CONVERGENCE OF A CELL-CENTERED FINITE VOLUME METHOD AND APPLICATION TO ELLIPTIC EQUATIONS

GUNG-MIN GIE AND ROGER TEMAM

Abstract. We study the consistency and convergence of the cell-centered Finite Volume (FV) external approximation of $H_0^1(\Omega)$, where a 2D polygonal domain Ω is discretized by a mesh of convex quadrilaterals. The discrete FV derivatives are defined by using the so-called Taylor Series Expansion Scheme (TSES). By introducing the Finite Difference (FD) space associated with the FV space, and comparing the FV and FD spaces, we prove the convergence of the FV external approximation by using the consistency and convergence of the FD method. As an application, we construct the discrete FV approximation of some typical elliptic equations, and show the convergence of the discrete FV approximations to the exact solutions.

Key words. Finite Volume method, Taylor Series Expansion Scheme (TSES), convergence and stability, convex quadrilateral meshes.

1. Introduction

In engineering, fluid dynamics, and physics more recently, the Finite Volume (FV) discretization method is widely used because of its local conservation property of the flux on each control volume. From the numerical analysis point of view, many different types of FV methods, depending on the way of computing the discrete fluxes, have been introduced and analyzed up to this day. Concerning the varieties of the FV methods and their applications, we refer the readers to, e.g., [52, 32, 51, 18, 24] for general references, and to [36, 1, 37, 47, 58, 60, 48, 38, 40, 49, 11] for the computational applications.

In proving the convergence of the cell-centered FV method, one specific difficulty is due to the *weak consistency* of the FV method. Namely, the companion discrete FV derivative arising in the discrete integration by parts does not usually converge strongly to the corresponding derivative of the limit function. To overcome this technical difficulty, in an important earlier work, the authors of [32] employed a discrete compactness argument for the FV space, even for linear problems. Since then, using this approach, further analysis of the cell-centered FV method has been made in, e.g, [28, 27, 12, 31, 41, 8, 14]. A different approach was introduced in our earlier works [39, 42] to prove the convergence of the cell-centered FV method. More precisely, we introduced there the Finite Difference (FD) space which is associated with the FV space, and compared the FV and FD spaces by defining a map between them. Then, thanks to the consistency and convergence of the FD method which are proven in a classical way, the convergence of the FV method is inferred. This approach was conducted in [39, 42] for the study of the cell-centered FV method when the domain considered has a rectangular mesh, whereas more general meshes are desirable for FV which are specifically aimed at handling complicated geometries. For a different type of FV methods other than the cell-centered FV, the convergence of, e.g., the cell-vertex FV method is well-studied in, e.g., [54, 55, 56, 10, 53, 59].

There is a broad class of FV methods which correspond to various equations and applications, and to various strategies of approximation. After choosing, for a given mesh, the nodal points and the reconstructed functions, one needs to define an approximation of the derivatives (which is not easy on a general mesh). For elliptic problems which admit a weak (variational) formulation, the gradient schemes (studied in, e.g., [34, 26]) consist in mimicking the variational formulation by replacing the exact derivatives by the approximate derivatives. Another more general direction applying to all classes of equations and conservation laws, consists in integrating the conservation law on the control volume and then looking for approximations of the fluxes. Our approach relates to the gradient schemes. Using cell-centered unknowns, we approach the spatial derivatives using the so-called Taylor Series Expansion Scheme (TSES) method that was introduced in the early 90's in the engineering literature, see, e.g., [46, 52, 45, 18]. Note that the TSES is commonly considered in the engineering fluid mechanics, whereas the MPFA approach, [9, 3], not studied here is considered in the petroleum and hydrogeology literatures. We can then define approximate variational problems and study the convergence of the approximate solutions to those of the exact ones. The construction of the TSES method is closely related to, e.g., that of the so-called *diamond scheme* in [22] or the Discrete Duality Finite Volume scheme in [44, 23] according to this terminology which was subsequently introduced; see Remark 3.2 below. See other related works as well in, e.g., [43, 30, 16, 2].

As we said, there is a substantial body of work related to the numerical analysis of the FV methods; see, e.g., the review articles [32], and more recently [34, 26]; see also, e.g., [31, 41] and the references quoted in these articles. Despite their importance and major interest, the existing works deal with objectives different than ours and do not cover our objectives. These works are generally motivated by reservoir (underground) flows and address the corresponding equations, and, on the mathematical side, they use compactness arguments to prove convergence, even for linear equations. Due to the growing importance of FV methods, and the considerable difficulties for proving their convergence, it is clear that there will be many more works in years to come on the numerical analysis of FV methods, and there is need to diversify the available tools. This article, like earlier works [39, 42], is generally motivated by classical or geophysical fluid mechanics; it uses a form of the FV method, the TSES method, which is not dealt with in the review articles previously mentioned; and it uses the comparison with a related Finite Difference method instead of compactness arguments. Another major difference between prior works and this article is that, in, e.g., [32, 34, 26], the FV method and its analysis is taylored to one specific equation in divergence form and, as far as we understand, the work needs to be redone or suitably adapted if we consider a different equation with, e.g., lower order terms as in equation (124) in this article. On the contrary, our approach consists in approximating the underlying function space of type $H^1(\Omega)$, leaving all flexibility for the equations whose coefficients can be nonhomogeneous and nonisotropic. Finally it is noteworthy that [24] emphasizes the use of the maximum principle which is mostly not relevant to classical or geophysical fluid mechanics (nor to multi-species underground reservoir flows which produce systems).

In this article, to prove the consistency and convergence of the TSES Finite Volume approximation of $H_0^1(\Omega)$, (which is equivalent to verifying the properties $(\mathcal{C}1)$ and $(\mathcal{C}2)$ below), we impose some conditions $(\mathcal{H}1)$-$(\mathcal{H}5)$ on the mesh. $(\mathcal{H}1)$-$(\mathcal{H}3)$ are standard hypotheses which guarantee that the mesh is not too distorted.

The hypotheses ($\mathcal{H}4$) and ($\mathcal{H}5$) on the mesh are specific to the discretization that we consider and comparable to similar hypotheses made in the literature. The hypothesis ($\mathcal{H}4$) is local while the hypothesis ($\mathcal{H}5$) is not too complicated, it relates to the function space that we approach and is then valid for all the corresponding equations, unlike some other similar hypotheses in the literature which relate the mesh to a particular equation.

Our work is organized as follows: We recall some elementary geometric notations in Section 2. Then we construct the discrete FV space in Section 3, and introduce the external approximation of $H_0^1(\Omega)$ by the FV space in Section 4, where the property ($\mathcal{C}1$) for the FV scheme is verified as well. As briefly explained before, due to the weak consistency for the FV method, we first introduce the FD space associated with the FV space in Section 5, before we verify the property ($\mathcal{C}2$) for the FV. Then the convergence of the FD approximation is proved in Section 6. Finally, by comparing the FV and FD spaces and using the convergence of the FD approximation, we finally obtain the convergence of the FV external approximation in Section 7. As an application of the convergent cell-centered FV approximation, we demonstrate in Section 8 how one can use the FV scheme to approximate the weak solution of some typical elliptic equations. The convergence of the discrete FV weak solution to the exact weak solution is proved as well.

2. Notations and preliminaries

For any point (x_*, y_*) in $\mathbb{R}^2$, we write $\boldsymbol{P}_* = (x_*, y_*)$. A vector from a point $\boldsymbol{P}_A$ to a point $\boldsymbol{P}_B$ is written as

$$\overrightarrow{\boldsymbol{P}_A \boldsymbol{P}_B} = (x_B - x_A, y_B - y_A). \tag{1}$$

Let K be a convex quadrilateral with the four vertices $\boldsymbol{P}_A$, $\boldsymbol{P}_B$, $\boldsymbol{P}_C$, and $\boldsymbol{P}_D$ which are ordered counter clockwise. Then the area of K, denoted by $|K|$, is classically written in the form,

$$|K| = \frac{1}{2}|\overrightarrow{\boldsymbol{P}_A \boldsymbol{P}_C} \times \overrightarrow{\boldsymbol{P}_B \boldsymbol{P}_D}| = \frac{1}{2}\left|\det\begin{pmatrix} x_C - x_A & x_D - x_B \\ y_C - y_A & y_D - y_B \end{pmatrix}\right|. \tag{2}$$

In the use of (2) below, a sign issue will occur, and to avoid it, we assume that the convex quadrilateral is not too distorted in the sense that the projection of each side onto the opposite side has a nonempty intersection with that side. Under this assumption, the determinant in (2) is always positive, and hence we write

$$|K| = \frac{1}{2}\det\begin{pmatrix} x_C - x_A & x_D - x_B \\ y_C - y_A & y_D - y_B \end{pmatrix}. \tag{3}$$

For the FV and the corresponding FD meshes in this article, thanks to the restrictions ($\mathcal{H}2$) and ($\mathcal{H}3$) below, we will mainly use the formula (3) to compute the area of certain convex quadrilateral cells.

The area of a triangle T with vertices $\boldsymbol{P}_A$, $\boldsymbol{P}_B$, and $\boldsymbol{P}_C$, ordered counter clockwise, is given by

$$|T| = \frac{1}{2}|\overrightarrow{\boldsymbol{P}_A \boldsymbol{P}_B} \times \overrightarrow{\boldsymbol{P}_A \boldsymbol{P}_D}| = \frac{1}{2}\left|\det\begin{pmatrix} x_B - x_A & x_D - x_A \\ y_B - y_A & y_D - y_A \end{pmatrix}\right|. \tag{4}$$

In this article, we denote by κ a generic constant, depending on the domain Ω and the other data, but independent of the mesh size. When we want to keep track of such a constant, we number it as κ_i.

3. Cell-centered Finite Volume setting

We consider the discretization of a 2D polygonal domain Ω by MN convex quadrilateral control volumes $K_{i,j}$, $1 \leq i \leq M$, $1 \leq j \leq N$, see, e.g., Figure 1 below, so that $\Omega = \cup_{i=1}^{M} \cup_{j=1}^{N} K_{i,j}$. Here the FV mesh is topologically equivalent to a rectangular mesh. The interiors of the $K_{i,j}$ are disjoint, but of course two adjacent control volumes share a (full) common edge.
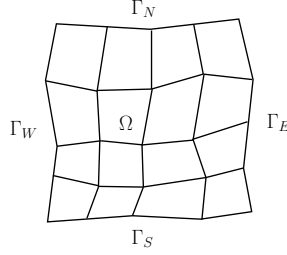

FIGURE 1. A polygonal domain Ω with boundary $\Gamma = \cup_{i=E,W,N,S} \Gamma_i$, which is discretized by convex quadrilaterals.

For each control volume $K_{i,j}$, we define the FV nodal points $\boldsymbol{P}_{i\pm1/2,\,j\pm1/2}$ as the corner points of $K_{i,j}$ as in Fig. 2 below. Then we obtain the $(M+1) \times (N+1)$ nodal points $\boldsymbol{P}_{i+1/2,\,j+1/2}$ such that

(5)
$$\boldsymbol{P}_{i+\frac{1}{2},\,j+\frac{1}{2}} = \left(x_{i+\frac{1}{2},\,j+\frac{1}{2}},\, y_{i+\frac{1}{2},\,j+\frac{1}{2}}\right) \in \begin{cases} \text{int } \Omega, & 1 \leq i \leq N-1,\ 1 \leq j \leq M-1, \\ \Gamma_W, & i = 0, \\ \Gamma_E, & i = M, \\ \Gamma_S, & j = 0, \\ \Gamma_N, & j = N. \end{cases}$$
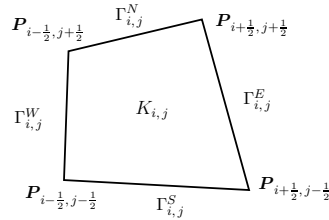

FIGURE 2. Control volume $K_{i,j}$ with the vertices $\boldsymbol{P}_{i\pm1/2,\,j\pm1/2}$ and the boundary $\Gamma_{i,j} = \cup_{k=E,W,N,S} \Gamma_{i,j}^{k}$.

On the boundary Γ of Ω, we define the flat control volumes:

(6)
$$\begin{cases} K_{0,j} = \text{ segment connecting } \boldsymbol{P}_{\frac{1}{2},\,j-\frac{1}{2}} \text{ and } \boldsymbol{P}_{\frac{1}{2},\,j+\frac{1}{2}},\, 1 \leq j \leq N, \\[2mm] K_{M+1,j} = \text{ segment connecting } \boldsymbol{P}_{M+\frac{1}{2},\,j-\frac{1}{2}} \text{ and } \boldsymbol{P}_{M+\frac{1}{2},\,j+\frac{1}{2}},\, 1 \leq j \leq N, \\[2mm] K_{i,0} = \text{ segment connecting } \boldsymbol{P}_{i-\frac{1}{2},\,\frac{1}{2}} \text{ and } \boldsymbol{P}_{i+\frac{1}{2},\,\frac{1}{2}},\, 1 \leq i \leq M, \\[2mm] K_{i,N+1} = \text{ segment connecting } \boldsymbol{P}_{i-\frac{1}{2},\,N+\frac{1}{2}} \text{ and } \boldsymbol{P}_{i+\frac{1}{2},\,N+\frac{1}{2}},\, 1 \leq i \leq M. \end{cases}$$

For convenience, we set

(7)
$$K_{0,0} = K_{M+1,0} = K_{0,N+1} = K_{M+1,N+1} = \emptyset.$$

Then we write the closure of Ω as the union of the control volumes:

(8)
$$\overline{\Omega} = \bigcup_{i=0}^{M+1} \bigcup_{j=0}^{N+1} K_{i,j}.$$

We introduce the barycenter of $K_{i,j}$:

(9) $\quad \boldsymbol{P}_{i,j} = \begin{cases} \text{midpoint of } K_{i,j}, \text{ for } i = 0 \text{ or } M+1, \text{ or } j = 0 \text{ or } N+1, \\ \text{barycenter of } K_{i,j}, \text{ for } 1 \le i \le M, \ 1 \le j \le N. \end{cases}$

Since each control volume is convex, we can write the barycenter $\boldsymbol{P}_{i,j}$ of $K_{i,j}$ as an interpolation of the four corners of $K_{i,j}$. That is, for $1 \le i \le M, 1 \le j \le N$,
(10)
$$\boldsymbol{P}_{i,j} = \sum_{l,m=\pm 1} \lambda_{l,m}^{i,j} \boldsymbol{P}_{i+\frac{l}{2},j+\frac{m}{2}} \quad \text{for some } \lambda_{l,m}^{i,j} \ge 0 \text{ such that } \sum_{j,m=\pm 1} \lambda_{l,m}^{i,j} = 1.$$

We write the boundary $\Gamma_{i,j}$ of each control volume $K_{i,j}$ in the form,

(11)
$$\Gamma_{i,j} = \bigcup_{k=E,W,N,S} \Gamma_{i,j}^k, \quad (\text{see Fig. 2}).$$

We name the four internal angles of $K_{i,j}$ as $\theta_{i,j}^m$, $m = WS, ES, EN, WN$, with an obvious notation.

