Journal of Computational Physics **172,** 401–425 (2001)
doi:10.1006/jcph.2001.6769, available online at http://www.idealibrary.com on IDEAL®

A Posteriori Finite-Element Output Bounds for the Incompressible Navier–Stokes Equations: Application to a Natural Convection Problem[1]

L. Machiels,* J. Peraire,† and A. T. Patera*

*Department of Mechanical Engineering, and †Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139
E-mail: machiels@mit.edu

Received April 30, 1999; revised June 9, 2000

We present a new Neumann subproblem *a posteriori* finite-element procedure for the efficient calculation of rigorous, constant-free, sharp lower and upper bounds for linear and nonlinear functional outputs of the incompressible Navier–Stokes equations. We first formulate the bound procedure; we derive and discuss a bound error expression; and we then demonstrate the capabilities of the method with numerical results obtained for natural convection problems. We also implement an optimal adaptive refinement strategy based on a local elemental decomposition of the bound gap.

1. INTRODUCTION AND MOTIVATION

In typical design problems, engineers are rarely interested in the entire field solution; only some selected characteristic metrics—or outputs—of the system are relevant. As an example, we consider the problem of cooling electronic components by natural convection of air in the enclosure represented in Fig. 1a, where the temperature θ is fixed on the boundary Γ_0, a heat flux q is imposed on segments Γ_1, Γ_2, and Γ_3, and $\partial\Omega\backslash\cup_{i=0}^{3}\Gamma_i$ is insulated. In our example, the Boussinesq approximation is applicable, and the flow field is described by the incompressible Navier–Stokes equations coupled to a temperature equation. Given an input heat flux q, we wish to determine whether the mean temperature over Γ_1, $s = \frac{1}{\Gamma_1}\int_{\Gamma_1}\theta\, ds$, is within an acceptable design interval $\mathcal{I}_{\text{des}} = [s_{\text{lo}}, s_{\text{up}}]$. In practice, many different fluxes q must be tested, so there is a premium on efficiency. Much more complex design questions can also be addressed.

[1] MIT FML Report 99-4.

0021-9991/01 $35.00

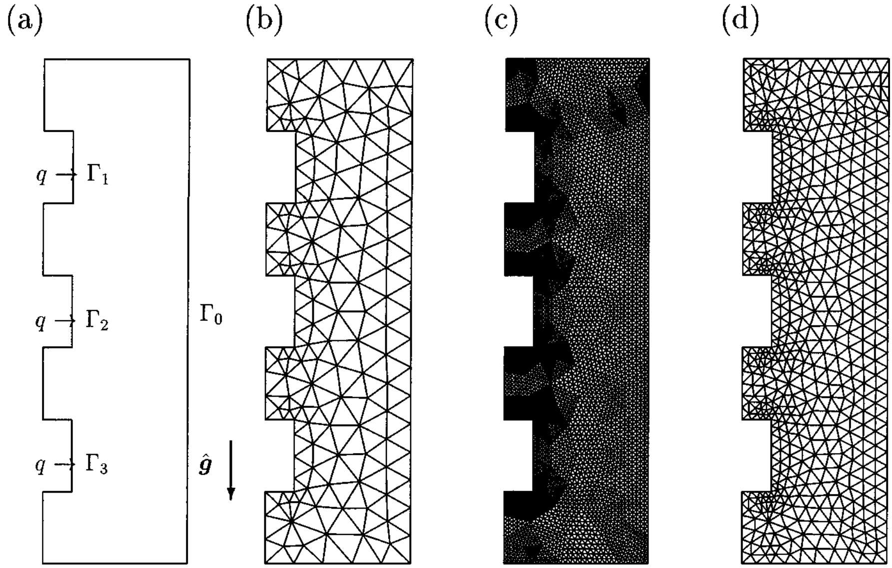

FIG. 1. (a) Domain Ω. (b) "Optimistic" coarse mesh, $\mathcal{T}_H$. (c) "Conservative" fine mesh, $\mathcal{T}_h = \mathcal{T}_{H/6}$. (d) Coarse mesh $\mathcal{T}_{H/2}$.

In a classical simulation-based design approach, the output of interest is evaluated from an approximate solution of the original problem. In the problem of cooling electronic components, Fig. 1a, we first evaluate the output of interest $s_\delta = S(\boldsymbol{u}_\delta, p_\delta, \theta_\delta)$ from an approximate field solution $(\boldsymbol{u}_\delta, p_\delta, \theta_\delta)$—here $\boldsymbol{u}_\delta$ is the velocity field, p_δ is the pressure field, θ_δ is the temperature field, and δ denotes the diameter of the discretization mesh $\mathcal{T}_\delta$. We then verify whether $s_\delta \in \mathcal{I}_{\text{des}}$, and the heat flux q is accepted or rejected accordingly. The shortcoming of the classical approach resides in the choice of the discretization mesh $\mathcal{T}_\delta$. If one chooses an "optimistic" coarse mesh $\mathcal{T}_H$, Fig. 1b, the calculation is *inexpensive* but also *uncertain* since neither $s_H \in \mathcal{I}_{\text{des}}$ implies $s \in \mathcal{I}_{\text{des}}$ nor $s_H \notin \mathcal{I}_{\text{des}}$ implies $s \notin \mathcal{I}_{\text{des}}$. If, instead, one chooses a "conservative" sufficiently fine mesh $\mathcal{T}_h$, Fig. 1c—this mesh is obtained by dividing each triangle of mesh $\mathcal{T}_H$ into 36 self-similar triangles—then $s_h \approx s$ with reasonable *certainty*, but s_h is now very *expensive* to compute.

This paper presents a new approach which offers great promise in reconciling these conflicting requirements. We propose to construct a pair of output bound approximations, the estimators s_H^+ and s_H^-, computed predominantly on the mesh $\mathcal{T}_H$ and with the following attributes:

A1. As $H \to h$, we have $s_H^+ \to s_h$ from above and $s_H^- \to s_h$ from below $\forall H \leq H^*$. Here H^* is an unknown threshold discretization parameter; a detailed discussion of H^* will be given subsequently.

A2. If we define the half bound gap $\Delta_H = \frac{1}{2}(s_H^+ - s_H^-)$, then $\Delta_H \leq \eta |s_H - s_h|$ as $H \to h$, with η independent of H. This property guarantees the optimal convergence rate and sharpness of the bounds, provided the effectivity factor η is not too large.

A3. The bound gap Δ_H admits elemental decomposition $\Delta_H = \sum_{T_H \in \mathcal{T}_H} \Delta_{T_H}$, with $\Delta_{T_H} \geq 0$ for all elements T_H in $\mathcal{T}_H$; this property will be used for adaptive refinement.

A4. The work required to evaluate s_H^+ and s_H^- is substantially less than that necessary for the computation of s_h, provided $H \ll h$.

Note that attribute A2 is important not only for efficiency but also for ensuring the "well-posedness" of the estimator formulation when $H^* < \infty$.

In our example, Fig. 1a, the bound-based design would proceed as follows. Given a heat flux q, we choose an initial mesh $\mathcal{T}_H$, and we compute s_H^+ and s_H^-. Next, we define $\mathcal{I}_\mathrm{b} = [s_H^-, s_H^+]$; according to attribute A1 (provided $H \leq H^*$), if $\mathcal{I}_\mathrm{b} \in \mathcal{I}_\mathrm{des}$, we accept q; if $\mathcal{I}_\mathrm{b} \in \mathbb{R}\backslash\mathcal{I}_\mathrm{des}$, we reject q; otherwise, we use property A2—that is, we narrow the bound gap by taking a finer mesh (a smaller H), and we repeat the procedure. In the last case, attribute A3 is important as it allows us to optimally refine the mesh through an adaptive procedure. Attribute A4 ensures that the complete procedure is much less expensive than directly computing s_h.

We illustrate these concepts by fixing a design interval: $\mathcal{I}_\mathrm{des} = [0.26, 0.28]$. For a particular choice of the flux q and of the parameters governing the problem, the outcome of the bound procedure performed on the coarse mesh $\mathcal{T}_H$ (Fig. 1b) is $s_h = 0.275 \pm 8\%$, or $s_h \in \mathcal{I}_\mathrm{b} = [0.253, 0.297]$. According to our specification of $\mathcal{I}_\mathrm{des}$, we cannot decide whether to accept or reject the heat flux q. Thus, we refine the mesh. The mesh $\mathcal{T}_{H/2}$ is obtained by dividing each triangle of mesh $\mathcal{T}_H$ into four triangles (by dividing each edge into two edges); see Fig. 1d. For this new mesh, the bound procedure yields $s_h = 0.276 \pm 1.4\%$, or $s_h \in \mathcal{I}_\mathrm{b} = [0.272, 0.280]$. Since now $\mathcal{I}_\mathrm{b} \in \mathcal{I}_\mathrm{des}$, we can safely accept the flux q. Note that, in this example, we have not used attribute A3 which allows optimal adaptive refinement. An example utilizing adaptivity will be given in Section 4.

The procedure is an extension of our recent general error-control strategy [8, 11, 13, 15, 17], and may be viewed as an implicit Aubin–Nitsche construction; for a review, see [9]. In [14], an early application of the technique to the Stokes equations is presented. Our method is indebted to, but considerably generalizes, earlier finite element *a posteriori* error-estimation techniques in that it provides a quantitative constant-free bound, in contrast to earlier explicit techniques [4], and the bounded quantity is the output of interest, in contrast to earlier implicit techniques [2, 3, 7]. Some aspects of the method are also analogous to certain domain decomposition techniques [6, 12].

The application of our general technique to the Navier–Stokes equation presents several new challenges. The purpose of the present paper is (i) to give a detailed and complete account of the bound algorithm, (ii) to derive and discuss a bound error expression, and (iii) to demonstrate the attributes of the procedure with numerical examples. A complete and rigorous numerical analysis of the method is relegated to a future paper.

The paper is organized as follows. Section 2 contains the formulation of the bound procedure. We start by formulating the finite-element discretization of the natural convection problem. We next introduce some preliminary definitions, and we describe the bound algorithm in detail. Finally, we derive a bound error expression. Section 3 presents numerical results for two natural convection problems and an implementation of an optimal adaptive refinement strategy based on the bound procedure. Section 4 concludes the paper with a review of the various attributes of the method. For the convenience of presentation, some mathematical developments have been presented in the Appendixes.

2. BOUND PROCEDURE

This section is divided into four parts. We first introduce the equations governing the natural convection problem considered, and we formulate the finite-element approximation. To ease the subsequent discussion, we also recall or introduce the necessary function spaces, finite-element spaces, and forms. In the second part, we complete our definitions by introducing the ingredients particular to the bound method, namely, the "broken" finite-element spaces—which are characterized by relaxation of the continuity constraint of the finite-element functions across edges of the coarse mesh—and the Taylor expansion and splitting of the nonlinear forms introduced in the first part. In the third part, after the introduction of the appropriate concepts and definitions, we detail and discuss the five steps of the bound construction. Finally, in the last part, we discuss the bound error expression.

2.1. *Problem Statement*

Given a two-dimensional domain $\Omega \in \mathbb{R}^2$, we consider a problem of natural convection which is, in the Boussinesq approximation, described by the system of equations

$$-\frac{\partial}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)+\frac{\partial (u_j u_i)}{\partial x_j}+\frac{\partial p}{\partial x_i}=-\beta\theta\hat{g}_i \quad \text{in } \Omega, \tag{1}$$

$$\frac{\partial u_i}{\partial x_i}=0 \quad \text{in } \Omega, \tag{2}$$

$$-\frac{1}{\alpha}\frac{\partial^2\theta}{\partial x_j \partial x_j}+\frac{\partial (u_j\theta)}{\partial x_j}=0 \quad \text{in } \Omega, \tag{3}$$

where the summation convention of repeated indices applies, $\boldsymbol{u}=(u_1,u_2)$ is the velocity field, p is the pressure field, θ is the temperature field, and $\hat{\boldsymbol{g}}=(\hat{g}_1,\hat{g}_2)$ is the unit vector indicating the direction of gravity. The flow is governed by two nondimensional parameters: the Prandtl number α and the Grashof number β. We supplement Eqs. (1), (2), and (3) with the boundary conditions

$$\boldsymbol{u}_{|\partial\Omega}=0 \quad \text{and} \quad \begin{cases} \theta = 0 & \text{on } \Gamma_0, \\ \frac{1}{\alpha}\frac{\partial\theta}{\partial n}=g_N & \text{on } \Gamma_k, \quad \text{for } k=1,2,\ldots,K, \\ \frac{\partial\theta}{\partial n}=0 & \text{on } \partial\Omega\backslash\bigcup_{k=0}^{K}\Gamma_k, \end{cases} \tag{4}$$

where $\Gamma_k \subset \partial\Omega,\ \Gamma_k\cap\Gamma_l=\emptyset$ when $k\neq l$, and the function $g_N : \bigcup_{k=0}^{K}\Gamma_k \to \mathbb{R}$ represents the heat flux.

