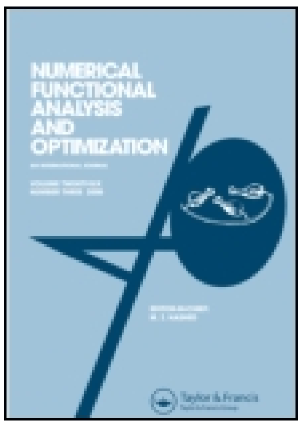


Numerical Functional Analysis and Optimization
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lnfa20

# A Posteriori Finite Element Error Analysis for Symmetric Positive Differential Equations

Jang Jou [a] & Jinn-Liang Liu [b]
[a] Department of Statistics , Ming Chuan University , Taipei, Taiwan E-mail:
[b] Department of Applied Mathematics , National Chiao Tung University , Hsinchu, Taiwan E-mail:
Published online: 15 May 2007.

NUMER. FUNCT. ANAL. AND OPTIMIZ., 20(5&6), 473–490 (1999)

# A POSTERIORI FINITE ELEMENT ERROR ANALYSIS FOR SYMMETRIC POSITIVE DIFFERENTIAL EQUATIONS*

JANG JOU† AND JINN-LIANG LIU‡

†Department of Statistics, Ming Chuan University, Taipei, Taiwan
(e-mail: Jouj@mcu.edu.tw)
‡Department of Applied Mathematics, National Chiao Tung University, Hsinchu, Taiwan
(e-mail: jinnliu@math.nctu.edu.tw)



**Abstract.** Based on the solution of local weak residual problems, conforming and nonconforming error estimators are presented and analyzed for finite element solutions of symmetric positive differential equations in the sense of Friedrichs. These estimators are devised to treat the Friedrichs system in a general setting in terms of application (hyperbolic as well as mixed-type problems), approximation ($h$-, $p$- and $hp$-version finite element methods), implementation (no local boundary conditions and no flux jumps across element boundaries) and a posteriori error analysis (very moderate conditions on the system and on the approximation). Three model problems of the Friedrichs system, namely, the neutron transport equation, the forward-backward heat equation and the Tricomi problem are used to illustrate the applicability of the weak residual error estimation.

**1. Introduction.** Over the last two decades, a posteriori error estimation in connection with adaptive finite element methods for partial differential equations (PDEs) has been a subject of active research. Although there are large amounts of literature concerning the error estimation for elliptic and parabolic PDEs [1, 2, 3, 7, 8, 9, 10, 11, 12, 13, 23, 24, 25, 26, 30, 31], comparatively very small have been addressed to the PDEs of mixed-type, such as the Tricomi equation and the forward-backward heat equation, or of hyperbolic type, such as the neutron transport equation.

The theory of Friedrichs' symmetric positive system [16] has been shown very useful for theoretical as well as numerical investigations for these types of problems [5, 6, 17, 18, 19, 27]. An error estimator, which is based on the solution of weak residual problems and is referred to as a conforming error estimator herein, was first proposed in [20] for the Friedrichs system and is applied primarily to the mixed-type problems. In [27], Süli presents another estimator which is applied to the hyperbolic problems and is based on a postprocess of residuals and normal flux on each element. The conforming estimator was shown to be bounded below and above by the true error in the norm induced by the bilinear form of Lesaint [19] whereas the postprocessing estimator was bounded below in $L^2$-norm.

In this article we propose two more estimators, one conforming and the other non-conforming, which are also based on the weak residual formulation with, however, a

* This work was supported by NSC under grant 87-2115-M-009-005, Taiwan.

different bilinear form. In comparison with the previous approach in [20], the present formulation offers simpler implementation and more complete error analysis for the resulting error estimators. It is shown that the present (conforming) estimator is identical to the previous one. However, the present approach is easier to be extended to the nonconforming formulation which is not given in the previous work. Furthermore, the weak residual error estimation is shown to be applicable to all the model problems, namely, the mixed-type as well as hyperbolic problems. Both conforming and nonconforming estimators are proved to have two-sided bounds by the true error in the norm induced by the new bilinear form with very moderate conditions on the system and on approximation that can be any one of $h$, $p$ and $hp$ finite element approximations.

The layout of the paper is as follows. In the next section, we briefly describe the Friedrichs system and its finite element approximation by means of Lesaint's formulation. The conforming and nonconforming error estimators are given in Sections 3 and 4, respectively. By examining all conditions made for the error analysis, the last section illustrates the applicability of the estimators to the neutron transport equation, the forward-backward heat equation and the Tricomi problem.

**2. Preliminaries.** Let $\Omega \subset R^2$ be a bounded region with a Lipschitz boundary $\partial\Omega$ and denote by $H^k(\Omega)$, $k \geq 0$ integers, the Sobolev spaces equipped with the norms $\|\cdot\|_k$. As usual, $H^0(\Omega) = L^2(\Omega)$. Define the product space $[H^k(\Omega)]^m := H^k(\Omega) \times \cdots \times H^k(\Omega)$ ($m$-times) with the corresponding norm denoted again by $\|\cdot\|_k$. We denote in particular $[H^1(\Omega)]^m$ by $H(\Omega)$ for simplicity. Consider the boundary value problem: Given a vector-valued function $\mathbf{f}(\mathbf{x}) := (f_1(\mathbf{x}), \cdots, f_m(\mathbf{x}))^t \in [L^2(\Omega)]^m$, $\mathbf{x} = (x_1, x_2) = (x, y)$, find a vector-valued function $\mathbf{u}(\mathbf{x}) := (u_1(\mathbf{x}), \cdots, u_m(\mathbf{x}))^t$ satisfying the system of equations

$$\begin{cases} \mathcal{L}\mathbf{u} = \sum_{i=1}^2 M_i \frac{\partial \mathbf{u}}{\partial x_i} + M_0 \mathbf{u} = \mathbf{f} \text{ in } \Omega \\ \mathcal{B}\mathbf{u} = (\mu - \beta)\mathbf{u} = \mathbf{0} \text{ on } \partial\Omega, \end{cases} \tag{2.1}$$

where $M_i(\mathbf{x})$, $0 \leq i \leq 2$, $\mu(\mathbf{x})$, and $\beta(\mathbf{x})$ are all $m \times m$ matrices and $\beta(\mathbf{x}) = \sum_{i=1}^2 \nu_i(\mathbf{x}) M_i(\mathbf{x})$ with $\nu_i$ being the components of the unit outward normal $\boldsymbol{\nu}$ on $\partial\Omega$. The matrices $M_i$, $0 \leq i \leq 2$, are Lipschitz-continuous in $x \in \bar{\Omega}$, and $M_0$ and $\mu$ are bounded in $\Omega$ and on $\partial\Omega$, respectively. We assume that the system is symmetric positive in the sense of Friedrichs [16], namely, that the following conditions hold:

- $M_1$ and $M_2$ are symmetric on $\bar{\Omega}$,
- the matrix $C := M_0 + M_0^t - \sum_{i=1}^2 \frac{\partial M_i}{\partial x_i}$ is positive definite in $\Omega$, i.e., there exists a positive constant $c$ such that $C \geq cI$ where $I$ is the identity matrix,
- the matrix $\mu + \mu^t$ is positive semidefinite on $\partial\Omega$ ($\mu$ is positive semidefinite for simplicity), and
- $\operatorname{Ker}(\mu - \beta) \oplus \operatorname{Ker}(\mu + \beta) = R^m$ on $\partial\Omega$.

The formal adjoints $\mathcal{L}^*$ and $\mathcal{B}^*$ of $\mathcal{L}$ and $\mathcal{B}$ are defined by

$$\begin{aligned} \mathcal{L}^*\mathbf{v} &= -\sum_{i=1}^2 \frac{\partial}{\partial x_i}(M_i\mathbf{v}) + M_0^t\mathbf{v} \\ \mathcal{B}^*\mathbf{v} &= (\mu^t + \beta)\mathbf{v}. \end{aligned}$$

Let

$$(\mathbf{g},\mathbf{h}) = \int_\Omega \mathbf{g}(\mathbf{x}) \cdot \mathbf{h}(\mathbf{x})\, d\mathbf{x}, \quad < \mathbf{g},\mathbf{h} > = \int_{\partial\Omega} \mathbf{g}(\mathbf{x}) \cdot \mathbf{h}(\mathbf{x})\, ds,$$

where $\mathbf{g} \cdot \mathbf{h} := \sum_{i=1}^m g_i h_i$.

