IMA Journal of Numerical Analysis (2016) **36**, 872–896
doi:10.1093/imanum/drv029
Advance Access publication on June 9, 2015

Unified analysis of higher-order finite volume methods for parabolic problems on quadrilateral meshes

Min Yang
Department of Mathematics, Yantai University, Yantai, China
yang@ytu.edu.cn

Jiangguo Liu
Department of Mathematics, Colorado State University, Fort Collins, CO 80523-1874, USA
liu@math.colostate.edu

and

Qingsong Zou*
Guangdong Province Key Laboratory of Computational Science, School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou, China
*Corresponding author: mcszqs@mail.sysu.edu.cn

[Received on 20 November 2014; revised on 11 May 2015]

In this paper, a unified analysis for higher-order finite volume methods for parabolic problems on quadrilateral meshes is presented. By studying the *quasi-symmetry* of the finite volume bilinear form, optimal-order error estimates in the $L^\infty(H^1)$- and $L^\infty(L^2)$-norms are derived. The theoretical estimates are validated by numerical experiments.

Keywords: error estimates; finite volume methods; Gaussian points; higher order; parabolic problems; quadrilateral meshes.

1. Introduction

Finite volume methods (FVMs) have been widely used in scientific computing and engineering due to their easy implementation and the local conservation property. Lower-order FVMs are tightly related to finite difference or finite element methods, and have been extensively studied for a long time; see, e.g., Angelini *et al.* (2013), Bank & Rose (1987), Chatzipantelidis *et al.* (2008), Chou & Ye (2007), Ewing *et al.* (2002), Eymard *et al.* (2000), Hackbusch (1989), Hajibeygi & Jenny (2009) and Li *et al.* (2000). Compared with lower-order methods, higher-order FVMs can produce more accurate solutions, and have been widely used in computational fluid dynamics to effectively resolve complex-flow features; see, e.g., Castro *et al.* (2006), Colella *et al.* (2011) and Zhang & Phillip (2012). However, the progress on the theoretical analysis of higher-order FVMs is rather slow.

Over the last few years, research on higher-order FVMs mainly focused on elliptic problems. The difficulty for the analysis lies in the establishment of stability (or inf–sup condition in general). Some earlier works in Li *et al.* (2000), Liebau (1996), Xu & Zou (2009), Lv & Li (2012) and Chen *et al.* (2012) adopt the so-called *elementwise stiffness matrix analysis*, which estimates the eigenvalues of the local stiffness matrix, and thus has to be proceeded scheme by scheme. To the best of our knowledge, only a few systematic works on high-order finite volume scheme appeared in the literature. For instance, a class of high-order finite volume schemes over rectangular meshes has been derived by Cai *et al.* (2003) from high-order finite element methods. One-dimensional high-order finite volume scheme was studied by Plexousakis & Zouraris (2004) and Cao *et al.* (2013). Very recently, a general framework for any order FVMs over quadrilateral meshes has been established in Zhang & Zou (2015).

For time-dependent problems, e.g., parabolic problems, little progress has been made until now. A main difficulty lies in measuring nonsymmetry of the discrete schemes. It is known from Smith (1985) and Thomée (2006) that the *symmetry* property plays a critical role in the error analysis of the numerical methods for parabolic problems. However, the finite volume schemes are usually not symmetric. Certain additional terms related to the deviation from symmetry appear in the error equations. In order to obtain the desired order for error estimates, these terms need some special treatments. The linear FVMs can be treated as small perturbations of the symmetric linear finite element methods, and thus these terms can be well estimated; see, e.g., Chatzipantelidis *et al.* (2008), Chou & Li (2000), Ma *et al.* (2003) and Sinha & Geiser (2007).

However, the higher-order FVMs differ considerably from the corresponding finite element methods, and the 'perturbation' technique successfully used for linear FVMs is not applicable. To our knowledge, limited progress has been made only on quadratic finite volume schemes for parabolic problems. For instance, based on a special dual partition related to the Simpson quadrature, Yang & Liu (2011) investigated techniques for controlling nonsymmetry of a quadratic FVM for parabolic problems on quadrilateral meshes. But only an optimal-order $L^2(H^1)$-error was obtained there. Later, a preprocessing technique based on an elementwise matrix analysis was adopted in Yang *et al.* (2013) to transform the unsymmetric discrete system into a symmetric one. Then with the help of a superconvergence argument, the optimal-order $L^\infty(H^1)$- and $L^\infty(L^2)$- errors were obtained. Note that the analysis developed in Yang & Liu (2011) and Yang *et al.* (2013) relies heavily on the meshes and the approximating polynomials. Thus, it lacks of generality and can hardly be extended to other higher-order cases.

This paper, which is a continuation of our previous work in Yang & Liu (2011), Yang *et al.* (2013) and Zhang & Zou (2015), intends to establish a *unified* analysis for an arbitrary rth ($r \geqslant 2$) order FVM on quadrilateral meshes for the following model parabolic problem:

$$\begin{cases} u_t - \nabla \cdot (a(\mathbf{x})\nabla u) = f(\mathbf{x}, t), \quad (\mathbf{x}, t) \in \Omega \times (0, T], \\ u = 0, \quad (\mathbf{x}, t) \in \partial\Omega \times (0, T], \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \end{cases} \tag{1.1}$$

where Ω is a convex bounded polygonal domain in $\mathbb{R}^2$ with boundary $\partial\Omega$ and $\mathbf{x} = (x, y)$. It is assumed that $f(\mathbf{x}, t) \in L^2(\Omega)$ for $t \in [0, T]$, and $a(\mathbf{x})$ is Lipschitz continuous and bounded almost everywhere with positive lower and upper bounds: a_* and a^*, respectively.

To present a unified analysis, we will adopt a special transfer operator introduced in Zhang & Zou (2015) for the purpose of systematically studying finite volume schemes for elliptic equations. A noticeable benefit of utilizing this operator is that the higher-order finite volume bilinear forms, after being preconditioned by this operator, are comparable to the corresponding finite element bilinear forms. Therefore, a case-by-case analysis can be avoided and the deviation of the bilinear forms from symmetry can be analysed in a systematic way. We will show that the deviations from symmetry is controlled by $\mathcal{O}(h^\gamma)$, which originates from the quadrilateral mesh deformation. We name such a phenomenon as 'quasi-symmetry'. Previously, 'quasi-symmetry' has only been studied for the spatial case $r = 2$ with $\gamma = 1$ in Yang *et al.* (2013). Here, it is the first time that quasi-symmetry is analysed for general higher-order quadrilateral FVMs. By handling quasi-symmetry, the nonsymmetric temporal terms arising from the error equation can be successfully analysed. The optimal-order errors in the L^2- and H^1-norms are then derived under suitable assumptions. The order of the errors is related to the regularities of the exact solution and the quadrilateral meshes being used. Roughly speaking, the errors are proportional to a 'product' of these two types of regularities: for a smoother solution, a less restrictive requirement on meshes can ensure the optimal-order errors; for meshes with better quality, optimal-order errors can be obtained with less requirements on solution regularity (see Lemma 4.2 and Theorems 4.5).

This paper is a first attempt to present a systematic analysis of higher-order finite volume schemes for time-dependent problems. The idea developed in this paper can be extended to numerical treatment of more general cases.

The rest of this paper is organized as follows. In Section 2, we introduce mesh assumptions and the construction of dual volumes based on the Gauss points. Semidiscrete and Crank–Nicolson fully discrete FVMs for parabolic problems on quadrilateral meshes are then presented. The *quasi-symmetry* of the bilinear forms is studied in Section 3. Section 4 derives an L^2-estimate for the elliptic projection, and then presents the convergence analysis of the developed finite volume schemes. Section 5 presents numerical results to illustrate the error estimates.

Throughout this paper, we use the standard notations for the Sobolev spaces $W^{m,p}(\Omega)$ with the norm $\|\cdot\|_{m,p,\Omega}$ and the seminorm $|\cdot|_{m,p,\Omega}$. We also denote $W^{m,2}(\Omega)$ by $H^m(\Omega)$ and skip the index $p=2$ and the domain Ω, when there is no ambiguity, i.e., $\|u\|_{m,p}=\|u\|_{m,p,\Omega}$, $\|u\|_m=\|u\|_{m,2,\Omega}$. The same convention is adopted for the seminorms. We will also use $A\lesssim B$ and $B\gtrsim A$ to denote $A\leqslant CB$, where C is an absolute constant that may take different values in different appearances, but is independent of spatial and temporal discretizations.

2. Finite volume schemes over quadrilateral meshes

We begin with a description of quadrilateral meshes. Let $\mathcal{T}_h=\{Q\}$ be a conforming shape-regular quadrilateral partition of Ω. We assume that quadrilateral partition is '$h^{1+\gamma}$ parallelogram' ($\gamma\geqslant 0$), which means, for any $Q\in\mathcal{T}_h$, the distance between the midpoints of two diagonals of Q is $\mathcal{O}(h_Q^{1+\gamma})$. Note that $\gamma=0$ represents arbitrary quadrilateral meshes, $\gamma=\infty$ represents parallelogram meshes.

REMARK 2.1 The '$h^{1+\gamma}$ parallelogram' assumption has been adopted in Arnold *et al.* (2002), Ewing *et al.* (1999), Süli (1992), Yang & Liu (2011) and Zhang & Zou (2015), although it takes several different forms in the literature. A detailed analysis on the relations of these different forms can be found in Chou & He (2002).

Let $\hat{Q}=[-1,1]^2$ be the reference element in the $\hat{x}\hat{y}$-plane. Assume that, for each element $Q\in\mathcal{T}_h$, there exists a bijective bilinear mapping $F_Q:\hat{Q}\longrightarrow Q$. Let $\mathcal{J}_{F_Q}$ be the Jacobian matrix of F_Q at $\hat{\mathbf{x}}$ and $J_{F_Q}=\det(\mathcal{J}_{F_Q})$, and, accordingly, $\mathcal{J}_{F_Q}^{-1}$ be the Jacobian matrix of F_Q^{-1} at $\mathbf{x}$ and $J_{F_Q^{-1}}=\det(\mathcal{J}_{F_Q}^{-1})$. The sign of J_{F_Q} changes if the local ordering of the vertices is taken in the opposite orientation. Therefore, we may assume that $J_{F_Q}>0$ for every Q.

For any integer $n\geqslant 1$, let

$$\mathbb{Z}_n=\{1,\ldots,n\},\quad \mathbb{Z}_n^0=\{0,1,\ldots,n\}.$$

Let $\{g_i\mid i\in\mathbb{Z}_r\}$ be the r Gauss points of degree r, i.e., zeros of L_r, the Legendre polynomial of degree r, on the interval $[-1,1]$. Let $\{l_i\mid i\in\mathbb{Z}_r^0\}$ be the $r+1$ Lobatto points of degree r in the interval $[-1,1]$, that is, $l_0=-1$, $l_r=1$ and $\{l_i\mid i\in\mathbb{Z}_{r-1}\}$ are the $r-1$ zeros of L_r'.

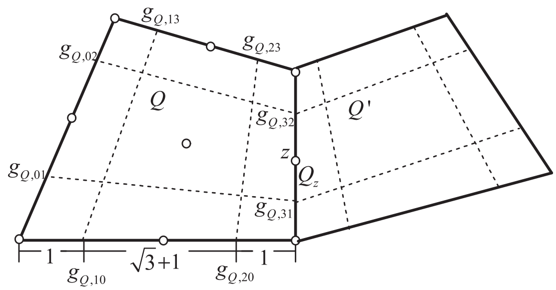

FIG. 1. Adjacent quadrilaterals and control volumes ($r = 2$).

The Gauss and Lobatto points in a quadrilateral Q are, respectively, defined by

$$G_Q = \{\mathbf{g}_{Q,ij} = F_Q(g_i, g_j), 0 \leqslant i, j \leqslant r+1, 1 \leqslant i+j \leqslant 2r+1\}$$

and

$$L_Q = \{\mathbf{l}_{Q,ij} = F_Q(l_i, l_j), 0 \leqslant i, j \leqslant r\}.$$

Moreover, let $\mathcal{G} = \bigcup_{Q\in\mathcal{T}_h} G_Q$ and $\mathcal{L} = \bigcup_{Q\in\mathcal{T}_h} L_Q$ be, respectively, the sets of all Gauss and Lobatto points on $\mathcal{T}_h$.

The dual partition is constructed with the Gauss points. We connect with a line segment each Gauss point on one edge to the one at the same position of its opposite edge. This way, each quadrilateral in $\mathcal{T}_h$ is divided into $(r+1)^2$ subquadrilaterals Q_z, $z \in L_Q$. For each point $z \in \mathcal{L}$, we can associate a control volume V_z, which is the union of the subregions Q_z containing the node z. Therefore, we obtain a collection of control volumes covering Ω. This is the dual partition $\mathcal{T}_h^*$. As an example, the dual partition of a quadrilateral for $r = 2$ is depicted in Fig. 1.

