A VERTEX-CENTERED LINEARITY-PRESERVING DISCRETIZATION OF DIFFUSION PROBLEMS ON POLYGONAL MESHES

Jiming Wu, Zhiming Gao, Zhihuan Dai

A VERTEX-CENTERED LINEARITY-PRESERVING DISCRETIZATION OF DIFFUSION PROBLEMS ON POLYGONAL MESHES*

JIMING WU*, ZHIMING GAO*†, AND ZIHUAN DAI*

Abstract. This paper introduces a vertex-centered linearity-preserving finite volume scheme for the heterogeneous anisotropic diffusion equations on general polygonal meshes. The unknowns of this scheme are purely the values at the mesh vertices, and no auxiliary unknowns are utilized. The scheme is locally conservative with respect to the dual mesh, captures exactly the linear solutions, leads to a symmetric positive definite matrix, and yields a nine-point stencil on structured quadrilateral meshes. The coercivity of the scheme is rigorously analyzed on arbitrary mesh size under some weak geometry assumptions. Also the relation with the finite volume element method is discussed. Finally some numerical tests show the optimal convergence rates for the discrete solution and flux on various mesh types and for various diffusion tensors.

Key words. diffusion equation, vertex-centered scheme, linearity-preserving

AMS subject classifications. 65M08, 35R05, 76S05

1. Introduction.

We consider the anisotropic steady-state diffusion problem

$$-\mathrm{div}(\Lambda\nabla u) = f \quad \text{in } \Omega, \tag{1.1}$$

$$u = g_D \quad \text{on } \Gamma_D, \tag{1.2}$$

$$-\Lambda\nabla u\cdot\boldsymbol{n} = g_N \quad \text{on } \Gamma_N, \tag{1.3}$$

where Ω is a bounded connected polygonal domain in $\mathbb{R}^2$, $\boldsymbol{n}$ denotes the exterior unit normal vector along the domain boundary $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, Λ is a 2×2 symmetric diffusion tensor, uniformly bounded above and below, *i.e.*, there exist positive constants $\overline{\kappa}$ and $\underline{\kappa}$, such that

$$\underline{\kappa}\|\boldsymbol{v}\|^2 \le \boldsymbol{v}^T\Lambda\boldsymbol{v} \le \overline{\kappa}\|\boldsymbol{v}\|^2, \quad \forall \boldsymbol{v}\in\mathbb{R}^2, \tag{1.4}$$

f, g_D and g_N denote the source term, the Dirichlet and flux boundary data, respectively. Anisotropic diffusion problem of this type arises in a wide range of scientific fields such as oil reservoir simulations, plasma physics, semiconductor modeling and so on. In accurate simulation of diffusion processes in these applications, finite volume (FV) method is among the most commonly used ones due to its simplicity, local conservation and some other good numerical properties.

*This work was supported by the National Natural Science Fund of China (Nos. 91330205, 11271053, 91330107, 11471047) and Foundation of President of China Academy of Engineering Physics (2014-1-042).

*Institute of Applied Physics and Computational Mathematics, P. O. Box 8009, Beijing 100088, P. R. China

†Corresponding author. E-mail:gao@iapcm.ac.cn

In recent decades, numerous efforts have been devoted to the construction of efficient FV schemes, which can be roughly classified as cell-centered schemes, hybrid schemes, mixed schemes, nonlinear positivity-preserving schemes and so on. The reader is referred to [14, 9] for some recent developments. All these FV schemes have cell-centered unknowns and some other possible ones defined on cell boundaries, and most of them satisfy the local conservation condition (continuity of the flux), which is very important in some practical applications such as radiation hydrodynamics. Cell-centered schemes have only one unknown for each cell and in this respect, they are more attractive than the hybrid or mixed ones. However, cell-centered schemes usually result in asymmetric linear systems and as a result, their coercivity and convergence analysis are quite difficult to obtain.

By comparison with cell-centered FV schemes, vertex-centered schemes have drawn less attention. A vertex-centered or nodal scheme based on mimetic method with a low-order accuracy was proposed in [2]. Its extensions to 2D and 3D cases with arbitrary order of accuracy were studied in [7] and [18], respectively. In these works, the lower-order methods reduce to the standard P_1 finite element method (FEM) in the case of simplicial meshes. The virtual element method[6] has been proposed recently which can be viewed as a further development of the nodal mimetic method, reformulating in a pure finite element framework of the two methods in [7] and [2]. In [11], a similar low-order vertex-centered scheme was suggested in terms of gradient schemes. Actually, all these methods can be considered as various extensions of Galerkin finite element method where the test function space coincides with the trial function space and as a result, symmetric and positive definite linear systems can be expected. However, the local conservation property is either lost or cannot be understood in the same sense as that of cell-centered schemes.

Another type of vertex-centered FV schemes comes from the finite volume element method (FVEM)[13, 4, 23] (sometimes called as generalized difference method[16] or box method[1]), which can be viewed as a certain Petrov-Galerkin finite element method and thus are mainly valid for triangular or quadrilateral meshes. The trial function space of FVEM is the same as that in FEM while the test function space consists of piecewise constant functions with respect to the so-called dual mesh, an adjoint mesh constructed from the original primary mesh. Generally speaking, FVEM are simple and locally conservative, however, they usually exhibit a different nature when compared with their counterparts in FEM having the same trial function space. For example, FVEM with trial function space of P_1 type (P_1-FVEM for short) on triangular meshes cannot have the standard $O(h^2)$ convergence rate in the L^2 norm when the source function only belongs to L^2 [15, 16, 5, 8], and optimal L^2 error estimates cannot be obtained directly by Aubin-Nitsche's duality technique[20]. Some higher order FVEM on quadrilateral meshes have recently been suggested [21, 28], which are not pure vertex-centered schemes, involving some other unknowns defined on the cell boundaries or inside the cells. Moreover, on quadrilateral meshes, FVEM is

valid only on $h^{1+\gamma}-$parallelogram meshes with $\gamma > 0$, at least it seems so theoretically.

The vertex-centered FV scheme proposed in this paper seems to be more related to the low-order FVEM, since its construction is based on the same primary and dual mesh in FVEM. However, the derivation of the scheme is performed along a different line of thought, i.e., the linearity-preserving approach developed in some exiting cell-centered or hybrid FV schemes[25, 27, 24, 26] where the concept of trial function space is not involved. The primary mesh here consists of arbitrary polygonal grids, instead of being confined to triangular or quadrilateral ones. Moreover, the symmetry and positive definiteness of the new scheme is always expected, which is not shared by FVEM except for some cases on triangular meshes. In summary, the new scheme has the following characteristics:

- It has only vertex-centered unknowns;
- It has a local stencil, a nine-point one on structured quadrilateral meshes;
- It is applicable to arbitrary polygonal grids, which may have concave cells or degenerate ones with hanging nodes;
- It allows heterogeneous full diffusion tensors;
- It reduces to a P_1-FVEM scheme on triangular meshes;
- It is locally conservative with respect to the dual mesh;
- It leads to symmetric and positive definite linear systems;
- It satisfies the linearity-preserving property in the sense that the scheme captures the exact solution if the diffusion tensor is piecewise constant and the solution is piecewise linear with respect to the primary mesh;
- It has approximately second-order accuracy on general meshes in case that the diffusion tensor is taken to be anisotropic and/or discontinuous.

More interesting is that the new scheme possesses simultaneously the three properties: the local conservation, the symmetry and positive definiteness, and the linearity-preserving, which is rarely seen in existing cell-centered or vertex-centered FV schemes.

The rest of the paper is organized as follows. In section 2, we describe the general construction algorithm for the new scheme, leaving the key ingredient, i.e., the construction of cell matrix to section 3. In section 4 we discuss the issues of the symmetry and coercivity. Numerical experiments are carried out in section 5 and some conclusions are given in the last section.

2. A new vertex-centered linearity-preserving scheme.

2.1. The primary and dual meshes. The construction of the primary and dual meshes is almost the same as that in FVEM. Suppose that Ω is partitioned into a number of non-overlapped polygonal cells that form the so-called *primary mesh*, see the mesh with solid line segments in Figure 2.1. The vertices of Ω and the possible joint points of Γ_D and Γ_N must be included in the set of primary vertices. For a primary cell, its cell center is defined at any point in the cell. Each primary cell is further partitioned into several quadrilateral subcells by connecting the cell center with the edge midpoints, see the dashed line segments in Figure 2.1. All subcells sharing a

same vertex of the primary mesh constitute a cell of the *dual mesh*. Throughout this paper, we shall always assume that

(H1) Each cell in the primary mesh is star-shaped with respect to its cell center.

By this assumption, the dual mesh makes sense.

