A POSTERIORI FINITE ELEMENT ERROR ESTIMATION FOR DIFFUSION PROBLEM

SLIMANE ADJERID†, BELKACEM BELGUENDOUZ†, AND JOSEPH E. FLAHERTY†

Abstract. Adjerid et al. [2] and Yu [19, 20] show that a posteriori estimates of spatial discretization errors of piecewise bi-p polynomial finite element solutions of elliptic and parabolic problems on meshes of square elements may be obtained from jumps in solution gradients at element vertices when \( p \) is odd and from local elliptic or parabolic problems when \( p \) is even. We show that these simple error estimates are asymptotically correct for other finite element spaces. The key requirement is that the trial space contain all monomial terms of degree \( p + 1 \) except for \( x_1^{p+1} \) and \( x_2^{p+1} \) in a Cartesian \((x_1, x_2)\) frame. Computational results show that the error estimates are accurate, robust, and efficient for a wide range of problems, including some that are not supported by the present theory. These involve quadrilateral-element meshes, problems with singularities, and nonlinear problems.

Key words. Finite element methods, a posteriori error estimation, \( p \)-refinement, hierarchical approximations, elliptic and parabolic partial differential equations.

AMS subject classifications. 65M60, 65M20, 65M15, 65M50.

1. Introduction. A posteriori error estimates are a standard ingredient of adaptive finite element software. They are used to appraise the accuracy of computed solutions and to control adaptive enrichment through \( h-\), \( p-\), and/or \( r\)-refinement. Successful techniques for estimating spatial discretization errors of finite element solutions of elliptic and parabolic problems are often based on residual correction with \( p\)-refinement [18]. Using this strategy, an error estimate is obtained in a space of piecewise polynomials having higher degree than used for the original solution by solving a finite element Galerkin problem with solution residuals as loading. The error estimation problem may be localized to the element level to avoid a global assembly and solution; hence, reducing computational cost. Localization typically involves estimating solution gradients at element boundaries [4, 10, 18] and the neglect of errors at certain points, lines, or surfaces [3, 9, 19, 20].

Babuska and Yu [9] considered the solution of linear two-dimensional elliptic problems on squares using piecewise bi-p polynomial spaces and showed that error estimates could be constructed from jumps in solution gradients at element vertices when \( p \) is odd and from local elemental solution residuals when \( p \) is even. Yu [19, 20] proved that error estimates computed in this manner are asymptotically exact; hence, they converge to zero under mesh refinement at the same rate and with the same constant as the actual finite element error. Adjerid et al. [2, 3] established similar results for the finite element method-of-lines solution of one- and two-dimensional parabolic problems.

Both odd- and even-order error estimation procedures are computationally simple. The odd-order estimates only require jumps in solution gradients at the four element vertices and neither element nor edge residuals are needed. Only nearest-
neighbor interaction is necessary; thus, simplifying implementation on a parallel computer. Gradient jumps may be shared between elements with a common vertex to halve the cost relative to element-by-element computation. The even-order elliptic and parabolic estimates are local to the element. No off-element communication is necessary; hence, there is no search for neighbor information and parallelization is perfect. Computations (§5) imply that the even-order estimates improve with increasing polynomial degree.

Piecewise tensor-product spaces are not as efficient as serendipity [21] or hierarchical [12, 17] approximations which have fewer degrees of freedom for the same order of accuracy. Herein, we show that the error estimates of Babuska and Yu [9, 19, 20] or Adjerid et al. [2, 3] converge to the actual error for a wider class of finite element spaces. The important consideration is that a solution space of order \( p \) contain all monomial terms \( x_1^{p+1-k}x_2^{k} \) of degree \( p + 1 \) except \( x_1^{p+1} \) and \( x_2^{p+1} \). These spaces have slightly larger dimension than the usual serendipity or hierarchical bases, but far less than the bi-p spaces (cf. Fig. 1). In return, for the modest increase in solution complexity relative to serendipity and hierarchical bases, the solution will be supported by a simple and asymptotically correct error estimate.

After stating the linear elliptic and parabolic Galerkin problems under consideration (§2), we review the odd and even error estimation procedures for piecewise bi-p polynomial spaces (§3). The procedure for constructing more general finite element spaces for which these error estimates apply is described in §4. Establishing asymptotic correctness of the error estimates for elliptic and parabolic problems using these new finite element spaces follows the earlier arguments used for bi-p polynomial spaces. Because these proofs are lengthy and involved, we have not duplicated the arguments but, rather, refer to the earlier analyses [2, 3, 9, 19, 20].

Although the odd and even error estimation procedures apply to many finite element spaces, we show how to modify standard hierarchical spaces [17] by the addition of certain interior ("bubble") modes (§4). The theory has only been developed for linear elliptic and parabolic problems on rectangular-element meshes; however, several examples (§5) reveal that the error estimates work more generally. In particular, they appear to be reliable and robust on some non-rectangular-element meshes, highly-graded meshes in the presence of singularities, and nonlinear problems. They even appear to apply to triangular elements [14]. Like many other error estimation techniques [5, 6, 7, 8, 16], the odd and even estimates may perform poorly on highly irregular meshes and when singularities introduce pollution errors.

2. Problem formulation. Consider the linear, scalar, two-dimensional parabolic partial differential equation

\[
\begin{align*}
\partial_t u + Lu &= f(x), \quad x = [x_1, x_2]^T \in \Omega, \\
Lu &= -\sum_{j=1}^{2}\sum_{k=1}^{2} \partial_{x_j}(a_{j,k}(x)\partial_{x_k} u) + b(x)u,
\end{align*}
\]

subject to the initial and Dirichlet boundary conditions

\[
\begin{align*}
u(x, 0) &= \varphi(x), \quad x \in \Omega \cup \partial\Omega, \\
u(x, t) &= 0, \quad x \in \partial\Omega, \quad t > 0.
\end{align*}
\]
The variables \( \mathbf{x} = [x_1, x_2]^T \) and \( t \) denote spatial and temporal coordinates, \( \partial_\alpha \) denotes partial differentiation with respect to \( \alpha \), and \( \Omega \) is a bounded piecewise rectangular domain with boundary \( \partial \Omega \). The functions \( a_{j,k}(\mathbf{x}) \), \( j, k = 1, 2 \), and \( b(\mathbf{x}) \) are smooth with \( L \) being a positive definite operator. Our results also hold for elliptic problems upon neglect of temporal dependence in (2.1).

