\mathcal{W}^{-1}(\Omega)$.

The average normal displacement linear functional has the same form as (22),

$$\ell^2(v) = \frac{1}{|\Gamma^O|} \int_{\Gamma^O} v_i \hat{n}_i \, ds \ , \tag{24}$$

where $\Gamma^O \subset \Gamma^3$. This linear functional is bounded for Γ^O of nonzero measure. Finally, given our two linear functionals, ℓ^1, ℓ^2, the corresponding outputs are then evaluated as $s^1 = \ell^1(u)$, $s^2 = \ell^2(u)$, respectively. □

3 Finite Element Discretization

We shall introduce here the general finite element ingredients required to construct the bounds.

3.1 Triangulations

We first introduce the geometrically conforming regular H–mesh triangulation $\mathcal{T}_H$ consisting of K triangles T_H such that

$$\overline{\Omega} = \bigcup_{T_H \in \mathcal{T}_H} \overline{T}_H \ . \tag{25}$$

We denote the set of all (open) edges γ of the triangulation as $\mathcal{E}(\mathcal{T}_H)$, and the set of three edges γ_{T_H} associated with each element T_H as $\mathcal{E}(T_H)$. We denote the set of interior edges as $\mathcal{E}_{int}(\mathcal{T}_H)$, and the sets of Dirichlet and Neumann edges — the edges that are part of Dirichlet and Neumann boundary segments — as $\mathcal{E}_D(\mathcal{T}_H)$ and $\mathcal{E}_N(\mathcal{T}_H)$, respectively (for $d = 1$), and $\mathcal{E}^n_D(\mathcal{T}_H)$, $\mathcal{E}^t_D(\mathcal{T}_H)$ and $\mathcal{E}^n_N(\mathcal{T}_H)$, $\mathcal{E}^t_N(\mathcal{T}_H)$, respectively (for $d = 2$). We require that $\mathcal{E}_D(\mathcal{T}_H)$ and $\mathcal{E}_N(\mathcal{T}_H)$, $\mathcal{E}^n_D(\mathcal{T}_H)$ and $\mathcal{E}^n_N(\mathcal{T}_H)$, and $\mathcal{E}^t_D(\mathcal{T}_H)$ and $\mathcal{E}^t_N(\mathcal{T}_H)$, are all pairwise mutually exclusive.

The triangulation and elemental edges are, of course, related. In particular, given an edge γ_{T_H} in $\mathcal{E}(T_H)$, we shall indicate the coincident edge γ in $\mathcal{E}(\mathcal{T}_H)$ as $\gamma = E(\gamma_{T_H})$. We next associate with each edge γ in $\mathcal{E}(\mathcal{T}_H)$ a unique normal $\hat{n}^\gamma$ such that, if γ lies on $\partial\Omega$, $\hat{n}^\gamma$ coincides with the outward normal $\hat{n}$ on $\partial\Omega$. Then, for all T_H in $\mathcal{T}_H$, and all edges γ_{T_H} in $\mathcal{E}(T_H)$, we define

$$\sigma_{T_H}^{\gamma_{T_H}} = \hat{n}_i^{E(\gamma_{T_H})} \hat{n}_i^{\gamma_{T_H}} \ , \tag{26}$$

where $\hat{n}^{\gamma_{T_H}}$ is the outward normal on γ_{T_H} with respect to T_H. In essense, $\sigma_{T_H}^{\gamma_{T_H}}$ is ± 1 on the two "sides" of an edge γ in $\mathcal{E}(\mathcal{T}_H)$.

We next introduce the h–mesh triangulation $\mathcal{T}_h$, consisting of triangles T_h such that

$$\overline{\Omega} = \bigcup_{T_h \in \mathcal{T}_h} \overline{T}_h \ . \tag{27}$$

We shall require that $\mathcal{T}_h$ be a refinement of $\mathcal{T}_H$, in that we can express each T_H in $\mathcal{T}_H$ as

$$\overline{T}_H = \bigcup_{T_h \in \mathcal{R}_{T_H}} \overline{T}_h \ , \tag{28}$$

where $\mathcal{R}_{T_H}$ is thus the set of h–mesh elements which comprise T_H. A uniform R refinement will denote an h–mesh in which $\mathcal{R}_{T_H}$ consists of R^2 triangles T_h similar to T_H.

3.2 Bilinear and Linear Forms

We define here the bilinear and linear forms for the convection–diffusion and elasticity applications. In many cases the latter are simply the former for $d = 2$.

Convection–Diffusion. We first define the "broken" $\mathcal{W}^1(\Omega)$ space as

$$\mathcal{W}^1_*(\Omega) = \{v \in \mathcal{W}^0(\Omega) \,|\, v|_{T_H} \in \mathcal{W}^1(T_H), \forall T_H \in \mathcal{T}_H\} \ , \tag{29}$$

which permits discontinuities across edges of the H–mesh. We then define the bilinear form associated with our operator as

$$a(w, v) = \sum_{T_H \in \mathcal{T}_H} a_{T_H}(w|_{T_H}, v|_{T_H}), \quad \forall w, v \in \mathcal{W}^1_*(\Omega) \ , \tag{30}$$

where for all T_H in $\mathcal{T}_H$

$$a_{T_H}(w, v) = \int_{T_H} \nu w_{,j} v_{,j} + w U_j v_{,j} \, dA, \quad \forall w, v \in \mathcal{W}^1(T_H) \ . \tag{31}$$

We shall also need the symmetric part of our operator,

$$a^S(w, v) = \sum_{T_H \in \mathcal{T}_H} a^S_{T_H}(w|_{T_H}, v|_{T_H}), \quad \forall w, v \in \mathcal{W}^1_*(\Omega) \ , \tag{32}$$

where for all T_H in $\mathcal{T}_H$

$$a^S_{T_H}(w,v) = \int_{T_H} \nu w_{,j} v_{,j}\, dA + \frac{1}{2}\int_{\partial T_H \cap \Gamma_N} w v U_j \hat{n}_j\, ds, \quad \forall w,v \in \mathcal{W}^1(T_H)\ . \quad (33)$$

The boundary term in (33) arises from the convection operator. Note that $a^S_{T_H}(\ ,\ )$ is not the symmetric part of $a_{T_H}(\ ,\ )$.

We next introduce a set of "jump" bilinear and linear forms which will be required in our variational relaxations. In particular, we define the bilinear form

$$b(v,q) = \sum_{T_H \in \mathcal{T}_H} \sum_{\gamma_{T_H} \in \mathcal{E}(T_H)} \sigma^{\gamma_{T_H}}_{T_H} \int_{\gamma_{T_H}} v|_{T_H} q|_{E(\gamma_{T_H})}\ ds,$$
$$\forall v \in \mathcal{W}^1_*(\Omega), \forall q \in \mathcal{H}^{-1/2}(\mathcal{E}(\mathcal{T}_H)), \quad (34)$$

where $v|_{T_H}$ is to be interpreted as the trace of $v|_{T_H}$ on γ_{T_H}. (More strictly, the L^2 inner products in (34) should be replaced by the proper duality pairings.) Note that q is defined only over the edges of the triangulation, and may, of course, be discontinuous; q may also be defined as the flux associated with a function $\mathcal{Q}$ in $\mathcal{H}(\mathrm{div},\Omega)$ [3]. Effectively, (34) computes the moments of the jumps in v over internal edges, and the moments of v over boundary edges. We also require the related linear functional associated with the Dirichlet conditions,

$$\ell^D(q) = \sum_{T_H \in \mathcal{T}_H} \sum_{\gamma_{T_H} \in \mathcal{E}(T_H)} \sigma^{\gamma_{T_H}}_{T_H} \int_{\gamma_{T_H} \cap \Gamma_D} g_D\, q|_{E(\gamma_{T_H})}\, ds, \quad \forall q \in \mathcal{H}^{-1/2}(\mathcal{E}(\mathcal{T}_H))\ , \quad (35)$$

which computes the moments of the imposed boundary conditions on Γ_D.

We need two final linear functionals. The first, associated with the volumetric and Neumann inhomogeneities, is given by

$$\ell^N(v) = \sum_{T_H \in \mathcal{T}_H} \ell^N_{T_H}(v|_{T_H}), \quad \forall v \in \mathcal{W}^1_*(\Omega)\ , \quad (36)$$

where for all T_H in $\mathcal{T}_H$,

$$\ell^N_{T_H}(v) = \int_{T_H} v f\, dA + \int_{\partial T_H \cap \Gamma_N} v g_N\, ds\ , \quad \forall v \in \mathcal{W}^1(T_H)\ . \quad (37)$$

The second, associated with our output functional, is given by

$$\ell^O(v) = \sum_{T_H \in \mathcal{T}_H} \ell^O_{T_H}(v|_{T_H}), \quad \forall v \in \mathcal{W}^1_*(\Omega)\ , \quad (38)$$

such that

$$\ell^O(v) = \ell(v), \quad \forall v \in \mathcal{W}^1(\Omega) \ . \tag{39}$$

Here $\ell(\,)$ is the formal output functional introduced in (21). Note that the construction (38)-(39) permits us to evaluate the output associated with a field variable which is not in $\mathcal{W}^1(\Omega)$. □

Elasticity. The space $\mathcal{W}^1_*(\Omega)$ and bilinear forms $a(\ ,\ )$ and $a^S(\ ,\ )$ are defined analogously to the scalar case, (29)-(33), however now $a_{T_H}(\ ,\ )$ is given by

$$a_{T_H}(w,v) = \int_{T_H} w_{i,j}\sigma_{ij}(v)\,dA, \quad \forall w,v \in \mathcal{W}^1(T_H)\ , \tag{40}$$

and

$$a^S_{T_H}(w,v) = a_{T_H}(w,v), \quad \forall w,v \in \mathcal{W}^1(T_H)\ , \tag{41}$$

since our operator is symmetric.

Turning to the "jump" bilinear and linear forms, in the vector case we now have

$$b(v,q) = \sum_{T_H\in\mathcal{T}_H} \sum_{\gamma_{T_H}\in\mathcal{E}(T_H)} \sigma^{\gamma_{T_H}}_{T_H} \int_{\gamma_{T_H}} v_i|_{T_H} q_i|_{E(\gamma_{T_H})}\ ds,$$
$$\forall v \in \mathcal{W}^1_*(\Omega), \forall q \in (\mathcal{H}^{-1/2}(\mathcal{E}(\mathcal{T}_H)))^d. \tag{42}$$

The related linear functional associated with the Dirichlet conditions is given by

$$\ell^D(q) = \sum_{T_H\in\mathcal{T}_H} \sum_{\gamma_{T_H}\in\mathcal{E}(T_H)} \sigma^{\gamma_{T_H}}_{T_H} \int_{\gamma_{T_H}\cap\Gamma^n_D} g^n_D \hat{n}_i\, q_i|_{E(\gamma_{T_H})}\, ds + \int_{\gamma_{T_H}\cap\Gamma^t_D} g^t_D \hat{t}_i\, q_i|_{E(\gamma_{T_H})}\, ds,$$
$$\forall q \in (\mathcal{H}^{-1/2}(\mathcal{E}(\mathcal{T}_H)))^d\ . \tag{43}$$

These two functionals play the same "continuity" role as in the scalar case.