2.2. The primary unknowns. At each vertex of the primary mesh in $\Omega \cup \Gamma_N$, we define a single primary unknown, see the solid points in Figure 2.1. A finite volume equation will be constructed associated with each primary unknown. Note that the primary vertices on $\bar{\Gamma}_D$ have no primary unknowns. Contrary to the cell-centered linearity-preserving FV schemes studied before, here we do not need to introduce any auxiliary unknowns so that there is no interpolation algorithm.

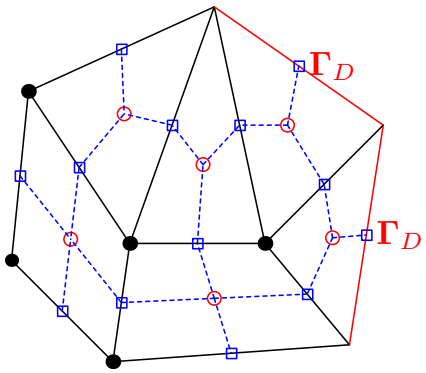


FIG. 2.1. *The primary mesh (solid line), dual mesh (dashed line) and the boundary* Γ_D *(red line).*

2.3. The flux discretization. From now on, all the derivations are conducted under the following assumptions:

1. The solution is smooth inside each primary cell and continuous on the whole domain $\bar{\Omega}$, while the diffusion tensor is constant on each primary cell.
2. The possible discontinuities of the solution gradient and the diffusion tensor are only allowed to appear on the edges of the primary mesh.
3. The normal component of the flux $\mathbf{F} = -\Lambda \nabla u$ is continuous across all interior edges of the dual mesh.

Obviously, the first and the second assumptions are standard and the same as those for cell-centered schemes. However, the last assumption is different from that of a cell-centered scheme, where the flux is assumed to be continuous across all interior edges of the primary mesh.

As shown in Figure 2.2, we introduce some notations.

- K, a generic primary cell with n_K edges whose cell center, measure and diameter are denoted as $\boldsymbol{x}_K$, $|K|$ and h_K, respectively.
- $\mathcal{M}$, the set of primary cells in $\bar{\Omega}$ and $h = \max_{K \in \mathcal{M}} h_K$ denotes the mesh size.
- $\mathcal{E}_K$, the set of edges in K.

- σ, a generic edge of K with measure denoted as $|\sigma|$, also the local numbering of edges in $\mathcal{E}_K$, depending on the context.
- $\boldsymbol{x}_\nu$ and $\boldsymbol{x}_{\nu^+}$, two generic vertices of the primary mesh, also the two endpoints of σ.
- $\boldsymbol{x}_\sigma$, the midpoint of σ.
- K_ν^*, a generic dual cell associated with $\boldsymbol{x}_\nu$ whose outward unit normal along the cell boundary is denoted as $\boldsymbol{n}_\nu^*$.
- σ_K^*, a generic dual edge connecting $\boldsymbol{x}_K$ and $\boldsymbol{x}_\sigma$.
- $\boldsymbol{n}_{K,\sigma}^*$, a unit vector normal to σ_K^* whose direction is fixed once and for all.
- u_ν, u_{ν^+}, the primary unknowns defined at $\boldsymbol{x}_\nu$ and $\boldsymbol{x}_{\nu^+}$, respectively.
- Λ_K, the constant restriction of Λ on K.
- F_{K,σ^*}, the discrete counterpart or approximation of $\int_{\sigma_K^*}(-\Lambda_K\nabla u)\cdot\boldsymbol{n}_{K,\sigma}^*\mathrm{d}s$.

In addition, for $\sigma\in\mathcal{E}_K$, we assume that its endpoint $\boldsymbol{x}_\nu$ is always pointed anticlockwisely to the other one $\boldsymbol{x}_{\nu^+}$. We also assume that all n_K unit vectors $\boldsymbol{n}_{K,\sigma}^*(\sigma\in\mathcal{E}_K)$ inside K are ordered clockwise (see Figure 2.2). As a result, $\boldsymbol{n}_{K,\sigma}^*\cdot\boldsymbol{n}_\nu^*$ may be either 1 or -1. Throughout, the hollow letters $\mathbb{A},\mathbb{F},\mathbb{U},\cdots$ will be used to denote rectangular matrices with a number of columns greater than one while the bold ones $\boldsymbol{F},\boldsymbol{x},\boldsymbol{n},\cdots$ will represent vectors or matrices with only one column.

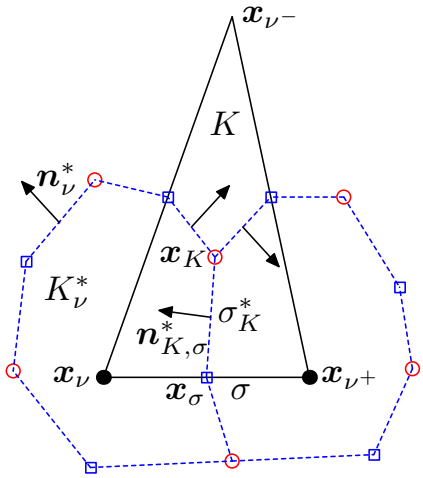


FIG. 2.2. *Notations for the flux discretization.*

Now we are ready to describe the construction of the flux approximation F_{K,σ^*}. First, we put all $F_{K,\sigma^*},\sigma\in\mathcal{E}_K$ in the same group and manipulate them together, instead of treating them one by one. Note that the flux approximations for the same dual cell may belong to different groups. Secondly, we seek the following local algebraic relation related to primary cell K,

$$\boldsymbol{F}_K=\mathbb{A}_K\delta\boldsymbol{U}_K, \tag{2.1}$$

where $\mathbb{A}_K$ is the so-called cell matrix of size $n_K \times n_K$, $\boldsymbol{F}_K = (F_{K,\sigma^*}, \sigma \in \mathcal{E}_K)^T$ and $\delta \boldsymbol{U}_K = (u_{\nu^+} - u_\nu, \sigma \in \mathcal{E}_K)^T$ are two vectors of size n_K, containing all the flux approximations in the same group and the successive differences of the primary unknowns on ∂K, respectively. Here we take for example the triangle cell K in Figure 2.2 to illustrate (2.1). For any discrete algorithm for F_{K,σ^*}, we can always expect the algebraic expression

$$F_{K,\sigma^*} = c u_\nu + c_+ u_{\nu^+} + c_- u_{\nu^-}. \tag{2.2}$$

If this formula is required to be exact for constant solutions, then we have $c+c_++c_- = 0$, leading to the following new expression

$$F_{K,\sigma^*} = c_+(u_{\nu^+} - u_\nu) + 0 \times (u_{\nu^-} - u_{\nu^+}) - c_-(u_\nu - u_{\nu^-}),$$

or equivalently the matrix form (2.1). Surely, for a specific discretization procedure, $\mathbb{A}_K$ may not be unique since the sum of the entries of $\delta \boldsymbol{U}_K$ is zero. However, F_{K,σ^*}, as the linear combination of the primary unknowns shown in (2.2), remains the same.

Finally, we establish the so-called linearity-preserving criterion for the construction of the cell matrix $\mathbb{A}_K$. Let $\mathbb{F}_K$ and $\mathbb{X}_K$ be two $n_K \times 2$ matrices defined respectively by

$$\mathbb{F}_K = \left(-|\sigma_K^*| \Lambda_K \boldsymbol{n}_{K,\sigma}^*,\ \sigma \in \mathcal{E}_K\right)^T, \quad \mathbb{X}_K = \left(\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_\nu,\ \sigma \in \mathcal{E}_K\right)^T. \tag{2.3}$$

We recall that, in the linearity-preserving method, all derivations are required to be exact whenever the solution is piecewise linear and the diffusion tensor is piecewise constant with respect to the primary mesh. By direct calculation, we find that (2.1) satisfies this linearity-preserving criterion if and only if

$$\mathbb{F}_K = \mathbb{A}_K \mathbb{X}_K. \tag{2.4}$$

How to construct $\mathbb{A}_K$ such that (2.4) is fulfilled is another issue and will be addressed in section 3.