The Galerkin form of (2.1) consists of determining \( u \in H^1_0 \) satisfying

\[
(2.2a) \quad (v, \partial_t u) + A(v, u) = (v, f), \quad t > 0,
\]

\[
(2.2b) \quad A(v, u) = A(v, u^0), \quad t = 0, \quad \forall v \in H^1_0,
\]

where the strain energy and \( L^2 \) inner products, respectively, are

\[
(2.2c) \quad A(v, u) = \int_\Omega \int_\Omega \sum_{j=1}^{2} \sum_{k=1}^{2} a_{j,k}(\mathbf{x}) \partial_\alpha x_j \partial_\alpha x_k u + b(\mathbf{x}) vu \, dx_1 dx_2
\]

and

\[
(2.2d) \quad (v, u) = (v, u)_0 = \int_\Omega \int_\Omega w \, dx_1 dx_2.
\]

As usual, functions in the Sobolev space \( H^s \), \( s \geq 0 \), have the inner product and norm

\[
(2.2e) \quad (v, u)_s = \sum_{|\alpha| \leq s} \langle \partial_\alpha x_1, \partial_\alpha x_2, \partial_\alpha x_1, \partial_\alpha x_2 u \rangle,
\]

\[
(2.2f) \quad \|u\|_s^2 = (u, u)_s,
\]

where \( |\alpha| = \alpha_1 + \alpha_2 \) with \( \alpha_1 \) and \( \alpha_2 \) being non-negative integers. A 0 subscript on \( H^1 \) implies that functions also satisfy (2.1d).

Finite element solutions of (2.2a, b) are obtained by approximating \( H^1 \) by a finite-dimensional subspace \( S^{N,p} \) and determining \( U \in S^{N,p}_0 \) such that

\[
(2.3a) \quad (V, \partial_t U) + A(V, U) = (V, f), \quad t > 0,
\]

\[
(2.3b) \quad A(V, U) = A(V, u^0), \quad t = 0, \quad \forall V \in S^{N,p}_0.
\]

Partitioning \( \Omega \) into a mesh of quadrilateral elements \( \Delta_i, i = 1, 2, \ldots, N \), define \( S^{N,p} \) as

\[
(2.4) \quad S^{N,p} = \{ w \in H^1 \mid w(\mathbf{x}) \in Q_p(\Delta_i), \quad \mathbf{x} \in \Delta_i, \quad i = 1, 2, \ldots, N \}
\]

where \( Q_p \) involves polynomials of order \( p \geq 1 \) that will be defined in §3 and §4 for bi-\( p \) and hierarchical approximations, respectively.
3. Error estimation procedure. The analyses of Yu [19, 20] and Adjerid et al. [2] on square elements suggest how to construct error estimates on quadrilateral elements and we develop such estimates for odd- and even-degree approximations in §3.1 and §3.2, respectively. To begin, let
\begin{equation}
e(x,t) = u(x,t) - U(x,t)
\end{equation}
denote the discretization error of the semi-discrete problem (2.3); \( \bar{\pi} \) denote a univariate operator of degree \( p \geq 1 \) that interpolates functions in \( H^1_0[-1,1] \) at the Lobatto points of degree \( p + 1 \); and
\begin{equation}
\hat{\psi}_{p+1}(\xi) = \xi^{p+1} - \bar{\pi}\xi^{p+1}.
\end{equation}
Thus, \( \hat{\psi}_{p+1}(\xi) \) vanishes at the Lobatto points and \( \hat{\psi}'_{p+1}(\xi) \) is proportional to the Legendre polynomial \( P_p(\xi) \) of degree \( p \), with \( (\cdot)' \) denoting ordinary differentiation.

Continuing, let \( \bar{\pi} \) denote the two-dimensional operator that interpolates functions as tensor products of the one-dimensional interpolants \( \bar{\pi} \) in the \( \xi_1 \) and \( \xi_2 \) directions; \( x(\xi_1,\xi_2) \) be the bilinear mapping from the “canonical element” \( \{ (\xi_1,\xi_2) \mid -1 \leq \xi_1, \xi_2 \leq 1 \} \) to \( \Delta_i \); and \( \hat{f}(\xi_1,\xi_2) = f(x(\xi_1,\xi_2)) \). Then define the interpolant \( \bar{\pi}_i \) on \( \Delta_i \) as \( \bar{\pi}_i f(x) = \bar{\pi} f(\xi_1(x),\xi_2(x)) \). We omit the elemental index \( i \) whenever confusion is unlikely.

3.1. Error estimates for odd-degree approximations. Error estimates of odd-degree approximations are constructed by assuming that the restriction of \( e(x) \) to element \( \Delta \) has the form
\begin{equation}
e(x,t) \approx E(x,t) = b_1(t)\psi_{p+1,1}(x) + b_2(t)\psi_{p+1,2}(x)
\end{equation}
with
\begin{equation}
\psi_{p+1,j}(x) = x_j^{p+1} - \bar{\pi}x_j^{p+1}, \quad j = 1, 2, \quad x \in \Delta.
\end{equation}

If \( u(x,t) \in C^1(\Omega) \), we may use (3.1) and (3.3a) to compute jumps in the spatial derivatives of \( e(x,t) \) at the vertices \( p_k \), \( k = 1, 2, 3, 4 \), of \( \Delta_i \) as
\[ [\partial_{x_j} e(p_k,t)]_j = -[\partial_{x_j} U(p_k,t)]_j \]
\[ \approx b_1(t)[\partial_{x_j}\psi_{p+1,1}(p_k)]_j + b_2(t)[\partial_{x_j}\psi_{p+1,2}(p_k)]_j, \]
\begin{equation}
\end{equation}
\[ j = 1, 2, \quad k = 1, 2, 3, 4, \quad x \in \Delta_i, \]
where \( [g(p)]_j \) denotes the jump in \( g \) at point \( p \) in the \( x_j \) direction.

The jumps \( [\partial_{x_j} U(p_k,t)]_j \) in the finite element solution derivatives are known in each coordinate direction \( (j = 1, 2) \) at each vertex \( (k = 1, 2, 3, 4) \) for elements not adjacent to \( \partial\Omega \). Thus, at each vertex \( p_k \), we may determine a solution \( (b_{1,k}, b_{2,k}), k = 1, 2, 3, 4, \) of (3.4) by solving
\begin{equation}
b_{1,k}(t)[\partial_{x_j}\psi_{p+1,1}(p_k)]_j + b_{2,k}(t)[\partial_{x_j}\psi_{p+1,2}(p_k)]_j = -[\partial_{x_j} U(p_k,t)]_j,
\end{equation}
\begin{equation}
\end{equation}
\[ j = 1, 2, \quad k = 1, 2, 3, 4. \]
These, in turn, can be used with (3.3a) to compute four error estimates on $\Delta_i$ which are averaged to obtain

\[
||E(\cdot, t)||_{1,i}^2 = \frac{1}{4} \sum_{k=1}^{4} ||b_{1,k}(t)\psi_{p+1,1}(\cdot) + b_{2,k}(t)\psi_{p+1,2}(\cdot)||_{1,i}^2,
\]

where the local $H^1$ norm $||\cdot||_{1,i}$ is defined like its global counterpart with $\Delta_i$ replacing $\Omega$ in (2.2). A global error estimate is obtained as

\[
||E(\cdot, t)||_{1}^2 = \sum_{i=1}^{N} ||E(\cdot, t)||_{1,i}^2.
\]

When $\Delta_i$ is adjacent to $\partial\Omega$ or (in the rare case) when (3.5) is singular, (3.6a) is obtained either by averaging (3.5) over the other vertices of $\Delta_i$ or by solving (3.5) in a least-squares sense using jumps at vertices across all edges of $\Delta_i$ except $\partial\Omega$. Such non-uniqueness may be expected and occurs in other error-estimation procedures [4] when local approximations are used to solve a global problem. Other algebraic solution techniques can be envisioned; however, the one described here provides acceptable accuracy.

