



**HAL**  
open science

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

Jiming Wu, Zhiming Gao, Zihuan Dai

► **To cite this version:**

Jiming Wu, Zhiming Gao, Zihuan Dai. A VERTEX-CENTERED LINEARITY-PRESERVING DISCRETIZATION OF DIFFUSION PROBLEMS ON POLYGONAL MESHES. 2015. hal-01098507

**HAL Id: hal-01098507**

**<https://inria.hal.science/hal-01098507>**

Preprint submitted on 8 Jan 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1                    **A VERTEX-CENTERED LINEARITY-PRESERVING**  
2                    **DISCRETIZATION OF DIFFUSION PROBLEMS ON POLYGONAL**  
3                    **MESHES\***

4                    JIMING WU\*, ZHIMING GAO\*<sup>†</sup>, AND ZIHUAN DAI\*

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

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

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

16                    **1. Introduction.** We consider the anisotropic steady-state diffusion problem

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

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

$$(1.3) \quad -\Lambda \nabla u \cdot \mathbf{n} = g_N \quad \text{on } \Gamma_N,$$

17 where  $\Omega$  is a bounded connected polygonal domain in  $\mathbb{R}^2$ ,  $\mathbf{n}$  denotes the exterior unit  
18 normal vector along the domain boundary  $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ ,  $\Lambda$  is a  $2 \times 2$  symmetric dif-  
19 fusion tensor, uniformly bounded above and below, *i.e.*, there exist positive constants  
20  $\bar{\kappa}$  and  $\underline{\kappa}$ , such that

$$(1.4) \quad \underline{\kappa} \|\mathbf{v}\|^2 \leq \mathbf{v}^T \Lambda \mathbf{v} \leq \bar{\kappa} \|\mathbf{v}\|^2, \quad \forall \mathbf{v} \in \mathbb{R}^2,$$

21  $f$ ,  $g_D$  and  $g_N$  denote the source term, the Dirichlet and flux boundary data, respec-  
22 tively. Anisotropic diffusion problem of this type arises in a wide range of scientific  
23 fields such as oil reservoir simulations, plasma physics, semiconductor modeling and  
24 so on. In accurate simulation of diffusion processes in these applications, finite vol-  
25 ume (FV) method is among the most commonly used ones due to its simplicity, local  
26 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

<sup>†</sup>Corresponding author. E-mail: gao@iapcm.ac.cn

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

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

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

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

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

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

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

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

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

101 **2.1. The primary and dual meshes.** The construction of the primary and  
 102 dual meshes is almost the same as that in FVEM. Suppose that  $\Omega$  is partitioned into  
 103 a number of non-overlapped polygonal cells that form the so-called *primary mesh*, see  
 104 the mesh with solid line segments in Figure 2.1. The vertices of  $\Omega$  and the possible joint  
 105 points of  $\Gamma_D$  and  $\Gamma_N$  must be included in the set of primary vertices. For a primary  
 106 cell, its cell center is defined at any point in the cell. Each primary cell is further  
 107 partitioned into several quadrilateral subcells by connecting the cell center with the  
 108 edge midpoints, see the dashed line segments in Figure 2.1. All subcells sharing a

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

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

112 By this assumption, the dual mesh makes sense.

113 **2.2. The primary unknowns.** At each vertex of the primary mesh in  $\Omega \cup \Gamma_N$ ,  
 114 we define a single primary unknown, see the solid points in Figure 2.1. A finite volume  
 115 equation will be constructed associated with each primary unknown. Note that the  
 116 primary vertices on  $\bar{\Gamma}_D$  have no primary unknowns. Contrary to the cell-centered  
 117 linearity-preserving FV schemes studied before, here we do not need to introduce any  
 118 auxiliary unknowns so that there is no interpolation algorithm.

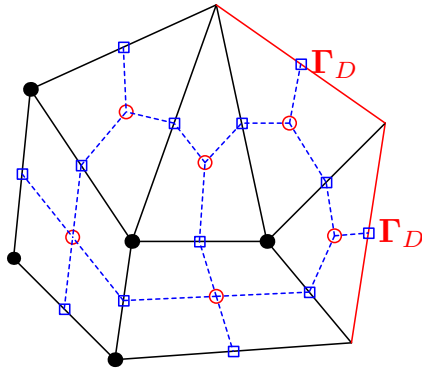


FIG. 2.1. The primary mesh (solid line), dual mesh (dashed line) and the boundary  $\Gamma_D$  (red line).