2.4. The final vertex-centered scheme. For an vertex $\boldsymbol{x}_\nu \in \Omega \cup \Gamma_N$, let $\mathcal{M}_\nu$ (resp. $\mathcal{E}_\nu$) be the set of primary cells (resp. edges) sharing $\boldsymbol{x}_\nu$. The finite volume equation associated with primary unknown u_ν is constructed as follows

$$\sum_{K \in \mathcal{M}_\nu} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_\nu} \left(\boldsymbol{n}_{K,\sigma}^* \cdot \boldsymbol{n}_\nu^*\right) F_{K,\sigma^*} = |K_\nu^*| f_{K_\nu^*} - \int_{\partial K_\nu^*} g_N ds, \tag{2.5}$$

where F_{K,σ^*} is given by (2.1), $|K_\nu^*|$ denotes the measure of the dual cell K_ν^* and

$$f_{K_\nu^*} = \frac{1}{|K_\nu^*|} \sum_{K \in \mathcal{M}_\nu} \int_{K_\nu^* \cap K} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

3. Construction of the cell matrix $\mathbb{A}_K$. We have seen that cell matrix $\mathbb{A}_K$ plays an important role in the new vertex-centered scheme and different cell matrices will result in different schemes. In this section, we shall make use of the special structure of 2D polygonal meshes to construct a symmetric and positive definite cell matrix $\mathbb{A}_K$ such that the linearity-preserving criterion (2.4) is satisfied.

LEMMA 3.1. *For the two matrices $\mathbb{F}_K$ and $\mathbb{X}_K$ defined in (2.3), we have*

$$\mathbb{F}_K^T\mathbb{X}_K = |K|\Lambda_K. \tag{3.1}$$

Proof. We start the argumentation by introducing the following formula (see, e.g., (2.17) in [10] or equivalently (3.7) in [27]):

$$\sum_{\sigma\in\mathcal{E}_K} |\sigma|(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)\boldsymbol{n}_{K,\sigma}^T = |K|\mathbb{I}_2, \tag{3.2}$$

where $\boldsymbol{n}_{K,\sigma}$ denotes the unit vector normal to σ outward from K and $\mathbb{I}_2$ the 2×2 identity matrix. Define

$$\mathcal{R} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then, it follows from (3.2) that

$$\sum_{\sigma\in\mathcal{E}_K} |\sigma|\Lambda_K\mathcal{R}(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)\boldsymbol{n}_{K,\sigma}^T\mathcal{R}^T = |K|\Lambda_K\mathcal{R}\mathcal{R}^T.$$

Noticing $\mathcal{R}(\boldsymbol{x}_\sigma - \boldsymbol{x}_K) = |\sigma_K^*|\boldsymbol{n}_{K,\sigma}^*$ and $|\sigma|\mathcal{R}\boldsymbol{n}_{K,\sigma} = -(\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_\nu)$, we arrive at

$$-\sum_{\sigma\in\mathcal{E}_K} |\sigma_K^*|\Lambda_K\boldsymbol{n}_{K,\sigma}^*(\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_\nu)^T = |K|\Lambda_K,$$

which leads to (3.1) and concludes the proof. ∎

Based on (3.1) and inspired by [3, 24, 26], we suggest that

$$\mathbb{A}_K = \frac{1}{|K|}\mathbb{F}_K\Lambda_K^{-1}\mathbb{F}_K^T + \gamma_K\mathbb{C}_K\mathbb{C}_K^T \quad \text{with} \quad \mathbb{C}_K = \mathbb{I}_K - \frac{1}{|K|}\mathbb{F}_K\Lambda_K^{-1}\mathbb{X}_K^T, \tag{3.3}$$

or equivalently,

$$\mathbb{A}_K = \frac{1}{|K|}\widetilde{\mathbb{F}}_K\Lambda_K\widetilde{\mathbb{F}}_K^T + \gamma_K\mathbb{C}_K\mathbb{C}_K^T \quad \text{with} \quad \mathbb{C}_K = \mathbb{I}_K - \frac{1}{|K|}\widetilde{\mathbb{F}}_K\mathbb{X}_K^T, \tag{3.4}$$

where $\mathbb{I}_K$ is an $n_K\times n_K$ identity matrix,

$$\widetilde{\mathbb{F}}_K = \left(-|\sigma_K^*|\boldsymbol{n}_{K,\sigma}^*, \sigma\in\mathcal{E}_K\right)^T, \tag{3.5}$$

and γ_K is a positive parameter. In practical computation, we choose

$$\gamma_K = \frac{1}{100}\text{trace}\left(\frac{1}{|K|}\mathbb{F}_K\Lambda_K^{-1}\mathbb{F}_K^T\right). \tag{3.6}$$

THEOREM 3.2. *The cell matrix $\mathbb{A}_K$, defined by (3.3), satisfies (2.4) and is symmetric and positive definite.*

Proof. (2.4) can be directly verified by (3.1) and (3.3). Obviously, $\mathbb{A}_K$ is symmetric and semi-positive definite. Suppose that there exists a vector $\boldsymbol{v} \in \mathbb{R}^{n_K}$, such that $\boldsymbol{v}^T \mathbb{A}_K \boldsymbol{v}$ equals to the zero vector $\mathbf{0}$. Then, from (3.3), we have

$$\mathbb{F}_K^T \boldsymbol{v} = \mathbf{0} \quad \text{and} \quad \mathbb{C}_K^T \boldsymbol{v} = \boldsymbol{v} - \frac{1}{|K|} \mathbb{X}_K \Lambda_K^{-1} \mathbb{F}_K^T \boldsymbol{v} = \mathbf{0}, \tag{3.7}$$

which implies $\boldsymbol{v} = \mathbf{0}$ and completes the proof. □

4. Symmetry and coercivity.

4.1. General results.

We first introduce two assumptions.

(H2) There exists a positive constant $\underline{\alpha}$, independent of mesh size h, such that

$$|K| \geq \underline{\alpha} h_K^2, \quad \forall K \in \mathcal{M}. \tag{4.1}$$

(H3) For the matrix $\mathbb{C}_K$ defined by (3.3), there exists a positive constant $\underline{\lambda}$, independent of h, such that

$$\|\mathbb{C}_K^T \mathbb{C}_K \boldsymbol{v}\|^2 \geq \underline{\lambda} \|\mathbb{C}_K \boldsymbol{v}\|^2, \quad \forall \boldsymbol{v} \in \mathbb{R}^{n_K}, \quad \forall K \in \mathcal{M}, \tag{4.2}$$

where $\|\cdot\|$ denotes the Euclidean vector norm.

Secondly, we define a discrete H^1 norm $|\cdot|_{1,\mathcal{M}}$, given by

$$|u_h|_{1,\mathcal{M}} = \left\{ \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} (u_{\nu^+} - u_\nu)^2 \right\}^{1/2} = \left\{ \sum_{K \in \mathcal{M}} \|\delta \boldsymbol{U}_K\|^2 \right\}^{1/2}, \tag{4.3}$$

where u_h denotes the discrete function whose nodal value at primary vertex $\boldsymbol{x}_\nu$ is u_ν.

LEMMA 4.1. *For the two matrices $\mathbb{X}_K$ and $\widetilde{\mathbb{F}}_K$ defined respectively in (2.3) and (3.5), we have*

$$\|\mathbb{X}_K \boldsymbol{v}\| \leq \sqrt{n_K}\, h_K \|\boldsymbol{v}\|, \quad \|\widetilde{\mathbb{F}}_K \boldsymbol{v}\| \leq \sqrt{n_K}\, h_K \|\boldsymbol{v}\|, \quad \forall \boldsymbol{v} \in \mathbb{R}^2, \quad \forall K \in \mathcal{M}. \tag{4.4}$$

Proof. For $\boldsymbol{v} \in \mathbb{R}^2$, we have by Cauchy inequality that

$$\|\mathbb{X}_K \boldsymbol{v}\|^2 = \sum_{\sigma \in \mathcal{E}_K} \left((\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_\nu)^T \boldsymbol{v} \right)^2 \leq \sum_{\sigma \in \mathcal{E}_K} \|\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_\nu\|^2 \|\boldsymbol{v}\|^2 \leq n_K h_K^2 \|\boldsymbol{v}\|^2,$$

which verifies the first part of (4.4). The second part can be proved analogously. □

LEMMA 4.2. *For the matrix $\mathbb{C}_K$ defined by (3.3), its column vectors span the null space of $\mathbb{X}_K^T$.*

Proof. By (3.3) and (3.1), it is easy to check that $\mathbb{X}_K^T\mathbb{C}_K = \mathbb{O}$ where $\mathbb{O}$ denotes a generic zero matrix. Since $\operatorname{rank}(\mathbb{X}_K) = 2$, we only need to prove that $\operatorname{rank}(\mathbb{C}_K) \geq n_K - 2$. Using the definition of $\mathbb{C}_K$ once again, we have

$$\operatorname{rank}(\mathbb{I}_K) \leq \operatorname{rank}(\mathbb{C}_K) + \operatorname{rank}\left(\frac{1}{|K|}\mathbb{F}_K\Lambda_K^{-1}\mathbb{X}_K^T\right) \leq \operatorname{rank}(\mathbb{C}_K) + \operatorname{rank}\left(\Lambda_K^{-1}\right),$$

which implies $\operatorname{rank}(\mathbb{C}_K) \geq n_K - 2$ and concludes the proof. ∎