Now, we formulate the FVM for the model problem (1.1). For any interior Lobatto point $z \in \mathcal{L}^0 = \mathcal{L} \setminus \partial\Omega$, integrating the first equation in (1.1) over an associated control volume V_z and applying the Green's formula, we obtain

$$\int_{V_z} u_t \, \mathrm{d}\mathbf{x} - \int_{\partial V_z} a\nabla u \cdot \mathbf{n} \, \mathrm{d}s = \int_{V_z} f(\mathbf{x}, t) \, \mathrm{d}\mathbf{x}, \tag{2.1}$$

where $\mathbf{n}$ denotes the unit outer normal vector on ∂V_z. The above formulation also states that we have a local conservation law on the control volume.

Let S_h be the standard Lagrange finite element space defined by

$$S_h = \{v \in H_0^1(\Omega) \cap C(\Omega) : v = \hat{v} \circ F_Q^{-1}, \hat{v} \in \mathbb{Q}_r(\hat{Q}), \forall Q \in \mathcal{T}_h\},$$

where $\mathbb{Q}_r(\hat{Q})$ is the set of all bi-polynomials on $\hat{Q}$ with degree no more than r ($r \geqslant 2$).

A semidiscrete finite volume scheme for (1.1) is defined as follows: Seek $u_h(t) \in S_h$ such that, for any $v \in S_h^*$,

$$\int_{V_z} u_{h,t} \, \mathrm{d}\mathbf{x} - \int_{\partial V_z} a\nabla u_h \cdot \mathbf{n} \, \mathrm{d}s = \int_{V_z} f(\mathbf{x}, t) \, \mathrm{d}\mathbf{x}, \quad \forall z \in \mathcal{L}^0, \; t \in (0, T], \tag{2.2}$$

with an initial approximation $u_h(0)$ given by $u_h(0) = R_h u_0$, where R_h is the elliptic (Ritz) projection to be defined in (4.5).

Let N be a positive integer. We consider a uniform time step $\Delta t = T/N$ and set $t_n = n\Delta t$ $(0 \leqslant n \leqslant N)$. For $n \geqslant 1$, let

$$\bar{\partial} u_h^n = \frac{u_h^n - u_h^{n-1}}{\Delta t}, \quad u_h^{n,1/2} = \frac{u_h^n + u_h^{n-1}}{2}.$$

A Crank–Nicolson fully discrete finite volume scheme for (1.1) seeks $u_h^n \in S_h$ such that for any $v \in S_h^*$,

$$\int_{V_z} \bar{\partial} u_h^n \,\mathrm{d}\mathbf{x} - \int_{\partial V_z} a\nabla u_h^{n,1/2} \cdot \mathbf{n} \,\mathrm{d}s = \int_{V_z} f^{n,1/2} \,\mathrm{d}\mathbf{x} \quad \forall z \in \mathcal{L}^0,\ n \geqslant 1, \tag{2.3}$$

with an initial approximation given by $u_h^0 = R_h u_0$.

3. Quasi-symmetry of FVMs

The *symmetry* property plays a critical role in the error analysis of numerical methods for parabolic problems. As the higher-order finite volume schemes are often not symmetric in a common sense, we discuss in this section the so-called *quasi-symmetry property* of the corresponding FVMs.

For this purpose, we shall follow Bank & Rose (1987), Li *et al.* (2000) and Xu & Zou (2009) to write our finite volume schemes into Petrov–Galerkin ones. A main advantage of the latter formulations is that we can use the setting of finite element methods for the analysis.

Let

$$S_h^* = \{v \in L^2(\Omega) : v|_{V_z} \text{ is constant},\ \forall z \in \mathcal{L}^0;\ v|_{V_z} = 0, \forall z \in \partial\Omega\}$$

be a piecewise constant function space defined on the control volumes.

We recall a transformation Π_h^* from the trial space S_h to the test space S_h^* introduced in Zhang & Zou (2015). For any $Q \in \mathcal{T}_h$ and $v \in S_h^*$, define by $v_{ij} = v(\mathbf{l}_{ij})$, $(i,j) \in \mathbb{Z}_r^0 \times \mathbb{Z}_r^0$. Let $[v]_{\hat{x},ij} = v_{ij} - v_{ij-1}$, $(i,j) \in \mathbb{Z}_r^0 \times \mathbb{Z}_r$ be the jump of v across the edge $\overline{\mathbf{g}_{ij}\mathbf{g}_{i+1j}}$ and $[v]_{\hat{y},ij} = v_{ij} - v_{i-1j}$, $(i,j) \in \mathbb{Z}_r \times \mathbb{Z}_r^0$ be the jump of v across the edge $\overline{\mathbf{g}_{ij}\mathbf{g}_{ij+1}}$. For $(i,j) \in \mathbb{Z}_r \times \mathbb{Z}_r$, let the double jump of v at the Gauss point $\mathbf{g}_{ij}$ be defined as

$$\lfloor v \rfloor_{ij} = v_{ij} + v_{i-1j-1} - v_{i-1j} - v_{ij-1}. \tag{3.1}$$

Let $\Pi_h^* : S_h \longrightarrow S_h^*$ be a linear mapping such that, for any $v_h \in S_h$, the coefficients of $\Pi_h^* v_h$ are determined by

$$\lfloor \Pi_h^* v_h \rfloor_{ij} = A_i A_j \frac{\partial^2 \hat{v}_h}{\partial \hat{x} \partial \hat{y}}(g_i, g_j) \quad \forall Q \in \mathcal{T}_h,\ (i,j) \in \mathbb{Z}_r \times \mathbb{Z}_r, \tag{3.2}$$

where $\hat{v}_h = v_h \circ F_Q \in \mathbb{Q}_r(\hat{Q})$ and A_i, $i \in \mathbb{Z}_r$ are the weights of the r-point Gauss quadrature for computing the integral $\int_{-1}^1 v(x)\,\mathrm{d}x$. The Gauss quadrature error is given by Davis & Rabinowitz (1984):

$$\int_{-1}^1 f \,\mathrm{d}x - \sum_{i=1}^r A_i f(g_i) = C_r f^{(2r)}(\zeta) \quad \text{for some } \zeta \in (-1,1), \tag{3.3}$$

where $C_r = 2^{2r+1}(r!)^4/(2r+1)((2r)!)^3$.

It has been proved in Zhang & Zou (2015) that Π_h^* is a well-defined bijection and satisfies the following properties.

LEMMA 3.1 For any $v_h \in S_h$ and any $Q = \Box P_1P_2P_3P_4 \in \mathcal{T}_h$,

$$(\Pi_h^* v_h)(P_i) = v_h(P_i), \quad 1 \leqslant i \leqslant 4, \tag{3.4}$$

$$[\Pi_h^* v_h]_{\hat{x},rj} = A_j \frac{\partial \hat{v}_h}{\partial \hat{y}}(1, g_j), \quad [\Pi_h^* v_h]_{\hat{y},ir} = A_i \frac{\partial \hat{v}_h}{\partial \hat{y}}(g_i, 1), \quad i,j \in \mathbb{Z}_r. \tag{3.5}$$

Moreover,

$$\|\Pi_h^* v_h\|_{0,Q} \lesssim \|v_h\|_{0,Q}. \tag{3.6}$$

Now, for any $v_h \in S_h$, we multiply (3.7) by the constant $(\Pi_h^* v_h)(z)$, and then sum the corresponding results over Ω_h^* to obtain the following Petrov–Galerkin formulation of the semidiscrete scheme: Seek $u_h(t) \in S_h$, $0 < t \leqslant T$ such that

$$(u_{h,t}, \Pi_h^* v_h) + a_h(u_h, \Pi_h^* v_h) = (f, \Pi_h^* v_h), \quad v_h \in S_h, \tag{3.7}$$

where the bilinear form $a_h(\cdot,\cdot)$ is defined as follows: for any $u \in H_0^1(\Omega), v \in S_h^*$,

$$a_h(u, v) = -\sum_{z \in \mathcal{L}^0} v(z) \int_{\partial V_z} a \nabla u \cdot \mathbf{n}\, \mathrm{d}s. \tag{3.8}$$

Similarly, the fully discrete finite volume (3.9) can be written as follows: seeks $u_h^n \in S_h$, $n \geqslant 1$, such that

$$(\bar{\partial} u_h^n, \Pi_h^* v_h) + a_h(u_h^{n,1/2}, \Pi_h^* v_h) = (f^{n,1/2}, \Pi_h^* v_h), \quad v_h \in S_h. \tag{3.9}$$

REMARK 3.2 Since Π_h^* is a bijective operator, the Petrov–Galerkin formulations (3.7) and (3.9) are equivalent to the integral formulations (2.2) and (2.3), respectively.

In the following, we study the *quasi-symmetry* of the bilinear forms $(\cdot, \Pi_h^* \cdot)$ and $a_h(\cdot, \Pi_h^* \cdot)$.

To investigate the symmetry property of $(\cdot, \Pi_h^* \cdot)$, we denote by $(\cdot,\cdot)_Q$ the local inner product on any $Q \in \mathcal{T}_h$ as follows:

$$(v, w)_Q = \int_Q vw \, \mathrm{d}x \, \mathrm{d}y, \quad v, w \in L^2(Q).$$

By using the transformation F_Q, we have

$$(v, w)_Q = \int_{\hat{Q}} \hat{v}\hat{w} J_{F_Q} \, \mathrm{d}\hat{x} \, \mathrm{d}\hat{y}.$$

We denote $\bar{J}_{F_Q}$ as the average of J_{F_Q} on Q and set

$$\overline{(v, w)_Q} = \int_{\hat{Q}} \hat{v}\hat{w} \bar{J}_{F_Q} \, \mathrm{d}\hat{x} \, \mathrm{d}\hat{y}.$$

For a function $\hat{v}(\hat{x}, \hat{y}) \in L^2(\hat{Q})$, we define

$$\hat{v}_{\hat{x}}^{-1}(\hat{x}, \hat{y}) = \int_{-1}^{\hat{x}} \hat{v}(\hat{x}, \hat{y})\, \mathrm{d}\hat{x}, \quad \hat{v}_{\hat{y}}^{-1}(\hat{x}, \hat{y}) = \int_{-1}^{\hat{y}} \hat{v}(\hat{x}, \hat{y})\, \mathrm{d}\hat{y},$$

as the primitive functions of $\hat{v}$ along the $\hat{x}$- and $\hat{y}$-directions, respectively. We also define

$$\hat{v}^{-2}(\hat{x}, \hat{y}) = \int_{-1}^{\hat{x}} \int_{-1}^{\hat{y}} \hat{v}(\hat{x}, \hat{y})\, \mathrm{d}\hat{x}\, \mathrm{d}\hat{y}.$$

THEOREM 3.3 If the mesh $\mathcal{T}_h$ is shape-regular and an $h^{1+\gamma}$-parallelogram, then

$$|(u_h, \Pi_h^* v_h) - (v_h, \Pi_h^* u_h)| \lesssim h^\gamma \|u_h\|_0 \|v_h\|_0 \quad \forall u_h, v_h in S_h. \tag{3.10}$$

When $\gamma > 0$ and h is small enough, there holds

$$(u_h, \Pi_h^* u_h) \gtrsim \|u_h\|_0^2 \quad \forall u_h \in S_h. \tag{3.11}$$

Proof. Since the mesh is regular, we have (Arnold *et al.*, 2002)

$$\|\hat{u}_h\|_{0,\hat{Q}} \leqslant \|J_{F_Q^{-1}}\|_{\infty,Q}^{1/2} \|u_h\|_{0,Q} \lesssim h_Q^{-1} \|u_h\|_{0,Q}.$$

Let T_i be the triangle formed by two edges sharing the vertex P_i, where $\{P_i\}_{i=1}^4$ denote the vertices of Q, labelled in an anticlockwise sequence. Note that

$$J_{F_Q} = 2|T_1| + 2(|T_2| - |T_1|)\hat{x} + 2(|T_4| - |T_1|)\hat{y}.$$

Note that $|T_2| - |T_1| \lesssim h_Q^{2+\gamma}$, $|T_4| - |T_1| \lesssim h_Q^{2+\gamma}$ under the $h^{1+\gamma}$-parallelogram assumption. Therefore,

$$|J_{F_Q} - \bar{J}_{F_Q}| \lesssim h_Q^{2+\gamma}.$$

Therefore,

$$|(u_h, \Pi_h^* v_h)_Q - \overline{(u_h, \Pi_h^* v_h)}_Q| \lesssim h^\gamma \|u_h\|_{0,Q} \|\Pi_h^* v_h\|_{0,Q} \lesssim h^\gamma \|u_h\|_{0,Q} \|v_h\|_{0,Q}, \tag{3.12}$$

where (3.6) has been used in the second inequality.

