Mathematical Models and Methods in Applied Sciences
Vol. 17, No. 1 (2007) 1–32

World Scientific
www.worldscientific.com

ON VERTEX RECONSTRUCTIONS FOR CELL-CENTERED FINITE VOLUME APPROXIMATIONS OF 2D ANISOTROPIC DIFFUSION PROBLEMS

ENRICO BERTOLAZZI

Dipartimento Ingegneria Meccanica e Strutturale,
Università di Trento, via Mesiano 77, Trento, Italy
enrico.bertolazzi@ing.unitn.it

GIANMARCO MANZINI

Istituto di Matematica Applicata e Tecnologie Informatiche,
Consiglio Nazionale delle Ricerche, via Ferrata 1, 27100 Pavia, Italy

Received 14 June 2005
Revised 9 December 2005
Communicated by N. Bellomo

The accuracy of the *diamond scheme* is experimentally investigated for anisotropic diffusion problems in two space dimensions. This finite volume formulation is cell-centered on unstructured triangulations and the numerical method approximates the cell averages of the solution by a suitable discretization of the flux balance at cell boundaries. The key ingredient that allows the method to achieve second-order accuracy is the reconstruction of vertex values from cell averages. For this purpose, we review several techniques from the literature and propose a new variant of the reconstruction algorithm that is based on linear Least Squares. Our formulation unifies the treatment of internal and boundary vertices and includes information from boundaries as linear constraints of the Least Squares minimization process. It turns out that this formulation is well-posed on those unstructured triangulations that satisfy a general regularity condition. The performance of the finite volume method with different algorithms for vertex reconstructions is examined on three benchmark problems having full Dirichlet, Dirichlet-Robin and Dirichlet–Neumann boundary conditions. Comparison of experimental results shows that an important improvement of the accuracy of the numerical solution is attained by using our Least Squares-based formulation. In particular, in the case of Dirichlet–Neumann boundary conditions and strongly anisotropic diffusions the good behavior of the method relies on the absence of locking phenomena that appear when other reconstruction techniques are used.

Keywords: Finite volumes; convection diffusion; unstructured meshes; PDE; numerical solvers.

AMS Subject Classification: 65C20, 65N22, 76R50

1. Introduction

Finite volume methods have been widely used for the numerical approximation of conservation laws and nonlinear elliptic and parabolic problems arising from many practical applications such as computer simulation of nonlinear heat transfer,[37] semiconductor devices,[20] compressible Navier–Stokes fluid dynamics,[26] petroleum reservoir models,[10] groundwater flow[39] and contaminant transport.[14,15] In the early literature, finite volumes were investigated as the so-called integral finite difference methods,[42,43] box schemes[6,29] and generalized finite difference schemes.[30,38] More recently, connections have been established between finite volumes and finite elements, yielding finite volume elements,[18,19,40,44] that seek for numerical approximations in finite element spaces. The degrees-of-freedom of finite volume elements are associated with the vertices of the computational grid and the discrete unknowns approximate the analytical solution over mesh control volumes with arbitrary shapes (Voronoi regions, median dual cells, etc.). Another class of finite volume methods, which has been developed in this last decade, is the one of cell-centered discretizations where the scheme unknowns are approximations of cell averages over simplices (triangles in 2D, tetrahedrons in 3D).[22] In the framework of cell-centered discretizations, we proposed a second-order accurate finite volume method that was successfully applied to steady transport models with scalar diffusion and homogeneous Dirichlet boundary condition.[13] A nonlinear variant of this scheme was also designed to ensure a discrete maximum principle to the numerical solution of diffusion[11] and advection-diffusion[16] problems. Recent developments also try to approximate the entire gradient as the adjoint of the divergence operator[32–35] by simultaneously using primal and dual meshes.[23,31] Other theoretical results were obtained for both classes of finite volume methods regarding *a priori*[24,25] and *a posteriori*[9] error estimates.

A number of engineering applications among those mentioned at the beginning of this paper may entail mathematical formulations with anisotropic coefficients in second derivative terms. Therefore, finite volume practitioners have focused their attention on the design of numerical methods that properly handle diffusion anisotropy.[1,2,36] For these reasons, in this paper we experimentally investigate the behavior of our cell-centered scheme when a strong diffusion anisotropy characterizes the model equation. In addition, by developing an idea that we have briefly presented in a preliminary publication,[17] we derive a new technique that is effective for implementing boundary conditions of very general type.

To these aims, let us consider the second-order elliptic problem:

Find a function u *satisfying*:

$$-\mathrm{div}(\mathcal{K}\nabla\mathsf{u}) = \mathsf{s}, \quad \text{in } \Omega, \tag{1.1a}$$

$$\mathsf{u} = \mathsf{g}^D, \quad \text{on } \Gamma^D, \tag{1.1b}$$

$$\tau\mathsf{u} + \mathbf{n}\cdot\mathcal{K}\nabla\mathsf{u} = \mathsf{g}^R, \quad \text{on } \Gamma^R, \tag{1.1c}$$

where $\Omega \subset \mathbb{R}^2$ is a polygonal domain with boundary $\Gamma = \Gamma^D \cup \Gamma^R$ and $\Gamma^D \cap \Gamma^R = \emptyset$. We assume that Γ^D and Γ^R are piecewise continuously differentiable. In the model problem described by (1.1a)–(1.1c), $\mathcal{K}$ is the 2×2 constant conductivity tensor, s the forcing term, $\mathbf{n}$ the unit vector almost everywhere orthogonal to Γ and pointing out from Ω, g^D and g^R the Dirichlet and Robin boundary functions respectively defined almost everywhere on Γ^D and Γ^R, and τ a suitable real factor. It is worth noting that the formulation in (1.1c) includes the case of pure Neumann condition for $\tau = 0$. Under suitable assumptions on the regularity of s and the boundary functions g^D and g^R, this model problem can be reformulated in weak form, and existence and uniqueness of solutions can be proved.[41]

In order to solve (1.1a)–(1.1c), we design a second-order accurate finite volume method on unstructured triangulations by first constructing the numerical diffusion flux of any mesh edge, and then defining the discrete "mass"-balance of any control volume by the divergence theorem. The numerical diffusion flux is defined by using both cell averages and vertex values of the solution, according to the formulation of the "diamond scheme".[13,22] Since in the framework of cell-centered methods we want to retain cell averages like scheme unknowns, we introduce a reconstruction algorithm for determining vertex values from cell averages. For any vertex of the mesh, we first define the reconstruction stencil as union of the cells that share the vertex; then, we calculate a cell weight for every element of the stencil. Cell weights are used in an interpolation formula that must be linearly consistent, i.e. exact on linear functions, to ensure that the resulting finite volume method be second-order accurate. It is important to realize that this reconstruction step is required by any vertex of the mesh, no matter the vertex being internal or on the boundary of the computational domain. In fact, the reconstructed value of a boundary vertex always gives a contribution to the numerical flux of the internal edges to which the vertex belongs. For the Dirichlet boundary Γ^D, the reconstruction is immediate, because, consistently with (1.1b), the vertex is assigned the value provided by evaluating g^D at the vertex position. For the Robin boundary Γ^R, the reconstruction is a much more difficult task, because the vertex value, although determined by the boundary condition (1.1c), is no longer available from the direct evaluation of a boundary function. We emphasize that an incorrect definition of the value of the boundary vertex may deteriorate the quality of the numerical solution, and produce a degeneration of the scheme performance to first-order accuracy[27,28] or prevent convergence at all due to locking phenomena.[5] As far as this issue is concerned, we derive the cell weights of the interpolation formula by solving an approximation problem on each reconstruction stencil by means of linear least squares. When this procedure is applied to a Robin boundary vertex, information from boundary condition (1.1c) is introduced in the evaluation of cell weights by suitable linear constraints. Thus, a suitable definition of cell weights allows us to apply the same interpolation formula to reconstruct the value of both internal and boundary vertices.

The outline of the paper is as follows. In Sec. 2 we introduce the grid notations used in the paper and state the regularity assumption that must be satisfied by

every computational mesh. In Sec. 3 we present the general formulation of our cell-centered finite volume method. In Sec. 4 we design the least squares-based reconstruction technique for internal and boundary vertices. For the sake of comparison, we also review several alternative weight constructions for the vertex interpolation formula which are taken from the literature. In Sec. 5 we evaluate the accuracy of several implementations of this finite volume method for a number of combinations of the least squares weights and the other cell weights of Sec. 4. The performance of the finite volume method is experimentally obtained by comparing the convergence rates of the numerical solution for three benchmark problems on a realistic range of diffusion anisotropy. To investigate the scheme accuracy in the treatment of complex boundary conditions, the three benchmark problems take, respectively, full Dirichlet, half Dirichlet-half Robin, and half Dirichlet-half Neumann boundary conditions. In Sec. 6, we discuss some final remarks and conclusions.

2. General Setting, Notations and Mesh Regularity Assumption

In this section, we introduce the notations adopted in this paper and, for ease of reference, collect the definitions of topological and geometrical entities. The polygonal domain $\Omega \subset \mathbb{R}^2$ is covered by a triangulation, which is a finite collection of two-dimensional closed simplices with non-overlapping interiors. The simplices of the domain mesh are denoted by the letter "T" and labeled by a Latin index like i $(j, k, \ldots)$. For example, T_i is the ith control volume (cell) of the mesh. The whole set of mesh control volumes is denoted by $\mathcal{T}_h = \{\mathsf{T}_i\}$, and these latter ones are such that $\overline{\Omega} = \cup_{\mathsf{T}_i \in \mathcal{T}_h} \mathsf{T}_i$. The parameter h that labels the mesh $\mathcal{T}_h$ is called *the mesh size*, and is formally given by the supremum of the mesh control volume diameters[21]: $h = \max_{\mathsf{T}_i \in \mathcal{T}_h} \mathsf{diam}\,\{\mathsf{T}_i\}$.

The mesh edges are denoted by the letter "e" and labeled by a couple of Latin indices, as, for example, e_{ij}. It is useful to distinguish between *internal* and *boundary* edges. Any internal edge e_{ij} is shared by two control volumes T_i and T_j such that $\mathsf{e}_{ij} = \mathsf{T}_i \cap \mathsf{T}_j$. When e_{ij} is a boundary edge, the index i always refers to the unique control volume T_i to which the edge belongs, while the index j is defined in accordance with a suitable boundary numbering system (*ghost cell*). The symbols $\mathcal{F}_h$, $\mathcal{F}_h^{\mathrm{int}}$ and $\mathcal{F}_h^{\mathrm{bnd}}$ denote respectively the set of all mesh edges, the set of *internal edges*, and the set of *boundary edges*. When dealing with internal edges, e_{ij} and e_{ji} are equivalent symbols that denote the same edge. In expressions like $\mathsf{e}_{ij} \in \mathcal{F}_h$ (or $\mathcal{F}_h^{\mathrm{int}}$) we assume that the edge labeled by i and j is considered only once (for example, by taking the representative with $i < j$). We have: $\mathcal{F}_h = \mathcal{F}_h^{\mathrm{int}} \cup \mathcal{F}_h^{\mathrm{bnd}}$ and $\mathcal{F}_h^{\mathrm{int}} \cap \mathcal{F}_h^{\mathrm{bnd}} = \emptyset$.

The mesh vertices are denoted by the symbol "v" and labeled by Greek letters like α $(\beta, \gamma, \ldots)$. The symbols $\mathcal{V}_h$, $\mathcal{V}_h^{\mathrm{int}}$ and $\mathcal{V}_h^{\mathrm{bnd}}$ denote respectively the set of all mesh vertices, the set of *internal* vertices, and the set of *boundary* vertices. This last set is split for convenience in the union of Dirichlet and Robin boundary vertices, respectively denoted by $\mathcal{V}_h^{\mathrm{Dir}}$ and $\mathcal{V}_h^{\mathrm{Rob}}$. We have: $\mathcal{V}_h^{\mathrm{bnd}} = \mathcal{V}_h^{\mathrm{Dir}} \cup \mathcal{V}_h^{\mathrm{Rob}}$, $\mathcal{V}_h^{\mathrm{Dir}} \cap \mathcal{V}_h^{\mathrm{Rob}} = \emptyset$, $\mathcal{V}_h = \mathcal{V}_h^{\mathrm{int}} \cup \mathcal{V}_h^{\mathrm{bnd}}$ and $\mathcal{V}_h^{\mathrm{int}} \cap \mathcal{V}_h^{\mathrm{bnd}} = \emptyset$.

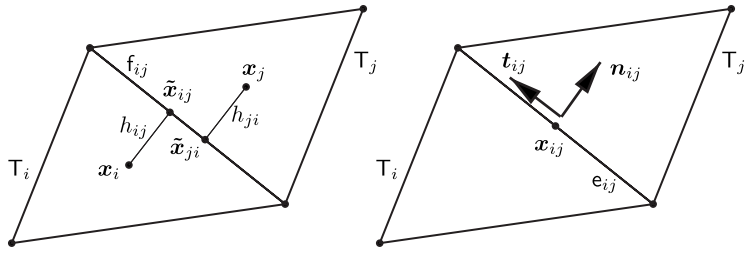


Fig. 1. Mesh entities for an internal edge.

The mesh entities related to the cell T_i are consistently labeled by the same index i:

— $|\mathsf{T}_i|$, the area of T_i,
— $\partial\mathsf{T}_i$, the boundary of T_i,
— $\mathbf{x}_i$, the barycenter of T_i.