LEMMA 4.3. *Under assumptions* **(H2)** *and* **(H3)**, *we have*

$$\boldsymbol{v}^T\mathbb{A}_K\boldsymbol{v} \geq \varrho_K\|\boldsymbol{v}\|^2, \quad \forall \boldsymbol{v} \in \mathbb{R}^{n_K}, \quad \forall K \in \mathcal{M}. \tag{4.5}$$

where $\mathbb{A}_K$ *is defined by* (3.3) *or* (3.4) *and* ϱ_K *is a positive constant, given by*

$$\varrho_K = \frac{\gamma_K\underline{\lambda}\,\underline{\alpha}^2\,\underline{\kappa}}{n_K(2n_K\underline{\kappa} + \gamma_K\underline{\alpha}\,\underline{\lambda})}. \tag{4.6}$$

Proof. For $\boldsymbol{v} \in \mathbb{R}^{n_K}$, by Lemma 4.2, there exist $\boldsymbol{v}_1 \in \mathbb{R}^2$ and $\boldsymbol{v}_2 \in \mathbb{R}^{n_K}$, such that

$$\boldsymbol{v} = \mathbb{X}_K\boldsymbol{v}_1 + \mathbb{C}_K\boldsymbol{v}_2 \quad \text{and} \quad \|\boldsymbol{v}\|^2 = \|\mathbb{X}_K\boldsymbol{v}_1\|^2 + \|\mathbb{C}_K\boldsymbol{v}_2\|^2. \tag{4.7}$$

Note that (3.1) implies $\widetilde{\mathbb{F}}_K^T\mathbb{X}_K = |K|\mathbb{I}_2$. Then, from (1.4), (3.4) and **(H3)**,

$$\begin{aligned}
\boldsymbol{v}^T\mathbb{A}_K\boldsymbol{v} &\geq \tfrac{\kappa}{|K|}\|\widetilde{\mathbb{F}}_K^T\boldsymbol{v}\|^2 + \gamma_K\|\mathbb{C}_K^T\boldsymbol{v}\|^2 \\
&= \tfrac{\kappa}{|K|}\||K|\boldsymbol{v}_1 + \widetilde{\mathbb{F}}_K^T\mathbb{C}_K\boldsymbol{v}_2\|^2 + \gamma_K\|\mathbb{C}_K^T\mathbb{C}_K\boldsymbol{v}_2\|^2 \\
&\geq \underline{\kappa}|K|(1-\varepsilon)\|\boldsymbol{v}_1\|^2 + \tfrac{\kappa}{|K|}(1-\tfrac{1}{\varepsilon})\|\widetilde{\mathbb{F}}_K^T\mathbb{C}_K\boldsymbol{v}_2\|^2 + \gamma_K\underline{\lambda}\,\|\mathbb{C}_K\boldsymbol{v}_2\|^2,
\end{aligned} \tag{4.8}$$

where $0 < \varepsilon < 1$. From (4.4) and **(H2)**, we have

$$\underline{\kappa}|K|\|\boldsymbol{v}_1\|^2 \geq \underline{\alpha}h_K^2\underline{\kappa}\,\|\boldsymbol{v}_1\|^2 \geq \frac{\underline{\alpha}\,\underline{\kappa}}{n_K}\|\mathbb{X}_K\boldsymbol{v}_1\|^2$$

and

$$\frac{\underline{\kappa}}{|K|}\|\widetilde{\mathbb{F}}_K^T\mathbb{C}_K\boldsymbol{v}_2\|^2 \leq \frac{\underline{\kappa}n_Kh_K^2}{|K|}\|\mathbb{C}_K\boldsymbol{v}_2\|^2 \leq \frac{\underline{\kappa}n_K}{\underline{\alpha}}\|\mathbb{C}_K\boldsymbol{v}_2\|^2.$$

Substituting these estimates into (4.8), we reach

$$\boldsymbol{v}^T\mathbb{A}_K\boldsymbol{v} \geq (1-\varepsilon)\,\frac{\underline{\alpha}\,\underline{\kappa}}{n_K}\|\mathbb{X}_K\boldsymbol{v}_1\|^2 + \left(1 - \frac{1}{\varepsilon}\right)\frac{\underline{\kappa}n_K}{\underline{\alpha}}\|\mathbb{C}_K\boldsymbol{v}_2\|^2 + \gamma_K\underline{\lambda}\,\|\mathbb{C}_K\boldsymbol{v}_2\|^2.$$

Finally, by choosing $\varepsilon = 2n_K\underline{\kappa}/(2n_K\underline{\kappa} + \gamma_K\underline{\alpha}\,\underline{\lambda})$ and by (4.7), we arrive at (4.5) with

$$\varrho_K = \min\left\{\frac{1}{2}\gamma_K\underline{\lambda},\ \frac{\gamma_K\underline{\lambda}\,\underline{\alpha}^2\,\underline{\kappa}}{n_K(2n_K\underline{\kappa} + \gamma_K\underline{\alpha}\,\underline{\lambda})}\right\} = \frac{\gamma_K\underline{\lambda}\,\underline{\alpha}^2\,\underline{\kappa}}{n_K(2n_K\underline{\kappa} + \gamma_K\underline{\alpha}\,\underline{\lambda})}.$$

∎

THEOREM 4.4. *For the scheme defined by* (2.5), (2.1) *and* (3.3), *we have*

(i) *(Symmetry) The resulting linear system is symmetric;*

(ii) *(Coercivity)Under assumptions* **(H2)**, **(H3)** *and* $g_D = 0$,

$$\sum_{\boldsymbol{x}_\nu \in \Omega} \sum_{K \in \mathcal{M}_\nu} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_\nu} \left(\boldsymbol{n}^*_{K,\sigma} \cdot \boldsymbol{n}^*_\nu\right) u_\nu F_{K,\sigma^*} \geq \left(\min_{K \in \mathcal{M}} \varrho_K\right) |u_h|^2_{1,\mathcal{M}}. \tag{4.9}$$

Proof. Since the symmetry of the resulting linear system does not depend on the concrete value of g_D, so we can always proceed the argumentation by assuming $g_D = 0$. Multiplying (2.5) with u_ν, summing over all the dual cells possessing primary unknowns and shifting the summation to primary cells, we have

$$\begin{aligned}\sum_{\boldsymbol{x}_\nu \in \Omega} \sum_{K \in \mathcal{M}_\nu} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_\nu} \left(\boldsymbol{n}^*_{K,\sigma} \cdot \boldsymbol{n}^*_\nu\right) u_\nu F_{K,\sigma^*} &= \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \left(u_{\nu^+} - u_\nu\right) F_{K,\sigma^*} \\ &= \sum_{K \in \mathcal{M}} \boldsymbol{F}_K^T \delta \boldsymbol{U}_K,\end{aligned}$$

where we have used

$$\left(\boldsymbol{n}^*_{K,\sigma} \cdot \boldsymbol{n}^*_\nu\right) u_\nu + \left(\boldsymbol{n}^*_{K,\sigma} \cdot \boldsymbol{n}^*_{\nu^+}\right) u_{\nu^+} = \left(\boldsymbol{n}^*_{K,\sigma} \cdot \boldsymbol{n}^*_{\nu^+}\right)\left(u_{\nu^+} - u_\nu\right) = u_{\nu^+} - u_\nu$$

and $u_\nu = 0$ (resp. $u_{\nu^+} = 0$) if $\boldsymbol{x}_\nu \in \Gamma_D$ (resp. $\boldsymbol{x}_{\nu^+} \in \Gamma_D$). Using (2.1), we obtain

$$\sum_{\boldsymbol{x}_\nu \in \Omega} \sum_{K \in \mathcal{M}_\nu} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_\nu} \left(\boldsymbol{n}^*_{K,\sigma} \cdot \boldsymbol{n}^*_\nu\right) u_\nu F_{K,\sigma^*} = \sum_{K \in \mathcal{M}} (\delta \boldsymbol{U}_K)^T \mathbb{A}_K^T \delta \boldsymbol{U}_K. \tag{4.10}$$

Thus the symmetry of the linear system comes from the symmetry of cell matrix $\mathbb{A}_K$, and by Lemma 4.3, we obtain (4.9). □

4.2. The case of triangular meshes. We have seen from the previous subsection that assumptions **(H2)** and **(H3)** play important roles in the coercivity analysis. For triangular meshes, **(H2)** is a little weaker than the standard regular assumption in finite element method. As for **(H3)**, if the cell center is chosen to be the barycenter, then by direct computation we find that

$$\mathbb{C}_K = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

which implies **(H3)** with $\underline{\lambda} = 1$. Through some straightforward but tedious calculations (see Appendix), we find that the new vertex-centered scheme reduces to the P_1-FVEM. Therefore, the result in this section can serve as an alternative coercivity analysis for P_1-FVEM.

