African Diaspora Journal of Mathematics
Volume 19, Number 2, pp. 1–28 (2016)

ISSN 1539-854X
www.math-res-pub.org/adjm

C
DDFV

N

A. K J *
The University of Buea
Higher Technical Teachers' Training College,
Department of Computer Science,
P.O. Box: 249 Kumba - Cameroon

A. N
The University of Douala
Faculty of Industrial Engineering
P.O. Box: 2701 Douala - Cameroon

J. N
The University of Yaounde I
National Advanced School of Engineering
P.O. Box: 8390 Yaounde - Cameroon

Abstract

Our purpose in this paper is to present a theoretical analysis of the Discrete Duality Finite Volume method (DDFV method) for 2D-flow problems in anisotropic heterogeneous porous media with full Neumann boundary conditions. We start with the derivation of the discrete problem, and then we give a result of existence and uniqueness of a solution for that problem. Their theoretical properties, namely stability and error estimates in discrete energy norms and L^2-norm are investigated. Numerical tests are provided.

AMS Subject Classification: (2013): 35J65, 65N12, 65N15, 74S10.

Keywords: Flow problems, Nonhomogeneous anisotropic media, Discrete duality finite volumes, Stability and error estimates, Numerical test.

*E-mail address: jeutsa2001@yahoo.fr

1 Introduction and the model problem

From the outset, the classical finite volume methods (see for instance [9]) were designed for ensuring the local mass conservation as well as the robustness for complex applications (multi-phase flow in geologically complex reservoirs for instance). But the imposed geometric constraints to mesh elements was an important handicap for those methods. Furthermore, the finite volume computation of anisotropic flows was a real challenge. To overcome these difficulties, several investigators have proposed variants finite volume methods. In these methods the key idea consists in approximating the fluxes using multi-point schemes known in the literature as Multi-Point Flux Approximation methods (see for instance [1], [2], [5], [8], [13], [15], [17] and [25]). The Discrete Duality Finite Volume (DDFV) methods (see for instance [6], [21], [22], [3] and [4]) combine the advantages of the above mentioned methods, i.e. local and global mass conservation principle, accurate for rough grids and coefficients. The DDFV methods can be considered as finite volume methods of new generation.

This work is a contribution to the theoretical analysis of the DDFV formulation presented in [14], [20], [18], [19], [10] and [11]. The matrix kernel analysis of our linear system is similar to [16] and [24]. Our analysis is focused on the case of anisotropic flow in heterogeneous media with full Neumann boundary conditions covered with a quadrilateral grid.

Let us consider the 2D diffusion problem consisting in finding a function φ in Ω that satisfies the following partial differential equation associated with homogeneous Neumann boundary conditions:

$$-div(D\ grad\ \varphi) = f \quad \text{in} \quad \Omega \tag{1.1}$$

$$-D\ grad\ \varphi \cdot \eta = g \quad \text{in} \quad \Gamma \tag{1.2}$$

where Ω is a given open square domain, Γ its boundary, f and g are given functions and η the unit normal vector} to Γ outward to Ω. $D = D(x)$, with $x = (x_1, x_2)^t \in \Omega$, is a full piecewise constant matrix describing the spatial variation of the diffusion coefficients. We assume that the matrix structure depends solely on the geological structure of Ω. Let us set

$$\mathcal{L}_\Omega = \{L;\ L\ is\ a\ lithologic\ component\ of\ \Omega\} \tag{1.3}$$

and let us consider

$$\sigma = \max\left\{\left|\frac{D_{12}^L}{D_{11}^L}\right| + \left|\frac{D_{12}^L}{D_{22}^L}\right|;\ L \in \mathcal{L}_\Omega\right\} \tag{1.4}$$

where D^L denotes the full diffusion matrix with constant coefficients corresponding to the lithologic component L. So σ is a strictly positive number that is mesh independent. As we will see later this number plays a key role in stability analysis.

Let us formulate some basic assumptions:

•**Symmetry:**

$$\forall 1 \le i, j \le 2 \qquad D_{ij}(x) = D_{ji}(x) \qquad a.e. \quad \text{in} \quad \Omega \tag{1.5}$$

•**Uniform ellipticity and boundedness:**

$$\begin{aligned} &\exists \gamma_{\min},\ \gamma_{\max} \in \mathbb{R}^{*+} \quad \text{such that} \quad \forall \xi \in \mathbb{R}^2,\quad \xi \neq 0 \\ &\gamma_{\min} |\xi|^2 \leq \xi^T D(x)\xi \leq \gamma_{\max} |\xi|^2 \qquad a.e. \ \text{ in } \quad \Omega \end{aligned} \tag{1.6}$$

where $|\cdot|$ denotes the euclidian norm in $\mathbb{R}^2$, and where $D_{ij}(\cdot)$ are components of D. We also suppose that $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$ and that the following compatibility condition is satisfied:

$$\int_\Omega f(x)dx - \int_\Gamma g(x)d\tau(x) = 0 \tag{1.7}$$

Under the previous assumptions the model problem (1.1)-(1.2) possesses a unique variational solution i.e. there exists a function u defined almost everywhere (a.e.) in Ω such that:

$$u \in V = \{v \in H^1(\Omega);\ \int_\Omega v(x)dx = 0\} \tag{1.8}$$

and

$$\int_\Omega D u(x)v(x)dx = \int_\Omega f(x)v(x)dx - \int_{\partial\Omega} g(x)\gamma_0(v)(x)d\tau \quad \forall v \in H^1(\Omega). \tag{1.9}$$

where γ_0 is the "trace" operator from $H^1(\Omega)$ into $L^2(\partial\Omega)$ (or onto $H^{\frac{1}{2}}(\partial\Omega)$).

Remark 1.1. The existence and uniqueness of solution of the problem (1.1), (1.2) and (1.5) is proved theoretically in quotient space $H^1(\Omega)/\mathbb{R}$ (see for instance [12]). So if u is this solution then, for any $c \in \mathbb{R}$, $\overline{u} = u + c$ is also a solution. So the constraint $\int_\Omega v dx = 0$ incorporated in the definition of V guarantees uniqueness of the solution to the problem (1.1), (1.2) and (1.5) in V.

This paper is organized as follows. The second section deals with a finite volume formulation of the model problem. Within this section we bring an affirmative answer to the well posedness issue concerning the discrete problem. In the third section we investigate the theoretical properties (stability and error estimates in convenient discrete norms) for the solution of the discrete problem. The fourth section is devoted to the numerical tests.

2 A DDFV formulation of the model problem

We are going to focus on the case of diffusion problems governed by piece wise constant full diffusion tensors. From the practical point of view this assumption is very realistic (see [7] and [23]). Indeed a subsurface area is made up of a collection of various geologic formations that may be characterized at intermediate scales by averaged full permeability tensors.

2.1 Formulation of the discrete problem

In what follows, we present the matrix form of a DDFV formulation for (1.1)-(1.2). Let us emphasize that this method applies for any convex polygonal domain covered with an unstructured primary grid. However we develop here the convergence analysis of that method on $\Omega =]0, 1[^2$ which is associated at square primary grid denoted $\mathcal{P}$ whose size is $h = \frac{1}{N}$, where N is a given strictly positive integer. On the other hand, we denote $K_{i,j}$ the primary grid-block defined by: $K_{i,j} = \left[x_1^{i-\frac{1}{2}}, x_1^{i+\frac{1}{2}}\right] \times \left[x_2^{j-\frac{1}{2}}, x_2^{j+\frac{1}{2}}\right]$ where $x_1^{i+\frac{1}{2}} = x_1^{i-\frac{1}{2}} + h$, $x_2^{j+\frac{1}{2}} = x_2^{j-\frac{1}{2}} + h,\ for\ i,j = 1,\ldots,N$ with $x_1^{\frac{1}{2}} = x_2^{\frac{1}{2}} = 0$.

Important assumption: The discontinuities of the diffusion coefficient D lie on gridblock interfaces and naturally divide Ω into a finite number of convex sub-domains $\{\Omega_s\}_{s\in S}$

From the boundary-value problem theory (see for instance [12]), the balance equation (1.1)-(1.2) possesses a unique variational solution in V under the assumption (1.5)-(1.7) and the condition $f \in L^2(\Omega)$.

We now make the additional assumption that the restriction over Ω_s of the exact solution to the system (1.1)-(1.2) denoted by $\varphi_{|\Omega_s}$, satisfies to

$$\varphi_{|\Omega_s} \in C^2(\overline{\Omega}_s) \qquad \forall\ s \in S$$

We should look for a finite volume formulation of the problem (1.1)-(1.2) in terms of a linear system which is derived from the elimination of auxiliary unknowns, namely interface pressures, in flux balance equations over grid-blocks. This linear system involves $\left\{u_{i,j}\right\}_{1\le i,j\le N}$ and $\left\{u_{i+\frac{1}{2},j+\frac{1}{2}}\right\}_{0\le i,j\le N}$ as discrete unknowns expected to be reasonable approximations of $\left\{\varphi_{i,j}\right\}_{1\le i,j\le N}$ (cell center pressures) and $\left\{\varphi_{i+\frac{1}{2},j+\frac{1}{2}}\right\}_{0\le i,j\le N}$ (cell corner pressures) respectively, where $\varphi_{i,j} = \varphi\left(x_1^i, x_2^j\right)$ and $\varphi_{i+\frac{1}{2},j+\frac{1}{2}} = \varphi\left(x_1^{i+\frac{1}{2}}, x_2^{j+\frac{1}{2}}\right)$, with:

$$x_1^i = \frac{x_1^{i-\frac{1}{2}} + x_1^{i+\frac{1}{2}}}{2}, \qquad x_2^j = \frac{x_2^{j-\frac{1}{2}} + x_2^{j+\frac{1}{2}}}{2} \qquad 1 \le i,j \le N \tag{2.1}$$

We also adopt the following conventions:

$$x_1^0 = x_1^{\frac{1}{2}}, \quad x_1^{N+1} = x_1^{N+\frac{1}{2}}, \quad x_2^0 = x_2^{\frac{1}{2}}, \quad x_2^{N+1} = x_2^{N+\frac{1}{2}} \tag{2.2}$$

We now give a summary description of the procedure leading to the linear discrete system. We integrate the balance equation (1.1) in the grid-block $K_{i,j}$, commonly called a control volume and centered at the point $\left(x_1^i, x_2^j\right)$. Applying Ostrogradski's theorem leads to integrate the flux on the boundary of $K_{i,j}$. This integration is performed using an adequate quadrature formula over each half-edge of $K_{i,j}$, and this leads to an expression involving the pressure value at edge mid-points. This pressure value is dropped away thanks to the flux continuity which is imposed over the grid-block interfaces.

Let us illustrate now our procedure for computing the fluxes across the gridblock boundaries. For this purpose, we consider the internal edge $[Now, Noe]$ associated with the grid-blocks $K_{i,j}$ and $K_{i,j+1}$ centered respectively at C and C' (see Figure 1 below).

In what follows, the restriction of φ over the closure of each gridblock, denoted again φ, is supposed to be C^2. From the definition of the gridblocks $K_{i,j}$, it is clear that $(x_1^i, x_2^j)^t$, $(x_1^{i+\frac{1}{2}}, x_2^{j+\frac{1}{2}})^t$, $(x_1^{i-\frac{1}{2}}, x_2^{j+\frac{1}{2}})^t$ and $(x_1^i, x_2^{j+\frac{1}{2}})^t$ are respectively the coordinates of the points C,

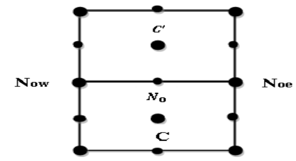

Figure 1. The edge $[Now, Noe]$ associated with the grid-blocks $K_{i,j}$ and $K_{i,j+1}$.

Noe, Now and No (see Figure 1 above). On the other hand, D^{ij} and D^{ij+1} denote respectively the diffusion tensor of the gridblocks $K_{i,j}$ and $K_{i,j+1}$.

