# A TRIANGULAR SPECTRAL ELEMENT METHOD USING FULLY TENSORIAL RATIONAL BASIS FUNCTIONS\*

JIE SHEN<sup>†</sup>, LI-LIAN WANG<sup>‡</sup>, AND HUIYUAN LI<sup>§</sup>

Abstract. A rational approximation in a triangle is proposed and analyzed in this paper. The rational basis functions in the triangle are obtained from the polynomials in the reference square through a collapsed coordinate transform. Optimal error estimates for the  $L^2$ - and  $H_0^1$ -orthogonal projections are derived with upper bounds expressed in the original coordinates in the triangle. It is shown that the rational approximation is as accurate as the polynomial approximation in the triangle. Illustrative numerical results, which are in agreement with the theoretical estimates, are presented.

 ${\bf Key}$  words. spectral method in a triangle, rational functions, error analysis, spectral element method

AMS subject classifications. 65N35, 65N22, 65F05, 35J05

#### DOI. 10.1137/070702023

1. Introduction. We study in this paper a spectral approximation using rational functions in a triangle. Besides its relevance in approximation theory, this work is motivated by the fact that triangular elements are more flexible for complex geometries and for adaptivity. Unlike the spectral methods in rectangular domains, spectral methods for triangular domains received only limited attentions. In general, spectral methods in a triangle can be classified into three categories: (i) approximations by polynomials in a triangle through mapping (cf. [22, 18, 8, 20, 26, 5, 17] and the references therein); (ii) approximations by nonpolynomial functions in a triangle through mapping (cf. [4, 15]); and (iii) approximation by polynomials in a triangle using special nodal points such as Fekete points (cf. [16, 27, 21]).

The motivation of using polynomials in triangles is obvious: after all, most spectral methods are based on polynomial (and Fourier) approximations. However, by restricting to polynomials, one also loses some flexibility and faces some complicated implementation issues. Take, for example, the polynomial basis:

(1.1) 
$$\mathcal{G}_{lm}(x,y) = (1-y)^l J_l^{0,0} \left(\frac{2x}{1-y} - 1\right) J_m^{2l+1,0}(2y-1), \quad \forall (x,y) \in \mathcal{T},$$

where  $\mathcal{T} := \{(x,y) : 0 < x, y < 1, 0 < x + y < 1\}$  is the reference triangle, and  $J_k^{\alpha,\beta}(z), z \in (-1,1)$  is the Jacobi polynomial of degree k. Note that  $\mathcal{G}_{lm}(x,y)$  are special cases of the polynomials considered in [22, 18] and were used in [8, 26] to construct their spectral-element methods in triangles. In contrast to the standard

<sup>\*</sup>Received by the editors September 5, 2007; accepted for publication (in revised form) December 11, 2008; published electronically April 22, 2009.

http://www.siam.org/journals/sinum/47-3/70202.html

<sup>&</sup>lt;sup>†</sup>Department of Mathematics, Purdue University, West Lafayette, IN, 47907 (shen@math.purdue. edu). The work of this author was partially supported by NFS grant DMS-0610646 and AFOSR grant FA9550-08-1-0416.

<sup>&</sup>lt;sup>‡</sup>Division of Mathematics, SPMS, Nanyang Technological University, 637616, Singapore (lilian@ ntu.edu.sg). The work of this author was partially supported by ACRF Tier 1 Grant RG58/08, Singapore MOE Grant T207B2202, and Singapore NRF2007 IDM002-010.

 $<sup>^{\</sup>S}$ Institute of Software, Chinese Academy of Sciences, Beijing 100080, China (lihuiyuan@gmail. com). The work of this author was partially supported by NSFC grants 10601056, 10431050, and 60573023.

tensor product basis functions, these basis functions involve a wrapped (since the two subscripts in  $\mathcal{G}_{lm}$  are dependent) tensor product and use the Jacobi polynomials with a variable parameter in  $J_m^{2l+1,0}$ , which make it difficult to analyze (cf. [12]) and somewhat difficult to implement. Moreover, no nodal basis corresponding to (1.1) is available, further complicating the implementation procedure.

In this paper, we take an approach in the second category. Namely, we consider rational functions generated by polynomials in the reference square through the Duffy transform (3.2)-(3.3) below (cf. [9]), that is,

(1.2) 
$$\mathcal{R}_{lm}(x,y) := J_l^{0,0} \left(\frac{2x}{1-y} - 1\right) J_m^{1,0}(2y-1), \quad \forall \ (x,y) \in \mathcal{T}.$$

These basis functions involve usual tensor product of Jacobi polynomials with indices (0,0) and (1,0) so they are easier to deal with in practice. Furthermore, one can construct a nodal basis corresponding to (1.2), making it suitable for implementations in a spectral-element framework.

An important question is whether this rational approximation in the triangle is as accurate as or better than the polynomial approximation? We shall answer this question by performing error analysis in the original coordinates in the triangle, and by presenting illustrative numerical results.

The difficulty in obtaining error bounds in the original coordinates is that the Duffy transform introduces a coordinate singularity, similar to the polar and spherical coordinate transforms. However, in polar and spherical geometries, one actually prefers to write the equations in polar and spherical coordinates rather than the original Cartesian coordinates, so it is natural to work in the "transformed" polar or spherical coordinates. But for triangular domains, one obviously prefers to work with the original coordinates.

We note that it is considerably easier to derive error estimates in the transformed coordinates. However, these error estimates would involve complicated norms in the reference square that cannot be easily quantified in terms of usual norms in the triangle, and consequently, cannot be used to compare directly with polynomial approximations in the triangle. However, sensible comparison can be performed if error estimates in the original coordinates are available.

While using a rational approximation in triangles enjoys the aforementioned advantages, it also has some drawbacks. Namely, the collocation points in the triangle are severely clustered near the singular vertex (cf. Figure 4.2, right), resulting in an unfavorable spectral radius for the derivative matrix, which in turns limits the size of allowable time steps for any explicit or semi-implicit time stepping schemes. However, the effect of this drawback will be limited in practice, since the method is to be used in the context of spectral-element methods in which the degree of polynomials used in each triangle will not be too large. This situation is reminiscent to the comparison between spherical harmonics and double Fourier series on the sphere (see a discussion in [4]).

The rest of the paper is organized as follows. In the next section, we present some preliminary results which will be used in the sequel. The main results on rational approximations are provided in section 3. In section 4, we provide error estimates and implementation details for the mono-domain rational approximation to a Poisson type equation, and describe briefly the result for the multidomain spectral element approximations. We present in section 5 some numerical results which are consistent with our error analysis and conclude with some remarks.

#### 2. Preliminaries.

- 2.1. Notations. Throughout this paper, we will use the following notations.
  - Let  $\Omega$  be a bounded domain, and  $\omega$  be a generic positive weight function which is not necessary in  $L^1(\Omega)$ . Denote by  $(u, v)_{\omega,\Omega} := \int_{\Omega} uv\omega d\Omega$  the inner product of  $L^2_{\omega}(\Omega)$  whose norm is denoted by  $\|\cdot\|_{\omega,\Omega}$ . We use  $H^m_{\omega}(\Omega)$  and  $H^m_{0,\omega}(\Omega)$  to denote the usual weighted Sobolev spaces, whose norms and seminorms are denoted by  $\|u\|_{m,\omega,\Omega}$  and  $|u|_{m,\omega,\Omega}$ , respectively. In cases where no confusion would arise,  $\omega$  (if  $\omega \equiv 1$ ) and  $\Omega$  may be dropped from the notations.
  - Let  $\mathbb{N}$  be the set of all nonnegative integers. For any  $N \in \mathbb{N}$ , we set I = (-1, 1)and denote by  $\mathcal{P}_N(I)$  the set of all polynomials of degree  $\leq N$ , and set  $\mathcal{P}_N^0(I) := \{\phi \in \mathcal{P}_N(I) : \phi(\pm 1) = 0\}.$
  - We denote by c a generic positive constant independent of any function and of any discretization parameters. We use the expression  $A \leq B$  to mean that  $A \leq cB$ .

2.2. One-dimensional approximation results. We derive and refine in this section some one-dimensional results on Jacobi polynomial approximations and Jacobi–Gauss-type interpolation approximations, which play important roles in the error analysis of rational approximation in the triangle.

**2.2.1. Jacobi polynomial approximations.** The classical Jacobi polynomials, denoted by  $J_k^{\alpha,\beta}(\zeta), \zeta \in I$  with  $\alpha, \beta > -1$ , are mutually orthogonal with respect to the Jacobi weight function  $\omega^{\alpha,\beta}(\zeta) = (1-\zeta)^{\alpha}(1+\zeta)^{\beta}$ :

(2.1) 
$$\int_{-1}^{1} J_{n}^{\alpha,\beta}(\zeta) J_{m}^{\alpha,\beta}(\zeta) \omega^{\alpha,\beta}(\zeta) d\zeta = \left\| J_{n}^{\alpha,\beta} \right\|_{\omega^{\alpha,\beta},I}^{2} \delta_{mn},$$

where  $\delta_{mn}$  is the Kronecker symbol, and

(2.2) 
$$\left\|J_{n}^{\alpha,\beta}\right\|_{\omega^{\alpha,\beta},I}^{2} = \frac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}.$$

Notice that the classical Jacobi polynomials are defined only for  $\alpha, \beta > -1$ , while in a recent work [11], the definition of Jacobi polynomials are extended to cases where  $\alpha$  and/or  $\beta$  are negative integers.

We define the orthogonal projection  $\pi_N^{\alpha,\beta}$  :  $L^2_{\omega^{\alpha,\beta}}(I) \to \mathcal{P}_N(I) \cap L^2_{\omega^{\alpha,\beta}}(I)$  by

(2.3) 
$$\int_{-1}^{1} \left( \pi_N^{\alpha,\beta} w - w \right) \varphi \omega^{\alpha,\beta} d\zeta = 0, \quad \forall \varphi \in \mathcal{P}_N(I) \cap L^2_{\omega^{\alpha,\beta}}(I).$$

Note that for  $\alpha, \beta > -1$ , we have  $\mathcal{P}_N(I) \cap L^2_{\omega^{\alpha,\beta}}(I) = \mathcal{P}_N(I)$ , but when  $\alpha$  and/or  $\beta$  are negative integers, suitable boundary conditions are involved. For example,  $\mathcal{P}_N(I) \cap L^2_{\omega^{-1,-1}}(I) = \mathcal{P}_N^0(I) = \mathcal{P}_N(I) \cap H^1_0(I)$ .

To describe the approximation errors, we introduce the nonuniformly weighted Sobolev space

(2.4) 
$$B^{\sigma}_{\alpha,\beta}(I) := \left\{ w \in L^2_{\omega^{\alpha,\beta}}(I) : w^{(k)}(\zeta) \in L^2_{\omega^{\alpha+k,\beta+k}}(I), \quad 0 \le k \le \sigma \right\}, \quad \forall \ \sigma \in \mathbb{N},$$

equipped with the norm and seminorm

(2.5) 
$$\|w\|_{B^{\sigma}_{\alpha,\beta}(I)} := \left(\sum_{k=0}^{\sigma} \left\|w^{(k)}\right\|_{\omega^{\alpha+k,\beta+k},I}^{2}\right)^{\frac{1}{2}}, \quad |w|_{B^{\sigma}_{\alpha,\beta}(I)} := \left\|w^{(\sigma)}\right\|_{\omega^{\alpha+\sigma,\beta+\sigma},I}.$$

Let  $\mathbb{I}$  be the identity operator. It is obvious that for any  $N, M \in \mathbb{N}$ , and  $N \ge M$ , the projection operator  $\pi_N^{\alpha,\beta}$  satisfies

$$(2.6) \quad \left\| \left( \pi_N^{\alpha,\beta} - \mathbb{I} \right) w \right\|_{\omega^{\alpha,\beta},I} \le \left\| \left( \pi_M^{\alpha,\beta} - \mathbb{I} \right) w \right\|_{\omega^{\alpha,\beta},I} \le \|w\|_{\omega^{\alpha,\beta},I}, \quad \forall w \in L^2_{\omega^{\alpha,\beta}}(I).$$

Another property is as follows.

LEMMA 2.1. If  $w \in L^2_{\omega^{\alpha,\beta}}(I)$  and  $w' \in L^2_{\omega^{\alpha+2,\beta+2}}(I)$  with  $\alpha, \beta > -1$ , then we have

(2.7) 
$$\left\| \left( \pi_N^{\alpha,\beta} - \mathbb{I} \right) w \right\|_{\omega^{\alpha,\beta},I} \lesssim \|w'\|_{\omega^{\alpha+2,\beta+2},I}.$$

*Proof.* For any  $w \in L^2_{\omega^{\alpha,\beta}}(I)$ , we write

(2.8) 
$$w(x) = \sum_{k=0}^{\infty} \hat{w}_k^{\alpha,\beta} J_k^{\alpha,\beta}(x), \qquad \hat{w}_k^{\alpha,\beta} = \frac{1}{\left\|J_k^{\alpha,\beta}\right\|_{\omega^{\alpha,\beta},I}^2} \int_{-1}^1 w(x) J_k^{\alpha,\beta}(x) \omega^{\alpha,\beta}(x) dx.$$

Define

$$g(\zeta) := w(\zeta) - \left(\pi_0^{\alpha,\beta} w\right)(\zeta) = w(\zeta) - \hat{w}_0^{\alpha,\beta}.$$

We next show that there exists  $\zeta_0 \in (-1, 1)$  such that  $g(\zeta_0) = 0$ . Indeed, one verifies readily that w is pointwise continuous in (-1, 1). Thus, by the orthogonality (2.1) and the intermediate-value theorem of integration,

$$0 = \int_{-1}^{1} \left( \sum_{k=1}^{\infty} \hat{w}_{k}^{\alpha,\beta} J_{k}^{\alpha,\beta}(\zeta) \right) J_{0}^{\alpha,\beta}(\zeta) \omega^{\alpha,\beta}(\zeta) d\zeta = \int_{-1}^{1} g(\zeta) \omega^{\alpha,\beta}(\zeta) d\zeta$$
$$= g(\zeta_{0}) \int_{-1}^{1} \omega^{\alpha,\beta}(\zeta) d\zeta \implies g(\zeta_{0}) = 0.$$

To proceed the proof, we recall the embedding result (see, e.g., [2, 13]): If  $v \in L^2_{\omega^{\alpha,\beta}}(I)$ and  $v' \in L^2_{\omega^{\alpha+2,\beta+2}}(I)$  with  $v(\zeta_0) = 0$  for certain  $\zeta_0 \in (-1, 1)$ , then we have

(2.9) 
$$\|v\|_{\omega^{\alpha,\beta},I} \lesssim \|v'\|_{\omega^{\alpha+2,\beta+2},I}, \quad \text{for } \alpha,\beta > -1.$$

Π

Therefore, by (2.6) and (2.9),

$$\begin{aligned} \left\| \left( \pi_N^{\alpha,\beta} - \mathbb{I} \right) w \right\|_{\omega^{\alpha,\beta},I} &\leq \left\| \left( \pi_0^{\alpha,\beta} - \mathbb{I} \right) w \right\|_{\omega^{\alpha,\beta},I} = \left\| g \right\|_{\omega^{\alpha,\beta},I} \lesssim \left\| g' \right\|_{\omega^{\alpha+2,\beta+2},I} \\ &= \left\| \left( \pi_0^{\alpha,\beta} w - w \right)' \right\|_{\omega^{\alpha+2,\beta+2},I} = \|w'\|_{\omega^{\alpha+2,\beta+2},I}, \end{aligned}$$

which ends the proof.

The following approximation result plays an fundamental role in the analysis in the forthcoming section. It is a generalization of the classical approximation results for Jacobi polynomials established in [10].

LEMMA 2.2. Let  $\alpha, \beta > -1$  or be negative integers. Then for any  $w \in B^{\sigma}_{\alpha,\beta}(I)$ with integers  $\sigma \geq \mu \geq 0$ ,

(2.10) 
$$\left\| \left( \pi_{N}^{\alpha,\beta} w - w \right)^{(\mu)} \right\|_{\omega^{\alpha+\mu,\beta+\mu},I} \lesssim N^{\mu-\sigma} \left\| \left( \pi_{N-\sigma}^{\alpha+\sigma,\beta+\sigma} - \mathbb{I} \right) w^{(\sigma)} \right\|_{\omega^{\alpha+\sigma,\beta+\sigma,I}} \\ \lesssim N^{\mu-\sigma} \left\| w^{(\sigma)} \right\|_{\omega^{\alpha+\sigma,\beta+\sigma},I}.$$

The proof of Lemma 2.2 will be postponed to Appendix A.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

**2.2.2.** Jacobi–Gauss-type interpolation approximation. In this paper, the interpolation in the triangle  $\mathcal{T}$  is based on a map of the tensorial product of Legendre– Gauss-Lobatto (LGL) and Jacobi-Gauss-Radau (JGR) interpolations in the rectangle  $\mathcal{Q}$ .