When M and N get large, we assume that the number of points in each direction remains comparable by imposing the analytic hypothesis below:

($\mathcal{H}1$) There exists $0 < \kappa_0 < 1$ such that

$$\kappa_0 \le \frac{M}{N} \le \kappa_0^{-1} \quad \text{as} \quad M, N \to \infty.$$

At each discretization level M and N, we consider the maximum and minimum lengthes of the edges of all the control volumes and assume that there exist $0 < \underline{h} \le \overline{h}$ such that

(12)
$$\underline{h} \le \min_{i,j} \min_{k=E,W,N,S} \left| \Gamma_{i,j}^k \right| \le \max_{i,j} \max_{k=E,W,N,S} \left| \Gamma_{i,j}^k \right| \le \overline{h},$$

where $\left| \Gamma_{i,j}^k \right|$ denotes the measure of $\Gamma_{i,j}^k$. We consider also the maximum and minimum sizes of the internal angles of all the control volumes and assume similarly that there exist $0 < \underline{\theta} \le \overline{\theta} < \pi$ such that

(13)
$$\underline{\theta} \le \min_{i,j} \min_{m=WS,ES,EN,WN} \theta_{i,j}^m \le \max_{i,j} \max_{m=WS,ES,EN,WN} \theta_{i,j}^m \le \overline{\theta}.$$

Following the suggestion in, e.g., [19, 57] and other references therein, we assume that the FV mesh is not highly distorted, that is:

($\mathcal{H}2$) There exists δ, $(2\sqrt{3})/9 \, (\approx 0.384) \le \delta < 1$ such that

(14)
$$\min \left(\sin \overline{\theta}, \sin \underline{\theta} \right) \ge \delta \frac{\overline{h}^2}{\underline{h}^2}.$$

Using ($\mathcal{H}2$) and by writing $|\boldsymbol{ABC}|$ the area of the triangle with vertices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, we find that
(15)
$$\frac{|K_{i,j}|}{|K_{k,l}|} \le \frac{|\boldsymbol{P}_{i-\frac{1}{2},j-\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}} \boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}}| + |\boldsymbol{P}_{i-\frac{1}{2},j-\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2},j-\frac{1}{2}}|}{|\boldsymbol{P}_{k-\frac{1}{2},l-\frac{1}{2}} \boldsymbol{P}_{k+\frac{1}{2},l+\frac{1}{2}} \boldsymbol{P}_{k-\frac{1}{2},l+\frac{1}{2}}| + |\boldsymbol{P}_{k-\frac{1}{2},l-\frac{1}{2}} \boldsymbol{P}_{k+\frac{1}{2},l+\frac{1}{2}} \boldsymbol{P}_{k+\frac{1}{2},l-\frac{1}{2}}|} \le \delta^{-1},$$

for $1 \leq i, k \leq M$ and $1 \leq j, l \leq N$. Moreover, since there are at most M (or N) control volumes in the horizonal (or vertical) direction, using ($\mathcal{H}2$), we notice that

$$(16) \qquad \max(M\overline{h},\, N\overline{h}) \leq \frac{\overline{h}}{\underline{h}}\, \max(M\underline{h},\, N\underline{h}) \leq \kappa_1 := \delta^{-\frac{1}{2}} \max_{k=E,W,N,S} \left|\Gamma_k\right|.$$

We infer from the lower bound of δ in ($\mathcal{H}2$) that

$$(17) \qquad\qquad\qquad\qquad \cos\underline{\theta} \leq \frac{h}{\overline{\overline{h}}}.$$

Thanks to (17), the area of each control volume $K_{i,j}$ is written in the form in (3). Hence the formula (3) can be made useful when we compute the area of $K_{i,j}$ (or $K_{i+1/2,j}$, $K_{i,j+1/2}$ in (25) below). The explicit expressions of the areas are given in (36) below. In addition, (17) also implies that

(18)
> If $K_{i,j}$ and $K_{i',j'}$ are two adjacent control volumes $(i - i' = \pm 1$ and $j = j'$, or $i = i'$ and $j - j' = \pm 1)$, then the vector $\overrightarrow{\boldsymbol{P}_{i,j}\boldsymbol{P}_{i',j'}}$ intersects the common boundary of $K_{i,j}$ and $K_{i',j'}$ (see Fig. 3 below),

and

(19)
> $\boldsymbol{P}_{i+1/2,\,j+1/2}$ is located inside the quadrilateral with vertices $\boldsymbol{P}_{i,j}$, $\boldsymbol{P}_{i+1,j}$, $\boldsymbol{P}_{i+1,j+1}$, and $\boldsymbol{P}_{i,j+1}$ for $1 \leq i \leq M - 1$, $1 \leq j \leq N - 1$ (see Fig. 3 below).

Thanks to (19), there exist $\mu_{l,m}^{i+1/2,\,j+1/2} \geq 0$ such that $\sum_{l,m=0,1} \mu_{l,m}^{i+1/2,\,j+1/2} = 1$ and

$$(20) \quad \boldsymbol{P}_{i+\frac{1}{2},\,j+\frac{1}{2}} = \sum_{l,m=0,1} \mu_{l,m}^{i+\frac{1}{2},\,j+\frac{1}{2}} \boldsymbol{P}_{i+l,\,j+m}, \quad 1 \leq i \leq M - 1,\ 1 \leq j \leq N - 1.$$

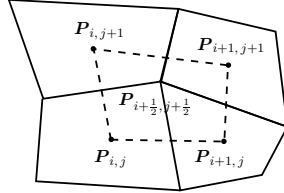This expression (20) will play an important role below when we define the discrete FV derivatives.

FIGURE 3. A nodal point $\boldsymbol{P}_{i+1/2,\,j+1/2}$ and the nearby barycenters.

Construction of the FV space. We now define the FV space of step functions that satisfy the homogeneous Dirichlet boundary condition in the form,

$$(21) \qquad V_h := \left\{ \begin{array}{l} \text{step functions } u_h \text{ on } \overline{\Omega} = \cup_{i=0}^{M+1} \cup_{j=0}^{N+1} K_{i,j} \text{ such that} \\[2mm] u_h\big|_{K_{i,j}} = \left\{ \begin{array}{l} u_{i,j}, \text{ for } 1 \leq i \leq M,\ 1 \leq j \leq N, \\[1mm] 0, \text{ for } i = 0, M+1,\ \text{or } j = 0, N+1. \end{array} \right. \end{array} \right\}.$$

Then, for any $u_h \in V_h$, we write

$$(22) \qquad\qquad\qquad\qquad u_h = \sum_{i=1}^{M} \sum_{j=1}^{N} u_{i,j}\, \chi_{K_{i,j}},$$

where $\chi_{K_{i,j}}$ is the characteristic function of $K_{i,j}$.

Remark 3.1. A Neumann boundary condition on, e.g., Γ_E can be implemented by setting

$$(23) \qquad\qquad u_{0,j} = u_{1,j}, \quad 0 \le j \le N+1.$$

Similarly, if the FV mesh in (5) is periodic near, e.g., Γ_W and Γ_E (that is, $K_{1,j}$ is the horizontal reflection of $K_{M,j}$ for $1 \le j \le N$), then one can enforce the periodic boundary condition on $\Gamma_W = \Gamma_E$ by setting

$$(24) \qquad\qquad u_{0,j} = u_{M+1,j} := \frac{1}{2}\big(u_{1,j} + u_{M,j}\big).$$

Using (23) and (24), many other mixed type boundary conditions can be implemented as well. Therefore all the analysis in this article is valid for any $2D$ polygonal domain, (including the one topologically equivalent to an annulus (i.e. a 2D torus)) under various boundary conditions.

To define the discrete FV derivatives on V_h, we apply the Taylor Series Expansion Scheme (TSES), which was introduced in, e.g., [45] and suitably modified in [42] for the case of a non-uniform mesh. The convergence of the FV method using the modified TSES scheme is proved in [42] when the domain is discretized by a mesh of rectangles.   Toward this end, we first introduce the quadrilaterals (diamond cells) $K_{i,j+1/2}$ and $K_{i+1/2,j}$ that will serve as the *domains of constancy* for the FV derivatives (see Fig. 4 and Remark 3.2):

(25)

$\quad K_{i+\frac{1}{2},j} =$ quadrilateral connecting four points $\boldsymbol{P}_{i,j}$, $\boldsymbol{P}_{i+1,j}$, and $\boldsymbol{P}_{i+\frac{1}{2},j\pm\frac{1}{2}}$,
$\qquad$ for $0 \le i \le M+1$, $0 \le j \le N$,

$\quad K_{i,j+\frac{1}{2}} =$ quadrilateral connecting four points $\boldsymbol{P}_{i,j}$, $\boldsymbol{P}_{i,j+1}$, and $\boldsymbol{P}_{i\pm\frac{1}{2},j+\frac{1}{2}}$,
$\qquad$ for $0 \le i \le M$, $0 \le j \le N+1$.

Thanks to (9), $K_{1/2,j}$, $K_{M+1/2,j}$, $K_{i,1/2}$ or $K_{i,N+1/2}$ near the boundary Γ becomes a triangle.
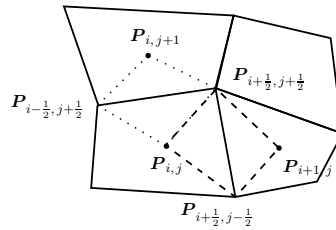
FIGURE 4. Dash-lined $K_{i+1/2,j}$ and dot-lined $K_{i,j+1/2}$ are the domains of constancy for the FV derivatives.

Then the domain Ω can be written in the form,

$$(26) \qquad\qquad \Omega = \Big(\bigcup_{i=0}^{M+1}\bigcup_{j=0}^{N} K_{i+\frac{1}{2},j}\Big)\bigcup\Big(\bigcup_{i=0}^{M}\bigcup_{j=0}^{N+1} K_{i,j+\frac{1}{2}}\Big).$$

Remark 3.2. The *diamond cells* $K_{i,j+1/2}$ and $K_{i+1/2,j}$ (or $\widetilde{K}_{i,j+1/2}$ and $\widetilde{K}_{i+1/2,j}$ defined below in (69)) were introduced and used in earlier works, e.g., [22, 23, 44] where some Finite Volume schemes, related to the current TSES method, are analyzed.

We consider the barycenters of two triangles with vertices $\boldsymbol{P}_{i-\frac{1}{2},j-\frac{1}{2}}$, $\boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}$, $\boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}}$ and $\boldsymbol{P}_{i+\frac{1}{2},j-\frac{1}{2}}$, $\boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}}$, $\boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}$, and name the one closer to the segment from $\boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}}$ to $\boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}$ as $\boldsymbol{Q}_{i,j}$. By definition of the barycenter, the distance between $\boldsymbol{P}_{i,j}$ and the segment from $\boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}}$ to $\boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}$ is bigger than that between $\boldsymbol{Q}_{i,j}$ and the segment from $\boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}}$ to $\boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}$ (see Fig. 5 below). Hence, using $(\mathcal{H}2)$ as well and by writing $|\boldsymbol{ABC}|$ the area of the triangle with vertices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, we find that

(27)
$$\left| \boldsymbol{P}_{i,j} \boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}} \right|$$
$$\geq \left| \boldsymbol{Q}_{i,j} \boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}} \right|$$
$$\geq \frac{1}{3} \min \left(\left| \boldsymbol{P}_{i-\frac{1}{2},j-\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}} \boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}} \right|, \left| \boldsymbol{P}_{i+\frac{1}{2},j-\frac{1}{2}} \boldsymbol{P}_{i-\frac{1}{2},j+\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}} \right| \right)$$
$$\geq \frac{1}{3} \underline{h}^2 \min \left(\sin \overline{\theta}, \sin \underline{\theta} \right)$$
$$\geq \frac{1}{3} \overline{h}^2 \delta.$$

Using this fact, we notice that

(28) $$\max \left(\frac{|K_{i,j}|}{|K_{k+\frac{1}{2},l}|} + \frac{|K_{i,j}|}{|K_{k,l+\frac{1}{2}}|} \right) \leq \frac{3}{2} \delta^{-1}, \quad 1 \leq i,k \leq M, \ 1 \leq j,l \leq N.$$
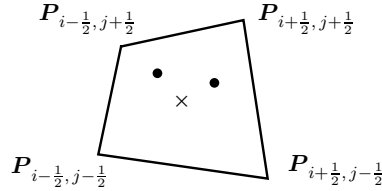

FIGURE 5. The X marked barycenter $\boldsymbol{P}_{i,j}$ of a control volume $K_{i,j}$ and the dotted barycenters of the two triangles, one with vertices $\boldsymbol{P}_{i-1/2,j-1/2}$, $\boldsymbol{P}_{i+1/2,j+1/2}$, and $\boldsymbol{P}_{i-1/2,j+1/2}$, and the other one with vertices $\boldsymbol{P}_{i+1/2,j-1/2}$, $\boldsymbol{P}_{i-1/2,j+1/2}$, and $\boldsymbol{P}_{i+1/2,j+1/2}$.

We introduce the (non-singular) geometric matrices $\mathbf{M}_{i,j+1/2}$ and $\mathbf{M}_{i+1/2,j}$ whose rows represent the diagonals of $K_{i,j+1/2}$ and $K_{i+1/2,j}$ respectively,

(29)
$$\mathbf{M}_{i+\frac{1}{2},j} = \begin{bmatrix} x_{i+1,j} - x_{i,j} & y_{i+1,j} - y_{i,j} \\ x_{i+\frac{1}{2},j+\frac{1}{2}} - x_{i+\frac{1}{2},j-\frac{1}{2}} & y_{i+\frac{1}{2},j+\frac{1}{2}} - y_{i+\frac{1}{2},j-\frac{1}{2}} \end{bmatrix},$$
$$0 \leq i \leq M, \ 0 \leq j \leq N+1,$$
$$\mathbf{M}_{i,j+\frac{1}{2}} = \begin{bmatrix} x_{i+\frac{1}{2},j+\frac{1}{2}} - x_{i-\frac{1}{2},j+\frac{1}{2}} & y_{i+\frac{1}{2},j+\frac{1}{2}} - y_{i-\frac{1}{2},j+\frac{1}{2}} \\ x_{i,j+1} - x_{i,j} & y_{i,j+1} - y_{i,j} \end{bmatrix},$$
$$0 \leq i \leq M+1, \ 0 \leq j \leq N.$$

Using (20), we define the intermediate value $u_{i+1/2,\,j+1/2}$ of $u_h \in V_h$ at $\boldsymbol{P}_{i+1/2,\,j+1/2}$ in the form,

(30)
$$
u_{i+\frac{1}{2},\,j+\frac{1}{2}} =
\begin{cases}
\displaystyle\sum_{l,m=0,1} \mu_{l,\,m}^{i+\frac{1}{2},\,j+\frac{1}{2}}\, u_{i+l,\,j+m}, & 1 \le i \le M-1,\ 1 \le j \le N-1, \\[2ex]
0, & i = 0, M, \text{ or } j = 0, N.
\end{cases}
$$

Thanks to (25)-(30), using the values of the function at the four vertices of $K_{i+1/2,\,j}$ (or $K_{i,\,j+1/2}$), we write two (discrete) directional derivatives of u_h in $K_{i+1/2,\,j}$ (or $K_{i,\,j+1/2}$) as a linear combination of the (discrete) gradient of u_h, where the coefficient vectors are identical to the rows of $\mathbf{M}_{i+1/2,\,j}$ (or $\mathbf{M}_{i,\,j+1/2}$). Then, by solving the linear system, we obtain the discrete FV derivatives on V_h:

For $u_h \in V_h$,

(31)
$$
\nabla_h u_h = \sum_{i=0}^{M} \sum_{j=0}^{N+1} \nabla_h u_h\big|_{K_{i+\frac{1}{2},\,j}}\, \chi_{K_{i+\frac{1}{2},\,j}} + \sum_{i=0}^{M+1} \sum_{j=0}^{N} \nabla_h u_h\big|_{K_{i,\,j+\frac{1}{2}}}\, \chi_{K_{i,\,j+\frac{1}{2}}},
$$

where

(32)
$$
\nabla_h u_h =
\begin{cases}
\mathbf{M}_{i+\frac{1}{2},\,j}^{-1}
\begin{bmatrix}
u_{i+1,\,j} - u_{i,\,j} \\[1.5ex]
u_{i+\frac{1}{2},\,j+\frac{1}{2}} - u_{i+\frac{1}{2},\,j-\frac{1}{2}}
\end{bmatrix}
& \text{on } K_{i+\frac{1}{2},\,j}, \\[4ex]
\mathbf{M}_{i,\,j+\frac{1}{2}}^{-1}
\begin{bmatrix}
u_{i+\frac{1}{2},\,j+\frac{1}{2}} - u_{i-\frac{1}{2},\,j+\frac{1}{2}} \\[1.5ex]
u_{i,\,j+1} - u_{i,\,j}
\end{bmatrix}
& \text{on } K_{i,\,j+\frac{1}{2}}.
\end{cases}
$$

We notice from (17) that the area formula (3) is valid for the $K_{i+1/2,j}$ and $K_{i,j+1/2}$. Hence we write

(33)
$$
\mathbf{M}_{i+\frac{1}{2},\,j}^{-1} = 2\big|K_{i+\frac{1}{2},\,j}\big|^{-1}
\begin{bmatrix}
y_{i+\frac{1}{2},\,j+\frac{1}{2}} - y_{i+\frac{1}{2},\,j-\frac{1}{2}} & -\big(y_{i+1,\,j} - y_{i,\,j}\big) \\[1.5ex]
-\big(x_{i+\frac{1}{2},\,j+\frac{1}{2}} - x_{i+\frac{1}{2},\,j-\frac{1}{2}}\big) & x_{i+1,\,j} - x_{i,\,j}
\end{bmatrix},
$$
$$
0 \le i \le M,\ 0 \le j \le N+1,
$$
$$
\mathbf{M}_{i,\,j+\frac{1}{2}}^{-1} = 2\big|K_{i,\,j+\frac{1}{2}}\big|^{-1}
\begin{bmatrix}
y_{i,\,j+1} - y_{i,\,j} & -\big(y_{i+\frac{1}{2},\,j+\frac{1}{2}} - y_{i-\frac{1}{2},\,j+\frac{1}{2}}\big) \\[1.5ex]
-\big(x_{i,\,j+1} - x_{i,\,j}\big) & x_{i+\frac{1}{2},\,j+\frac{1}{2}} - x_{i-\frac{1}{2},\,j+\frac{1}{2}}
\end{bmatrix},
$$
$$
0 \le i \le M+1,\ 0 \le j \le N.
$$

