

THE ACCURACY OF CELL VERTEX FINITE VOLUME METHODS ON QUADRILATERAL MESHES

ENDRE SÜLI

ABSTRACT. For linear first-order hyperbolic equations in two dimensions we restate the cell vertex finite volume scheme as a finite element method. On structured meshes consisting of distorted quadrilaterals, the global error is shown to be of second order in various mesh-dependent norms, provided that the quadrilaterals are close to parallelograms in the sense that the distance between the midpoints of the diagonals is of the same order as the measure of the quadrilateral. On tensor product nonuniform meshes, the cell vertex scheme coincides with the familiar box scheme. In this case, second-order accuracy is shown without any additional assumption on the regularity of the mesh, which explains the insensitivity of the cell vertex scheme to mesh stretching in the coordinate directions, observed in practice.

1. INTRODUCTION

Over the last two decades, finite volume methods have enjoyed great popularity in the computational aerodynamics community and, since their independent introduction by McDonald [8] and MacCormack and Paullay [7], they have been widely used for the numerical simulation of transonic flows governed by conservation laws. The basic idea behind the construction of finite volume schemes is to exploit the divergence form of the equation by integrating it over finite volumes, and to use Gauss' theorem to convert the volume integrals into contour integrals, which are then discretized. Finite volume methods based on central differences have become particularly popular, following the work of Jameson et al. [5]. In this formulation, usually referred to as the *cell center* scheme, the flow variables are associated with the centers of the computational cells, which are quadrilaterals in two dimensions. An alternative scheme has been introduced by Ni [11], where the flow variables are kept at the vertices of the computational cells. The resulting method is called the *cell vertex* scheme, and it presents a natural generalization of the familiar box scheme to quadrilateral meshes.

In spite of a significant progress on numerical modelling of complex fluid flow problems by cell center and cell vertex finite volume methods, the accuracy of these schemes on distorted multidimensional partitions has not been rigorously investigated. We note, however, that relevant preliminary work based on truncation error analysis has been carried out recently by Giles [4], Morton and Paisley [9], and Roe [13, 14]. The practical evidence presented in [13] suggests

Received by the editor October 12, 1990 and, in revised form, October 2, 1991.

1991 *Mathematics Subject Classification*. Primary 65N06, 65N12, 65N15, 65N30.

Key words and phrases. Finite volume methods, stability, error estimates.

©1992 American Mathematical Society
0025-5718/92 \$1.00 + \$.25 per page

that the cell vertex finite volume scheme has a marked advantage over the cell center scheme in terms of accuracy on distorted quadrilateral meshes, which is in agreement with the findings of Morton and Paisley in [9]. For Friedrichs systems in the plane, Lesaint and Raviart [6] have considered a general class of finite element collocation methods which includes, as a special case, the cell vertex scheme. Their error analysis in the case of first-order hyperbolic equations is, however, restricted to regular rectangular partitions.

In this paper, a theoretical framework is introduced which provides a new interpretation of the cell vertex finite volume scheme and embeds it into the class of finite element methods. This approach enables us to investigate its stability and accuracy, and to obtain optimal error bounds on distorted quadrilateral meshes.

The outline of the paper is as follows. In the next section some notational conventions are introduced. We formulate our model hyperbolic initial boundary value problem and construct its finite volume discretization. Section 3 is devoted to the derivation of a discrete Gårding inequality, which forms the basis of the stability proof. In §4, optimal error bounds are derived on quadrilateral partitions under minimum smoothness requirements on the solution. Our results indicate that both stability and accuracy depend on the distortion of the mesh. More specifically, the scheme is second-order accurate if the quadrilaterals are close to parallelograms in the sense that the distance between the midpoints of the diagonals is of the same order as the measure of the element (Theorem 4). Moreover, on rectangular partitions, the scheme is shown to be second-order accurate without any additional hypothesis on the regularity of the mesh (Theorem 5). In particular, the regularity requirements of Lesaint and Raviart [6] are not necessary in this instance, which explains the insensitivity of the cell vertex scheme to mesh stretching in the coordinate directions, observed in [13].

2. THE MODEL PROBLEM AND ITS DISCRETIZATION

For a complex Banach space V , $\omega \geq 0$, and $p \in [1, \infty]$, we denote by $L_{p,\omega}(V)$ the weighted Bochner space, consisting of all strongly measurable mappings $v: (0, \infty) \rightarrow V$ such that $e^{-\omega t}v \in L_p((0, \infty); V)$. We equip $L_{p,\omega}(V)$ with the norm

$$\|v\|_{L_{p,\omega}(V)} = \begin{cases} (\int_0^\infty e^{-p\omega t} \|v(t)\|_V^p dt)^{1/p} & \text{if } 1 \leq p < \infty, \\ \text{ess sup}_{t \geq 0} e^{-\omega t} \|v(t)\|_V & \text{if } p = \infty. \end{cases}$$

We denote by $H_\omega^m(V)$ the weighted Sobolev space of order m , $m \geq 0$, i.e.,

$$H_\omega^m(V) = \left\{ v \in L_{2,\omega}(V) \mid \frac{d^k v}{dt^k} \in L_{2,\omega}(V), \quad 0 \leq k \leq m \right\}.$$

Let Ω denote the open unit square $(0, 1) \times (0, 1)$. For m , a nonnegative integer, and $p \in [1, \infty]$, we denote by $W_p^m(\Omega)$ the complex Sobolev space of order m , equipped with the usual norm $\|\cdot\|_{W_p^m(\Omega)}$ and seminorm $|\cdot|_{W_p^m(\Omega)}$ (cf. [1]). In particular, when $p = 2$, $W_2^m(\Omega)$ is denoted by $H^m(\Omega)$.

For a measurable set, G , we denote by $m(G)$ the Lebesgue measure of G and by χ_G its characteristic function; \overline{G} denotes the closure of G . For two points in \mathbf{R}^2 , P and Q , say, $\text{dist}(P, Q)$ denotes the Euclidean distance between P and Q , and $\text{diam}(G) = \sup_{P,Q \in G} \text{dist}(P, Q)$.

Suppose that \mathbf{a} is a two-component real vector function with continuously differentiable entries a_1 and a_2 defined on $\overline{\Omega}$. We introduce the following subsets of $\partial\Omega$:

$$\begin{aligned}\partial_{-}\Omega &= \{\mathbf{x} \in \partial\Omega \mid \mathbf{a}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) < 0\}, \\ \partial_{+}\Omega &= \{\mathbf{x} \in \partial\Omega \mid \mathbf{a}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \geq 0\},\end{aligned}$$

where $\mathbf{n}(\mathbf{x})$ denotes the unit outward normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$; when \mathbf{x} is a vertex of Ω , $\mathbf{n}(\mathbf{x})$ is taken to be the unit vector along the axis of the normal cone at $\mathbf{x} \in \partial\Omega$ (see [2, Definition 4.1.3]). With \mathbf{a} , we associate the space $H_{-}^1(\Omega)$ consisting of all v in $H^1(\Omega)$ whose trace on $\partial_{-}\Omega$ is zero.

Given $f \in H_{\omega}^1(L_2(\Omega))$, $\omega \geq 0$, and $u_0 \in H_{-}^1(\Omega)$, consider the following initial boundary value problem:

$$(1) \quad \begin{aligned}\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{a}u) &= f && \text{in } \Omega \times (0, \infty), \\ u &= 0 && \text{on } \partial_{-}\Omega \times (0, \infty), \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}) && \text{in } \Omega.\end{aligned}$$

With the help of semigroup theory, this initial boundary value problem can be shown to possess a unique strong solution (see Pazy [12, Corollary 4.2.10]).

In order to transform (1) into its variational formulation, we introduce the sesquilinear form $B: H_{-}^1(\Omega) \times L_2(\Omega) \rightarrow \mathbf{C}$ defined by

$$B(u, p) = (\nabla \cdot (\mathbf{a}u), p).$$

Now we can restate (1) as follows: find $u \in H_{\omega}^1(L^2(\Omega)) \cap L_{2,\omega}(H_{-}^1(\Omega))$ satisfying

$$(2) \quad \begin{aligned}\left(\frac{\partial u}{\partial t}, p \right) + B(u, p) &= (f, p) \quad \forall p \in L_2(\Omega), \\ (u(\cdot, 0) - u_0, p) &= 0 \quad \forall p \in L_2(\Omega).\end{aligned}$$

The construction of the finite volume method is based on this formulation.

Let $\mathcal{T} = \{\mathcal{T}^h\}$, $h > 0$, be a family of partitions $\mathcal{T}^h = \{K_i^h\}$, where each K_i^h is a convex open quadrilateral. We assume that, with $\mathcal{T}^h = \{K_i^h\}$, $i = 1, 2, \dots, m_h$, $\overline{\Omega} = \bigcup_{i=1}^{m_h} \overline{K_i^h}$, and that each pair $\overline{K_i^h}, \overline{K_j^h}$, $i \neq j$, has either an entire side or a vertex in common, or has empty intersection.

Let $h_{K_i^h}$ denote the diameter of K_i^h , and let $\rho_{K_i^h}$ denote the maximum diameter of circles contained in $\overline{K_i^h}$. We denote by $P_{K_i^h}$ and $Q_{K_i^h}$ the midpoints of the diagonals of K_i^h . The family \mathcal{T} will be assumed to possess the following regularity properties:

Hypothesis H1. The family \mathcal{T} is structured, i.e., for each $h > 0$, \mathcal{T}^h is topologically equivalent to a rectangular partition of Ω .

Hypothesis H2. The quantity $h = \max\{h_{K_i^h} | K_i^h \in \mathcal{T}^h\}$ approximates zero, and there exist two constants $c_0 \geq 0$ and $c_1 > 0$, independent of h , such that for all $K_i^h \in \mathcal{T}^h$, $\mathcal{T}^h \in \mathcal{F}$,

- (i) $\text{dist}(P_{K_i^h}, Q_{K_i^h}) \leq c_0 m(K_i^h),$
- (ii) $h_{K_i^h} \leq c_1 \rho_{K_i^h}.$

We note that hypotheses H2(i) and H2(ii) are independent of each other: H2(i) demands that the quadrilaterals are close to parallelograms, whereas H2(ii) is the usual local regularity condition (cf. [3, §3.1]).

The finite volume discretization of (2) is performed on a family of partitions satisfying hypotheses H1 and H2. In order to introduce the relevant approximation spaces, we define the reference square $\hat{K} = (0, 1) \times (0, 1)$ and denote by $F_{K_i^h}$ the bilinear function which maps \hat{K} onto K_i^h . Since each K_i^h is convex, the determinant $J_{K_i^h}$ of the matrix $DF_{K_i^h}$, the Jacobian matrix of $F_{K_i^h}$, can be assumed to be positive on the closure of \hat{K} . Further, let $Q_1(\hat{K})$ be the set of bilinear functions on \hat{K} , and $Q_0(\hat{K})$ the set of constant functions of \hat{K} . We define

$$\begin{aligned}\mathcal{U}^h &= \{v \in H^1(\Omega) | v = \hat{v} \circ F_{K_i^h}^{-1}, \hat{v} \in Q_1(\hat{K}), K_i^h \in \mathcal{T}^h\}, \\ \mathcal{M}^h &= \{p \in L_2(\Omega) | p = \hat{p} \circ F_{K_i^h}^{-1}, \hat{p} \in Q_0(\hat{K}), K_i^h \in \mathcal{T}^h\},\end{aligned}$$

as well as $\mathcal{U}_-^h = \mathcal{U}^h \cap H_-^1(\Omega)$. Let $P^h: L_2(\Omega) \rightarrow \mathcal{M}^h$ denote the orthogonal projector in $L_2(\Omega)$ onto \mathcal{M}^h , and let $I^h: C(\bar{\Omega}) \rightarrow \mathcal{U}^h$ be the interpolation projector onto \mathcal{U}^h . For a two-component vector function, $w = (w_1, w_2)$, we define $I^h w = (I^h w_1, I^h w_2)$. The discrete analogue of the sesquilinear form B is given by

$$B^h(v, p) = (\nabla \cdot I^h(\mathbf{a}v), p) \quad \forall v \in \mathcal{U}_-^h, \forall p \in \mathcal{M}^h.$$

We define the cell vertex finite volume approximation of (2) as follows: find u^h in $H_\omega^1(\mathcal{U}_-^h)$ satisfying

$$(3) \quad \begin{aligned}\left(\frac{\partial u^h}{\partial t}, p \right) + B^h(u^h, p) &= (f, p) \quad \forall p \in \mathcal{M}^h, \\ (u^h(\cdot, 0) - u_0, p) &= 0 \quad \forall p \in \mathcal{M}^h.\end{aligned}$$

In particular, when p is chosen to be the characteristic function of a quadrilateral K from the partition, an elementary calculation reveals that the spatial discretization, induced by the sesquilinear form B^h in (3), gives rise to a four-point finite difference scheme involving the values of the approximate solution at the four vertices of K . This establishes the connection between (3) and the usual finite difference formulation of the cell vertex scheme.