2.1.1. Variational formulation. Our point of departure for a finite-element approximation of Eqs. (1), (2), and (3) is a variational formulation. We first recall the definitions of some function spaces and their associated norms and seminorms. For $1\le p<+\infty$, we define the spaces

$$L^p(\Omega)=\left\{v:\Omega\to\mathbb{R}\mid v \text{ is measurable and } \int_\Omega |v|^p\,d\Omega<+\infty\right\}$$

and their associated norm

$$\|v\|_{L^p(\Omega)}=\left(\int_\Omega |v|^p\,d\Omega\right)^{1/p}.$$

We also define

$$L^\infty(\Omega)=\{v:\Omega\to\mathbb{R}\mid \operatorname{ess\,sup}\{|v(\boldsymbol{x})|;\boldsymbol{x}\in\Omega\}<+\infty\}$$

and the norm $\|v\|_\infty = \text{ess sup}\{|v(\boldsymbol{x})|; \boldsymbol{x} \in \Omega\}$. Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ with α_1 and α_2 being non-negative integers; we define

$$D^{\boldsymbol{\alpha}} v = \frac{\partial^{|\boldsymbol{\alpha}|} v}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2}},$$

where $|\boldsymbol{\alpha}| = \alpha_1 + \alpha_2$. The Sobolev space $H^k(\Omega)$ (e.g., see [1]), where k is a non-negative integer, is the space

$$H^k(\Omega) = \{v \in L^2(\Omega) \mid D^{\boldsymbol{\alpha}} v \in L^2(\Omega), \forall |\boldsymbol{\alpha}| \leq k\}.$$

We will also use the spaces

$$H_0^k(\Omega) = \{v \in H^k(\Omega) \mid v_{|\partial\Omega} = 0\}.$$

We finally introduce the norms and seminorms associated with the spaces $H^k(\Omega)$,

$$\|v\|_{H^k(\Omega)} = \left(\sum_{|\boldsymbol{\alpha}| \leq k} \int_\Omega |D^{\boldsymbol{\alpha}} v|^2 \, d\Omega \right)^{1/2}, \quad |v|_{H^k(\Omega)} = \left(\sum_{|\boldsymbol{\alpha}| = k} \int_\Omega |D^{\boldsymbol{\alpha}} v|^2 \, d\Omega \right)^{1/2}.$$

We now define the function spaces $X = H_0^1(\Omega) \times H_0^1(\Omega)$, $M = L^2(\Omega)$,

$$X^\theta = \left\{ v \in H^1(\Omega) \mid v_{|\Gamma_0} = 0 \right\},$$

and $Y = X \times M \times X^\theta$. To facilitate the variational formulation of the problem, we define the bilinear and trilinear forms

$$\begin{aligned}
a : X \times X \to \mathbb{R}, &\quad a(\boldsymbol{v}, \boldsymbol{w}) = \int_\Omega \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \frac{\partial w_i}{\partial x_j} \, d\Omega, \\
a^\theta : X^\theta \times X^\theta \to \mathbb{R}, &\quad a^\theta(\phi, \rho) = \frac{1}{\alpha} \int_\Omega \frac{\partial \phi}{\partial x_i} \frac{\partial \rho}{\partial x_i} \, d\Omega, \\
b : X \times M \to \mathbb{R}, &\quad b(\boldsymbol{v}, r) = -\int_\Omega \frac{\partial v_i}{\partial x_i} r \, d\Omega, \\
c : X \times X \times X \to \mathbb{R}, &\quad c(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}) = -\int_\Omega (v_{1i} v_{2j}) \frac{\partial w_i}{\partial x_j} \, d\Omega, \\
c^\theta : X^\theta \times X \times X^\theta \to \mathbb{R}, &\quad c^\theta(\phi, \boldsymbol{v}, \rho) = -\int_\Omega (\phi v_j) \frac{\partial \rho}{\partial x_j} \, d\Omega, \\
d : X^\theta \times X \to \mathbb{R}, &\quad d(\phi, \boldsymbol{w}) = \beta \int_\Omega \phi w_i \hat{g}_i \, d\Omega.
\end{aligned}$$

Finally, we introduce the "natural convection form,"

$$\begin{aligned}
\mathcal{A}((\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho)) &= a(\boldsymbol{v}, \boldsymbol{w}) + a^\theta(\phi, \rho) + b(\boldsymbol{v}, r) + b(\boldsymbol{w}, q) \\
&\quad + c(\boldsymbol{v}, \boldsymbol{v}, \boldsymbol{w}) + c^\theta(\phi, \boldsymbol{v}, \rho) + d(\phi, \boldsymbol{w}) - \langle g_N, \rho \rangle_N, \qquad (5)
\end{aligned}$$

where

$$\langle g_N, \rho \rangle_N = \sum_{0<k\leq K} \int_{\Gamma_k} g_N \rho_{|\Gamma_k} \, d\Gamma$$

(we assume $g_N \in L^2(\cup_{k=1,\ldots,K} \Gamma_k)$). The variational formulation of Eqs. (1), (2), and (3), with boundary conditions (4), then consists of finding $(\boldsymbol{u}, p, \theta) \in Y$ such that

$$\mathcal{A}((\boldsymbol{u}, p, \theta), (\boldsymbol{v}, q, \phi)) = 0, \quad \forall (\boldsymbol{v}, q, \phi) \in Y. \tag{6}$$

As indicated in the Introduction, we assume that we are not directly interested in the complete field solution $(\boldsymbol{u}, p, \theta)$, but that we wish to evaluate an output of interest $s = S(\boldsymbol{u}, p, \theta)$. We assume that the form S can be expressed as

$$S(\boldsymbol{v}, q, \phi) = c_0 + \ell(\boldsymbol{v}, q, \phi) + m(\boldsymbol{v}, \boldsymbol{v}), \tag{7}$$

where $c_0 \in \mathbb{R}$, $\ell : Y \to \mathbb{R}$ is a bounded linear functional, and $m : X \times X \to \mathbb{R}$ is a continuous, symmetric, bilinear functional such that

$$m(\boldsymbol{v}, \boldsymbol{w}) \leq C \|\boldsymbol{v}\|_{L^2(\Omega)} \|\boldsymbol{w}\|_{L^2(\Omega)},$$

with $C > 0$. The norm in the product space $L^2(\Omega) \times L^2(\Omega)$ is defined by $\|\boldsymbol{v}\|^2_{L^2(\Omega)} = \|v_1\|^2_{L^2(\Omega)} + \|v_2\|^2_{L^2(\Omega)}$. In addition, since in this formulation the pressure p is defined only to within a constant, we require that

$$\ell(\boldsymbol{0}, 1, 0) = 0. \tag{8}$$

2.1.2. Finite-element formulation. We will now define a finite-element approximation of (6). We consider a (regular, uniform) triangulation $\mathcal{T}_\delta$ of the domain Ω; that is, $\mathcal{T}_\delta$ is a set of triangles T_δ such that

$$\bar{\Omega} = \cup_{T_\delta \in \mathcal{T}_\delta} \bar{T}_\delta \quad \text{and} \quad T_\delta \cap T'_\delta = \emptyset \quad \text{if } T_\delta \neq T'_\delta,$$

where $\bar{A}$ denotes the closure of $A \subset \mathbb{R}^2$, and δ is the diameter of the triangulation $\mathcal{T}_\delta$,

$$\delta = \max_{T_\delta \in \mathcal{T}_\delta} \operatorname{diam}(T_\delta), \quad \text{where } \operatorname{diam}(T_\delta) = \max_{\boldsymbol{x},\boldsymbol{y} \in T_\delta} |\boldsymbol{x} - \boldsymbol{y}|.$$

For a given triangulation $\mathcal{T}_\delta$, we will denote by $\mathcal{T}_{\delta/2}, \mathcal{T}_{\delta/3}, \ldots$ the refinements of $\mathcal{T}_\delta$ obtained by dividing each triangle of T_δ into $2^2, 3^2, \ldots$ self-similar triangles. In particular, we will say that $\mathcal{T}_{\delta'}$ is a refinement of $\mathcal{T}_\delta$ if there is an integer $n > 0$ such that $\mathcal{T}_{\delta'} = \mathcal{T}_{\delta/n}$.

In this work, we choose the Crouzeix–Raviart finite-element spaces for the velocity and the pressure [5]; that is, the approximate velocity $\boldsymbol{u}_\delta$ is in X_δ and the approximate pressure p_δ is in M_δ, where

$$X_\delta = \left\{ \boldsymbol{v} \in X \mid v_{i|T_\delta} \in \mathbb{P}_2(T_\delta) \oplus B_3(T_\delta), \forall T_\delta \in \mathcal{T}_\delta \right\},$$
$$M_\delta = \left\{ p \in L^2(\Omega) \mid p_{|T_\delta} \in \mathbb{P}_1(T_\delta), \forall T_\delta \in \mathcal{T}_\delta \right\}.$$

Here $\mathbb{P}_p(T_\delta)$ is the space of all polynomials of degree $\leq p$ defined on T_δ, and $B_3(T_\delta)$ denotes the space of bubble functions of degree 3 on T_δ. We use the standard quadratic elements for the temperature field, $\theta_\delta \in X_\delta^\theta$, where

$$X_\delta^\theta = \left\{\phi \in X^\theta \mid \phi_{|T_\delta} \in \mathbb{P}_2(T_\delta), \forall T_\delta \in \mathcal{T}_\delta\right\}.$$

Finally, we define the solution space as the product space $Y_\delta = X_\delta \times M_\delta \times X_\delta^\theta$.

The finite-element method for Eqs. (1), (2), and (3) with the boundary conditions (4) consists of the construction of an approximate solution $(\boldsymbol{u}_\delta, p_\delta, \theta_\delta) \in Y_\delta$ such that

$$\mathcal{A}((\boldsymbol{u}_\delta, p_\delta, \theta_\delta), (\boldsymbol{v}, q, \phi)) = 0, \quad \forall(\boldsymbol{v}, q, \phi) \in Y_\delta.$$

We now assume that we are given two triangulations, $\mathcal{T}_H$, the coarse mesh, and $\mathcal{T}_h$, the fine mesh, with $\mathcal{T}_h = \mathcal{T}_{H/n}$ a refinement of $\mathcal{T}_H$ and $h \ll H$. Therefore, we have $Y_H \subset Y_h$, where Y_H is the finite-element space associated with the triangulation $\mathcal{T}_H$, and Y_h is the finite-element space associated with $\mathcal{T}_h$. This inclusion, although not required for the subsequent formulation of the bound method, follows immediately if we note that the fine mesh is obtained by dividing each triangle $T_H \in \mathcal{T}_H$ into self-similar triangles and, thus, the trace of the cubic bubble function on any edge of the fine mesh is at most quadratic. The associated coarse-space and fine-space approximations, $(\boldsymbol{u}_H, p_H, \phi_H)$ and $(\boldsymbol{u}_h, p_h, \phi_h)$, exhibit complementary advantages and disadvantages. The fine-space solution, $(\boldsymbol{u}_h, p_h, \phi_h) \in Y_h$, which satisfies the discrete equations

$$\mathcal{A}((\boldsymbol{u}_h, p_h, \theta_h), (\boldsymbol{v}, q, \phi)) = 0, \quad \forall(\boldsymbol{v}, q, \phi) \in Y_h, \tag{9}$$

yields a very good approximation, $s_h = S(\boldsymbol{u}_h, p_h, \theta_h)$, to the exact output s; nevertheless, the computational effort required to obtain $(\boldsymbol{u}_h, p_h, \theta_h)$ will typically be prohibitive. In contrast, the coarse-space solution $(\boldsymbol{u}_H, p_H, \theta_H) \in Y_H$, which satisfies the discrete equations

$$\mathcal{A}((\boldsymbol{u}_H, p_H, \theta_H), (\boldsymbol{v}, q, \phi)) = 0, \quad \forall(\boldsymbol{v}, q, \phi) \in Y_H,$$

can be obtained with relatively modest computational effort; nevertheless the fidelity of the corresponding approximate output, $s_H = S(\boldsymbol{u}_H, p_H, \theta_H)$, is no longer ensured.

2.2. *Bounds Preliminaries*

The formulation of the bound algorithm requires the definition of particular finite element spaces, the "broken" spaces, which are obtained as extensions of the classical finite-element spaces by relaxing the continuity of the finite-element basis functions across edges of the coarse mesh. Moreover, due to the nonlinearity, noncoercivity, and nonsymmetry of the problem, we must also expand the nonlinear forms $\mathcal{A}$ and S into other linear, bilinear, and trilinear forms. We then further decompose some of these new forms into coercive—with respect to divergence-free functions—and noncoercive parts and symmetric and antisymmetric components.

