Journal of Computational Physics 252 (2013) 219–249

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Monotonic solution of heterogeneous anisotropic diffusion problems

Costanza Aricò *, Tullio Tucciarelli

Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, dei Materiali, Università di Palermo, Viale delle Scienze, 90128 Palermo, Italy

ARTICLE INFO

Article history:
Received 14 February 2012
Received in revised form 1 March 2013
Accepted 6 June 2013
Available online 27 June 2013

Keywords:
Anisotropic diffusion
Heterogeneous medium
M-matrix
Delaunay mesh
Affine transformation
Edge swap

ABSTRACT

Anisotropic problems arise in various areas of science and engineering, for example groundwater transport and petroleum reservoir simulations. The pure diffusive anisotropic time-dependent transport problem is solved on a finite number of nodes, that are selected inside and on the boundary of the given domain, along with possible internal boundaries connecting some of the nodes. An unstructured triangular mesh, that attains the Generalized Anisotropic Delaunay condition for all the triangle sides, is automatically generated by properly connecting all the nodes, starting from an arbitrary initial one. The control volume of each node is the closed polygon given by the union of the midpoint of each side with the "anisotropic" circumcentre of each final triangle. A structure of the flux across the control volume sides similar to the standard Galerkin Finite Element scheme is derived. A special treatment of the flux computation, mainly based on edge swaps of the initial mesh triangles, is proposed in order to obtain a stiffness *M*-matrix system that guarantees the monotonicity of the solution. The proposed scheme is tested using several literature tests and the results are compared with analytical solutions, as well as with the results of other algorithms, in terms of convergence order. Computational costs are also investigated.

1. Introduction

Diffusion equation with anisotropic coefficients arises in many environmental topics, for example heat transfer, groundwater flow and transport problems, petroleum reservoir simulations, hydrodynamic simulations, semiconductor modelling, biology problems, ... These problems are characterized by a full rank diffusive tensor, that is diagonal only if the reference system is aligned with the principal direction of anisotropy [7].

Steady-state diffusion problems satisfy the Maximum Principle (MP), which states that the solution cannot have a maximum or a minimum within the interior of the domain. The numerical solvers are aimed at satisfying the discrete counterpart of the MP, the so-called Discrete Maximum Principle (DMP), computing numerical solution free of spurious oscillations. A number of sufficient conditions are given for a class of linear elliptic Partial Differential Equations (PDEs) problems [40,10,11,37,38]. The most common one is that the problem stiffness matrix is an *M*-matrix. An *M*-matrix is an irreducible matrix with diagonal positive coefficients, strictly diagonally dominant, or weakly diagonally dominant with strict inequality for at least one row and non-positive off-diagonal coefficients (see for example [27,33]).

Numerical schemes satisfying the DMP have been developed according to the sufficient conditions, by either properly discretizing the governing PDE or by employing a suitable mesh. Most success has been obtained in the isotropic case, where the diffusive tensor reduces to a scalar value. It has been shown [9,11] that the linear Finite Element (FE) method

* Corresponding author.
E-mail addresses: costanza.arico@unipa.it (C. Aricò), tullio.tucciarelli@unipa.it (T. Tucciarelli).

0021-9991/$ – see front matter
http://dx.doi.org/10.1016/j.jcp.2013.06.017

satisfies the DMP when the mesh is a simplicial and satisfies the so-called non-obtuse angle condition, that is the dihedral angles of all mesh elements are non-obtuse [27]. In the 2D and homogeneous case this condition can be replaced by the weaker Delaunay condition, that the sum of any pair of angles opposite to a common edge is less than or equal to π [26,39].

The anisotropic case is more difficult. Draganescu et al. [12] proved that the non-obtuse angle condition fails to guarantee DMP satisfaction in the anisotropic diffusion problem. Several techniques have been proposed: local adjustments have been apported to the underlying numerical scheme by Liska and Shashkov [30] and Kuzmin et al. [23], Sharma and Hammett [36] used a slop limiter function in the discretization of the governing PDEs, Li et al. [28] optimized a triangular mesh for an FE scheme in order to reduce unphysical oscillations. Le Potier [24,25] and Lipnikov et al. [29] proposed a non-linear first-order Finite Volume (FV) method that obtains an M-matrix on arbitrary meshes in cases of parabolic PDEs but does not satisfy DMP for steady-state problems. Mlacnik and Durlofsky [33] optimized the mesh for a Multi-Point Flux Approximation (MPFA) FV scheme. The authors in [33] map the anisotropic physical problem into an isotropic computational one and optimize in this space the original grid of the physical problem. They solve the isotropic problem in the computational space. The authors in [33] predict some limitations of the proposed optimization technique for heterogeneous media. Li and Huang [27] studied a linear FE method for steady-state diffusion problems and developed a generalization of the non-obtuse angle condition for anisotropic cases. The condition requires that the dihedral angles of all mesh elements, measured in a metric depending on the diffusive tensor, be non-obtuse. This requirement reduces to the non-obtuse angle condition for isotropic problems. The authors in [27] derived also a metric tensor to use for mesh generation based on the anisotropic non-obtuse angle condition. They adopted the so-called M-uniform mesh approach [20], where an anisotropic mesh is generated as an M-uniform mesh or a uniform mesh in the metric specified by a tensor. The metric tensor is symmetric and positive definite and provides information on the size, shape and orientation of mesh elements. M-uniform meshes generated with the metric tensor satisfy the anisotropic non-obtuse angle condition and are aligned with the diffusive tensor. Edwards and Zheng [13,14] developed a family of flux-continuous FV schemes for the solution of the general tensor pressure equation for subsurface flows. These schemes have full pressure continuity imposed across control volume faces, while the early family of flux-continuous schemes present a point-wise pressure and flux continuity. One of the main advantages of these schemes is that, due to continuity conditions, they retain a single degree of freedom per control volume. The authors in [14] assert that optimal support can be achieved by anisotropy favoring triangulation (where the sign of triangulation angle equals the sign of dominant principal direction angle). The authors in [14] triangulate each primal quadrilateral element of a structured grid, according to the local sign variation of the off-diagonal coefficient of the diffusive tensor. The scheme in [14] minimizes spurious oscillations in the computed solutions and leads to more robust quasi-positive families of flux-continuous schemes applicable to generally discontinuous full tensor problems.

Mesh locking effects may arise for strong anisotropy problems and can be experimentally observed when the discretization error does not decrease at the expected rate for the limiting anisotropy ratio [18,19]. Havu [18] and Havu and Pitkäranta [19] introduced a modification to the bilinear Galerkin scheme to solve locking problems for high anisotropy ratios. These modifications could alter the consistency of the original scheme for the isotropic case and factors tending to zero, proportionally to the expected convergence rate of the scheme, have to be introduced to recover the consistency. A completely different approach is to try to adapt the domain, or the mesh, or both, to anisotropy of the problem [2,3]. This approach can be used if the anisotropy ratio is constant. Manzini and Putti [31] presented a FV diamond scheme with second-order accurate spatial reconstruction for both tangential and normal cell interface gradients. This scheme maintains second-order convergence rate on unstructured triangular grids and the only locking effects appear for very high anisotropy ratios and for quasi-purely Neumann boundary conditions problems.

A novel methodology for the solution of the anisotropic heterogeneous diffusion problem is presented in this paper. The algorithm is aimed at solving the problem on a finite number of irregular points arbitrary selected within the computational domain, as done by the MPFA computational scheme. On the opposite, the resulting spatial discretization of the fluxes across the control volume is similar to the one occurring in the standard linear ($P1$) Galerkin FE scheme, the number of unknowns is restricted to the number of nodes and a procedure for flux coefficients formulation is suggested in order to prevent from spurious oscillations in the solution, that avoid mesh locking effects. This algorithm acts directly on the physical mesh and does not deal with computational space, nor with metrics depending on the anisotropy tensor and the number and the location of the given input nodes remain unchanged. The proposed approach represents a grid adjustment algorithm leading to anisotropy favoring grids, eventhough, unlike the optimal anisotropy favoring triangulation schemes proposed in [14], grid adjustment is obtained by checking the sign of the global stiffness coefficients.

The proposed algorithm is mainly finalized to the solution of the diffusive problem where tensor coefficients are constant in time, though it has been easily extended to problems, like groundwater simulations, with time-dependent coefficients, function of the velocity field [5]. In such problems (where diffusive coefficients depend on the velocity field), the proposed procedure has to be combined with methods achieving inter-element flux continuity, like Mixed Finite Elements and their hybridized formulation (see [32] and cited references), or MPFA schemes (see [33] and cited references). In [5] flux continuity has been obtained by a modified formulation of an element-lumping Mixed Hybrid FE scheme with unknowns number equal to the number of elements.

The work is organized as follows. In Section 2, the governing PDEs along with the initial and boundary conditions and the space/time integration leading to the solving system are presented. In Section 3, the computation of the so-called "anisotropic" circumcentre of each triangle and the general formulation of the flux coefficients are given. The flux coefficients are computed: in Section 4 for the case of isotropic medium and Generalized Delaunay (GD) mesh; in Section 5 for the case

of directionally homogeneous anisotropic medium and Anisotropic Generalized Delaunay (GAD) mesh; in Section 6 for the most general case of anisotropic heterogeneous medium. Finally, in Section 7 several numerical tests are proposed. The first test is aimed at investigating the difference between the spectral condition number of the stiffness matrix of the proposed method and the one computed by a standard linear Galerkin FE scheme. The other tests are aimed at investigating the DMP property, the convergence order for imposed arbitrary analytical solution, the possible existence of mesh locking effects and the computational costs required by the adopted procedure for stiffness matrix computation satisfying the M-property with respect to the cost due to the solution of the system.

2. Governing equations, spatial discretization and input data structure

Let Ω be a physical domain in $\mathbb{R}^2$ (in the vertical or horizontal plane), Γ the boundary of Ω and $\mathbf{x} = [x_1, x_2]^T$ the spatial co-ordinate vector. Let $H^1(\Omega)$ be the Sobolev space of square-integrable functions with square-integrable first-order derivatives over Ω.

Assume the following diffusive time-dependent problem in the unknown variable $u(\mathbf{x}, t) \in H^1(\Omega)$:

$$\begin{cases} \zeta \dfrac{\partial u}{\partial t} + \nabla \cdot \mathbf{q} = f \quad \text{in } \Omega \times [0, T], \\ \mathbf{q} = -\mathbf{D}\nabla u, \end{cases} \tag{1}$$

subject to the initial and boundary conditions:

$$\begin{cases} u(\mathbf{x}, 0) = u_0(\mathbf{x}) \quad \text{in } \Omega \text{ at } t = 0, \ \mathbf{x} \in \Omega, \\ u = u_D(\mathbf{x}, t), \quad \mathbf{x} \in \Gamma_D, \quad \mathbf{q}(\mathbf{x}, t) \cdot \mathbf{n} = g_N(\mathbf{x}, t), \quad \mathbf{x} \in \Gamma_N, \end{cases} \tag{2}$$

where t is time, T is the total simulation time, ζ is a non-negative piecewise constant function, $\mathbf{q}$ is the unit diffusive flux vector, $\mathbf{D}(\mathbf{x})$ is the (2×2) diffusive matrix, assumed symmetric and positive definite, Γ_D and Γ_N are the portions of Γ where Dirichlet and Neumann conditions respectively hold, u_D is a fixed Dirichlet value on Γ_D, g_N is the assigned Neumann flux (with $\mathbf{n}$ the unit outward vector normal to the boundary), u_0 is the initial state and $f = f(\mathbf{x}, t) \in L_2(\Omega)$ is a given function (source term).

Spatial discretization of the governing PDE (1) is based on a generally unstructured triangular mesh. Assume a polygonal approximation Ω_h of Ω and an unstructured triangulation T_h of Ω_h to be available. N_T is the number of triangle T of T_h. Triangle T is called *primary element* and $|T|$ is the area of T. Let $\{i, \ i = 1, \ldots, N\}$ be the set of all vertices (or nodes) of all $T \in T_h$ and N the number of nodes. A dual mesh $E_h = \{e_i, \ i = 1, \ldots, N\}$ is constructed over the triangulation T_h and the dual finite control volume e_i associated with node i is the closed polygon given by the union of the midpoint of each side with the "anisotropic" circumcentre of each triangle T sharing i, further defined. In the following of the paper the dual volumes e_i are called also cells and $|e_i|$ is the area of e_i. Cells e_i satisfy:

$$\Omega = \bigcup e_i. \tag{3}$$

The following input data are assumed to be known: (1) the initial condition, (2) a finite number of points (nodes), inside the computational domain and along its internal and external boundaries. An initial T_h domain triangulation, connecting all the given points (nodes), must also be available. (3) The external numerical boundary, given as a closed set of straight lines connecting the boundary nodes, (4) the boundary conditions at the external boundary nodes, (5) the interior numerical boundaries, given as finite number of straight lines connecting other couple of nodes. The interior boundaries divide the domain in zones with strong different medium parameters (e.g., diffusion coefficients). Even in heterogeneous media diffusion or permeability parameters are assumed constant within each element. In stochastic simulations, a minimum number of 5–10 computational cells within the correlation length is strongly suggested [8]. An internal boundary is needed if abrupt discontinuities in the diffusion tensor occur. (6) The diffusive tensor in all the nodes. Nodes on the interior boundaries can have two different diffusive tensors, one for each of the two zones lying on each side of the interior boundary.

Tensor $\mathbf{D}$ is assumed as constant inside each resulting triangular element and computed on the base of the nodal values, as further specified in Sections 5 and 6.

The above mentioned initial triangulation T_h connecting the given input nodes, is not a pre-established mesh to support the sought after numerical solution, but the same mesh will be the result of a proposed edge swapping procedure. In the following we call the initial triangulation T_h initial (or starting) domain triangulation. The mesh input data are only the node locations. Moreover, the reason to assign tensor $\mathbf{D}$ to nodes and not to triangles will be more clear in the following.

Storage capacity is assumed to be concentrated in the cells in the measure of 1/3 of the area of all triangles sharing the node and a linear variation is assumed for the variable u inside each triangle.

After space and fully implicit time integration, as well as application of the Green's theorem, the ith equation of the diffusive system (1) in the u_i unknowns can be generally written as:

$$\widehat{\zeta}_i \frac{u_i^{t+1} - u_i^t}{\Delta t} |e_i| + \sum_{k \neq i} G_{i,k} \delta_{i,k} = \widehat{f}_i |e_i|, \quad i, k = 1, \ldots, N, \tag{4a}$$

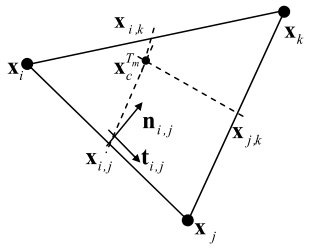

Fig. 1. Element notation.

where $\widehat{\zeta}_i$ and $\widehat{f}_i$ are the mean values of ζ and f in node i, indices t and $t+1$ mark the beginning and the end of the time step and the summation on the l.h.s. represents the spatial discretization of the diffusive operator $-\nabla\cdot(\mathbf{D}\nabla u)$. $\delta_{i,k}$ is equal to 1 or 0 according if node k is linked to node i or not; summation is taken over all the nodes different from i and $G_{i,k}$ represents the total flux through the two sides of the control volume sharing the midpoint between i and k. We will show in the following that, due to the given definition of "anisotropic" circumcentre, it is possible to write Eq. (4a) in the form:

$$\widehat{\zeta}_i\frac{u_i^{t+1}-u_i^t}{\Delta t}|e_i|+\sum_{k\neq i}F_{i,k}\big(u_k^{t+1}-u_i^{t+1}\big)\delta_{i,k}=\widehat{f}_i|e_i|,\quad i,k=1,\ldots,N, \tag{4b}$$

where $F_{i,k}$ is the global stiffness matrix coefficient and can be viewed as the flux from k to i due to a difference $(u_k-u_i)=1$. $F_{i,k}$ is obtained by summation over all the elements T_m sharing node i, of the element stiffness coefficients $F_{i,k}^{T_m}$, in the following called also flux coefficients and further defined, and Eq. (4b) can be written as:

$$\sum_{T_m=1,N_T}\frac{|T_m|\zeta_m\delta_i^{T_m}}{3}\frac{u_i^{t+1}-u_i^t}{\Delta t}+\sum_{T_m=1,N_T}\big(F_{i,j}^{T_m}\big(u_j^{t+1}-u_i^{t+1}\big)+F_{i,k}^{T_m}\big(u_k^{t+1}-u_i^{t+1}\big)\big)\delta_i^{T_m}$$
$$=\sum_{T_m=1,N_T}\frac{|T_i|f_m\delta_i^{T_m}}{3},\quad i=1,\ldots,N, \tag{4c}$$

where ζ_m and f_m are respectively the values of ζ and f in triangle T_m, $\delta_i^{T_m}$ is equal to 1 or 0 according if i is a node of element T_m or not, j and k are the other two nodes of T_m. The order of the associated linear system is equal to the nodes number. A preconditioned conjugate gradient method is used for its solution.

In the following, we look for the location of the "anisotropic" circumcentres that allows to move from formulation (4a) to formulation (4b), such that the flux $G_{i,k}$ through each element is function only of the difference between the u variables at the nodes i and k.

3. Stiffness and flux coefficients general structure

3.1. The anisotropic circumcentre

Assume a medium with full diffusion tensor $\mathbf{D}$ known in each point. The medium is treated as homogeneous within each triangular element and the element diffusion tensor is obtained as function of the nodal values, as further specified in Sections 5 and 6. Let T_m be a triangular element of the triangulation T_h with nodes i, j and k, where j and k are the nodes following and preceding respectively node i in counterclockwise direction. Let p be one of the T_m nodes and q the following one in counterclockwise direction. The edge vector $\mathbf{r}_{p,q}$ connects nodes p and q, oriented from p to q and $\mathbf{r}_{q,p}=-\mathbf{r}_{p,q}$. Call c_{T_m} the anisotropic circumcentre of T_m and $\mathbf{x}_c^{T_m}$ its spatial co-ordinate vector. Call $P_{p,q}$, the midpoint of edge $\mathbf{r}_{p,q}$, with $\mathbf{x}_{p,q}$ the corresponding co-ordinate vectors. Call $(\mathbf{x}_c^{T_m}-\mathbf{x}_{p,q})$ the vector linking c_{T_m} with $P_{p,q}$. Call $\mathbf{n}_{p,q}$ the inward unit vector orthogonal to the edge $\mathbf{r}_{p,q}$ (see Fig. 1), and $\mathbf{n}_{p,q}^{c_{T_m}}$ the unit vector orthogonal to $(\mathbf{x}_c^{T_m}-\mathbf{x}_{p,q})$, oriented from p to q so that c_{T_m} follows midpoint $P_{p,q}$ in counterclockwise order on side $(\mathbf{x}_c^{T_m}-\mathbf{x}_{p,q})$. The anisotropic circumcentre is computed in order to set to zero fluxes $Fn_{i,j}^{T_m}$ and $Fn_{i,k}^{T_m}$, respectively across segments $|P_{i,j}c_{T_m}|$ and $|P_{i,k}c_{T_m}|$, due to the component of the vector $\mathbf{D}\nabla u$ orthogonal to edges $\mathbf{r}_{i,j}$ and $\mathbf{r}_{i,k}(=-\mathbf{r}_{k,i})$. This is equivalent to set:

$$Fn_{i,j}^{T_m}=-\mathbf{D}\mathbf{n}_{i,j}\cdot\mathbf{n}_{i,j}^{c_{T_m}}=0, \tag{5a}$$

$$Fn_{i,k}^{T_m}=-\mathbf{D}\mathbf{n}_{i,k}\cdot\mathbf{n}_{i,k}^{c_{T_m}}=0, \tag{5b}$$

with $\mathbf{n}_{i,k}=-\mathbf{n}_{k,i}$ and $\mathbf{n}_{i,k}^{c_{T_m}}=-\mathbf{n}_{k,i}^{c_{T_m}}$ and symbol $(\cdot)$ is the dot product. Since $\mathbf{D}$ is a full tensor, sides $\mathbf{r}_{i,j}$ and $\mathbf{r}_{i,k}$ are generally not orthogonal to vectors $-\mathbf{D}\mathbf{n}_{i,j}$ and $-\mathbf{D}\mathbf{n}_{i,k}$. In Appendix B it is shown that Eqs. (5a)–(5b) imply also:

$$Fn_{j,k}^{T_m}=-\mathbf{D}\mathbf{n}_{j,k}\cdot\mathbf{n}_{j,k}^{c_{T_m}}=0, \tag{5c}$$

with the above specified symbols. An isotropic problem can be viewed as a particular case of the anisotropic one, where full tensor $\mathbf{D}$ becomes a positive scalar value D and sides $\mathbf{r}_{i,j}$ and $\mathbf{r}_{i,k}$ are orthogonal to the vectors linking the circumcentre of the triangle with the midpoints of its edges.