We first consider the one-dimensional Legendre-Gauss-Lobatto interpolation. Let  $\{\xi_i^L\}_{i=0}^N$  be the LGL interpolation points (i.e., the zeros of  $(1-\xi^2)L'_N(\xi)$ ). For any  $w \in C(\overline{I})$ , the LGL interpolant  $\mathcal{I}_N^L w \in \mathcal{P}_N(I)$  and satisfies

(2.11) 
$$\left(\mathcal{I}_N^L w\right)\left(\xi_j^L\right) = w\left(\xi_j^L\right), \quad 0 \le j \le N.$$

We have the following approximation results. Lemma 2.3. If  $w \in L^2(I)$  and  $w' \in B_{0,0}^{\sigma-1}(I)$  with integers  $\sigma \ge \mu \ge 0$  and  $\sigma \ge 1$ , then we have

(2.12) 
$$\left\| \left( \mathcal{I}_N^L w - w \right)^{(\mu)} \right\|_{\omega^{\mu-1,\mu-1},I} \lesssim N^{\mu-\sigma} \left\| \left( \pi_{N-\sigma}^{\sigma-1,\sigma-1} - \mathbb{I} \right) w^{(\sigma)} \right\|_{\omega^{\sigma-1,\sigma-1},I} \\ \lesssim N^{\mu-\sigma} \left\| w^{(\sigma)} \right\|_{\omega^{\sigma-1,\sigma-1},I}.$$

*Proof.* Let

$$w_b(\xi) = \frac{1+\xi}{2}w(1) + \frac{1-\xi}{2}w(-1)$$

Clearly,  $(w - w_b)(\pm 1) = 0$  and  $w - w_b \in B^1_{-1,-1}(I)$ . Further, define

$$w_N = \pi_N^{-1,-1}(w - w_b) + w_b \in \mathcal{P}_N(I).$$

Observe that  $w_N - w = (\pi_N^{-1,-1} - \mathbb{I})(w - w_b)$ . Thus, by Lemma 2.2 with  $\alpha = \beta = -1$ ,

In the last step, we used the fact  $\pi_N^{\alpha,\beta}\phi = \phi$  for any  $\phi \in \mathcal{P}_N(I)$  with  $N \ge 1$  and  $\alpha, \beta > -1$ , which implies

$$\left(\pi_{N-\sigma}^{\sigma-1,\sigma-1}-\mathbb{I}\right)\left(w^{(\sigma)}-w_{b}^{(\sigma)}\right)=\left(\pi_{N-\sigma}^{\sigma-1,\sigma-1}-\mathbb{I}\right)w^{(\sigma)},\quad\sigma\geq1.$$

We now recall the inverse inequality (see, e.g., Lemma 3.3 of [13]):

(2.14) 
$$\left\|\phi^{(\mu)}\right\|_{\omega^{\mu-1,\mu-1},I} \lesssim N^{\mu} \|\phi\|_{\omega^{-1,-1},I}, \quad \forall \phi \in \mathcal{P}_{N}^{0}(I).$$

We also use the stability result of the LGL interpolation operator (see Theorem 4.9 of [13]):

(2.15) 
$$\left\|\mathcal{I}_{N}^{L}v\right\|_{\omega^{-1,-1},I} \lesssim \|v\|_{\omega^{-1,-1},I} + N^{-1}\|v'\|_{I}, \quad \forall v \in B^{1}_{-1,-1}(I).$$

Therefore, by the triangular inequality, (2.13)-(2.15) and (2.6),

$$\begin{split} \left\| \left( \mathcal{I}_{N}^{L} w - w \right)^{(\mu)} \right\|_{\omega^{\mu-1,\mu-1},I} &\leq \left\| \left( \mathcal{I}_{N}^{L} (w_{N} - w) \right)^{(\mu)} \right\|_{\omega^{\mu-1,\mu-1},I} + \left\| (w_{N} - w)^{(\mu)} \right\|_{\omega^{\mu-1,\mu-1},I} \\ &\stackrel{(2.14)}{\lesssim} N^{\mu} \left\| \mathcal{I}_{N}^{L} (w_{N} - w) \right\|_{\omega^{-1,-1},I} + \left\| (w_{N} - w)^{(\mu)} \right\|_{\omega^{\mu-1,\mu-1},I} \\ &\stackrel{(2.15)}{\lesssim} N^{\mu} \| w_{N} - w \|_{\omega^{-1,-1},I} + N^{\mu-1} \| (w_{N} - w)' \|_{I} + \left\| (w_{N} - w)^{(\mu)} \right\|_{\omega^{\mu-1,\mu-1},I} \\ &\stackrel{(2.13)}{\lesssim} N^{\mu-\sigma} \left\| \left( \pi_{N-\sigma}^{\sigma-1,\sigma-1} - \mathbb{I} \right) w^{(\sigma)} \right\|_{\omega^{\sigma-1,\sigma-1},I} \stackrel{(2.6)}{\lesssim} N^{\mu-\sigma} \left\| w^{(\sigma)} \right\|_{\omega^{\sigma-1,\sigma-1},I}. \end{split}$$

This ends the proof.  $\Box$ 

We now turn to the Jacobi–Gauss–Radau (JGR) interpolation associated with the weight function  $\omega^{1,0}(\eta) = 1 - \eta$ . Let  $\{\eta_j^R\}_{j=0}^N$  be the JGR interpolation points (i.e., the zeros of  $(1 + \eta)J_N^{0,1}(\eta)$ ). For any  $w \in C([-1,1))$ , the JGR interpolant is defined by  $\mathcal{I}_N^R w \in \mathcal{P}_N(I)$  and

(2.16) 
$$\left(\mathcal{I}_N^R w\right) \left(\eta_j^R\right) = w\left(\eta_j^R\right), \quad 0 \le j \le N.$$

In order to establish the error estimate, we use the stability result of the JGR interpolation operator (cf. Theorem 4.6 of [14]). Letting  $w_N(\eta) = \int_{-1}^{\eta} \pi_{N-1}^{2,1} w'(\eta) d\eta + w(-1)$ and using Lemma 2.2 of [14] and a similar argument to Lemma 2.3, we can derive the following result for  $\mathcal{I}_N^R$ .

LEMMA 2.4. If  $w \in B^{\sigma}_{1,0}(I)$  with integers  $\sigma \geq \mu \geq 0$  and  $\sigma \geq 1$ , then

$$\begin{aligned} & (2.17) \\ & \left\| \left( \mathcal{I}_N^R w - w \right)^{(\mu)} \right\|_{\omega^{\mu+1,\mu},I} \lesssim N^{\mu-\sigma} \left\| \left( \pi_{N-\sigma}^{\sigma+1,\sigma} - \mathbb{I} \right) w^{(\sigma)} \right\|_{\omega^{\sigma+1,\sigma},I} \lesssim N^{\mu-\sigma} \left\| w^{(\sigma)} \right\|_{\omega^{\sigma+1,\sigma},I}. \end{aligned}$$

3. Rational approximations in the triangle. In this section, we introduce a rational orthogonal system in the triangle  $\mathcal{T}$  and study the approximation property of the associated  $L^2$ - and  $H_0^1$ -projection operators.

**3.1. Coordinate transform.** We consider two coordinate systems: the Cartesian coordinate (x, y)-system for the reference triangle:

(3.1) 
$$\mathcal{T} := \{ (x, y) : 0 < x, y < 1, 0 < x + y < 1 \},\$$

and the  $(\xi, \eta)$ -system for the square:  $\mathcal{Q} = (-1, 1)^2$ . The one-to-one transformation between  $\mathcal{T}$  and  $\mathcal{Q}$  is given by

(3.2) 
$$x = \frac{1}{4}(1+\xi)(1-\eta), \quad y = \frac{1+\eta}{2}, \quad \forall (\xi,\eta) \in \mathcal{Q},$$

with the inversion

(3.3) 
$$\xi = \frac{2x}{1-y} - 1, \qquad \eta = 2y - 1, \quad \forall (x,y) \in \mathcal{T}.$$

This mapping collapses one edge:  $\eta = 1, -1 \leq \xi \leq 1$ , of the rectangle  $\mathcal{Q}$  into a vertex (0,1) of the triangle  $\mathcal{T}$ , so it is referred to as a *collapsed coordinate system* in [17] (also called the Duffy's system (cf. [9])). We point out that this type of coordinate

singularity also occurs in polar and spherical coordinate transformations. Notice that there holds

(3.4) 
$$0 < \frac{x}{1-y} = \frac{1+\xi}{2} \le 1, \quad \forall (x,y) \in \mathcal{T} \text{ and } (\xi,\eta) \in \mathcal{Q}.$$

The following formulas associated with the transformation (3.2)–(3.3) will be used frequently throughout this paper:

(3.5) 
$$\frac{\partial\xi}{\partial x} = \frac{2}{1-y} = \frac{4}{1-\eta}, \quad \frac{\partial\xi}{\partial y} = \frac{2x}{(1-y)^2} = \frac{2(1+\xi)}{1-\eta}, \quad \frac{\partial\eta}{\partial x} = 0, \quad \frac{\partial\eta}{\partial y} = 2,$$

and

$$(3.6) \qquad \frac{\partial x}{\partial \xi} = \frac{1-y}{2} = \frac{1-\eta}{4}, \quad \frac{\partial x}{\partial \eta} = \frac{x}{2(1-y)} = -\frac{1+\xi}{4}, \quad \frac{\partial y}{\partial \xi} = 0, \quad \frac{\partial y}{\partial \eta} = \frac{1}{2}.$$

From the above, one easily finds that the determinant of the Jacobian for (3.2)–(3.3) is given by

(3.7) 
$$\det\left(\frac{\partial(x,y)}{\partial(\xi,\eta)}\right) = \frac{1-\eta}{8} = \frac{1-y}{4}.$$

Throughout the paper, we shall associate a function u in  $\mathcal{T}$  with a function v in  $\mathcal{Q}$  through

(3.8) 
$$v(\xi,\eta) = u(x,y), \quad \forall (x,y) \in \mathcal{T}, \ \forall (\xi,\eta) \in \mathcal{Q}.$$

One verifies readily that

(3.9) 
$$\nabla u = (\partial_x u, \ \partial_y u)^T = \left(\frac{4}{1-\eta}\partial_\xi v, \ \frac{2(1+\xi)}{1-\eta}\partial_\xi v + 2\partial_\eta v\right)^T,$$

and conversely,

(3.10) 
$$\tilde{\nabla}v := \left(\partial_{\xi}v, \ \partial_{\eta}v\right)^{T} = \left(\frac{1-y}{2}\partial_{x}u, \ \frac{x}{2(1-y)}\partial_{x}u + \frac{1}{2}\partial_{y}u\right)^{T}.$$

In particular, we have

(3.11) 
$$\partial_{\xi} v(\xi, 1) = 0$$
 a.e. if  $\partial_x u$  is a measurable function.

We note that the above corresponds to the pole conditions in the polar and spherical coordinates.

**3.2. Error estimates of the**  $L^2$ **-orthogonal projection.** We now consider the approximation of functions in  $L^2(\mathcal{T})$  by using series of the rational basis  $\{\mathcal{R}_{lm}\}$  defined in (1.2). We first notice that  $\{\mathcal{R}\}_{lm}$  are mutually orthogonal in  $L^2(\mathcal{T})$ . Define

(3.12) 
$$\widetilde{\mathcal{R}}_{lm}(\xi,\eta) := \mathcal{R}_{lm}(x,y) = J_l^{0,0}(\xi) J_m^{1,0}(\eta), \quad \forall \ (\xi,\eta) \in \mathcal{Q}.$$

Since the determinant of the Jacobian for the mapping (3.2) is  $\frac{1-\eta}{8}$ , we derive from the orthogonality of the Jacobi polynomials (2.1) and (3.7) that

$$\iint_{\mathcal{T}} \mathcal{R}_{lm}(x,y) \mathcal{R}_{l'm'}(x,y) dx dy = \frac{1}{8} \int_{-1}^{1} J_{l}^{0,0}(\xi) J_{l'}^{0,0}(\xi) d\xi \int_{-1}^{1} J_{m}^{1,0}(\eta) J_{m'}^{1,0}(\eta) (1-\eta) d\eta$$
$$= \gamma_{lm} \delta_{ll'} \delta_{mm'}, \quad \text{with} \quad \gamma_{lm} = \frac{1}{2(m+1)(2l+1)}.$$

For any function  $u \in L^2(\mathcal{T})$ , we write

(3.14) 
$$u(x,y) = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \hat{u}_{lm} \mathcal{R}_{lm}(x,y)$$

with

(3.15) 
$$\hat{u}_{lm} = \frac{1}{\gamma_{lm}} \iint_{\mathcal{T}} u(x, y) \mathcal{R}_{lm}(x, y) dx dy.$$

Letting  $v(\xi, \eta) = u(x, y)$ , we have  $v \in L^2_{\varpi}(\mathcal{Q})$  with  $\varpi = \frac{1-\eta}{8}$ , and

(3.16) 
$$u(x,y) = v(\xi,\eta) = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \hat{u}_{lm} \widetilde{\mathcal{R}}_{lm}(\xi,\eta),$$

where  $\{\widetilde{\mathcal{R}}_{lm}\}$  are defined in (3.12).

Define the space of rational functions in  ${\cal T}$ 

(3.17) 
$$\Re_{LM} := \operatorname{span} \left\{ \mathcal{R}_{lm} : 0 \le l \le L, \ 0 \le m \le M \right\},$$

which corresponds to the space of polynomials  $P_L(I_{\xi}) \times P_M(I_{\eta})$  in  $\mathcal{Q}$ . The projection operator  $\Pi_{LM} : L^2(\mathcal{T}) \to \Re_{LM}$  is defined by

(3.18) 
$$\iint_{\mathcal{T}} (\Pi_{LM} u - u) \, \Phi dx dy = 0, \quad \forall \Phi \in \Re_{LM},$$

and is given by

$$\Pi_{LM}u(x,y) = \sum_{l=0}^{L} \sum_{m=0}^{M} \hat{u}_{lm} \mathcal{R}_{lm}(x,y).$$

Define the Jacobi weight function in the triangle  $\mathcal{T}: \chi^{\alpha,\beta,\gamma} = \chi^{\alpha,\beta,\gamma}(x,y) := (1-x-y)^{\alpha}x^{\beta}y^{\gamma}$ , and introduce the Jacobi-weighted space in  $\mathcal{T}:$  for integers  $r, s \ge 0$ ,

(3.19) 
$$\mathcal{H}^{r,s}(\mathcal{T}) = \left\{ u \in L^2(\mathcal{T}) : \|u\|_{\mathcal{H}^{r,s}(\mathcal{T})} < \infty \right\},$$

where

$$\|u\|_{\mathcal{H}^{r,s}(\mathcal{T})} = \left(\|u\|_{L^{2}(\mathcal{T})}^{2} + \|\partial_{x}^{r}u\|_{L^{2}_{\chi^{r,r,0}}(\mathcal{T})}^{2} + \sum_{j=0}^{s} \left\|\partial_{y}^{j}(\partial_{y} - \partial_{x})^{s-j}u\|_{L^{2}_{\chi^{j,s-j,s}}(\mathcal{T})}^{2}\right)^{\frac{1}{2}}.$$

Observe that the weight functions  $\chi^{r,r,0}$  and  $\chi^{j,s-j,s}$  are uniformly bounded on  $\overline{\mathcal{T}}$ , and hence,  $H^r(\mathcal{T}) \subset \mathcal{H}^{r,r}(\mathcal{T})$ , where  $H^r(\mathcal{T})$  is the usual Sobolev space.

Remark 3.1. It is worthwhile to point out that  $\|\partial_x^r u\|_{L^2_{\chi^{r,r,0}}(\mathcal{T})}^2$  and  $\|\partial_y^j(\partial_y - \partial_x)^{s-j}u\|_{L^2_{\chi^{j,s-j,s}}(\mathcal{T})}^2$  are induced by self-adjoint differential operators  $(-1)^r \partial_x^r(\chi^{r,r,0}\partial_x^r)$  and  $(-1)^s \partial_y^j(\partial_y - \partial_x)^{s-j}(\chi^{j,s-j,s}\partial_y^j(\partial_y - \partial_x)^{s-j})$ , respectively. In fact, for any sufficiently smooth function u on  $\mathcal{T}$ ,

$$\partial_x^k \chi^{r,r,0} \partial_x^r u = \sum_{i=0}^k \sum_{j=0}^{k-i} (-1)^i \frac{k!}{(k-i-j)!} \binom{r}{i} \binom{r}{j} \chi^{r-i,r-j,0} \partial_x^{r+k-i-j} u,$$

### Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

and for any integers i, j < r,  $\chi^{r-i, r-j, 0}(0, y) = \chi^{r-i, r-j, 0}(1-y, y) = 0$ , thus, for any  $0 \le k \le r-1$ ,

$$\left(\partial_x^k \chi^{r,r,0} \partial_x^r u\right)(0,y) = 0, \qquad \left(\partial_x^k \chi^{r,r,0} \partial_x^r u\right)(1-y,y) = 0.$$

As a consequence, we derive, by successive integration by parts, that

$$\iint_{\mathcal{T}} |\partial_x^r u|^2 \chi^{r,r,0} dx dy = \int_0^1 dy \int_0^{1-y} (\partial_x^r u)^2 \chi^{r,r,0} dx$$
  
=  $-\int_0^1 dy \int_0^{1-y} (\partial_x^{r-1} u) \cdot \partial_x (\chi^{r,r,0} \partial_x^r u) dx$   
=  $\int_0^1 dy \int_0^{1-y} (\partial_x^{r-2} u) \cdot \partial_x^2 (\chi^{r,r,0} \partial_x^r u) dx$   
=  $\cdots = (-1)^r \int_0^1 dy \int_0^{1-y} u \partial_x^r (\chi^{r,r,0} \partial_x^r u) dx,$ 

which implies that

$$\left\|\partial_x^r u\right\|_{L^2_{\chi^{r,r,0}}(\mathcal{T})}^2 = \left(u, (-1)^r \partial_x^r \chi^{r,r,0} \partial_x^r u\right)_{\mathcal{T}}.$$

Similarly, we have

$$\left\|\partial_y^j(\partial_y - \partial_x)^{s-j}u\right\|_{L^2_{\chi^{j,s-j,s}}(\mathcal{T})}^2 = \left(u, (-1)^s \partial_y^j(\partial_y - \partial_x)^{s-j}\chi^{j,s-j,s}\partial_y^j(\partial_y - \partial_x)^{s-j}u\right)_{\mathcal{T}}.$$

The estimates of the  $L^2$  errors are stated in the following theorem.