Next, we will show that the bilinear form $\overline{(\cdot, \Pi_h^* \cdot)}_Q$ is symmetric. For any $u_h, v_h \in S_h$ and $(\hat{x}, \hat{y}) \in \hat{Q}$, let

$$\Upsilon(\hat{x}, \hat{y}) = \frac{\partial^2 \hat{v}_h}{\partial \hat{x} \partial \hat{y}}(\hat{x}, \hat{y}) \hat{u}_h^{-2}(\hat{x}, \hat{y}), \quad K(\hat{y}) = \int_{-1}^{1} \Upsilon(\hat{x}, \hat{y})\, \mathrm{d}\hat{x} - \sum_{i \in \mathbb{Z}_r} A_i \Upsilon(g_i, \hat{y}),$$

$$K_1(\hat{x}) = -\frac{\partial \hat{v}_h}{\partial \hat{x}}(\hat{x}, 1) \hat{u}_h^{-2}(\hat{x}, 1), \quad K_2(\hat{y}) = -\frac{\partial \hat{v}_h}{\partial \hat{y}}(1, \hat{y}) \hat{u}_h^{-2}(1, \hat{y}).$$

Noting $\Pi_h^* v_h \in S_h^*$, we regroup the sum to obtain

$$\begin{aligned}
\overline{(u_h, \Pi_h^* v_h)_Q}\bar{J}_{F_Q}^{-1} &= \sum_{(i,j)\in\mathbb{Z}_r^0\times\mathbb{Z}_r^0} (\Pi_h^* v_h)_{ij} \int_{g_i}^{g_{i+1}} \int_{g_j}^{g_{j+1}} \hat{u}_h \,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y} \\
&= \sum_{(i,j)\in\mathbb{Z}_r^0\times\mathbb{Z}_r^0} (\Pi_h^* v_h)_{ij} \{\hat{u}_h^{-2}(g_{i+1}, g_{j+1}) - \hat{u}_h^{-2}(g_{i+1}, g_j) - \hat{u}_h^{-2}(g_i, g_{j+1}) + \hat{u}_h^{-2}(g_i, g_j)\} \\
&= \sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r} \lfloor \Pi_h^* v_h \rfloor_{ij} \hat{u}_h^{-2}(g_i, g_j) + (\Pi_h^* v_h)_{rr} \hat{u}_h^{-2}(g_{r+1}, g_{r+1}) \\
&\quad - \sum_{j\in\mathbb{Z}_r} [\Pi_h^* v_h]_{\hat{x},rj} \hat{u}_h^{-2}(g_{r+1}, g_j) - \sum_{i\in\mathbb{Z}_r} [\Pi_h^* v_h]_{\hat{y},ir} \hat{u}_h^{-2}(g_i, g_{r+1}).
\end{aligned}$$

Then, by Lemma 3.1,

$$\overline{(u_h, \Pi_h^* v_h)_Q}\bar{J}_{F_Q}^{-1} = \sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r} A_i A_j \Upsilon(g_i, g_j) + (v_h \hat{u}_h^{-2})(g_{r+1}, g_{r+1}). + \sum_{i\in\mathbb{Z}_r} A_i K_1(g_i) + \sum_{j\in\mathbb{Z}_r} A_j K_2(g_j).$$

On the other hand, it follows from integration by parts that

$$\overline{(u_h, v_h)_Q}\bar{J}_{F_Q}^{-1} = \int_{-1}^{1}\int_{-1}^{1} \Upsilon \,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y} + (\hat{w}_h \hat{u}_h^{-2})(1,1) + \int_{-1}^{1} K_1(\hat{x})\,\mathrm{d}\hat{x} + \int_{-1}^{1} K_2(\hat{y})\,\mathrm{d}\hat{y}.$$

Noting $g_{r+1} = 1$, we have

$$\overline{(u_h, \Pi_h^* v_h)_Q}(\bar{J}_{F_Q})^{-1} - \overline{(u_h, v_h)_Q}(\bar{J}_{F_Q})^{-1} = T_1 + T_2 + T_3, \tag{3.13}$$

where

$$\begin{aligned}
T_1 &= \sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r} A_i A_j \Upsilon(g_i, g_j) - \int_{\hat{Q}} \Upsilon \,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}, \\
T_2 &= \sum_{i\in\mathbb{Z}_r} A_i K_1(g_i) - \int_{-1}^{1} K_1(\hat{x})\,\mathrm{d}\hat{x}, \\
T_3 &= \sum_{j\in\mathbb{Z}_r} A_j K_2(g_j) - \int_{-1}^{1} K_2(\hat{y})\,\mathrm{d}\hat{y}.
\end{aligned}$$

A straightforward calculation yields that

$$T_1 = T_{11} + T_{12} + T_{13},$$

where

$$T_{11} = \int_{-1}^{1} K_1(\hat{x})\,\mathrm{d}\hat{x} - \sum_{i\in\mathbb{Z}_r} A_i K_1(g_i),$$

$$T_{12} = \int_{-1}^{1} K_2(\hat{y})\,\mathrm{d}\hat{y} - \sum_{j\in\mathbb{Z}_r} A_j K_2(g_j),$$

$$T_{13} = \int_{-1}^{1} K(\hat{y})\,\mathrm{d}\hat{y} - \sum_{j\in\mathbb{Z}_r} A_j K(g_j).$$

It follows from the residual formula (3.3) and integration by parts that

$$T_{13} = C_r^2 \frac{\partial^{2r}\hat{u}_h}{\partial\hat{x}^r\partial\hat{y}^r}\frac{\partial^{2r}\hat{v}_h}{\partial\hat{x}^r\partial\hat{y}^r}$$

and

$$T_{11} = C_r \int_{-1}^{1} \frac{\partial^r\hat{u}_h}{\partial\hat{x}^r}\frac{\partial^r\hat{v}_h}{\partial\hat{x}^r}\,\mathrm{d}\hat{y} - T_2, \quad T_{12} = C_r \int_{-1}^{1} \frac{\partial^r\hat{u}_h}{\partial\hat{y}^r}\frac{\partial^r\hat{v}_h}{\partial\hat{y}^r}\,\mathrm{d}\hat{x} - T_3.$$

Consequently,

$$T_1 + T_2 + T_3 = C_r \int_{-1}^{1} \frac{\partial^r\hat{u}_h}{\partial\hat{x}^r}\frac{\partial^r\hat{v}_h}{\partial\hat{x}^r}\,\mathrm{d}\hat{y} + C_r \int_{-1}^{1} \frac{\partial^r\hat{u}_h}{\partial\hat{y}^r}\frac{\partial^r\hat{v}_h}{\partial\hat{y}^r}\,\mathrm{d}\hat{x} + C_r^2 \frac{\partial^{2r}\hat{u}_h}{\partial\hat{x}^r\partial\hat{y}^r}\frac{\partial^{2r}\hat{v}_h}{\partial\hat{x}^r\partial\hat{y}^r} \tag{3.14}$$

is symmetric with respect to u_h and v_h. By (3.13), $\overline{(u_h, \Pi_h^* v_h)_Q}$ is also symmetric with respect to u_h and v_h. Consequently, (3.10) follows from (3.12).

Finally, it is trivial from (3.14) that $T_1 + T_2 + T_3 \geqslant 0$ when $v_h = u_h$. Then, by (3.12–3.14),

$$\begin{aligned}(u_h, \Pi_h^* u_h)_Q &\gtrsim \overline{(u_h, \Pi_h^* u_h)_Q} - h^\gamma \|u_h\|_{0,Q}^2 \\ &\gtrsim \overline{(u_h, u_h)_Q} - h^\gamma \|u_h\|_{0,Q}^2 \\ &\gtrsim \|u_h\|_{0,Q}^2 - h^\gamma \|u_h\|_{0,Q}^2.\end{aligned}$$

The desired result (3.11) follows for sufficiently small h. □

Let $\mathcal{E}_h^*$ denote the set of edges of the dual partition $\mathcal{T}_h^*$. For all $u \in H_0^1(\Omega)$ and $v \in S_h^*$, we write

$$a_h(u,v) = \sum_{Q\in\mathcal{T}_h} a_Q(u,v),$$

where the elementwise bilinear form is defined as

$$a_Q(u,v) = \sum_{e\in\mathcal{E}_h^*\cap Q} [v]_e \int_e a\nabla u\cdot\mathbf{n}\,\mathrm{d}s.$$

Note that $[v]_e = v(z_2) - v(z_1)$ denotes the jump of v across the common edge $e = \partial V_{z_2} \cap \partial V_{z_1}$ with $z_1, z_2 \in \mathcal{L}$, and $\mathbf{n}$ denotes the outer normal on ∂V_{z_1}. For any $(x, y) \in Q$, we define

$$(\partial_x^{-1} u)(x,y) = \int_{F_Q(-1,\hat{y})(x,y)} a\nabla u \cdot \mathbf{n}\,\mathrm{d}s, \quad (\partial_y^{-1} u)(x,y) = \int_{F_Q(\hat{x},-1)(x,y)} a\nabla u \cdot \mathbf{n}\,\mathrm{d}s.$$

With these notations, the elementwise bilinear form $a_Q(u, v)$ can be rewritten as

$$\begin{aligned} a_Q(u,v) &= \sum_{(i,j)\in\mathbb{Z}_r^0\times\mathbb{Z}_r} [v]_{\hat{x},ij} \int_{\mathbf{g}_{ij}\mathbf{g}_{i+1j}} a\nabla u \cdot \mathbf{n}\,\mathrm{d}s + \sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r^0} [v]_{\hat{y},ij} \int_{\mathbf{g}_{ij}\mathbf{g}_{ij+1}} a\nabla u \cdot \mathbf{n}\,\mathrm{d}s \\ &= -\sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r} \lfloor v \rfloor_{ij} (\partial_x^{-1} u)(\mathbf{g}_{ij}) + \sum_{j\in\mathbb{Z}_r} [v]_{\hat{x},rj} (\partial_x^{-1} u)(\mathbf{g}_{r+1j}) \\ &\quad - \sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r} \lfloor v \rfloor_{ij} (\partial_y^{-1} u)(\mathbf{g}_{ij}) + \sum_{i\in\mathbb{Z}_r} [v]_{\hat{y},ir} (\partial_y^{-1} u)(\mathbf{g}_{ir+1}) \\ &:= a_Q^1(u,v) + a_Q^2(u,v). \end{aligned}$$

Applying the transformation F_Q, we have

$$\begin{aligned} \partial_x^{-1} u = \partial_{\hat{x}}^{-1} \hat{u} &= \int_{-1}^{\hat{x}} \hat{a} \left(b_{12} \frac{\partial \hat{u}}{\partial \hat{x}} + b_{11} \frac{\partial \hat{u}}{\partial \hat{y}} \right) \mathrm{d}\hat{x}, \\ \partial_y^{-1} u = \partial_{\hat{y}}^{-1} \hat{u} &= \int_{-1}^{\hat{y}} \hat{a} \left(b_{22} \frac{\partial \hat{u}}{\partial \hat{x}} + b_{21} \frac{\partial \hat{u}}{\partial \hat{y}} \right) \mathrm{d}\hat{y}, \end{aligned}$$

where

$$\begin{aligned} b_{11} &= J_{F_Q}^{-1} \left[\left(\frac{\partial x}{\partial \hat{x}} \right)^2 + \left(\frac{\partial y}{\partial \hat{x}} \right)^2 \right], \quad b_{22} = J_{F_Q}^{-1} \left[\left(\frac{\partial x}{\partial \hat{y}} \right)^2 + \left(\frac{\partial y}{\partial \hat{y}} \right)^2 \right], \\ b_{12} &= b_{21} = -J_{F_Q}^{-1} \left[\frac{\partial x}{\partial \hat{x}} \frac{\partial x}{\partial \hat{y}} + \frac{\partial y}{\partial \hat{x}} \frac{\partial y}{\partial \hat{y}} \right]. \end{aligned}$$

Since Q is a regular $h^{1+\gamma}$-parallelogram, we have $D^\alpha J_{F_Q} = \mathcal{O}(h_Q^{2+|\alpha|\gamma})$. Differentiating the identity $J_{F_Q} J_{F_Q^{-1}} = 1$ and using mathematical induction, we have $D^\alpha J_{F_Q^{-1}} = \mathcal{O}(h_Q^{-2+|\alpha|\gamma})$. Then, the estimate

$$|D^\alpha b_{ij}| \lesssim h_Q^{|\alpha|\gamma}, \quad \alpha \geqslant 0,\ i,j = 1,2 \tag{3.15}$$

follows from the Leibniz rule (Zlámal, 1978), where γ is the parameter in the $h^{1+\gamma}$ mesh assumption.

