Numer. Math. (2011) 117:65–88
DOI 10.1007/s00211-010-0335-2

Numerische
Mathematik

# The effect of numerical integration on the finite element approximation of linear functionals

**Ivo Babuška · Uday Banerjee · Hengguang Li**

Received: 21 September 2009 / Revised: 6 June 2010 / Published online: 30 September 2010


**Abstract** In this paper, we have studied the effect of numerical integration on the finite element method based on piecewise polynomials of degree $k$, in the context of approximating linear functionals, which are also known as "quantities of interest". We have obtained the optimal order of convergence, $\mathcal{O}(h^{2k})$, of the error in the computed functional, when the integrals in the stiffness matrix and the load vector are computed with a quadrature rule of algebraic precision $2k-1$. However, this result was obtained under an increased regularity assumption on the data, which is more than required to obtain the optimal order of convergence of the energy norm of the error in the finite element solution with quadrature. We have obtained a lower bound of the error in the computed functional for a particular problem, which indicates the necessity of the increased regularity requirement of the data. Numerical experiments have been presented indicating that over-integration may be necessary to accurately approximate the functional, when the data lack the increased regularity.

**Mathematics Subject Classification (2000)** 65N15 · 65N30

## 1 Introduction

Determination of various quantities of interest is one of the major goals in scientific computation. For example in elasticity computations, typical quantities of interest are

The research of I. Babuška was partially supported by NSF grant DMS 0611094. The research of U. Banerjee was partially supported by NSF grant DMS 0610778.

I. Babuška
Institute for Computational Engineering and Sciences,
University of Texas at Austin, ACE 6.412, Austin, TX 78712, USA

U. Banerjee (✉) · H. Li
Department of Mathematics, Syracuse University, Syracuse, NY 13244, USA
e-mail: banerjee@syr.edu

various resultants, stress intensity factors, etc. These quantities are values of functionals evaluated at the solution of the underlying problems, e.g., the solution of the system of partial differential equations modeling elasticity (Lamé equations).

Suppose the solution of the problem is characterized by the solution $u \in V \subset H^1(\Omega)$ of the variational problem

$$a(u, v) = f(v), \quad \forall\, v \in V,$$

where $a(\cdot,\cdot)$ is a continuous bilinear form satisfying the inf-sup condition on $V \times V$ and $f(\cdot)$ is a continuous linear form on $V$. Then the quantity of interest is given by $G(u)$, where $G(\cdot)$ is a bounded linear functional on $V$, characterized by $g \in V$ satisfying $G(v) = a(v, g),\ \forall\, v \in V$. $G(u)$ is approximated by the quantity $G(u_h)$, where $u_h \in S_h \subset V$ is the Galerkin approximation of $u$. Using the Galerkin orthogonality $a(u - u_h, v) = 0,\ \forall\, v \in S_h$, the error in the computed quantity of interest $G(u_h)$ is given by

$$G(u) - G(u_h) = a(u - u_h, g - v), \quad \forall\, v \in S_h. \tag{1}$$

In particular, if $u_h$ is the solution of the finite element method (FEM) based on piecewise polynomials of degree $k$, and if $u, g \in H^{k+1}(\Omega)$, then

$$|G(u) - G(u_h)| \le C\|u - u_h\|_{H^1(\Omega)}\|g - g_h\|_{H^1(\Omega)} = \mathcal{O}(h^{2k}).$$

We note that the enhanced accuracy (which is same as $\|u - u_h\|^2_{H^1(\Omega)}$) in the computed quantity of interest is the consequence of the Galerkin orthogonality.

In many applications, the quantity of interest, e.g., stress intensity factors or resultants as mentioned before, is given as $F(u)$, where the functional $F(\cdot)$ is *not* bounded on $V$ but $F(u)$ is finite. In such situations, $F(\cdot)$ could be expressed in terms of a bounded linear functional $G(\cdot)$ on $V$, and $F(u)$ is approximated by a quantity $F_{u_h}$, which is computed using the values of $G(u_h)$, such that

$$F(u) - F_{u_h} \approx G(u) - G(u_h).$$

We will illustrate a procedure to write $F(u)$ in terms of a bounded linear functional $G(u)$ in Sect. 4. These ideas were first systematically studied in [1,2]. A detailed discussion on computing various quantities of interest can also be found in Chapter 11 of [15]. We note however that $F_{u_h}$ will have the enhanced accuracy again as a consequence of the Galerkin orthogonality.

In a FEM, the definite integrals in the elements of the stiffness matrix and the load vector are approximated by numerical integration; as a consequence, the Galerkin orthogonality is violated and (1) does not hold. In this paper, we will address the consequences of the violation of the Galerkin orthogonality and study the effect of numerical integration on the accuracy of the computed quantity of interest. We will show that the loss of accuracy in the computed quantity of interested due to numerical integration could be more than the loss of accuracy in the energy norm of the finite element solution due to numerical integration. We will also give sufficient conditions on the quadrature as well as on the data of the problem that will yield accurate values of the computed quantity of interest.

The effect of numerical integration on the finite element solution $u_h$ have been studied by various authors; we mention [5,7,10–13] ([7] is for the $p$-version). An excellent exposition of this problem is given in [9] and we will use various ideas from [9] in this study. It is well-established (see Theorem 4.1.6 in [9]) that the error in $u_h$ in the energy norm yields optimal order of convergence, $\mathcal{O}(h^k)$, provided

- the numerical integration rule used on each triangle of the finite element triangulation is exact for all polynomials of degree $2k - 2$,
- the exact solution $u \in H^{k+1}(\Omega)$ and the data of the underlying problem, i.e., the variable coefficients and the load function of the PDE, have sufficient regularity depending on $k$,

where the FEM is based on piecewise polynomials of degree $k$.

In this paper, we have considered the FEM based on a quasi-uniform mesh on a polygonal domain and have obtained the following results:

- We have shown that the error in the computed quantity of interest, under numerical integration, is $\mathcal{O}(h^{2k})$. We obtained this result under the assumption that $u, g \in H^{k+1}(\Omega)$, the numerical integration rule on each triangle is exact for all polynomials of degree $2k - 1$, and the data of the underlying problem have more regularity than is required to obtain the optimal order of convergence of the energy norm of the error in $u_h$, as mentioned in the second bullet of the last paragraph.
- We have obtained a lower bound of the error in the computed quantity of interest for a particular problem. This result indicates that increased regularity assumptions on the data of the problem, as mention in the last bullet, is in general necessary.

We have also presented computational results in this paper that show that "over-integration" may be required to obtain accurate value of the quantity of interest, when the data of the problem do not have the required increased regularity.

We mention that the error in the computed functional is similar to the error in approximate eigenvalues [3], obtained from the FEM with exact integration; both are $\mathcal{O}(h^{2k})$, where $k$ is the degree of the polynomials used in the FEM. The same result for the error in approximate eigenvalues, i.e., $\mathcal{O}(h^{2k})$, was obtained in [6] in the presence of numerical integration. However, it required an increased regularity assumption on the data—similar to what we require to obtain the error in the computed functional (i.e., $\mathcal{O}(h^{2k})$) under numerical integration, mentioned before. In contrast to [6], we show that this increased regularity of the data is also in general necessary in the context of approximation of functionals.

We note that the use of a numerical integration rule that is exact for the polynomials of degree $2k - 1$ is not restrictive, as the quadrature rules used in most finite element codes satisfy this condition. But the increased regularity requirement on the data of the problem may have serious consequences. This requirement indicates that the Jacobian of the mapping for a curved element in a finite element triangulation (in the case of a non-polygonal domain) should be sufficiently smooth—a requirement that is restrictive especially in the $h$–$p$ version of the FEM, where a curved element could be large and the Jacobian of the associated mapping may not have the required increased regularity. In such situations, over-integration may become necessary to control the energy norm of the error in the finite element solution; but even more over-integration may be required to control the error in the computed quantity of interest. We have not studied curved elements or $h$–$p$ version of FEM in this paper, but the results presented in this paper could be helpful in the analysis of such problems.

The paper is organized as follows. In Sect. 2, we have given the notation and the model problem, and defined the FEM with numerical integration. We then defined the quantity of interest as a bounded linear functional $G(u)$, and presented the optimal error estimate for the computed value of $G(u)$ in the case of exact integration. In Sect. 3, we presented one of our main results, where we obtained the optimal error estimate for the computed value of $G(u)$ in the presence of numerical integration. We derived a lower bound of the error in the computed value of a linear functional for a particular problem in the presence of numerical integration in Sect. 4, which is another main result of this paper. In Sect. 5, we gave some computational results that illuminate the results obtained in Sects. 3 and 4.

## 2 Preliminaries and notation

### 2.1 The model problem

Let $\Omega \subset \mathbb{R}^n$ be a bounded $n$-dimensional domain with piecewise straight faces. We consider standard Sobolev spaces $H^m(\Omega)$ with norm $\|\cdot\|_{H^m(\Omega)}$ and semi-norm $|\cdot|_{H^m(\Omega)}$. In addition, we denote by $H_0^1(\Omega) \subset H^1(\Omega)$ the subspace consisting of zero-trace functions.

We consider the following elliptic boundary value problem:

$$\begin{cases} -\mathrm{div} A\nabla u = f \quad \text{in} \quad \Omega \\ u = 0 \quad \text{on} \quad \partial\Omega, \end{cases} \tag{2}$$

where $A = (a_{ij}(x))$ is a symmetric matrix, and $a_{ij}(x)$ are smooth functions. We will refer to the functions $a_{ij}(x)$ and $f(x)$ as the *data of the problem*. We assume that $A$ satisfies the uniform ellipticity condition, i.e., there exists a constant $C > 0$, such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq C\sum_{i=1}^{n} \xi_i^2, \quad \forall\, \xi \in \mathbb{R}^n, \quad \forall\, x \in \bar{\Omega}.$$

The variational form of the problem (2) is given as

$$\begin{cases} u \in H_0^1(\Omega), \\ a(u, v) := \int_\Omega \sum_{i,j=1}^n a_{ij} \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} dx = \int_\Omega f v dx, \quad \forall\, v \in H_0^1(\Omega). \end{cases} \tag{3}$$

It is well-known that the bilinear form $a(\cdot, \cdot)$ is coercive and bounded, and $\|v\|_a := \sqrt{a(v, v)}$ defines an equivalent norm on $H_0^1(\Omega)$; the problem (3) has a unique solution.