THEOREM 3.1. For any  $u \in \mathcal{H}^{r,s}(\mathcal{T})$ , with integers  $r, s \geq 0$ , we have

$$\|\Pi_{LM}u - u\|_{L^{2}(\mathcal{T})} \lesssim L^{-r} \|\partial_{x}^{r}u\|_{L^{2}_{\chi^{r,r,0}}(\mathcal{T})} + M^{-s} \left(\sum_{j=0}^{s} \|\partial_{y}^{j}(\partial_{y} - \partial_{x})^{s-j}u\|_{L^{2}_{\chi^{j,s-j,s}}(\mathcal{T})}^{2}\right)^{\frac{1}{2}},$$

which implies that

(3.22) 
$$\|\Pi_{LM}u - u\|_{L^{2}(\mathcal{T})} \lesssim \left(L^{-r} + M^{-s}\right) \|u\|_{\mathcal{H}^{r,s}(\mathcal{T})}$$

In particular, if r = s, L = M, and  $\Pi_M := \Pi_{MM}$ , then for any  $u \in H^r(\mathcal{T})$  with integer  $r \ge 0$ ,

(3.23) 
$$\|\Pi_M u - u\|_{L^2(\mathcal{T})} \lesssim M^{-r} \|u\|_{H^r(\mathcal{T})}.$$

*Proof.* By the definition of the operator  $\Pi_{LM}$ ,

$$\|\Pi_{LM}u - u\|_{L^2(\mathcal{T})} \le \|\Phi - u\|_{L^2(\mathcal{T})}, \quad \forall \Phi \in \Re_{LM}.$$

Let  $v(\xi,\eta) = u(x,y)$  and  $\Psi(\xi,\eta) = \Phi(x,y)$  (the coordinates (x,y) and  $(\xi,\eta)$  are connected by the transformation (3.2)–(3.3)). By taking  $\Psi = \pi_{L,\xi}^{0,0} \pi_{M,\eta}^{1,0} v$  (the subscript  $\xi$  indicates that the operator  $\pi_L^{0,0}$  acts on the variable  $\xi$  and likewise for  $\eta$ ), a direct calculation, together with (3.7), leads to

$$\|\Pi_{LM}u - u\|_{\mathcal{T}}^{2} \leq \iint_{\mathcal{Q}} \left(\pi_{L,\xi}^{0,0} \pi_{M,\eta}^{1,0} v - v\right)^{2} \frac{1 - \eta}{8} d\xi d\eta$$
  
$$\lesssim \iint_{\mathcal{Q}} \left(\pi_{L,\xi}^{0,0} \left(\pi_{M,\eta}^{1,0} v - v\right)\right)^{2} (1 - \eta) d\xi d\eta$$
  
$$+ \iint_{\mathcal{Q}} \left(\pi_{L,\xi}^{0,0} v - v\right)^{2} (1 - \eta) d\xi d\eta.$$

By Lemma 2.2 with  $\alpha = \beta = 0, \mu = 0$ , and r = 0, we have

$$\iint_{\mathcal{Q}} \left( \pi_{L,\xi}^{0,0} \left( \pi_{M,\eta}^{1,0} v - v \right) \right)^2 (1-\eta) d\xi d\eta \lesssim \iint_{\mathcal{Q}} \left( \pi_{M,\eta}^{1,0} v - v \right)^2 (1-\eta) d\xi d\eta.$$

Using Lemma 2.2 again yields

$$(3.24)$$

$$\|\Pi_{LM}u - u\|_{\mathcal{T}}^{2} \lesssim \iint_{\mathcal{Q}} \left(\pi_{L,\xi}^{0,0}v - v\right)^{2} (1 - \eta) d\xi d\eta + \iint_{\mathcal{Q}} \left(\pi_{M,\eta}^{1,0}v - v\right)^{2} (1 - \eta) d\xi d\eta$$

$$\lesssim \iint_{\mathcal{Q}} \left(L^{-2r} |\partial_{\xi}^{r}v|^{2} \left(1 - \xi^{2}\right)^{r} + M^{-2s} |\partial_{\eta}^{s}v|^{2} \left(1 - \eta^{2}\right)^{s}\right) (1 - \eta) d\xi d\eta.$$

Now, we bound the norms of v in the right-hand side of (3.24) by the norms of u on  $\mathcal{T}$ . By (3.5)–(3.6) and a recursive calculation, we have that for integer  $k \geq 1$ ,

$$\begin{aligned} \partial_{\xi} v(\xi,\eta) &= \frac{1-\eta}{4} \partial_x u(x,y), \\ \partial_{\xi}^2 v(\xi,\eta) &= \partial_{\xi} \left( \frac{1-\eta}{4} \partial_x u(x,y) \right) = \frac{1-\eta}{4} \partial_{\xi} \partial_x u(x,y) = \left( \frac{1-\eta}{4} \right)^2 \partial_x^2 u(x,y), \\ &\vdots \end{aligned}$$

(3.25)

$$\begin{aligned} \partial_{\xi}^{k} v(\xi, \eta) &= \partial_{\xi} \left( \partial_{\xi}^{k-1} v(\xi, \eta) \right) = \partial_{\xi} \left( \left( \frac{1-\eta}{4} \right)^{k-1} \partial_{x}^{k-1} u(x, y) \right) \\ &= \left( \frac{1-\eta}{4} \right)^{k-1} \partial_{\xi} \partial_{x}^{k-1} u(x, y) = \left( \frac{1-\eta}{4} \right)^{k} \partial_{x}^{k} u(x, y), \end{aligned}$$

and

(3.26)

$$\begin{aligned} \partial_{\eta} v(\xi,\eta) &= \left(\frac{1}{2}\partial_{y} - \frac{1+\xi}{4}\partial_{x}\right)u(x,y) = \left(\frac{1-\xi}{4}\partial_{y} + \frac{1+\xi}{4}(\partial_{y} - \partial_{x})\right)u(x,y),\\ \partial_{\eta}^{2}v(\xi,\eta) &= \partial_{\eta}\left(\partial_{\eta}v(\xi,\eta)\right) = \partial_{\eta}\left(\frac{1-\xi}{4}\partial_{y} + \frac{1+\xi}{4}(\partial_{y} - \partial_{x})\right)u(x,y)\\ &= \left(\frac{1-\xi}{4}\partial_{\eta}\partial_{y} + \frac{1+\xi}{4}\partial_{\eta}(\partial_{y} - \partial_{x})\right)u(x,y)\\ &= \left(\frac{(1-\xi)^{2}}{16}\partial_{y}^{2} + \frac{(1-\xi)(1+\xi)}{8}\partial_{y}(\partial_{y} - \partial_{x}) + \frac{(1+\xi)^{2}}{16}(\partial_{y} - \partial_{x})^{2}\right)u(x,y),\\ &\vdots \end{aligned}$$

$$\begin{split} \partial_{\eta}^{k} v(\xi,\eta) &= \partial_{\eta} \left( \partial_{\eta}^{k-1} v(\xi,\eta) \right) \\ &= \partial_{\eta} \left( \sum_{j=0}^{k-1} \binom{k-1}{j} \left( \frac{1-\xi}{4} \right)^{j} \left( \frac{1+\xi}{4} \right)^{k-j-1} \partial_{y}^{j} (\partial_{y} - \partial_{x})^{k-j-1} u(x,y) \right) \\ &= \sum_{j=0}^{k-1} \binom{k-1}{j} \left( \frac{1-\xi}{4} \right)^{j} \left( \frac{1+\xi}{4} \right)^{k-j-1} \partial_{\eta} \partial_{y}^{j} (\partial_{y} - \partial_{x})^{k-j-1} u(x,y) \\ &= \sum_{j=0}^{k} \binom{k}{j} \left( \frac{1-\xi}{4} \right)^{j} \left( \frac{1+\xi}{4} \right)^{k-j} \partial_{y}^{j} (\partial_{y} - \partial_{x})^{k-j-1} u(x,y). \end{split}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Moreover, by (3.3),

(3.27)

$$\left(\frac{1-\eta}{4}\right)^{2r} \left(1-\xi^2\right)^r = x^r (1-x-y)^r,$$
$$\left(\frac{1-\xi}{4}\right)^{2j} \left(\frac{1+\xi}{4}\right)^{2s-2j} \left(1-\eta^2\right)^s = 2^{-s} (1-\xi)^j (1+\xi)^{s-j} \cdot (1-x-y)^j x^{s-j} y^s$$
$$\leq (1-x-y)^j x^{s-j} y^s, \quad \text{for } s \geq j \geq 0.$$

Thanks to (3.7) and (3.25)-(3.27),

$$\iint_{\mathcal{Q}} |\partial_{\xi}^{r} v|^{2} \left(1-\xi^{2}\right)^{r} (1-\eta) d\xi d\eta = 8 \iint_{\mathcal{T}} |\partial_{x}^{r} u|^{2} \left(\frac{1-\eta}{4}\right)^{2r} \left(1-\xi^{2}\right)^{r} dx dy$$
$$\lesssim \iint_{\mathcal{T}} |\partial_{x}^{r} u|^{2} x^{r} (1-x-y)^{r} dx dy,$$

and

$$\begin{split} \iint_{\mathcal{Q}} \left| \partial_{\eta}^{s} v \right|^{2} \left( 1 - \eta^{2} \right)^{s} (1 - \eta) d\xi d\eta \\ &= 8 \iint_{\mathcal{T}} \left| \sum_{j=0}^{s} \binom{s}{j} \left( \frac{1 - \xi}{4} \right)^{j} \left( \frac{1 + \xi}{4} \right)^{s-j} \partial_{y}^{j} (\partial_{y} - \partial_{x})^{s-j} u(x, y) \right|^{2} (1 - \eta^{2})^{s} dx dy \\ &\lesssim \sum_{j=0}^{s} \iint_{\mathcal{T}} \left| \partial_{y}^{j} (\partial_{y} - \partial_{x})^{s-j} u(x, y) \right|^{2} \left( \frac{1 - \xi}{4} \right)^{2j} \left( \frac{1 + \xi}{4} \right)^{2s-2j} \left( 1 - \eta^{2} \right)^{s} dx dy \\ &\lesssim \sum_{j=0}^{s} \iint_{\mathcal{T}} \left| \partial_{y}^{j} (\partial_{y} - \partial_{x})^{s-j} u(x, y) \right|^{2} (1 - x - y)^{j} x^{s-j} y^{s} dx dy. \end{split}$$

Inserting the above two estimates into (3.24) leads to

$$(3.28)$$

$$\|\Pi_{LM}u - u\|_{L^{2}(\mathcal{T})}^{2} \lesssim L^{-2r} \iint_{\mathcal{T}} |\partial_{x}^{r}u|^{2} (1 - x - y)^{r} x^{r} dx dy$$

$$+ M^{-2s} \sum_{j=0}^{s} \iint_{\mathcal{T}} |\partial_{y}^{j} (\partial_{y} - \partial_{x})^{s-j} u|^{2} (1 - x - y)^{j} x^{s-j} y^{s} dx dy,$$

which implies (3.21). By the definition (3.20), the estimate (3.22) follows from (3.21) immediately. Finally, (3.23) is a direct consequence of (3.22) due to the fact  $H^r(\mathcal{T}) \subset \mathcal{H}^{r,r}(\mathcal{T})$ .  $\Box$ 

**3.3. Error estimates for the**  $H_0^1$ **-orthogonal projection.** We now consider the  $H_0^1$ -orthogonal projection which is essential for the analysis of Galerkin approximation of elliptic equations.

We define the finite-dimensional approximation space

$$(3.29) X_{LM}^0 = \Re_{LM} \cap H_0^1(\mathcal{T})$$

It is clear that the transform (3.2) maps  $X_{LM}^0$  in  $\mathcal{T}$  to  $\mathcal{P}_L^0(I_{\xi}) \times \mathcal{P}_M^0(I_{\eta})$  in  $\mathcal{Q}$ .

To simplify the presentation, we only consider the case L = M and set  $X_M^0 := X_{MM}^0$ .

Define the  $H_0^1$ -orthogonal projection  $\Pi_M^{1,0}: H_0^1(\mathcal{T}) \to X_M^0$  by

(3.30) 
$$\left(\nabla\left(\Pi_M^{1,0}u-u\right), \ \nabla\Phi\right)_{\mathcal{T}} = 0, \quad \forall\Phi \in X_M^0.$$

We first point out that under the mapping (3.2)–(3.3), the space  $H^1(\mathcal{T})$  corresponds to the weighted space

(3.31) 
$$\widetilde{H}^{1}_{\varpi}(\mathcal{Q}) := \left\{ v \in L^{2}_{\varpi}(\mathcal{Q}) : \partial_{\xi} v \in L^{2}_{\varpi^{-1}}(\mathcal{Q}) \text{ and } \partial_{\eta} v \in L^{2}_{\varpi}(\mathcal{Q}) \right\},$$

with the weight  $\varpi = (1 - \eta)/8$  being the Jacobian in (3.7), and the norm

(3.32) 
$$\|v\|_{\widetilde{H}^{1}_{\varpi}(\mathcal{Q})} = \left( \|\partial_{\xi}v\|_{L^{2}_{\varpi^{-1}}(\mathcal{Q})}^{2} + \|\partial_{\eta}v\|_{L^{2}_{\varpi}(\mathcal{Q})}^{2} + \|v\|_{L^{2}_{\varpi}(\mathcal{Q})}^{2} \right)^{\frac{1}{2}}.$$

One verifies the equivalence by using (3.4) and (3.9)-(3.10):

$$(3.33) \|v\|_{\widetilde{H}^1_{\varpi}(\mathcal{Q})} \lesssim \|u\|_{H^1(\mathcal{T})} \lesssim \|v\|_{\widetilde{H}^1_{\varpi}(\mathcal{Q})}$$

We have the following results on error estimates of the  $H_0^1$ -orthogonal projection. THEOREM 3.2. For any  $u \in H_0^1(\mathcal{T}) \cap H^r(\mathcal{T})$  with  $r \ge 1$ ,

(3.34)

$$\begin{split} \left\| \Pi_{M}^{1,0} u - u \right\|_{H^{\mu}(\mathcal{T})} &\lesssim M^{\mu-r} \bigg( \sum_{(i,j) \in \Lambda_{r}} \left\| \partial_{x}^{i} \partial_{y}^{j} (\partial_{y} - \partial_{x})^{r-i-j} u \right\|_{L^{2}_{\chi^{i+j-1,r-j-1,r-i-1}}(\mathcal{T})} \\ &+ \left\| \partial_{x}^{r} u \right\|_{L^{2}_{\chi^{r-1,r-1,0}}(\mathcal{T})} + \left\| \partial_{y}^{r} u \right\|_{L^{2}_{\chi^{\tilde{r},0,r-1}}(\mathcal{T})} + \left\| (\partial_{y} - \partial_{x})^{r} u \right\|_{L^{2}_{\chi^{0,\tilde{r},r-1}}(\mathcal{T})} \bigg), \quad \mu = 0, 1, \end{split}$$

where the weight function  $\chi^{\alpha,\beta,\gamma} = (1-x-y)^{\alpha}x^{\beta}y^{\gamma}$ ,  $\tilde{r} = \max\{0,r-2\}$  and the index set  $\Lambda_r = \{(i,j) \in \mathbb{Z}^2 : 0 \leq i,j; i+j \leq r\} \setminus \{(0,0), (0,r), (r,0)\}$ . In particular, the above estimate implies

(3.35) 
$$\left\| \Pi_{M}^{1,0} u - u \right\|_{H^{\mu}(\mathcal{T})} \lesssim M^{\mu-r} \| u \|_{H^{r}(\mathcal{T})}, \quad \mu = 0, 1.$$

*Proof.* We first prove (3.34) with  $\mu = 1$ . By the projection theorem,

(3.36) 
$$\left\| \nabla \left( \Pi_M^{1,0} u - u \right) \right\|_{L^2(\mathcal{T})} \le \left\| \nabla (\Phi - u) \right\|_{L^2(\mathcal{T})}, \quad \forall \Phi \in X_M^0.$$

Taking  $\Phi = 0$ , we get that

$$\|\nabla \left(\Pi_{M}^{1,0}u - u\right)\|_{L^{2}(\mathcal{T})} \leq \|\nabla u\|_{L^{2}(\mathcal{T})} \leq \|\partial_{x}u\|_{L^{2}(\mathcal{T})} + \|\partial_{y}u\|_{L^{2}(\mathcal{T})} + \|(\partial_{y} - \partial_{x})u\|_{L^{2}(\mathcal{T})}$$

which leads to (3.34) with  $\mu = r = 1$ .