2.2.1. "Broken" spaces. We denote by $\Gamma(\mathcal{T}_H)$ the set of edges γ of $\mathcal{T}_H$. We define the "broken" finite-element spaces as

$$\begin{aligned}\hat{X}_H &= \left\{\boldsymbol{v} \in L^2(\Omega) \times L^2(\Omega) \mid v_{i|T_H} \in \mathbb{P}_2(T_H) \oplus B_3(T_H), \forall T_H \in \mathcal{T}_H\right\},\\ \hat{X}_h &= \left\{\boldsymbol{v} \in L^2(\Omega) \times L^2(\Omega) \mid v_{i|T_H} \in H^1(T_h), \forall T_H \in \mathcal{T}_H;\right.\\ &\qquad \left. v_{i|T_h} \in \mathbb{P}_2(T_h) \oplus B_3(T_h), \forall T_h \in \mathcal{T}_h\right\},\\ \hat{X}^\theta_H &= \left\{\phi \in L^2(\Omega) \mid \phi_{|T_H} \in \mathbb{P}_2(T_H), \forall T_H \in \mathcal{T}_H\right\},\\ \hat{X}^\theta_h &= \left\{\phi \in L^2(\Omega) \mid \phi_{|T_H} \in H^1(T_H), \forall T_H \in \mathcal{T}_H; \phi_{|T_h} \in \mathbb{P}_2(T_h), \forall T_h \in \mathcal{T}_h\right\}.\end{aligned}$$

If $\hat{Y}_H = \hat{X}_H \times M_H \times \hat{X}^\theta_H$ and $\hat{Y}_h = \hat{X}_h \times M_h \times \hat{X}^\theta_h$, we have $Y_H \subset \hat{Y}_H$ and $Y_h \subset \hat{Y}_h$.

We also define the "hybrid flux" spaces

$$\begin{aligned}Q_H &= \left\{z \in L^2(\Gamma(\mathcal{T}_H)) \times L^2(\Gamma(\mathcal{T}_H)) \mid z_{i|\gamma} \in \mathbb{P}_4(\gamma), \forall \gamma \in \Gamma(\mathcal{T}_H)\right\},\\ Q^\theta_H &= \left\{z \in L^2(\Gamma(\mathcal{T}_H)) \mid z_{|\gamma} \in \mathbb{P}_4(\gamma), \forall \gamma \in \Gamma(\mathcal{T}_H); z_{|\gamma} = 0, \forall \gamma \in \partial\Omega\backslash\Gamma_0\right\}.\end{aligned}$$

The reason for using quartic polynomials ($\mathbb{P}_4$) will become apparent in the bound procedure. (In fact, $\mathbb{P}_2$ polynomials would be sufficient.) We define the product space $Z_H = Q_H \times Q^\theta_H$ and the form $B : (X_h \times M_h) \times Z_H \to \mathbb{R}$,

$$B((\boldsymbol{v}, \theta), (\boldsymbol{z}, z^\theta)) = \sum_{\gamma \in \Gamma(\mathcal{T}_H)} \int_\gamma \left([v_i]_\gamma z_{i|\gamma} + [\theta]_\gamma z^\theta_{|\gamma}\right) d\gamma,$$

where $[v]_\gamma$ denotes the jump in v across γ when $\gamma \in \Omega$ and the trace of v when $\gamma \in \partial\Omega$. Note that we have the following equivalence condition:

$$Y_H = \left\{(\boldsymbol{v}, q, \phi) \in \hat{Y}_H \mid B((\boldsymbol{v}, \phi), (\boldsymbol{z}, z^\theta)) = 0, \forall (\boldsymbol{z}, z^\theta) \in Z_H\right\}.$$

2.2.2. Form expansions. We define

$$\begin{aligned}\mathcal{E}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho)) &= \mathcal{A}((\bar{\boldsymbol{u}} + \boldsymbol{v}, \bar{p} + q, \bar{\theta} + \phi), (\boldsymbol{w}, r, \rho))\\ &\quad - \mathcal{A}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}), (\boldsymbol{w}, r, \rho)).\end{aligned} \tag{10}$$

Since $\mathcal{A}((\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho))$ is quadratic in $(\boldsymbol{v}, \phi)$, we can express $\mathcal{E}$ as

$$\begin{aligned}\mathcal{E}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho)) &= E((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho))\\ &\quad + F((\boldsymbol{v}, \phi), (\boldsymbol{v}, \phi), (\boldsymbol{w}, \rho)),\end{aligned} \tag{11}$$

where E is linear in $(\boldsymbol{v}, q, \phi)$ and $(\boldsymbol{w}, r, \rho)$, and F is trilinear. The form E represents the first variation of the natural convection form (5). We now proceed with the decomposition

$$E((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho)) = E_0((\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho)) + E_1((\bar{u}, \bar{\theta}); (\boldsymbol{v}, \phi), (\boldsymbol{w}, \rho)),$$

where

$$E_0((\boldsymbol{v}, q, \phi), (\boldsymbol{w}, r, \rho)) = a(\boldsymbol{v}, \boldsymbol{w}) + a^\theta(\phi, \rho) + b(\boldsymbol{v}, r) + b(\boldsymbol{w}, q), \tag{12}$$

and

$$E_1((\bar{\boldsymbol{u}}, \bar{\theta}); (\boldsymbol{v}, \phi), (\boldsymbol{w}, \rho)) = c(\bar{\boldsymbol{u}}, \boldsymbol{v}, \boldsymbol{w}) + c(\boldsymbol{v}, \bar{\boldsymbol{u}}, \boldsymbol{w}) + d(\phi, \boldsymbol{w}) + c^\theta(\bar{\theta}, \boldsymbol{v}, \rho) + c^\theta(\phi, \bar{\boldsymbol{u}}, \rho).$$

The form F is expressed as

$$\begin{aligned} &F((\boldsymbol{v}_1, \phi_1), (\boldsymbol{v}_2, \phi_2), (\boldsymbol{w}, \rho)) \\ &\quad = \frac{1}{2}\left[c(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}) + c(\boldsymbol{v}_2, \boldsymbol{v}_1, \boldsymbol{w}) + c^\theta(\phi_1, \boldsymbol{v}_2, \rho) + c^\theta(\phi_2, \boldsymbol{v}_1, \rho)\right], \end{aligned}$$

and is then symmetric in $(\boldsymbol{v}_1, \phi_1)$ and $(\boldsymbol{v}_2, \phi_2)$. Note that, if we define the space of discretely divergence-free functions

$$V_h = \{\boldsymbol{v} \in X_h \mid b(\boldsymbol{v}, q) = 0, \forall q \in M_h\},$$

then E_0 is positive definite with respect to $V_h \times M_h \times X_h^\theta$ since

$$E_0((\boldsymbol{v}, q, \phi), (\boldsymbol{v}, q, \phi)) = a(\boldsymbol{v}, \boldsymbol{v}) + a^\theta(\phi, \phi) > 0,$$

for all $\boldsymbol{v} \in V_h, q \in M_h$, and $\phi \in X_h^\theta$ with $(\boldsymbol{v}, \phi) \neq (\boldsymbol{0}, 0)$.

Now, we expand the output, Eq. (7), as

$$S(\bar{\boldsymbol{u}} + \boldsymbol{v}, \bar{p} + q, \bar{\theta} + \phi) - S(\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}) = L((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi)) + M(\boldsymbol{v}, \boldsymbol{v}),$$

where

$$L((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi)) = \ell(\boldsymbol{v}, q, \phi) + m(\bar{\boldsymbol{u}}, \boldsymbol{v}) + m(\boldsymbol{v}, \bar{\boldsymbol{u}}) \tag{13}$$

and $M(\boldsymbol{v}, \boldsymbol{v}) = m(\boldsymbol{v}, \boldsymbol{v})$.

We then define the following primal residual:

$$\mathcal{R}^{\mathrm{pr}}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi)) = -\mathcal{A}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}), (\boldsymbol{v}, q, \phi)). \tag{14}$$

Finally, we define, for $(\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}) \in Y$, the linear dual problem: find the adjoint $(\boldsymbol{\psi}, \lambda, \mu) \in Y$ solution of

$$E((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi), (\boldsymbol{\psi}, \lambda, \mu)) = -L((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi)), \quad \forall (\boldsymbol{v}, q, \phi) \in Y.$$

We also define the generalized dual residual

$$\begin{aligned} &\mathcal{R}^{\mathrm{du}}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{\psi}_0, \lambda_0, \mu_0); (\boldsymbol{\psi}_1, \mu_1); (\boldsymbol{v}, q, \phi)) = -L((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi)) \\ &\quad - E_0((\boldsymbol{v}, q, \phi), (\boldsymbol{\psi}_0, \lambda_0, \mu_0)) - E_1((\bar{\boldsymbol{u}}, \bar{\theta}); (\boldsymbol{v}, \phi), (\boldsymbol{\psi}_1, \mu_1)). \end{aligned} \tag{15}$$

The importance of the dual problem and the associate dual residual will appear shortly. Note that, if we define $\bar{\boldsymbol{e}} = \boldsymbol{u}_h - \bar{\boldsymbol{u}}$, $\bar{\varepsilon} = p_h - \bar{p}$, and $\bar{e}^\theta = \theta_h - \bar{\theta}$, then it follows, from Eqs. (9), (10), and (14), that

$$\mathcal{E}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\bar{\boldsymbol{e}}, \bar{\varepsilon}, \bar{e}^\theta), (\boldsymbol{v}, q, \phi)) = \mathcal{R}^{\mathrm{pr}}((\bar{\boldsymbol{u}}, \bar{p}, \bar{\theta}); (\boldsymbol{v}, q, \phi)), \quad \forall (\boldsymbol{v}, q, \phi) \in Y_h. \tag{16}$$

2.3. *Algorithm*

The bound algorithm proceeds in five steps. The computation is initiated by two global solves on the coarse mesh $\mathcal{T}_H$: one for the initial nonlinear (primal) problem, Step 1, and one for the adjoint linear (dual) problem, Step 2. In Step 3, we compute the hybrid fluxes which will serve as boundary conditions for the local Neumann subproblems in Step 4 after the elimination, by fine mesh projections, of the indefinite terms associated with the incompressibility constraint. In Step 5 we compute the bounds using the "reconstructed errors" obtaine as the solutions to the local subproblems in Step 4.

2.3.1. Step 1. We compute $(\boldsymbol{u}_H, p_H, \theta_H) \in Y_H$ as the solution of the *primal* problem

$$\mathcal{A}((\boldsymbol{u}_H, p_H, \theta_H), (\boldsymbol{v}, q, \phi)) = 0, \quad \forall (\boldsymbol{v}, q, \phi) \in Y_H.$$

In practice, we find $(\boldsymbol{u}_H, p_H, \theta_H) = (\boldsymbol{u}_H^k, p_H^k, \theta_H^k)$ using Newton iteration, which is conveniently expressed as follows:

1. Select an initial guess $(\boldsymbol{u}_H^0, p_H^0, \theta_H^0) \in Y_H$ and set $k = 0$.
2. Find $(\boldsymbol{v}_H, q_H, \phi_H) \in Y_H$ such that

$$E\left((\boldsymbol{u}_H^k, p_H^k, \theta_H^k); (\boldsymbol{v}_H, q_H, \phi_H), (\boldsymbol{w}, r, \rho)\right) \\ = \mathcal{R}^{\mathrm{pr}}\left((\boldsymbol{u}_H^k, p_H^k, \theta_H^k); (\boldsymbol{w}, r, \rho)\right), \quad \forall (\boldsymbol{w}, r, \rho) \in Y_H.$$

Note that the solvability of the divergence-free constraint system is ensured by the homogeneous Dirichlet boundary condition.

3. Set $k = k + 1$ and $(\boldsymbol{u}_H^k, p_H^k, \theta_H^k) = (\boldsymbol{u}_H^{k-1}, p_H^{k-1}, \theta_H^{k-1}) + (\boldsymbol{v}_H, q_H, \phi_H)$; then go to 1 or stop, according to an appropriate stopping criterion.

2.3.2. Step 2. We compute the adjoint $(\psi_H, \lambda_H, \mu_H) \in Y_H$ as the solution of the following *dual* problem:

$$E((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{v}, q, \phi), (\psi_H, \lambda_H, \mu_H)) \\ = -L((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{v}, q, \phi)), \quad \forall (\boldsymbol{v}, q, \phi) \in Y_H. \tag{17}$$

Again, the solvability of the constraint system is ensured by the homogeneous Dirichlet boundary condition and the condition

$$L((\boldsymbol{u}_H, p_H, \theta_H); (\mathbf{0}, 1, 0)) = 0,$$

which follows from our definition of the output, Eq. (8), and from Eq. (13).