Following Lesaint's formulation [19], the weak version of (2.1) is to find $\mathbf{u} \in H(\Omega)$ such that

$$B(\mathbf{u},\mathbf{v}) = F(\mathbf{v}) \;\; \forall \mathbf{v} \in H(\Omega), \tag{2.2}$$

where

$$\begin{aligned} B(\mathbf{u},\mathbf{v}) &= \frac{1}{2}(\mathcal{L}\mathbf{u},\mathbf{v}) + \frac{1}{2}(\mathbf{u},\mathcal{L}^*\mathbf{v}) + \frac{1}{2} < \mu\mathbf{u},\mathbf{v} > \\ F(\mathbf{v}) &= (\mathbf{f},\mathbf{v}). \end{aligned}$$

The finite element approximation of (2.2) is to find $\mathbf{u}_h \in S_h$ such that

$$B(\mathbf{u}_h,\mathbf{v}) = F(\mathbf{v}) \;\; \forall \mathbf{v} \in S_h, \tag{2.3}$$

where $S_h$ is a finite element subspace of $H(\Omega)$, which is associated with a mesh $T_h = \{\tau_i | i = 1,2,\cdots,n\}$ on $\bar{\Omega}$. Meshes are characterized by the mesh size parameter $h$. For any two distinct elements (triangles or rectangles or both) $\tau_i$ and $\tau_j$ in $T_h$, $\tau_i \cap \tau_j$ is either empty, a single vertex or a common edge. Two elements are said to be adjacent if they have a common edge. For a given rectangle element let $h_{max}$ and $h_{min}$ denote the largest and smallest edge lengths, respectively. Then the element edge ratio is defined by $h_{min}/h_{max}$. We always assume that the mesh $T_h$ belongs to a *regular* family of meshes on $\bar{\Omega}$. Recall that, see e.g. [4, 14], the family is regular if all angles of its triangular elements and all edge ratios of rectangular elements are bounded below by some constant $\sigma > 0$. Shape regularity does not require a mesh to be globally quasi-uniform, but it does imply local quasi-uniformity of the mesh. We require all finite element spaces to have *locally affine* bases [13].

One of very important properties that distinguishes the symmetric positive system from the elliptic system of PDEs is that the bilinear form $B$ is coercive in the $[L^2(\Omega)]^m$ norm but bounded by the $[H^1(\Omega)]^m$ norm, i.e.,

$$C_1 \|\mathbf{w}\|_0^2 \le B(\mathbf{w},\mathbf{w}) \le C_2 \|\mathbf{w}\|_1^2 \;\; \forall \mathbf{w} \in H(\Omega), \tag{2.4}$$

where $C_1$ and $C_2$ are positive constants depending only on $\Omega$. In fact, the error estimators proposed here are primarily motivated by this property.

Note that, by (2.4), the symmetric bilinear form

$$\frac{1}{2}B(\mathbf{w},\mathbf{v}) + \frac{1}{2}B(\mathbf{v},\mathbf{w}), \;\; \forall \mathbf{w},\mathbf{v} \in H(\Omega),$$

defines an inner product for the space $H(\Omega) = [H^1(\Omega)]^m$ and thus a norm

$$\|\mathbf{w}\|_B = \sqrt{B(\mathbf{w},\mathbf{w})} \ \ \forall \mathbf{w} \in H(\Omega).$$

Since $\mathcal{L} + \mathcal{L}^* = C$, we see that

$$(\mathcal{L}\mathbf{w},\mathbf{w}) + (\mathbf{w},\mathcal{L}^*\mathbf{w}) = (C\mathbf{w},\mathbf{w}) \ \ \forall \mathbf{w} \in H(\Omega)$$

and

$$\|\mathbf{w}\|_B^2 = B(\mathbf{w},\mathbf{w}) = \frac{1}{2}(C\mathbf{w},\mathbf{w}) + \frac{1}{2} < \mu\mathbf{w},\mathbf{w} > \quad \forall \mathbf{w} \in H(\Omega).$$

Note that the matrix $C$ is symmetric and positive definite. Define the bilinear form

$$A(\mathbf{w},\mathbf{v}) = A_1(\mathbf{w},\mathbf{v}) + A_0(\mathbf{w},\mathbf{v}) \tag{2.5}$$

with

$$\begin{aligned} A_1(\mathbf{w},\mathbf{v}) &= \frac{1}{2}\int_\Omega C\mathbf{w}\cdot\mathbf{v}\,d\mathbf{x} \\ A_0(\mathbf{w},\mathbf{v}) &= \frac{1}{4}\int_{\partial\Omega} (\mu\mathbf{w}\cdot\mathbf{v} + \mathbf{w}\cdot\mu\mathbf{v})\,ds. \end{aligned}$$

Then

$$A(\mathbf{w},\mathbf{w}) = B(\mathbf{w},\mathbf{w})$$

Obviously, the bilinear form $A$ induces a norm, denoted by $\|\cdot\|_A$, on $H(\Omega)$, which is identical to the $B$-norm, i.e.,

$$\|\mathbf{w}\|_A = \|\mathbf{w}\|_B \ \ \forall \mathbf{w} \in H(\Omega). \tag{2.6}$$

**3. A Conforming Error Estimator.** Our objective now is to estimate the true error $\mathbf{e} = \mathbf{u} - \mathbf{u}_h \in H(\Omega)$ between the exact solution $\mathbf{u} \in H(\Omega)$ of (2.2) and the approximate solution $\mathbf{u}_h \in S_h$ of (2.3). Substituting $\mathbf{u} = \mathbf{u}_h + \mathbf{e}$ into Eq. (2.2), we have

$$B(\mathbf{e},\mathbf{v}) = F(\mathbf{v}) - B(\mathbf{u}_h,\mathbf{v}) \quad \forall \mathbf{v} \in H(\Omega). \tag{3.1}$$

Weak residual error estimators are derived by some approximation of (3.1). Since $\mathbf{u}_h \in S_h$ satisfies (2.3), we have the orthogonality

$$B(\mathbf{e},\mathbf{v}) = F(\mathbf{v}) - B(\mathbf{u}_h,\mathbf{v}) = 0 \quad \forall \mathbf{v} \in S_h. \tag{3.2}$$

Consequently, in order to obtain a nontrivial approximation to $\mathbf{e}$, we should consider the approximation of (3.1) in a richer space $S_{\bar{h}}$, $S_h \subset S_{\bar{h}} \subset H(\Omega)$, i.e., consider the problem: Determine $\bar{\mathbf{e}} \in S_{\bar{h}}$ such that

$$B(\bar{\mathbf{e}},\mathbf{v}) = F(\mathbf{v}) - B(\mathbf{u}_h,\mathbf{v}) \qquad \forall \mathbf{v} \in S_{\bar{h}}. \tag{3.3}$$

However, the bilinear form $B$ is not symmetric and it is more complicated and expensive in terms of both implementation and computation than the bilinear form $A$. We propose here a more efficient formula for error estimation, that is, determine $\mathbf{e}_a \in H(\Omega)$ such that

$$A(\mathbf{e}_a, \mathbf{v}) = F(\mathbf{v}) - B(\mathbf{u}_h, \mathbf{v}) \quad \forall \mathbf{v} \in H(\Omega). \tag{3.4}$$

Note that the existence and uniqueness of the problem (3.4) follows from the definition of the bilinear form $A$ and that the orthogonality (3.2) still holds for this error problem. Similarly, there exists a unique $\bar{\mathbf{e}}_a \in S_{\bar{h}}$ such that

$$A(\bar{\mathbf{e}}_a, \mathbf{v}) = F(\mathbf{v}) - B(\mathbf{u}_h, \mathbf{v}) \qquad \forall \mathbf{v} \in S_{\bar{h}}. \tag{3.5}$$

Since (3.5) is a standard finite element approximation of (3.4) and the bilinear form $A$ is symmetric, it follows from Céa's Lemma [14] that

$$\|\mathbf{e}_a - \bar{\mathbf{e}}_a\|_A = \inf_{\mathbf{v} \in S_{\bar{h}}} \|\mathbf{e}_a - \mathbf{v}\|_A. \tag{3.6}$$