We equip the FV space V_h with the inner products $(\cdot,\,\cdot)_{V_h}$ and $((\cdot,\,\cdot))_{V_h}$, which mimic respectively those of $L^2(\Omega)$ and $H_0^1(\Omega)$:

For $u_h,\, v_h$ in V_h, we define

(34)
$$
\big(u_h,\, v_h\big)_{V_h} = \big(u_h,\, v_h\big)_{L^2(\Omega)} = \sum_{i=1}^{M} \sum_{j=1}^{N} u_{i,\,j}\, v_{i,\,j}\, |K_{i,\,j}|,
$$

(35)
$$
\begin{aligned}
\big((u_h,\, v_h)\big)_{V_h} &= \big(\nabla_h u_h,\, \nabla_h v_h\big)_{L^2(\Omega)} \\
&= \sum_{i=0}^{M} \sum_{j=0}^{N+1} \nabla_h u_h\big|_{K_{i+\frac{1}{2},\,j}} \cdot \nabla_h v_h\big|_{K_{i+\frac{1}{2},\,j}}\, |K_{i+\frac{1}{2},\,j}| \\
&\quad + \sum_{i=0}^{M+1} \sum_{j=0}^{N} \nabla_h u_h\big|_{K_{i,\,j+\frac{1}{2}}} \cdot \nabla_h v_h\big|_{K_{i,\,j+\frac{1}{2}}}\, |K_{i,\,j+\frac{1}{2}}|.
\end{aligned}
$$

Here we write the measure $|K_*|$ of K_* as

$$|K_{i,j}| = \frac{1}{2}\det\begin{bmatrix} x_{i+\frac{1}{2},j+\frac{1}{2}} - x_{i-\frac{1}{2},j-\frac{1}{2}} & x_{i-\frac{1}{2},j+\frac{1}{2}} - x_{i+\frac{1}{2},j-\frac{1}{2}} \\ y_{i+\frac{1}{2},j+\frac{1}{2}} - y_{i-\frac{1}{2},j-\frac{1}{2}} & y_{i-\frac{1}{2},j+\frac{1}{2}} - y_{i+\frac{1}{2},j-\frac{1}{2}} \end{bmatrix},$$

$$(36) \qquad |K_{i+\frac{1}{2},j}| = \frac{1}{2}\det\begin{bmatrix} x_{i+1,j} - x_{i,j} & x_{i+\frac{1}{2},j+\frac{1}{2}} - x_{i+\frac{1}{2},j-\frac{1}{2}} \\ y_{i+1,j} - y_{i,j} & y_{i+\frac{1}{2},j+\frac{1}{2}} - y_{i+\frac{1}{2},j-\frac{1}{2}} \end{bmatrix},$$

$$|K_{i,j+\frac{1}{2}}| = \frac{1}{2}\det\begin{bmatrix} x_{i+\frac{1}{2},j+\frac{1}{2}} - x_{i-\frac{1}{2},j+\frac{1}{2}} & x_{i,j+1} - x_{i,j} \\ y_{i+\frac{1}{2},j+\frac{1}{2}} - y_{i-\frac{1}{2},j+\frac{1}{2}} & y_{i,j+1} - y_{i,j} \end{bmatrix}.$$

We denote by $|\cdot|_{V_h}$ and $\|\cdot\|_{V_h}$ the norms associated with these scalar products $(\cdot,\cdot)_{V_h}$ and $((\cdot,\cdot))_{V_h}$ respectively.

4. External approximation of $H_0^1(\Omega)$ via the FV space V_h

To approximate the Sobolev space $H_0^1(\Omega)$ in the sense of [17] (see also [57]), we consider an external approximation as in Fig. 6 where the maps between the spaces are defined in (37):
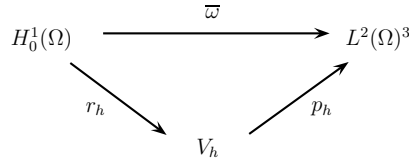

FIGURE 6. External approximation of $H_0^1(\Omega)$ via V_h.

$$(37)$$
$$\begin{cases} \overline{\omega}(u) = (u, \nabla u), \quad u \in H_0^1(\Omega), \\[2mm] r_h(u) = \displaystyle\sum_{i=0}^{M+1}\sum_{j=0}^{N+1} (r_h u)_{i,j}\, \chi_{K_{i,j}}, \quad u \in C_0^\infty(\Omega) \subset H_0^1(\Omega), \\[2mm] \qquad \text{where } (r_h u)_{i,j} = \begin{cases} \dfrac{1}{|K_{i,j}|}\displaystyle\int_{K_{i,j}} u\,d\boldsymbol{x},\ 1 \leq i \leq M,\ 1 \leq j \leq N, \\[2mm] 0,\ i = 0, M+1,\ \text{or } j = 0, N+1, \end{cases} \\[2mm] p_h(u_h) = (u_h, \nabla_h u_h), \quad u_h \in V_h. \end{cases}$$

We state and prove the discrete Poincaré inequality for this FV space:

Lemma 4.1. Under the assumptions ($\mathcal{H}$1) and ($\mathcal{H}$2), we have

$$(38) \qquad\qquad |u_h|_{V_h} \leq \kappa_P \|u_h\|_{V_h}, \quad u_h \in V_h,$$

for a constant $\kappa_P = 2\sqrt{6}\,\kappa_1^{-1}\delta^{-1/2}$, independent of the mesh size.

Proof. Considering u_h in V_h, since $u_{i,0} = 0$ for $1 \leq i \leq M$, we write

$$(39) \qquad\qquad u_{i,j} = \sum_{k=0}^{j-1}\left(u_{i,k+1} - u_{i,k}\right), \quad 1 \leq i \leq M,\ 1 \leq j \leq N.$$

Then, using the Schwarz inequality, we find

$$(40) \qquad u_{i,j}^2 \leq N\sum_{k=0}^{N}\left(u_{i,k+1} - u_{i,k}\right)^2, \quad 1 \leq i \leq M,\ 1 \leq j \leq N.$$

On the other hand, we infer from (29) and (32) that

$$(41) \qquad \overrightarrow{\boldsymbol{P}_{i,k}\boldsymbol{P}_{i,k+1}} \cdot \nabla_h u_h \big|_{K_{i,k+\frac{1}{2}}} = u_{i,k+1} - u_{i,k}.$$

It is clear that

$$(42) \qquad \left| \overrightarrow{\boldsymbol{P}_{i,k}\boldsymbol{P}_{i,k+1}} \right| \le \left| \overrightarrow{\boldsymbol{P}_{i,k}\boldsymbol{P}_{i+\frac{1}{2},k+\frac{1}{2}}} \right| + \left| \overrightarrow{\boldsymbol{P}_{i+\frac{1}{2},k+\frac{1}{2}}\boldsymbol{P}_{i,k+1}} \right| \le 4\overline{h}.$$

Combining (40)-(42), we find that

$$(43) \qquad u_{i,j}^2 \le 16 N \overline{h}^2 \sum_{k=0}^{N} \left| \nabla_h u_h \big|_{K_{i,k+\frac{1}{2}}} \right|^2, \quad 1 \le i \le M, \ 1 \le j \le N.$$

Now, using (43), we write

$$(44) \quad |u_h|_{V_h}^2 = \sum_{i=1}^{M} \sum_{j=1}^{N} u_{i,j}^2 |K_{i,j}| \le 16 N \overline{h}^2 \sum_{i=1}^{M} \sum_{j=1}^{N} \left\{ \sum_{k=0}^{N} \left| \nabla_h u_h \big|_{K_{i,k+\frac{1}{2}}} \right|^2 |K_{i,j}| \right\}.$$

Using (28), we deduce from (44) that

$$(45) \qquad |u_h|_{V_h}^2 \le 24 N^2 \overline{h}^2 \delta^{-1} \sum_{i=1}^{M} \sum_{k=0}^{N} \left| \nabla_h u_h \big|_{K_{i,k+\frac{1}{2}}} \right|^2 |K_{i,k+\frac{1}{2}}| \le 24 \kappa_1^2 \delta \|u_h\|_{V_h}^2.$$

The proof of (38) is now complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Thanks to the discrete Poincaré inequality (38), the stability (uniform boundedness) of the operators p_h, defined in (37), follows:
$$(46)$$
$$\|p_h\|_{L(V_h, L^2(\Omega)^3)}^2 = \sup_{u_h \in V_h} \frac{|u_h|_{L^2(\Omega)}^2 + |\nabla u_h|_{L^2(\Omega)}^2}{\|u_h\|_{V_h}^2} = \sup_{u_h \in V_h} \frac{|u_h|_{V_h}^2 + \|u_h\|_{V_h}^2}{\|u_h\|_{V_h}^2} \le (1 + \kappa_P^2).$$

Convergence and consistency of FV. To prove the convergence and consistency of the FV method, we need to prove the following two properties (see [17] or Sections 3 and 4 of Chapter 1 in [57]):

($\mathcal{C}$1) $(p_h \circ r_h)(u) \to \overline{\omega}(u)$ in $L^2(\Omega)^3$ as $\overline{h} \to 0$, $\forall u \in C_0^\infty(\Omega)$,

($\mathcal{C}$2) If $u_h \in V_h$ and $p_h(u_h) \rightharpoonup \phi$ weakly in $L^2(\Omega)^3$ as $\overline{h} \to 0$, then $\phi \in \overline{\omega}\big(H_0^1(\Omega)\big)$.

We recall some elementary lemmas which can be easily verified by using Taylor expansions (see also [42]):

Lemma 4.2. Let K be a convex polygon in $\mathbb{R}^2$ with barycenter $\boldsymbol{\xi}_K$. Then,

$$(47) \qquad \frac{1}{|K|} \int_K \phi \, d\boldsymbol{x} = \phi(\boldsymbol{\xi}_K) + \mathcal{O}\big(|K|\big), \quad \phi \in C^2(\overline{K}),$$

where $\mathcal{O}\big(|K|\big) \le \|\phi\|_{C^2(K)} |K|$.

Lemma 4.3. Let K be a convex polygon in $\mathbb{R}^2$ with vertices $\boldsymbol{\xi}_i$, $1 \le i \le p$. Then, for any point $\boldsymbol{\xi}$ inside K of the form,

$$\boldsymbol{\xi} = \sum_{i=1}^{p} \gamma_i \, \boldsymbol{\xi}_i, \quad \gamma_i \ge 0, \quad \sum_{i=1}^{p} \gamma_i = 1,$$

we have

$$(48) \qquad \sum_{i=1}^{p} \gamma_i \, \phi(\boldsymbol{\xi}_i) = \phi(\boldsymbol{\xi}) + \mathcal{O}\big(|K|\big), \quad \phi \in C^2(\overline{K}),$$

where $\mathcal{O}\big(|K|\big) \leq \|\phi\|_{C^2(K)}|K|$.

Lemma 4.4. Let $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, and $\boldsymbol{\xi}$ be three (not necessarily aligned) points in $\mathbb{R}^2$. Then we have

$$(49) \qquad \phi(\boldsymbol{\xi}_2) - \phi(\boldsymbol{\xi}_1) = \nabla\phi(\boldsymbol{\xi}) \cdot (\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1) + \mathcal{O}\Big(\sum_{i=1}^{2} |\boldsymbol{\xi}_i - \boldsymbol{\xi}|^2\Big), \quad \phi \in C^2(\mathbb{R}^2),$$

where $\mathcal{O}\big(\sum_{i=1}^{2} |\boldsymbol{\xi}_i - \boldsymbol{\xi}|^2\big) \leq \|\phi\|_{C^2(\mathbb{R}^2)} \sum_{i=1}^{2} |\boldsymbol{\xi}_i - \boldsymbol{\xi}|^2$.

We recall the big order $\mathcal{O}\big(\overline{h}^{\,\gamma}\big)$ for $\gamma \geq 0$ with respect to the mesh size $\overline{h}$ such that

$$(50) \qquad \big|\mathcal{O}\big(\overline{h}^{\,\gamma}\big)\big| \leq \kappa\,\overline{h}^{\,\gamma},$$

for a generic constant $\kappa > 0$ which is independent of i, j, or $\overline{h}$. The small order $o\big(\overline{h}^{\,\gamma}\big)$, $\gamma \geq 0$, with respect to the mesh size $\overline{h}$ is defined as well so that

$$(51) \qquad \lim_{\overline{h} \to 0} \left|\frac{o\big(\overline{h}^{\,\gamma}\big)}{\overline{h}^{\,\gamma}}\right| = 0.$$

4.1. Proof of $(\mathcal{C}1)$ **for FV.** To verify the property $(\mathcal{C}1)$ for this FV space, we first choose a smooth function $u \in C_0^\infty(\Omega)$ and want to show that

$$(52) \qquad r_h(u) \to u \quad \text{strongly in } L^2(\Omega) \text{ as } \overline{h} \to 0.$$

For a point (x, y) in Ω (up to a set of measure zero), we choose i and j so that $(x, y) \in K_{i,j}$. Using the definition of r_h in (37), Lemma 4.2, and the Taylor expansion, we infer that

$$(53) \qquad \begin{aligned} \big|r_h(u)(x, y) - u(x, y)\big| &= \left|\frac{1}{|K_{i,j}|} \int_{K_{i,j}} u \, d\boldsymbol{x} - u(x, y)\right| \\ &\leq \big|u(\boldsymbol{P}_{i,j}) - u(x, y)\big| + \mathcal{O}\big(\overline{h}^2\big) \\ &\leq \|\nabla u\|_{L^\infty(\Omega)}\,\overline{h}. \end{aligned}$$

Then we deduce that

$$r_h(u) \to u \quad \text{strongly in } L^\infty(\Omega) \text{ as } \overline{h} \to 0,$$

and hence (52) follows.

As a next step, we need to verify that

$$(54) \qquad \nabla_h r_h(u) \to \nabla u \quad \text{strongly in } L^2(\Omega) \text{ as } \overline{h} \to 0.$$

From (26), we notice that any point in Ω (up to a set of measure zero) is located in exactly one of the $K_{i,j+1/2}$ or $K_{i+1/2,j}$. Without loss of generality, we assume that an arbitrary chosen (but fixed) point (x, y) is inside of $K_{i,j+1/2}$ for some i and j; the other case when $(x, y) \in K_{i+1/2,j}$ can be treated in the same manner. Then, using (32) and (37), we write

$$(55) \qquad \nabla_h r_h(u)(x, y) = \mathbf{M}_{i,j+\frac{1}{2}}^{-1} \begin{bmatrix} (r_h u)_{i+\frac{1}{2},j+\frac{1}{2}} - (r_h u)_{i-\frac{1}{2},j+\frac{1}{2}} \\ (r_h u)_{i,j+1} - (r_h u)_{i,j} \end{bmatrix},$$

where $(r_h u)_{i+1/2,j+1/2}$ is defined by (30) with $u_{i,j}$ replaced by $(r_h u)_{i,j}$.

Using the definition of $(r_h u)_{i,j}$ in (37), and using Lemmas 4.2 and 4.3, we notice that

(56)
$$\begin{cases} (r_h u)_{i,j} = u(\boldsymbol{P}_{i,j}) + \mathcal{O}(\overline{h}^2), & 1 \le i \le M, \ 1 \le j \le N, \\ (r_h u)_{i+\frac{1}{2},j+\frac{1}{2}} = u(\boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}) + \mathcal{O}(\overline{h}^2), & 1 \le i \le M-1, \ 1 \le j \le N-1. \end{cases}$$

Thanks to (29), (33), (56), and Lemma 4.4, we find that

(57)
$$\begin{bmatrix} (r_h u)_{i+\frac{1}{2},j+\frac{1}{2}} - (r_h u)_{i-\frac{1}{2},j+\frac{1}{2}} \\ (r_h u)_{i,j+1} - (r_h u)_{i,j} \end{bmatrix} = \mathbf{M}_{i,j+\frac{1}{2}} \nabla u(x,y) + \mathcal{O}(\overline{h}^2).$$

Combining (55) and (57), we obtain

(58)
$$|\nabla_h r_h(u)(x,y) - \nabla u(x,y)| \le \mathcal{O}(\overline{h}).$$

Then we deduce that $\nabla_h r_h(u)$ converges to ∇u in $L^\infty(\Omega)$ as $\overline{h} \to 0$, and hence (54) follows as well.

Thanks to (52) and (54), the property $(\mathcal{C}1)$ of the FV space is obtained.

As we already recalled, the FV method is weakly consistent, and hence proving the $(\mathcal{C}2)$ property for the FV is not as direct as for the other classical methods such as Finite Differences or Finite Elements; see, e.g., [32, 39, 42]. In this article, to verify the property $(\mathcal{C}2)$ for the FV, we follow the approach introduced in [39, 42]. More precisely, we will first construct in Section 3 the Finite Differences space which is associated with the mesh corresponding to the FV space, and prove the stability and convergence (the properties $(\mathcal{C}1)$ and $(\mathcal{C}2)$) of the Finite Differences. Then, by comparing the FV and FD spaces, we will finally deduce that the property $(\mathcal{C}2)$ holds true for the FV.