In view of Lemma 3.1, we have that, for $u \in H_0^1(\Omega)$ and $v_h \in S_h$,

$$a_Q^1(u, \Pi_h^* v_h) = -\sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r} A_i A_j \left(\frac{\partial^2 \hat{v}_h}{\partial \hat{x} \partial \hat{y}} \partial_{\hat{x}}^{-1} \hat{u} \right) (g_i, g_j) + \sum_{j\in\mathbb{Z}_r} A_j \left(\frac{\partial \hat{v}_h}{\partial \hat{y}} \partial_{\hat{x}}^{-1} \hat{u} \right) (1, g_j),$$

and

$$a_Q^2(u,\Pi_h^* v_h) = -\sum_{(i,j)\in\mathbb{Z}_r\times\mathbb{Z}_r} A_i A_j \left(\frac{\partial^2 \hat{v}_h}{\partial\hat{x}\partial\hat{y}}\partial_{\hat{y}}^{-1}\hat{u}\right)(g_i,g_j) + \sum_{i\in\mathbb{Z}_r} A_i \left(\frac{\partial \hat{v}_h}{\partial\hat{x}}\partial_{\hat{y}}^{-1}\hat{u}\right)(g_i,1).$$

We note that $a_Q^1(\cdot,\Pi_h^*\cdot)$ and $a_Q^2(\cdot,\Pi_h^*\cdot)$ are, respectively, the Gauss quadratures of the following bilinear forms:

$$\tilde{a}_Q^1(u,v_h) = -\int_{\hat{Q}} \frac{\partial^2 \hat{v}_h}{\partial\hat{x}\partial\hat{y}}\partial_{\hat{x}}^{-1}\hat{u}\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y} + \int_{-1}^{1}\left(\frac{\partial \hat{v}_h}{\partial\hat{y}}\partial_{\hat{x}}^{-1}\hat{u}\right)(1,\hat{y})\,\mathrm{d}\hat{y} \tag{3.16}$$

and

$$\tilde{a}_Q^2(u,v_h) = -\int_{\hat{Q}} \frac{\partial^2 \hat{v}_h}{\partial\hat{x}\partial\hat{y}}\partial_{\hat{y}}^{-1}\hat{u}\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y} + \int_{-1}^{1}\left(\frac{\partial \hat{v}_h}{\partial\hat{x}}\partial_{\hat{y}}^{-1}\hat{u}\right)(\hat{x},1)\,\mathrm{d}\hat{x}. \tag{3.17}$$

Integrating by parts for $\tilde{a}_Q^1(\cdot,\cdot)$ and $\tilde{a}_Q^2(\cdot,\cdot)$, we know that

$$\begin{aligned}\tilde{a}_Q(u,v_h) &= \tilde{a}_Q^1(u,v_h) + \tilde{a}_Q^2(u,v_h)\\ &= \int_{\hat{Q}} \hat{a}\left(b_{11}\frac{\partial\hat{u}}{\partial\hat{y}}\frac{\partial\hat{v}_h}{\partial\hat{y}} + b_{22}\frac{\partial\hat{u}}{\partial\hat{x}}\frac{\partial\hat{v}_h}{\partial\hat{x}} + b_{12}\frac{\partial\hat{u}}{\partial\hat{x}}\frac{\partial\hat{v}_h}{\partial\hat{y}} + b_{12}\frac{\partial\hat{u}}{\partial\hat{y}}\frac{\partial\hat{v}_h}{\partial\hat{x}}\right)\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}\\ &= \int_Q a\nabla u\cdot\nabla v_h\,\mathrm{d}x\,\mathrm{d}y \qquad (3.18)\end{aligned}$$

is a symmetric bilinear form.

In the following, we analyse the difference between $a_Q(\cdot,\Pi_h^*\cdot)$ and $\tilde{a}_Q(\cdot,\cdot)$. To this end, we set

$$\begin{aligned}\Phi_j(\hat{x}) &= \frac{\partial^2\hat{v}_h}{\partial\hat{x}\partial\hat{y}}(\hat{x},g_j)(\partial_{\hat{x}}^{-1}\hat{u}_h)(\hat{x},g_j) \quad \forall j\in\mathbb{Z}_r;\\ \Psi(\hat{x},\hat{y}) &= \hat{a}\frac{\partial\hat{v}_h}{\partial\hat{y}}\left(\frac{\partial\hat{u}_h}{\partial\hat{x}}b_{12} + \frac{\partial\hat{u}_h}{\partial\hat{y}}b_{11}\right).\end{aligned}$$

We also define the quadrature error as follows:

$$E_{1,j}^Q(u,v_h) = \int_{-1}^{1}\Phi_j(\hat{x})\,\mathrm{d}\hat{x} - \sum_{i\in\mathbb{Z}_r} A_i\Phi_j(g_i) \quad \forall j\in\mathbb{Z}_r, \tag{3.19}$$

$$E_{2,\hat{x}}^Q(u,v_h) = \int_{-1}^{1}\Psi(\hat{x},\hat{y})\,\mathrm{d}\hat{y} - \sum_{j\in\mathbb{Z}_r} A_j\Psi(\hat{x},g_j). \tag{3.20}$$

LEMMA 3.4 If $Q\in\mathcal{T}_h$ is a regular $h^{1+\gamma}$-parallelogram and $a(\mathbf{x})$ is a constant in Q, then

$$|E_{1,j}^Q(u_h,v_h) - E_{1,j}^Q(v_h,u_h)| \lesssim h^\gamma |u_h|_{1,Q}|v_h|_{1,Q} \quad \forall j\in\mathbb{Z}_r, \tag{3.21}$$

$$\left|\int_{-1}^{1}(E_{2,\hat{x}}^Q(u_h,v_h) - E_{2,\hat{x}}^Q(v_h,u_h))\,\mathrm{d}\hat{x}\right| \lesssim h^\gamma |u_h|_{1,Q}|v_h|_{1,Q}. \tag{3.22}$$

Proof. We prove (3.21) only, since (3.22) can be established by similar arguments. By using the Gauss quadrature error (3.3), we have

$$E^{Q}_{1,j}(u_h, v_h) = C_r \Phi^{(2r)}_j(\zeta), \quad \zeta \in (-1, 1).$$

Note that

$$\begin{aligned}\Phi^{(2r)}_j(\zeta) &= a\sum_{k=0}^{r-1}\binom{2r}{k}\frac{\partial^{k+2}\hat{v}_h}{\partial\hat{x}^{k+1}\partial\hat{y}}(\zeta, g_j)\{(\hat{u}_{h,\hat{x}}b_{12} + \hat{u}_{h,\hat{y}}b_{11})|_{\hat{y}=g_j}\}^{(2r-k-1)}(\zeta)\\ &= T_1 + T_2 + T_3,\end{aligned}$$

where

$$\begin{aligned}T_1 &= a\binom{2r}{r-1}\frac{\partial^{r+1}\hat{v}_h}{\partial\hat{x}^r\partial\hat{y}}(\zeta, g_j)\frac{\partial^r\hat{u}_h}{\partial\hat{x}^r}(\zeta, g_j)\frac{\partial b_{12}}{\partial\hat{x}}(\zeta, g_j),\\ T_2 &= a\binom{2r}{r-1}\frac{\partial^{r+1}\hat{v}_h}{\partial\hat{x}^r\partial\hat{y}}(\zeta, g_j)\frac{\partial^{r+1}\hat{u}_h}{\partial\hat{x}^r\partial\hat{y}}(\zeta, g_j)b_{11}(\zeta, g_j),\\ T_3 &= a\sum_{k=0}^{r-2}\binom{2r}{k}\frac{\partial^{k+2}\hat{v}_h}{\partial\hat{x}^{k+1}\partial\hat{y}}(\zeta, g_j)\left\{(\hat{u}_{h,\hat{x}}b_{12} + \hat{u}_{h,\hat{y}}b_{11})|_{\hat{y}=g_j}\right\}^{(2r-k-1)}(\zeta).\end{aligned}$$

By (3.15) and a scaling argument,

$$|T_1| \lesssim h^{\gamma}|\hat{u}_h|_{1,\hat{Q}}|\hat{v}_h|_{1,\hat{Q}}.$$

For T_3, the Leibnitz rule implies that

$$\begin{aligned}&\{(\hat{u}_{h,\hat{x}}b_{12} + \hat{u}_{h,\hat{y}}b_{11})|_{\hat{y}=g_j}\}^{(2r-k-1)}(\zeta)\\ &\quad= \sum_{l=0}^{2r-k-1}\binom{2r-k-1}{l}\{\hat{u}_{h,\hat{x}}|^{(l)}_{\hat{y}=g_j}b_{12}|^{(2r-k-1-l)}_{\hat{y}=g_j} + \hat{u}_{h,\hat{y}}|^{(l)}_{\hat{y}=g_j}b_{11}|^{(2r-k-1-l)}_{\hat{y}=g_j}\}(\zeta).\end{aligned}$$

Since $2r-k-1 \geqslant r+1$ when $0 \leqslant k \leqslant r-2$, both b_{11} and b_{12} will be differentiated at least once. By a scaling argument and (3.15), we have

$$|T_3| \lesssim h^{\gamma}|\hat{u}_h|_{1,\hat{Q}}|\hat{v}_h|_{1,\hat{Q}}.$$

Combining the estimates for T_1, T_3 and noting the fact that T_2 is symmetric with respect to u_h and v_h, we obtain

$$|E^{Q}_{1,j}(u_h, v_h) - E^{Q}_{1,j}(v_h, u_h)| \lesssim h^{\gamma}|\hat{u}_h|_{1,\hat{Q}}|\hat{v}_h|_{1,\hat{Q}}.$$

Since the mesh is shape-regular, a scaling argument gives $|\hat{v}_h|_{1,\hat{Q}} \lesssim |v_h|_{1,Q}$ for $v_h \in S_h$. From the above two estimates, we obtain the desired estimate (3.21). □

We are now ready to estimate the symmetry of $a_h(\cdot, \Pi^*_h\cdot)$.

THEOREM 3.5 Assume that the mesh $\mathcal{T}_h$ is shape-regular and an $h^{1+\gamma}$-parallelogram, and the coefficient $a(\mathbf{x})$ is piecewise continuous with respect to $\mathcal{T}_h$. Then,

$$|a_h(u_h, \Pi_h^* v_h) - a_h(v_h, \Pi_h^* u_h)| \lesssim (m(a,h)_{\mathcal{T}_h} + h^\gamma)|u_h|_1|v_h|_1, \tag{3.23}$$

where the piecewise modulus of continuity is defined as

$$m_{\mathcal{T}_h}(a,h) = \sup\{|a(\mathbf{x}_1) - a(\mathbf{x}_2)| : \mathbf{x}_1, \mathbf{x}_2 \in Q, \forall Q \in \mathcal{T}_h\}.$$

Proof. Integrating by parts, we obtain

$$a_Q^1(u_h, \Pi_h^* v_h) - \tilde{a}_Q^1(u_h, v_h) = \sum_{j\in\mathbb{Z}_r} A_j E_{1,j}^Q(u_h, v_h) - \int_{-1}^{1} E_{2,\hat{x}}^Q(u_h, v_h)\,\mathrm{d}\hat{x}. \tag{3.24}$$

The difference between $a_Q^2(u_h, \Pi_h^* v_h)$ and $\tilde{a}_Q^2(u_h, v_h)$ takes a similar form. When $a(\mathbf{x})$ is piecewise constant with respect to Ω, $m_{\mathcal{T}_h}(a,h) = 0$, (3.23) is a direct consequence of the estimates in Lemma 3.4 and the symmetry of $\tilde{a}_Q(\cdot,\cdot)$. When $a(\mathbf{x})$ is only piecewise continuous with respect to $\mathcal{T}_h$, (3.23) is a consequence of the result in the piecewise constant case and the continuity property (4.1). □

REMARK 3.6 If a is piecewise Lipschtiz continuous with respect to $\mathcal{T}_h$, $m(a,h) \lesssim h$, then

$$|a_h(u_h, \Pi_h^* v_h) - a_h(v_h, \Pi_h^* u_h)| \lesssim h^{\min\{1,\gamma\}}|u_h|_1|v_h|_1. \tag{3.25}$$

REMARK 3.7 If $\mathcal{T}_h$ is a parallelogram mesh, $(u_h, \Pi_h^* v_h)$ is symmetric. If $\mathcal{T}_h$ is a parallelogram mesh and the coefficient a is piecewise constant with respect to $\mathcal{T}_h$, $a_h(u_h, \Pi_h^* v_h)$ is symmetric.