## 2.2 The finite element method

Let $\mathcal{T}$ be a quasi-uniform triangulation of $\Omega$, and let

$$S_h = \{v \in \mathcal{C}^0(\bar{\Omega}),\ v|_{\partial\Omega} = 0,\ v|_K \in \mathcal{P}_k(K),\ \forall\ K \in \mathcal{T}\} \subset H_0^1(\Omega)$$

be the finite element subspace, where $\mathcal{P}_k$ is the space containing all polynomials of degree $k$. The finite element approximation to the solution $u \in H_0^1(\Omega)$ of (3) is given by the FEM,

$$u_h \in S_h,\ a(u_h, v_h) = f(v_h), \quad \forall\ v_h \in S_h,$$

where $f(v_h) = \int_\Omega f v_h dx$. It is well known that

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^k \|u\|_{H^{k+1}(\Omega)}. \tag{4}$$

Note that $a(u_h, v_h)$ and $f(v_h)$ contain definite integrals that are computed numerically. Consequently, the FEM with numerical integration is given by

$$u_h^* \in S_h,\ a^*(u_h^*, v_h) = f^*(v_h), \quad \forall\ v_h \in S_h, \tag{5}$$

where

$$a^*(u, v) = \sum_{K\in\mathcal{T}} \sum_{l=1}^{L} \omega_{l,K}^{(1)} \sum_{i,j=1}^{n} \left[a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j}\right] (b_{l,K}^{(1)}),$$

$$f^*(v) = \sum_{K\in\mathcal{T}} \sum_{l=1}^{L} \omega_{l,K}^{(0)} [f v] \left(b_{l,K}^{(0)}\right).$$

The two sets $\{\omega_{l,K}^{(1)},\ b_{l,K}^{(1)}\}_{l=1}^L$ and $\{\omega_{l,K}^{(0)},\ b_{l,K}^{(0)}\}_{l=1}^L$ determine two quadrature rules (possibly different), on the triangle $K \in \mathcal{T}$.

We define the error functionals,

$$E_K^{(i)}(\varphi) = \int_K \varphi(x)dx - \sum_{l=1}^{L} \omega_{l,K}^{(i)} \varphi(b_{l,K}^{(i)}), \quad i = 0, 1.$$

Let $F_K : \hat{K} \to K$, $F_K(\hat{x}) := B_K\hat{x} + b_K$, be the affine mapping that maps the reference element $\hat{K}$ onto $K$. Then the error functional on the reference element $\hat{K}$ is,

$$\hat{E}^{(i)}(\hat{\varphi}) := \int_{\hat{K}} \hat{\varphi}(\hat{x})d\hat{x} - \sum_{l=1}^{L} \hat{\omega}_{l,\hat{K}}^{(i)} \hat{\varphi}\left(\hat{b}_{l,\hat{K}}^{(i)}\right),$$

where $b_{l,K}^{(i)} = F_K(\hat{b}_{l,\hat{K}}^{(i)})$, $\omega_{l,K}^{(i)} = \det(B_K)\hat{\omega}_{l,\hat{K}}^{(i)}$, and $\hat{\varphi}(\hat{x}) = \varphi(x)$ for any $x = F_K(\hat{x})$, $\hat{x} \in \hat{K}$. Note that the sets $\{\hat{\omega}_{l,\hat{K}}^{(i)},\ \hat{b}_{l,\hat{K}}^{(i)}\}_{l=1}^{L}$, $i = 0, 1$, define numerical quadrature rules (possibly different) on the reference element $\hat{K}$. Thus, by the standard scaling argument, it is clear that

$$E_K^{(i)}(\varphi) = \det(B_K)\hat{E}^{(i)}(\hat{\varphi}).$$

We define the *algebraic precision* of a quadrature rule to be $m$, if the rule is exact on all polynomials of degree $\leq m$. Throughout this paper we will assume that the algebraic precision of the quadrature rules given by $\{\hat{\omega}_{l,\hat{K}}^{(i)},\ \hat{b}_{l,\hat{K}}^{(i)}\}_{l=1}^{L}$, $i = 0, 1$ is $2k - 1$, i.e., $\hat{E}^{(i)}(\hat{\varphi}) = 0,\ \forall \hat{\varphi} \in \mathcal{P}_{2k-1}$.

We now state a result similar to the Theorem 4.1.6 in [9] that we will use later.

**Theorem 2.1** *Suppose $\hat{E}^{(i)}(\hat{\varphi}) = 0\ \forall\ \hat{\varphi} \in \mathcal{P}_{2k-1}(\hat{K})$ and let $a_{ij} \in W^{k,\infty}(\Omega)$ and $f \in H^{k+1}(\Omega)$. If $u \in H^{k+1}(\Omega)$, then for $n = 1, 2, 3$ we have*

$$\|u - u_h^*\|_{H^1(\Omega)} \leq Ch^k \left( \sum_{i,j=1}^{n} \|a_{ij}\|_{W^{k,\infty}(\Omega)} \|u\|_{H^{k+1}(\Omega)} + \|f\|_{H^{k+1}(\Omega)} \right) \tag{6}$$

*where the constant $C$ is independent of $u$ and $h$, but may depend on $k$.*

The proof of this result can be obtained by following the arguments of the proof of Theorem 4.1.6 in [9] and we do not give the proof in this paper. We note that this result can be extended for $n \geq 4$ provided $f \in W^{k+1,q}(\Omega)$ with a suitable $q \geq 2$ depending on $k$ and $n$.

## 2.3 The linear functionals

Let $G : H_0^1(\Omega) \to \mathbb{R}$ be a bounded linear functional on $H_0^1(\Omega)$. We are interested in approximating the quantity of interest, $G(u)$, where $u$ is the solution of Eq. (3).

Recall that $\|\cdot\|_a$ is equivalent to $\|\cdot\|_{H^1(\Omega)}$ for any function in $H_0^1(\Omega)$. Then, by the Riesz representation theorem, there exists a unique $g \in H_0^1(\Omega)$, such that

$$G(v) = a(v, g), \quad \forall\ v \in H_0^1(\Omega). \tag{7}$$

We approximate $G(u)$ with $G(u_h)$, and the associated error bound is given in the following lemma.

**Lemma 2.2** *Suppose that $g \in H^{s+2}(\Omega)$, $0 \leq s \leq k-1$, and let $u \in H^{k+1}(\Omega)$. Then, there is a constant $C > 0$ independent of $g$, $u$, and the mesh size $h$, such that*

$$|G(u) - G(u_h)| \leq Ch^{k+s+1} \|g\|_{H^{s+2}(\Omega)} \|u\|_{H^{k+1}(\Omega)}. \tag{8}$$

*Proof* Let $g_h \in S_h$ be the projection of $g$ with respect to $a(\cdot, \cdot)$, namely,

$$a(g - g_h, v_h) = 0, \quad \forall\, v_h \in S_h. \tag{9}$$

Now using the Galerkin orthogonality $a(u-u_h, v) = 0$ for all $v \in S_h$, the boundedness of $a(\cdot, \cdot)$, a standard approximation result, and (4), we get

$$\begin{aligned} |G(u) - G(u_h)| &= |a(g, u - u_h)| = |a(g - g_h, u - u_h)| \\ &\le C\|g - g_h\|_{H^1(\Omega)}\|u - u_h\|_{H^1(\Omega)} \\ &\le Ch^{k+s+1}\|g\|_{H^{s+2}(\Omega)}\|u\|_{H^{k+1}(\Omega)}, \end{aligned}$$

which is the desired result. □

*Remark 2.3* It is clear from (8) in Lemma 2.2 that when $s = k - 1$, i.e., when $g \in H^{k+1}(\Omega)$, we have

$$|G(u) - G(u_h)| \le Ch^{2k}\|g\|_{H^{k+1}(\Omega)}\|u\|_{H^{k+1}(\Omega)}.$$

This is the optimal order of convergence that can be obtained for the error in computing $G(u)$ by $G(u_h)$, where $u_h$ is the finite element solution based on piecewise polynomials of degree $k$ and $u \in H^{k+1}(\Omega)$.

*Remark 2.4* The result (8) of Lemma 2.2 also holds for problems with other boundary conditions (e.g., Robin boundary conditions), as long as the associated bilinear form $a(\cdot, \cdot)$ induces an equivalent norm $\|\cdot\|_a$ on a subspace $V \subset H^1(\Omega)$, $G : V \to \mathbb{R}$ is a bounded linear functional, and $S_h \subset V$. For the problem (2), $V = H_0^1(\Omega)$.

*Remark 2.5* As mentioned in the introduction, often a quantity of interest $F(u)$ is finite, but $F : V \to \mathbb{R}$ is not a bounded linear functional on $V$. In such situations, $F(u)$ is written in terms of a bounded linear functional $G(u)$ on $V$, and $F(u)$ is approximated by a quantity $F_{u_h}$ such that $|F(u) - F_{u_h}| \approx |G(u) - G(u_h)|$ (see [4]). Let $g \in V$ be the unique solution of $a(v, g) = G(v),\ \forall\, v \in V$. Then it is clear from (8) of Lemma 2.2 that

$$|F(u) - F_{u_h}| = \mathcal{O}(h^{k+s+1}),$$

provided $u \in H^{k+1}(\Omega)$ and $g \in H^{s+2}(\Omega)$, $0 \le s \le k-1$. In Sect. 4, we will consider a particular quantity of interest $F(u)$, where $F(\cdot)$ is not a bounded functional on $V$.

