



**HAL**  
open science

## Comparison of DDFV and DG methods for flow in anisotropic heterogeneous porous media

Vincent Baron, Yves Coudière, Pierre Sochala

► **To cite this version:**

Vincent Baron, Yves Coudière, Pierre Sochala. Comparison of DDFV and DG methods for flow in anisotropic heterogeneous porous media. *Oil & Gas Science and Technology - Revue d'IFP Energies nouvelles*, 2014, 69 (4), pp.673-686. 10.2516/ogst/2013157 . hal-00829898

**HAL Id: hal-00829898**

**<https://hal.science/hal-00829898>**

Submitted on 4 Jun 2013

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

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

# Comparison of DDFV and DG methods for flow in anisotropic heterogeneous porous media

V. Baron<sup>1,3\*</sup>, Y. Coudière<sup>2</sup> and P. Sochala<sup>3</sup>

<sup>1</sup> Laboratoire de mathématiques Jean Leray, 2 rue de la Houssinière, 44322 Nantes, France

<sup>2</sup> IMB, 351 cours de la libération, 33405 Talence cedex, France

<sup>3</sup> BRGM, 3 avenue Claude Guillemin, 45060 Orléans, France

e-mail: [vincent.baron@univ-nantes.fr](mailto:vincent.baron@univ-nantes.fr) — [yves.coudiere@inria.fr](mailto:yves.coudiere@inria.fr) — [p.sochala@brgm.fr](mailto:p.sochala@brgm.fr)

\*Corresponding author

**Résumé** — *Comparaison des méthodes DDFV et DG pour des écoulements en milieu poreux hétérogène anisotrope* — Nous présentons un travail préliminaire visant à simuler l'injection de gaz dans les aquifères profonds. Des écoulements instationnaires monophasiques sont considérés. Nous comparons des schémas volumes finis en dualité discrète (DDFV) et de Galerkin discontinu (DG) utilisés pour discrétiser le terme diffusif. Une formule de différentiation rétrograde d'ordre deux est utilisée pour la discrétisation en temps. D'une part, les schémas DDFV sont simples à implémenter, préservent les propriétés physiques et présentent souvent une super convergence en norme  $L^2$ . D'autre part, les méthodes DG sont flexibles, permettent des ordres de convergence arbitrairement élevés et leur fondement théorique en font un choix adapté à de nombreux problèmes. La méthode de Galerkin à pénalisation intérieure pondérée symétrique est choisie dans ce travail. La précision et la robustesse de ces deux schémas sont testées et comparées sur des cas tests variés, notamment en milieu hétérogène anisotrope.

**Abstract** — *Comparison of DDFV and DG methods for flow in anisotropic heterogeneous porous media* — We present a preliminary work to simulate gas injection in deep aquifers. Unsteady single-phase flows are considered. We compare Discrete Duality Finite Volume (DDFV) and Discontinuous Galerkin (DG) schemes applied to discretize the diffusive term. The second-order Backward Differentiation Formula is used for the time-stepping method. On the one hand, the DDFV methods are easy to implement, ensure a preservation of physical properties and offer superconvergence in the  $L^2$ -norm on a regular basis. On the other hand, the DG methods are flexible, allow arbitrary order of accuracy, and their ample theoretical foundation make them a reliable choice for many computational problems. We consider here the Symmetric Weighted Interior Penalty Galerkin method. Accuracy and robustness of these two schemes are tested and compared on various test cases, especially in anisotropic heterogeneous media.

## 1 INTRODUCTION

Prediction of fluid flows in geological subsurface is a key process in many applications like groundwater hydrology, oil recovery, and more recently CO<sub>2</sub> geological storage and deep geothermal science. Several works are devoted to design numerical schemes (listed below) for simulating these types of flows. This article is a comparison of two conservative methods which are accurate (*i.e.* having a second-order convergence), robust in anisotropic heterogeneous media (*i.e.* with a diffusivity tensor inducing a space-dependent preferential direction for the flow) and valid for general meshes (*i.e.* not only for K-orthogonal meshes). We present a preliminary work dedicated to single-phase flows. An extension of this study is to consider two phases (gas and water) in the frame of CO<sub>2</sub> injection in deep aquifers.

A various choice of methods is available to approximate nonlinear diffusion equations. For instance to discretize the Richards equation, Celia et al. (1990) consider finite element (FE), Narasimhan and Witherspoon (1976) apply finite volume (FV) with two-point flux approximation, Manzini and Ferraris (2004) use diamond cell FV, and Knabner and Schneid (2002) employ mixed finite element (MFE). Several other FV methods and variants exist, including the MPFA, SUSHI and VAG schemes, see Eymard et al. (2012) and references therein for details. In this work, we propose an alternative approach by applying and comparing DDFV and DG methods. The DDFV methodology was introduced about ten years ago by Hermeline (2000). It has been applied to a variety of problems over the past few years, from isotropic scalar diffusion with Domelevo and Omnes (2005) to electrocardiology, see Coudière et al. (2009), among others. This FV method is a natural choice because it has proven robustness when applied to anisotropic and heterogeneous diffusion problems on general, most notably non-matching and distorted meshes. It is locally conservative and preserves the symmetry of the continuous problem, which allows efficient iterative solvers to be employed. Besides it offers a  $L^2$ -norm superconvergence on various examples, together with a very accurate approximation of the gradient, as was noted in Herbin et al. (2008). In this paper we adapt a version of the DDFV scheme described in Coudière et al. (2009), Krell (2010) which allows discontinuities of the permeability tensor along faces. DG methods have appeared thirty years ago with Lesaint and Raviart (1974) and are efficient to approximate equations present in engineering sciences. Advantages of these methods are multiple especially the local conservation (as with FV and MFE). The order of accuracy (as FE and MFE) can be different for each element of the mesh, facilitating the  $p$ -refinement. The flexibility in the use of non-matching meshes (as FV) is particularly adapted for the  $h$ -refinement. Several DG methods can be used for one-phase and two-phase flows in porous media, such as the local DG method described in Bastian et al. (2007), Fagherazzi et al. (2004) and the non-symmetric or the sym-

metric interior penalty Galerkin method used in Bastian (2003), Klieber and Rivière (2006), Sochala et al. (2009). Here, we choose the symmetric weighted interior penalty DG method (SWIP), because it preserves the natural symmetry in the discrete diffusion operator and uses diffusion-dependent weighted averages to reduce the amount of stabilization required, Di Pietro and Ern (2012). Regarding time discretization, we propose to use the second-order backward differentiation formula (BDF2) from Curtiss and Hirschfelder (1952). This implicit scheme avoids the CFL condition particularly restrictive for explicit schemes when the diffusion term is present. Furthermore, BDF2 is more efficient than implicit Runge-Kutta (DIRK) schemes, which require several stages (three for the second-order version). Besides, the last stage of DIRK schemes is ill-posed where the soil is saturated, as was underlined in Sochala et al. (2009). Finally, BDF2 is preferred to Crank-Nicolson which is not L-stable, Hairer and E. (2010).

The aim of this work is to test and compare the DDFV and SWIP schemes for solving flows in anisotropic heterogeneous porous media. We propose to consider the conservative form of the Richards equation in a two-dimensional domain  $\Omega$ ,

$$\begin{cases} \partial_t \theta(\psi) - \nabla \cdot (K(\psi)(\nabla \psi + e_z)) = f & \text{in } \Omega \times ]0, T], \\ \psi = \psi^0 & \text{in } \Omega \times \{0\}, \\ \psi = \psi_D & \text{on } \partial\Omega^D \times ]0, T], \\ -K(\psi)(\nabla \psi + e_z) \cdot n_\Omega = v_N & \text{on } \partial\Omega^N \times ]0, T], \end{cases} \quad (1)$$

where  $T$  is the final time of simulation,  $\partial\Omega^D$  (resp.  $\partial\Omega^N$ ) the boundary of  $\Omega$  where a Dirichlet condition  $\psi_D$  (resp. Neumann condition  $v_N$ ) is applied, and  $n_\Omega$  the outward unit normal of  $\partial\Omega = \partial\Omega^D \cup \partial\Omega^N$ . The vector  $e_z = (0, 1)^t$  is the gravity contribution. The pressure head  $\psi$  is the unknown and is related to the velocity  $v$  through Darcy's law  $v(\psi) = -K(\psi)(\nabla \psi + e_z)$ . Two constitutive relationships  $\psi \mapsto \theta(\psi)$  and  $\psi \mapsto K(\psi)$  are necessary to close the model. Examples of water content  $\theta$  and conductivity  $K$  are indicated in subsections 4.1 and 4.2. For the sake of simplicity, we suppose that the source term  $f$  is equal to zero (except for the validation test case described in subsection 4.1). We investigate three test cases (TC) to compare DDFV and SWIP methods in several configurations. TC1 and TC2 are column infiltration problems in an isotropic homogeneous soil. TC1 has an analytical solution and permits to validate the convergence rate of the methods. The stencil of the DDFV method produces a sparser matrix than SWIP. In return, the latter leads to a narrower bandwidth and smaller condition numbers. TC2 features a stiff pressure front which highlights significant differences for coarse meshes. DDFV underestimates the front propagation speed while SWIP presents non-physical oscillations. TC3 is a quarter five-spot configuration in an anisotropic heterogeneous soil. The two methods only differ in the evaluation of the mass error, depending on the choice of the nonlinear solver tolerance.