Eqs. (5a)–(5b) can be written in the scalar form as:

$$\begin{aligned}&\left(-D_{x_1x_1}(x_{2,j}-x_{2,i})+D_{x_1x_2}(x_{1,j}-x_{1,i})\right)\left(x_{2,c}^{T_m}-x_{2,i,j}\right)\\&\quad-\left(-D_{x_2x_1}(x_{2,j}-x_{2,i})+D_{x_2x_2}(x_{1,j}-x_{1,i})\right)\left(x_{1,c}^{T_m}-x_{1,i,j}\right)=0,\end{aligned}\tag{6a}$$

$$\begin{aligned}&\left(-D_{x_1x_1}(x_{2,k}-x_{2,i})+D_{x_1x_2}(x_{1,k}-x_{1,i})\right)\left(x_{2,c}^{T_m}-x_{2,i,k}\right)\\&\quad-\left(-D_{x_2x_1}(x_{2,k}-x_{2,i})+D_{x_2x_2}(x_{1,k}-x_{1,i})\right)\left(x_{1,c}^{T_m}-x_{1,i,k}\right)=0.\end{aligned}\tag{6b}$$

3.2. Computation of the flux and stiffness coefficients

According to the computation of c_{T_m}, fluxes across $|P_{i,j}c_{T_m}|$ and $|P_{i,k}c_{T_m}|$ are entirely due to the component of the vector $\mathbf{D}\nabla u$ along the $\mathbf{r}_{i,j}$ and $\mathbf{r}_{i,k}$ directions and do not depend on the u values of the opposite nodes, respectively nodes k and j.

With the above specified symbols p and q for triangle T_m, we define:

$$\mathbf{t}_{p,q}=\frac{\mathbf{r}_{p,q}}{|\mathbf{r}_{p,q}|}=\frac{\mathbf{x}_q-\mathbf{x}_p}{|\mathbf{x}_q-\mathbf{x}_p|},\tag{7}$$

the unit vector parallel to side $\mathbf{r}_{p,q}$ (see Fig. 1 and [35]) and we compute the flux coefficient relative to nodes i and j (the flux across $|P_{i,j}c_T|$ due to $(u_j-u_i)=1$) as:

$$F_{i,j}^{T_m}=-\int_{P_{i,j}}^{c_{T_m}}\frac{1}{|\mathbf{r}_{i,j}|}\mathbf{D}\mathbf{t}_{i,j}\cdot d\mathbf{n}=-\frac{1}{|\mathbf{r}_{i,j}|}\mathbf{D}\mathbf{t}_{i,j}\cdot\mathbf{n}_{i,j}^{c_{T_m}}\left|\mathbf{x}_c^{T_m}-\mathbf{x}_{i,j}\right|.\tag{8a}$$

Similarly, flux coefficient relative to nodes k and i (flux across $|P_{i,k}c_T|$ due to $(u_i-u_k)=1$) is:

$$F_{k,i}^{T_m}=-\int_{P_{i,k}}^{c_{T_m}}\frac{1}{|\mathbf{r}_{k,i}|}\mathbf{D}\mathbf{t}_{k,i}\cdot d\mathbf{n}=-\int_{c_{T_m}}^{P_{i,k}}\frac{1}{|\mathbf{r}_{i,k}|}\mathbf{D}\mathbf{t}_{i,k}\cdot d\mathbf{n}=-\frac{1}{|\mathbf{r}_{i,k}|}\mathbf{D}\mathbf{t}_{i,k}\cdot\mathbf{n}_{i,k}^{c_{T_m}}\left|\mathbf{x}_c^{T_m}-\mathbf{x}_{i,k}\right|.\tag{8b}$$

Finally, coefficients $F_{i,j}^{T_m}$ and $F_{i,k}^{T_m}$ can be computed as:

$$F_{i,j}^{T_m}=-\mathbf{D}(\mathbf{x}_j-\mathbf{x}_i)\cdot\mathbf{n}_{i,j}^{c_{T_m}}\left|\mathbf{x}_c^{T_m}-\mathbf{x}_{i,j}\right|\frac{1}{|\mathbf{r}_{i,j}|^2},\tag{9a}$$

$$F_{i,k}^{T_m}=F_{k,i}^{T_m}=-\mathbf{D}(\mathbf{x}_k-\mathbf{x}_i)\cdot\mathbf{n}_{i,k}^{c_{T_m}}\left|\mathbf{x}_c^{T_m}-\mathbf{x}_{i,k}\right|\frac{1}{|\mathbf{r}_{i,k}|^2}.\tag{9b}$$

Eqs. (9) in scalar form can be written as:

$$\begin{aligned}F_{i,j}^{T_m}=\frac{1}{|\mathbf{r}_{i,j}|^2}\Big[&\left(D_{x_1x_1}(x_{1,j}-x_{1,i})+D_{x_1x_2}(x_{2,j}-x_{2,i})\right)\left(x_{2,c}^{T_m}-x_{2,i,j}\right)\\&-\left(D_{x_2x_1}(x_{1,j}-x_{1,i})+D_{x_2x_2}(x_{2,j}-x_{2,i})\right)\left(x_{1,c}^{T_m}-x_{1,i,j}\right)\Big],\end{aligned}\tag{10a}$$

$$\begin{aligned}F_{i,k}^{T_m}=\frac{1}{|\mathbf{r}_{i,k}|^2}\Big[&\left(D_{x_1x_1}(x_{1,i}-x_{1,k})+D_{x_1x_2}(x_{2,i}-x_{2,k})\right)\left(x_{2,c}^{T_m}-x_{2,i,k}\right)\\&-\left(D_{x_2x_1}(x_{1,i}-x_{1,k})+D_{x_2x_2}(x_{2,i}-x_{2,k})\right)\left(x_{1,c}^{T_m}-x_{1,i,k}\right)\Big].\end{aligned}\tag{10b}$$

Call T_1 and T_2 two neighboring triangles of T_h sharing side $\mathbf{r}_{i,k}$. The nodes of T_1 are i, j and k, the nodes of T_2 are k, l and i (in counterclockwise order, see Fig. 2). Anisotropic circumcentres of T_1 and T_2 are c_{T_1} and c_{T_2}.

As specified in Section 2, in the resolving system (4), the stiffness coefficient $F_{i,k}$ corresponding to the connected nodes i and k is given by the sum of the contributions of the flux coefficients in elements T_1 and T_2 related to sides $|P_{i,k}c_{T_1}|$ and $|P_{i,k}c_{T_2}|$ (see Fig. 2):

$$F_{i,k}=F_{i,k}^{T_1}+Fd_{i,k}^{T_2},\tag{11}$$

with coefficients $F_{i,k}^{T_1}$ and $F_{i,k}^{T_2}$ defined according to Eqs. (8)–(9). Eq. (11) can be written as:

$$F_{i,k}=-\frac{1}{|\mathbf{r}_{i,k}|^2}\left[\mathbf{D}^{T_1}(\mathbf{x}_k-\mathbf{x}_i)\cdot\mathbf{n}_{i,k}^{c_{T_1}}\left|\mathbf{x}_c^{T_1}-\mathbf{x}_{i,k}\right|+\mathbf{D}^{T_2}(\mathbf{x}_k-\mathbf{x}_i)\cdot\mathbf{n}_{i,k}^{c_{T_2}}\left|\mathbf{x}_c^{T_2}-\mathbf{x}_{i,k}\right|\right],\tag{12}$$

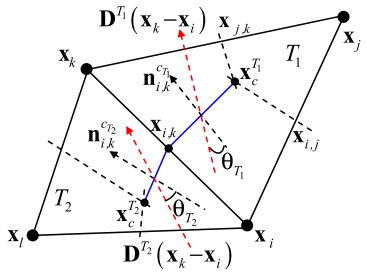

Fig. 2. Computation of matrix coefficient $F_{i,k}$.

where $\mathbf{D}^{T_{1(2)}}$ is the diffusive tensor of element $T_{1(2)}$. Eq. (12) can be written as:

$$F_{i,k} = -\frac{(d_{i,k}^{T_1} c_{i,k}^{T_1} \cos\theta_{T_1} + d_{i,k}^{T_2} c_{i,k}^{T_2} \cos\theta_{T_2})}{|\mathbf{r}_{i,k}|^2}, \tag{13}$$

$$\text{where } d_{i,k}^{T_q} = \left|\mathbf{D}^{T_q}(\mathbf{x}_k - \mathbf{x}_i)\right|, \text{ with } q = 1, 2, \tag{14}$$

$c_{i,k}^{T_q}$ is the distance between c_{T_q} and midpoint $P_{i,k}$ in triangle T_q and θ_{T_q} is the angle between vectors $\mathbf{D}^{T_q}(\mathbf{x}_k - \mathbf{x}_i)$ and $\mathbf{n}_{i,k}^{c_{T_q}}$ in T_q (see Fig. 2). Observe, in Fig. 2, that vectors $(\mathbf{x}_c^{T_1} - \mathbf{x}_{i,k})$ and $(\mathbf{x}_c^{T_2} - \mathbf{x}_{i,k})$ have different directions since generally $\mathbf{D}^{T_1} \neq \mathbf{D}^{T_2}$.

The flux coefficients adopted in the conforming linear ($P1$) Galerkin can be derived in this way (see [35] for example), for the specific case of isotropic medium and non-obtuse triangle meshes.

According to formulation of the solving system (4), stiffness coefficients $F_{i,k}$ have to be always negative to guarantee the M-property of the system matrix.

4. Stiffness and flux coefficients computation in the case of isotropic medium and Generalized Delaunay (GD) mesh

Starting from the definition of anisotropic circumcentre, in the isotropic case the dual finite volume e_i associated with node P_i is the closed polygon given by the union of the midpoint of each side with the circumcentre of each triangle T. This is also the control volume of the mass balance of the standard Galerkin FE stiffness equation for node P_i (see for example [34,35]).

If $\mathbf{D}^{T_q} = D^{T_q}\mathbf{I}$ ($q = 1, 2$, $\mathbf{I}$ is the identity matrix and D^{T_q} is a positive scalar value) are the diffusive tensors of triangles T_1 and T_2 (the way we compute element tensors from the assigned nodal values is explained in the next section), coefficients $F_{i,j}^{T_1}$ and $F_{i,k}^{T_1}$ in the standard linear ($P1$) Galerkin FE approach are computed as (see for example [35]):

$$F_{i,j}^{T_1} = -\frac{D^{T_1} c_{i,j}^{T_1}}{|\mathbf{r}_{i,j}|} \quad \text{and} \quad F_{i,k}^{T_1} = -\frac{D^{T_1} c_{i,k}^{T_1}}{|\mathbf{r}_{i,k}|}, \tag{15}$$

with the above specified symbols and the stiffness coefficient $F_{i,k}$ of the stiffness matrix corresponding to the connected nodes i and k is equal to:

$$F_{i,k} = -\frac{(D^{T_1} c_{i,k}^{T_1} + D^{T_2} c_{i,k}^{T_2})}{|\mathbf{r}_{i,k}|}. \tag{16}$$

Distance $c_{i,k}^{T_{1(2)}}$ can be computed as:

$$c_{i,k}^{T_q} = \frac{(x_{1,i} - x_{1,k})(x_{2,c}^{T_q} - x_{2,i,k}) - (x_{2,i} - x_{2,k})(x_{1,c}^{T_q} - x_{1,i,k})}{\sqrt{(x_{1,i} - x_{1,k})^2 + (x_{2,i} - x_{2,k})^2}} \delta_q, \quad q = 1, 2, \tag{17}$$

where $\delta_q = -1$ or 1 if direction of $\mathbf{r}_{i,k}$ is respectively counterclockwise or not in triangle T_q.

A Delaunay triangulation in $\mathbb{R}^2$ is defined by the condition that all the nodes in the mesh are not interior to the circles defined by the three nodes of each triangle. Elements T_1 and T_2 in Fig. 3(a) satisfy the Delaunay condition, while in Fig. 3(b) they do not [21]. In Fig. 3(a), the control volumes boundaries portions in T_1 and T_2 for nodes i, k, j and l are respectively $\overline{P_{i,j}c_{T_1}c_{T_2}P_{l,i}}$, $\overline{P_{j,k}c_{T_1}c_{T_2}P_{k,l}}$, $\overline{P_{i,j}c_{T_1}P_{j,k}}$ and $\overline{P_{l,i}c_{T_2}P_{k,l}}$. It can be shown [15,35] that each Delaunay triangulation satisfies the following condition:

$$c_{i,k}^{T_1} + c_{i,k}^{T_2} \geqslant 0, \tag{18}$$

for each interior edge connecting nodes i and k.

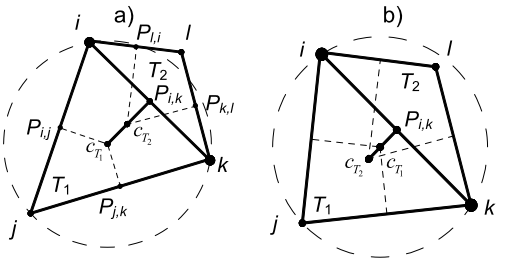

Fig. 3. (a) Elements T_1 and T_2 satisfy the Delaunay condition; (b) elements T_1 and T_2 do not satisfy the Delaunay condition.

Given a Delaunay mesh, we propose to change the stiffness coefficient $F_{i,k}$ in Eq. (16) as:

$$F_{i,k} = -\frac{(D^{T_1}c_1 + D^{T_2}c_2)}{|\mathbf{r}_{i,k}|}, \tag{19}$$

with c_q $(q = 1, 2)$ defined as:

$$c_1 = c_{i,k}^{T_1}, \qquad c_2 = c_{i,k}^{T_2} \quad \text{if } c_{i,k}^{T_1} \geqslant 0 \text{ and } c_{i,k}^{T_2} \geqslant 0, \tag{20a}$$

$$c_1 = c_{i,k}^{T_1} + c_{i,k}^{T_2}, \qquad c_2 = 0 \quad \text{if } c_{i,k}^{T_2} < 0 \text{ and } c_{i,k}^{T_1} \geqslant |c_{i,k}^{T_2}|, \tag{20b}$$

$$c_1 = 0, \qquad c_2 = c_{i,k}^{T_1} + c_{i,k}^{T_2} \quad \text{if } c_{i,k}^{T_1} < 0 \text{ and } c_{i,k}^{T_2} \geqslant |c_{i,k}^{T_1}|, \tag{20c}$$

and distance $c_{i,k}^{T_q}$ given by Eq. (17). According to Eqs. (20a)–(20c), c_q is never smaller than zero.

Assuming two different element parameters D^{T_q}, the stiffness coefficient $F_{i,k}$ from node i to node k given by the standard Galerkin FE discretization, can loose consistency with the u difference, even if the mesh satisfies the Delaunay property and the sum of the distances $c_{i,k}^{T_1} + c_{i,k}^{T_2}$ is positive. On the opposite, the formulation provided by Eqs. (19)–(20) still guarantees the negative sign of the same stiffness coefficient defined by Eq. (19).

If the two triangles sharing nodes i and k are acute triangles, formulations (19)–(20) and (16) overlap. If one of the two triangles is obtuse, as in Fig. 3(a), where $c_{i,k}^{T_1} > 0$, $c_{i,k}^{T_2} < 0$ (see Eq. (17)) and $c_{i,k}^{T_1} > |c_{i,k}^{T_2}|$, $F_{i,k}$ computed according to Eqs. (19)–(20) is still equal to the flux due to $(u_k - u_i) = 1$, across the side having the vertices in the two triangle circumcentres (side $\overline{c_{i,k}^{T_1} c_{i,k}^{T_2}}$) and computed according to the $\mathbf{q}$ vector defined in Eq. (1). This flux is consistent with the geometry of the Voronoi cells associated to nodes i and k, since it is function of the diffusion tensor of the acute triangle where the side $\overline{c_{i,k}^{T_1} c_{i,k}^{T_2}}$ is entirely located (see Fig. 3(a)) and equal to $F_{i,k} = -(D^{T_1}c_1)/|\mathbf{r}_{i,k}| < 0$, with $c_1 = c_{i,k}^{T_1} + c_{i,k}^{T_2} > 0$. In the same case, $F_{i,k}$ computed according to the Galerkin formulation in Eq. (16) is different and could not be consistent with the $\mathbf{q}$ vector occurring in the acute triangle if $D^{T_2} > D^{T_1}$. Eq. (16) leads in fact to $F_{i,k} = -(D^{T_1}c_{i,k}^{T_1} + D^{T_2}c_{i,k}^{T_2})/|\mathbf{r}_{i,k}| > 0$, if $D^{T_1}c_{i,k}^{T_1} + D^{T_2}c_{i,k}^{T_2} < 0$.

Most of the today available mesh-generators satisfy the Delaunay property, even if some exceptions occur around internal boundaries, or when the mesh density is forced to change in given sub-domains. If the starting domain triangulation T_h does not satisfy the Delaunay property (as in the example shown in the previous Fig. 3(b)), it is still possible [15] to obtain a new mesh satisfying condition in Eq. (18) for all the internal edges, starting from the original one and without changing the location of the original nodes. This can be done by a series of local edge swaps, where two elements sharing the same edge are changed in a new couple, sharing the same nodes but having a different edge, connecting the two nodes opposite to the previous edge. See for example the new triangles obtained in Fig. 4(b) by the original ones of Fig. 4(a). It can be shown [15] that the common edge satisfies the Delaunay property in at least one of the two configurations. By iterating the same control for all the edges, the Delaunay property is quickly attained for all the edges of the mesh that are shared by two triangles.

If element T_1 is a boundary element and $\mathbf{r}_{i,k}$ is a boundary edge opposite to an obtuse angle, the flux coefficient, proportional to $c_{i,k}^{T_1}$, is positive, even if the mesh satisfies the Delaunay property, because the distance of the circumcentre from the boundary edge, computed by Eq. (17), is negative.

We define Generalized Delaunay (GD) mesh a Delaunay mesh where the condition:

$$c_{i,k}^{T_1} \geqslant 0 \tag{21}$$

holds for all the boundary edges.