The flux expression over the edge $[Noe, Now]$ satisfies to the relation: (for more details see [19])

$$\begin{aligned}\int_{[Noe,Now]}\left[-D^{ij} grad\varphi\cdot n\right] ds = 2D^{ij}_{22}\left[\varphi_{i,j}-\varphi_{i,j+\frac{1}{2}}\right] - D^{ij}_{21}\left[\varphi_{i+\frac{1}{2},j+\frac{1}{2}}-\varphi_{i,j+\frac{1}{2}}\right] \\ - D^{ij}_{21}\left[\varphi_{i,j+\frac{1}{2}}-\varphi_{i-\frac{1}{2},j+\frac{1}{2}}\right] + hR^{e}_{i,j}\end{aligned} \tag{2.3}$$

with $\left|R^{e}_{i,j}\right| \le Ch$, where C depends exclusively on Ω, $\frac{\partial^2\varphi}{\partial x_2^2}$ and the lithologic structure of the porous medium.

Furthermore the flux continuity over the interface between the gridblocks $K_{i,j}$ and $K_{i,j+1}$ leads to the eliminating the edge mid-point pressure $\varphi_{i,j+\frac{1}{2}}$ in (2.3). Hence we have the following approximation of the flux over the edge $[Noe, Now]$:

$$\begin{aligned}\int_{[Noe,Now]}\left[-D^{ij}\ grad\ \varphi\ \cdot\ n\right]\ ds\ \approx\ \frac{2D^{ij}_{22}D^{ij+1}_{22}}{D^{ij}_{22}+D^{ij+1}_{22}}\left[\varphi_{i,j}-\varphi_{i,j+1}\right] \\ + \frac{D^{ij}_{22}D^{ij+1}_{21}+D^{ij+1}_{22}D^{ij}_{21}}{D^{ij}_{22}+D^{ij+1}_{22}}\left[\varphi_{i-\frac{1}{2},j+\frac{1}{2}}-\varphi_{i+\frac{1}{2},j+\frac{1}{2}}\right]\end{aligned}$$

Note that in the case of a boundary-edge satisfying Neumann conditions, the flux over this edge is equal to the imposed flux leads to an easy elimination of the corresponding edge mid-point pressure.

Let us introduce the fictitious gridblocks $K_{i,0}$; $K_{0,j}$; $K_{i,N+1}$; $K_{N+1,j}$ $i,j = 0,...,N+1$ associated with the null permeability. The use of such fictitious gridblocks leads to a synthetic formulation of discrete balance equation valid for any primary gridblock (without discriminating between internal and boundary primary gridblocks): see Figure 2 below.

It is then clear that this procedure applies to the boundary of any gridblock $K_{i,j}$, with

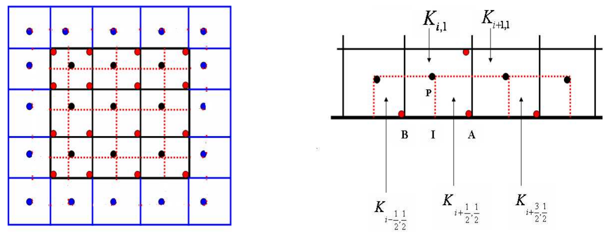

Figure 2. Left: Primary grid in black lines, dual grid in red dashed lines and fictitious gridblocks in blue lines.
Right: Degenerated dual gridblocks in red dashed lines.

$1 \leq i,\ j \leq N$, and leads to the following system of relations:

$$
\begin{aligned}
& D_{22,22}^{ij,ij+1}\left[\varphi_{i,j}-\varphi_{i,j+1}\right]+D_{22,21}^{ij,ij+1}\left[\varphi_{i-\frac{1}{2},j+\frac{1}{2}}-\varphi_{i+\frac{1}{2},j+\frac{1}{2}}\right] \\
&+ D_{22,22}^{ij,ij-1}\left[\varphi_{i,j}-\varphi_{i,j-1}\right]+D_{22,21}^{ij,ij-1}\left[\varphi_{i+\frac{1}{2},j-\frac{1}{2}}-\varphi_{i-\frac{1}{2},j-\frac{1}{2}}\right] \\
&+D_{11,11}^{ij,i+1j}\left[\varphi_{i,j}-\varphi_{i+1,j}\right]+\ D_{11,12}^{ij,i+1j}\left[\varphi_{i+\frac{1}{2},j-\frac{1}{2}}-\varphi_{i+\frac{1}{2},j+\frac{1}{2}}\right] \\
&+D_{11,11}^{ij,i-1j}\left[\varphi_{i,j}-\varphi_{i-1,j}\right]+D_{11,12}^{ij,i-1j}\left[\varphi_{i-\frac{1}{2},j+\frac{1}{2}}-\varphi_{i-\frac{1}{2},j-\frac{1}{2}}\right] \\
&\approx \int_{K_{i,j}} f(x)dx \ - \ \int_{\Gamma\cap\partial K_{i,j}} g(x)d\tau(x) \quad \forall\ \ 1\leq i,j\leq N
\end{aligned}
\tag{2.4}
$$

where

$$
D_{22,22}^{ij,ij\pm1} = \frac{2D_{22}^{ij}D_{22}^{ij\pm1}}{D_{22}^{ij}+D_{22}^{ij\pm1}}, \quad D_{22,21}^{ij,ij\pm1} = \frac{D_{22}^{ij}D_{21}^{ij\pm1}+D_{22}^{ij\pm1}D_{21}^{ij}}{D_{22}^{ij}+D_{22}^{ij\pm1}} \tag{2.5}
$$

$$
D_{11,11}^{ij,i\pm1j} = \frac{2D_{11}^{ij}D_{11}^{i\pm1j}}{D_{11}^{ij}+D_{11}^{i\pm1j}}, \quad D_{11,12}^{ij,i\pm1j} = \frac{D_{11}^{ij}D_{21}^{i\pm1j}+D_{11}^{i\pm1j}D_{21}^{ij}}{D_{11}^{ij}+D_{11}^{i\pm1j}} \tag{2.6}
$$

Note that since D is a symmetric tensor we have: $D_{11,12}^{ij,i\pm1j} = D_{11,21}^{ij,i\pm1j}$ and $D_{22,21}^{ij,ij\pm1} = D_{22,12}^{ij,ij\pm1}$. The discrete system (2.4) is not closed since the number of unknowns is greater than the number of equations. Indeed there are $[N^2+(N+1)^2]$ unknowns and only N^2 equations. Therefore we should look for $(N+1)^2$ supplementary equations for closing that system. For this purpose let us introduce the dual grid $\mathcal{D}$ made of gridblocks $K_{i+\frac{1}{2},j+\frac{1}{2}}$ defined by $K_{i+\frac{1}{2},j+\frac{1}{2}} = \left]x_1^i, x_1^{i+1}\right[\times\left]x_2^j, x_2^{j+1}\right[$ for $i,\ j = 0, 1, ..., N$ where it is set

$$
x_1^0 = x_1^{\frac{1}{2}} = 0,\ \ x_2^0 = x_2^{\frac{1}{2}} = 0,\ \ x_1^{N+1} = x_1^{N+\frac{1}{2}} = 1,\ \ x_2^{N+1} = x_1^{\frac{1}{2}} = 1.
$$

Let us denote by D the generic name of dual gridblocks and x_D its center. The centers of some dual gridblocks are located on Γ the domain boundary. Such dual gridblocks are named degenerated dual gridblocks defining a set denoted by $\mathcal{D}_{deg}$ (see figure 2 above).

In the same order of idea, let us denote by $\mathcal{V}(D)$ the set of the vertices of the dual gridblock D, $\mathcal{E}$ the set of edges associated with the dual grid $\mathcal{D}$, $\mathcal{E}^{int}$ the set of edges E from

$\mathcal{E}$ such that $E \subset \Omega$, $\mathcal{E}^{ext}$ the set of edges E from $\mathcal{E}$ such that $E \subset \Gamma$, $\mathcal{E}^D$ the set of edges from $\mathcal{E}$ such that $\cup_{E\in\mathcal{E}^D} E = \partial D$, $\mathcal{N}(D)$ the set of dual gridblocks sharing a common edge with $D \in \mathcal{D}$.

Remark 2.1. One should note that:

- $\forall D, C \in \mathcal{D}$ such that $mes(\overline{D} \cap \overline{C}) \neq 0$, there exists two gridblocks P and L of primary gird such that $\overline{D} \cap \overline{C} = [x_P, x_L]$.
- $\forall D, C \in \mathcal{D}_{deg}$ *such that* $mes(\overline{D} \cap \overline{C}) \neq 0$, *there exists a unique* $P \in \mathcal{P}$ satisfying to $\overline{C} \cap \overline{D} = [x_P, x_{CD}]$, where x_{CD} is the midpoint of $[x_C, x_D] \subset \partial P \cap \Gamma$.
- The boundary of each dual gridblock is a union of a finite number of edges of the form $[x_P, x_L]$ or $[x_P, x_{CD}]$. $\diamond$

Let us now look for supplementary equations that should help to close the discrete system (2.4). For this purpose we introduce some useful notations: $\mathcal{E}_h$ is the set of mid-edge points (note that this set is the same for the primary and the dual grids); let $D \in \mathcal{D}$ (be a dual gridblock), $\mathcal{E}_h^D$ is a subset of $\mathcal{E}_h$ made of mid-edge points lying on Γ_D the boundary of D. For a given half-edge $[x_P, x_I]$ of D, where $P \in \mathcal{P}$ and $I \in \mathcal{E}_h^D$, it is natural to introduce $\xi^D_{[x_P,x_I]}$ the corresponding unit normal vector exterior to the half-plane from $\mathbb{R}^2$ containing the point x_D and bordered by the straight line $(x_P\ x_I)$, and $\xi^P_{[x_C,x_D]}$ the unit normal vector to $[x_C\,, x_D]$ exterior to the gridblock $P \in \mathcal{P}$, where $C \in \mathcal{N}(D)$ is such that $[x_C\,, x_D]$ is an edge of P. Let D^P denote the permeability tensor of any primary gridblock P. Then, it is easily seen that the following decomposition holds:

$$D^P \xi^P_{[x_C,x_D]} = a_h(D^P)\xi^P_{[x_C,x_D]} - b_h(D^P)\xi^D_{[x_P,x_I]} \tag{2.7}$$

$$D^P \xi^D_{[x_P,x_I]} = c_h(D^P)\xi^P_{[x_C,x_D]} - d_h(D^P)\xi^D_{[x_P,x_I]} \tag{2.8}$$

where the real numbers $a_h(D^P)$, $b_h(D^P)$, $c_h(D^P)$ and $d_h(D^P)$ are given by the relations:

$$a_h(D^P) = \left(\xi^P_{[x_C,x_D]}\right)^t D^P \left(\xi^P_{[x_C,x_D]}\right), \qquad b_h(D^P) = \left(\xi^D_{[x_P,x_I]}\right)^t D^P \left(\xi^P_{[x_C,x_D]}\right)$$

$$c_h(D^P) = \left(\xi^P_{[x_C,x_D]}\right)^t D^P \left(\xi^D_{[x_P,x_I]}\right), \qquad d_h(D^P) = \left(\xi^D_{[x_P,x_I]}\right)^t D^P \left(\xi^D_{[x_P,x_I]}\right).$$