The outline is as follows. Section 2 introduces the space discretization associated with DDFV and SWIP methods. Section 3 describes the time discretization performed by the BDF2 formula and the Picard fixed point algorithm used to solve the nonlinear discrete system. Section 4 details the results on the three TC described above. Conclusions and perspectives are drawn in Section 5.

## 2 SPACE DISCRETIZATION

### 2.1 Notations

Let  $\{\mathcal{T}_h\}_{h>0}$  be a shape-regular family of unstructured meshes of  $\Omega$  consisting for simplicity of affine triangles. For an element  $\tau \in \mathcal{T}_h$ , let  $\partial\tau$  denote its boundary and  $n_\tau$  its outward unit normal. The set  $\mathcal{F}_h$  of mesh faces is partitioned into  $\mathcal{F}_h^i \cup \mathcal{F}_h^D \cup \mathcal{F}_h^N$ , where  $\mathcal{F}_h^i$  is the set of internal faces and  $\mathcal{F}_h^D$  (resp.  $\mathcal{F}_h^N$ ) is the set of Dirichlet (resp. Neumann) faces. For a face  $\sigma \in \mathcal{F}_h^i$ , there exist  $\tau^+$  and  $\tau^-$  in  $\mathcal{T}_h$  such that  $\sigma = \partial\tau^+ \cap \partial\tau^-$ . We denote by  $F_\tau \subset \mathcal{F}_h$  the set of the faces  $\sigma$  such that  $\partial\tau = \cup_{\sigma \in F_\tau} \sigma$ , and by  $F_s \subset \mathcal{F}_h$  the set of the faces  $\sigma$  that share  $s$  as a vertex. We define  $n_\sigma$  as the unit normal to  $\sigma$  pointing from  $\tau^-$  to  $\tau^+$  (see Fig 1).

The set  $\mathcal{S}_h$  of mesh vertices is partitioned into  $\mathcal{S}_h^i \cup \mathcal{S}_h^D \cup \mathcal{S}_h^N$  where  $\mathcal{S}_h^i$  are the interior vertices,  $\mathcal{S}_h^N$  gathers all the endpoints of the edges in  $\mathcal{F}_h^N$  and  $\mathcal{S}_h^D$  are the remaining boundary vertices. A secondary mesh is associated to this set  $\mathcal{S}_h$  as follows: any vertex  $s \in \mathcal{S}_h^i \cup \mathcal{S}_h^N$  is associated to a unique polygonal cell  $S$  whose vertices are the centers of the triangles  $\tau$  and the midpoints of the faces  $\sigma$  that share  $s$  as a vertex; and which contains  $s$  (see Fig 2). Consequently, each face  $\sigma$  is associated to its two neighbouring cells  $\tau^+$  and  $\tau^-$  and its two vertices, denoted by  $s^+$  and  $s^-$ . The following orientation assumption is made for these notations:

$$\det(x_{\tau^+} - x_{\tau^-}, x_{s^+} - x_{s^-}) > 0,$$

although the complete orientation is arbitrary and irrelevant in the sequel. Here,  $x_\tau$  is the center of gravity of  $\tau \in \mathcal{T}_h$  and  $x_s$  are the coordinates of  $s \in \mathcal{S}_h$ . We also denote by  $x_\sigma$  the midpoints of the faces  $\sigma \in \mathcal{F}_h$  (see Fig 3). Note that a face  $\sigma \subset \partial\Omega$  has only one neighbouring triangle which is denoted by  $\tau^-$  in the sequel, so that the normal  $n_\sigma$  is also the unit normal to  $\partial\Omega$  outward of  $\Omega$ .

The DDFV method couples, through the flux computations, two finite volume schemes written on a primary mesh  $\mathcal{T}_h$  (considering the cells  $\tau$  as control volumes) and on a secondary mesh  $\mathcal{S}_h$  (considering the cells  $S$  as control volumes). Hence the DDFV discrete unknown is a set of degrees of freedom  $\Psi_h$ , that eventually defines a function  $\psi_h$  (section 2.2). The DG method directly defines a discrete unknown function  $\psi_h$  (section 2.3) and is based on a discrete variational formulation.

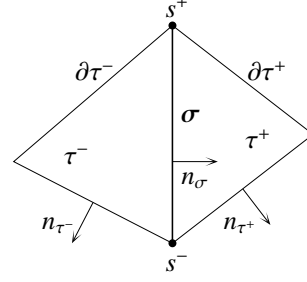


Figure 1: Basic notations.

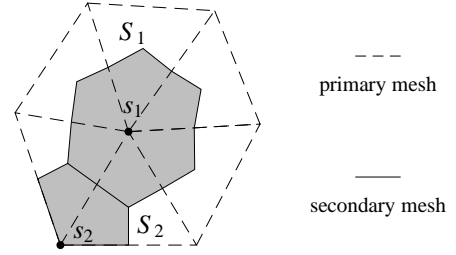


Figure 2: Cells (in gray) of the secondary mesh used in the DDFV method.

### 2.2 Discrete Duality Finite Volume method

The DDFV method approximates the average of  $\psi$  on each control volume  $\tau \in \mathcal{T}_h$  and on each control volume  $S$  associated to a vertex  $s \in \mathcal{S}_h^i \cup \mathcal{S}_h^N$ . Hence we define the discrete DDFV unknown as

$$Q_h \stackrel{\text{def}}{=} \{\Psi_h = (\psi_\tau, \psi_s)_{\tau \in \mathcal{T}_h, s \in \mathcal{S}_h^i \cup \mathcal{S}_h^N}\}.$$

Several functions can be associated to these degrees of freedom, for instance the usual piecewise constant function on each  $\tau \in \mathcal{T}_h$  (or its sibling piecewise constant on each  $S$  for  $s \in \mathcal{S}_h$ ), but also the function  $\psi_h$  that is piecewise linear on the triangles with base  $\sigma$  and vertex  $x_\tau$  (for all  $\tau \in \mathcal{T}_h$  and  $\sigma \in F_\tau$ ) and such that  $\psi_h(x_\tau) = \psi_\tau$  and  $\psi_h(x_s) = \psi_s$ .

If we integrate equation (1) on a control volume  $\tau \in \mathcal{T}_h$  we get :

$$\frac{d}{dt} \int_\tau \theta(\psi) - \sum_{\sigma \in F_\tau^*} \int_\sigma K(\psi)(\nabla\psi + e_z) \cdot n_\tau - \sum_{\sigma \in F_\tau^N} \int_\sigma v_N = 0$$

where  $F_\tau^N = F_\tau \cap \mathcal{F}_h^N$  and  $F_\tau^* = F_\tau \setminus F_\tau^N$ . Note that the flux  $-K(\psi)(\nabla\psi + e_z) \cdot n_\tau$  is continuous through each face  $\sigma$  of  $\tau$ . The DDFV scheme reads, for any control volume  $\tau \in \mathcal{T}_h$ ,

$$\frac{d}{dt} |\tau| \theta(\psi_\tau) - \sum_{\sigma \in F_\tau^*} K(\psi_{\sigma,\tau}) (\nabla_{\sigma,\tau} \Psi_h + e_z) \cdot N_\sigma - \sum_{\sigma \in F_\tau^N} V_\sigma = 0, \quad (2)$$

where  $\psi_{\sigma,\tau}$  (resp.  $\nabla_{\sigma,\tau} \Psi_h$ ) approximates  $\psi$  (resp.  $\nabla\psi$ ) on the side  $\tau$  of the face  $\sigma$ , the normal  $N_\sigma = |\sigma|n_\sigma$  is the normal

to  $\sigma$  outward of  $\tau$  accounting for the length of  $\sigma$  and  $V_\sigma$  approximates  $\int_\sigma v_N$ . Note that in the TC presented in section 4, the boundary data is piecewise constant and we can take exactly  $V_\sigma = \int_\sigma v_N$ . There are several possible choices for the value  $\psi_{\sigma,\tau}$ . We take here

$$\psi_{\sigma,\tau} = \frac{1}{3} (\psi_\tau + \psi_{s^-} + \psi_{s^+}),$$

because it seems to provide the best accuracy and robustness. For each vertex  $s \in \mathcal{S}_h^D$ , we choose to take  $\psi_s = \psi_D(x_s)$  from the Dirichlet boundary data. For any control volume  $S$  associated to a vertex  $s \in \mathcal{S}_h^i \cup \mathcal{S}_h^N$ , the discrete equation is