Let $\mathbf{e}_0 \in S_h$ be the solution of (3.5) in the original finite element space $S_h$. By (3.2), it is a trivial solution, i.e., $\mathbf{e}_0 = 0$. Hence, we have

$$\|\mathbf{e}_a - \bar{\mathbf{e}}_a\|_A \le \rho \, \|\mathbf{e}_a\|_A \tag{3.7}$$

with $\rho \in [0, 1]$. The inequality will assume the equality with $\rho = 1$ if and only if

$$A(\bar{\mathbf{e}}_a, \bar{\mathbf{e}}_a) = 2A(\mathbf{e}_a, \bar{\mathbf{e}}_a) = 2A(\bar{\mathbf{e}}_a, \bar{\mathbf{e}}_a) = 0, \tag{3.8}$$

which implies that $\rho \in [0, 1)$ provided $S_{\bar{h}} \neq S_h$. This suggests that the enlarged space $S_{\bar{h}}$ can be defined on the current mesh $T_h$. By this we mean that more basis functions that do not belong to $S_h$ are constructed on the present mesh without any re-meshing. These functions constitute a complementary subspace $S^c$ to $S_h$ in $S_{\bar{h}}$. Apparently, for any fixed mesh or equivalently any fixed mesh parameter $h$, the constant $\rho$ is independent of the $h$ and depends only on how many or how these complementary basis functions are constructed as long as $S^c \neq \emptyset$. We thus define the enlarged space by

$$S_{\bar{h}} = S_h \oplus S^c, \; S_h \cap S^c = \{0\}, \; S^c \neq \{0\}. \tag{3.9}$$

LEMMA 3.1. *Let* $\mathbf{e}$, $\mathbf{e}_a$ *and* $\bar{\mathbf{e}}_a$ *be the solutions of* (3.1), (3.4) *and* (3.5), *respectively. Then*

$$\begin{aligned} \|\mathbf{e}_a\|_A &= \|\mathbf{e}\|_A & (3.10)\\ (1-\rho)\|\mathbf{e}\|_A &\le \|\bar{\mathbf{e}}_a\|_A \le \|\mathbf{e}\|_A, & (3.11) \end{aligned}$$

*where* $\rho \in [0, 1)$ *is a constant independent of the mesh size* $h$ *provided* (3.9) *holds.*

*Proof.* We first observe that

$$\begin{aligned} \|\mathbf{e}_a\|_A{}^2 &= A(\mathbf{e}_a, \mathbf{e}_a)\\ &= F(\mathbf{e}_a) - B(\mathbf{u}_h, \mathbf{e}_a)\\ &= B(\mathbf{e}, \mathbf{e}_a) \end{aligned}$$

$$\begin{aligned} &\leq \|\mathbf{e}\|_B \|\mathbf{e}_a\|_B \\ &= \|\mathbf{e}\|_A \|\mathbf{e}_a\|_A , \end{aligned}$$

which implies $\|\mathbf{e}_a\|_A \leq \|\mathbf{e}\|_A$. On the other hand,

$$\begin{aligned} \|\mathbf{e}\|_B{}^2 &= B(\mathbf{e},\mathbf{e}) \\ &= F(\mathbf{e}) - B(\mathbf{u}_h,\mathbf{e}) \\ &= A(\mathbf{e}_a,\mathbf{e}) \\ &\leq \|\mathbf{e}_a\|_A \|\mathbf{e}\|_A \\ &= \|\mathbf{e}_a\|_A \|\mathbf{e}\|_B, \end{aligned}$$

which implies $\|\mathbf{e}\|_A \leq \|\mathbf{e}_a\|_A$. We thus have (3.10). Analogously, we have the right inequality of (3.11). The left inequality follows immediately from (3.7) and (3.10). □

Let $\gamma = \sup\{A(\mathbf{w},\mathbf{v}) : \mathbf{w} \in S_h,\ \|\mathbf{w}\|_A = 1,\ \mathbf{v} \in S^c,\ \|\mathbf{v}\|_A = 1\}$. We then readily have $\gamma \in [0,1)$ by the definition of the space $S^c$ in (3.9). However, it is not clear whether the constant is uniformly independent of the mesh parameter $h$ for all adaptive meshes. It is shown, for example, in [10] and [15] that this is indeed the case for both $L^2(\Omega)$- and $H^1(\Omega)$-inner product or their equivalence. Obviously, the inner product defined by the bilinear form $A_1$ in (2.5) is equivalent to $[L^2(\Omega)]^m$-inner product since the symmetric matrix $C$ is positive definite. This implies that there a constant $\gamma_C \in [0,1)$ independent of $h$ such that

$$|A_1(\mathbf{w},\mathbf{v})| \leq \gamma_C\ [A_1(\mathbf{w},\mathbf{w})]^{\frac{1}{2}}[A_1(\mathbf{v},\mathbf{v})]^{\frac{1}{2}}\ \forall \mathbf{w} \in S_h,\ \mathbf{v} \in S_h^c. \tag{3.12}$$

However, the corresponding strengthened Cauchy-Schwarz inequality to the bilinear form $A_0$ does not appear to be such evident due to the fact that the boundary matrix $\mu$ may not be positive definite on its entire domain, i.e., on $\partial\Omega$. It seems to be convenient to treat the boundary condition in a more specific way. For this, we postpone our proof of the inequality to the last section and assume for the moment that there exists a constant $\gamma_\mu \in [0,1)$ independent of $h$ such that for any $\mathbf{w} \in S_h,\ \mathbf{v} \in S_h^c$

$$|A_0(\mathbf{w},\mathbf{v})| \leq \gamma_\mu\, [A_0(\mathbf{w},\mathbf{w})]^{\frac{1}{2}}[A_0(\mathbf{v},\mathbf{v})]^{\frac{1}{2}}. \tag{3.13}$$

The above two strengthened Cauchy-Schwarz inequalities imply that

$$\begin{aligned} A(\mathbf{w},\mathbf{v})^2 &= [A_1(\mathbf{w},\mathbf{v}) + A_0(\mathbf{w},\mathbf{v})]^2 \\ &\leq \gamma^2 \left([A_1(\mathbf{w},\mathbf{w})]^{\frac{1}{2}}[A_1(\mathbf{v},\mathbf{v})]^{\frac{1}{2}} + [A_0(\mathbf{w},\mathbf{w})]^{\frac{1}{2}}[A_0(\mathbf{v},\mathbf{v})]^{\frac{1}{2}}\right)^2 \\ &\leq \gamma^2\ [A_1(\mathbf{w},\mathbf{w}) + A_0(\mathbf{w},\mathbf{w})][A_1(\mathbf{v},\mathbf{v}) + A_0(\mathbf{v},\mathbf{v})] \\ &= \gamma^2\ A(\mathbf{w},\mathbf{w})A(\mathbf{v},\mathbf{v}), \end{aligned} \tag{3.14}$$

where $\gamma = \max\{\gamma_C,\ \gamma_\mu\}$. We summarize as follows:

LEMMA 3.2. *If there exists a constant $\gamma_\mu \in [0,1)$ independent of $h$ such that (3.13) holds for all $\mathbf{w} \in S_h$ and $\mathbf{v} \in S_h^c$ in the enlarged space $S_{\bar{h}}$, then there exists a constant $\gamma \in [0,1)$ independent of $h$ such that, for all $\mathbf{w} \in S_h$ and $\mathbf{v} \in S^c$,*

$$|A(\mathbf{w},\mathbf{v})| \leq \gamma\ \|\mathbf{w}\|_A \|\mathbf{v}\|_A .$$

A conforming error estimator can be derived by a further reduction of the approximation (3.5): Determine $\mathbf{e}_a^c \in S^c$ such that

$$A(\mathbf{e}_a^c, \mathbf{v}) = F(\mathbf{v}) - B(\mathbf{u}_h, \mathbf{v}) \quad \forall \mathbf{v} \in S^c. \tag{3.15}$$

Lemmas 3.1 and 3.2 then yield the following well-known result for the conforming error estimator [10, 11, 12, 13, 20, 23].