When the mesh is rectangular and $Q_p$ is a tensor product of one-dimensional polynomials through degree $p$, the error estimates (3.5, 6) reduce to a case similar to the one considered by Adjerid et al. [2] and Yu [20] and, hence, they may be proven to be asymptotically correct as indicated by the following theorem.

**Theorem 3.1.** Let $\Omega$ be a rectangle that has been partitioned into $N$ rectangular $h_{i,1} \times h_{i,2}$ elements $\Delta_i$, $i = 1, 2, ..., N$. Let positive constants $c$ and $C$ exist such that

\[
c \leq \frac{h_{adj(i,k),j}}{h_{i,j}} \leq C, \quad j = 1, 2, \quad k = 1, 2, 3, 4, \quad i = 1, 2, ..., N,
\]

(3.7)

where $h_{adj(i,k),j}$ is the length of the edge of the element adjacent to $\Delta_i$ in the $x_j$ direction and sharing vertex $k$. Further let $u \in H^1 \cap H^{p+2}$ and $U \in S^N_p$ be solutions of (2.2) and (2.3), respectively, where $p \geq 1$ is an odd integer. Then

\[
||e(\cdot, t)||_{1i}^2 = ||E(\cdot, t)||_{1i}^2 + O(h^{2p+1})
\]

where $h = \max_{i=1,2,...,N} \max_{j=1,2} (h_{i,j})$ and

\[
||E(\cdot, t)||_{1i}^2 = \frac{1}{4(2p+1)} \sum_{i=1}^{N} h_{i,1}h_{i,2} \sum_{j=1}^{4} \sum_{k=1}^{4} \left( \frac{[\partial_{x_j}U(p_i,k,t)]_{j}}{1 + (h_{adj(i,k),j}/h_{i,j})^p} \right)^2.
\]

(3.8b)

**Proof.** Adjerid et al. [2] and Yu [20] established convergence, respectively, for parabolic and elliptic problems on square domains. The extension of their results to rectangular domains and elements is straight forward. \(\square\)
3.2. Error estimates for even-degree approximations. When $p$ is even, gradient jumps of $\psi_{p+1,j}(p)$ vanish and error estimates cannot be obtained using (3.8). Instead, we construct a Galerkin problem for $e$ by replacing $u$ in (2.2a,b) by $U + e$ to obtain

\begin{align}
(3.9a) \quad (v, \partial_t e) + A(v, e) &= g(t, v), \quad t > 0,
\end{align}

\begin{align}
A(v, e) &= A(v, u^0 - U), \quad t = 0, \quad \forall v \in \mathcal{H}_0^1,
\end{align}

with

\begin{align}
(3.9c) \quad g(t, v) &= (v, f) - (v, \partial_t U) - A(v, U).
\end{align}

The trial function (3.3) is again used to approximate $e$ while the approximation of the test function $v$ is selected as

\begin{align}
(3.10a) \quad V_j(x) &= \psi_{p+1,j}(x)\delta(\xi_{j \text{ mod } 2} + 1(x)), \quad j = 1, 2, \quad x \in \Delta_i,
\end{align}

where

\begin{align}
(3.10b) \quad \delta(\xi) &= \frac{\tilde{\psi}_{p+1}(\xi)}{\xi},
\end{align}

\begin{align}
(3.10c) \quad \sigma(\xi) &= \frac{\tilde{\psi}_{p+1}'(\xi)}{\xi}, \quad \xi \in [-1,1],
\end{align}

and $\xi_j(x)$, $j = 1, 2$, is a bilinear mapping of $\Delta_i$ to $[-1,1] \times [-1,1]$.

The functions $\psi_{p+1,j}(x)$ and $V_j(x)$, $j = 1, 2$, vanish on the edges of $\Delta_i$; thus, the computation of $E(x,t)$ is local to $\Delta_i$ and is obtained as the solution of

\begin{align}
(3.11a) \quad (V_j, \partial_t E)_i + A_i(V_j, E) &= g_i(t, V_j), \quad t > 0,
\end{align}

\begin{align}
(3.11b) \quad A_i(V_j, E) &= A_i(V_j, u^0 - U), \quad t = 0, \quad j = 1, 2,
\end{align}

where the subscript $i$ denotes a local inner product whose domain is restricted to $\Delta_i$. This problem may be further simplified by (i) neglecting the off-diagonal diffusion coefficients $a_{j,k}$, $j \neq k$, and the reaction term $b(x)$ as being higher-order and (ii) approximating the diagonal diffusion coefficients $a_{j,j}$, by their values at element centroids $\tilde{a}_{j,j}$, $j = 1, 2$. Thus, $E$ may be obtained by solving (3.11a,b) with $A_i$ on the left of (3.11a) and in (3.11b) replaced by the simpler strain energy

\begin{align}
(3.11c) \quad \tilde{A}_i(v, u) &= \int \int_{\Delta_i} \sum_{j=1}^{2} \tilde{a}_{j,j} \partial_{x_j} v \partial_{x_j} u \, dx_1 \, dx_2.
\end{align}

When the mesh is rectangular, the symmetry of $\psi_{p+1,j}(x)$ and $V_j(x)$, $j = 1, 2$, lead to the uncoupled constant-coefficient initial-value problem on $\Delta_i$

\begin{align}
(3.12a) \quad b_j'(t) + r_j b_j(t) &= G_j(t), \quad t > 0,
\end{align}
\[ b_j(0) = \frac{(h_i,j/2)^{-2p+3}}{\bar{a}_{j,j}(h_i,j \mod 2+1/2)} \int_{-1}^1 \int_{-1}^1 \sigma^2(\xi) \delta(\xi_{j (j \mod 2)+1}) d\xi_1 d\xi_2 \frac{\hat{A}_i(V_j,u^0(\cdot) - U(\cdot,0))}{\int_{-1}^1 \int_{-1}^1 \tilde{\psi}_{p+1}^2(\xi_j) \delta(\xi_{(j \mod 2)+1}) d\xi_1 d\xi_2}. \]

\text{(3.12b)} \quad j = 1, 2,

where

\[ r_j = \frac{\bar{a}_{j,j}}{(h_i,j/2)^2} \int_{-1}^1 \int_{-1}^1 \sigma^2(\xi_j) \delta(\xi_{j (j \mod 2)+1}) d\xi_1 d\xi_2 \int_{-1}^1 \int_{-1}^1 \tilde{\psi}_{p+1}^2(\xi_j) \delta(\xi_{(j \mod 2)+1}) d\xi_1 d\xi_2. \]