119 **2.3. The flux discretization.** From now on, all the derivations are conducted  
 120 under the following assumptions:

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

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

131 As shown in Figure 2.2, we introduce some notations.

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

- 136 •  $\sigma$ , a generic edge of  $K$  with measure denoted as  $|\sigma|$ , also the local numbering
- 137 of edges in  $\mathcal{E}_K$ , depending on the context.
- 138 •  $\mathbf{x}_\nu$  and  $\mathbf{x}_{\nu+}$ , two generic vertices of the primary mesh, also the two endpoints
- 139 of  $\sigma$ .
- 140 •  $\mathbf{x}_\sigma$ , the midpoint of  $\sigma$ .
- 141 •  $K_\nu^*$ , a generic dual cell associated with  $\mathbf{x}_\nu$  whose outward unit normal along
- 142 the cell boundary is denoted as  $\mathbf{n}_\nu^*$ .
- 143 •  $\sigma_K^*$ , a generic dual edge connecting  $\mathbf{x}_K$  and  $\mathbf{x}_\sigma$ .
- 144 •  $\mathbf{n}_{K,\sigma}^*$ , a unit vector normal to  $\sigma_K^*$  whose direction is fixed once and for all.
- 145 •  $u_\nu, u_{\nu+}$ , the primary unknowns defined at  $\mathbf{x}_\nu$  and  $\mathbf{x}_{\nu+}$ , respectively.
- 146 •  $\Lambda_K$ , the constant restriction of  $\Lambda$  on  $K$ .
- 147 •  $F_{K,\sigma^*}$ , the discrete counterpart or approximation of  $\int_{\sigma_K^*} (-\Lambda_K \nabla u) \cdot \mathbf{n}_{K,\sigma}^* ds$ .

148 In addition, for  $\sigma \in \mathcal{E}_K$ , we assume that its endpoint  $\mathbf{x}_\nu$  is always pointed anticlock-

149 wisesly to the other one  $\mathbf{x}_{\nu+}$ . We also assume that all  $n_K$  unit vectors  $\mathbf{n}_{K,\sigma}^*$  ( $\sigma \in \mathcal{E}_K$ )

150 inside  $K$  are ordered clockwise (see Figure 2.2). As a result,  $\mathbf{n}_{K,\sigma}^* \cdot \mathbf{n}_\nu^*$  may be either 1

151 or  $-1$ . Throughout, the hollow letters  $\mathbb{A}, \mathbb{F}, \mathbb{U}, \dots$  will be used to denote rectangular

152 matrices with a number of columns greater than one while the bold ones  $\mathbf{F}, \mathbf{x}, \mathbf{n}, \dots$

153 will represent vectors or matrices with only one column.

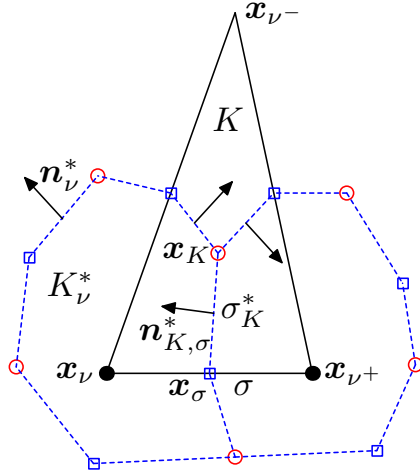


FIG. 2.2. Notations for the flux discretization.

154 Now we are ready to describe the construction of the flux approximation  $F_{K,\sigma^*}$ .

155 First, we put all  $F_{K,\sigma^*}, \sigma \in \mathcal{E}_K$  in the same group and manipulate them together,

156 instead of treating them one by one. Note that the flux approximations for the

157 same dual cell may belong to different groups. Secondly, we seek the following local

158 algebraic relation related to primary cell  $K$ ,

$$(2.1) \quad \mathbf{F}_K = \mathbb{A}_K \delta \mathbf{U}_K,$$

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

$$(2.2) \quad F_{K,\sigma^*} = cu_\nu + c_+u_{\nu^+} + c_-u_{\nu^-}.$$

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

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

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

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

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

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

$$(2.4) \quad \mathbb{F}_K = \mathbb{A}_K \mathbb{X}_K.$$

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

177 **2.4. The final vertex-centered scheme.** For an vertex  $\mathbf{x}_\nu \in \Omega \cup \Gamma_N$ , let  $\mathcal{M}_\nu$   
 178 (resp.  $\mathcal{E}_\nu$ ) be the set of primary cells (resp. edges) sharing  $\mathbf{x}_\nu$ . The finite volume  
 179 equation associated with primary unknown  $u_\nu$  is constructed as follows