It is still possible to obtain a GD mesh, also saving the internal and/or external boundaries, if condition in Eq. (21) does not hold for one or more boundary edges in the starting domain triangulation T_h of the input nodes. To this aim the triangle(s) sharing the external (internal) boundary is (are) divided in new triangles by adding new nodes along the original boundary edge. The new triangles have the same height as the original one(s) with respect to the boundary edge, but the

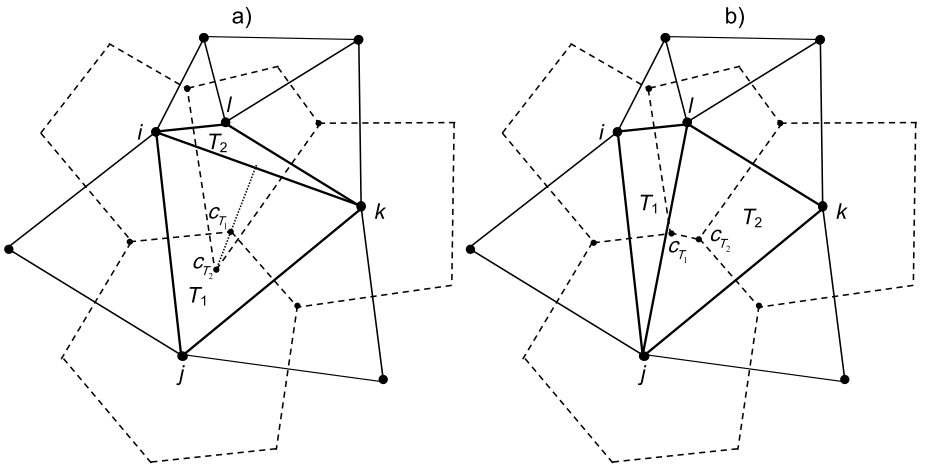

Fig. 4. (a) Original not Delaunay triangulation; (b) Delaunay triangulation after edge swap.

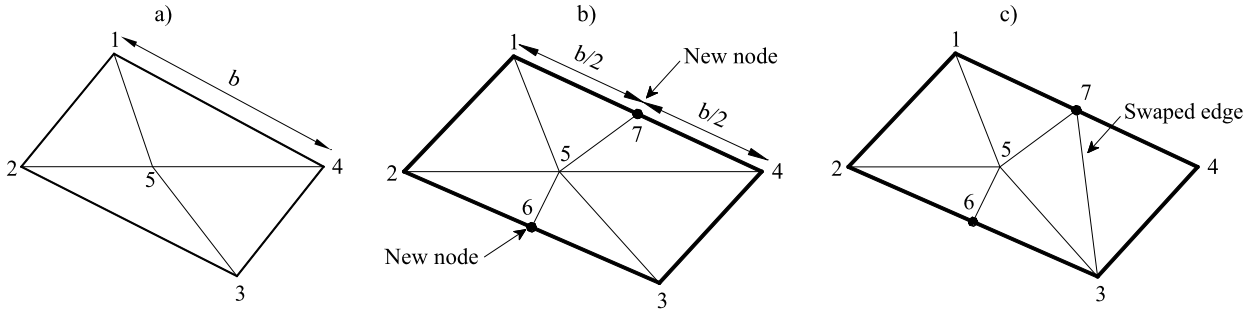

Fig. 5. (a) Original triangulation; (b) new triangulation after adding two new nodes; (c) Generalized Delaunay (GD) triangulation after applying the swap technique.

base length will be a fraction of the original one. The triangulation obtained by adding nodes along the boundary edge will be the new starting domain triangulation. After this change, the same edge swap iterative procedure can be applied to the new resulting starting triangulation. It can be easily shown that the resulting mesh will satisfy the GD property, if the number of new nodes is large enough (see [4] and Appendix A). See an example in Fig. 5, where the boundary edges 2–3 and 1–4 do not satisfy the GD property (Fig. 5(a)). The mesh is first changed in a new one by setting a new node in both edges 2–3 and 1–4 (Fig. 5(b)), and then changed in a GD mesh by applying the swap technique to the edge 4–5, that is changed with the new edge 3–7 (Fig. 5(c)).

Observe that GD property is not a sufficient condition for the M-property according to the formulation of flux and stiffness coefficients given in the Galerkin $P1$ scheme by Eqs. (15)–(16), because of a possible difference between D^{T_1} and D^{T_2}. Observe also that, in the isotropic case, GD condition depends only on the mesh geometry and it is not function of the spatial distribution of the scalar diffusion coefficients.

The Galerkin approach guarantees the positive definite condition (all the eigenvalues greater than zero) of the final linear system matrix, even if the GD condition does not hold, but does not guarantee the M-property [17,35]. On the other hand, if a non-GD mesh is used with the proposed algorithm, the iterative methods used for the solution of the linear system could fail, because of the negative eigenvalues. This restricts the use of the proposed algorithm to triangular meshes that satisfy conditions in Eqs. (18) and (21) in all the edges (see also [4]).

5. Generalized Anisotropic Delaunay (GAD) mesh. Stiffness and flux coefficients computation in the directionally homogeneous anisotropic case

Full tensor **D** can be written as:

$$\mathbf{D} = d_0 \mathbf{D}' = d_0 \begin{pmatrix} D'_{11} & D'_{12} \\ D'_{21} & D'_{22} \end{pmatrix}, \tag{22a}$$

where

$$d_0 = D_{11} + D_{22} \quad \text{and} \quad D'_{rs} = \frac{D_{rs}}{d_0}, \quad r, s = 1, 2. \tag{22b}$$

Observe that d_0 is the first invariant of tensor $\mathbf{D}$. Tensor $\mathbf{D}'$, as well as $\mathbf{D}$, is symmetric and positive definite. In the following, we call d_0 and $\mathbf{D}'$ respectively the scalar and the directional components of $\mathbf{D}$ and we assume the component $\mathbf{D}'$ to be constant throughout the computational domain, due to the hypothesis of directionally homogeneous anisotropy. Eigenvectors of $\mathbf{D}'$ are the same of $\mathbf{D}$, except for the scale factor d_0.

Tensor $\mathbf{D}$ is originally assigned to the mesh nodes, as mentioned in Section 2. According to Eqs. (22), it is possible to compute, for each node, its scalar and directional components. The element tensor $\mathbf{D}$ in each triangle is obtained as product of the mean of the scalar component at its nodes times $\mathbf{D}'$.

In the isotropic case it has been shown [15] that, if the edge shared by two triangles does not respect the Delaunay condition, the new edge obtained by swapping the old edge in a new one will do it. Similarly to the previous isotropic case, it is also possible to save the M-property in the more general anisotropic directionally homogeneous case, even if the scalar diffusion d_0 changes in space but the matrix $\mathbf{D}'$ is constant over all the domain.

Let $\mathbf{D}^{T_1}$ and $\mathbf{D}^{T_2}$ be two diffusion tensors in the two neighboring triangles T_1 and T_2 sharing side $\mathbf{r}_{i,k}$. Since the two tensors have the same directional component $\mathbf{D}'$, we set:

$$\mathbf{D}^{T_q} = d_0^{T_q}\mathbf{D}', \quad q = 1, 2. \tag{23}$$

Using the same approach of the isotropic case, if condition in Eq. (18) is satisfied, the stiffness coefficient $F_{i,k}$ can be computed by changing Eq. (13) in the following one:

$$F_{i,k} = -\frac{(d_{i,k}^{T_1} c_1 \cos\theta_{T_1} + d_{i,k}^{T_2} c_2 \cos\theta_{T_2})}{|\mathbf{r}_{i,k}|^2}, \tag{24a}$$

that, according to the hypothesis of directional homogeneity, becomes:

$$F_{i,k} = -\frac{(d_{i,k}^{T_1} c_1 + d_{i,k}^{T_2} c_2)\cos\theta}{|\mathbf{r}_{i,k}|^2}, \tag{24b}$$

where θ is the angle between vectors $d_0^{T_1}\mathbf{D}'^{T_1}(\mathbf{x}_k - \mathbf{x}_i)$ and $\mathbf{n}_{i,k}^{c_{T_1}}$ (equal to the angle between $d_0^{T_2}\mathbf{D}'^{T_2}(\mathbf{x}_k - \mathbf{x}_i)$ and $\mathbf{n}_{i,k}^{c_{T_2}}$), c_1 and c_2 are given by Eqs. (20) and, according to Eqs. (14) and (23), $d_{i,k}^{T_{1(2)}}$ in Eqs. (24) are set as:

$$d_{i,k}^{T_q} = d_0^{T_q}\left|\mathbf{D}'^{T_q}(\mathbf{x}_k - \mathbf{x}_i)\right|, \quad q = 1, 2. \tag{25}$$

Observe that, even if $d_{i,k}^{T_1}$ and $d_{i,k}^{T_2}$ have different values, the stiffness coefficient $F_{i,k}$ computed by Eq. (24b) is negative – and M-matrix property is preserved – if angle $-\pi/2 \leqslant \theta \leqslant \pi/2$.

We say a mesh to satisfy the Generalized Anisotropic Delaunay (GAD) property if the stiffness coefficient $F_{i,k}$ computed for each edge of the mesh by Eq. (24b) is always negative. Observe that the GAD condition depends not only on the mesh geometry, but also on the particular choice of the directional diffusion matrix $\mathbf{D}'$ that affects the θ values.

In the following we shall prove that, also in the anisotropic case, if the GAD property is satisfied for a given edge, it is not for the new "swapped" one and vice versa. The proof is based on the proof of the following conditions: (1) an affine transformation does exist that maps the actual physical space in a computational space and the anisotropic circumcentres of the physical space are mapped in the isotropic circumcentres of the mapped triangles, (2) the sign of the stiffness coefficient $F_{i,k}$ between the linked nodes in the physical and in the computational space remains the same. Condition (1) implies that an actual and a "swapped" edges in the computational space correspond to an actual and a "swapped" edges in the physical space. Condition (2) implies that the sign of the product of the two stiffness coefficients, along the actual and the "swapped" edges, must be negative both in the physical and in the computational space. This is equivalent to say that the GAD condition holds either for the actual edge or for the swapped edge in the physical space, as it holds for the GD condition in the computational isotropic one. Of course there is no practical need, in the computation, to move into the computational space.

The reason to initially assign diffusive tensor to nodes, instead of to triangles, relies in the iterative edge swap procedure applied to guarantee the GAD mesh property, which can change triangles in the mesh along the computation.

Starting from Eqs. (5), the orthogonality in triangle T_1 between $\mathbf{Dn}_{i,q}$ and $\mathbf{n}_{i,q}^{T_1}$ ($q = k$ or j), implies the following condition:

$$\left(\mathbf{x}_c^{T_1} - \frac{1}{2}(\mathbf{x}_q + \mathbf{x}_i)\right)_\perp \cdot \mathbf{D}'(\mathbf{x}_q - \mathbf{x}_i)_\perp = 0, \tag{26}$$

where symbol $(\cdot)_\perp$ indicates the vector direction orthogonal to vector $(\cdot)$, since $\mathbf{D}'(\mathbf{x}_q - \mathbf{x}_i)_\perp$ has the same direction of $\mathbf{Dn}_{i,q}$. By changing $\mathbf{x}_q$ with the co-ordinate vector $\mathbf{x}$ of a generic point, Eq. (26) can be written in the form:

$$\left(\mathbf{x}_c^{T_1} - \frac{1}{2}(\mathbf{x} + \mathbf{x}_i)\right)_\perp \cdot \mathbf{D}'(\mathbf{x} - \mathbf{x}_i)_\perp = 0. \tag{27}$$

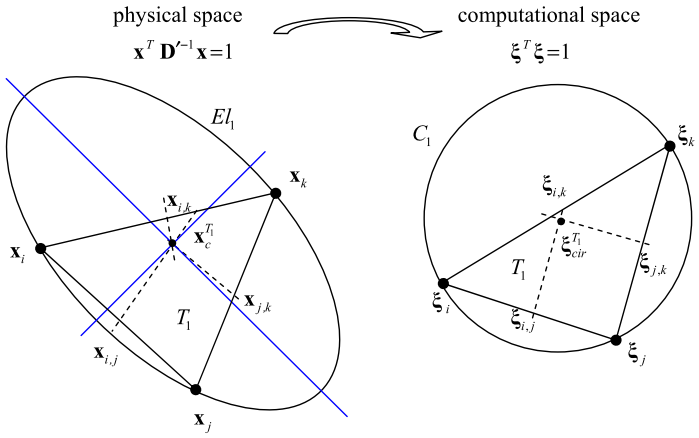

Fig. 6. (left) Geometrical interpretation of a diffusion tensor in an anisotropic medium. Match between physical and computational space. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Eq. (27) is the equation of an ellipse, El_1, with centre c_{T_1}, passing through the three vertices of element T_1 (see the proof in Appendix B). The axes of the ellipse (see Fig. 6, left) have the direction of the eigenvectors of tensor $\mathbf{D}'$ and the ratio between the axes lengths is equal to the ratio between the eigenvalues of $\mathbf{D}'$. Matrix $\mathbf{D}'$ is equal to:

$$\mathbf{D}' = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^T = \mathbf{H}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\mathbf{H}^T, \tag{28}$$

where the columns of matrix $\mathbf{H}$ are the eigenvectors $\mathbf{h}_1$ and $\mathbf{h}_2$ of $\mathbf{D}'$, $\boldsymbol{\Lambda}$ is the diagonal matrix with diagonal elements equal to the eigenvalues λ_1 and λ_2 of $\mathbf{D}'$ and matrix $\mathbf{H}^T$ is the transposed of $\mathbf{H}$.

According to Eq. (28), Eq. (27) can be written as:

$$\left(\mathbf{x}_c^{T_1} - \frac{1}{2}(\mathbf{x}+\mathbf{x}_i)\right)_\perp \mathbf{C}\mathbf{C}^T(\mathbf{x}-\mathbf{x}_i)_\perp = 0, \tag{29}$$

where

$$\mathbf{C} = \mathbf{H}\boldsymbol{\Lambda}^{1/2}, \tag{30}$$

and $\mathbf{C}^T$ is the transposed of $\mathbf{C}$. Observe that matrix $\mathbf{C}$ acts on the co-ordinate vectors $\mathbf{x}$ with a rotation (with respect to the Cartesian directions) and a distortion transformation: $\mathbf{H}$ accounts for the rotation and $\boldsymbol{\Lambda}^{1/2}$ for the distortion. It is then possible to define a new co-ordinate vector:

$$\boldsymbol{\xi} = \mathbf{F}\mathbf{x} \quad \text{where } \mathbf{F} = \mathbf{C}^T. \tag{31}$$

$\mathbf{x}$ and $\boldsymbol{\xi}$ represent the co-ordinates vectors respectively in the physical and in the computational space. According to Eq. (31), Eq. (29) in the computational space can be written as:

$$\left(\boldsymbol{\xi}_{cir}^{T_1} - \frac{1}{2}(\boldsymbol{\xi}+\boldsymbol{\xi}_i)\right)_\perp^T \cdot (\boldsymbol{\xi}-\boldsymbol{\xi}_i)_\perp = 0. \tag{32}$$

Eq. (32) shows that vector $(\boldsymbol{\xi}_{cir}^{T_1} - \frac{1}{2}(\boldsymbol{\xi}+\boldsymbol{\xi}_i))_\perp$ is orthogonal to vector $(\boldsymbol{\xi}-\boldsymbol{\xi}_i)_\perp$. Therefore, Eq. (32) represents a circle equation (see also Appendix B); this implies that the ellipse El_1 in the physical space is transformed into a circle C_1 in the computational space by matrix $\mathbf{F}$ (see Fig. 6). The original physical anisotropic problem becomes in the computational space an isotropic one.

Observe that the sign of the flux coefficient $F_{i,k}^{T_{1(2)}}$ depends, in the physical space, on the angle between vector $\mathbf{t}' = \mathbf{D}'(\mathbf{x}_k - \mathbf{x}_i)$ and vector $\mathbf{s}' = \mathbf{n}_{i,k}^{c_{T_{1(2)}}}$. Matrix $\mathbf{F} = \boldsymbol{\Lambda}^{1/2}\mathbf{H}^T$ provides an affine transformation between the physical and the computational space and matrix $\mathbf{H}^T$ can be viewed as an intermediate transformation of the co-ordinates $\mathbf{x}$ in a new space $\boldsymbol{\xi}' = \mathbf{H}^T\mathbf{x}$. The original figure is simply rotated in the new space $\boldsymbol{\xi}'$, therefore, the flux coefficient in the original $\mathbf{x}$ space and in the new one $\boldsymbol{\xi}'$ is the same. The diagonal matrix $\boldsymbol{\Lambda}^{1/2}$ provides a second transformation from $\boldsymbol{\xi}'$ to $\boldsymbol{\xi}$, that is a simple contraction along the $\boldsymbol{\xi}'$ principal axis. The corresponding stiffness coefficient in the computational space is:

$$F_{i,k}^{\xi} = F_{i,k}\lambda_1^{1/2}\lambda_2^{1/2}. \tag{33}$$

Since λ_1 and λ_2 are both positives, the sign of two coefficients $F_{i,k}^{\xi}$ and $F_{i,k}$, respectively in the physical and the computational space, is the same.

Similarly to the previous isotropic case, GAD mesh property can be achieved by applying the above swapping procedure, after adding to the starting domain triangulation extra nodes along the internal and/or external boundary edges, so that nodes and boundary edges satisfy the MD property (see Appendix A). On the other hand, the number of new nodes that should be added depends on the particular $\mathbf{D}'$ direction matrix. Because the proposed algorithm is aimed to be applied also to transport problems, where diffusive matrix $\mathbf{D}'$ is subject to change also in time, we prefer to maintain the same number of input nodes in the starting domain triangulation (including the nodes added to the boundaries) and to relax the anisotropy condition along the boundaries as follows. Let T be an internal or external boundary element and i and k its boundary nodes. If the stiffness coefficient $F_{i,k}$ is positive, we enforce the coefficient to be zero. To do that, we replace the diffusion matrix $\mathbf{D}$ in T with the following one:

$$\tilde{\mathbf{D}} = \mathbf{H}_D \begin{pmatrix} \lambda_{D1} & 0 \\ 0 & \alpha\lambda_{D1} \end{pmatrix} \mathbf{H}_D^T, \tag{34}$$

where $\mathbf{H}_D$ and λ_{D1} are respectively the eigenvector matrix and the first eigenvalue of the original $\mathbf{D}$ tensor, and α is chosen to set to zero the flux between nodes i and k. This is equivalent to set:

$$\tilde{\mathbf{D}}(\mathbf{x}_k - \mathbf{x}_i)_\perp \cdot \tilde{\mathbf{D}}\mathbf{n}_{i,k} = 0, \tag{35}$$

with the above specified symbols. Observe that the l.h.s. of Eq. (35) is a continuous function of α and has the same coefficient sign. If we assume that the nodes and the boundary edges satisfy the MD property (see Appendix A), the isotropic case corresponding to $\alpha = 1$ will provide a negative stiffness coefficient. On the other hand, we know that the value $\alpha = \lambda_{D2}/\lambda_{D1}$ corresponds to a positive stiffness coefficient and therefore at least a positive α value exists that satisfies Eq. (35). This implies that the corresponding tensor will satisfy the positive definite condition.

6. Stiffness and flux coefficients computation in the heterogeneous anisotropic case

We compute the element tensor, also in the most general heterogeneous anisotropic case, as $\mathbf{D} = d_0\mathbf{D}'$, where d_0 is the mean of the tensor scalar component at the three nodes of the element. Unfortunately, it is not possible to guarantee the M-property in this most general case, unless some "smoothing" is applied to the directional $\mathbf{D}'$ element component. To do that, we initialize the directional element component $\mathbf{D}'$ as the mean of the tensors at the three nodes of the element. When an internal edge is swapped, the $\mathbf{D}'$ element component is updated in both the new triangles with the average value of the two old ones, following the procedure depicted in the flow chart shown in Fig. 7.

Terms adopted in Fig. 7 have the following meanings:

Call "marked" the edges connecting two nodes i, k such that the condition $F_{i,k} < 0$ is known and "unmarked" the edges connecting two nodes i, k such that the condition $F_{i,k} < 0$ has to be checked.

Set initially all the edges as "unmarked" and let L_u be the ensemble of "unmarked" edges.