5. Corresponding Finite Difference setting

In this section, we construct the Finite Differences (FD) space which is associated with the FV space in Section 3.

As a first step, we first choose the FD nodal points along the boundary Γ to be the same as those of the FV,

(59) $\widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j+\frac{1}{2}} = (\widetilde{x}_{i+\frac{1}{2},j+\frac{1}{2}}, \widetilde{y}_{i+\frac{1}{2},j+\frac{1}{2}}) = \boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}, \quad i = 0, M, \text{ or } j = 0, M.$

Then the boundary cells of the FD mesh are naturally defined as those of the FV mesh,

(60) $\widetilde{K}_{i,j} = K_{i,j}, \quad i = 0, M+1, \text{ or } j = 0, M+1.$

We keep the FV barycenters as the FD points,

(61) $\widetilde{\boldsymbol{P}}_{i,j} = (\widetilde{x}_{i,j}, \widetilde{y}_{i,j}) = \boldsymbol{P}_{i,j}, \quad 0 \le i \le M+1, \ 0 \le j \le N+1.$

We define the inner FD nodal points as the average of the nearby $\widetilde{\boldsymbol{P}}_{i,j}$,

(62)
$$\widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{1}{4} \sum_{l,m=0,1} \widetilde{\boldsymbol{P}}_{i+l,j+m}, \quad 1 \le i \le M-1, \ 1 \le j \le N-1, \text{ (see Fig. 7 below)}.$$

In general, the average defined on the right-hand side of (62) is different from the barycenter (center of mass) of $\widetilde{\boldsymbol{P}}_{i+l,j+m}$, $l, m = 0, 1$. They coincide only when the quadrilateral with vertices $\widetilde{\boldsymbol{P}}_{i+l,j+m}$, $l, m = 0, 1$, is a parallelogram.

Using (60) and (62), we obtain the FD discretization of the domain Ω in the form,

$$(63) \qquad \overline{\Omega} = \bigcup_{i=0}^{M+1} \bigcup_{j=0}^{N+1} \widetilde{K}_{i,j},$$

where

$$(64) \qquad \widetilde{K}_{i,j} = \text{ convex quadrilateral connecting the four points } \widetilde{\boldsymbol{P}}_{i\pm\frac{1}{2},\,j\pm\frac{1}{2}}.$$
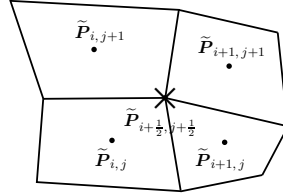

FIGURE 7. The X marked FD nodal point $\widetilde{\boldsymbol{P}}_{i+1/2,\,j+1/2}$ is the average of the nearby four points $\widetilde{\boldsymbol{P}}_{i+l,\,j+m}$, $j, m = 0, 1$; see Fig. 3 to compare with the corresponding FV mesh, in which $\boldsymbol{P}_{i+1/2,\,j+1/2}$ is not the average (nor the barycenter) of the points $\widetilde{\boldsymbol{P}}_{i+l,\,j+m}$ around it; see (62).

For a cell $\widetilde{K}_{i,j}$, we write its boundary $\widetilde{\Gamma}_{i,j}$ as

$$(65) \qquad \widetilde{\Gamma}_{i,j} = \bigcup_{k=E,W,N,S} \widetilde{\Gamma}_{i,j}^k.$$

The four internal angles of $\widetilde{K}_{i,j}$ are denoted $\widetilde{\theta}_{i,j}^m$, $m = WS, ES, EN, WN$. This setting of the FD mesh appears in Fig. 2 with $\boldsymbol{P}$, K, and Γ respectively replaced by $\widetilde{\boldsymbol{P}}$, $\widetilde{K}$, and $\widetilde{\Gamma}$.

Since $\widetilde{\boldsymbol{P}}_{i,j} \in \widetilde{K}_{i,j}$, we write $\widetilde{\boldsymbol{P}}_{i,j}$ as an interpolation of the nearby nodal points (the vertices of $\widetilde{K}_{i,j}$):

For $1 \leq i \leq M$, $1 \leq j \leq N$,

$$(66)$$
$$\widetilde{\boldsymbol{P}}_{i,j} = \sum_{l,m=\pm 1} \widetilde{\lambda}_{l,m}^{i,j} \widetilde{\boldsymbol{P}}_{i+\frac{l}{2},\,j+\frac{m}{2}} \quad \text{for some } \widetilde{\lambda}_{l,m}^{i,j} \geq 0 \text{ such that } \sum_{j,m=\pm 1} \widetilde{\lambda}_{l,m}^{i,j} = 1.$$

Construction of the FD space. We define the FD space of step functions that satisfy the homogeneous Dirichlet boundary condition,

$$(67) \qquad \widetilde{V}_h := \left\{ \begin{array}{l} \text{step functions } \widetilde{u}_h \text{ on } \overline{\Omega} = \cup_{i=0}^{M+1} \cup_{j=0}^{N+1} \widetilde{K}_{i,j} \text{ such that} \\[2mm] \widetilde{u}_h\big|_{\widetilde{K}_{i,j}} = \left\{ \begin{array}{l} \widetilde{u}_{i,j}, \text{ for } 1 \leq i \leq M,\ 1 \leq j \leq N, \\ 0, \text{ for } i = 0, M+1, \text{ or } j = 0, N+1. \end{array} \right. \end{array} \right\}.$$

Then a FD step function $\widetilde{u}_h \in \widetilde{V}_h$ is written in the form,

$$(68) \qquad \widetilde{u}_h = \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j}\, \chi_{\widetilde{K}_{i,j}}.$$

To define the discrete FD derivatives on $\widetilde{V}_h$, we apply the classical Taylor Series Expansion Scheme (TSES):

We introduce the quadrilaterals $\widetilde{K}_{i,\,j+1/2}$ and $\widetilde{K}_{i+1/2,\,j}$ which will serve as the *domains of constancy* for the FD derivatives (see Fig. 8),

(69)
$$\widetilde{K}_{i+\frac{1}{2},\,j} = \text{quadrilateral connecting the four points } \widetilde{\boldsymbol{P}}_{i,\,j},\ \widetilde{\boldsymbol{P}}_{i+1,\,j},\text{ and } \widetilde{\boldsymbol{P}}_{i+\frac{1}{2},\,j\pm\frac{1}{2}},$$
$$\text{for } 0 \leq i \leq M+1,\ 0 \leq j \leq N,$$

$$\widetilde{K}_{i,\,j+\frac{1}{2}} = \text{quadrilateral connecting the four points } \widetilde{\boldsymbol{P}}_{i,\,j},\ \widetilde{\boldsymbol{P}}_{i,\,j+1},\text{ and } \widetilde{\boldsymbol{P}}_{i\pm\frac{1}{2},\,j+\frac{1}{2}},$$
$$\text{for } 0 \leq i \leq M,\ 0 \leq j \leq N+1.$$

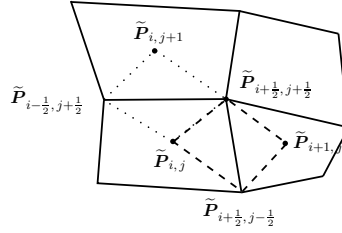Then the domain Ω can be written in the form (70):

FIGURE 8. Dash-lined $\widetilde{K}_{i+1/2,\,j}$ and dot-lined $\widetilde{K}_{i,\,j+1/2}$ as the domains of constancy for the FD derivatives.

(70)
$$\Omega = \Big(\bigcup_{i=0}^{M+1} \bigcup_{j=0}^{N} \widetilde{K}_{i+\frac{1}{2},\,j} \Big) \bigcup \Big(\bigcup_{i=0}^{M} \bigcup_{j=0}^{N+1} \widetilde{K}_{i,\,j+\frac{1}{2}} \Big).$$

We define the (non-singular) geometric matrices $\widetilde{\mathbf{M}}_{i,\,j+1/2}$ and $\widetilde{\mathbf{M}}_{i+1/2,\,j}$ whose rows represent the diagonals of $\widetilde{K}_{i,\,j+1/2}$ and $\widetilde{K}_{i+1/2,\,j}$ respectively,

(71)
$$\widetilde{\mathbf{M}}_{i+\frac{1}{2},\,j} = \begin{bmatrix} \widetilde{x}_{i+1,\,j} - \widetilde{x}_{i,\,j} & \widetilde{y}_{i+1,\,j} - \widetilde{y}_{i,\,j} \\ \widetilde{x}_{i+\frac{1}{2},\,j+\frac{1}{2}} - \widetilde{x}_{i+\frac{1}{2},\,j-\frac{1}{2}} & \widetilde{y}_{i+\frac{1}{2},\,j+\frac{1}{2}} - \widetilde{y}_{i+\frac{1}{2},\,j-\frac{1}{2}} \end{bmatrix},$$
$$0 \leq i \leq M,\ 0 \leq j \leq N+1,$$
$$\widetilde{\mathbf{M}}_{i,\,j+\frac{1}{2}} = \begin{bmatrix} \widetilde{x}_{i+\frac{1}{2},\,j+\frac{1}{2}} - \widetilde{x}_{i-\frac{1}{2},\,j+\frac{1}{2}} & \widetilde{y}_{i+\frac{1}{2},\,j+\frac{1}{2}} - \widetilde{y}_{i-\frac{1}{2},\,j+\frac{1}{2}} \\ \widetilde{x}_{i,\,j+1} - \widetilde{x}_{i,\,j} & \widetilde{y}_{i,\,j+1} - \widetilde{y}_{i,\,j} \end{bmatrix},$$
$$0 \leq i \leq M+1,\ 0 \leq j \leq N.$$

Using (62), we define the intermediate value $\widetilde{u}_{i+1/2,\,j+1/2}$ of $\widetilde{u}_h \in \widetilde{V}_h$, for $0 \leq i \leq M$, $0 \leq j \leq N$, in the form,

(72)
$$\widetilde{u}_{i+\frac{1}{2},\,j+\frac{1}{2}} = \begin{cases} \dfrac{1}{4} \displaystyle\sum_{l,m=0,1} \widetilde{u}_{i+l,\,j+m}, & 1 \leq i \leq M-1,\ 1 \leq j \leq N-1, \\[2mm] 0, & i = 0, M, \text{ or } j = 0, N. \end{cases}$$

Thanks to (69)-(72), we define the discrete FD derivatives on $\widetilde{V}_h$:
For $\widetilde{u}_h \in \widetilde{V}_h$,

(73)
$$\widetilde{\nabla}_h \widetilde{u}_h = \sum_{i=0}^{M} \sum_{j=0}^{N+1} \widetilde{\nabla}_h \widetilde{u}_h \big|_{\widetilde{K}_{i+\frac{1}{2},\,j}} \chi_{\widetilde{K}_{i+\frac{1}{2},\,j}} + \sum_{i=0}^{M+1} \sum_{j=0}^{N} \widetilde{\nabla}_h \widetilde{u}_h \big|_{\widetilde{K}_{i,\,j+\frac{1}{2}}} \chi_{\widetilde{K}_{i,\,j+\frac{1}{2}}},$$

where

$$(74) \qquad \widetilde{\nabla}_h \widetilde{u}_h = \begin{cases} \widetilde{\mathbf{M}}_{i+\frac{1}{2},j}^{-1} \begin{bmatrix} \widetilde{u}_{i+1,j} - \widetilde{u}_{i,j} \\ \widetilde{u}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{u}_{i+\frac{1}{2},j-\frac{1}{2}} \end{bmatrix} & \text{on } \widetilde{K}_{i+\frac{1}{2},j}, \\[3mm] \widetilde{\mathbf{M}}_{i,j+\frac{1}{2}}^{-1} \begin{bmatrix} \widetilde{u}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{u}_{i-\frac{1}{2},j+\frac{1}{2}} \\ \widetilde{u}_{i,j+1} - \widetilde{u}_{i,j} \end{bmatrix} & \text{on } \widetilde{K}_{i,j+\frac{1}{2}}. \end{cases}$$

Using (3), we write

$$\widetilde{\mathbf{M}}_{i+\frac{1}{2},j}^{-1} = 2 \big| \widetilde{K}_{i+\frac{1}{2},j} \big|^{-1} \begin{bmatrix} \widetilde{y}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{y}_{i+\frac{1}{2},j-\frac{1}{2}} & -\big(\widetilde{y}_{i+1,j} - \widetilde{y}_{i,j}\big) \\ -\big(\widetilde{x}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{x}_{i+\frac{1}{2},j-\frac{1}{2}}\big) & \widetilde{x}_{i+1,j} - \widetilde{x}_{i,j} \end{bmatrix},$$

$$0 \le i \le M, \ 0 \le j \le N+1,$$

$$(75)$$

$$\widetilde{\mathbf{M}}_{i,j+\frac{1}{2}}^{-1} = 2 \big| \widetilde{K}_{i,j+\frac{1}{2}} \big|^{-1} \begin{bmatrix} \widetilde{y}_{i,j+1} - \widetilde{y}_{i,j} & -\big(\widetilde{y}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{y}_{i-\frac{1}{2},j+\frac{1}{2}}\big) \\ -\big(\widetilde{x}_{i,j+1} - \widetilde{x}_{i,j}\big) & \widetilde{x}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{x}_{i-\frac{1}{2},j+\frac{1}{2}} \end{bmatrix},$$

$$0 \le i \le M+1, \ 0 \le j \le N.$$

The FD space $\widetilde{V}_h$ is equipped with the inner products $(\cdot,\,\cdot)_{\widetilde{V}_h}$ and $((\cdot,\,\cdot))_{\widetilde{V}_h}$ which mimic respectively those of $L^2(\Omega)$ and $H_0^1(\Omega)$:

For $\widetilde{u}_h$, $\widetilde{v}_h$ in $\widetilde{V}_h$, we define

$$(76) \qquad \big(\widetilde{u}_h,\,\widetilde{v}_h\big)_{\widetilde{V}_h} = \big(\widetilde{u}_h,\,\widetilde{v}_h\big)_{L^2(\Omega)} = \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j}\,\widetilde{v}_{i,j} |\widetilde{K}_{i,j}|,$$

$$\big(\!\big(\widetilde{u}_h,\,\widetilde{v}_h\big)\!\big)_{\widetilde{V}_h} = \big(\widetilde{\nabla}_h \widetilde{u}_h,\,\widetilde{\nabla}_h \widetilde{v}_h\big)_{L^2(\Omega)}$$

$$(77) \qquad = \sum_{i=0}^{M} \sum_{j=0}^{N+1} \widetilde{\nabla}_h \widetilde{u}_h \big|_{\widetilde{K}_{i+\frac{1}{2},j}} \cdot \widetilde{\nabla}_h \widetilde{v}_h \big|_{\widetilde{K}_{i+\frac{1}{2},j}} |\widetilde{K}_{i+\frac{1}{2},j}|$$

$$+ \sum_{i=0}^{M+1} \sum_{j=0}^{N} \widetilde{\nabla}_h \widetilde{u}_h \big|_{\widetilde{K}_{i,j+\frac{1}{2}}} \cdot \widetilde{\nabla}_h \widetilde{v}_h \big|_{\widetilde{K}_{i,j+\frac{1}{2}}} |\widetilde{K}_{i,j+\frac{1}{2}}|,$$

where the measure $|\widetilde{K}_*|$ of $\widetilde{K}_*$ is given by (36) in which K, x, and y are replaced by $\widetilde{K}$, $\widetilde{x}$, and $\widetilde{y}$ respectively. The corresponding norms $|\cdot|_{\widetilde{V}_h}$ and $\|\cdot\|_{\widetilde{V}_h}$ are naturally deduced from (76) and (77) as well.