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

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

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

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

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

$$(3.1) \quad \mathbb{F}_K^T \mathbb{X}_K = |K| \mathbb{I}_K.$$

186

187 *Proof.* We start the argumentation by introducing the following formula (see, e.g.,  
 188 (2.17) in [10] or equivalently (3.7) in [27]):

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

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

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

Then, it follows from (3.2) that

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

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

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

189 which leads to (3.1) and concludes the proof.  $\square$

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

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

191 or equivalently,

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

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

$$(3.5) \quad \tilde{\mathbb{F}}_K = (-|\sigma_K^*| \mathbf{n}_{K,\sigma}^*, \sigma \in \mathcal{E}_K)^T,$$

193 and  $\gamma_K$  is a positive parameter. In practical computation, we choose

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



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

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

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

199 which implies  $\mathbf{v} = \mathbf{0}$  and completes the proof.  $\square$

#### 200 4. Symmetry and coercivity.

201 **4.1. General results.** We first introduce two assumptions.

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

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

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

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

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

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

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

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

208 LEMMA 4.1. *For the two matrices  $\mathbb{X}_K$  and  $\tilde{\mathbb{F}}_K$  defined respectively in (2.3) and*  
 209 *(3.5), we have*

$$(4.4) \quad \|\mathbb{X}_K \mathbf{v}\| \leq \sqrt{n_K} h_K \|\mathbf{v}\|, \quad \|\tilde{\mathbb{F}}_K \mathbf{v}\| \leq \sqrt{n_K} h_K \|\mathbf{v}\|, \quad \forall \mathbf{v} \in \mathbb{R}^2, \quad \forall K \in \mathcal{M}.$$

210

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

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

211 which verifies the first part of (4.4). The second part can be proved analogously.  $\square$

212 LEMMA 4.2. *For the matrix  $\mathbb{C}_K$  defined by (3.3), its column vectors span the null*  
 213 *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  $\text{rank}(\mathbb{X}_K) = 2$ , we only need to prove that  $\text{rank}(\mathbb{C}_K) \geq n_K - 2$ . Using the definition of  $\mathbb{C}_K$  once again, we have

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

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

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

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

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

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

217

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

219

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

220 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),

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

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

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

and

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

Substituting these estimates into (4.8), we reach

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

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

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

221  $\square$

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

- 223 (i) (Symmetry) The resulting linear system is symmetric;  
 224 (ii) (Coercivity) Under assumptions **(H2)**, **(H3)** and  $g_D = 0$ ,

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

*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  $u_\nu$ , summing over all the dual cells possessing primary unknowns and shifting the summation to primary cells, we have

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

where we have used

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

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

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

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

**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

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

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

232 **4.3. The case of quadrilateral meshes.** For general quadrilateral meshes, we  
 233 have the following result.

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

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

$$(4.11) \quad \mathbf{a} = \mathbf{x}_{\nu^+} - \mathbf{x}_\nu, \quad \mathbf{b} = \mathbf{x}_{\nu^-} - \mathbf{x}_\nu, \quad \mathbf{c} = \mathbf{x}_\nu + \mathbf{x}_{\nu^{++}} - \mathbf{x}_{\nu^-} - \mathbf{x}_{\nu^+}.$$

238 We have the following formula

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

By assumption, the cell center of  $K$  is defined by

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

239 Then, from (3.4), we have

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

240 and further,

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

where

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

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

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

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

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

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

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

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

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

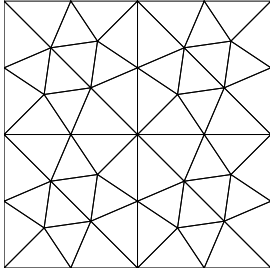
$$E_u = \frac{\|U - U^{ex}\|}{\|U^{ex}\|}, \quad E_q = \frac{\|F - F^{ex}\|}{\|F^{ex}\|},$$