In order to simplify the presentation, in the rest of the paper, \mathbf{a} will be assumed to be a constant vector. We can also assume, without restricting generality, that both entries of \mathbf{a} are positive, in which case the inflow boundary $\partial_- \Omega$ coincides with the intersection of $\partial \Omega$ with the coordinate axes. Our results can be extended, at the expense of some technical difficulties, to problems with variable coefficients, provided that the components of \mathbf{a} are of constant sign.

3. STABILITY ANALYSIS

The cell vertex scheme (3) will be shown to be stable in mesh-dependent versions of the norms of $L^2(\Omega)$ and $L^2(\partial_+ \Omega)$. The precise definition of these is given below.

3.1. Mesh-dependent norms. For a partition $\mathcal{T}^h \in \mathcal{F}$, we define

$$\|v\|_{l_2(\Omega)} = \left\{ \sum_{K \in \mathcal{T}^h} m(K) \left| \frac{1}{m(K)} \int_K v(x) dx \right|^2 \right\}^{1/2}$$

and, denoting $\partial_+ K = \partial_+ \Omega \cap \overline{K}$,

$$\|v\|_{l_2(\partial_+ \Omega)} = \left\{ \sum_{K \in \mathcal{T}^h, \partial_+ K \neq \emptyset} m(\partial_+ K) \left| \frac{1}{m(\partial_+ K)} \int_{\partial_+ K} v ds \right|^2 \right\}^{1/2}.$$

Clearly, $\|\cdot\|_{l_2(\Omega)}$ is a seminorm on $L_2(\Omega)$, and it is a norm on \mathcal{M}^h . Under the assumption H1, $\|\cdot\|_{l_2(\Omega)}$ is also a norm on \mathcal{U}_-^h .

Let us note that, by virtue of H1, for each $\mathcal{T}^h \in \mathcal{F}$ there exists a pair of positive integers $(M(h), N(h))$ such that m_h , the cardinality of \mathcal{T}^h , is equal to $M(h)N(h)$, and with each $K^h \in \mathcal{T}^h$ we can associate a pair (i, j) , $0 \leq i \leq M(h)-1$, $0 \leq j \leq N(h)-1$. Thus, we label K^h by the subscript ij and write K_{ij}^h instead. The vertices of K_{ij}^h will be denoted $\mathbf{x}_{ij}^h, \mathbf{x}_{i+1,j}^h, \mathbf{x}_{i+1,j+1}^h, \mathbf{x}_{i,j+1}^h$, starting with the lower-left corner and labeling anticlockwise.

For a partition $\mathcal{T}^h = \{K_{ij}^h | 0 \leq i \leq M(h)-1, 0 \leq j \leq N(h)-1\}$, we define the sets

$$\begin{aligned} \Omega_{kl}^h &= \bigcup_{i=0}^{k-1} \bigcup_{j=0}^{l-1} \overline{K_{ij}^h}, \quad k = 1, \dots, M(h), \quad l = 1, \dots, N(h), \\ \partial_+ \Omega_{kl}^h &= \partial \Omega_{kl}^h \setminus \partial_- \Omega, \quad k = 1, \dots, M(h), \quad l = 1, \dots, N(h). \end{aligned}$$

Clearly, $\Omega_{M(h), N(h)}^h = \overline{\Omega}$ and $\partial_+ \Omega_{M(h), N(h)}^h = \partial_+ \Omega$ for all h . For a continuous complex function v defined on $\overline{\Omega}$, let

$$\begin{aligned} \mu v_{ij} &= \frac{1}{m(K_{ij}^h)} \int_{K_{ij}^h} v dx, \\ \mu_1 v_{ij} &= \frac{1}{|\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h|} \int_{[\mathbf{x}_{ij}^h, \mathbf{x}_{i+1,j}^h]} v ds, \\ \mu_2 v_{ij} &= \frac{1}{|\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h|} \int_{[\mathbf{x}_{ij}^h, \mathbf{x}_{i,j+1}^h]} v ds. \end{aligned}$$

We introduce the following mesh-dependent norms:

$$\|v\|_{l_2(\Omega_{kl}^h)} = \left\{ \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} m(K_{ij}^h) |\mu v_{ij}|^2 \right\}^{1/2},$$

$$\|v\|_{l_2(\partial_+ \Omega_{kl}^h)} = \left\{ \sum_{i=0}^{k-1} |\mathbf{x}_{i+1,i}^h - \mathbf{x}_{il}^h| |\mu_1 v_{il}|^2 + \sum_{j=0}^{l-1} |\mathbf{x}_{k,j+1}^h - \mathbf{x}_{kj}^h| |\mu_2 v_{kj}|^2 \right\}^{1/2}.$$

When $k = M(h)$ and $l = N(h)$, these coincide with $\|\cdot\|_{l_2(\Omega)}$ and $\|\cdot\|_{l_2(\partial_+ \Omega)}$,

respectively. In addition to these, we shall also need the mesh-dependent norm

$$\|v\|_{l_2^o(\Omega_{kl}^h)} = \left\{ \sum_{i=0}^{k-1} \sum_{j=1}^{l-1} m(K_{ij}^h \cup K_{i,j-1}^h) |\mu_1 v_{ij}|^2 + \sum_{i=1}^{k-1} \sum_{j=0}^{l-1} m(K_{ij}^h \cup K_{i-1,j}^h) |\mu_2 v_{ij}|^2 \right\}^{1/2}.$$

In the definition of $\|\cdot\|_{l_2^o(\Omega_{kl}^h)}$, we adopt the convention that empty sums, corresponding to $k=1$ or $l=1$, are equal to zero.

3.2. Discrete Gårding inequality. The main result of this subsection is the discrete Gårding inequality stated in Theorem 1 below. The proof of this relies on some technical lemmas, and proving these is our first objective.

The following lemma establishes the connection between hypothesis H2 and some familiar regularity conditions from the theory of finite element methods, and will play an important role in the subsequent analysis.

Lemma 1. Suppose that the family $\mathcal{T} = \{\mathcal{T}^h\}$ satisfies H2. For $K \in \mathcal{T}^h$, let h'_K denote the length of the shortest side of K , and let α_K^j , $j = 1, 2, 3, 4$, denote the interior angles in K . Then there exist two positive constants σ and τ , independent of h , such that for all $K_i^h \in \mathcal{T}^h$, $\mathcal{T}^h \in \mathcal{F}$,

$$(4) \quad h_{K_i^h}/h'_{K_i^h} \leq \sigma$$

and

$$(5) \quad |\cos \alpha_{K_i^h}^j| \leq 1 - \tau, \quad j = 1, 2, 3, 4.$$

Proof. Let us first prove (4). Consider a partition $\mathcal{T}^h \in \mathcal{F}$ and a quadrilateral $K_i^h \in \mathcal{T}^h$. If $h'_{K_i^h} \geq \rho_{K_i^h}$, then (4) immediately follows from H2(ii) with $\sigma = c_1$. If, on the other hand, $h'_{K_i^h} < \rho_{K_i^h}$ (and therefore $c_0 > 0$), then, denoting by $h''_{K_i^h}$ the length of the side of K_i^h opposite the shortest, we have that $h''_{K_i^h} \geq \rho_{K_i^h}$. However, thanks to H2(i), $h''_{K_i^h} \leq h'_{K_i^h} + 2c_0 m(K_i^h)$, and therefore

$$(6) \quad h'_{K_i^h} < \rho_{K_i^h} \leq h''_{K_i^h} \leq h'_{K_i^h} + 2c_0 m(K_i^h).$$

Since $h = \max\{h_{K_i^h} | K_i^h \in \mathcal{T}^h\}$ approximates zero, we can assume, without restricting generality, that $h \leq 1/(4c_0 c_1)$. Thus, $\rho_{K_i^h} - 2c_0 m(K_i^h) \geq \rho_{K_i^h}/2$, and, according to (6),

$$\frac{h_{K_i^h}}{h'_{K_i^h}} \leq \frac{h_{K_i^h}}{\rho_{K_i^h} - 2c_0 m(K_i^h)} \leq 2c_1 \quad (= \sigma).$$

In order to establish (5), let us denote by A , B , C , and D (labeling in the anticlockwise direction) the vertices of $K_i^h \in \mathcal{T}^h$, and let \mathcal{C} denote a circle contained in $\overline{K_i^h}$ with diameter equal to $\rho_{K_i^h}$. Let O denote the center of this circle. We can assume, without restricting generality, that the smallest interior

angle in the quadrilateral K_i^h is at vertex A . Let us construct the two pairs of tangents from A and C to the circle \mathcal{C} and denote their intersections by B' and D' . Since $\angle DAB \geq \angle D'AB'$ and, by assumption, the smallest interior angle in K_i^h is at vertex A , we have thus obtained a lower bound on the smallest angle. It remains to bound the cosine of the angle $\angle D'AB'$ from below in terms of $h_{K_i^h}$ and $\rho_{K_i^h}$.

Let Z denote the point at which the tangent AB' touches the circle, and note that $\angle AZO$ is a right angle. Then

$$\cos \angle OAZ = \frac{|\overline{AZ}|}{|\overline{OA}|} = \frac{|\overline{AZ}|}{(|\overline{ZO}|^2 + |\overline{AZ}|^2)^{1/2}},$$

and, thanks to H2(ii), it follows that $\cos \angle OAZ \leq 2c_1/(1 + 4c_1^2)^{1/2}$. Since $\angle DAB \leq \pi/2$ and $\angle D'AB' = 2\angle OAZ$, this implies that

$$\begin{aligned} 0 &\leq \cos \angle DAB \leq \cos \angle D'AB' \\ &\leq \cos 2 \left(\arccos \frac{2c_1}{(1 + 4c_1^2)^{1/2}} \right) = \frac{4c_1^2 - 1}{4c_1^2 + 1}. \end{aligned}$$

Assuming, without restricting generality, that $\alpha_{K_i^h}^k \leq \alpha_{K_i^h}^l$ for $1 \leq k \leq l \leq 4$, we have thus proved that

$$0 \leq \cos \alpha_{K_i^h}^1 \leq 1 - \frac{2}{1 + 4c_1^2}.$$

Clearly, $\alpha_{K_i^h}^2 \leq 2\pi/3$ and $\alpha_{K_i^h}^3 \leq \pi - \alpha_{K_i^h}^1$. Hence,

$$\begin{aligned} -\frac{1}{2} &\leq \cos \alpha_{K_i^h}^2 \leq \cos \alpha_{K_i^h}^1 \leq 1 - \frac{2}{1 + 4c_1^2}, \\ -1 + \frac{2}{1 + 4c_1^2} &\leq \cos \alpha_{K_i^h}^3 \leq \cos \alpha_{K_i^h}^1 \leq 1 - \frac{2}{1 + 4c_1^2}. \end{aligned}$$

In order to estimate $\cos \alpha_{K_i^h}^4$, let us first note that if $\alpha_{K_i^h}^k$ and $\alpha_{K_i^h}^l$ are two interior angles at diagonally opposite vertices of K_i^h , then, by virtue of H2,

$$|\cos \alpha_{K_i^h}^k - \cos \alpha_{K_i^h}^l| \leq 8c_0 m(K_i^h)(\rho_{K_i^h})^{-1} \leq 8c_0 c_1 h_{K_i^h}.$$