## 3 The effect of numerical integration

In almost all finite element computations, numerical integration is unavoidable. Consequently, $u_h^*$ instead of $u_h$ is available, and $G(u)$ is approximated by $G(u_h^*)$. We shall concentrate on estimating $|G(u) - G(u_h^*)|$ in this section, and start with the following Strang-type lemma.

**Lemma 3.1** *Let $u_h^* \in S_h$ be the finite element solution with numerical integration, given in (5). Let $G$, $g$, $a^*(\cdot,\cdot)$, and $f^*(\cdot)$ be defined as in the last section. Then,*

$$G(u) - G(u_h^*) = a(g - g_h, u - u_h) + f(g_h) - f^*(g_h) + a^*(u_h^*, g_h) - a(u_h^*, g_h), \tag{10}$$

*where $g_h \in S_h$ is the projection of $g$ onto $S_h$, given by $a(g - g_h, v_h) = 0$, $\forall v \in S_h$.*

*Proof* Recall $u_h \in S_h$ is the finite element solution with exact integration. We first note that

$$\begin{aligned} G(u) - G(u_h^*) = a(g, u - u_h^*) &= a(g, u - u_h) + a(g, u_h - u_h^*) \\ &= a(g - g_h, u - u_h) + a(g, u_h - u_h^*), \end{aligned} \tag{11}$$

where we used the well-known fact that $u_h$ is the projection of $u$ on $S_h$ with respect to $a(\cdot,\cdot)$.

Now, using the definition of $g_h$ and (5), we have

$$\begin{aligned} a(g, u_h - u_h^*) = a(g_h, u_h - u_h^*) &= a(g_h, u_h) - a(g_h, u_h^*) \\ &= f(g_h) - a(g_h, u_h^*) \\ &= f(g_h) - f^*(g_h) + f^*(g_h) - a(g_h, u_h^*) \\ &= f(g_h) - f^*(g_h) + a^*(u_h^*, g_h) - a(u_h^*, g_h). \end{aligned}$$

Combining the above with (11) completes the proof. □

It is clear from Lemma 3.1, that the effect of numerical integration on the approximation of $G(u)$, i.e., on the error $G(u) - G(u_h^*)$, depends on $f(g_h) - f^*(g_h)$ and $a^*(u_h^*, g_h) - a(u_h^*, g_h)$. In fact, if the integration is exact, these differences are zeros. To estimate these terms, we need the following two lemmas.

We now present the first result towards estimating the terms in (10). This result is similar to Lemma 6.2 in [6], where the analysis was given for the one-dimensional case; we prove it here for higher dimensions.

**Lemma 3.2** *Suppose $\hat{E}^{(0)}(\hat{\chi}) = 0$, $\forall \hat{\chi} \in \mathcal{P}_{2k-1}(\hat{K})$ and let $0 \le s \le k-1$. Then, for $1 \le n \le 3$,*

$$|E_K^{(0)}(f\phi)| \le C h^{k+s+1} \|f\|_{H^{k+s+1}(K)} \|\phi\|_{H^{s+2}(K)}, \quad \forall\, \phi \in \mathcal{P}_k(K),$$

*where the constant $C$ is independent of $h$ and $K$, but may depend on $k$.*

*Proof* We first prove it for $k > 1$. Let $\hat{\Pi}$ be the $L^2$-projection onto $\mathcal{P}_{s+1}(\hat{K})$. We write

$$\hat{E}^{(0)}(\hat{f}\hat{\phi}) = \hat{E}^{(0)}(\hat{f}\hat{\Pi}\hat{\phi}) + \hat{E}^{(0)}(\hat{f}\hat{\phi} - \hat{f}\hat{\Pi}\hat{\phi}). \tag{12}$$

Since $k + s + 1 \ge k + 1 > n/2$, we use Sobolev's inequality (see 1.4.6 in [8]) to get

$$|\hat{E}^{(0)}(\hat{f}\hat{\Pi}\hat{\phi})| \le C \|\hat{f}\hat{\Pi}\hat{\phi}\|_{L^\infty(\hat{K})} \le C \|\hat{f}\hat{\Pi}\hat{\phi}\|_{H^{k+s+1}(\hat{K})}.$$

We recall that $\hat{E}^{(0)}(\hat{\chi}) = 0, \forall \hat{\chi} \in \mathcal{P}_{k+s}(\hat{K})$. Therefore,

$$\begin{aligned}|\hat{E}^{(0)}(\hat{f}\hat{\Pi}\hat{\phi})| &\le C|\hat{f}\hat{\Pi}\hat{\phi}|_{H^{k+s+1}(\hat{K})} \le C\sum_{i=0}^{s+1}|\hat{f}|_{H^{k+s+1-i}(\hat{K})}|\hat{\Pi}\hat{\phi}|_{W^{i,\infty}(\hat{K})}\\ &\le C\sum_{i=0}^{s+1}|\hat{f}|_{H^{k+s+1-i}(\hat{K})}|\hat{\Pi}\hat{\phi}|_{H^{i}(\hat{K})},\end{aligned} \tag{13}$$

where we used the Bramble–Hilbert Lemma and the equivalence of norms on finite dimensional spaces. Since $\hat{\Pi}v = v,\ \forall\, v \in \mathcal{P}_{i-1}(\hat{K}),\ 1 \le i \le s+1$, by the inverse inequality and the Bramble–Hilbert Lemma, we get

$$|\hat{\phi} - \hat{\Pi}\hat{\phi}|_{H^i(\hat{K})} \le C\|\hat{\phi} - \hat{\Pi}\hat{\phi}\|_{L^2(\hat{K})} \le C|\hat{\phi}|_{H^i(\hat{K})},$$

where $C$ may depend on $k$. Therefore,

$$|\hat{\Pi}\hat{\phi}|_{H^i(\hat{K})} \le C|\hat{\phi}|_{H^i(\hat{K})} + |\hat{\phi} - \hat{\Pi}\hat{\phi}|_{H^i(\hat{K})} \le C|\hat{\phi}|_{H^i(\hat{K})}.$$

Also since $\hat{\Pi}$ is the $L^2$-projection, the above inequality is true for $i = 0$. Thus from (13) and using a standard scaling argument, we obtain

$$\begin{aligned}|\hat{E}^{(0)}(\hat{f}\hat{\Pi}\hat{\phi})| &\le C\sum_{i=0}^{s+1}|\hat{f}|_{H^{k+s+1-i}(\hat{K})}|\hat{\phi}|_{H^{i}(\hat{K})}\\ &\le C\det(B_K)^{-1}h^{k+s+1}\sum_{i=0}^{s+1}|f|_{H^{k+s+1-i}(K)}|\phi|_{H^{i}(K)}\\ &\le C\det(B_K)^{-1}h^{k+s+1}\|f\|_{H^{k+s+1}(K)}\|\phi\|_{H^{s+1}(K)}.\end{aligned} \tag{14}$$

We now estimate the term $\hat{E}^{(0)}(\hat{f}\hat{\phi} - \hat{f}\hat{\Pi}\hat{\phi})$ in (12). Note that $\hat{f}\hat{\phi} - \hat{f}\hat{\Pi}\hat{\phi} = 0$ if $s = k-1$. So assume $0 \le s < k-1$. Since $k > n/2$, again using the Sobolev inequality, we obtain

$$\begin{aligned}|\hat{E}^{(0)}(\hat{f}\hat{\phi} - \hat{f}\hat{\Pi}\hat{\phi})| &\le C\|\hat{f}\hat{\phi} - \hat{f}\hat{\Pi}\hat{\phi}\|_{L^\infty(\hat{K})} \le C\|\hat{f}\|_{L^\infty(\hat{K})}\|\hat{\phi} - \hat{\Pi}\hat{\phi}\|_{L^\infty(\hat{K})}\\ &\le C\|\hat{f}\|_{H^k(\hat{K})}\|\hat{\phi} - \hat{\Pi}\hat{\phi}\|_{L^\infty(\hat{K})}.\end{aligned}$$

Hence, for a fixed $\hat{\phi} \in \mathcal{P}_k(\hat{K})$, the linear functional $H^k(\hat{K}) \ni \hat{f} \mapsto \hat{E}^{(0)}(\hat{f}\hat{\phi} - \hat{f}\hat{\Pi}\hat{\phi})$ is bounded and vanishes over the space $\mathcal{P}_{k-1}(\hat{K})$. Now using the Bramble–Hilbert Lemma, the equivalence of norms on finite dimensional spaces, the scaling argument, and the fact that $\hat{\Pi}$ leaves $\mathcal{P}_{s+1}(\hat{K})$ invariant, we obtain

$$\begin{aligned}|\hat{E}^{(0)}(\hat{f}\hat{\phi} - \hat{f}\hat{\Pi}\hat{\phi})| &\le C|\hat{f}|_{H^k(\hat{K})}\|\hat{\phi} - \hat{\Pi}\hat{\phi}\|_{L^2(\hat{K})} \le C|\hat{f}|_{H^k(\hat{K})}|\hat{\phi}|_{H^{s+2}(\hat{K})}\\ &\le C\det(B_K)^{-1}h^{k+s+2}|f|_{H^k(K)}|\phi|_{H^{s+2}(K)}\end{aligned} \tag{15}$$

Finally combining (12), (14), and (15), we have

$$|E_K^{(0)}(f\phi)| = |\det(B_K)\hat{E}^{(0)}(\hat{f}\hat{\phi})| \le Ch^{k+s+1}\|f\|_{H^{k+s+1}(K)}\|\phi\|_{H^{s+2}(K)},$$

which completes the proof for $k > 1$.

The proof for $k = 1$ does not use the projection $\hat{\Pi}$ and is simpler. We omit the proof for $k = 1$. □

*Remark 3.3* The result in Lemma 3.2 has been obtained for $1 \le n \le 3$. A similar result could be obtained for $n \ge 4$ provided $f \in W^{k+s+1,q}(K)$, where $q > 2$ is a real number satisfying $k > n/q$. We do not prove the result in this paper.