2.3.3. Step 3. We compute the primal and dual hybrid fluxes, $(\mathbf{z}, z^\theta)^{\mathrm{pr}} \in Z_H$ and $(\mathbf{z}, z^\theta)^{\mathrm{du}} \in Z_H$, which satisfy the equations

$$B((\boldsymbol{v}, \phi), (\mathbf{z}, z^\theta)^{\mathrm{pr}}) = \mathcal{R}^{\mathrm{pr}}((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{v}, q, \phi)), \quad \forall (\boldsymbol{v}, q, \phi) \in \hat{Y}_H \tag{18}$$

and

$$B((\boldsymbol{v}, \phi), (\mathbf{z}, z^\theta)^{\mathrm{du}}) = \mathcal{R}^{\mathrm{du}}((\boldsymbol{u}_H, p_H, \theta_H); (\psi_H, \lambda_H, \mu_H); (\psi_H, \mu_H); (\boldsymbol{v}, q, \phi)), \\ \forall (\boldsymbol{v}, q, \phi) \in \hat{Y}_H. \tag{19}$$

To find a suitable solution of Eq. (18), we first form the approximation, $(\boldsymbol{z}_0, z_0^\theta)^{\text{pr}} \in Z_H$,

$$z^{\text{pr}}_{0i|\gamma} = -\frac{1}{2}\left(\tau^{\text{pr}}_{ij|\gamma} n^+_{\gamma j} - \tau^{\text{pr}}_{ij|\gamma} n^-_{\gamma j}\right),$$
$$z^{\theta,\text{pr}}_{0\,|\,\gamma} = -\frac{1}{2}\left(\tau^{\theta,\text{pr}}_{j|\gamma} n^+_{\gamma j} - \tau^{\theta,\text{pr}}_{j|\gamma} n^-_{\gamma j}\right),$$

where $\boldsymbol{n}^+_\gamma$ (respectively $\boldsymbol{n}^-_\gamma$) denote the outward normal from the element on the arbitrarily chosen "positive" (respectively "negative") side of the edge γ, and

$$\tau^{\text{pr}}_{ij} = \left(\frac{\partial u_{Hi}}{\partial x_j} + \frac{\partial u_{Hj}}{\partial x_i}\right) - p_H \delta_{ij} - u_{Hi} u_{Hj},$$
$$\tau^{\theta,\text{pr}}_{j} = \frac{1}{\alpha}\frac{\partial \theta_H}{\partial x_j} - u_{Hj}\theta_H.$$

Note that, since $u_{Hi|\gamma} \in \mathbb{P}_2(\gamma)$, the approximation space for the hybrid fluxes includes quartic polynomials, $\mathbb{P}_4(\gamma)$, to correctly represent $z^{\text{pr}}_{0i|\gamma}$ and $z^{\theta,\text{pr}}_{0|\gamma}$. An alternative approach is to interpolate the quartic term onto $\mathbb{P}_2$, and thereby use only $\mathbb{P}_2$ hybrid fluxes. For $(z_0, z_0^\theta)^{\text{du}}$, we obtain similar expressions,

$$\tau^{\text{du}}_{i,j} = \left(\frac{\partial \psi_{Hi}}{\partial x_j} + \frac{\partial \psi_{Hj}}{\partial x_i}\right) - \lambda_H \delta_{ij},$$
$$\tau^{\theta,\text{du}}_{j} = \frac{1}{\alpha}\frac{\partial \mu_H}{\partial x_j}.$$

Now, to satisfy Eqs. (18) and (19), we correct these initial approximations as follows,

$$(\boldsymbol{z}, z^\theta)^{\text{pr}} = \left(\boldsymbol{z}_0, z_0^\theta\right)^{\text{pr}} + \left(\boldsymbol{z}_l, z_l^\theta\right)^{\text{pr}} + \left(\boldsymbol{z}_q, z_q^\theta\right)^{\text{pr}},$$
$$(\boldsymbol{z}, z^\theta)^{\text{du}} = \left(\boldsymbol{z}_0, z_0^\theta\right)^{\text{du}} + \left(\boldsymbol{z}_l, z_l^\theta\right)^{\text{du}} + \left(\boldsymbol{z}_q, z_q^\theta\right)^{\text{du}},$$

where $(\boldsymbol{z}_l, z_l^\theta)^{\text{pr}}$ and $(\boldsymbol{z}_l, z_l^\theta)^{\text{du}}$ are $\mathbb{P}_1$ corrections, and $(\boldsymbol{z}_q, z_q^\theta)^{\text{pr}}$ and $(\boldsymbol{z}_q, z_q^\theta)$ are $\mathbb{P}_2$ corrections. To obtain these corrections, we adapt a procedure developed in the context of energy-norm implicit Neumann subproblem indicators [2, 7, 10].

2.3.4. Step 4. We compute the "incompressible local projections": $(\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H) \in Y_h$ for the primal problem and $(\tilde{\psi}_H, \tilde{\lambda}_H, \tilde{\mu}_H) \in Y_h$ for the dual problem. These projections are constructed to satisfy the following equations on the fine mesh,

$$b(\tilde{\boldsymbol{u}}_H, q) = 0, \quad \forall q \in M_h, \tag{20}$$
$$b(\tilde{\psi}_H, q) = -L((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{0}, q, 0)), \quad \forall q \in M_h, \tag{21}$$

which eliminate the indefinite terms associated with the incompressibility constraint. We choose $\tilde{\theta}_H = \theta_H$, and, to satisfy Eq. (20), we write $(\tilde{\boldsymbol{u}}_H, \tilde{p}_H) = (\boldsymbol{u}_H, p_H) + (\boldsymbol{\Delta}_h, \delta_h)$, where, for all $T_H \in \mathcal{T}_H$, we compute $(\boldsymbol{\Delta}_{h|T_H}, \delta_{h|T_H}) \in X_{T_H} \times M_{T_H}$ as the solution of

$$a_{T_H}\left(\boldsymbol{\Delta}_{h|T_H}, \boldsymbol{v}\right) + b_{T_H}\left(\boldsymbol{v}, \delta_{h|T_H}\right) = 0, \quad \forall \boldsymbol{v} \in X_{T_H},$$
$$b_{T_H}\left(\boldsymbol{\Delta}_{h|T_H}, q\right) = -b\left(\boldsymbol{u}_{H|T_H}, q\right), \quad \forall q \in M_{T_H},$$

with

$$a_{T_H}(\boldsymbol{v}, \boldsymbol{w}) = \int_{T_H} \frac{\partial v_i}{\partial x_j} \frac{\partial w_i}{\partial x_j}\, d\Omega, \quad b_{T_H}(\boldsymbol{v}, r) = -\int_{T_H} \frac{\partial v_i}{\partial x_i} r\, d\Omega,$$

and

$$X_{T_H} = \left\{\boldsymbol{v} \in H_0^1(T_H) \times H_0^1(T_H) \mid v_{i|T_h} \in \mathbb{P}_2(T_h) \oplus B_3(T_h), \forall T_h \in \mathcal{T}_h\right\},$$
$$M_{T_H} = \left\{p \in L^2(T_H) \mid p_{|T_h} \in \mathbb{P}_1(T_h), \forall T_h \in \mathcal{T}_h\right\}.$$

This problem is solvable since $b(\boldsymbol{u}_{H|T_H}, 1) = b(\boldsymbol{u}_H, 1_{|T_H}) = 0$ for all $T_H \in \mathcal{T}_H$, thanks to the discontinuous pressure space.

Similarly, we choose $\tilde{\mu}_H = \mu_H$ and, to satisfy Eq. (21), we write $(\tilde{\boldsymbol{\psi}}_H, \tilde{\mu}_H) = (\boldsymbol{\psi}_H, \mu_H) + (\boldsymbol{\Delta}_h, \delta_h)$, where, for all $T_H \in \mathcal{T}_H$, we compute $(\boldsymbol{\Delta}_{h|T_H}, \delta_{h|T_H}) \in X_{T_H} \times M_{T_H}$ as the solution of

$$a_{T_H}\left(\boldsymbol{\Delta}_{h|T_H}, \boldsymbol{v}\right) + b_{T_H}\left(\boldsymbol{v}, \delta_{h|T_H}\right) = 0, \quad \forall \boldsymbol{v} \in X_{T_H},$$
$$b_{T_H}\left(\boldsymbol{\Delta}_{h|T_H}, q\right) = -b\left(\boldsymbol{\psi}_{H|T_H}, q\right) - L((\boldsymbol{v}_H, p_H, \theta_H); (\boldsymbol{0}, q, 0)), \quad \forall q \in M_{T_H}.$$

These equations form a solvable system since, from Eq. (17),

$$b\left(\boldsymbol{\psi}_H, 1_{|T_H}\right) = -L((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{0}, 1_{|T_H}, 0)).$$

A priori estimates for the projections $\tilde{\boldsymbol{u}}_H$ and $\tilde{\boldsymbol{\psi}}_H$ are derived in Appendix 1.

We now compute the primal and dual "reconstructed errors," $(\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\text{pr}} \in \hat{Y}_h$ and $(\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\text{du}} \in \hat{Y}_h$, which satisfy the following equations:

$$2E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\text{pr}}, (\boldsymbol{v}, q, \phi)) = \mathcal{R}^{\text{pr}}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{v}, q, \phi)) - B((\boldsymbol{v}, \phi), (\mathbf{z}, z^\theta)^{\text{pr}}), \quad \forall(\boldsymbol{v}, q, \phi) \in \hat{Y}_h, \tag{22}$$

and

$$2E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\text{du}}, (\boldsymbol{v}, q, \phi)) = \mathcal{R}^{\text{du}}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\boldsymbol{\psi}_H, \mu_H); (\boldsymbol{v}, q, \phi)) - B((\boldsymbol{v}, \phi), (\mathbf{z}, z^\theta)^{\text{du}}), \quad \forall(\boldsymbol{v}, q, \phi) \in \hat{Y}_h. \tag{23}$$

According to our definition of the space $\hat{Y}_h$, Eqs. (22) and (23) correspond to local (elemental) decoupled Neumann problems; the solvability of these subproblems is discussed in Appendix 2. Note that both the incompressible projection and the Neumann subproblems are local and linear and, therefore, inexpensive to compute compared to the global fine mesh problem, Eq. (9). This issue is further discussed in Section 4.

2.3.5. Step 5. Finally, we compute the bounds, $s_H^\pm$, according to

$$s_H^\pm = S(\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H) - \mathcal{R}^{\text{pr}}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\mu}_H, \tilde{\lambda}_H)) \pm \kappa^u E_0((\hat{\boldsymbol{e}}^\pm, \hat{\varepsilon}^\pm, 0), (\hat{\boldsymbol{e}}^\pm, \hat{\varepsilon}^\pm, 0)) \pm \kappa^\theta E_0((\boldsymbol{0}, 0, \hat{e}^{\theta\pm}), (\boldsymbol{0}, 0, \hat{e}^{\theta\pm})), \tag{24}$$

where

$$(\hat{\boldsymbol{e}}^\pm, \hat{\varepsilon}^\pm, \hat{e}^{\theta,\pm}) = (\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\text{pr}} \mp \frac{1}{\kappa^u}(\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\text{du}} \mp \frac{1}{\kappa^\theta}(\boldsymbol{0}, 0, \hat{e}^\theta)^{\text{du}}. \tag{25}$$

In this expression, κ^u and κ^θ are two strictly positive real number. The choice of these parameters will affect the half bound gap

$$\begin{aligned}\Delta_H &\equiv \frac{1}{2}(S_H^+ - S_H^-) \\ &= \kappa^u E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{pr}}, (\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{pr}}) + \frac{1}{\kappa^u} E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{du}}, (\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{du}}) \\ &\quad + \kappa^\theta E_0((\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{pr}}, (\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{pr}}) + \frac{1}{\kappa^\theta} E_0((\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{du}}, (\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{du}}).\end{aligned}$$

Since $(\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\mathrm{pr}}$ and $(\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\mathrm{du}}$ do not depend on the choice of κ^u and κ^θ, we can readily find κ^u and κ^θ that minimize the bound gap, and hence render the lower and upper bound as sharp as possible. We find

$$\kappa^u = \sqrt{\frac{E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{du}}, (\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{du}})}{E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{pr}}, (\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\mathrm{pr}})}},$$

$$\kappa^\theta = \sqrt{\frac{E_0((\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{du}}, (\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{du}}}{E_0((\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{pr}}, (\mathbf{0}, 0, \hat{e}^\theta)^{\mathrm{pr}})}}.$$

2.4. *Bounding Properties*

If we define the errors $\tilde{e} = \boldsymbol{u}_h - \tilde{\boldsymbol{u}}_H$, $\tilde{\varepsilon} = p_h - \tilde{p}_H$, and $\tilde{e}^\theta = \theta_h - \tilde{\theta}_H$, we can derive the bound error expression

$$s_H^\pm = s_h \pm D^\pm + I^\pm, \tag{26}$$

where

$$D^\pm = \kappa^u \mathcal{F}_0(\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm, \tilde{\varepsilon} - \hat{\varepsilon}^\pm, 0) + \kappa^\theta \mathcal{F}_0(\mathbf{0}, 0, \tilde{e}^\theta - \hat{e}^{\theta,\pm}),$$

with $\mathcal{F}_0(\boldsymbol{v}, q, \phi) = E_0((\boldsymbol{v}, q, \phi), (\boldsymbol{v}, q, \phi))$. Recall that $s_h = S(\boldsymbol{u}_h, p_h, \theta_h)$ is the fine mesh (truth) output. The derivation of the bound error expression and the definition of $I^\pm$ are given in Appendix 3.