$$\begin{aligned} \frac{d}{dt} |S| \theta(\psi_s) - \sum_{\sigma \in F_s^*} \left( K(\psi_{\sigma,\tau}) (\nabla_{\sigma,\tau} \Psi_h + e_z) \cdot N_{e^-} \right. \\ \left. + K(\psi_{\sigma,\tau^+}) (\nabla_{\sigma,\tau^+} \Psi_h + e_z) \cdot N_{e^+} \right) \\ - \sum_{\sigma \in F_s^N} \left( K(\psi_{\sigma,\tau}) (\nabla_{\sigma,\tau} \Psi_h + e_z) \cdot N_e + V_{\sigma s} \right) = 0, \quad (3) \end{aligned}$$

where  $F_s^N = F_s \cap \mathcal{F}_h^N$  and  $F_s^* = F_s \setminus F_s^N$ . In these summations, if  $\sigma \in F_s^*$  then the vertex  $s$  is identified to  $s^-$ , the neighbouring triangles are  $\tau^\pm$  and the interface between the control volumes associated to  $s^-$  and  $s^+$  has two parts  $e^+$  and  $e^-$ , of which the corresponding normals are  $N_{e^\pm} = n_{e^\pm} |e^\pm|$  as depicted in Fig. 3. If  $\sigma \in F_s^N$ , then there is only one triangle  $\tau$  neighbouring  $\sigma$  but the flux still has two terms, one along the edge  $e$  defined as in the previous case (with the normal  $N_e = n_e |e|$ ) and one along the line  $(x_s, x_\sigma) \subset \sigma$ . The number  $V_{\sigma s}$  approximates the integral of the Neumann boundary data along this line. Given a face  $\sigma \in \mathcal{F}_h^i$ , we denote by

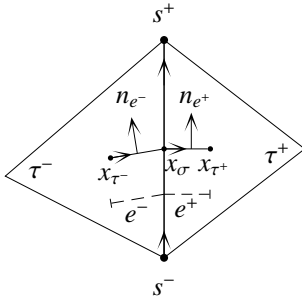


Figure 3: Notations for DDFV method.

$$-F_{\sigma,\tau^\pm} = \pm K(\psi_{\sigma,\tau^\pm}) (\nabla_{\sigma,\tau^\pm} \Psi_h + e_z) \cdot N_\sigma$$

the fluxes out of  $\tau^-$  and  $\tau^+$  through the face  $\sigma$  and by

$$\begin{aligned} F_e = K(\psi_{\sigma,\tau^-}) (\nabla_{\sigma,\tau^-} \Psi_h + e_z) \cdot N_{e^-} \\ + K(\psi_{\sigma,\tau^+}) (\nabla_{\sigma,\tau^+} \Psi_h + e_z) \cdot N_{e^+} \end{aligned}$$

the total flux from  $s^-$  to  $s^+$  through the interface  $S^- \cap S^+$ . Given a face  $\sigma \in \mathcal{F}_h^N \cap F_\tau$  and an endpoint  $s$  of  $\sigma$ ,

$$F_e = K(\psi_{\sigma,\tau}) (\nabla_{\sigma,\tau} \Psi_h + e_z) \cdot N_e$$

is the flux from  $s$  to the opposite endpoint of  $\sigma$ . The flux  $F_e$  is continuous through the interface between  $S^-$  and  $S^+$  by construction. We ensure the continuity of the fluxes  $F_{\sigma,\tau^\pm}$  through  $\sigma$  by introducing an auxiliary unknown  $\psi_\sigma$  used to compute the gradients  $\nabla_{\sigma,\tau^-} \Psi_h$  and  $\nabla_{\sigma,\tau^+} \Psi_h$ , following the DDFV strategy from Coudière et al. (2009) :

$$\begin{aligned} \nabla_{\sigma,\tau^-} \Psi_h &= \frac{1}{2D_{\sigma,\tau^-}} \left( (\psi_\sigma - \psi_{\tau^-}) N_\sigma + (\psi_{s^+} - \psi_{s^-}) N_{e^-} \right) \\ \nabla_{\sigma,\tau^+} \Psi_h &= \frac{1}{2D_{\sigma,\tau^+}} \left( (\psi_{\tau^+} - \psi_\sigma) N_\sigma + (\psi_{s^+} - \psi_{s^-}) N_{e^+} \right), \end{aligned}$$

where  $2D_{\sigma,\tau^\pm} = \pm \det(x_{\tau^\pm} - x_\sigma, x_{s^+} - x_{s^-})$ , meaning that  $D_{\sigma,\tau^\pm}$  are the areas of the triangles with base  $\sigma$  and vertices  $x_{\tau^\pm}$ . In the boundary case  $\sigma \in \mathcal{F}_h^N$ , there is only one gradient, namely  $\nabla_{\sigma,\tau^-} \Psi_h$ . The auxiliary unknown is then the unique solution to the linear equation  $F_{\sigma,\tau^-} + F_{\sigma,\tau^+} = 0$  in the general interior case and to the linear equation  $F_{\sigma,\tau^-} = V_\sigma$  in the boundary case. Then the flux between the primary cells  $\tau^\pm$  is  $F_\sigma = F_{\sigma,\tau^-} = -F_{\sigma,\tau^+}$  and the scheme (2)-(3) reads :

$$\begin{aligned} \frac{d}{dt} |\tau| \theta(\psi_\tau) - \sum_{\sigma \in F_\tau^*} F_\sigma - \sum_{\sigma \in F_\tau^N} V_\sigma = 0, \\ \frac{d}{dt} |S| \theta(\psi_s) - \sum_{\sigma \in F_s^*} F_e - \sum_{\sigma \in F_s^N} V_{\sigma s} = 0. \end{aligned}$$

In view of the expression of the fluxes and the gradients, we adopt some additional notations in the general interior case. On each side  $\tau^\pm$  of  $\sigma$ , we define the following Gram matrices :

$$\begin{pmatrix} a_\pm & b_\pm \\ b_\pm & c_\pm \end{pmatrix} \stackrel{\text{def}}{=} \frac{1}{2D_{\sigma,\tau^\pm}} K(\psi_{\sigma,\tau^\pm}) \begin{pmatrix} N_\sigma \\ N_{e^\pm} \end{pmatrix} \begin{pmatrix} N_\sigma & N_{e^\pm} \end{pmatrix},$$

and  $\alpha_\pm, \beta_\pm$  are defined similarly by

$$\begin{pmatrix} \alpha_\pm \\ \beta_\pm \end{pmatrix} \stackrel{\text{def}}{=} K(\psi_{\sigma,\tau^\pm}) e_z \cdot \begin{pmatrix} N_\sigma \\ N_{e^\pm} \end{pmatrix}.$$

Hence the equation in  $\psi_\sigma$ ,  $F_{\sigma,\tau^-} + F_{\sigma,\tau^+} = 0$  now reads  $a_-(\psi_\sigma - \psi_{\tau^-}) + b_-(\psi_{s^+} - \psi_{s^-}) + \alpha_- = a_+(\psi_{\tau^+} - \psi_\sigma) + b_+(\psi_{s^+} - \psi_{s^-}) + \alpha_+$  and can be easily solved. Afterwards, we replace  $\psi_\sigma$  by its expression and find that the calculation of the fluxes  $F_\sigma$  and  $F_e$  can be carried out as follows:

$$\begin{pmatrix} F_\sigma \\ F_e \end{pmatrix} = A_\sigma(\Psi_h) \begin{pmatrix} \psi_{\tau^+} - \psi_{\tau^-} \\ \psi_{s^+} - \psi_{s^-} \end{pmatrix} - b_\sigma(\Psi_h),$$

with

$$A_\sigma(\Psi_h) = \begin{pmatrix} \frac{a_- a_+}{a_- + a_+} & \frac{a_+ b_- + a_- b_+}{a_- + a_+} \\ \frac{a_+ b_- + a_- b_+}{a_- + a_+} & c_+ + c_- - \frac{(b_+ - b_-)^2}{a_- + a_+} \end{pmatrix}$$

$$b_\sigma(\Psi_h) = - \begin{pmatrix} \frac{a_+ \alpha_- + a_- \alpha_+}{a_- + a_+} \\ \beta_- + \beta_+ - \frac{b_+ - b_-}{a_- + a_+} (\alpha_+ - \alpha_-) \end{pmatrix}.$$

In the case of a face  $\sigma \in \mathcal{F}_h^N \cap \tau$ , we assume that the neighbouring triangle  $\tau$  is  $\tau^-$  and the equation on  $\psi_\sigma$  reads  $a_-(\psi_\sigma - \psi_{\tau^-}) + b_-(\psi_{s^+} - \psi_{s^-}) + \alpha_- = V_\sigma$ . It is solved easily and results in the expression

$$F_e = \left( c_- - \frac{b_-^2}{a_-} \right) (\psi_{s^+} - \psi_{s^-}) + \beta_- - \frac{b_-}{a_-} \alpha_- + \frac{b_-}{a_-} V_\sigma.$$

Lastly, if  $\sigma \in \mathcal{F}_h^D$ , the quantity  $F_e$  does not need to be defined, as only  $F_\sigma$  adds a contribution to the system. In this case, the auxiliary  $\psi_\sigma$  is chosen as  $\psi_\sigma = \frac{1}{2}(\psi_{s^+} + \psi_{s^-})$ .