The mesh entities related to the edge e_{ij}, which are displayed in Fig. 1, are consistently labeled by the same couple of indices ij:

— $|\mathsf{e}_{ij}|$, the length of $\mathsf{e}_{ij} \in \mathcal{F}_h$;
— $\mathbf{x}_{ij}$, the center of $\mathsf{e}_{ij} \in \mathcal{F}_h$;
— $\mathbf{n}_{ij}$, the unit vector orthogonal to $\mathsf{e}_{ij} \in \mathcal{F}_h$ and oriented from T_i to T_j if $\mathsf{e}_{ij} \in \mathcal{F}_h^{\mathrm{int}}$, pointing out of Ω if $\mathsf{e}_{ij} \in \mathcal{F}_h^{\mathrm{bnd}}$;
— $\mathbf{t}_{ij}$, the unit vector parallel to $\mathsf{e}_{ij} \in \mathcal{F}_h$ and such that the vector pair $(\mathbf{n}_{ij}, \mathbf{t}_{ij})$ forms a counterclockwise oriented basis of $\mathbb{R}^2$;
— $\tilde{\mathbf{x}}_{ij}$, the orthogonal projection of $\mathbf{x}_i$ on the line containing the edge $\mathsf{e}_{ij} \in \mathcal{F}_h$;
— $h_{ij} = (\tilde{\mathbf{x}}_{ij} - \mathbf{x}_i) \cdot \mathbf{n}_{ij}$, the distance between the center of T_i and the edge $\mathsf{e}_{ij} \in \mathcal{F}_h$;
— $H_{ij} = (\mathbf{x}_j - \mathbf{x}_i) \cdot \mathbf{n}_{ij} = h_{ij} + h_{ji}$, the *effective distance* between the centers of T_i and T_j when $\mathsf{e}_{ij} \in \mathcal{F}_h^{\mathrm{int}}$.

The mesh entities related to the mesh vertex v_α are consistently labeled by the same index α:

— $\mathbf{x}_\alpha$, the position vector of v_α;
— $\mathcal{V}_\alpha$, the co-volume associated to v_α and defined in the median dual mesh.

According to the vertex notation (and with slight exception to the edge notation), we indicate the two edges that are incident to the boundary vertex v_α by the symbols $\mathsf{e}_\alpha^\pm$. The labels "$\pm$" are set by assuming that the boundary Γ is counterclockwise oriented; the edge labeled by the minus sign "$-$" is encountered before the vertex v_α and the edge labeled by the plus sign "$+$" is encountered after the vertex v_α. We define:

— $\mathbf{n}_\alpha^\pm$, unit vectors orthogonal to $\mathsf{e}_\alpha^\pm$;

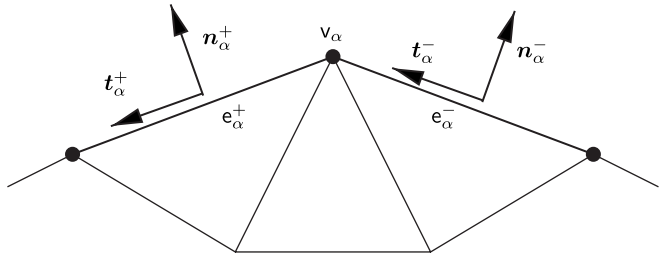


Fig. 2. Mesh entities for a boundary vertex.

— $\mathbf{t}_\alpha^\pm$, unit vectors parallel to $\mathsf{e}_\alpha^\pm$ and such that the vector pairs $(\mathbf{n}_\alpha^\pm, \mathbf{t}_\alpha^\pm)$ form two possibly distinct counterclockwise oriented basis of $\mathbb{R}^2$.

The mesh entities associated to the boundary vertex v_α are shown in Fig. 2.

We denote the barycentric coordinates of $\tilde{\mathbf{x}}_{ij}$ with respect to the vertices v_α and v_β of e_{ij} by $\tilde{\lambda}_\alpha^{ij}$ and $\tilde{\lambda}_\beta^{ij}$. By definition, it must hold that $\tilde{\lambda}_\alpha^{ij} + \tilde{\lambda}_\beta^{ij} = 1$ and $\tilde{\lambda}_\alpha^{ij}\mathbf{x}_\alpha + \tilde{\lambda}_\beta^{ij}\mathbf{x}_\beta = \tilde{\mathbf{x}}_{ij}$. The following algebraic vectors, which are associated to the vertex v_α, will be used in the formulation of the least squares reconstruction of vertex values:

— $\mathbf{u}_\alpha = (u_1, \ldots, u_{\mathsf{N}_\alpha})^T \in \mathbb{R}^{\mathsf{N}_\alpha}$, the cell averages of the triangles surrounding v_α;
— $\mathbf{w}_\alpha = (\mathrm{w}_{\alpha,1}, \ldots, \mathrm{w}_{\alpha,\mathsf{N}_\alpha})^T$, the cell weights for the elements of $\mathbf{u}_\alpha$;
— $\mathbf{g}_\alpha = (\mathrm{g}_\alpha^-, \mathrm{g}_\alpha^+)^T \in \mathbb{R}^2$, the data associated to the boundary edges $\mathsf{e}_\alpha^\pm$ incident to the boundary vertex v_α:

$$\mathrm{g}_\alpha^\pm = \begin{cases} 0, & \text{if } \mathsf{v}_\alpha \in \mathcal{V}_h^{\mathrm{int}}, \\ \mathrm{g}^D(\mathbf{x}_\alpha), & \text{if } \mathsf{v}_\alpha \in \mathcal{V}_h^{\mathrm{Dir}}, \\ \lim_{\mathbf{X}\to\mathbf{x}_\alpha^\pm} \mathrm{g}^R(\mathbf{x}), & \text{if } \mathsf{v}_\alpha \in \mathcal{V}_h^{\mathrm{Rob}}, \end{cases} \tag{2.1}$$

where the limits to $\mathbf{x}_\alpha^\pm$ in the third expression are taken along the edges $\mathsf{e}_\alpha^\pm$;
— $\mathbf{w}_\alpha^b = (\mathrm{w}_\alpha^-, \mathrm{w}_\alpha^+)^T$, the edge weights for the elements of $\mathbf{g}_\alpha$.

Finally, summations are taken over the following sets of indices:

— σ_i, indices of the cells sharing an edge with T_i,
— σ_α, indices of the cells surrounding the vertex v_α,
— σ_i', indices of the boundary edges of T_i,
— ν_{ij}, indices of the vertices of the edge e_{ij}, i.e. α and β.

Coherently with edge notations, e_{ij} is a boundary edge of T_i for every $j \in \sigma_i'$ and the internal edge shared by T_i and T_j for $j \in \sigma_i$. Thus, the index $j \in \sigma_i \cup \sigma_i'$ labels all the edges forming the boundary of T_i.

2.1. *Mesh regularity assumption*

Let us denote by ρ_i the maximum radius of the balls contained in the cell T_i. The finite volume method described in this paper is formulated on a family of two-dimensional grids $\{\mathcal{T}_h\}$ that are *regular* if:

(i) there exists a *mesh regularity constant* $C_{\text{reg}} > 0$ that is independent of h_i and such that $h_i/\rho_i \leq C_{\text{reg}}$ for any $h \leq h_0$;
(ii) the triangulations $\{\mathcal{T}_h\}$ are weakly acute for any $h \leq h_0$.

The first property is normally required in the convergence analysis of conforming finite element methods[21] to ensure that the shape of triangles does not degenerate when $h \to 0$. Although the analysis of our finite volume method is beyond the scope of this paper, we take this condition for the same reason. The second property ensures that $\tilde{\mathbf{x}}_{ij}$ is an internal point of the edge e_{ij}.

2.2. *The cell average operator*

Let $\mathbb{P}^0(\mathcal{T}_h)$ be the subspace of the Hilbert space $L^2(\Omega)$ consisting of the functions having constant restriction on the control volumes of the triangulation $\mathcal{T}_h$. We denote the orthogonal projector from $L^2(\Omega)$ onto $\mathbb{P}^0(\mathcal{T}_h)$ by $\mathcal{A}_h(\cdot)$ and call it *the cell-average operator*. This operator is used in Sec. 5 to define convergence errors on cell averages and gradients of the analytical solution. The dimension of $\mathbb{P}^0(\mathcal{T}_h)$ is equal to $\mathsf{N_T}$ and the identification between $\mathbb{P}^0(\mathcal{T}_h)$ and $\mathbb{R}^{\mathsf{N_T}}$ is readily established. Therefore, we can interpret this operator as the vector-valued functional $\mathcal{A}_h : L^2(\Omega) \mapsto \mathbb{R}^{\mathsf{N_T}}$ such that

$$\mathcal{A}_h(\mathsf{u})|_i = \frac{1}{|\mathsf{T}_i|} \int_{\mathsf{T}_i} \mathsf{u}(\mathbf{x})\, dV, \quad \text{for } 0 \leq i \leq \mathsf{N_T}.$$

Throughout the paper, we will also use the shortcut $\mathcal{A}_i(\mathsf{u}) = \mathcal{A}_h(\mathsf{u})|_i$.

3. Finite Volume Formulation

The cell-centered finite volume method that approximates the solution of (1.1a)–(1.1c) is derived by exploiting the local conservation property

$$-\int_{\partial \mathsf{T}_i} \mathbf{n} \cdot \mathcal{K} \nabla \mathsf{u}\; dS = \int_{\mathsf{T}_i} \mathsf{s}\; dV, \quad \text{for } \mathsf{T}_i \in \mathcal{T}_h. \tag{3.1}$$

Let us denote the approximate cell average of u on the cell T_i by u_i, and the $\mathsf{N_T}$-sized vector containing all the approximate cell averages of u associated to the control volumes of $\mathcal{T}_h$ by $\mathsf{u}_h \in \mathbb{R}^{\mathsf{N_T}}$. The vector space $\mathbb{R}^{\mathsf{N_T}}$ is isomorphic to the subspace of $L^2(\Omega)$ of the functions defined on Ω and having a constant restriction on any cell of $\mathcal{T}_h$. For this reason, throughout the paper we will exploit the identification between these two spaces by indicating corresponding elements with the same symbol. The

cell-centered finite volume discretization approximates the local conservation property (3.1) by the *discrete* balance equations:

$$\frac{1}{|\mathsf{T}_i|} \sum_{j \in \sigma_i \cup \sigma_i'} |\mathsf{e}_{ij}|\, \Phi_{ij}(\mathsf{u}_h) = \mathsf{s}_i, \quad \text{for } \mathsf{T}_i \in \mathcal{T}_h. \tag{3.2}$$

The ith equation in (3.2) correlates the approximate cell average u_i to the integral balance of the numerical fluxes $\Phi_{ij}(\mathsf{u}_h)$ across the edges e_{ij} forming the control volume boundary $\partial\mathsf{T}_i$ and the cell average s_i of the source term s in (1.1a). The proper definition of the numerical flux $\Phi_{ij}(\mathsf{u}_h)$ from cell averages is crucial to provide second-order accuracy of the cell-centered finite volume method. To this aim, let $\boldsymbol{\mathcal{G}}_{ij}^{\diamond}(\mathsf{u}_h)$ and $\boldsymbol{\mathcal{G}}_{ij}^{\mathrm{Dir}}(\mathsf{u}_h)$ respectively denote the discrete constant gradients on the internal edges $\mathcal{F}_h^{\mathrm{int}}$ and the boundary edges of Dirichlet type $\mathcal{E}_h^{\mathrm{Dir}}$.

The numerical diffusive flux is given by

$$\Phi_{ij}(\mathsf{u}_h) = \begin{cases} -\mathbf{n}_{ij} \cdot \mathcal{K}_{ij} \boldsymbol{\mathcal{G}}_{ij}^{\diamond}(\mathsf{u}_h), & \text{if } \mathsf{e}_{ij} \in \mathcal{F}_h^{\mathrm{int}}, \\ -\mathbf{n}_{ij} \cdot \mathcal{K}_{ij} \boldsymbol{\mathcal{G}}_{ij}^{\mathrm{Dir}}(\mathsf{u}_h), & \text{if } \mathsf{e}_{ij} \in \mathcal{E}_h^{\mathrm{Dir}}, \\ \mathsf{g}_{ij}^{R} - \tau(u_\alpha + u_\beta)/2, & \text{if } \mathsf{e}_{ij} \in \mathcal{E}_h^{\mathrm{Rob}}. \end{cases} \tag{3.3}$$

The first two formulas of (3.3) approximate the co-normal derivative of u that appears in (3.1), i.e. $\mathbf{n} \cdot \mathcal{K}\nabla\mathsf{u}$, by the discrete counterpart that is derived by first decomposing the flux integral (3.1) on every edge $\mathsf{e}_{ij} \in \partial\mathsf{T}_i$ and then applying the second-order mid-point quadrature rule. We define $\mathcal{K}_{ij}$ by evaluating the diffusion tensor at $\mathbf{x}_{ij} \in \mathsf{e}_{ij}$, i.e. $\mathcal{K}_{ij} = \mathcal{K}(\mathbf{x}_{ij})$, or by averaging $\mathcal{K}$ on e_{ij} in a componentwise sense, that is

$$\mathcal{K}_{ij} = \frac{1}{|\mathsf{e}_{ij}|} \int_{\mathsf{e}_{ij}} \mathcal{K}(\mathbf{x})\, dS.$$

The third formula of the numerical flux definition (3.3) accounts for the Robin boundary conditions expressed by the function g^R on the right-hand side of (1.1c). In analogy with the definition of $\mathcal{K}_{ij}$, the discrete term g_{ij}^R can be given by a direct evaluation of the boundary function $\mathsf{g}_{ij}^R = \mathsf{g}^R(\mathbf{x}_{ij})$, or by the line-integral average

$$\mathsf{g}_{ij}^R = \frac{1}{|\mathsf{e}_{ij}|} \int_{\mathsf{e}_{ij}} \mathsf{g}^R(\mathbf{x})\, dS.$$