Suppose that $\alpha_{K_i^h}^k$, $k \in \{1, 2, 3\}$, is diagonally opposite to $\alpha_{K_i^h}^4$. Then it follows that

$$\begin{aligned} \cos \alpha_{K_i^h}^4 &= \cos \alpha_{K_i^h}^k + (\cos \alpha_{K_i^h}^4 - \cos \alpha_{K_i^h}^k) \\ &\geq -1 + \left(\frac{2}{1 + 4c_1^2} - 8c_0 c_1 h_{K_i^h} \right). \end{aligned}$$

We can assume, without restricting generality, that

$$h \leq \frac{1}{8c_0 c_1 (1 + 4c_1^2)} \quad \left(\leq \frac{1}{4c_0 c_1} \right).$$

Then, noting that $\alpha_{K_i^h}^4 \geq \pi/2$, we obtain

$$-1 + \frac{1}{1 + 4c_1^2} \leq \cos \alpha_{K_i^h}^4 \leq 0.$$

Thus, we have proved (5) for $j = 1, 2, 3, 4$, with $\tau = (1 + 4c_1^2)^{-1}$. \square

Let us recall from §3.1 that, for a partition $\mathcal{T}^h \in \mathcal{F}$, H1 implies the existence of two integers $M(h)$ and $N(h)$ such that \mathcal{T}^h consists of $M(h)$ columns and $N(h)$ rows of quadrilaterals. Thus, each quadrilateral in the partition can be labeled by an index ij , and we shall write K_{ij}^h , $0 \leq i \leq M(h) - 1$, $0 \leq j \leq N(h) - 1$. Let $K_{ij}^h \in \mathcal{T}^h$ be a quadrilateral with vertices \mathbf{x}_{ij} , $\mathbf{x}_{i+1,j}$, $\mathbf{x}_{i+1,j+1}$, $\mathbf{x}_{i,j+1}$, and denote by \mathbf{n}_{ij}^E , \mathbf{n}_{ij}^W , \mathbf{n}_{ij}^N , \mathbf{n}_{ij}^S the unit outward normals to the East, West, North, and South side of K_{ij}^h , respectively. Let us define

$$\begin{aligned} c_{ij} &= -|\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h| \mathbf{a} \cdot \mathbf{n}_{ij}^W, \quad i = 0, \dots, M(h), \quad j = 0, \dots, N(h) - 1, \\ e_{ij} &= -|\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h| \mathbf{a} \cdot \mathbf{n}_{ij}^S, \quad i = 0, \dots, M(h) - 1, \quad j = 0, \dots, N(h), \end{aligned}$$

with the convention that $\mathbf{n}_{M(h),j}^W = (-1, 0)$ for $j = 0, \dots, N(h) - 1$, and $\mathbf{n}_{i,N(h)}^S = (0, -1)$ for $i = 0, \dots, M(h) - 1$.

Lemma 2. *Suppose that the family \mathcal{F} satisfies hypotheses H1 and H2. Then*

$$(7) \quad |c_{i+1,j} - c_{ij}| \leq 2c_0 |\mathbf{a}| m(K_{ij}^h),$$

$$(8) \quad |e_{i,j+1} - e_{ij}| \leq 2c_0 |\mathbf{a}| m(K_{ij}^h),$$

$$(9) \quad |c_{i+1,j} - c_{i-1,j}| \leq 2c_0 |\mathbf{a}| m(K_{ij}^h \cup K_{i-1,j}^h),$$

$$(10) \quad |e_{i,j+1} - e_{i,j-1}| \leq 2c_0 |\mathbf{a}| m(K_{ij}^h \cup K_{i,j-1}^h).$$

Proof. We begin by establishing (7); the proof of (8) is analogous. Thanks to hypothesis H2(i),

$$\begin{aligned} |c_{i+1,j} - c_{ij}| &\leq |\mathbf{a}| \left| |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h| \mathbf{n}_{i+1,j}^W - |\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h| \mathbf{n}_{ij}^W \right| \\ &= 2|\mathbf{a}| \text{dist}(P_{K_{ij}^h}, Q_{K_{ij}^h}) \leq 2c_0 |\mathbf{a}| m(K_{ij}^h). \end{aligned}$$

Inequalities (9) and (10) follow from (7) and (8) by the triangle inequality. \square

Lemma 3. *Suppose that the family \mathcal{F} satisfies hypotheses H1 and H2. Then*

$$(11) \quad |\mathbf{n}_{ij}^E \cdot \mathbf{a} - \mathbf{n}_{i-1,j}^E \cdot \mathbf{a}| \leq 8c_0 c_1^2 |\mathbf{a}| h'_{K_{ij}^h}$$

for $i = 1, \dots, M(h)$ and $j = 0, \dots, N(h)$.

Proof. Let us consider the quadrilateral $K_{ij}^h \in \mathcal{T}^h$. Then, by H2(i),

$$\begin{aligned} |\mathbf{n}_{ij}^E - \mathbf{n}_{i-1,j}^E| &\leq \frac{1}{|\mathbf{x}_{ij}^h - \mathbf{x}_{i,j+1}^h|} \left| |\mathbf{x}_{ij}^h - \mathbf{x}_{i,j+1}^h| - |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h| \right| \\ (12) \quad &\quad + \frac{1}{|\mathbf{x}_{ij}^h - \mathbf{x}_{i,j+1}^h|} 2 \text{dist}(P_{K_{ij}^h}, Q_{K_{ij}^h}) \\ &\leq \frac{4}{|\mathbf{x}_{ij}^h - \mathbf{x}_{i,j+1}^h|} \text{dist}(P_{K_{ij}^h}, Q_{K_{ij}^h}). \end{aligned}$$

Analogously,

$$(13) \quad |\mathbf{n}_{ij}^E - \mathbf{n}_{i-1,j}^E| \leq \frac{4}{|\mathbf{x}_{i+1,j}^h - \mathbf{x}_{i+1,j+1}^h|} \text{dist}(P_{K_{ij}^h}, Q_{K_{ij}^h}).$$

Combining (12) and (13), and using hypothesis H2, we obtain

$$\begin{aligned} |\mathbf{n}_{ij}^E \cdot \mathbf{a} - \mathbf{n}_{i-1,j}^E \cdot \mathbf{a}| &\leq \frac{4|\mathbf{a}| \operatorname{dist}(P_{K_{ij}^h}, Q_{K_{ij}^h})}{\max(|\mathbf{x}_{ij}^h - \mathbf{x}_{i,j+1}^h|, |\mathbf{x}_{i+1,j}^h - \mathbf{x}_{i+1,j+1}^h|)} \\ &\leq 4c_0|\mathbf{a}| \frac{m(K_{ij}^h)}{\rho_{K_{ij}^h}} \leq 4c_0c_1|\mathbf{a}|h_{K_{ij}^h}. \end{aligned}$$

Applying (4) with $\sigma = 2c_1$ (as in the proof of Lemma 1), we obtain (11). \square

Lemma 4. *Let $c_a = \min(a_1, a_2)/|\mathbf{a}|$, and assume that hypotheses H1 and H2 hold with $c_2 := c_a - 8c_0c_1^2(1 + 2c_0) > 0$. Then*

$$(14) \quad c_{i+1,j} + c_{ij} \geq 2c_2|\mathbf{a}| |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h|,$$

$$(15) \quad e_{i,j+1} + e_{ij} \geq 2c_2|\mathbf{a}| |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i,j+1}^h|$$

for $i = 0, \dots, M(h) - 1$ and $j = 0, \dots, N(h) - 1$.

Proof. We shall only prove the first of the two inequalities. It is clear that $\mathbf{a} \cdot \mathbf{n}_{M(h),j}^W = -\mathbf{a} \cdot \mathbf{n}_{M(h)-1,j}^E$ and

$$(16) \quad \mathbf{a} \cdot \mathbf{n}_{M(h)-1,j}^E = (a_1, a_2) \cdot (1, 0) = a_1 \geq c_a|\mathbf{a}| \quad (> 0).$$

For $0 \leq i \leq M(h) - 2$ and $0 \leq j \leq N(h) - 1$, it follows from (16) and (11) that

$$\begin{aligned} \mathbf{a} \cdot \mathbf{n}_{ij}^E &= \mathbf{a} \cdot \mathbf{n}_{M(h)-1,j}^E - \sum_{k=i+1}^{M(h)-1} (\mathbf{a} \cdot \mathbf{n}_{kj}^E - \mathbf{a} \cdot \mathbf{n}_{k-1,j}^E) \\ &\geq |\mathbf{a}| \left(c_a - 8c_0c_1^2 \sum_{k=i+1}^{M(h)-1} h'_{K_{kj}^h} \right). \end{aligned}$$

Thus, using (7), we conclude that

$$\begin{aligned} c_{i+1,j} + c_{ij} &= 2c_{i+1,j} - (c_{i+1,j} - c_{ij}) \\ &\geq 2\mathbf{a} \cdot \mathbf{n}_{ij}^E |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h| - 2|\mathbf{a}|c_0m(K_{ij}^h) \\ &\geq 2|\mathbf{a}| |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h| \left(c_a - 8c_0c_1^2 \sum_{k=i+1}^{M(h)-1} h'_{K_{kj}^h} - c_0 \frac{h_{K_{ij}^h}^2}{h'_{K_{ij}^h}} \right). \end{aligned}$$

By virtue of (4) with $\sigma = 2c_1$, the last term in the brackets can be bounded from below by $-4c_0c_1^2h'_{K_{ij}^h}$. Hence,

$$c_{i+1,j} + c_{ij} \geq 2|\mathbf{a}| |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h| \left(c_a - 8c_0c_1^2 \sum_{k=0}^{M(h)-1} |\mathbf{x}_{kj}^h - \mathbf{x}_{k+1,j}^h| \right).$$

However by H2(i),

$$\begin{aligned} |\mathbf{x}_{ij}^h - \mathbf{x}_{i+1,j}^h| &\leq |\mathbf{x}_{i0}^h - \mathbf{x}_{i+1,0}^h| + \sum_{l=0}^{j-1} \left| |\mathbf{x}_{i,l+1}^h - \mathbf{x}_{i+1,l+1}^h| - |\mathbf{x}_{il}^h - \mathbf{x}_{i+1,l}^h| \right| \\ &\leq |\mathbf{x}_{i0}^h - \mathbf{x}_{i+1,0}^h| + 2c_0 \sum_{l=0}^{N(h)-1} m(K_{il}^h), \end{aligned}$$

and therefore,

$$(17) \quad \sum_{i=0}^{M(h)-1} |\mathbf{x}_{ij}^h - \mathbf{x}_{i+1,j}^h| \leq 1 + 2c_0, \quad j = 0, \dots, N(h).$$

Employing inequality (17) yields

$$(18) \quad c_{i+1,j} + c_{ij} \geq 2|\mathbf{a}| |\mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h| (c_a - 8c_0c_1^2(1 + 2c_0))$$

for $i = 0, \dots, M(h) - 2$, $j = 0, \dots, N(h) - 1$. In fact, according to (16), the result (18) is also valid for $i = M(h) - 1$ and $j = 0, \dots, N(h) - 1$. Setting $c_2 := c_a - 8c_0c_1^2(1 + 2c_0)$, we obtain (14). \square

Recalling the definition of the set Ω_{kl}^h from §3.1, we denote by χ_{kl} the characteristic function of Ω_{kl}^h , i.e., $\chi_{kl}(x) = \chi_{\Omega_{kl}^h}(x)$. The proof of the stability of the finite volume method (3) is based on the following discrete sharp Gårding inequality.