We now show that $D^\pm \geq 0$. We have $\mathcal{F}_0(\mathbf{0}, 0, \tilde{e}^\theta - \hat{e}^{\theta\pm}) \geq 0$ since, for all $\phi \in X_h^\theta$, $\mathcal{F}_0(\mathbf{0}, 0, \phi) = a^\theta(\phi, \phi) \geq 0$. Moreover, using the definition of E_0, Eq. (12), we can write

$$\mathcal{F}_0(\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm, \tilde{\varepsilon} - \hat{\varepsilon}^\pm, 0) = a(\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm, \tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm) + 2b(\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm, \tilde{\varepsilon} - \hat{\varepsilon}^\pm).$$

For all $q \in M_h$, we have, from (22) and (23), taking $(\boldsymbol{v}, q, \phi) = (\mathbf{0}, q, 0)$,

$$2\kappa^u b(\hat{\boldsymbol{e}}^\pm, q) = -\kappa^u b(\tilde{\boldsymbol{u}}_H, q) \pm [b(\tilde{\psi}_H, q) + L((\boldsymbol{u}_H, p_H, \theta_H); (\mathbf{0}, q, 0))],$$

and, by virtue of Eqs. (20) and (21), we find that $2\kappa^u b(\hat{\boldsymbol{e}}^\pm, q) = 0$ and $b(\tilde{\boldsymbol{e}}, q) = b(\boldsymbol{u}_h - \tilde{\boldsymbol{u}}_H, q) = 0, \forall q \in M_h$; therefore,

$$\mathcal{F}_0(\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm, \tilde{\varepsilon} - \hat{\varepsilon}^\pm, 0) = a(\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm, \tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm) \geq 0.$$

We now examine the bound error expression, Eq. (26), and we see that the upper bound (respectively, the lower bound) departs from s_h, by a positive (respectively negative) definite contribution D^+ (respectively D^-) and an indefinite contribution I^+ (respectively I^-). The essential point is that $D^\pm$ involves gradients only and is expected to converge at best like H^4 [10]. More precisely, we assume that there is a positive constant, $C_D > 0$, such that $C_D H^4 \leq D^\pm$—a variant of the method, based on a hypothesis easier to verify, is introduced in [16]. As argued in Appendix 4, the indefinite contribution $I^\pm$ involves only weaker norms and converges like H^5: there exists a positive constant C_I such that $I^\pm \leq C_I H^5$. Therefore, if we define $H^* = C_D/C_I$, according to Eq. (26), we obtain bounds for all $H \leq H^*$. The numerical examples presented hereafter will show that in practice bounds are always obtained provided the mesh even barely resolves the flow. A similar argument has been used successfully in earlier applications of the bound method to a variety of other noncoercive or nonlinear problems [8, 11]. This issue will be further discussed in Section 4.

3. NUMERICAL EXAMPLES

We first present numerical results for natural convection in a square domain. The problem is defined by Eqs. (1), (2), and (3), the geometry of the problem is shown in Fig. 2a, and the boundary conditions are $\boldsymbol{u}|_{\partial\Omega} = 0$, $\theta = 0$ on Γ_0, $\frac{\partial\theta}{\partial n} = 1$ on Γ_I, and $\frac{\partial\theta}{\partial n} = 0$ otherwise. The Prandtl number and the Grashof number are $\alpha = 1$ and $\beta = 30{,}000$, respectively. For this choice of parameters, the Nusselt number is $\mathrm{Nu} = 2.49$ ($\mathrm{Nu}^{-1} = \frac{1}{|\Gamma_I|}\int_{\Gamma_I} \theta_h \, d\Gamma$, computed with the fine mesh solution); for a pure conduction problem in this geometry, $\mathrm{Nu} = 1$, and

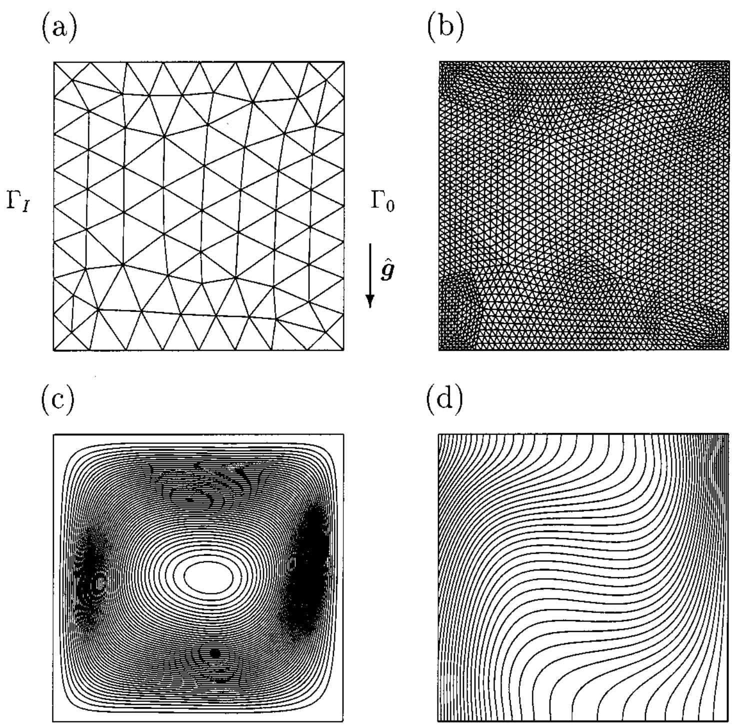

FIG. 2. (a) Domain and coarse mesh, $\mathcal{T}_{H_0}$. (b) Fine "truth" mesh, $\mathcal{T}_h$. (c) Streamlines. (d) Isotherms.

thus Nu $= 2.49$ indicates the existence of velocity and thermal boundary layers near Γ_I and Γ_0. The elements of the coarsest mesh, $\mathcal{T}_{H_0}$, are chosen so that the diameter of the mesh, H_0, is approximately the size of the thickness of the boundary layers. Figures 2b, 2c, and 2d represent the fine mesh $\mathcal{T}_h$, the streamlines, and the isotherms obtained on the fine mesh, respectively.

Figures 3a and 3b show the convergence of the bounds for the nonlinear output,

$$S_1(\boldsymbol{u}, p, \theta) = \int_\Omega |\boldsymbol{u}|^2 \, d\Omega,$$

corresponding to the kinetic energy of the flow. In Fig. 3a, we have represented the normalized lower and upper bounds, s_H^+/s_h and s_H^-/s_h, and the normalized coarse mesh output, s_H/s_h, for different meshes $\mathcal{T}_H$. The mesh diameters vary from $H = H_0$ to $H = H_0/6$; $\mathcal{T}_h = \mathcal{T}_{H_0/6}$ is the fine mesh. For this problem, bounds are obtained on all the meshes considered. Figure 3b demonstrates that the convergence rate of the bound gap is optimal, $O(H^4)$; the measured effectivity factor η, defined as $\eta = \Delta_H/|s_h - s_H|$, is $7.3 \leq \eta \leq 7.7$.

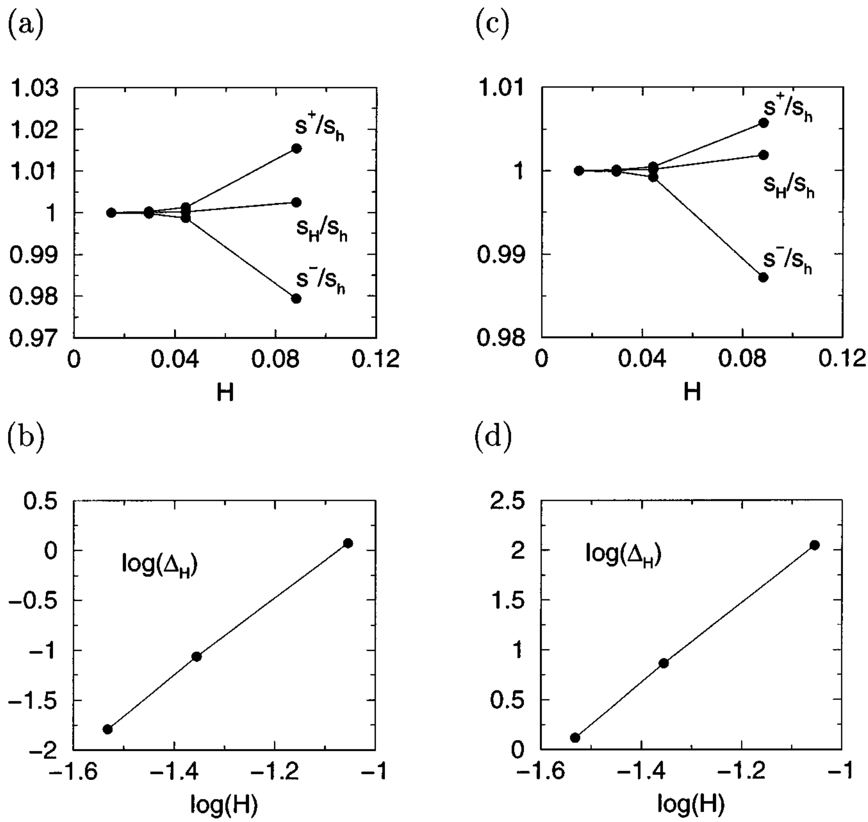

FIG. 3. Convergence of the (normalized) bounds $s_H^\pm/s_h$, for (a) the kinetic energy, S_1, and (c) the mean temperature on the heated boundary, S_2; (b) and (d) show the convergence of the bound gap, Δ_H, for S_1 and S_2, respectively.

Similar results are obtained for the output

$$S_2(\boldsymbol{u}, p, \theta) = \frac{1}{|\Gamma_I|} \int_{\Gamma I} \theta \, d\Gamma.$$

Figures 3c and 3d show the convergence of the bounds and of the bound gap; again optimality is achieved. For this output the effectivity factor is $3.6 \leq \eta \leq 5$.

We next present results obtained for the model problem defined by Eqs. (1), (2) and (3), the boundary conditions (4), and the domain represented in Fig. 1a. The output of interest is

$$S(\boldsymbol{u}, p, \theta) = \frac{1}{|\Gamma_1|} \int_{\Gamma_1} \theta \, d\Gamma.$$

For this case the (nondimensional) heat flux is $q = 1$, and the Prandtl and Grashof numbers are $\alpha = 1$ and $\beta = 50{,}000$; the resulting streamlines are shown in Fig. 4a.

Since the cost of computing the bounds is essentially a function of the number of elements T_H in $\mathcal{T}_H$, it is desirable to construct optimized triangulations that maximize the bound accuracy (minimize the bound gap) for a given number of degrees of freedom. As shown in [17], the bound gap, Δ_H, can be expressed as a sum of local elemental positive contributions: $\Delta_H = \sum_{T_H \in \mathcal{T}_H} \Delta_{T_H}$, with $\Delta_{T_H} \geq 0$. We can, therefore, implement a simple adaptive strategy. Starting from an initial grid, $\mathcal{T}_H^0$, with bound gap Δ_H^0, we generate a sequence of triangulations $\{\mathcal{T}_H^k, k = 1, 2, \ldots\}$ with corresponding bound gaps $\{\Delta_H^k, k = 1, 2, \ldots\}$, such that each triangulation $\mathcal{T}_H^k$ is obtained by refinement of the selected triangles T_H^{k-1} with $\Delta_{T_H}^{k-1} > \alpha \max_{T_H} \Delta_{T_H}^{k-1}$, for a specified parameter $0 < \alpha < 1$. The approach ensures that, for a sufficiently large k, $\Delta_H^k \leq \Delta_{\text{targ}}$, where $\Delta_{\text{targ}} > 0$ is a gap target.

In practice, the refined mesh is obtained by dividing each selected triangle into four self-similar triangles (1 : 4 division). The adjacent elements are divided into two elements

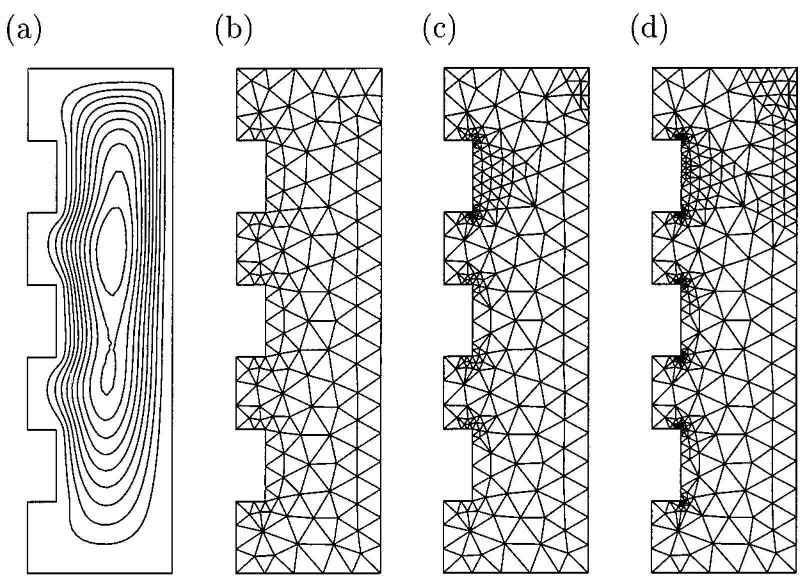

FIG. 4. Model problem. (a) Streamlines. (b) Mesh $\mathcal{T}_H^0$. (c) Mesh $\mathcal{T}_H^{k_1}$. (d) Mesh $\mathcal{T}_H^{k_2}$.