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

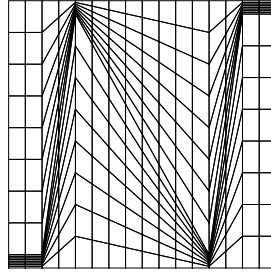
The rate of convergence  $R_\alpha$  ( $\alpha = u, q$ ) is obtained by a least squares fit on the ones computed on each two successive meshes by the following formula

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

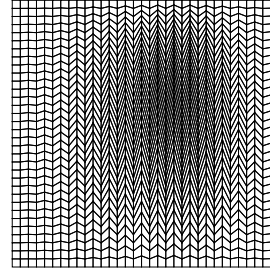
where  $h_1, h_2$  denote the mesh sizes of the two successive meshes, and  $E_\alpha(h_1), E_\alpha(h_2)$  the corresponding discrete errors.



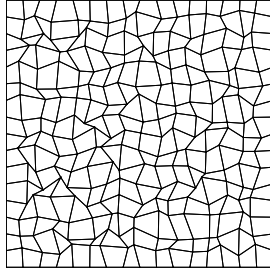
(a) Mesh1: triangular mesh



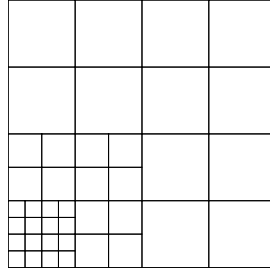
(b) Mesh2: Kershaw mesh



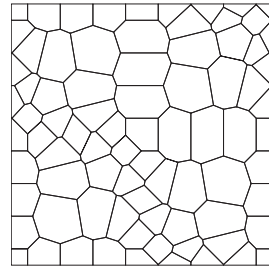
(c) Mesh3: skewed quadrilateral mesh



(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

260 linearity-preserving, with the solution and flux errors in the  $L^2$ -norm of the order of  
 261 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:

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

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

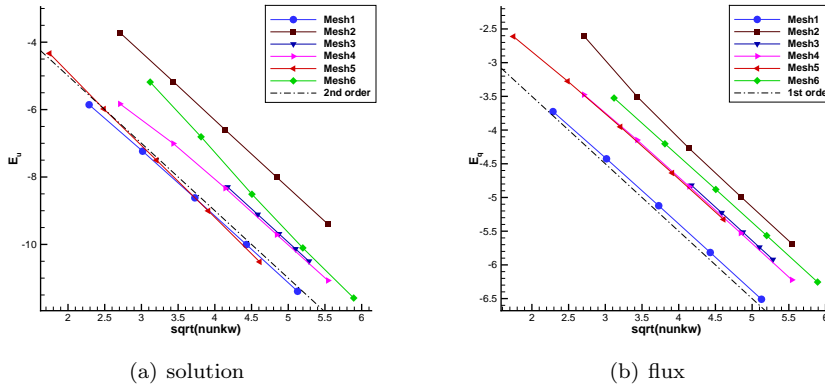


FIG. 5.2.  $L_2$  errors of the solution and its flux versus `nunkw` on Mesh1–Mesh6.

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

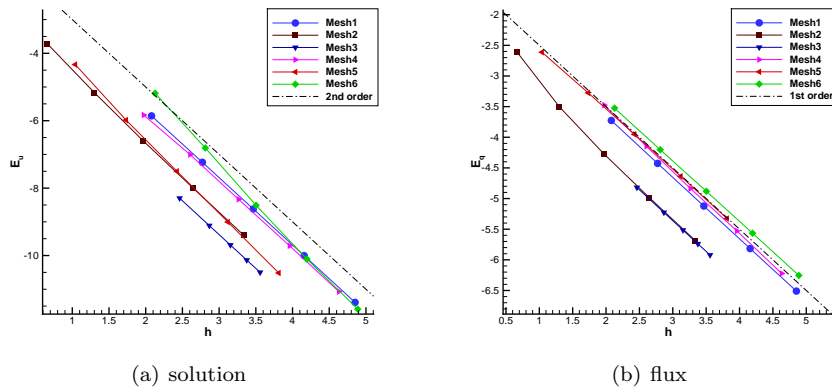


FIG. 5.3.  $L_2$  errors of the solution and its flux versus mesh size  $h$  on Mesh1–Mesh6.