Theorem 1. *Let $c_a = \min(a_1, a_2)/|\mathbf{a}|$, and assume that hypotheses H1 and H2 hold with $c_2 := c_a - 8c_0c_1^2(1 + 2c_0) > 0$. Then, for all k and l , $1 \leq k \leq M(h)$, $1 \leq l \leq N(h)$,*

$$(19) \quad \begin{aligned} \operatorname{Re} B^h(v, \chi_{kl} P^h v) &\geq -4c_0|\mathbf{a}| \|v\|_{L_2(\Omega_{kl}^h)}^2 \\ &\quad - \frac{1}{2}c_0|\mathbf{a}| \|v\|_{L_2(\Omega_{kl}^h)}^2 + \frac{1}{2}c_2|\mathbf{a}| \|v\|_{L_2(\partial_+ \Omega_{kl}^h)}^2 \quad \forall v \in \mathcal{U}_-^h. \end{aligned}$$

Proof. Let $K_{ij}^h \in \mathcal{T}^h$, and denote by $\mathbf{n}_{ij}^E, \mathbf{n}_{ij}^W, \mathbf{n}_{ij}^N, \mathbf{n}_{ij}^S$ the unit outward normals to the East, West, North, and South side of K_{ij}^h , respectively.

We shall first consider the case when $2 \leq k \leq M(h)$, $2 \leq l \leq N(h)$. Recalling the definition of the sesquilinear form B^h , we can write

$$\begin{aligned} B^h(v, \chi_{kl} P^h v) &= \int_{\Omega} [\nabla \cdot I^h(\mathbf{a}v)] \chi_{kl} P^h \bar{v} \, dx = \int_{\Omega_{kl}^h} [\nabla \cdot I^h(\mathbf{a}v)] P^h \bar{v} \, dx \\ &= \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} \left(\int_{\partial K_{ij}^h} I^h(\mathbf{a}v) \cdot \mathbf{n} \, ds \right) \left(\frac{1}{m(K_{ij}^h)} \int_{K_{ij}^h} \bar{v} \, dx \right), \end{aligned}$$

where \bar{v} denotes the complex conjugate of v . Since $I^h(\mathbf{a}v)$ is a linear function along each side of the quadrilateral K_{ij}^h , the contour integral appearing in the last expression can be evaluated by the trapezium rule, thus yielding

$$\begin{aligned} B^h(v, \chi_{kl} P^h v) &= \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} \left\{ \left| \mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i+1,j}^h \right| \frac{(\mathbf{a}v)_{i+1,j+1} \cdot \mathbf{n}_{ij}^E + (\mathbf{a}v)_{i+1,j} \cdot \mathbf{n}_{ij}^E}{2} \right. \\ &\quad + \left| \mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h \right| \frac{(\mathbf{a}v)_{i,j+1} \cdot \mathbf{n}_{ij}^W + (\mathbf{a}v)_{ij} \cdot \mathbf{n}_{ij}^W}{2} \\ &\quad + \left| \mathbf{x}_{i+1,j+1}^h - \mathbf{x}_{i,j+1}^h \right| \frac{(\mathbf{a}v)_{i+1,j+1} \cdot \mathbf{n}_{ij}^N + (\mathbf{a}v)_{i,j+1} \cdot \mathbf{n}_{ij}^N}{2} \\ &\quad \left. + \left| \mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h \right| \frac{(\mathbf{a}v)_{i+1,j} \cdot \mathbf{n}_{ij}^S + (\mathbf{a}v)_{ij} \cdot \mathbf{n}_{ij}^S}{2} \right\} \tilde{v}_{ij}, \end{aligned}$$

where $\tilde{v}_{ij} = \mu \bar{v}_{ij}$. Shifting the indices in the i -summation in the first term and in the j -summation in the third term, and noting that $(\mathbf{av})_{0,j+1} = (\mathbf{av})_{0,j} = (\mathbf{av})_{i+1,0} = (\mathbf{av})_{i,0} = 0$, $\mathbf{n}_{ij}^E = -\mathbf{n}_{i+1,j}^W$, and $\mathbf{n}_{ij}^N = -\mathbf{n}_{i,j+1}^S$, we obtain

$$\begin{aligned} B^h(v, \chi_{kl} P^h v) &= \sum_{i=1}^{k-1} \sum_{j=0}^{l-1} |\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h| \frac{(\mathbf{av})_{i,j+1} + (\mathbf{av})_{ij}}{2} \cdot \mathbf{n}_{ij}^W (\tilde{v}_{ij} - \tilde{v}_{i-1,j}) \\ &\quad + \sum_{i=0}^{k-1} \sum_{j=1}^{l-1} |\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h| \frac{(\mathbf{av})_{i+1,j} + (\mathbf{av})_{ij}}{2} \cdot \mathbf{n}_{ij}^S (\tilde{v}_{ij} - \tilde{v}_{i,j-1}) \\ &\quad - \sum_{j=0}^{l-1} |\mathbf{x}_{k,j+1}^h - \mathbf{x}_{kj}^h| \frac{(\mathbf{av})_{k,j+1} + (\mathbf{av})_{kj}}{2} \cdot \mathbf{n}_{kj}^W \tilde{v}_{k-1,j} \\ &\quad - \sum_{i=0}^{k-1} |\mathbf{x}_{i+1,l}^h - \mathbf{x}_{il}^h| \frac{(\mathbf{av})_{i+1,l} + (\mathbf{av})_{il}}{2} \cdot \mathbf{n}_{il}^S \tilde{v}_{i,l-1} \\ &= -\frac{1}{2} \sum_{i=1}^{k-1} \sum_{j=0}^{l-1} c_{ij} \mu_2 v_{ij} (\mu_2 \bar{v}_{i+1,j} - \mu_2 \bar{v}_{i-1,j}) \\ &\quad - \frac{1}{2} \sum_{i=0}^{k-1} \sum_{j=1}^{l-1} e_{ij} \mu_1 v_{ij} (\mu_1 \bar{v}_{i,j+1} - \mu_1 \bar{v}_{i,j-1}) \\ &\quad + \sum_{j=0}^{l-1} c_{kj} \mu_2 v_{kj} \tilde{v}_{k-1,j} + \sum_{i=0}^{k-1} e_{il} \mu_1 v_{il} \tilde{v}_{i,l-1}. \end{aligned}$$

Hence, shifting indices again and observing that

$$\mu v_{ij} = \frac{1}{2} \mu_1 (v_{ij} + v_{i,j+1}) = \frac{1}{2} \mu_2 (v_{ij} + v_{i+1,j}),$$

we obtain

$$\begin{aligned} B^h(v, \chi_{kl} P^h v) &= \frac{1}{2} \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} \{-c_{ij} \mu_2 v_{ij} \mu_2 \bar{v}_{i+1,j} + c_{i+1,j} \mu_2 \bar{v}_{ij} \mu_2 v_{i+1,j}\} \\ &\quad + \frac{1}{2} \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} \{-e_{ij} \mu_1 v_{ij} \mu_1 \bar{v}_{i,j+1} + e_{i,j+1} \mu_1 \bar{v}_{ij} \mu_1 v_{i,j+1}\} \\ &\quad + \frac{1}{2} \sum_{j=0}^{l-1} c_{kj} |\mu_2 v_{kj}|^2 + \frac{1}{2} \sum_{i=0}^{k-1} e_{il} |\mu_1 v_{il}|^2. \end{aligned}$$

Since $\operatorname{Re}(\alpha \bar{\beta}) = \operatorname{Re}(\bar{\alpha} \beta)$, this yields

$$\begin{aligned} \operatorname{Re} B^h(v, \chi_{kl} P^h v) &= \frac{1}{2} \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} (c_{i+1,j} - c_{ij}) \operatorname{Re}(\mu_2 v_{ij} \mu_2 \bar{v}_{i+1,j}) \\ &\quad + \frac{1}{2} \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} (e_{i,j+1} - e_{ij}) \operatorname{Re}(\mu_1 v_{ij} \mu_1 \bar{v}_{i,j+1}) \\ &\quad + \frac{1}{2} \sum_{j=0}^{l-1} c_{kj} |\mu_2 v_{kj}|^2 + \frac{1}{2} \sum_{i=0}^{k-1} e_{il} |\mu_1 v_{il}|^2. \end{aligned}$$

Using the identity

$$\operatorname{Re}(\alpha\bar{\beta}) = 2 \left| \frac{\alpha + \beta}{2} \right|^2 - \frac{1}{2}(|\alpha|^2 + |\beta|^2),$$

we obtain

$$\begin{aligned} & \operatorname{Re} B^h(v, \chi_{kl} P^h v) \\ &= \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} (c_{i+1,j} - c_{ij}) |\mu v_{ij}|^2 + \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} (e_{i,j+1} - e_{ij}) |\mu v_{ij}|^2 \\ &\quad - \frac{1}{4} \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} (c_{i+1,j} - c_{ij}) (|\mu_2 v_{ij}|^2 + |\mu_2 v_{i+1,j}|^2) \\ &\quad - \frac{1}{4} \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} (e_{i,j+1} - e_{ij}) (|\mu_1 v_{ij}|^2 + |\mu_1 v_{i,j+1}|^2) \\ &\quad + \frac{1}{2} \sum_{j=0}^{l-1} c_{kj} |\mu_2 v_{kj}|^2 + \frac{1}{2} \sum_{i=0}^{k-1} e_{il} |\mu_1 v_{il}|^2 \\ &= \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} [(c_{i+1,j} - c_{ij}) + (e_{i,j+1} - e_{ij})] |\mu v_{ij}|^2 \\ &\quad - \frac{1}{4} \left\{ \sum_{i=1}^{k-1} \sum_{j=0}^{l-1} (c_{i+1,j} - c_{i-1,j}) |\mu_2 v_{ij}|^2 + \sum_{i=0}^{k-1} \sum_{j=1}^{l-1} (e_{i,j+1} - e_{i,j-1}) |\mu_1 v_{ij}|^2 \right\} \\ &\quad + \frac{1}{4} \left\{ \sum_{j=0}^{l-1} (c_{kj} + c_{k-1,j}) |\mu_2 v_{kj}|^2 + \sum_{i=0}^{k-1} (e_{il} + e_{i,l-1}) |\mu_1 v_{il}|^2 \right\}. \end{aligned}$$

Employing (7)–(10), we find

$$\begin{aligned} & \operatorname{Re} B^h(v, \chi_{kl} P^h v) \\ &\geq -4c_0 |\mathbf{a}| \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} m(K_{ij}^h) |\mu v_{ij}|^2 \\ &\quad - \frac{1}{2} c_0 |\mathbf{a}| \left\{ \sum_{i=0}^{k-1} \sum_{j=1}^{l-1} m(K_{ij}^h \cup K_{i,j-1}^h) |\mu_1 v_{ij}|^2 \right. \\ &\quad \left. + \sum_{i=1}^{k-1} \sum_{j=0}^{l-1} m(K_{ij}^h \cup K_{i-1,j}^h) |\mu_2 v_{ij}|^2 \right\} \\ &\quad + \frac{1}{4} \left\{ \sum_{j=0}^{l-1} (c_{kj} + c_{k-1,j}) |\mu_2 v_{kj}|^2 + \sum_{i=0}^{k-1} (e_{il} + e_{i,l-1}) |\mu_1 v_{il}|^2 \right\}. \end{aligned} \tag{20}$$

Applying (14) and (15) to bound the last two sums in (20) from below, and recalling the definitions of the mesh-dependent norms $\|\cdot\|_{L_2(\Omega_{kl}^h)}$, $\|\cdot\|_{L_2^0(\Omega_{kl}^h)}$, and $\|\cdot\|_{L_2(\partial_\nu \Omega_{kl}^h)}$, we obtain (19) for $2 \leq k \leq M(h)$ and $2 \leq l \leq N(h)$.

Now let us consider the case when at least one of the two indices k and l equals 1. Suppose $k = 1$, $2 \leq l \leq N(h)$; the case $l = 1$, $2 \leq k \leq M(h)$ is dealt with similarly, and the case when both $k = 1$ and $l = 1$ will be considered separately.