6. External approximation of $H_0^1(\Omega)$ via the FD space $\widetilde{V}_h$

To approximate the Sobolev space $H_0^1(\Omega)$ via a FD space, we consider an external approximation as in Fig. 6 where r_h, p_h, and V_h are respectively replaced by $\widetilde{r}_h$, $\widetilde{p}_h$, and $\widetilde{V}_h$, and $\widetilde{\omega} = \overline{\omega}$. The maps $\widetilde{r}_h$ and $\widetilde{p}_h$ are defined by

$$(78) \qquad \begin{cases} \widetilde{r}_h(u) = \displaystyle\sum_{i=0}^{M+1} \sum_{j=0}^{N+1} (\widetilde{r}_h u)_{i,j}\, \chi_{\widetilde{K}_{i,j}}, \quad u \in C_0^\infty(\Omega) \subset H_0^1(\Omega), \\[3mm] \qquad \text{where } (\widetilde{r}_h u)_{i,j} = \begin{cases} u\big(\widetilde{\boldsymbol{P}}_{i,j}\big), \ 1 \le i \le M, \ 1 \le j \le N, \\ 0, \ i = 0, M+1, \ \text{or } j = 0, N+1, \end{cases} \\[3mm] \widetilde{p}_h(\widetilde{u}_h) = (\widetilde{u}_h,\, \widetilde{\nabla}_h \widetilde{u}_h), \quad \widetilde{u}_h \in \widetilde{V}_h. \end{cases}$$

For the analysis below of the FD space, we assume that the FD mesh as well is not too distorted in the sense that

($\mathcal{H}3$) The constants $\underline{h}$, $\overline{h}$, $\underline{\theta}$, and $\overline{\theta}$ in (12) and (13) are chosen so that the analogues of (12) and (13) hold with Γ and θ replaced by $\widetilde{\Gamma}$ and $\widetilde{\theta}$.

Remark 6.1. In fact the condition ($\mathcal{H}3$) can be verified as a consequence of the condition ($\mathcal{H}2$) and the condition ($\mathcal{H}5$) that we impose below. We choose to state ($\mathcal{H}3$) here in this section to complete the analysis of the FD approximation in a self-contained manner.

Then the discrete Poincaré inequality for the FD space is proven exactly as for the FV space:

Lemma 6.2. Under the assumptions ($\mathcal{H}1$)-($\mathcal{H}3$), we have

$$(79) \qquad\qquad |\widetilde{u}_h|_{\widetilde{V}_h} \leq \kappa_P \|\widetilde{u}_h\|_{\widetilde{V}_h}, \quad \widetilde{u}_h \in \widetilde{V}_h,$$

for a constant $\kappa_P = 2\sqrt{6}\,\kappa_1^{-1}\delta^{-1/2}$, independent of the mesh size.

Following the same computations as in (46), one can prove the stability of $\widetilde{p}_h$:

$$(80) \qquad\qquad \|\widetilde{p}_h\|_{L(\widetilde{V}_h,\, L^2(\Omega)^3)} \leq \sqrt{1 + \kappa_p^2}.$$

Convergence and consistency of FD. To prove the convergence and consistency of the FD method, we need to prove the following two properties:

($\mathcal{C}1$) $(\widetilde{p}_h \circ \widetilde{r}_h)(u) \to \overline{\varpi}(u)$ in $L^2(\Omega)^3$ as $\overline{h} \to 0$, $u \in C_0^\infty(\Omega)$,

($\mathcal{C}2$) If $\widetilde{u}_h \in \widetilde{V}_h$ and $\widetilde{p}_h(\widetilde{u}_h) \rightharpoonup \phi$ weakly in $L^2(\Omega)^3$ as $\overline{h} \to 0$, then $\phi \in \overline{\varpi}\big(H_0^1(\Omega)\big)$.

The proof of ($\mathcal{C}1$) for the FD is almost the same as (and even easier than) that of the FV in Section 4.1. Hence we omit this proof and prove ($\mathcal{C}2$) for the FD only in Section 6.1.

6.1. Proof of ($\mathcal{C}2$) for FD . We assume that the corresponding FD mesh is sufficiently regular in the sense below:

($\mathcal{H}4$) The FD points $\widetilde{\boldsymbol{P}}_{i,j}$, (which are identical to the cell centers $\boldsymbol{P}_{i,j}$ of FV mesh), satisfy that

$$\begin{cases} \dfrac{\widetilde{\boldsymbol{P}}_{i+1,j} + \widetilde{\boldsymbol{P}}_{i-1,j}}{2} = \widetilde{\boldsymbol{P}}_{i,j} + \mathcal{O}\big(\overline{h}^2\big), & \dfrac{\widetilde{\boldsymbol{P}}_{i,j+1} + \widetilde{\boldsymbol{P}}_{i,j-1}}{2} = \widetilde{\boldsymbol{P}}_{i,j} + \mathcal{O}\big(\overline{h}^2\big), \\[3mm] \dfrac{\widetilde{\boldsymbol{P}}_{i+1,j+1} + \widetilde{\boldsymbol{P}}_{i-1,j-1}}{2} = \widetilde{\boldsymbol{P}}_{i,j} + \mathcal{O}\big(\overline{h}^2\big), & \dfrac{\widetilde{\boldsymbol{P}}_{i+1,j-1} + \widetilde{\boldsymbol{P}}_{i-1,j+1}}{2} = \widetilde{\boldsymbol{P}}_{i,j} + \mathcal{O}\big(\overline{h}^2\big), \end{cases}$$

for $2 \leq i \leq M - 1$ and $2 \leq j \leq N - 1$.

Remark 6.3. The condition ($\mathcal{H}4$), which is inspired by the earlier work [42], is satisfied in particular by a typical problematic mesh with the alternating sizes of h and $2h$ in the analysis of FV. The simple examples A and B in Fig. 9 satisfy the assumption ($\mathcal{H}4$) as well, because the inner cells of the corresponding FD mesh for A (or B) become identical to a parallelogram (or a rectangle).
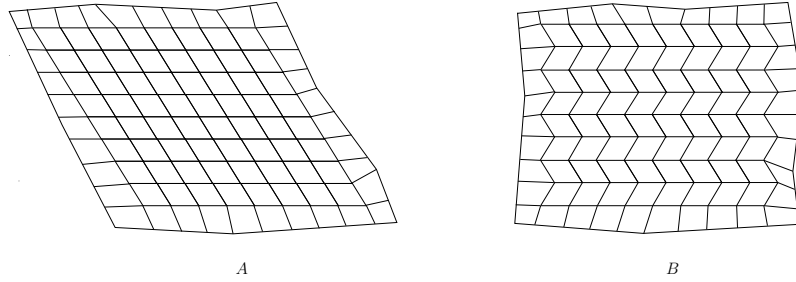
FIGURE 9. Some particular FV meshes.

Now, to prove the property $(\mathcal{C}2)$ for the FD, we choose a family $\widetilde{u}_h \in \widetilde{V}_h$ such that

$$(81) \qquad \widetilde{u}_h \rightharpoonup \phi_0, \quad \widetilde{\nabla}_h \widetilde{u}_h \rightharpoonup (\phi_1, \phi_2) \quad \text{weakly in } L^2(\Omega) \text{ as } \overline{h} \to 0,$$

and we wish to show that $(\phi_0, \phi_1, \phi_2) \in \overline{\omega}(H_0^1(\Omega))$, that is:

$$(82) \qquad \int_{\mathbb{R}^2} (\overline{\phi}_1, \overline{\phi}_2)\, \theta\, d\boldsymbol{x} = -\int_{\mathbb{R}^2} \overline{\phi}_0\, \nabla\theta\, d\boldsymbol{x}, \quad \forall \theta \in C_0^\infty(\mathbb{R}^2).$$

Here $\overline{\phi}_*$ is the function equal to ϕ_* in Ω and to 0 outside of Ω.

Setting,

$$(83) \qquad I_h = \int_\Omega \widetilde{\nabla}_h \widetilde{u}_h\, \theta\, d\boldsymbol{x},$$

we infer from (81) that

$$(84) \qquad I_h \to \int_\Omega (\phi_1, \phi_2)\, \theta\, d\boldsymbol{x} = \int_{\mathbb{R}^2} (\overline{\phi}_1, \overline{\phi}_2)\, \theta\, d\boldsymbol{x} \quad \text{as } \overline{h} \to 0.$$

Therefore, to prove (82), (and hence $(\mathcal{C}2)$ for the FD), it is enough to verify that

$$(85) \qquad I_h \to -\int_\Omega \phi_0\, \nabla\theta\, d\boldsymbol{x} \quad \text{as } \overline{h} \to 0,$$

which is the same as the right-hand side of (82). Hereafter our task is to verify (85).

Using the definition of the FD derivatives (73), we write I_h in (83) in the form,

$$(86) \qquad I_h = I_h^H + I_h^V,$$

where

$$(87) \qquad \begin{cases} I_h^H = \displaystyle\sum_{i=0}^{M} \sum_{j=0}^{N+1} \widetilde{\mathbf{M}}_{i+\frac{1}{2}, j}^{-1} \begin{bmatrix} \widetilde{u}_{i+1, j} - \widetilde{u}_{i, j} \\ \widetilde{u}_{i+\frac{1}{2}, j+\frac{1}{2}} - \widetilde{u}_{i+\frac{1}{2}, j-\frac{1}{2}} \end{bmatrix} \displaystyle\int_{\widetilde{K}_{i+\frac{1}{2}, j}} \theta\, d\boldsymbol{x}, \\[3ex] I_h^V = \displaystyle\sum_{i=0}^{M+1} \sum_{j=0}^{N} \widetilde{\mathbf{M}}_{i, j+\frac{1}{2}}^{-1} \begin{bmatrix} \widetilde{u}_{i+\frac{1}{2}, j+\frac{1}{2}} - \widetilde{u}_{i-\frac{1}{2}, j+\frac{1}{2}} \\ \widetilde{u}_{i, j+1} - \widetilde{u}_{i, j} \end{bmatrix} \displaystyle\int_{\widetilde{K}_{i, j+\frac{1}{2}}} \theta\, d\boldsymbol{x}. \end{cases}$$

Setting

$$(88) \qquad \mathcal{M}_{i+\frac{1}{2}, j} = \left(m_{i+\frac{1}{2}, j}^{kl} \right)_{1 \le k, l \le 2} := |\widetilde{K}_{i+\frac{1}{2}, j}|\, \widetilde{\mathbf{M}}_{i+\frac{1}{2}, j}^{-1},$$

and

$$(89) \qquad \widetilde{\theta}_{i+\frac{1}{2}, j} := \text{value of } \widetilde{\theta} \text{ evaluated at } \widetilde{\boldsymbol{P}}_{i+\frac{1}{2}, j} = \frac{1}{2}\left(\widetilde{\boldsymbol{P}}_{i, j} + \widetilde{\boldsymbol{P}}_{i+1, j} \right),$$

we rewrite I_h^H,
(90)
$$I_h^H = \sum_{i=0}^{M} \sum_{j=0}^{N+1} \mathcal{M}_{i+\frac{1}{2},j} \begin{bmatrix} \widetilde{u}_{i+1,j} - \widetilde{u}_{i,j} \\ \widetilde{u}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{u}_{i+\frac{1}{2},j-\frac{1}{2}} \end{bmatrix} \widetilde{\theta}_{i+\frac{1}{2},j} + \|\widetilde{u}_h\|_{\widetilde{V}_h} \mathcal{O}(\overline{h}) =$$
$$\sum_{i=0}^{M} \sum_{j=0}^{N+1} \begin{bmatrix} m_{i+\frac{1}{2},j}^{11}(\widetilde{u}_{i+1,j} - \widetilde{u}_{i,j}) + m_{i+\frac{1}{2},j}^{12}(\widetilde{u}_{i+1,j+1} + \widetilde{u}_{i,j+1} - \widetilde{u}_{i+1,j-1} - \widetilde{u}_{i,j-1}) \\ m_{i+\frac{1}{2},j}^{21}(\widetilde{u}_{i+1,j} - \widetilde{u}_{i,j}) + m_{i+\frac{1}{2},j}^{22}(\widetilde{u}_{i+1,j+1} + \widetilde{u}_{i,j+1} - \widetilde{u}_{i+1,j-1} - \widetilde{u}_{i,j-1}) \end{bmatrix}$$
$$\times \widetilde{\theta}_{i+\frac{1}{2},j} + \|\widetilde{u}_h\|_{\widetilde{V}_h} \mathcal{O}(\overline{h}).$$

Integrating by parts, we find that

(91)
$$I_h^H = I_h^{H,I} + I_h^{H,II} + \|\widetilde{u}_h\|_{\widetilde{V}_h} \mathcal{O}(\overline{h}),$$

where
(92)
$$I_h^{H,I} = -\frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i+\frac{1}{2},j}^{11} + m_{i-\frac{1}{2},j}^{11} \\ m_{i+\frac{1}{2},j}^{21} + m_{i-\frac{1}{2},j}^{21} \end{bmatrix} \left(\widetilde{\theta}_{i+\frac{1}{2},j} - \widetilde{\theta}_{i-\frac{1}{2},j} \right)$$
$$-\frac{1}{8} \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i+\frac{1}{2},j+1}^{12} + m_{i+\frac{1}{2},j-1}^{12} \\ m_{i+\frac{1}{2},j+1}^{22} + m_{i+\frac{1}{2},j-1}^{22} \end{bmatrix} \left(\widetilde{\theta}_{i+\frac{1}{2},j+1} - \widetilde{\theta}_{i+\frac{1}{2},j-1} \right)$$
$$-\frac{1}{8} \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i-\frac{1}{2},j+1}^{12} + m_{i-\frac{1}{2},j-1}^{12} \\ m_{i-\frac{1}{2},j+1}^{22} + m_{i-\frac{1}{2},j-1}^{22} \end{bmatrix} \left(\widetilde{\theta}_{i-\frac{1}{2},j+1} - \widetilde{\theta}_{i-\frac{1}{2},j-1} \right),$$

and
(93)
$$I_h^{H,II} = -\frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i+\frac{1}{2},j}^{11} - m_{i-\frac{1}{2},j}^{11} \\ m_{i+\frac{1}{2},j}^{21} - m_{i-\frac{1}{2},j}^{21} \end{bmatrix} \left(\widetilde{\theta}_{i+\frac{1}{2},j} + \widetilde{\theta}_{i-\frac{1}{2},j} \right)$$
$$-\frac{1}{8} \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i+\frac{1}{2},j+1}^{12} - m_{i+\frac{1}{2},j-1}^{12} \\ m_{i+\frac{1}{2},j+1}^{22} - m_{i+\frac{1}{2},j-1}^{22} \end{bmatrix} \left(\widetilde{\theta}_{i+\frac{1}{2},j+1} + \widetilde{\theta}_{i+\frac{1}{2},j-1} \right)$$
$$-\frac{1}{8} \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i-\frac{1}{2},j+1}^{12} - m_{i-\frac{1}{2},j-1}^{12} \\ m_{i-\frac{1}{2},j+1}^{22} - m_{i-\frac{1}{2},j-1}^{22} \end{bmatrix} \left(\widetilde{\theta}_{i-\frac{1}{2},j+1} + \widetilde{\theta}_{i-\frac{1}{2},j-1} \right).$$

Thanks to the assumption ($\mathcal{H}4$), we simplify some expressions in (92) and (93) which are related to $\mathcal{M}_{1+1/2,j}$ defined in (75) and (88): Using (72), we find that,

for $2 \leq i \leq M - 1$ and $1 \leq j \leq N$,

(94)
$$
\begin{cases}
\begin{aligned}
m^{11}_{i+\frac{1}{2},j} + m^{11}_{i-\frac{1}{2},j} &= \frac{1}{2}\big(\widetilde{y}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{y}_{i+\frac{1}{2},j-\frac{1}{2}}\big) + \frac{1}{2}\big(\widetilde{y}_{i-\frac{1}{2},j+\frac{1}{2}} - \widetilde{y}_{i-\frac{1}{2},j-\frac{1}{2}}\big) \\
&= \frac{1}{2}\big(\widetilde{y}_{i,j+1} - \widetilde{y}_{i,j-1}\big) + o\big(\overline{h}^2\big), \\
m^{12}_{i+\frac{1}{2},j+1} + m^{12}_{i+\frac{1}{2},j-1} &= \frac{1}{2}\big(-\widetilde{y}_{i+1,j+1} + \widetilde{y}_{i,j+1}\big) + \frac{1}{2}\big(-\widetilde{y}_{i+1,j-1} + \widetilde{y}_{i,j-1}\big) \\
&= -\widetilde{y}_{i+1,j} + \widetilde{y}_{i,j} + o\big(\overline{h}^2\big) \\
&= -\frac{1}{2}\big(\widetilde{y}_{i+1,j} - \widetilde{y}_{i-1,j}\big) + o\big(\overline{h}^2\big), \\
m^{12}_{i-\frac{1}{2},j+1} + m^{12}_{i-\frac{1}{2},j-1} &= -\widetilde{y}_{i,j} + \widetilde{y}_{i-1,j} + o\big(\overline{h}^2\big) \\
&= -\frac{1}{2}\big(\widetilde{y}_{i+1,j} - \widetilde{y}_{i-1,j}\big) + o\big(\overline{h}^2\big),
\end{aligned}
\end{cases}
$$

and

(95)
$$
\begin{cases}
\begin{aligned}
m^{11}_{i+\frac{1}{2},j} - m^{11}_{i-\frac{1}{2},j} &= \frac{1}{2}\big(\widetilde{y}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{y}_{i+\frac{1}{2},j-\frac{1}{2}}\big) - \frac{1}{2}\big(\widetilde{y}_{i-\frac{1}{2},j+\frac{1}{2}} - \widetilde{y}_{i-\frac{1}{2},j-\frac{1}{2}}\big) \\
&= o\big(\overline{h}^2\big), \\
m^{12}_{i+\frac{1}{2},j+1} - m^{12}_{i+\frac{1}{2},j-1} &= \frac{1}{2}\big(-\widetilde{y}_{i+1,j+1} + \widetilde{y}_{i,j+1}\big) + \frac{1}{2}\big(\widetilde{y}_{i+1,j-1} - \widetilde{y}_{i,j-1}\big) \\
&= \frac{1}{2}\big(-\widetilde{y}_{i,j+1} + \widetilde{y}_{i-1,j+1}\big) + \frac{1}{2}\big(\widetilde{y}_{i+1,j-1} - \widetilde{y}_{i,j-1}\big) + o\big(\overline{h}^2\big) \\
&= o\big(\overline{h}^2\big).
\end{aligned}
\end{cases}
$$