4. Error estimates

We begin this section by investigating well-posedness of the finite volume schemes (2.2) and (2.3). We list the following properties of $a_h(\cdot, \Pi_h^* \cdot)$ that have been proved in Zhang & Zou (2015).

LEMMA 4.1 For any $v_h, w_h \in S_h$, there holds

$$|a_h(v_h, \Pi_h^* w_h)| \lesssim |v_h|_1|w_h|_1. \tag{4.1}$$

Furthermore, when the mesh parameter $\gamma > 0$, coercivity holds as shown below:

$$a_h(v_h, \Pi_h^* v_h) \gtrsim |v_h|_1^2. \tag{4.2}$$

Let $\{\Phi_z : z \in Z_h^0\}$ and $\{\Psi_z : z \in Z_h^0\}$ be the standard basis functions of S_h and S_h^*, respectively. Then, the semidiscrete scheme (2.2) can be rewritten as a system of ordinary differential equations

$$\mathcal{M}\alpha'(t) + \mathcal{S}\alpha(t) = \tilde{f}(t), \quad 0 \leqslant t \leqslant T; \quad \alpha(0) = \beta, \tag{4.3}$$

where $\mathcal{M} = ((\Phi_z, \Psi_w))_{zw}$ and $\mathcal{S} = (a_h(\Phi_z, \Psi_w))_{zw}$ are the mass and stiffness matrices, respectively, and $\alpha(t)$ and β are vectors of the nodal values of $u_h(t)$ and $R_h u_0$, respectively. Let $\mathcal{T}$ denote the transfer matrix determined by Π_h^*. From (3.11) and (4.2), we know that both $\mathcal{T}\mathcal{M}$ and $\mathcal{T}\mathcal{S}$ are invertible. We conclude that $\mathcal{M}$ and $\mathcal{S}$ are also invertible. This implies that there exists a unique solution $u_h(\cdot,t)$ on $\Omega \times [0,T]$.

The fully discrete scheme (2.3) takes a matrix form as

$$(\mathcal{M} + \tfrac{1}{2}\Delta t\mathcal{S})\alpha^n = (\mathcal{M} - \tfrac{1}{2}\Delta t\mathcal{S})\alpha^{n-1} + \Delta t\tilde{f}^{n,1/2}. \tag{4.4}$$

By (3.11) and (4.2) again, we know that both $\mathcal{T}\mathcal{M} + (\mathcal{T}\mathcal{M})^{\mathrm{T}}$ and $\mathcal{T}\mathcal{S} + (\mathcal{T}\mathcal{S})^{\mathrm{T}}$ are positive-definite. So, for any nonzero vector $\mathbf{x}$,

$$\mathbf{x}^{\mathrm{T}}(\mathcal{T}\mathcal{M} + \tfrac{1}{2}\Delta t\mathcal{T}\mathcal{S})\mathbf{x} = \tfrac{1}{2}\mathbf{x}^{\mathrm{T}}(\mathcal{T}\mathcal{M} + (\mathcal{T}\mathcal{M})^{\mathrm{T}})\mathbf{x} + \tfrac{1}{4}\Delta t\mathbf{x}^{\mathrm{T}}(\mathcal{T}\mathcal{S} + (\mathcal{T}\mathcal{S})^{\mathrm{T}})\mathbf{x} > 0,$$

which means $\mathcal{T}\mathcal{M} + \frac{1}{2}\Delta t\mathcal{T}\mathcal{S}$ is invertible. Hence, $\mathcal{M} + \frac{1}{2}\Delta t\mathcal{S}$ is also invertible. Therefore, the fully discrete scheme can be solved uniquely at each time step.

4.1 *L^2-error estimate of the elliptic projection*

The initial approximations should be determined by the following elliptic projection R_h:

$$a_h(R_h u, v) = a_h(u, v) \quad \forall v \in S_h^*. \tag{4.5}$$

It has been proved in Zhang & Zou (2015) that

$$|u - R_h u|_1 \lesssim h^r \|u\|_{r+1}, \tag{4.6}$$

when the '$h^{1+\gamma}$ mesh assumption holds with $\gamma > 0$.

To derive the error estimates, we shall further study the L^2- error of the Ritz projection $R_h u$. Consider an auxiliary problem

$$-\nabla \cdot (a(\mathbf{x})\nabla w) = u - R_h u, \quad w \in H_0^1(\Omega) \cap H^2(\Omega). \tag{4.7}$$

It is known that

$$\|w\|_2 \lesssim \|u - R_h u\|_0. \tag{4.8}$$

Let $w_h = I_h^1 w \in S_h$ be a bilinear Lagrange interpolation of w, which satisfies

$$|w - I_h^1 w|_s \lesssim h^{2-s}|w|_2, \quad 0 \leqslant s \leqslant 2. \tag{4.9}$$

Testing (4.7) by $u - R_h u$ and using the fact that $a_h(u - R_h u, \Pi_h^* v_h) = 0, \forall v_h \in S_h$, we have

$$\|u - R_h u\|_0^2 = a(u - R_h u, w - I_h^1 w) + [a(u - R_h u, I_h^1 w) - a_h(u - R_h u, \Pi_h^*(I_h^1 w))]. \tag{4.10}$$

Here,

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x}$$

denotes a standard finite element bilinear form. It is obvious that the first term is bounded by $h|u - R_h u|_1 \|w\|_2$. The second term can be estimated by applying the following lemma.

Lemma 4.2 Let $w_h = I_h^1 w \in S_h$ be the bilinear Lagrange interpolation of w. Then, for any $u \in H_0^1(\Omega) \cap H^{m+1}(\Omega)$, $r \leqslant m \leqslant 2r$ and $u_h \in S_h$, we have

$$|a_h(u - u_h, \Pi_h^* w_h) - a(u - u_h, w_h)| \lesssim h^{\min\{1,(m-r)\gamma\}}(|u - u_h|_1 + h^r\|u\|_{m+1})\|w_h\|_2. \tag{4.11}$$

Proof. First, we assume that the coefficient a is piecewise constant with respect to Ω. By (3.18) and (3.24), for any $v \in H_0^1(\Omega)$,

$$a_h(v, \Pi_h^* w_h) - a(v, w_h) = \sum_{Q \in \mathcal{T}_h} \sum_{i=1}^{2} a_Q^i(v, \Pi_h^* w_h) - \tilde{a}_Q^i(v, w_h), \tag{4.12}$$

where

$$a_Q^1(v, \Pi_h^* w_h) - \tilde{a}_Q^1(v, w_h) = \sum_{j \in \mathbb{Z}_r} A_j E_{1,j}^Q(v, w_h) - \int_{-1}^{1} E_2^Q(v, w_h)\, \mathrm{d}\hat{x},$$

and $a_Q^2(v, \Pi_h^* w_h) - \tilde{a}_Q^2(v, w_h)$ takes a similar form. In view of (3.19), $E_{1,j}^Q(v, w_h)$ is a Gauss quadrature residual, that can be estimated as Davis & Rabinowitz (1984)

$$|E_{1,j}^Q(v, w_h)| \lesssim \int_{-1}^{1} \left| \frac{\mathrm{d}^m \Phi_j(v, w_h)}{\mathrm{d}\hat{x}^m} \right| \mathrm{d}\hat{x}, \quad r \leqslant m \leqslant 2r.$$

An application of the Leibnitz rule together with (3.15) yields

$$\begin{aligned}
\left| \frac{\mathrm{d}^m \Phi_j(v, w_h)}{\mathrm{d}\hat{x}^m} \right| &= |a| \left| \frac{\partial^2 \hat{w}_h}{\partial \hat{x} \partial \hat{y}} \left\{ (\hat{v}_{\hat{x}} b_{12} + \hat{v}_{\hat{y}} b_{11})|_{\hat{y}=g_j} \right\}^{(m-1)} \right| \\
&\lesssim \left| \frac{\partial^2 \hat{w}_h}{\partial \hat{x} \partial \hat{y}} \right| \sum_{l=0}^{m-1} \left\{ \left| \frac{\partial^{l+1} \hat{v}}{\partial \hat{x}^{l+1}} \right| + \left| \frac{\partial^{l+1} \hat{v}}{\partial \hat{x}^l \hat{y}} \right| \right\} |_{\hat{y}=g_j} h^{(m-1-l)\gamma} \\
&\lesssim \left| \frac{\partial^2 \hat{w}_h}{\partial \hat{x} \partial \hat{y}} \right| \sum_{l=0}^{m-1} \left\{ \left| \frac{\partial^{l+1} \hat{v}}{\partial \hat{x}^{l+1}} \right| + \left| \frac{\partial^{l+1} \hat{v}}{\partial \hat{x}^l \hat{y}} \right| \right\} |_{\hat{y}=g_j}.
\end{aligned}$$

By taking $m = r$, we deduce from the above two estimates that

$$\left| \sum_{j \in \mathbb{Z}_r} A_j E_{1,j}^Q(v, w_h) \right| \lesssim \left\| \frac{\partial^2 \hat{w}_h}{\partial \hat{x} \partial \hat{y}} \right\|_{0,\hat{Q}} \sum_{l=0}^{r-1} (|\hat{v}|_{l+1,\hat{Q}} + |\hat{v}|_{l+2,\hat{Q}}), \tag{4.13}$$

where a trace theorem has been used. Similarly, by (3.20),

$$\begin{aligned}
\int_{-1}^{1} |E_2^Q(v, w_h)|\, \mathrm{d}\hat{x} &\lesssim \int_{-1}^{1} \int_{-1}^{1} \left| \frac{\partial^m \Phi_j(v, w_h)}{\partial \hat{y}^m} \right| \mathrm{d}\hat{y}\, \mathrm{d}\hat{x} \\
&\lesssim \left\| \frac{\partial \hat{w}_h}{\partial \hat{y}} \right\|_{0,\hat{Q}} \sum_{l=0}^{m} \left\{ \left\| \frac{\partial^{l+1} \hat{v}}{\partial \hat{x} \partial \hat{y}^l} \right\|_{0,\hat{Q}} + \left\| \frac{\partial^{l+1} \hat{v}}{\partial \hat{y}^{l+1}} \right\|_{0,\hat{Q}} \right\} h^{(m-l)\gamma}, \quad r \leqslant m \leqslant 2r.
\end{aligned} \tag{4.14}$$

Let $I_h u \in S_h$ be the standard Lagrange interpolation of u, which satisfies

$$|u - I_h u|_s \lesssim h^{r+1-s} |w|_{r+1}, \quad 0 \leqslant s \leqslant r+1. \tag{4.15}$$

If we set $v = I_h u - u_h$ in (4.13) and (4.14), then an inverse estimate and a scaling argument combined yield

$$\begin{aligned}
&\left|\sum_{j\in\mathbb{Z}_r} A_j E^{Q}_{1,j}(I_h u - u_h, w_h)\right| + \int_{-1}^{1} |E^{Q}_{2,\hat{x}}(I_h u - u_h, w_h)|\,\mathrm{d}\hat{x} \\
&\quad \lesssim |\hat{w}_h|_{2,\hat{Q}} |\widehat{I_h u - u_h}|_{1,\hat{Q}} + |\hat{w}_h|_{1,\hat{Q}} |\widehat{I_h u - u_h}|_{1,\hat{Q}} \sum_{l=0}^{r} h^{(m-l)\gamma} \\
&\quad \lesssim h\|w_h\|_{2,Q} |I_h u - u_h|_{1,Q} + |w_h|_{1,Q} |I_h u - u_h|_{1,Q} h^{(m-r)\gamma}, \quad r \leqslant m \leqslant 2r.
\end{aligned}$$

If we set $v = u - I_h u$ in (4.13) and (4.14), then a scaling argument and the interpolation estimate (4.15) combined lead to

$$\begin{aligned}
&\left|\sum_{j\in\mathbb{Z}_r} A_j E^{Q}_{1,j}(u - I_h u, w_h)\right| + \int_{-1}^{1} |E^{Q}_{2,\hat{x}}(u - I_h u, w_h)|\,\mathrm{d}\hat{x} \\
&\lesssim h^{r+1}\|w_h\|_{2,Q}\|u\|_{r+1,Q} + |w_h|_{1,Q}\left(h^r\|u\|_{r+1,Q}\sum_{l=0}^{r-1} h^{(m-l)\gamma} + \sum_{l=r}^{m} h^{l+(m-l)\gamma}\|u\|_{l+1,Q}\right).
\end{aligned}$$

It follows from the above two estimates and a triangle inequality that

$$\begin{aligned}
&|a^1_Q(u-u_h, \Pi^*_h w_h) - \tilde{a}^1_Q(u-u_h, w_h)| \\
&\quad \lesssim (h^{\min\{1,(m-r)\gamma\}}|u-u_h|_{1,Q} + h^{r+1}\|u\|_{r+1,Q} + h^{r+(m-r)\gamma}\|u\|_{m+1,Q})\|w_h\|_{2,Q}.
\end{aligned}$$

The term $a^2_Q(u-u_h, \Pi^*_h w_h) - \tilde{a}^2_Q(u-u_h, w_h)$ can be similarly estimated. Then, a summation over $\mathcal{T}_h$ yields the desired result.