Integrating the balance equation (1.1) in each dual gridblock $D = K_{i+\frac{1}{2},j+\frac{1}{2}}$ and applying the Ostrogradski's theorem to the left hand-side of the equality leads to diffusion flux computations over the boundary of D. Thanks to a suitable quadrature formula, one derives (as in the case of primary gridblocks) the following discrete balance equations in $D = K_{i+\frac{1}{2},j+\frac{1}{2}}$:

$$\begin{aligned}
&D^{ij+1,i+1j+1}_{11,12}\left[\varphi_{i,j+1} - \varphi_{i+1,j+1}\right] + \Delta^{ij+1,i+1j+1}_{22}\left[\varphi_{i+\frac{1}{2},j+\frac{1}{2}} - \varphi_{i+\frac{1}{2},j+\frac{3}{2}}\right]\\
&+D^{ij,i+1j}_{11,12}\left[\varphi_{i+1,j} - \varphi_{i,j}\right] + \Delta^{ij,i+1j}_{22}\left[\varphi_{i+\frac{1}{2},j+\frac{1}{2}} - \varphi_{i+\frac{1}{2},j-\frac{1}{2}}\right]\\
&+D^{i+1j,i+1j+1}_{22,21}\left[\varphi_{i+1,j} - \varphi_{i+1,j+1}\right] + \Delta^{i+1j,i+1j+1}_{11}\left[\varphi_{i+\frac{1}{2},j+\frac{1}{2}} - \varphi_{i+\frac{3}{2},j+\frac{1}{2}}\right]\\
&+D^{ij,ij+1}_{22,21}\left[\varphi_{i,j+1} - \varphi_{i,j}\right] + \Delta^{ij,ij+1}_{11}\left[\varphi_{i+\frac{1}{2},j+\frac{1}{2}} - \varphi_{i-\frac{1}{2},j+\frac{1}{2}}\right]\\
&\approx \int_{K_{i+\frac{1}{2},j+\frac{1}{2}}} f(x)dx \;-\int_{\Gamma\cap\partial K_{i+\frac{1}{2},j+\frac{1}{2}}} g d\tau \;+\sum_{C\in\mathcal{N}(D)\cap\mathcal{D}_{deg}} \frac{c_h(D^P)}{a_h(D^P)}\int_{[x_C,x_D]} g(x)d\tau\\
&\forall \quad 0 \leq\ i,\ j \leq N
\end{aligned} \tag{2.9}$$

where $P \in \mathcal{P}$ is a primary grid accepting x_D as its vertex and where we have set:

$$\begin{aligned}
\Delta_{11}^{ij,ij+1} &= -\frac{\left(D_{12}^{ij+1}-D_{12}^{ij}\right)^2}{2\left(D_{22}^{ij}+D_{22}^{ij+1}\right)} + \frac{D_{11}^{ij}+D_{11}^{ij+1}}{2} \\
\Delta_{11}^{i+1j,i+1j+1} &= -\frac{\left(D_{12}^{i+1j+1}-D_{12}^{i+1j}\right)^2}{2\left(D_{22}^{i+1j}+D_{22}^{i+1j+1}\right)} + \frac{D_{11}^{i+1j}+D_{11}^{i+1j+1}}{2}
\end{aligned} \tag{2.10}$$

$$\begin{aligned}
\Delta_{22}^{ij,i+1j} &= -\frac{\left(D_{21}^{i+1j}-D_{21}^{ij}\right)^2}{2\left(D_{11}^{ij}+D_{11}^{i+1j}\right)} + \frac{D_{22}^{ij}+D_{22}^{i+1j}}{2} \\
\Delta_{22}^{ij+1,i+1j+1} &= -\frac{\left(D_{21}^{i+1j+1}-D_{21}^{ij+1}\right)^2}{2\left(D_{11}^{ij+1}+D_{11}^{i+1j+1}\right)} + \frac{D_{22}^{ij}+D_{22}^{i+1j+1}}{2}
\end{aligned} \tag{2.11}$$

It is important to emphasize that the coefficients above are simply null if $i\, or\, j \in \{0, N+1\}$. Recall that the coefficients $D_{11,12}^{kl,mn}$ and $D_{22,21}^{kl,mn}$ involved in the above discrete balance equations i.e. (2.9) are defined by relations (2.5)-(2.6)

Note that the exact solution φ does not satisfy (2.4) and (2.9) with equalities everywhere. We derive the discrete system from (2.4)-(2.9) replacing φ and " $\approx$ " with u and " $=$ " respectively. Therefore the discrete problem consists in finding $\left\{u_{i,j}\right\}_{1\le i,j\le N}$ and $\left\{u_{i+\frac{1}{2},j+\frac{1}{2}}\right\}_{0\le i,j\le N}$ real unknowns such that:

$$\begin{aligned}
& D_{22,22}^{ij,ij+1}\left[u_{i,j}-u_{i,j+1}\right] + D_{22,21}^{ij,ij+1}\left[u_{i-\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{1}{2},j+\frac{1}{2}}\right] \\
&+ D_{22,22}^{ij,ij-1}\left[u_{i,j}-u_{i,j-1}\right] + D_{22,21}^{ij,ij-1}\left[u_{i+\frac{1}{2},j-\frac{1}{2}} - u_{i-\frac{1}{2},j-\frac{1}{2}}\right] \\
&+ D_{11,11}^{ij,i+1j}\left[u_{i,j}-u_{i+1,j}\right] + D_{11,12}^{ij,i+1j}\left[u_{i+\frac{1}{2},j-\frac{1}{2}} - u_{i+\frac{1}{2},j+\frac{1}{2}}\right] \\
&+ D_{11,11}^{ij,i-1j}\left[u_{i,j}-u_{i-1,j}\right] + D_{11,12}^{ij,i-1j}\left[u_{i-\frac{1}{2},j+\frac{1}{2}} - u_{i-\frac{1}{2},j-\frac{1}{2}}\right] \\
&= \int_{K_{i,j}} f(x)dx \;-\; \int_{\Gamma\cap\partial K_{i,j}} g(x)d\tau \quad \forall\ 1 \le i,j \le N
\end{aligned} \tag{2.12}$$

$$\begin{aligned}
& D_{11,12}^{ij+1,i+1j+1}\left[u_{i,j+1}-u_{i+1,j+1}\right] + \Delta_{22}^{ij+1,i+1j+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{1}{2},j+\frac{3}{2}}\right] \\
&+ D_{11,12}^{ij,i+1j}\left[u_{i+1,j}-u_{i,j}\right] + \Delta_{22}^{ij,i+1j}\left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{1}{2},j-\frac{1}{2}}\right] \\
&+ D_{22,21}^{i+1j,i+1j+1}\left[u_{i+1,j}-u_{i+1,j+1}\right] + \Delta_{11}^{i+1j,i+1j+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{3}{2},j+\frac{1}{2}}\right] \\
&+ D_{22,21}^{ij,ij+1}\left[u_{i,j+1}-u_{i,j}\right] + \Delta_{11}^{ij,ij+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i-\frac{1}{2},j+\frac{1}{2}}\right] \\
&= \int_{K_{i+\frac{1}{2},j+\frac{1}{2}}} f(x)dx \;-\int_{\Gamma\cap\partial K_{i+\frac{1}{2},j+\frac{1}{2}}} g d\tau \;+ \sum_{C\in\mathcal{N}(D)\cap\mathcal{D}_{deg}} \frac{c_h(D^P)}{a_h(D^P)} \int_{[x_C,\,x_D]} g(x)d\tau \\
&\forall\ \ 0 \le\ i,\ j \le N
\end{aligned} \tag{2.13}$$

2.2 Existence of discrete solutions and conditions for uniqueness

We are going to deal now with the existence and uniqueness of a solution for the discrete problem (2.12)-(2.13). Let us assume that the discrete unknowns are numbered from 1 to $m = N^2 + (N+1)^2$. Therefore the matrix form of this discrete system may be expressed as follows:

$$\begin{pmatrix} A & B \\ E & C \end{pmatrix}\begin{pmatrix} U_{cc} \\ U_{vc} \end{pmatrix} = \begin{pmatrix} F_{cc} \\ F_{vc} \end{pmatrix} \tag{2.14}$$

where we have set:

$$U_{cc} = \left\{u_{(j-1)N+i}\right\}_{1\le i,j\le N}\ ,\ \ U_{vc} = \left\{u_{j(N+1)+i+1+N^2}\right\}_{0\le i,j\le N} \tag{2.15}$$

$$F_{cc} = \left\{F_{(j-1)N+i}\right\}_{1\le i,j\le N} \quad \text{and} \quad F_{vc} = \left\{F_{j(N+1)+i+1+N^2}\right\}_{0\le i,j\le N} \tag{2.16}$$

where

$$F_{(j-1)N+i} \equiv F_{i,j} = \int_{K_{i,j}} f(x)dx \ - \ \int_{\Gamma\cap\partial K_{i,j}} g(x)d\tau \quad \forall\ 1 \le i,j \le N$$

$$F_{j(N+1)+i+1+N^2} \equiv F_{i+\frac{1}{2},j+\frac{1}{2}} = \int_{K_{i+\frac{1}{2},j+\frac{1}{2}}} f(x)dx \ - \int_{\Gamma\cap\partial K_{i+\frac{1}{2},j+\frac{1}{2}}} g d\tau$$

$$+ \sum_{C\in\mathcal{N}(D)\cap\mathcal{D}_{deg}} \frac{c_h(D^P)}{a_h(D^P)} \int_{[x_C,x_D]} g(x)d\tau \qquad \forall\ 0 \le i,j \le N$$

A is a N^2 symmetric matrix, associated to the classical grid-centered finite volume when $D(.)$ is diagonal i.e.$D_{12}(.) = D_{21}(.) \equiv 0$.

C is a $(N+1)^2$ symmetric matrix, associated to the classical vertex-centered finite volume when D is diagonal.

B is a $N^2 \times (N+1)^2$ matrix and E is a $(N+1)^2 \times N^2$ matrix. The following remarks will play a key role in what follows.

Remark 2.2.

$$\sum_{1\le i,j\le N} F_{i,j} = \int_\Omega f(x)dx - \int_\Gamma g d\tau \ = 0$$

$$\sum_{0\le i,j\le N} F_{i+\frac{1}{2},j+\frac{1}{2}} \ = \int_\Omega f(x)dx - \int_\Gamma g d\tau \ = 0. \quad \diamond$$

Remark 2.3.

$$\sum_{D\in\mathcal{D}} \sum_{C\in\mathcal{N}(D)\cap\mathcal{D}_{deg}} \frac{c_h(D^P)}{a_h(D^P)} \int_{[x_C,x_D]} g(x)d\tau = 0. \quad \diamond$$

Remark 2.4. Setting

$$M = \begin{bmatrix} A & B \\ E & C \end{bmatrix} = \left(M_{sp}\right)_{1\le s,p\le m}, \quad m = N^2 + (N+1)^2.$$

The matrix A, B, C, E and M satisfy the following properties:

1. M is symmetric matrix

2. $\forall\ 1 \leq s \leq N^2 \quad \sum_{p=1}^{N^2} A_{sp} = 0 \quad$ and $\quad \sum_{p=1}^{(N+1)^2} B_{sp} = 0$

3. $\forall\ 1 \leq s \leq (N+1)^2 \quad \sum_{p=1}^{N^2} E_{sp} = 0 \quad$ and $\quad \sum_{p=1}^{(N+1)^2} C_{sp} = 0$

4.
$$\forall\ 1 \leq s \leq m \quad \sum_{p=1}^{m} M_{sp} = 0 \tag{2.17}$$
$$\forall\ 1 \leq p \leq m \quad \sum_{s=1}^{m} M_{sp} = 0 \ \diamond$$

Proposition 2.5. *Let Λ_{cc} and Λ_{vc} be two vectors of $\mathbb{R}^m$ such that*

$$(\Lambda_{cc})_s = \begin{cases} 1 \text{ if } 1 \leq s \leq N^2 \\ 0 \quad \text{otherwise} \end{cases} \quad \text{and} \quad (\Lambda_{vc})_s = \begin{cases} 1 \text{ if } N^2+1 \leq s \leq m \\ 0 \quad \text{otherwise} \end{cases}$$

. Then $dim(Ker(M)) = 2$ and the family of vectors $\{\Lambda_{cc}, \Lambda_{vc}\}$ defines a basis of $Ker(M)$. $\diamond$

Proof. First remark that the vectors Λ_{cc} and Λ_{vc} are in $Ker(M)$ and that: for any $v \in \mathbb{R}^m$ $v \in Ker(M)$ implies $v^t M v = 0$. It remains to show that $v^t M v = 0$ implies $v \in Ker(M)$.