Thus, we assume that $k = 1$, $2 \leq l \leq N(h)$. By the same argument as in the case when $k \geq 2$, we obtain

$$\begin{aligned} \operatorname{Re} B^h(v, \chi_{1l} P^h v) &= \sum_{j=0}^{l-1} (e_{0,j+1} - e_{0,j}) |\mu v_{0j}|^2 - \frac{1}{4} \sum_{j=1}^{l-1} (e_{0,j+1} - e_{0,j-1}) |\mu_1 v_{0j}|^2 \\ &\quad + \frac{1}{2} \sum_{j=0}^{l-1} c_{1j} |\mu_2 v_{1j}|^2 + \frac{1}{4} (e_{0l} + e_{0,l-1}) |\mu_1 v_{0l}|^2. \end{aligned}$$

However,

$$\begin{aligned} c_{1j} &\geq c_2 |\mathbf{a}| |\mathbf{x}_{1,j+1}^h - \mathbf{x}_{1j}^h|, \\ |e_{0,j+1} - e_{0,j}| &\leq 2c_0 |\mathbf{a}| m(K_{0j}^h), \\ |e_{0,j+1} - e_{0,j-1}| &\leq 2c_0 |\mathbf{a}| m(K_{0j}^h \cup K_{0,j-1}^h), \\ e_{0l} + e_{0,l-1} &\geq 2c_2 |\mathbf{a}| |\mathbf{x}_{1l}^h - \mathbf{x}_{0l}^h|, \end{aligned}$$

and therefore,

$$\begin{aligned} \operatorname{Re} B^h(v, \chi_{1l} P^h v) &\geq -2c_0 |\mathbf{a}| \sum_{j=0}^{l-1} m(K_{0j}^h) |\mu v_{0j}|^2 - \frac{1}{2} c_0 |\mathbf{a}| \sum_{j=1}^{l-1} m(K_{0j}^h \cup K_{0,j-1}^h) |\mu_1 v_{0j}|^2 \\ &\quad + \frac{1}{2} c_2 |\mathbf{a}| \left\{ |\mathbf{x}_{1l}^h - \mathbf{x}_{0l}^h| |\mu_1 v_{0l}|^2 + \sum_{j=0}^{l-1} |\mathbf{x}_{1,j+1}^h - \mathbf{x}_{1j}^h| |\mu_2 v_{1j}|^2 \right\}. \end{aligned}$$

With the definitions of the mesh-dependent norms $\|\cdot\|_{l_2(\Omega_{kl}^h)}$, $\|\cdot\|_{l_2^o(\Omega_{kl}^h)}$, and $\|\cdot\|_{l_2(\partial_+ \Omega_{kl}^h)}$ in mind, this yields

$$\begin{aligned} \operatorname{Re} B^h(v, \chi_{1l} P^h v) &\geq -2c_0 |\mathbf{a}| \|v\|_{l_2(\Omega_{1l}^h)}^2 \\ &\quad - \frac{1}{2} c_0 |\mathbf{a}| \|v\|_{l_2^o(\Omega_{1l}^h)}^2 + \frac{1}{2} c_2 |\mathbf{a}| \|v\|_{l_2(\partial_+ \Omega_{1l}^h)}^2 \quad \forall v \in \mathcal{U}_-^h \end{aligned}$$

for $2 \leq l \leq N(h)$. Analogously, when $l = 1$ and $2 \leq k \leq M(h)$,

$$\begin{aligned} \operatorname{Re} B^h(v, \chi_{k1} P^h v) &\geq -2c_0 |\mathbf{a}| \|v\|_{l_2(\Omega_{k1}^h)}^2 \\ &\quad - \frac{1}{2} c_0 |\mathbf{a}| \|v\|_{l_2^o(\Omega_{k1}^h)}^2 + \frac{1}{2} c_2 |\mathbf{a}| \|v\|_{l_2(\partial_+ \Omega_{k1}^h)}^2 \quad \forall v \in \mathcal{U}_-^h. \end{aligned}$$

Let us finally consider the case $k = 1$, $l = 1$. In this instance,

$$\operatorname{Re} B^h(v, \chi_{11} P^h v) = \frac{1}{2} c_{10} |\mu_2 v_{10}|^2 + \frac{1}{2} e_{01} |\mu_1 v_{01}|^2.$$

Noting that

$$c_{10} \geq c_2 |\mathbf{a}| |\mathbf{x}_{11}^h - \mathbf{x}_{10}^h|, \quad e_{01} \geq c_2 |\mathbf{a}| |\mathbf{x}_{11}^h - \mathbf{x}_{01}^h|,$$

we obtain

$$\operatorname{Re} B^h(v, \chi_{11} P^h v) \geq \frac{1}{2} c_2 |\mathbf{a}| \|v\|_{l_2(\partial_+ \Omega_{11}^h)}^2 \quad \forall v \in \mathcal{U}_-^h. \quad \square$$

3.3. Stability. In order to complete the stability proof initiated in the previous subsection, we need a bivariate discrete Gronwall inequality asserted in Lemma 6 below. Its proof relies on the following univariate discrete Gronwall lemma which is easily proved by induction.

Lemma 5. Suppose that (a_i) , (b_i) , (c_i) , and (d_i) are four sequences of nonnegative real numbers such that the sequence (c_i) is nondecreasing, and

$$a_i + b_i \leq c_i + \sum_{j=0}^{i-1} d_j a_j, \quad i \geq 1, \quad a_0 + b_0 \leq c_0.$$

Then

$$a_i + b_i \leq c_i \exp\left(\sum_{j=0}^{i-1} d_j\right), \quad i \geq 1.$$

As a consequence of this, we obtain the following result.

Lemma 6. Suppose that C is a nonnegative constant, and (a_{ij}) , (b_{ij}) , (c_{ij}) , (α_i) , and (β_i) , $i, j \geq 0$, are five sequences of nonnegative real numbers such that

$$a_{kl} + b_{kl} + c_{kl} \leq C + \sum_{j=0}^{l-1} \alpha_j b_{kj} + \sum_{i=0}^{k-1} \beta_i c_{il}, \quad k, l \geq 0,$$

with the convention that empty sums are equal to zero. Then

$$(21) \quad a_{kl} + b_{kl} + c_{kl} \leq C \min\{\exp(A_l + B_k \exp(A_l)), \exp(B_k + A_l \exp(B_k))\}$$

for $k, l \geq 1$, where

$$A_l = \sum_{j=0}^{l-1} \alpha_j, \quad B_k = \sum_{i=0}^{k-1} \beta_i, \quad k, l \geq 1.$$

Proof. Let us fix k , and apply Lemma 5 to obtain

$$a_{kl} + b_{kl} + c_{kl} \leq \left(C + \sum_{i=1}^{k-1} \beta_i c_{il}\right) \exp(A_l), \quad k, l \geq 1.$$

Now fixing l in this inequality and applying Lemma 5 again, we get

$$a_{kl} + b_{kl} + c_{kl} \leq C \exp(A_l + B_k \exp(A_l)).$$

Repeating the process by first fixing l and then k , we obtain

$$a_{kl} + b_{kl} + c_{kl} \leq C \exp(B_k + A_l \exp(B_k)).$$

Finally, (21) is arrived at by taking the minimum of the right-hand side terms in the last two inequalities. \square

Lemma 7. Suppose that the family $\mathcal{T}^h \in \mathcal{F}$ satisfies hypotheses H1 and H2.

Then

$$\begin{aligned} \sum_{j=1}^{l-1} \max_{0 \leq i \leq k-1} \frac{m(K_{ij}^h \cup K_{i,j-1}^h)}{|\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h|} &\leq 8c_1^2(1 + 2c_0), \quad 1 \leq k \leq M(h), \\ &\quad 2 \leq l \leq N(h), \\ \sum_{i=1}^{k-1} \max_{0 \leq j \leq l-1} \frac{m(K_{ij}^h \cup K_{i-1,j}^h)}{|\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h|} &\leq 8c_1^2(1 + 2c_0), \quad 2 \leq k \leq M(h), \\ &\quad 1 \leq l \leq N(h). \end{aligned}$$

Proof. We shall only prove the first inequality; the proof of the second is analogous. Let us consider a partition $\mathcal{T}^h = \{K_{ij}^h | 0 \leq i \leq M(h) - 1, 0 \leq j \leq N(h) - 1\}$. Then, by Lemma 1,

$$\begin{aligned} m(K_{ij}^h \cup K_{i,j-1}^h) &= m(K_{ij}^h) + m(K_{i,j-1}^h) \leq \left\{ \frac{h_{K_{ij}^h}^2}{h'_{K_{ij}^h}} + \frac{h_{K_{i,j-1}^h}^2}{h'_{K_{i,j-1}^h}} \right\} |\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h| \\ &\leq \sigma^2 (|\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h| + |\mathbf{x}_{ij}^h - \mathbf{x}_{i,j-1}^h|) |\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h| \\ &\leq \sigma^2 \left(\max_{0 \leq i \leq M(h)} |\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h| + \max_{0 \leq i \leq M(h)} |\mathbf{x}_{ij}^h - \mathbf{x}_{i,j-1}^h| \right) |\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h|. \end{aligned}$$

Therefore,

$$\begin{aligned} \max_{0 \leq i \leq k-1} \frac{m(K_{ij}^h \cup K_{i,j-1}^h)}{|\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h|} &\leq \sigma^2 \left(\max_{0 \leq i \leq M(h)} |\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h| + \max_{0 \leq i \leq M(h)} |\mathbf{x}_{ij}^h - \mathbf{x}_{i,j-1}^h| \right). \end{aligned}$$

Summing through $j = 1, \dots, l-1$, we obtain

$$(22) \quad \sum_{j=1}^{l-1} \max_{0 \leq i \leq k-1} \frac{m(K_{ij}^h \cup K_{i,j-1}^h)}{|\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h|} \leq 2\sigma^2 \sum_{j=0}^{N(h)-1} \max_{0 \leq i \leq M(h)} |\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h|$$

for $2 \leq l \leq N(h)$. The sum appearing on the right-hand side of this inequality is estimated in the same way as (17) was arrived at:

$$\sum_{j=0}^{N(h)-1} \max_{0 \leq i \leq M(h)} |\mathbf{x}_{i,j+1}^h - \mathbf{x}_{ij}^h| \leq 1 + 2c_0.$$

Substituting this into the right-hand side of (22) yields

$$\sum_{j=1}^{l-1} \max_{0 \leq i \leq k-1} \frac{m(K_{ij}^h \cup K_{i,j-1}^h)}{|\mathbf{x}_{i+1,j}^h - \mathbf{x}_{ij}^h|} \leq 2\sigma^2(1 + 2c_0)$$

for $1 \leq k \leq M(h)$ and $2 \leq l \leq N(h)$. Since $\sigma = 2c_1$, this gives the desired inequality. \square

Now we are ready to prove the main result of this section.

Theorem 2. Assuming H1, problem (3) has a unique solution $u^h(\cdot, t) \in \mathcal{U}_-^h$. If, in addition, H2 holds with $c_2 := c_a - 8c_0c_1^2(1+2c_0) > 0$, then

$$(23) \quad \begin{aligned} & \|u^h\|_{L_{\infty,\omega}(l_2(\Omega))}^2 + c_2|\mathbf{a}|\|u^h\|_{L_{2,\omega}(l_2(\partial_+\Omega))}^2 \\ & \leq c_3(\|u^h(0)\|_{l_2(\Omega)}^2 + \|f\|_{L_{2,\omega}(l_2(\Omega))}^2), \end{aligned}$$

where $\omega \geq \frac{1}{2}(1+8c_0|\mathbf{a}|)$, $c_3 = \exp\{c_4 + c_4 \exp(c_4)\}$, and $c_4 = 16c_0c_1^2(1+2c_0)/c_2$.

Proof. The existence of a unique solution is proved constructively. Let $K_{ij}^h \in \mathcal{T}^h$, $\mathcal{T}^h \in \mathcal{F}$, and choose $p = \chi_{K_{ij}^h}$ in (3). Then (3) becomes a single ordinary differential equation involving $u^h(\mathbf{x}_{ij}^h, t)$, $u^h(\mathbf{x}_{i+1,j}^h, t)$, $u^h(\mathbf{x}_{i+1,j+1}^h, t)$, and $u^h(\mathbf{x}_{i,j+1}^h, t)$, the values of $u^h(\cdot, t)$ at the four vertices of K_{ij}^h .