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

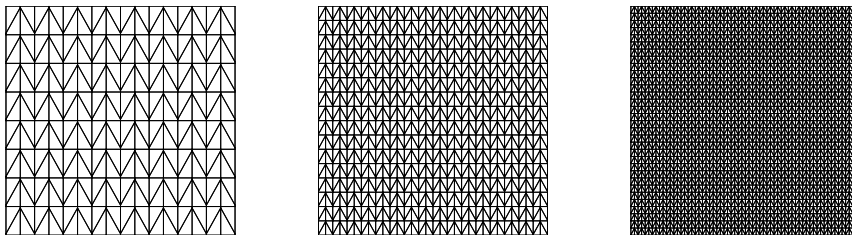


FIG. 5.4. Three samples of family of the uniform trapezoidal meshes Mesh7 with `ratio=1:19`.

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

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

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

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

278 This test is inspired by a steady numerical test in [19]. Convergence investigation  
 279 is conducted on two mesh types **Mesh1** and **Mesh2**. [Table 5.4](#) and [Table 5.5](#) show  
 280 the numbers of unknowns **nunkw**, the relative errors of the solution and flux, and the  
 281 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  $\Omega$  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

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



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

284 Table 5.6–Table 5.9 present the numerical convergence of the scheme on the uni-  
 285 form triangular mesh (see Figure 5.5(b)), uniform square mesh, locally refined quadri-  
 286 lateral mesh Mesh5 and uniform trapezoidal mesh Mesh7 (see Figure 5.4). We have  
 287 the following results:

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

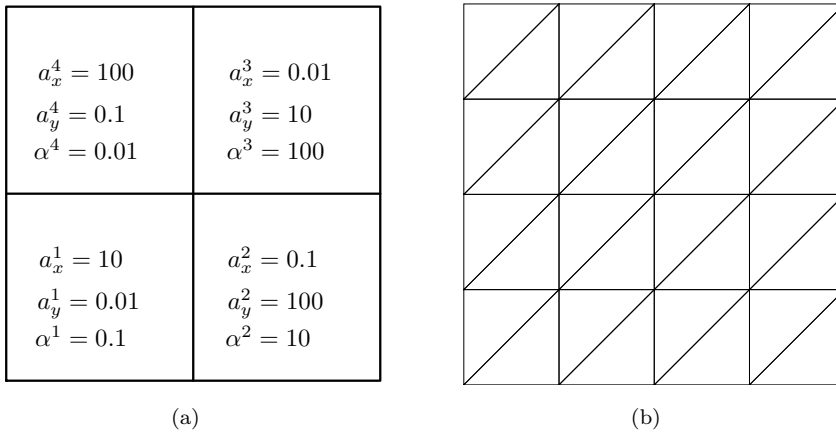


FIG. 5.5. Coefficients in the definition of diffusion tensor and exact solution (left) and the uniform triangular mesh (right).

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

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

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

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

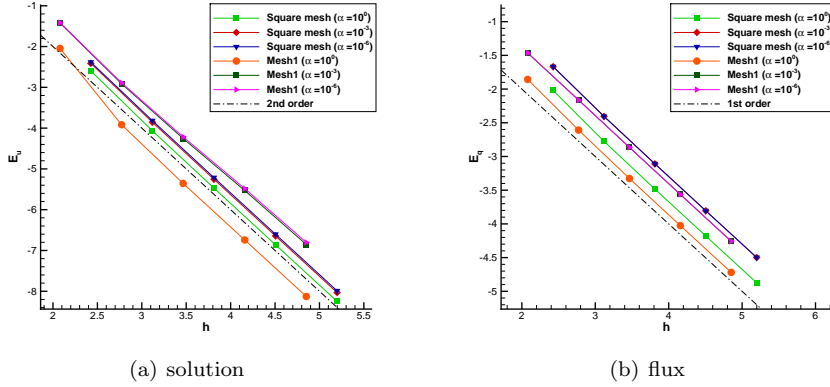


FIG. 5.6.  $L_2$  errors of the solution and flux versus mesh size  $h$  on the uniform square mesh and triangular mesh **Mesh1** with  $\alpha = 1, 10^{-3}$  and  $10^{-6}$ .