\text{(3.12c)}

\[ G_j(t) = \frac{(h_i,j/2)^{-2p+3}}{(h_i,j \mod 2+1/2)} \int_{-1}^1 \int_{-1}^1 \tilde{\psi}_{p+1}^2(\xi_j) \delta(\xi_{(j \mod 2)+1}) d\xi_1 d\xi_2. \]

\text{(3.12d)}

Finally, the time derivative in (3.12a) may be neglected to obtain

\[ b_j(t) = \frac{(h_i,j/2)^{-2p+1}}{\bar{a}_{j,j}(h_i,j \mod 2+1/2)} \int_{-1}^1 \int_{-1}^1 \sigma^2(\xi_j) \delta(\xi_{j (j \mod 2)+1}) d\xi_1 d\xi_2 \frac{g_i(t, V_j)}{\int_{-1}^1 \int_{-1}^1 \tilde{\psi}_{p+1}^2(\xi_j) \delta(\xi_{(j \mod 2)+1}) d\xi_1 d\xi_2}. \]

\text{(3.13)} \quad t > 0, \quad j = 1, 2.

Thus, error estimates are determined as solutions of either local parabolic (3.12) or elliptic (3.13) problems. Both methods produce asymptotically correct results as indicated by the following theorem.

**Theorem 3.2.** Let the mesh and solution structure be as described in Theorem 3.1 and let \( p \geq 2 \) be an even integer. Let \( b_j, j = 1, 2, \) be solutions of either (3.12) or (3.13) that are used to obtain an error estimate according to (3.3). Then

\[ ||e(\cdot,t)||_1^2 = ||E(\cdot,t)||_1^2 + O(h^{2p+1}), \quad t > 0, \]

where

\[ ||E(\cdot,t)||_1^2 = \sum_{i=1}^N \sum_{j=1}^2 \tilde{\psi}_{p+1}^2(t)(h_i,j/2)^{2p+1}(h_i,j \mod 2+1/2) \int_{-1}^1 \int_{-1}^1 \sigma^2(\xi_j) d\xi_1 d\xi_2. \]

\text{(3.14b)}

**Proof.** cf. Adjerid et al. [2, 3] for parabolic problems and Yu [19] for elliptic problems on square domains. Again, the extension to rectangular meshes is straightforward. \( \square \)
4. Error estimation for other finite element bases. With sufficient smoothness, the convergence rate of finite element solutions is determined by the highest degree polynomial that can be interpolated exactly. Thus, piecewise bi-$p$ polynomial approximations contain many higher-order terms that do not increase the convergence rate. Different bases of order $p$, such as serendipity [21] or hierarchical [12, 17] approximations, typically lead to (i) better conditioned stiffness matrices, (ii) reduced computational complexity, and (iii) simpler implementations. Unfortunately, the error estimation procedures (3.6, 8, 11-14) are not asymptotically correct when used with these spaces.

The terms $x_1^{p+1}$ and $x_2^{p+1}$ are the only monomials missing from a bi-$p$ polynomial approximation for it to contain a complete $(p+1)$-degree polynomial (cf. Fig. 1). As indicated by the following theorem, these are the only monomial terms needed to make the error estimates of §3 asymptotically correct when the other monomial terms of degree $p+1$ are present in the solution space $S^{N:p}$.

**Theorem 4.1.** Under the conditions of Theorems 3.1 and 3.2, let $Q_p(\Delta_i)$ be the restriction of $S^{N:p}$ to $\Delta_i$ (cf. (2.4)) and let $M_p(\Delta_i)$ be a space of complete polynomials of degree $p$ on $\Delta_i$. If $Q_p$ satisfies

$$M_p \subset Q_p \subset M_{p+1}, \quad M_{p+1} \subset Q_p \cup \{x_1^{p+1}, x_2^{p+1}\}$$

then the error estimates (3.8) or (3.14) apply when $p$ is odd or even, respectively.

**Proof.** Again, the proof closely parallels those of Adjerid et al. [2, 3] and Yu [19, 20].

Conditions (4.1) may be used to construct many solution spaces where the error estimates of §3 are asymptotically correct; however, let us focus on the hierarchical basis of Szabo and Babuska [17]. Letting

$$\tilde{\phi}_1(\xi) = (1 \pm \xi)/2, \quad \tilde{\phi}_0(\xi) = \sqrt{2k - 1}/2 \int_{-1}^{1} P_{k-1}(\zeta) d\zeta, \quad k = 2, 3, ..., p,$$

be the one-dimensional basis, then the two-dimensional hierarchical basis with respect to the canonical square element contains

i the four vertex shape functions

$$N_1(\xi_1, \xi_2) = \tilde{\phi}_{-1}(\xi_1)\tilde{\phi}_{-1}(\xi_2),$$

$$N_2(\xi_1, \xi_2) = \tilde{\phi}_{1}(\xi_1)\tilde{\phi}_{1}(\xi_2),$$

$$N_3(\xi_1, \xi_2) = \tilde{\phi}_{1}(\xi_1)\tilde{\phi}_{-1}(\xi_2),$$

$$N_4(\xi_1, \xi_2) = \tilde{\phi}_{-1}(\xi_1)\tilde{\phi}_{1}(\xi_2)$$

ii the $4(p-1)$ edge shape functions

$$N_5^k(\xi_1, \xi_2) = \tilde{\phi}_{-1}(\xi_2)\tilde{\phi}_0(\xi_1),$$
(4.4b) \[ N^k_2(\xi_1, \xi_2) = \tilde{\phi}_1^k(\xi_1)\tilde{\phi}_0^k(\xi_2). \]

(4.4c) \[ N^k_0(\xi_1, \xi_3) = \tilde{\phi}_1^k(\xi_1)\tilde{\phi}_0^k(\xi_3). \]

(4.4d) \[ N^k_4(\xi_1, \xi_2) = \tilde{\phi}_{-1}^k(\xi_1)\tilde{\phi}_0^k(\xi_2), \quad k = 2, 3, ..., p; \]

iii \((p - 2)(p - 3)/2\) internal shape functions

(4.5a) \[ N^1_0(\xi_1, \xi_2) = 4\tilde{\phi}_{-1}^1(\xi_1)\tilde{\phi}_1^1(\xi_1)\tilde{\phi}_{-1}^1(\xi_2)\tilde{\phi}_1^1(\xi_2), \]

(4.5b) \[ N^2_0(\xi_1, \xi_2) = N^1_0(\xi_1, \xi_2) P_1(\xi_1), \]

(4.5c) \[ N^3_0(\xi_1, \xi_2) = N^1_0(\xi_1, \xi_2) P_1(\xi_2), \]

(4.5d) \[ N^4_0(\xi_1, \xi_2) = N^1_0(\xi_1, \xi_2) P_2(\xi_1), \]

(4.5e) \[ N^5_0(\xi_1, \xi_2) = N^1_0(\xi_1, \xi_2) P_1(\xi_1) P_1(\xi_2). \]

(4.5f) \[ N^6_0(\xi_1, \xi_2) = N^1_0(\xi_1, \xi_2) P_3(\xi_2), ..., \]

(4.5g) \[ N^{(p-2)(p-3)/2}_0(\xi_1, \xi_2) = N^1_0(\xi_1, \xi_2) P_{p-4}(\xi_2). \]

The four vertex shape functions (4.3) are the usual bilinear shape functions that vanish on the two edges opposite the vertex to which they are associated. Edge functions (4.4) are present in the basis for \(p \geq 2\). They are nonzero only on one element edge and decrease linearly in a direction normal to this edge. The interior “bubble” functions (4.5) are present when \(p \geq 4\) and vanish on all element edges. The monomial terms that are present in this hierarchical approximation when \(p = 4\) are shown in Fig. 1.