Now let  $r \geq 2$ . For  $u \in H_0^1(\mathcal{T})$  and  $\Phi \in X_M^0$ , we set  $v(\xi, \eta) = u(x, y)$  and  $\Psi(\xi, \eta) = \Phi(x, y)$ , where the coordinates  $(\xi, \eta)$  and (x, y) are associated with the mapping (3.2)–(3.3). Then, we have  $v \in H_{0,\varpi}^1(\mathcal{Q}) \cap \tilde{H}_{\varpi}^1(\mathcal{Q})$  and  $\Psi \in \mathcal{P}_M^0(I_{\xi}) \times \mathcal{P}_M^0(I_{\eta})$  (where  $I_{\xi} = I_{\eta} = (-1, 1)$ ). By (3.7), (3.9), and (3.36),

(3.38)

$$\left\|\nabla\left(\Pi_{M}^{1,0}u-u\right)\right\|_{L^{2}(\mathcal{T})} \lesssim \left\|\partial_{\xi}(\Psi-v)\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} + \left\|\partial_{\eta}(\Psi-v)\right\|_{L^{2}_{\omega^{1,0}}(I_{\eta};L^{2}(I_{\xi}))}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Let  $\pi_M^{-1,-1}$  be the  $L^2_{\omega^{-1,-1}}$ -orthogonal projection operator as defined in section 2. Using Lemma 2.2 with  $\alpha = \beta = -1$  and  $r = \mu = 1$  leads to

$$(3.39) \\ \left\| \partial_{\xi} \pi_{M,\xi}^{-1,-1} \left( \pi_{M,\eta}^{-1,-1} - \mathbb{I} \right) v \right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} \lesssim \left\| \partial_{\xi} \left( \pi_{M,\eta}^{-1,-1} - \mathbb{I} \right) v \right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))},$$

and

$$(3.40) \quad \left\| \partial_{\eta} \pi_{M,\eta}^{-1,-1} \left( \pi_{M,\xi}^{-1,-1} - \mathbb{I} \right) v \right\|_{L^{2}(I_{\eta};L^{2}(I_{\xi}))} \lesssim \left\| \partial_{\eta} \left( \pi_{M,\xi}^{-1,-1} - \mathbb{I} \right) v \right\|_{L^{2}(I_{\eta};L^{2}(I_{\xi}))}.$$

Hence, taking

(3.41) 
$$\Psi(\xi,\eta) = \pi_{M,\eta}^{-1,-1} \pi_{M,\xi}^{-1,-1} v = \pi_{M,\xi}^{-1,-1} \pi_{M,\eta}^{-1,-1} v \in \mathcal{P}_M^0(I_\xi) \times \mathcal{P}_M^0(I_\eta)$$

in (3.38), we derive from the triangular inequality, (3.39), and Lemma 2.2 that

$$\begin{split} \|\partial_{\xi}(\Psi-v)\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} &= \left\|\partial_{\xi}\left(\pi^{-1,-1}_{M,\eta}\pi^{-1,-1}_{M,\xi}v-v\right)\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} \\ &\leq \left\|\partial_{\xi}\pi^{-1,-1}_{M,\xi}\left(\pi^{-1,-1}_{M,\eta}-\mathbb{I}\right)v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} + \left\|\partial_{\xi}\left(\pi^{-1,-1}_{M,\xi}-\mathbb{I}\right)v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} \\ &\stackrel{(3.39)}{\lesssim} \left\|\partial_{\xi}\left(\pi^{-1,-1}_{M,\eta}-\mathbb{I}\right)v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} + \left\|\partial_{\xi}\left(\pi^{-1,-1}_{M,\xi}-\mathbb{I}\right)v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} \\ &\lesssim \left\|\left(\pi^{-1,-1}_{M,\eta}-\mathbb{I}\right)\partial_{\xi}v\right\|_{L^{2}_{\omega^{-1,-1}}(I_{\eta};L^{2}(I_{\xi}))} + \left\|\partial_{\xi}\left(\pi^{-1,-1}_{M,\xi}-\mathbb{I}\right)v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} \\ &\lesssim \left\|\left(\pi^{-1,-1}_{M,\eta}-\mathbb{I}\right)\partial_{\xi}v\right\|_{L^{2}_{\omega^{-1,-1}}(I_{\eta};L^{2}(I_{\xi}))} + \left\|\partial_{\xi}\left(\pi^{-1,-1}_{M,\xi}-\mathbb{I}\right)v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}(I_{\xi}))} \\ &+ M^{1-r}\left\|\partial_{\xi}^{r}v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}_{\omega^{r-1,r-1}}(I_{\xi}))\cdot \end{split}$$

To further simplify the first term in the right-hand side of the above estimate, we define

(3.42) 
$$v^* = (1-\eta)^{-1} \partial_{\xi} v \stackrel{(3.9)}{=} \frac{1}{4} \partial_x u \implies \partial_{\xi} v = (1-\eta) v^*.$$

Therefore,

(3.43) 
$$\partial_{\eta}^{r-1}\partial_{\xi}v = \partial_{\eta}^{r-1}\left((1-\eta)v^*\right) = (1-\eta)\partial_{\eta}^{r-1}v^* - (r-1)\partial_{\eta}^{r-2}v^*.$$

Hence, using the above identity, (2.6) and (2.7), gives

$$\begin{split} & \left\| \left( \pi_{M-r+1,\eta}^{r-2,r-2} - \mathbb{I} \right) \partial_{\eta}^{r-1} \partial_{\xi} v \right\|_{L^{2}_{\omega^{r-2,r-2}}(I_{\eta};L^{2}(I_{\xi}))} \\ & \lesssim \left\| \left( \pi_{M-r+1,\eta}^{r-2,r-2} - \mathbb{I} \right) (1-\eta) \partial_{\eta}^{r-1} v^{*} \right\|_{L^{2}_{\omega^{r-2,r-2}}(I_{\eta};L^{2}(I_{\xi}))} \\ & + \left\| \left( \pi_{M-r+1,\eta}^{r-2,r-2} - \mathbb{I} \right) \partial_{\eta}^{r-2} v^{*} \right\|_{L^{2}_{\omega^{r-2,r-2}}(I_{\eta};L^{2}(I_{\xi}))} \\ & \lesssim \left\| (1-\eta) \partial_{\eta}^{r-1} v^{*} \right\|_{L^{2}_{\omega^{r-2,r-2}}(I_{\eta};L^{2}(I_{\xi}))} + \left\| \partial_{\eta}^{r-1} v^{*} \right\|_{L^{2}_{\omega^{r,r}}(I_{\eta};L^{2}(I_{\xi}))} \\ & \lesssim \left\| \partial_{\eta}^{r-1} v^{*} \right\|_{L^{2}_{\omega^{r,r-2}}(I_{\eta};L^{2}(I_{\xi}))} \cdot \end{split}$$

Therefore, we have the upper bound for the first term in the right-hand side of (3.38):

We now turn to the second term in the right-hand side of (3.38) with  $\Psi$  given by (3.41). Similarly, by the triangular inequality, (3.40) and Lemma 2.2, we derive that

$$\begin{split} \|\partial_{\eta}(\Psi - v)\|_{L^{2}_{\omega^{1,0}}(I_{\eta};L^{2}(I_{\xi}))} &\lesssim \|\partial_{\eta}(\Psi - v)\|_{L^{2}(I_{\eta};L^{2}(I_{\xi}))} \\ \stackrel{(3.41)}{\lesssim} \|\partial_{\eta}\left(\pi^{-1,-1}_{M,\eta} - \mathbb{I}\right)v\|_{L^{2}(I_{\eta};L^{2}(I_{\xi}))} + \|\partial_{\eta}\pi^{-1,-1}_{M,\eta}\left(\pi^{-1,-1}_{M,\xi} - \mathbb{I}\right)v\|_{L^{2}(I_{\eta};L^{2}(I_{\xi}))} \\ \stackrel{(3.40)}{\lesssim} \|\partial_{\eta}\left(\pi^{-1,-1}_{M,\eta} - \mathbb{I}\right)v\|_{L^{2}(I_{\eta};L^{2}(I_{\xi}))} + \|\left(\pi^{-1,-1}_{M,\xi} - \mathbb{I}\right)\partial_{\eta}v\|_{L^{2}(I_{\eta};L^{2}(I_{\xi}))} \\ \stackrel{(2.10)}{\lesssim} M^{1-r} \|\partial_{\eta}^{r}v\|_{L^{2}_{\omega^{r-1},r-1}(I_{\eta};L^{2}(I_{\xi}))} \\ + M^{1-r} \left\|\left(\pi^{r-2,r-2}_{M-r+1,\xi} - \mathbb{I}\right)\partial_{\xi}^{r-1}\partial_{\eta}v\right\|_{L^{2}\left(I_{\eta};L^{2}_{\omega^{r-2},r-2}(I_{\xi})\right)}. \end{split}$$

To deal with the second term in the right-hand side of the above estimate, we use (3.9), (3.10), and (3.25) to derive that

$$\begin{aligned} (3.45) \\ \partial_{\xi}^{r-1} \partial_{\eta} v \stackrel{(3.10)}{=} \partial_{\xi}^{r-1} \left( \frac{1}{2} \partial_{y} u - \frac{1+\xi}{4} \partial_{x} u \right) \\ &= \frac{1}{2} \partial_{\xi}^{r-1} \partial_{y} u - \frac{1+\xi}{4} \partial_{\xi}^{r-1} \partial_{x} u - \frac{r-1}{4} \partial_{\xi}^{r-2} \partial_{x} u \\ \stackrel{(3.9)}{=} \left[ \frac{1-\xi}{4} \partial_{\xi}^{r-1} \partial_{y} u + \frac{1+\xi}{4} \partial_{\xi}^{r-1} (\partial_{y} - \partial_{x}) u \right] - \frac{r-1}{1-\eta} \partial_{\xi}^{r-1} v \\ \stackrel{(3.25)}{=} \left[ \frac{(1-\xi)(1-\eta)^{r-1}}{4^{r}} \partial_{x}^{r-1} \partial_{y} u + \frac{(1+\xi)(1-\eta)^{r-1}}{4^{r}} \partial_{x}^{r-1} (\partial_{y} - \partial_{x}) u \right] \\ &- \frac{r-1}{1-\eta} \partial_{\xi}^{r-1} v. \end{aligned}$$

For simplicity, we define

(3.46)  
$$v^{\star} = \partial_{\xi}^{r-1} \partial_{\eta} v + \frac{r-1}{1-\eta} \partial_{\xi}^{r-1} v$$
$$= \frac{(1-\xi)(1-\eta)^{r-1}}{4^{r}} \partial_{x}^{r-1} \partial_{y} u + \frac{(1+\xi)(1-\eta)^{r-1}}{4^{r}} \partial_{x}^{r-1} (\partial_{y} - \partial_{x}) u$$

Hence, by (2.6) and (2.7), we have

$$\begin{split} \left\| \left( \pi_{M-r+1,\xi}^{r-2,r-2} - \mathbb{I} \right) \partial_{\xi}^{r-1} \partial_{\eta} v \right\|_{L^{2} \left( I_{\eta}; L_{\omega}^{2} r^{-2,r-2}(I_{\xi}) \right)} \leq \left\| \left( \pi_{M-r+1,\xi}^{r-2,r-2} - \mathbb{I} \right) v^{\star} \right\|_{L^{2} \left( I_{\eta}; L_{\omega}^{2} r^{-2,r-2}(I_{\xi}) \right)} \\ &+ (r-1) \left\| (1-\eta)^{-1} \left( \pi_{M-r+1,\xi}^{r-2,r-2} - \mathbb{I} \right) \partial_{\xi}^{r-1} v \right\|_{L^{2} \left( I_{\eta}; L_{\omega}^{2} r^{-2,r-2}(I_{\xi}) \right)} \\ & \stackrel{(2.6)}{\underset{(2.7)}{\overset{($$

# Copyright ${\tt O}$ by SIAM. Unauthorized reproduction of this article is prohibited.

Therefore, we obtain the upper bound of the second term in the right-hand side of (3.38):

$$(3.47) \|\partial_{\eta}(\Psi - v)\|_{L^{2}_{\omega^{1,0}}(I_{\eta}; L^{2}(I_{\xi}))} \lesssim M^{1-r} \Big( \|\partial_{\eta}^{r}v\|_{L^{2}_{\omega^{r-1,r-1}}(I_{\eta}; L^{2}(I_{\xi}))} + \|v^{\star}\|_{L^{2}(I_{\eta}; L^{2}_{\omega^{r-2,r-2}}(I_{\xi}))} + \|\partial_{\xi}^{r}v\|_{L^{2}_{\omega^{-2,0}}(I_{\eta}; L^{2}_{\omega^{r,r}}(I_{\xi}))} \Big).$$

A combination of (3.38), (3.44), and (3.47) yields

$$\begin{aligned} \|\nabla\left(\Pi_{M}^{1,0}u-u\right)\|_{L^{2}(\mathcal{T})} &\lesssim M^{1-r}\left(\left\|\partial_{\eta}^{r-1}v^{*}\right\|_{L^{2}_{\omega^{r,r-2}}(I_{\eta};L^{2}(I_{\xi}))} \\ &+ \left\|\partial_{\xi}^{r}v\right\|_{L^{2}_{\omega^{-1,0}}(I_{\eta};L^{2}_{\omega^{r-1,r-1}}(I_{\xi}))} + \left\|\partial_{\xi}^{r}v\right\|_{L^{2}_{\omega^{-2,0}}\left(I_{\eta};L^{2}_{\omega^{r,r}}(I_{\xi})\right)} \\ &+ \left\|v^{*}\right\|_{L^{2}\left(I_{\eta};L^{2}_{\omega^{r-2,r-2}}(I_{\xi})\right)} + \left\|\partial_{\eta}^{r}v\right\|_{L^{2}_{\omega^{r-1,r-1}}(I_{\eta};L^{2}(I_{\xi}))}\right). \end{aligned}$$

We now bound the terms in the right-hand side by the norms of u on  $\mathcal{T}$ . Using (3.4), (3.7), and (3.25)–(3.26), we find

$$\begin{split} \left\|\partial_{\eta}^{r-1}v^{*}\right\|_{L^{2}_{\omega^{r,r-2}}(I_{\eta};L^{2}(I_{\xi}))}^{2} &= \iint_{\mathcal{Q}} \left|\partial_{\eta}^{r-1}v^{*}\right|^{2} (1-\eta)^{r} (1+\eta)^{r-2} d\xi d\eta \\ &\lesssim \sum_{j=0}^{r-1} \iint_{\mathcal{T}} \left|\partial_{x}\partial_{y}^{j}(\partial_{y}-\partial_{x})^{r-j-1}u\right|^{2} \chi^{j,r-j-1,r-2}(x,y) \,\omega^{j,r-j-1}(\xi) dx dy \\ &\lesssim \sum_{j=0}^{r-1} \iint_{\mathcal{T}} \left|\partial_{x}\partial_{y}^{j}(\partial_{y}-\partial_{x})^{r-j-1}u\right|^{2} \chi^{j,r-j-1,r-2} dx dy, \\ &\left\|\partial_{\xi}^{r}v\right\|_{L^{2}_{\omega^{-1,0}}\left(I_{\eta};L^{2}_{\omega^{r-1,r-1}(I_{\xi})}\right) = \iint_{\mathcal{Q}} \left|\partial_{\xi}^{r}v\right|^{2} (1-\eta)^{-1} (1-\xi^{2})^{r-1} d\xi d\eta \\ &\lesssim \iint_{\mathcal{T}} \left|\partial_{x}^{r}u\right|^{2} \chi^{r-1,r-1,0}(x,y) dx dy, \\ &\partial_{\xi}^{r}v\right\|_{L^{2}_{-2,0}\left(I_{\eta};L^{2}_{\omega,r,r}(I_{\xi})\right)} = \iint_{\mathcal{Q}} \left|\partial_{\xi}^{r}v\right|^{2} (1-\eta)^{-2} \left(1-\xi^{2}\right)^{r} d\xi d\eta \end{split}$$

$$\begin{split} \left\| \partial_{\xi}^{r} v \right\|_{L^{2}_{\omega^{-2,0}}\left(I_{\eta}; L^{2}_{\omega^{r,r}}(I_{\xi})\right)}^{2} &= \iint_{\mathcal{Q}} \left| \partial_{\xi}^{r} v \right|^{2} (1-\eta)^{-2} \left(1-\xi^{2}\right)^{r} d\xi d\eta \\ &\lesssim \iint_{\mathcal{Q}} \left| \partial_{x}^{r-1} \left( \partial_{y} u - \left( \partial_{y} - \partial_{x} \right) u \right) \right|^{2} (1-\eta)^{2r-2} \left(1-\xi^{2}\right)^{r} d\xi d\eta \\ &\lesssim \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} \partial_{y} u \right|^{2} \chi^{r-1,r-2,0}(x,y) \, \omega^{1,2}(\xi) dx dy \\ &+ \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} \left( \partial_{y} - \partial_{x} \right) u \right|^{2} \chi^{r-2,r-1,0}(x,y) \, \omega^{2,1}(\xi) dx dy \\ &\lesssim \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} \partial_{y} u \right|^{2} \chi^{r-1,r-2,0} dx dy + \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} \left( \partial_{y} - \partial_{x} \right) u \right|^{2} \chi^{r-2,r-1,0} dx dy, \end{split}$$