By assembling these local contributions, the semi-discrete DDFV scheme finally reads

$$M_h \frac{d}{dt} \Theta(\Psi_h) + A_h(\Psi_h) \Psi_h = b_h(\Psi_h), \quad (4)$$

where  $\Theta(\Psi_h) = (\theta(\Psi_\tau), \theta(\Psi_s))_{\tau \in \mathcal{T}_h, s \in \mathcal{S}_h^i \cup \mathcal{S}_h^N}$ ,  $M_h$  is a diagonal mass matrix and  $A_h(\Psi_h)$  is a symmetric and positive matrix.

### 2.3 Discontinuous Galerkin method

The Discontinuous Galerkin methods are based on variational formulations (as for the classical finite element). The discontinuous finite element space  $V_h$  is defined as

$$V_h \stackrel{\text{def}}{=} \{ \phi \in L^2(\Omega), \forall \tau \in \mathcal{T}_h, \phi|_\tau \in \mathbb{P}_p(\tau) \},$$

where  $\mathbb{P}_p(\tau)$  is the set of polynomials of degree less than or equal to  $p$  on an element  $\tau$ . We observe that the functions in  $V_h$  are not necessarily continuous. This fact is exploited by selecting basis functions which are locally supported in a single mesh element. As a consequence, when a triangular mesh and a  $\mathbb{P}_1$ -Lagrange approximation are chosen, the unknowns are localized at the vertices of each triangle as illustrated on Fig. 4. A linear approximation yields the same order of convergence as DDFV, though higher-order elements could also be considered. We define the weighted average operator  $\{\cdot\}_{\omega,\sigma}$  and the jump operator  $[\![\cdot]\!]_\sigma$  as follows: for a function  $\xi$  which is possibly two-valued on  $\sigma$ ,

$$\{\xi\}_{\omega,\sigma} \stackrel{\text{def}}{=} \omega^- \xi^- + \omega^+ \xi^+ \quad \text{and} \quad [\![\xi]\!]_\sigma \stackrel{\text{def}}{=} \xi^- - \xi^+,$$

where  $\xi^\pm = \xi|_{\tau^\pm}$ . For boundary faces,  $\{\xi\}_{\omega,\sigma} = \xi|_\sigma$  and  $[\![\xi]\!]_\sigma = \xi|_\sigma$ . For vector-valued functions, average and jump operators are defined componentwise. For the Richards equation, the SWIP method can be written as

$$\frac{d}{dt} \int_\Omega \theta(\psi_h) \phi + a_h(\psi_h, \phi) = b_h(\psi_h, \phi),$$

where for  $(\psi, \phi) \in V_h \times V_h$ ,

$$\begin{aligned} a_h(\psi, \phi) &\stackrel{\text{def}}{=} \int_\Omega K(\psi) \nabla \psi \cdot \nabla \phi + \sum_{\sigma \in \mathcal{F}_h^i \cup \mathcal{F}_h^D} \frac{\eta \gamma_K}{d_\sigma} \int_\sigma [\![\psi]\!]_\sigma [\![\phi]\!]_\sigma \\ &- \sum_{\sigma \in \mathcal{F}_h^i \cup \mathcal{F}_h^D} \int_\sigma ([\![\psi]\!]_\sigma \{K(\psi|_\tau) \nabla \phi\}_{\omega,\sigma} + [\![\phi]\!]_\sigma \{K(\psi|_\tau) \nabla \psi\}_{\omega,\sigma}) \cdot n_\sigma, \end{aligned}$$

$$\begin{aligned} b_h(\psi, \phi) &\stackrel{\text{def}}{=} \int_\Omega \nabla \cdot (K(\psi) e_z) \phi - \sum_{\sigma \in \mathcal{F}_h^N} (v_N + K(\psi|_\tau) e_z \cdot n_\Omega) \phi \\ &+ \sum_{\sigma \in \mathcal{F}_h^D} \int_\sigma \left( -K(\psi|_\tau) \nabla \phi \cdot n_\Omega + \frac{\eta \gamma_K}{d_\sigma} \phi \right) \psi_D. \end{aligned}$$

A specific choice of the weights is made to yield robust error estimates with respect to the diffusivity,

$$\omega^- = \frac{\delta_{Kn}^+}{\delta_{Kn}^- + \delta_{Kn}^+} \quad \text{and} \quad \omega^+ = \frac{\delta_{Kn}^-}{\delta_{Kn}^- + \delta_{Kn}^+},$$

with  $\delta_{Kn}^\pm = K^\pm n_\sigma \cdot n_\sigma$  for  $\sigma \in \mathcal{F}_h^i$ . In the penalty term,  $\eta$  is a positive parameter (to be taken larger than a minimal threshold depending on the shape-regularity of  $\mathcal{T}_h$ ), and  $d_\sigma$  is the diameter of the face  $\sigma$ , *i.e.* the largest diameter of the triangle(s) of which  $\sigma$  is a face. The coefficient  $\gamma_K$  is defined as

$$\gamma_K \stackrel{\text{def}}{=} \begin{cases} \frac{2\delta_{Kn}^- \delta_{Kn}^+}{\delta_{Kn}^- + \delta_{Kn}^+} & \text{if } \sigma \in \mathcal{F}_h^i, \\ \delta_{Kn} & \text{if } \sigma \in \mathcal{F}_h^D, \end{cases}$$

with  $\delta_{Kn} = K n_\sigma \cdot n_\sigma$ . The above choice of  $\omega^-$  and  $\omega^+$  leads to the harmonic mean of the diffusivity tensor,

$$\{K(\psi|_\tau) \nabla \psi\}_{\omega,\sigma} \cdot n_\sigma = \frac{2\delta_{Kn}^- \delta_{Kn}^+}{\delta_{Kn}^- + \delta_{Kn}^+} (\nabla \psi|_{\tau^-} + \nabla \psi|_{\tau^+}).$$

Using harmonic means is particularly adapted at an interface between poorly and highly conductive media. In this case, the harmonic mean value of the diffusivity is close to the value in the poorly conductive medium, Di Pietro and Ern (2012).

Again, by assembling the local contributions, the semi-discrete DG scheme takes the form of (4), where  $M_h$  is a block diagonal matrix,  $A_h$  is a block matrix and  $\Psi_h$  is the set of the values of  $\psi_h$  at each degree of freedom.

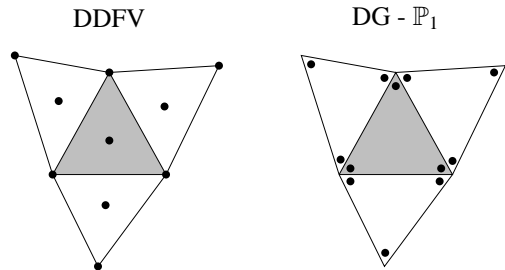


Figure 4: Unknowns localization for each method.

### 3 TIME DISCRETIZATION AND NONLINEAR RESOLUTION

#### 3.1 Backward Differentiation Formula

The next step consists in approximating the time derivative in (4). We propose the second order Backward Differentiation Formula (BDF2) :

$$\frac{d}{dt}y^n \simeq \frac{1}{\delta t} \left( \frac{3}{2}y^n - 2y^{n-1} + \frac{1}{2}y^{n-2} \right),$$

where the superscript  $n$  denotes the time level, and  $\delta t$  is the constant time step, chosen such that  $T/\delta t$  is an integer. Substituting the last expression in (4) finally yields the following fully discrete scheme:

$$\begin{aligned} \forall n \geq 2, \quad & \frac{3}{2\delta t} M_h \Theta(\Psi_h^n) + A_h(\Psi_h^n) \Psi_h^n \\ & = b_h(\Psi_h^n) + \frac{2}{\delta t} M_h \Theta(\Psi_h^{n-1}) - \frac{1}{2\delta t} M_h \Theta(\Psi_h^{n-2}). \end{aligned}$$

The vector  $\Psi_h^1$  is computed by the Crank-Nicolson scheme.