3.1. *Approximation of the solution gradient at mesh edges*

The derivation of the discrete gradient formulas for internal and Dirichlet boundary edges starts from defining the *one-sided* gradient $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$ associated to the

edge e_{ij} from the interior of T_i. The discrete gradient $\boldsymbol{\mathcal{G}}^{\diamond}_{ij}(\mathsf{u}_h)$ of the internal edge $\mathsf{e}_{ij} \in \mathcal{F}_h^{\mathrm{int}}$ is uniquely defined by averaging the two one-sided contributions that are built within the cells T_i and T_j sharing e_{ij}. Instead, the discrete gradient $\boldsymbol{\mathcal{G}}^{\mathrm{Dir}}_{ij}(\mathsf{u}_h)$ at each boundary edge of Dirichlet type $\mathsf{e}_{ij} \in \mathcal{E}_h^{\mathrm{Dir}}$ coincides with the gradient defined within the (unique) cell T_i to which edge e_{ij} belongs. In order to formulate the one-sided edge gradient $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$, we apply the Gauss–Green relation to the volume integral of $\mathcal{K}\nabla\mathsf{u}$ on the triangular domain defined by the center of gravity of T_i and the vertices of edge e_{ij}. Let us recall that α and β denote the indices labeling the vertices of e_{ij}, i.e. $\nu_{ij} = \{\alpha, \beta\}$, and $\tilde{\mathbf{x}}_{ij} \in \mathsf{e}_{ij}$ is the orthogonal projection of the barycenter $\mathbf{x}_i$ onto the edge e_{ij}. The vertex values u_α and u_β at vertices v_α and v_β are provided by one of the reconstruction algorithms that will be described in Sec. 4. The approximation of the solution at $\tilde{\mathbf{x}}_{ij}$ is given by the linear interpolation:

$$\tilde{u}_{ij} = \tilde{\lambda}^{ij}_{\alpha} u_\alpha + \tilde{\lambda}^{ij}_{\beta} u_\beta. \tag{3.4}$$

As $\tilde{\mathbf{x}}_{ij}$ is an internal point of e_{ij} because of the mesh regularity assumption, its barycentric coordinates $\tilde{\lambda}^{ij}_{\alpha}$ and $\tilde{\lambda}^{ij}_{\beta}$ are non-negative numbers bounded from above by unity. Thus, $\tilde{u}_{ij}$ is a *convex* linear interpolation of u_α and u_β. The one-sided gradient at e_{ij} within T_i is finally written as

$$\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h) = \frac{\tilde{u}_{ij} - u_i}{h_{ij}} \mathbf{n}_{ij} + \frac{u_\beta - u_\alpha}{|\mathsf{e}_{ij}|} \mathbf{t}_{ij}. \tag{3.5}$$

3.1.1. *Internal edges*

A unique definition of the numerical edge flux is required to obtain a *conservative* formulation of the numerical diffusive flux $\Phi_{ij}(\mathsf{u}_h)$. To this aim, given an internal edge e_{ij}, the two contributions arising from the one-sided numerical gradients, $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$ built in T_i and $\boldsymbol{\mathcal{G}}_{ji}(\mathsf{u}_h)$ built in T_j, are averaged by:

$$\boldsymbol{\mathcal{G}}^{\diamond}_{ij}(\mathsf{u}_h) = W_{ij} \boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h) + W_{ji} \boldsymbol{\mathcal{G}}_{ji}(\mathsf{u}_h), \tag{3.6}$$

where the non-negative weights are:

$$W_{ij} = 1 - W_{ji} = \frac{h_{ij}}{H_{ij}}.$$

This results in the following definition of the numerical flux:

$$\Phi_{ij}(\mathsf{u}_h) = \left(\frac{u_j - u_i}{H_{ij}} \right) \kappa^{(n)}_{ij} + \left(\frac{u_\beta - u_\alpha}{|\mathsf{e}_{ij}|} \right) \kappa^{(t)}_{ij},$$

where the coefficients $\kappa_{ij}^{(n)}$ and $\kappa_{ij}^{(t)}$ contain the normal and tangential projections of the co-normal vector $\mathcal{K}\mathbf{n}_{ij}$, i.e.

$$\kappa_{ij}^{(n)} = \mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{n}_{ij}, \tag{3.7a}$$

$$\kappa_{ij}^{(t)} = \left(\tilde{\lambda}_{\alpha}^{ij} + \tilde{\lambda}_{\beta}^{ji} - 1\right) \frac{|\mathsf{e}_{ij}|}{H_{ij}} \mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{n}_{ij} + \mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{t}_{ij}. \tag{3.7b}$$

Note that the tangential term $\mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{t}_{ij}$ in (3.7b) disappears when diffusion is isotropic, i.e. the diffusion tensor is a scalar multiple of the identity matrix. As a consequence, it is this tangential contribution that bears the responsibility for a consistent approximation of the anisotropic flux. The accuracy of the vertex reconstruction is thus of fundamental importance to achieve full accuracy of the tangential component of the numerical flux and good performance for anisotropic diffusion.

3.1.2. *Dirichlet boundary edges*

When $\mathsf{e}_{ij} \in \mathcal{E}_h^{\mathrm{Dir}}$, the edge gradient $\boldsymbol{\mathcal{G}}_{ij}^{\mathrm{Dir}}(\mathsf{u}_h)$ is set equal to the unique one-sided edge gradient $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$. Thus, we have

$$\boldsymbol{\mathcal{G}}_{ij}^{\mathrm{Dir}}(\mathsf{u}_h) = \frac{\tilde{u}_{ij} - u_i}{h_{ij}} \mathbf{n}_{ij} + \frac{u_\beta - u_\alpha}{|\mathsf{e}_{ij}|} \mathbf{t}_{ij},$$

so that

$$\Phi_{ij}(\mathsf{u}_h) = -\frac{u_i}{h_{ij}} \kappa_{ij}^{(n)} + \frac{\kappa_{\beta,ij}^{(t)} u_\beta - \kappa_{\alpha,ij}^{(t)} u_\alpha}{|\mathsf{e}_{ij}|}.$$

Tangential contributions involve the gradient coefficients

$$\kappa_{\alpha,ij}^{(t)} = -\frac{\tilde{\lambda}_{\beta}^{ij} |\mathsf{e}_{ij}|}{h_{ij}} \mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{n}_{ij} - \mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{t}_{ij},$$

$$\kappa_{\beta,ij}^{(t)} = +\frac{\tilde{\lambda}_{\alpha}^{ij} |\mathsf{e}_{ij}|}{h_{ij}} \mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{n}_{ij} + \mathbf{n}_{ij} \cdot \mathcal{K}_{ij}\mathbf{t}_{ij}.$$

Since at least one of the two boundary vertices v_α and v_β is of Dirichlet type, we may argue that the relation $u_\alpha = \mathsf{g}_\alpha^D$ or $u_\beta = \mathsf{g}_\beta^D$ (or both, simultaneously) must be true. We remark that the tangential component of the diffusive flux on the boundary edge is proportional to $u_\beta - u_\alpha$. This clearly implies that this flux component is completely determined by the Dirichlet boundary function g^D when both $u_\alpha = \mathsf{g}_\alpha^D$ and $u_\beta = \mathsf{g}_\beta^D$ are used.

3.2. *Source term contribution*

The discretization of the right-hand side source term s in (1.1a) is taken into account by the $\mathsf{N_T}$-sized vector $\mathbf{s} = \{\mathsf{s}_i\}$. The components of $\mathbf{s}$ are estimated by averaging the right-hand side analytical source function $\mathsf{s}(\mathbf{x})$ on T_i. We have

$$\mathsf{s}_i = \frac{1}{|\mathsf{T}_i|} \int_{\mathsf{T}_i} \mathsf{s}(\mathbf{x})\ dV = \frac{1}{3} \sum_{j \in \sigma_i \cup \sigma_i'} \mathsf{s}(\mathbf{x}_{ij}).$$

3.3. *The finite volume linear system*

The final system of linear equations is written as:

$$\mathbf{G}\mathsf{u}_h = \mathsf{s} + \mathbf{g},$$

where $\mathbf{G} = \{\mathsf{G}_{ij}\}$ is the $\mathsf{N_T} \times \mathsf{N_T}$ stiffness matrix and $\mathbf{g} = \{\mathsf{g}_i\} \in \mathbb{R}^{\mathsf{N_T}}$ is the vector of boundary conditions. The matrix $\mathbf{G}$ and the vector $\mathbf{g}$ are built in the following assembly process. First, we express $\tilde{u}_{ij}$ as a linear combination of vertex values by using (3.4), and these latter as combination of cell averages by means of a vertex reconstruction formula (see Sec. 4). Then, we use these expressions in (3.5) to obtain the one-sided gradient $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$ for every $\mathsf{T}_i \in \mathcal{T}_h$ and $j \in \sigma_i \cup \sigma_i'$. It is worth noting that the gradient $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$ is defined in terms of cell averages of the triangles sharing v_α and all triangles sharing v_β. In accordance with (3.6), we define the gradient $\boldsymbol{\mathcal{G}}_{ij}^{\diamond}(\mathsf{u}_h)$ by averaging the two one-sided gradients $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$ and $\boldsymbol{\mathcal{G}}_{ji}(\mathsf{u}_h)$. On the boundary, we take the one-sided gradient $\boldsymbol{\mathcal{G}}_{ij}(\mathsf{u}_h)$ for the edge gradient $\boldsymbol{\mathcal{G}}_{ij}^{\mathrm{Dir}}(\mathsf{u}_h)$. Finally, we reformulate the discrete flux balance (3.2) as

$$\frac{1}{|\mathsf{T}_i|} \sum_{j \in \sigma_i \cup \sigma_i'} |\mathsf{e}_{ij}|\, \Phi_{ij}(\mathsf{u}_h) = (\mathbf{G}\mathbf{u}_h - \mathbf{g})|_i = \mathsf{s}_i, \quad \text{for } \mathsf{T}_i \in \mathcal{T}_h,$$

by employing the expressions of the internal and boundary edge gradients in the definition of the numerical flux (3.3). More details of this assembly process and the formulas of $\mathbf{G}$ and $\mathbf{g}$ are reported in a previous publication by Bertolazzi and Manzini.[13] The resulting non-symmetric linear system arising from this discretization is solved by using the routine `ma41`[3] from the HSL collection of FORTRAN linear algebraic solvers for sparse matrices.

4. Vertex Reconstructions

In this section, we formulate our least squares-based algorithm and review, for the sake of comparison, a number of different algorithms from the literature for determining the vertex value from cell averages. Basically, we want to express the vertex value u_α as a weighted mean of the cell-averaged data $\{u_k,\ k \in \sigma_\alpha\}$ that are

associated to the cells surrounding v_α. All the techniques considered in this section essentially differ for the weight construction that provides the coefficients of the vertex interpolation formula. In particular, we propose a derivation of cell weights that is based on the resolution of an *unconstrained* least squares problem if the vertex is internal and a *constrained* least squares problem if the vertex is on the boundary with a condition of Robin (Neumann) type. We show in Proposition 4.1 that recovering vertex values from cell averages by using this technique is a well-posed problem on weakly acute triangulations. Such a treatment of Robin (Neumann) conditions requires the introduction of two additional terms in the interpolation formula to take into account contributions from the boundary edges $\mathsf{e}_\alpha^\pm$ that are incident to the boundary vertex v_α. No reconstruction is required by a Dirichlet boundary vertex, because the vertex value can be directly derived from the boundary function g^D. However, we will include this last case in the general formulation by suitably modifying the cell weights.

4.1. *Least squares reconstruction*

We identify the vertex value u_α with the value provided at the vertex position $\mathbf{x}_\alpha$ by the following linear approximation of the solution within the co-volume $\mathcal{V}_\alpha$:

$$\mathcal{R}_\alpha(\mathbf{x};\mathsf{u}_h) = a + \mathbf{b}\cdot(\mathbf{x}-\mathbf{x}_\alpha), \quad \mathbf{x}\in\mathcal{V}_\alpha. \tag{4.1}$$

Thus, $u_\alpha = \mathcal{R}_\alpha(\mathbf{x}_\alpha;\mathsf{u}_h)$. In the case of a Robin boundary vertex, the boundary condition (1.1c) is taken into account by introducing the additional constraints:

$$\tau a + \mathbf{n}_\alpha^\pm \cdot \Big(\mathcal{K}_\alpha \mathbf{b}\Big) = \mathsf{g}_\alpha^\pm, \tag{4.2}$$

where $\mathsf{g}_\alpha^\pm$ is defined in (2.1). Let $\{\lambda_k,\, k\in\sigma_\alpha\}$ denote a generic set of weighting coefficients that are supposed to be bounded, i.e. $0<\lambda_k<1$ for $k\in\sigma_\alpha$, and satisfy the normalization condition $\sum_{k\in\sigma_\alpha}\lambda_k = 1$. For practical calculations in Sec. 5, we will take $\lambda_k = |\mathsf{T}_k| \,/\, \sum_{k\in\sigma_\alpha}|\mathsf{T}_k|$. The coefficients $(a,\mathbf{b}^T)^T$ in (4.1)–(4.2) are solution of the least squares problem

$$(a,\mathbf{b}^T)^T = \mathrm{argmin}_{(a',\mathbf{b}'^{,T})^T\in\mathbb{V}} \left(\sum_{k\in\sigma_\alpha} \lambda_k \left[a' + \mathbf{b}'\cdot(\mathbf{x}_k-\mathbf{x}_\alpha) - u_k\right]^2 \right),$$

where

$$\mathbb{V} = \begin{cases} \Big\{ (a',\mathbf{b}'^{,T})^T\in\mathbb{R}^3 \text{ such that } \tau a' + \mathbf{n}_\alpha^\pm\cdot\big(\mathcal{K}_\alpha\mathbf{b}'\big) = \mathsf{g}_\alpha^\pm \Big\} & \text{for } \mathsf{v}_\alpha\in\mathcal{V}_h^{\mathrm{Rob}}, \\ \mathbb{R}^3 & \text{otherwise.} \end{cases}$$

4.1.1. *Technicalities*

For ease of presentation, we list below the definition of the principal matrix operators and vectors used in Proposition 4.1. All these symbols logically refer to the vertex v_α to which the reconstruction algorithm is applied and should consistently take the sub-index α. However, we shall drop α from the matrix symbols $\mathbf{\Lambda}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{B}^\dagger$ and $\mathbf{K}$ to simplify notations. We also use the vector symbol $\mathbf{0} = (0, \ldots, 0)^T$ to indicate the null vector having contextually consistent size. The matrices $\mathbf{\Lambda}$ and $\mathbf{A}$ are associated to the least squares algorithm and are defined for both internal and boundary vertices as follows:

- $\mathbf{\Lambda} = \mathsf{diag}\{\lambda_1, \ldots, \lambda_{\mathsf{N}_\alpha}\} \in \mathbb{R}^{\mathsf{N}_\alpha \times \mathsf{N}_\alpha}$ is the diagonal matrix collecting the least squares coefficients;
- $\mathbf{A} \in \mathbb{R}^{\mathsf{N}_\alpha \times 3}$ is the least squares matrix given by

$$\mathbf{A} = \begin{pmatrix} 1 & \mathsf{x}_1 - \mathsf{x}_\alpha & \mathsf{y}_1 - \mathsf{y}_\alpha \\ \vdots & \vdots & \vdots \\ 1 & \mathsf{x}_{\mathsf{N}_\alpha} - \mathsf{x}_\alpha & \mathsf{y}_{\mathsf{N}_\alpha} - \mathsf{y}_\alpha \end{pmatrix}.$$

The matrix $\mathbf{B}$ is used to write the boundary constraint equations. The matrix $\mathbf{B}^\dagger$ is the generalized inverse[8] of $\mathbf{B}$ normalized to have $\mathbf{B}^\dagger\mathbf{B} = \mathbf{I}$. The columns of the matrix $\mathbf{K}$ form a basis for the kernel of $\mathbf{B}^T$, i.e. $\mathbf{B}^T\mathbf{K} = 0$. The definitions of $\mathbf{B}$, $\mathbf{B}^\dagger$ and $\mathbf{K}$ are different for $\mathbf{n}_\alpha^- = \mathbf{n}_\alpha^+$ and $\mathbf{n}_\alpha^- \neq \mathbf{n}_\alpha^+$, and these two cases must be separately treated. We take $\mathbf{n}_\alpha^\pm$ as algebraic column vectors and the symbol "$\times$" to indicate the vector product in $\mathbb{R}^3$. We have:

- $\mathbf{n}_\alpha^- = \mathbf{n}_\alpha^+$;

$$\mathbf{B} = \begin{bmatrix} \tau \\ \mathcal{K}_\alpha \mathbf{n}_\alpha \end{bmatrix}, \quad \mathbf{B}^{\dagger,T} = \begin{bmatrix} 0 \\ \mathcal{K}_\alpha^{-1} \mathbf{n}_\alpha \end{bmatrix} \in \mathbb{R}^{3\times 1}, \quad \mathbf{K} = \begin{bmatrix} 0 & -1 \\ \mathcal{K}_\alpha^{-1}\mathbf{t}_\alpha & \tau\mathcal{K}_\alpha^{-1}\mathbf{n}_\alpha \end{bmatrix} \in \mathbb{R}^{3\times 2},$$

  where $\mathbf{n}_\alpha = \mathbf{n}_\alpha^\pm$ and $\mathbf{t}_\alpha = \mathbf{t}_\alpha^\pm$ indicate the (unique) normal and tangential vectors.

- $\mathbf{n}_\alpha^- \neq \mathbf{n}_\alpha^+$;

$$\mathbf{B} = \begin{bmatrix} \tau & \tau \\ \mathcal{K}_\alpha \mathbf{n}_\alpha^- & \mathcal{K}_\alpha \mathbf{n}_\alpha^+ \end{bmatrix}, \qquad \mathbf{B}^{\dagger,T} = \begin{bmatrix} 0 & 0 \\ \dfrac{\mathcal{K}_\alpha^{-1}\mathbf{t}_\alpha^+}{\mathbf{n}_\alpha^- \cdot \mathbf{t}_\alpha^+} & \dfrac{\mathcal{K}_\alpha^{-1}\mathbf{t}_\alpha^-}{\mathbf{n}_\alpha^+ \cdot \mathbf{t}_\alpha^-} \end{bmatrix} \in \mathbb{R}^{3\times 2},$$

$$\mathbf{K} = \left[\begin{pmatrix} \tau \\ \mathcal{K}_\alpha \mathbf{n}_\alpha^- \end{pmatrix} \times \begin{pmatrix} \tau \\ \mathcal{K}_\alpha \mathbf{n}_\alpha^+ \end{pmatrix} \right] \in \mathbb{R}^{3\times 1}.$$

  Note that the condition $\mathbf{n}_\alpha^\pm \cdot \mathbf{t}_\alpha^\mp \neq 0$ is true from the mutual orthogonality of these vectors.

4.1.2. *The reconstruction coefficients and the displacement factor*

Let $\mathbf{w}_\alpha = \left(\mathrm{w}_{\alpha,k}\right)^T$ be the algebraic vector containing the cell-centered weights, and $\mathbf{w}^b_\alpha = \left(\mathrm{w}^\pm_\alpha\right)^T$ be the algebraic vector containing the boundary edge weights.

Proposition 4.1. *Let $\mathcal{T}_h$ be a weakly acute triangulation. The least squares approximation of the solution value at v_α is given by*

$$u_\alpha = \sum_{k\in\sigma_\alpha} \mathrm{w}_{\alpha,k} u_k + \mathrm{w}^-_\alpha \mathrm{g}^-_\alpha + \mathrm{w}^+_\alpha \mathrm{g}^+_\alpha, \tag{4.3}$$

where

(i) *if v_α is an internal vertex,*

$$\mathbf{w}_\alpha = \boldsymbol{\Lambda}\mathbf{A}(\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A})^{-1}\begin{pmatrix}1\\ \mathbf{0}\end{pmatrix} \quad \textit{and} \quad \mathbf{w}^b_\alpha = \mathbf{0}; \tag{4.4}$$

(ii) *if v_α is a boundary vertex with a Robin-type condition,*

$$\mathbf{w}_\alpha = \boldsymbol{\Lambda}\mathbf{A}\mathbf{K}(\mathbf{K}^T\,\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}\,\mathbf{K})^{-1}\mathbf{K}^T\begin{pmatrix}1\\ \mathbf{0}\end{pmatrix}, \tag{4.5a}$$

and

$$\mathbf{w}^b_\alpha = \begin{cases} \mathbf{B}^\dagger(\mathbf{I} - \mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}\mathbf{K}(\mathbf{K}^T\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}\mathbf{K})^{-1}\mathbf{K}^T)\begin{pmatrix}1\\ \mathbf{0}\end{pmatrix} & \textit{if } \mathbf{n}^-_\alpha \neq \mathbf{n}^+_\alpha, \\ \frac{1}{2}\mathbf{B}^\dagger(\mathbf{I} - \mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}\mathbf{K}(\mathbf{K}^T\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}\mathbf{K})^{-1}\mathbf{K}^T)\begin{pmatrix}1\\ \mathbf{0}\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix} & \textit{if } \mathbf{n}^-_\alpha = \mathbf{n}^+_\alpha; \end{cases} \tag{4.5b}$$

(iii) *if v_α is a boundary vertex with a Dirichlet-type condition, $\mathbf{w}_\alpha = \mathbf{0}$ and $\mathbf{w}^b_\alpha = (1/2, 1/2)^T$.*

The proof of this proposition is postponed after the two technical lemmas that follow below. The first lemma describes two geometrical facts that are consequence of the assumption that the mesh is weakly acute. These facts are used to demonstrate the second lemma.

Lemma 4.2. *Let $\mathcal{T}_h$ be a weakly acute triangulation. Then,*

(i) *if v_α is an internal vertex the barycenters of the triangular cells of the stencil σ_α are not collinear;*
(ii) *if v_α is a boundary vertex and $\mathbf{n}^-_\alpha = \mathbf{n}^+_\alpha$, the barycenters of the triangular cells of the stencil σ_α are not aligned with $\boldsymbol{n}^+_\alpha$.*

Proof. Both items are proved by contradiction arguments.

(i) Without loss of generality, we take the vertex v_α coincident with the origin, i.e. $\mathbf{x}_\alpha = \mathbf{0}$, and assume that all the barycenters of the triangles T_k for $k \in \sigma_\alpha$ are on the same vertical line in the strictly negative half-plane. As the triangulation is weakly acute and v_α is an internal vertex, there must exist a triangle in the stencil σ_α that is entirely contained in the positive half-plane. This triangle contains the vertex v_α which coincides with the origin and at least another vertex with a strictly positive abscissa. Therefore, its barycenter must be located in the positive half-plane and cannot be on a line in the strictly negative half-plane thus resulting in the contradiction.

(ii) As $\mathbf{n}_\alpha^- = \mathbf{n}_\alpha^+$, we can consider the triangles incident to $\mathsf{v}_\alpha \in \mathcal{V}_h^{\mathrm{bnd}}$ and the triangles built by specular reflection with respect to the boundary line. If (ii) was true, all these triangles would have aligned gravity centers, thus violating (i). □

Note that Lemma 4.2(ii) does not prevent the barycenters of the triangles incident to a boundary vertex to be aligned, but just to be aligned along a direction that is orthogonal to the boundary.

Lemma 4.3. *Let $\mathcal{T}_h$ be a weakly acute triangulation. Then,*

(i) *if* $\mathsf{v}_\alpha \in \mathcal{V}_h^{\mathrm{int}}$, $\mathsf{rank}\,\mathbf{A} = 3$;
(ii) *if* $\mathsf{v}_\alpha \in \mathcal{V}_h^{\mathrm{bnd}}$, $\mathsf{ker}\,(\mathbf{B}^T) \cap \mathsf{ker}\,(\mathbf{A}) = \{\mathbf{0}\}$.

Proof. (i) As the mesh is weakly acute, the condition $\mathsf{card}\,\{\sigma_\alpha\} \geq 4$ must hold and, in view of Lemma 4.2, at least three triangles in σ_α must have non-collinear barycenters. Therefore, the matrix $\mathbf{A}$ has at least one non-null 3×3-minor.

(ii) Let us note that condition $\mathsf{ker}\,(\mathbf{B}^T) \cap \mathsf{ker}\,(\mathbf{A}) = \{\mathbf{0}\}$ is equivalent to $\mathsf{range}\,(\mathbf{B}) + \mathsf{range}\,(\mathbf{A}^T) = \{\mathbb{R}^3\}$, and the latter to

$$\mathsf{rank}\begin{pmatrix}\mathbf{A}\\ \mathbf{B}^T\end{pmatrix} = 3.$$

We will show that the equivalent condition on the matrix rank is always satisfied if $\mathbf{A}$ and $\mathbf{B}$ are built on a weakly acute triangulation. If $\mathsf{rank}\,\mathbf{A} = 3$ then (ii) is trivially true. Let us assume that $\mathsf{rank}\,\mathbf{A} < 3$, and distinguish the two following sub cases.

- $\mathbf{n}_\alpha^- = \mathbf{n}_\alpha^+$. Then, the boundary is "locally flat" near v_α, and $\mathsf{card}\,\{\sigma_\alpha\} \geq 2$ because the mesh is weakly acute. In view of Lemma 4.2, there exist at least two triangles in σ_α with barycenters $\mathbf{x}_{k_1}$ and $\mathbf{x}_{k_2}$ not aligned with $\mathbf{n}_\alpha^+ =$

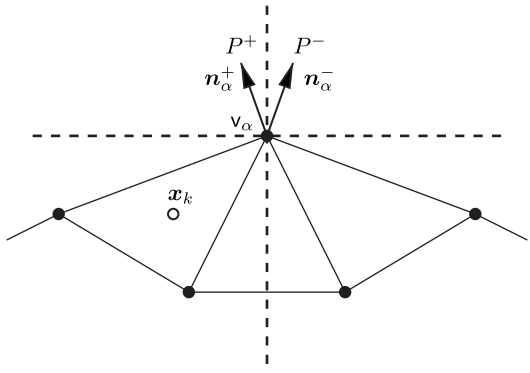


Fig. 3. Proof of Lemma 4.3; the picture exemplifies the case $\mathbf{n}_\alpha^- \neq \mathbf{n}_\alpha^+$; note that $\mathbf{x}_k$ does not belong to the same half-plane of the points $P^\pm$ by construction and, hence, cannot be aligned to them.

$(n_{\alpha,x}^+, n_{\alpha,y}^+)^T$. Thus, the minor

$$\begin{vmatrix} 1 & \mathsf{x}_{k_1} - \mathsf{x}_\alpha & \mathsf{y}_{k_2} - \mathsf{y}_\alpha \\ 1 & \mathsf{x}_{k_1} - \mathsf{x}_\alpha & \mathsf{y}_{k_2} - \mathsf{y}_\alpha \\ \tau & n_{\alpha,x}^+ & n_{\alpha,y}^+ \end{vmatrix}$$

is non-null.

- $\mathbf{n}_\alpha^- \neq \mathbf{n}_\alpha^+$. Let us take the two vectors $\mathbf{n}_\alpha^\pm$ centered at v_α and denote their other extremal points by P^+ and P^- as depicted in Fig. 3. As the mesh is weakly acute, there must exist a triangle T_k that is incident to v_α and entirely contained within the half-plane delimited by the orthogonal line to the bisector of the angle between $\mathbf{n}_\alpha^\pm$ and not containing P^+ and P^-. The gravity center of T_k, i.e. $\mathbf{x}_k$, cannot be aligned to P^+ and P^-, because it belongs to a different half-plane (see Fig. 3); thus, the minor

$$\begin{vmatrix} 1 & \mathsf{x}_k - \mathsf{x}_\alpha & \mathsf{y}_k - \mathsf{y}_\alpha \\ \tau & n_{\alpha,x}^+ & n_{\alpha,y}^+ \\ \tau & n_{\alpha,x}^- & n_{\alpha,y}^- \end{vmatrix}$$

is non-null. □

Proof of Proposition 4.1. For compactness of notation, we introduce the vectors $\boldsymbol{\zeta} = (a, \mathbf{b}^T)^T$ and $\boldsymbol{\xi}(\boldsymbol{\zeta}) = \mathbf{A}\boldsymbol{\zeta} - \mathbf{u}_\alpha$; note that $\nabla_{\boldsymbol{\zeta}}\boldsymbol{\xi}(\boldsymbol{\zeta}) = \mathbf{A}$. For the sake of reference, we rewrite Eq. (4.3) as

$$u_\alpha = \mathbf{w}_\alpha^T \mathbf{u}_\alpha + \mathbf{w}_\alpha^{b,T} \mathbf{g}_\alpha, \tag{4.6}$$

and evaluate the reconstructed value by $u_\alpha = a = (1, \mathbf{0}^T)\boldsymbol{\zeta}$ from (4.1). The proof of Proposition 4.1 only takes into consideration the cases of an internal vertex

and a Robin boundary vertex, because the case of a Dirichlet boundary vertex is obvious.