4.3. The case of quadrilateral meshes. For general quadrilateral meshes, we have the following result.

THEOREM 4.5. *Assume that K is a convex or concave quadrilateral. If the center of K is chosen to be the geometric center, then* **(H3)** *holds with $\underline{\lambda} = 1$.*

Proof. Denote the four vertices of K by $\boldsymbol{x}_\nu, \boldsymbol{x}_{\nu^+}, \boldsymbol{x}_{\nu^{++}}$ and $\boldsymbol{x}_{\nu^-}$, which are ordered anticlockwisely. Set

$$\boldsymbol{a} = \boldsymbol{x}_{\nu^+} - \boldsymbol{x}_\nu, \quad \boldsymbol{b} = \boldsymbol{x}_{\nu^-} - \boldsymbol{x}_\nu, \quad \boldsymbol{c} = \boldsymbol{x}_\nu + \boldsymbol{x}_{\nu^{++}} - \boldsymbol{x}_{\nu^-} - \boldsymbol{x}_{\nu^+}. \tag{4.11}$$

We have the following formula

$$\left(2\boldsymbol{b}^T + \boldsymbol{c}^T\right) \mathcal{R}^T \left(2\boldsymbol{a} + \boldsymbol{c}\right) = 2\left(\boldsymbol{x}_{\nu^{++}}^T - \boldsymbol{x}_\nu^T\right) \mathcal{R}^T \left(\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_{\nu^-}\right) = 4|K|. \tag{4.12}$$

By assumption, the cell center of K is defined by

$$\boldsymbol{x}_K = \frac{1}{4}\left(\boldsymbol{x}_\nu + \boldsymbol{x}_{\nu^+} + \boldsymbol{x}_{\nu^{++}} + \boldsymbol{x}_{\nu^-}\right).$$

Then, from (3.4), we have

$$\mathbb{C}_K = \mathbb{I}_K - \frac{1}{4|K|} \begin{pmatrix} 2\boldsymbol{b}^T + \boldsymbol{c}^T \\ -2\boldsymbol{a}^T - \boldsymbol{c}^T \\ -2\boldsymbol{b}^T - \boldsymbol{c}^T \\ 2\boldsymbol{a}^T + \boldsymbol{c}^T \end{pmatrix} \mathcal{R}^T \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b}+\boldsymbol{c} & -\boldsymbol{a}-\boldsymbol{c} & -\boldsymbol{b} \end{pmatrix} \tag{4.13}$$

and further,

$$\mathbb{C}_K = \frac{1}{2} \begin{pmatrix} 2-t & t-1 & 2-t & t-1 \\ s & 1-s & s & 1-s \\ t & 1-t & t & 1-t \\ -s & s+1 & -s & s+1 \end{pmatrix}, \tag{4.14}$$

where

$$t = \frac{1}{2|K|}\left(2\boldsymbol{b}^T + \boldsymbol{c}^T\right) \mathcal{R}^T \boldsymbol{a}, \quad s = \frac{1}{2|K|}\left(2\boldsymbol{a}^T + \boldsymbol{c}^T\right) \mathcal{R}^T \boldsymbol{a}.$$

Let $\boldsymbol{v} = (v_1, v_2, v_3, v_4)^T$ and $w = s^2 + (t-1)^2$. By direct calculations, we have

$$\|\mathbb{C}_K^T \mathbb{C}_K \boldsymbol{v}\|^2 - \|\mathbb{C}_K \boldsymbol{v}\|^2 = \frac{1}{2}(2w^2 + w)\left(v_1 + v_3 - v_2 - v_4\right)^2 \geq 0,$$

which implies **(H3)** with $\underline{\lambda} = 1$.□

Thanks to Theorem 4.5, the coercivity of the present scheme can be established on quadrilateral meshes with arbitrary mesh size only under assumptions **(H1)** and **(H2)**. We remark that the coercivity of Q_1-FVEM requires also **(H1)**, **(H2)** and some other additional assumptions, such as the mesh should be an $h^{1+\gamma}-$parallelogram one with $\gamma > 0$ and the mesh size h should be small enough[17, 28].

5. Numerical Examples. The new vertex-centered linearity-preserving scheme is denoted as VLPS for short. We study several numerical tests to demonstrate that VLPS satisfies the nice features mentioned in the introduction.

Let us define the following mesh-dependent norms for the solution vector $\boldsymbol{U} = \{u_\nu, \boldsymbol{x}_\nu \in \Omega\}$ and a vector $\boldsymbol{F}$ of edge-based fluxes on the dual mesh:

$$|||\boldsymbol{U}|||^2 = \sum_{\boldsymbol{x}_\nu \in \Omega} |K_\nu^*||u_\nu|^2, \qquad |||\boldsymbol{F}|||^2 = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} S_\sigma |F_{K,\sigma^*}|^2,$$

where S_σ is an area associated with σ (for example, $S_\sigma = |K| n_K^{-1}$). We now give the following discrete relative erorrs

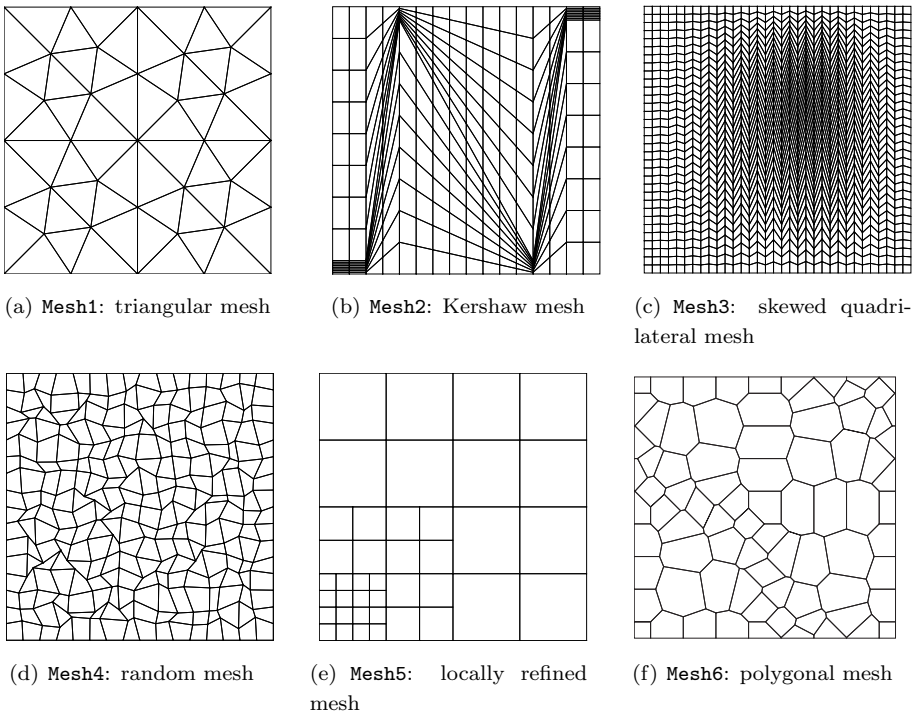
$$E_u = \frac{|||\boldsymbol{U} - \boldsymbol{U}^{ex}|||}{|||\boldsymbol{U}^{ex}|||}, \qquad E_q = \frac{|||\boldsymbol{F} - \boldsymbol{F}^{ex}|||}{|||\boldsymbol{F}^{ex}|||},$$

where $\boldsymbol{U}^{ex}$ and $\boldsymbol{F}^{ex}$ denote the exact solutions and fluxes, and $\boldsymbol{F}^{ex}$ can be evaluated by the mid-point rule.

The rate of convergence R_α ($\alpha = u, q$) is obtained by a least squares fit on the ones computed on each two successive meshes by the following formula

$$\frac{\log[E_\alpha(h_2)/E_\alpha(h_1)]}{\log(h_2/h_1)},$$

where h_1, h_2 denote the mesh sizes of the two successive meshes, and $E_\alpha(h_1), E_\alpha(h_2)$ the corresponding discrete errors.



(a) `Mesh1`: triangular mesh (b) `Mesh2`: Kershaw mesh (c) `Mesh3`: skewed quadrilateral mesh

(d) `Mesh4`: random mesh (e) `Mesh5`: locally refined mesh (f) `Mesh6`: polygonal mesh

FIG. 5.1. *Six mesh types used in the numerical tests.*

5.1. Continuous solutions. In this section, we study the convergence of VLPS for problems with a smooth solution on the domain $\Omega = (0,1)^2$, and use a sequence of six mesh types shown in Figure 5.1.