#### 3.2 Linearization algorithm

The above is obviously a nonlinear system, and may thus be solved through an iterative procedure, which we describe in this section. The iteration level will be referred to as the superscript  $m$ . As suggested in Manzini and Ferraris (2004), we use the Picard fixed point method to obtain :

$$\begin{aligned} & \frac{3}{2\delta t} M_h \Theta(\Psi_h^{n,m+1}) + A_h(\Psi_h^{n,m}) \Psi_h^{n,m+1} \\ & = b_h(\Psi_h^{n,m}) - \frac{1}{\delta t} \left( -2M_h \Theta(\Psi_h^{n-1}) + \frac{1}{2} M_h \Theta(\Psi_h^{n-2}) \right). \end{aligned}$$

To complete the linearization, we now use the following Taylor series expansion :

$$\Theta(y) \simeq \Theta(x) + \frac{\partial \Theta}{\partial \Psi}(x) \cdot (y - x),$$

which we apply with  $x = \Psi_h^{n,m}$  and  $y = \Psi_h^{n,m+1}$ . Finally, the system to solve is :

$$\begin{aligned} & \left( \frac{3}{2} M_h \partial_\psi \Theta(\Psi_h^{n,m}) + \delta t A_h(\Psi_h^{n,m}) \right) \delta \Psi_h^{n,m} = \delta t b_h(\Psi_h^{n,m}) \\ & - \delta t A_h(\Psi_h^{n,m}) \Psi_h^{n,m} - M_h \left( \frac{3}{2} \Theta(\Psi_h^{n,m}) - 2\Theta(\Psi_h^{n-1}) + \frac{1}{2} \Theta(\Psi_h^{n-2}) \right), \end{aligned}$$

where  $\delta \Psi_h^{n,m} = \Psi_h^{n,m+1} - \Psi_h^{n,m}$  is the unknown of the system. In this form, the algorithm produces at each time level

$n$ , a sequence  $(\Psi_h^{n,m})_{m \geq 0}$  defined by the recurrence relation  $\Psi_h^{n,m+1} = \Psi_h^{n,m} + \delta \Psi_h^{n,m}$ , together with a second-order initialization :  $\Psi_h^{n,0} = 3\Psi_h^{n-1} - 3\Psi_h^{n-2} + \Psi_h^{n-3}$  for  $n \geq 3$ , (and  $\Psi_h^{1,0} = \Psi_h^0$ ,  $\Psi_h^{2,0} = 2\Psi_h^1 - \Psi_h^0$ ). We use the LU decomposition for solving each linear system and our stopping criterion is  $\|\delta \Psi_h^{n,m}\|_2 \|\Psi_h^{n-1}\|_2^{-1} \leq \epsilon$ , where  $\epsilon$  is the user-defined tolerance.

#### 3.3 Mass conservation properties

The study of the mass evolution and repartition (*i.e.* the separation over time of the mass variation in the domain, the mass inflow and the mass outflow) helps to understand the behavior of the system. It also allows to quantify the mass defect generated by the linearization algorithm. The volume of water in the domain  $\Omega$  at time  $n\delta t$  is obtained by integrating the volumetric water content in  $\Omega$ ,

$$V^n \stackrel{\text{def}}{=} \int_{\Omega} \theta(\psi_h^n).$$

By using the BDF2 formula, we have

$$\frac{3}{2} V^n - 2V^{n-1} + \frac{1}{2} V^{n-2} = (F_D^n + F_N^n) \delta t + \epsilon^n, \quad (5)$$

where the quantities  $F_D^n$  and  $F_N^n$  are defined as

$$F_D^n \stackrel{\text{def}}{=} - \int_{\partial\Omega^D} v_h^{*,n}, \quad \text{and} \quad F_N^n \stackrel{\text{def}}{=} - \int_{\partial\Omega^N} v_h^n,$$

and  $\epsilon^n$  stands for the numerical error in the resolution of the nonlinear system. The reconstructed normal velocity  $v_h^{*,n}$  on a Dirichlet face  $\sigma$  is estimated from the pressure,

$$v_h^{*,n}|_{\sigma} = \begin{cases} -K(\psi_{\sigma,\tau}^n) (\nabla_{\sigma,\tau} \Psi_h^n + e_z) \cdot n_{\sigma} & \text{for DDFV,} \\ v(\psi_h^n|_{\sigma}) \cdot n_{\sigma} + \frac{\eta \gamma_K}{d_{\sigma}} (\psi_h^n|_{\sigma} - \psi_D) & \text{for SWIP.} \end{cases}$$

The variation of the volume of water over the time step  $[(n-1)\delta t, n\delta t]$  is deduced from a reformulation of (5), yielding

$$V^n - V^{n-1} = (\Phi_D^n + \Phi_N^n) \delta t + e^n, \quad (6)$$

where the flux  $\Phi^n \in \{\Phi_D^n, \Phi_N^n\}$  and the error  $e^n$  over the time step  $[(n-1)\delta t, n\delta t]$  are defined as

$$\Phi^n \stackrel{\text{def}}{=} \frac{2}{3} F^n + \frac{1}{3} \Phi^{n-1}, \quad \text{and} \quad e^n \stackrel{\text{def}}{=} \frac{2}{3} \epsilon^n + \frac{1}{3} e^{n-1}.$$

The initialization of the fluxes recursive formulas depends on the scheme used at the first time step. For the Crank-Nicolson scheme, the global mass conservation is written as

$$V^1 - V^0 = (\Phi_D^1 + \Phi_N^1) \delta t + e^1 \quad \text{with} \quad \Phi^1 = \frac{1}{2} (F^0 + F^1). \quad (7)$$

## 4 RESULTS

We propose three TC with increasing difficulties corresponding to a more general form of the conductivity defined as

$$K(\psi, x) = \frac{\rho g}{\nu} k(\psi) \mathbb{K}(x),$$

where  $k(\psi)$  is the relative permeability,  $\mathbb{K}(x)$  the intrinsic tensor permeability of the soil;  $\rho$ ,  $g$  and  $\nu$  are respectively the fluid density, the gravity constant and the dynamic viscosity. We first consider two downward infiltrations in an isotropic homogeneous medium (*i.e.*  $\mathbb{K} = \mathbb{I}$ ). The first TC presents an analytical solution to verify the theoretical convergence rates and the second TC is the propagation of a stiff pressure front. We also study the quarter five-spot problem in an anisotropic heterogeneous medium (*i.e.*  $\mathbb{K} \neq \mathbb{I}$  and is space-dependent).

### 4.1 Isotropic homogeneous validation test case

We propose a TC with an analytical solution to compare the two methods in terms of matrix properties and convergence. The domain is  $\Omega = [0, 4] \times [0, 20]$  (in *cm*) and the final time  $T = 2 \text{min}$ . The analytical solution

$$\psi(z, t) = 20.4 \tanh \left[ 0.5 \left( z + \frac{t}{12} - 15 \right) \right] - 41.1$$

is used to determine adequate source term and Dirichlet boundary condition enforced on  $\partial\Omega$ . The water content and the conductivity are defined by Haverkamp's constitutive relationships which are derived in Haverkamp et al. (1977) :

$$\theta(\psi) = \frac{\theta_s - \theta_r}{1 + |\tilde{\alpha}\psi|^\beta} + \theta_r \quad \text{and} \quad K(\psi) = \frac{K_s}{1 + |\tilde{A}\psi|^\gamma},$$

with parameters

$$\theta_s = 0.287, \quad \theta_r = 0.075, \quad \tilde{\alpha} = 0.0271 \text{ cm}^{-1}, \quad \beta = 3.96, \\ K_s = 9.44 \cdot 10^{-3} \text{ cm.s}^{-1}, \quad \tilde{A} = 0.0524 \text{ cm}^{-1}, \quad \gamma = 4.74.$$

For a mesh, the number of unknowns  $N_u$  is

$$N_u = \begin{cases} N_t + N_n - N_n^D & \text{for DDFV,} \\ 3N_t & \text{for DG} - \mathbb{P}_1, \end{cases}$$

where  $N_t$  is the number of triangles,  $N_n$  the number of nodes and  $N_n^D$  the number of nodes where a Dirichlet condition is enforced (and which is equal to the number of nodes located on  $\partial\Omega$ ). Consequently, the number of unknowns for DDFV is about half the number of unknowns for SWIP when the mesh is sufficiently fine (*i.e.*  $N_n \gg N_n^D$ ) :

$N_t + N_n - N_n^D \simeq N_t + N_n \simeq 3/2 N_t$ , because  $N_t \simeq 2N_n$  from the

Euler relations. We have therefore constructed two families of unstructured meshes  $\{M_i\}_{1 \leq i \leq 6}$  and  $\{M_i\}_{1 \leq i \leq 6}$  respectively for DDFV and SWIP, and such that the number of unknowns for  $M_i$  and  $M_i$  are close. For each mesh, Tab. 1 refers to  $N_n, N_t$  and  $N_u$ . The number of nonzero elements  $nnz$ , the mean number of nonzero elements per row  $\overline{nnz}$  ( $= nnz/N_u$ ) and the bandwidth  $bw$  of the corresponding matrix (obtained after a reverse Cuthill–McKee ordering, see Cuthill and McKee (1969)) are also indicated. The first remark is that the number of nonzero elements is different for each method since it depends on the stencil. By considering fine meshes (typically  $M_5, M_6, M_5$  and  $M_6$ ), we have for DDFV,

$$nnz \simeq 7N_t + 13N_n \Rightarrow \overline{nnz} \simeq \frac{7N_t + 13N_n}{N_t + N_n - N_n^D} \simeq 9,$$

while we have for SWIP,

$$nnz = 9(5N_t - 2N_n + 2) \Rightarrow \overline{nnz} = \frac{3(5N_t - 2N_n + 2)}{N_t} \simeq 12.$$

The second remark is that the bandwidth is smaller for SWIP which has a more simple connectivity graph due to the block structure of its matrix. To sum up, DDFV requires less storage capacity whereas SWIP seems more efficient for solving each linear system.

TABLE 1: TC1 - Meshes used for DDFV on the top and SWIP on the bottom.