We now state the next result, which is similar to Lemma 6.1 in [6] (it was also proved for the one-dimensional case). We do not give a proof as it could be obtained by following exactly the arguments in the proof of Lemma 6.1 in [6].

**Lemma 3.4** *Suppose $\hat{E}^{(1)}(\hat{\chi}) = 0$, $\forall\hat{\chi} \in \mathcal{P}_{2k-1}(\hat{K})$ and for $0 \le s \le k-1$, let $a(x) \in W^{k+s+1,\infty}(\Omega)$. Then,*

$$|E_K^{(1)}(a\varphi\phi)| \le Ch^{k+s+1}\|a\|_{W^{k+s+1,\infty}(\Omega)}\|\varphi\|_{H^k(K)}\|\phi\|_{H^{s+1}(K)}, \quad \forall\, \varphi, \phi \in \mathcal{P}_k(K),$$

*where the constant $C$ is independent of $K$.*

*Remark 3.5* We note that in Lemma 6.1 of [6], it was assumed that $a(x) \in W^{2k,\infty}(\Omega)$. A careful reading of the proof of Lemma 6.1 of [6] reveals that $a(x) \in W^{k+s+1,\infty}(\Omega)$, $0 \le s \le k-1$, is only needed; we have used this assumption on $a(x)$ in the above Lemma 3.4.

With Lemmas 3.1, 3.2 and 3.4, we are now ready to give the upper bound of $|G(u) - G(u_h^*)|$ in the following theorem.

**Theorem 3.6** *Suppose $\hat{E}^{(i)}(\hat{\chi}) = 0$, $\forall\chi \in \mathcal{P}_{2k-1}(\hat{K})$, $i = 0, 1$. For $0 \le s \le k-1$, let $f \in H^{k+s+1}(\Omega)$, $g \in H^{s+2}(\Omega)$, and $a_{ij} \in W^{k+s+1,\infty}(\Omega)$. Then for $1 \le n \le 3$, we have*

$$|G(u) - G(u_h^*)| \le Ch^{k+s+1}(\|u\|_{H^{k+1}(\Omega)} + \|f\|_{H^{k+s+1}(\Omega)})\|g\|_{H^{s+2}(\Omega)}, \tag{16}$$

*where the constant $C$ does not depend on the mesh size $h$. In particular, for $s = k-1$, we have*

$$|G(u) - G(u_h^*)| \le Ch^{2k}(\|u\|_{H^{k+1}(\Omega)} + \|f\|_{H^{2k}(\Omega)})\|g\|_{H^{k+1}(\Omega)}. \tag{17}$$

*Proof* We first note from Lemma 3.1 that

$$G(u) - G(u_h^*) = a(g - g_h, u - u_h) + f(g_h) - f^*(g_h) + a^*(u_h^*, g_h) - a(u_h^*, g_h), \tag{18}$$

where $g_h \in S_h$ satisfies $a(g - g_h, v_h) = 0, \ \forall \ v_h \in S_h$. It is clear from (4) and a approximation result that

$$|a(g - g_h, u - u_h)| \leq C\|g - g_h\|_{H^1(\Omega)}\|u - u_h\|_{H^1(\Omega)}$$
$$\leq Ch^{k+s+1}\|g\|_{H^{s+2}(\Omega)}\|u\|_{H^{k+1}(\Omega)}. \tag{19}$$

We will now obtain upper bounds for $|f(g_h) - f^*(g_h)|$ and $|a^*(u_h^*, g_h) - a(u_h^*, g_h)|$. For $v \in L^2(\Omega)$ such that $v \in H^i(K)$, for all $K \in \mathcal{T}$, $0 \leq i \leq k$, we define

$$\|v\|_{i,\mathcal{T}} := \left(\sum_{K \in \mathcal{T}} \|v\|^2_{H^i(K)}\right)^{1/2}, \quad i = 0, 1, \ldots, k.$$

Thus from Lemma 3.2, we get

$$|f(g_h) - f^*(g_h)| \leq \sum_{K \in \mathcal{T}} |E_K^{(0)}(f\, g_h)|$$
$$\leq Ch^{k+s+1} \sum_{K \in \mathcal{T}} \|f\|_{H^{k+s+1}(K)}\|g_h\|_{H^{s+2}(K)}$$
$$\leq Ch^{k+s+1}\|f\|_{H^{k+s+1}(\Omega)}\|g_h\|_{s+2,\mathcal{T}}. \tag{20}$$

Let $\mathcal{I}_h v \in S_h$ be the $S_h$-interpolant of $v \in H^{s+2}(\Omega)$. Then it is well known (see [9]) that

$$\|v - \mathcal{I}_h v\|_{i,\mathcal{T}} \leq Ch^{s+2-i}\|v\|_{H^{s+2}(\Omega)}, \quad i = 0, 1, \ldots, s+2. \tag{21}$$

Therefore from the triangle inequality, the inverse inequality, and a standard approximation result, we have

$$\|g_h\|_{s+2,\mathcal{T}} \leq \|g\|_{H^{s+2}(\Omega)} + \|g - \mathcal{I}_h g\|_{s+2,\mathcal{T}} + \|\mathcal{I}_h g - g_h\|_{s+2,\mathcal{T}}$$
$$\leq C\|g\|_{H^{s+2}(\Omega)} + Ch^{-(s+1)}\|\mathcal{I}_h g - g_h\|_{1,\mathcal{T}}$$
$$\leq C\|g\|_{H^{s+2}(\Omega)} + Ch^{-(s+1)}(\|\mathcal{I}_h g - g\|_{H^1(\Omega)} + \|g - g_h\|_{H^1(\Omega)})$$
$$\leq C\|g\|_{H^{s+2}(\Omega)}.$$

Hence from (20), we get

$$|f(g_h) - f^*(g_h)| \leq Ch^{k+s+1}\|f\|_{H^{k+s+1}(\Omega)}\|g\|_{H^{s+2}(\Omega)}. \tag{22}$$

For the last term $|a^*(u_h^*, g_h) - a(u_h^*, g_h)|$, we use Lemma 3.4 and follow the argument leading to (20) to get

$$|a^*(u_h^*, g_h) - a(u_h^*, g_h)|$$
$$\leq Ch^{k+s+1} \sum_{i,j=1}^{n} \|a_{i,j}\|_{W^{k+s+1,\infty}(\Omega)}\|u_h^*\|_{k+1,\mathcal{T}}\|g_h\|_{s+2,\mathcal{T}}. \tag{23}$$

Again by the triangle inequality, the inverse inequality, (21) with $v = u$, $s = k - 1$, and (6), we obtain

$$\begin{aligned}\|u_h^*\|_{k+1,\mathcal{T}} &\le \|u\|_{H^{k+1}(\Omega)} + \|u - \mathcal{I}_h u\|_{k+1,\mathcal{T}} + \|\mathcal{I}_h u - u_h^*\|_{k+1,\mathcal{T}} \\ &\le C\|u\|_{H^{k+1}(\Omega)} + Ch^{-k}\|\mathcal{I}_h u - u_h^*\|_{1,\mathcal{T}} \\ &\le C\|u\|_{H^{k+1}(\Omega)} + Ch^{-k}(\|\mathcal{I}_h u - u\|_{H^1(\Omega)} + \|u - u_h^*\|_{H^1(\Omega)}) \\ &\le C\left(\sum_{i,j=1}^{n} \|a_{ij}\|_{W^{k,\infty}(\Omega)}\|u\|_{H^{k+1}(\Omega)} + \|f\|_{H^{k+1}(\Omega)}\right).\end{aligned}$$

Hence from (23) we have,

$$\begin{aligned}&|a^*(u_h^*, g_h) - a(u_h^*, g_h)| \\ &\quad\le Ch^{k+s+1}\left[\sum_{i,j=1}^{n} \|a_{ij}\|_{W^{k+s+1,\infty}(\Omega)}\right]\left(\sum_{i,j=1}^{n} \|a_{ij}\|_{W^{k,\infty}(\Omega)}\|u\|_{H^{k+1}(\Omega)}\right. \\ &\qquad\left.+\|f\|_{H^{k+1}(\Omega)}\right)\|g\|_{H^{s+2}(\Omega)}.\end{aligned} \tag{24}$$

Finally, combining (18), (19), (22), and (24), we obtain

$$\begin{aligned}&|G(u) - G(u_h^*)| \\ &\quad\le Ch^{k+s+1}\Bigg[\|g\|_{H^{s+2}(\Omega)}\|u\|_{H^{k+1}(\Omega)} + \|f\|_{H^{k+s+1}(\Omega)}\|g\|_{H^{s+2}(\Omega)} \\ &\qquad+\left(\sum_{i,j=1}^{n} \|a_{ij}\|_{W^{k+s+1,\infty}(\Omega)}\right)\left(\sum_{i,j=1}^{n} \|a_{ij}\|_{W^{k,\infty}(\Omega)}\|u\|_{H^{k+1}(\Omega)}\right. \\ &\qquad\left.+\|f\|_{H^{k+1}(\Omega)}\right)\|g\|_{H^{s+2}(\Omega)}\Bigg] \\ &\quad\le Ch^{k+s+1}(\|u\|_{H^{k+1}(\Omega)} + \|f\|_{H^{k+s+1}(\Omega)})\|g\|_{H^{s+2}(\Omega)},\end{aligned}$$

which is the desired result. □

*Remark 3.7* We note that we get the optimal order of convergence (17) under an increased regularity requirement on the data, i.e., $f \in H^{2k}(\Omega) \cap \mathcal{C}^0(\Omega)$ and $a_{ij} \in W^{2k,\infty}(\Omega)$; this increased regularity is not necessary to obtain (8)—an optimal-order error estimate with exact integration.