Call l the generic edge with nodes i and k. Let $\mathbf{D}^{T_1} \neq \mathbf{D}^{T_2}$ be the two tensors of triangles T_1 and T_2 sharing l, obtained by averaging the nodal scalar and directional components.

Operation I. Compute stiffness coefficient $F_{i,k}$ according to Eqs. (24)–(25).

Operation II. Assign $\mathbf{D}'^{T_1} = \frac{\mathbf{D}'^{T_1} + \mathbf{D}'^{T_2}}{2}$ and $\mathbf{D}'^{T_2} = \frac{\mathbf{D}'^{T_1} + \mathbf{D}'^{T_2}}{2}$.

Operation A. Mark edge l and unmark all the other edges of the triangles sharing edge l, update stiffness coefficients of all the edges of the triangles sharing edge l.

Operation B. Set $F_{i,k} = 0$ and update $\mathbf{D}'$ in the boundary element by solving Eq. (35) in the α unknown.

Observe that the check of the stiffness coefficient sign and the possible consequent edge swapping depicted in the previous flow chart (in Fig. 7) have to be applied at each possible change of the directional tensors $\mathbf{D}'$. To avoid the need of totally regenerating the mesh starting from the GD one after a minimum change of $\mathbf{D}'$, it is more convenient to regenerate the mesh starting from the last one and to initialize only the nodal $\mathbf{D}'$ values.

If internal boundaries with parameter discontinuities are present, two different values of tensor $\mathbf{D}$ at the nodes of the internal boundary are given as input. Call $\mathbf{r}_{i,j}$ a side of an internal boundary, oriented from node i to node j according to a fixed direction on the same side. Let $\mathbf{D}_i^1$ and $\mathbf{D}_i^2$ be the two diffusive tensors assigned to node i. Call T_e the triangle of the initial GD mesh, or the mesh computed at the previous iteration of the procedure depicted in the flow chart of Fig. 7, sharing side $\mathbf{r}_{i,j}$. Let $\mathbf{l}_e$ be the oriented vector from midpoint of $\mathbf{r}_{i,j}$ to the centre of mass of T_e. Compute the product $p_e = \mathbf{r}_{i,j} \wedge \mathbf{l}_e$. The average diffusive tensor in element T_e will be computed using one of the two $\mathbf{D}_i^q$ $(q = 1, 2)$ node tensors: $\mathbf{D}_i^1$ will be related to the case $p_e \geqslant 0$, $\mathbf{D}_i^2$ to the case $p_e < 0$.

7. Numerical tests

We present 7 numerical tests, organized as follows. Tests 1 to 6 are steady-state Boundary Values Problems (BVP). In the first test we study the differences between the spectral condition number of the final stiffness matrix computed by the proposed procedure and by the standard $P1$ Galerkin FE scheme. Tests 2 and 3 are proposed to verify the performance of the proposed method in satisfying the DMP. Tests 4 and 5 are presented to investigate the convergence order for assigned analytical solution. Test 6 is related to mesh locking effects. Finally, in test 7 a time-dependent problem with a given

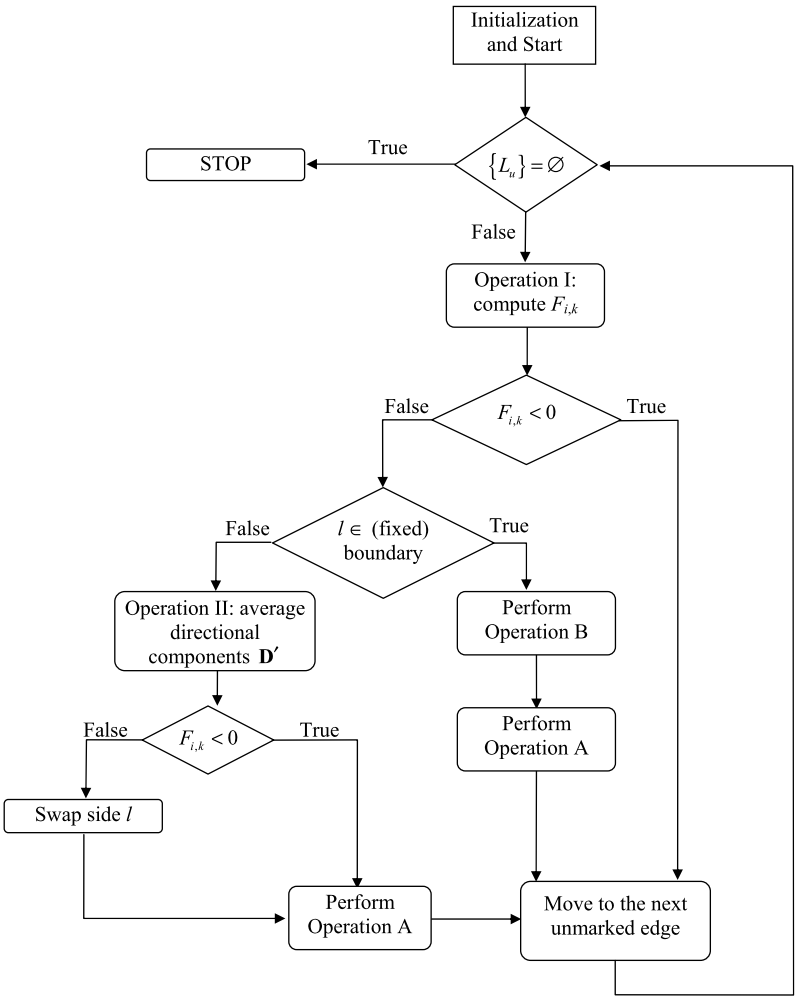

Fig. 7. Flow chart of the procedure for flux computation in heterogeneous anisotropic medium.

analytical solution is presented and the convergence order is estimated. For many of the presented tests, comparison with other literature models are provided too.

7.1. Test 1. Analysis of the system matrix condition number and comparison with the standard $P1$ Galerkin scheme

Assume the unitary $[0,1]^2$ square domain and a diagonal diffusion tensor $\mathbf{D}$. Coefficient D_{11} is kept constant and equal to 1 while D_{22} ranges from 1 to 1.d−10. The starting triangulation, with 14 acute triangles, is shown in Fig. 8(a) [32]. Dirichlet conditions are imposed on the left and right boundary sides.

The Gershgorin theorem (see [32]) states that each eigenvalue λ of a square matrix A_{nn} satisfies the following constraint:

$$|\lambda - A_{ii}| \leqslant \sum_{i \neq j} |A_{ij}|, \quad i, j = 1, 2, \ldots, n. \tag{36}$$

Eq. (36) implies that each eigenvalue of the matrix A must be within a distance d_i from A_{ii}, where $d_i = \sum_{j \neq i} |A_{ij}|$ is called Gershgorin radius. Fig. 9 shows, for all the investigated D_{22} range, the ratio between the sum of the absolute values of the off-diagonal ($\sum_{j \neq i} |A_{ij}|$) and the diagonal (A_{ii}) coefficients for a generic row of the stiffness matrix of the proposed and of the standard $P1$ Galerkin schemes. In the proposed procedure, the mentioned ratio and the Gershgorin radius scale as D_{22} and, according to the Gershgorin theorem, eigenvalues concentrate around to the diagonal coefficients and the minimum eigenvalue is bounded by D_{11}. In $P1$ Galerkin scheme the Gershgorin radius tends to a constant finite value.

Tables 1a and 1b show the maximum and the minimum stiffness matrix eigenvalues (λ_{max} and λ_{min} respectively), as well as their ratio, for different decreasing D_{22} values, for both the numerical schemes. The spectral condition number of the stiffness matrix is the ratio $\lambda_{max}/\lambda_{min}$. This ratio is closer to 1 in the proposed algorithm than in the $P1$ Galerkin scheme, and this implies a better conditioning of the final system matrix. Because of the anisotropy, some of the stiffness

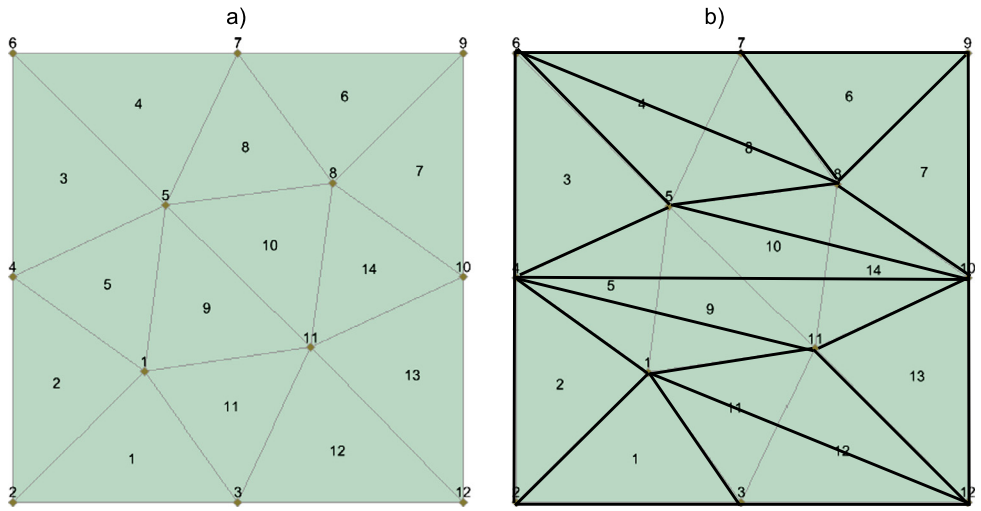

Fig. 8. Test 1. (a) The non-uniform original GD unstructured mesh; (b) final GAD mesh obtained by the proposed procedure after edge swap (D_{22} 1.d−06 and 1.d−10).

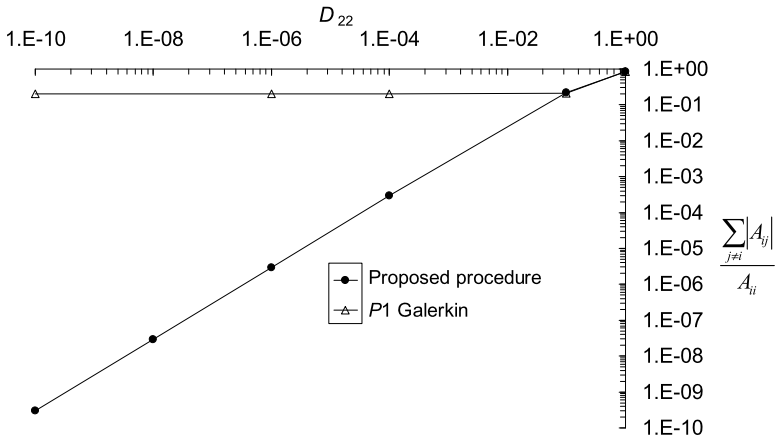

Fig. 9. Test 1. Ratio between the sum of the absolute values of the off-diagonal coefficients and the diagonal coefficient of matrix system versus D_{22} values for proposed and $P1$ Galerkin schemes.

Table 1a
Test 1. Extreme eigenvalues and spectral condition number computed by the proposed procedure for different D_{22} values.

D_{22}	λ_{min}	λ_{max}	$\lambda_{max}/\lambda_{min}$
1.d+00	0.88030	4.8407	5.4989
1.d−01	0.72124	2.49489	3.4592
1.d−04	0.57819	1.87076	3.2355
1.d−06	0.57799	1.87010	3.2355
1.d−10	0.57799	1.87009	3.2355

Table 1b
Test 1. Extreme eigenvalues and spectral condition number computed by the $P1$ Galerkin scheme for different D_{22} values (from [32]).

D_{22}	λ_{min}	λ_{max}	$\lambda_{max}/\lambda_{min}$
1.d+00	0.8803	4.8407	5.4989
1.d−01	0.72129	3.0363	4.2095
1.d−04	0.6348	2.954	4.6534
1.d−06	0.63468	2.9539	4.6542
1.d−10	0.63468	2.9539	4.6542

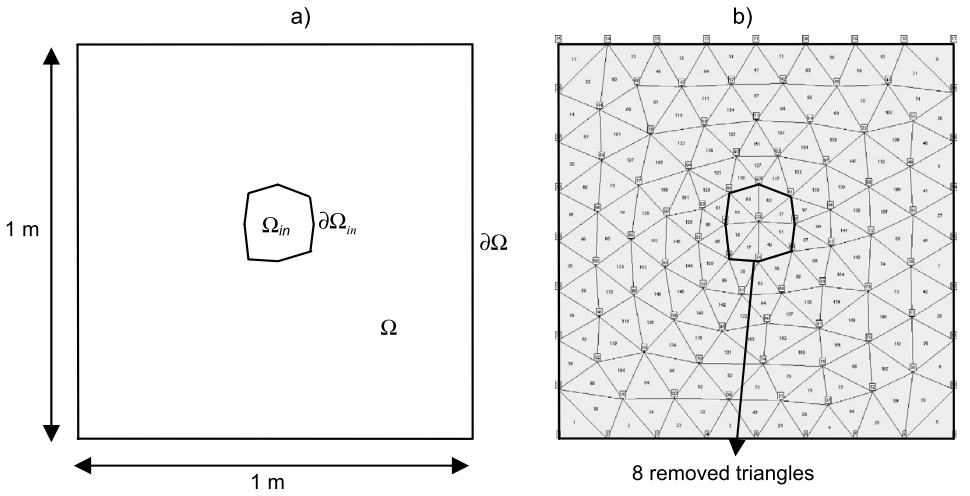

Fig. 10. Test 2. (a) Spatial domain; (b) coarse computational mesh.

coefficients of the stiffness matrix computed by the $P1$ Galerkin scheme are positive and M-property is lost. This could explain the higher condition numbers in comparison with the corresponding ones of the proposed procedure.

In Fig. 8(b) the final GAD mesh obtained by the proposed procedure is reported after edge swaps for $D_{22} = 1.d{-}06$ and $1.d{-}10$. Observe that triangles tend to align along the principal anisotropy direction (x_1 in this case).

In Appendix C it is proved that the proposed procedure maintains Positive Transmissibility (PT) condition (non-positive stiffness matrix coefficients), before performing the edge swap, for more general anisotropic conditions than the standard $P1$ Galerkin scheme. PT condition is a prerequisite for the existence of an M-matrix (see for example [35]).

7.2. Test 2. DMP property: part 1

We solve the following steady-state diffusion problem over the domain $\Omega = \{[0,1]^2/\Omega_{in}\}$, where Ω_{in} is an internal hole of Ω (see Fig. 10(a))

$$-\nabla\cdot\mathbf{D}\nabla u = 0, \qquad u_D = 0, \quad \mathbf{x}\in\Gamma, \qquad u_D = 2, \quad \mathbf{x}\in\Gamma_{in}, \tag{37}$$

and Γ_{in} is the boundary of Ω_{in}. The internal hole has been obtained by removing eight triangles of a coarse starting domain triangulation of the square domain $[0,1]^2$ (164 triangles and 99 nodes, shown in Fig. 10(b)). This test is inspired by [27], where Ω_{in} is the square domain $[4/9, 5/9]^2$. Starting from the above coarse triangulation, three refinements have been performed by halving each side. Diffusion coefficient is given by its eigen-decomposition [27]:

$$\mathbf{D} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{38}$$

where $\lambda_{1(2)}$ are the eigenvalues of $\mathbf{D}$ and θ is the angle of primary diffusion direction with x_1-axis (parallel to the first eigenvectors of $\mathbf{D}$). The eigenvalues are assumed constant ($\lambda_1 = 1000$, $\lambda_2 = 1$) and the problem is solved using both a constant and a variable θ angles [27].

In the constant $\mathbf{D}$ case, $\theta = \pi/4$. From the coarse to the most refined mesh level, the algorithm always satisfies the DMP condition, since numerical solution is bounded between 0 and 2 and no unphysical oscillation occurs. Figs. 11(a) and 11(b) show the contours of the iso-u lines and the 3D u profile computed by the proposed procedure using the starting triangulation obtained after the 2nd refinement level (2496 triangles and 1328 nodes). Fig. 12 shows the final GAD mesh obtained by the proposed algorithm after the edge swaps iterative procedure. Observe that triangles dispose along the principal anisotropy direction (they are rotated of angle θ with respect to x_1).

In paper [27], the authors develop a mesh adaptation under which a linear FE approximation of the anisotropic diffusion problem in Eq. (37) satisfies the DMP. This condition is the extension of the non-obtuse angle condition for isotropic diffusion problems and requires that the dihedral angles of mesh elements, measured in a metric depending on the diffusion tensor, be non-obtuse. The authors in [27] develop a metric tensor to use in anisotropic mesh generation based on the anisotropic non-obtuse angle condition and present also two variants of the anisotropic non-obtuse angle condition satisfying the DMP. Applying the two variants to a uniform anisotropic mesh metric, the authors in [27] obtain numerical solutions without overshoots ($u > 2$) and undershoots ($u < 0$). In a uniform mesh metric, solution is characterized by undershoots decreasing with mesh refinement.

In the variable $\mathbf{D}$ case, $\theta = \pi\,\mathrm{sen}(x_1)\cos(x_2)$. Numerical solution computed by the proposed procedure is also in this case bounded by 0 and 2, without unphysical oscillations, similarly to the solution computed by Li and Huang [27] applying

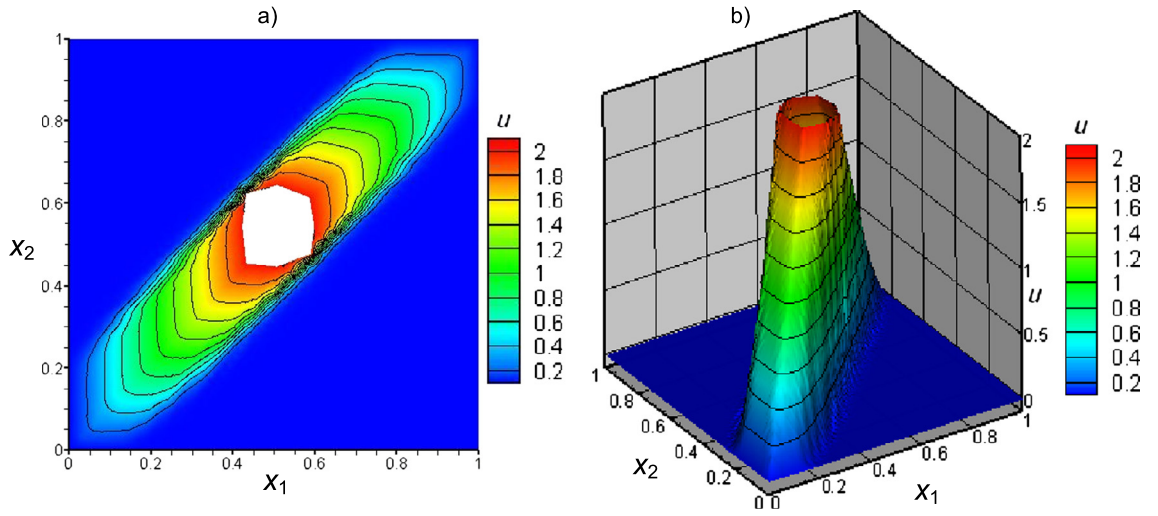

Fig. 11. Test 2. $\theta = \pi/4$: (a) contours of the iso-u lines; (b) 3D u profile (2nd refinement level, 2496 triangles and 1328 nodes).