To begin with, we consider the constant diffusion tensor and the exact solution $u = 1 - 2x - 3y$. On the coarsest meshes of six mesh types, relative errors of the solution and flux are given in Table 5.1. We find that the new scheme VLPS is always

linearity-preserving, with the solution and flux errors in the L^2-norm of the order of machine precision.

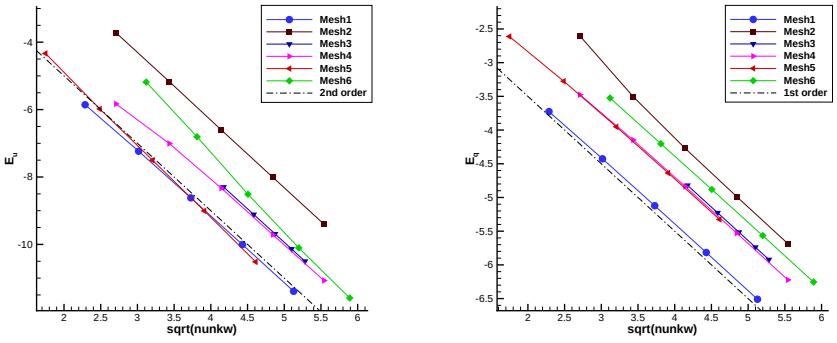
TABLE 5.1
Relative errors of the solution and flux on the coarsest meshes.

Error	Mesh1	Mesh2	Mesh3	Mesh4	Mesh5	Mesh6
E_u	3.73E-16	1.49E-15	2.17E-15	2.48E-16	2.77E-16	1.32E-15
E_q	1.31E-15	6.40E-15	2.28E-14	1.11E-15	7.50E-16	3.50E-15

Secondly, we consider a test on the fifth conference on discretization schemes for anisotropic diffusion problems on general grids [14] (FVCA V for short). We consider the problem (1.1)–(1.2) and $\Omega = [0,1]^2$. A homogeneous anisotropic tensor and the exact solution are given below:

$$\Lambda = \begin{pmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{pmatrix}, \quad u(x,y) = \sin((1-x)(1-y)) + (1-x)^3(1-y)^2.$$

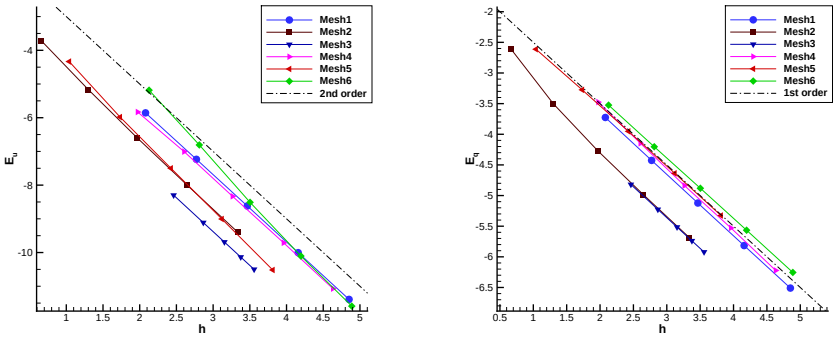
The convergence rates for the discrete L^2-norm of solution errors and flux errors are graphically depicted in Figure 5.2–Figure 5.3 as log-log plots of the errors versus the square root of the number of unknowns `nunkw`, and the inverse of mesh size h on six mesh types. The actual convergence orders are reflected by the slopes of the experimental error curves. Three figures show a second order convergence rate with respect to the solution errors and first order convergence rate with respect to the flux errors on six mesh types. Note that when computing the errors with respect to `nunkw`, convergence rates remain the same as expected but the relative position varies.



(a) solution (b) flux

FIG. 5.2. *L_2 errors of the solution and its flux versus* `nunkw` *on* `Mesh1`–`Mesh6`.

In addition, we also give the numerical results on the uniform trapezoidal meshes `Mesh7` which are composed of elements of right-angled trapezoids, and the ratios of lengths of two bases of all trapezoidal elements `ratio` are fixed (see Figure 5.4 with `ratio`=1:19). The meshes do not satisfy the nearly parallelogram condition when the mesh size decreases, and the standard Q_1-conforming quadrilateral finite volume



(a) solution (b) flux

FIG. 5.3. *L_2 errors of the solution and its flux versus mesh size h on* `Mesh1`–`Mesh6`.

method or finite element method performs not optimal [12]. Table 5.2–Table 5.3 show the optimal convergence of the scheme VLPS on the uniform trapezoidal meshes `Mesh7` with `ratio`=1:19 and 1:199, respectively.

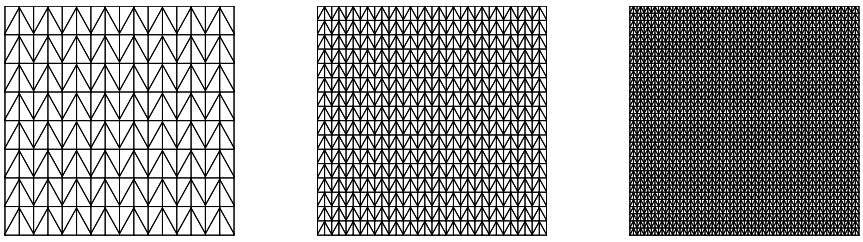


FIG. 5.4. *Three samples of family of the uniform trapezoidal meshes* `Mesh7` *with* `ratio`*=1:19.*

TABLE 5.2
Relative errors on the uniform trapezoidal mesh `Mesh7` *with* `ratio`*=1:19.*

`nunkw`	E_u	Rate	E_q	Rate
225	5.58E-4	–	4.30E-2	–
961	1.38E-4	2.01	2.15E-2	1.00
3969	3.46E-5	2.00	1.07E-2	1.00
16129	8.65E-6	2.00	5.36E-3	1.00
65025	2.16E-6	2.00	2.68E-3	1.00

5.2. Discontinuous solutions. We deal with the problem (1.1)–(1.2) on $\Omega = [0,1]^2$, and choose an heterogeneous medium such that

$$\Lambda(x,y) = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & x \le 0.5, \\ \begin{pmatrix} 10 & 3 \\ 3 & 1 \end{pmatrix}, & x > 0.5. \end{cases}$$

TABLE 5.3
Relative errors on the uniform trapezoidal mesh `Mesh7` *with* `ratio`*=1:199.*

`nunkw`	E_u	Rate	E_q	Rate
225	5.49E-4	–	5.93E-2	–
961	1.33E-4	2.04	2.95E-2	1.01
3969	3.31E-5	2.01	1.47E-2	1.00
16129	8.28E-6	2.00	7.33E-3	1.00
65025	2.07E-6	2.00	3.66E-3	1.00

We choose the exact solution

$$u(x,y) = \begin{cases} 1 - 2y^2 + 4xy + 6x + 2y, & x \le 0.5, \\ -2y^2 + 1.6xy - 0.6x + 3.2y + 4.3, & x > 0.5. \end{cases}$$

This test is inspired by a steady numerical test in [19]. Convergence investigation is conducted on two mesh types `Mesh1` and `Mesh2`. Table 5.4 and Table 5.5 show the numbers of unknowns `nunkw`, the relative errors of the solution and flux, and the convergence rates.

TABLE 5.4
Behaviors on the triangular mesh `Mesh1`.

`nunkw`	E_u	Rate	E_q	Rate
97	4.20E-4	–	3.84E-2	–
417	1.09E-4	1.95	1.91E-2	1.01
1729	2.76E-5	1.98	9.50E-3	1.01
7041	6.95E-6	1.99	4.74E-3	1.00
28417	1.74E-6	2.00	2.37E-3	1.00

TABLE 5.5
Behaviors on Kershaw mesh `Mesh2`.