The hierarchical basis (4.3-5) does not satisfy conditions (4.1) for \(p \geq 3\); however, the space \(Q_p\) obtained by adding the interior shape functions associated with the hierarchical basis of order \(p + 1\) and the hierarchical basis of order \(p\) does satisfy (4.1). As shown in Fig. 1, the only terms missing from this modified hierarchical space are \(\xi_1^{p+1}\) and \(\xi_2^{p+1}\). As an example, we list the minimal sets \(Q_p, p = 1, 2, 3, 4\), that satisfy (4.1) in terms of the hierarchical basis (4.3-5)

(4.6a) \[ Q_1 = \{ N_1, N_2, N_3, N_4 \}, \]

(4.6b) \[ Q_2 = Q_1 \cup \{ N_1^2, N_2^2, N_3^2, N_4^2 \}, \]

(4.6c) \[ Q_3 = Q_2 \cup \{ N_1^3, N_2^3, N_3^3, N_4^3 \} \cup \{ N_0^1 \}, \]

(4.6d) \[ Q_4 = Q_3 \cup \{ N_1^4, N_2^4, N_3^4, N_4^4 \}. \]
FIG. 1. Pascal triangle showing the monomial terms present in a bi-p polynomial, a p-degree hierarchical polynomial, and a modified p-degree hierarchical polynomial satisfying conditions (4.1) for p = 4.

\[ Q_4 = Q_3 \cup \{N_1^4, N_2^4, N_3^4, N_4^4\} \cup \{N_2^3, N_3^3\} \]

As noted, the sets \( Q_1 \) and \( Q_2 \) are identical to the usual hierarchical basis \((4.3,4)\). The set \( Q_3 \) differs from the usual hierarchical basis by the bubble function \( N_1^3 \), which is normally associated with the hierarchical basis of degree four. Likewise \( Q_4 \) contains the internal modes \( N_2^3 \) and \( N_3^3 \), which are usually associated with a fifth-degree basis. The addition of these internal modes in a finite element software system is simple and results in a minor loss in efficiency relative to the standard hierarchical basis. In return for this extra effort, the solution will be supported by a simple asymptotically correct error estimate.

5. **Examples.** We consider five examples that illustrate the performance of the error estimation procedures for both odd- and even-degree approximations by solving elliptic and parabolic problems having (i) smooth solutions, (ii) solutions with line and point singularities, (iii) nonuniform and highly-graded quadrilateral meshes, and (iv) nonlinearity. The assumptions of Theorems 3.1-4.1 are violated for all examples; thus, indicating that the estimation procedures apply more widely than the theory suggests.

Accuracy of the error estimates is measured by the global and local *effectivity indices*

\[ \theta = \frac{\|E(\cdot,t)\|_1}{\|e(\cdot,t)\|_1}, \quad \theta_i = \frac{\|E(\cdot,t)\|_{1,i}}{\|e(\cdot,t)\|_{1,i}}, \quad i = 1, 2, \ldots, N. \]

If \( E \) is an asymptotically correct estimate of \( e \) then \( \theta \) should converge to unity as the mesh is refined. The estimate is, furthermore, robust if \( \theta \) does not appreciably differ from unity for a wide range of mesh spacings and polynomial degrees.

*Example 1.* Consider Poisson's equation

\[ \Delta u = f(x), \quad x \in \Omega, \]

\[ \Delta u = f(x), \quad x \in \Omega, \]
on the quadrilateral domain $\Omega$ with vertices at $(0.5,0.5)$, $(2.0,0.4)$, $(2.5,1.2)$, and $(0.0,2.0)$. Let $f(x)$ and the Dirichlet boundary conditions be selected such that the exact solution is

$$u(x_1, x_2) = e^{-2x_1x_2}. \quad (5.2b)$$

We solve (5.2) on uniform quadrilateral-element meshes having 100, 225, 400, 625, and 900 elements using bi-$p$, modified hierarchical (cf. (4.6)), and standard hierarchical (cf. (4.5)) piecewise polynomial approximations of orders 1 to 4. Errors and global effectivity indices in the $H^1$ norm are presented in Tables 1-3 for piecewise bi-$p$, modified hierarchical, and standard hierarchical approximations, respectively. Solutions by all three methods are identical with $p = 1$ and results are only shown with the bi-$p$ approximations of Table 1. Likewise, the modified and standard hierarchical bases agree when $p \geq 3$ (cf. §4.) and these results are not duplicated in Table 3.

Although the theory is not developed for non-rectangular meshes, we expect good results with bi-$p$ and modified hierarchical approximations since the solution is smooth and the mesh is not severely distorted. Indeed, effectivity indices for the bi-$p$ and modified hierarchical bases are in excess of 0.85 for virtually all mesh-order combinations. The effectivity indices of both of these approximations appear to converge to unity under mesh refinement. The data of Tables 1 and 2 indicate a better performance for even-degree polynomials than for odd. On the contrary, the results of Table 3 for the standard hierarchical basis do not indicate asymptotic correctness of the error estimates for $p = 3, 4$.

We also solve this problem on a chevron-patterned mesh obtained by mapping the vertices

$$\zeta_{1,j} = \frac{j}{n}, \; \zeta_{2,k} =$$

$$\frac{k}{n} + \frac{(-1)^{j+k}}{3n} \begin{cases} 
1 & \text{if } k \in [1, n - 1] \\
0 & \text{if } k = 0, n 
\end{cases}, \quad j, k = 0, 1, ..., n = \sqrt{N},$$

(5.3)

of $[0,1] \times [0,1]$ (cf. Fig. 3.) onto corresponding vertices in $\Omega$ by a bilinear transformation and forming quadrilateral elements. We solve problems with $p = 1, 2, 3, 4$ and $N = 100, 225, 400, 625$ using bi-$p$ approximations. Errors and global effectivity indices are displayed in Table 4. Results for piecewise modified hierarchical approximations are similar. While effectivity indices are in excess of 0.8, convergence under mesh refinement on these quadrilateral meshes is less clear than on parallel quadrilateral-element meshes.