If $a(x)$ is not piecewise constant, we can have a perturbation argument by taking the piecewise constant approximation in each element Q and the result will still hold. □

Note that $\|I^1_h w\|_2 \leqslant \|w - I^1_h w\|_2 + \|w\|_2 \lesssim \|w\|_2$. The following lemma is a direct consequence of (4.10), Lemma 4.2 and (4.6).

LEMMA 4.3 Let $R_h u$ be the elliptic projection defined as in (4.5). Then,

$$\|u - R_h u\|_0 \lesssim h^{r+\min\{1,(m-r)\gamma\}}\|u\|_{m+1}, \quad r \leqslant m \leqslant 2r. \tag{4.16}$$

Consequently, if $(m-r)\gamma \geqslant 1$, then we have the following optimal-order L^2-error:

$$\|u - R_h u\|_0 \lesssim h^{r+1}\|u\|_{m+1}. \tag{4.17}$$

REMARK 4.4 The same optimal-order L^2-estimate for the special case $r = 2$ has been recently obtained by Lv & Li (2012) under the condition $\gamma = 1$ and $m = r + 1$.

4.2 *Analysis of the semidiscrete finite volume scheme*

Since the finite volume bilinear forms are usually not symmetric, some additional terms reflecting the nonsymmetry will appear in the error equation. More specifically, the terms look like $(w_h, \Pi_h^* v_h) - (v_h, \Pi_h^* w_h)$ or $a_h(w_h, \Pi_h^* v_h) - a_h(v_h, \Pi_h^* w_h)$. Now, owing to the quasi-symmetry properties (3.10) and (3.25), these terms can be well controlled in the convergence analysis.

THEOREM 4.5 Let u be the solution of (1.1) and u_h be the numerical solution of the semidiscrete finite volume scheme (2.2). Assume that the mesh parameter satisfies $\gamma \geqslant 2/3$. If $u \in L^\infty(0,T; H^{r+1}) \cap H^1(0,T; H^{m+1})$, $r \leqslant m \leqslant 2r$, then, for $0 \leqslant t \leqslant T$,

$$|u(t) - u_h(t)|_1 \lesssim h^r + h^{r+\gamma-1+\min\{1,(m-r)\gamma\}}. \tag{4.18}$$

If $u \in L^\infty(0,T; H^{\tilde{m}+1}) \cap H^1(0,T; H^{m+1})$, $r+1 \leqslant m, \tilde{m} \leqslant 2r$, then, for $0 \leqslant t \leqslant T$,

$$\|u(t) - u_h(t)\|_0 \lesssim h^{r+\min\{1,(\tilde{m}-r)\gamma\}} + (1 + h^{\gamma-1}) h^{r+\min\{1,(m-r)\gamma\}}. \tag{4.19}$$

Proof. We decompose the error as $u_h - u = \xi - \eta$, where $\xi = u_h - R_h u$ and $\eta = u - R_h u$. From (2.1) and (3.7), we have the following error equation:

$$(\xi_t, \Pi_h^* v_h) + a_h(\xi, \Pi_h^* v_h) = (\eta_t, \Pi_h^* v_h), \quad v_h \in S_h. \tag{4.20}$$

We take $v = \xi$ in (2.1) and use (3.6) and (3.10) to obtain

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\xi, \Pi_h^*\xi) + a_h(\xi, \Pi_h^*\xi) &= \frac{1}{2}((\xi, \Pi_h^*\xi_t) - (\xi_t, \Pi_h^*\xi)) + (\eta_t, \Pi_h^*\xi) \\
&\lesssim h^\gamma \|\xi\|_0 \|\xi_t\|_0 + \|\eta_t\|_0 \|\Pi_h^*\xi\|_0 \\
&\lesssim h^\gamma \|\xi\|_0 \|\xi_t\|_0 + \|\eta_t\|_0 \|\xi\|_0.
\end{aligned}$$

Integrating the above estimate on $[0,t]$, noting that $\xi(0) = 0$, and using coercivity of the bilinear forms, we have

$$\begin{aligned}
\|\xi(t)\|_0^2 + \int_0^t |\xi|_1^2 \,\mathrm{d}t &\lesssim h^\gamma \int_0^t \|\xi\|_0 \|\xi_t\|_0 \,\mathrm{d}t + \int_0^t \|\eta_t\|_0 \|\xi\|_0 \,\mathrm{d}t \\
&\leqslant Ch^\gamma \int_0^t |\xi|_1 \|\xi_t\|_0 \,\mathrm{d}t + C\int_0^t \|\eta_t\|_0^2 \,\mathrm{d}t + \epsilon \int_0^t |\xi|_1^2 \,\mathrm{d}t,
\end{aligned} \tag{4.21}$$

where we have used the Poincáre's inequality $\|u\|_0 \lesssim |u|_1$ for $u \in H_0^1(\Omega)$ in the last step. We choose ϵ small enough to obtain

$$\int_0^t |\xi|_1^2 \,\mathrm{d}t \lesssim h^\gamma \int_0^t |\xi|_1 \|\xi_t\|_0 \,\mathrm{d}t + \int_0^t |\eta_t|_1^2 \,\mathrm{d}t. \tag{4.22}$$

On the other hand, taking $v = \xi_t$ in (4.20) and using (3.23) and an inverse estimate, we obtain

$$\begin{aligned}
(\xi_t, \Pi_h^*\xi_t) + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} a_h(\xi, \Pi_h^*\xi) &= \frac{1}{2}(a_h(\xi_t, \Pi_h^*\xi) - a_h(\xi, \Pi_h^*\xi_t)) + (\eta_t, \Pi_h^*\xi_t) \\
&\leqslant Ch^{\min\{1,\gamma\}} |\xi|_1 |\xi_t|_1 + \|\eta_t\|_0 \|\Pi_h^*\xi_t\|_0 \\
&\lesssim h^{\min\{1,\gamma\}-1} |\xi|_1 \|\xi_t\|_0 + \|\eta_t\|_0 \|\xi_t\|_0.
\end{aligned}$$

Integration on $[0,t]$ gives

$$\int_0^t \|\xi_t\|_0^2\,\mathrm{d}t + |\xi(t)|_1^2 \lesssim \int_0^t (\xi_t, \Pi_h^*\xi_t)\,\mathrm{d}t + \frac{1}{2}a_h(\xi, \Pi_h^*\xi)$$
$$\leqslant Ch^{2\min\{1,\gamma\}-2}\int_0^t |\xi|_1^2\,\mathrm{d}t + C\int_0^t \|\eta_t\|_0^2\,\mathrm{d}t + \epsilon\int_0^t \|\xi_t\|_0^2\,\mathrm{d}t.$$

Using (4.22) to the first term on the right-hand side above, we have

$$\begin{aligned}\int_0^t \|\xi_t\|_0^2\,\mathrm{d}t + |\xi(t)|_1^2 &\\ \leqslant Ch^{2\min\{1,\gamma\}+\gamma-2}\int_0^t |\xi|_1\|\xi_t\|_0\,\mathrm{d}t + Ch^{2\min\{1,\gamma\}-2}\int_0^t \|\eta_t\|_0^2\,\mathrm{d}t + \epsilon\int_0^t \|\xi_t\|_0^2\,\mathrm{d}t &\\ \leqslant Ch^{4\min\{1,\gamma\}+2\gamma-4}\int_0^t |\xi|_1^2\,\mathrm{d}t + Ch^{2\min\{1,\gamma\}-2}\int_0^t \|\eta_t\|_0^2\,\mathrm{d}t + 2\epsilon\int_0^t \|\xi_t\|_0^2\,\mathrm{d}t. & \qquad (4.23)\end{aligned}$$

Taking $\epsilon \leqslant 1/2$ in (4.23) gives

$$|\xi(t)|_1^2 \lesssim h^{4\min\{1,\gamma\}+2\gamma-4}\int_0^t |\xi|_1^2\,\mathrm{d}t + h^{2\min\{1,\gamma\}-2}\int_0^t \|\eta_t\|_0^2\,\mathrm{d}t. \qquad (4.24)$$

Thus, $\gamma \geqslant 2/3$ ensures the convergence. Using the Gronwall's inequality and the Ritz projection (4.16), we have

$$|\xi(t)|_1 \lesssim (1+h^{\gamma-1})h^{r+\min\{1,(m-r)\gamma\}}\|u\|_{H^1(H^{m+1})}, \quad r \leqslant m \leqslant 2r. \qquad (4.25)$$

Now, we use (4.6), (4.16) and (4.25) to obtain

$$|u(t)-u_h(t)|_1 \leqslant |u(t)-R_hu(t)|_1 + |\xi(t)|_1 \lesssim (1+h^{\gamma-1})h^{r+\min\{1,(m-r)\gamma\}}, \quad r \leqslant m \leqslant 2r.$$

Similarly,

$$\begin{aligned}\|u(t)-u_h(t)\|_0 &\lesssim \|u(t)-R_hu(t)\|_0 + |\xi(t)|_1\\ &\lesssim h^{r+\min\{1,(\tilde{m}-r)\gamma\}} + (1+h^{\gamma-1})h^{r+\min\{1,(m-r)\gamma\}} \quad r+1 \leqslant m, \tilde{m} \leqslant 2r,\end{aligned}$$

which gives the desired results (4.18) and (4.19). □

4.3 *Analysis of the fully discrete finite volume scheme*

THEOREM 4.6 Let u be the solution of (1.1) and u_h^n the numerical solution of the Crank–Nicolson fully discrete finite volume scheme (2.3). Assume that the mesh parameter satisfies $\gamma \geqslant 2/3$. If $u \in L^\infty(0,T;H^{r+1}) \cap H^1(0,T;H^{m-1}) \cap H^3(0,T;L^2)$, $r \leqslant m \leqslant 2r$, then, for $0 \leqslant M \leqslant N$,

$$|u^M - u_h^M|_1 \lesssim h^r + h^{r+\gamma-1+\min\{1,(m-r)\gamma\}} + h^{\gamma-1}\Delta t^2. \qquad (4.26)$$

If $u \in L^\infty(0,T;H^{\tilde{m}-1}) \cap H^1(0,T;H^{m-1}) \cap H^3(0,T;L^2)$, $r+1 \leqslant m, \tilde{m} \leqslant 2r$, then for $0 \leqslant M \leqslant N$,

$$\|u^M - u_h^M\|_0 \lesssim h^{r+\min\{1,(s-r)\gamma\}} + (1+h^{\gamma-1})h^{r+\min\{1,(m-r)\gamma\}} + h^{\gamma-1}\Delta t^2. \qquad (4.27)$$

Proof. Let $\xi = u_h - R_h u$ and $\eta = u - R_h u$. It is obvious that $\xi^0 = 0$. We have the following error equation:

$$(\bar{\partial}\xi^n, \Pi_h^* v_h) + a_h(\xi^{n,1/2}, \Pi_h^* v_h) = (\omega^n, \Pi_h^* v_h), \quad v_h \in S_h, \quad n \geqslant 1, \tag{4.28}$$

where

$$\omega^n = \bar{\partial}\eta^n + u_t^{n,1/2} - \bar{\partial}u^n.$$

We choose $v = 2\Delta t \xi^{n,1/2}$ in (4.28) to obtain

$$\begin{aligned}((\xi^n, \Pi_h^*\xi^n) - (\xi^{n-1}, \Pi_h^*\xi^{n-1})) + 2\Delta t a_h(\xi^{n,1/2}, \Pi_h^*\xi^{n,1/2})\\ = \Delta t((\xi^n, \Pi_h^*\bar{\partial}\xi^n) - (\bar{\partial}\xi^n, \Pi_h^*\xi^n)) + 2\Delta t(\omega^n, \Pi_h^*\xi^{n,1/2}).\end{aligned}$$

Summing from $n = 1$ to M and using (3.10) and (4.2) yield

$$\begin{aligned}(\xi^M, \Pi_h^*\xi^M) + \Delta t\sum_{n=1}^M |\xi^{n,1/2}|_1^2 &\lesssim \Delta t\sum_{n=1}^M h^\gamma \|\xi^n\|_0\|\bar{\partial}\xi^n\|_0 + \Delta t\sum_{n=1}^M \|\omega^n\|_0\|\xi^{n,1/2}\|_0\\ &\lesssim \Delta t\sum_{n=1}^M h^\gamma |\xi^n|_1\|\bar{\partial}\xi^n\|_0 + \Delta t\sum_{n=1}^M \|\omega^n\|_0^2 + \epsilon\Delta t\sum_{n=1}^M \|\xi^{n,1/2}\|_0^2.\end{aligned}$$