*Remark 3.8* It is well known that to obtain the optimal order of convergence of the $H^1$-norm of the error in the finite element solution under numerical integration, one needs an increased regularity of the data, namely, $a_{ij} \in W^{k,\infty}(\Omega)$ and $f \in H^k(\Omega) \cap \mathcal{C}^0(\Omega)$ for $1 \le n \le 3$ (see [9]). Moreover, the algebraic precision of the numerical integration rule is required to be $2k-2$. In a standard finite element software, often the algebraic precision of the numerical quadrature is $2k-1$, which is the same as the assumption used in Theorem 3.6. In Theorem 3.6, however, we required a higher regularity of the data, i.e., $a_{ij} \in W^{2k,\infty}(\Omega)$ and $f \in H^{2k}(\Omega) \cap \mathcal{C}^0(\Omega)$ for the optimal order of convergence of $G(u_h^*)$. Thus, the error in approximating linear functionals may be more sensitive to numerical integration than the error (in $H^1$-norm) in the finite element solution, when the data is not sufficiently smooth.

*Remark 3.9* The result of Theorem 3.6 also holds when the exact solution $u$ satisfies other boundary conditions, provided the functions in $S_h$ satisfy the essential boundary conditions (see Remark 2.4).

## 4 A lower bound on the error in approximating functionals

In this section, our main goal is to show that the smoothness assumption on the data of the problem (see Theorem 3.6) is necessary, in general, to obtain the optimal order of convergence of the computed quantity of interest. Also in this section, we will consider a quantity of interest $F(u)$, which is finite but $F(\cdot)$ is not a bounded linear functional on the energy space. We will first write $F(u)$ in terms of a bounded linear functional $G(u)$ and use the results in Sect. 3 to obtain an upper bound of the error in the computed quantity of interest, when the data has enough smoothness. We will then obtain a lower bound of the error in computed quantity of interest for a particular problem. This result will indicate that the increased regularity of the data is, in general, necessary to obtain the optimal order of convergence of the computed quantity of interest.

Consider the one-dimensional problem on $\Omega = (0, 1)$ with the Robin boundary condition

$$\begin{cases} -u''(x) = f(x) & x \in \Omega, \\ u(0) = 0, & u(1) + u'(1) = 0, \end{cases} \tag{25}$$

Suppose we want to approximate the quantity of interest $F(u) = u'(0)$.

The variational formulation of (25) is given by

$$\begin{cases} u \in H_D^1 := \{v \in H^1(\Omega), \quad v(0) = 0\}, \\ a(u, v) := \int_0^1 u'v'dx + u(1)v(1) = \int_0^1 fv dx, \quad \forall\, v \in H_D^1, \end{cases} \tag{26}$$

where we assume that $f \in L^2(\Omega)$. Since $a(\cdot,\cdot)$ is coercive and bounded, it is immediate that $\|v\|_a := \sqrt{a(v, v)}$ is an equivalent norm in $H_D^1$ with the inner product $a(\cdot,\cdot)$.

We first note that $F(u) = u'(0)$ is finite, but $F(\cdot)$ is not a bounded linear functional on $H_D^1$, and therefore we cannot use the framework developed in Sect. 3. Moreover, $F(u_h)$ may give less accurate approximation of $F(u)$, where $u_h$ is the finite element solution (see Remark 4.1). To address this issue (see Section 6.2 in [4]), we choose a function $\psi \in H^1(\Omega)$, such that $\psi(0) = 1$ and $\psi(1) = 0$. Then using integration by parts, we have $F(u) = \int_0^1 f\psi\, dx - \int_0^1 \psi' u' dx$. Let

$$G(u) := \int_0^1 \psi' u' dx.$$

The functional $G(\cdot)$, which depends on $\psi$, is bounded on $H^1_D$. We write

$$F(u) = \int_0^1 f\psi\, dx - G(u). \tag{27}$$

Note that this expression of $F(u)$ is meaningful, since $f \in L^2(\Omega)$. Furthermore, $F(u)$ can be written in this form for any $\psi \in H^1(\Omega)$ with $\psi(0) = 1$ and $\psi(1) = 0$. We consider a particular $\psi$ given by

$$\psi(x) := 1 - x + \epsilon \sin(\pi x), \tag{28}$$

for a fixed $0 < \epsilon \leq 1$. In fact, we could have considered $\epsilon = 0$ in the definition of $\psi(x)$, but we chose $\epsilon > 0$ to show a particular feature at the end of this section.

We further note that by the Riesz representation theorem, there exists a unique $g \in H^1_D$ such that

$$a(v, g) = G(v), \quad \forall\, v \in H^1_D. \tag{29}$$

An easy calculation using integration by parts yields

$$g(x) = \epsilon \sin(\pi x) - 0.5x \in H^3(\Omega).$$

To approximate the solution of (26), we consider the points $x_j = jh$, $j = 0, 1, \ldots, n$, where $n$ is a natural number and $h = 1/n$. Let $S_h \subset H^1_D$ be the piecewise quadratic ($k = 2$) finite element space with uniform mesh of size $h$. The FEM to approximate the solution $u \in H^1_D$ of (26) is given by

$$u_h \in S_h,\ a(u_h, v_h) = \int_0^1 f v_h dx, \quad \forall\, v_h \in S_h. \tag{30}$$

We define the quantity

$$F_{u_h} := \int_0^1 f\psi\, dx - G(u_h). \tag{31}$$

If $\int_0^1 f\psi\, dx$ is computed exactly, then the quantity of interest $F(u)$ can be approximated by $F_{u_h}$ and from (27), we get $F(u) - F_{u_h} = G(u) - G(u_h)$. If $u \in H^3(\Omega)$, then from (8) of Lemma 2.2 with $k = 2$ and $s = k - 1 = 1$, we get

$$|F(u) - F_{u_h}| = |G(u) - G(u_h)| \le Ch^4 \|g\|_{H^3(\Omega)} \|u\|_{H^3(\Omega)}. \tag{32}$$

But $u_h$ is often not available, as the integrals in (30) are computed by numerical integration, and therefore we cannot compute $F_{u_h}$ to approximate $F(u)$.

We use the 2-point Gaussian quadrature, determined by the set $\{\omega_{l,i}, b_{l,i}\}_{l=1}^2$, on the interval $(x_{i-1}, x_i)$, $i = 1, 2, \ldots, n$, to approximate the integrals in (30). This quadrature rule is exact for polynomials of degree $2k - 1 = 3$, as required in Theorem 3.6. The FEM under numerical integration is given by

$$u_h^* \in S_h,\ a(u_h^*, v_h) = f^*(v_h), \quad \forall\, v_h \in S_h, \tag{33}$$

where

$$f^*(v) = \sum_{i=1}^{n} \sum_{l=1}^{2} \omega_{l,i} f(b_{l,i}) v(b_{l,i}).$$

We note that with the 2-point Gaussian quadrature, we have $a^*(u, v) = a(u, v)$ for all $u, v \in S_h$.

We approximate $F(u)$ by the quantity

$$F_{u_h^*} := \int_0^1 f\psi\, dx - G(u_h^*). \tag{34}$$

If $\int_0^1 f\psi\, dx$ is computed exactly, then from (27) we get

$$F(u) - F_{u_h^*} = G(u) - G(u_h^*). \tag{35}$$

Suppose $f \in H^4(\Omega)$. Then from (17) of Theorem 3.6 (considering $k = 2$ and $s = k - 1 = 1$), we have

$$|F(u) - F_{u_h^*}| = |G(u) - G(u_h^*)| \le Ch^4 (\|u\|_{H^3(\Omega)} + \|f\|_{H^4(\Omega)}) \|g\|_{H^3(\Omega)}. \tag{36}$$

We note however that in practice, the definite integral $\int_0^1 f\psi\, dx$ in (34) is also computed by numerical integration. Therefore this definite integral has to be computed accurately to obtain $|F(u) - F_{u_h^*}| = \mathcal{O}(h^4)$.

*Remark 4.1* It is possible to approximate $F(u)$ directly without using the function $\psi(x)$ as in (31). In particular, $F(u) = u'(0)$ in our example could be well approximated by directly evaluating $F(u_h) = u_h'(0)$ (assuming exact integration), as it is well known (see [14,16]) that $|F(u) - F(u_h)| \le \|u - u_h\|_{W^{1,\infty}(\Omega)} = \mathcal{O}(h^k)$. But computing

$F(u_h)$ in terms of the function $\psi(x)$ as in (31) enables us to to get $|F(u) - F(u_h)| = \mathcal{O}(h^{2k}) = \|u - u_h\|^2_{H^1(\Omega)}$. We mention however that computing $F(u_h)$ using (31) (or $F(u_h^*)$ using (34)) is costlier than a direct computation, since an appropriate $\psi(x)$ has to chosen and the integrals $\int_0^1 f\psi\, dx$ and $G(u_h)$ (or $G(u_h^*)$) have to be computed accurately. On the other hand, computing $F(u_h)$ using (31) (or $F(u_h^*)$ using (34)) will allow the use of a coarser mesh to approximate $F(u)$ within a given tolerance.

*Remark 4.2* In many problems $F(u)$ could be infinite and thus meaningless, e.g., say when $F(u)$ is the value of the stresses at a re-entrant corner in the elasticity problem. Thus $|F(u) - F(u_h^*)|$ is meaningless and any conclusion based on the value of $F_{u_h^*}$ (which could be finite) would be misleading.

We will now obtain a lower bound of the error $|F(u) - F_{u_h^*}|$ for a particular problem, where the data does not have the increased regularity assumed in Theorem 3.6. This result will indicate that the increased regularity assumption on the data in Theorem 3.6 may be necessary to obtain (17). We consider a particular problem (25) with

$$f(x) = (1 - x)^{5/3}. \tag{37}$$

Clearly, $f \in H^2(\Omega)$, but $f \notin H^3(\Omega)$. Thus $f$ has less regularity than required in Theorem 3.6 with $k = 2$ and $s = k - 1 = 1$, and $|F(u) - F_{u_h^*}|$ may not have the optimal order of convergence $\mathcal{O}(h^4)$ in contrast to (36) (see also Remarks 3.7 and 3.8). We note however that the regularity $f \in H^2(\Omega)$ is sufficient to obtain the optimal order of convergence, $\mathcal{O}(h^2)$, for $\|u - u_h^*\|_{H^1(\Omega)}$ (see [9]).