THEOREM 3.3. *Let* $\mathbf{u} \in H(\Omega)$ *and* $\mathbf{u}_h \in S_h$ *be the solutions of* (2.2) *and* (2.3), *respectively, and let* $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$. *If the conditions of Lemma 3.2 hold, then the reduced conforming error problem* (3.15) *has a unique solution* $\mathbf{e}_a^c \in S^c$ *such that*

$$(1-\rho)\sqrt{1-\gamma^2}\,\|\mathbf{e}\|_A \le \|\mathbf{e}_a^c\|_A \le \|\mathbf{e}\|_A\,, \tag{3.16}$$

*where the constants* $\gamma \in [0,1)$ *and* $\rho \in [0,1)$ *are independent of the mesh parameter* $h$.

We briefly remark on the estimated error $\mathbf{e}_a^c$ and that of [20], which was derived simply by replacing the bilinear form $A$ by the original $B$ form in the left side of (3.15) and is denoted by $\mathbf{e}_b^c$. Following the proof of (3.10), it can be easily shown that $\|\mathbf{e}_a^c\|_A = \|\mathbf{e}_b^c\|_A$. This shows that the present estimator is exactly the same as that of [20]. However, in terms of the definition of both bilinear forms, the implementation of the present estimator is simpler. Furthermore, in terms of error analysis, we remove the saturation assumption of [20] and prove explicitly the strengthened Cauchy-Schwarz inequality. Finally, the present approach is easier to be extended to the nonconforming formulation which, not given in the previous work, will be introduced in the next section.

**4. A Nonconforming Error Estimator.** The complementary subspace $S^c$ would result (3.15) in a global system of linear equations if its basis functions have supports across element boundaries. This apparently is not suitable for adaptive computation in practice. On the other hand, if the basis functions have supports only on their individual elements, the resulting error estimator may not be effective to estimate errors that occur on the boundaries of elements if they are dominant errors.

Many estimators, primarily for the elliptic PDEs, have been proposed to handle these errors (flux jumps) across elements, see e.g. [3, 7, 8, 9, 10, 12, 25, 26, 30, 31]. All those estimators involve the jumps and some may require certain boundary conditions to hold for local problems. We present another way to treat the jumps for the symmetric positive system. The objective of the treatment is to retain the weak residual term of (3.15) without explicitly involving the jumps. The jumps are handled indirectly by a proper construction of the basis functions of the complementary space $S^c$.

For simplicity, we assume that the approximation is linear, i.e., $S_h$ consists of piecewise linear functions. The results in what follows hold for more general approximation with some technical modifications. We first introduce some notation. For any fixed mesh $T_h$, let $N(t_i)$ denote the index set of $j \neq i$ such that $t_j \in T_h$ is an adjacent element of $t_i \in T_h$. With the nodal point at the center of the common (interior) edge of the adjacent pair $t_i$ and $t_j$ in $T_h$, we construct a basis function $\phi_{ij}$ so that it has the support on $t_i \cup t_j$. Note that $\phi_{ij} = \phi_{ji}$. Let

$$\Phi_{ij}^1 = \begin{pmatrix} \phi_{ij} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \Phi_{ij}^2 = \begin{pmatrix} 0 \\ \phi_{ij} \\ 0 \\ \vdots \end{pmatrix}, \cdots, \ \Phi_{ij}^m = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \phi_{ij} \end{pmatrix}. \tag{4.1}$$

Define

$$S^c = \text{span}\{\Phi_{ij}^l \,|\, 1 \le l \le m, \ 1 \le i \le n, \ j \in N(t_i)\} \subset H(\Omega). \tag{4.2}$$

Based on this conforming subspace, we then construct a nonconforming subspace as follows. For each element $t_i \in T_h$, let

$$\tilde{\phi}_{ij}^i = \begin{cases} \phi_{ij}, \text{ on } t_i, \\ 0, \text{ otherwise.} \end{cases} \tag{4.3}$$

Note that the halved-function is identified by a shape function on a fixed reference triangle or rectangle $\hat{t}$ via an affine transformation that maps the reference element $\hat{t}$ one-to-one and onto $t_i$. We assume that the mesh $T_h$ contains at least two elements. Let

$$\tilde{\Phi}_{ij}^{1,i} = \begin{pmatrix} \tilde{\phi}_{ij}^i \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \tilde{\Phi}_{ij}^{2,i} = \begin{pmatrix} 0 \\ \tilde{\phi}_{ij}^i \\ 0 \\ \vdots \end{pmatrix}, \text{ etc..} \tag{4.4}$$

Define

$$\begin{aligned} S^n(t_i) &= \text{span}\{\tilde{\Phi}_{ij}^{l,i} \,|\, 1 \le l \le m, \ j \in N(t_i)\} \qquad (4.5) \\ &= \{\tilde{\mathbf{v}}_i \mid \tilde{\mathbf{v}}_i = \sum_{j \in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i}\, \tilde{\Phi}_{ij}^{l,i}\} \not\subseteq H(\Omega) \\ S^n &= S^n(t_1) \oplus S^n(t_2) \oplus \cdots \oplus S^n(t_n) \not\subseteq H(\Omega) \qquad (4.6) \\ S^c(t_i)^+ &= \{\mathbf{v}_i^+ \mid \mathbf{v}_i^+ = \sum_{j \in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i}\, \Phi_{ij}^l\} \qquad (4.7) \\ &= \text{span}\{\Phi_{ij}^l \,|\, 1 \le l \le m, \ j \in N(t_i)\} \subset S^c \subset H(\Omega) \end{aligned}$$

Note that the coefficients $d_{ij}^{l,i}$ of (4.5) and (4.7) are the same. The basis functions of $S^c(t_i)^+$ have support on the patch subdomain

$$T_h(i) = t_i \cup \left(\cup_{j \in N(t_i)} t_j\right). \tag{4.8}$$

The correspondence

$$\tilde{\mathbf{v}}_i = \sum_{j \in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i}\, \tilde{\Phi}_{ij}^{l,i} \in S^n(t_i) \longrightarrow \mathbf{v}_i^+ = \sum_{j \in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i}\, \Phi_{ij}^l \in S^c(t_i)^+$$

defines a one-to-one and onto mapping from $S^n(t_i)$ to $S^c(t_i)^+$. Hence $\mathbf{v}_i^+|_{t_i} = \tilde{\mathbf{v}}_i$. For any connected open subset $D$ of $\Omega$, define

$$A_D(\mathbf{w},\mathbf{v}) = A_{D,1}(\mathbf{w},\mathbf{v}) + A_{D,0}(\mathbf{w},\mathbf{v}) \qquad \forall \mathbf{v},\ \mathbf{w} \in (H^1(D))^m, \tag{4.9}$$

the restriction of $A(\cdot,\cdot)$ to $D$, with

$$\begin{aligned} A_{D,1}(\mathbf{w},\mathbf{v}) &= \frac{1}{2}\int_D C\mathbf{w}\cdot\mathbf{v}\,d\mathbf{x} \\ A_{D,0}(\mathbf{w},\mathbf{v}) &= \frac{1}{4}\int_{\partial\Omega\cap\partial D} (\mu\mathbf{w}\cdot\mathbf{v} + \mathbf{w}\cdot\mu\mathbf{v})\,ds, \end{aligned}$$

and let $\|\cdot\|_{A,\,D}$ be the restriction of the $A$-norm to the closure of the set $D$. The following two lemmas are direct consequences of the shape regularity property of the mesh $T_h$, the affine property of the basis functions of $S^n(t_i)$ and the finite dimensionality of $S^n(t_i)$ (cf. [13]).