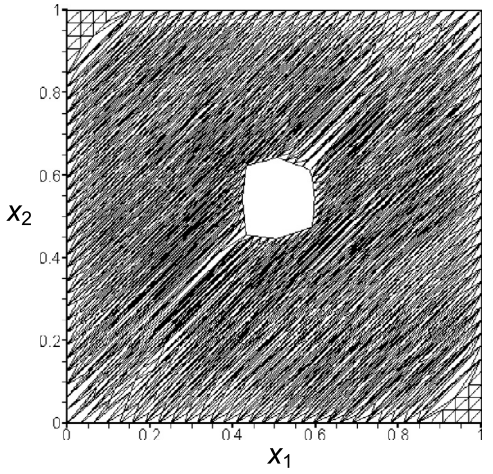

Fig. 12. Test 2. $\theta = \pi/4$: final computed GAD mesh after the edge swap (2nd refinement level).

the two variants to the uniform metric. Figs. 13(a) and 13(b) show the contours of the iso-u lines and the 3D u profile computed using the 2nd mesh refinement level and Fig. 14 the final computed GAD mesh after the edge swaps for the same refinement level.

7.3. Test 3. DMP property: part 2

The following BVP is solved over the square domain $\Omega = [0,1]^2$:

$$\begin{cases} -\nabla \cdot \mathbf{D}\nabla u = 0, \\ u_D(0,0) = 0, \quad u_D(1,1) = 200, \quad u_D|_{\Gamma_D/[(0,0),(1,1)]} = 100, \end{cases} \tag{39}$$

and the diffusion tensor is discontinuous: in the sub-regions $(0 \leqslant x_1 \leqslant 1) \times (0 \leqslant x_2 \leqslant 1/3)$ and $(0 \leqslant x_1 \leqslant 1) \times (2/3 < x_2 \leqslant 1)$ we have:

$$\mathbf{D} = \begin{pmatrix} 2464.36002 & 1148.683643 \\ 1148.683643 & 536.6399794 \end{pmatrix}, \tag{40a}$$

while in the sub-region $(0 \leqslant x_1 \leqslant 1) \times (1/3 < x_2 \leqslant 2/3)$ we have:

$$\mathbf{D} = \begin{pmatrix} 2464.36002 & -1148.683643 \\ -1148.683643 & 536.6399794 \end{pmatrix}. \tag{40b}$$

In each sub-region principal axes directions are oriented at 25 degrees (plus, minus and plus 25 degrees) with respect to the (x_1, x_2) reference system and tensor $\mathbf{D}$ has a principal anisotropy ratio equal to 3000:1. This test has been proposed

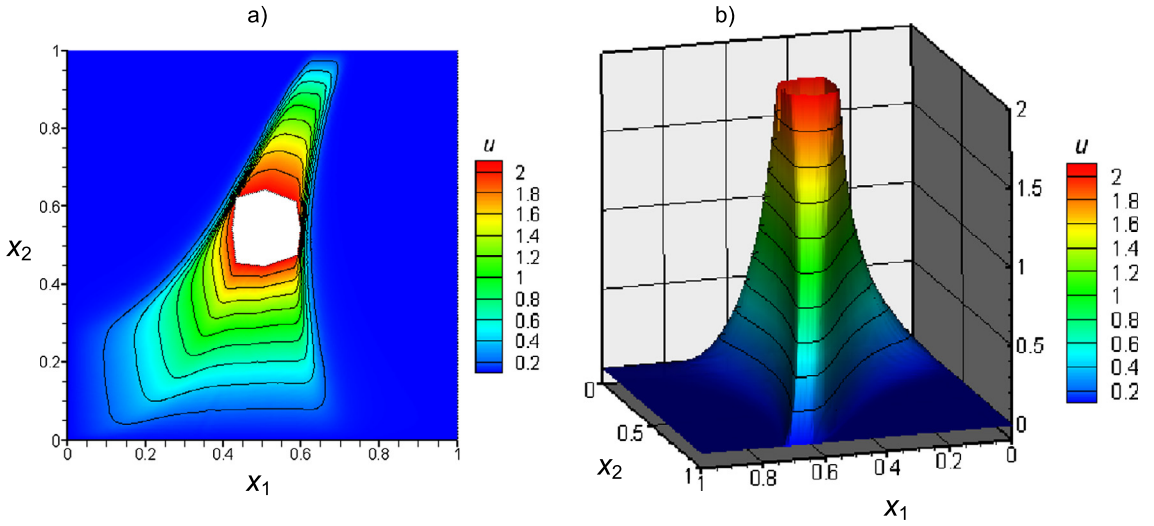

Fig. 13. Test 2. $\theta = \pi\,\mathrm{sen}(x_1)\cos(x_2)$: (a) contours of the iso-u lines; (b) 3D u profile (2nd refinement level, 2496 triangles and 1328 nodes).

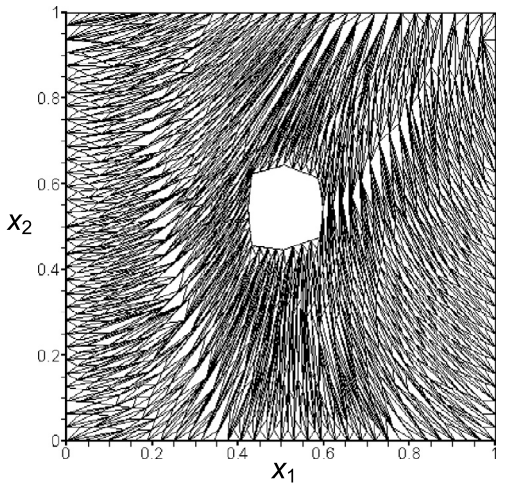

Fig. 14. Test 2. $\theta = \pi\,\mathrm{sen}(x_1)\cos(x_2)$: final computed GAD mesh after the edge swap (2nd refinement level).

by Edwards and Zheng [14]. The authors in [14] present results computed by two different Multi-Point Flux Approximations (MPFA): a linear Triangle Pressure Support (TPS) scheme and a Full-Pressure Support (FPS) scheme. The authors in [14] discretize the domain with a (65×65) grid resolution, using triangular (for TPS scheme) and quadrilateral (for FPS scheme) elements. When quadrilateral elements are used, the control volume boundaries are aligned on the interfaces between jumps in the diffusion tensor while, when triangular elements are used, the control volume faces are aligned along interior boundaries crossing the diffusion tensor jumps [14]. The TPS scheme violates the MP property, producing unphysical oscillations in the computed solutions. The FPS scheme computes solutions almost free of oscillations (see [14]).

For the proposed approach, the domain is discretized using a starting triangulation with 2560 triangles and 1353 nodes. This has been obtained by refining twice a coarse triangulation with 160 triangles and 99 nodes. Triangulation nodes are aligned along the lines separating the domain sub-regions.

Two different input data sets have been tested: in the first one, no internal boundaries have been assigned; in the second one, internal boundaries overlapping the jumps of the diffusion tensors have been assigned. The computed contours of the iso-u, the 3D u profile and the final GAD mesh after the edge swaps are respectively shown in Figs. 15(a), 15(b) and 16 for the first case, in Figs. 17(a), 17(b) and 18 for the second case.

Computed solutions are in both cases free of spurious oscillations, but iso-u profiles obtained by setting the internal boundaries are a little bit sharper than the ones computed without internal boundaries. Compare results in Figs. 15(a) and 17(a). Assuming a mean value for the directional components of the diffusion tensor for heterogeneous medium (Operation II in Section 6) provides a smoothing in the computed results.

We have shown at the end of Section 5 that we enforce a possible artificial reduction of the anisotropy ratio along the boundaries, in order to guarantee the M-property. The amount of this reduction can be easily lowered by increasing

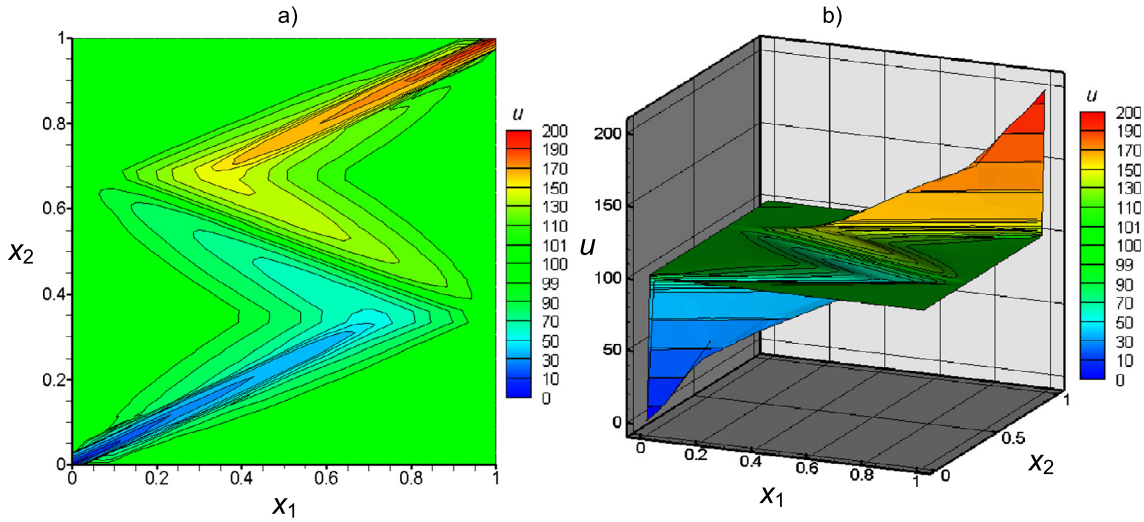

Fig. 15. Test 3. Situation 1 (no internal boundaries). (a) Contours of the iso-u; (b) 3D u profile (2560 triangles and 1353 nodes).

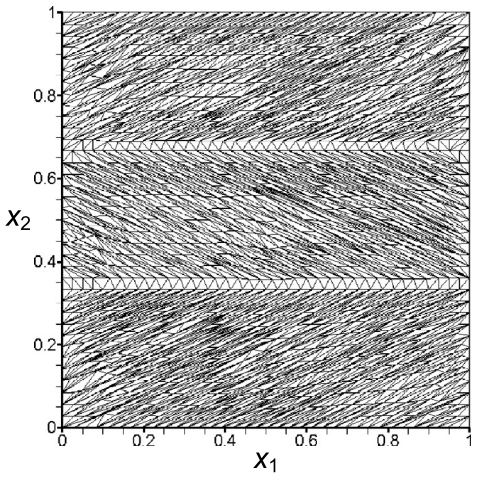

Fig. 16. Test 3. Situation 1 (no internal boundaries). Final computed GAD mesh after the edge swap.

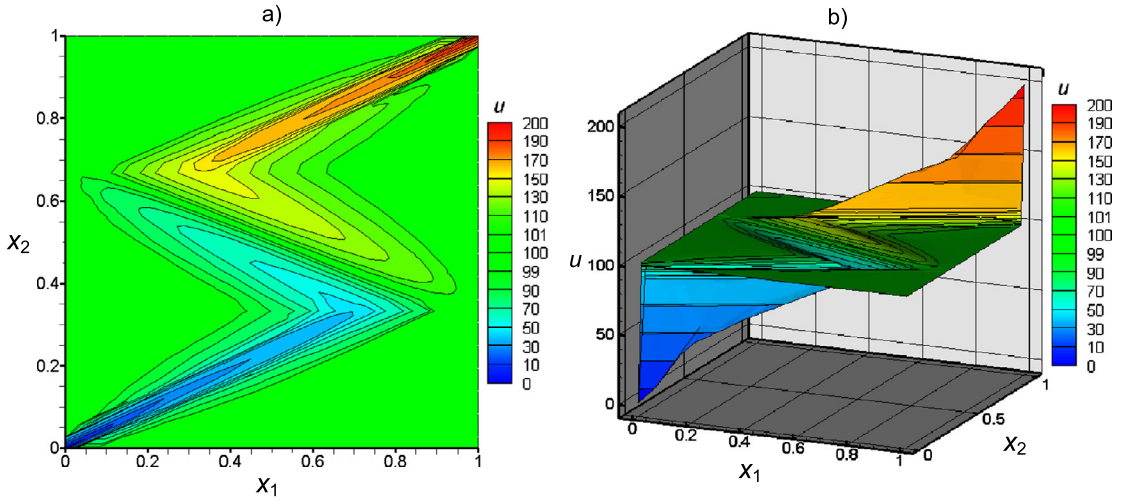

Fig. 17. Test 3. Situation 2 (internal fixed boundaries at **D** jumps). (a) Contours of the iso-u; (b) 3D u profile (2560 triangles and 1353 nodes).

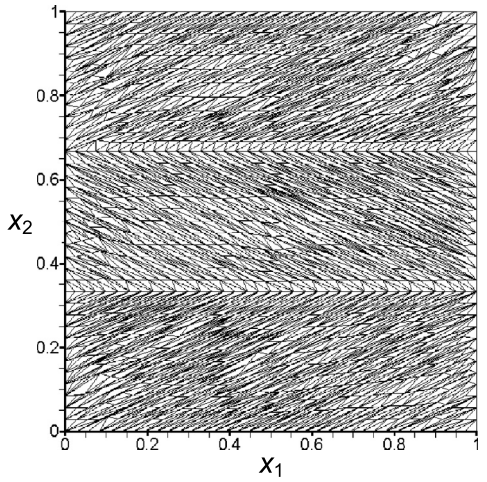

Fig. 18. Test 3. Situation 2 (internal boundaries at **D** jumps). Final computed GAD mesh after the edge swap.

the number of nodes along the boundaries in the starting domain triangulations. The reason is that, by increasing the number of nodes along the boundary, the angles at the vertices opposite to the boundary side of the boundary elements become smaller and smaller. This implies that the anisotropic circumcentre falls closer and closer to the midpoints of the two interior edges and the distance between the circumcentre and the midpoint of the boundary edge becomes closer and closer to a strictly positive value (that is half the triangle height).

To verify this assumption, the following numerical experiment has been carried out. We run the simulations over four different starting domain triangulations. The first triangulation is the same coarse triangulation with 160 triangles and 99 nodes, the other three triangulations have been obtained refining the coarse one by halving the edges only along the internal boundaries dividing two zones with different diffusion tensor. A GAD mesh has been obtained for each refinement. The normalized changes of the diffusion tensor components, with respect to the original ones, due to the reduction of the anisotropy ratio at the internal boundary (Eqs. (34)–(35)) have been computed. As expected, these changes decrease with hyperbolic law by increasing the number of nodes along the fixed boundary of the starting domain triangulations and asymptotically go to zero.

7.4. Test 4. Estimation of the convergence order: part 1

Convergence rate is estimated by comparing an analytical solution to the results obtained with a series of refined triangulations. We solve over the square domain $\Omega = [0,1]^2$ the following full Dirichlet BVP previously presented by Gao and Wu in [16],

$$-\nabla \cdot \mathbf{D}\nabla u = f, \qquad u_D = u_{ex}, \quad \mathbf{x} \in \Gamma, \tag{41}$$

where the source term f on the r.h.s. is computed by space differentiating the same solution on the l.h.s. of the same Eq. (41). The imposed arbitrary analytical solution u_{ex} presents two opposite steep slopes in the (x_1, x_2) plane and is given by:

$$u_{ex} = \begin{cases} \exp(-20\pi\,(x_2 - 0.5)^2)[1 + (x_1 - 0.5)(0.1 + 8\pi\,(x_2 - 0.5))], & x_1 \leqslant 0.5, \\ \exp(-20\pi\,(x_2 - 0.5)^2)\exp(x_1 - 0.5), & x_1 > 0.5, \end{cases} \tag{42}$$

and it is represented in Fig. 7 of paper [16]. The diffusion tensor **D** changes across $x_1 = 0.5$ as:

$$\mathbf{D} = \begin{cases} \begin{pmatrix} 10 & 2 \\ 2 & 5 \end{pmatrix} & \text{if } x_1 \leqslant 0.5, \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{if } x_1 > 0.5. \end{cases} \tag{43}$$

Computational domain has been discretized using the starting triangulation shown in Fig. 19(a) (278 elements and 164 nodes) and five refinements have been performed by halving each side.

L_2 norm of the relative error corresponding to the lth refinement level is evaluated as:

$$err_l = \frac{\sqrt{\sum_{i=1,N}(u_i - u_{ex,i})^2}}{\sqrt{\sum_{i=1,N}(u_{ex,i})^2}}, \tag{44}$$

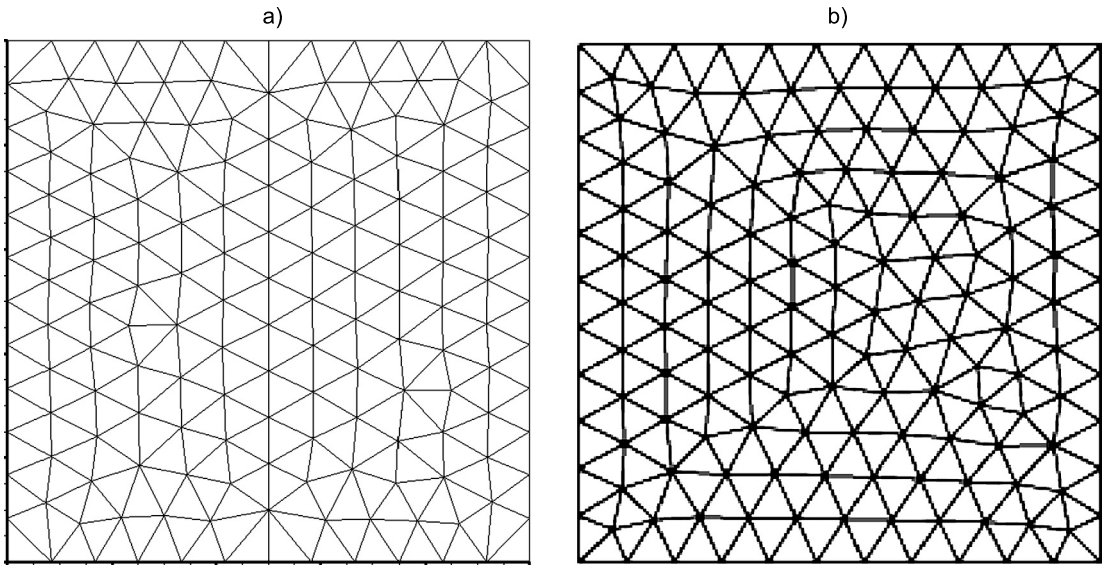

Fig. 19. Coarse GD mesh: (a) mesh for test 4; (b) mesh for test 6.

Table 2
Test 4. Number of elements, L_2 norm of relative errors and convergence order.

Refinement level l	Nel	er_l	r_l
0	278	6.56d−02	–
1	1112	2.48d−02	1.4
2	4448	8.06d−03	1.62
3	17792	2.15d−03	1.9
4	71168	4.95d−04	2.11
5	284672	1.12d−04	2.14

where u_i and $u_{ex,i}$ are the computed and the exact solutions at node i. The rate of convergence is defined by comparing the relative errors of two consecutive refinement levels. We assume the relative error computed for mesh level l proportional to a power of the linear size of the area of the mean triangle in the mesh,

$$err_l = \left(\sqrt{|T|_l}\right)^{r_c}, \tag{45}$$

where $|T|_l$ is the area of the mean triangle at refinement level l and $\sqrt{|T|_l}$ represents a measure of its linear size. The rate of convergence r_c is computed by comparing the relative errors of two successive refinement levels l and $l+1$:

$$r_c = \frac{\log(\frac{err_l}{err_{l+1}})}{\log(2)}. \tag{46}$$

Table 2 shows the L_2 norms of the relative errors and the convergence order. Convergence order increases with mesh refinement and reaches values higher than 2. The growth of the convergence order along with the mesh density is very important, because it is indicative of stable results also when a coarse mesh is used instead of a very refined one.