Step 1: We show that there exist a real number ρ only depending on Ω such that

$$\begin{bmatrix} U_{cc}^t \ U_{vc}^t \end{bmatrix} \begin{bmatrix} A & B \\ E & C \end{bmatrix} \begin{bmatrix} U_{cc} \\ U_{vc} \end{bmatrix} \geq \rho \left(|u_h|_{1,\mathcal{L}}^* \right)^2 \tag{2.18}$$

where

$$\begin{aligned} |u_h|_{1,\mathcal{L}}^* \quad = \quad & \left\{ \sum_{\substack{1 \leq i \leq N \\ 1 \leq j \leq N-1}} \left\{ \left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i-\frac{1}{2},j+\frac{1}{2}} \right]^2 + \left[u_{i,j+1} - u_{i,j} \right]^2 \right\} \right. \\ & + \sum_{\substack{1 \leq i \leq N-1 \\ 1 \leq j \leq N}} \left\{ \left[u_{i+1,j} - u_{i,j} \right]^2 + \left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{1}{2},j-\frac{1}{2}} \right]^2 \right\} \\ & + \sum_{j=1}^{N} \left\{ \left[u_{\frac{1}{2},j+\frac{1}{2}} - u_{\frac{1}{2},j-\frac{1}{2}} \right]^2 + \left[u_{N+\frac{1}{2},j+\frac{1}{2}} - u_{N+\frac{1}{2},j-\frac{1}{2}} \right]^2 \right\} \\ & \left. + \sum_{i=1}^{N} \left\{ \left[u_{i-\frac{1}{2},\frac{1}{2}} - u_{i+\frac{1}{2},\frac{1}{2}} \right]^2 + \left[u_{i+\frac{1}{2},N+\frac{1}{2}} - u_{i-\frac{1}{2},N+\frac{1}{2}} \right]^2 \right\} \right\}^{\frac{1}{2}}. \end{aligned} \tag{2.19}$$

Let us emphasize that notation $|\cdot|_{1,\mathcal{L}}^*$ will be made clear in Section 3.

Multiplying (2.12) by $u_{i,j}$ and (2.13) by $u_{i+\frac{1}{2},j+\frac{1}{2}}$ and summing leads to (here notations (2.15) are utilized)

$$\begin{bmatrix} U_{cc}^t & U_{vc}^t \end{bmatrix}\begin{bmatrix} A & B \\ E & C \end{bmatrix}\begin{bmatrix} U_{cc} \\ U_{vc} \end{bmatrix} = RHS\,1 + RHS\,2 + RHS\,3 \tag{2.20}$$

where we have set

$$\begin{aligned} RHS\,1 = \sum_{\substack{1\le i\le N \\ 1\le j\le N-1}} \Big\{ & D_{22,22}^{ij,ij+1}\left[u_{i,j+1}-u_{i,j}\right]^2 + \Delta_{11}^{ij,ij+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i-\frac{1}{2},j+\frac{1}{2}}\right]^2 \\ & + \; 2D_{22,21}^{ij,ij+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i-\frac{1}{2},j+\frac{1}{2}}\right]\left[u_{i,j+1}-u_{i,j}\right]\Big\} \end{aligned} \tag{2.21}$$

$$\begin{aligned} RHS\,2 = \sum_{\substack{1\le i\le N-1 \\ 1\le j\le N}} \Big\{ & D_{11,11}^{ij,i+1j}\left[u_{i+1,j}-u_{i,j}\right]^2 + \Delta_{22}^{ij,i+1j}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i+\frac{1}{2},j-\frac{1}{2}}\right]^2 \\ + \quad & 2\,D_{11,12}^{ij,i+1j}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i+\frac{1}{2},j-\frac{1}{2}}\right]\left[u_{i+1,j}-u_{i,j}\right]\Big\} \end{aligned} \tag{2.22}$$

$$\begin{aligned} RHS\,3 \;\; = \;\; & \sum_{j=1}^{N}\nabla_{11}^{1j}\left[u_{\frac{1}{2},j+\frac{1}{2}}-u_{\frac{1}{2},j-\frac{1}{2}}\right]^2 + \sum_{j=1}^{N}\nabla_{11}^{Nj}\left[u_{N+\frac{1}{2},j+\frac{1}{2}}-u_{N+\frac{1}{2},j-\frac{1}{2}}\right]^2 \\ & + \sum_{i=1}^{N}\nabla_{22}^{i1}\left[u_{i-\frac{1}{2},\frac{1}{2}}-u_{i+\frac{1}{2},\frac{1}{2}}\right]^2 + \sum_{i=1}^{N}\nabla_{22}^{iN}\left[u_{i+\frac{1}{2},N+\frac{1}{2}}-u_{i-\frac{1}{2},N+\frac{1}{2}}\right]^2 \end{aligned}$$

where

$$\nabla_{11}^{kj} = \frac{D_{22}^{kj}D_{11}^{kj}-\left(D_{12}^{kj}\right)^2}{2D_{11}^{kj}} \quad \forall\; 1\le j\le N \quad k=1,N$$

and

$$\nabla_{22}^{ik} = \frac{D_{22}^{ik}D_{11}^{ik}-\left(D_{12}^{ik}\right)^2}{2D_{22}^{ik}} \quad \forall\; 1\le i\le N \quad k=1,N$$

Let us note $D^{ij,ij+1}$ and $D^{ij,i+1j}$ the homogenized symmetric permeability tensors which appear respectively into the left hand-side of (2.21) and (2.22). It is easy to check that

$$det(D^{ij,i+1j}) = \frac{D_{11}^{i+1j}}{D_{11}^{ij}+D_{11}^{i+1j}}\times det(D^{ij}) + \frac{D_{11}^{ij}}{D_{11}^{ij}+D_{11}^{i+1j}}\times det(D^{i+1j}) \tag{2.23}$$

and

$$det(D^{ij,ij+1}) = \frac{D_{22}^{ij}}{D_{22}^{ij}+D_{22}^{ij+1}}\times det(D^{ij+1}) + \frac{D_{22}^{ij+1}}{D_{22}^{ij}+D_{22}^{ij+1}}\times det(D^{ij}) \tag{2.24}$$

Due to the positive definiteness of $D(x)$ and the relations (2.23) and (2.24) the symmetric matrices $D^{ij,ij+1}$ and $D^{ij,i+1j}$ are positive definite. Therefore each of these matrices possesses two strictly positive eigenvalues. Let $\lambda_{\min}^{ij,ij+1}$ and $\lambda_{\min}^{ij,i+1j}$ be respectively their lowest eigenvalues. So we have

$$RHS\,1 \geq \sum_{\substack{1\leq i\leq N\\1\leq j\leq N-1}} \lambda_{\min}^{ij,ij+1} \left\{\left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i-\frac{1}{2},j+\frac{1}{2}}\right]^2 + \left[u_{i,j+1} - u_{i,j}\right]^2\right\}$$

$$RHS\,2 \geq \sum_{\substack{1\leq i\leq N-1\\1\leq j\leq N}} \lambda_{\min}^{ij,i+1j} \left\{\left[u_{i+1,j} - u_{i,j}\right]^2 + \left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{1}{2},j-\frac{1}{2}}\right]^2\right\}$$

$$\begin{aligned} RHS\,3 \geq\ & \delta \sum_{j=1}^{N} \left\{\left[u_{\frac{1}{2},j+\frac{1}{2}} - u_{\frac{1}{2},j-\frac{1}{2}}\right]^2 + \left[u_{N+\frac{1}{2},j+\frac{1}{2}} - u_{N+\frac{1}{2},j-\frac{1}{2}}\right]^2\right\} \\ & +\beta \sum_{i=1}^{N} \left\{\left[u_{i-\frac{1}{2},\frac{1}{2}} - u_{i+\frac{1}{2},\frac{1}{2}}\right]^2 + \left[u_{i+\frac{1}{2},N+\frac{1}{2}} - u_{i-\frac{1}{2},N+\frac{1}{2}}\right]^2\right\} \end{aligned}$$

where

$$\begin{aligned} \delta &= \min\left\{\min_{1\leq j\leq N}(\nabla_{11}^{1j}), \min_{1\leq j\leq N}(\nabla_{11}^{Nj})\right\} > 0 \\ \beta &= \min\left\{\min_{1\leq j\leq N}(\nabla_{22}^{iN}), \min_{1\leq j\leq N}(\nabla_{22}^{iN})\right\} > 0 \end{aligned}$$

Thus

$$\begin{aligned} & RHS\,1 + RHS\,2 + RHS\,3 \\ \geq\ & \sum_{\substack{1\leq i\leq N\\1\leq j\leq N-1}} \lambda_{\min}^{ij,ij+1} \left\{\left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i-\frac{1}{2},j+\frac{1}{2}}\right]^2 + \left[u_{i,j+1} - u_{i,j}\right]^2\right\} \\ & + \sum_{\substack{1\leq i\leq N-1\\1\leq j\leq N}} \lambda_{\min}^{ij,i+1j} \left\{\left[u_{i+1,j} - u_{i,j}\right]^2 + \left[u_{i+\frac{1}{2},j+\frac{1}{2}} - u_{i+\frac{1}{2},j-\frac{1}{2}}\right]^2\right\} \\ & +\delta \sum_{j=1}^{N} \left\{\left[u_{\frac{1}{2},j+\frac{1}{2}} - u_{\frac{1}{2},j-\frac{1}{2}}\right]^2 + \left[u_{N+\frac{1}{2},j+\frac{1}{2}} - u_{N+\frac{1}{2},j-\frac{1}{2}}\right]^2\right\} \\ & +\beta \sum_{i=1}^{N} \left\{\left[u_{i-\frac{1}{2},\frac{1}{2}} - u_{i+\frac{1}{2},\frac{1}{2}}\right]^2 + \left[u_{i+\frac{1}{2},N+\frac{1}{2}} - u_{i-\frac{1}{2},N+\frac{1}{2}}\right]^2\right\} \\ & \geq \rho\left(|u_h|_{1,\mathcal{L}}^{*}\right)^2 \end{aligned}$$

where

$$\rho = \min\left\{\beta, \delta, \min_{\substack{1\leq i\leq N\\1\leq j\leq N-1}} \lambda_{\min}^{ij,ij+1}, \min_{\substack{1\leq i\leq N-1\\1\leq j\leq N}} \lambda_{\min}^{ij,i+1j}\right\}$$

is actually a real positive number depending exclusively on the geological structure of the medium. According to what precedes we have

$${}^{t}vMv = 0 \Longrightarrow 0 \leq \rho\left(|v|_{1,\mathcal{L}}^{*}\right)^2 \leq 0 \Longrightarrow |v|_{1,\mathcal{L}}^{*} = 0$$

Thus there exists two constant real numbers θ and μ such that

$$\forall 1 \le s \le N^2 \quad (v_{cc})_s = \theta \qquad \text{and} \qquad \forall N^2+1 \le s \le m \quad (v_{vc})_s = \mu.$$

For $1 \le s \le N^2$

$$\begin{aligned}(Mv)_s &= \sum_{p=1}^{N^2} A_{sp}\,(v_{cc})_p + \sum_{p=N^2+1}^{m} B_{s,p-N^2}\,(v_{vc})_p \\ &= \theta \sum_{p=1}^{N^2} A_{sp} + \mu \sum_{p=1}^{(N+1)^2} B_{sp} = 0\end{aligned}$$

and for $N^2+1 \le s \le m$

$$\begin{aligned}(Mv)_s &= \sum_{p=1}^{N^2} E_{sp}\,(v_{cc})_p + \sum_{p=N^2+1}^{m} C_{s,p-N^2}\,(v_{vc})_p \\ &= \theta \sum_{p=1}^{N^2} E_{sp} + \mu \sum_{p=1}^{(N+1)^2} C_{sp} = 0\end{aligned}$$

by using the remark 2.4. Therefore $Mv = 0$ and $v \in Ker(M)$.

Step 2: Deduce from Step 1 that for $v \in \mathbb{R}^m$:

$$v^t M v = 0 \Leftrightarrow \exists\, \theta,\mu \in \mathbb{R} \text{ such that } v = \theta\Lambda_{cc} + \mu\Lambda_{vc} \tag{2.25}$$

For all $v \in \mathbb{R}^m, v^t M v = 0$ implies (according to Step1) that

$$\exists\, \theta,\mu \in \mathbb{R} \text{ such that } v = \theta\Lambda_{cc} + \mu\Lambda_{vc} \tag{2.26}$$

Reversely, if v satisfies (2.26) then $Mv = 0$ thanks to Remark 2.4. From Step 1 we deduce that $v^t M v = 0$. Then $dim(Ker(M)) = 2$ and the family of vectors $\{\Lambda_{cc}, \Lambda_{vc}\}$ defines a basis of $Ker(M)$. The proof of Proposition 2.5 is ended.