LEMMA 4.1. *For any $\tilde{\mathbf{v}}_i = \sum_{j\in N(t_i)}\sum_{l=1}^m d_{ij}^{l,i}\,\tilde{\Phi}_{ij}^{l,i} \in S^n(t_i)$, $t_i \in T_h$, there exist two positive constants $C_3$ and $C_4$ independent of $h$ such that*

$$C_3 \sum_{j\in N(t_i)}\sum_{l=1}^m \left\| d_{ij}^{l,i}\,\tilde{\Phi}_{ij}^{l,i}\right\|_{A,\,t_i}^2 \le \|\tilde{\mathbf{v}}_i\|_{A,\,t_i}^2 \le C_4 \sum_{j\in N(t_i)}\sum_{l=1}^m \left\| d_{ij}^{l,i}\,\tilde{\Phi}_{ij}^{l,i}\right\|_{A,\,t_i}^2. \tag{4.10}$$

LEMMA 4.2. *For any $t_i \in T_h$ and $j \in N(t_i)$, there exists a positive constant $C_5$ independent of $h$ such that*

$$\left\|\Phi_{ij}^l\right\|_{A,\,t_j}^2 \le C_5 \left\|\Phi_{ij}^l\right\|_{A,\,t_i}^2. \tag{4.11}$$

The equation (3.15) can now be localized on each element $t_i$ as follows. The test and trial functions on the left side are taken to be the halved-functions of $S^n(t_i)$. On the other hand, these functions are extended from the element to form continuous basis functions on the patch subdomain $T_h(i)$ in order to account for both interior and interface residuals on the element. More specifically, the conforming problem (3.15) is modified as to determine $\tilde{\mathbf{e}}_i \in S^n(t_i)$ such that, for each $t_i \in T_h$,

$$A_{t_i}(\tilde{\mathbf{e}}_i, \tilde{\mathbf{v}}_i) = \frac{1}{2}\left(F(\mathbf{v}_i^+) - B(\mathbf{u}_h, \mathbf{v}_i^+)\right) \quad \forall \tilde{\mathbf{v}}_i \in S^n(t_i), \tag{4.12}$$

where $\mathbf{v}_i^+ \in S^c(t_i)^+$. Since each basis function $\Phi_{ij}^l$ of $S^c(t_i)^+$ has a support covering two elements $t_i$ and $t_j$, the factor $\frac{1}{2}$ appeared on the right side of equation (4.12) reflects the average weight taken on each element provided that any two neighboring elements do not differ too much, i.e., the mesh $T_h$ is locally quasi-uniform. Of course, the factor can be tuned for general meshes. As can be seen below, our error analysis is not affected by this factor so long as it remains as a constant independent of the mesh size. Moreover, we do not need any extra boundary conditions for the localized problems (4.12) since the construction of the conforming and nonconforming subspaces $S^c(t_i)^+$ and $S^n(t_i)$ takes not only both interior and flux errors but also the solvability of the problems into consideration. Using the basis functions, we can rewrite (4.12) as

$$A_{t_i}(\tilde{\mathbf{e}}_i, \tilde{\Phi}_{ij}^{l,i}) = \frac{1}{2}\left(F(\Phi_{ij}^l) - B(\mathbf{u}_h, \Phi_{ij}^l)\right) \quad \forall \tilde{\Phi}_{ij}^{l,i} \in S^n(t_i). \tag{4.13}$$

The uniqueness and existence of $\tilde{\mathbf{e}}_i$ is guaranteed since the bilinear form $A$ is an inner product of the finite dimensional space $S^n(t_i)$. Define

$$\|\mathbf{e}_a^n\|_A := \left(\sum_{i=1}^{n} \|\tilde{\mathbf{e}}_i\|_{A,\, t_i}^2\right)^{\frac{1}{2}}, \quad \mathbf{e}_a^n = \tilde{\mathbf{e}}_1 \oplus \tilde{\mathbf{e}}_2 \oplus \cdots \oplus \tilde{\mathbf{e}}_n. \tag{4.14}$$

THEOREM 4.3. *Let* $\mathbf{u} \in H(\Omega)$ *and* $\mathbf{u}_h \in S_h$ *be the solutions of* (2.2) *and* (2.3), *respectively and let* $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$. *If the conditions of Lemma 3.2 hold, then the estimated error* $\mathbf{e}_a^n$ *obtained by solving* (4.13) *has the following bounds*

$$(1-\rho)\sqrt{1-\gamma^2}\,\|\mathbf{e}\|_A \le \|\mathbf{e}_a^n\|_A \le C_6\,\|\mathbf{e}\|_A, \tag{4.15}$$

*where the constants* $\gamma \in [0,1)$, $\rho \in [0,1)$ *and* $C_6$ *are all independent of the mesh parameter* $h$.

*Proof.* Let $\mathbf{e}_a^c \in S^c$ be the solution of (3.15). By the definition of $S^c$ in (4.2), the solution can be written as

$$\mathbf{e}_a^c = \frac{1}{2}\sum_{i=1}^{n}\sum_{j\in N(t_i)}\sum_{l=1}^{m} c_{ij}^l\,\Phi_{ij}^l,$$

where $c_{ij}^l$ are the $l$-th components of the midpoint nodal values of $\mathbf{e}_a^c$ associated with the adjacent pair $t_i$ and $t_j$. Note that $c_{ij}^l = c_{ji}^l$ and $\Phi_{ij}^l = \Phi_{ji}^l \;\; \forall j \in N(t_i)$ and the summations will visit each element twice. From (3.15) and (4.13), for any fixed $i = 1, 2, \cdots, n$, we have

$$\begin{aligned} A_{t_i}(\tilde{\mathbf{e}}_i, \tilde{\Phi}_{ij}^{l,i}) &= \frac{1}{2}\left(F(\Phi_{ij}^l) - B(\mathbf{u}_h, \Phi_{ij}^l)\right) \\ &= \frac{1}{2}A(\mathbf{e}_a^c, \Phi_{ij}^l) \end{aligned}$$

for all $j \in N(t_i)$ and $1 \le l \le m$. Define

$$\mathbf{e}_i^c = \begin{cases} \mathbf{e}_a^c, & \text{on } t_i, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\mathbf{e}_i^c = \sum_{j\in N(t_i)}\sum_{l=1}^{m} c_{ij}^l\,\tilde{\Phi}_{ij}^{l,i},$$

and

$$A_{t_i}(\tilde{\mathbf{e}}_i, \mathbf{e}_i^c) = A_{t_i}(\tilde{\mathbf{e}}_i, \sum_{j\in N(t_i)}\sum_{l=1}^{m} c_{ij}^l\tilde{\Phi}_{ij}^{l,i}) = \frac{1}{2}A(\mathbf{e}_a^c, \sum_{j\in N(t_i)}\sum_{l=1}^{m} c_{ij}^l\Phi_{ij}^l).$$

It follows that

$$\begin{aligned} \|\mathbf{e}_a^c\|_A^2 &= A(\mathbf{e}_a^c, \mathbf{e}_a^c) \\ &= A(\mathbf{e}_a^c, \frac{1}{2}\sum_{i=1}^{n}\sum_{j\in N(t_i)}\sum_{l=1}^{m} c_{ij}^l\,\Phi_{ij}^l) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} A_{t_i}(\tilde{\mathbf{e}}_i, \mathbf{e}_i^c) \\
&\leq \sum_{i=1}^{n} \|\tilde{\mathbf{e}}_i\|_{A,t_i} \|\mathbf{e}_i^c\|_{A,t_i} \\
&\leq \frac{1}{2}\sum_{i=1}^{n} \left(\|\tilde{\mathbf{e}}_i\|_{A,t_i}^2 + \|\mathbf{e}_i^c\|_{A,t_i}^2\right) \\
&= \frac{1}{2}\left(\|\mathbf{e}_a^n\|_A^2 + \|\mathbf{e}_a^c\|_A^2\right).
\end{aligned}
$$

Hence

$$\|\mathbf{e}_a^c\|_A \leq \|\mathbf{e}_a^n\|_A ,$$

which together with (3.16) proves the left inequality of (4.15).

For the right inequality, we extend $\tilde{\mathbf{e}}_i = \sum_{j\in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i}\, \tilde{\Phi}_{ij}^{l,i} \in S^n(t_i)$ to a new function $\mathbf{e}_i^+$ by

$$\mathbf{e}_i^+ = \sum_{j\in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i}\, \Phi_{ij}^l,$$

which has support on the closure of the extended subdomain $T_h(i)$. Note that $\mathbf{e}_i^+ \in S^c$ and

$$
\begin{aligned}
\|\tilde{\mathbf{e}}_i\|_{A,t_i}^2 &= A_{t_i}(\tilde{\mathbf{e}}_i, \tilde{\mathbf{e}}_i) \\
&= A_{t_i}(\tilde{\mathbf{e}}_i, \sum_{j\in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i}\, \tilde{\Phi}_{ij}^{l,i}) \\
&= \frac{1}{2} \sum_{j\in N(t_i)} \sum_{l=1}^{m} d_{ij}^{l,i} A(\mathbf{e}_a^c, \Phi_{ij}^l) \\
&= \frac{1}{2} A(\mathbf{e}_a^c, \mathbf{e}_i^+) \\
&\leq \frac{1}{2} \|\mathbf{e}_a^c\|_{A,T_h(i)} \left\|\mathbf{e}_i^+\right\|_{A,T_h(i)}. \qquad (4.16)
\end{aligned}
$$