By (3.11), we can take ϵ small enough to obtain

$$\Delta t\sum_{n=1}^M |\xi^{n,1/2}|_1^2 \lesssim \Delta t\sum_{n=1}^M h^\gamma |\xi^n|_1\|\bar{\partial}\xi^n\|_0 + \Delta t\sum_{n=1}^M \|\omega^n\|_0^2. \tag{4.29}$$

On the other hand, we choose $v = 2\Delta t\bar{\partial}\xi^n$ in (4.28) to get

$$\begin{aligned}2\Delta t(\bar{\partial}\xi^n, \Pi_h^*\bar{\partial}\xi^n) + a_h(\xi^n, \Pi_h^*\xi^n) - a_h(\xi^{n-1}, \Pi_h^*\xi^{n-1})\\ = \Delta t(a_h(\xi^{n-1}, \Pi_h^*\bar{\partial}\xi^n) - a_h(\bar{\partial}\xi^n, \Pi_h^*\xi^{n-1})) + 2\Delta t(\omega^n, \Pi_h^*\bar{\partial}\xi^n).\end{aligned} \tag{4.30}$$

Note that $\xi^{n-1} = \xi^{n,1/2} - \Delta t\bar{\partial}\xi^n/2$. Then, we use (3.23) and an inverse estimate to find that the first term on the right-hand side of (4.30) can be estimated as

$$\begin{aligned}a_h(\xi^{n-1}, \Pi_h^*\bar{\partial}\xi^n) - a_h(\bar{\partial}\xi^n, \Pi_h^*\xi^{n-1}) &= a_h(\xi^{n,1/2}, \Pi_h^*\bar{\partial}\xi^n) - a_h(\bar{\partial}\xi^n, \Pi_h^*\xi^{n,1/2}) - \Delta t a_h(\bar{\partial}\xi^n, \Pi_h^*\bar{\partial}\xi^n)\\ &\lesssim h^{\min\{1,\gamma\}-1}\|\bar{\partial}\xi^n\|_0|\xi^{n,1/2}|_1 - \Delta t a_h(\bar{\partial}\xi^n, \Pi_h^*\bar{\partial}\xi^n).\end{aligned}$$

Summing (4.30) from $n = 1$ to M yields

$$\begin{aligned}&2\Delta t\|\bar{\partial}\xi^n\|_0^2 + a_h(\xi^M, \Pi_h^*\xi^M) + \Delta t\sum_{n=1}^M a_h(\bar{\partial}\xi^n, \Pi_h^*\bar{\partial}\xi^n)\\ &\lesssim \Delta t\sum_{n=1}^M h^{\min\{1,\gamma\}-1}\|\bar{\partial}\xi^n\|_0|\xi^{n,1/2}|_1 + \Delta t\sum_{n=1}^M \|\omega^n\|_0\|\bar{\partial}\xi^n\|_0\\ &\lesssim \Delta t\sum_{n=1}^M h^{2\min\{1,\gamma\}-2}|\xi^{n,1/2}|_1^2 + \Delta t\sum_{n=1}^M \|\omega^n\|_0^2 + \epsilon\Delta t\sum_{n=1}^M \|\bar{\partial}\xi^n\|_0^2.\end{aligned}$$

Using (4.29) to replace the first term on the right-hand side above gives

$$
\begin{aligned}
&2\Delta t\|\bar{\partial}\xi^n\|_0^2 + a_h(\xi^M, \Pi_h^*\xi^M) + \Delta t\sum_{n=1}^{M} a_h(\bar{\partial}\xi^n, \Pi_h^*\bar{\partial}\xi^n) \\
&\lesssim h^{2\min\{1,\gamma\}+\gamma-2}\Delta t\sum_{n=1}^{M}|\xi^n|_1\|\bar{\partial}\xi^n\|_0 + (1+h^{2\gamma-2})\Delta t\sum_{n=1}^{M}\|\omega^n\|_0^2 + \epsilon\Delta t\sum_{n=1}^{M}\|\bar{\partial}\xi^n\|_0^2 \\
&\lesssim h^{4\min\{1,\gamma\}+2\gamma-4}\Delta t\sum_{n=1}^{M}|\xi^n|_1^2 + h^{2\min\{1,\gamma\}-2}\Delta t\sum_{n=1}^{M}\|\omega^n\|_0^2 + 2\epsilon\Delta t\sum_{n=1}^{M}\|\bar{\partial}\xi^n\|_0^2.
\end{aligned}
$$

Taking ϵ small enough and using the coercivity in (4.2), we have

$$
|\xi^M|_1^2 \lesssim h^{4\min\{1,\gamma\}+2\gamma-4}\Delta t\sum_{n=1}^{M}|\xi^n|_1^2 + h^{2\min\{1,\gamma\}-2}\Delta t\sum_{n=1}^{M}\|\omega^n\|_0^2. \tag{4.31}
$$

Applying the discrete Gronwall's lemma yields

$$
|\xi^M|_1^2 \lesssim (1+h^{2\gamma-2})\Delta t\sum_{n=1}^{M}\|\omega^n\|_0^2. \tag{4.32}
$$

Then, the estimates for the Ritz projection and the Taylor expansion remainder together lead to

$$
\begin{aligned}
\Delta t\sum_{n=1}^{M}\|\omega^n\|_0^2 &\leqslant 2\Delta t\sum_{n=1}^{M}(|\bar{\partial}\eta^n|_0^2 + \|u_t^{n,1/2} - \bar{\partial}u^n\|_0^2) \\
&\lesssim \left(h^{2r+2\min\{1,(m-r)\gamma\}}\int_0^{t_M}\|u_t\|_{m+1}^2\,\mathrm{d}t + \Delta t^4\int_0^{t_M}\|u_{ttt}\|_0^2\,\mathrm{d}t,\right),
\end{aligned}
$$

for $r \leqslant m \leqslant 2r$. The desired result follows from (4.6) and (4.16). □

REMARK 4.7 Theorems 4.5 and 4.6 reveal the relationships between the errors and the regularities of the meshes and the exact solution. It is shown that, for L^2-error or H^1-error of the fully discrete scheme, the mesh parameter γ needs to be chosen as $\gamma = 1$ to obtain the optimal-order errors. This observation is consistent with a recent result derived in Yang *et al.* (2013) for the special case $r = 2$. On the contrary, for H^1-error of the semidiscrete scheme, we find that if $u \in H^1(0,T;H^{r+2})$, a less restrictive condition $\gamma = 2/3$ will ensure the optimal order. Moreover, if $\gamma = 1$, we see that $u \in H^1(0,T;H^{r+1}(\Omega))$ can ensure the optimal-order H^1-errors for both semidiscrete and fully discrete schemes, which is better than the regularity assumption $u \in H^1(0,T;H^4(\Omega))$ that was used for the special case $r = 2$ in Yang *et al.* (2013).

5. Numerical experiments

In this section, we present numerical results to illustrate the theoretical estimates in the previous sections.

TABLE 1 *Example* 1 : *Numerical results of quadratic finite volumes* ($r = 2$) *on uniform rectangular meshes combined with Crank–Nicolson time-marching;* $\Delta t = \frac{1}{2}h$

$1/h$	$\|u^N - u_h^N\|_0$	Rate	$\|u^N - u_h^N\|_1$	Rate
4	9.684e−4	—	2.549e−2	—
8	1.226e−4	2.98	6.381e−3	1.99
16	1.538e−5	2.99	1.595e−3	1.99
32	1.926e−6	2.99	3.989e−4	1.99
64	2.416e−7	2.99	9.974e−5	2.00

To balance spatial and temporal truncation errors in the FVMs, we employ the Crank–Nicolson scheme for time-marching when ($r = 2$) quadratic shape functions are used for finite volumes, whereas the modified BDF3 temporal scheme (Iserles, 1996) is utilized when ($r = 3$) cubic finite volume elements are used for spatial discretization. In particular, the modified BDF3 scheme discretizes the variational form as

$$\int_{V_z} \bar{\partial}_3 u_h^n \, \mathrm{d}\mathbf{x} - \int_{\partial V_z} a \nabla u_h^n \cdot \mathbf{n} \, \mathrm{d}s = \int_{V_z} f^n \, \mathrm{d}\mathbf{x} \quad \forall z \in \mathcal{L}^0, \; n \geqslant 3,$$

where

$$\bar{\partial}_3 u_h^n = \frac{1}{\Delta t}\left(\frac{11}{6} u_h^n - 3u_h^{n-1} + \frac{3}{2} u_h^{n-2} - \frac{1}{3} u_h^{n-3}\right).$$

In this case, we use the back-Euler and modified BDF2 as starter schemes, as suggested in Thomée (2006) and adopted in Yang & Liu (2011) and Yang *et al.* (2013).

We have implemented these higher-order FVMs as a `Matlab` package so that the linear solver built in `Matlab` could be readily used. Certain data structures and programming techniques in this finite volume package are similar to those in the finite element package `iFEM` Chen (2009). For both the line and double integrals in the numerical scheme, we use the fifth-order Gaussian quadratures.

EXAMPLE 5.1 We consider the unit square $\Omega = [0, 1]^2$ and $a(\mathbf{x}) = 1$. A known exact solution $u(\mathbf{x}, t) = u(x, y, t) = e^{-(\ln 2)t} \sin(\pi x) \sin(\pi y)$ is chosen so that it satisfies the homogeneous Dirichlet boundary condition for all time. We set the final time as $T = 1$. Obviously, the solution is infinitely smooth, so the accuracy of the numerical solution relies on the order of the finite volume shape functions and the regularity of the quadrilateral meshes being used. We report results for $r = 2$ and $r = 3$ on rectangular and quadrilateral meshes, in particular, the L_2-norm and the H^1-seminorm of the error at the final time step: $\|u^N - u_h^N\|_0$ and $|u^N - u_h^N|_1$.

Shown in Table 1 are the results of quadratic finite volumes ($r = 2$) on uniform rectangular meshes. It can be clearly observed that $\|u^N - u_h^N\|_0$ demonstrates close to third-order convergence, whereas $|u^N - u_h^N|_1$ shows close to second-order convergence.

Shown in Table 2 are the results of cubic finite volumes ($r = 3$) on uniform rectangular meshes. Similarly, it is clear that $\|u^N - u_h^N\|_0$ exhibits close to fourth-order convergence, whereas $|u^N - u_h^N|_1$ displays close to third-order convergence.

TABLE 2 *Example 1: Numerical results of cubic finite volumes ($r = 3$) on uniform rectangular meshes combined with the modified BDF3 temporal scheme; $\Delta t = \frac{1}{2}h$*

$1/h$	$\|u^N - u_h^N\|_0$	Rate	$\|u^N - u_h^N\|_1$	Rate
4	4.471e−5	—	1.690e−3	—
8	2.798e−6	3.99	2.117e−4	2.99
16	1.760e−7	3.99	2.647e−5	2.99
32	1.128e−8	3.96	3.310e−6	2.99

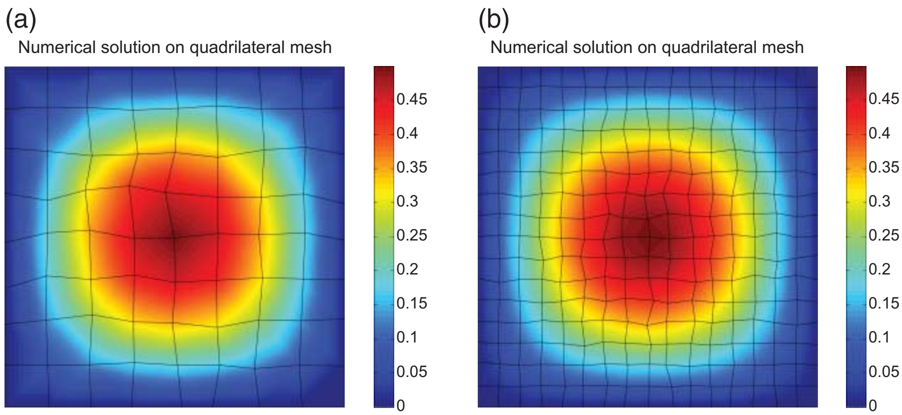

FIG. 2. Example 1: Profiles of two random quadrilateral meshes and the numerical solutions at time $T = 1$ for cubic finite volumes combined with modified BDF3: (a) an 8×8 quadrilateral mesh; (b) a 16×16 quadrilateral mesh.