TABLE I
Convergence of the Bounds for the Problem of Fig. 1a

	$\mathcal{T}_H$	$\mathcal{T}_H^{k_1}$	$\mathcal{T}_H^{k_2}$	$\mathcal{T}_{H/2}$
Number of elements	264	417	680	1056
s^-	0.253	0.270	0.274	0.272
s^+	0.297	0.282	0.279	0.280
Δ_H	0.0223	0.0058	0.0025	0.0037

if necessary to avoid any hanging nodes (1 : 2 division). To preserve shape regularity, we do not allow the 1 : 2 divided elements to be divided at a subsequent level of refinement—should further division be required, we first produce a 1 : 4 division of the original element, and then proceed with the required subsequent divisions.

The effectiveness of the adaptive procedure is summarized in Table I, in which $\mathcal{T}_H^0$, $\mathcal{T}_H^{k_1}$, and $\mathcal{T}_H^{k_2}$ denote successive adapted meshes corresponding to Figs. 4b, 4c, and 4d, respectively; the mesh $\mathcal{T}_{H/2}$ is a *uniform* refinement of $\mathcal{T}_H^0$. Table I shows that a reduction of the relative bound gap from 8% for mesh $\mathcal{T}_H^0$ to less than 1% for mesh $\mathcal{T}_H^{k_2}$ is achieved by optimal refinement with a final mesh which contains only slightly more than twice the number of elements of the mesh $\mathcal{T}_H^0$. Note that, for each of the adapted coarse meshes, the truth fine mesh is obtained by dividing each element of the adapted coarse mesh into 6^2 elements, implying that the truth mesh is also adapted; nevertheless, our choice of a conservative initial fine mesh guarantees that the truth solution is insensitive to the adaptation of the fine mesh.

4. ATTRIBUTES OF THE METHOD

We review here the four attributes of the method defined in the Introduction.

4.1. *Attribute A1*

To summarize the previous discussion, we have shown that bounds are obtained for all mesh diameters $H \leq H^*$, where H^* is an unknown threshold. Note that we expect the singular perturbation prefactors, $1/\alpha$ and $1/\beta$ (recall that α and β are the Prandtl and the Grashof number, respectively), to unfortunately favor the indefinite terms; nevertheless, the corresponding *numerical* singular perturbation parameter will be small as soon as the coarse space even roughly resolves the structure of the solution, and thus the bound property will be preserved except perhaps on very crude meshes.

We conclude that, even if H^* is not known *a priori*, bounds are obtained once the solution is marginally resolved. In numerical examples, we have always observed bounds, and thus the uncertainty associated with H^* is not an important practical issue, although it constitutes a real theoretical issue. Note that H^* is a threshold parameter (bounds are obtained for all $H \leq H^*$); therefore, even if H^* is not known, the method presents a significant advantage over previous explicit *a posteriori* estimation procedures in which the estimators themselves involved unknown constants and functions. Note also that, for linear coercive problems (and the Stokes problem), bounds are obtained for any H ($H^* = \infty$) [13, 15, 17].

There is another source of uncertainty, namely, the choice of the fine mesh $\mathcal{T}_h$, which should be selected fine enough to ensure that $|s - s_h|$ is negligible (recall that s is the exact solution). In practice, the mesh $\mathcal{T}_h$ is chosen conservatively by estimating *a priori* the size of the smallest structures anticipated and by ensuring that $\mathcal{T}_h$ provides for an extremely accurate representation of such structures. Note that the work to compute the bounds is not overly sensitive to the fineness of the truth mesh (see Section 4.4).

4.2. *Attribute A2*

The bounds are sharp. As demonstrated by the numerical results, the bound gap converges at an optimal rate; for sufficiently regular problems we have observed that $\Delta_H \leq CH^4$. The effectivity factor, $\eta = \Delta_H/|s_h - s_H|$, has also been measured and typically $1 \leq \eta \leq 10$, though this will be problem— and output—dependent.

4.3. *Attribute A3*

We have used the local decomposition of the bound gap into elemental positive contributions to implement an adaptive refinement strategy. We have shown that a straightforward implementation yields a very efficient procedure which allows us to obtain an error for the output below a specified threshold with an “optimal” (small) number of elements.

4.4. *Attribute A4*

We recall that $(\boldsymbol{u}_h, p_h, \theta_h)$ and $s_h = S(\boldsymbol{u}_h, p_h, \theta_h)$ are the field variables and the output that are effectively indistinguishable from the exact field solution and output and correspondingly expensive. Our lower and upper bounds are only interesting if they can be obtained at a considerably lower expense than the computation of s_h—and preferably at only a slightly higher expense than the computation of s_H, the coarse approximation. From our definitions of Y_h and $\hat{Y}_h$, we see that Eqs. (22) and (23) correspond to many small, local, linear Neumann subproblems, whereas Eq. (9) corresponds to a single large, global, nonlinear problem. It follows that Eqs. (22) and (23) present a smaller bandwidth—yielding substantial savings in both memory and computational time of direct solution strategies—and a smaller condition number—yielding faster convergence in iterative methods. In addition, Eqs. (22) and (23) are symmetric, positive semi-definite, and completely decoupled—the last making straightforward parallel implementations possible. Furthermore, the dual problem, Eq. (17), and the subproblems, Eqs. (22) and (23), are linear, leading to additional savings compared to the original nonlinear fine-mesh problem. The relative savings are, of course, reduced as the global solver improves.

APPENDIX 1. INCOMPRESSIBLE PROJECTION

Recall that $(\tilde{\boldsymbol{u}}_H, \tilde{p}_H) = (\boldsymbol{u}_H, p_H) + (\boldsymbol{\Delta}_h, \delta_H)$, where for all $T_H \in \mathcal{T}_H$, we have $(\boldsymbol{\Delta}_{h|T_H}, \delta_{h|T_H}) \in X_{T_H} \times M_{T_H}$, which satisfies

$$a_{T_H}(\boldsymbol{\Delta}_{h|T_H}, \boldsymbol{v}) + b_{T_H}(\boldsymbol{v}, \delta_{h|T_H}) = 0, \quad \forall \boldsymbol{v} \in X_{T_H}, \tag{27}$$

$$b_{T_H}(\boldsymbol{\Delta}_{h|T_H}, q) = -b(\boldsymbol{u}_{H|T_H}, q), \quad \forall q \in M_{T_H}. \tag{28}$$

This problem is solvable since $b(\boldsymbol{u}_{H|T_H}, 1) = b(\boldsymbol{u}_H, 1_{|T_H}) = 0$ for all $T_H \in \mathcal{T}_H$.

Given any $T_H \in \mathcal{T}_H$, we first note that the inf-sup parameter, β_{T_H}, for Eqs. (27) and (28) is independent of h and H (β_{T_H} depends only on the shape of T_H and not on its size); the latter follows from the scale invariance of β_{T_H} when defined with respect to the $H^1(T_H)$ seminorm (which is equivalent to the $H^1(T_H)$ norm, as we have Dirichlet conditions on

∂T_H). Therefore, $\exists \boldsymbol{v} \in X_{T_H}$ such that

$$\beta_{T_H} |\boldsymbol{v}|_{H^1(T_H)} \|\delta_h\|_{L^2(T_H)} \leq |b_{T_H}(\boldsymbol{v}, \delta_{h|T_H})| = |a_{T_H}(\boldsymbol{\Delta}_{h|T_H}, \boldsymbol{v})| \leq |\boldsymbol{\Delta}_h|_{H^1(T_H)} |\boldsymbol{v}|_{H^1(T_H)},$$

and thus

$$\|\delta_h\|_{L^2(T_H)} \leq \frac{1}{\beta_{T_H}} |\boldsymbol{\Delta}_h|_{H^1(T_H)}.$$

We then note that

$$\begin{aligned} a_{T_H}(\boldsymbol{\Delta}_h, \boldsymbol{\Delta}_h) = |\boldsymbol{\Delta}_h|^2_{H^1(T_H)} &= -b_{T_H}(\boldsymbol{\Delta}_h, \delta_h) = b(\boldsymbol{u}_{H|T_H}, \delta_h) \\ &= -b(\boldsymbol{u}_{h|T_H} - \boldsymbol{u}_{H|T_H}, \delta_h) \\ &\leq C |\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(T_H)} \|\delta_h\|_{L^2(T_H)} \\ &\leq \frac{C}{\beta_{T_H}} |\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(T_H)} |\boldsymbol{\Delta}_h|_{H^1(T_H)}. \end{aligned}$$

Therefore, we have, for all $T_H \in \mathcal{T}_H$,

$$|\boldsymbol{\Delta}_h|_{H^1(T_H)} \leq \frac{C}{\beta_{T_H}} |\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(T_H)};$$

moreover, since $\boldsymbol{\Delta}_{h|\partial T_H} = 0$, we can apply the Poincaré inequality in conjunction with a scaling argument to find

$$\|\boldsymbol{\Delta}_h\|_{L^2(T_H)} \leq \frac{C}{\beta_{T_H}} H |\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(T_H)}.$$

Finally, squaring and summing over all $T_H \in \mathcal{T}_H$ yields the estimates

$$\begin{aligned} |\tilde{\boldsymbol{u}}_H - \boldsymbol{u}_H|_{H^1(\Omega)} &\leq C |\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(\Omega)}, \\ \|\tilde{\boldsymbol{u}}_H - \boldsymbol{u}_H\|_{L^2(\Omega)} &\leq CH |\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(\Omega)}. \end{aligned}$$

The procedure for $(\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H)$ is identical except that the right-hand side in Eq. (28) is now $-b(\boldsymbol{\psi}_H \mid_{T_H}, q) - L((\boldsymbol{u}_H, p_H, \tilde{\theta}_H); (\mathbf{0}, q, 0))$. If we denote the adjoint obtained on the fine mesh by $\boldsymbol{\psi}_h$, then we have

$$b(\boldsymbol{\psi}_{h|_{T_H}}, q) = -L((\boldsymbol{u}_H, p_H, \tilde{\theta}_H); (\mathbf{0}, q, 0)), \quad \forall q \in M_{T_H};$$

therefore,

$$a_{T_H}(\boldsymbol{\Delta}_h, \boldsymbol{\Delta}_h) = -b_{T_H}(\boldsymbol{\psi}_h - \boldsymbol{\psi}_H, \delta_h),$$

and the rest of the proof follows. We thus obtain the estimates

$$\begin{aligned} |\tilde{\boldsymbol{\psi}}_H - \boldsymbol{\psi}_H|_{H^1(\Omega)} &\leq C |\boldsymbol{\psi}_h - \boldsymbol{\psi}_H|_{H^1(\Omega)}, \\ \|\tilde{\boldsymbol{\psi}}_H - \boldsymbol{\psi}_H\|_{L^2(\Omega)} &\leq CH |\boldsymbol{\psi}_h - \boldsymbol{\psi}_H|_{H^1(\Omega)}. \end{aligned}$$

APPENDIX 2. SOLVABILITY OF THE LOCAL NEUMANN SUBPROBLEMS

To ensure solvability of Eqs. (22) and (23), we must prove

$$\mathcal{R}^{\mathrm{pr}}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{v}^s, q^s, \phi^s)) = B((\boldsymbol{v}^s, \phi^s), (\boldsymbol{z}, z^\theta)^{\mathrm{pr}})$$

and

$$\mathcal{R}^{\mathrm{du}}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\psi}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\psi_H, \mu_H); (\boldsymbol{v}^s, q^s, \phi^s)) = B((\boldsymbol{v}^s, \phi^s), (\boldsymbol{z}, z^\theta)^{\mathrm{du}}),$$

for all the singular modes $(\boldsymbol{v}^s, q^s, \phi^s) \in \hat{Y}_h$ defined by

$$E_0((\boldsymbol{w}, r, \rho), (\boldsymbol{v}^s, q^s, \phi^s)) = 0, \quad \forall (\boldsymbol{w}, r, \rho) \in \hat{Y}_h.$$

The velocity singular modes are $\boldsymbol{v}^s = (1_{|T_H}, 0), (0, 1_{|T_H}), (y_{|T_H}, -x_{|T_H})$, for all $T_H \in \mathcal{T}_H$, corresponding respectively to (elemental) translations in the x direction, translations in the y direction, and rotations. We also have the temperature modes $\phi^s = 1_{|T_H}$ and the pressure modes $q^s = 1_{|T_H}$. Since all these modes are in $\hat{Y}_H$, in view of Eqs. (18) and (19), we have to prove that

$$\mathcal{R}^{\mathrm{pr}}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{v}^s, q^s, \phi^s)) = \mathcal{R}^{\mathrm{pr}}((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{v}^s, q^s, \phi^s)) \tag{29}$$

and

$$\begin{aligned} &\mathcal{R}^{\mathrm{du}}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\psi}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\psi_H, \mu_H); (\boldsymbol{v}^s, q^s, \phi^s)) \\ &\quad = \mathcal{R}^{\mathrm{du}}((\boldsymbol{u}_H, p_H, \theta_H); (\psi_H, \lambda_H, \mu_H); (\psi_H, \mu_H); (\boldsymbol{v}^s, q^s, \phi^s)). \end{aligned} \tag{30}$$