By Lemma 4.2, we have

$$\left\|d_{ij}^{l,i}\Phi_{ij}^l\right\|_{A,\,t_j}^2 \leq C_5 \left\|d_{ij}^{l,i}\Phi_{ij}^l\right\|_{A,\,t_i}^2. \qquad (4.17)$$

Hence,

$$
\begin{aligned}
\left\|\mathbf{e}_i^+\right\|_{A,T_h(i)}^2 &= \left\|\mathbf{e}_i^+\right\|_{A,\,t_i}^2 + \sum_{j\in N(t_i)} \left\|\mathbf{e}_i^+\right\|_{A,\,t_j}^2 \\
&= \|\tilde{\mathbf{e}}_i\|_{A,\,t_i}^2 + \sum_{j\in N(t_i)} \left\|\sum_{l=1}^{m} d_{ij}^{l,i}\, \Phi_{ij}^l\right\|_{A,\,t_j}^2
\end{aligned}
$$

$$
\begin{aligned}
&\leq \|\tilde{\mathbf{e}}_i\|^2_{A,\,t_i} + C_5 \sum_{j\in N(t_i)} \sum_{l=1}^{m} \left\| d^{l,i}_{ij}\, \Phi^l_{ij} \right\|^2_{A,\,t_i} \\
&= \|\tilde{\mathbf{e}}_i\|^2_{A,\,t_i} + C_5 \sum_{j\in N(t_i)} \sum_{l=1}^{m} \left\| d^{l,i}_{ij}\, \tilde{\Phi}^{l,i}_{ij} \right\|^2_{A,\,t_i} \\
&\leq \|\tilde{\mathbf{e}}_i\|^2_{A,\,t_i} + \frac{C_5}{C_3}\, \|\tilde{\mathbf{e}}_i\|^2_{A,\,t_i} \\
&\leq C_7\, \|\tilde{\mathbf{e}}_i\|^2_{A,\,t_i}\,, \qquad (4.18)
\end{aligned}
$$

where $C_7 = 1 + \frac{C_5}{C_3}$. From (4.16) and (4.18), we have

$$
\|\tilde{\mathbf{e}}_i\|_{A,t_i} \leq \frac{\sqrt{C_7}}{2}\, \|\mathbf{e}^c_a\|_{A,T_h(i)}\,.
$$

Therefore,

$$
\begin{aligned}
\|\mathbf{e}^n_a\|^2_A &= \sum_{i=1}^{n} \|\tilde{\mathbf{e}}_i\|^2_{A,\,t_i} \\
&\leq \frac{C_7}{4} \sum_{i=1}^{n} \|\mathbf{e}^c_a\|^2_{A,T_h(i)} \\
&\leq \frac{C_7}{4} \max_i\{|T_h(i)|\}\, \|\mathbf{e}^c_a\|^2_A \\
&\leq \frac{5C_7}{4}\, \|\mathbf{e}^c_a\|^2_A\,, \qquad (4.19)
\end{aligned}
$$

where $|T_h(i)|$ denotes the number of elements of the subdomain, which is less than or equal to 5 (a rectangle having 4 adjacent elements at most). Combining (3.16) and (4.19), the right inequality of (4.15) then follows with $C_6 = \frac{\sqrt{5C_7}}{2}$ which is independent of $h$. This completes the proof of the theorem. □

**5. Model Problems.** The purpose of this section is to apply the weak residual error estimation to the following three classes of problems which exemplify the importance of the Friedrichs theory. Verification of the corresponding symmetric positiveness to these problems can be found in the cited references. We only verify the conditions posed in Lemma 3.2. As for numerical results, we refer to [20, 21].

The condition (3.9) holds for all hierarchical shape functions (for the complementary spaces $S^c$) such as those of [28]. It also holds for some particular shape functions such as those of [20, 21], which are particularly constructed for the forward-backward heat equation. If both conforming and nonconforming formulas (3.15) and (4.13) are used, we obtain a general error estimation for all $h$, $p$ and $hp$ finite element approximations. For example, assuming that the current approximation is of order $p$, if the next hierarchical shape functions of degree $p+1$ are internal modes, we use (3.15) to calculate the error estimator. Otherwise, we use (4.13) for side modes.

We are now only left to verify the strengthened Cauchy-Schwarz inequality (3.13). By a similar argument of (3.14), we can partition the bilinear form $A_0$ into a finite number of bilinear forms $A_{\Gamma_i}$, $\partial\Omega = \cup\overline{\Gamma}_i$, by restricting $A_0$ to each one of the finite number boundaries $\Gamma_i$ and prove the inequality for each individual bilinear form. Since the matrix $\mu$ may or may not be positive definite on these individual boundaries, the proof can be categorized into two cases according to the following two types of the matrix $\mu$. We say that the matrix $\mu$ is *invertible* on $\Gamma$ if the determinant of the matrix $\mu + \mu^t$ is nonzero for all $(x,y) \in \Gamma$ and is *noninvertible* if the determinant is zero for some $(x,y) \in \Gamma$.

If the matrix $\mu$ is invertible on $\Gamma$, the corresponding bilinear form $A_\Gamma$ defines an inner product, denoted by $A_\Gamma$-inner product, for the space $[L^2(\Gamma)]^m$, which is equivalent to the $L^2(\Gamma)$-inner product that induces the usual $[L^2(\Gamma)]^m$ norm. The previous argument in [10, 15] for the $L^2(\Omega)$- or $H^1(\Omega)$-inner product to prove the strengthened Cauchy-Schwarz inequality goes word for word for the $L^2(\Gamma)$-inner product and consequently for the equivalent $A_\Gamma$-inner product. We therefore only have to prove for the case that the matrix is not invertible.

Let $q(\tau)$ be independent of the mesh parameter $h$ and be a positive and bounded continuous function for all $\tau$ in the interval $(0,1)$. Let the bilinear form $\widehat{A}$ be defined by

$$\widehat{A}(w,v) = \int_0^1 q(\tau)w(\tau)v(\tau)\,d\tau \quad \forall w,v \in L^2(0,1). \tag{5.1}$$

LEMMA 5.1. *Let $S_1$ and $S_2$ be two nonempty finite dimensional subspaces of $L^2(0,1)$ such that $S_1 \cap S_2 = \{0\}$. Then there exists a constant $\widehat{\gamma} \in [0,1)$ independent of $h$ such that*

$$\left|\widehat{A}(w,v)\right| \le \widehat{\gamma}\left[\widehat{A}(w,w)\right]^{1/2}\left[\widehat{A}(v,v)\right]^{1/2} \quad \forall w \in S_1,\ v \in S_2. \tag{5.2}$$

*Proof.* The bilinear form obviously defines an inner product for $L^2(0,1)$, which is a fixed geometric operator on $L^2(0,1)$ depending only on the function $q$ and determines an "angle" between any two functions in the space. Since the basis functions of $S_1$ and of $S_2$ are finite and linearly independent, there must be a nonzero angle between any two functions $w \in S_1$ and $v \in S_2$, i.e., there exists a constant $\widehat{\gamma} \in [0,1)$ independent of $h$ such that (5.2) holds. □

**Example 5.1. The Neutron Transport Equation.**

$$\begin{cases} \nabla u \cdot \mathbf{d} + u = f & \text{in} \quad \Omega = (0,1) \times (0,1) \\ u = 0 & \text{on} \quad \partial\Omega_-, \end{cases}$$

where $\mathbf{d} = (1,1)$ and $\partial\Omega_-$ is the inflow boundary defined by

$$\begin{aligned} \partial\Omega_- &= \{(x,y) \in \partial\Omega : \boldsymbol{\nu}(x,y) \cdot \mathbf{d} < 0\} \\ &= \{\,(0,y)^t : y \in (0,1)\,\} \cup \{\,(x,0)^t : x \in (0,1)\,\}. \end{aligned}$$

This problem is symmetric positive [17] with $m = 1$, $M_1 = M_2 = M_0 = 1$, $\beta = \boldsymbol{\nu} \cdot \mathbf{d}$ and $\mu = |\beta| = 1$. Since the matrix $\mu$ is invertible, the inequality (3.13) holds for this example.