To analyze the error $G(u) - G(u_h^*)$, and consequently the error $F(u) - F_{u_h^*}$ (see (35)), we first consider the projection $g_h \in S_h$ of $g$ given by

$$a(g - g_h, v_h) = 0, \quad \forall\, v_h \in S_h. \tag{38}$$

Since $g_h \in S_h$, it is piecewise quadratic and can be written as

$$g_h(x) = \sum_{i=1}^{n} [c_i\phi_i(x) + b_i B_i(x)], \tag{39}$$

where $\phi_i(x)$ is the usual “hat function” centered at $x_i$, and $B_i(x) = \sqrt{6}h^{-2}(x - x_{i-1})(x - x_i)$ is the quadratic “bubble function” on the interval $(x_{i-1}, x_i)$. It is well known (in one-dimension) that $c_i = g_h(x_i) = g(x_i)$, and therefore

$$g_h(x) = \frac{g(x_i) - g(x_{i-1})}{h}(x - x_{i-1}) + g(x_{i-1}) + b_i B_i(x), \quad x \in (x_{i-1}, x_i). \tag{40}$$

*Remark 4.3* We note that if we define $\psi$ in (28) with $\epsilon = 0$, then $g(x) = -0.5x$ and consequently, $g_h = g$.

In the following lemma, we estimate the constant $b_i$ in (40). The proof is simple and we give a brief sketch of the proof.

**Lemma 4.4** *Let $g_h(x) = \sum_{i=1}^{n} [c_i\phi_i(x) + b_i B_i(x)]$ be the projection of $g$ onto $S_h$ given by (38). Then,*

$$b_i = \frac{\int_{x_{i-1}}^{x_i} \psi' B_i' dx}{\int_{x_{i-1}}^{x_i} (B_i')^2 dx} = -\frac{\pi^2\sqrt{6}}{12}\epsilon h^2 \sin\left(\frac{\pi}{2}(x_{i-1}+x_i)\right) + \mathcal{O}(\epsilon h^4).$$

*Proof* We first note that with $v = B_i$ in (29) and using (40), we get

$$\int_0^1 \psi' B_i' dx = a(g, B_i) = a(g_h, B_i) = b_i \int_{x_{i-1}}^{x_i} [B_i']^2 dx$$

Now directly computing $\int_0^1 \psi' B_i' dx$ and $\int_{x_{i-1}}^{x_i} [B_i']^2 dx$, and using the Taylor's theorem, we get the desired result. □

We are now ready to provide a lower bound of $|G(u) - G(u_h^*)|$.

**Theorem 4.5** *Let $u$ be the solution of (26) with $f$ as in (37), and let $u_h^*$ be the solution of (33) with $k = 2$. Then there exists a constant $C > 0$ such that*

$$|G(u) - G(u_h^*)| \geq \frac{0.25}{27} C_g h^{8/3} + Ch^{11/3}.$$

*where $C_g > 0$ is a constant depending on the numerical integration rule.*

*Proof* Since $a^*(u, v) = a(u, v), \forall u, v \in S_h$, using Lemma 3.1, we have

$$|G(u) - G(u_h^*)| \geq \left| |f(g_h) - f^*(g_h)| - |a(g - g_h, u - u_h)| \right|, \tag{41}$$

where $a(\cdot,\cdot)$ is as defined in (26). We first address the term $|f(g_h) - f^*(g_h)|$.

We recall that $g_h$ is piecewise quadratic. Therefore, by the standard error formula for the 2-point Gaussian quadrature, there is $\xi_i \in (x_{i-1}, x_i)$, such that

$$\begin{aligned} f(g_h) - f^*(g_h) &= \sum_{i=1}^{n} \left( \int_{x_{i-1}}^{x_i} [fg_h] dx - \sum_{l=1}^{2} \omega_{i,l} [fg_h](b_{i,l}) \right) = C_g h^5 \sum_{i=1}^{n} [fg_h]^{(4)}(\xi_i) \\ &= C_g h^5 \sum_{i=1}^{n} \left\{ [f^{(4)} g_h](\xi_i) + 4[f^{(3)} g_h'](\xi_i) + 6[f^{(2)} g_h^{(2)}](\xi_i) \right\} \\ &= C_g h^5 \sum_{i=1}^{n} \left[ \frac{40}{81}(1-\xi_i)^{-7/3} g_h(\xi_i) + \frac{40}{27}(1-\xi_i)^{-4/3} g_h'(\xi_i) \right. \\ &\quad \left. + \frac{20}{3}(1-\xi_i)^{-1/3} g_h''(\xi_i) \right] := \mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3, \end{aligned} \tag{42}$$

where $C_g > 0$ is a fixed constant depending on the numerical integration. Also, $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$ are defined as the first, second, and the third term in (42), respectively.

We first estimate the term $\mathcal{S}_1$ in (42). Recall that $g_h(x_i) = g(x_i)$, and note that for the bubble function $B_i(x_{i-1}) = B_i(x_i) = 0$. Thus, from the definition of $g_h$ in (40), we have

$$g_h(\xi_i) = \left[\frac{g(x_i) - g(x_{i-1})}{h}(\xi_i - x_{i-1}) + g(x_{i-1}) + b_i B_i(\xi_i)\right]. \tag{43}$$

Recall $g(x) = \epsilon \sin(\pi x) - 0.5x$. It is easy to show using Taylor's theorem that

$$\left|\frac{g(x_i) - g(x_{i-1})}{h}(\xi_i - x_{i-1})\right| \le h \left|\epsilon\pi \cos\left(\frac{\pi}{2}(x_{i-1} + x_i)\right) - 0.5 + \mathcal{O}(\epsilon h^2)\right| \le C_1 h. \tag{44}$$

Moreover, using Lemma 4.4 and noting that $|B_i(\xi_i)| < \sqrt{6}$, we get

$$|b_i B_i(\xi_i)| \le \left|\frac{\pi^2}{2}\epsilon \sin\left(\frac{\pi}{2}(x_{i-1} + x_i)\right) + \mathcal{O}(\epsilon h^2)\right| h^2 \le C_2 h^2. \tag{45}$$

We will now find upper bounds for the term $g(x_{i-1})$ in (43) for $1 \le i \le [\frac{7n}{8}] - 1$ and $[\frac{7n}{8}] \le i \le n$, where $[x]$ denotes the largest integer that is less than or equal to $x$. We first note that $g(x) < 1$ for $0 \le x \le 1$ and for all $0 < \epsilon \le 1$; consequently

$$g(x_{i-1}) < 1, \quad \text{for } 1 \le i \le \left[\frac{7n}{8}\right] - 1. \tag{46}$$

Furthermore, an easy calculation shows that for all $0 < \epsilon \le 1$, $g(x) < -0.035$ for $0.87 \le x \le 1$. Since, for $h$ small (in fact $h < 0.0025$), $x_{i-1} > 0.87$ for $i \ge [\frac{7n}{8}]$, it is clear that

$$g(x_{i-1}) < -0.035, \quad \text{for } \left[\frac{7n}{8}\right] \le i \le n. \tag{47}$$

Therefore, for $1 \le i \le [7n/8] - 1$ and $h$ small, using (44), (45) and (46) in (43), we have

$$g_h(\xi_i) \le 1 + C_1 h + C_2 h^2 \le 2. \tag{48}$$

Similarly for $i \ge [7n/8]$ and $h$ small, using (44), (45) and (47) in (43), we get

$$g_h(\xi_i) \le -0.035 + C_1 h + C_2 h^2 \le -0.025. \tag{49}$$

We next note that the function $(1 - x)^{-7/3}$ is positive and increasing on $(0, 1)$. Therefore, recalling that $x_{i-1} \le \xi_i \le x_i$, we easily see that for $h$ small,

$$h^5 \sum_{i=1}^{[7n/8]-1} (1 - \xi_i)^{-7/3} \le h^4 \int_0^{7/8} (1 - x)^{-7/3} dx = \frac{3}{4} h^4 (8^{4/3} - 1), \tag{50}$$

and

$$h^5 \sum_{i=[7n/8]}^{n} (1-\xi_i)^{-7/3} \geq h^4 \int_{7/8}^{1-h} (1-x)^{-7/3} dx = \frac{3}{4} h^4 (h^{-4/3} - 8^{4/3}). \tag{51}$$

We now split the summation in $\mathcal{S}_1$ (see (42)) and use (48)–(51), to get

$$\begin{aligned} \mathcal{S}_1 &= \frac{40}{81} C_g h^5 \left[ \sum_{i=1}^{[7n/8]-1} (1-\xi_i)^{-7/3} g_h(\xi_i) + \sum_{i=[7n/8]}^{n} (1-\xi_i)^{-7/3} g_h(\xi_i) \right] \\ &\leq \frac{40}{81} C_g h^5 \left[ 2 \sum_{i=1}^{[7n/8]-1} (1-\xi_i)^{-7/3} - 0.025 \sum_{i=[7n/8]}^{n} (1-\xi_i)^{-7/3} \right] \\ &\leq -\frac{0.25}{27} C_g h^4 (h^{-4/3} - 8^{4/3}) + \frac{20}{27} C_g h^4 (8^{4/3} - 1). \end{aligned} \tag{52}$$

We use similar arguments as above to to estimate the terms $\mathcal{S}_2$, $\mathcal{S}_3$ in (42). For all $0 < \epsilon \leq 1$ and $h$ small, we get

$$\mathcal{S}_2 \leq -\frac{20}{9} C_g h^4 (h^{-1/3} - 8^{1/3}) + \frac{140}{9} C_g h^4 (8^{1/3} - 1) \tag{53}$$

and

$$\mathcal{S}_3 \leq -5 C_g \epsilon h^5 (8^{-2/3} - h^{2/3}) + 10 C_g h^4 (1 - 8^{-2/3}). \tag{54}$$