Finally, of our two linear functionals, $\ell^N(v)$ of (36), and $\ell^O(v)$ of (38), only the definition of $\ell^N(v)$ must be modified. To wit, for the elasticity problem

$$\ell^N_{T_H}(v) = \int_{T_H} vf\,dA + \int_{\partial T_H\cap\Gamma^n_N} v_i g^n_N \hat{n}_i\,ds + \int_{\partial T_H\cap\Gamma^t_N} v_i g^t_N \hat{t}_i\,ds\ , \quad \forall v \in \mathcal{W}^1(T_H)\ . \tag{44}$$

The global functional $\ell^N(v)$ defined for $v \in \mathcal{W}^1_*(\Omega)$ then follows directly from (36). □

3.3 Function Spaces

Our function spaces may be defined generally for both the scalar ($d = 1$) and vector ($d = 2$) cases. We first introduce the standard conforming linear approximation spaces,

$$X_\delta = \{v|_{T_\delta} \in (\mathbf{P}_1(T_\delta))^d, \forall T_\delta \in \mathcal{T}_\delta\} \bigcap \mathcal{W}_0^1(\Omega) \ , \tag{45}$$

and

$$X_\delta^{D\pm} = \{v \pm u_L \,|\, v \in X_\delta\} \ , \tag{46}$$

where $u_L \in \mathcal{W}_D^1(\Omega)$ is any lifting of the Dirichlet boundary data, and $\mathbf{P}_1(T_\delta)$ is the space of linear polynomials over T_δ. We shall need only $\delta = H$ and $\delta = h$, corresponding to our "working" and "truth" discretizations, respectively: from our refinement hypothesis, $X_H^{D\pm} \subset X_h^{D\pm}$ and $X_H \subset X_h$. Note that, for the problems described in Section 2.1, u_L may (and will) be chosen such that $u_L|_{T_H} \in (\mathbf{P}_1(T_H))^d, \forall T_H \in \mathcal{T}_H$.

We next introduce standard *nonconforming* linear spaces [11, 20],

$$Y_{H,k}^D = \{v|_{T_H} \in (\mathbf{P}_1(T_H))^d, \forall T_H \in \mathcal{T}_H \,|\, b(v,q) = \ell^D(q), \forall q \in Q_k\} \ , \tag{47}$$

$$Y_{H,k} = \{v|_{T_H} \in (\mathbf{P}_1(T_H))^d, \forall T_H \in \mathcal{T}_H \,|\, b(v,q) = 0, \forall q \in Q_k\} \ , \tag{48}$$

where

$$\begin{aligned} Q_k = \{q|_\gamma \in (\mathbf{P}_k(\gamma))^d, \forall \gamma \in \mathcal{E}(\mathcal{T}_H) \ , \text{ and} & \\ & q|_{\Gamma_N} = 0 \, (d=1) \ , \text{ or} \\ & q_i \hat{n}_i|_{\Gamma_N^n} = q_i \hat{t}_i|_{\Gamma_N^t} = 0 \, (d=2) \ \} \ . \end{aligned} \tag{49}$$

The nonconforming space requires that the jumps in v over internal edges be orthogonal to polynomials of order k: this implies that members of $Y_{H,0}$ need only be continuous at the midpoints of edges; for members of $Y_{H,1}$, given our linear approximation of v, the conforming approximation is recovered (in the interior). Note that, in addition to weakly imposing the continuity requirements, (47) and (48) also weakly impose the Dirichlet conditions.

Finally, we introduce our subdomain–local spaces: for all T_H in $\mathcal{T}_H$,

$$Z_H(T_H) = (\mathbf{P}_1(T_H))^d \ , \tag{50}$$

and

$$Z_h(T_H) = \{v|_{T_h} \in (\mathbf{P}_1(T_h))^d, \forall T_h \in \mathcal{R}_{T_H}\} \bigcap \mathcal{W}^1(T_H) \ , \tag{51}$$

where we recall that $\mathcal{R}_{T_H}$ is the set of h–mesh elements that constitute T_H. We can then define the associated "product" spaces as

$$V_\delta = \{v \in \mathcal{W}^1_*(\Omega) \,|\, v|_{T_H} \in Z_\delta(T_H), \forall T_H \in \mathcal{T}_H\} \ , \tag{52}$$

for $\delta = H$ and $\delta = h$. In essence, the $Z_\delta(T_H)$ are Neumann spaces over each T_H, for which V_δ is the corresponding global representation.

3.4 The h–Mesh Problem

We define here the quantity for which we wish to obtain a bound. In particular, we look for a $u_h \in X_h^{D+}$ such that

$$a(v, u_h) = \ell^N(v), \quad \forall v \in X_h \ , \tag{53}$$

and then compute the output (or outputs) of interest,

$$s_h = \ell^O(u_h) \ . \tag{54}$$

We consider here only a standard Galerkin approximation. However, for the convection–diffusion problem, stabilization techniques are certainly of interest, in particular for the H–mesh calculations; stabilization techniques are, of course, *required* if we wish to address the hyperbolic limit $\nu \to 0$ [23].

As regards convergence rate, we expect $\|e_h\|_{1,\Omega} \le C h^{\overline{m}} \|u\|_{\overline{m}+1,\Omega}$ for $u \in \mathcal{W}^{m+1}(\Omega)$, where $e_h = u - u_h$, $\overline{m} = \min(m, 1)$, and $C = C(\Omega, \nu, U)$ for the convection-diffusion problem and $C = C(\Omega, \hat{\lambda}^1, \hat{\lambda}^2)$ for the elasticity problem. For bounded linear functionals — such as the average of u_i over a region or line in Ω, or the integral of the flux or traction over a boundary strip bordered by Neumann segments — we know from the Riesz representation theorem that we can identify a $u^* \in \mathcal{W}^{m^*+1}(\Omega) \cap \mathcal{W}^1_0(\Omega), m^* \ge 0$, such that $\ell^O(v) = a(u^*, v), \forall v \in \mathcal{W}^1_0(\Omega)$. By the usual Aubin–Nitsche procedure [9], we then obtain

$$|\ell^O(e)| \le C h^{\overline{m}+\overline{m}^*} \|u\|_{\overline{m}+1,\Omega} \|u^*\|_{\overline{m}^*+1,\Omega} \ , \tag{55}$$

where $\overline{m}^* = \min(m^*, 1)$. For example, for a convex domain, smooth data, and a smooth functional, we expect $|s - s_h| \approx O(h^2)$ as $h \to 0$, assuming that no problems arise due to mixed boundary conditions.

4 Bound Procedure

The bound procedure involves three steps. The first step determines the adjoint; the second step the hybrid flux; and the third step the local h–mesh solution from which we construct the bounds.

4.1 The H–Mesh Adjoint Calculation

Here we look for $u_H \in X_H^{D+}$ and the adjoints $\psi_H^{\pm} \in X_H^{D-}$ such that

$$a(w, u_H) = \ell^N(w), \quad \forall w \in X_H \ , \tag{56}$$

and

$$a(\psi_H^{\pm}, w) = -(2a^S(w, u_H) - \ell^N(w) \pm \ell^O(w)) \, , \quad \forall w \in X_H \ . \tag{57}$$

(Note that the $\pm$ in $X^{D\pm}$ refers to the Dirichlet boundary conditions on the field variable and adjoint; all other $\pm$ refer to the *pair* of solutions required to ultimately generate lower $(+)$ and upper $(-)$ bounds.) In fact, (56)-(57) is a classical, albeit somewhat trivial, saddle problem [11]. For the convection–diffusion equation the flow is "reversed" for the adjoint problem, however the system is still well–posed, as the Neumann condition is transformed into a weak Dirichlet requirement. The numerical solution of (56)-(57) requires the "inversion" of nonsymmetric stiffness matrices $\underline{A}_H$ and $\underline{A}_H^T$, one (first) of the former in (56), and two (subsequently) of the latter in (57) — one each for ψ_H^+ and ψ_H^-; here $(\ )^T$ refers to the algebraic transpose. For direct strategies the work required to solve all three systems (56)-(57) is only slightly larger than the work required to solve the single system (56), as only a single $\underline{A}_H = \underline{L}\,\underline{U}$ decomposition need be performed; for iterative strategies, such economies of scale are more difficult to realize.

For purposes of future comparison we rewrite (57) in two ways, in both of which we reinterpret (57) as an equation for u_H *given* $\psi_H^{\pm}$. In the first reformulation, we look for $u_H \in X_H^{D+}$ such that

$$2a^S(w, u_H) = -F^{\pm}(w; \psi_H^{\pm}) \ , \quad \forall w \in X_H \ . \tag{58}$$

Here, for any function $\mathcal{F}$ in $\mathcal{W}_*^1(\Omega)$,

$$F^{\pm}(v; \mathcal{F}) = \sum_{T_H \in \mathcal{T}_H} F_{T_H}^{\pm}(v|_{T_H}; \mathcal{F}), \quad \forall v \in \mathcal{W}_*^1(\Omega) \ , \tag{59}$$

where for all T_H in $\mathcal{T}_H$,

$$F^{\pm}_{T_H}(v;\mathcal{F}) = -\ell^N_{T_H}(v) \pm \ell^O_{T_H}(v) + a_{T_H}(\mathcal{F}|_{T_H}, v), \quad \forall v \in \mathcal{W}^1(T_H)\ . \tag{60}$$

In some sense, $-F^{\pm}(v;\psi^{\pm}_H)$ represents the sum of the "forces" in (57). In the second reformulation, we look for $u_H \in X^{D+}_H$ such that

$$B^{\pm}(w, u_H) = 0\ , \quad \forall w \in X_H\ . \tag{61}$$

Here, for any function $\mathcal{G}$ in $\mathcal{W}^1_*(\Omega)$,

$$B^{\pm}(v,\mathcal{G}) = \sum_{T_H \in \mathcal{T}_H} B^{\pm}_{T_H}(v|_{T_H}, \mathcal{G}), \quad \forall v \in \mathcal{W}^1_*(\Omega)\ , \tag{62}$$

where for all T_H in $\mathcal{T}_H$,

$$B^{\pm}_{T_H}(v,\mathcal{G}) = 2a^S_{T_H}(v, \mathcal{G}|_{T_H}) + F^{\pm}_{T_H}(v;\psi^{\pm}_H), \quad \forall v \in \mathcal{W}^1(T_H)\,. \tag{63}$$

In effect, (61) is simply a restatement of equilibrium, or balance.

4.2 The H–Mesh Hybrid Flux Calculation

4.2.1 Nonconforming Approach

In this approach, we first look for new field variables $\overline{u}^{\pm}_H \in Y^D_{H,0}$ such that

$$2a^S(w, \overline{u}^{\pm}_H) = -F^{\pm}(w;\psi^{\pm}_H), \quad \forall w \in Y_{H,0}\ , \tag{64}$$

or equivalently,

$$B^{\pm}(w, \overline{u}^{\pm}_H) = 0, \quad \forall w \in Y_{H,0}\ , \tag{65}$$

which represents a standard linear nonconforming discretization [11, 20]. The numerical solution of (64) requires two "inversions" of the *symmetric* nonconforming stiffness matrix $\underline{A}^S_H$, one for $\overline{u}^+_H$ and one for $\overline{u}^-_H$. As discussed above, for direct strategies the work required to solve both systems is only slightly larger than the work required to solve the first. For iterative strategies, inspection of (64) and (58) indicates that $\overline{u}^{\pm}_H$ and u_H are, respectively, nonconforming and conforming approximations of the same quantity; u_H can thus serve as a good initial iterate for $\overline{u}^{\pm}_H$.