Let us assume, to begin with, that $i = j = 0$. Since $u^h(\mathbf{x}_{00}^h, t)$, $u^h(\mathbf{x}_{10}^h, t)$, and $u^h(\mathbf{x}_{01}^h, t)$ are determined by the boundary condition for all $t \geq 0$, they can be eliminated, thus yielding a single linear first-order ordinary differential equation for $u^h(\mathbf{x}_{11}^h, t)$. This uniquely determines $u^h(\mathbf{x}_{11}^h, t)$ for all $t \geq 0$.

In the same way, sweeping through all K_{ij}^h from left to right and bottom to top, we can determine $u^h(\mathbf{x}_{ij}^h, t)$ uniquely for any pair (i, j) of nonnegative integers, $1 \leq i \leq M(h)$, $1 \leq j \leq N(h)$. The nodal values uniquely determine u^h on the whole of $\overline{\Omega}$.

In order to establish (23), we first note that for any k and l , $1 \leq k \leq M(h)$, $1 \leq l \leq N(h)$,

$$(24) \quad \|P^h v\|_{L_2(\Omega_{kl}^h)} = \|v\|_{l_2(\Omega_{kl}^h)} \quad \forall v \in L_2(\Omega),$$

$$(25) \quad \operatorname{Re} \left(\frac{\partial v}{\partial t}, \chi_{kl} P^h v \right) = \frac{1}{2} \frac{d}{dt} \|v\|_{l_2(\Omega_{kl}^h)}^2 \quad \forall v \in H_\omega^1(L_2(\Omega)).$$

Choosing $p = \chi_{kl} P^h v$ in (3), and employing (19), (24), and (25), we obtain

$$\begin{aligned} & \frac{d}{dt} \|u^h(t)\|_{l_2(\Omega_{kl}^h)}^2 - (1 + 8c_0|\mathbf{a}|) \|u^h(t)\|_{l_2(\Omega_{kl}^h)}^2 \\ & - c_0|\mathbf{a}|\|u^h(t)\|_{l_2^o(\Omega_{kl}^h)}^2 + c_2|\mathbf{a}|\|u^h(t)\|_{l_2(\partial_+\Omega_{kl}^h)}^2 \leq \|f(t)\|_{l_2(\Omega_{kl}^h)}^2. \end{aligned}$$

Letting $\omega \geq \frac{1}{2}(1+8c_0|\mathbf{a}|)$, multiplying by $\exp(-2\omega t)$, and integrating over the interval $(0, t)$, gives

$$\begin{aligned} & e^{-2\omega t} \|u^h(t)\|_{l_2(\Omega_{kl}^h)}^2 + c_2|\mathbf{a}| \int_0^t e^{-2\omega\tau} \|u^h(\tau)\|_{l_2(\partial_+\Omega_{kl}^h)}^2 d\tau \\ & \leq \|u^h(0)\|_{l_2(\Omega_{kl}^h)}^2 + \int_0^t e^{-2\omega\tau} \|f(\tau)\|_{l_2(\Omega_{kl}^h)}^2 d\tau \\ & + c_0|\mathbf{a}| \int_0^t e^{-2\omega\tau} \|u^h(\tau)\|_{l_2^o(\Omega_{kl}^h)}^2 d\tau. \end{aligned}$$

Now take the supremum over $t \geq 0$ to get

$$(26) \quad \begin{aligned} & \frac{1}{2} \sup_{t \geq 0} e^{-2\omega t} \|u^h(t)\|_{l_2(\Omega_{kl}^h)}^2 + \frac{1}{2} c_2 |\mathbf{a}| \int_0^\infty e^{-2\omega t} \|u^h(t)\|_{l_2(\partial_+\Omega_{kl}^h)}^2 dt \\ & \leq \|u^h(0)\|_{l_2(\Omega_{kl}^h)}^2 + \int_0^\infty e^{-2\omega t} \|f(t)\|_{l_2(\Omega_{kl}^h)}^2 dt \\ & + c_0 |\mathbf{a}| \int_0^\infty e^{-2\omega t} \|u^h(t)\|_{l_2^o(\Omega_{kl}^h)}^2 dt. \end{aligned}$$

Letting

$$\begin{aligned}
C &= 2\|u^h(0)\|_{l_2(\Omega)}^2 + 2 \int_0^\infty e^{-2\omega t} \|f(t)\|_{l_2(\Omega)}^2 dt, \\
\alpha_0 &= 0, \quad \beta_0 = 0, \\
\alpha_l &= \frac{2c_0}{c_2} \max_{0 \leq k \leq M(h)-1} \frac{m(K_{kl}^h \cup K_{k,l-1}^h)}{|\mathbf{x}_{k+1,l}^h - \mathbf{x}_{kl}^h|}, \\
\beta_k &= \frac{2c_0}{c_2} \max_{0 \leq l \leq N(h)-1} \frac{m(K_{kl}^h \cup K_{k-1,l}^h)}{|\mathbf{x}_{k,l+1}^h - \mathbf{x}_{kl}^h|}, \\
a_{kl} &= \sup_{t \geq 0} e^{-2\omega t} \|u^h(t)\|_{l_2(\Omega_{kl}^h)}^2, \\
b_{kl} &= c_2 |\mathbf{a}| \sum_{i=0}^{k-1} |\mathbf{x}_{i+1,l}^h - \mathbf{x}_{il}^h| \int_0^\infty e^{-2\omega t} |\mu_1 u^h(\mathbf{x}_{il}, t)|^2 dt, \\
c_{kl} &= c_2 |\mathbf{a}| \sum_{j=0}^{l-1} |\mathbf{x}_{k,j+1}^h - \mathbf{x}_{kj}^h| \int_0^\infty e^{-2\omega t} |\mu_2 u^h(\mathbf{x}_{kj}, t)|^2 dt
\end{aligned}$$

for $1 \leq k \leq M(h)$ and $1 \leq l \leq N(h)$, and setting $a_{k0} = a_{0l} = b_{k0} = b_{0l} = c_{k0} = c_{0l} = 0$ for $0 \leq k \leq M(h)$ and $0 \leq l \leq N(h)$, we can write (26) as

$$a_{kl} + b_{kl} + c_{kl} \leq C + \sum_{j=0}^{l-1} \alpha_j b_{kj} + \sum_{i=0}^{k-1} \beta_i c_{il}, \quad 0 \leq k \leq M(h), \quad 0 \leq l \leq N(h)$$

(empty sums are equal to zero). Applying Lemma 6, in tandem with Lemma 7, and letting $c_4 = 16c_0c_1^2(1 + 2c_0)/c_2$, we obtain

$$a_{kl} + b_{kl} + c_{kl} \leq C \cdot \exp\{c_4 + c_4 \exp(c_4)\}.$$

Recalling the definition of C , a_{kl} , b_{kl} , and c_{kl} , we obtain the inequality (23) with $c_3 = \exp\{c_4 + c_4 \exp(c_4)\}$. \square

On a tensor-product nonuniform mesh, H1 is automatically satisfied and, since each $K_{ij}^h \in \mathcal{T}^h$ is a rectangle, H2(i) also holds with $c_0 = 0$. In this case, Theorem 2 can be improved: stability can be shown to hold without assuming H2(ii). More precisely, we have the following result.

Theorem 3. Suppose that $\mathcal{F} = \{\mathcal{T}^h\}$ is a family of rectangular partitions of $\overline{\Omega}$ and let $\omega \geq \frac{1}{2}(1 + 8c_0|\mathbf{a}|)$ ($= \frac{1}{2}$). Then

$$\begin{aligned}
(27) \quad &\|u^h\|_{L_{\infty,\omega}(l_2(\Omega))}^2 + c_a |\mathbf{a}| \|u^h\|_{L_{2,\omega}(l_2(\partial_+\Omega))}^2 \\
&\leq 2\|u^h(0)\|_{l_2(\Omega)}^2 + 2\|f\|_{L_{2,\omega}(l_2(\Omega))}^2.
\end{aligned}$$

Proof. Choosing $p = P^h u^h$ in (3), noting that (20) implies

$$\operatorname{Re} B^h(v, P^h v) \geq \frac{1}{2} c_a |\mathbf{a}| \|v\|_{l_2(\partial_+\Omega)}^2 \quad \forall v \in \mathcal{U}_-^h,$$

and using (24) and (25), we obtain

$$\frac{1}{2} \frac{d}{dt} \|u^h(t)\|_{l_2(\Omega)}^2 + \frac{1}{2} c_a |\mathbf{a}| \|u^h(t)\|_{l_2(\partial_+\Omega)}^2 \leq \frac{1}{2} \|u^h(t)\|_{l_2(\Omega)}^2 + \frac{1}{2} \|f(t)\|_{l_2(\Omega)}^2$$

for all $t \geq 0$. Multiplying by $e^{-2\omega t}$ and integrating, we obtain (27). \square

4. CONVERGENCE

In §3, the stability of the finite volume method has been proved under the hypotheses H1 and H2. Here we investigate the accuracy of the scheme (3). We begin by stating some preliminary results.

Lemma 8 [3, Theorem 4.3.2]. *Let D and \widehat{D} be two bounded open subsets of \mathbb{R}^n such that $D = F(\widehat{D})$, where F is a sufficiently smooth bijection with a sufficiently smooth inverse $F^{-1}: D \rightarrow \widehat{D}$.*

Then, if the function $v: D \rightarrow \mathbf{C}$ belongs to the space $W_p^l(D)$ for some integer $l \geq 0$ and some $p \in [1, \infty]$, then the function $\hat{v} = v \circ F: \widehat{D} \rightarrow \mathbf{C}$ belongs to $W_p^l(\widehat{D})$, and there exists a constant C such that

$$\begin{aligned} \|\hat{v}\|_{L_p(\widehat{D})} &\leq \|J_{F^{-1}}\|_{L_\infty(D)}^{1/p} \|v\|_{L_p(D)}, \quad v \in L_p(D), \\ |\hat{v}|_{W_p^1(\widehat{D})} &\leq C \|J_{F^{-1}}\|_{L_\infty(D)}^{1/p} |F|_{W_\infty^1(\widehat{D})} |v|_{W_p^1(D)}, \quad v \in W_p^1(D), \\ |\hat{v}|_{W_p^2(\widehat{D})} &\leq C \|J_{F^{-1}}\|_{L_\infty(D)}^{1/p} (|F|_{W_\infty^1(\widehat{D})}^2 |v|_{W_p^2(D)} + |F|_{W_\infty^2(\widehat{D})} |v|_{W_p^1(D)}), \\ &\quad v \in W_p^2(D), \\ |\hat{v}|_{W_p^3(\widehat{D})} &\leq C \|J_{F^{-1}}\|_{L_\infty(D)}^{1/p} (|F|_{W_\infty^1(\widehat{D})}^3 |v|_{W_p^3(D)} + |F|_{W_\infty^1(\widehat{D})} |F|_{W_\infty^2(\widehat{D})} |v|_{W_p^2(D)} \\ &\quad + |F|_{W_\infty^3(\widehat{D})} |v|_{W_p^1(D)}), \\ &\quad v \in W_p^3(D). \end{aligned}$$

Lemma 9. *Suppose that the family of partitions $\mathcal{F} = \{\mathcal{T}^h\}$ satisfies H2, and, for $K \in \mathcal{T}^h$, let F_K denote the bilinear isoparametric mapping from the reference square $\widehat{K} = (0, 1)^2$ onto K . Then F_K is a bijection, and there exists a constant $C = C(c_0, c_1)$, independent of h_K , such that*

$$\begin{aligned} |F_K|_{W_\infty^1(\widehat{K})} &\leq Ch_K, \quad |F_K|_{W_\infty^2(\widehat{K})} \leq Ch_K^2, \quad |F_K|_{W_\infty^3(\widehat{K})} = 0, \\ |F_K^{-1}|_{W_\infty^1(K)} &\leq Ch_K^{-1}, \quad |J_{F_K}|_{W_\infty^1(\widehat{K})} \leq Ch_K^3, \\ \|J_{F_K}\|_{L_\infty(\widehat{K})} &\leq Ch_K^2, \quad \|J_{F_K^{-1}}\|_{L_\infty(K)} \leq Ch_K^{-2}. \end{aligned}$$

Proof. In §4.3 of [3], the same result is stated but assuming (4) and (5) instead of H2. According to Lemma 1, H2 implies both (4) and (5). Hence the result. \square

According to this lemma, a family of partitions \mathcal{F} satisfying hypothesis H2 is 1-strongly regular in the sense of Zlámal [17].