By symmetry, we also find that, for $2 \leq i \leq M - 1$ and $1 \leq j \leq N$,

(96)
$$
\begin{cases}
m^{21}_{i+\frac{1}{2},j} + m^{21}_{i-\frac{1}{2},j} = -\frac{1}{2}\big(\widetilde{x}_{i,j+1} - \widetilde{x}_{i,j-1}\big) + o\big(\overline{h}^2\big), \\
m^{22}_{i+\frac{1}{2},j+1} + m^{22}_{i+\frac{1}{2},j-1} = \frac{1}{2}\big(\widetilde{x}_{i+1,j} - \widetilde{x}_{i-1,j}\big) + o\big(\overline{h}^2\big), \\
m^{22}_{i-\frac{1}{2},j+1} + m^{22}_{i-\frac{1}{2},j-1} = \frac{1}{2}\big(\widetilde{x}_{i+1,j} - \widetilde{x}_{i-1,j}\big) + o\big(\overline{h}^2\big), \\
m^{21}_{i+\frac{1}{2},j} - m^{21}_{i-\frac{1}{2},j} = o\big(\overline{h}^2\big), \\
m^{22}_{i+\frac{1}{2},j+1} - m^{22}_{i+\frac{1}{2},j-1} = o\big(\overline{h}^2\big).
\end{cases}
$$

Using (93), (95), and (96), we notice that

(97)
$$
\big|I^{H,II}_h\big| \leq \|\widetilde{u}_h\|_{\widetilde{V}_h}\, o(1).
$$

Now using the Taylor expansion, we write $I_h^{H,I}$ in the form,

(98)
$$I_h^{H,I} = -\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i+\frac{1}{2},j}^{1\,1} + m_{i-\frac{1}{2},j}^{1\,1} \\ m_{i+\frac{1}{2},j}^{2\,1} + m_{i-\frac{1}{2},j}^{2\,1} \end{bmatrix} \nabla\widetilde{\theta}_{i,j}\cdot\big(\widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j} - \widetilde{\boldsymbol{P}}_{i-\frac{1}{2},j}\big)$$

$$-\frac{1}{8}\sum_{i=1}^{M}\sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i+\frac{1}{2},j+1}^{1\,2} + m_{i+\frac{1}{2},j-1}^{1\,2} \\ m_{i+\frac{1}{2},j+1}^{2\,2} + m_{i+\frac{1}{2},j-1}^{2\,2} \end{bmatrix} \nabla\widetilde{\theta}_{i,j}\cdot\big(\widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j+1} - \widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j-1}\big)$$

$$-\frac{1}{8}\sum_{i=1}^{M}\sum_{j=1}^{N} \widetilde{u}_{i,j} \begin{bmatrix} m_{i-\frac{1}{2},j+1}^{1\,2} + m_{i-\frac{1}{2},j-1}^{1\,2} \\ m_{i-\frac{1}{2},j+1}^{2\,2} + m_{i-\frac{1}{2},j-1}^{2\,2} \end{bmatrix} \nabla\widetilde{\theta}_{i,j}\cdot\big(\widetilde{\boldsymbol{P}}_{i-\frac{1}{2},j+1} - \widetilde{\boldsymbol{P}}_{i-\frac{1}{2},j-1}\big)$$

$$+\|\widetilde{u}_h\|_{\widetilde{V}_h}\,\mathcal{O}(\overline{h}).$$

Using the assumption $(\mathcal{H}4)$, we simplify the vectors appearing in (98):

(99)
$$\begin{cases} \widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j} - \widetilde{\boldsymbol{P}}_{i-\frac{1}{2},j} &= \frac{1}{2}\big(\widetilde{\boldsymbol{P}}_{i+1,j} - \widetilde{\boldsymbol{P}}_{i-1,j}\big), \\[2mm] \widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j+1} - \widetilde{\boldsymbol{P}}_{i+\frac{1}{2},j-1} &= \frac{1}{2}\big(\widetilde{\boldsymbol{P}}_{i+1,j+1} + \widetilde{\boldsymbol{P}}_{i,j+1} - \widetilde{\boldsymbol{P}}_{i+1,j-1} - \widetilde{\boldsymbol{P}}_{i,j-1}\big), \\[2mm] &= \widetilde{\boldsymbol{P}}_{i,j+1} - \widetilde{\boldsymbol{P}}_{i,j-1} + \mathcal{O}(\overline{h}^2), \\[2mm] \widetilde{\boldsymbol{P}}_{i-\frac{1}{2},j+1} - \widetilde{\boldsymbol{P}}_{i-\frac{1}{2},j-1} &= \frac{1}{2}\big(\widetilde{\boldsymbol{P}}_{i,j+1} + \widetilde{\boldsymbol{P}}_{i-1,j+1} - \widetilde{\boldsymbol{P}}_{i,j-1} - \widetilde{\boldsymbol{P}}_{i-1,j-1}\big), \\[2mm] &= \widetilde{\boldsymbol{P}}_{i,j+1} - \widetilde{\boldsymbol{P}}_{i,j-1} + \mathcal{O}(\overline{h}^2). \end{cases}$$

Using (94), (96), and (99), we rewrite $I_h^{H,I}$ in (98) in the form,

(100)
$$I_h^{H,I} = -\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N} \widetilde{u}_{i,j}\,\nabla\widetilde{\theta}_{i,j}\,\mathbf{A}_{i,j} + \|\widetilde{u}_h\|_{\widetilde{V}_h}\,\mathcal{O}(1),$$

where the matrix $\mathbf{A}_{i,j}$ is defined by

(101)
$$\mathbf{A}_{i,j} = \frac{1}{4}\big\{\big(\widetilde{y}_{i,j+1} - \widetilde{y}_{i,j-1}\big)\big(\widetilde{x}_{i+1,j} - \widetilde{x}_{i-1,j}\big)$$

$$-\big(\widetilde{y}_{i+1,j} - \widetilde{y}_{i-1,j}\big)\big(\widetilde{x}_{i,j+1} - \widetilde{x}_{i,j-1}\big)\big\}\,\mathbf{I}_{2\times 2}$$

$$= \frac{1}{2}(\text{area of quadrilateral connecting vertices }\widetilde{\boldsymbol{P}}_{i\pm 1,j}\text{ and }\widetilde{\boldsymbol{P}}_{i,j\pm 1})\,\mathbf{I}_{2\times 2}.$$

On the other hand, using (3) and $(\mathcal{H}4)$, we observe that

(102)
$$|\widetilde{K}_{i,j}| = \frac{1}{2}\det\begin{bmatrix} \widetilde{x}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{x}_{i-\frac{1}{2},j-\frac{1}{2}} & \widetilde{x}_{i+\frac{1}{2},j-\frac{1}{2}} - \widetilde{x}_{i-\frac{1}{2},j+\frac{1}{2}} \\ \widetilde{y}_{i+\frac{1}{2},j+\frac{1}{2}} - \widetilde{y}_{i-\frac{1}{2},j-\frac{1}{2}} & \widetilde{y}_{i+\frac{1}{2},j-\frac{1}{2}} - \widetilde{y}_{i-\frac{1}{2},j+\frac{1}{2}} \end{bmatrix}$$

$$= \frac{1}{4}\big\{\big(\widetilde{y}_{i,j+1} - \widetilde{y}_{i,j-1}\big)\big(\widetilde{x}_{i+1,j} - \widetilde{x}_{i-1,j}\big)$$

$$-\big(\widetilde{y}_{i+1,j} - \widetilde{y}_{i-1,j}\big)\big(\widetilde{x}_{i,j+1} - \widetilde{x}_{i,j-1}\big)\big\} + \mathcal{O}(\overline{h}^2).$$

Therefore we finally deduce from (91), (97), and (100)-(102) that

$$
(103) \quad
\begin{aligned}
I_h^H &= -\frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \widetilde{u}_{i,j} \, \nabla\theta\big(\widetilde{\boldsymbol{P}}_{i,j}\big) \, \big|\widetilde{K}_{i,j}\big| + \|\widetilde{u}_h\|_{\widetilde{V}_h} \, \mathcal{O}(1) \\
&= -\frac{1}{2} \int_\Omega \widetilde{u}_h \, \nabla\theta \, d\boldsymbol{x} + \|\widetilde{u}_h\|_{\widetilde{V}_h} \, \mathcal{O}(1).
\end{aligned}
$$

Thanks to the symmetry between the expressions of I_h^H and I_h^V, using the same method as for I_h^H, one can verify that I_h^V in (86) can be written as the right-hand side of (103). Therefore, we finally have

$$
(104) \qquad I_h = - \int_\Omega \widetilde{u}_h \, \nabla\theta \, d\boldsymbol{x} + \|\widetilde{u}_h\|_{\widetilde{V}_h} \, \mathcal{O}(1).
$$

We deduce from (81) that the term I_h in (104) attains the limit announced in (85). Hence the property $(\mathcal{C}2)$ for the FD discretization follows. $\square$

Now, we conclude that the FD approximation, constructed in Section 5, is stable and convergent in the sense of [17, 57]. We summarize the result below as a proposition:

Proposition 6.4. *Under the assumptions $(\mathcal{H}1)$-$(\mathcal{H}4)$, the FD discretization method under consideration is stable and convergent. More precisely, we have*

$(\mathcal{S})$ $\|\widetilde{p}_h\|_{L(\widetilde{V}_h,\, L^2(\Omega)^3)} \le \kappa,$

$(\mathcal{C}1)$ $(\widetilde{p}_h \circ \widetilde{r}_h)(u) \to \overline{\omega}(u)$ in $L^\infty(\Omega)^3$ (hence in $L^2(\Omega)^3$) as $\overline{h} \to 0$, $\forall u \in C_0^\infty(\Omega)$,

$(\mathcal{C}2)$ If $\widetilde{u}_h \in \widetilde{V}_h$ and $\widetilde{p}_h(\widetilde{u}_h) \rightharpoonup \phi$ weakly in $L^2(\Omega)^3$ as $\overline{h} \to 0$, then $\phi \in \overline{\omega}\big(H_0^1(\Omega)\big)$.

7. Comparison between Finite Volumes and Finite Differences

The Finite Difference and Finite Volume spaces, $\widetilde{V}_h$ and V_h, are related by a bijective map Λ_h from $\widetilde{V}_h$ to V_h defined by:

$$
(105) \qquad \Lambda_h \widetilde{u}_h = \sum_{i=1}^{N} \sum_{j=1}^{M} \widetilde{u}_{i,j} \chi_{K_{i,j}}, \quad \text{for} \quad \widetilde{u}_h = \sum_{i=1}^{N} \sum_{j=1}^{M} \widetilde{u}_{i,j} \chi_{\widetilde{K}_{i,j}} \in \widetilde{V}_h.
$$

The inverse Λ_h^{-1} of Λ_h is defined as well:

$$
(106) \qquad \Lambda_h^{-1} u_h = \sum_{i=1}^{N} \sum_{j=1}^{M} u_{i,j} \chi_{\widetilde{K}_{i,j}}, \quad \text{for} \quad u_h = \sum_{i=1}^{N} \sum_{j=1}^{M} u_{i,j} \chi_{K_{i,j}} \in V_h.
$$

To prove the property $(\mathcal{C}2)$ for FV, we further assume:

$(\mathcal{H}5)$ The FV and FD nodal points are close to each other in the sense that

$$
\boldsymbol{P}_{i+1/2,\,j+1/2} = \widetilde{\boldsymbol{P}}_{i+1/2,\,j+1/2} + \mathcal{O}\big(\overline{h}\big),
$$

for $0 \le i \le M$ and $0 \le j \le N$; see Remark 7.3 below as well.

Remark 7.1. In Remark 7.3 below, we introduce a condition weaker than $(\mathcal{H}5)$ (but a bit more complicated) which is sufficient to prove the Lemma 7.2 below and hence the convergence of the FV method (the main result in this article) stated in Theorem 7.4. However, for simplicity, we stay mainly with the condition $(\mathcal{H}5)$ throughout this article because it is easily computationally verified for a given mesh in many practical applications.

Our next task is to prove the following lemma:

Lemma 7.2. *Under the assumptions* ($\mathcal{H}1$)-($\mathcal{H}3$) *and* ($\mathcal{H}5$), *we have*

$$(107) \quad \left| \int_\Omega \left(\nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h \right) \varphi \, d\boldsymbol{x} \right| \leq \|u_h\|_{V_h} \, \mathcal{O}(1), \quad \forall u_h \in V_h, \ \forall \varphi \in C_0^\infty(\mathbb{R}^2).$$

Proof. We first write
$$(108)$$
$$\left| \int_\Omega \left(\nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h \right) \varphi \, d\boldsymbol{x} \right| \leq \left| \sum_{i=0}^{M} \sum_{j=1}^{N} \int_{K_{i+\frac{1}{2}, j}} \left(\nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h \right) \varphi \, d\boldsymbol{x} \right|$$
$$+ \left| \sum_{i=1}^{M} \sum_{j=0}^{N} \int_{K_{i, j+\frac{1}{2}}} \left(\nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h \right) \varphi \, d\boldsymbol{x} \right|.$$

On $K_{i+\frac{1}{2}, j} \cap \widetilde{K}_{i+\frac{1}{2}, j}$, using (30), (32), (72), (74), and (106), we notice that

$$(109) \qquad\qquad \nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h = J_h^1 + J_h^2,$$

where
$$(110)$$
$$\begin{cases} J_h^1 = \left(\mathbf{M}_{i+\frac{1}{2}, j}^{-1} - \widetilde{\mathbf{M}}_{i+\frac{1}{2}, j}^{-1} \right) \begin{bmatrix} u_{i+1, j} - u_{i, j} \\ u_{i+\frac{1}{2}, j+\frac{1}{2}} - u_{i+\frac{1}{2}, j-\frac{1}{2}} \end{bmatrix} \\[4mm] J_h^2 = \\[2mm] \widetilde{\mathbf{M}}_{i+\frac{1}{2}, j}^{-1} \begin{bmatrix} 0 \\ \left(u_{i+\frac{1}{2}, j+\frac{1}{2}} - \dfrac{1}{4} \sum\limits_{l,m=0}^{1} u_{i+l, j+m} \right) - \left(u_{i+\frac{1}{2}, j-\frac{1}{2}} - \dfrac{1}{4} \sum\limits_{l,m=0}^{1} u_{i+l, j+m-1} \right) \end{bmatrix} \end{cases}.$$

We notice from ($\mathcal{H}5$) that

$$(111) \qquad\qquad \left| \overrightarrow{\boldsymbol{P}_{i+\frac{1}{2}, j-\frac{1}{2}} \boldsymbol{P}_{i+\frac{1}{2}, j+\frac{1}{2}}} - \overrightarrow{\widetilde{\boldsymbol{P}}_{i+\frac{1}{2}, j-\frac{1}{2}} \widetilde{\boldsymbol{P}}_{i+\frac{1}{2}, j+\frac{1}{2}}} \right| = \mathcal{o}(\overline{h}),$$

for $0 \leq i \leq M$ and $1 \leq j \leq N$. Then, using (111) and the fact that $K_{i+1/2, j}$ and $\widetilde{K}_{i+1/2, j}$ share a diagonal connecting $\boldsymbol{P}_{i, j}$ and $\boldsymbol{P}_{i+1, j}$, we observe that

$$(112) \qquad \left| K_{i+\frac{1}{2}, j} \right| = \left| \widetilde{K}_{i+\frac{1}{2}, j} \right| + \mathcal{o}(\overline{h}^2), \quad 0 \leq i \leq M, \ 1 \leq j \leq N.$$

Concerning the first term J_h^1 in (110), we first use (33), (75), and (112), and notice that

$$(113) \qquad \left| \mathbf{M}_{i+\frac{1}{2}, j}^{-1} - \widetilde{\mathbf{M}}_{i+\frac{1}{2}, j}^{-1} \right| \leq \frac{\left| K_{i+\frac{1}{2}, j} \right| - \left| \widetilde{K}_{i+\frac{1}{2}, j} \right|}{\underline{h}^4} \, \overline{h} \leq \frac{\mathcal{o}(1)}{\underline{h}}.$$