To compute the hybrid flux, we look for $p^{\pm} \in Q_0$ such that

$$b(v, p^{\pm}) = B^{\pm}(v, \overline{u}^{\pm}_H), \quad \forall v \in V_H\ ; \tag{66}$$

note that we test on $v \in V_H$ in (66), not on $v \in Y_{H,0}$ as in (64). Assuming that (66) has a solution, we directly obtain the important property that, for all T_H in $\mathcal{T}_H$,

$$\sum_{\gamma_{T_H} \in \mathcal{E}(T_H)} \sigma_{T_H}^{\gamma_{T_H}} \int_{\gamma_{T_H}} \sum_{i=1}^{d} v_i\, p_i^{\pm}|_{E(\gamma_{T_H})}\, ds = B_{T_H}^{\pm}(v, \overline{u}_H^{\pm}), \quad \forall v \in Z_H(T_H)\ . \tag{67}$$

The condition (67) will ensure solvability of the subsequent subdomain–local truth calculations. In particular, (67) applies to *all* those cases $v = v^s$ for which our bilinear form $a_{T_H}^S(\ ,\ )$ is singular, giving

$$\sum_{\gamma_{T_H} \in \mathcal{E}(T_H)} \sigma_{T_H}^{\gamma_{T_H}} \int_{\gamma_{T_H}} \sum_{i=1}^{d} v_i^s\, p_i^{\pm}|_{E(\gamma_{T_H})}\, ds = F_{T_H}^{\pm}(v^s; \psi_H^{\pm})\ , \tag{68}$$

where in the scalar case ($d = 1$) $v^s = 1$, and in the vector case ($d = 2$) $v = e_1, e_2$, and r. Here $e_1 = (1,0)$ and $e_2 = (0,1)$ are the translation modes, and $r = -x_2 e_1 + x_1 e_2$ is the rigid–body mode; it is crucial that these modes are contained in $Z_H(T_H)$. (Note that, due to the presence of the "Robin" boundary term in (33), $a_{T_H}^S(\ ,\ )$ is not singular for the convection–diffusion problem in those elements T_H with an edge on Γ_N for which $U_i \hat{n}_i \neq 0$; in what follows, this exception will be understood, and not stated explicitly in each instance.)

It remains to show that (66) does, indeed, have a solution. In fact, the left nullspace of (66) is rather large, however the system is solvable, and the solution is unique; the latter follows from the fact that the inf–sup condition is satisfied for the associated hybrid discretization [11, 20]. To demonstrate this less abstractly, we simply *construct* the solution following the standard nonconforming postprocessing treatment [20]. Towards this end, we first introduce the "nonconforming" basis functions for $\mathbf{P}_1(T_H)$, $\zeta_{T_H}^{\gamma_{T_H}}$, where for all T_H in $\mathcal{T}_H$,

$$\zeta_{T_H}^{\gamma_{T_H}} \in \mathbf{P}_1(T_H), \quad \forall \gamma_{T_H} \in \mathcal{E}(T_H)\ , \tag{69}$$

and

$$\int_{\gamma'_{T_H}} \zeta_{T_H}^{\gamma_{T_H}}\, ds = \begin{cases} |\gamma_{T_H}| & \gamma'_{T_H} = \gamma_{T_H} \\ 0 & \gamma'_{T_H} \neq \gamma_{T_H} \end{cases}, \quad \forall \gamma_{T_H}, \gamma'_{T_H} \in \mathcal{E}(T_H)\ , \tag{70}$$

where $|\gamma_{T_H}|$ is the length of γ_{T_H}. The function $\zeta_{T_H}^{\gamma_{T_H}}$ is $+1$ at the two endpoints of γ_{T_H}, and -1 at the remaining vertex of T_H.

We first consider the case of an interior edge, $\gamma \in \mathcal{E}_{int}(\mathcal{T}_H)$. Given the orthogonality property (70) and our requirement that $p^\pm \in Q_0$, it is clear that the two equations for $p^\pm|_\gamma$ are

$$p_i^\pm|_\gamma = \sigma_{T_H}^{\gamma_{T_H}} \frac{1}{|\gamma|} B_{T_H}^\pm(\zeta_{T_H}^{\gamma_{T_H}} e_i, \overline{u}_H^\pm) \ , \tag{71}$$

and

$$p_i^\pm|_\gamma = \sigma_{T'_H}^{\gamma_{T'_H}} \frac{1}{|\gamma|} B_{T'_H}^\pm(\zeta_{T'_H}^{\gamma_{T'_H}} e_i, \overline{u}_H^\pm) \ , \tag{72}$$

where $E(\gamma_{T_H}) = E(\gamma_{T'_H}) = \gamma$, and $|\gamma|$ is the length of γ. (Note that we define $e_1 = 1$ for $d = 1$.) To demonstrate that (71) and (72) are consistent, we first introduce the function $\xi^\gamma = \mathcal{P}\zeta_{T_H}^{\gamma_{T_H}} + \mathcal{P}\zeta_{T'_H}^{\gamma_{T'_H}}$, where for any domain $\mathcal{D} \subset \Omega$, and function $\mathcal{F}\colon \mathcal{D} \to \mathbf{R}$, $\mathcal{P}\mathcal{F}\colon \Omega \to \mathbf{R}$ is the extension of $\mathcal{F}$ defined by

$$\mathcal{P}\mathcal{F} = \begin{cases} \mathcal{F} & \text{in } \mathcal{D} \\ 0 & \text{in } \Omega \setminus \mathcal{D} \end{cases} \ . \tag{73}$$

We now subtract (72) from (71) to obtain the consistency condition

$$B^\pm(\xi^\gamma e_i, \overline{u}_H^\pm) = 0 \ . \tag{74}$$

However, from our orthogonality condition (70), $\xi^\gamma e_i$ is a member of $Y_{H,0}$, and thus from the equilibrium requirement, (65), (74) is clearly satisfied. We can thus choose *either* (71) or (72), or

$$p_i^\pm|_\gamma = \frac{1}{2|\gamma|}(\sigma_{T_H}^{\gamma_{T_H}} B_{T_H}^\pm(\zeta_{T_H}^{\gamma_{T_H}} e_i, \overline{u}_H^\pm) + \sigma_{T'_H}^{\gamma_{T'_H}} B_{T'_H}^\pm(\zeta_{T'_H}^{\gamma_{T'_H}} e_i, \overline{u}_H^\pm)) \ , \tag{75}$$

which is an equivalent averaged flux. An algebraic interpretation of (75) is given in [18].

We next consider the case in which γ is an edge on $\partial\Omega$. In this case, (66) yields only a single equation, say (71). However, we see from (49) that, if γ is in $\mathcal{E}_N(\mathcal{T}_H)$ ($d = 1$) or $\mathcal{E}_N^n(\mathcal{T}_H) \cup \mathcal{E}_N^t(\mathcal{T}_H)$ ($d = 2$), there are additional "boundary" conditions that must also be satisfied. It can readily be shown that these two requirements are, in fact, consistent. For example, for $d = 2$ and $\gamma \in \mathcal{E}_N^n(\mathcal{T}_H)$ such that $\gamma = E(\gamma_{T_H})$, $\zeta_{T_H}^{\gamma_{T_H}} \hat{n}_i$ is clearly a member of $Y_{H,0}$, and thus $B^{\pm}_{T_H}(\zeta^{\gamma_{T_H}}_{T_H}\hat{n}_i, \overline{u}^{\pm}_H) = 0$. We may therefore write our edge hybrid fluxes as

$$p^{\pm} = \begin{cases} \sigma^{\gamma_{T_H}}_{T_H} \frac{1}{|\gamma|} B^{\pm}_{T_H}(\zeta^{\gamma_{T_H}}_{T_H}, \overline{u}^{\pm}_H) & \gamma \in \mathcal{E}_D(\mathcal{T}_H) \\ 0 & \gamma \in \mathcal{E}_N(\mathcal{T}_H) \end{cases} , \tag{76}$$

for $d = 1$, and

$$p^{\pm}_i \hat{z}_i = \begin{cases} \sigma^{\gamma_{T_H}}_{T_H} \frac{1}{|\gamma|} B^{\pm}_{T_H}(\zeta^{\gamma_{T_H}}_{T_H}\hat{z}_i, \overline{u}^{\pm}_H) & \gamma \in \mathcal{E}^z_D(\mathcal{T}_H) \\ 0 & \gamma \in \mathcal{E}^z_N(\mathcal{T}_H) \end{cases} , \tag{77}$$

for $d = 2$. Note that in (77) we consider both $z =$ 'n' and $z =$ 't'.

We close this subsection by summarizing the advantages and disadvantages of this approach. The advantages are that, first, we obtain a uniquely defined hybrid flux, and second, the associated procedure is relatively simply implemented. The disadvantages are that, first, we must perform additional global solves on the H-mesh, and second, for the vector case, the coercivity of the nonconforming system (64) is *not* guaranteed [11]. Given the latter concerns, we now present a second, though related, approach which computes the hybrid fluxes directly from the conforming solution of Section 4.1.

4.2.2 Conforming Approach

To compute the hybrid flux, we first define a $\overline{p}^{\pm} \in Q_0$ as

$$\overline{p}^{\pm}_i|_{\gamma} = \frac{1}{2|\gamma|}(\sigma^{\gamma_{T_H}}_{T_H} B^{\pm}_{T_H}(\zeta^{\gamma_{T_H}}_{T_H} e_i, u_H) + \sigma^{\gamma_{T'_H}}_{T'_H} B^{\pm}_{T'_H}(\zeta^{\gamma_{T'_H}}_{T'_H} e_i, u_H)) , \tag{78}$$

for any interior edge $\gamma \in \mathcal{E}_{int}(\mathcal{T}_H)$ such that $\gamma = E(\gamma_{T_H}) = E(\gamma_{T'_H})$. Similarly, for edges γ on the boundary $\partial\Omega$, we write, for $d = 1$,

$$\overline{p}^{\pm} = \begin{cases} \sigma^{\gamma_{T_H}}_{T_H} \frac{1}{|\gamma|} B^{\pm}_{T_H}(\zeta^{\gamma_{T_H}}_{T_H}, u_H) & \gamma \in \mathcal{E}_D(\mathcal{T}_H) \\ 0 & \gamma \in \mathcal{E}_N(\mathcal{T}_H) \end{cases} , \tag{79}$$

and for $d = 2$,

$$\overline{p}^{\pm}_i \hat{z}_i = \begin{cases} \sigma^{\gamma_{T_H}}_{T_H} \frac{1}{|\gamma|} B^{\pm}_{T_H}(\zeta^{\gamma_{T_H}}_{T_H}\hat{z}_i, u_H) & \gamma \in \mathcal{E}^z_D(\mathcal{T}_H) \\ 0 & \gamma \in \mathcal{E}^z_N(\mathcal{T}_H) \end{cases} , \tag{80}$$

with $z =$ 'n' and $z =$ 't'.