Looking first at Eq. (29), by the definition of $\mathcal{R}^{\mathrm{pr}}$ in Eq. (14),

$$\begin{aligned} \mathcal{R}^{\mathrm{pr}}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{v}^s, 0, 0)) &= -a(\tilde{\boldsymbol{u}}_H, \boldsymbol{v}^s) - b(\boldsymbol{v}^s, \tilde{p}_H) - c(\tilde{\boldsymbol{u}}_H, \tilde{\boldsymbol{u}}_H, \boldsymbol{v}^s) - d(\tilde{\theta}_H, \boldsymbol{v}^s) \\ &= -c(\tilde{\boldsymbol{u}}_H, \tilde{\boldsymbol{u}}_H, \boldsymbol{v}^s) - d(\tilde{\theta}_H, \boldsymbol{v}^s). \end{aligned}$$

Using the definition of c (note the importance of the conservative form), and since the support of $\boldsymbol{v}^s$ is an element $T_H \in \mathcal{T}_H$, we can write

$$c(\tilde{\boldsymbol{u}}_H, \tilde{\boldsymbol{u}}_H, \boldsymbol{v}^s) = -\int_{T_H} \tilde{u}_{Hi}\tilde{u}_{Hj} \frac{\partial v_i^s}{\partial x_j}\, d\Omega = -\frac{1}{2}\int_{T_H} \tilde{u}_{Hi}\tilde{u}_{Hj} \left(\frac{\partial v_i^s}{\partial x_j} + \frac{\partial v_j^s}{\partial x_i}\right) d\Omega = 0,$$

and since the same argument applies to $c(\boldsymbol{u}_H, \boldsymbol{u}_H, \boldsymbol{v}^s)$ and $\tilde{\theta}_H = \theta_H$, we have

$$\mathcal{R}^{\mathrm{pr}}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{v}^s, 0, 0)) = \mathcal{R}^{\mathrm{pr}}((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{v}^s, 0, 0)).$$

We now consider

$$\begin{aligned} \mathcal{R}^{\mathrm{pr}}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{0}, 0, \phi^s)) &= -a^\theta(\tilde{\theta}_H, \phi^s) - c^\theta(\tilde{\theta}_H, \tilde{\boldsymbol{u}}_H, \phi^s) - \langle g_N, \phi^s\rangle_N \\ &= -\langle g_N, \phi^s\rangle_N \\ &= \mathcal{R}^{\mathrm{pr}}((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{0}, 0, \phi^s)). \end{aligned}$$

Then, since for $q^s = 1_{|T_H}$,

$$\mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{0}, q^s, 0)) = 0 = \mathcal{R}^{\rm pr}((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{0}, q^s, 0)),$$

Eq. (29) is satisfied.

Turning now to the dual problem, the definition of $\mathcal{R}^{\rm du}$, Eq. (15), gives

$$\begin{aligned}
&\mathcal{R}^{\rm du}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\boldsymbol{\psi}_H, \mu_H); (\boldsymbol{v}^s, q^s, \phi^s))\\
&\quad = -L((\boldsymbol{u}_H, p_H, \theta_H); (\boldsymbol{v}^s, q^s, \phi^s)) - a(\tilde{\boldsymbol{\psi}}_H, \boldsymbol{v}^s) - a^\theta(\tilde{\mu}_h, \phi^s)\\
&\qquad - E_1((\boldsymbol{u}_H, \theta_H); (\boldsymbol{v}^s, \phi^s), (\boldsymbol{\psi}_H, \mu_H)).
\end{aligned}$$

Since $a(\tilde{\boldsymbol{\psi}}_H, \boldsymbol{v}^s) = 0$ and $a(\boldsymbol{\psi}_H, \boldsymbol{v}^s) = 0$, and since $a^\theta(\tilde{\mu}_h, \phi^s) = 0$ and $a^\theta(\mu_h, \phi^s) = 0$, Eq. (30) is satisfied. Note that, thanks to our definition of the generalized dual residual, Eq. (15), the last argument of E_1 is $(\boldsymbol{\psi}_H, \mu_H)$ and not $(\tilde{\boldsymbol{\psi}}_H, \tilde{\mu}_H)$ in the above expression; this is essential to ensure the solvability of the dual subproblems.

APPENDIX 3. BOUND ERROR EXPRESSION

We define the errors $\tilde{\boldsymbol{e}} = \boldsymbol{u}_h - \tilde{\boldsymbol{u}}_H$, $\tilde{\varepsilon} = p_h - \tilde{p}_H$, and $\tilde{e}^\theta = \theta_h - \tilde{\theta}_H$. We note that, since $(\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta) \in Y_h$,

$$B((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\boldsymbol{z}, z^\theta)^{\rm pr}) = 0 \quad \text{and} \quad B((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\boldsymbol{z}, z^\theta)^{\rm du}) = 0.$$

Therefore, from Eqs. (22) and (23), and from the definition of the form E_0, Eq. (12), we have

$$2E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, 0)^{\rm pr}, (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)) = \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)), \tag{31}$$

$$2E_0((\boldsymbol{0}, 0, \hat{e}^\theta)^{\rm pr}, (\boldsymbol{0}, 0, \tilde{e}^\theta)) = \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{0}, 0, \tilde{e}^\theta)), \tag{32}$$

and similarly

$$2E_0((\hat{\boldsymbol{e}}, \hat{\varepsilon}, \hat{e}^\theta)^{\rm du}, (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) = \mathcal{R}^{\rm du}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\boldsymbol{\psi}_H, \mu_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)). \tag{33}$$

Then, using Eq. (25), we combine Eqs. (31), (32), and (33) to write

$$\begin{aligned}
&2\kappa^u E_0((\hat{\boldsymbol{e}}^\pm, \hat{\varepsilon}^\pm, 0), (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)) + 2\kappa^\theta E_0((\boldsymbol{0}, 0, \hat{e}^{\theta\pm}), (\boldsymbol{0}, 0, \tilde{e}^\theta))\\
&\quad = \kappa^u \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)) + \kappa^\theta \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\boldsymbol{0}, 0, \tilde{e}^\theta))\\
&\qquad \mp \mathcal{R}^{\rm du}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\boldsymbol{\psi}_H, \mu_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)).
\end{aligned} \tag{34}$$

We will now expand the right-hand side of Eq. (34). Using Eq. (16) and the expansion, Eq. (11), one can write

$$\begin{aligned}
&\kappa^u \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)) = \kappa^u E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0))\\
&\quad + \kappa^u E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, 0)) + \kappa^u F((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, 0))
\end{aligned} \tag{35}$$

and

$$\kappa^\theta \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\mathbf{0}, 0, \tilde{e}^\theta)) = \kappa^\theta E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta), (\mathbf{0}, 0, \tilde{e}^\theta)) \\ + \kappa^\theta E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\mathbf{0}, \tilde{e}^\theta)) + \kappa^\theta F((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{e}^\theta)(\mathbf{0}, \tilde{e}^\theta)). \tag{36}$$

From the definition of the dual residual, Eq. (15), we also have

$$\mathcal{R}^{\rm du}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\boldsymbol{\psi}_H, \mu_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) \\ = -L((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) - E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta), (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H)) \\ - E_1((\boldsymbol{u}_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}), (\boldsymbol{\psi}_H, \mu_H)),$$

which, by virtue of Eqs. (16) and (11), can be rewritten as

$$\mathcal{R}^{\rm du}((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H); (\boldsymbol{\psi}_H, \mu_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) \\ = -L((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) - \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H)) \\ + F((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{\psi}}_H, \tilde{\mu}_H)) + [E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{\psi}}_H, \tilde{\mu}_H)) \\ - E_1((\boldsymbol{u}_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\boldsymbol{\psi}_H, \mu_H))] + [L((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) \\ - L((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta))]. \tag{37}$$

Now, combining Eqs. (34)–(37), we write

$$0 = \mp 2\kappa^u E_0((\hat{\boldsymbol{e}}^\pm, \hat{\varepsilon}^\pm, 0), (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)) \mp 2\kappa^\theta E_0((\mathbf{0}, 0, \hat{e}^{\theta\pm}), (\mathbf{0}, 0, \tilde{e}^\theta)) \\ \pm \kappa^u E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)) \pm \kappa^u E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, 0)) \\ \pm \kappa^\theta E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta), (\mathbf{0}, 0, \tilde{e}^\theta)) \pm \kappa^\theta E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\mathbf{0}, \tilde{e}^\theta)) \\ \pm \kappa^u F((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, 0)) \pm \kappa^\theta F((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\mathbf{0}, \tilde{e}^\theta)) \\ + L((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) + \mathcal{R}^{\rm pr}((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{\psi}}_H, \tilde{\lambda}_H, \tilde{\mu}_H)) \\ - F((\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{\psi}}_H, \tilde{\mu}_H)) - [E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\tilde{\boldsymbol{\psi}}_H, \tilde{\mu}_H)) \\ - E_1((\boldsymbol{u}_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{e}^\theta), (\boldsymbol{\psi}_H, \mu_H))] - [L((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta)) \\ -L((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta))]. \tag{38}$$

If we add Eq. (38) to the bound equation, Eq. (24), noting that

$$E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta), (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)) = E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0), (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, 0)), \\ E_0((\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^\theta), (\mathbf{0}, 0, \tilde{e}^\theta)) = E_0((\mathbf{0}, 0, \tilde{e}^\theta)), (\mathbf{0}, 0, \tilde{e}^\theta)),$$

we find the bound error expression

$$s_H^\pm = s_h \pm D^\pm + I^\pm,$$

where

$$D^\pm = \kappa^u \mathcal{F}_0(\tilde{\boldsymbol{e}} - \hat{\boldsymbol{e}}^\pm, \tilde{\varepsilon} - \hat{\varepsilon}^\pm, 0) + \kappa^\theta \mathcal{F}_0(\mathbf{0}, 0, \tilde{e}^\theta - \hat{e}^\theta)$$

and $I^{\pm} = \sum_{j=1}^{6} I_j^{\pm}$. In these expressions, $\mathcal{F}_0(\boldsymbol{v}, q, \phi) = E_0((\boldsymbol{v}, q, \phi), (\boldsymbol{v}, q, \phi))$ and

$$\begin{aligned}
I_1^{\pm} &= -F((\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\tilde{\psi}_H, \tilde{\mu}_H)),\\
I_2^{\pm} &= \pm\kappa^u E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\tilde{\boldsymbol{e}}, 0)) \pm \kappa^{\theta} E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\boldsymbol{0}, \tilde{e}^{\theta})),\\
I_3^{\pm} &= \pm\kappa^u F((\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\tilde{\boldsymbol{e}}, 0)) \pm \kappa^{\theta} F((\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\boldsymbol{0}, \tilde{e}^{\theta})),\\
I_4^{\pm} &= -M(\tilde{\boldsymbol{e}}, \tilde{\boldsymbol{e}}),\\
I_5^{\pm} &= E_1((\boldsymbol{u}_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\psi_H, \mu_H)) - E_1((\tilde{\boldsymbol{u}}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{e}^{\theta}), (\tilde{\psi}_H, \tilde{\mu}_H)),\\
I_6^{\pm} &= L((\boldsymbol{u}_H, p_H, \theta_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^{\theta})) - L((\tilde{\boldsymbol{u}}_H, \tilde{p}_H, \tilde{\theta}_H); (\tilde{\boldsymbol{e}}, \tilde{\varepsilon}, \tilde{e}^{\theta})).
\end{aligned}$$

APPENDIX 4. CONVERGENCE OF THE INDEFINITE TERMS

We give here explicit expressions for the indefinite terms, $I_{1,\ldots,6}$, defined in Appendix 3,

$$\begin{aligned}
|I_1^{\pm}| &= \left| \int_{\Omega} \tilde{e}_i \tilde{e}_j \frac{\partial \tilde{\psi}_{H_i}}{\partial x_j}\, d\Omega + \int_{\Omega} \tilde{e}^{\theta} \tilde{e}_j \frac{\partial \tilde{\mu}_H}{\partial x_j}\, d\Omega \right|,\\
|I_2^{\pm}| &= \left| \kappa^u \left[\int_{\Omega} \tilde{u}_{H_i} \tilde{e}_j \frac{\partial \tilde{e}_i}{\partial x_j}\, d\Omega + \int_{\Omega} \tilde{e}_i \tilde{u}_{H_j} \frac{\partial \tilde{e}_i}{\partial x_j}\, d\Omega \right] \right.\\
&\quad \left. + \kappa^{\theta} \left[\int_{\Omega} \tilde{\theta}_H \tilde{e}_j \frac{\partial \tilde{e}^{\theta}}{\partial x_j}\, d\Omega + \int_{\Omega} \tilde{e}^{\theta} \tilde{u}_{H_j} \frac{\partial \tilde{e}^{\theta}}{\partial x_j}\, d\Omega \right] - \beta\kappa^u \int_{\Omega} \hat{g}_i \tilde{e}_i \tilde{e}^{\theta} d\Omega \right|,\\
|I_3^{\pm}| &= \left| \kappa^u \int_{\Omega} \tilde{e}_i \tilde{e}_j \frac{\partial \tilde{e}_i}{\partial x_j}\, d\Omega + \kappa^{\theta} \int_{\Omega} \tilde{e}^{\theta} \tilde{e}_j \frac{\partial \tilde{e}^{\theta}}{\partial x_j}\, d\Omega \right|,\\
|I_4^{\pm}| &= |m(\tilde{\boldsymbol{e}}, \tilde{\boldsymbol{e}})|,\\
|I_5^{\pm}| &= \left| \int_{\Omega} u_{Hi} \tilde{e}_j \frac{\partial \psi_{Hi}}{\partial x_j}\, d\Omega + \int_{\Omega} \tilde{e}_i u_{Hj} \frac{\partial \psi_{Hi}}{\partial x_j}\, d\Omega + \int_{\Omega} \theta_H \tilde{e}_j \frac{\partial \mu_H}{\partial x_j}\, d\Omega \right.\\
&\quad + \int_{\Omega} \tilde{e}^{\theta} u_{Hj} \frac{\partial \mu_H}{\partial x_j}\, d\Omega - \int_{\Omega} \tilde{u}_{Hi} \tilde{e}_j \frac{\partial \tilde{\psi}_{Hi}}{\partial x_j}\, d\Omega - \int_{\Omega} \tilde{e}_i \tilde{u}_{Hj} \frac{\partial \tilde{\psi}_{Hi}}{\partial x_j}\, d\Omega\\
&\quad \left. - \int_{\Omega} \tilde{\theta}_H \tilde{e}_j \frac{\partial \tilde{\mu}_H}{\partial x_j}\, d\Omega - \int_{\Omega} \tilde{e}^{\theta} \tilde{u}_{Hj} \frac{\partial \tilde{\mu}_H}{\partial x_j}\, d\Omega - \beta \int_{\Omega} \tilde{e}^{\theta} (\psi_{Hi} - \tilde{\psi}_{Hi}) \hat{g}_i\, d\Omega \right|,\\
|I_6^{\pm}| &= 2|m(\boldsymbol{u}_H - \tilde{\boldsymbol{u}}_H, \tilde{\boldsymbol{e}})|.
\end{aligned}$$