**Example 5.2. The Forward-Backward Heat Equation.**

$$\begin{cases} x\phi_y - \phi_{xx} = f_1 & \text{in } \Omega = (-1,1) \times (0,1) \\ \phi(\pm 1, y) = 0 & \forall y \in [0,1] \\ \phi(x,0) = 0 & \forall x \in [0,1] \\ \phi(x,1) = 0 & \forall x \in [-1,0]. \end{cases}$$

Note that the equation changes its type as $x$ changes sign in $\Omega$. There have been a number of papers addressing to this kind of mixed-type heat equations, for further references see [6, 29]. After the change of variables

$$\mathbf{u} = (u_1,\, u_2)^t = \left(e^{-0.1y}\phi,\; e^{-0.1y}\phi_x\right)^t,$$

its corresponding system (2.1) is symmetric positive [6] with

$$M_1 = \begin{pmatrix} -x & -1 \\ -1 & 0 \end{pmatrix},\; M_2 = \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix},\; M_0 = \begin{pmatrix} 0.1x & x \\ 0 & 1 \end{pmatrix},\; \mathbf{f} = \begin{pmatrix} e^{0.1y} f_1 \\ 0 \end{pmatrix},$$

and $\mu$ shown in Table 5.1 where the boundary $\partial\Omega = \Gamma_1 \cup \cdots \cup \Gamma_6$. For the matrix $\beta$, we refer to [6].

Note that the boundary matrix $\mu$ is noninvertible on all the boundary segments. On the boundary segment $\Gamma_1$, for any $\mathbf{w} = (w_1,\, w_2)^t \in S_h$ and $\mathbf{v} = (v_1,\, v_2)^t \in S_h^c$, we have

$$\frac{1}{2}(\mu\mathbf{w}\cdot\mathbf{v} + \mathbf{w}\cdot\mu\mathbf{v}) = -x w_1 v_1.$$

Assume that $w_1 \neq 0$ and $v_1 \neq 0$ since, otherwise, the strengthened Cauchy-Schwarz inequality is trivially satisfied. Let $t_i \in T_h$ be such an element that $\partial t_i \cap \Gamma_1$ is not empty, i.e., the intersection is an interval $(a,b)$ for some $a$ and $b$ such that $-1 \leq a < b \leq 0$. By Lemma 5.1, there exist two constants $\widehat{\gamma}_i \in [0,1)$, $i = 1,2$, independent of $h$ such that

$$\begin{aligned} & \left| \int_{\partial t_i \cap \Gamma_1} \frac{1}{2}(\mu\mathbf{w}\cdot\mathbf{v} + \mathbf{w}\cdot\mu\mathbf{v})\, ds \right| \\ = & \left| \int_a^b -x w_1(x)\, v_1(x)\, dx \right| \\ = & \left| \int_0^1 (-a(1-\tau) - b\tau)\; \hat{w}_1 \hat{v}_1\, d\tau \right| \\ = & -a \left| \int_0^1 (1-\tau)\hat{w}_1\hat{v}_1\, d\tau \right| - b \left| \int_0^1 \tau\; \hat{w}_1\hat{v}_1\, d\tau \right| \\ \leq & -a\widehat{\gamma}_1 \left[ \int_0^1 (1-\tau)\hat{w}_1^2\, d\tau \right]^{\frac{1}{2}} \left[ \int_0^1 (1-\tau)\hat{v}_1^2\, d\tau \right]^{\frac{1}{2}} \\ & -b\widehat{\gamma}_2 \left[ \int_0^1 \tau\hat{w}_1^2\, d\tau \right]^{\frac{1}{2}} \left[ \int_0^1 \tau\hat{v}_1^2\, d\tau \right]^{\frac{1}{2}} \end{aligned}$$

Table 5.1. The boundary matrix $\mu$ of Example 5.2.

| | $2\mu$ |
|---|---|
| $\Gamma_1 = \{x \in [-1,0],\ y = 0\}$ | $\begin{pmatrix} -x & 0 \\ 0 & 0 \end{pmatrix}$ |
| $\Gamma_2 = \{x = -1,\ y \in [0,1]\}$ | $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ |
| $\Gamma_3 = \{x \in [-1,0],\ y = 1\}$ | $\begin{pmatrix} -x & 0 \\ 0 & 0 \end{pmatrix}$ |
| $\Gamma_4 = \{x \in [0,1],\ y = 1\}$ | $\begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}$ |
| $\Gamma_5 = \{x = 1,\ y \in [0,1]\}$ | $\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ |
| $\Gamma_6 = \{x \in [0,1],\ y = 0\}$ | $\begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}$ |

$$
\begin{aligned}
&\le \max\{\widehat{\gamma}_1, \widehat{\gamma}_2\} \left[\int_0^1 (-a(1-\tau) - b\tau)\, \hat{w}_1^2\, d\tau\right]^{\frac{1}{2}} \left[\int_0^1 (-a(1-\tau) - b\tau)\, \hat{v}_1^2\, d\tau\right]^{\frac{1}{2}} \\
&\le \gamma_2 \left[\int_{\partial t_i \cap \Gamma_2} \mu \mathbf{w} \cdot \mathbf{w}\, ds\right]^{\frac{1}{2}} \left[\int_{\partial t_i \cap \Gamma_2} \mu \mathbf{v} \cdot \mathbf{v}\, ds\right]^{\frac{1}{2}},
\end{aligned}
$$

where $\widehat{\gamma} = \max\{\widehat{\gamma}_1, \widehat{\gamma}_2\} \in [0,1)$. For all other segments, the proof proceeds in the same way. Therefore, we have (3.13) for this example.

**Example 5.3. The Tricomi Equation.**

$$
\begin{cases}
y\phi_{xx} - \phi_{yy} = f_1 & \text{in } \Omega \\
\frac{d\phi}{ds} = 0 & \text{on } \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \\
\phi_x = \phi_y & \text{on } \Gamma_4.
\end{cases}
$$

Here the domain $\Omega$ is bounded by the five curve segments $\Gamma_i$, $i = 1, \cdots, 5$, which are given in Table 5.2. The equation is hyperbolic in the region for $y > 0$, parabolic for $y = 0$ and elliptic for $y < 0$. This is a classical mixed-type problem for which Friedrichs formulated it as a symmetric positive system in his renowned paper [16]. Using the change of variables

$$
\mathbf{u} = (u_1, u_2)^t = (\phi_x, \phi_y)^t,
$$

the corresponding system can be obtained with

$$
M_1 = \begin{pmatrix} -\frac{x+3}{2}y & y \\ y & -\frac{x+3}{2} \end{pmatrix},\ M_2 = \begin{pmatrix} -y & \frac{x+3}{2} \\ \frac{x+3}{2} & -1 \end{pmatrix},\ M_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\ \mathbf{f} = \begin{pmatrix} -\frac{x+3}{2}f_1 \\ 0 \end{pmatrix},
$$

and $\mu$ shown in Table 5.2. For the matrix $\beta$, we refer to [18].

Table 5.2. The boundary matrix $\mu$ of Example 5.3.

| | $\mu$ |
|---|---|
| $\Gamma_1 = \{x \in [0,1],\ y^3 = \frac{9}{4}(x-1)^2\}$ | $\frac{x+3+2\sqrt{y}}{2\sqrt{1+y}} \begin{pmatrix} y & -\sqrt{y} \\ -\sqrt{y} & 1 \end{pmatrix}$ |
| $\Gamma_2 = \{x = 1,\ y \in [-1,0]\}$ | $\begin{pmatrix} -2y & y \\ y & 2-y \end{pmatrix}$ |
| $\Gamma_3 = \{x \in [-1,1],\ y = -1\}$ | $\frac{1}{2} \begin{pmatrix} (x+3)^2+2 & -(x+3) \\ -(x+3) & 2 \end{pmatrix}$ |
| $\Gamma_4 = \{x = -1,\ y \in [-1,0]\}$ | $\begin{pmatrix} -y & y \\ y & 1-2y \end{pmatrix}$ |
| $\Gamma_5 = \{x \in [-1,0],\ y^3 = \frac{9}{4}(x+1)^2\}$ | $\frac{x+3+2\sqrt{y}}{2\sqrt{1+y}} \begin{pmatrix} y & \sqrt{y} \\ \sqrt{y} & 1 \end{pmatrix}$ |

The matrix $\mu$ is invertible on $\Gamma_3$ and is noninvertible on the other segments. With some transformations on $\Gamma_1$ and on $\Gamma_5$ to straight lines, the proof of the strengthened inequality for all segments other than $\Gamma_3$ is similar. We only prove for $\Gamma_2$.