We expect that, in analogy to the corresponding nonconforming result (75) and (76)-(77), (78) and (79)-(80) should give a good approximation to the hybrid flux. However, unlike (75) and (76)-(77), (78) and (79)-(80) will not lead to solvability in the subsequent subdomain–local truth calculations: u_H is not equilibrated in the nonconforming sense, and thus, for example, the average in (78) brings into element T_H information that is inconsistent, however slightly, with force balance for the singular modes $v^s = 1$ (for $d = 1$) and $v^s = e_1, e_2, r$ (for $d = 2$). We thus look for $p^\pm \in Q_1$, and in particular write

$$p^\pm = \overline{p}^\pm + \hat{p}^\pm \ , \tag{81}$$

where $\hat{p}^\pm \in Q_1$. We then require that

$$b(v, p^\pm) = B^\pm(v, u_H), \quad \forall v \in V_H \ , \tag{82}$$

which then gives, for all T_H in $\mathcal{T}_H$,

$$\sum_{\gamma_{T_H} \in \mathcal{E}(T_H)} \sigma_{T_H}^{\gamma_{T_H}} \int_{\gamma_{T_H}} \sum_{i=1}^{d} v_i \, p_i^\pm|_{E(\gamma_{T_H})} \, ds = B_{T_H}^\pm(v, u_H), \quad \forall v \in Z_H(T_H) \ . \tag{83}$$

In particular, (83) applies for the singular modes of $a_{T_H}^S(\ ,\ )$, which thus yields

$$\sum_{\gamma_{T_H} \in \mathcal{E}(T_H)} \sigma_{T_H}^{\gamma_{T_H}} \int_{\gamma_{T_H}} \sum_{i=1}^{d} v_i^s p_i^\pm|_{E(\gamma_{T_H})} \, ds = F_{T_H}^\pm(v^s; \psi_H^\pm) \ , \tag{84}$$

for $v^s = 1$ in the scalar case and $v^s = e_1, e_2$, and r in the vector case.

The system (82) is singular, but solvable. However, unlike our nonconforming system (66), (82) will not yield a unique solution. To treat this indeterminacy we assume that $\overline{p}^\pm$ is, in fact, a good approximation, and look for that solution $p^\pm$ for which $\hat{p}^\pm$ is as small as possible. (The notion of solvability "correction" also permits the exploitation of disparate discretizations, such as Proper Orthogonal Decomposition approximations [12].) The resulting procedure is thus very similar to those originally proposed in [15], [8], and [2]. A first difference is that, for our complex Lagrangian, the hybrid "flux" contains more than simply the gradient of u_H — there are boundary terms that arise due to the $a_{T_H}(\psi_H^\pm, v)$ contribution — and thus $\overline{p}^\pm$ must be computed directly from the underlying linear form, as in (78). A second difference is that the flux is computed on all edges — interior, Neumann, and

Dirichlet — in (almost) the same fashion; this is discussed further in Section 4.3.

The system (82) is well understood [15, 8, 2, 3], and we shall thus only briefly summarize the characteristics and computational aspects of the discrete problem. To begin, we identify the function $\varphi^n_{T_H}$ as the "conforming" basis function for $\mathbf{P}_1(T_H)$ associated with (global) node x_n of the triangulation $\mathcal{T}_H$: $\varphi^n_{T_H}(x)$ is unity at $x = x_n$, and vanishes at the other two vertices of T_H. Then, for any edge γ in $\mathcal{E}(\mathcal{T}_H)$, we write for $\hat{p}|_\gamma \equiv \hat{p}^+|_\gamma$ (say), and for $d = 1$,

$$\hat{p}|_\gamma = \alpha^\gamma_n\, \theta^\gamma_n(x) + \alpha^\gamma_m\, \theta^\gamma_m(x)\ , \tag{85}$$

where x_n, x_m are the nodes of $\mathcal{T}_H$ which define γ, α^γ_n and α^γ_m are real coefficients, and $(\theta^\gamma_n, \theta^\gamma_m)$ are linear functions on γ that are *bi–orthogonal* to $(\varphi^n_{T_H}|_\gamma, \varphi^m_{T_H}|_\gamma)$ for any element T_H for which there is an edge $\gamma_{T_H} \in \mathcal{E}(T_H)$ such that $\gamma = E(\gamma_{T_H})$ [15, 8]. More precisely,

$$\int_\gamma \theta^\gamma_m\, \varphi^n_{T_H}|_\gamma\, ds = |\gamma|\delta_{mn}\ , \tag{86}$$

where δ_{mn} is the Kronecker–delta symbol.

By virtue of the bi–orthogonality, the determination of the α^γ_n reduces to $\mathcal{N}$ smaller systems, one for each node of $\mathcal{T}_H$. For an *interior* node x_n with element multiplicity M_n, the local system is of size M_n, given by (for simplicity, for the case $M_n = 3$),

$$\begin{bmatrix} \sigma_{T_H}^{\gamma^1_{T_H}}|\gamma^1| & \sigma_{T_H}^{\gamma^2_{T_H}}|\gamma^2| & 0 \\ 0 & \sigma_{T'_H}^{\gamma^1_{T'_H}}|\gamma^2| & \sigma_{T'_H}^{\gamma^2_{T'_H}}|\gamma^3| \\ \sigma_{T''_H}^{\gamma^2_{T''_H}}|\gamma^1| & 0 & \sigma_{T''_H}^{\gamma^1_{T''_H}}|\gamma^3| \end{bmatrix} \begin{bmatrix} \alpha^{\gamma^1}_n + \frac{1}{2}\overline{p}^+|_{\gamma^1} \\ \alpha^{\gamma^2}_n + \frac{1}{2}\overline{p}^+|_{\gamma^2} \\ \alpha^{\gamma^3}_n + \frac{1}{2}\overline{p}^+|_{\gamma^3} \end{bmatrix} = \begin{bmatrix} B^+_{T_H}(\varphi^n_{T_H}, u_H) \\ B^+_{T'_H}(\varphi^n_{T'_H}, u_H) \\ B^+_{T''_H}(\varphi^n_{T''_H}, u_H) \end{bmatrix}\ , \tag{87}$$

as is readily derived from (82). Here T_H, T'_H, T''_H and $\gamma^1, \gamma^2, \gamma^3$ are, respectively, the $M_n = 3$ elements and $M_n = 3$ edges which share x_n; furthermore, $\gamma^1 = E(\gamma^1_{T_H}) = E(\gamma^2_{T''_H}), \gamma^2 = E(\gamma^1_{T'_H}) = E(\gamma^2_{T_H}), \gamma^3 = E(\gamma^1_{T''_H}) = E(\gamma^2_{T'_H})$. From the definition of $\sigma_{T_H}^{\gamma_{T_H}}$ in (26), the left nullspace of the coefficient matrix in (87) is $(1,1,1)$. However, the function $\mathcal{P}\varphi^n_{T_H} + \mathcal{P}\varphi^n_{T'_H} + \mathcal{P}\varphi^n_{T''_H}$ is clearly in X_H, and thus from equilibrium, (61), the right–hand side of (87) is orthogonal to $(1,1,1)$, and the system is therefore solvable. Finally, since by assumption $\hat{p}^+$ is a correction to $\overline{p}^+$, we take the pseudoinverse in (87) to control the size of the α_n^γ [15, 16].

We now consider nodes x_n on the *boundary*, $\partial\Omega$. The local system will now be of size $M_n \times (M_n+1)$, as the edge multiplicity will be greater than the element multiplicity. For simplicity, we consider the case $M_n = 2$, the local system for which comprises the first two equations of (87); we assume that γ^1 and γ^3 are on $\partial\Omega$, and $\gamma^2 \in \mathcal{E}_{int}(\mathcal{T}_H)$. We first address the situation in which γ^1 and γ^3 are both in $\mathcal{E}_D(\mathcal{T}_H)$. In this case, there are no boundary conditions to be satisfied; furthermore, the left nullspace is now empty, and solvability is, thus, not an issue. The indeterminacy in the α_n^γ is "removed," as before, by application of the pseudoinverse. Next, we address the situation in which $\gamma^1 \in \mathcal{E}_N(\mathcal{T}_H)$ and $\gamma^3 \in \mathcal{E}_D(\mathcal{T}_H)$. Now, to honor the boundary conditions on p^+, we set $\alpha_n^{\gamma^1} = 0$; it is clear that the coefficients $\alpha_n^{\gamma^2}, \alpha_n^{\gamma^3}$ are then uniquely determined. Finally, we address the situation in which both γ^1 and γ^3 are in $\mathcal{E}_N(\mathcal{T}_H)$. We must now set $\alpha_n^{\gamma^1} = \alpha_n^{\gamma^3} = 0$ to honor the boundary conditions. Fortunately, the remaining $M_n \times (M_n - 1)$ system is solvable: equilibrium, (61), applies at a boundary node x_n with two Neumann edges, and thus the (first two members of the) right–hand side of (87) will be orthogonal to $(1, 1)$, the left nullspace of the coefficient matrix. It follows that $\alpha_n^{\gamma^2}$ may be uniquely determined.

In the vector ($d = 2$) case, interior nodes pose no problem, as the two components are effectively independent. Boundary nodes, however, are more troublesome, since the components couple through the boundary conditions

5 Proof of Bounding Properties

To begin, we derive a simple energy equality that will provide the stabilization in our augmented Lagrangian. For the h–mesh solution u_h of (53) we know that

$$a(u_h - G, u_h) = \ell^N(u_h - G) \ , \tag{88}$$

where G is any function in X_h^{D+}. Expanding (94), and observing that

$$a(u_h, u_h) = a^S(u_h, u_h) + c_U \ , \tag{89}$$

we thus obtain

$$a^S(u_h, u_h) + c_U - \ell^N(u_h) - [a(G, u_h) - \ell^N(G)] = 0 \ , \tag{90}$$

as the desired identity.

We now introduce the set of functions $\mathcal{S}$,

$$\mathcal{S} = \{v \in X_h^{D+} \,|\, a(w,v) = \ell^N(w),\ \forall w \in X_h\} \ . \tag{91}$$

It is clear that this set of functions is not particularly interesting, consisting only of the singleton $v = u_h$. Continuing with the trivialities, it thus follows that

$$\pm\, s_h = \min_{v \in \mathcal{S}} (a^S(v,v) + c_U - \ell^N(v) - [a(G,v) - \ell^N(G)] \pm \ell^O(v)) \ , \tag{92}$$

since for $v = u_h$ the objective reduces to $\pm\ell^O(u_h)$. We next define the set of functions $\tilde{\mathcal{S}}_k$,

$$\tilde{\mathcal{S}}_k = \{v \in V_h \,|\, a(w,v) = \ell^N(w),\ \forall w \in X_h;\ b(v,q) = \ell^D(q),\ \forall q \in Q_k\} \ , \tag{93}$$

and the quantity $\tilde{s}_h^\pm \in \mathbf{R}$ as

$$\pm\, \tilde{s}_h^\pm = \min_{v \in \tilde{\mathcal{S}}_k} (a^S(v,v) + c_U - \ell^N(v) - [a(G,v) - \ell^N(G)] \pm \ell^O(v)) \ . \tag{94}$$

Now it is clear that any function (in fact, *the* function) v in $\mathcal{S}$ will also be in $\tilde{\mathcal{S}}_k$, and thus $\mathcal{S} \subset \tilde{\mathcal{S}}_k$ for all k. It then immediately follows that

$$\pm\, \tilde{s}_h^\pm \leq \pm s_h \ , \tag{95}$$

since (100) is a relaxation of (98). Note that (100) relaxes both the internal continuity requirements and the strong imposition of the Dirichlet boundary conditions.