Then, using the Schwarz inequality as well, we write
$$(114)$$
$$\left| \sum_{i=0}^{M} \sum_{j=1}^{N} \int_{K_{i+\frac{1}{2}, j} \cap \widetilde{K}_{i+\frac{1}{2}, j}} J_h^1 \, \varphi \, d\boldsymbol{x} \right|$$
$$\leq \mathcal{o}(1) \sum_{i=0}^{M} \sum_{j=1}^{N} \left\{ \left| u_{i+1, j} - u_{i, j} \right| + \left| u_{i+\frac{1}{2}, j+\frac{1}{2}} - u_{i+\frac{1}{2}, j-\frac{1}{2}} \right| \right\} \overline{h} \leq \|u_h\|_{V_h} \, \mathcal{o}(1).$$

Thanks to ($\mathcal{H}5$) and the definition of the intermediate value $u_{i+1/2,\,j+1/2}$, we see that

(115)
$$\left| u_{i+\frac{1}{2},\,j+\frac{1}{2}} - \frac{1}{4} \sum_{j,\,m=0}^{1} u_{i+l,\,j+m} \right| \le \mathcal{o}(\overline{h}) \sum_{j,\,m=0}^{1} |u_{i+l,\,j+m}|, \quad 0 \le i \le M,\ 0 \le j \le N.$$

Using (115), Lemma 4.2, and the discrete Poincaré inequality, we write

(116)
$$\left| \sum_{i=0}^{M} \sum_{j=1}^{N} \int_{K_{i+\frac{1}{2},\,j} \cap \widetilde{K}_{i+\frac{1}{2},\,j}} J_h^2\, \varphi\, d\boldsymbol{x} \right| \le \frac{\mathcal{o}(\overline{h})}{\underline{h}} \sum_{i=1}^{M} \sum_{j=1}^{N} |u_{i,\,j}| \|\varphi\|_{L^\infty(\Omega)}\, \overline{h}^2$$
$$\le \|u_h\|_{V_h}\, \mathcal{o}(1).$$

Combining (114) and (116), we find that

(117)
$$\left| \sum_{i=0}^{M} \sum_{j=1}^{N} \int_{K_{i+\frac{1}{2},\,j} \cap \widetilde{K}_{i+\frac{1}{2},\,j}} \left(\nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h \right) \varphi\, d\boldsymbol{x} \right| \le \|u_h\|_{V_h}\, \mathcal{o}(1).$$

On $K_{i+\frac{1}{2},\,j}^* := K_{i+\frac{1}{2},\,j} \setminus \widetilde{K}_{i+\frac{1}{2},\,j}$, we notice from ($\mathcal{H}5$) that

(118)
$$\left| K_{i+\frac{1}{2},\,j}^* \right| = \mathcal{o}(\overline{h}^2), \quad 0 \le i \le M,\ 1 \le j \le N.$$

Then we find that

(119)
$$\left| \sum_{i=0}^{M} \sum_{j=1}^{N} \int_{K_{i+\frac{1}{2},\,j}^*} \left(\nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h \right) \varphi\, d\boldsymbol{x} \right|$$
$$\le \sum_{i=0}^{M} \sum_{j=1}^{N} \left\{ \left| \nabla_h u_h \big|_{K_{i+\frac{1}{2},\,j}} \right| + \frac{1}{\underline{h}} |u_{i+1,\,j} - u_{i,\,j}| + \frac{1}{\underline{h}} |u_{i,\,j+1} - u_{i,\,j}| \right\} \left| K_{i+\frac{1}{2},\,j}^* \right|$$
$$\le \|u_h\|_{V_h} \frac{1}{\underline{h}} \left(\sum_{i=0}^{M} \sum_{j=1}^{N} \left| K_{i+\frac{1}{2},\,j}^* \right|^2 \right)^{\frac{1}{2}} \le \|u_h\|_{V_h}\, \mathcal{o}(1).$$

We deduce from (117) and (119) that

(120)
$$\left| \sum_{i=0}^{M} \sum_{j=1}^{N} \int_{K_{i+\frac{1}{2},\,j}} \left(\nabla_h u_h - \widetilde{\nabla}_h \Lambda_h^{-1} u_h \right) \varphi\, d\boldsymbol{x} \right| \le \|u_h\|_{V_h}\, \mathcal{o}(1).$$

By symmetry, one can verify that (120) with $K_{i+1/2,\,j}$ replaced by $K_{i,\,j+1/2}$ holds true as well. Then, finally (107) follows from (108) and (120), and the proof of Lemma 7.2 is complete. $\square$

Remark 7.3. With some additional assumptions, the condition ($\mathcal{H}5$) can be relaxed to:

$$\boldsymbol{P}_{i+1/2,\,j+1/2} = \widetilde{\boldsymbol{P}}_{i+1/2,\,j+1/2} + \mathcal{o}(1), \quad 0 \le i \le M,\ 0 \le j \le N.$$

In fact, the Lemma 107 can be verified under the relaxed condition above together with (112), (118), and a technical assumption,

$$\begin{cases} \left| K_{i+\frac{1}{2},\,j} \cap \widetilde{K}_{i+\frac{1}{2},\,j} \right| = \left| K_{i+\frac{1}{2},\,j+1} \cap \widetilde{K}_{i+\frac{1}{2},\,j+1} \right| + \mathcal{O}(\overline{h}^3), \\ \left| K_{i,\,j+\frac{1}{2}} \cap \widetilde{K}_{i,\,j+\frac{1}{2}} \right| = \left| K_{i+1,\,j+\frac{1}{2}} \cap \widetilde{K}_{i+1,\,j+\frac{1}{2}} \right| + \mathcal{O}(\overline{h}^3), \end{cases}$$

for $1 \leq i \leq M - 1$, $1 \leq j \leq N - 1$. See Remark 7.1 above as well.

7.1. Proof of $(\mathcal{C}2)$ for FV. We choose a family u_h in V_h such that

$$(121) \qquad u_h \rightharpoonup \phi_0, \quad \nabla_h u_h \rightharpoonup (\phi_1, \phi_2) \quad \text{weakly in } L^2(\Omega) \text{ as } \overline{h} \to 0.$$

To prove the property $(\mathcal{C}2)$ for the FV, since $(\mathcal{C}_2)$ has been already proven for the FD, it is enough to show that

$$(122) \qquad \begin{cases} \Lambda_h^{-1} u_h \rightharpoonup \phi_0 \quad \text{weakly in } L^2(\Omega) \text{ as } \overline{h} \to 0, \\ \widetilde{\nabla}_h \Lambda_h^{-1} u_h \rightharpoonup (\phi_1, \phi_2) \quad \text{weakly in } L^2(\Omega) \text{ as } \overline{h} \to 0. \end{cases}$$

Thanks to the definition of the FV derivatives, it is easy to verify that

$$(123) \qquad |u_h - \Lambda_h^{-1} u_h|_{L^2(\Omega)} \leq \|u_h\|_{V_h} \, \mathcal{O}(\overline{h});$$

hence $(122)_1$ follows from (121).

Thanks to Lemma 7.2, $(122)_2$ follows from (121) as well, and then the property $(\mathcal{C}2)$ for the FV space is finally inferred.

Together with the stability and the property $(C1)$, which were proved in Section 4, we now conclude that the FV approximation, constructed in Section 3, is stable and convergent in the sense of [17, 57]. We summarize this below as the main result of this article:

Theorem 7.4. *Under the assumptions $(\mathcal{H}1)$-$(\mathcal{H}5)$, the present cell-centered FV discretization method is stable and convergent. More precisely, we have*

$(\mathcal{S})$ $\|p_h\|_{L(V_h, L^2(\Omega)^3)} \leq \kappa,$

$(\mathcal{C}1)$ $(p_h \circ r_h)(u) \to \overline{\omega}(u)$ in $L^\infty(\Omega)^3$ (hence in $L^2(\Omega)^3$) as $\overline{h} \to 0$, $\forall u \in C_0^\infty(\Omega)$,

$(\mathcal{C}2)$ If $u_h \in V_h$ and $p_h(u_h) \rightharpoonup \phi$ weakly in $L^2(\Omega)^3$ as $\overline{h} \to 0$, then $\phi \in \overline{\omega}(H_0^1(\Omega))$.

8. An application

In this section, we construct the FV approximation of a class of classical coercive elliptic equations. Then, thanks to the convergence results, that is the properties $(\mathcal{C}1)$ and $(\mathcal{C}2)$ of the FV method, we prove that a discrete solution to the FV weak formulation converges to the weak solution of the original problem. More precisely, we consider an elliptic equation, supplemented with the homogeneous Dirichlet boundary condition in the form,

$$(124) \qquad \begin{cases} -\operatorname{div}\left(\boldsymbol{D}(x, y) \cdot \nabla u\right) + \operatorname{div}\left(\boldsymbol{b}(x, y)\, u\right) + g(x, y)u = f(x, y) \quad \text{in } \Omega, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad u = 0 \quad \text{on } \Gamma, \end{cases}$$

where Ω is a 2D polygonal domain as considered in the previous sections, and g and f are given smooth functions in Ω. The smooth data $\boldsymbol{D}$ and $\boldsymbol{b}$ are defined by

$$(125) \qquad \boldsymbol{D}(x, y) := \left[a^{\alpha\beta}(x, y)\right]_{1 \leq \alpha, \beta \leq 2}, \quad \boldsymbol{b}(x, y) := \left(b^1(x, y),\, b^2(x, y)\right).$$

Using (125), the equation $(124)_1$ can be written in the form,

$$(126) \qquad -\sum_{\alpha, \beta=1}^{2} \partial_\alpha \left(a^{\alpha\beta}(x, y) \partial_\beta u\right) + \sum_{\alpha=1}^{2} \partial_\alpha \left(b^\alpha(x, y)u\right) + g(x, y)u = f(x, y),$$

where $\partial_1 = \partial/\partial x$ and $\partial_2 = \partial/\partial y$.

We introduce the function spaces,

$$(127) \qquad H := L^2(\Omega), \quad V := H_0^1(\Omega), \quad V' := (H_0^1(\Omega))' = H^{-1}(\Omega).$$

Then, using (126) and (127), we classically write the weak formulation of (124):

Given $f \in H$, find u in V such that

$$(128) \qquad \mathrm{a}(u,\, v) = l(v), \quad \forall v \in V,$$

where

$$(129) \qquad \mathrm{a}(u,\, v) = \sum_{\alpha,\beta=1}^{2} \int_{\Omega} a^{\alpha\beta}\, \partial_{\beta} u\, \partial_{\alpha} v\, d\boldsymbol{x} - \sum_{\alpha=1}^{2} \int_{\Omega} b^{\alpha}\, u\, \partial_{\alpha} v\, d\boldsymbol{x} + \int_{\Omega} guv\, d\boldsymbol{x},$$

and

$$(130) \qquad l(v) = (f,\, v)_H = \int_{\Omega} f\, v\, d\boldsymbol{x}.$$

For the simplicity of our analysis below, we assume that, for each $\alpha, \beta = 1, 2$,

$$(131) \qquad a^{\alpha\beta},\, b^{\alpha},\, g \in C(\overline{\Omega}),$$

and we guarantee the coercivity by assuming that

$$(132) \qquad \begin{cases} \displaystyle\sum_{\alpha,\beta=1}^{2} a^{\alpha\beta}(x,y)\, \xi_\alpha \xi_\beta \geq \kappa_{\mathrm{a}1} |\boldsymbol{\xi}|^2, \quad \boldsymbol{\xi} = (\xi_1, \xi_2) \in \mathbb{R}^2, \quad (x,\, y) \in \Omega, \\[2mm] \|g\|_{L^\infty(\Omega)} - \dfrac{1}{\kappa_{\mathrm{a}1}} \|\boldsymbol{b}\|_{L^\infty(\Omega)}^2 \geq \kappa_{\mathrm{a}2}, \end{cases}$$

for some strictly positive constants $\kappa_{\mathrm{a}1}$ and $\kappa_{\mathrm{a}2}$. In fact, one can verify that, under the assumptions (131) and (132), the continuous bilinear form $\mathrm{a}(\cdot,\, \cdot)$ on $V \times V$ is coercive:

$$(133) \qquad \mathrm{a}(u,\, u) \geq \kappa_{\mathrm{a}} \|u\|_V^2, \quad u \in V,$$

for a strictly positive constant $\kappa_{\mathrm{a}} := \min(\kappa_{\mathrm{a}1}/2,\, \kappa_{\mathrm{a}2})$.

Then, thanks to the Lax-Milgram theorem, there exists a unique weak solution u to the variational problem (128). In what follows, we will study the FV approximation of the weak solution u.

8.1. FV scheme for the problem (124). To define the discretized variational form of (124) (or (126)), we restrict the equation to a fixed control volume $K_{i,j}$ and integrate it over $K_{i,j}$. Then, by using the definition of the FV space in Section 3, we write the discrete bilinear form associated with (128) in the form,

$$(134) \qquad \mathrm{a}_h(u_h,\, v_h) = \mathrm{a}_h^D(u_h,\, v_h) + \mathrm{a}_h^{\boldsymbol{b}}(u_h,\, v_h) + \mathrm{a}_h^g(u_h,\, v_h),$$

where

$$(135) \qquad \begin{cases} \mathrm{a}_h^D(u_h,\, v_h) = \displaystyle\sum_{\alpha,\beta=1}^{2} \sum_{i=0}^{M} \sum_{j=1}^{N} a_{i+\frac{1}{2},\,j}^{\alpha\beta} \left(\nabla_h^\beta u_h \nabla_h^\alpha v_h\right)\big|_{K_{i+\frac{1}{2},\,j}} \left|K_{i+\frac{1}{2},\,j}\right| \\[3mm] \qquad\qquad + \displaystyle\sum_{\alpha,\beta=1}^{2} \sum_{i=1}^{M} \sum_{j=0}^{N} a_{i,\,j+\frac{1}{2}}^{\alpha\beta} \left(\nabla_h^\beta u_h \nabla_h^\alpha v_h\right)\big|_{K_{i,\,j+\frac{1}{2}}} \left|K_{i,\,j+\frac{1}{2}}\right|, \\[3mm] \mathrm{a}_h^{\boldsymbol{b}}(u_h,\, v_h) = \displaystyle\sum_{\alpha=1}^{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=E,W,N,S} b_{i,\,j}^\alpha\, u_{i,\,j}\, \nabla_h^\alpha v_h\big|_{K_{i,\,j}^k} \left|K_{i,\,j}^k\right|, \\[3mm] \mathrm{a}_h^g(u_h,\, v_h) = \displaystyle\sum_{i=1}^{M} \sum_{j=1}^{N} g_{i,\,j}\, u_{i,\,j}\, v_{i,\,j} \left|K_{i,\,j}\right|. \end{cases}$$

Here we set

$$(136) \quad \begin{cases} a^{\alpha\beta}_{i+\frac{1}{2},\,j} = a^{\alpha\beta}\Big(\frac{1}{2}(\boldsymbol{P}_{i,\,j} + \boldsymbol{P}_{i+1,\,j})\Big), & 0 \le i \le M,\ 1 \le j \le N, \\[2mm] a^{\alpha\beta}_{i,\,j+\frac{1}{2}} = a^{\alpha\beta}\Big(\frac{1}{2}(\boldsymbol{P}_{i,\,j} + \boldsymbol{P}_{i,\,j+1})\Big), & 1 \le i \le M,\ 0 \le j \le N, \\[2mm] b^{\alpha}_{i,\,j} = b^{\alpha}\big(\boldsymbol{P}_{i,\,j}\big), \quad g_{i,\,j} = g\big(\boldsymbol{P}_{i,\,j}\big), & 1 \le i \le M,\ 1 \le j \le N, \end{cases}$$

and

$$(137) \quad \begin{cases} K^{E}_{i,\,j} = K_{i,\,j} \cap K_{i+\frac{1}{2},\,j}, & K^{W}_{i,\,j} = K_{i,\,j} \cap K_{i-\frac{1}{2},\,j}, \\[2mm] K^{N}_{i,\,j} = K_{i,\,j} \cap K_{i,\,j+\frac{1}{2}}, & K^{S}_{i,\,j} = K_{i,\,j} \cap K_{i,\,j-\frac{1}{2}}, \end{cases}$$

for $1 \le i \le M$, $1 \le j \le N$, so that $\sum_{k=E,W,N,S} K^{k}_{i,\,j} = K_{i,\,j}$. Using the definition of the FV derivative, we also set $\nabla^{1}_{h} = \nabla^{x}_{h} = (1,\,0)\cdot\nabla_{h}$ and $\nabla^{2}_{h} = \nabla^{y}_{h} = (0,\,1)\cdot\nabla_{h}$.