and

$$\begin{split} \|v^{\star}\|_{L^{2}(I_{\eta};L^{2}_{\omega^{r-2},r-2}(I_{\xi}))} \\ \stackrel{(3.46)}{=} \frac{1}{8^{r}} \iint_{\mathcal{Q}} \left| (1-\xi) \, \partial_{x}^{r-1} \partial_{y} u + (1+\xi) \, \partial_{x}^{r-1} (\partial_{y} - \partial_{x}) u \right|^{2} (1-\eta)^{2r-3} \left( 1-\xi^{2} \right)^{r-2} dx dy \\ \lesssim \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} \partial_{y} u \right|^{2} \chi^{r-1,r-2,0} (x,y) \, \omega^{1,0} (\xi) dx dy \\ &+ \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} \partial_{y} u \right|^{2} \chi^{r-2,r-1,0} (x,y) \, \omega^{0,1} (\xi) dx dy \\ \lesssim \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} \partial_{y} u \right|^{2} \chi^{r-1,r-2,0} dx dy + \iint_{\mathcal{T}} \left| \partial_{x}^{r-1} (\partial_{y} - \partial_{x}) u \right|^{2} \chi^{r-2,r-1,0} dx dy, \\ \left\| \partial_{\eta}^{r} v \right\|_{L^{2}_{\omega^{r-1},r-1}(I_{\eta};L^{2}(I_{\xi}))}^{2} = \iint_{\mathcal{Q}} \left| \partial_{\eta}^{r} v \right|^{2} (1-\eta^{2})^{r-1} \, d\xi d\eta \\ \lesssim \sum_{j=1}^{r-1} \iint_{\mathcal{T}} \left| \partial_{y}^{j} (\partial_{y} - \partial_{x})^{r-j} u \right|^{2} \chi^{j-1,r-j-1,r-1} (x,y) \, \omega^{j+1,r+j+1} (\xi) dx dy \\ &+ \iint_{\mathcal{T}} \left| \partial_{y}^{r} u \right|^{2} \chi^{r-2,0,r-1} (x,y) (1-\xi) \, \omega^{r+2,0} (\xi) dx dy \\ &+ \iint_{\mathcal{T}} \left| \partial_{y}^{j} (\partial_{y} - \partial_{x})^{r-j} u \right|^{2} \chi^{j-1,r-j-1,r-1} dx dy \\ &\leq \sum_{j=1}^{r-1} \iint_{\mathcal{T}} \left| \partial_{y}^{j} (\partial_{y} - \partial_{x})^{r-j} u \right|^{2} \chi^{j-1,r-j-1,r-1} dx dy \\ &+ \iint_{\mathcal{T}} \left| \partial_{y}^{r} u \right|^{2} \chi^{r-2,0,r-1} dx dy + \iint_{\mathcal{T}} \left| (\partial_{y} - \partial_{x})^{r} u \right|^{2} \chi^{0,r-2,r-1} dx dy. \end{split}$$

We conclude (3.34) from (3.38) and the above inequalities. This ends the proof of (3.34) with  $\mu = 1$ .

The case with  $\mu = 0$  can be proved by using a duality argument. Let  $e_M = \prod_M^{1,0} u - u$  and consider the auxiliary problem:

(3.49) Find  $w \in H_0^1(\mathcal{T})$  such that  $(\nabla \phi, \nabla w)_{\mathcal{T}} = (\phi, e_M)_{\mathcal{T}}, \quad \forall \phi \in H_0^1(\mathcal{T}),$ 

which admits a unique solution  $w \in H^1_0(\mathcal{T})$  with the regularity

$$(3.50) ||w||_{H^2(\mathcal{T})} \lesssim ||e_M||_{\mathcal{T}}$$

Taking  $\phi = e_M$  in (3.49), we have from (3.30) and (3.35) with  $\mu = 1$  that

$$\begin{aligned} \|e_M\|_{\mathcal{T}}^2 &= (\nabla e_M, \nabla w) = \left(\nabla e_M, \nabla \left(w - \Pi_M^{1,0} w\right)\right) \le \|\nabla e_M\|_{\mathcal{T}} \left\|\nabla \left(w - \Pi_M^{1,0} w\right)\right\|_{\mathcal{T}} \\ &\lesssim \|\nabla e_M\|_{\mathcal{T}} M^{-1} \|w\|_{H^2(\mathcal{T})} \lesssim M^{-1} \|\nabla e_M\|_{\mathcal{T}} \|e_M\|_{\mathcal{T}}, \end{aligned}$$

which implies  $||e_M||_{\mathcal{T}} \leq M^{-1} ||\nabla e_M||_{\mathcal{T}}$ . By using (3.34) with  $\mu = 1$ , we finally get the estimate (3.34) with  $\mu = 0$ .  $\Box$ 

**3.4. Error estimates for the interpolation.** Let  $\{\xi_l^L\}_{l=0}^M$  and  $\{\eta_l^R\}_{l=0}^M$  be the LGL and JGR points defined in section 2, respectively. The interpolation grid in  $\mathcal{T}$  is given by

(3.51) 
$$x_{lm} = \frac{1}{4} \left( 1 + \xi_l^L \right) \left( 1 - \eta_m^R \right), \qquad y_{lm} = \frac{1}{2} \left( 1 + \eta_m^R \right), \qquad 0 \le l, m \le M.$$

We now define the rational interpolation operator. For any  $u \in C(\overline{\mathcal{T}})$ , the interpolant  $\mathcal{I}_M u \in \Re_{MM}$  such that

(3.52) 
$$(\mathcal{I}_M u)(x_{lm}, y_{lm}) = u(x_{lm}, y_{lm}), \qquad 0 \le l, m \le M.$$

Let  $v(\xi, \eta) = u(x, y)$ . One verifies readily that

$$(\mathcal{I}_M u)(x,y) = \left(\mathcal{I}_{M,\xi}^L \mathcal{I}_{M,\eta}^R v\right)(\xi,\eta) = \left(\mathcal{I}_{M,\eta}^R \mathcal{I}_{M,\xi}^L v\right)(\xi,\eta),$$

where  $\mathcal{I}_{M,\xi}^L$  and  $\mathcal{I}_{M,\eta}^R$  are the Legendre–Gauss–Lobatto interpolation operator in  $\xi$ and the Jacobi–Gauss–Radau interpolation operator in  $\eta$ , respectively (cf. section 2).

THEOREM 3.3. For any  $u \in H^r(\mathcal{T})$  with integers  $r \geq 2$ , we have

$$\begin{aligned} \|\mathcal{I}_{M}u - u\|_{L^{2}(\mathcal{T})} &\lesssim M^{-r} \Big( \sum_{j=0}^{r} \left\| \partial_{y}^{j} (\partial_{y} - \partial_{x})^{r-j} u \right\|_{L^{2}_{\chi^{j,r-j,r}}(\mathcal{T})} + \left\| \partial_{x}^{r} u \right\|_{L^{2}_{\chi^{r-1,r-1,0}}(\mathcal{T})} \\ &+ \left\| \partial_{x}^{r-1} \partial_{y} u \right\|_{L^{2}_{\chi^{r,r-2,1}}(\mathcal{T})} + \left\| \partial_{x}^{r-1} (\partial_{y} - \partial_{x}) u \right\|_{L^{2}_{\chi^{r-2,r,1}}(\mathcal{T})} \Big). \end{aligned}$$

In particular, the above estimate implies

(3.54) 
$$\|\mathcal{I}_M u - u\|_{L^2(\mathcal{T})} \lesssim M^{-r} \|u\|_{H^r(\mathcal{T})}.$$

4. Application to a Poisson-type equation. Let  $\Omega$  be an open bounded domain, and consider the Poisson-type equation:

(4.1) 
$$-\Delta u + \gamma u = f, \quad \text{in } \mathcal{T}; \quad u|_{\partial\Omega} = 0, \quad \gamma \ge 0.$$

The variational formulation of (4.1) is to find  $u \in H^1_0(\Omega)$  such that

(4.2) 
$$a(u,v) = (\nabla u, \nabla v)_{\Omega} + \gamma(u,v)_{\Omega} = (f,v)_{\Omega}, \quad \forall v \in H^1_0(\Omega),$$

which, thanks to the Lax-Milgram lemma, admits a unique solution satisfying

(4.3) 
$$\|\nabla u\|_{\Omega} + \sqrt{\gamma} \|u\|_{\Omega} \lesssim \|f\|_{\Omega}.$$

**4.1. Mono-domain case.** Let  $\Omega = \mathcal{T}$ . The rational spectral-Galerkin approximation to (4.2) is: find  $u_M \in X_M^0$  such that

(4.4) 
$$a(u_M,\varphi) = (\mathcal{I}_M f,\varphi)_{\mathcal{T}}, \quad \forall \varphi \in X_M^0,$$

which has a unique solution satisfying (4.2) with  $u_M$  and  $\mathcal{I}_M f$  in place of u and f, respectively.

**4.1.1. Error estimates.** Thanks to the approximation result in Theorems 3.2 and 3.3, we immediately derive the following convergence result using a standard procedure.

THEOREM 4.1. Let u and  $u_M$  be the solutions of (4.2) and (4.4), respectively. If  $u \in H_0^1(\mathcal{T}) \cap H^r(\mathcal{T})$  and  $f \in H^s(\mathcal{T})$  with  $r \ge 1$  and  $s \ge 2$ , then we have

$$\begin{aligned} \|u_{M} - u\|_{H^{\mu}(\mathcal{T})} &\lesssim M^{\mu - r} \Biggl( \sum_{(i,j) \in \Lambda_{r}} \|\partial_{x}^{i} \partial_{y}^{j} (\partial_{y} - \partial_{x})^{r - i - j} u\|_{L^{2}_{\chi^{i} + j - 1, r - j - 1, r - i - 1}(\mathcal{T})} \\ &+ \|\partial_{x}^{r} u\|_{L^{2}_{\chi^{r - 1, r - 1, 0}(\mathcal{T})}} + \|\partial_{y}^{r} u\|_{L^{2}_{\chi^{\tilde{r}, 0, r - 1}(\mathcal{T})}} \\ &+ \|(\partial_{y} - \partial_{x})^{r} u\|_{L^{2}_{\chi^{0, \tilde{r}, r - 1}(\mathcal{T})}} \Biggr) \\ &+ M^{-s} \Biggl( \sum_{j = 0}^{s} \|\partial_{y}^{j} (\partial_{y} - \partial_{x})^{s - j} f\|_{L^{2}_{\chi^{j, r - j, r}(\mathcal{T})}} + \|\partial_{x}^{s} f\|_{L^{2}_{\chi^{s - 1, s - 1, 0}(\mathcal{T})}} \\ &+ \|\partial_{x}^{s - 1} \partial_{y} f\|_{L^{2}_{\chi^{s, s - 2, 1}(\mathcal{T})}} + \|\partial_{x}^{s - 1} (\partial_{y} - \partial_{x}) f\|_{L^{2}_{\chi^{s - 2, s, 1}(\mathcal{T})}} \Biggr), \\ \mu = 0, 1, \end{aligned}$$

which implies that

(4.6) 
$$||u_M - u||_{H^{\mu}(\mathcal{T})} \lesssim M^{\mu - r} ||u||_{H^r(\mathcal{T})} + M^{-s} ||f||_{H^s(\mathcal{T})}, \quad \mu = 0, 1.$$

*Proof.* One derives immediately from (4.2) and (4.4) that

(4.7) 
$$\|u_M - u\|_{H^1(\mathcal{T})} \lesssim \inf_{v_M \in X_M^0} \|v_M - u\|_{H^1(\mathcal{T})} + \|\mathcal{I}_M f - f\|_{L^2(\mathcal{T})}.$$

Then, applying Theorems 3.2 and 3.3 to the above, we obtain immediately (4.5) with  $\mu = 1$ . The result for  $\mu = 0$  can then be derived by using a standard duality argument as in the proof of Theorem 3.2.

Remark 4.1. Alternative to (4.4), we can also consider the following rational spectral-Galerkin approximation with numerical integration: Find  $u_M \in X_M^0$  such that

(4.8) 
$$a_M(u_M,\varphi) := \gamma(u_M,\varphi)_{M,\mathcal{T}} + (\nabla u_M,\nabla\varphi)_{M,\mathcal{T}} = (f,\varphi)_{M,\mathcal{T}}, \quad \forall \varphi \in X^0_M,$$

where

(4.9) 
$$(u,v)_{M,\mathcal{T}} := \sum_{i=0}^{M} \sum_{j=0}^{M} u(x_{ij}, y_{ij}) v(x_{ij}, y_{ij}) \frac{1}{8} \omega_i^L \omega_j^R, \quad \forall u, v \in C(\bar{\mathcal{T}}),$$

with  $\{x_{ij}, y_{ij}\}_{i,j=0}^{M}$  being the collocation points defined in (3.51),  $\{\omega_i^L\}_{i=0}^{M}$  being the weights for the LGL nodes  $\{\xi_i^L\}_{i=0}^{M}$ , and  $\{\omega_j^R\}_{j=0}^{M}$  being the weights for JGR nodes  $\{\eta_j^R\}_{j=0}^{M}$  (cf. section 2). Then, it can be easily shown that the results of Theorem 4.1 hold for (4.8) as well.

Remark 4.2. The simple forms of the error estimates (3.22), (3.35), (3.54), and (4.6) are obtained at the expense of more precise error bounds. Consider, for instance, that (4.1) has the solution

(4.10) 
$$u(x,y) = x^{\alpha} y^{\beta} (1-x-y), \quad \forall \ (x,y) \in \mathcal{T}, \ \alpha, \beta > 0,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

1636

(4.5)

with  $\alpha$  and/or  $\beta$  being nonintegers. By using the error bound in (4.6) and ignoring the interpolation error for f, one can only derive that for any  $\epsilon > 0$ ,

$$(4.11) ||u_M - u||_{\mu,\mathcal{T}} \lesssim M^{\mu - r + \epsilon}$$

with

(4.12) 
$$r = \begin{cases} \min\left\{\alpha + \frac{1}{2}, \beta + \frac{1}{2}\right\}, & \text{if } \alpha, \beta \notin \mathbb{N}, \\ \alpha + \frac{1}{2}, & \text{if } \alpha \notin \mathbb{N}, \beta \in \mathbb{N}, \\ \beta + \frac{1}{2}, & \text{if } \alpha \in \mathbb{N}, \beta \notin \mathbb{N}. \end{cases}$$

However, by directly calculating all the terms in (4.5) and ignoring the interpolation error for f, we can show that

$$(4.13) ||u_M - u||_{\mu,\mathcal{T}} \lesssim M^{\mu - 2r + \epsilon}.$$

This estimate is in agreement with our numerical experiments presented below.

**4.1.2. Modal basis functions.** In order to treat nonhomogeneous boundary conditions and/or to enforce continuity across the interfaces in a triangular spectral-element method, we need to construct basis functions for

$$(4.14) X_{LM} = \Re_{LM} \cap H^1(\mathcal{T})$$

which, through the transformation (3.2), corresponds to

(4.15)

$$\widetilde{X}_{LM} := \left( P_L(I_{\xi}) \times P_M(I_{\eta}) \right) \cap \widetilde{H}^1_{\varpi}(\mathcal{Q}) = \left\{ \Psi \in P_L(I_{\xi}) \times P_M(I_{\eta}) : \partial_{\xi} \Psi(\xi, 1) = 0 \right\}.$$

Note that the condition  $\partial_{\xi}\Psi(\xi, 1) = 0$  is needed to ensure that  $\Psi(\xi, 1) \equiv c$  for  $\xi \in [-1, 1]$ . To simplify the presentation, we shall take L = M and denote  $X_M = X_{MM}$  and  $\widetilde{X}_M = \widetilde{X}_{MM}$ .

We start with a commonly used  $C^0$ -model basis for  $P_L(I_{\xi}) \times P_M(I_{\eta})$ :

(4.16) 
$$\Psi_{lm}(\xi,\eta) = \psi_l(\xi)\psi_m(\eta), \quad 0 \le l, m \le M,$$

where

(4.17) 
$$\psi_k(z) = \begin{cases} \frac{1-z}{2}, & k = 0, \\ \frac{1-z^2}{4} J_{k-1}^{1,1}(z), & 1 \le k \le M-1, \\ \frac{1+z}{2}, & k = M. \end{cases}$$

We observe that

$$\partial_{\xi} \Psi_{lm}(\xi, 1) = 0, \text{ for } 0 \le l \le M, \ 0 \le m \le M - 1,$$

while

$$\partial_{\xi} \Psi_{lM}(\xi, 1) \neq 0$$
, for  $0 \le l \le M$ , but  $\partial_{\xi} (\Psi_{0M}(\xi, 1) + \Psi_{MM}(\xi, 1)) = 0$ .

Thus, by redefining  $\Psi_{0M}(\xi,\eta) := \frac{1+\eta}{2}$ , we have that

(4.18) 
$$X_M = \operatorname{span} \left\{ \Psi_{lm}(\xi, \eta) : (l, m) \in \Upsilon_M \right\},$$

where the index set

(4.19) 
$$\Upsilon_M := \{(l,m) : 0 \le l \le M, 0 \le m \le M - 1\} \cup \{(0,M)\}.$$

Let  $\Phi_{lm}(x,y) = \Psi_{lm}(\xi,\eta)$  (under the mapping (3.2)–(3.3)), and

(4.20) 
$$X_M = \operatorname{span} \left\{ \Phi_{lm}(x, y) : (l, m) \in \Upsilon_M \right\},$$

and  $\dim(X_M) = M(M+1) + 1$ .