We now form the Lagrangian associated with the constrained minimization problem (100), $\tilde{\mathcal{L}}_G^\pm : V_h \times X_h \times Q_k \to \mathbf{R}$, as

$$\begin{aligned} \tilde{\mathcal{L}}_G^\pm(v,\mu',q) = (a^S(v,v) + c_U - \ell^N(v) - [a(G,v) - \ell^N(G)] \pm \ell^O(v)) \\ +(a(\mu',v) - \ell^N(\mu')) - (b(v,q) - \ell^D(q)) \ . \end{aligned} \tag{96}$$

We next note that we can absorb the "G" terms into the adjoint μ' to define an equivalent but simpler Lagrangian, $\tilde{\mathcal{L}}^\pm : V_h \times X_h^{D-} \times Q_k \to \mathbf{R}$, as

$$\begin{aligned} \tilde{\mathcal{L}}^\pm(v,\mu,q) = (a^S(v,v) + c_U - \ell^N(v) \pm \ell^O(v)) \\ +(a(\mu,v) - \ell^N(\mu)) - (b(v,q) - \ell^D(q)) \ . \end{aligned} \tag{97}$$

(An alternative derivation of this boundary condition is given in [18].) Finally, it shall prove convenient to write (103) as

$$\tilde{\mathcal{L}}^{\pm}(v,\mu,q) = [-a^S(v,v) + c_U - \ell^N(\mu) + \ell^D(q)] \\ +[2a^S(v,v) + F^{\pm}(v;\mu) - b(v,q)] \ , \tag{98}$$

for $F^{\pm}(v;\mu)$ as defined in (59)-(60).

We now show that $\eta^{\pm}$ of (92) can be expressed as

$$\eta^{\pm} = \tilde{\mathcal{L}}^{\pm}(\hat{\mathcal{U}}_h^{\pm}, \psi_H^{\pm}, p^{\pm}) \ , \tag{99}$$

and that

$$\tilde{\mathcal{L}}^{\pm}(\hat{\mathcal{U}}_h^{\pm}, \psi_H^{\pm}, p^{\pm}) = \min_{v \in V_h} \tilde{\mathcal{L}}^{\pm}(v, \psi_H^{\pm}, p^{\pm}) \ . \tag{100}$$

Note that, as required by the definition of our Lagrangian, $\psi_H^{\pm} \in X_h^{D-}$ and $p^{\pm} \in Q_k$; the former is a result of our refinement hypothesis, the latter of our hybrid flux construction procedure. It then follows from classical quadratic–linear duality theory [21, 18] that

$$\eta^{\pm} \leq \pm \tilde{s}_h^{\pm} \ . \tag{101}$$

To conclude, we evoke (101), which proves that $(s_h)_{LB}$ and $(s_h)_{UB}$ of (90), (91) do, indeed satisfy (1), that is, provide lower and upper bounds for the output of interest. (A perhaps simpler proof replaces Q_k with all functions in $(\mathcal{H}^{-1/2}(\mathcal{E}(\mathcal{T}_H)))^d$ that vanish on Neumann boundaries. We can then proceed from (98) directly to (102), and thus from (105),(106) directly to $\eta^{\pm} \leq \pm s_h$. This approach also permits higher–order approximations for the hybrid flux.)

It thus only remains to demonstrate (105) and (106). To this end, we expand our Lagrangian (104) for $v = \hat{\mathcal{U}}_h^{\pm} + w, \mu = \psi_H^{\pm}, q = p^{\pm}$, to obtain

$$\begin{aligned} \tilde{\mathcal{L}}^{\pm}(\hat{\mathcal{U}}_h^{\pm} + w, \psi_H^{\pm}, p^{\pm}) &= \tilde{\mathcal{L}}^{\pm}(\hat{\mathcal{U}}_h^{\pm}, \psi_H^{\pm}, p^{\pm}) \\ &+ 2a^S(w, \hat{\mathcal{U}}_h^{\pm}) + F^{\pm}(w; \psi_H^{\pm}) - b(w, p^{\pm}) \\ &+ a^S(w,w) \ , \quad \forall w \in V_h \ . \end{aligned} \tag{102}$$

However, from (89) we observe that the terms linear in w collectively vanish, and thus (106) directly follows since $a^S(\ ,\ )$ is positive semi-definite over V_h. Finally, to prove (105), we simply set $v = \hat{\mathcal{U}}_h^{\pm}, \mu = \psi_H^{\pm}$, and $q = p^{\pm}$ in $\tilde{\mathcal{L}}^{\pm}(v,\mu,q)$ of (104), and then evoke (89); note that, due to weak imposition of the Dirichlet conditions, all the subdomain problems in (88)-(89) are *natural*, and thus $v = \hat{\mathcal{U}}_h^{\pm}$ is an admissible test function.

We close this section with three remarks. First, we note that solvability is essential in the construction of the bounds [3]: without solvability, in a "singular" element T_H, $F_{T_H}^{\pm}(cv^s; \psi_H^{\pm}) - b(c\mathcal{P}v^s, p^{\pm}) \to -\infty$ for either $c \to \infty$ or $c \to -\infty$; since $a_{T_H}^S(cv^s, cv^s) = 0$,

$$\min_{v \in V_h} \tilde{L}^{\pm}(v, \psi_H^{\pm}, p^{\pm}) = -\infty \ , \tag{103}$$

and the resulting bounds are meaningless. Solvability ensures that, in a singular element T_H, $F_{T_H}^{\pm}(v^s; \psi_H^{\pm}) - b(\mathcal{P}v^s, p^{\pm}) = 0$, and thus the modes not controlled by the quadratic stabilization term can not contribute. The choice of the *level* and rotation of $\hat{\mathcal{U}}_h^{\pm}$ in the singular elements T_H is, of course, irrelevant. Second, we note that we have proven here only that $(s_h)_{LB}$ and $(s_h)_{UB}$ are bounds for s_h; we expect to obtain *good* bounds since $\psi_H^{\pm}$ and $p^{\pm}$ are the saddlepoints of H–mesh approximations to our Lagrangian (103) [18]. Third, we note that we can clearly replace our energy equality of (94) with $\kappa a(u_h - G, u_h) = \kappa \ell^N(u_h - G)$ for any non–negative real κ; we may then choose κ to optimize our bounds, that is, to maximize the $\eta^{\pm}$. The details of this procedure are described in [18] and will not be repeated here; the bounds presented in the next section correspond to $\kappa = 1$, and can thus be further improved.

6 Numerical Results

6.1 Convection–Diffusion Equation

We first present results for the convection diffusion equation, for $\nu = .1$, and for the output segment Γ^O shown in Figure 1. We apply here the *nonconforming* hybrid flux formulation of Section 4.2.1. The triangulations investigated, $\mathcal{T}_{(H_0,R)}$, are uniform R refinements of the coarsest mesh, $\mathcal{T}_{(H_0,1)}$. The H–meshes $\mathcal{T}_H$ correspond to $\mathcal{T}_{(H_0,R)}$, $R = 1, 2, 3, 4, 6$, and the truth h–mesh corresponds to $\mathcal{T}_h = \mathcal{T}_{(H_0,12)}$; $\mathcal{T}_{(H_0,1)}$ and $\mathcal{T}_h$ are shown in Figures 3a and 3b, respectively. Note that, for the refinement values R chosen, $X_H \subset X_h$, as required by the theory. We shall denote the effective working–approximation element size associated with triangulation $\mathcal{T}_H = \mathcal{T}_{(H_0,R)}$ as $H \equiv 1/R$.

In all the meshes $\mathcal{T}_{(H_0,R)}$ elements are concentrated near the "left" and "top" walls in anticipation of the boundary layers that will form on these surfaces. The purpose of the bounds is to ascertain whether this H–mesh grading, and the coarser discretization in the interior, is sufficient to capture the engineering output of interest, s — here the average boundary temperature near the top left corner, (22) — to within some acceptable accuracy. More precisely, we are interested in establishing, through the *intermediary* of the H-mesh discretizations, inexpensive yet sharp lower and upper bounds, $(s_h)_{LB}$ and $(s_h)_{UB}$, for the output on the truth mesh, s_h. We note that our meshes are also graded near the "bottom" and "right" walls, in anticipation of boundary layers in the *adjoint*, $\psi_H^{\pm}(x)$, for which the flow velocity is reversed; this refinement is not required for the original problem, and is thus a "hidden" cost associated with the bound procedure. The quantitative effect of inadequate resolution of the adjoint problem has not yet been fully investigated.

We plot in Figure 4a $(s_h)_{UB}/s_h$, $(s_h)_{LB}/s_h$, $(s_h)_{pre}/s_h$, and s_H/s_h as a function of H. Here s_h is the output on the truth mesh, s_H is the output on the H–mesh, $(s_h)_{LB}$ and $(s_h)_{UB}$ are the bounds (90)-(91), and $(s_h)_{pre}$ is the average of the bounds, as defined in (2). As a predictor of the true output, s_h, we see that the most accurate estimator is s_H, followed by $(s_h)_{pre}$, $(s_h)_{UB}$, and $(s_h)_{LB}$. We contend that the degradation in the accuracy of $(s_h)_{LB}$ and $(s_h)_{UB}$ is more than balanced by the certainty that these quantities *bound* s_h from below and above; no such two–sided, or even one–sided, assurances are associated with either s_H or $(s_h)_{pre}$. Nevertheless, as an error estimator for s_H, the bounds are (at least for this non–optimal κ) less than impressive.

There are, however, instances — in particular for unbounded functionals [17] — in which $(s_h)_{pre}$, and even the bounds $(s_h)_{LB}$ and $(s_h)_{UB}$, provide better predictions than s_H for the true output s_h. Furthermore, in all cases studied to date, we find that the bounds converge to s_h at least as fast as s_H converges to s_h as $H \to h$. In order to demonstrate this in the current case, we plot in Figure 4b $e_{UB} \equiv \log|(s_h)_{UB} - s_h|$, $e_{LB} \equiv \log|(s_h)_{LB} - s_h|$, $e_{pre} \equiv \log|(s_h)_{pre} - s_h|$, and $e_H \equiv \log|s_H - s_h|$ as a function of $\log H$. We observe that $s_H, (s_h)_{LB}$, and $(s_h)_{UB}$ all converge to s_h as $O(H^2)$, as might be expected from the discussion of Section 3.4.

It is clear that, if the bounds are tight, the approximations $\hat{\mathcal{U}}_h^{\pm}$ must be close to u_h. We compare, in Figures 5a and 5b, $\hat{\mathcal{U}}_h^{+}$ associated with the working mesh $\mathcal{T}_H = \mathcal{T}_{(H_0,3)}$, and u_h for the truth mesh $\mathcal{T}_h = \mathcal{T}_{(H_0,12)}$, respectively. (In order to render visual comparison more transparent, we choose the level of $\hat{\mathcal{U}}_h^+$ in all singular subdomains T_H so that $\hat{\mathcal{U}}_h^+$ and u_H agree at one vertex of T_H.) We observe that $\hat{\mathcal{U}}_h^+$ is, indeed, quite close to u_h; furthermore, the relatively small discontinuities in $\hat{\mathcal{U}}_h^+$ across subdomain boundaries is an indication of the accuracy of the hybrid fluxes. Note that, for the working approximation $\mathcal{T}_H = \mathcal{T}_{(H_0,1)}$, $\hat{\mathcal{U}}_h^+$ is *not* close to u_h [17], as for this H–mesh the grid Peclet number is "supercritical"; the resulting oscillations are also responsible for the poor quality of the bounds for $H = 1$ in Figure 4a. Stabilization techniques [23] applied on the H–mesh should improve the bounds for the larger H.

6.2 Elasticity Equation

We now present results for the linear elasticity equation, for $E = 1$ and $\nu_p = .3$. We apply here the *conforming* hybrid flux formulation of Section 4.2.2. The triangulations investigated, $\mathcal{T}_{(H_0,R)}$, are uniform R refinements of the coarsest mesh, $\mathcal{T}_{(H_0,1)}$. The working meshes $\mathcal{T}_H$ correspond to $\mathcal{T}_{(H_0,R)}$, $R = 1, 2, 4, 8$, and the truth mesh to $\mathcal{T}_h = \mathcal{T}_{(H_0,16)}$; $\mathcal{T}_{(H_0,1)}$ and $\mathcal{T}_{(H_0,8)}$ are shown in Figures 6a and 6b, respectively. Note that, for the refinement values R chosen, $X_H \subset X_h$, as required by the theory. As before, we shall denote the effective working–approximation element size associated with triangulation $\mathcal{T}_H = \mathcal{T}_{(H_0,R)}$ as $H \equiv 1/R$.