**Example 2.** Consider Poisson’s equation (5.2a) on a unit square with $f(x)$, the Neumann boundary conditions on $x_2 = 0, 1$, and the Dirichlet boundary conditions on $x_1 = 0, 1$ specified so that the exact solution is

$$u(x_1, x_2) = a_{x_2}^{3/2}. \quad (5.4)$$

This one-dimensional solution has a line singularity at $x_2 = 0$. Numerical results using piecewise bi-$p$ and modified hierarchical approximations are virtually identical, so results are only presented for bi-$p$ spaces (of degrees 1 to 4). Computations are
TABLE 1 
Errors and effectivity indices for Example 1 using piecewise bi-p polynomial approximations.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
<td></td>
<td>|</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.294</td>
<td>0.823</td>
<td>0.283(-1)</td>
<td>0.972</td>
</tr>
<tr>
<td>225</td>
<td>0.199</td>
<td>0.896</td>
<td>0.230(-1)</td>
<td>0.987</td>
</tr>
<tr>
<td>400</td>
<td>0.150</td>
<td>0.926</td>
<td>0.738(-2)</td>
<td>0.992</td>
</tr>
<tr>
<td>625</td>
<td>0.120</td>
<td>0.947</td>
<td>0.475(-2)</td>
<td>0.995</td>
</tr>
<tr>
<td>900</td>
<td>0.100</td>
<td>0.960</td>
<td>0.331(-2)</td>
<td>0.996</td>
</tr>
</tbody>
</table>

TABLE 2 
Errors and effectivity indices for Example 1 using piecewise modified hierarchical polynomial approximations.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
<td></td>
<td>|</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.284(-1)</td>
<td>0.966</td>
<td>0.237(-2)</td>
<td>0.483</td>
</tr>
<tr>
<td>225</td>
<td>0.130(-1)</td>
<td>0.983</td>
<td>0.725(-3)</td>
<td>0.673</td>
</tr>
<tr>
<td>400</td>
<td>0.739(-2)</td>
<td>0.990</td>
<td>0.310(-3)</td>
<td>0.779</td>
</tr>
<tr>
<td>625</td>
<td>0.475(-2)</td>
<td>0.994</td>
<td>0.159(-3)</td>
<td>0.841</td>
</tr>
<tr>
<td>900</td>
<td>0.331(-2)</td>
<td>0.996</td>
<td>0.925(-4)</td>
<td>0.881</td>
</tr>
</tbody>
</table>

TABLE 3 
Errors and effectivity indices for Example 1 using piecewise hierarchical polynomial approximations.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
<td></td>
<td>|</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.586(-2)</td>
<td>0.940</td>
<td>0.665(-3)</td>
<td>0.289</td>
</tr>
<tr>
<td>225</td>
<td>0.140(-2)</td>
<td>1.090</td>
<td>0.143(-3)</td>
<td>0.506</td>
</tr>
<tr>
<td>400</td>
<td>0.580(-3)</td>
<td>1.208</td>
<td>0.452(-4)</td>
<td>0.313</td>
</tr>
<tr>
<td>625</td>
<td>0.291(-3)</td>
<td>1.288</td>
<td>0.183(-4)</td>
<td>0.321</td>
</tr>
<tr>
<td>900</td>
<td>0.165(-3)</td>
<td>1.344</td>
<td>0.879(-5)</td>
<td>0.227</td>
</tr>
</tbody>
</table>

TABLE 4 
Errors and effectivity indices for Example 1 using piecewise bi-p polynomial approximations on chevron-patterned meshes.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
<td></td>
<td>|</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.342</td>
<td>0.888</td>
<td>0.301(-1)</td>
<td>0.907</td>
</tr>
<tr>
<td>225</td>
<td>0.246</td>
<td>0.932</td>
<td>0.156(-1)</td>
<td>0.895</td>
</tr>
<tr>
<td>400</td>
<td>0.187</td>
<td>0.936</td>
<td>0.587(-2)</td>
<td>0.864</td>
</tr>
<tr>
<td>625</td>
<td>0.153</td>
<td>0.950</td>
<td>0.578(-2)</td>
<td>0.861</td>
</tr>
</tbody>
</table>

performed on uniform meshes and meshes graded near the singularity having 100, 225, 400, 625, and 900 elements. The graded meshes are uniform in the \( x_1 \) direction and have vertices in the \( x_2 \) direction at

\[
x_{2,j} = (j/n)^\beta, \quad j = 0, 1, ..., n,
\]

with \( n = \sqrt{N} \). As suggested by Szabo and Babuska \[17\], we select \( \beta = (p + 1/2)/(3/2 - 1/2) \) to match the singularity of the solution and recover the optimal
Errors and effectivity indices on elements adjacent to improve performance. Global effectivity indices appear to converge to unity and local solution and result in inaccurate error estimates throughout the domain. Results is either nonexistent or very slow. As shown in the upper portion of Fig. 2, large errors and local effectivity indices on elements adjacent to \(x_2 = 0\) “pollute” the solution and result in inaccurate error estimates throughout the domain. Results with highly-graded meshes (cf. Table 6 and the lower portion of Fig. 2) substantially improve performance. Global effectivity indices appear to converge to unity and local errors are much closer to an equilibrated state.

**Example 3.** Consider a Dirichlet problem for Laplace’s equation on a unit square with the data selected so that the exact solution in polar coordinates is

\[
E_{2/3} = \frac{2}{3}\sin\left(\frac{2}{3}\phi\right).
\]

This solution behaves as \(O(r^{2/3})\) near the origin and this singularity limits the \(h\)-convergence rate. Unless the singularity is resolved by, e.g., grading the mesh, it will again pollute the solution and error estimate globally. Like others [7], our local error estimates fail to recognize such pollution errors and tend to give poor performance in their presence. Reasonable accuracy returns when the singularity is resolved to the point where the pollution errors are small relative to the local errors.

We initially solve this problem using piecewise bi-p approximations on uniform meshes having 100, 225, 400, and 625 elements. Exact errors and global effectivity indices are presented in Table 7. Local errors and effectivity indices on a 400-element mesh are presented in the upper portion of Fig. 4. As with Example 2, solutions on uniform meshes concentrate errors in the element adjacent to the singularity. The a posteriori error estimates cannot perform well under these conditions and poor global effectivity indices result.