Mesh	$N_n$	$N_t$	$N_u$	$nnz$	$\overline{nnz}$	$bw$
$M_1$	28	34	42	234	5.6	6
$M_2$	79	118	159	1127	7.1	16
$M_3$	253	430	609	4897	8.0	34
$M_4$	917	1688	2461	21009	8.5	58
$M_5$	3380	6474	9570	83742	8.8	145
$M_6$	13233	25896	39031	348595	8.9	267
$M_1$	13	16	48	504	10.5	12
$M_2$	40	54	162	1728	10.7	15
$M_3$	128	204	612	6894	11.3	24
$M_4$	464	826	2478	28836	11.6	57
$M_5$	1702	3198	9594	113292	11.8	96
$M_6$	6591	12780	38340	456480	11.9	196

Tab. 2 presents the errors on the hydraulic head  $e_{\psi, \Omega}$  and on the velocity  $e_{v, \Omega}$  for all the meshes,

$$e_{\psi, \Omega} = \frac{\max_{1 \leq n \leq N_T} \|\psi^n - \psi_h^n\|_{L^2(\Omega)}}{\max_{1 \leq n \leq N_T} \|\psi^n\|_{L^2(\Omega)}}, \quad e_{v, \Omega} = \frac{\|v - v_h\|_{L^2(\Omega \times [0, T])}}{\|v\|_{L^2(\Omega \times [0, T])}}.$$

For the two methods, the results confirm theoretical estimations since a second-order convergence is observed on the



TABLE 2: TC1 - Convergence results for DDFV on the top and SWIP on the bottom.

Mesh	$\delta t$	$e_{\psi,\Omega}$		$e_{v,\Omega}$	
		error	rate	error	rate
$M_1$	8	1.00e-2		1.15e-1	
$M_2$	4	3.34e-3	2.46	4.18e-2	2.26
$M_3$	2	1.00e-3	1.90	1.91e-2	1.23
$M_4$	1	2.51e-4	2.23	9.41e-3	1.14
$M_5$	1/2	6.29e-5	2.03	4.59e-3	1.05
$M_6$	1/4	1.58e-5	1.99	2.28e-3	1.01
$M_1$	8	4.51e-2		3.25e-1	
$M_2$	4	4.40e-3	2.60	1.17e-1	0.70
$M_3$	2	1.10e-3	2.00	5.91e-2	0.99
$M_4$	1	2.70e-4	2.17	2.94e-2	1.08
$M_5$	1/2	6.41e-5	2.26	1.42e-2	1.14
$M_6$	1/4	1.59e-5	1.92	7.09e-3	0.96

hydraulic head and a first-order convergence is verified on the velocity.

Tab. 3 reports the mean number  $\bar{N}_{it}$  of iterations in the linearization algorithm and the mean condition number  $\bar{\kappa}_2^1$ ,

$$\bar{N}_{it} = \frac{1}{N_T} \sum_n N_{it}^n \quad \text{and} \quad \bar{\kappa}_2^1 = \left( \sum_n N_{it}^n \right)^{-1} \left( \sum_{n,m} \frac{\lambda_{\max}^{n,m}}{\lambda_{\min}^{n,m}} \right),$$

where  $N_{it}^n$  is the number of iterations performed at time level  $n$  and  $\lambda_{\max}^{n,m}$  (resp.  $\lambda_{\min}^{n,m}$ ) is the maximal (resp. minimal) eigenvalue of the matrix at the  $m$ -th nonlinear iteration of the  $n$ -th time step. We observe that the mean condition number is inversely proportional to the mesh size  $h$ . This is consistent with a finite difference result which states that the condition number  $\kappa_2$  is proportional to  $h^{-2}$ . Furthermore, the SWIP method significantly reduces  $e_{\psi,\Omega}$  (by a factor of 1.5) and  $\bar{\kappa}_2^1$  (by a factor of 6) obtained with the most classical symmetric interior penalty galerkin method (which uses  $\omega^+ = \omega^- = 0.5$ ). To balance the various error sources (due to space and time discretizations as well as nonlinear systems resolution), an adaptive inexact Newton method and adaptive time-stepping based on *a posteriori* estimation could be implemented, as was done in Ern and Vohralík (2010).

#### 4.2 Isotropic homogeneous stiff case

This TC is a stiff infiltration problem proposed by Celia et al. (1990). The domain is  $\Omega = [0, 20] \times [0, 100]$  (in *cm*) and the final time is  $T = 48h$ . A constant initial condition is considered. A Dirichlet condition is imposed on the top  $\mathcal{T}$  and the

TABLE 3: TC1 - Mean number of iterations and mean condition number.

Mesh	DDFV		SWIP	
	$\bar{N}_{it}$	$\bar{\kappa}_2^1$	$\bar{N}_{it}$	$\bar{\kappa}_2^1$
1	2.9	65.1	2.3	2.0
2	1.8	120.5	1.2	2.1
3	1.0	203.6	1.1	4.6
4	1.0	432.1	1.0	11.6
5	1.0	813.7	1.0	24.8
6	1.0	1625.4	1.0	54.5

bottom  $\mathcal{B}$  of the column. A homogeneous Neumann condition (corresponding to a zero flux) is imposed on the lateral parts  $\mathcal{L}$  of the domain (see Fig. 5),

$$\begin{cases} \psi^0 = -10 \text{ m} & \text{in } \Omega, \\ \psi^D = -10 \text{ m} & \text{on } \mathcal{B} \times ]0, T], \\ \psi^D = -75 \text{ cm} & \text{on } \mathcal{T} \times ]0, T], \\ v(\psi) \cdot n_\Omega = 0 & \text{on } \mathcal{L} \times ]0, T]. \end{cases}$$

The water content and the conductivity are defined by Van Genuchten's constitutive relationships, see Van Genuchten (1980), and plotted on Fig. 6 :

$$\theta(\psi) = \frac{(\theta_s - \theta_r)}{(1 + (\xi|\psi|)^\beta)^\gamma} + \theta_r,$$

$$K(\psi) = K_s \frac{(1 - (\xi|\psi|)^{\beta-1} (1 + (\xi|\psi|)^\beta)^{-\gamma})^2}{(1 + (\xi|\psi|)^\beta)^\frac{\gamma}{2}},$$

with parameters

$$\begin{aligned} \theta_s &= 0.368, & \theta_r &= 0.102, & K_s &= 9.22 \cdot 10^{-3} \text{ cm} \cdot \text{s}^{-1}, \\ \xi &= 0.0335 \text{ cm}^{-1}, & \beta &= 2, & \gamma &= 1 - \beta^{-1}. \end{aligned}$$

The evoked ‘‘stiffness’’ of this TC is related to the strong constrained overpressure (equal to  $9.25m$ ) imposed on the top of the column. A sharp variation of the conductivity ( $K(-75\text{cm})/K(-10\text{m}) = 8.92 \cdot 10^4$ ) is induced, see Fig 6.

We focus here on the horizontal mean value of the pressure head  $\bar{\psi}_h(z)$  :

$$\forall z \in [0, 100], \quad \bar{\psi}_h(z) = \frac{1}{20} \int_0^{20} \psi_h(x, z) dx.$$

Fig. 7 plots  $\bar{\psi}_h(z)$  at  $24h$  and  $48h$  obtained with the meshes ( $M_4 - M_4$ ), ( $M_5 - M_5$ ) and ( $M_6 - M_6$ ). The DDFV pressure profiles are depicted by solid lines and the SWIP pressure profiles are depicted by dashed lines. Each method presents one drawback when coarse meshes are used. For DDFV, a delay of the pressure front is observed, especially when  $M_4$  is employed since a significant shift is observed between the

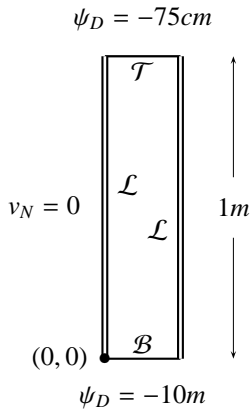


Figure 5: TC2 - Polmann test case.

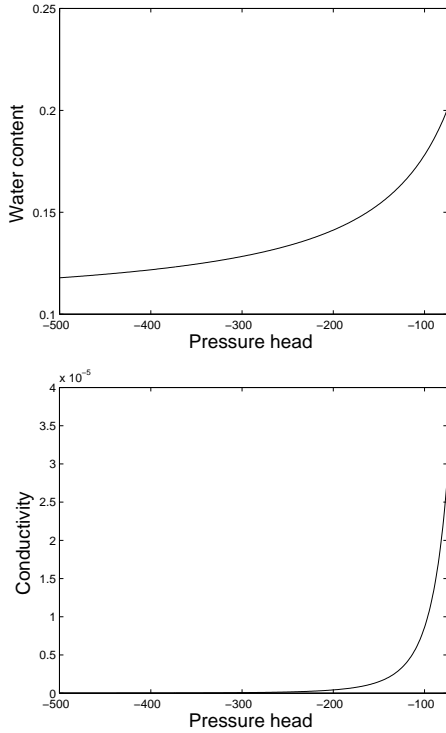


Figure 6: TC2 - Constitutive curves  $\theta(\psi)$  and  $K(\psi)$ .