As in the *p*-version of finite elements, we can split this model basis into *interior* and *boundary* modes (including *vertex* and *edge* modes). All interior modes are zero on the triangle boundary. The vertex modes have a unit magnitude at one vertex and are zero at all other vertices, and the edge modes only have magnitude along one edge and are zero at all other vertices and edges.

• Interior modes  $(1 \le l, m \le M - 1)$ :

(4.21) 
$$\Phi_{lm}(x,y) = xy \left(1 - \frac{x}{1-y}\right) J_{l-1}^{1,1} \left(\frac{2x}{1-y} - 1\right) J_{m-1}^{1,1}(2y-1),$$
$$\Psi_{lm}(\xi,\eta) = \frac{\left(1 - \xi^2\right) \left(1 - \eta^2\right)}{16} J_{l-1}^{1,1}(\xi) J_{m-1}^{1,1}(\eta).$$

• Edge modes:

$$\begin{cases} 4.22 \\ y = 0: & \Phi_{l0}(x, y) = x \left( 1 - \frac{x}{1 - y} \right) J_{l-1}^{1,1} \left( \frac{2x}{1 - y} - 1 \right), \\ \eta = -1: & \Psi_{l0}(\xi, \eta) = \frac{\left( 1 - \xi^2 \right) (1 - \eta)}{8} J_{l-1}^{1,1}(\xi), \quad 1 \le l \le M - 1; \\ \begin{cases} x = 0: & \Phi_{0m}(x, y) = (1 - x - y) y J_{m-1}^{1,1}(2y - 1), \\ \xi = -1: & \Psi_{0m}(\xi, \eta) = \frac{\left( 1 - \xi \right) \left( 1 - \eta^2 \right)}{8} J_{m-1}^{1,1}(\eta), \quad 1 \le m \le M - 1; \\ \begin{cases} x + y = 1: & \Phi_{Mm}(x, y) = x y J_{m-1}^{1,1}(2y - 1), \\ \xi = 1: & \Psi_{Mm}(\xi, \eta) = \frac{\left( 1 + \xi \right) \left( 1 - \eta^2 \right)}{8} J_{m-1}^{1,1}(\eta), \quad 1 \le m \le M - 1. \end{cases}$$

• Vertex modes:

(4.23) 
$$\begin{cases} (x,y) = (0,0): & \Phi_{00}(x,y) = 1 - x - y, \\ (\xi,\eta) = (-1,-1): & \Psi_{00}(\xi,\eta) = \frac{(1-\xi)(1-\eta)}{4}; \\ (x,y) = (1,0): & \Phi_{M0}(x,y) = x, \\ (\xi,\eta) = (1,-1): & \Psi_{M0}(\xi,\eta) = \frac{(1+\xi)(1-\eta)}{4}; \\ (x,y) = (0,1): & \Phi_{0M}(x,y) = y, \\ (\xi,\eta) = (-1,1): & \Psi_{0M}(\xi,\eta) = 1 \cdot \frac{1+\eta}{2}. \end{cases}$$

Note that the total number of modal basis functions in the above is  $(M-1)^2 + 3(M-1) + 3 = M(M+1) + 1 = \dim X_M$ .

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

**4.1.3. Linear system.** We now examine the linear system associated to the spectral-Galerkin approximation (4.4).

Let  $\{\Phi_{lm} = \Psi_{lm}\}$  be the basis of  $X_M^0$  as defined in (4.21). By (3.7), (3.9)–(3.10), and (4.16), we have that

$$\begin{aligned} &(4.24) \\ \iint_{\mathcal{T}} \nabla \Phi_{lm} \nabla \Phi_{l'm'} dx dy = \frac{1}{2} \iint_{\mathcal{Q}} \left[ \frac{4 + (1+\xi)^2}{1-\eta} \partial_{\xi} \Psi_{lm} \partial_{\xi} \Psi_{l'm'} + (1+\xi) \partial_{\xi} \Psi_{lm} \partial_{\eta} \Psi_{l'm'} \right] \\ &+ (1+\xi) \partial_{\eta} \Psi_{lm} \partial_{\xi} \Psi_{l'm'} + (1-\eta) \partial_{\eta} \Psi_{lm} \partial_{\eta} \Psi_{l'm'} \right] d\xi d\eta \\ &= \frac{1}{2} \left( \int_{-1}^{1} \left( 4 + (1+\xi)^2 \right) \psi_{l}'(\xi) \psi_{l'}'(\xi) d\xi \right) \left( \int_{-1}^{1} \psi_m(\eta) \psi_{m'}(\eta) \frac{1}{1-\eta} d\eta \right) \\ &+ \frac{1}{2} \left( \int_{-1}^{1} (1+\xi) \psi_{l}'(\xi) \psi_{l'}(\xi) d\xi \right) \left( \int_{-1}^{1} \psi_m(\eta) \psi_{m'}'(\eta) d\eta \right) \\ &+ \frac{1}{2} \left( \int_{-1}^{1} (1+\xi) \psi_{l}(\xi) \psi_{l'}'(\xi) d\xi \right) \left( \int_{-1}^{1} \psi_m'(\eta) \psi_{m'}(\eta) d\eta \right) \\ &+ \frac{1}{2} \left( \int_{-1}^{1} \psi_{l}(\xi) \psi_{l'}(\xi) d\xi \right) \left( \int_{-1}^{1} \psi_m'(\eta) \psi_{m'}'(\eta) d\eta \right) ,\end{aligned}$$

and

(4.25) 
$$\iint_{\mathcal{T}} \Phi_{lm} \Phi_{l'm'} dx dy = \frac{1}{8} \iint_{\mathcal{Q}} \Psi_{lm} \Psi_{l'm'} (1-\eta) d\xi d\eta \\ = \frac{1}{8} \left( \int_{-1}^{1} \psi_l(\xi) \psi_{l'}(\xi) d\xi \right) \left( \int_{-1}^{1} \psi_m(\eta) \psi_{m'}(\eta) (1-\eta) d\eta \right).$$

Setting

(4.26)

$$a_{ij} = \int_{-1}^{1} \left(4 + (1+z)^2\right) \psi'_j(z)\psi'_i(z)dz, \qquad b_{ij} = \int_{-1}^{1} (1-z)\psi'_j(z)\psi'_i(z)dz,$$

$$c_{ij}^{(k)} = \int_{-1}^{1} (1+z)^k \psi'_j(z)\psi_i(z)dz, \quad k = 0, 1,$$

$$d_{ij}^{(k)} = \int_{-1}^{1} (1-z)^k \psi_j(z)\psi_i(z)dz, \quad k = 0, \pm 1,$$

$$u_M(x,y) = \sum_{l=1}^{M-1} \sum_{m=1}^{M-1} \hat{u}_{lm} \Phi_{lm}(x,y), \qquad f_{lm} = \iint_{\mathcal{T}} \mathcal{I}_M f(x,y) \Phi_{lm}(x,y) dxdy,$$

and

$$\mathbf{A} = (a_{ij})_{1 \le i,j \le M-1}, \quad \mathbf{C}_k = \left(c_{ij}^{(k)}\right)_{1 \le i,j \le M-1}, \quad \mathbf{U} = (\hat{u}_{ij})_{1 \le i,j \le M-1}$$

(and likewise for  $\mathbf{B}, \mathbf{D}_k$ , and  $\mathbf{F}$ ), the linear system associated to (4.4) becomes

(4.27) 
$$\mathbf{AUD}_{-1} + \mathbf{C}_1 \mathbf{UC}_0^t + \mathbf{C}_1^t \mathbf{UC}_0 + \mathbf{D}_0 \mathbf{UB} + \frac{\gamma}{4} \mathbf{D}_0 \mathbf{UD}_1 = 2\mathbf{F}.$$

We can also rewrite the above equation in the following form using tensor product notation:

(4.28)

1640

$$(\mathbf{S} + \gamma \mathbf{M}) \mathbf{u} = \left[\frac{1}{2} \left( \mathbf{A} \otimes \mathbf{D}_{-1} + \mathbf{C}_1 \otimes \mathbf{C}_0 + \mathbf{C}_1^t \otimes \mathbf{C}_0^t + \mathbf{D}_0 \otimes \mathbf{B} \right) + \frac{\gamma}{8} \mathbf{D}_0 \otimes \mathbf{D}_1 \right] \mathbf{u} = \mathbf{f},$$

where  ${\bf u}$  and  ${\bf f}$  are  ${\bf U}$  and  ${\bf F}$  written in the form of a column vector, i.e.,

 $\mathbf{u} = (\hat{u}_{11}, \dots, \hat{u}_{1,M-1}; \hat{u}_{21}, \dots, \hat{u}_{2,M-1}; \dots; \hat{u}_{M-1,1}, \dots, \hat{u}_{M-1,M-1})^t,$ 

and  $\otimes$  denotes the tensor product of matrices, i.e.,  $\mathbf{A} \otimes \mathbf{B} = (\mathbf{A}b_{ij})_{i,j=0,\ldots,M-1}$ . Since the triangle  $\mathcal{T}$  is not a separable domain, the linear system (4.27) cannot be solved by the usual matrix diagonalization method (cf. [19, 24]).

By using

(4.29) 
$$(1-z^2) J_{k-1}^{1,1}(z) = \frac{2k}{2k+1} \left( J_{k-1}^{0,0}(z) - J_{k+1}^{0,0}(z) \right) = -2k \int_{-1}^{z} J_{k}^{0,0}(t) dt,$$

and the three-term recurrence formulas

(4.30) 
$$L_{n+1}(z) = \frac{2n+1}{n+1} z L_n(z) - \frac{n}{n+1} L_{n-1}(z); \quad L_0(z) = 1, \quad L_1(z) = z,$$

(4.31)

$$J_{n+1}^{1,1}(z) = \frac{(n+2)(2n+3)}{(n+1)(n+3)} z J_n^{1,1}(z) - \frac{n+2}{n+3} J_{n-1}^{1,1}(z); \quad J_0^{1,1}(z) = 1, \quad J_1^{1,1}(z) = 2z,$$

and the orthogonality of Jacobi polynomials, we can easily verify that

$$\begin{split} a_{ij} &= 0, \quad \forall \ |i-j| > 2, \qquad b_{ij} = 0, \quad \forall \ |i-j| > 1, \\ c_{ij}^{(k)} &= 0, \quad \forall \ |i-j| > 1+k, \quad k = 0, 1, \\ d_{ij}^{(k)} &= 0, \quad \forall \ |i-j| > 2+k, \quad k = 0, \pm 1. \end{split}$$

The nonzero entries of these matrices can be evaluated exactly. The structures of the mass matrix M and the stiffness matrix S are depicted in Figure 4.1. Therefore, the linear system can be effectively solved by using an efficient sparse solver such as SPARSEPACK.

Alternatively, one may attempt to use a suitable iterative solver. As is typical in a spectral method, the matrix  $\mathbf{S} + \gamma \mathbf{M}$  is usually very ill-conditioned so it is necessary to construct a suitable preconditioner. We now examine the condition numbers of the stiffness and mass matrices as well as the effect of diagonal preconditioner. For this purpose, let  $\mathbf{\Lambda} = (\operatorname{diag}(\mathbf{S}))^{-1/2}$ , and let  $\alpha_A$  be an index such that  $\operatorname{Cond}(A) = M^{\alpha_A}$ , i.e.,  $\alpha_A = \log_M(\operatorname{Cond}(A))$  for a given matrix A. We tabulate below the condition numbers of the matrices  $\mathbf{S}$ ,  $\mathbf{\tilde{S}} = \mathbf{\Lambda} \mathbf{S} \mathbf{\Lambda}$ ,  $\mathbf{T}_{\gamma} := \mathbf{S} + \gamma \mathbf{M}$ , and  $\mathbf{\tilde{T}}_{\gamma} = \mathbf{\Lambda} \mathbf{T}_{\gamma} \mathbf{\Lambda}$  for various M and  $\gamma = 10^4$ .

We observe from Table 4.1 that the condition numbers of the stiffness matrix **S** and  $\mathbf{T}_{\gamma} := \mathbf{S} + \gamma \mathbf{M}$  behave like  $O(M^3)$ , while their preconditioned (by diagonal preconditioner) counterparts behave like  $O(M^2)$ . How to construct a simple and optimal preconditioner in this case is still an open question. We refer to [15] for an attempt using finite difference.

**4.1.4. Nodal basis functions.** In practice, it is often more convenient to use a nodal basis. We now construct a nodal basis for  $X_M$ . Let  $\{\xi_i = \eta_i\}_{0 \le i \le M}$  (with

TABLE 4.1					
Condition numbers.					

M				$\gamma = 10000$				
	$\mathrm{cond}(\mathbf{S})$	$\alpha_{\mathbf{s}}$	$\mathrm{cond}(\widetilde{\mathbf{S}})$	$\alpha_{\widetilde{\mathbf{s}}}$	$\operatorname{cond}(\mathbf{T}_{\gamma})$	$\alpha_{{\bf T}_\gamma}$	$\operatorname{cond}(\widetilde{\mathbf{T}}_{\gamma})$	$\alpha_{\widetilde{\mathbf{T}}\gamma}$
2	3.9	1.97	2.8	1.49	6.9	2.79	1.8	0.85
4	19.3	2.14	9.4	1.62	117.9	3.44	19.3	2.14
8	119.5	2.30	31.5	1.66	1727.6	3.58	143.8	2.39
16	843.2	2.43	108.8	1.69	8369.7	3.26	298.2	2.06
32	6389.5	2.53	394.1	1.72	31234.0	2.99	629.1	1.86
64	50152.4	2.60	1646.5	1.78	119320.9	2.81	1964.7	1.82



FIG. 4.1. The structure of the mass matrix (left), and the stiffness matrix (right), where M = 8.

 $\xi_M = \eta_M = 1$ ) be the Legendre–Gauss–Lobatto points, and let  $\{h_i(\xi), h_i(\eta)\}_{0 \le i \le M}$ be the Lagrange basis associated with  $\{\xi_i = \eta_i\}_{0 \le i \le M}$ . We define

(4.32) 
$$\begin{aligned} \Phi_{lm}(x,y) &= \Psi_{lm}(\xi,\eta) = h_l(\xi)h_m(\eta), \quad 0 \le l \le M, \quad 0 \le m \le M-1, \\ \tilde{\Phi}_{0M}(x,y) &= \tilde{\Psi}_{0M}(\xi,\eta) = h_M(\eta), \end{aligned}$$

where (x, y) and  $(\xi, \eta)$  are related by the mapping (3.2)–(3.3). Then, we have

$$\widetilde{X}_M = \operatorname{span}\left\{\widetilde{\Psi}_{lm}(\xi,\eta) : (l,m) \in \Upsilon_M\right\}, \quad X_M = \operatorname{span}\left\{\widetilde{\Phi}_{lm}(x,y) : (l,m) \in \Upsilon_M\right\},$$

where the index set  $\Upsilon_M$  is given in (4.19).

Note that the set of collocation points (Figure 4.2, right) used here to construct the nodal basis for  $X_M$  is different from the set of interpolation points (Figure 4.2, left) in section 3. This approach is reminiscent of the Chebyshev–Legendre method in [7] (see, also [25]).

**4.2.** Multidomain case. We now briefly describe how to set up a multidomain spectral-element method for (4.2) following the approach in [3]. For more details in this regard, we refer to the books [3, 17, 6].

Let  $\Omega$  be an open bounded domain with Lipschitz boundaries which is decomposed as follows:

(4.34) 
$$\bar{\Omega} = \bigcup_{k=1}^{K} \bar{\Omega}_k, \ \Omega_i \cap \Omega_j = \emptyset, \ i \neq j.$$



FIG. 4.2. Distributions of the collocation points (M = 6). Left: Legendre-Gauss-Lobatto (in  $\xi$ ) and Jacobi-Gauss-Radau (in  $\eta$ ); Right: Legendre-Gauss-Lobatto in both directions.

We assume that the decomposition satisfies the following properties:

- each  $\Omega_k$  is spectrally admissible in the sense that there exists  $F_k$ , a bijection of class  $C^{\infty}$ , which maps  $\overline{\hat{\Omega}}$  onto  $\overline{\Omega}_k$ , where  $\hat{\Omega}$  is either the reference triangle  $\mathcal{T}$  or the reference square  $(-1,1)^2$ , such that its inverse  $F_k^{-1}$  is of class  $C^{\infty}$  on  $\overline{\Omega}_k$ ;
- the decomposition is conforming in the sense that the intersection  $\bar{\Omega}_k \cap \bar{\Omega}_j$  $(1 \le k < j \le K)$  is either empty or a node or a whole side of  $\Omega_k$  and  $\Omega_j$ ;
- for each common side  $\Gamma_{kj}$  of  $\overline{\Omega}_k$  and  $\overline{\Omega}_j$ , which is an image of a side  $\Gamma$  of  $\overline{\Omega}$ through  $F_k$  and  $\Gamma'$  of  $\Omega'$  through  $F_j$ , we have the identity

$$F_k^{-1}(x) = G \circ F_j^{-1}(x), \ \forall x \in \Gamma_{kj}$$

where G is the rotation or translation which maps  $\hat{\Gamma}'$  to  $\hat{\Gamma}.$  We denote

(4.35) 
$$\tilde{X}_{M}^{k} = \left\{ v_{M} = \hat{v}_{M} \circ F_{k}^{-1} : \hat{v}_{M} \in \Re_{MM}(\mathcal{T}) \right\} \quad \text{if } \Omega_{k} \text{ is a triangle,}$$

or

(4.36)  $\tilde{X}_M^k = \left\{ v_M = \hat{v}_M \circ F_k^{-1} : \hat{v}_M \in \mathcal{P}_M(I) \times \mathcal{P}_M(I) \right\}$  if  $\Omega_k$  is a quadrilateral.