LEMMA 5.2. *For any given constants c and d, denote the matrix*

$$\mu_{c,d} = \begin{pmatrix} c^2 & cd \\ cd & d^2 \end{pmatrix}. \tag{5.3}$$

*Then the strengthened inequality holds for the bilinear form $A_{c,d}$ defined by*

$$A_{c,d}(\mathbf{w},\mathbf{v}) = \int_a^b q(s)\mu_{c,d}\mathbf{w}\cdot\mathbf{v}\,ds \quad \forall \mathbf{w},\mathbf{v} \in \left[L^2(a,b)\right]^2, \tag{5.4}$$

*where $q(s)$ is a positive and bounded continuous function on the interval $(a,b)$.*

*Proof.* For any $\mathbf{w} = (w_1,\ w_2)^t \in S_h$ and $\mathbf{v} = (v_1,\ v_2)^t \in S_h^c$, (5.4) can be written as

$$A_{c,d}(\mathbf{w},\mathbf{v}) = \int_a^b q(s)\,(cw_1 + dw_2)\,(cv_1 + dv_2)\,ds,$$

where $(cw_1 + dw_2)$ and $(cv_1 + dv_2)$ are linearly independent. The proof then proceeds in a similar way as that in the previous example. □

Again, let $t_i \in T_h$ be such an element that $\partial t_i \cap \Gamma_2$ is not empty. For any $\mathbf{w} = (w_1,\ w_2)^t \in S_h$ and $\mathbf{v} = (v_1,\ v_2)^t \in S_h^c$, we have, for $-1 \le a < b \le 0$,

$$\begin{aligned} \int_{\partial t_i\cap\Gamma_2} \frac{1}{2}(\mu\mathbf{w}\cdot\mathbf{v} + \mathbf{w}\cdot\mu\mathbf{v})\,dy &= \int_a^b \left[(-y)\mu_{1,0} + 2\,\mu_{0,1} + (-y)\,\mu_{1,-1}\right]\,\mathbf{w}\cdot\mathbf{v}\,dy \\ &= A_{1,0}(\mathbf{w},\mathbf{v}) + A_{0,1}(\mathbf{w},\mathbf{v}) + A_{1,-1}(\mathbf{w},\mathbf{v}), \end{aligned} \tag{5.5}$$

where the corresponding functions $q(y) = -y$, $q(y) = 2$ and $q(y) = -y$ are all positive and bounded continuous functions on $\Gamma_2$. Consequently, the bilinear form (5.5) satisfies the strengthened inequality. We therefore conclude that (3.13) holds for the Tricomi equation.

## REFERENCES


[1] S. Adjerid and J. E. Flaherty, *Second-order finite element approximations and a posteriori error estimation for two-dimensional parabolic systems,* Numer. Math., 53 (1988), pp. 183-198.

[2] S. Adjerid, J. E. Flaherty and Y. J. Wang, *A posteriori error estimation with finite element methods of lines for one-dimensional parabolic systems,* Numer. Math., 65 (1993), pp. 1-21.

[3] 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.

[4] O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems, Theory and Computation,* Academic Press Inc. (1984), pp. 218-226.

[5] A. K. Aziz and S. H. Leventhal, *Numerical solution of linear partial differential equations of elliptic-hyperbolic type,* in Numerical Solution of Partial Differential Equations III, B. E. Hubbard, ed., Academic Press (1976), pp. 55-88.

[6] A. K. Aziz and J. -L. Liu, *A weighted least squares method for the backward-forward heat equation,* SIAM J. Numer. Anal., 28 (1991), pp. 156-167.

[7] I. Babuška and A. Miller, *The post-processing approach in the finite element methods, Part 1: Calculation of displacements, stresses and other higher derivatives of the displacements,* Int. J. Numer. Meth. Eng., 20 (1984), pp. 1085-1109, *Part 2: The calculation of stress intensity factors,* Int. J. Numer. Meth. Eng., 20 (1984), pp. 1111-1129, *Part 3: A posteriori error estimates and adaptive mesh selection,* Int. J. Numer. Meth. Eng., 20 (1984), pp. 2311-2324.

[8] I. Babuška and W. C. Rheinboldt, *Error estimates for adaptive finite element computations,* SIAM J. Numer. Anal., 15 (1978), pp. 736-754.

[9] I. Babuška and D. Yu, *Asymptotically exact a posteriori error estimator for bi-quadratic elements,* Fin. Elem. in. Anal. & Design, 3 (1987), pp. 341-354.

[10] R. E. Bank, *Hierarchical bases and the finite element method,* Acta Numerica 1996 (A. Iserles, ed.), Cambridge University Press (1996), pp. 1-43.

[11] R. E. Bank and R. K. Smith, *A posteriori error estimates based on hierarchical bases,* SIAM J. Numer. Anal. 30 (1993), pp. 921-935.

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

[13] F. A. Bornemann, B. Erdmann and R. Kornhuber, *A posteriori error estimates for elliptic problems in two and three space dimensions,* SIAM J. Numer. Anal., 33 (1996), pp. 1188-1204.

[14] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems,* North-Holland, New York, 1978

[15] V. Eijkhout and P. Vassilevski, *The role of the strengthened Cauchy-Buniakowskii-Schwarz inequality in multilevel methods,* SIAM Rev., 33 (1991), pp. 405-419.

[16] K. O. Friedrichs, *Symmetric positive differential equations,* Comm. Pure Appl. Math., 11 (1958), pp. 333-418.

[17] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method,* Cambridge University Press, 1990.

[18] T. Katsanis, *Numerical solution of Tricomi equation using theory of symmetric positive differential equations,* SIAM J. Numer. Anal., 6 (1969), pp. 236-253.

[19] P. Lesaint, *Finite element methods for symmetric hyperbolic equations,* Numer. Math., 21 (1973), pp. 244-255

[20] J.-L. Liu, *Weak residual error estimates for symmetric positive systems,* Numer. Funct. Anal. Optimiz., 14 (1993), pp. 607-619.

[21] J.-L. Liu, *On weak residual error estimation,* SIAM J. on Sci. Comput., 17 (1996), pp. 1249-1268.

[22] J.-L. Liu, I.-J. Lin, M.-Z. Shih, R.-C. Chen and M.-C. Hsieh, *Object oriented programming of adaptive finite element and finite volume methods,* Appl. Numer. Math., 21 (1996) pp. 439-467.

[23] J.-L. Liu and W. C. Rheinboldt, *A posteriori finite element error estimators for indefinite elliptic boundary value problems,* Numer. Funct. Anal. Optimiz., 15 (1994), pp. 335-356.

[24] J.-L. Liu and W. C. Rheinboldt, *A posteriori finite element error estimators for parametrized nonlinear boundary value problems,* Numer. Funct. Anal. Optimiz., 17 (1996), pp.605-637.

[25] J. T. Oden, L. Demkowicz, W. Rachowicz, and T. A. Westermann, *Toward a universal h-p adaptive finite element strategy, Part 2. A posteriori error estimation,* Compt. Meth. Appl. Mech. Engng., 77 (1989), pp. 113-180.

[26] J. T. Oden, W. Wu, and M. Ainsworth, *An a posteriori error estimate for finite elmenent approximations of the Navier-Stokes equations,* Compt. Meth. Appl. Mech. Engng., 111 (1994), pp. 185-202.

[27] E. Süli, *A posteriori error analysis and global error control for adaptive finite element approximations of hyperbolic problems,* In: D. F. Griffiths and G. A. Watson, eds., Proceedings of the 16th International Conference on Numerical Analysis, Dundee, 1995, Pitman Research Notes in Mathematics, Longman Scientific and Technical, Harlow.

[28] B. A. Sźabo and I. Babuška, *Finite Element Analysis,* Wiley, New York, 1991.

[29] V. Vanaja and R. B. Kellogg, *Iterative methods for a forward-backward heat equation,* SIAM J. Numer. Anal., 27 (1990), pp. 622-635.

[30] R. Verfürth, *A posteriori error estimates for the Stokes equations,* Numer. Math., 55 (1989) pp. 309-325.

[31] O. C. Zienkiewicz and J. Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis,* Int. J. Numer. Meth. Engrg., 24 (1987), pp. 337-357.