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

308 **Appendix.** We show that the new vertex-centered scheme reduces to the  $P_1$ -FVEM  
 309 on triangular meshes. Still take the triangle cell  $K$  in [Figure 2.2](#) for exposition. By  
 310 the definition of  $\delta \mathbf{U}_K$ , we get

$$(6.1) \quad \delta \mathbf{U}_K = \mathbb{T}_K \mathbf{U}_K,$$

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

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

It follows from [\(2.1\)](#), [\(3.4\)](#) and [\(6.1\)](#) that

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

and further,

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

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

$$(6.2) \quad \begin{aligned} c &= \frac{1}{|K|} (\mathbf{x}_\sigma - \mathbf{x}_K)^T \mathcal{R}^T \Lambda_K \mathcal{R}(\mathbf{x}_{\sigma^-} - \mathbf{x}_\sigma), \\ c_+ &= \frac{1}{|K|} (\mathbf{x}_\sigma - \mathbf{x}_K)^T \mathcal{R}^T \Lambda_K \mathcal{R}(\mathbf{x}_\sigma - \mathbf{x}_{\sigma^+}), \\ c_- &= \frac{1}{|K|} (\mathbf{x}_\sigma - \mathbf{x}_K)^T \mathcal{R}^T \Lambda_K \mathcal{R}(\mathbf{x}_{\sigma^+} - \mathbf{x}_{\sigma^-}). \end{aligned}$$

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

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

with

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

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

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

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

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

313 By recalling the definitions of  $\mathbf{x}_\sigma, \mathbf{x}_{\sigma^+}$  and  $\mathbf{x}_{\sigma^-}$ , we find that (6.2) is equivalent to  
314 (6.3). Hence, on triangular grids, the present vertex-centered scheme is identical to  
315  $P_1$ -FVEM in the sense of yielding the same discrete flux expression.

#### REFERENCES

- 316
- 317 [1] R. E. BANK AND D. J. ROSE, *Some error estimates for the box method*, SIAM J. Numer. Anal.,  
318 24 (1987), pp. 777–787.
- 319 [2] F. BREZZI, A. BUFFA, AND K. LIPNIKOV, *Mimetic finite differences for elliptic problems*, E-  
320 SAIM: M2AN, 43 (2009), pp. 277–295.
- 321 [3] F. BREZZI, K. LIPNIKOV, AND V. SIMONCINI, *A family of mimetic finite difference methods on*  
322 *polygonal and polyhedral meshes*, Math. Mod. Methods Appl. Sci., 15 (2005), pp. 1533–  
323 1551.
- 324 [4] Z. CAI, *On the finite volume element method*, Numer. Math., 58 (1991), pp. 713–735.
- 325 [5] Z. CHEN, R. LI, AND A. ZHOU, *A note on the optimal  $L_2$ -estimate of the finite volume element*  
326 *method*, Adv. Comput. Math., 16 (2002), pp. 291–303.
- 327 [6] L. BEIRAO DA VEIGA, F. BREZZI, A. CANGIANI, G. MANZINI, L. D. MARINI, AND A. RUSSO,  
328 *Basic principles of virtual element methods*, Math. Models Methods Appl. Sci., 23 (2013),  
329 pp. 119–214.
- 330 [7] L. BEIRAO DA VEIGA, K. LIPNIKOV, AND G. MANZINI, *Arbitrary-order nodal mimetic dis-*  
331 *cretizations of elliptic problems on polygonal meshes*, SIAM J. Numer. Anal., 49 (2011),  
332 pp. 1737–1760.
- 333 [8] R. E. EWING, T. LIN, AND Y. LIN, *On the accuracy of the finite volume element method based*  
334 *on piecewise linear polynomials*, SIAM J. Numer. Anal., 39 (2002), pp. 1865–1888.
- 335 [9] R. EYMARD, G. HENRY AND R. HERBIN, F. HUBERT, R. KLOFKORN, AND G. MANZINI, *3d bench-*  
336 *mark on discretization schemes for anisotropic diffusion problems on general grids*, in Fi-  
337 *nite Volumes for Complex Applications VI - Problems and Perspectives*, Jaroslav Fort, Jiri  
338 Furst, Jan Halama, Raphael Herbin, and Florence Hubert, eds., Springer, 2011, pp. 893–  
339 929.