Hence, combining (42) and (52)–(54), we conclude that, for $h$ small, there exists a constant $C > 0$ such that

$$f(g_h) - f^*(g_h) \leq -\frac{0.25}{27} C_g h^{8/3} - C h^{11/3}.$$

Consequently,

$$|f(g_h) - f^*(g_h)| \geq \frac{0.25}{27} C_g h^{8/3} + C h^{11/3}.$$

Now, since

$$|a(g - g_h, u - u_h)| \leq C \|g - g_h\|_{H^1(\Omega)} \|u - u_h\|_{H^1(\Omega)} \leq C h^4 \|g\|_{H^3(\Omega)} \|u\|_{H^3(\Omega)},$$

using (41) we conclude that there exists a constant $C > 0$, such that

$$\begin{aligned} |G(u) - G(u_h^*)| &\geq \big| |f(g_h) - f^*(g_h)| - |a(g - g_h, u - u_h)| \big| \\ &\geq \frac{0.25}{27} C_g h^{8/3} + C h^{11/3}, \end{aligned}$$

which completes the proof of the desired result. □

We recall the definition of $F_{u_h^*}$ in (34), and we assume that the definite integral $\int_0^1 f\psi\,dx$ is evaluated exactly. We approximate the quantity of interest $F(u)$ by $F_{u_h^*}$. A lower bound of the error $|F(u)-F_{u_h^*}|$ is easily obtained using (35) and Theorem 4.5, which we state below.

**Theorem 4.6** *Let $u$ be the solution of (25) with $f(x)=(1-x)^{5/3}$, and $u_h^*$ be the finite element solution ($k=2$) obtained with the 2-point Gaussian quadrature. Suppose the integral $\int_0^1 f\psi\,dx$ in the definition of $F_{u_h^*}$ is computed exactly. Then, for the quantity of interest $F(u)=u'(0)$, there exists a constant $C>0$ such that*

$$|F(u)-F_{u_h^*}| = |G(u)-G(u_h^*)| \geq \frac{0.25}{27}C_g h^{8/3} + Ch^{11/3},$$

*for $h$ small, where the constant $C_g>0$ depends on the 2-point Gaussian quadrature.*

*Remark 4.7* If we choose $f(x)=(1-x)^{2/3}$ in our example (25), then $f\in H^1(\Omega)$ but $f\notin H^2(\Omega)$. Using the FEM with $k=1$ and 1-point or the 2-point Gaussian rule, we can show that $|F(u)-F_{u_h^*}|\geq Ch^{5/3}$. For the FEM with $k=2$, we can also obtain the same lower bound $|F(u)-F_{u_h^*}|\geq Ch^{5/3}$ with the 2-point Gaussian rule. We do not present the analysis here, but will illuminate these results through numerical examples presented in the next section.

*Remark 4.8* We note that the use of the Gaussian quadrature with more integration points will result into a smaller value of $C_g$, and thus the computed value of $|G(u)-G(u_h^*)|$ will be smaller. More Gaussian points, however, will not increase the order of convergence as $h\to 0$. We show this in our examples in the next section.

We recall that $g(x)=\epsilon\sin(\pi x)-0.5x\in H^3(\Omega)$. Moreover, since $f(x)=(1-x)^{5/3}\in H^2(\Omega)$, it is well known that $u\in H^4(\Omega)$. Therefore (32) holds for our particular problem, i.e., $|F(u)-F_{u_h}|\leq Ch^4$. Then, we can easily conclude from Theorem 4.6, that the ratio

$$R_h(u) := \frac{|F(u)-F_{u_h^*}|}{|F(u)-F_{u_h}|} = \frac{|G(u)-G(u_h^*)|}{|G(u)-G(u_h)|} \geq Ch^{-4/3}\to\infty, \quad \text{as} \quad h\to 0. \qquad (55)$$

Thus the ratio $R_h(u)$, which could be interpreted as a relative error in the computed quantity of interest due to numerical integration, becomes unbounded as $h\to 0$. In contrast, the ratio of $\|u-u_h^*\|_{H^1(\Omega)}$ to $\|u-u_h\|_{H^1(\Omega)}$ is bounded as $h\to 0$. We remark that a careful analysis with $\epsilon=0$ in (28) shows (in fact, the analysis is simpler) that the results in Theorem 4.5 and Theorem 4.6 are still true (with different constants). However, from the definition of $g$ in (29) with $\epsilon=0$, we have

$$F(u)-F_{u_h} = G(u)-G(u_h) = a(g, u-u_h) = a(g-g_h, u-u_h) = 0,$$

since $g=g_h$ (see Remark 4.3). Therefore, we will not be able to draw a meaningful conclusion about the ratio $R_h(u)$ as $h\to 0$. This is one of the reasons we used $\epsilon\neq 0$ in (28).

## 5 Numerical illustrations

In this section, we will present numerical results that will illuminate the results obtained in the last section. We will show that the optimal order of convergence in the approximation of functionals, i.e., the quantities of interest, is not obtained when the data do not have the increased regularity as specified in Theorem 3.6. We will further show that over-integration may yield the optimal order of convergence, at least in the pre-asymptotic range, i.e., when the mesh parameter $h$ is not too small.

We consider two boundary value problems, namely,

$$\begin{aligned} -u'' &= (1-x)^{5/3} \quad \text{in} \quad \Omega, \\ u(0) &= 0, \quad u(1) + u'(1) = 0, \end{aligned} \tag{56}$$

and

$$\begin{aligned} -u'' &= (1-x)^{2/3} \quad \text{in} \quad \Omega, \\ u(0) &= 0, \quad u(1) + u'(1) = 0, \end{aligned} \tag{57}$$

where $\Omega = (0, 1)$. We note that in (56), the data $f(x) = (1-x)^{5/3} \in H^2(\Omega)$, whereas in (57), the data $f(x) = (1-x)^{2/3} \in H^1(\Omega)$ but $\notin H^2(\Omega)$. The goal in both problems is to approximate the bounded linear functional $G(u) = \int_0^1 \psi' u' \, dx$, where $\psi(x) = 1 - x + \sin \pi x$. We recall that the functional $G(u)$ is related to the functional $F(u) = u'(0)$ as discussed in the last section.

In this section, we will consider the FEM with nodes $x_j = jh$, $0 \leq j \leq n$, with $nh = 1$, as described in the last section. We will denote the finite element solution with numerical integration by $u^*_{h,\ell}$, where the integrals in the load vector are computed by the $\ell$-point Gaussian integration rule. The order of convergence of $u^*_{h,\ell}$, for a fixed $\ell$, will be approximated by, rate $= \frac{\ln(|e_n/e_{2n}|)}{\ln(2)}$, where $e_n = |G(u) - G(u^*_{h,\ell})|$, $h = 1/n$.

We first consider the problem (56). We computed the finite element solutions $u_h$ and $u^*_{h,1}$, with $k = 1$, to approximate the solution $u$ of (56), where $u_h$ is computed with exact integration. We note that the algebraic precision of 1-point Gauss rule is 1, as required in Theorem 3.6 for $k = 1$. Also $f$ has the increased regularity for the case $k = 1$, i.e., $f \in H^2(\Omega)$. We have presented the values of $|G(u) - G(u_h)|$ and $|G(u) - G(u^*_{h,1})|$ for several values of $h$ in Table 1, where $n = 1/h$. We observe the optimal order of convergence $\mathcal{O}(h^2)$ for $|G(u) - G(u_h)|$ and $|G(u) - G(u^*_{h,1})|$, which illuminates the results in Lemma 2.2 and Theorem 3.6, respectively.

We next computed the finite element solutions $u_h$, $u^*_{h,2}$ and $u^*_{h,6}$, with $k = 2$, to approximate the solution $u$ of the same problem (56). We note that the algebraic precision of the 2-point Gauss rule is 3, as required in Theorem 3.6 for $k = 2$. But the data $f \notin H^4(\Omega)$, and thus does not have the increased regularity required in Theorem 3.6 for the case $k = 2$. We mention that $u^*_{h,6}$ is the finite element solution with "over-integration". In Table 2, we presented the computed values of $|G(u) - G(u_h)|$, $|G(u) - G(u^*_{h,2})|$ and $|G(u) - G(u^*_{h,6})|$.

1. We observe that $|G(u) - G(u_h)|$ converges with $\mathcal{O}(h^4)$, which is optimal. But $|G(u) - G(u^*_{h,2})|$ is not converging with the optimal order. In fact, the order of

**Table 1** $u_h$ and $u^*_{h,1}$ are piecewise linear ($k = 1$) finite element solutions with exact integration and with 1-point Gaussian quadrature, respectively, approximating the solution of (56)

| $n$ | $\|G(u) - G(u_h)\|$ | Rate | $\|G(u) - G(u^*_{h,1})\|$ | Rate |
|---|---|---|---|---|
| 10 | 1.833E–03 | | 1.663E–03 | |
| 20 | 4.569E–04 | 2.004 | 4.122E–04 | 2.012 |
| 40 | 1.142E–04 | 2.001 | 1.027E–04 | 2.005 |
| 80 | 2.853E–05 | 2.000 | 2.565E–05 | 2.002 |
| 160 | 7.133E–06 | 2.000 | 6.407E–06 | 2.001 |
| 320 | 1.783E–06 | 2.000 | 1.513E–06 | 2.001 |
| 640 | 4.458E–07 | 2.000 | 4.002E–07 | 2.000 |

The data has the increased regularity for $k = 1$

**Table 2** $u_h$, $u^*_{h,2}$, and $u^*_{h,6}$ are piecewise quadratic ($k = 2$) finite element solutions with exact integration, with 2-point Gaussian quadrature, and with 6-point Gaussian quadrature (over-integration), respectively, approximating the solution of the problem (56)

| $n$ | $\|G(u) - G(u_h)\|$ | Rate | $\|G(u) - G(u^*_{h,2})\|$ | Rate | $\|G(u) - G(u^*_{h,6})\|$ | Rate |
|---|---|---|---|---|---|---|
| 10 | 1.293E–06 | | 4.264E–06 | | 1.297E–06 | |
| 20 | 8.124E–08 | 3.992 | 3.840E–07 | 3.473 | 8.188E–08 | 3.985 |
| 40 | 5.086E–09 | 3.998 | 4.110E–08 | 3.224 | 5.186E–09 | 3.981 |
| 80 | 3.180E–10 | 3.999 | 5.154E–09 | 2.995 | 3.337E–10 | 3.958 |
| 160 | 1.987E–11 | 4.000 | 7.209E–10 | 2.838 | 2.234E–11 | 3.901 |
| 320 | 1.194E–12 | 4.000 | 1.072E–10 | 2.750 | 1.581E–12 | 3.820 |

The data does not have the increased regularity for $k = 2$

convergence is monotonically decreasing, and it appears that it is getting closer to $\mathcal{O}(h^{8/3})$, which is the order of the lower bound of the error given in Theorem 4.5.