The next lemma is a simple consequence of the continuity of the trace operator $T: H^1(\widehat{K}) \rightarrow L_2(\partial\widehat{K})$ (cf. [1, Lemma 5.19]).

Lemma 10. *Let $\widehat{K} = (0, 1) \times (0, 1)$. Then there is a positive constant C such that for every $w \in H^3(\widehat{K})$*

$$|w|_{H^2(\partial\widehat{K})} \leq C(|w|_{H^2(\widehat{K})} + |w|_{H^3(\widehat{K})}).$$

The main result of this section is the following theorem.

Theorem 4. *Suppose that the family of partitions $\mathcal{F} = \{\mathcal{T}^h\}$ satisfies hypotheses H1 and H2 with $c_2 := c_a - 8c_0c_1^2(1 + 2c_0) > 0$, and let $u \in H_\omega^1(H^2(\Omega)) \cap$*

$L_{2,\omega}(H^3(\Omega))$, where $\omega \geq \frac{1}{2}(1 + 8c_0|\mathbf{a}|)$. Then

$$(28) \quad \begin{aligned} & \|u - u^h\|_{L_{\infty,\omega}(l_2(\Omega))} + \|u - u^h\|_{L_{2,\omega}(l_2(\partial_+\Omega))} \\ & \leq Ch^2(\|u\|_{H_\omega^1(H^2(\Omega))} + \|u\|_{L_{2,\omega}(H^3(\Omega))}), \end{aligned}$$

where $C = C(c_0, c_1, c_2, |\mathbf{a}|)$.

Proof. Let $\eta = u - I^h u$ and $\xi = I^h u - u^h$. Then $u - u^h = \eta + \xi$. We begin by estimating ξ . Clearly,

$$\left(\frac{\partial \xi}{\partial t}, p \right) + B^h(\xi, p) = - \left(\frac{\partial \eta}{\partial t} + \nabla \cdot (\mathbf{a}\eta), p \right)$$

for all p in \mathcal{M}^h . Thus, ξ is the solution of problem (3) with f replaced by $-(\partial \eta / \partial t + \nabla \cdot (\mathbf{a}\eta))$. By virtue of Theorem 2,

$$\begin{aligned} & \|\xi\|_{L_{\infty,\omega}(l_2(\Omega))}^2 + c_2|\mathbf{a}|\|\xi\|_{L_{2,\omega}(l_2(\partial_+\Omega))}^2 \\ & \leq C(c_0, c_1, c_2) \left(\|\xi(0)\|_{l_2(\Omega)}^2 + \left\| \frac{\partial \eta}{\partial t} + \nabla \cdot (\mathbf{a}\eta) \right\|_{L_{2,\omega}(l_2(\Omega))}^2 \right). \end{aligned}$$

Noting that $\|\xi(0)\|_{l_2(\Omega)} = \|\eta(0)\|_{l_2(\Omega)}$, and using the triangle inequality, yields

$$(29) \quad \begin{aligned} & \|u - u^h\|_{L_{\infty,\omega}(l_2(\Omega))} + \|u - u^h\|_{L_{2,\omega}(l_2(\partial_+\Omega))} \\ & \leq C(c_0, c_1, c_2, |\mathbf{a}|) \left(\|\eta\|_{L_{\infty,\omega}(l_2(\Omega))} + \|\eta\|_{L_{2,\omega}(l_2(\partial_+\Omega))} + \|\eta(0)\|_{l_2(\Omega)} \right. \\ & \quad \left. + \left\| \frac{\partial \eta}{\partial t} \right\|_{L_{2,\omega}(l_2(\Omega))} + \|\nabla \cdot (\mathbf{a}\eta)\|_{L_{2,\omega}(l_2(\Omega))} \right). \end{aligned}$$

The first four terms on the right are easily estimated by using the Bramble-Hilbert lemma [3, Theorem 4.1.3] and Lemma 8, together with the bounds in Lemma 9:

$$(30) \quad \|\eta\|_{L_{\infty,\omega}(l_2(\Omega))} \leq C(c_0, c_1)h^2\|u\|_{L_{\infty,\omega}(H^2(\Omega))},$$

$$(31) \quad \|\eta\|_{L_{2,\omega}(l_2(\partial_+\Omega))} \leq C(c_0, c_1)h^2\|u\|_{L_{2,\omega}(H^2(\partial_+\Omega))},$$

$$(32) \quad \|\eta(0)\|_{l_2(\Omega)} \leq C(c_0, c_1)h^2\|u_0\|_{H^2(\Omega)},$$

$$(33) \quad \left\| \frac{\partial \eta}{\partial t} \right\|_{L_{2,\omega}(l_2(\Omega))} \leq C(c_0, c_1)h^2\|u\|_{H_\omega^1(H^2(\Omega))},$$

where $h = \max_{K \in \mathcal{F}^h} h_K$.

The nontrivial part of the proof consists of estimating the last term on the right-hand side of (29). Thus, we consider

$$(34) \quad \frac{1}{m(K)} \int_K \nabla \cdot (\mathbf{a}\eta) dx = \frac{1}{m(K)} \int_{\partial K} J_{F_K}(DF_K)^{-1} \mathbf{a} \cdot \mathbf{n} \hat{\eta} d\hat{s}.$$

The expression on the right-hand side of (34) can be decomposed as follows:

$$(35) \quad \begin{aligned} & \frac{1}{m(K)} \int_{\partial K} J_{F_K}(DF_K)^{-1} \mathbf{a} \cdot \mathbf{n} \hat{\eta} d\hat{s} \\ & = \int_{\partial K} \left\{ \frac{J_{F_K}(\hat{s})}{m(K)} (DF_K)^{-1}(\hat{s}) - \frac{J_{F_K}(\hat{x}_0)}{m(K)} (DF_K)^{-1}(\hat{x}_0) \right\} \mathbf{a} \cdot \mathbf{n} \hat{\eta} d\hat{s} \\ & \quad + \int_{\partial K} \frac{J_{F_K}(\hat{x}_0)}{m(K)} (DF_K)^{-1}(\hat{x}_0) \mathbf{a} \cdot \mathbf{n} \hat{\eta} d\hat{s} \\ & \equiv T_1 + T_2, \end{aligned}$$

where \hat{x}_0 is an arbitrary but fixed (interior) point in \hat{K} . We begin by estimating T_1 . Let us define the following two matrices:

$$A = \frac{m(K)}{J_{F_K}(\hat{s})}(DF_K)(\hat{s}), \quad B = \frac{m(K)}{J_{F_K}(\hat{x}_0)}(DF_K)(\hat{x}_0).$$

Recalling the well-known inequality

$$\|A^{-1} - B^{-1}\| \leq \|A - B\| \|A^{-1}\| \|B^{-1}\|,$$

where $\|\cdot\|$ denotes the matrix norm subordinate to the Euclidean norm on \mathbf{R}^2 , and denoting by \hat{s}_0 the point $\hat{s} \in \partial\hat{K}$ at which $\|A^{-1} - B^{-1}\|$ (being a continuous function on $\partial\hat{K}$) reaches its maximum value, we obtain

$$\begin{aligned} |T_1| &\leq \left| \left[\frac{DF_K(\hat{x}_0)}{J_{F_K}(\hat{x}_0)} - \frac{DF_K(\hat{s}_0)}{J_{F_K}(\hat{s}_0)} \right] \right| \frac{J_{F_K}(\hat{x}_0) J_{F_K}(\hat{s}_0)}{m(K)} \\ &\quad \times \|[(DF_K)^{-1}(\hat{x}_0)]\| \|(DF_K)^{-1}(\hat{s}_0)\| |\mathbf{a}| \|\hat{\eta}\|_{L_1(\partial\hat{K})} \\ &\leq \{|J_{F_K}(\hat{s}_0) - J_{F_K}(\hat{x}_0)| |F_K|_{W_\infty^1(\hat{K})} \\ &\quad + |[DF_K(\hat{x}_0) - DF_K(\hat{s}_0)]| \|J_{F_K}\|_{L_\infty(\hat{K})}\} |\mathbf{a}| \frac{|F_K^{-1}|_{W_\infty^1(K)}^2}{m(K)} \|\hat{\eta}\|_{L_1(\partial\hat{K})}, \end{aligned}$$

because $(DF_K)^{-1}(\hat{x}) = D(F_K^{-1}(x))$. Now by virtue of Lemma 9,

$$\begin{aligned} (36) \quad |T_1| &\leq C(c_0, c_1) \frac{|\mathbf{a}|}{h_K^2 m(K)} \{ h_K |J_{F_K}(\hat{s}_0) - J_{F_K}(\hat{x}_0)| \\ &\quad + h_K^2 |[DF_K(\hat{x}_0) - DF_K(\hat{s}_0)]| \} \|\hat{\eta}\|_{L_1(\partial\hat{K})} \\ &\leq C(c_0, c_1) |\mathbf{a}| \frac{h_K^2}{m(K)} \|\hat{\eta}\|_{L_1(\partial\hat{K})}. \end{aligned}$$

By H2(ii),

$$|T_1| \leq C(c_0, c_1) |\mathbf{a}| \|\hat{\eta}\|_{L_1(\partial\hat{K})}.$$

The application of the Bramble-Hilbert lemma to the expression on the right gives

$$|T_1| \leq C(c_0, c_1) |\mathbf{a}| \|\hat{u}\|_{H^2(\partial\hat{K})},$$

and thence, by Lemma 10,

$$|T_1| \leq C(c_0, c_1) |\mathbf{a}| (\|\hat{u}\|_{H^2(\hat{K})} + \|\hat{u}\|_{H^3(\hat{K})}).$$

Returning to the original variables, using Lemmas 8 and 9, we obtain

$$(37) \quad |T_1| \leq C(c_0, c_1) |\mathbf{a}| h_K \|u\|_{H^3(K)},$$

which is our final bound on T_1 .