(i) *Internal vertex case*
The unknown vector of the reconstruction coefficients $\boldsymbol{\zeta} = (a, \mathbf{b})^T$ is the minimizer of the least squares functional

$$\mathcal{J}(\boldsymbol{\zeta}) = \frac{1}{2}\boldsymbol{\xi}(\boldsymbol{\zeta})^T \boldsymbol{\Lambda}\boldsymbol{\xi}(\boldsymbol{\zeta}). \tag{4.7}$$

We impose the null gradient condition on the least squares functional (4.7), i.e.

$$\nabla_{\boldsymbol{\zeta}}\mathcal{J}(\boldsymbol{\zeta}) = (\nabla_{\boldsymbol{\zeta}}\boldsymbol{\xi}(\boldsymbol{\zeta}))^T \boldsymbol{\Lambda}\xi(\boldsymbol{\zeta}) = \mathbf{A}^T\boldsymbol{\Lambda}\,(\mathbf{A}\boldsymbol{\zeta} - \mathbf{u}_\alpha) = \mathbf{0},$$

in order to obtain the algebraic system of the *weighted normal equations*

$$\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}\,\boldsymbol{\zeta} = \mathbf{A}^T\boldsymbol{\Lambda}\mathbf{u}_\alpha. \tag{4.8}$$

It turns out that the 3×3 real matrix $\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}$ of the left-hand side of (4.8) is non-singular. To see this, we first note that the matrix $\mathbf{A}$ has maximum rank, i.e. $\mathsf{rank}\,\mathbf{A} = 3$, due to the mesh regularity assumption and Lemma 4.3. A standard result of matrix theory implies that $\mathsf{rank}\,(\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}) = \mathsf{rank}\,\mathbf{A}$; thus, $\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A}$, which is a square matrix of order 3, has rank 3 and is non-singular. The solution of (4.8) is formally written as

$$\boldsymbol{\zeta} = (\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{A})^{-1}\mathbf{A}^T\boldsymbol{\Lambda}\mathbf{u}_\alpha. \tag{4.9}$$

Finally, we set the boundary weights $\mathbf{w}_\alpha^b$ to zero because no boundary data are to be taken into account for an internal vertex, and derive the formula (4.4) for the weights $\mathbf{w}_\alpha$ by comparing (4.9) and (4.6).

(ii) *Robin boundary vertex case*
We introduce the two boundary constraints (4.2) in the least squares functional by using the Lagrangian multipliers $\boldsymbol{\mu} = (\mu^-, \mu^+)^T$. We also consider the special solution of the linear algebraic problem $\mathbf{B}^T\hat{\boldsymbol{\zeta}} = \mathbf{g}_\alpha$ that is given by $\hat{\boldsymbol{\zeta}} = (\hat{a}, \hat{\mathbf{b}})^T = \mathbf{B}^{\dagger,T}\mathbf{g}_\alpha$. The vector of the reconstruction coefficients $\boldsymbol{\zeta} = (a, \mathbf{b})^T$ and the vector of the Lagrangian multipliers $\boldsymbol{\mu}$ are the minimizers of the constrained least squares functional

$$\mathcal{J}(\boldsymbol{\zeta}, \boldsymbol{\mu}) = \frac{1}{2}\boldsymbol{\xi}(\boldsymbol{\zeta})^T \boldsymbol{\Lambda}\boldsymbol{\xi}(\boldsymbol{\zeta}) + \boldsymbol{\mu}^T(\mathbf{B}^T\boldsymbol{\zeta} - \mathbf{g}_\alpha).$$

By imposing the null gradient conditions

$$\begin{aligned}
\nabla_{\boldsymbol{\zeta}}\mathcal{J}(\boldsymbol{\zeta}, \boldsymbol{\mu}) &= (\nabla_{\boldsymbol{\zeta}}\boldsymbol{\xi}(\boldsymbol{\zeta}))^T \boldsymbol{\Lambda}\xi(\boldsymbol{\zeta}) + \mathbf{B}\boldsymbol{\mu} = \mathbf{0}, \\
\nabla_{\boldsymbol{\mu}}\mathcal{J}(\boldsymbol{\zeta}, \boldsymbol{\mu}) &= \mathbf{B}^T\boldsymbol{\zeta} - \mathbf{g}_\alpha = \mathbf{0},
\end{aligned}$$

we obtain the augmented algebraic system

$$\begin{bmatrix} \mathbf{A}^T\mathbf{\Lambda}\mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{\Lambda}\mathbf{u}_\alpha \\ \mathbf{g}_\alpha \end{bmatrix}. \tag{4.10}$$

We formally write the solution to (4.10) by the null space technique. We substitute the general expression of solution $\boldsymbol{\zeta} = \hat{\boldsymbol{\zeta}} + \mathbf{K}\boldsymbol{\rho}$ into (4.10) and project onto the null-space of $\mathbf{B}^T$ in order to derive the reduced algebraic system

$$\mathbf{K}^T\,\mathbf{A}^T\mathbf{\Lambda}\mathbf{A}\,\mathbf{K}\,\boldsymbol{\rho} = \mathbf{K}^T\,\mathbf{A}^T\mathbf{\Lambda}(\mathbf{u}_\alpha - \mathbf{A}\hat{\boldsymbol{\zeta}}) \tag{4.11}$$

for the vector $\boldsymbol{\rho}$. Mesh regularity assumption and Lemma 4.3 imply that

$$\ker(\mathbf{B}^T) \cap \ker(\mathbf{A}) = \{\mathbf{0}\}.$$

Therefore, the matrix $\mathbf{K}^T\mathbf{A}^T\mathbf{\Lambda}\mathbf{A}\mathbf{K}$ on the left-hand side of (4.11) is non-singular and the vector $\boldsymbol{\rho}$ can be formally written as

$$\boldsymbol{\rho} = (\mathbf{K}^T\,\mathbf{A}^T\mathbf{\Lambda}\mathbf{A}\,\mathbf{K})^{-1}\mathbf{K}^T\mathbf{A}^T\mathbf{\Lambda}(\mathbf{u}_\alpha - \mathbf{A}\hat{\boldsymbol{\zeta}}). \tag{4.12}$$

Then, we substitute (4.12) for $\boldsymbol{\rho}$ and use $\hat{\boldsymbol{\zeta}} = \mathbf{B}^{\dagger,T}\mathbf{g}_\alpha$ in $\boldsymbol{\zeta} = \hat{\boldsymbol{\zeta}} + \mathbf{K}\boldsymbol{\rho}$ to get

$$\begin{aligned}\boldsymbol{\zeta} = {} & \mathbf{K}(\mathbf{K}^T\big(\mathbf{A}^T\mathbf{\Lambda}\mathbf{A}\big)\mathbf{K})^{-1}\mathbf{K}^T\mathbf{A}^T\mathbf{\Lambda}\mathbf{u}_\alpha \\ & + [\mathbf{I} - \mathbf{K}(\mathbf{K}^T(\mathbf{A}^T\mathbf{\Lambda}\mathbf{A})\mathbf{K})^{-1}\mathbf{K}^T\mathbf{A}^T\mathbf{\Lambda}\mathbf{A}]\mathbf{B}^{\dagger,T}\mathbf{g}_\alpha.\end{aligned}$$

Finally, we derive formulae (4.5a)–(4.5b) by comparing this last expression to (4.6).

(iii) *Dirichlet boundary vertex case*
Obvious. □

4.1.3. *Remarks on the reconstruction process*

Remark 4.4. A simple calculation allows the reader to verify that the coefficient vectors $\mathbf{w}_\alpha$ and $\mathbf{w}_\alpha^b$ are the unique solution of

$$\mathbf{A}^T\mathbf{w}_\alpha + \tilde{\mathbf{B}}^T\mathbf{w}_\alpha^b = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \tag{4.13}$$

where $\tilde{\mathbf{B}}$ is the generalized inverse[8] of $\mathbf{B}^\dagger$ normalized to have $\mathbf{B}^{\dagger,T}\tilde{\mathbf{B}} = \mathbf{I}$. Equation (4.13) is consistent with both requirements that the reconstruction be exact for linear polynomials on the co-volume region $\mathcal{V}_\alpha$ and have a constrained gradient on the Robin boundary edges. For an internal vertex, relation (4.13) can be simplified as

$$\mathbf{A}^T\mathbf{w}_\alpha = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \tag{4.14}$$

which is still expression of the linear consistency of the interpolation formula using the weights $\mathbf{w}_\alpha$.

Remark 4.5. The algebraic problem resulting from the weighted normal equations may be ill-conditioned when the mesh is highly stretched.[4,7] As the sensitivity of the solution of both (4.8) and (4.10) formally depends on the square of the condition number of $\mathbf{A}$, a loss of accuracy may occur in the computation of least squares weights. This situation takes place when the gravity centers of the triangles are nearly collinear, that is, when a situation of numerical rank deficiency occurs in the resolution of least squares problems. A possible strategy to overcome this difficulty consists in enforcing the regularity requirements on the computational grid by a *minimal angle condition* at the mesh generation level. This makes it possible to avoid excessive mesh stretching and reduce the alignment of control volume centers. A further improvement[4,7] consists in applying the orthogonal decomposition $\mathbf{A} = \mathbf{QR}$, where $\mathbf{Q}$ is the orthogonal matrix and $\mathbf{R}$ the upper triangular matrix of the Gram–Schmidt process.

4.2. *Alternative reconstruction algorithms*

In this subsection, we review the formulation of several algorithms from the literature for determining the cell weights $\{\mathrm{w}_{\alpha,k}\}$ of the vertex reconstruction formula:

$$u_\alpha = \sum_{k\in\sigma_\alpha} \mathrm{w}_{\alpha,k} u_k. \tag{4.15}$$

We shall apply (4.15) to all mesh vertices without any particular care of the boundary ones. Possible choices of reconstruction weights $\{\mathrm{w}_{\alpha,k}\}_{k\in\sigma_\alpha}$ are given by[4,22,27,28]:

- $\mathrm{w}_{\alpha,k} = 1/\mathsf{N}_\alpha$;
- $\mathrm{w}_{\alpha,k} = |\mathbf{x}_k - \mathbf{x}_\alpha|^{-1} \Big/ \sum\limits_{k\in\sigma_\alpha} |\mathbf{x}_k - \mathbf{x}_\alpha|^{-1}$;
- $\mathrm{w}_{\alpha,k} = |\mathsf{T}_k| \,/ \sum\limits_{k\in\sigma_\alpha} |\mathsf{T}_k|$.

To study the order of formal consistency of the interpolation formula (4.15) with the above cell weights, we first rewrite condition (4.14) in the form of the explicit relations:

$$\sum_{k\in\sigma_\alpha} \mathrm{w}_{\alpha,k} = 1, \tag{4.16a}$$

$$\sum_{k\in\sigma_\alpha} \mathrm{w}_{\alpha,k}(\mathbf{x}_k - \mathbf{x}_\alpha) = 0. \tag{4.16b}$$

If the cell weights satisfy relation (4.16a), the vertex interpolation formula (4.15) is formally exact for constant functions, and we will say that these weights are

first-order accurate. This condition is sufficient to make the finite volume method at least first-order accurate. If, in addition, the cell weights satisfy relation (4.16b), the vertex interpolation formula (4.15) is formally exact for linear functions, and we will say that these weights are second-order accurate. The simultaneous satisfaction of (4.16a)–(4.16b) is enough to make the finite volume method linearly consistent, and, thus, to achieve formal second order of accuracy.

It is easy to check that all the weights listed above only verify (4.16a), so that the reconstruction formula (4.15) is at most first-order accurate. However, cell weights providing second-order accuracy can be constructed from first-order ones by the following strategy. Let $\{W^{\mathrm{1st}}_{\alpha,k}\}$ indicate a set of first-order accurate weights, and k_1 and k_2 be two indices in σ_α such that the vectors $(\mathbf{x}_\alpha - \mathbf{x}_{k_1})$ and $(\mathbf{x}_\alpha - \mathbf{x}_{k_2})$ are not parallel. In order to get condition (4.16b) satisfied, we solve the algebraic system of the two linear equations for the unknowns s_1 and s_2 that arises from the vector relation:

$$(\mathbf{x}_\alpha - \mathbf{x}_{k_1})s_1 + (\mathbf{x}_\alpha - \mathbf{x}_{k_2})s_2 = \sum_{\substack{k\in\sigma_\alpha,\\ k\neq k_1,k_2}} W^{\mathrm{1st}}_{\alpha,k}(\mathbf{x}_k - \mathbf{x}_\alpha).$$

Then, we evaluate the normalization factor

$$W^{\mathrm{2nd}}_{\alpha,\mathrm{TOT}} = s_1 + s_2 + \sum_{\substack{k\in\sigma_\alpha,\\ k\neq k_1,k_2}} W^{\mathrm{1st}}_{\alpha,k},$$

and, for every k in σ_α, we set

$$W^{\mathrm{2nd}}_{\alpha,k} = \frac{1}{W^{\mathrm{2nd}}_{\alpha,\mathrm{TOT}}} \times \begin{cases} W^{\mathrm{1st}}_{\alpha,k} & \text{for } k \neq k_1,\, k_2, \\ s_1 & \text{for } k = k_1, \\ s_2 & \text{for } k = k_2. \end{cases}$$

As can be readily seen, the new set of weights $\{W^{\mathrm{2nd}}_{\alpha,k}\}$ satisfies both conditions (4.16a)–(4.16b), thus achieving the sought linear consistency in the reconstruction formula (4.15). It is worth noting that the choice of the couple of indices k_1 and k_2 is not unique. A possible strategy for selecting them is as follows:

(a) we select the index k_1 that maximizes $\left|W^{\mathrm{1st}}_{\alpha,k}\right|$ over the index set σ_α;
(b) we select the index $k_2 \neq k_1$ that maximizes $\left|W^{\mathrm{1st}}_{\alpha,k}\right|$ over the index set $\sigma_\alpha \backslash \{k_1\}$;
(c) if the vectors $(\mathbf{x}_\alpha - \mathbf{x}_{k_1})$ and $(\mathbf{x}_\alpha - \mathbf{x}_{k_2})$ are not aligned we stop the procedure; otherwise, we discard k_2 from σ_α and repeat the procedure from step (b).