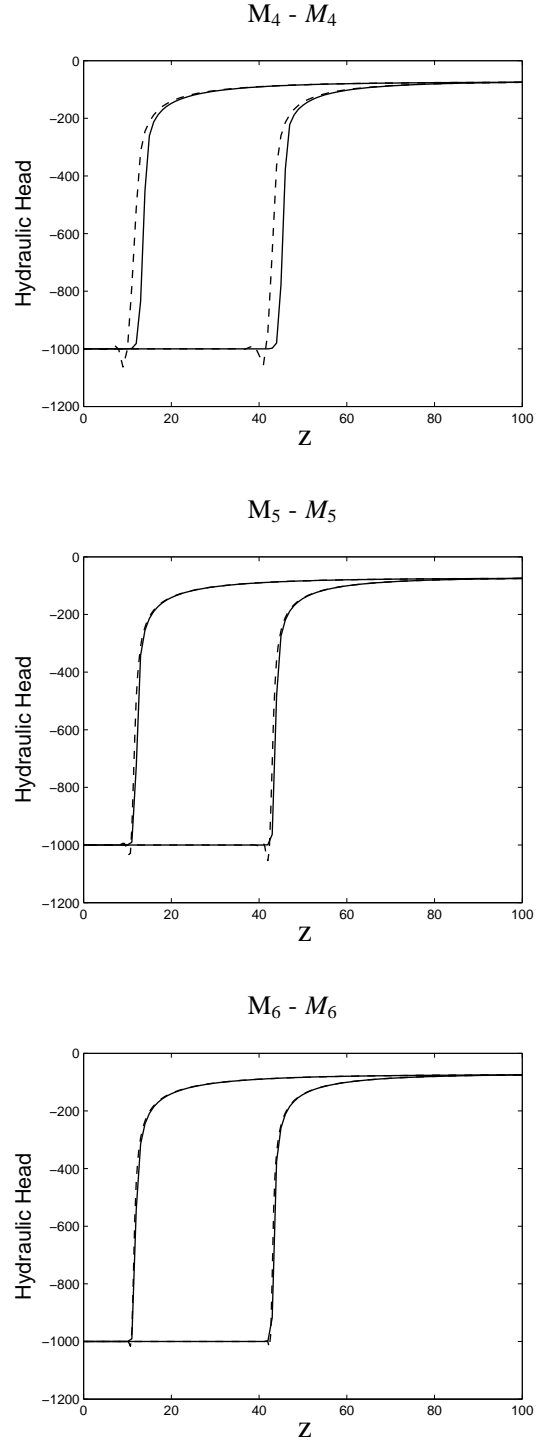


Figure 7: TC2 - Pressure head  $\bar{\psi}_h(z)$  at 24h and 48h obtained with different meshes for DDFV (solid line) and SWIP (dashed line).

two methods. For SWIP, non-physical oscillations appear and could be reduced with the help of a slope limiter as in Cockburn and Shu (1998). In this stiff case, only DDFV satisfies the discrete maximum principle whereas SWIP has a better estimation of the propagation speed of the pressure front. When the meshes are sufficiently fine, the two methods yield the same pressure in agreement with Celia et al. (1990) and Manzini and Ferraris (2004).

### 4.3 Anisotropic heterogeneous five-spot problem

This TC is inspired by the quarter five-spot problem studied in Simmons et al. (1959), which reproduces an elementary cell of a periodic network consisting of sources and sinks. The domain  $\Omega = [0, 1]^2$  (in  $m$ ) is divided into four parts (see Fig. 8),

$$\begin{aligned}\Omega_1 &= \Omega \cap \{x + z \leq 0.5\}, & \Omega_2 &= \Omega \cap \{0.5 < x + z \leq 1\}, \\ \Omega_3 &= \Omega \cap \{1 < x + z \leq 1.5\}, & \Omega_4 &= \Omega \cap \{1.5 < x + z\}.\end{aligned}$$

The soil properties are the same as those of the previous TC except for the piecewise constant intrinsic permeability,

$$\mathbb{K}(x) = \sum_{i=1}^4 \mathbb{1}_{\Omega_i}(x) R_{\omega_i} D R_{\omega_i}^t,$$

where  $\mathbb{1}_{\Omega_i}(x)$  is the indicator function of the set  $\Omega_i$ ,  $D$  is a diagonal matrix and  $R_{\omega_i}$  is the rotation matrix associated to  $\Omega_i$ :

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 10^{-3} \end{bmatrix} \quad \text{and} \quad R_{\omega} = \begin{bmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{bmatrix}.$$

The angles of rotation are :

$$\omega_1 = \pi/4, \quad \omega_2 = 0, \quad \omega_3 = \pi/2, \quad \omega_4 = \pi/4.$$

The matrix  $R_{\omega}^t$  denotes the transpose of  $R_{\omega}$ . This permeability induces a preferential direction for the flow on each part of the domain as illustrated on Fig 8.

The final time is  $T = 6h$ . An hydrostatic initial condition is considered, an homogeneous Neumann condition is imposed on the boundary of the domain except on the top right corner  $C$  where an incoming flux  $v_N$  is enforced and on the line  $\mathcal{L} = [0, 0.025] \times \{0\}$  where a zero pressure condition is applied,

$$\begin{cases} \psi^0 = -z & \text{in } \Omega, \\ \psi^D = 0 & \text{on } \mathcal{L} \times ]0, T], \\ v(\psi) \cdot n_{\Omega} = v_N & \text{on } C \times ]0, T], \\ v(\psi) \cdot n_{\Omega} = 0 & \text{on } \Gamma \times ]0, T], \end{cases}$$

where  $C = \{1\} \times [0.975, 1] \cup [0.975, 1] \times \{1\}$  and  $\Gamma = \partial\Omega \setminus (\mathcal{L} \cup C)$ . The function  $v_N$  (in  $cm.s^{-1}$ ) is defined as

$$v_N(t) = \begin{cases} -5 \cdot 10^{-3} \frac{t}{1800} & \text{if } t \leq 0.5h, \\ -5 \cdot 10^{-3} & \text{if } 0.5h < t \leq 4h, \\ 0 & \text{if } t > 4h. \end{cases}$$

The condition  $v_N = -5 \cdot 10^{-3} cm.s^{-1}$  on  $C$  corresponds to a water injection of  $9kg$  per hour.

Fig 9 shows that the triangles never cross the interface between each subdomain. The meshes are obtained from the

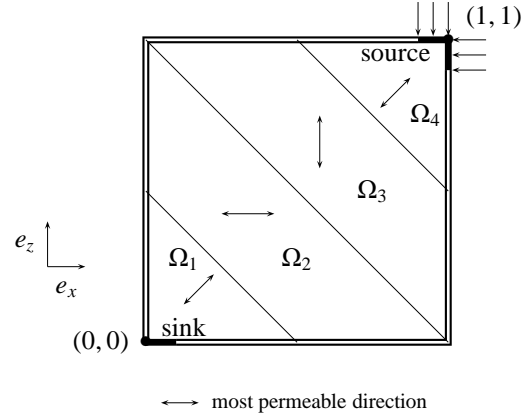


Figure 8: TC3 - Five-spot problem in anisotropic heterogeneous media.

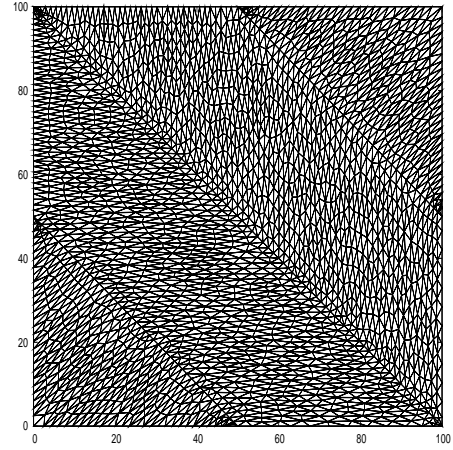


Figure 9: TC3 - Example of anisotropic heterogeneous mesh.

FreeFem<sup>©</sup> software, which allows to define anisotropic heterogeneous metrics. Fig 10 plots isolines of the overpressure  $\psi_h^T - \psi_h^0$  at the final time of the simulation. The two methods yield the same results for fine meshes (8760 triangles for DG, 16446 for DDFV). Concerning the convergence of the solution as  $h$  tends to 0, SWIP needs further refinement than DDFV when isotropic meshes are used. Therefore, DDFV is less sensitive than SWIP to the choice of the mesh when an anisotropic permeability is considered.