---

**TABLE 5**

<table>
<thead>
<tr>
<th>(p)</th>
<th>(h^p)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>(||H^1||)</td>
<td>(\gamma)</td>
<td>(||H^1||)</td>
<td>(\gamma)</td>
<td>(||H^1||)</td>
</tr>
<tr>
<td>100</td>
<td>0.480(1)</td>
<td>0.885</td>
<td>0.663(1)</td>
<td>0.988</td>
<td>0.238(1)</td>
</tr>
<tr>
<td>225</td>
<td>0.353(1)</td>
<td>0.885</td>
<td>0.446(1)</td>
<td>0.988</td>
<td>0.199(1)</td>
</tr>
<tr>
<td>400</td>
<td>0.256(1)</td>
<td>0.891</td>
<td>0.333(1)</td>
<td>0.988</td>
<td>0.140(1)</td>
</tr>
<tr>
<td>625</td>
<td>0.200(1)</td>
<td>0.896</td>
<td>0.266(1)</td>
<td>0.988</td>
<td>0.112(1)</td>
</tr>
<tr>
<td>900</td>
<td>0.177(1)</td>
<td>0.899</td>
<td>0.222(1)</td>
<td>0.988</td>
<td>0.093(1)</td>
</tr>
</tbody>
</table>

**TABLE 6**

<table>
<thead>
<tr>
<th>(p)</th>
<th>(h^p)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>(||H^1||)</td>
<td>(\gamma)</td>
<td>(||H^1||)</td>
<td>(\gamma)</td>
<td>(||H^1||)</td>
</tr>
<tr>
<td>100</td>
<td>0.359(1)</td>
<td>1.142</td>
<td>0.131(1)</td>
<td>0.963</td>
<td>0.129(1)</td>
</tr>
<tr>
<td>225</td>
<td>0.263(1)</td>
<td>1.001</td>
<td>0.592(1)</td>
<td>0.997</td>
<td>0.396(1)</td>
</tr>
<tr>
<td>400</td>
<td>0.198(1)</td>
<td>1.001</td>
<td>0.336(1)</td>
<td>0.979</td>
<td>0.170(1)</td>
</tr>
<tr>
<td>625</td>
<td>0.158(1)</td>
<td>1.000</td>
<td>0.216(1)</td>
<td>0.985</td>
<td>0.087(1)</td>
</tr>
<tr>
<td>900</td>
<td>0.132(1)</td>
<td>1.000</td>
<td>0.151(1)</td>
<td>0.987</td>
<td>0.056(1)</td>
</tr>
</tbody>
</table>
FIG. 2. Local errors (upper-left) and the difference between the local effectivity indices and unity (upper-right) for Example 2 on a uniform 900-element mesh using piecewise bi-p approximations with $p = 2$. Similar data for computations performed on a highly-graded mesh are shown at the bottom.

TABLE 7

Errors and effectivity indices for Example 3 using piecewise bi-p polynomial approximations on uniform meshes.

<table>
<thead>
<tr>
<th>N</th>
<th>$|\varepsilon|_H$</th>
<th>$|\varepsilon|_{H^{-1}}$</th>
<th>$|\varepsilon|_H$</th>
<th>$|\varepsilon|_{H^{-1}}$</th>
<th>$|\varepsilon|_H$</th>
<th>$|\varepsilon|_{H^{-1}}$</th>
<th>$|\varepsilon|_H$</th>
<th>$|\varepsilon|_{H^{-1}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.166(-1)</td>
<td>0.648(-1)</td>
<td>0.191</td>
<td>0.310</td>
<td>0.166(-1)</td>
<td>0.648(-1)</td>
<td>0.191</td>
<td>0.310</td>
</tr>
<tr>
<td>225</td>
<td>0.357(-1)</td>
<td>0.682(-1)</td>
<td>0.191</td>
<td>0.311</td>
<td>0.124(-1)</td>
<td>0.225(-1)</td>
<td>0.191</td>
<td>0.311</td>
</tr>
<tr>
<td>400</td>
<td>0.296(-1)</td>
<td>0.6512</td>
<td>0.191</td>
<td>0.311</td>
<td>0.124(-1)</td>
<td>0.225(-1)</td>
<td>0.191</td>
<td>0.311</td>
</tr>
<tr>
<td>625</td>
<td>0.256(-1)</td>
<td>0.6531</td>
<td>0.191</td>
<td>0.311</td>
<td>0.092(-2)</td>
<td>0.660(-2)</td>
<td>0.191</td>
<td>0.311</td>
</tr>
</tbody>
</table>
Errors and effectivity indices for Example 3 using piecewise bi-p polynomial approximations on highly-graded meshes.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>418</td>
<td>0.161(-1)</td>
<td>1.062</td>
<td>0.272(-3)</td>
<td>1.986</td>
</tr>
<tr>
<td>653</td>
<td>0.138(-1)</td>
<td>1.053</td>
<td>0.999</td>
<td>1.811</td>
</tr>
<tr>
<td>938</td>
<td>0.123(-1)</td>
<td>1.046</td>
<td>1.022</td>
<td>1.712</td>
</tr>
<tr>
<td>1273</td>
<td>0.111(-1)</td>
<td>1.040</td>
<td>0.864(-3)</td>
<td>1.654</td>
</tr>
</tbody>
</table>

Results are also obtained using piecewise bi-p and modified hierarchical approximations on locally graded meshes that were generated by refining the element closest to the singularity of a uniform $N$-element mesh. Refinement of the element nearest the singularity consists of generating $n$ elements along each coordinate axis according to the distribution (5.5) and generating a diagonal from the upper right vertex of the smallest square element to that of the original square element. As shown in right portion of Fig. 3 for $N = 25$ and $n = 4$, this process creates a mesh with $N$ square and $2(n - 1)$ trapezoidal elements. Results in Tables 8 and 9 use the following combinations of $N$ and $n$: 400, 10; 625, 15; 900, 20; and 1225, 50. Values of $\beta$ are selected as $1/(2/3)$ for $p = 1$ and $(p/2)/(2/3)$ for $p > 1$. Local errors and effectivity indices are presented for piecewise bi-p approximations with $p = 4$ in the lower portion of Fig. 4.

The severe grading has reduced the local error on the element adjacent to the singularity and this has substantially improved the performance of the global and local effectivity indices. As with Examples 1 and 2, results for $p = 3$ are poorer than those for other orders. Effectivity indices are closer to unity with bi-p approximations than with modified hierarchical approximations. Additional equilibration of loading on the edges of odd-order approximations may be necessary to improve the performance of the error estimate [16]. A similar degradation of performance was observed by Babuška and Yu [9] with first- and second-order approximations in the presence of singularities.

**Example 4.** Consider the convection-diffusion problem

\[
(5.7a) \quad u_t - \Delta u + u_{x_1} + u_{x_2} = f(x, t), \quad x \in [0, 1] \times [0, 1], \quad t > 0,
\]

with $f(x, t)$ and the initial and Dirichlet boundary conditions specified so that the exact solution is

\[
(5.7b) \quad u(x_1, x_2, t) = \frac{1}{2}[1 - \tanh(10x_1 + 2x_2 - 10t - 2)].
\]

We solve this problem on $0 < t \leq 0.5$ using uniform meshes having 100, 400, 900, and 1600 square elements with piecewise bi-p and modified hierarchical polynomial
approximations of orders 1 to 4. Temporal integration utilizes the backward-difference software system DASSL [15] with error tolerances of $10^{-6}$ for $p = 1$, $10^{-8}$ for $p = 3$, and $10^{-10}$ for $p = 4$. Such small tolerances minimize temporal discretization errors relative to the spatial errors that we are studying.