Setting

$$(4.37) \quad \widehat{X}_M = \left\{ v_M \in H^1(\Omega) : v_M|_{\Omega_k} \in \widetilde{X}_M^k, \ 1 \le k \le K \right\}, \quad \widehat{X}_M^0 = \widehat{X}_M \cap H_0^1(\Omega),$$

the combined rational function-Legendre spectral-element approximation to (4.2) is:

Find  $u_M \in \widehat{X}_M^0$  such that

$$(4.38) \quad \sum_{k=1}^{K} \left( \gamma(u_M, v_M)_{\Omega_k} + (\nabla u_M, \nabla v_M)_{\Omega_k} \right) = \sum_{k=1}^{K} \left( \mathcal{I}_M^k f, v_M \right)_{\Omega_k}, \quad \forall v_M \in \widehat{X}_M^0$$

where  $\mathcal{I}_{M}^{k} f \in \tilde{X}_{M}^{k}$  such that  $\mathcal{I}_{M}^{k} f(\xi_{i}^{k}, \eta_{j}^{k}) = f(\xi_{i}^{k}, \eta_{j}^{k}), \quad 0 \leq i, j \leq M$ , and  $(\xi_{i}^{k}, \eta_{j}^{k}) = F_{k}^{-1}(\hat{\xi}_{i}^{k}, \hat{\eta}_{j}^{k})$  with  $\{\hat{\xi}_{i}^{k}\}_{0 \leq i \leq M}$  and  $\{\hat{\eta}_{i}^{k}\}_{0 \leq i \leq M}$  being, respectively, the Legendre–Gauss–Lobatto points and Jacobi-(with index (1,0))–Gauss–Radau points considered in section 2.

Then, using the standard error estimates of the projection and interpolation operators for the Legendre approximation (see, for instance, [2]) and the error estimates of the projection and interpolation operators for the rational approximation established in section 3, and following a procedure similar to that in sections 6.1 and 6.2 of [3], we can prove the following result.

PROPOSITION 4.1. Assuming that the solution of (4.2)  $u \in H_0^1(\Omega)$  and  $u|_{\Omega_k} \in H^r(\Omega_k)$  for  $1 \leq k \leq K$  with  $r \geq 1$ , and that  $f|_{\Omega_k} \in H^s(\Omega_k)$  for  $1 \leq k \leq K$  with  $s \geq 2$ . Then, the approximate solution  $u_M$  of (4.38) satisfies the following error estimate:

(4.39) 
$$\|u - u_M\|_{1,\Omega} \lesssim \sum_{k=1}^{K} \left( M^{1-r} \|u\|_{H^r(\Omega_k)} + M^{-s} \|f\|_{H^s(\Omega_k)} \right).$$

*Remark* 4.3. It can also be shown that Proposition 4.1 holds if we replace the inner product in (4.38) in each subdomain by the discrete inner product (cf. Remark 4.1).

5. Numerical results and discussions. To illustrate the convergence rate of our rational approximation, we implemented the rational approximation to the model equation (4.1) with two exact solutions and we report our numerical results below.

For a given M, we denote the discrete  $L^2$ -error by

(5.1) 
$$E_M = \left(\sum_{l=0}^{M} \sum_{m=0}^{M} \left(u_M(\xi_l, \eta_m) - u(\xi_l, \eta_m)\right)^2 \rho_{lm}\right)^{\frac{1}{2}}$$

*Example* 1. We consider the equation (4.1) with  $\gamma = 1$  and the exact solution:

(5.2) 
$$u(x,y) = xy \left(e^{x+y} - e\right), \quad (x,y) \in \mathcal{T}.$$

Since  $u \in H^r(\mathcal{T})$  for any r > 0, Theorem 4.1 indicates that the error will converge faster than any algebraic order. Indeed, as shown in Figure 5.1 (left), the errors decay exponentially, typical for a spectral approximation to an analytic function.

*Example 2.* We consider the equation (4.1) with  $\gamma = 1$  and the exact solution:

(5.3) 
$$u(x,y) = x^{\alpha} y^{\beta} (1-x-y), \quad (x,y) \in \mathcal{T}, \ \alpha, \beta > 0.$$



FIG. 5.1. Maximum pointwise errors (marked by " $\circ$ ") and  $L^2$ -errors (marked by " $\diamond$ ") against various M. Left: Example 1; Right: Example 2 with  $\alpha = 3.2$  and  $\beta = 2.8$ .

TABLE 5.1						
$L^2$ -errors	and	convergence	rate.			

	$\alpha = \beta$	= 2.5	$\alpha = 3.3$ and $\beta = 2$		
M	$E_M$	rate	$E_M$	rate	
8	5.935656e - 03	—	3.956740e - 04	_	
16	1.209618e - 04	5.62	4.921643e - 07	9.65	
24	1.535837e - 05	5.09	3.218648e - 08	6.73	
32	3.634080e - 06	5.01	4.817740e - 09	6.60	
48	4.814852e - 07	4.99	$3.349831e{-10}$	6.58	
56	2.234979e - 07	4.98	$1.216786e{-10}$	6.57	

As discussed in Remark 4.2, we expect the error estimate (4.13) with r given by (4.12) with  $\alpha$  and/or  $\beta$  being noninteger, assuming that the integrals in (4.26) are computed exactly. If  $\alpha, \beta \in \mathbb{N}$ , we expect that our numerical solution converges exponentially to the exact solution. In Figure 5.1 (right), we plot the maximum and  $L^2$ -errors versus various M with  $\alpha = 3.2$  and  $\beta = 2.8$  in the log-log scale. The algebraic convergence rate is evidenced by the near straight lines in the plot.

We also tabulated in Table 5.1 the discrete  $L^2$ -errors and the approximate convergence rate defined as

$$-\frac{\ln E_{M_k} - \ln E_{M_{k+1}}}{\ln M_k - \ln M_{k+1}}.$$

We note from the table that the asymptotic convergence rate is about 5 in the case of  $\alpha = \beta = 2.5$  and about 6.6 in the case of  $\alpha = 3.3$  and  $\beta = 2$ . Since we used a quadrature rule to compute the integrals in (4.26), these rates of convergence are about one-order less than the rate given in (4.13). This difference can be attributed to the interpolation error which is not taken into account in (4.13). Hence, our numerical results are in good agreement with the error estimates in Theorem 4.1 (cf. Remark 4.2).

**Concluding remarks.** We introduced and analyzed in this paper a rational approximation in the triangle. The rational basis functions in the triangle are obtained from tensor product of 1D polynomials in the reference square through the collapsed coordinate transform (3.2). We derived optimal error estimates for the  $L^2$ - and  $H_0^1$ -orthogonal projections with upper bounds expressed in the original coordinates in the triangle. These fundamental approximation results are then used to derive error estimates for the spectral-element method using a combination of rational functions in triangles and polynomials in quadrilaterals.

Our error analysis indicates that the rational approximation leads to error estimates which are as accurate as the polynomial approximation in the triangle. Furthermore, these rational basis functions appear to be easier to deal with, both in analysis and in practice, than the polynomial basis functions in (1.1). Hence, this rational approximation represents a viable alternative to the polynomial approximation for triangular domains, despite the drawback induced by the clustering of collocation points around the singular vertex.

We provided implementation detail for the rational spectral-Galerkin approximation to a Poisson-type equation, and show that the resultant linear system is sparse and can be efficiently solved, for example, by a sparse solver. We also presented illustrative numerical results which are in agreement with the theoretical estimates.

This work is the first step towards developing a spectral-element method for complex geometries using rational functions in triangles.

Appendix A. Proof of Lemma 2.2. We collect below some properties of Jacobi and generalized Jacobi polynomials. Recall that the classical Jacobi polynomials with indexes  $\alpha, \beta > -1$  satisfy the following recurrence relations (see [1] and [23]),

(A.1)  $J_n^{\alpha,\beta}(\zeta) = (-1)^n J_n^{\beta,\alpha}(-\zeta), \quad \alpha,\beta > -1, \ n \ge 0,$ (A.2)  $J_{n-1}^{\alpha,\beta}(\zeta) = J_n^{\alpha,\beta-1}(\zeta) - J_n^{\alpha-1,\beta}(\zeta), \quad \alpha,\beta > 0, \ n \ge 0,$ 

(A.3)

$$J_n^{\alpha,\beta}(\zeta) = \frac{n+\beta}{n+\alpha+\beta} J_n^{\alpha,\beta-1}(\zeta) + \frac{n+\alpha}{n+\alpha+\beta} J_n^{\alpha-1,\beta}(\zeta), \quad \alpha,\beta > 0, \ n \ge 0,$$

$$\begin{aligned} (A.4)\\ (1-\zeta)J_n^{\alpha+1,\beta}(\zeta)\\ &= \frac{2(n+\alpha+1)}{2n+\alpha+\beta+2}J_n^{\alpha,\beta}(\zeta) - \frac{2(n+1)}{2n+\alpha+\beta+2}J_{n+1}^{\alpha,\beta}(\zeta), \quad \alpha,\beta > -1, \ n \ge 0, \end{aligned}$$

(A.5)

$$\partial_{\zeta} J_n^{\alpha,\beta}(\zeta) = \frac{n+\alpha+\beta+1}{2} J_{n-1}^{\alpha+1,\beta+1}(\zeta), \quad \alpha,\beta > -1, \ n \ge 1,$$

(A.6)  

$$-\omega^{-\alpha,-\beta}(\zeta)\partial_{\zeta}\left(\omega^{\alpha+1,\beta+1}(\zeta)\partial_{\zeta}J_{n}^{\alpha,\beta}(\zeta)\right)$$

$$=n(n+\alpha+\beta+1)J_{n}^{\alpha,\beta}(\zeta), \quad \alpha,\beta>-1, \ n\geq 0.$$

By (A.5), (A.4), (A.2), and (A.3), we also derive that for any  $\alpha>0,\beta>-1,$   $n\geq0,$ 

$$\begin{aligned} (A.7)\\ \partial_{\zeta} \left( (1-\zeta)^{\alpha} J_{n}^{\alpha,\beta}(\zeta) \right) &= (1-\zeta)^{\alpha-1} \left( -\alpha J_{n}^{\alpha,\beta}(\zeta) + (1-\zeta) \partial_{\zeta} J_{n}^{\alpha,\beta}(\zeta) \right) \\ &= (1-\zeta)^{\alpha-1} \left( -\alpha J_{n}^{\alpha,\beta}(\zeta) + \frac{n+\alpha+\beta+1}{2} (1-\zeta) J_{n-1}^{\alpha+1,\beta+1}(\zeta) \right) \\ &= (1-\zeta)^{\alpha-1} \left( -\alpha J_{n}^{\alpha,\beta}(\zeta) + \frac{n+\alpha+\beta+1}{2n+\alpha+\beta+1} \left( (n+\alpha) \right) \right) \\ &= (1-\zeta)^{\alpha-1} \left( -\alpha J_{n}^{\alpha,\beta}(\zeta) + \frac{n+\alpha+\beta+1}{2n+\alpha+\beta+1} \left( (n+\alpha) \right) \right) \\ &= (1-\zeta)^{\alpha-1} \left( -\alpha J_{n}^{\alpha,\beta}(\zeta) + \frac{n+\alpha+\beta+1}{2n+\alpha+\beta+1} \right) J_{n}^{\alpha,\beta}(\zeta) + \frac{n+\alpha+\beta+1}{2n+\alpha+\beta+1} \\ &\qquad \left( -(n+\alpha) - \frac{(n+\alpha)n}{n+\alpha+\beta+1} \right) J_{n}^{\alpha-1,\beta+1}(\zeta) \right) \\ &= -(n+\alpha)(1-\zeta)^{\alpha-1} J_{n}^{\alpha-1,\beta+1}(\zeta). \end{aligned}$$

Now we resort to [11], and define the generalized Jacobi polynomials

$$J_{n}^{\alpha,\beta}(\zeta) = \begin{cases} \left(\frac{\zeta-1}{2}\right)^{-\alpha} \left(\frac{\zeta+1}{2}\right)^{-\beta} J_{n+\alpha+\beta}^{-\alpha,-\beta}(\zeta), & \alpha \in \mathbb{Z}^{-}, \ \beta \in \mathbb{Z}^{-}, \ n \ge n_{0}(\alpha,\beta) := -\alpha - \beta, \\ c_{n}^{\alpha,\beta} \left(\frac{\zeta-1}{2}\right)^{-\alpha} J_{n+\alpha}^{-\alpha,\beta}(\zeta), & \alpha \in \mathbb{Z}^{-}, \ \beta > -1, \ n \ge n_{0}(\alpha,\beta) := -\alpha, \\ c_{n}^{\alpha,\beta} \left(\frac{\zeta+1}{2}\right)^{-\beta} J_{n+\beta}^{\alpha,-\beta}(\zeta), & \alpha > -1, \ \beta \in \mathbb{Z}^{-}, \ n \ge n_{0}(\alpha,\beta) := -\beta, \\ J_{n}^{\alpha,\beta}(\zeta), & \alpha > -1, \ \beta > -1, \ n \ge n_{0}(\alpha,\beta) := 0, \end{cases}$$

with the normalization coefficient

$$(A.9) \quad c_n^{\alpha,\beta} = \begin{cases} \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}, & \alpha > -1, \ \beta \in \mathbb{Z}^-; \ \text{or} \ \beta > -1, \ \alpha \in \mathbb{Z}^-, \\ 1, & \text{otherwise.} \end{cases}$$

As it is known in [11],  $J_n^{\alpha,\beta}(\zeta), n \ge n_0(\alpha,\beta)$  are mutually orthogonal with respect to the weight  $(1-\zeta)^{\alpha}(1+\zeta)^{\beta}$  for any  $\alpha,\beta > -1$  or being negative integers. By such a definition, the generalized Jacobi polynomials maintain most of the recurrence relations of the classic Jacobi polynomials.

LEMMA A.1. Let  $\alpha, \beta > -1$  or negative integers. Then for any  $n \ge n_0(\alpha, \beta)$ , (A.10)

$$J_{n}^{\alpha,\beta}(\zeta) = (-1)^{n} J_{n}^{\beta,\alpha}(-\zeta),$$
(A.11)  

$$\frac{d^{\varrho}}{d\zeta^{\varrho}} J_{n}^{\alpha,\beta}(\zeta) = \frac{\Gamma(n+\alpha+\beta+\varrho+1)}{2^{\varrho}\Gamma(n+\alpha+\beta+1)} J_{n-\varrho}^{\varrho+\alpha,\varrho+\beta}(\zeta), \qquad n \ge \varrho \ge 1,$$
(A.12)  

$$(-1)^{\varrho} \frac{d^{\varrho}}{d\zeta^{\varrho}} \left( \omega^{\varrho+\alpha,\varrho+\beta}(\zeta) \frac{d^{\varrho}}{d\zeta^{\varrho}} J_{n}^{\alpha,\beta}(\zeta) \right) = \frac{\Gamma(n+1)\Gamma(n+\alpha+\beta+\varrho+1)}{\Gamma(n-\varrho+1)\Gamma(n+\alpha+\beta+1)} \omega^{\alpha,\beta}(\zeta) J_{n}^{\alpha,\beta}(\zeta).$$

*Proof.* (A.10) is an immediate consequence of the definition (A.8). Note that for each admissible integer n,

$$(A.13) J_n^{\alpha,\beta}(\zeta) = \begin{cases} \left(\frac{\zeta-1}{2}\right)^{-\alpha} \left(\frac{\zeta+1}{2}\right)^{-\beta} J_{n+\alpha+\beta}^{-\alpha,-\beta}(\zeta), & \alpha,\beta \in \mathbb{Z}^- \cup \{0\}, \\ c_n^{\alpha,\beta} \left(\frac{\zeta-1}{2}\right)^{-\alpha} J_{n+\alpha}^{-\alpha,\beta}(\zeta), & \alpha \in \mathbb{Z}^- \cup \{0\}, \beta > -1, \\ c_n^{\alpha,\beta} \left(\frac{\zeta+1}{2}\right)^{-\beta} J_{n+\beta}^{\alpha,-\beta}(\zeta), & \alpha > -1, \beta \in \mathbb{Z}^- \cup \{0\}, \\ J_n^{\alpha,\beta}(\zeta), & \alpha > -1, \beta > -1. \end{cases}$$

Thus, by (A.5), (A.6), and (A.13), we get that for  $\alpha, \beta \in \mathbb{Z}^-$ ,

$$\begin{aligned} \partial_{\zeta} J_{n}^{\alpha,\beta}(\zeta) &= (-1)^{\alpha} 2^{\alpha+\beta} \partial_{\zeta} \left( (1-\zeta)^{-\alpha} (1-\zeta)^{-\beta} J_{n+\alpha+\beta}^{-\alpha,-\beta}(\zeta) \right) \\ &= \frac{(-1)^{\alpha} 2^{\alpha+\beta+1}}{n} \partial_{\zeta} \left( (1-\zeta)^{-\alpha} (1-\zeta)^{-\beta} \partial_{\zeta} J_{n+\alpha+\beta+1}^{-\alpha-1,-\beta-1}(\zeta) \right) \\ &= (-1)^{\alpha+1} 2^{\alpha+\beta+1} (n+\alpha+\beta+1) (1-\zeta)^{-\alpha-1} (1-\zeta)^{-\beta-1} J_{n+\alpha+\beta+1}^{-\alpha-1,-\beta-1}(\zeta) \\ &= \frac{n+\alpha+\beta+1}{2} J_{n-1}^{\alpha+1,\beta+1}(\zeta), \end{aligned}$$

## Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

while by (A.7), (A.13), we have that for  $\alpha \in \mathbb{Z}^-, \beta > -1$ ,

(A.15) 
$$\partial_{\zeta} J_n^{\alpha,\beta}(\zeta) = 2^{\alpha} c_n^{\alpha,\beta} \partial_{\zeta} \left( (\zeta - 1)^{-\alpha} J_{n+\alpha}^{-\alpha,\beta}(\zeta) \right)$$
$$= 2^{\alpha} n c_n^{\alpha,\beta} (\zeta - 1)^{-\alpha - 1} J_{n+\alpha}^{-\alpha - 1,\beta + 1}(\zeta) = \frac{n}{2} J_{n-1}^{\alpha + 1,\beta + 1}(\zeta)$$

we also get from (A.10) and (A.15) that for  $\alpha > -1, \beta \in \mathbb{Z}^-$ ,

(A.16) 
$$\partial_{\zeta} J_n^{\alpha,\beta}(\zeta) = (-1)^n \partial_{\zeta} J_n^{\beta,\alpha}(-\zeta) = \frac{(-1)^{n+1}n}{2} J_{n-1}^{\beta+1,\alpha+1}(-\zeta) = \frac{n}{2} J_{n-1}^{\alpha+1,\beta+1}(\zeta).$$

A combination of (A.14), (A.15), (A.16), (A.5) yields that for  $\alpha, \beta > -1$  or being negative integers,

(A.17) 
$$\partial_{\zeta} J_n^{\alpha,\beta}(\zeta) = \frac{n+\alpha+\beta+1}{2} J_{n-1}^{\alpha+1,\beta+1}(\zeta), \quad n \ge n_0(\alpha,\beta).$$

Hence, a recursive using of (A.17) eventually leads to (A.11).