Other procedures may be envisaged to determine k_1 and k_2, such as, for example, searching the indices that minimize the absolute value of the weights, or the indices providing the greatest and the smallest absolute values, or again randomly choosing two distinct indices within σ_α, and so on. Regarding this issue, we

experimentally tested a number of variants without noting significant differences in the final performance of the finite volume method.

5. Numerical Experiments

The performance of our finite volume method for different vertex reconstructions is experimentally investigated for the numerical resolution of the elliptic problem (1.1a)–(1.1c) with the anisotropic diffusion tensor

$$\mathcal{K} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \tag{5.1}$$

The diffusion tensor $\mathcal{K}$ depends on two parameters: the rotation angle θ and the anisotropy ratio ε. The rotation angle θ controls the orientation of the principal directions of anisotropy, i.e. the eigenvectors of $\mathcal{K}$, which are parallel to the axis of an orthogonal reference frame rotated by this angle with respect to the coordinate directions. The anisotropy ratio ε, which is the ratio between the eigenvalues of $\mathcal{K}$, controls the relative strength of diffusion along the principal directions of anisotropy. In this section, we present the numerical results obtained for $\theta = \pi/6$ and $\varepsilon = 1$, 10^{-2}, 10^{-4}, which are well representative of the performance of our finite volume method.

We consider the following three cases:

Case I: Dirichlet boundary data with $\Gamma^{\mathrm{Rob}} = \emptyset$ and

$$\mathsf{u} = \mathsf{g}^D \text{ on } \Gamma^{\mathrm{Dir}} = \{(x,y) \,|\, x = 0, 1 \text{ or } y = 0, 1\};$$

Case II: mixed Dirichlet and Robin boundary data with

$$\begin{aligned} \mathsf{u} &= \mathsf{g}^D \text{ on } \Gamma^{\mathrm{Dir}} = \{(x,y) \,|\, x = 0 \text{ or } y = 0\}; \\ \mathsf{u} + \mathbf{n} \cdot \mathcal{K} \nabla \mathsf{u} &= \mathsf{g}^R \text{ on } \Gamma^{\mathrm{Rob}} = \{(x,y) \,|\, x = 1 \text{ or } y = 1\}; \end{aligned}$$

Case III: mixed Dirichlet and Neumann boundary data with

$$\begin{aligned} \mathsf{u} &= \mathsf{g}^D \text{ on } \Gamma^{\mathrm{Dir}} = \{(x,y) \,|\, x = 0 \text{ or } y = 0\}; \\ \mathbf{n} \cdot \mathcal{K} \nabla \mathsf{u} &= \mathsf{g}^R \text{ on } \Gamma^{\mathrm{Rob}} = \{(x,y) \,|\, x = 1 \text{ or } y = 1\}. \end{aligned}$$

Case I allows us to measure and compare the influence of different algorithms for the *internal* vertex reconstruction on the performance of the finite volume scheme with an assigned value to the boundary vertices. Instead, Case II and Case III make it possible to investigate the effect of the boundary vertex treatment on the scheme perfomance for the Robin and Neumann case, respectively. It is worth noting that, in Cases II and III, the value that is reconstructed at a boundary vertex affects the

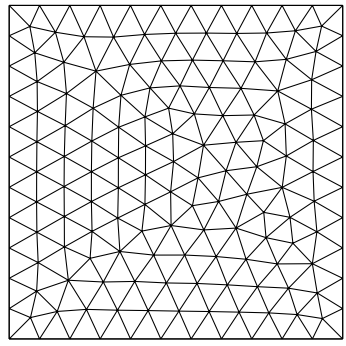


Fig. 4. Base computational grid for numerical experiments formed by 272 triangular cells, 430 edges, 159 vertices, and 44 boundary edges and vertices; all angles are strictly acute.

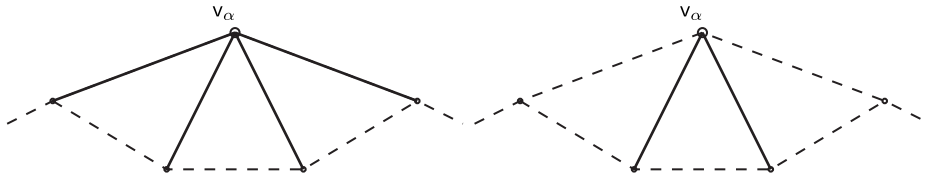


Fig. 5. The numerical flux of the edges in bold line takes a contribution from the reconstructed value of the boundary vertex v_α, which is assigned a Robin boundary condition on the left and a Neumann boundary condition on the right.

evaluation of the numerical flux of each internal edge incident to this vertex in a very different way. This situation is illustrated in Fig. 5, and the features of these three test cases are summarized as follows:

- in Case I, only Dirichlet boundary conditions are considered and the approximate value of each boundary vertex is derived from a direct evaluation of the boundary function g^D at the vertex position;
- in Case II, we set $\tau = 1$ in the third equation of (3.3), and, consequently, the vertex value u_α (respectively, u_β) gives a contribution to the numerical flux of all the edges incident to v_α (respectively, v_β) including the *boundary* edges;
- in Case III, we set $\tau = 0$ in the third equation of (3.3), and u_α (respectively, u_β) can only give contribution to the numerical flux of the *internal* edges incident to v_α (respectively, v_β).

Finally, in order to compare the numerical and the analytical solution, the source term and the boundary functions g^D and g^R are specialized so that the exact solution to model problem (1.1a)–(1.1c) with conductivity tensor (5.1) is given by

$$\mathsf{u}(x, y) = \sin(2\pi x)\sin(2\pi y) + x^2 + y^2 + 1$$

on the domain $\Omega = [0, 1] \times [0, 1]$.

5.1. *Evaluation of the scheme performance*

The scheme performance is evaluated by measuring the convergence rate of the numerical approximation on a sequence of successively refined meshes starting from the base grid shown in Fig. 4. The base grid consists of 272 triangular cells, 430 edges, 159 vertices, and 44 boundary edges and vertices; all its angles are acute and, therefore, it satisfies the mesh regularity assumption. The mesh data structures are managed by P2MESH,[12] a C++ public domain library designed for fast and efficient implementation of partial differential equation solvers. The refinement process for all grids progresses by connecting together the midpoints of the edges of each triangle. This ensures that the mesh size parameter h is halved at every level of refinement and that the newly generated triangles maintain the same aspect ratio of the original ones. Moreover, all the derived grids are also regular because no new angles are generated in this process.

The rate of convergence is measured by comparing relative errors at two consecutive mesh levels with mesh parameters h and $h/2$. We found it convenient to measure the relative error for the approximation of cell averages by using L^2 norms:

$$\mathcal{E}_h = \frac{\|\mathsf{u}_h - \mathcal{A}_h(\mathsf{u})\|_h}{\|\mathsf{u}\|_{L^2(\Omega)}} = \frac{\left(\sum_{\mathsf{T}_i \in \mathcal{T}_h} |\mathsf{T}_i| \, |u_i - \mathcal{A}_i(\mathsf{u})|^2\right)^{\frac{1}{2}}}{\|\mathsf{u}\|_{L^2(\Omega)}},$$

and the relative error for the approximation of cell-averaged gradients by using H^1 semi-norms:

$$\mathcal{E}_{\mathsf{G},h} = \frac{\|\boldsymbol{\mathcal{G}}_h(\mathsf{u}_h) - \mathcal{A}_h(\nabla \mathsf{u})\|_h}{|\mathsf{u}|_{H^1(\Omega)}} = \frac{\left(\sum_{\mathsf{T}_i \in \mathcal{T}_h} |\mathsf{T}_i| \, |\boldsymbol{\mathcal{G}}_i - \mathcal{A}_i(\nabla \mathsf{u})|^2\right)^{\frac{1}{2}}}{|\mathsf{u}|_{H^1(\Omega)}}. \tag{5.2}$$

The error $\mathcal{E}_{\mathsf{G},h}$ is defined in (5.2) by comparing the piecewise-constant finite volume gradient $\boldsymbol{\mathcal{G}}_h(\mathsf{u}_h) = \{\boldsymbol{\mathcal{G}}_i\} \in \mathbb{P}^0(\mathcal{T}_h) \times \mathbb{P}^0(\mathcal{T}_h)$ with the cell average of the solution gradient $\mathcal{A}_h(\nabla \mathsf{u}) = (\mathcal{A}_h(\partial \mathsf{u}/\partial x), \mathcal{A}_h(\partial \mathsf{u}/\partial y)) \in \mathbb{P}^0(\mathcal{T}_h) \times \mathbb{P}^0(\mathcal{T}_h)$. The value of $\boldsymbol{\mathcal{G}}_i$ is calculated by taking the gradient of the linear function interpolating the numerical solution at the three vertices of T_i.

By *second-order convergence rate* we imply an $\mathcal{O}(h^2)$ decrease of the relative error $\mathcal{E}_h$ of the cell-average approximation, and an $\mathcal{O}(h)$ decrease of the relative error $\mathcal{E}_{\mathsf{G},h}$ of the gradient approximation.

5.2. *Weights used in vertex reconstructions*

We denote the reconstruction based on the least squares weights $\{W^{\mathrm{LS}}_{\alpha,k}]\} = \{W_{\alpha,k}, W^b_{\alpha,k}]\}$ by $\mathcal{R}^{\mathrm{LS}}(\cdot)$ (Sec. 4.1), the reconstruction based on the weights $\{W^{\mathrm{1st}}_{\alpha,k}\} = \{|\mathsf{T}_k| \, / \sum_{j \in \sigma_\alpha} |\mathsf{T}_j|\}$ by $\mathcal{R}^{\mathrm{1st}}(\cdot)$, and the reconstruction based on the second-order weights $\{W^{\mathrm{2nd}}_{\alpha,k}\}$ by $\mathcal{R}^{\mathrm{2nd}}(\cdot)$ (Sec. 4.2). As the reconstruction algorithm is readily identified by the set of weights that is actually in use, we will also

refer to the cell weights by the same symbol $\mathcal{R}^*(\cdot)$, where $* = \mathrm{LS}, 1\mathrm{st}, 2\mathrm{nd}$. The "dot" argument denotes the set of vertices to which the reconstruction algorithm is applied: $\mathcal{V}_h^{\mathrm{int}}$ for internal vertices and $\mathcal{V}_h^{\mathrm{bnd}}$ for boundary vertices. For example, $\mathcal{R}^{\mathrm{LS}}(\mathcal{V}_h^{\mathrm{bnd}})$ indicates both the reconstruction algorithm and the set of weights based on least squares algorithm and applied to the boundary vertices.

The results of the numerical experiments presented in Sec. 5.3 are obtained by using variants of the finite volume method based on the following combinations of reconstruction algorithms for internal and boundary vertices of Robin and Neumann type:

— $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$, which uses $\{W_{\alpha,k}^{\mathrm{LS}}\}$ for both internal and boundary vertices;
— $\mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$, which uses $\{W_{\alpha,k}^{1\mathrm{st}}\}$ for both internal and boundary vertices;
— $\mathcal{R}^{2\mathrm{nd}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{2\mathrm{nd}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$, which uses $\{W_{\alpha,k}^{2\mathrm{nd}}\}$ for both internal and boundary vertices;
— $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$, which uses $\{W_{\alpha,k}^{\mathrm{LS}}\}$ for internal vertices and $\{W_{\alpha,k}^{1\mathrm{st}}\}$ for boundary vertices;
— $\mathcal{R}^{2\mathrm{nd}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$, which uses $\{W_{\alpha,k}^{2\mathrm{nd}}\}$ for internal vertices and $\{W_{\alpha,k}^{1\mathrm{st}}\}$ for boundary vertices.

All these reconstruction algorithms except $\mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ are linearly consistent on internal vertices. Thus, we expect that the corresponding finite volume implementations display second-order convergence in Case I where no reconstruction is required for the numerical treatment of the boundary. On the other hand, only the reconstructions $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ and $\mathcal{R}^{2\mathrm{nd}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{2\mathrm{nd}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ are linearly consistent on the entire mesh vertex set. Thus, we expect that only the finite volume implementations based on them provide second-order accurate numerical approximations to the solutions of Cases II and III. In fact, a deterioration of the global accuracy of the finite volume approximation may occur when the reconstruction of boundary vertices is less accurate than the reconstruction of internal vertices. We will investigate this effect by using the two combinations $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ and $\mathcal{R}^{2\mathrm{nd}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$. Finally, $\mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{int}}\big) - \mathcal{R}^{1\mathrm{st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$, which is first-order accurate for all the vertices, is expected to produce a first-order accurate finite volume method and is included in our investigation list for the sake of comparison.

5.3. *Analysis of numerical results*

Figures 6–8 illustrate the convergence rates of the finite volume approximations produced by the various combinations of reconstruction algorithms of Sec. 5.2. These three figures graphically present the log–log plots of the relative error norms ($\mathcal{E}_h$ on the left, $\mathcal{E}_{\mathsf{G},h}$ on the right) versus the characteristic mesh size h. The "real" order of accuracy is reflected in the figures by the slopes of the experimental curves,

and can be approximately evaluated by comparison with the "theoretical" first- and second-order slopes in the bottom-left corner of each graph.