The function χ required for evaluation of the force linear–functional output (23) is chosen such that $\chi_2 = 0$, and

$$\chi_1 = \begin{cases} 1 - 2x_1 & (x_1, x_2) \in \Omega' \\ 0 & (x_1, x_2) \in \Omega \setminus \Omega' \end{cases} . \tag{104}$$

Here $\Omega' = 0 \le x_1 \le .5, .8 \le x_2 \le 1.0$, corresponding to the region directly above the hole. As discussed in [18], and readily demonstrated by integration by parts in (57), there will be a jump in the gradient of the adjoint at points at which χ is not C^1–continuous — in our example, the right boundary of Ω', $x_1 = .5, .8 \le x_2 \le 1.0$. In practice, the particular choice for χ affects the value of the bounds very little.

We plot in Figures 7a, 7b, and 7c, $(s_h)_{UB}/s_h$, $(s_h)_{LB}/s_h$, $(s_h)_{pre}/s_h$, and s_H/s_h as a function of (effective) H for, respectively, s^1, s^2 for Γ^O as indicated in Figure 2, and s^2 for $\Gamma^O = \Gamma^3$. Recall that s^1 is the normal force on Γ^5, defined by the linear functional (23), while s^2 is the normal average displacement on Γ^O, as defined by the linear functional (24). (Note that the ratio $(s_h)_{UB}/s_h$ ($(s_h)_{LB}/s_h$) for s^1 in Figure 7a approaches unity from below (above) because s^1 is negative.) For the output s^1, $(s_h)_{pre}$ and s_H are the most accurate predictors of s_h, followed by the bounds $(s_h)_{LB}$ and $(s_h)_{UB}$; as always, the bounds enjoy the additional advantage of "certainty." For the output s^2, in particular for $\Gamma^O = \Gamma^3$, s_H, $(s_h)_{LB}$, and $(s_h)_{pre}$ are all of commensurate accuracy, for reasons that will be discussed below. We note that the relative errors in the different outputs are, not unexpectedly, different; our method can efficiently and quantitatively assess these variations — a marked improvement over current practice, in which a single error measure is typically (implicitly) assumed to apply to all quantities of interest.

In Figures 8a, 8b, and 8c, we plot $e_{UB} \equiv \log|(s_h)_{UB} - s_h|$, $e_{LB} \equiv \log|(s_h)_{LB} - s_h|$, $e_{pre} \equiv \log|(s_h)_{pre} - s_h|$, and $e_H \equiv \log|s_H - s_h|$ as a function of $\log H$ for, respectively, s^1, s^2 for Γ^O as indicated in Figure 2, and s^2 for $\Gamma^O = \Gamma^3$. The quantities $s_H, (s_h)_{UB}$, and $(s_h)_{LB}$ appear to converge to s_h as $O(H^2)$ as $H \to h$; $(s_h)_{pre}$ converges perhaps even faster. Although the effect of the corner singularities might be observed on finer meshes, there is no evidence of significant pollution for the discretizations and outputs considered here. As in all *a posteriori* analyses, our bounds reflect the particularities of the problem, not the worse–case behavior. Indeed, our approach is even more "application–sensitive" than earlier proposals [15, 8, 2], since our error estimates are tailored to the specific engineering outputs of interest.

Finally, we note that, for $\Gamma^O = \Gamma^3$, s^2 is simply the compliance. In this particular case, it follows from simple variational arguments that s_H must be a lower bound for s_h, and that $s_h - s_H$ is proportional to $a(u_h - u_H, u_h - u_H)$; for a smooth problem, and the linear approximations of this paper, quadratic convergence is thus expected. It can further be shown that, within our bounds framework, $(s_h)_{LB} = s_H$. Figures 7c and 8c demonstrate these results empirically. In such "variationally (very) special" circumstances — a symmetric problem and energy output — our new bound procedure is less compelling: the only advantage is that an *upper* bound for the output may also be readily obtained (without recourse to complementary energy principles). Unfortunately, or fortunately, most problems and outputs are less accommodating than the compliance.

The engineering relevance of the results presented in this paper should be clear. We obtain rigorous, quantitative, and relatively sharp lower and upper bounds for the specific outputs of interest in a particular engineering design study — at a fraction of the effort that would be required by direct appeal to the "truth." In effect, one obtains truth–mesh certainty at working–mesh cost, permitting a much faster and more reliable design framework. Much work remains, however, not only in extending the approach to different problems and different numerical formulations, but also in considering larger classes of output functionals, and in providing local information for purposes of adaptivity.

Acknowledgements

This work was supported by DARPA and ONR under Grant N00014-91-J-1889, AFOSR under Grant F49620-94-1-0121, and NASA under Grants NAG 1-1613, NAG 4-105, and NAG 1-1587.

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*, TICAM Report 96-19, 1996.

[4] N. Alexandrov, J.E. Dennis, Jr., R. M. Lewis, and V. Torczon, *A trust region framework for managing the use of approximation models in optimization*, in preparation.

[5] I. Babuška, *The finite element method with Lagrangian multipliers*, Numer. Math., **20** (1973), pp. 179–192.

[6] I. Babuška and W.C. Rheinboldt, *A posteriori error estimates for the finite element method*, Int. J. Numer. Methods Engrg., **12** (1978), pp. 1597–1615.

[7] I. Babuška, T. Strouboulis, C.S. Upadhyay, and S.K. Gangaraj, A posteriori *estimation and adaptive control of the pollution error in the h–version of the finite element method*, Int. J. Numer. Methods Engrg., **38** (1995), pp. 4207–4235.

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

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

[10] 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.

[11] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer–Verlag, New York, 1991.

[12] V. Colmar, Diplome, EPFL, Lausanne, in progress (M.I.T.).

[13] K. Eriksson and C. Johnson, *An adaptive finite element method for linear elliptic problems*, Math. Comp., **50** (1988), pp. 361-383.

[14] P.F. Fischer and A. T. Patera, *Parallel simulation of viscous incompressible flows*, Annu. Rev. Fluid Mech., **26** (1994), pp. 483-527.

[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] P. Ladeveze, P. Marin, J.P. Pelle, and J.L. Gastine, *Accuracy and optimal meshes in finite element computation for nearly incompressible materials*, Comp. Meth. Appl. Mech. Engrg., **94**:3 (1992), pp. 303–315.

[17] M. Paraschivoiu, Ph.D. Thesis, Department of Mechanical Engineering, M.I.T., in progress.

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

[19] 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.

[20] J.E. Roberts and J-M. Thomas, *Mixed and Hybrid Methods*, in Handbook of Numerical Analysis, Volume II: Finite Element Methods (Part 1), eds. P.G. Ciarlet and J.L. Lions, North–Holland, Amsterdam, 1991.

[21] G. Strang, *Introduction to Applied Mathematics*, Wellesley–Cambridge Press, Wellesley, Massachusetts, 1986.

[22] J. Teichman, S.M. Thesis, Department of Mechanical Engineering, M.I.T., in progress.

[23] H. Vailong, A Posteriori *Bounds for Linear–Functional Outputs of Hyperbolic Partial Differential Equations*, S.M. Thesis, Department of Aeronautics and Astronautics, M.I.T., 1997.

[24] H. Vailong, J. Peraire, M. Paraschivoiu, and A.T. Patera, in preparation.

[25] S. Yeşilyurt and A. T. Patera, *Surrogates for numerical simulations; optimization of eddy–promoter heat exchangers*, Comp. Methods Appl. Mech. Engrg., **121** (1995), pp. 231-257.

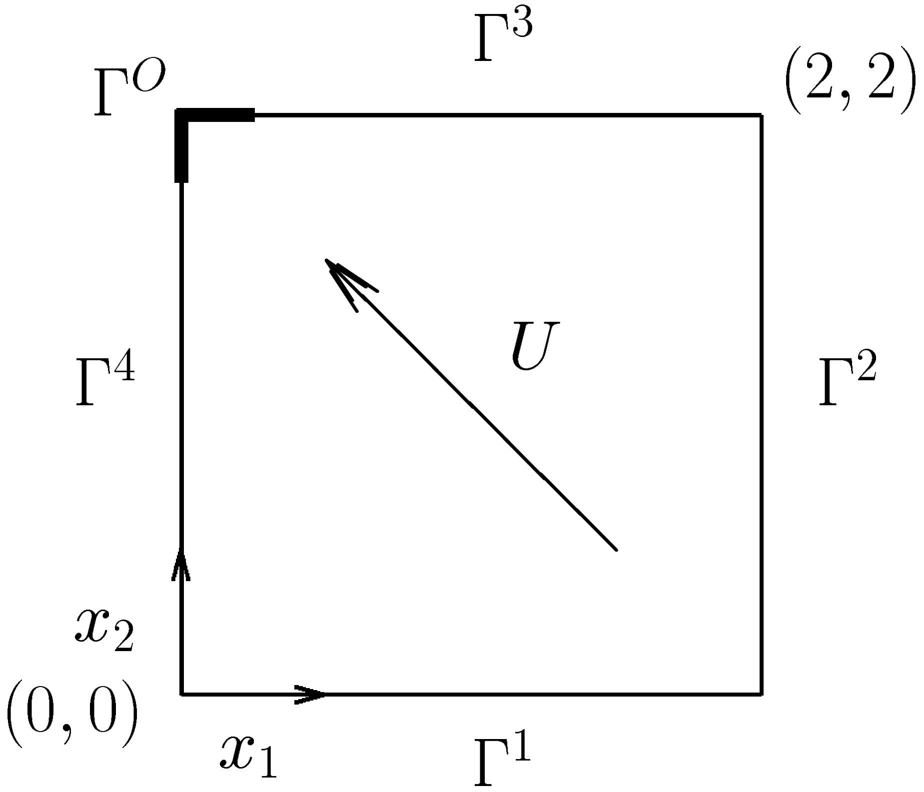


Figure 1: Geometry for convection–diffusion example: Γ^1 and Γ^2 are Dirichlet boundaries, and Γ^3 and Γ^4 are Neumann boundaries. The flow is specified as $U = (-1, 1)$, creating a boundary layer on the Neumann sides of the domain; the output of interest is the temperature in the vicinity of $(x_1, x_2) = (0, 2)$, in particular, the average temperature over the boundary segment Γ^O.