Moreover, by (A.17) and (A.13), we get that for any  $\alpha, \beta \in \mathbb{Z}^-$ ,

$$\begin{aligned} \text{(A.18)} \\ \partial_{\zeta} \Big( \omega^{\alpha+1,\beta+1}(\zeta) \partial_{\zeta} J_{n}^{\alpha,\beta}(\zeta) \Big) &= \frac{n+\alpha+\beta+1}{2} \partial_{\zeta} \left( \omega^{\alpha+1,\beta+1}(\zeta) J_{n-1}^{\alpha+1,\beta+1}(\zeta) \right) \\ &= (-1)^{\alpha+1} 2^{\alpha+\beta+1} (n+\alpha+\beta+1) \partial_{\zeta} J_{n+\alpha+\beta+1}^{-\alpha-1,-\beta-1}(\zeta) \\ &= (-1)^{\alpha+1} 2^{\alpha+\beta} (n+\alpha+\beta+1) n J_{n+\alpha+\beta}^{-\alpha,-\beta}(\zeta) \\ &= -n(n+\alpha+\beta+1) \omega^{\alpha,\beta}(\zeta) J_{n}^{\alpha,\beta}(\zeta). \end{aligned}$$

Further by (A.17), (A.7), and (A.13), we have that for any  $\alpha > -1$  and  $\beta \in \mathbb{Z}^-$ ,

$$\begin{aligned} \partial_{\zeta} \Big( \omega^{\alpha+1,\beta+1}(\zeta) \partial_{\zeta} J_{n}^{\alpha,\beta}(\zeta) \Big) \\ &= \frac{n+\alpha+\beta+1}{2} \partial_{\zeta} \left( \omega^{\alpha+1,\beta+1}(\zeta) J_{n-1}^{\alpha+1,\beta+1}(\zeta) \right) \\ (A.19) &= 2^{\beta} \frac{\Gamma(n+\alpha+1)(n+\beta+1)}{\Gamma(n)\Gamma(n+\alpha+\beta+1)} \partial_{\zeta} \left( (1-\zeta)^{\alpha+1} J_{n+\beta}^{\alpha+1,-\beta-1}(\zeta) \right) \\ &= -2^{\beta} \frac{\Gamma(n+\alpha+1)(n+\beta+1)}{\Gamma(n)\Gamma(n+\alpha+\beta+1)} (n+\alpha+\beta+1)(1-\zeta)^{\alpha} J_{n+\beta}^{\alpha,-\beta}(\zeta) \\ &= -n(n+\alpha+\beta+1)\omega^{\alpha,\beta}(\zeta) J_{n}^{\alpha,\beta}(\zeta). \end{aligned}$$

By using (A.10) and (A.19), we also have that for any  $\beta > -1$  and  $\alpha \in \mathbb{Z}^-$ ,

(A.20)  
$$\partial_{\zeta} \left( \omega^{\alpha+1,\beta+1}(\zeta) \partial_{\zeta} J_n^{\alpha,\beta}(\zeta) \right) = (-1)^n \partial_{\zeta} \left( \omega^{\beta+1,\alpha+1}(-\zeta) \partial_{\zeta} J_n^{\beta,\alpha}(-\zeta) \right)$$
$$= (-1)^{n+1} n (n+\alpha+\beta+1) \omega^{\beta,\alpha}(-\zeta) J_n^{\beta,\alpha}(-\zeta)$$
$$= -n(n+\alpha+\beta+1) \omega^{\alpha,\beta}(\zeta) J_n^{\alpha,\beta}(\zeta).$$

Now we derive from (A.18), (A.20), (A.19), and (A.6) that for any  $\alpha, \beta > -1$  or being negative integers,

(A.21)  
$$\partial_{\zeta} \Big( \omega^{\alpha+1,\beta+1}(\zeta) \partial_{\zeta} J_n^{\alpha,\beta}(\zeta) \Big) = -n(n+\alpha+\beta+1)\omega^{\alpha,\beta}(\zeta) J_n^{\alpha,\beta}(\zeta), \quad n \ge n_0(\alpha,\beta).$$

Finally, we deduce from (A.21) and (A.11) that

$$\begin{split} \partial_{\zeta}^{\varrho} \Big( \omega^{\alpha+\varrho,\beta+\varrho}(\zeta) \partial_{\zeta}^{\varrho} J_{n}^{\alpha,\beta}(\zeta) \Big) \\ &= \frac{\Gamma(n+\alpha+\beta+\varrho)}{2^{\varrho-1}\Gamma(n+\alpha+\beta+1)} \partial_{\zeta}^{\varrho-1} \partial_{\zeta} \left( \omega^{\alpha+\varrho,\beta+\varrho}(\zeta) \partial_{\zeta} J_{n-\varrho+1}^{\alpha+\varrho-1,\beta+\varrho-1}(\zeta) \right) \\ &= -(n-\varrho+1) \frac{\Gamma(n+\alpha+\beta+\varrho+1)}{2^{\varrho-1}\Gamma(n+\alpha+\beta+1)} \partial_{\zeta}^{\varrho-1} \left( \omega^{\alpha+\varrho-1,\beta+\varrho-1}(\zeta) J_{n-\varrho+1}^{\alpha+\varrho-1,\beta+\varrho-1}(\zeta) \right) \\ &= -(n-\varrho+1)(n+\alpha+\beta+\varrho) \partial_{\zeta}^{\varrho-1} \left( \omega^{\alpha+\varrho-1,\beta+\varrho-1}(\zeta) \partial_{\zeta}^{\varrho-1} J_{n}^{\alpha,\beta}(\zeta) \right), \end{split}$$

and conclude by induction that

(A.22) 
$$\partial_{\zeta}^{\varrho} \left( \omega^{\alpha+\varrho,\beta+\varrho}(\zeta) \partial_{\zeta}^{\varrho} J_{n}^{\alpha,\beta}(\zeta) \right)$$
$$= \frac{(-1)^{\varrho-1} \Gamma(n) \Gamma(n+\alpha+\beta+\varrho+1)}{\Gamma(n-\varrho+1) \Gamma(n+\alpha+\beta+2)} \partial_{\zeta} \left( \omega^{\alpha+1,\beta+1}(\zeta) \partial_{\zeta} J_{n}^{\alpha,\beta}(\zeta) \right)$$
$$= (-1)^{\varrho} \frac{\Gamma(n+1) \Gamma(n+\alpha+\beta+\varrho+1)}{\Gamma(n-\varrho+1) \Gamma(n+\alpha+\beta+1)} \omega^{\alpha,\beta}(\zeta) J_{n}^{\alpha,\beta}(\zeta).$$

This ends the proof.  $\hfill \square$ 

Now let us concentrate on the proof of Lemma 2.2. From the expansion

$$w(\xi) = \sum_{k=0}^{\infty} \hat{w}_k^{\alpha,\beta} J_k^{\alpha,\beta}(\xi),$$

we derive

(A.23) 
$$\frac{d^{\mu}}{d\zeta^{\mu}} \left(\pi_N^{\alpha,\beta} - \mathbb{I}\right) w(\zeta) = -\sum_{k=N+1}^{\infty} \hat{w}_k^{\alpha,\beta} \frac{d^{\mu}}{d\zeta^{\mu}} J_k^{\alpha,\beta}(\zeta).$$

Hence, by (A.11) and the orthogonality of the Jacobi polynomials,

$$\begin{split} \left\| \frac{d^{\mu}}{d\zeta^{\mu}} \left( \pi_{N}^{\alpha,\beta} - \mathbb{I} \right) w \right\|_{\omega^{\mu+\alpha,\mu+\beta},I}^{2} \\ &= \sum_{k=N+1}^{\infty} \left| \hat{w}_{k}^{\alpha,\beta} \right|^{2} \left\| \frac{d^{\mu}}{d\zeta^{\mu}} J_{k}^{\alpha,\beta} \right\|_{\omega^{\mu+\alpha,\mu+\beta},I}^{2} \\ &\leq \sum_{k=N+1}^{\infty} \frac{\Gamma(k-\mu+1)\Gamma(k+\alpha+\beta+\varrho+1)}{\Gamma(k-\varrho+1)\Gamma(k+\alpha+\beta+\mu+1)} \left| \hat{w}_{k}^{\alpha,\beta} \right|^{2} \left\| \frac{d^{\sigma}}{d\zeta^{\sigma}} J_{k}^{\alpha,\beta} \right\|_{\omega^{\sigma+\alpha,\sigma+\beta},I}^{2} \\ &\lesssim N^{2\mu-2\sigma} \sum_{k=N+1}^{\infty} \left| \hat{w}_{k}^{\alpha,\beta} \right|^{2} \left\| \frac{d^{\sigma}}{d\zeta^{\sigma}} J_{k}^{\alpha,\beta} \right\|_{\omega^{\sigma+\alpha,\sigma+\beta},I}^{2}. \end{split}$$

Using the orthogonality and the definition of the projection operator leads to we get that

$$\begin{split} \sum_{k=N+1}^{\infty} \left| \hat{w}_{k}^{\alpha,\beta} \right|^{2} \left\| \frac{d^{\sigma}}{d\zeta^{\sigma}} J_{k}^{\alpha,\beta} \right\|_{\omega^{\sigma+\alpha,\sigma+\beta},I}^{2} &= \left\| \sum_{k=N+1}^{\infty} \hat{w}_{k}^{\alpha,\beta} \frac{d^{\sigma}}{d\zeta^{\sigma}} J_{k}^{\alpha,\beta} \right\|_{\omega^{\sigma+\alpha,\sigma+\beta},I}^{2} \\ &= \left\| \left( \pi_{N-\sigma}^{\sigma+\alpha,\sigma+\beta} - \mathbb{I} \right) \frac{d^{\sigma}w}{d\zeta^{\sigma}} \right\|_{\omega^{\sigma+\alpha,\sigma+\beta},I}^{2}, \end{split}$$

Copyright  ${\tt O}$  by SIAM. Unauthorized reproduction of this article is prohibited.

which gives the first inequality in (2.10). The second inequality is an immediate consequence of (2.6). Therefore, we complete the proof of (2.10).

Acknowledgments. The authors would like to thank Professor Claudio Canuto for stimulating discussions and constructive suggestions regarding this work. The third author would like to thank Nanyang Technological University for the hospitality during his visit.

#### REFERENCES

- [1] R. ASKEY, Orthogonal Polynomials and Special Functions, SIAM, Philadelphia, 1975.
- [2] C. BERNARDI AND Y. MADAY, Spectral methods, in Handbook of Numerical Analysis, Vol. V, P. G. Ciarlet and J. L. Lions, ed., Techniques of Scientific Computing (Part 2), Elsevier, Amsterdam, 1997, pp. 209–486.
- [3] C. BERNARDI, Y. MADAY, AND F. RAPETTI, Discrétisations variationnelles de problèmes aux limites elliptiques, Math. Appl. 45, Springer-Verlag, Berlin, 2004.
- [4] J. P. BOYD, Chebyshev and Fourier Spectral Methods, 2nd ed., Dover Publications, Mineola, NY, 2001.
- [5] D. BRAESS AND C. SCHWAB, Approximation on simplices with respect to weighted Sobolev norms, J. Approx. Theory, 103 (2000), pp. 329–337.
- [6] C. CANUTO, M. Y. HUSSAINI, A. QUARTERONI, AND T. A. ZANG, Spectral methods, Scientific Computation, Springer-Verlag, Berlin, 2006. Fundamentals in single domains.
- [7] W. S. DON AND D. GOTTLIEB, The Chebyshev-Legendre method: Implementing Legendre methods on Chebyshev points, SIAM J. Numer. Anal., 31 (1994), pp. 1519–1534.
- [8] M. DUBINER, Spectral methods on triangles and other domains, J. Sci. Comput., 6 (1991), pp. 345–390.
- M. G. DUFFY, Quadrature over a pyramid or cube of integrands with a singularity at a vertex, SIAM J. Numer. Anal., 19 (1982), pp. 1260–1262.
- [10] D. FUNARO, Polynomial Approximations of Differential Equations, Springer-Verlag, Berlin, 1992.
- [11] B.-Y. GUO, J. SHEN, AND L.-L. WANG, Optimal spectral-Galerkin methods using generalized Jacobi polynomials, J. Sci. Comput., 27 (2006), pp. 305–322.
- [12] B. GUO AND L.-L. WANG, Error analysis of spectral method on a triangle, Adv. Comput. Math., 26 (2007), pp. 473–496.
- [13] B. GUO AND L.-L. WANG, Jacobi interpolation approximations and their application to singular differential equations, Adv. Comput. Math., 14 (2001), pp. 227–276.
- [14] B. GUO AND L.-L. WANG, Jacobi approximations in non-uniformly Jacobi-weighted Sobolev spaces, J. Approximation Theory, 128 (2004), pp. 1–41.
- [15] W. HEINRICHS AND B. I. LOCH, Spectral schemes on triangular elements, J. Comput. Phys., 173 (2001), pp. 279–301.
- [16] J. S. HESTHAVEN, From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex, SIAM J. Numer. Anal., 35 (1998), pp. 655–676.
- [17] G. E. KARNIADAKIS AND S. J. SHERWIN, Spectral/hp element methods for computational fluid dynamics, Numerical Mathematics and Scientific Computation, 2nd ed., Oxford University Press, New York, 2005.
- [18] T. KOORNWINDER, Two-variable analogues of the classical orthogonal polynomials, in Theory and application of special functions (Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975), pp. 435–495. Math. Res. Center, Univ. Wisconsin, Publ. No. 35. Academic Press, New York, 1975.
- [19] R. E. LYNCH, J. R. RICE, AND D. H. THOMAS, Direct solution of partial differential equations by tensor product methods, Numer. Math., 6 (1964), pp. 185–199.
- [20] R. G. OWENS, Spectral approximations on the triangle, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci., 454 (1998), pp. 857–872.
- [21] R. PASQUETTI AND F. RAPETTI, Spectral element methods on unstructured meshes: Comparisons and recent advances, J. Sci. Comput., 27 (2006), pp. 377–387.
- J. PRORIOL, Sur une famille de polynomes à deux variables orthogonaux dans un triangle, C. R. Acad. Sci. Paris, 245 (1957), pp. 2459–2461.
- [23] E. D. RAINVILLE, Special Functions, Macmillan, New York, 1960.
- [24] J. SHEN, Efficient spectral-Galerkin method. I. Direct solvers of second- and fourth-order equations using Legendre polynomials, SIAM J. Sci. Comput., 15 (1994), pp. 1489–1505.

#### JIE SHEN, LI-LIAN WANG, AND HUIYUAN LI

- [25] J. SHEN, Efficient Chebyshev-Legendre Galerkin methods for elliptic problems, in Proceedings of ICOSAHOM'95, A. V. Ilin and R. Scott, eds., Houston J. Math., 1996, pp. 233–240.
- [26] S. J. SHERWIN AND G. E. KARNIADAKIS, A new triangular and tetrahedral basis for high-order (hp) finite element methods, Internat. J. Numer. Methods Engrg., 38 (1995), pp. 3775– 3802.
- [27] M. A. TAYLOR, B. A. WINGATE, AND R. E. VINCENT, An algorithm for computing Fekete points in the triangle, SIAM J. Numer. Anal., 38 (2000), pp. 1707–1720.