2. We observe that the order of convergence of $|G(u) - G(u^*_{h,6})|$ is approximately $\mathcal{O}(h^4)$ initially for smaller values of $n$, but the convergence rate is slowing down for larger values of $n$, i.e., for small $h$.
3. We observe that the values of $|G(u) - G(u^*_{h,6})|$ are smaller than the values of $|G(u) - G(u^*_{h,2})|$, as indicated in Remark 4.8. Thus, over-integration yields smaller absolute values of the error, but does not affect the asymptotic convergence rate.

We remark that for $k = 2$, the data $f$ of the problem (56) has adequate regularity for the optimal order of convergence of the finite element solution in the energy norm (see Theorem 4.1.6 in [9]) and both $\|u - u^*_{h,2}\|_{H^1(\Omega)}$ and $\|u - u^*_{h,6}\|_{H^1(\Omega)}$ yield the optimal order of convergence, i.e., $\mathcal{O}(h^2)$. We have not presented numerical results to illuminate this fact in this paper.

We now consider the problem (57). First, we computed the finite element solutions $u_h$, $u^*_{h,1}$, and $u^*_{h,2}$, with $k = 1$, to approximate the solution $u$ of the problem (57). Note that $u^*_{h,2}$ is the finite element solution with over-integration. We have summarized the results in Table 3. We then computed the finite element solutions $u_h$, $u^*_{h,2}$, and $u^*_{h,50}$, with $k = 2$, to approximate the solution $u$ of the problem (57); $u^*_{h,50}$ is the finite element solution with over-integration. We have summarized the results for $k = 2$ in Table 4. We note that the Gauss rules employed to compute $u^*_{h,1}$ in Table 3 and $u^*_{h,2}$

**Table 3** $u_h$, $u^*_{h,1}$, and $u^*_{h,2}$ are piecewise linear ($k = 1$) finite element solutions with exact integration, with 1-point Gaussian quadrature, and with 2-point Gaussian quadrature (over-integration), respectively, approximating the solution of the problem (57)

| $n$ | $\lvert G(u) - G(u_h)\rvert$ | Rate | $\lvert G(u) - G(u^*_{h,1})\rvert$ | Rate | $\lvert G(u) - G(u^*_{h,2})\rvert$ | Rate |
|---|---|---|---|---|---|---|
| 10 | 3.220E–03 | | 4.270E–03 | | 3.274E–03 | |
| 20 | 8.047E–04 | 2.001 | 1.114E–03 | 1.938 | 8.203E–04 | 1.997 |
| 40 | 2.011E–04 | 2.000 | 2.924E–04 | 1.930 | 2.059E–04 | 1.994 |
| 80 | 5.028E–05 | 2.000 | 7.725E–05 | 1.920 | 5.175E–05 | 1.992 |
| 160 | 1.257E–05 | 2.000 | 2.060E–05 | 1.907 | 1.303E–05 | 1.990 |
| 320 | 3.413E–06 | 2.000 | 5.549E–06 | 1.892 | 3.286E–06 | 1.987 |
| 640 | 7.856E–07 | 2.000 | 1.513E–06 | 1.875 | 8.306E–07 | 1.984 |

The data does not have the increased regularity for $k = 1$

**Table 4** $u_h$, $u^*_{h,2}$, and $u^*_{h,50}$ are piecewise quadratic ($k = 2$) finite element solutions with exact integration, with 2-point Gaussian quadrature, and with 50-point Gaussian quadrature (over-integration), respectively, approximating the solution of the problem (57)

| $n$ | $\lvert G(u) - G(u_h)\rvert$ | Rate | $\lvert G(u) - G(u^*_{h,2})\rvert$ | Rate | $\lvert G(u) - G(u^*_{h,50})\rvert$ | Rate |
|---|---|---|---|---|---|---|
| 10 | 8.647E–07 | | 5.492E–05 | | 8.665E–07 | |
| 20 | 5.695E–08 | 3.925 | 1.575E–05 | 1.802 | 5.751E–08 | 3.913 |
| 40 | 3.670E–09 | 3.956 | 4.750E–06 | 1.730 | 3.847E–09 | 3.902 |
| 80 | 2.337E–10 | 3.973 | 1.465E–06 | 1.697 | 2.895E–10 | 3.732 |
| 160 | 1.478E–11 | 3.982 | 4.566E–07 | 1.682 | 3.237E–11 | 3.161 |

The data does not have the increased regularity for $k = 2$

in Table 4 have the algebraic precision, as required in Theorem 3.6. But the data $f$ of problem (57) does not have the increased regularity required in Theorem 3.6, either for $k = 1$ or for $k = 2$.

1. We observe from Tables 3 and 4 that $|G(u) - G(u_h)|$ converges with the optimal order, $\mathcal{O}(h^{2k})$, $k = 1, 2$, as expected in Lemma 2.2, since $u \in H^3(\Omega)$.
2. It is clear from Table 3 that the rate of convergence of $|G(u) - G(u^*_{h,1})|$ is decreasing with increasing values of $n$. We recall that the lower bound of the error is $\mathcal{O}(h^{5/3})$, as mentioned in Remark 4.7. The rate of convergence of $|G(u) - G(u^*_{h,2})|$ in Table 3 is close to optimal, i.e., $\mathcal{O}(h^2)$, for smaller values of $n$, but the rate slows down as $n$ increases. This indicates the over-integration does not yield the optimal rate of convergence. However, the over-integration reduces the size of the error.
3. Table 4 shows that the rate of convergence of $|G(u) - G(u^*_{h,2})|$ decreases with increasing $n$ and indicates that the rate is approaching $\mathcal{O}(h^{5/3})$—the lower bound of the error as mentioned in Remark 4.7. Also the third column of Table 4 clearly indicates that over-integration does not yield optimal order of convergence, but certainly reduces the size of the error.

These experiments indicate that the increased regularity assumptions to obtain the optimal order of convergence in Theorem 3.6 is, in general, necessary. Also, over-integration reduces the size of the error, and may even yield the optimal order of convergence in the pre-asymptotic range. But over-integration does not yield the asymptotic optimal rate of convergence.

## References


1. Babuška, I., Miller, A.: The post-processing approach in the finite element method. Part I: calculation of displacements, stresses and higher order derivatives of displacement. Int. J. Numer. Methods Eng. **20**, 1085–1109 (1984)
2. Babuška, I., Miller, A.: The post-processing approach in the finite element method. Part II: calculation of stress intensity factors. Int. J. Numer. Methods Eng. **20**, 1111–1129 (1984)
3. Babuška, I., Osborn, J.: Eigenvalue problems. In: Handbook of Numerical Analysis, vol. II, pp. 641–787. North-Holland, Amsterdam (1991)
4. Babuška, I., Strouboulis, T.: The Finite Element Method and Its Reliability. Numerical Mathematics and Scientific Computation. Oxford University Press, New York (2001)
5. Babuška, I., Aziz, A.K.: Survey lectures on the mathematical foundation of the finite element method. In: Aziz, A.K. (ed.) The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, pp. 3–359. Academic Press, New York (1972)
6. Banerjee, U., Osborn, J.E.: Estimation of the effect of numerical integration in finite element eigenvalue approximation. Numer. Math. **56**, 735–762 (1990)
7. Banerjee, U., Suri, M.: The effect of numeical quadrature in the $p$-version of the finite element method. Math. Comput. **59**, 1–20 (1992)
8. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics. Springer, New York (2002)
9. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Studies in Mathematics and Its Applications, vol. 4. North-Holland, Amsterdam (1978)
10. Ciarlet, P.G., Raviart, P.-A.: The combined effect of curved boundaries and numerical integration in isoparametric finite element methods. In: Aziz, A.K. (ed.) The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, pp. 404–474. Academic Press, New York (1972)
11. Fix, G.J.: Effects of quadrature errors in finite element approximation of steady state, eigenvalue and parabolic problems. In: Aziz, A.K. (ed.) The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, pp. 525–556. Academic Press, New York (1972)
12. Herbold, R.J., Schultz, M.H., Varga, R.S.: The effect of quadrature errors in the numerical solution of boundary value problems by variational techniques. Aequ. Math. **3**, 247–270 (1969)
13. Herbold, R.J., Varga, R.S.: The effect of quadrature errors in the numerical solution of two-dimensional boundary value problems by variational techniques. Aeq. Math. **7**, 36–58 (1972)
14. Schatz, A.H., Wahlbin, L.B.: Interior maximum-norm estimates for finite element methods, part II. Math. Comput. **64**(221), 907–928 (1995)
15. Szabó, B., Babuška, I.: Finite Element Analysis. Wiley, New York (1991)
16. Wahlbin, L.B.: Superconvergence in the Galerkin Finite Element Methods. Lecture Notes in Mathematics, vol. 1605. Springer, New York (1991)