Next, we examine the performance of the quadratic and cubic FVMs on quadrilateral meshes that are random perturbations of rectangular meshes. In particular, we have the node coordinates of the quadrilateral meshes on the unit square as follows:

$$\begin{aligned} x_{ij} &= \frac{i}{M} + 0.1\frac{1}{M}\sin\left(\frac{j\pi}{M}\right) \text{ randn()}, \\ y_{ij} &= \frac{j}{M} + 0.1\frac{1}{M}\sin\left(\frac{i\pi}{M}\right) \text{ randn()}, \\ 0 &\leqslant i, j \leqslant M, \end{aligned}$$

where $M = 4, 8, 16, 32$ or 64 is the number of partitions in both x- and y-directions, and `randn()` is a built-in random number generator that usually produces a uniformly distributed random number in $(0, 1)$. The profiles of these quadrilateral meshes could be found in Fig. 2. Note that the random quadrilateral meshes used in Tables 3 or 4 are not nested as M increases from 4 to 64. The family of the quadrilateral meshes used in Table 3 is also different from that used in Table 4, since, for each run, the random numbers are different. What we know is that the maximal distortion in the meshes is about 20% of the uniform mesh size $h = 1/M$. It should be pointed out that this family of quadrilateral meshes satisfy the $h^{1+\gamma}$ mesh requirement with $\gamma \approx 0.84$.

TABLE 3 *Example* 1 : *Numerical results of quadratic finite volumes* ($r = 2$) *on quadrilateral meshes combined with Crank–Nicolson time-marching;* $\Delta t = \frac{1}{2}h$

$1/h$	$\|u^N - u_h^N\|_0$	Rate	$\|u^N - u_h^N\|_1$	Rate
4	1.020e−3	—	2.664e−2	—
8	1.315e−4	2.95	6.720e−3	1.98
16	1.690e−5	2.95	1.717e−3	1.96
32	2.096e−6	3.01	4.266e−4	2.00
64	2.619e−7	3.00	1.062e−4	2.00

TABLE 4 *Example* 1 : *Numerical results of cubic finite volumes* ($r = 3$) *on quadrilateral meshes combined with the modified BDF3 temporal scheme;* $\Delta t = \frac{1}{2}h$

$1/h$	$\|u^N - u_h^N\|_0$	Rate	$\|u^N - u_h^N\|_1$	Rate
4	5.302e−5	—	1.966e−3	—
8	3.340e−6	3.98	2.432e−4	3.01
16	2.309e−7	3.85	3.240e−5	2.90
32	1.442e−8	4.00	4.015e−6	3.01
64	9.30e−10	3.95	4.954e−7	3.01

Shown in Table 3 are the numerical results of the quadratic finite volumes ($r = 2$) on a family of random quadrilateral meshes combined with the Crank–Nicolson marching scheme. One can observe an average convergence rate 2.97 in $\|u^N - u_h^N\|_0$ and an average convergence rate 1.98 in $|u^N - u_h^N|_1$.

Shown in Table 4 are the numerical results of the cubic finite volumes ($r = 3$) on a family of random quadrilateral meshes combined with the modified BDF3 temporal scheme. One can observe an average convergence rate 3.94 in $\|u^N - u_h^N\|_0$ and an average convergence rate 2.98 in $|u^N - u_h^N|_1$.

EXAMPLE 5.2 Finite element methods and FVMs for elliptic problems with low regularity have been investigated in Chen (2010), Wihler & Riviére (2011), Liu *et al.* (2012a) and Liu *et al.* (2012b). The following parabolic problem with low regularity is derived from an elliptic problem tested in Wihler & Riviére (2011) and Liu *et al.* (2012b). Here, $\Omega = [0, 1]^2, T = 1$ and a known analytical solution is specified as

$$u(x, y, t) = e^{\alpha t} x(1 - x)y(1 - y)(x^2 + y^2)^{(\beta-2)/2}$$

with $\alpha = -\ln(2), \beta = 0.5$. A nonzero right-hand side $f(x, y, t)$ for Equation (1.1) can be derived accordingly. It is clear that, for any fixed t, $u(x, y, t) \in H_0^1(\Omega) \cap H^{1+\beta-\varepsilon}(\Omega)$, where ε is any small positive number (Liu *et al.*, 2012b). In other words, the spatial regularity of the exact solution is almost of order $(1 + \beta)$.

Similar to Example 5.1, we use quadrilateral meshes obtained by randomly perturbing rectangular meshes with $h = 1/2^m, m = 3, 4, 5, 6$, respectively. Due to the low regularity in spatial variables, $r = 2$ is chosen for finite volume discretization. Accordingly, $\Delta = h/2$, and the Crank–Nicolson scheme is used for time-marching. Listed in Table 5 are the errors of the numerical solution at the final time $T = 1$. It can be observed that optimal convergence rates in the L_2-norm $\|u^N - u_h^N\|_0$ and H^1-seminorm $|u^N - u_h^N|_1$, respectively, around 1.48 and 0.49, are obtained. This example also reveals that approximation accuracy is mainly determined by the low regularity of the problem, even though higher-order approximants ($r = 2$) are used.

TABLE 5 *Example 2: Numerical results of quadratic finite volumes ($r = 2$) on quadrilateral meshes combined with Crank–Nicolson time-marching; $\Delta t = \frac{1}{2}h$*

$1/h$	$\|u^N - u_h^N\|_0$	Rate	$\|u^N - u_h^N\|_1$	Rate
8	9.908e−4	—	6.511e−2	—
16	3.591e−4	1.46	4.652e−2	0.48
32	1.262e−4	1.50	3.292e−2	0.49
64	4.563e−5	1.46	2.346e−2	0.48

REMARK 5.3 The analysis of multistep temporal discretization for finite element schemes is based on an eigendecomposition (see Thomée, 2006), where an essential technical part is that all eigenvalues of a self-adjoint finite element operator are positive. But the eigenvalues of a finite volume operator are much more complicated due to the nonsymmetry of the bilinear form. Therefore, new techniques must be developed to successfully analyse general multistep finite volume schemes.

Funding

The first author is supported by National Natural Science Foundation of China (11201405) and Shandong Province Natural Science Foundation (ZR2014AM003). The third author is partially supported by Natural Science Foundation of China under grants 11171359 and 11428103.

REFERENCES

ANGELINI, O., BRENNER, K. & HILHORST, D. (2013) A finite volume method on general meshes for a degenerate parabolic convection–reaction–diffusion equation. *Numer. Math.*, **123**, 219–257.

ARNOLD, D., BOFFI, D. & FALK, R. (2002) Approximation by quadrilateral finite elements. *Math. Comput.*, **71**, 909–922.

BANK, R. E. & ROSE, D. J. (1987) Some error estimates for the box method. *SIAM J. Numer. Anal.*, **24**, 777–787.

CAI, Z., DOUGLAS JR, J. & PARK, M. (2003) Development and analysis of higher-order finite volume methods over rectangles for elliptic equations. *Adv. Comput. Math.*, **19**, 3–33.

CAO, W., ZHANG, Z. & ZOU, Q. (2013) Superconvergence of any order finite volume schemes for 1D general elliptic equations. *J. Sci. Comput.*, **56**, 566–590.

CASTRO, M., GALLARDO, J. & PARÉS, C. (2006) High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with nonconservative products, applications to shallow-water systems. *Math. Comp.*, **75**, 1103–1134.

CHATZIPANTELIDIS, P., LAZAROV, R. D. & THOMÉE, V. (2008) Parabolic finite volume element equations in nonconvex polygonal domains. *Numer. Methods Partial Differential Equations*, **25**, 507–525.

CHEN, L. (2009) iFEM: an integrated finite element methods package in MATLAB. *Technical Report*, University of California at Irvine.

CHEN, L. (2010) A new class of high order finite volume methods for second-order elliptic equations. *SIAM J. Numer. Anal.*, **47**, 4021–4043.

CHEN, Z., WU, J. & XU, Y. (2012) Higher-order finite volume methods for elliptic boundary value problems. *Adv. Comput. Math.*, **37**, 191–253.

CHOU, S. H. & HE, S. (2002) On the regularity and uniformness conditions on quadrilateral grids. *Comput. Meth. Appl. Mech. Engng.*, **191**, 5149–5158.

CHOU, S. H. & LI, Q. (2000) Error estimates in L^2, H^1 and L^∞ in covolume methods for elliptic and parabolic problems: a unified approach. *Math. Comput.*, **69**, 103–120.

Chou, S. H. & Ye, X. (2007) Unified analysis of finite volume methods for second-order elliptic problems. *SIAM J. Numer. Anal.*, **45**, 1639–1653.

Colella, P., Dorr, M. R., Hittinger, J. A. & Martin, D. F. (2011) Higher-order finite-volume methods in mapped coordinates. *J. Comput. Phys.*, **230**, 2952–2976.

Davis, P. J. & Rabinowitz, P. (1984) *Methods of Numerical Integration*, 2nd edn. Boston: Academic Press.

Ewing, R. E., Lin, T. & Lin, Y. (2002) On the accuracy of the finite volume element method based on piecewise linear polynomials. *SIAM J. Numer. Anal.*, **39**, 1865–1888.

Ewing, R. E., Liu, M. & Wang, J. (1999) Superconvergence of mixed finite element approximations over quadrilaterals. *SIAM J. Numer. Anal.*, **36**, 772–787.

Eymard, R., Gallouët, T. & Herbin, R. (2000) *Finite Volume Methods: Handbook of Numerical Analysis*. Amsterdam: North-Holland.

Hackbusch, W. (1989) On first and second-order box schemes. *Computing*, **41**, 277–296.

Hajibeygi, H. & Jenny, P. (2009) Multiscale finite-volume method for parabolic problems arising from compressible multiphase flow in porous media. *J. Comput. Phys.*, **228**, 5129–5147.

Iserles, A. (1996) *A First Course in the Numerical Analysis of Differential Equations*. Cambridge: Cambridge University Press.

Li, R., Chen, Z. & Wu, W. (2000) *Generalized Difference Methods for Differential Equations: Numerical Analysis of Finite Volume Methods*. New York: Marcel Dekker.

Liebau, F. (1996) The finite volume element method with quadratic basis functions. *Computing*, **57**, 281–299.

Liu, J., Mu, L. & Ye, X. (2012a) L2 error estimation for DGFEM for elliptic problems with low regularity. *Appl. Math. Lett.*, **25**, 1614–1618.

Liu, J., Mu, L., Ye, X. & Jari, R. (2012b) Convergence of the discontinuous finite volume method for elliptic problems with minimal regularity. *J. Comput. Appl. Math.*, **236**, 4537–4546.

Lv, J. & Li, Y. (2012) Optimal biquadratic finite volume element methods on quadrilateral meshes. *SIAM J. Numer. Anal.*, **50**, 2379–2399.

Ma, X., Shu, S. & Zhou, A. (2003) Symmetric finite volume discretizations for parabolic problems. *Comput. Meth. Appl. Mech. Engrg.*, **192**, 4467–4485.

Plexousakis, M. & Zouraris, G. E. (2004) On the construction and analysis of high order locally conservative finite volume-type methods for one-dimensional elliptic problems. *SIAM J. Numer. Anal.*, **42**, 1226–1260.

Sinha, R. K. & Geiser, J. (2007) Error estimates for finite volume element methods for convection–diffusion–reaction equations. *Appl. Numer. Math.*, **57**, 59–72.

Smith, G. D. (1985) *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford: Oxford University Press.

Süli, E. (1992) The accuracy of cell vertex finite volume methods on quadrilateral meshes. *Math. Comput.*, **59**, 359–382.

Thomée, V. (2006) *Galerkin Finite Element Methods for Parabolic Problems*. Berlin: Springer.

Wihler, T. P. & Riviére, B. (2011) Discontinuous Galerkin methods for second-order elliptic PDE with low-regularity solutions. *J. Sci. Comput.*, **46**, 151–165.

Xu, J. C. & Zou, Q. (2009) Analysis of linear and quadratic simplicial finite volume methods for elliptic equations. *Numer. Math.*, **111**, 469–492.

Yang, M. & Liu, J. (2011) A quadratic finite volume element method for parabolic problems on quadrilateral meshes. *IMA J. Numer. Anal.*, **31**, 1038–1061.

Yang, M., Liu, J. & Lin, Y. (2013) Quadratic finite-volume methods for elliptic and parabolic problems on quadrilateral meshes: optimal-order errors based on Barlow points. *IMA J. Numer. Anal.*, **33**, 1342–1364.

Zhang, Q. & Phillip, C. (2012) A fourth-order accurate finite-volume method with structured adaptive mesh refinement for solving the advection–diffusion equation. *SIAM J. Sci. Comput.*, **34**, 179–201.

Zhang, Z. & Zou, Q. (2015) Vertex-centered finite volume schemes of any order over quadrilateral meshes for elliptic boundary value problems. *Numer. Math.*, **130**, 363–393.

Zlámal, M. (1978) Superconvergence and reduced integration in the finite element method. *Math. Comp.*, **143**, 663–685.