`nunkw`	E_u	Rate	E_q	Rate
225	4.65E-3	–	7.86E-2	–
961	1.13E-3	2.21	3.21E-2	1.40
3969	2.78E-4	2.10	1.61E-2	1.04
16129	6.89E-5	2.05	8.23E-3	0.98
65025	1.72E-5	2.02	4.17E-3	0.99

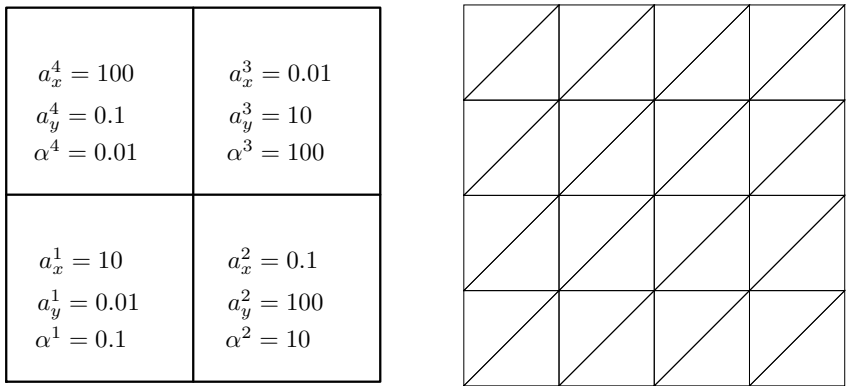
In the next group of tests, the domain Ω is split into four subdomains $\Omega = \cup_{i=1}^{4}\Omega_i$ (see Figure 5.5(a)), and the homogeneous Dirichlet boundary condition is imposed in this test. The diffusion tensor and exact solution are given by

$$\Lambda(x,y) = \begin{pmatrix} a_x^i & 0 \\ 0 & a_y^i \end{pmatrix}, \quad u(x,y) = \alpha_i \sin(2\pi x)\sin(2\pi y), \text{ for } (x,y) \in \Omega_i,$$

where the value of coefficients a_x^i, a_y^i and α^i can be found in Figure 5.5(a), and the diffusion tensor Λ is discontinuous across the lines $x = 0.5$ and $y = 0.5$.

Table 5.6–Table 5.9 present the numerical convergence of the scheme on the uniform triangular mesh (see Figure 5.5(b)), uniform square mesh, locally refined quadrilateral mesh `Mesh5` and uniform trapezoidal mesh `Mesh7` (see Figure 5.4). We have the following results:

- On the uniform triangular mesh, the convergence rate of solution errors is less than $h^{1.7}$.
- On the uniform square mesh, `Mesh5` and `Mesh7`, the optimal convergence rates for the solution and flux errors are observed.



(a) (b)

FIG. 5.5. *Coefficients in the definition of diffusion tensor and exact solution (left) and the uniform triangular mesh (right).*

TABLE 5.6
Behaviors on the uniform triangular mesh.

`nunkw`	E_u	Rate	E_q	Rate
9	1.03E-1	–	5.11E-1	–
49	3.84E-2	1.43	2.63E-1	0.96
225	1.40E-2	1.45	1.32E-1	0.99
961	4.88E-3	1.52	6.66E-2	0.99
3969	1.55E-3	1.65	3.34E-2	0.99

5.3. Heterogeneous rotating anisotropy. Problem (1.1)–(1.2) is defined in $\Omega = [0, 1]^2$ with a rotating anisotropic diffusion tensor:

$$\Lambda(x, y) = \frac{1}{x^2 + y^2} \begin{pmatrix} \alpha x^2 + y^2 & (\alpha - 1)xy \\ (\alpha - 1)xy & x^2 + \alpha y^2 \end{pmatrix},$$

TABLE 5.7
Behaviors on the uniform square mesh.

nunkw	E_u	Rate	E_q	Rate
225	2.69E-2	–	7.08E-2	–
961	6.69E-3	2.01	3.54E-2	1.00
3969	1.68E-3	2.00	1.76E-2	1.00
16129	4.21E-4	2.00	8.84E-3	1.00
65025	1.05E-4	2.00	4.42E-3	1.00

TABLE 5.8
Behaviors on the locally refined quadrilateral mesh `Mesh5`.

nunkw	E_u	Rate	E_q	Rate
33	3.10E-1	–	3.29E-1	–
145	7.52E-2	2.04	1.50E-1	1.14
609	1.87E-2	2.01	7.24E-2	1.05
2497	4.71E-3	1.99	3.58E-2	1.02
10113	1.19E-3	1.99	1.78E-2	1.01

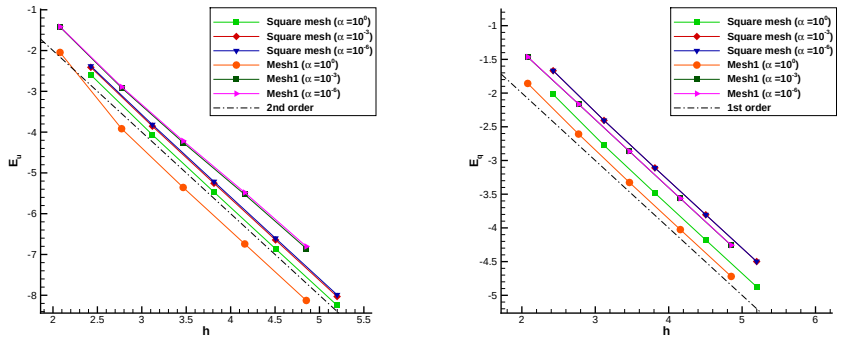
TABLE 5.9
Behaviors on the uniform trapezoidal mesh `Mesh7`.

nunkw	E_u	Rate	E_q	Rate
225	4.21E-2	–	1.81E-1	–
961	1.06E-2	1.99	9.27E-2	0.97
3969	2.68E-3	1.98	4.67E-2	0.99
16129	6.80E-4	1.98	2.34E-2	1.00
65025	1.72E-4	1.99	1.17E-2	1.00

where α characterizes the level of anisotropy. We consider the smooth exact solution $u(x, y) = \sin(\pi x)\sin(\pi y)$. This test is inspired from Le Potier's work[22], and we use families of the uniform square mesh and the triangular mesh `Mesh1` with 5 mesh levels in this test.

For various anisotropy $\alpha = 1, 10^{-3}$ and 10^{-6}, plots of the rate of convergence on both two meshes are depicted in Figure 5.6. We observe that the new scheme VLPS delivers the optimal rate of convergence in the L^2-norm of the solution errors and flux errors.

6. Conclusion. In this article, we have presented a symmetric vertex-centered linearity-preserving finite volume scheme (VLPS) for two dimensional diffusion problems. The discretization takes into account the general polygonal meshes and arbitrary heterogeneous anisotropic diffusion tensors. VLPS is proved to be symmetric and coercive under general assumptions. Many numerical tests using meshes with



(a) solution (b) flux

FIG. 5.6. *L_2 errors of the solution and flux versus mesh size h on the uniform square mesh and triangular mesh* `Mesh1` *with $\alpha = 1, 10^{-3}$ and 10^{-6}.*

right-angled trapezoidal elements, severely distorted elements and nonmatching elements and arbitrary (continuous or discontinuous) anisotropic diffusion tensors show the robustness of the scheme.

Appendix. We show that the new vertex-centered scheme reduces to the P_1-FVEM on triangular meshes. Still take the triangle cell K in Figure 2.2 for exposition. By the definition of $\delta \boldsymbol{U}_K$, we get

$$\delta \boldsymbol{U}_K = \mathbb{T}_K \boldsymbol{U}_K, \tag{6.1}$$

where $\boldsymbol{U}_K = (u_\nu, u_{\nu^+}, u_{\nu^-})^T$ and

$$\mathbb{T}_K = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

Now, on the one hand, by (2.3) and through straightforward calculations, we have

$$\mathbb{X}_K \widetilde{\mathbb{F}}_K^T \mathbb{T}_K = |K| \mathbb{T}_K, \quad \mathbb{C}_K^T \mathbb{T}_K = \mathbb{O}.$$

It follows from (2.1), (3.4) and (6.1) that

$$F_{K,\sigma^*} = (1,0,0)\, \mathbb{A}_K \mathbb{T}_K \boldsymbol{U}_K = \frac{1}{|K|} (1,0,0)\, \widetilde{\mathbb{F}}_K \Lambda_K \widetilde{\mathbb{F}}_K^T \mathbb{T}_K \boldsymbol{U}_K$$

and further,

$$F_{K,\sigma^*} = \frac{1}{|K|} (\boldsymbol{x}_\sigma - \boldsymbol{x}_K)^T \mathcal{R}^T \Lambda_K \left(\mathcal{R}(\boldsymbol{x}_{\sigma^-} - \boldsymbol{x}_\sigma), \mathcal{R}(\boldsymbol{x}_\sigma - \boldsymbol{x}_{\sigma+}), \mathcal{R}(\boldsymbol{x}_{\sigma^+} - \boldsymbol{x}_{\sigma^-}) \right) \boldsymbol{U}_K,$$

where $\boldsymbol{x}_{\sigma^+} = (\boldsymbol{x}_{\nu^+} + \boldsymbol{x}_{\nu^-})/2$ and $\boldsymbol{x}_{\sigma^-} = (\boldsymbol{x}_{\nu^-} + \boldsymbol{x}_\nu)/2$. Then, we arrive at (2.2) with

$$\begin{aligned} c &= \tfrac{1}{|K|} \left(\boldsymbol{x}_\sigma - \boldsymbol{x}_K \right)^T \mathcal{R}^T \Lambda_K \mathcal{R} \left(\boldsymbol{x}_{\sigma^-} - \boldsymbol{x}_\sigma \right), \\ c_+ &= \tfrac{1}{|K|} \left(\boldsymbol{x}_\sigma - \boldsymbol{x}_K \right)^T \mathcal{R}^T \Lambda_K \mathcal{R} \left(\boldsymbol{x}_\sigma - \boldsymbol{x}_{\sigma^+} \right), \\ c_- &= \tfrac{1}{|K|} \left(\boldsymbol{x}_\sigma - \boldsymbol{x}_K \right)^T \mathcal{R}^T \Lambda_K \mathcal{R} \left(\boldsymbol{x}_{\sigma^+} - \boldsymbol{x}_{\sigma^-} \right). \end{aligned} \tag{6.2}$$