Case I: Dirichlet boundary conditions
The convergence curves for anisotropy ratios $\varepsilon = 1$, 10^{-2} and 10^{-4} are shown from top to bottom in Fig. 6. From these plots the following facts are experimentally observed:

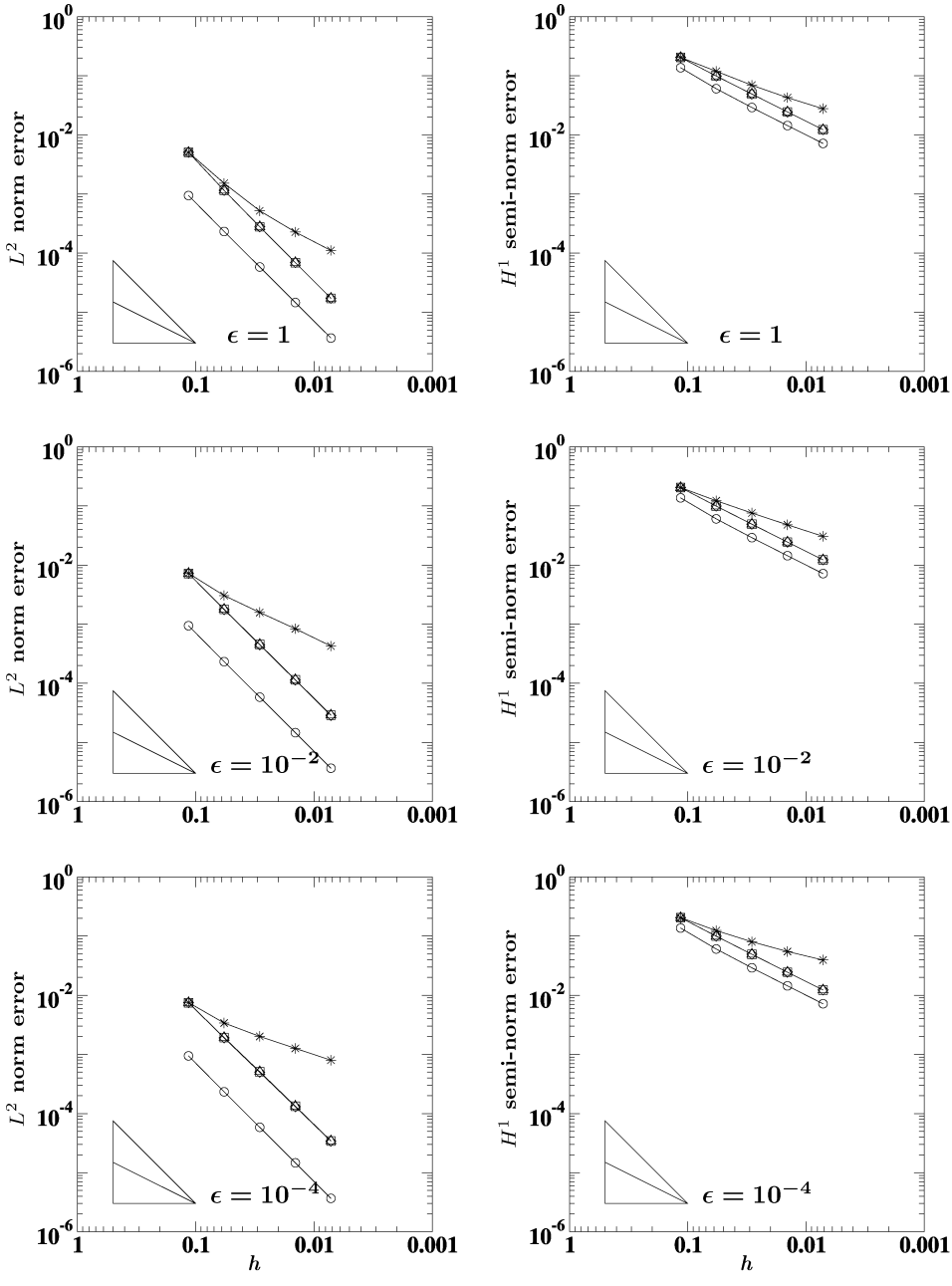
(i) all the finite volume methods except the one using $\mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ show second-order convergence rates;
(ii) the error curves of the scheme using $\mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ show a linear convergence rate in the log–log plots of $\mathcal{E}_h$ and a sub-linear convergence rate in the log–log plots of $\mathcal{E}_{\mathsf{G},h}$;
(iii) the error curves for both $\mathcal{E}_h$ and $\mathcal{E}_{\mathsf{G},h}$ are independent of the values of the anisotropy ratio ε in the range $[10^{-4}, 1]$;
(iv) for the $\mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{int}}\big)$–$\mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ scheme, the $\mathcal{E}_h$ errors are one order of magnitude smaller than the corresponding errors of the other schemes; for the gradients this reduction decreases to a factor of two.

The expectations reported in Sec. 5.2 that all the methods based on linearly accurate reconstructions for internal vertices provide second-order convergence is confirmed by the numerical results. Note that this behavior is independent of the anisotropy parameter ε. The numerical experiments show that the magnitude of the errors for a fixed h is strongly dependent on the reconstruction algorithm, in particular for cell-averages. The best numerical approximation is indeed provided by the finite volume scheme using least squares weights.

Case II: Mixed Dirichlet and Robin boundary conditions
In this case, we evaluate the influence of the different reconstructions of boundary vertices on the convergence behavior of the finite volume scheme (Fig. 7). From these plots the following facts are experimentally observed:

(i) the finite volume methods using $\mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ and $\mathcal{R}^{2\text{nd}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{2\text{nd}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ show second-order convergence rate;
(ii) the $\mathcal{E}_h$ and $\mathcal{E}_{\mathsf{G},h}$ error curves for $\mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ and $\mathcal{R}^{2\text{nd}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ are very close to the corresponding error curves for $\mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{1\text{st}}\big(\mathcal{V}_h^{\text{bnd}}\big)$, which manifest first-order accuracy;
(iii) both the relative errors $\mathcal{E}_h$ and $\mathcal{E}_{\mathsf{G},h}$ of the scheme using $\mathcal{R}^{2\text{nd}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{2\text{nd}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ increase when ε decreases from 1 to 10^{-4}, while the corresponding relative errors for $\mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{int}}\big) - \mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ are independent of ε.

As is clearly visible in Fig. 7 and pointed out by items (i)–(iii), the reconstruction of Robin vertices has a strong impact on the accuracy of these finite volume approximations. Optimal convergence rates are attained by the scheme using $\mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{int}}\big)$–$\mathcal{R}^{\text{LS}}\big(\mathcal{V}_h^{\text{bnd}}\big)$ and $\mathcal{R}^{2\text{nd}}\big(\mathcal{V}_h^{\text{int}}\big)$–$\mathcal{R}^{2\text{nd}}\big(\mathcal{V}_h^{\text{bnd}}\big)$. Instead, the inaccurate treatment of the boundary vertices of Robin type when using first-order reconstruction



Vertex reconstructions: $\bigcirc : \mathcal{R}^{LS}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{LS}\left(\mathcal{V}_h^{\text{bnd}}\right)$; $\star : \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{bnd}}\right)$;
$\square : \mathcal{R}^{2nd}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{2nd}\left(\mathcal{V}_h^{\text{bnd}}\right)$; $\diamond : \mathcal{R}^{LS}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{bnd}}\right)$; $\triangle : \mathcal{R}^{2nd}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{bnd}}\right)$.

Fig. 6. Case (i): Dirichlet boundary conditions; relative error curves in the approximation of $\mathsf{u}(x,y) = \sin(2\pi x)\sin(2\pi y) + x^2 + y^2 + 1$ on the domain $\Omega = (0,1) \times (0,1)$ by using permeability ratios $\varepsilon = \{1, 10^{-2}, 10^{-4}\}$ (from top to bottom) with principal diffusive directions rotated of $\theta = \pi/6$ with respect to the coordinate axis.

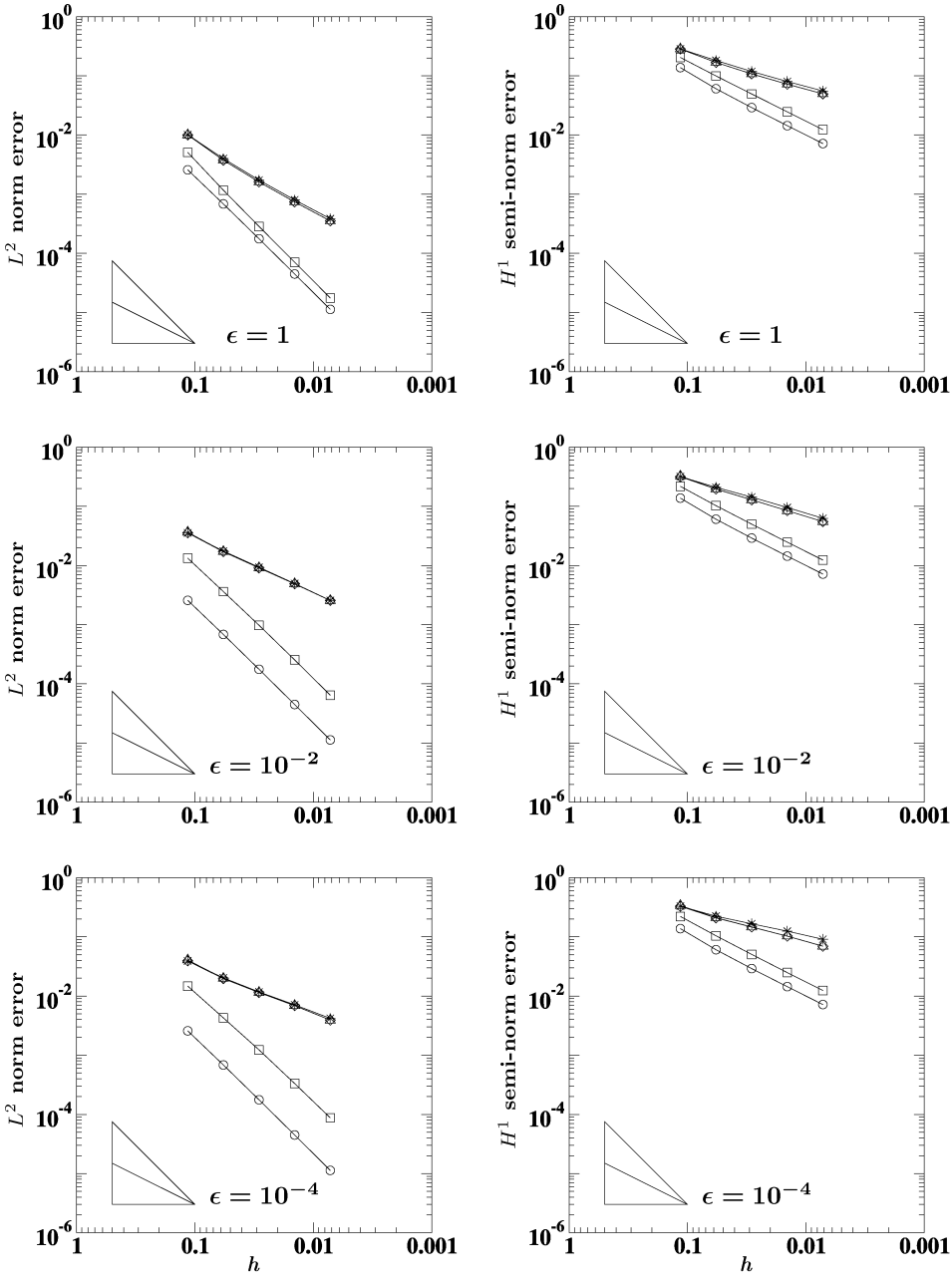
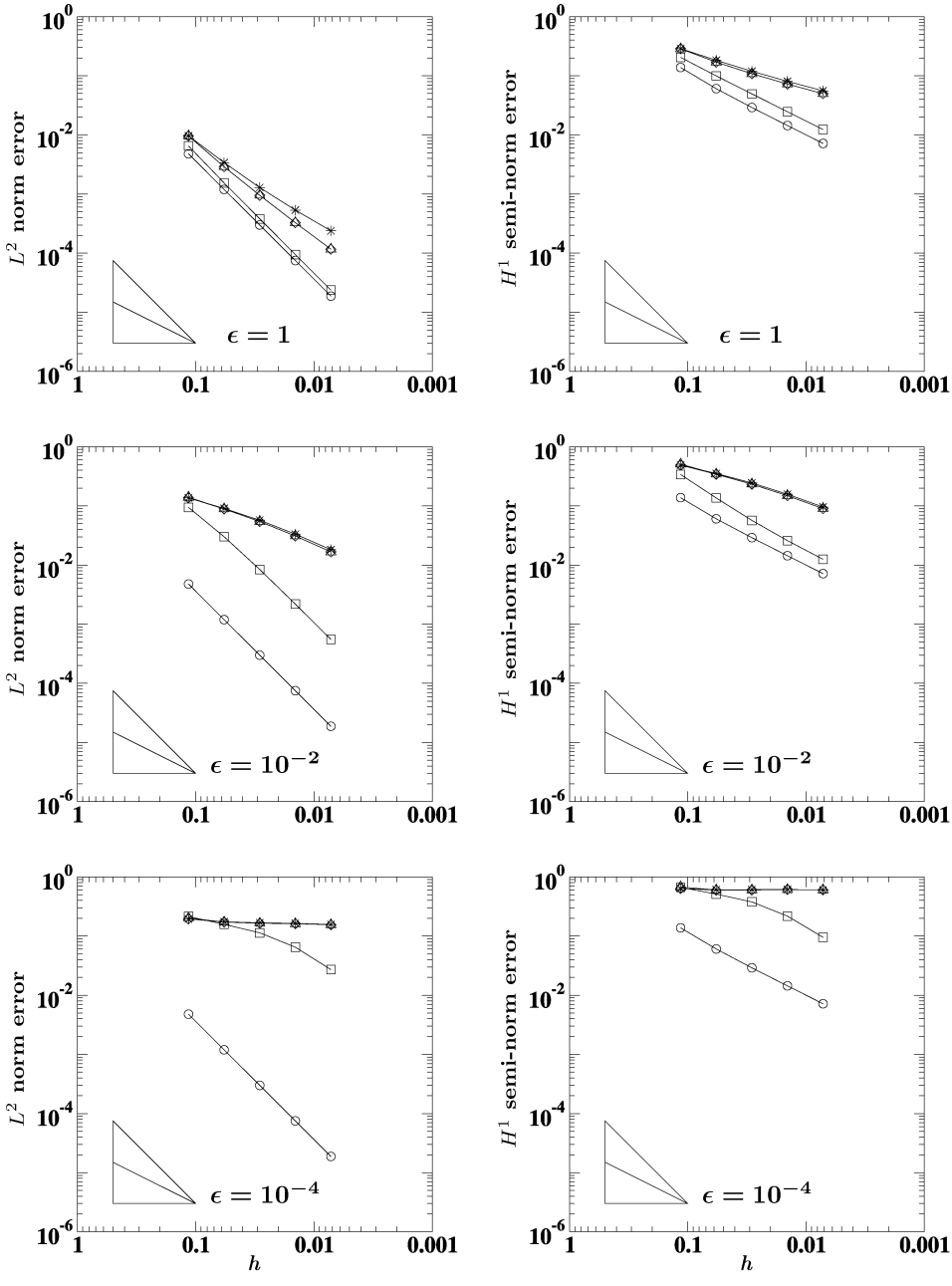


Fig. 7. Case (ii): half Dirichlet and half Robin boundary conditions; relative error curves in the approximation of $\mathsf{u}(x,y) = \sin(2\pi x)\sin(2\pi y) + x^2 + y^2 + 1$ on the domain $\Omega = (0,1) \times (0,1)$ by using permeability ratios $\varepsilon = \{1, 10^{-2}, 10^{-4}\}$ (from top to bottom) with principal diffusive directions rotated of $\theta = \pi/6$ with respect to the coordinate axis.