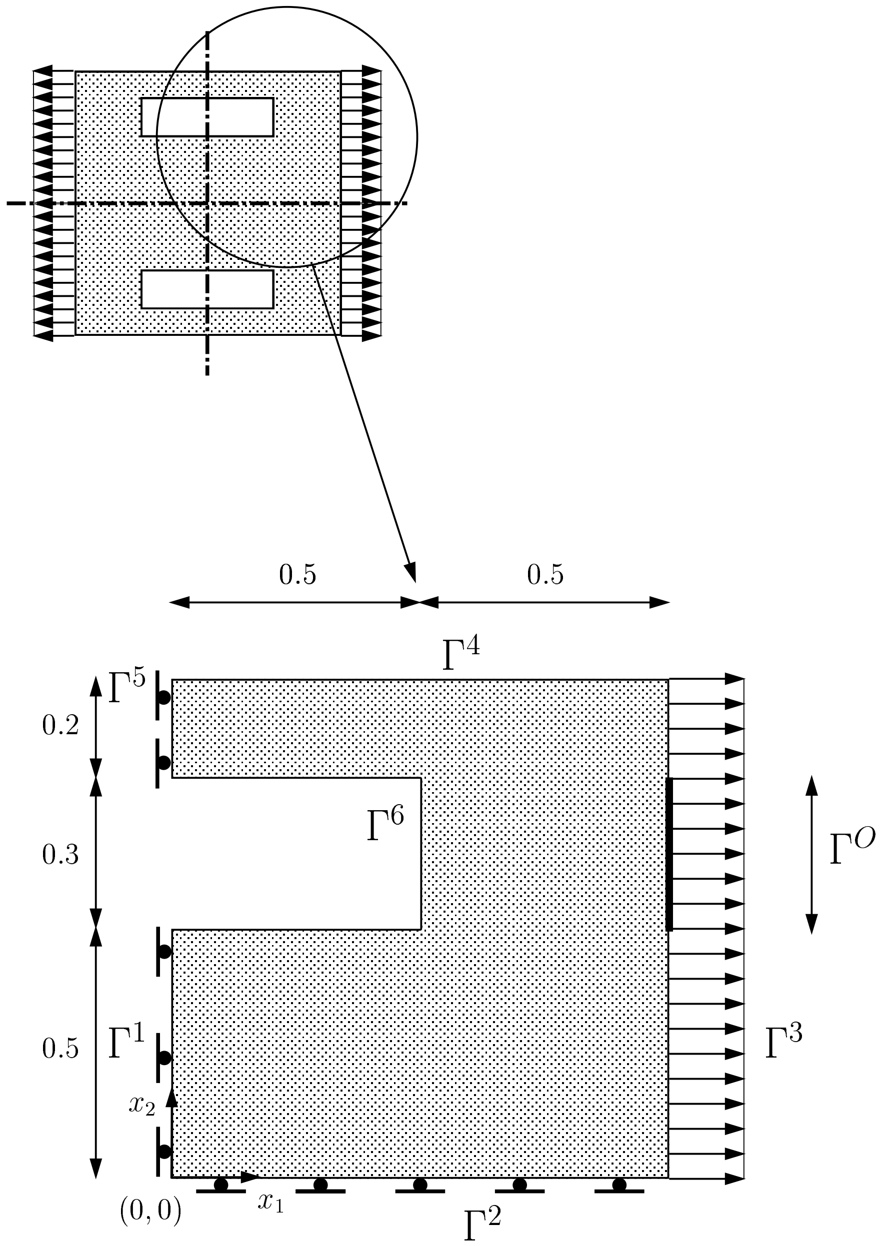
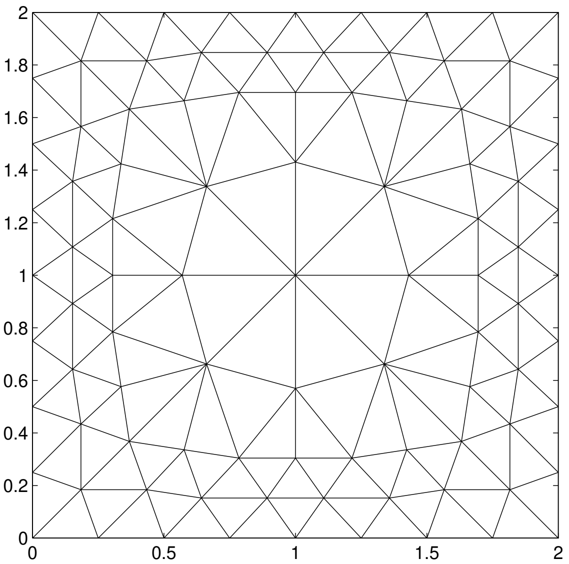


Figure 2: Geometry for linear elasticity example: one–fourth of a symmetrically loaded plate with two square holes. A normal traction is applied on Γ^3; the outputs of interest are the normal force on Γ^5, and the average normal displacement over the segment $\Gamma^O \subset \Gamma^3$.



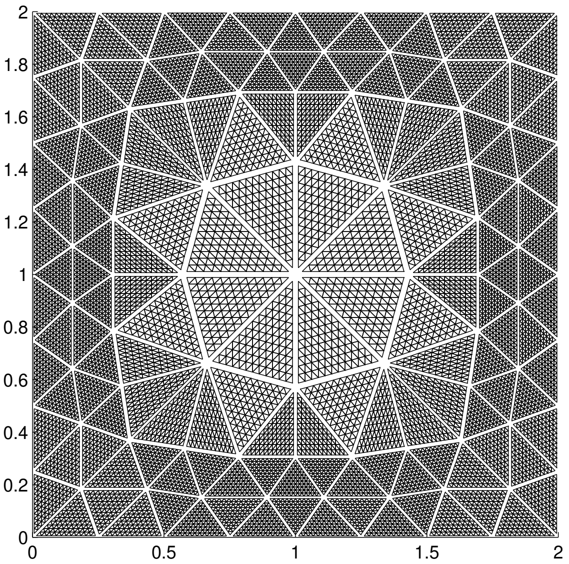


Figure 3: (a) Coarsest working mesh $\mathcal{T}_H = \mathcal{T}_{(H_0,1)}$, and (b) truth mesh $\mathcal{T}_h = \mathcal{T}_{(H_0,12)}$, for the convection–diffusion example.

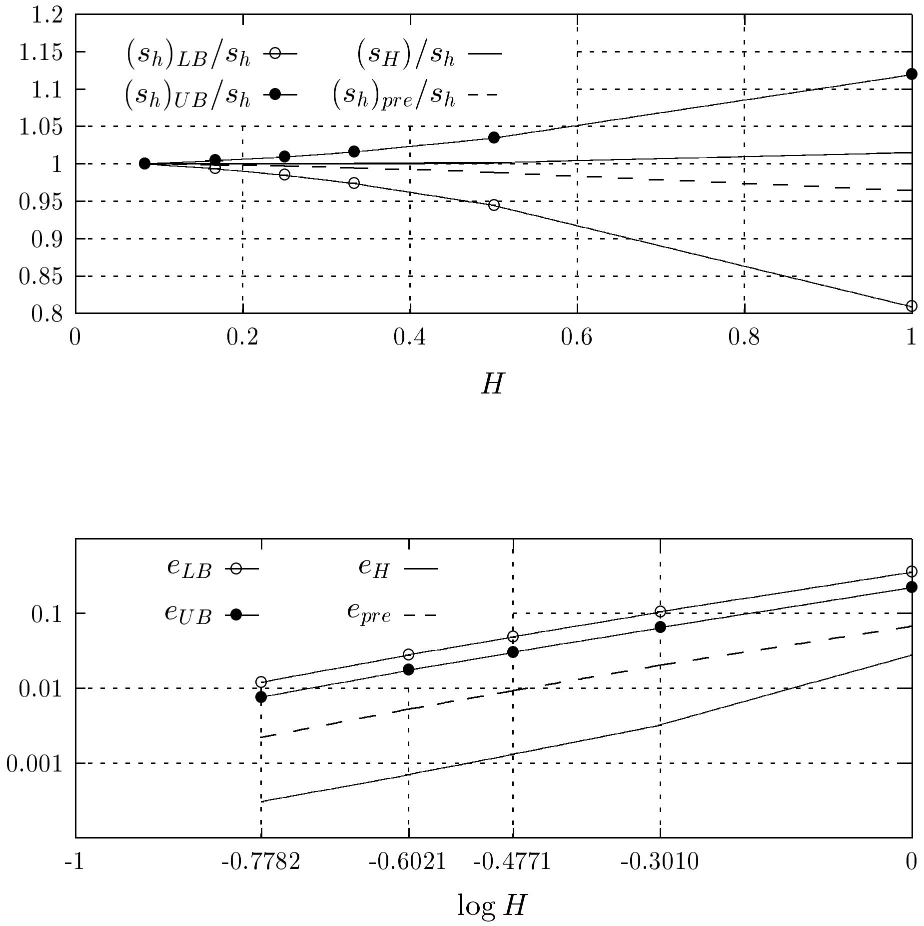


Figure 4: Bound predictions for s_h in the convection–diffusion example; the output s is the average temperature over the segment Γ^O shown in Figure 1. (a) Plot of $(s_h)_{UB}/s_h$, $(s_h)_{LB}/s_h$, $(s_h)_{pre}/s_h$, and s_H/s_h as a function of (effective) working–mesh element size H. (b) Plot of $e_{UB} \equiv \log|(s_h)_{UB} - s_h|$, $e_{LB} \equiv \log|(s_h)_{LB} - s_h|$, $e_{pre} \equiv \log|(s_h)_{pre} - s_h|$, and $e_H \equiv \log|s_H - s_h|$ as a function of $\log H$.

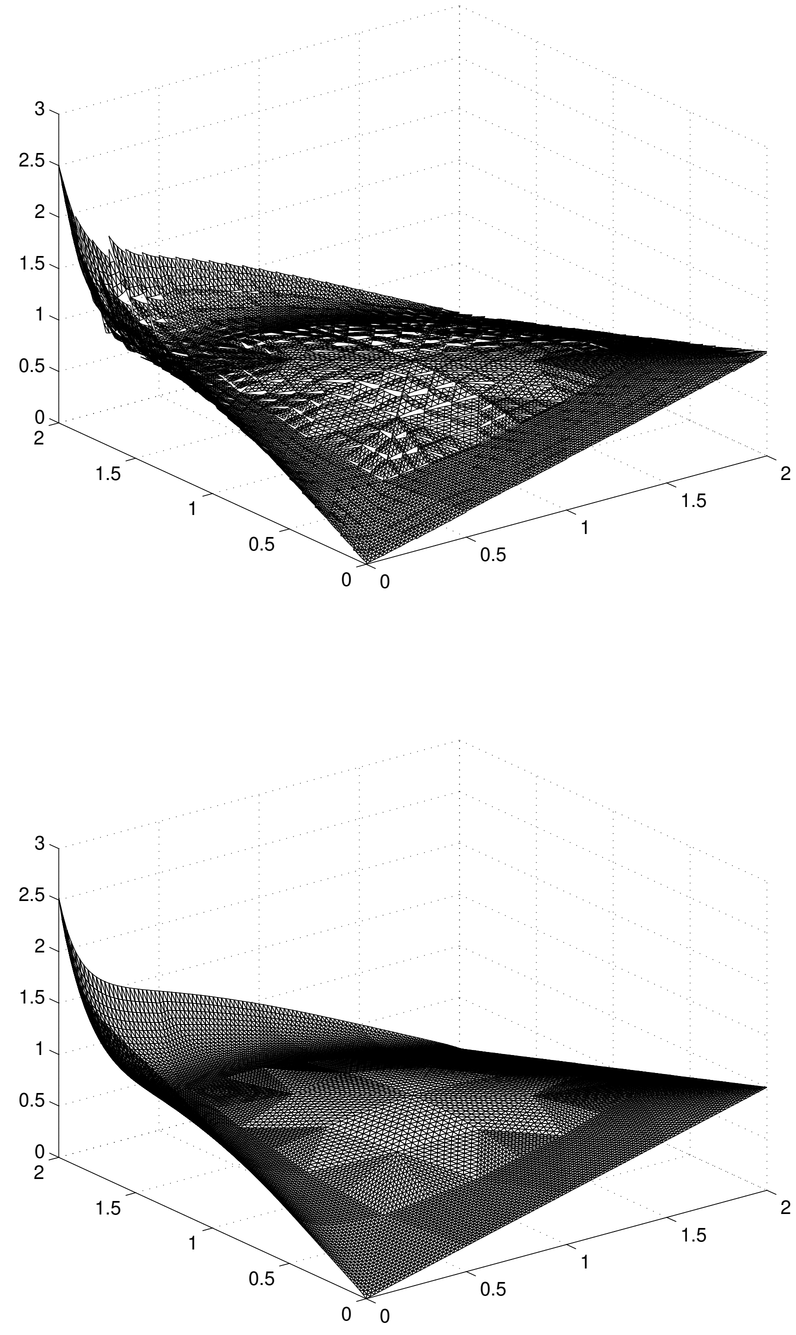
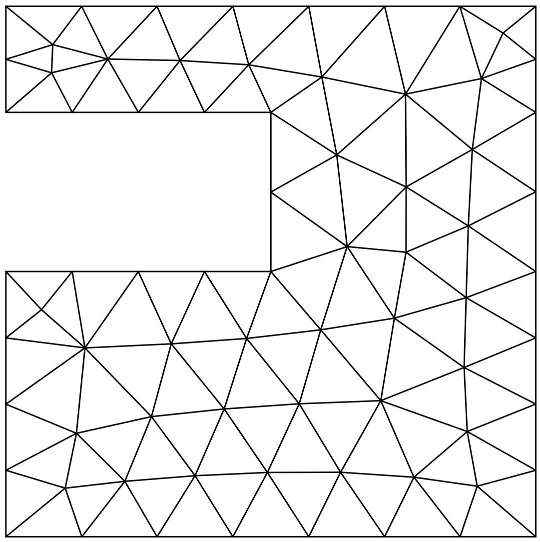