On the other hand, for P_1-FVEM, the discrete flux across σ_K^* is given by

$$F_{K,\sigma^*} = \int_{\sigma_K^*} (-\Lambda_K \nabla \tilde{u}_h) \cdot \boldsymbol{n}_{K,\sigma}^* ds$$

with

$$\tilde{u}_h = u_\nu \phi_\nu + u_{\nu^+} \phi_{\nu^+} + u_{\nu^-} \phi_{\nu^-}$$

where ϕ_ν, ϕ_{ν^+} and ϕ_{ν^-} denote the P_1 basis functions at $\boldsymbol{x}_\nu, \boldsymbol{x}_{\nu^+}$ and $\boldsymbol{x}_{\nu^-}$, respectively. By straightforward calculations, we have

$$\nabla\phi_\nu = \frac{-1}{2|K|}\mathcal{R}\left(\boldsymbol{x}_{\nu^-} - \boldsymbol{x}_{\nu^+}\right), \ \nabla\phi_\nu = \frac{-1}{2|K|}\mathcal{R}\left(\boldsymbol{x}_{\nu} - \boldsymbol{x}_{\nu^-}\right), \ \nabla\phi_\nu = \frac{-1}{2|K|}\mathcal{R}\left(\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_{\nu}\right).$$

Noting once again that $|\sigma_K^*|\boldsymbol{n}_{K,\sigma^*} = \mathcal{R}(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)$, we reach (2.2) with

$$\begin{aligned} c &= \tfrac{1}{2|K|}\left(\boldsymbol{x}_\sigma - \boldsymbol{x}_K\right)^T \mathcal{R}^T \Lambda_K \mathcal{R}\left(\boldsymbol{x}_{\nu^-} - \boldsymbol{x}_{\nu^+}\right), \\ c_+ &= \tfrac{1}{2|K|}\left(\boldsymbol{x}_\sigma - \boldsymbol{x}_K\right)^T \mathcal{R}^T \Lambda_K \mathcal{R}\left(\boldsymbol{x}_{\nu} - \boldsymbol{x}_{\nu^-}\right), \\ c_- &= \tfrac{1}{2|K|}\left(\boldsymbol{x}_\sigma - \boldsymbol{x}_K\right)^T \mathcal{R}^T \Lambda_K \mathcal{R}\left(\boldsymbol{x}_{\nu^+} - \boldsymbol{x}_{\nu}\right). \end{aligned} \tag{6.3}$$

By recalling the definitions of $\boldsymbol{x}_\sigma, \boldsymbol{x}_{\sigma^+}$ and $\boldsymbol{x}_{\sigma^-}$, we find that (6.2) is equivalent to (6.3). Hence, on triangular grids, the present vertex-centered scheme is identical to P_1- FVEM in the sense of yielding the same discrete flux expression.

REFERENCES

[1] R. E. Bank and D. J. Rose, *Some error estimates for the box method*, SIAM J. Numer. Anal., 24 (1987), pp. 777–787.

[2] F. Brezzi, A. Buffa, and K. Lipnikov, *Mimetic finite differences for elliptic problems*, E-SAIM: M2AN, 43 (2009), pp. 277–295.

[3] F. Brezzi, K. Lipnikov, and V. Simoncini, *A family of mimetic finite difference methods on polygonal and polyhedral meshes*, Math. Mod. Methods Appl. Sci., 15 (2005), pp. 1533–1551.

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

[5] Z. Chen, R. Li, and A. Zhou, *A note on the optimal L2-estimate of the finite volume element method*, Adv. Comput. Math., 16 (2002), pp. 291–303.

[6] L. Beirao da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo, *Basic principles of virtual element methods*, Math. Models Methods Appl. Sci., 23 (2013), pp. 119–214.

[7] L. Beirao da Veiga, K. Lipnikov, and G. Manzini, *Arbitrary-order nodal mimetic discretizations of elliptic problems on polygonal meshes*, SIAM J. Numer. Anal., 49 (2011), pp. 1737–1760.

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

[9] R. Eymard, G. Henry and. R. Herbin, F. Hubert, R.Klofkorn, and G.Manzini, *3d benchmark on discretization schemes for anisotropic diffusion problems on general grids*, in Finite Volumes for Complex Applications VI - Problems and Perspectives, Jaroslav Fort, Jiri Furst, Jan Halama, Raphaele Herbin, and Florence Hubert, eds., Springer, 2011, pp. 893–929.

[10] 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), pp. 1009–1043.

[11] R. Eymard, R. Herbin, and C. Guichard, *Small-stencil 3D schemes for diffusive flows in porous media*, ESAIM: M2AN, 46 (2012), pp. 265–290.

[12] Xinlong Feng, Rongfei Li, Yinnian He, and Demin Liu, *P1-nonconforming quadrilateral finite volume methods for the semilinear elliptic equations*, J. Sci. Comput., 52 (2012), pp. 519–545.

[13] V. Girault, *Theory of a finite difference method on irregular networks*, SIAM J. Numer. Anal., 11 (1974), pp. 260–282.

[14] R. Herbin and F. Hubert, *Benchmark on discretization schemes for anisotropic diffusion problems on general grids*, in Finite Volumes for Complex Applications V-Problems and Perspectives, Robert Eymard and Jean-Marc Herard, eds., Wiley press, 2008, pp. 659–692.

[15] J. Huang and S. Xi, *On the finite volume element method for general self-adjoint elliptic problem*, SIAM J. Numer. Anal., 35 (1998), pp. 1762–1774.

[16] R. Li, Z. Chen, and W. Wu, *Generalized difference methods for differential equations*, Marcel Dekker, New York, 2000.

[17] Y. Li and R. Li, *Generalized difference methods on arbitrary quadrilateral networks*, J. Comput. Math., 17 (1999), pp. 653–672.

[18] K. Lipnikov and G. Manzini, *A high-order mimetic method on unstructured polyhedral meshes for the diffusion equation*, J. Comput. Phys., 272 (2014), pp. 360–385.

[19] K. Lipnikov, D. Svyatskiy, and Y. Vassilevski, *Interpolation-free monotone finite volume method for diffusion equations on polygonal meshes*, J. Comput. Phys., 228 (2009), pp. 703–716.

[20] Junliang Lv and Yonghai Li, *L2 error estimate of the finite volume element methods on quadrilateral meshes*, Adv. Comput. Math., 33 (2010), pp. 129–148.

[21] ———, *Optimal biquadratic finite volume element methods on quadrilateral meshes*, SIAM J. Numer. Anal., 50 (2012), pp. 2379–2399.

[22] Christophe Le Potier, *schema volumes finis monotone pour des operateurs de diffusion fortement anisotropes sur des maillages de triangle non structures*, C. R. Math. Acad. Sci. Paris, 341 (2005), pp. 787–792.

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

[24] W. Sun, J. Wu, and X. Zhang, *A family of linearity-preserving schemes for anisotropic diffusion problems on arbitrary polyhedral grids*, Comput. Methods Appl. Mech. Engrg., 267 (2013), pp. 418–433.

[25] Jiming Wu, Zihuan Dai, Zhiming Gao, and Guangwei Yuan, *Linearity preserving nine-point schemes for diffusion equation on distorted quadrilateral meshes*, J. Comput. Phys., 229 (2010), pp. 3382–3401.

[26] L. Yin, J. Wu, and Z. Gao, *The cell functional minimization scheme for the anisotropic diffusion problems on arbitrary polygonal grids*, ESAIM: M2AN, published online (2014).

[27] Li Yin, Jiming Wu, and Yanzhong Yao, *A cell functional minimization scheme for parabolic problem*, J. Comput. Phys., 229 (2010), pp. 8935–8951.

[28] Z. Zhang and Q. Zou, *Vertex-centered finite volume schemes of any order over quadrilateral meshes for elliptic boundary value problems*, Numer. Math., published online (2014).