Vertex reconstructions: $\bigcirc : \mathcal{R}^{LS}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{LS}\left(\mathcal{V}_h^{\text{bnd}}\right)$; $\star : \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{bnd}}\right)$;
$\square : \mathcal{R}^{2nd}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{2nd}\left(\mathcal{V}_h^{\text{bnd}}\right)$; $\diamond : \mathcal{R}^{LS}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{bnd}}\right)$; $\triangle : \mathcal{R}^{2nd}\left(\mathcal{V}_h^{\text{int}}\right) - \mathcal{R}^{1st}\left(\mathcal{V}_h^{\text{bnd}}\right)$.

Fig. 8. Case (iii): half Dirichlet and half Neumann boundary conditions; relative error curves in the approximation of $\mathtt{u}(x,y) = \sin(2\pi x)\sin(2\pi y) + x^2 + y^2 + 1$ on the domain $\Omega = (0,1) \times (0,1)$ by using permeability ratios $\varepsilon = \{1, 10^{-2}, 10^{-4}\}$ (from top to bottom) with principal diffusive directions rotated of $\theta = \pi/6$ with respect to the coordinate axis.

bears the responsibility of the accuracy deterioration mentioned in (ii). Finally, (iii) reveals that the $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ scheme again yields a strong reduction of the error with respect to the use of $\mathcal{R}^{\mathrm{2nd}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{2nd}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ in the case of diffusion anisotropy.

Case III: Mixed Dirichlet and Neumann boundary conditions
In this case, we show the effects of the Neumann boundary conditions on the convergence rates. The results are reported in Fig. 8. From these plots the following facts are experimentally observed:

(i) for $\varepsilon = 1$, the error curves of the schemes using $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ and $\mathcal{R}^{\mathrm{2nd}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{2nd}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ show second-order convergence rates. Instead, the error curves of the schemes using $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{1st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ and $\mathcal{R}^{\mathrm{2nd}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{1st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ are very close to the first-order error curves of the scheme using $\mathcal{R}^{\mathrm{1st}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{1st}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$;
(ii) for $\varepsilon = 10^{-2}$, all the error curves for both $\mathcal{E}_h$ and $\mathcal{E}_{\mathsf{G},h}$ except the ones of the scheme that uses $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ show a deterioration of the error constant;
(iii) for $\varepsilon = 10^{-4}$, all the error curves for both $\mathcal{E}_h$ and $\mathcal{E}_{\mathsf{G},h}$ except the ones of the scheme using $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ are affected by a lack of convergence on the coarsest grids.

From (i)–(iii) it follows that the numerical treatment of Neumann boundaries strongly influences the convergence rates of these finite volume approximations and this influence increases with the degree of diffusion anisotropy. When the ratio of anisotropic diffusion is decreased below unity, a locking phenomenon[5] takes place for all the error curves except the ones produced by the scheme using $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$. As pointed out in (ii), this locking phenomenon is already visible for the intermediate value $\varepsilon = 10^{-2}$ as a deterioration of the error constant. This effect is much stronger for the smallest value $\varepsilon = 10^{-4}$ as the convergence of some of the numerical approximations is prevented on the considered grid refinement levels. On the other hand, the finite volume method based on $\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{int}}\big)-\mathcal{R}^{\mathrm{LS}}\big(\mathcal{V}_h^{\mathrm{bnd}}\big)$ is experimentally locking-free within the considered range of values of the parameter ε and, thus, robust in the sense of Babuska and Suri.[5] As for the case of mixed Dirichlet and Robin boundary vertices of the previous paragraph, these results allow us to conclude that the accuracy of the boundary vertex reconstruction plays a major role in providing a high quality numerical approximation to the solution of a Poisson problem with strong diffusion anisotropy.

6. Conclusions

We developed a finite volume method for two-dimensional elliptic problems with strong diffusion anisotropy. The method is cell-centered on unstructured domain triangulations and approximates the cell averages of the analytical solution by a numerical formulation of the control volume balance of the diffusion flux. The

numerical definition of the diffusion flux is performed at any mesh edge by using both cell averages and vertex values of the solution.

In the cell-centered formulation, cell averages are directly evaluated by the numerical scheme and a special strategy, the vertex reconstruction, is needed to define vertex values. Every vertex value is expressed as a linear interpolation of the cell-averaged solutions of the cells surrounding that vertex, but particular care must be taken in the treatment of a vertex located on a domain boundary with conditions of Neumann or Robin type. The accuracy of the vertex reconstruction, including the treatment of boundary vertices, is indeed crucial for the global accuracy of the finite volume method. Regarding this issue, we proposed a general approach for every kind of mesh vertex, e.g. internal or boundary one, which is based on the resolution of a local least squares problem for the evaluation of the cell weights of the vertex interpolation formula. Our formulation was designed to include information from boundary conditions as linear constraints in the algorithm of weight construction while maintaining the formal consistency that is required to achieve second-order accuracy of the final approximation scheme.

We investigated the performance of this approach by considering three benchmark problems with boundary conditions of full Dirichlet type, half Dirichlet–half Robin type, and half Dirichlet–half Neumann type, respectively. In these test cases, the diffusion anisotropy was controlled by changing the value of the anisotropy ratio, which is the ratio between the eigenvalues of the diffusion tensor. Numerical results were produced for anisotropy ratios varying over a range of representative values of the level of anisotropy that may be found in engineering applications. The accuracy of the method was evaluated by measuring the approximation errors of cell averages and cell-averaged gradients: optimal convergence rates were reflected by the slopes of the experimental convergence curves in log-log plots. By comparison, we also showed that the numerical solutions provided by our least squares-based finite volume method are more accurate than the numerical solutions provided by scheme implementations using other reconstruction algorithms.

References

1. I. Aavatsmark, T. Barkve, O. Boe and T. Mannseth, Discretization on unstructured grids for inhomogeneous, anisotropic media. Part I: Derivation of the methods, *SIAM J. Sci. Comput.* **19** (1998) 1700–1716.
2. I. Aavatsmark, T. Barkve, O. Boe and T. Mannseth, Discretization on unstructured grids for inhomogeneous, anisotropic media. Part II: Discussion and numerical results, *SIAM J. Sci. Comput.* **19** (1998) 1717–1736.
3. P. Amestoy and C. Puglisi, An unsymmetrized LU multifrontal factorization, *SIAM J. Matrix Anal. Appl.* **24** (2002) 553–569.
4. W. K. Anderson and D. L. Bonhaus, An implicit upwind algorithm for computing turbulent flows on unstructured grids, *Computers Fluids* **23** (1994) 1–21.
5. I. Babuška and M. Suri, On locking and robustness in the finite element method, *SIAM J. Numer. Anal.* **29** (1992) 1261–1293.
6. R. E. Bank and D. J. Rose, Some error estimates for the box method, *SIAM J. Numer. Anal.* **24** (1987) 777–787.

7. T. J. Barth and D. C. Jespersen, The design and applications of upwind schemes on unstructured meshes, in *27th Aerospace Sciences Meeting* (AIAA, 1989), pp. 1–12.
8. A. Ben-Israel and T. N. E. Greville, *Generalized Inverses*: *Theory and Applications* (Springer-Verlag, 2003).
9. A. Bergam, Z. Mghazli and R. Verfürth, Estimations a posteriori d'un schéma volumes finis pour un problème non linéaire, *Numer. Math.* **95** (2003) 599–624.
10. L. Bergamaschi, S. Mantica and G. Manzini, A mixed finite element–finite volume formulation of the black-oil model, *SIAM J. Sci. Comput.* **20** (1998) 970–997.
11. E. Bertolazzi, Discrete conservation and discrete maximum principle for elliptic PDEs, *Math. Mod. Meth. Appl. Sci.* **8** (1998) 685–711.
12. E. Bertolazzi and G. Manzini, Algorithm 817 P2MESH: Generic object-oriented interface between 2-D unstructured meshes and FEM/FVM-based PDE solvers, *ACM Trans. Math. Softw.* **28** (2002) 101–132.
13. E. Bertolazzi and G. Manzini, A cell-centered second-order accurate finite volume method for convection-diffusion problems on unstructured meshes, *Math. Mod. Meth. Appl. Sci.* **8** (2004) 1235–1260.
14. E. Bertolazzi and G. Manzini, A finite volume method for transport of contaminants in porous media, *Appl. Numer. Math.* **49** (2004) 291–305.
15. E. Bertolazzi and G. Manzini, Least square-based finite volumes for solving the advection-diffusion of contaminants in porous media, *Appl. Numer. Math.* **51** (2004) 451–461.
16. E. Bertolazzi and G. Manzini, A second-order maximum principle preserving finite volume method for steady convection-diffusion problems, to appear in *SIAM J. Numer. Anal.* 2005.
17. E. Bertolazzi and G. Manzini, A unified treatment of boundary conditions in least square-based finite volume methods, 2005, IAN-CNR Tech. Rep. 1203 (2002), to appear in *Comput. Math. Appl.*
18. Z. Cai, On the finite volume element method, *Numer. Math.* **58** (1991) 713–735.
19. Z. Cai, J. Mandel and S. McCormick, The finite volume element method for diffusion equations on composite grids, *SIAM J. Numer. Anal.* **28** (1991) 392–402.
20. C. Chainais-Hillairet and Y.-J Peng, Finite volume approximation for degenerate drift-diffusion system in several space dimensions, *Math. Mod. Meth. Appl. Sci.* **14** (2004) 461–481.
21. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems* (North-Holland, 1980).
22. Y. Coudière, J.-P Vila and P. Villedieu, Convergence rate of a finite volume scheme for a two dimensional convection-diffusion problem, *Math. Mod. Numer. Anal.* **33** (1999) 493–516.
23. K. Domelevo and P. Omnes, A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids, to appear in *Math. Mod. Numer. Anal.*
24. R. E. Ewing, T. Lin and Y. Lin, On the accuracy of the finite volume element method based on piecewise linear polynomial, *SIAM J. Numer. Anal.* **39** (2002) 1865–1888.
25. R. Eymard, T. Gallouët and R. Herbin, in *Handbook of numerical analysis*, (North-Holland, 2000), Vol. VII, pp. 713–1020.
26. M. Feistauer, J. Felcman and I. Straškraba, *Mathematical and Computational Methods for Compressible Flow* (Oxford Univ. Press, 2003).
27. N. T. Frink, Recent progress toward a three-dimensional unstructured Navier-Stokes solver, in *32nd Aerospace Sciences Meeting*, (AIAA, 1994), pp. 1–20.
28. N. T. Frink, P. Parikh and S. Pirzadeh, A fast upwind solver for the Euler equations on three-dimensional unstructured meshes, in *Computational Fluid Dynamic Conference*, (AIAA, 1991) pp. 1–6.
29. W. Hackbusch, On first and second order box scheme, *Computing* **41** (1989) 277–296.

30. B. Heinrich, Finite difference methods on irregular networks, 1987.
31. F. Hermeline, A finite volume method for the approximation of diffusion operators on distorted meshes, *J. Comput. Phys.* **160** (2000) 481–499.
32. J. Hyman and M. Shashkov, Solving diffusion equations with rough coefficients in rough grids, *J. Comput. Phys.* **129** (1997) 383–405.
33. J. Hyman, M. Shashkov and S. Steinberg, The numerical solution of diffusion problems in strongly heterogeneous non-isotropic materials, *J. Comput. Phys.* **132** (1997) 130–148.
34. J. M. Hyman and M. Shashkov, Adjoint operators for the natural discretizations of the divergence, gradient, and curl on logically rectangular grids, *Appl. Numer. Math.* **25** (1997) 413–442.
35. J. M. Hyman and M. Shashkov, Natural discretizations for the divergence, gradient, and curl on logically rectangular grids, *Comput. Math. Appl.* **33** (1997) 81–104.
36. P. A. Jayantha and I. W. Turner, A second order control-volume finite-element least-squares strategy for simulating diffusion in strongly anisotropic media, *J. Comput. Math.* **23** (2005) 1–16.
37. O. Klein and P. Philip, Transient conductive-radiative heat transfer: discrete existence and uniqueness for a finite volume scheme, *Math. Mod. Meth. Appl. Sci.* **15** (2005) 227–258.
38. R. H. Li, Z. Y. Chen and W. Wu, *Generalized Difference Methods for Differential Equations* (Marcel Dekker, 2000).
39. G. Manzini and S. Ferraris, Mass-conservative finite volume methods on 2-D unstructured grids for the Richards' equation, *Adv. Water Res.* **27** (2004) 1199–1215.
40. M. Putti and C. Cordes, Finite element approximation of the diffusion operator on tetrahedra, *SIAM J. Sci. Comput.* **19** (1998) 1154–1168.
41. A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations* (Springer-Verlag, 1994).
42. A. N. Tikhonov and A. A. Samarskii, Homogeneous difference schemes, *Zh. Vychisl. Mat. Mat. Fiz.* **1** (1961) 5–63.
43. A. N. Tikhonov and A. A. Samarskii, Homogeneous difference schemes on non-uniform nets, *Zh. Vychisl. Mat. Mat. Fiz.* **2** (1962) 812–832.
44. A. Weiser and M. F. Wheeler, On convergence of block-centered finite differences for elliptic problems, *SIAM J. Numer. Anal.* **25** (1988) 351–375.