For the right-hand side of (124) (or (126)), we define the continuous linear functional l_{h} on V_{h} by setting:

$$(138) \qquad l_{h}(v_{h}) = (r_{h}f,\, v_{h})_{V_{h}} = \sum_{i=1}^{M} \sum_{j=1}^{N} f_{i,\,j}\, v_{i,\,j}\big|K_{i,\,j}\big|,$$

where

$$(139) \qquad f_{i,\,j} = \frac{1}{\big|K_{i,\,j}\big|} \int_{K_{i,\,j}} f\, d\boldsymbol{x}, \quad 1 \le i \le M,\ 1 \le j \le N.$$

Since $f \in H$, we see that the linear functionals l_{h} are uniformly continuous with respect to the mesh size $\overline{h}$.

Now, using (134) and (138), we introduce the discrete FV variational approximation of (128):

Given $f \in H$ (hence $r_{h}f \in V_{h}$), find u_{h} in V_{h} such that

$$(140) \qquad \mathrm{a}_{h}(u_{h},\, v_{h}) = l_{h}(v_{h}), \quad \forall v_{h} \in V_{h}.$$

8.2. Convergence of the FV approximation of (124).

We aim to prove first the existence of a unique solution to the discrete FV variational problem (140), and second the convergence of the discrete FV weak solution u_{h} to the weak solution u of (128) strongly in $H^{1}(\Omega)$ as the mesh size $\overline{h}$ tends to zero, using the results proven in Theorem 7.4.

We first prove the uniform continuity and coercivity of the bilinear forms a_{h}:

Lemma 8.1. *The bilinear forms* a_{h} *are continuous and coercive uniformly with respect to the mesh size* $\overline{h}$.

Proof. Using the Schwarz inequality, we estimate $a_h^{\boldsymbol{b}}(u_h,\, v_h)$ in $(135)_2$ in the form,

$$\left| a_h^{\boldsymbol{b}}(u_h,\, v_h) \right|$$

$$\leq \left| \sum_{\alpha=1}^{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=E,W,N,S} b_{i,\,j}^{\alpha}\, u_{i,\,j}\, \nabla_h^{\alpha} v_h|_{K_{i,\,j}^{k}} \left| K_{i,\,j}^{k} \right| \right|$$

(141)
$$\leq \|\boldsymbol{b}\|_{L^\infty(\Omega)} \left(\sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=E,W,N,S} u_{i,\,j}^2 \left| K_{i,\,j}^{k} \right| \right)^{\frac{1}{2}}$$

$$\times \left(\sum_{\alpha=1}^{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=E,W,N,S} \left(\nabla_h^{\alpha} v_h|_{K_{i,\,j}^{k}} \right)^2 \left| K_{i,\,j}^{k} \right| \right)^{\frac{1}{2}}$$

$$\leq \|\boldsymbol{b}\|_{L^\infty(\Omega)} |u_h|_{V_h} \|v_h\|_{V_h}.$$

Then the uniform continuity of a_h follows from (131), (134), (135) and (141).

The uniform coercivity of a_h follows from (132), (134), (135) and (141) with v_h replaced by u_h. Lemma 8.1 is now proved. □

Using Lemma 8.1, the Lax-Milgram theorem asserts that the discrete FV variational problem (140) has a unique solution u_h in V_h.

In view of proving the convergence of the FV approximate solution u_h to the weak solution u of (128) as the mesh size $\overline{h}$ tends to zero, we recall the following consistency lemma whose proof can be found in [42]:

Lemma 8.2. *If v_h converges to v strongly in V as $\overline{h}$ tends to zero, and if w_h converges to w weakly in V as $\overline{h}$ tends to zero, then we have*

(142)
$$\begin{cases} \lim_{\overline{h} \to 0} a_h(v_h, w_h) = a(v, w), \\[2mm] \lim_{\overline{h} \to 0} a_h(w_h, v_h) = a(w, v), \\[2mm] \lim_{\overline{h} \to 0} l_h(v_h) = l(v). \end{cases}$$

Thanks to the general convergence theorem in [17] and [57], the convergence of the FV weak solution u_h to the original weak solution u finally follows from Lemma 8.2:

Theorem 8.3. *Under the assumptions $(\mathcal{H}1)$-$(\mathcal{H}5)$, the FV approximate solution u_h of (140) converges to the weak solution u of (128) strongly in V as the mesh size $\overline{h}$ tends to zero, that is,*

(143)
$$\left(u_h,\, \nabla_h u_h \right) \to \left(u,\, \nabla u \right) \quad \text{strongly in } L^2(\Omega)^3 \text{ as } \overline{h} \to 0.$$

Remark 8.4. Equations similar to (124) have been considered in the literature, see, e.g., [20, 33, 4, 21, 5, 7, 35]; however the analysis depends often on the equation considered. Because of the high generality of our construction based on hypotheses on the mesh which can be easily computationally verified, and which are independent of the problem under consideration and other existing results, Theorem 8.3 can be extended to many linear and nonlinear problems. In particular, nonlinear elliptic operators of the monotone type can be considered. We can also consider nonlinear operators of the Leray-Lions type, also called pseudo-monotone operators [50, 15], as in, e.g., [26, 25, 29, 6, 13, 7], at the price of proving a discrete compactness theorem [32]. Note that the Leray-Lions equations that we can consider are not necessarily of the divergence form as in, e.g., [26] and thus are more general than those of [26]. We refrain to do so to avoid making the article too long.

Acknowledgements

References

[1] K. Adamy and D. Pham, A finite volume implicit Euler scheme for the linearized shallow water equations: stability and convergence, Numer. Funct. Anal. Optim., 27 (2006) 757–783.

[2] M. Afif and B. Amaziane, Convergence of finite volume schemes for a degenerate convection-diffusion equation arising in flow in porous media, Comput. Methods Appl. Mech. Engrg., 191 (2002) 5265–5286.

[3] Léo Agélas, Daniele A. Di Pietro, and Jérôme Droniou. The G method for heterogeneous anisotropic diffusion on general meshes, M2AN Math. Model. Numer. Anal., 44 (2010) 597–625.

[4] B. Andreianov, M. Bendahmane, and K. H. Karlsen, Discrete duality finite volume schemes for doubly nonlinear degenerate hyperbolic-parabolic equations, J. Hyperbolic Differ. Equ., 7 (2010) 1–67.

[5] B. Andreianov, R. Eymard, M. Ghilani, and N. Marhraoui, Finite volume approximation of degenerate two-phase flow model with unlimited air mobility, Numer. Methods Partial Differential Equations, 29 (2013) 441–474.

[6] B. Andreianov, F. Boyer, and F. Hubert, Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes. Numer. Methods Partial Differential Equations, 23 (2007) 145–195.

[7] B. Andreianov, M. Bendahmane, and F. Hubert, On 3D DDFV discretization of gradient and divergence operators: discrete functional analysis tools and applications to degenerate parabolic problems, Comput. Methods Appl. Math., 13 (2013) 369–410.

[8] O. Angelini, K. Brenner, and D. Hilhorst, A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation, Numer. Math., 123 (2013) 219–257.

[9] I. Aavatsmark, T. Barkve, Ø. Bøe, and T. Mannseth, Discretization on unstructured grids for inhomogeneous, anisotropic media. I. Derivation of the methods, SIAM J. Sci. Comput., 19 (1998) 1700–1716 (electronic).

[10] P. Balland and E. Süli, Analysis of the cell-vertex finite volume method for hyperbolic problems with variable coefficients, SIAM J. Numer. Anal., 34 (1997) 1127–1151.

[11] A. Bousquet, M. Marion, and R. Temam, Finite volume multilevel approximation of the shallow water equations, Chin. Ann. Math. Ser. B, 34 (2013) 1–28.

[12] F. Boyer, Analysis of the upwind finite volume method for general initial- and boundary-value transport problems, IMA J. Numer. Anal., 32 (2012) 1404–1439.

[13] F. Boyer and F. Hubert, Finite volume method for 2D linear and nonlinear elliptic problems with discontinuities, SIAM J. Numer. Anal., 46 (2008) 3032–3070.

[14] A. Bradji. Convergence analysis of some high-order time accurate schemes for a finite volume method for second order hyperbolic equations on general nonconforming multidimensional spatial meshes, Numer. Methods PDEs, 29 (2013) 1278–1321.

[15] H. Brezis, Équations et inéquations non linéaires dans les espaces vectoriels en dualité, Ann. Inst. Fourier (Grenoble), 18 (1968) 115–175.

[16] Z. Q. Cai, On the finite volume element method, Numer. Math., 58 (1991) 713–735.

[17] J. Céa, Approximation variationnelle des problèmes aux limites, Ann. Inst. Fourier (Grenoble), 14 (1964) 345–444.

[18] T. J. Chung, Computational fluid dynamics, Cambridge University Press, Cambridge, 2002.

[19] P. G. Ciarlet and P. A. Raviart, The combined effect of curved boundaries and numerical integration in isoparametric finite element methods, In The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972), pages 409–474. Academic Press, New York, 1972.

[20] Y. Coudière and G. Manzini, The discrete duality finite volume method for convection-diffusion problems, SIAM J. Numer. Anal., 47 (2010) 4163–4192.

[21] Y. Coudière, C. Pierre, O. Rousseau, and R. Turpault, A 2D/3D discrete duality finite volume scheme. Application to ECG simulation, Int. J. Finite Vol., 6 (2009) 1634–0655.

[22] Y. Coudière, J. Vila, and P. Villedieu, Convergence rate of a finite volume scheme for a two-dimensional convection-diffusion problem, M2AN Math. Model. Numer. Anal., 33 (1999) 493–516.

[23] K. Domelevo and P. Omnes, A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids, M2AN Math. Model. Numer. Anal., 39 (2005) 1203–1249.

[24] J. Droniou, Finite volume schemes for diffusion equations: Introduction to and review of modern methods, Math. Model. Methods Appl. Sci., pages 1–45.

[25] J. Droniou, Finite volume schemes for fully non-linear elliptic equations in divergence form, M2AN Math. Model. Numer. Anal., 40 (2006) 1069–1100 (2007).

[26] J. Droniou, R. Eymard, T. Gallouet, and R. Herbin, Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic equations, Math. Models Methods Appl. Sci., 23 (2013) 2395–2432.

[27] A. Evgrafov, M. M. Gregersen, and M. P. Sørensen, Convergence of cell based finite volume discretizations for problems of control in the conduction coefficients, ESAIM Math. Model. Numer. Anal., 45 (2011) 1059–1080.

[28] R. Eymard, T. Gallouët, and R. Herbin, A cell-centered finite-volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension, IMA J. Numer. Anal., 26 (2006) 326–353.

[29] R. Eymard, T. Gallouët, and R. Herbin, Cell centred discretisation of non linear elliptic problems on general multidimensional polyhedral grids, J. Numer. Math., 17 (2009) 173–193.

[30] R. Eymard, T. Gallouët, and R. Herbin, Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces, IMA J. Numer. Anal., 30 (2010) 1009–1043.

[31] R. Eymard, T. Gallouët, R. Herbin, and A. Linke, Finite volume schemes for the biharmonic problem on general meshes, Math. Comp., 81 (2012) 2019–2048.

[32] R. Eymard, T. Gallouët, and R. Herbin, Finite volume methods. In Handbook of numerical analysis, Vol. VII, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.

[33] R. Eymard, T. Gallouët, Ra. Herbin, and A. Michel, Convergence of a finite volume scheme for nonlinear degenerate parabolic equations, Numer. Math., 92 (2002) 41–82.

[34] R. Eymard, C. Guichard, and R. Herbin, Small-stencil 3D schemes for diffusive flows in porous media, ESAIM Math. Model. Numer. Anal., 46 (2012) 265–290.

[35] R. Eymard, A. Handlovičová, and K. Mikula, Study of a finite volume scheme for the regularized mean curvature flow level set equation, IMA J. Numer. Anal., 31 (2011) 813–846.

[36] S. Faure, J. Laminie, and R. Temam, Finite volume discretization and multilevel methods in flow problems, J. Sci. Comput., 25 (2005) 231–261.

[37] S. Faure, J. Laminie, and R. Temam, Colocated finite volume schemes for fluid flows, Commun. Comput. Phys., 4 (2008) 1–25.

[38] S. Faure, M. Petcu, R. Temam, and J. Tribbia, On the inaccuracies of some finite volume discretizations of the linearized shallow water problem, Int. J. Numer. Anal. Model., 8 (2011) 518–541.

[39] S. Faure, D. Pham, and R. Temam, Comparison of finite volume and finite difference methods and application, Anal. Appl. (Singap.), 4 (2006) 163–208.

[40] X. Feng, R., Y. He, and D. Liu, P_1-nonconforming quadrilateral finite volume methods for the semilinear elliptic equations, J. Sci. Comput., 52 (2012) 519–545.

[41] T. Gallouët, A. Larcher, and J. C. Latché, Convergence of a finite volume scheme for the convection-diffusion equation with L^1 data, Math. Comp., 81 (2012) 1429–1454.

[42] G.-M. Gie and R. Temam, Cell centered finite volume methods using Taylor series expansion scheme without fictitious domains, Int. J. Numer. Anal. Model., 7 (2010) 1–29.

[43] R. Herbin and F. Hubert, Benchmark on discretization schemes for anisotropic diffusion problems on general grids, In Finite volumes for complex applications V, pages 659–692. ISTE, London, 2008.

[44] F. Hermeline, A finite volume method for the approximation of diffusion operators on distorted meshes, J. Comput. Phys., 160 (200) 481–499.

[45] W. Huang and A. M. Kappen. A study of cell–center finite volume methods for diffusion equations, electronic, 1998.

[46] Y. Nen Jeng and J. L. Chen, Geometric conservation law of the finite-volume method for the simpler algorithm and a proposed upwind scheme, Numerical Heat Transfer, Part B: Fundamentals, 22 (1992) 211–234.

[47] C.-Y. Jung and R. Temam, Finite volume approximation of one-dimensional stiff convection-diffusion equations. J. Sci. Comput., 41 (2009) 384–410.

[48] C.-Y. Jung and R. Temam, Finite volume approximation of two-dimensional stiff problems, Int. J. Numer. Anal. Model., 7 (2010) 462–476.

[49] G. S. B. Lebon, M. K. Patel, and K. A. Pericleous, Investigation of instabilities arising with non-orthogonal meshes used in cell centred elliptic finite volume computations, J. Algorithms Comput. Technol., 6 (2012) 129–152.

[50] J. Leray and J.-L. Lions, Quelques résulatats de Višik sur les problèmes elliptiques non-linéaires par les méthodes de Minty-Browder, Bull. Soc. Math. France, 93 (1965) 97–107.

[51] R. J. LeVeque, Finite volume methods for hyperbolic problems. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.

[52] K. W. Morton. Numerical solution of convection-diffusion problems, volume 12 of Applied Mathematics and Mathematical Computation, Chapman & Hall, London, 1996.

[53] K. W. Morton, M. Stynes, and E. Süli, Analysis of a cell-vertex finite volume method for convection-diffusion problems, Math. Comp., 66 (1997) 1389–1406.

[54] K. W. Morton and E. Süli, Finite volume methods and their analysis, IMA J. Numer. Anal., 11 (1991) 241–260.

[55] E. Süli, Convergence of finite volume schemes for Poisson's equation on nonuniform meshes, SIAM J. Numer. Anal., 28 (1991) 1419–1430.

[56] E. Süli, The accuracy of cell vertex finite volume methods on quadrilateral meshes, Math. Comp., 59 (1992) 359–382.

[57] R. Temam. Navier-Stokes equations, AMS Chelsea Publishing, Providence, RI, 2001. Theory and numerical analysis, Reprint of the 1984 edition.

[58] R. M. Temam and J. J. Tribbia, editors, Handbook of numerical analysis. Vol. XIV. Special volume: computational methods for the atmosphere and the oceans, volume 14 of Handbook of Numerical Analysis, Elsevier/North-Holland, Amsterdam, 2009.

[59] M. Vanmaele, K. W. Morton, E. Süli, and A. Borzì, Analysis of the cell vertex finite volume method for the Cauchy-Riemann equations, SIAM J. Numer. Anal., 34 (1997) 2043–2062.

[60] Q. Zhao and G. Yuan, Analysis and construction of cell-centered finite volume scheme for diffusion equations on distorted meshes, Comput. Methods Appl. Mech. Engrg., 198 (2009) 3039–3050.

Department of Mathematics, University of Louisville, Louisville, KY 40292, U.S.A.
E-mail: `gungmin.gie@louisville.edu`

Department of Mathematics and Institute for Scientific Computing and Applied Mathematics, Indiana University, 831 East 3rd St., Bloomington, IN 47405, U.S.A.
E-mail: `temam@indiana.edu`