Figure 5: Plot of (a) $\hat{\mathcal{U}}_h^+$ for the working mesh $\mathcal{T}_H = \mathcal{T}_{(H_0,3)}$, and (b) u_h for the truth mesh $\mathcal{T}_h = \mathcal{T}_{(H_0,12)}$, for the convection–diffusion example.



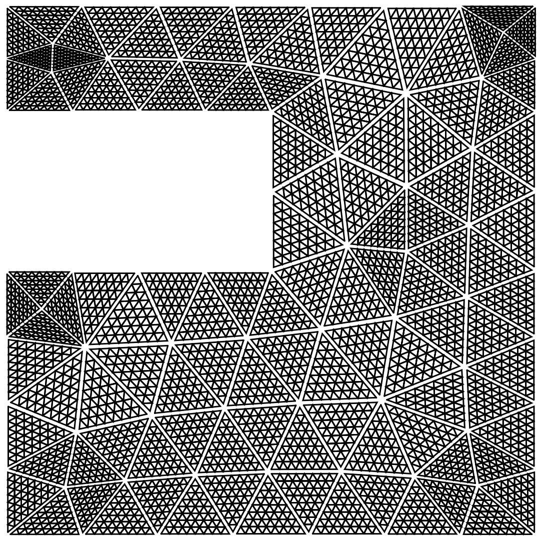


Figure 6: Two examples of H–meshes for the elasticity example: (a) coarsest H–mesh $\mathcal{T}_H = \mathcal{T}_{(H_0,1)}$, and (b) finest H–mesh $\mathcal{T}_H = \mathcal{T}_{(H_0,8)}$. The truth mesh is given by $\mathcal{T}_h = \mathcal{T}_{(H_0,16)}$.

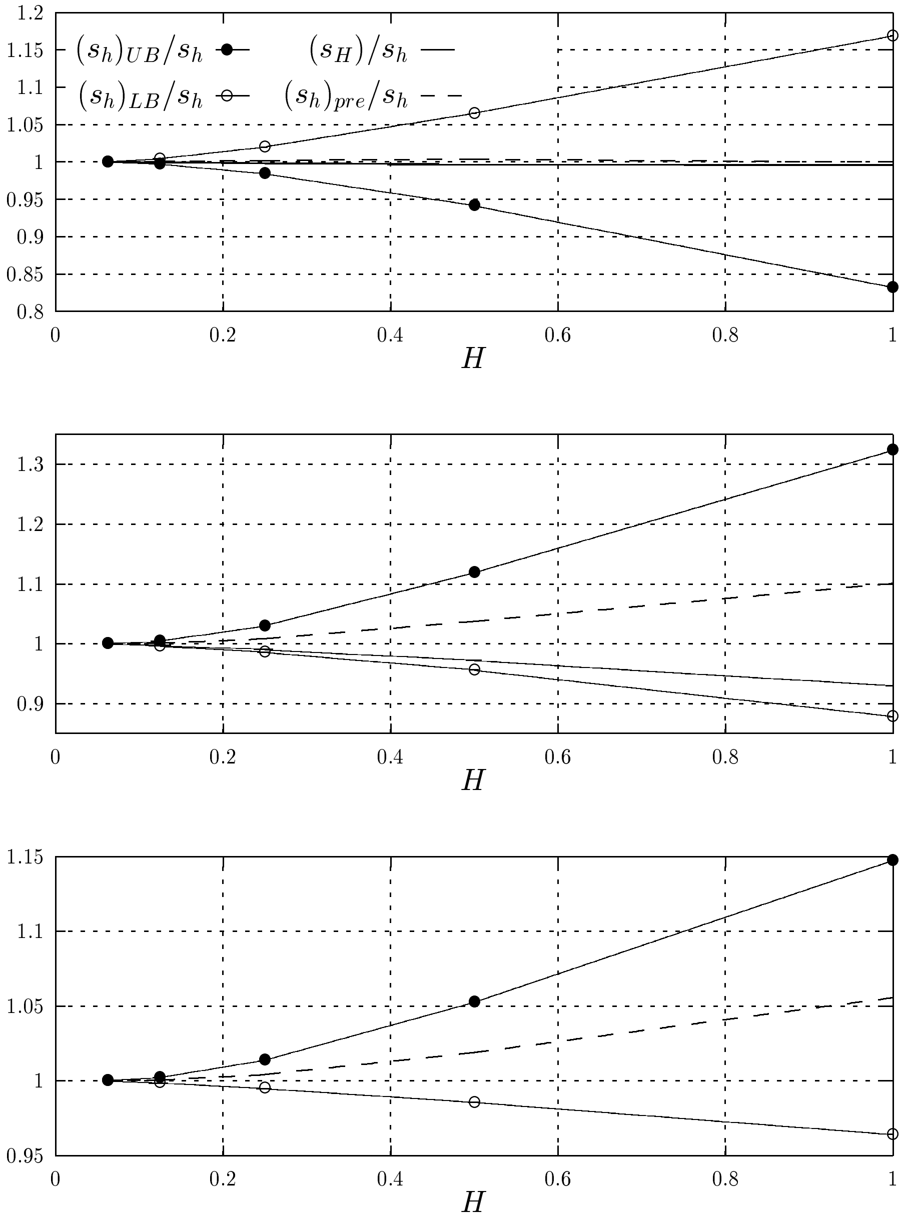


Figure 7: Plots of $(s_h)_{UB}/s_h$, $(s_h)_{LB}/s_h$, $(s_h)_{pre}/s_h$, and s_H/s_h as a function of (effective) H for (a) s^1, the force over Γ^5, (b) s^2, the average normal displacement over Γ^O indicated in Figure 2, and (c) s^2, the average normal displacement over $\Gamma^O = \Gamma^3$.

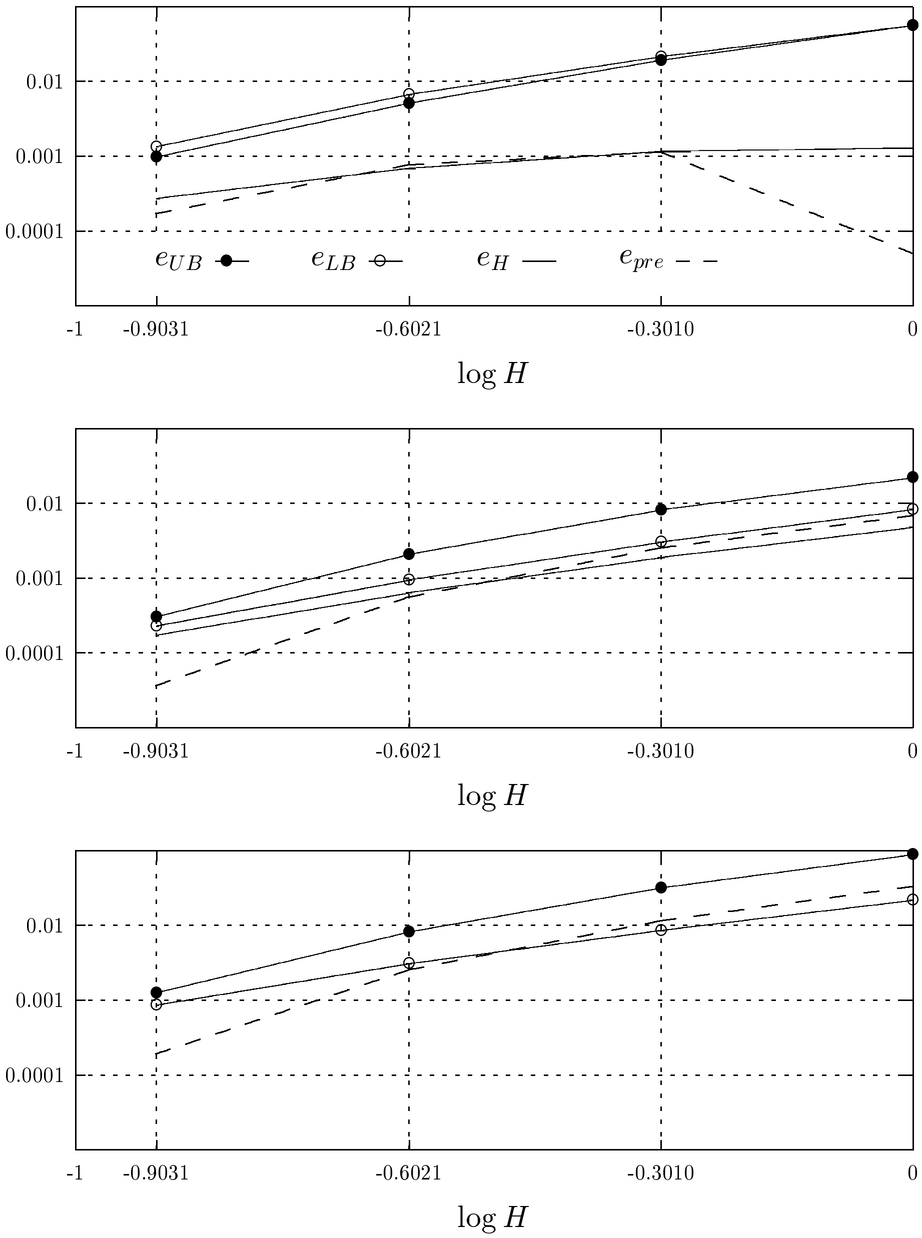


Figure 8: Plots of $e_{UB} \equiv \log |(s_h)_{UB} - s_h|$, $e_{LB} \equiv \log |(s_h)_{LB} - s_h|$, $e_{pre} \equiv \log |(s_h)_{pre} - s_h|$, and $e_H \equiv \log |s_H - s_h|$ as a function of $\log H$ for (a) s^1, the force over Γ^5, (b) s^2, the average normal displacement over Γ^O indicated in Figure 2, and (c) s^2, the average normal displacement over $\Gamma^O = \Gamma^3$.