Exact errors and effectiveness indices in $H^1$ obtained using (3.8, 14) at $t = 0.5$ are presented in Tables 10 and 11, respectively, for piecewise bi-$p$ and modified hierarchical approximations. The results again indicate convergence of effectiveness indices to unity under mesh refinement. The error estimates have a good range of applicability with effectiveness indices in excess of 0.9 for virtually all computations. Once again, performance is poorer when $p = 3$. Results are slightly better with modified hierarchical than with bi-$p$ approximations.

Example 5. Consider the nonlinear reaction-diffusion equation

\begin{equation}
(5.8a) \quad u_t - \frac{1}{2} \Delta u = q u^2 (1 - u), \quad x \in [0, 1] \times [0, 3/2], \quad t > 0,
\end{equation}

with $q \geq 0$ and the initial and Dirichlet boundary conditions specified so that the exact solution is

\begin{equation}
(5.8b) \quad u(x_1, x_2) = [1 + e^{\sqrt{q/2} (x_1 + x_2 - t \sqrt{q/2})}]^{-1}.
\end{equation}

This solution represents a wave front moving normal to the line $x_1 = -x_2$ with speed $\sqrt{q/2}$.

We solve (5.8) with $q = 20$ on $0 < t \leq 0.5$ using the meshes and piecewise polynomial approximations specified in Example 4. Temporal tolerances are selected as $10^{-6}$ for $p = 1$, $10^{-10}$ for $p = 2$, and $10^{-13}$ for $p = 3, 4$.

Exact errors and effectiveness indices at $t = 0.5$ appear in Tables 12 and 13, respectively for piecewise bi-$p$ and modified hierarchical approximations. Results are comparable to those of Example 4, except that effectiveness indices for $p = 3$ are much closer to unity here than there.

6. Discussion. We show that simple a posteriori estimates of spatial discretization errors of piecewise bi-$p$ polynomial finite element solutions of two-dimensional
elliptic and parabolic problems [2, 3, 9, 19, 20] extend to other spaces of order $p$. The finite element space must contain all monomial terms of degree $p + 1$ except the principal terms $x_1^{p+1}$ and $x_2^{p+1}$. If so, then estimates involving jumps in solution gradients at element vertices when $p$ is odd and from the solution of local elliptic or parabolic problems when $p$ is even are asymptotically correct on rectangular-element grids. The error estimates are stated for arbitrarily structured grids of quadrilateral elements and computational results of §5 show that they may be asymptotically cor-

FIG. 4. Local errors (upper-left) and the difference between the local effectivity indices and unity (upper-right) for Example 3 on a uniform 400-element mesh using piecewise bi-$p$ approximations with $p = 4$. Similar data for computations performed on a highly-graded 418-element mesh are shown at the bottom.
The error estimates are simple to construct and require at most nearest-neighbor information from the finite element solution; hence, they are efficient for both serial and parallel computation. Temporal superconvergence appears to be robust; thus, there is little advantage of using the parabolic error estimation procedure relative to the elliptic procedure for even-order approximations. An exception might occur when using error estimates to control mesh motion (r-refinement).

TABLE 10
Errors and effectiveness indices for Example 4 using piecewise bi-p polynomial approximations.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.543</td>
<td>0.867</td>
<td>0.965</td>
<td>-1</td>
</tr>
<tr>
<td>400</td>
<td>0.233</td>
<td>0.959</td>
<td>0.252</td>
<td>-1</td>
</tr>
<tr>
<td>900</td>
<td>0.157</td>
<td>0.981</td>
<td>0.113</td>
<td>-1</td>
</tr>
<tr>
<td>1600</td>
<td>0.118</td>
<td>0.989</td>
<td>0.640</td>
<td>-2</td>
</tr>
</tbody>
</table>

TABLE 11
Errors and effectiveness indices for Example 4 using piecewise modified hierarchical polynomial approximations.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.957</td>
<td>-1</td>
<td>0.957</td>
<td>-1</td>
</tr>
<tr>
<td>400</td>
<td>0.252</td>
<td>-1</td>
<td>0.990</td>
<td>0.251</td>
</tr>
<tr>
<td>900</td>
<td>0.113</td>
<td>-1</td>
<td>0.995</td>
<td>0.755</td>
</tr>
<tr>
<td>1600</td>
<td>0.640</td>
<td>-2</td>
<td>0.997</td>
<td>0.319</td>
</tr>
</tbody>
</table>

TABLE 12
Errors and effectiveness indices for Example 5 using piecewise bi-p polynomial approximations.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.446</td>
<td>-1</td>
<td>0.989</td>
<td>0.224</td>
</tr>
<tr>
<td>400</td>
<td>0.220</td>
<td>-1</td>
<td>0.977</td>
<td>0.559</td>
</tr>
<tr>
<td>900</td>
<td>0.146</td>
<td>-2</td>
<td>0.998</td>
<td>0.249</td>
</tr>
<tr>
<td>1600</td>
<td>0.110</td>
<td>-2</td>
<td>0.999</td>
<td>0.140</td>
</tr>
</tbody>
</table>

TABLE 13
Errors and effectiveness indices for Example 5 using piecewise modified hierarchical polynomial approximations.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.229</td>
<td>-1</td>
<td>0.967</td>
<td>0.107</td>
</tr>
<tr>
<td>400</td>
<td>0.561</td>
<td>-2</td>
<td>0.990</td>
<td>0.128</td>
</tr>
<tr>
<td>900</td>
<td>0.249</td>
<td>-3</td>
<td>0.966</td>
<td>0.377</td>
</tr>
<tr>
<td>1600</td>
<td>0.148</td>
<td>-3</td>
<td>0.998</td>
<td>0.160</td>
</tr>
</tbody>
</table>
Focusing on spatial error estimation, we ensured that temporal errors were negligible relative to spatial errors. In a practical computational system, however, the temporal and spatial errors must be related. One way of doing this is to maintain the local temporal error per step at a small percentage of the total error [11, 13].

Several theoretical extensions of the error estimation procedures described herein are necessary. For example, the performance of the error estimates in the presence of singularities and singular perturbations needs investigation. The latter situation involves small diffusivities [1] whereas these error estimates rely on diffusion dominance. At the very least, error estimates become less robust near singularities and in the presence of singular perturbations. A theory is needed for nonlinear problems and vector systems. Likewise, convergence analyses are needed on meshes of arbitrary triangular and quadrilateral elements. Although the computations reported here are promising, mesh shape and gradation can greatly affect the accuracy of an error estimate [16].

The present error estimates can be easily extended to three-dimensional cubic elements and the present theory holds in this case. However, analyses are needed for meshes of tetrahedral and hexahedral elements. Additional work will also be necessary to extend the error estimates to spatially varying polynomial degrees (e.g., adaptive p- and hp-refinement). Nothing is known about the convergence of error estimation procedures under such situations and both gradient jumps and internal residuals may be necessary.

REFERENCES

[14] A.V. Illin, B. Bagheri, R.W. Metcalfe, and L.R. Scott, Error control and mesh optimiza-