7.5. Test 5. Estimation of the convergence order: part 2

The convergence of the proposed method is studied for the BVP in Eq. (41) over square domain $[0,1]^2$ with full Dirichlet conditions and heterogeneous anisotropic diffusion tensor. This problem has been proposed in [16]. The imposed analytical solution is:

$$u_{ex} = \exp\left(-20\pi\left((x_1-0.5)^2+(x_2-0.5)^2\right)\right), \tag{47}$$

and the diffusion tensor is continuously space dependent as:

$$\mathbf{D} = \begin{pmatrix} \alpha x_1^2 + x_2^2 & (\alpha-1)x_1x_2 \\ (\alpha-1)x_1x_2 & \alpha x_2^2 + x_1^2 \end{pmatrix}, \tag{48}$$

Table 3a
Test 5. L_2 norm of relative errors, convergence order and extreme values: $\alpha = 1000$, structured mesh.

Refinement level l	α 1000			
	err_l	r_l	u_{min}	u_{max}
0	3.6d−01		−2.3d−02	1.37d+00
1	9.5d−02	1.92	−1.d−02	1.12d+00
2	2.2d−02	2.11	−8.d−03	1.002d+00
3	5.d−03	2.14	−3.d−03	1.d+00
4	1.2d−03	2.06	−3.d−04	1.d+00
5	2.1d−04	2.51	−1.7d−04	1.d+00

Table 3b
Test 5. L_2 norm of relative errors, convergence order and extreme values: $\alpha = 100$, structured mesh.

Refinement level l	α 100			
	err_l	r_l	u_{min}	u_{max}
0	2.8d−01		−1.4d−02	1.31d+00
1	8.1d−02	1.79	−6.d−03	1.01d+00
2	2.d−02	2.02	−2.1d−03	1.0071d+00
3	4.3d−03	2.22	−5.d−04	1.d+00
4	9.4d−04	2.19	−1.d−04	1.d+00
5	1.6d−04	2.55	−4.7d−05	1.d+00

Table 3c
Test 5. L_2 norm of relative errors, convergence order and extreme values: $\alpha = 10$, structured mesh.

Refinement level l	α 10			
	err_l	r_l	u_{min}	u_{max}
0	9.75d−02		−1.97d−02	1.07d+00
1	3.d−02	1.7	−1.d−02	1.002d+00
2	8.7d−03	1.79	−2.3d−03	1.0005d+00
3	2.d−03	2.12	−7.4d−04	1.d+00
4	4.5d−04	2.15	−7.6d−05	1.d+00
5	1.d−04	2.17	0.d+00	1.d+00

where α is a real number greater than 1. Source term f, computed as for the previous test 4, corresponds to an injection at the centre of the domain located between two sinks and its magnitude depends on the value of α coefficient [16]. $\alpha = 1000$, 100 and 10 have been tested. Spatial source term distribution for $\alpha = 1000$ is shown in Fig. 14 of paper [16]. The exact solution is bounded from 0 and 1.

Two coarse starting domain triangulations have been considered for the presented model simulations: the first is structured with isosceles rectangle elements (512 triangles and 289 nodes), the second is the unstructured one shown in Fig. 10(b) (164 triangles and 99 nodes). Five refinement levels have been performed for both triangulations by halving each element side.

Tables 3a–3c show the L_2 norms of relative errors and convergence orders computed by the proposed procedure using the structured mesh. Overshoots and undershoots ($u_{min} = -2.3\text{d}{-}02$ and $u_{max} = 1.37$) affect results obtained using the coarse mesh and $\alpha = 1000$. These differences from the exact extreme values significantly decrease for the same coarse mesh by reducing α or refining the mesh and keeping $\alpha = 1000$. Convergence order is about 2 for all the investigated α values.

Gao and Wu [16] compared the efficiency of four numerical schemes: the lowest-order Raviart–Thomas Mixed Finite Element Method (MFEM), a MPFA and two FV linearity preserving schemes (LPEW1 and LPEW2). The authors in [16] used a MPFA for quadrilateral grids proposed in [1,22], where the domain of the governing PDEs is covered by non-overlapping quadrilateral interaction regions, obtained by linking the mass centres of each cell with the midpoint of cell interfaces. This method is characterized by linear variation of the unknown variables and by a single condition for flux and unknown variable at each of the four edges of the interaction region. LPEW1 and LPEW2 are both linearity preserving schemes, where the flux on each cell edge is expressed as function of two cell-centred unknowns defined on the cells sharing that edge and two vertex unknowns at the two edge endpoints [16]. The vertex unknowns are expressed as a linear combination of the neighboring cell-centred unknowns [16]. Gao and Wu [16] discretize the square domain with the same structured mesh adopted for the simulations of the present model. The MPFA used in [16] (proposed in [1,22]) shows no convergence for $\alpha = 100$ and 1000 on the most refined mesh and LPEW1 shows no convergence for $\alpha = 1000$ for the most refined meshes. For $\alpha = 1000$, MFEM computes significant overshoots and undershoots with the coarse mesh ($u_{min} = -4.1$ and $u_{max} = 9.3$, see [16]). LPEW1 and LPEW2 provide solutions free of spurious oscillations and convergence order increases from approximately 1.5 to 2 with α decreasing from 1000 to 10. MPFA and MFEM schemes have convergence order 2, as expected.

Table 4a
Test 5. L_2 norm of relative errors, convergence order and extreme values: $\alpha = 1000$, unstructured mesh.

Refinement level l	α 1000			
	err_l	r_l	u_{min}	u_{max}
0	4.1d+00	–	−9.d−01	4.36d+00
1	4.4d−01	3.22	−1.7d−01	1.31d+00
2	8.d−02	2.46	−1.01d−01	1.15d+00
3	1.5d−02	2.42	−1.07d−02	1.04d+00
4	3.5d−03	2.1	−8.2d−03	1.002d+00
5	7.d−04	2.32	−7.2d−04	1.d+00

Table 4b
Test 5. L_2 norm of relative errors, convergence order and extreme values: $\alpha = 100$, unstructured mesh.

Refinement level l	α 100			
	err_l	r_l	u_{min}	u_{max}
0	1.74d+00	–	−6.d−01	1.6d+00
1	3.1d−01	2.49	−1.d−01	1.2d+00
2	5.8d−02	2.42	−4.2d−02	1.03d+00
3	1.3d−02	2.16	−8.2d−03	1.0063d+00
4	2.5d−03	2.38	−4.2d−03	1.d+00
5	5.d−04	2.32	−5.41d−04	1.d+00

Table 4c
Test 5. L_2 norm of relative errors, convergence order and extreme values: $\alpha = 10$, unstructured mesh.

Refinement level l	α 10			
	err_l	r_l	u_{min}	u_{max}
0	2.4d−01	–	−5.1d−02	1.19d+00
1	7.79d−02	1.62	−2.d−02	1.03d+00
2	1.9d−02	2.04	−4.d−03	1.001d+00
3	4.79d−03	1.99	−2.d−03	1.d+00
4	1.2d−03	2.0	−1.02d−04	1.d+00
5	2.6d−04	2.21	−2.5d−05	1.d+00

Higher overshoots and undershoots affect the solutions of the proposed procedure computed with the unstructured mesh in Fig. 10(b), as shown in Tables 4a–4c, but, similarly to the simulations over the structured mesh, overshoots and undershoots drop along with the mesh refinement for $\alpha = 1000$, or for the same coarse mesh by reducing α. Refining the mesh, convergence order remains 2 for all the range of the α anisotropy coefficient. The differences between the exact and the numerical solution of the proposed model can be explained with the spatial discretization of the source term, function, in his turn, of the spatial discretization of the components of the diffusion tensor. Due to the heterogeneity of the medium, this spatial discretization is also influenced by the smoothing process (see procedure in Section 6), required to maintain the stiffness matrix "nice" characteristics.

The smoothing operation for heterogeneous medium, described in Section 6, can change the initial average $\mathbf{D}'$ values inside each triangle. The higher is this change, the higher is the angle between the eigenvectors of the original and the "smoothed" $\mathbf{D}'$. Fig. 20 shows (for different triangulation refinement levels), the angles between the eigenvectors of the original $\mathbf{D}'$ computed in each element according to Eq. (48) and the eigenvectors of the $\mathbf{D}'$ components computed after performing the smoothing procedure depicted in Fig. 7. Observe that the computed angles become smaller and smaller along with the mesh refinement. This could also justify the reduction of the overshoots and undershoots of the numerical solution, before discussed. Fig. 20 is related to the case of $\alpha = 1000$ and the unstructured mesh. Similar results have also been obtained for $\alpha = 100$ and $\alpha = 10$, and for the structured mesh, but are not shown here for brevity.

7.6. Test 6. Mesh locking investigation

The performance of the proposed model is evaluated by studying the following parametric BVP over the square domain $[0,1]^2$:

$$\begin{cases} -\nabla \cdot \mathbf{D}\nabla u = f, \\ u = u_D, \quad \mathbf{x} \in \Gamma_D, \quad \mathbf{q}(\mathbf{x},t)\cdot\mathbf{n} = g_N(\mathbf{x},t), \quad \mathbf{x} \in \Gamma_N, \end{cases} \tag{49}$$

where the diffusion tensor is:

$$\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \end{pmatrix}, \tag{50}$$

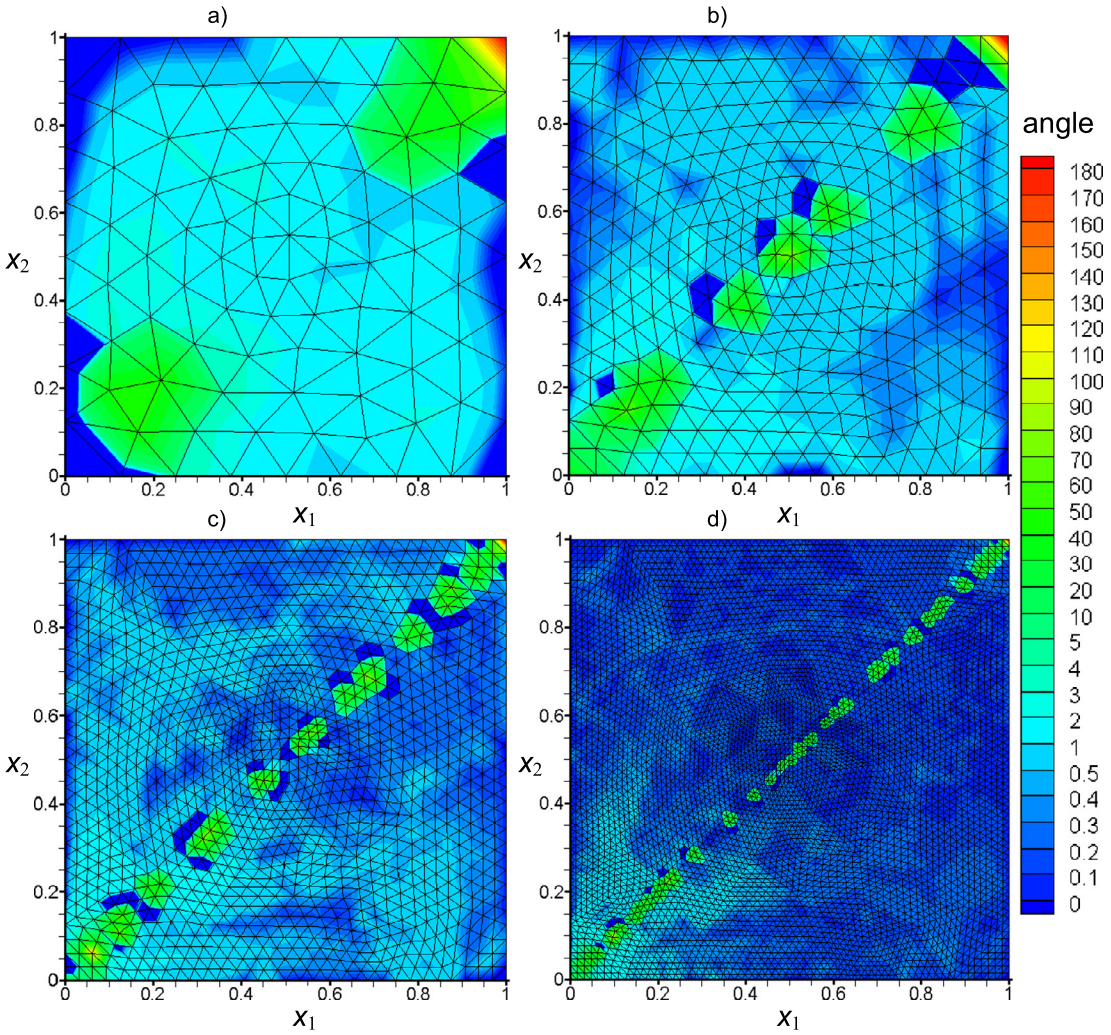

Fig. 20. Test 4. Angles between the eigenvectors of the original average element $\mathbf{D}'$ and the eigenvectors of the $\mathbf{D}'$ components computed after performing "smoothing" operation. (a) Coarse mesh; (b) 1st refinement level; (c) 2nd refinement level; (d) 3rd refinement level.

with parameter ε a positive real number in the range $[10^{-6}, 1]$. Definition of portions of boundary Γ_D and Γ_N depends on the particular selected case, as further specified. We assume the exact solution

$$u_{ex} = \exp(-2\pi\sqrt{\varepsilon}x_1)\sin(2\pi x_2), \tag{51}$$

parametrically dependent on ε in the x_1 direction so that the corresponding source term f is zero. Observe that maximum and minimum values are located on the domain boundary and that the solution becomes almost constant in x_1 direction for small ε values. This test has been proposed in [6,31] and [16].

Three types of boundary condition have been considered:

Case A. Dirichlet boundary conditions, with $\Gamma_N = \emptyset$ and $u = u_D$ on $\Gamma = \Gamma_D = \{(x_1, x_2) \mid x_1 = 0, 1 \text{ or } x_2 = 0, 1\}$.

Case B. Mixed Dirichlet and Neumann boundary conditions, with $u = u_D$ on $\Gamma_D = \{(x_1, x_2) \mid x_1 = 0 \text{ or } x_2 = 0\}$ and $\mathbf{q} \cdot \mathbf{n} = g_N$ on $\Gamma_N = \{(x_1, x_2) \mid x_1 = 1 \text{ or } x_2 = 1\}$.

Case C. Nearly pure Neumann boundary conditions, with $u = u_D$ on $\Gamma_D = \{(x_1, x_2) \mid 1 - h \leqslant x_1 \leqslant 1,\ x_2 = 1 \text{ or } x_1 = 1,\ 1 - h \leqslant x_2 \leqslant 1\}$ (with h the characteristic linear mesh size) and $\mathbf{q} \cdot \mathbf{n} = g_N$ on $\Gamma_N = \Gamma / \Gamma_D$.

Case C can be regarded as a modification of the fully Neumann boundary condition. This last leads to a singular discrete Laplacian operator, while the insertion of the Dirichlet condition on the two boundary edges located at the upper right

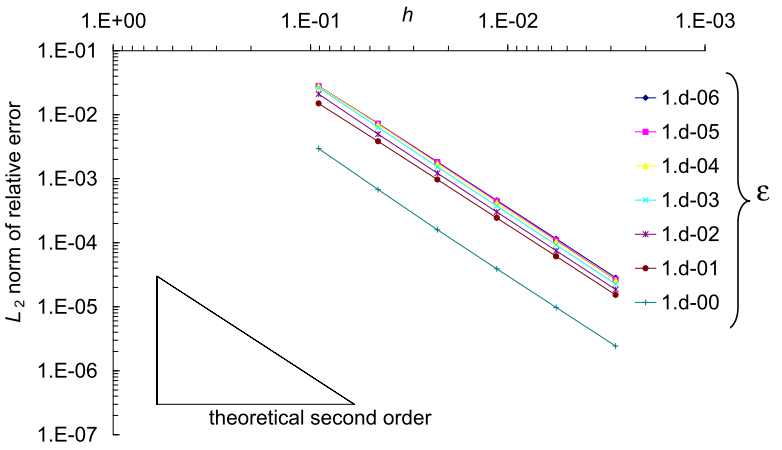

Fig. 21a. Test 6, Case A. L_2 norms of relative errors versus characteristic length size h computed by the proposed procedure for different ε values.

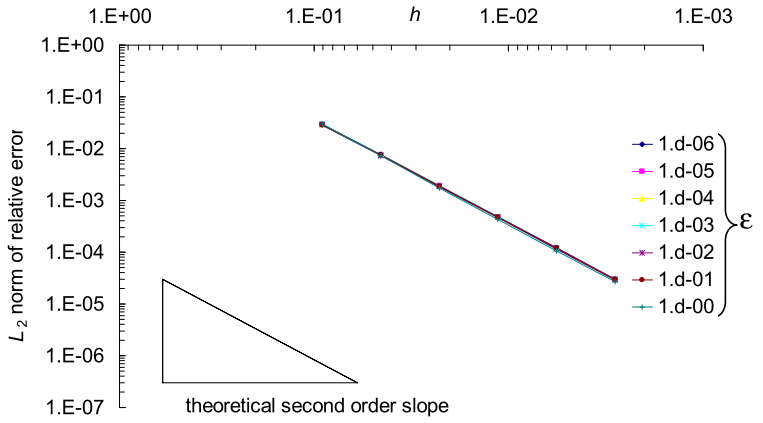

Fig. 21b. Test 6, Case B. L_2 norms of relative errors versus characteristic length size h computed by the proposed procedure for different ε values.

corner is used to remove this singularity. The length of these edges goes to zero as $h \to 0$. Observe that maximum and minimum solution values, located at the boundaries, are difficult to obtain with Neumann boundary conditions.

As stated in [19], locking, or parametric error amplification, is a well-known phenomenon that may arise when solving parametric elliptic problems with the FE method. Locking usually effectively prohibits the convergence of low-order FE schemes when the parameter associated with the problem, ε, approaches an asymptotic value such as zero. The solution may fall in two different asymptotic states, depending on the boundary conditions and on the load. These states could be named as the *cool* state and the *hot* state, respectively. In the cool state, the main conduction occurs in the direction of high conductivity (as normally expected), whereas in the hot state, the conduction in the direction of low conductivity dominates [19]. Cool state is induced by boundary condition as in Cases A and B, while hot state is induced by boundary condition of type C [19,31].

The square domain has been discretized with the starting triangulation in Fig. 19(b) and five refinements have been performed by halving each triangle side. Figs. 21(a)–21(c) show the L_2 norms of relative errors versus characteristic length size h for Cases A to C. In the same figures, the theoretical second-order convergence curves are shown too on the bottom left side.

In Figs. 22(a)–22(c) the analogous results computed by Manzini and Putti [31] are shown. The authors propose a second-order accurate cell-centered FV scheme based on edge-centered piecewise constant definition of the solution gradients that takes into account both normal and tangential components of the gradients in the formulation of the local flux balance equations.

Mesh locking effects on convergence curves, for decreasing ε values, are essentially two: (1) they move upward, with a corresponding error increment and (2) they become flat or near flat, with a loss in the convergence rate [31].

For Case A (full Dirichlet problem), both the compared schemes do not show locking effect. The isotropic case ($\varepsilon = 1$) provides the highest errors in the FV scheme of Manzini and Putti [31], while the highest errors in the proposed method are computed for the lowest ε. The reasons could be the following. Manzini and Putti [31] affirm that the tangential flux terms (due to anisotropy and equal to zero if $\varepsilon = 1$) are essential to maintain the consistency of their scheme [31]. Since

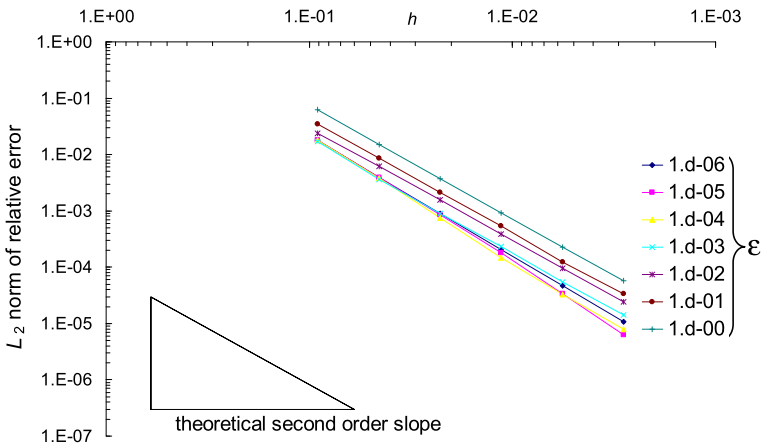

Fig. 21c. Test 6, Case C. L_2 norms of relative errors versus characteristic length size h computed by the proposed procedure for different ε values.