□

Proposition 2.6. *The discrete problem (2.12)-(2.13) possesses an infinite number of solutions.* ◇

Proof. It is based upon the discrete version of Fredholm Alternative and Remark 2.2. □

Proposition 2.7. *There is a unique vector of $u_h \in \mathbb{R}^m$ satisfying the discrete equations (2.12)-(2.13) together with the following constraints*

$$\sum_{P\in\mathcal{P}} \frac{mes(P)}{5} \sum_{x^P \in T} u(x^P) = 0 \tag{2.27}$$

$$\sum_{T\in\mathcal{T}} \frac{mes(T)}{3} \sum_{x^T \in T} u(x^T) = 0. \quad \diamond \tag{2.28}$$

Proof. It suffices to show that for $v \in \mathbb{R}^m$
[$v^t M v = 0$ and v satisfies (2.27)-(2.28)] $\Rightarrow v = 0$.
Consider $v \in \mathbb{R}^m$ such that for all $v \in \mathbb{R}^m$ $v^t M v = 0$. Then (according to what precedes):

$$\exists\, \theta, \mu \in \mathbb{R} \text{ such that } v = \theta\Lambda_{cc} + \mu\Lambda_{vc}$$

If v satisfies (2.27) and (2.28) then

$$\theta + 4\mu = 0 \qquad \textit{and} \qquad \theta + 2\mu = 0$$

Therefore $\theta = \mu = 0$ and thus $v = 0$. □

3 Stability and error estimates

In what precedes we have shown the existence and uniqueness of a solution to the discrete problem (2.12)-(2.13) with the constraints (2.27)-(2.28). In what follows we derive from that solution some approximate solutions for the model problem (1.1)-(1.2).

3.1 A piecewise constant approximate solution

Let us start with introducing a new grid $\mathcal{L}$ made up of rhombi recovering Ω as indicated in Figure 3 below. With the grid $\mathcal{L}$ is associated a space $E(\mathcal{L})$ made up of functions v defined

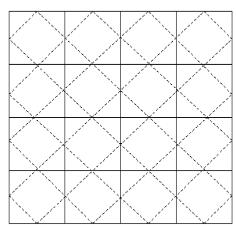

Figure 3. An example of grid $\mathcal{L}$ made up of rhombi associated with a primary square mesh.

almost every in Ω such that v is constant in $L \cap \Omega$, for any $L \in \mathcal{L}$. We equip the space $E(\mathcal{L})$ with the following discrete $H^1(\Omega)$ norm:

$$\forall v \in E(\mathcal{L}) \qquad ||v||_{\mathbf{E}(\mathcal{L})} = \left[\sum_{s \in \mathbf{V}(\mathcal{L})} (\Delta_s v)^2 + \left(\int_\Omega v\, dx \right)^2 \right]^{\frac{1}{2}} \tag{3.1}$$

where $\mathbf{V}(\mathcal{L})$ is the set of vertices associated with the grid $\mathcal{L}$ and where we have set:

$$\Delta_s v = \sum_{\substack{L,K \in \mathcal{L} \text{ such that} \\ \Gamma_K \cap \Gamma_L = \{s\}}} |v_L - v_K|^2$$

with Γ_E denoting the boundary of any mesh element E. Let $S_h(\mathcal{L})$ be the subspace of $E(\mathcal{L})$ made up of functions v that satisfy to (2.27)-(2.28). $S_h(\mathcal{L})$ is not empty as it contains the cell-wise constant function corresponding to the solution to the discrete problem (2.12)-(2.13). In what follows this function is denoted by $u_{\mathcal{L}}$ and is called "piecewise constant approximate solution" to the model problem.

Remark 3.1. The discrete H^1 semi-norm

$$v \mapsto |v|^*_{1,\mathcal{L}} = \left[\sum_{s \in \mathbf{V}(\mathcal{L})} (\Delta_s v)^2 \right]^{\frac{1}{2}} \tag{3.2}$$

defines a norm over $S_h(\mathcal{L})$ and this norm is equivalent to the discrete H^1 norm defined by relation (3.1).

3.2 A piecewise linear approximate solution

One can easily define a piecewise linear approximate solution for the model problem (1.1)-(1.2) by using an adequate triangulation $\mathcal{T}$ of the domain Ω (see Figure 4 below). As in P_1 finite element theory this solution denoted by $u_{\mathcal{T}}$ is continuous over $\overline{\Omega} = [0,1] \times [0,1]$. Recall that the quadrature formulae (2.27) is exact for functions of class P_1 over triangular elements from $\mathcal{T}$. Since $u_{\mathcal{T}}$ is in $H^1(\Omega)$, it is clear that $u_{\mathcal{T}}$ satisfies to the constraint (1.8) which is imposed to the exact solution i.e.

$$u_{\mathcal{T}} \in V = \{v \in H^1(\Omega); \ \int_{\Omega} v(x)dx = 0\} \tag{3.3}$$

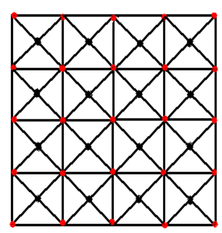

Figure 4. Triangulation of the domain associated with the primary grid

3.3 Stability of the piecewise constant approximate solution

Let us start with the following important remark.

Remark 3.2. One can prove that there exists a constant C without dependence on the spatial discretization such that (see [9]):

$$\forall v \in E(\mathcal{L}) \qquad ||v||^2_{L^2(\Omega)} \le C||v||^2_{E(\mathcal{L})}. \quad \diamond \tag{3.4}$$

An immediate consequence of the preceding remark is that the mappings

$$||v||_{E(\mathcal{L})} = \left[\sum_{s \in V(\mathcal{L})} (\Delta_s v)^2 + \left(\int_{\Omega} v dx \right)^2 \right]^{\frac{1}{2}} \tag{3.5}$$

and

$$|||v|||_{E(\mathcal{L})} = \left[\sum_{s \in V(\mathcal{L})} (\Delta_s v)^2 + \int_\Omega v^2 dx \right]^{\frac{1}{2}} \tag{3.6}$$

define two equivalent norms over the space $E(\mathcal{L})$.

Let us introduce now two projection operators defined as follows:

$$v \in E(\mathcal{L}) \longmapsto \Pi_{\mathcal{P}} v \equiv v^{\mathcal{P}} = \{v_P\}_{P \in \mathcal{P}} \in E(\mathcal{P})$$

and

$$v \in E(\mathcal{L}) \longmapsto \Pi_{\mathcal{D}} v \equiv v^{\mathcal{D}} = \{v_D\}_{D \in \mathcal{D}} \in E(\mathcal{D})$$

where $E(\mathcal{P})$ and $E(\mathcal{D})$ are respectively the spaces of constant functions over primary and dual grid blocks. These spaces are endowed respectively with the following norms:

$$\|w\|_{E(\mathcal{P})} = \left[\sum_{i \in I(\mathcal{L})} \sum_{\substack{K,L \in \mathcal{P} \text{ with} \\ i \in \Gamma_K \cap \Gamma_L}} |w_K - w_L|^2 \right]^{\frac{1}{2}} \tag{3.7}$$

and

$$\|w\|_{E(\mathcal{D})} = \left[\sum_{i \in I(\mathcal{L})} \sum_{\substack{K^*,L^* \in \mathcal{D} \text{ with} \\ i \in \Gamma_{K^*} \cap \Gamma_{L^*}}} |w_{K^*} - w_{L^*}|^2 \right]^{\frac{1}{2}} \tag{3.8}$$

where $I(\mathcal{L})$ is the set of gridblock centres with respect to the grid $\mathcal{L}$.

Remark 3.3. Note that:

$$|v|^*_{1,\mathcal{L}} = \left\{ \left\| v^{\mathcal{P}} \right\|^2_{E(\mathcal{P})} + \left\| v^{\mathcal{D}} \right\|^2_{E(\mathcal{D})} \right\}^{\frac{1}{2}} \qquad \forall v \in \mathbf{E}(\mathcal{L}). \tag{3.9}$$

and that

$$\left\| v^{\mathcal{P}} \right\|_{E(\mathcal{P})} \le \|v\|_{E(\mathcal{L})}, \qquad \left\| v^{\mathcal{D}} \right\|_{E(\mathcal{D})} \le \|v\|_{E(\mathcal{L})} \qquad \forall v \in E(\mathcal{L}) \tag{3.10}$$

These obvious relations play a key-role in what follows. $\diamond$

The following result ([9]) plays a central role in the proof of the stability of the piecewise constant approximate solution to the model problem.

Lemma 3.4. *Let $\mathcal{F}$ be a rectangular grid defined over Ω and $E(\mathcal{F})$ the space of functions v defined on Ω such that $v|_M$ is a constant for all $M \in \mathcal{F}$. For every $v \in E(\mathcal{F})$, define: $[\gamma_0(v)](x) = v_M$ for almost every (in the sense of 1D-Lebesgue measure) $x \in \Gamma \cap \Gamma_M$, where M is a gridblock from $\mathcal{F}$ adjacent to the domain boundary Γ. Then we have*

$$\|\gamma_0(v)\|_{L^2(\Gamma)} \le C \|v\|_{E(\mathcal{F})} \qquad \forall v \in E(\mathcal{F})$$

where

$$\|v\|_{E(\mathcal{F})} = \left[\sum_{s \in V(\mathcal{F})} (\Delta_s v)^2 + \left(\int_\Omega v dx \right)^2 \right]^{\frac{1}{2}} \qquad \forall v \in E(\mathcal{F}). \qquad \diamond$$

All the ingredients are gathered for proving the following important result.

Theorem 3.5. *(Stability result)*
Let us assume that the data f and g are sufficiently regular and satisfy to the compatibility condition (1.7). Then the piecewise constant approximate solution to the model problem (1.1)-(1.2) obeys to the following inequality:

$$||u_{\mathcal{L}}||_{E(\mathcal{L})} \leq C\Big[||f||_{L^2(\Omega)} + ||g||_{L^2(\Gamma)}\Big],$$

or equivalently

$$|u_{\mathcal{L}}|^*_{1,\mathcal{L}} \leq C\Big[||f||_{L^2(\Omega)} + ||g||_{L^2(\Gamma)}\Big]$$

where the strictly positive number C represents diverse constants mesh independent. ◇

Proof.