Let us consider

$$T_2 = \frac{J_{F_K}(\hat{x}_0)}{m(K)} (DF_K)^{-1}(\hat{x}_0) \mathbf{a} \cdot \int_{\partial\hat{K}} \mathbf{n} \hat{\eta} d\hat{s},$$

the second term on the right-hand side of (35). Define the vector

$$A := \frac{J_{F_K}(\hat{x}_0)}{m(K)} (DF_K)^{-1}(\hat{x}_0) \mathbf{a}.$$

Then, with A_1 and A_2 denoting the components of A ,

$$(38) \quad |T_2| = \left| A_1 \left(\int_{\partial \hat{K}_E} \hat{\eta} d\hat{x}_2 - \int_{\partial \hat{K}_W} \hat{\eta} d\hat{x}_2 \right) + A_2 \left(\int_{\partial \hat{K}_N} \hat{\eta} d\hat{x}_1 - \int_{\partial \hat{K}_S} \hat{\eta} d\hat{x}_1 \right) \right|,$$

where $\partial \hat{K}_E$, $\partial \hat{K}_W$ (resp. $\partial \hat{K}_N$, $\partial \hat{K}_S$) denote the East, West (resp. North, South) side of \hat{K} , oriented in the positive y (resp. positive x) direction. Applying the Bramble-Hilbert lemma, we obtain

$$\begin{aligned} \left| \int_{\partial \hat{K}_E} \hat{\eta} d\hat{x}_2 - \int_{\partial \hat{K}_W} \hat{\eta} d\hat{x}_2 \right| &\leq C \left\| \frac{\partial^3 \hat{u}}{\partial \hat{x}_1 \partial \hat{x}_2^2} \right\|_{L_2(\hat{K})}, \\ \left| \int_{\partial \hat{K}_N} \hat{\eta} d\hat{x}_1 - \int_{\partial \hat{K}_S} \hat{\eta} d\hat{x}_1 \right| &\leq C \left\| \frac{\partial^3 \hat{u}}{\partial \hat{x}_1^2 \partial \hat{x}_2} \right\|_{L_2(\hat{K})}, \end{aligned}$$

where $C = C(c_0, c_1)$. Changing variables, using Lemmas 8 and 9, and employing H2(ii), we get

$$(39) \quad \left| \int_{\partial \hat{K}_E} \hat{\eta} d\hat{x}_2 - \int_{\partial \hat{K}_W} \hat{\eta} d\hat{x}_2 \right| \leq Ch_K^2 \|u\|_{H^3(K)},$$

$$(40) \quad \left| \int_{\partial \hat{K}_N} \hat{\eta} d\hat{x}_1 - \int_{\partial \hat{K}_S} \hat{\eta} d\hat{x}_1 \right| \leq Ch_K^2 \|u\|_{H^3(K)}.$$

Using again Lemma 9 and hypothesis H2(ii) gives

$$(41) \quad |A_1| \leq C|\mathbf{a}|h_K^{-1}, \quad |A_2| \leq C|\mathbf{a}|h_K^{-1}.$$

Thus, from (38)–(41) we obtain

$$(42) \quad |T_2| \leq C|\mathbf{a}|h_K \|u\|_{H^3(K)}.$$

Now (37) and (42) provide the desired bound for the expression appearing on the right-hand side of (34):

$$(43) \quad \left| \frac{1}{m(K)} \int_{\partial \hat{K}} J_{F_K}(DF_K)^{-1} \mathbf{a} \cdot \mathbf{n} \hat{\eta} d\hat{s} \right| \leq C|\mathbf{a}|h_K \|u\|_{H^3(K)}.$$

From (34) and (43) we get

$$\left| \frac{1}{m(K)} \int_K \nabla \cdot (\mathbf{a}\eta) dx \right| \leq C|\mathbf{a}|h_K \|u\|_{H^3(K)},$$

and so

$$(44) \quad \|\nabla \cdot (\mathbf{a}\eta)\|_{l_2(\Omega)} \leq Ch^2 |\mathbf{a}| \|u\|_{H^3(\Omega)},$$

where $h = \max_{K \in \mathcal{T}^h} h_K$. Substituting (30)–(33) and (44) into (29), we obtain (28). \square

For a family of tensor-product nonuniform partitions, hypotheses H1 and H2(i) are automatically fulfilled, and the error estimate (28) holds with $c_0 = 0$. However, the constant C appearing on the right-hand side of the estimate is still dependent on c_1 , which might suggest that allowing the constituent rectangles in the partition to be thin and elongated damages the accuracy of the scheme. In the next theorem we show that this is not the case by removing the dependence of C on c_1 .

Theorem 5. Suppose that $\mathcal{T} = \{\mathcal{T}^h\}$ is a family of rectangular partitions of $\overline{\Omega}$, and let $u \in H_\omega^1(H^2(\Omega)) \cap L_{2,\omega}(H^3(\Omega))$, where $\omega \geq \frac{1}{2}$. Then

$$(45) \quad \begin{aligned} & \|u - u^h\|_{L_{\infty,\omega}(l_2(\Omega))} + \|u - u^h\|_{L_{2,\omega}(l_2(\partial_+\Omega))} \\ & \leq C(|\mathbf{a}|)h^2(|u|_{H_\omega^1(H^2(\Omega))} + |u|_{L_{2,\omega}(H^3(\Omega))}). \end{aligned}$$

Proof. Following the same route as in the beginning of the proof of Theorem 4, but using Theorem 3 instead of Theorem 2, we conclude that

$$(46) \quad \begin{aligned} & \|u - u^h\|_{L_{\infty,\omega}(l_2(\Omega))} + \|u - u^h\|_{L_{2,\omega}(l_2(\partial_+\Omega))} \\ & \leq C(|\mathbf{a}|) \left(\|\eta\|_{L_{\infty,\omega}(l_2(\Omega))} + \|\eta\|_{L_{2,\omega}(l_2(\partial_+\Omega))} \right. \\ & \quad \left. + \|\eta(0)\|_{l_2(\Omega)} + \left\| \frac{\partial \eta}{\partial t} \right\|_{L_{2,\omega}(l_2(\Omega))} + \|\nabla \cdot (\mathbf{a}\eta)\|_{L_{2,\omega}(l_2(\Omega))} \right). \end{aligned}$$

It remains to estimate the terms on the right-hand side. Consider $\|\eta\|_{l_2(\Omega)}$, and note that

$$(47) \quad \frac{1}{m(K)} \int_K \eta \, dx = \int_{\widehat{K}} \hat{\eta} \, d\hat{x}.$$

Using the Bramble-Hilbert lemma, we get

$$\left| \int_{\widehat{K}} \hat{\eta} \, d\hat{x} \right| \leq C |\hat{u}|_{H^2(\widehat{K})}.$$

Returning to our original variables, we obtain

$$\left| \int_{\widehat{K}} \hat{\eta} \, d\hat{x} \right| \leq C[m(K)]^{-1/2} h_K^2 |u|_{H^2(K)}.$$

Thus, by (47) and the definition of the norm $\|\cdot\|_{l_2(\Omega)}$,

$$(48) \quad \|\eta\|_{l_2(\Omega)} \leq Ch^2 |u|_{H^2(\Omega)},$$

where C is a uniform constant. Similarly,

$$(49) \quad \|\eta\|_{l_2(\partial_+\Omega)} \leq Ch^2 |u|_{H^2(\partial_+\Omega)},$$

$$(50) \quad \left\| \frac{\partial \eta}{\partial t} \right\|_{l_2(\Omega)} \leq Ch^2 \left| \frac{\partial u}{\partial t} \right|_{H^2(\Omega)}.$$

The last term on the right-hand side of (46) is handled as follows:

$$(51) \quad \begin{aligned} \frac{1}{m(K)} \int_K \nabla \cdot (\mathbf{a}\eta) \, dx &= \frac{a_1}{m(K)} \int_K \frac{\partial \eta}{\partial x_1} \, dx + \frac{a_2}{m(K)} \int_K \frac{\partial \eta}{\partial x_2} \, dx \\ &= \frac{a_1}{h_K^{(1)}} \int_{\widehat{K}} \frac{\partial \hat{\eta}}{\partial \hat{x}_1} \, d\hat{x} + \frac{a_2}{h_K^{(2)}} \int_{\widehat{K}} \frac{\partial \hat{\eta}}{\partial \hat{x}_2} \, d\hat{x}, \end{aligned}$$

where $h_K^{(1)}$ and $h_K^{(2)}$ denote the lengths of the horizontal and vertical sides of the rectangle K , respectively. By virtue of the Bramble-Hilbert lemma,

$$\left| \int_{\widehat{K}} \frac{\partial \hat{\eta}}{\partial \hat{x}_1} \, d\hat{x} \right| \leq C \left| \frac{\partial \hat{u}}{\partial \hat{x}_1} \right|_{H^2(\widehat{K})}, \quad \left| \int_{\widehat{K}} \frac{\partial \hat{\eta}}{\partial \hat{x}_2} \, d\hat{x} \right| \leq C \left| \frac{\partial \hat{u}}{\partial \hat{x}_2} \right|_{H^2(\widehat{K})},$$

where C is an absolute constant. Returning to our original variables, we get

$$(52) \quad \left| \int_{\hat{K}} \frac{\partial \hat{\eta}}{\partial \hat{x}_1} d\hat{x} \right| \leq C[m(K)]^{-1/2} \{(h_K^{(1)})^3 + h_K^{(1)}(h_K^{(2)})^2\} |u|_{H^3(K)},$$

$$(53) \quad \left| \int_{\hat{K}} \frac{\partial \hat{\eta}}{\partial \hat{x}_2} d\hat{x} \right| \leq C[m(K)]^{-1/2} \{(h_K^{(2)})^3 + h_K^{(2)}(h_K^{(1)})^2\} |u|_{H^3(K)}.$$

From (51)–(53) we obtain that

$$\left| \frac{1}{m(K)} \int_K \nabla \cdot (\mathbf{a}\eta) dx \right| \leq C|\mathbf{a}|[m(K)]^{-1/2} h_K^2 |u|_{H^3(K)}.$$

Hence,

$$(54) \quad \|\nabla \cdot (\mathbf{a}\eta)\|_{L_2(\Omega)} \leq C|\mathbf{a}|h^2 |u|_{H^3(\Omega)}.$$

Substituting (48)–(50) and (54) into (46), we obtain (45). \square

This result is consistent with the experimental evidence presented in [13], which suggests that the accuracy of the cell vertex scheme is insensitive to mesh stretching in the coordinate directions.

5. CONCLUSION

We have demonstrated that the cell vertex finite volume method for a time-dependent linear hyperbolic equation in two spatial dimensions is second-order accurate on a quadrilateral partition, provided that each quadrilateral is an $O(h^2)$ perturbation of a parallelogram, and that the partition is regular in the usual sense. Moreover, it has been shown that on a tensor-product nonuniform partition, second-order accuracy can be maintained without assuming regularity. Similar results hold for cell center finite volume approximations of elliptic equations [16] and cell vertex approximations of steady hyperbolic equations [10, 15]. The extension of the developments presented in this paper to linear hyperbolic systems will be a subject of future investigation.

BIBLIOGRAPHY

1. R. A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.
2. J.-P. Aubin and I. Ekeland, *Applied nonlinear analysis*, Wiley-Interscience, New York, 1988.
3. P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, New York, Oxford, 1978.
4. M. Giles, *Accuracy of node based solutions on irregular meshes*, Proceedings of the 11th International Conference on Numerical Methods in Fluid Dynamics (D. L. Dwyer, M. Y. Hussaini, and R. G. Voigt, eds.), Springer-Verlag, Berlin, 1988, pp. 273–277.
5. A. Jameson, W. Schmidt, and E. Turkel, *Numerical solutions of the Euler equations by finite volume methods using Runge-Kutta time-marching schemes*, Paper 81-1259, AIAA, New York, 1981.
6. P. Lesaint and P.-A. Raviart, *Finite element collocation methods for first order systems*, Math. Comp. **33** (1979), 891–918.
7. R. W. MacCormack and A. J. Paullay, *Computational efficiency achieved by time splitting of finite difference operators*, Paper 72-154, AIAA, New York, 1972.
8. P. W. McDonald, *The computation of transonic flow through two-dimensional gas turbine cascades*, Paper 71-GT-89, ASME, New York, 1971.
9. K. W. Morton and M. F. Paisley, *A finite volume scheme with shock fitting for the steady Euler equations*, J. Comp. Phys. **80** (1989), 168–203.

10. K. W. Morton and E. Süli, *Finite volume methods and their analysis*, IMA J. Numer. Anal. **11** (1991), 241–260.
11. R.-H. Ni, *A multiple grid scheme for solving the Euler equations*, AIAA J. **20** (1981), 1565–1571.
12. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York, 1983.
13. P. Roe, *The influence of mesh quality on solution accuracy*, preprint, Cranfield Institute of Technology, 1989.
14. ———, *Error estimates for cell-vertex solutions of the compressible Euler equations*, ICASE Report, No. 87-6, 1987.
15. E. Süli, *The accuracy of finite volume methods on distorted partitions*, Mathematics of Finite Elements and Applications VII (J. R. Whiteman, ed.), Academic Press, 1991, pp. 253–260.
16. ———, *Convergence of finite volume schemes for Poisson’s equation on nonuniform meshes*, SIAM J. Numer. Anal. **28** (1991), 1419–1430.
17. M. Zlámal, *Superconvergence and reduced integration in the finite element method*, Math. Comp. **32** (1978), 663–685.

OXFORD UNIVERSITY COMPUTING LABORATORY, NUMERICAL ANALYSIS GROUP, OXFORD OX1
3QD, GREAT BRITAIN

E-mail address: endre.suli@uk.ac.ox.comlab