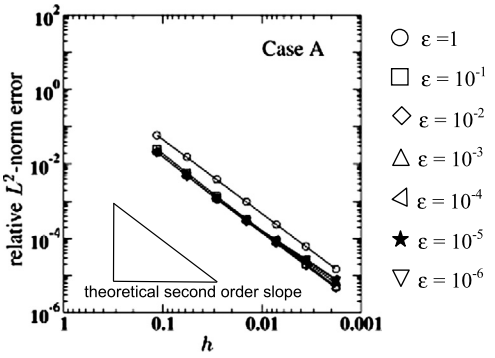

Fig. 22a. Test 6, Case A. L_2 norms of relative errors versus characteristic length size h computed by Manzini and Putti [31] for different ε values (from [31]).

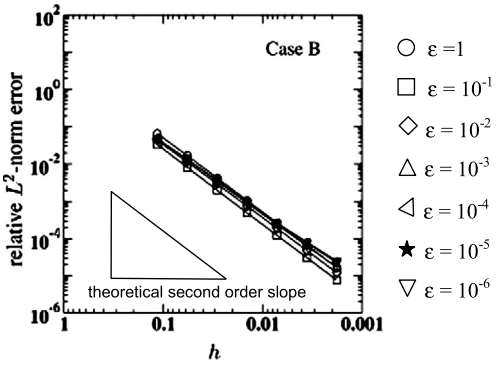

Fig. 22b. Test 6, Case B. L_2 norms of relative errors versus characteristic length size h computed by Manzini and Putti [31] for different ε values (from [31]).

the starting domain triangulation is acute, in the proposed methodology the most regular situation corresponds to isotropic condition. Reducing ε, triangular elements tend to lengthen and to align along the principal anisotropy direction (x_1) and, in some cases, circumcentres of two triangles sharing the same edge fall on the same side with respect to that edge and swapping is necessary. This introduces an error in the boundary edges, where Dirichlet conditions are imposed. Both numerical procedures show second convergence order.

For Case B (Dirichlet and Neumann boundary conditions), especially for the proposed methodology, errors seem to be quite independent of the anisotropy ratio. Both compared schemes do not present locking phenomena and the convergence order is the same as in Case A.

For Case C (quasi-Neumann boundary conditions), only the solution computed by the proposed method is free of locking effects and the scheme maintains second convergence order for all the investigated anisotropy ratios. Solution of the FV

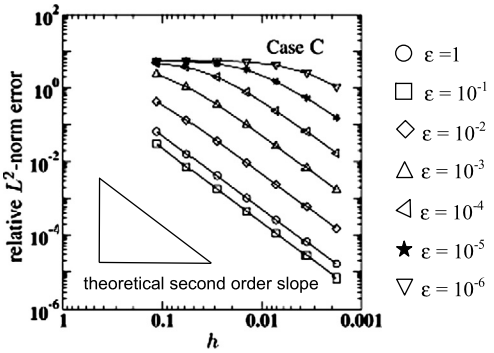

Fig. 22c. Test 6, Case C. L_2 norms of relative errors versus characteristic length size h computed by Manzini and Putti [31] for different ε values (from [31]).

Table 5
Test 7. Number of elements, L_2 norm of relative errors and convergence order.

Refinement level l	*Nel*	er_l	r_l
0	278	6.35d−02	
1	1112	2.45d−02	1.37
2	4448	8.02d−03	1.61
3	17792	2.15d−03	1.9
4	71168	4.94d−04	2.12
5	284672	1.12d−04	2.14

scheme [31], especially for the highest anisotropy ratios, is affected by mesh locking. In the worse case the errors increase by a factor greater than 100 for the coarse mesh and the convergence curves become flat.

Gao and Wu [16] present an analogous test case where the principal anisotropy direction is x_2 instead of x_1. They compute approximately a second convergence order for their proposed LPEW1 and LPEW2 FV schemes for Cases A and B, but also observe locking effects for the quasi-Neumann case for the highest anisotropy ratios.

7.7. Test 7. Time-dependent problem

In this final test we investigate the convergence order computed by the proposed method for a time-dependent problem with discontinuous diffusion tensor and internal boundaries. We assume the same physical case as in test 4 and the imposed analytical solution u_{ex} is a time-dependent variation of the one in test 4:

$$u_{ex} = \left(1 - \exp\left(\frac{-t}{T_0}\right)\right)\breve{u}_{ex}, \tag{52}$$

where $\breve{u}_{ex}$ is the analytical solution given for test 4 (Eq. (42)) and T_0 is a constant time value equal to 60 s. We use the same starting domain triangulations as for test 4 (the coarse one in Fig. 19(a)).

We distinguish two different cases, the first without and the second with internal boundaries along the jumps of the diffusion tensor. L_2 norms of relative errors for the case without internal boundaries are essentially coincident with the values computed in Table 2 for the steady-state case and the convergence order increases from 1.4 up to 2.14. For brevity results are not reported here.

In Table 5, the L_2 norms of the relative errors and the convergence order are reported along with elements number for each mesh refinement for the scenario with internal boundaries. Computed errors are a bit smaller than the ones computed without internal boundaries (see Table 2), especially in the coarse triangulations. The reason of the bit higher convergence order of the first case can be explained in the following way. Refining the mesh without internal boundaries, the possible areal extension of the error of the tensor coefficients is reduced (because the error is proportional to the maximum distance of a node from the tensor coefficients discontinuity line). This implies a further reduction of the error, that is missing when an internal boundary is applied.

8. Investigation of the computational costs

Test 2 with variable θ has been used for investigation of the computational cost (CPU times). Starting from the coarse domain triangulation shown in Fig. 10(b) (164 triangles and 99 nodes), three refinements have been performed by halving each side. In Table 6, the mean computational time in seconds required by the procedure for flux coefficients estimation that satisfies the M-property in the general heterogeneous anisotropic case, as well as the mean computational time required for the solution of the final system are reported. The mean CPU time required for flux coefficients estimation has been obtained

Table 6
Mean CPU times (test 2 – variable θ).

Refinement level	N_T	N	L	Swapped edges	CPU M-property [s]	CPU system solution [s]
0	156	98	254	82	0.d+00	0.d+00
1	624	352	976	611	6.35d−07	4.43d−08
2	2496	1328	3824	1561	1.08d−06	4.58d−07
3	9984	5152	15 136	12 325	2.06d−06	7.57d−07

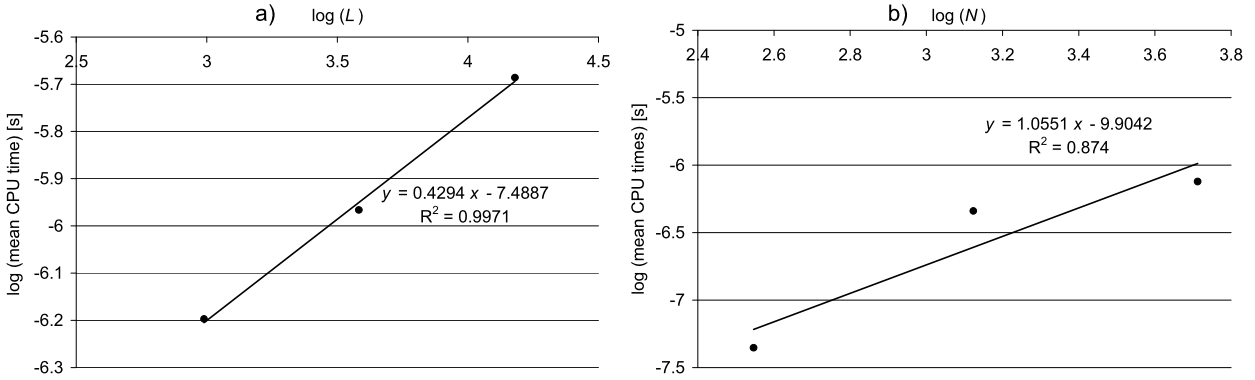

Fig. 23. (a) Mean CPU times for flux coefficients computation satisfying the M-property of the stiffness system matrix; (b) mean CPU times for system solution.

dividing the total time required by the procedure by the number of mesh edges L, while the mean CPU time for the solution of the final system has been obtained by dividing the total time required for the system solution by the number of nodes. A single processor Intel T 9400, 2.53 GHz has been used.

The computation of the stiffness matrix coefficients is the most demanding one and its mean CPU time is approximately twice the cost for the system solution.

The CPU time per single cell required for the solution of the diffusive system increases with the element number. In fact this step requires the solution of a large linear system of the order of the nodes number. On the opposite, the procedure of the stiffness matrix computation satisfying the M-property, requires a CPU time per single side that grows with the sides number much less than linearly.

The growth rate β, measured as the exponent of the relationship:

$$\overline{CP} = (N_u)^{\beta} \quad \Rightarrow \quad \log(\overline{CP}) = \log(N_u) \cdot \beta + c, \tag{53}$$

where N_u is the number of edges or nodes, $\overline{CP}$ is the mean CPU time and c is an arbitrary constant, has been investigated for the two components of the algorithm.

The mean CPU time required for the solution of the system increases with nodes number only a bit more than linearly, while the mean CPU time for flux coefficients computation increases with the sides number much less than linearly. In fact in Figs. 23(a)–23(b) the growth rate β is 1.055 for the diffusive step and only 0.43 for the procedure of the stiffness matrix computation satisfying the M-property.

9. Conclusions

A new methodology for the numerical solution of the most general anisotropic heterogeneous diffusion problem has been presented. The goal of the algorithm is to compute the unknown variable in a given set of nodes, located by the user inside the computational domain, along with an external and some internal boundary lines. A starting domain triangulation connecting the nodes is assumed, where internal/external boundaries are preserved by possibly adding new nodes with respect to the given ones. A series of edge swaps is applied by the algorithm to the starting domain triangulation (including the extra nodes added to preserve the boundaries), in order to get a final mesh satisfying the Generalized Delaunay, or Generalized Anisotropic Delaunay condition and that ensures a monotonic, oscillation-free solution.

The main advantages of the proposed procedure with respect to other similar ones are: (1) the algorithm acts directly on the physical mesh, without dealing with a different computational space; (2) the number and the location of the given input nodes of the starting domain triangulation are not changed; (3) the algorithm can be easily included in the solution of a more general convection – diffusion transport problem, where diffusion matrix coefficients are subject to change in time (see [5]).

Several numerical tests have been presented. The proposed methodology always satisfies the DMP, has shown a convergence order close to 2 and no mesh locking effects have been observed, also for strong anisotropy ratios. Numerical results

have also shown that both the relaxing anisotropy procedure across an internal boundary and the "smoothing" procedure for heterogeneous anisotropic medium without internal boundaries, introduce an error of the tensor coefficients with respect to the initially given values that decreases by refining the triangulation.

Appendix A. Addition of new nodes on internal/external boundary

In a given triangulation, call 1 and 2 two triangles with nodes i, j, k and i, j, m, sharing side $\mathbf{r}_{i,j}$. A sufficient, but not necessary, condition for the two triangles to satisfy the Delaunay condition with respect to side $\mathbf{r}_{i,j}$ is (see also [4]):

(1) the distances $d_{i,j}^{k,1}$ and $d_{i,j}^{m,2}$ of both nodes k and m with respect to the midpoint of $\mathbf{r}_{i,j}$ are greater than $d_{ij}/2$, where d_{ij} is the distance between nodes i and j.

The proof of the previous statement is the following. Observe that according to condition (1) nodes k and m will be external to the circle having the centre on the midpoint of $\mathbf{r}_{i,j}$ and radius $d_{ij}/2$. All the points on this circle form with nodes i and j right triangles, all the external nodes acute triangles, all the internal nodes obtuse triangles. Because $d_{i,j}^{k,1} > d_{ij}/2$ and $d_{i,j}^{m,2} > d_{ij}/2$, both triangles with nodes i, j, k and i, j, m are acute triangles and edge $\mathbf{r}_{i,j}$ satisfies the Delaunay property.

For a similar reason, the distance of the circumcentre of each triangle with nodes i, j, k from its side $\mathbf{r}_{i,j}$ on the external boundary is always positive if:

(2) the distance $d_{i,j}^{k,1}$ of node k with respect to the midpoint of $\mathbf{r}_{i,j}$ is greater than $d_{ij}/2$.

If condition (1) ((2)) holds, we say nodes k and m (node k) satisfy (satisfies) the Minimum Distance (MD) property.

To guarantee the convergence of the swapping iterative process to a domain triangulation satisfying conditions in Eqs. (18) and (21), without changing the edges overlapping internal boundaries, it will be sufficient to: (a) start with a triangulation with all the internal and external boundaries overlapping one or more element edges $\mathbf{r}_{i,j}$, and (b) guarantee that condition (1) (or (2)) is satisfied for all the potential triangles formed by $\mathbf{r}_{i,j}$ and a third node. This is true if all the nodes in the mesh different from i and j have a distance from the midpoint of $\mathbf{r}_{i,j}$ greater than $d_{ij}/2$ and this can be easily obtained by reducing the length of the edges $\mathbf{r}_{i,j}$.

Appendix B. Properties of the ellipse through the triangle nodes

The geometric interpretation in $\mathbb{R}^2$ of a diffusion tensor $\mathbf{D}$ is an ellipse [7], which axes are rotated with respect to a fixed Cartesian reference system $\mathbf{x} = (x_1, x_2)$ as the eigenvectors of $\mathbf{D}$, representing the principal directions and the axes ratio is equal to the ratio between the eigenvalues of $\mathbf{D}$.

Assume tensor $\mathbf{D}$ given by (see Eq. (22a)):

$$\mathbf{D} = d_0 \mathbf{D}', \tag{B.1}$$

with the symbols specified in Section 5.

Nodes of triangle T_1 are i, j and k (see Fig. 6, left). According to Eqs. (5) we have:

(1) the anisotropic circumcentre c_{T_1} of T_1 is the intersection of the two lines with directions $\mathbf{D}\mathbf{n}_{i,j}$ and $\mathbf{D}\mathbf{n}_{i,k}$, equal respectively to the directions $\mathbf{D}'\mathbf{n}_{i,j}$ and $\mathbf{D}'\mathbf{n}_{i,k}$,
(2) vector $\mathbf{D}'\mathbf{n}_{i,j(k)}$ is parallel to vector $(\mathbf{x}_c^{T_1} - \mathbf{x}_{i,j(k)})$, connecting c_{T_1} to midpoint of side $\mathbf{r}_{i,j}(\mathbf{r}_{i,k})$. Symbols have been specified in Section 3.

Call $\mathbf{h}_1$ and $\mathbf{h}_2$ the eigenvectors of $\mathbf{D}'$. These coincide with $\mathbf{D}$ eigenvectors except for the scaling factor d_0. We will prove that triangle nodes i, j and $k \in$ ellipse El_1 (see Fig. 6, left) and the following properties hold for El_1:

(a) centre of $El_1 c_{el}^{T_1}$ (with vector spatial co-ordinates $\mathbf{x}_{el}^{T_1}$) is equal to c_{T_1},
(b) axes of El_1 (the blue lines in Fig. 6, left) are respectively aligned along $\mathbf{h}_1$ and $\mathbf{h}_2$ directions ($\mathbf{h}_1$ and $\mathbf{h}_2$ are principal directions of the ellipse),
(c) axes ratio is equal to the ratio between the eigenvalues of $\mathbf{D}'$.

If property (b) holds for ellipse $El_1 \Rightarrow c_{el}^{T_1} = c_{T_1}$.

Proof. Let $\tilde{c}_1$ and $\tilde{c}_2$ be two cords of El_1, respectively orthogonal to $\mathbf{h}_1$ and $\mathbf{h}_2$. According to the definition of anisotropic circumcentre, the intersection point of the two diameters conjugate to $\tilde{c}_1$ and $\tilde{c}_2$ is c_{T_1}. These two diameters are aligned along the principal directions. Then $c_{el}^{T_1} = c_{T_1}$. If nodes i, j and $k \in El_1$, the two triangle sides $\mathbf{r}_{i,j}$ and $\mathbf{r}_{i,k}$ are cords of El_1 and vectors $(\mathbf{x}_c^{T_1} - \mathbf{x}_{i,j})$ and $(\mathbf{x}_c^{T_1} - \mathbf{x}_{i,k})$ are aligned along the directions of two ellipse diameters. □

If $c_{el}^{T_1} = c_{T_1} \Rightarrow$ nodes i, j and $k \in El_1$.

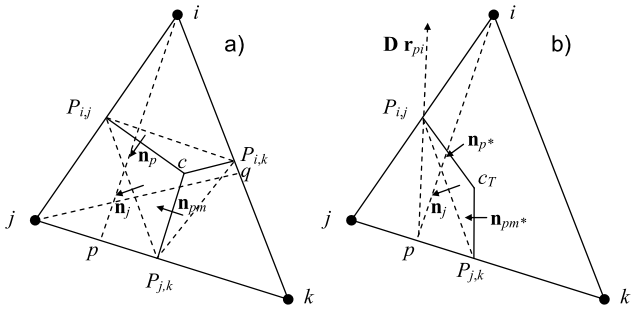

Fig. C1. (a) Isotropic case. Notations for the element shape functions and element stiffness Galerkin coefficients $F_{i,j}^{T,G}$. (b) Anisotropic case. Notations for the element shape functions and element stiffness Galerkin coefficients $F_{i,j}^{T,G}$.

Proof. According to Eqs. (5), centre $c_{el}^{T_1}$ is the intersection of the two vectors $(\mathbf{x}_{el}^{T_1} - \mathbf{x}_{i,j})$ and $(\mathbf{x}_{el}^{T_1} - \mathbf{x}_{i,k})$ and, on the base of properties of $c_{el}^{T_1}$, the two vectors are aligned along two diameters of El_1. Following the definition of ellipse diameter, the two triangle sides $\mathbf{r}_{i,j}$ and $\mathbf{r}_{i,k}$ are cords of El_1 and points i, j and $k \in El_1$. □

Also $\mathbf{r}_{j,k}$ is cord of El_1, then vector $(\mathbf{x}_c^{T_1} - \mathbf{x}_{j,k})$ is aligned along one ellipse diameter direction and, according to definition of anisotropic circumcentre, vector $\mathbf{D}\mathbf{n}_{j,k}$ is parallel to vector $(\mathbf{x}_c^{T_1} - \mathbf{x}_{j,k})$, that is equivalent to Eq. (5c).

Let's assume a vector connecting points p and q, where $q \in El_1$, with $q = i$, or $q = j$ or $q = k$. The following equation,

$$\left(\mathbf{x}_{el}^{T_1} - \frac{1}{2}(\mathbf{x}_p + \mathbf{x}_q)\right)_{\perp} \cdot \mathbf{D}'(\mathbf{x}_p - \mathbf{x}_q)_{\perp} = 0, \tag{B.2}$$

where symbol $(\cdot)_{\perp}$ is specified in Section 5, states the orthogonality of vectors $\mathbf{D}'\mathbf{n}_{pq}$ and $(\mathbf{x}_{el}^{T_1} - \frac{1}{2}(\mathbf{x}_p + \mathbf{x}_q))_{\perp}$ ($\mathbf{n}_{q,p}$ is the unit normal component to $(\mathbf{x}_p - \mathbf{x}_q)$), so that $(\mathbf{x}_p - \mathbf{x}_q)$ is a cord of El_1, then also point $p \in El_1$. Eq. (B.2) represents the locus of points p equal to the ellipse El_1 with properties (a) to (c), above defined. Observe that Eq. (B.2) is equivalent to Eq. (26).