$$[U_{cc}\ Uvc]\begin{bmatrix} A & B \\ E & C \end{bmatrix}\begin{bmatrix} U_{cc} \\ U_{vc} \end{bmatrix} = [U_{cc}\ U_{vc}]\begin{bmatrix} F_{cc} \\ F_{vc} \end{bmatrix} \tag{3.11}$$

$$LHS\ = [U_{cc}\ Uvc]\begin{bmatrix} A & B \\ E & C \end{bmatrix}\begin{bmatrix} U_{cc} \\ U_{vc} \end{bmatrix}$$

$$RHS\ = [U_{cc}\ Uvc]\begin{bmatrix} F_{cc} \\ F_{vc} \end{bmatrix}$$

From the step 1 of the proof of Proposition 2.5 we know that the left hand side of (3.11) satisfies to the following inequality:

$$\rho\Big(|u_h|^*_{1,\mathcal{L}}\Big)^2 \leq LHS \tag{3.12}$$

In addition, the right hand side of (3.11) obeys to the following relation

$$\begin{aligned}
RHS \quad &= \quad \sum_{P\in\mathcal{P}}\int_P f(x)u_P\,dx + \sum_{D\in\mathcal{D}}\int_D f(x)u_D\,dx \\
&+ \quad \sum_{P\in\mathcal{P}}\int_{\Gamma\cap\Gamma_P} g(x)u_P\,d\tau(x) - \sum_{D\in\mathcal{D}}\int_{\Gamma\cap\Gamma_D} g(x)u_D\,d\tau(x) \\
&+ \quad \sum_{i=1}^{N}\Big[u_{i-\frac{1}{2},\frac{1}{2}} - u_{i+\frac{1}{2},\frac{1}{2}}\Big]\frac{D_{12}^{i1}}{D_{22}^{i1}}\int_{K_{i1}\cap\Gamma} gd\tau \\
&+ \quad \sum_{i=1}^{N}\Big[u_{i+\frac{1}{2},N+\frac{1}{2}} - u_{i-\frac{1}{2},N+\frac{1}{2}}\Big]\frac{D_{12}^{iN}}{D_{22}^{iN}}\int_{K_{iN}\cap\Gamma} gd\tau \\
&+ \quad \sum_{j=1}^{N}\Big[u_{\frac{1}{2},j+\frac{1}{2}} - u_{\frac{1}{2},j-\frac{1}{2}}\Big]\frac{D_{12}^{1j}}{D_{11}^{i1}}\int_{K_{1j}\cap\Gamma} gd\tau \\
&+ \quad \sum_{j=1}^{N}\Big[u_{N+\frac{1}{2},j-\frac{1}{2}} - u_{N+\frac{1}{2},j+\frac{1}{2}}\Big]\frac{D_{12}^{1j}}{D_{22}^{i1}}\int_{K_{Nj}\cap\Gamma} gd\tau
\end{aligned} \tag{3.13}$$

where the convention that the integral term is zero if $\Gamma\cap\Gamma_P$ or $\Gamma\cap\Gamma_D$ are empty sets. Recall that $\mathcal{P}$ and $\mathcal{D}$ are respectively the (set of gridblocks defining the) primary grid and the (set of gridblocks defining the) dual grid. By a double application of Cauchy-Schwarz inequality and thanks to Remark 3.1 and Remark 3.3 one can see that on one hand we have

$$\left|\sum_{P\in\mathcal{P}}\int_P f(x)u_P\,dx+\sum_{D\in\mathcal{D}}\int_D f(x)u_D\,dx\right|\le\sqrt{2}\,\|f\|_{L^2(\Omega)}\,|u_h|^*_{1,\mathcal{L}} \tag{3.14}$$

and

$$\begin{aligned}&\left|\sum_{P\in\mathcal{P}}\int_{\Gamma\cap\Gamma_P} g(x)u_P\,d\tau(x)-\sum_{D\in\mathcal{D}}\int_{\Gamma\cap\Gamma_D} g(x)u_D\,d\tau(x)\right|\\ \le\;& \|g\|_{L^2(\Gamma)}\left\|\gamma_0[(u_{\mathcal{L}})^{\mathcal{P}}]\right\|_{L^2(\Gamma)}+\|g\|_{L^2(\Gamma)}\left\|\gamma_0[(u_{\mathcal{L}})^{\mathcal{D}}]\right\|_{L^2(\Gamma)}\\ \le\;& C\|g\|_{L^2(\Gamma)}\left\|(u_{\mathcal{L}})^{\mathcal{P}}\right\|_{E(\mathcal{P})}+C\|g\|_{L^2(\Gamma)}\left\|(u_{\mathcal{L}})^{\mathcal{D}}\right\|_{E(\mathcal{D})} \qquad (3.15)\\ \le\;& C\|g\|_{L^2(\Gamma)}\,|u_h|^*_{1,\mathcal{L}}\end{aligned}$$

where C represents diverse strictly positive number mesh independent. On the other hand, we have

$$\begin{aligned}&\left|\sum_{i=1}^{N}\left[u_{i-\frac12,\frac12}-u_{i+\frac12,\frac12}\right]\frac{D_{12}^{i1}}{D_{22}^{i1}}\int_{K_{i1}\cap\Gamma} g\,d\tau\right|\\ \le\;& \sigma\sum_{i=1}^{N} mes(K_{i1}\cap\Gamma)\left|u_{i-\frac12,\frac12}-u_{i+\frac12,\frac12}\right|\left(\int_{K_{i1}\cap\Gamma} g^2\,d\tau\right)^{\frac12}\\ \le\;& \sigma\, mes(\Gamma)\left[\sum_{i=1}^{N}\left(u_{i-\frac12,\frac12}-u_{i+\frac12,\frac12}\right)^2\right]^{\frac12}\left[\sum_{i=1}^{N}\int_{K_{i1}\cap\Gamma} g^2\,d\tau\right]^{\frac12}\\ \le\;& \sigma\, mes(\Gamma)\,\|g\|_{L^2(\Gamma)}\,|u_h|^*_{1,\mathcal{L}} \qquad (3.16)\end{aligned}$$

where σ a real positive number given by (1.4). From the above estimates it follows that

$$RHS\le C\left[\|f\|_{L^2(\Omega)}+\|g\|_{L^2(\Gamma)}\right]|u_h|^*_{1,\mathcal{L}} \tag{3.17}$$

Thanks to (3.12) and (3.17) we get

$$|u_h|^*_{1,\mathcal{L}}\le C\left[\|f\|_{L^2(\Omega)}+\|g\|_{L^2(\Gamma)}\right].$$

This ends the proof. □

3.4 Error estimates for piecewise constant/linear solutions

Theorem 3.6. *(Error estimates for piecewise constant solution $u_{\mathcal{L}}$)*
Under the assumptions (1.5)-(1.7) and the condition $f\in L^2(\Omega)$, the unique variational solution to the model problem (1.1)-(1.2) is such that its restriction to any primary gridblock $P\in\mathcal{P}$ lies in $C^2(\overline{P})$ and the error function $\varepsilon_h=\varphi_h-u_{\mathcal{L}}$ obeys to the following inequalities:

$$\|\varepsilon_h\|_{L^2(\Omega)}+|\varepsilon_h|^*_{1,\mathcal{L}}\le Ch \tag{3.18}$$

where C represents diverse positive constants mesh independent. Recall that $u_{\mathcal{L}} = \{u_N\}_{N\in\mathcal{L}}$ and $\varphi_h = \{\varphi(x_N)\}_{N\in\mathcal{L}}$, x_N being the center of the gridblock, $N \in \mathcal{L}$ and h being the size of $\mathcal{L}$. $\diamond$

Proof. Taking account truncation errors, the equations (2.12)-(2.13) are transformed as follows :

$$\begin{aligned}
&D_{22,22}^{ij,ij+1}\left[u_{i,j}-u_{i,j+1}\right]+D_{22,21}^{ij,ij+1}\left[u_{i-\frac{1}{2},j+\frac{1}{2}}-u_{i+\frac{1}{2},j+\frac{1}{2}}\right]\\
&+D_{22,22}^{ij,ij-1}\left[u_{i,j}-u_{i,j-1}\right]+D_{22,21}^{ij,ij-1}\left[u_{i+\frac{1}{2},j-\frac{1}{2}}-u_{i-\frac{1}{2},j-\frac{1}{2}}\right]\\
&+D_{11,11}^{ij,i+1j}\left[u_{i,j}-u_{i+1,j}\right]+D_{11,12}^{ij,i+1j}\left[u_{i+\frac{1}{2},j-\frac{1}{2}}-u_{i+\frac{1}{2},j+\frac{1}{2}}\right]\\
&+D_{11,11}^{ij,i-1j}\left[u_{i,j}-u_{i-1,j}\right]+D_{11,12}^{ij,i-1j}\left[u_{i-\frac{1}{2},j+\frac{1}{2}}-u_{i-\frac{1}{2},j-\frac{1}{2}}\right]\\
&=\int_{K_{i,j}}f(x)dx+\sum_{e\in E_{i,j}}hR_{i,j}^{e}\quad\forall\quad 1\le i,j\le N
\end{aligned}\tag{3.19}$$

$$\begin{aligned}
&D_{11,12}^{ij+1,i+1j+1}\left[u_{i,j+1}-u_{i+1,j+1}\right]+\Delta_{22}^{ij+1,i+1j+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i+\frac{1}{2},j+\frac{3}{2}}\right]\\
&+D_{11,12}^{ij,i+1j}\left[u_{i+1,j}-u_{i,j}\right]+\Delta_{22}^{ij,i+1j}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i+\frac{1}{2},j-\frac{1}{2}}\right]\\
&+D_{22,21}^{i+1j,i+1j+1}\left[u_{i+1,j}-u_{i+1,j+1}\right]+\Delta_{11}^{i+1j,i+1j+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i+\frac{3}{2},j+\frac{1}{2}}\right]\\
&+D_{22,21}^{ij,ij+1}\left[u_{i,j+1}-u_{i,j}\right]+\Delta_{11}^{ij,ij+1}\left[u_{i+\frac{1}{2},j+\frac{1}{2}}-u_{i-\frac{1}{2},j+\frac{1}{2}}\right]\\
&=\int_{K_{i+\frac{1}{2},j+\frac{1}{2}}}f(x)dx-\int_{\Gamma\cap\partial K_{i+\frac{1}{2},j+\frac{1}{2}}}gd\tau+\sum_{C\in\mathcal{N}(D)\cap\mathcal{D}_{deg}}\frac{c_h(D^P)}{a_h(D^P)}\int_{[x_C,x_D]}g(x)d\tau\\
&+\sum_{e\in E_{i+\frac{1}{2},j+\frac{1}{2}}}hR_{i+\frac{1}{2},j+\frac{1}{2}}^{e}\quad\text{for all}\quad 0\le\ i,\ j\le N
\end{aligned}\tag{3.20}$$

where $E_{i,j}$ and $E_{i+\frac{1}{2},j+\frac{1}{2}}$ are sets of edges associated respectively with $K_{i,j}$ and $K_{i+\frac{1}{2},j+\frac{1}{2}}$, and where $R^e_{i,j}$ and $R^e_{i+\frac{1}{2},j+\frac{1}{2}}$ denote the truncation error associated with the approximation of the flux over the edges $e_{i,j}$ and $e_{i+\frac{1}{2},j+\frac{1}{2}}$ respectively. Moreover, under the assumption $\varphi \in C^2$ over the closure of primary grid-blocks, the truncation error satisfy the following inequalities :

$$\left|R^e_{i,j}\right| \le Ch \qquad \text{and} \qquad \left|R^e_{i+\frac{1}{2},j+\frac{1}{2}}\right| \le Ch \tag{3.21}$$

In what follows, the notation R^e_K will be used to denote the truncation error for the approximation of the flux over the edge e_K of any control volume K. Due to the conservatively property of the proposed finite volume formulation, we have

$$R^e_K \ + \ R^e_I = 0 \tag{3.22}$$

where K and I are two adjacent control volumes such that $e = \Gamma_K \cap \Gamma_I$.

Let us define a function ε_h almost everywhere in $\overline{\Omega}$ in the following way:

$$\varepsilon_h(x) = \varepsilon_L \quad \text{if } x \in Int(L) \text{ with } L \in \mathcal{L} \tag{3.23}$$

where we have set $\varepsilon_L = \varphi_L - u_L$ for all $L \in \mathcal{L}$. Note that the element L of the additive mesh $\mathcal{L}$ is necessary centered on a point whose cartesian coordinates are of the form $\left(x_1^i, x_2^j\right)$ or $\left(x_1^{i+\frac{1}{2}}, x_2^{j+\frac{1}{2}}\right)$. ε_L is a generic name of $\varepsilon_{i,j}$ or $\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}$.