We also note that

$$\begin{aligned}
|\tilde{\boldsymbol{e}}|_{H^1(\Omega)} &= |\boldsymbol{u}_h - \tilde{\boldsymbol{u}}_H|_{H^1(\Omega)}\\
&\le |\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(\Omega)} + |\boldsymbol{u}_H - \tilde{\boldsymbol{u}}_H|_{H^1(\Omega)}\\
&\le C|\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(\Omega)} \le CH^2;
\end{aligned}$$

the last line follows form standard *a priori* estimates assuming that $\boldsymbol{u} \in H^3(\Omega) \times H^3(\Omega)$. Similarly, we have

$$\|\tilde{\boldsymbol{e}}\|_{L^2(\Omega)} \le \|\boldsymbol{u}_h - \boldsymbol{u}_H\|_{L^2(\Omega)} + CH|\boldsymbol{u}_h - \boldsymbol{u}_H|_{H^1(\Omega)} \le CH^3,$$

from the standard Aubin–Nitsche estimate. This result is directly useful to prove, for instance, that, in term $I_6^{\pm}$,

$$|m(\boldsymbol{u}_H - \tilde{\boldsymbol{u}}_H, \tilde{\boldsymbol{e}})| \leq C\|\boldsymbol{u}_H - \tilde{\boldsymbol{u}}_H\|_{L^2(\Omega)}\|\tilde{\boldsymbol{e}}\|_{L^2(\Omega)} \leq CH^6.$$

We will now estimate the convergence rates of two other (representative) selected terms in $\sum_{i=1}^{6}|I_i^{\pm}|$. The convergence rates of the remaining terms are obtained by similar arguments.

We first consider a term in $|I_1^{\pm}|$,

$$\begin{aligned}\left|\int_\Omega \tilde{e}_i\tilde{e}_j\frac{\partial\tilde{\psi}_{Hi}}{\partial x_j}\,d\Omega\right| &\leq \left|\int_\Omega \tilde{e}_i\tilde{e}_j\frac{\partial(\tilde{\psi}_{Hi}-\psi_i)}{\partial x_j}\,d\Omega\right| + \left|\int_\Omega \tilde{e}_i\tilde{e}_j\frac{\partial\psi_i}{\partial x_j}\,d\Omega\right|,\\ &\leq \|\tilde{\boldsymbol{e}}\|_{L^4(\Omega)}\|\tilde{\boldsymbol{e}}\|_{L^4(\Omega)}\|\nabla(\tilde{\boldsymbol{\psi}}_H-\boldsymbol{\psi})\|_{L^2(\Omega)}\\ &\quad+\|\tilde{\boldsymbol{e}}\|_{L^2(\Omega)}\|\tilde{\boldsymbol{e}}\|_{L^4(\Omega)}\|\nabla\boldsymbol{\psi}\|_{L^4(\Omega)},\end{aligned}$$

where we have used the Cauchy–Schwartz inequality. By virtue of the Sobolev inequality, $\|v\|_{L^4(\Omega)} \leq C\|v\|_{H^1(\Omega)}, \forall v \in H^1(\Omega)$ (e.g., see [1]), we now write

$$\begin{aligned}\left|\int_\Omega \tilde{e}_i\tilde{e}_j\frac{\partial\tilde{\psi}_{Hi}}{\partial x_j}d\Omega\right| &\leq C_1\|\tilde{\boldsymbol{e}}\|_{H^1(\Omega)}\|\tilde{\boldsymbol{e}}\|_{H^1(\Omega)}|\tilde{\boldsymbol{\psi}}_H-\boldsymbol{\psi}|_{H^1(\Omega)} + C_2\|\tilde{\boldsymbol{e}}\|_{L^2(\Omega)}\|\tilde{\boldsymbol{e}}\|_{H^1(\Omega)}\|\boldsymbol{\psi}\|_{H^2(\Omega)}\\ &\leq C_1H^6 + C_2H^5,\end{aligned}$$

where we have exploited our previous estimates; here C_1 and C_2 denote two generic positive constants independent of H, and we have assumed $\boldsymbol{u}, \boldsymbol{\psi} \in H^3(\Omega)\times H^3(\Omega)$.

For our second example, we rewrite $I_5^{\pm}$ as

$$\begin{aligned}|I_5^{\pm}| = \Bigg|&\int_\Omega (u_{Hi}-\tilde{u}_{Hi})\tilde{e}_j\frac{\partial\psi_{Hi}}{\partial x_j}\,d\Omega + \int_\Omega \tilde{u}_{Hi}\tilde{e}_j\frac{\partial(\psi_{Hi}-\tilde{\psi}_{Hi})}{\partial x_j}\,d\Omega\\ &+\int_\Omega \tilde{e}_i(u_{Hj}-\tilde{u}_{Hj})\frac{\partial\psi_{Hi}}{\partial x_j}\,d\Omega + \int_\Omega \tilde{e}_i\tilde{u}_{Hj}\frac{\partial(\psi_{Hi}-\tilde{\psi}_{Hi})}{\partial x_j}\,d\Omega\\ &+\int_\Omega e^\theta(u_{Hj}-\tilde{u}_{Hj})\frac{\partial\mu_H}{\partial x_j}\,d\Omega - \beta\int_\Omega \tilde{e}^\theta(\psi_{Hi}-\tilde{\psi}_{Hi})\hat{g}_i\,d\Omega\Bigg|,\end{aligned}$$

and we consider, for instance,

$$\begin{aligned}\left|\int_\Omega(\tilde{u}_{Hi}-u_{Hi})\tilde{e}_j\frac{\partial\psi_{Hi}}{\partial x_j}\,d\Omega\right| &\leq \left|\int_\Omega(\tilde{u}_{Hi}-u_{Hi})\tilde{e}_j\frac{\partial\psi_i}{\partial x_j}\,d\Omega\right|\\ &\quad+\left|\int_\Omega(\tilde{u}_{Hi}-u_{Hi})\tilde{e}_j\frac{\partial(\tilde{\psi}_{Hi}-\psi_i)}{\partial x_j}\,d\Omega\right|\\ &\leq C_1\|\tilde{\boldsymbol{e}}\|_{L^2(\Omega)}\|\boldsymbol{u}_H-\tilde{\boldsymbol{u}}_H\|_{L^4(\Omega)}\|\nabla\boldsymbol{\psi}\|_{L^4(\Omega)}\\ &\quad+C_2\|\tilde{\boldsymbol{\psi}}_H-\boldsymbol{\psi}\|_{H^1(\Omega)}\|\tilde{\boldsymbol{e}}\|_{L^4(\Omega)}\|\tilde{\boldsymbol{u}}_H-\boldsymbol{u}_H\|_{L^4(\Omega)}\\ &\leq C_1\|\tilde{\boldsymbol{e}}\|_{L^2(\Omega)}\|\boldsymbol{u}_H-\tilde{\boldsymbol{u}}_H\|_{H^1(\Omega)}\|\boldsymbol{\psi}\|_{H^2(\Omega)}\\ &\quad+C_2\|\tilde{\boldsymbol{\psi}}_H-\boldsymbol{\psi}\|_{H^1(\Omega)}\|\tilde{\boldsymbol{e}}\|_{H^1(\Omega)}\|\tilde{\boldsymbol{u}}_H-\boldsymbol{u}_H\|_{H^4(\Omega)}\\ &\leq C_1H^5 + C_2H^6,\end{aligned}$$

where the second inequality follows from the application of the Cauchy–Schwartz inequality. In the third inequality, we have again used the standard Sobolev inequality $\|v\|_{L^4(\Omega)} \leq$

$C\|v\|_{H^1(\Omega)}$ for all $v \in H^1(\Omega)$. The last inequality is a direct consequence of our previous estimates.

ACKNOWLEDGMENTS

This work was supported by NASA Grants NAG1-1978, NAG1-1587, and NAG4-105, DARPA and ONR Grant N00014-91-J-1-1889, and AFOSR Grant F49620-97-1-0052. L.M. was partially supported by Fulbright, BAEF, and NATO Fellowships. We acknowledge our longstanding and very fruitful collaboration with Professor Y. Maday of Université Pierre et Marie Curie, Paris, France.

REFERENCES

1. R. A. Adams, *Sobolev Spaces* (Academic Press, New York, 1975).
2. M. Ainsworth and J. T. Oden, A posteriori error estimation in finite element analysis, *Comput. Methods Appl. Mech. Eng*. **142**, 1 (1997).
3. R. E. Bank and A. Weiser, Some a posteriori error estimators for elliptic partial differential equations, *Math. Comput.* **44** (170), 283 (1985).
4. R. Becker and R. Rannacher, A feedback approach to error control in finite element method: Basic analysis and examples, *East–West J. Numer. Math*. **4**, 237 (1996).
5. M. Crouzeix and P.-A. Raviart, Conforming and non-conforming finite element methods for solving the stationary Stokes equation, *RAIRO Anal. Numer*. **7**, 33 (1973).
6. C. Farhat and J. Mandel, Scalable substructuring by Lagrange multipliers in theory and practice, in *Domain Decomposition Methods for Partial Differential Equations*, edited by P. Bjorstad, M. Espedal, and D. Keyes (Wiley, New York, 1998).
7. P. Ladeveze and D. Leguillon, Error estimation procedures in the finite element method and applications, *SIAM J. Numer. Anal*. **20**, 485 (1983).
8. 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 *Proceedings of ICFD Conference, Oxford*, 1998.
9. L. Machiels, J. Peraire, and A. T. Patera, Output bound approximations for partial differential equations: Application to the incompressible Navier–Stokes equations, in *Proceedings of the Istanbul Workshop on Industrial and Environmental Applications of Direct and Large Eddy Numerical Simulation*, edited by S. Biringen (Springer-Verlag, Berlin, 1998).
10. Y. Maday and A. T. Patera, Numerical analysis of a posteriori finite element bounds for linear-functional outputs, *Math. Models Methods Appl. Sci*. **10**(5), 785 (2000).
11. 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 Sér. I* **328**, 823 (1999).
12. J. Mandel and M. Brezina, *Balancing Domain Decomposition*: *Theory and Performance in Two and Three Dimensions*, Technical report, Center for Computational Mathematics (University of Colorado at Denver, 1993).
13. M. Paraschivoiu and A. T. Patera, A hierarchical duality approach to bounds for the outputs of partial differential equation, *Comput. Methods. Appl. Mech. Eng*. **158**, 389 (1998).
14. M. Paraschiviou and A. T. Patera, A posteriori bounds for linear-functional outputs of Crouzeix–Raviart finite element discretizations of the incompressible Stokes problem, *Int. J. Numer. Methods Fluids* **32**, 823 (2000).
15. M. Paraschivoiu, J. Peraire, and A. T. Patera, A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations, *Comput. Methods Appl. Mech. Eng*. **150**, 289 (1997).
16. A. T. Patera and E. M. Ronquist, A general output bound result: Application to discretization and iteration error estimation and control, *Math. Models Methods Appl. Sci*. **11**(4), 685 (2001).
17. J. Peraire and A. T. Patera, Bounds for linear-functional outputs of coercive partial differential equations: Local indictors and adaptive refinement, in *On New Advances in Adaptive Computational Methods in Mechanics*, edited by P. Ladeveze and J. T. Oden (Elsevier Science Ltd., Oxford, 1998).