The isotropic medium is a particular case of the anisotropic one, with $\mathbf{D} = d_0\mathbf{I}$ ($\mathbf{I}$ is the identity matrix). In this case we have only one eigenvalue, ellipse El_1 becomes a circle C_1 and every direction in the space $\mathbf{x} = (x_1, x_2)$ is principal direction. Triangle circumcentre $c_{T_1} = c_{cir}^{T_1}$, centre of circle C_1 (circumcircle of T_1). Triangle nodes i, j and $k \in C_1$. According to Eq. (B.2), a generic point $p \neq i$, j, k, with p:

$$\left(\mathbf{x}_{cir}^{T_1} - \frac{1}{2}(\mathbf{x}_p + \mathbf{x}_q)\right)_{\perp} \cdot (\mathbf{x}_p - \mathbf{x}_q)_{\perp} = 0, \quad q = i, \text{ or } j \text{ or } k, \tag{B.3}$$

belongs to C_1. Condition (B.3) is equivalent to Eq. (32).

Appendix C. Preserving Positive Transmissibility (PT) condition in anisotropic problems. Proposed procedure vs conforming linear ($P1$) Galerkin scheme

The condition for a stiffness matrix system to have non-positive stiffness coefficients is called Positive Transmissibility (PT) and it is a prerequisite for the matrix M-property [35].

In this section we will focus on the discretization of the general equation for diffusive problem:

$$-\nabla \cdot \mathbf{D}\nabla u = 0. \tag{C.1}$$

Assume an isotropic homogeneous medium where $\mathbf{D} = d_0\mathbf{I}$ (d_0 a scalar value and $\mathbf{I}$ the unit matrix). Let i, j and k be the nodes of triangle T and c its circumcentre (Fig. C1(a)).

It is well known (see for example [35]) that the linear ($P1$) conforming Galerkin stiffness coefficient for triangle T is:

$$F_{i,j}^{T,G} = d_0 \int_T \nabla N_i \cdot \nabla N_j \, dT, \tag{C.2}$$

where N_i is the basis function and apices T and G mark respectively the quantities related to triangle T and Galerkin method. The global stiffness coefficient $F_{i,j}^{G}$ is the sum of $F_{i,j}^{T,G}$ over all triangles and the Galerkin equation for node i can be written as [35]:

$$\sum_{j \neq i} F_{i,j}^{G}(u_j - u_i) = 0. \tag{C.3}$$

In order to preserve the M-property for the stiffness matrix, the $F_{i,j}^{G}$ have to be non-positive.

By definition, basis function N_j is a linear function equal to 1 on node j and 0 on nodes i and k, so that gradient of N_j is a constant vector orthogonal to side $\mathbf{r}_{i,k}$,

$$\nabla N_j = \frac{\mathbf{n}_j}{|\mathbf{r}_{q,j}|} = \frac{\mathbf{r}_{q,j}}{\mathbf{r}_{q,j}\cdot\mathbf{r}_{q,j}}, \tag{C.4}$$

where $\mathbf{n}_j$ is the unit vector normal to edge $\mathbf{r}_{i,k}$, pointing toward node j (see Fig. C1(a)). Substituting Eq. (C.4) in Eq. (C.2), one gets [35]:

$$\int_T \nabla N_i \cdot \nabla N_j \, dT = \nabla N_i \cdot \frac{\mathbf{n}_j}{|\mathbf{r}_{q,j}|}|T| = \nabla N_i \cdot \mathbf{n}_j |\mathbf{r}_{P_{j,k},P_{i,j}}|, \tag{C.5}$$

and the r.h.s. of Eq. (C.5) can be interpreted as the flux of ∇N_i across segment $\overline{P_{j,k}P_{i,j}}$. Starting from the Gauss theorem, this flux can be computed as the line integral over any path Γ_p connecting midpoints $P_{j,k}$ and $P_{i,j}$. It is well known [35] that the (i,j)th element Galerkin stiffness coefficient is given by:

$$F_{i,j}^{T,G} = d_0 \int_{\Gamma_p} \nabla N_i \cdot d\mathbf{n}, \tag{C.6}$$

where Γ_p is the boundary of the nodal control volume of node j. Choosing Γ_p as the Thiessen polygon boundary $\overline{P_{j,k}cP_{i,j}}$ (Fig. C1(a)), from Eq. (C.6) one gets:

$$F_{i,j}^{T,G} = d_0 \frac{\mathbf{r}_{p,i}}{\mathbf{r}_{p,i}\cdot\mathbf{r}_{p,i}} \left(\int_{P_{i,j}c} d\mathbf{n}_p + \int_{cP_{j,k}} d\mathbf{n}_{pm} \right), \tag{C.7}$$

where the unit normals $\mathbf{n}_p$ and $\mathbf{n}_{pm}$ are shown in Fig. C1(a) and the second integral in Eq. (C.7) is null since $\mathbf{r}_{p,i} \perp \mathbf{n}_{pm}$.

From simple geometric considerations (see also [35]), it can be shown that:

$$F_{i,j}^{T,G} = \mp d_0 \frac{|\mathbf{r}_{P_{i,j},c}|}{|\mathbf{r}_{i,j}|}, \tag{C.8}$$

where sign $-$ holds if the angle in k is $< \pi/2$, otherwise sign $+$ holds. Similarly, the element coefficients $F_{j,k}^{T,G}$ and $F_{i,k}^{T,G}$ are respectively equal to:

$$F_{j,k}^{T,G} = \mp d_0 \frac{|\mathbf{r}_{P_{j,k},c}|}{|\mathbf{r}_{j,k}|} \quad \text{and} \quad F_{i,k}^{T,G} = \mp d_0 \frac{|\mathbf{r}_{P_{j,k},c}|}{|\mathbf{r}_{i,k}|}. \tag{C.9}$$

According to Eq. (C.3), PT condition is much more assured, the more non-positive are the $F_{i,j}^{G}$ coefficients.

Assume now an anisotropic medium. In the following we show that, for given element geometry and diffusive tensor, the flux coefficient sign (i.e. the elemental stiffness coefficient sign) is more likely to be negative according to our algorithm than to the Galerkin one.

Let c_T be the anisotropic circumcentre of T, computed as in Eqs. (5). According to analogous consideration to the ones in the isotropic case and applying the Gauss theorem, the (i,j)th element Galerkin stiffness coefficient is:

$$F_{i,j}^{T,G} = \int_{\Gamma_p} \mathbf{D}\nabla N_i \cdot d\mathbf{n}, \tag{C.10}$$

which represents the flux of $\mathbf{D}\nabla N_i$ across any path Γ_p connecting midpoints $P_{i,j}$ and $P_{j,k}$. Assuming $\Gamma_p = \overline{P_{j,k}c_T P_{i,j}}$ (the part of the boundary of the control volume of the proposed procedure) (see Fig. C1(b)), according to Eq. (C.7), Eq. (C.10) becomes:

$$F_{i,j}^{T,G} = \frac{\mathbf{D}\mathbf{r}_{p,i}}{\mathbf{r}_{p,i}\cdot\mathbf{r}_{p,i}} \left(\int_{P_{i,j}c_T} d\mathbf{n}_{p*} + \int_{c_T P_{j,k}} d\mathbf{n}_{pm*} \right), \tag{C.11}$$

where $\mathbf{n}_{p*}$ and $\mathbf{n}_{pm*}$ are the unit vectors orthogonal to $\overline{P_{i,j}c_T}$ and $\overline{c_T P_{j,k}}$ respectively (in Fig. C1(b)). The second integral in Eq. (C.11) is null, since vector $\mathbf{D}\mathbf{r}_{p,i} \perp \mathbf{n}_{pm*}$ (see Fig. C1(b), where direction of vector $\mathbf{D}\mathbf{r}_{p,i}$ is in dashed line), while the sign of $\mathbf{D}\mathbf{r}_{p,i}.\mathbf{n}_{p*}$ is negative if c_T lies on the opposite side to node j with respect to $\overline{P_{j,k}P_{i,j}}$ (see the two cases depicted in Fig. C2(a)). If c_T lies on the same side of j with respect to $\overline{P_{j,k}P_{i,j}}$ (see the two cases in Fig. C2(b)), $\mathbf{D}\mathbf{r}_{p,i}.\mathbf{n}_{p*}$ becomes positive, also if T is acute and $c_T \in T$. $\mathbf{D}\mathbf{r}_{p,i}.\mathbf{n}_{p*}$ is equal to zero if $c_T \in \overline{P_{j,k}P_{i,j}}$.

Observe in Fig. C3(a) that c_T is on the opposite side to j with respect to $\overline{P_{j,k}P_{i,j}}$ and on the opposite side to k with respect to $\overline{P_{j,k}P_{i,k}}$, but on the same side of i with respect to $\overline{P_{i,j}P_{i,k}}$. From simple geometrical considerations one gets that

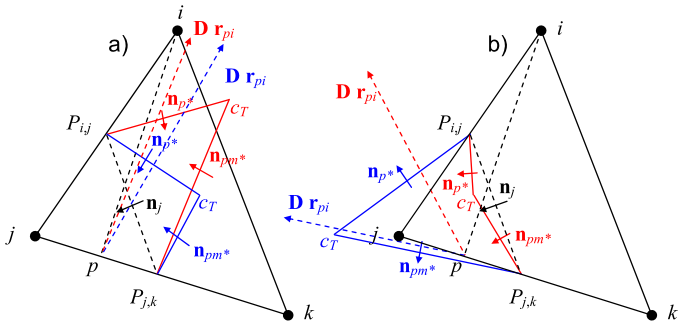

Fig. C2. (a) Anisotropic case $F_{i,j}^{T,G} < 0$; (b) anisotropic case $F_{i,j}^{T,G} > 0$.

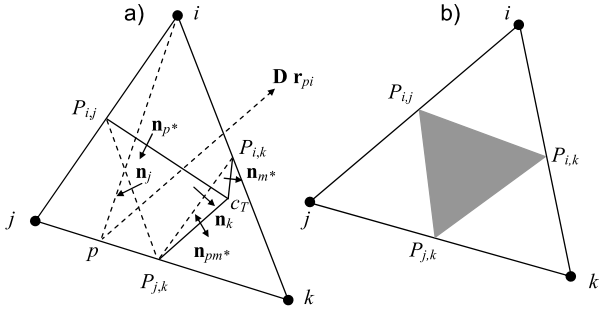

Fig. C3. (a) Anisotropic case $Fd_{i,j}^{T,G}$, $Fd_{j,k}^{T,G} < 0$ and $Fd_{i,k}^{T,G} > 0$; (b) triangle $\widehat{T}$ (shaded area).

coefficients $F_{i,j}^{T,G}, F_{j,k}^{T,G} < 0$ and $F_{i,k}^{T,G} > 0$. Defining $\widehat{T}$ the triangle obtained connecting the three midpoints edge of T (see Fig. C3(b)), the following result is obtained:

$$F_{i,j}^{T,G} \leqslant 0 \quad \text{and} \quad F_{j,k}^{T,G} \leqslant 0 \quad \text{and} \quad F_{i,k}^{T,G} \leqslant 0 \quad \Leftrightarrow \quad c_T \in \widehat{T}, \tag{C.12}$$

where symbol $\Leftrightarrow$ means "if and only if". It is easy to recognize that flux coefficient $F_{i,j}^{T}$ of the proposed methodology is positive if c^T is external to triangle T and lies on the opposite side with respect to $\mathbf{r}_{i,j}$. Applying the same considerations to coefficients $F_{j,k}^{T}$ and $F_{i,k}^{T}$, we can conclude that:

$$F_{i,j}^{T} \leqslant 0 \quad \text{and} \quad F_{j,k}^{T} \leqslant 0 \quad \text{and} \quad F_{i,k}^{T} \leqslant 0 \quad \Leftrightarrow \quad c_T \in T. \tag{C.13}$$

From Eqs. (C.12)–(C.13), one gets that the proposed procedure, compared with the standard $P1$ Galerkin formulation, maintains PT condition for more general anisotropic conditions before performing the edge swap.

References

[1] I. Aavatsmark, T. Barkve, Ø. Bøe, T. Mannseth, Discretization on non-orthogonal, quadrilateral grids for inhomogeneous, anisotropic media, J. Comput. Phys. 127 (1996) 2–14.
[2] I. Aavatsmark, T. Barkve, Ø. Bøe, T. Mannseth, Discretization on unstructured grids for inhomogeneous, anisotropic media. I. Derivation of the methods, SIAM J. Sci. Comput. 19 (5) (1998) 1700–1716.
[3] I. Aavatsmark, T. Barkve, Ø. Bøe, T. Mannseth, Discretization on unstructured grids for inhomogeneous, anisotropic media. II. Discussion and numerical results, SIAM J. Sci. Comput. 19 (5) (1998) 1717–1736.
[4] C. Aricò, M. Sinagra, L. Begnudelli, T. Tucciarelli, MAST-2D diffusive model for flood prediction on domains with triangular Delaunay unstructured meshes, Adv. Water Resour. 34 (2011) 1427–1449.
[5] C. Aricò, M. Sinagra, T. Tucciarelli, Monotonic solution of flow and transport problems in heterogeneous media using Delaunay unstructured triangular meshes, Adv. Water Resour. 52 (2013) 132–150.
[6] I. Babuška, M. Suri, On locking and robustness in the finite element method, SIAM J. Numer. Anal. 29 (5) (1992) 1261–1293.
[7] J. Bear, Dynamics of Fluids in Porous Media, Dover, New York, 1972.
[8] A. Bellin, P. Salandin, A. Rinaldo, Simulation of dispersion in heterogeneous porous formations: Statistics, first order theories, convergence of computations, Water Resour. Res. 28 (9) (1992) 2211–2227.
[9] J. Brandts, S. Korotov, M. Křížek, The discrete maximum principle for linear simplicial finite element approximations of a reaction–diffusion problem, Linear Algebra Appl. 429 (2008) 2344–2357.
[10] P.G. Ciarlet, Discrete maximum principle for finite difference operators, Aequationes Math. 4 (1970) 338–352.
[11] P.G. Ciarlet, P.-A. Raviart, Maximum principle and uniform convergence for the finite element method, Comput. Meth. Appl. Mech. Engrg. 2 (1973) 17–31.
[12] A. Draganescu, T.F. Dupont, L.R. Scott, Failure of the discrete maximum principle for an elliptic finite element problem, Math. Comput. 74 (2004) 1–23.

[13] M.G. Edwards, H. Zheng, A quasi-positive family of continuous Darcy-flux finite-volume schemes with full pressure support, J. Comput. Phys. 227 (22) (2008) 9333–9364.
[14] M.G. Edwards, H. Zheng, Double-families of quasi-positive Darcy-flux approximations with highly anisotropic tensors on structured and unstructured grids, J. Comput. Phys. 229 (3) (2010) 594–625.
[15] P.A. Forsyth, A control volume finite element approach to NAPL groundwater contamination, SIAM J. Sci. Statist. Comput. 12 (1991) 1029–1057.
[16] Z. Gao, J. Wu, A linearity-preserving cell-centered scheme for the heterogeneous and anisotropic diffusion equations on general meshes, Int. J. Numer. Meth. Fluids 67 (12) (2011) 2157–2183.
[17] G. Gambolati, Diagonally dominant matrices for the finite element method in hydrology, Int. J. Numer. Meth. Engrg. 6 (1973) 587–591.
[18] V. Havu, An analysis of asymptotic consistency error in a parameter dependent model problem, Calcolo (ISSN 0008-0624) 40 (2) (2003) 121–130.
[19] V. Havu, J. Pitkäranta, An analysis of finite element locking in a parameter dependent model problem, Numer. Math. (ISSN 0029-599X) 89 (4) (2001) 691–714.
[20] W. Huang, Metric tensors for anisotropic mesh generation, J. Comput. Phys. 204 (2) (2005) 633–665.
[21] B. Joe, Delaunay triangular meshes in convex polygons, SIAM J. Sci. Statist. Comput. 7 (1986) 514–539.
[22] R.A. Klausen, R. Winther, Robust convergence of multipoint flux approximation on rough grids, Numer. Math. 104 (2006) 317–337.
[23] D. Kuzmin, M.J. Shashkov, D. Svyatskiy, A constrained finite element method satisfying the discrete maximum principle for anisotropic diffusion problems, J. Comput. Phys. 228 (2009) 3448–3463.
[24] C. Le Potier, Schéma volumes finis monotone pour des opérateurs de diffusion fortement anisotropes sur des maillages de triangles non-structurés, C. R. Math. Acad. Sci. Paris 341 (2005) 787–792.
[25] C. Le Potier, A nonlinear finite volume scheme satisfying maximum and minimum principles for diffusion operators, Int. J. Finite Vol. 6 (2009) 20.
[26] F.W. Letniowski, Three-dimensional Delaunay triangulations for finite element approximations to a second-order diffusion operator, SIAM J. Sci. Statist. Comput. 13 (1992) 765–770.
[27] X. Li, W. Huang, An anisotropic mesh adaptation method for the finite element solution of heterogeneous anisotropic diffusion problems, J. Comput. Phys. 229 (21) (2010) 8072–8094.
[28] X.P. Li, D. Svyatskiy, M. Shashkov, Mesh adaptation and discrete maximum principle for 2D anisotropic diffusion problems, LANL Technical report, LAUR 10-01227, 2007.
[29] K. Lipnikov, M. Shashkov, D. Svyatskiy, Yu. Vassilevski, Monotone finite volume schemes for diffusion equations on unstructured triangular and shape-regular polygonal meshes, J. Comput. Phys. 227 (1) (2007) 492–512.
[30] R. Liska, M. Shashkov, Enforcing the discrete maximum principle for linear finite element solutions of second-order elliptic problems, Commun. Comput. Phys. 3 (2008) 852–877.
[31] G. Manzini, M. Putti, Mesh locking effects in the finite volume solution of 2-D anisotropic diffusion equations, J. Comput. Phys. 220 (2) (2007) 751–771.
[32] A. Mazzia, G. Manzini, M. Putti, Bad behaviour of Godunov Mixed Methods for strongly anisotropic advection-dispersion equations, J. Comput. Phys. 230 (23) (2011) 8410–8426.
[33] M.J. Mlacnik, L.J. Durlofsky, Unstructured grid optimization for improved monotonicity of discrete solutions of elliptic equations with highly anisotropic coefficients, J. Comput. Phys. 216 (1) (2006) 337–361.
[34] T.N. Narasimhan, P.A. Witherspoon, An integrated finite difference method for analyzing fluid flow in porous media, Water Resour. Res. 12 (1976) 57–64.
[35] M. Putti, C. Cordes, Finite element approximation of the diffusion operator on tetrahedral, SIAM J. Sci. Comput. 19 (4) (1998) 1154–1168.
[36] P. Sharma, G.W. Hammett, Preserving monotonicity in anisotropic diffusion, J. Comput. Phys. 227 (1) (2007) 123–142.
[37] G. Stoyan, On a maximum principle for matrices, and on conservation of monotonicity. With applications to discretization methods, Z. Angew. Math. Mech. 62 (1982) 375–381.
[38] G. Stoyan, On maximum principles for monotone matrices, Linear Algebra Appl. 78 (1986) 147–161.
[39] G. Strang, G.J. Fix, An Analysis of the Finite Element Method, Prentice Hall, Englewood Cliffs, NJ, 1973.
[40] R.S. Varga, On a discrete maximum principle, SIAM J. Numer. Anal. 3 (1966) 355–359.