Remark 3.7. From the relation (3.22) we see that the function ε_h is actually in the space $\mathbf{E}(\mathcal{L})$. This function expresses the error in some sense (i.e. the difference between the exact and the weak approximate solution u_h) and certain estimates of this error are given in what follows ·

We immediately should show that the following quantities $\left\{\varepsilon_{i,\,j}\right\}_{1\le i,\ j\le N}$ and $\left\{\varepsilon_{i+\frac{1}{2},\,j+\frac{1}{2}}\right\}_{1\le i,\ j\le N-1}$ are a solution of a discrete problem of the form (2.12)-(2.13). Subtracting (2.12) from (3.19) and (2.13) from (3.20), and reordering the terms yields:

$$\begin{aligned}
& D_{22,22}^{ij,ij+1}\left[\varepsilon_{i,j}-\varepsilon_{i,j+1}\right]+D_{22,21}^{ij,ij+1}\left[\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}\right] \\
&+ D_{22,22}^{ij,ij-1}\left[\varepsilon_{i,j}-\varepsilon_{i,j-1}\right]+D_{22,21}^{ij,ij-1}\left[\varepsilon_{i+\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{i-\frac{1}{2},j-\frac{1}{2}}\right] \\
&+D_{11,11}^{ij,i+1j}\left[\varepsilon_{i,j}-\varepsilon_{i+1,j}\right]+ D_{11,12}^{ij,i+1j}\left[\varepsilon_{i+\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}\right] \\
&+D_{11,11}^{ij,i-1j}\left[\varepsilon_{i,j}-\varepsilon_{i-1,j}\right]+D_{11,12}^{ij,i-1j}\left[\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i-\frac{1}{2},j-\frac{1}{2}}\right] \\
&= \sum_{e\in E_{i,j}} hR_{i,j}^e \quad \forall \quad 1\le i,j\le N
\end{aligned} \tag{3.24}$$

and

$$\begin{aligned}
& D_{11,12}^{ij+1,i+1j+1}\left[\varepsilon_{i,j+1}-\varepsilon_{i+1,j+1}\right]+\Delta_{22}^{ij,i+1j+1}\left[\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j+\frac{3}{2}}\right] \\
&+D_{11,12}^{ij+1,i+1j}\left[\varepsilon_{i+1,j}-\varepsilon_{i,j}\right]+\Delta_{22}^{ij,i+1j}\left[\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j-\frac{1}{2}}\right] \\
&+D_{22,21}^{i+1j,i+1j+1}\left[\varepsilon_{i+1,j}-u_{i+1,j+1}\right]+\Delta_{11}^{i+1j,i+1j+1}\left[\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i+\frac{3}{2},j+\frac{1}{2}}\right] \\
&+D_{22,21}^{ij,ij+1}\left[\varepsilon_{i,j+1}-\varepsilon_{i,j}\right]+\Delta_{11}^{i+1j,i+1j+1}\left[\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}}\right] \\
&= \sum_{e\in E_{i+\frac{1}{2},j+\frac{1}{2}}} hR_{i+\frac{1}{2},j+\frac{1}{2}}^e \quad \text{for all} \quad 1\le\ i,\ j\le N-1
\end{aligned} \tag{3.25}$$

Multiplying (3.24) by $\varepsilon_{i,j}$ and (3.25) by $\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}$ and summing over i, j and reordering the terms of the left hand side after summation side by side of the results obtained, leads to the following inequality, thanks to (1.6) and (3.22):

$$\gamma\left(|\varepsilon_h|^*_{1,\mathcal{L}}\right)^2 \leq$$

$$\sum_{\substack{1\leq i\leq N\\ 1\leq j\leq N-1}} h\left\{\left(\varepsilon_{i,j}-\varepsilon_{i,j+1}\right)R_{i,j}^{e^{ij,ij+1}}+\left(\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}\right)R_{i-\frac{1}{2},j+\frac{1}{2}}^{e^{i-\frac{1}{2}j+\frac{1}{2},i+\frac{1}{2}j+\frac{1}{2}}}\right\}$$

$$+\sum_{\substack{1\leq i\leq N-1\\ 1\leq j\leq N}} h\left\{\left(\varepsilon_{i,j}-\varepsilon_{i+1,j}\right)R_{i,j}^{e^{ij,i+1j}}+\left(\varepsilon_{i+\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}\right)R_{i+\frac{1}{2},j-\frac{1}{2}}^{e^{i+\frac{1}{2}j-\frac{1}{2},i+\frac{1}{2}j+\frac{1}{2}}}\right\}$$

$$+\sum_{i=1}^{N} h\left\{\left(\varepsilon_{i-\frac{1}{2},\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},\frac{1}{2}}\right)R_{i-\frac{1}{2},\frac{1}{2}}^{e^{i-\frac{1}{2}\frac{1}{2},i+\frac{1}{2}\frac{1}{2}}}+\left(\varepsilon_{i-\frac{1}{2},N+\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},N+\frac{1}{2}}\right)R_{i-\frac{1}{2},N+\frac{1}{2}}^{e^{i-\frac{1}{2}N+\frac{1}{2},i+\frac{1}{2}N+\frac{1}{2}}}\right\}$$

$$+\sum_{j=1}^{N} h\left\{\left(\varepsilon_{\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{\frac{1}{2},j+\frac{1}{2}}\right)R_{\frac{1}{2},j-\frac{1}{2}}^{e^{\frac{1}{2}j-\frac{1}{2},\frac{1}{2}j+\frac{1}{2}}}+\left(\varepsilon_{N+\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{N+\frac{1}{2},j+\frac{1}{2}}\right)R_{i+\frac{1}{2},N-\frac{1}{2}}^{e^{N+\frac{1}{2}j-\frac{1}{2},N+\frac{1}{2}j+\frac{1}{2}}}\right\}$$

where $\qquad e^{K,L}=\Gamma_K\cap\Gamma_L.$

Therefore

$$\gamma\left(|\varepsilon_h|^*_{1,h}\right)^2 \leq h\sum_{1\leq i\leq N,\ 1\leq j\leq N-1} a_{ij}\left[\left|\varepsilon_{i,j}-\varepsilon_{i,j+1}\right|+\left|\varepsilon_{i-\frac{1}{2},j+\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}\right|\right]$$

$$+h\sum_{1\leq i\leq N-1,\ 1\leq j\leq N} b_{ij}\left[\left|\varepsilon_{i,j}-\varepsilon_{i+1,j}\right|+\left|\varepsilon_{i+\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},j+\frac{1}{2}}\right|\right]$$

$$+h\sum_{i=1}^{N}\left[C_{i,\frac{1}{2}}\left|\varepsilon_{i-\frac{1}{2},\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},\frac{1}{2}}\right|+C_{i,N+\frac{1}{2}}\left|\varepsilon_{i-\frac{1}{2},N+\frac{1}{2}}-\varepsilon_{i+\frac{1}{2},N+\frac{1}{2}}\right|\right]$$

$$+h\sum_{j=1}^{N}\left[C_{\frac{1}{2},j}\left|\varepsilon_{\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{\frac{1}{2},j+\frac{1}{2}}\right|+C_{N+\frac{1}{2},j}\left|\varepsilon_{N+\frac{1}{2},j-\frac{1}{2}}-\varepsilon_{N+\frac{1}{2},j+\frac{1}{2}}\right|\right]$$

where we have set, for $1\leq i\leq N$ and $0\leq j\leq N-1$:

$$a_{i,j}=\max\left\{R_{i,j}\ ,\ R_{i-\frac{1}{2},j+\frac{1}{2}}\right\}$$

with

$$R_{i,j}=\max_e\left|R_{i,j}^e\right|,\ R_{i-\frac{1}{2},j+\frac{1}{2}}=\max_e\left|R_{i-\frac{1}{2},j+\frac{1}{2}}^e\right|$$

for $1\leq i\leq N-1$ and $1\leq j\leq N$:

$$b_{i,j}=\max\left\{R_{i,j}\ ,\ R_{i+\frac{1}{2},j-\frac{1}{2}}\right\}$$

with

$$R_{i,j}=\max_e\left|R_{i,j}^e\right|,\ R_{i+\frac{1}{2},j-\frac{1}{2}}=\max_e\left|R_{i+\frac{1}{2},j-\frac{1}{2}}^e\right|$$

for $1 \leq i \leq N$ and $k \in \{0, N\}$:

$$C_{i-\frac{1}{2},k+\frac{1}{2}} = \max_e \left| R^e_{i-\frac{1}{2},k+\frac{1}{2}} \right|$$

for $1 \leq j \leq N$ and $k \in \{0, N\}$:

$$C_{k+\frac{1}{2},j-\frac{1}{2}} = \max_e \left| R^e_{k+\frac{1}{2},j-\frac{1}{2}} \right|$$

.

By applications of Cauchy-Schwarz inequality we have:

$$\begin{aligned} \gamma\left(|\varepsilon_h|^*_{1,h}\right)^2 \leq 2h \Bigg[& \sum_{1\leq i\leq N,\ 1\leq j\leq N-1} a^2_{i,j} + \sum_{1\leq i\leq N-1,\ 1\leq j\leq N} b^2_{i,j} \\ & + \sum_{i=1}^{N}\left(C^2_{i,\frac{1}{2}} + C^2_{i,N+\frac{1}{2}}\right) + \sum_{j=1}^{N}\left(C^2_{\frac{1}{2},j} + C^2_{N+\frac{1}{2},j}\right)\Bigg]^{\frac{1}{2}} |\varepsilon_h|^*_{1,\mathcal{L}} \end{aligned} \tag{3.26}$$

Therefore, we deduce thanks to (3.21) that if $\varphi \in C^2\left(\overline{K}\right)$ for any grid-block K, we have

$$|\varepsilon_h|^*_{1,\mathcal{L}} \leq Ch \tag{3.27}$$

where C is a positive real number depending exclusively on φ, Ω and γ.
Thanks to (3.4), we have

$$||\varepsilon_h||_{L^2(\Omega)} \leq \sqrt{C}\, h$$

□

An L^2-Error estimate can be derived for the piecewise linear approximate solution to the model problem (1.1)-(1.2).

Proposition 3.8. *(L^2-Error estimate for piecewise linear solution $u_{\mathcal{T}}$)*
Under the same assumptions as those of Theorem 3.6, the difference between the exact solution φ and the piecewise linear approximate solution $u_{\mathcal{T}}$ satisfies to the following estimate:

$$||\varphi - u_{\mathcal{T}}||_{L^2(\Omega)} \leq Ch \tag{3.28}$$

where C represents a positive constant which is mesh independent. ◇

4 Test simulation

We deal in what follows with a test case of diffusion problems in anisotropic heterogeneous media.

Notations

- **nunkw**: number of unknowns
- **nnmat**: number of nonzero terms in the matrix
- **sumflux** the discrete flux balance, that is: $sumflux = flux0 + flux1 + fluy0 + fluy1$, where $flux0$, $flux1$, $fluy0$ and $fluy1$ are respectively the outward numerical fluxes through the boundaries $x = 0$, $x = 1$, $y = 0$ and $y = 1$ (for instance $flux0$ is an approximation of $\int_{x=0} K\nabla u \cdot n\, ds$) and $sumf = \sum_{K\in\mathcal{T}} |K|\ f(x_K)$ where x_K denotes some point of the control volume K. Note that the residual $sumflux$ is a measure of the global conservativity of the scheme.
- **umin**: value of the minimum of the approximate solution.
- **umax**: value of the maximum of the approximate solution.
- **ener1**, **ener2**, where **ener1** and **ener2** are approximations of the energy following the two expressions

$$E_1 = \int_\Omega K\nabla u \cdot \nabla u dx,\ E_1 = \int_\Gamma K\nabla u \cdot n\ u ds$$

Let us denote by u the exact solution, by $\mathcal{T}$ the mesh and by $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ the piecewise constant approximate solution.

- **erl2**, relative discrete L^2 norm of the error, that is, for instance:

$$erl2 = \left(\frac{\sum\limits_{K\in\mathcal{T}} |K|\,(u(x_K) - u_K)^2}{\sum\limits_{K\in\mathcal{T}} |K|\,u(x_K)^2} \right)^{1/2}$$

- **ergrad** relative L^2 norm of the error on the gradient, if available
- **ratiol2**: for $i \geq 2$,

$$ratiol2(i) = -2\frac{\ln(erl2(i)) - \ln(erl2(i-1))}{\ln(numkw(i)) - \ln(numkw(i-1))}$$

- **ratiograd**: for $i \geq 2$, same formula as above with $ergrad$ instead of $erl2$.
- **erflx0, erflx1, erfly0, erfly1** relative error between $flux0$, $flux1$, $fluy0$, $fluy1$ and the corresponding flux of the exact solution:

$$erflx0 = \left| \frac{flux0 + \int_{x=0} K\nabla u \cdot n\ ds}{\int_{x=0} K\nabla u \cdot n\ ds} \right|$$

- **ocvl2** order of convergence of the method for the L^2 norm of the solution as defined by $errl2$ with respect to the mesh size:

$$ocvl2 = \frac{\ln(erl2(i\max)) - \ln(erl2(i\max-1))}{\ln(h(i\max)) - \ln(h(i\max-1))}$$

where h is the maximum of the diameter of the control volume.