- 340 [10] R. EYMARD, T. GALLOUËT, AND R. HERBIN, *Discretization of heterogeneous and anisotropic*  
 341 *diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization*  
 342 *and hybrid interfaces*, IMA J. Numer. Anal., 30 (2010), pp. 1009–1043.
- 343 [11] R. EYMARD, R. HERBIN, AND C. GUICHARD, *Small-stencil 3D schemes for diffusive flows in*  
 344 *porous media*, ESAIM: M2AN, 46 (2012), pp. 265–290.
- 345 [12] XINLONG FENG, RONGFEI LI, YINNAN HE, AND DEMIN LIU, *P1-nonconforming quadrilateral*  
 346 *finite volume methods for the semilinear elliptic equations*, J. Sci. Comput., 52 (2012),  
 347 pp. 519–545.
- 348 [13] V. GIRAULT, *Theory of a finite difference method on irregular networks*, SIAM J. Numer. Anal.,  
 349 11 (1974), pp. 260–282.
- 350 [14] R. HERBIN AND F. HUBERT, *Benchmark on discretization schemes for anisotropic diffusion*  
 351 *problems on general grids*, in Finite Volumes for Complex Applications V-Problems and  
 352 Perspectives, Robert Eymard and Jean-Marc Herard, eds., Wiley press, 2008, pp. 659–692.
- 353 [15] J. HUANG AND S. XI, *On the finite volume element method for general self-adjoint elliptic*  
 354 *problem*, SIAM J. Numer. Anal., 35 (1998), pp. 1762–1774.
- 355 [16] R. LI, Z. CHEN, AND W. WU, *Generalized difference methods for differential equations*, Marcel  
 356 Dekker, New York, 2000.
- 357 [17] Y. LI AND R. LI, *Generalized difference methods on arbitrary quadrilateral networks*, J. Com-  
 358 put. Math., 17 (1999), pp. 653–672.
- 359 [18] K. LIPNIKOV AND G. MANZINI, *A high-order mimetic method on unstructured polyhedral meshes*  
 360 *for the diffusion equation*, J. Comput. Phys., 272 (2014), pp. 360–385.
- 361 [19] K. LIPNIKOV, D. SVYATSKIY, AND Y. VASSILEVSKI, *Interpolation-free monotone finite volume*  
 362 *method for diffusion equations on polygonal meshes*, J. Comput. Phys., 228 (2009), pp. 703–  
 363 716.
- 364 [20] JUNLIANG LV AND YONGHAI LI, *L2 error estimate of the finite volume element methods on*  
 365 *quadrilateral meshes*, Adv. Comput. Math., 33 (2010), pp. 129–148.
- 366 [21] ———, *Optimal biquadratic finite volume element methods on quadrilateral meshes*, SIAM J.  
 367 Numer. Anal., 50 (2012), pp. 2379–2399.
- 368 [22] CHRISTOPHE LE POTIER, *schema volumes finis monotone pour des operateurs de diffusion*  
 369 *fortement anisotropes sur des maillages de triangle non structures*, C. R. Math. Acad. Sci.  
 370 Paris, 341 (2005), pp. 787–792.
- 371 [23] E. SÜLI, *Convergence of finite volume schemes for poisson’s equation on nonuniform meshes*,  
 372 SIAM J. Numer. Anal., 28 (1991), pp. 1419–1430.
- 373 [24] W. SUN, J. WU, AND X. ZHANG, *A family of linearity-preserving schemes for anisotropic*  
 374 *diffusion problems on arbitrary polyhedral grids*, Comput. Methods Appl. Mech. Engrg.,  
 375 267 (2013), pp. 418–433.
- 376 [25] JIMING WU, ZIHUAN DAI, ZHIMING GAO, AND GUANGWEI YUAN, *Linearity preserving nine-point*  
 377 *schemes for diffusion equation on distorted quadrilateral meshes*, J. Comput. Phys., 229  
 378 (2010), pp. 3382–3401.
- 379 [26] L. YIN, J. WU, AND Z. GAO, *The cell functional minimization scheme for the anisotropic*  
 380 *diffusion problems on arbitrary polygonal grids*, ESAIM: M2AN, published online (2014).
- 381 [27] LI YIN, JIMING WU, AND YANZHONG YAO, *A cell functional minimization scheme for parabolic*  
 382 *problem*, J. Comput. Phys., 229 (2010), pp. 8935–8951.
- 383 [28] Z. ZHANG AND Q. ZOU, *Vertex-centered finite volume schemes of any order over quadrilateral*  
 384 *meshes for elliptic boundary value problems*, Numer. Math., published online (2014).