- **ocvenerdisc** order of convergence of the method for the norm $\|.\|_{E(\mathcal{L})}$ defined by (3.1).

- **ocvgradl2** order of convergence of the method in the L^2 norm of the gradient as defined by $ergradl2$ with respect to the mesh size, same formula as above with ergrad instead of $erl2$.

Test problem

We consider a diffusion problem formulated as (1.1)-(1.2), where the permeability tensor is defined as:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 10^5 \end{pmatrix}.$$

Consider

$$f = (2\pi)^2(10^5 - 10^{-5})\sin(2\pi y)\exp(-2\pi x 10^{-5/2})$$

and

$$g = \begin{cases} -2\pi 10^5 \cos(2\pi y)\exp(-2\pi x 10^{-5/2}) & if \quad 0 \le x \le 1 \quad and \quad y = 1 \\ -2\pi 10^{-5/2} \sin(2\pi y)\exp(-2\pi x 10^{-5/2}) & if \quad x = 0 \quad and \quad 0 \le y \le 1 \\ 2\pi 10^{-5/2} \sin(2\pi y)\exp(-2\pi x 10^{-5/2}) & if \quad x = 1 \quad and \quad 0 \le y \le 1 \\ 2\pi 10^5 \cos(2\pi y)\exp(-2\pi x 10^{-5/2}) & if \quad 0 \le x \le 1 \quad and \quad y = 0 \end{cases} \tag{4.1}$$

The exact solution

$$\varphi(x,y) = \sin(2\pi y)e^{-2\pi x\sqrt{1/10^5}} \quad \text{on} \quad \Omega =]0,1[\times]0,1[.$$

of the problem (1.1)-(1.2) satisfies the following null average condition

$$\int_\Omega \varphi(x)\,dx = 0$$

and the following compatibility condition is checked

$$\int_\Omega f(x)dx - \int_\Gamma g(x)d\gamma(x) = 0.$$

Let us consider the following square meshes (see for instance the figure below) The approximate solution of the problem (1.1)-(1.2) satisfies the following discrete null average conditions:

$$\sum_{P\in\mathcal{P}} U_P + 2\sum_{D\in\mathcal{D}^{ext}} U_D + 4\sum_{D\in\mathcal{D}^{int}} U_D + 2\sum_{i=1}^{4} U_{Di} = 0$$

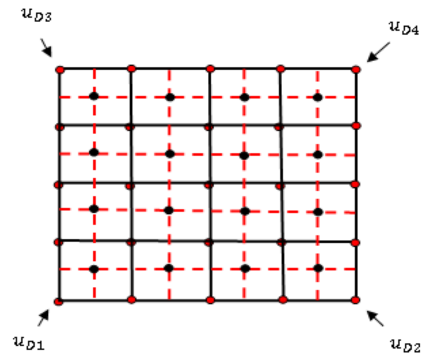

Figure 5. A primary mesh (full black lines) and the dual mesh (dotted red lines), including cell-points and vertices respectively in black and red colors.

and

$$\sum_{D\in\mathcal{D}^{int}} U_D + \frac{1}{2}\sum_{D\in\mathcal{D}^{ext}} U_D + \frac{1}{3}(U_{D2}+U_{D3}) + \frac{1}{6}(U_{D1}+U_{D4}) = 0$$

where $\mathcal{D}^{int}$ is the subset of $\mathcal{D}$ made up cells strictly included into Ω and $\mathcal{D}^{ext}$ is the subset of $\mathcal{D}$ made up cells having only an adjacent edge with the boundary of $\overline{\Omega}$.

Comments about numerical simulations:

Numerical simulations of a diffusion phenomenon governed by full Neumann boundary conditions together with the permeability tensor defined above have been performed.
For that purpose we have utilized the Discrete Duality Finite Volumes exposed and analyzed in this work. The different quantities arising from computations confirm our theoretical results, namely error estimates given in Theorem 3.6 and Proposition 3.8 (see Table 1 - (d) below).
Indeed, according to these results a convergence of order one at least was expected for both L^2-norm and the discrete energy norm $||.||_{E(\mathcal{L})}$. We have obtained much better in terms of order of convergence: a quadratic convergence for L^2-norm and a quasi-quadratic convergence for the discrete energy norm $||.||_{E(\mathcal{L})}$.
Note that there is no contradictions between the theoretical results and the numerical computations obtained here. In fact the theoretical results are established for a permeability tensor with piecewise constant coefficients. In this case study the physical medium is taken to be homogeneous.
Let us emphasize the fact the Discrete Duality Finite Volume method under consideration here has shown a strong capability for honoring the mass balance law as depicted in Table 1: (a), (b) and (c) below. Since our Finite Volume Scheme is conservative, the results from Tables 4.3: (a), (b) and (c) were expected.

i	flux0	flux1	Flux0	fluy1	ener1	ener2	eren	enerdisc
1	6.22E00	-6.22E00	0E00	0E00	4.15E01	1.12E01	7.30E-01	2.99E00
2	6.22E00	-6.22E00	-5.82E-11	-5.82E-11	3.09E01	2.38E01	2.31E-01	3.23E-01
3	6.22E00	-6.22E00	0E00	0E00	3.62E01	3.37E01	6.77E-02	2.23E-01
4	6.22E00	-6.22E00	0E00	0E00	3.80E01	3.73E01	1.81E-02	6.04E-02
5	6.22E00	-6.22E00	-2.18E-11	0E00	3.85E01	3.83E01	4.67E-03	1.71E-02

(a)

ocvl2	ocvgradl2	ocvenerdisc
2.00E00	1.04E00	1.82E00

(b)

i	erflux0	erflux1	erflux0	erfluy1	Erflm	umin	umax
1	1.43E-16	1.43E-16	1E00	1E00	4.14E-01	-1.79E00	1.79E00
2	1.43E-16	0E00	5.11E02	1.02E03	2.37E-02	-1.04E00	1.04E00
3	1.43E-16	1.43E-16	1E00	1E00	5.66E-03	-1.01E00	1.01E00
4	-0E00	0E00	1E00	1E00	4.72E-03	-1E00	1E00
5	-0E00	0E00	1.26E02	1E00	4.24E-03	-1E00	1E00

(c)

i	numkw	nnmat	sumflux	erl2	ergrad	Ratiol2	ratiograd
1	21	68	0E00	4.7E-01	1.62E-01	0E00	0E00
2	57	236	-1.16E-10	1.06E-01	7.69E-02	2.99E00	6.1E00
3	177	860	0E00	2.59E-02	2.48E-03	2.48E00	2E00
4	609	3260	0E00	6.44E-03	1.97E-03	2.25E00	3.72E-01
5	2241	12668	-2.18E-11	1.96E-03	1.96E-03	2.13E00	3.78E-03

(d)

Table 1. Diverse numerical results for the test problem:
(a) is devoted to flux and energy computations combined with error on energy computations;
(b) is dealing with error on flux computations and min-max principle validation;
(c) is concerned with numerical flux balance;
(d) is giving orders of convergence for L^2-norm and for the discrete energy norm $||.||_{E(\mathcal{L})}$.

References

[1] I. Aavatsmark.(2002), An introduction to multipoint flux approximations for quadrilateral grids, Comput. Geosci.6 (2002), pp. 405−432.

[2] Aavatsmark, I.: Multi-point flux approximation methods for quadrilateral grids. In: 9th International Forum on Reservoir Simulation, Abu Dhabi (2007)

[3] Andreianov, B., Boyer, F., Hubert, F. (2007), Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes. Numer. Methods Partial Differ. Equations 23, No 1, 145−195.

[4] Boyer, F., Hubert, F.: Finite volume method for 2D linear and non-linear elliptic problems with discontinuities. SIAM J. Numer. Anal. 46(6), 3032−3070 (2008)

[5] Chen, Q.Y., Wan, J., Yang, Y., Mifflin, R.T.: Enriched multipoint flux approximation for general grids. J. Comput. Phys. 227(3), 1701−1721(2008)

[6] Domelevo, K., Omnes, P.: A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids. ESAIM Math. Model. Numer. Anal. 39(6), 1203−1249 (2005)

[7] Durlofsky, L. J. (2005), Upscaling and Gridding of Geologically Complex Systems, Department of Petroleum Engineering, Stanford University Chevron Texaco E & P Technology Company.

[8] M.G. Edwards and C.F. Rogers (1994), A flux continuous scheme for the full tensor pressure equation, Proceedings of the Fourth European Conference on the Mathematics of Oil Recovery, Roros, Norway, 7−10 June 1994.

[9] R. Eymard, T. Gallouet and R. Herbin, Finite volume methods, Editors: P.G. Ciarlet, J-L. Lions. (2000), Handbook of Numerical Analysis, VII, Amsterdam (North-Holland), 2000, PP 713−1020.

[10] Eymard, R., Gallouet, T., Herbin, R.: Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes. SUSHI: a scheme using stabilisation and hybrid interfaces. IMA J. Numer. Anal. 30(4), 1009−1043 (2010)

[11] Eymard, R., Gallouet, T., and Herbin, R.: A new finite volume scheme for anisotropic diffusion problems on general grids: convergence analysis. C.R. Math., Acad. Sci. Paris 344(6), 40340 (2007)

[12] H Brezis. (1983), Analyse fonctionnelle , theorie et applications, Ed. Masson

[13] F. Hermeline (2000), A finite volume method for the approximation of diffusion operators on distorted meshes. J. Comput. Phys., 160 (2000), No 2, 481−499.

[14] F. Hermeline (2003), Approximation of diffusion operators with discontinuous tensor coefficients on distorted meshes. Comput. Methods Appl. Mech. Eng. 192, No 16−18, 1939−1959.

[15] Klausen, R.A., Radu, F.A., Eigestad, G.T.: Convergence of MPFA on triangulations and for Richards equation. Int. J. Numer. Methods Fluids 58(12), 1327−1351 (2008)

[16] D. Y. Le Roux, Spurious inertial oscillations in shallow-water models", Journal of Computational Physics, 2012, 231, pp. 7959−7987.

[17] I. Nguena Moukouop and A. Njifenjou (2005), A new finite volume formulation for diffusion problems in anisotropic non-homogeneous media, Proceedings of the 4th International Symposium on Finite Volume for Complex Applications, Marraketch (Morocco), July 4−8, 2005, pp. 435−446 (Hermes edition)

[18] A. Njifenjou, I. Moukouop-Nguena (2001), "Traitement des anisotropies de perméabilité en simulation d'écoulement en milieu poreux par les volumes finis", Proceedings of an International Conference on "Systèemes Informatiques pour la Gestion de l'Environnement", (M. Tchuente, Ed.), Douala (Cameroon), pp. 245−259.

[19] A. Njifenjou, A. Kinfack Jeutsa (2008): Convergence analysis of an MPFA method for flow problems in anisotropic heterogeneous porous media. Int. J. Finite Vol. 5(1), 17−56 (published online)

[20] A. Njifenjou H. Donfack I. Moukouop-Nguena (2013), Analysis on general meshes of a discrete duality finite volume method for subsurface flow problems. Comput Geosci (2013) 17:391−415

[21] Omnes, P.: Error estimates for a finite volume method for the Laplace equation in dimension one through discrete Green functions. Int. J. Finite Vol. 6(1), 24−41 (2009) (online publication)

[22] Omnes, P.: On the second-order convergence of a function reconstructed from finite volume approximations of the Laplace equation on Delaunay-Voronoi meshes. ESAIM: Mathematical Modelling and Numerical Analysis 45(N. 04), 627−650 (2011)

[23] Renard, P. (2005), Averaging methods for permeability fields, Development, Protection, Management and Sequestration of Subsurface. Fluid Flow and Transport in Porous and Fractured Media, Summer School, Cargèse (France) 2005, july 27

[24] V. Rostand, D. Y. Le Roux, G. F. Carey, Kernel analysis of the discretized finite difference and finite element shallow-water models", Siam J. SCI. Comput., 31, 2008, pp. 531−556.

[25] Rogers, M.G., Rogers, C.F.: A flux continuous scheme for the full tensor pressure equation. In: Proceedings of the Fourth ECMOR, Roros, Norway (1994)