Fig 11 and Fig 12 present results on mass conservation. Multiplying equations (6) and (7) by the water density  $\rho$ , and summing over the time intervals in  $[0, n\delta t]$  leads to

$$\underbrace{\rho(V^n - V^0)}_{\Delta M^n} = \underbrace{\sum_{i=1}^n \rho \delta t \Phi_N^i}_{M_{in}^n} + \underbrace{\sum_{i=1}^n \rho \delta t \Phi_D^i}_{M_{out}^n} + \underbrace{\sum_{i=1}^n \rho e^n}_{E^n}.$$

$\Delta M^n$  is the total mass variation over the time interval  $[0, n\delta t]$  and the quantities  $\sum_{i=1}^n M_{in}^i$ ,  $\sum_{i=1}^n M_{out}^i$  and  $\sum_{i=1}^n |E^n|$  are respectively the total water inflow, the total water outflow and the total mass balance defect cumulated at time  $n\delta t$ . The quantities  $\Delta M^n$ ,  $\sum_{i=1}^n M_{in}^i$  and  $\sum_{i=1}^n M_{out}^i$  are close for the two methods and three phases are clearly recognizable on Fig 11. The first phase  $[0, 1h]$  corresponds to the soil saturation : the total mass variation equals the total water inflow. The second phase  $[1h, 4h]$  features soil saturation and drainage : the total mass variation increases more slowly than during the previous phase because exfiltration occurs. The last phase  $[4h, 6h]$  is the soil drainage only since the injection is stopped, inducing a diminution of the total mass variation.

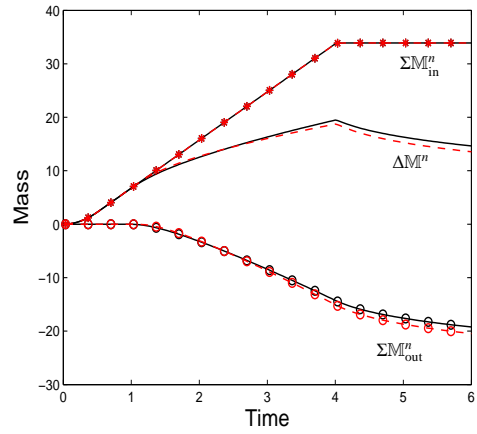


Figure 11: TC3 - Mass repartition for DDFV (solid lines) and SWIP (dashed lines).

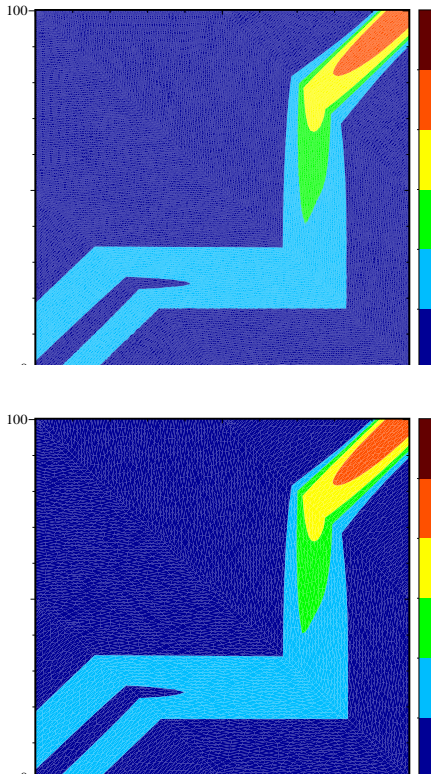


Figure 10: TC3 - Isolines of overpressure at 6h for DDFV (top) and SWIP (bottom).

The total mass balance defect (or mass error)  $\sum_{i=1}^n |E^n|$  is plotted on Fig 12, where various tolerances  $\epsilon$  in the linearization algorithm are tested. When  $\epsilon = 10^{-3}$ , the defects look similar, but lowering the tolerance has no improving effect with SWIP. Mass error can be reduced by using higher order approximations instead. Meanwhile, about two orders of magnitude can still be gained with DDFV before convergence, which is achieved for  $\epsilon = 10^{-5}$ .

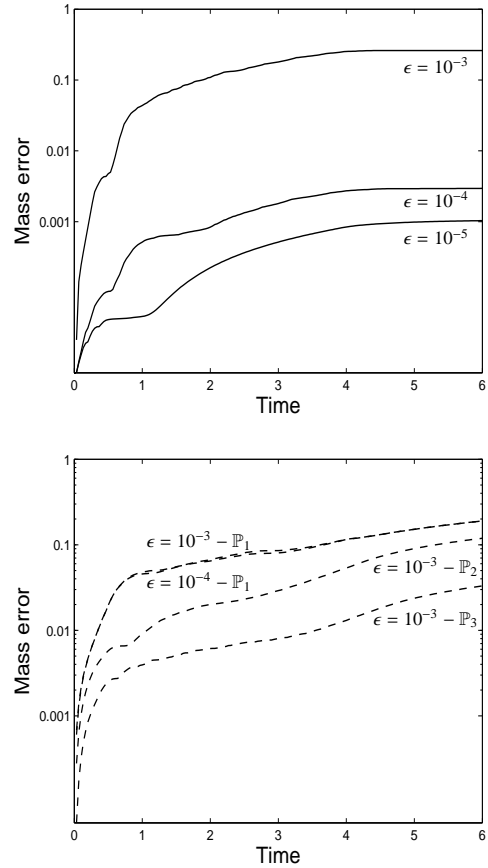


Figure 12: TC3 - Mass error for DDFV (solid lines) and SWIP (dashed lines).

## 5 CONCLUSION

We presented a comparison between a discontinuous finite element method and a more recent finite volume scheme on various test cases. As was presented in section 2, the approaches are quite different and the numerical results highlight different behaviors. First, DDFV provides a sparser structure than SWIP which produces in return a narrower bandwidth and a better condition number. In the stiff case, DDFV verifies the discrete maximum principle while SWIP evaluates better the speed of propagation for coarse meshes. Finally, in the quarter five-spot problem, the two methods are in accordance for the overpressure and the mass repartition over time. The mass balance defect can be considerably lowered with DDFV by requiring a more restrictive tolerance in the Picard algorithm.

## REFERENCES

- P. Bastian. Higher order discontinuous Galerkin methods for flow and transport in porous media. *Challenges in Scientific Computing-CISC 2002*, pages 1–22, 2003.
- P. Bastian, O. Ippisch, F. Rezanezhad, HJ Vogel, and K. Roth. Numerical simulation and experimental studies of unsaturated water flow in heterogeneous systems. *Reactive Flows, Diffusion and Transport*, pages 579–597, 2007.
- M.A. Celia, E.T. Bouloutas, and R.L. Zarba. A general mass-conservative numerical solution for the unsaturated flow equation. *Water Resources Research*, 26(7):1483–1496, 1990.
- B. Cockburn and C.W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM Journal on Numerical Analysis*, 35(6):2440–2463, 1998.
- Y. Coudière, C. Pierre, O. Rousseau, R. Turpault, et al. A 2D/3D discrete duality finite volume scheme. Application to ECG simulation. *IJFV*, 6(1):1–24, 2009.
- CF Curtiss and J.O. Hirschfelder. Integration of stiff equations. *Proceedings of the National Academy of Sciences of the United States of America*, 38(3):235, 1952.
- E. Cuthill and J. McKee. Reducing the bandwidth of sparse symmetric matrices. In *Proceedings of the 1969 24th national conference*, pages 157–172. ACM, 1969.
- D.A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Springer, 2012.
- K. Domelevo and P. Omnes. A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids. *ESAIM: Mathematical Modelling and Numerical Analysis*, 39(06):1203–1249, 2005.
- A. Ern and M. Vohralík. A posteriori error estimation based on potential and flux reconstruction for the heat equation. *SIAM Journal on Numerical Analysis*, 48(1):198–223, 2010.
- R. Eymard, C. Guichard, R. Herbin, and R. Masson. Vertex-centred discretization of multiphase compositional Darcy flows on general meshes. *Computational Geosciences*, pages 1–19, 2012.
- S. Fagherazzi, D.J. Furbish, P. Rasetarinera, and M.Y. Hussaini. Application of the discontinuous spectral Galerkin method to groundwater flow. *Advances in water resources*, 27(2):129–140, 2004.
- G. Hairer and Wanner E. *Solving ordinary differential equations II*. Springer, 2010.
- R. Haverkamp, M. Vauclin, J. Touma, PJ Wierenga, and G. Vachaud. A comparison of numerical simulation models for one-dimensional infiltration. *Soil Science Society of America Journal*, 41(2):285–294, 1977.
- R. Herbin, F. Hubert, et al. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. *Finite volumes for complex applications V*, pages 659–692, 2008.
- F. Hermeline. A finite volume method for the approximation of diffusion operators on distorted meshes. *Journal of computational Physics*, 160(2):481–499, 2000.
- W. Klieber and B. Rivière. Adaptive simulations of two-phase flow by discontinuous Galerkin methods. *Computer methods in applied mechanics and engineering*, 196(1):404–419, 2006.
- P. Knabner and E. Schneid. Adaptive hybrid mixed finite element discretization of instationary variably saturated flow in porous media. *High Performance Scientific and Engineering Computing*, 29:37–44, 2002.
- S. Krell. *Schémas Volumes Finis en mécanique des fluides complexes*. PhD thesis, Université de Provence-Aix-Marseille I, 2010.
- P. Lesaint and P.A. Raviart. On a finite element method for solving the neutron transport equations. 1974.
- G. Manzini and S. Ferraris. Mass-conservative finite volume methods on 2D unstructured grids for the Richards' equation. *Advances in Water Resources*, 27(12):1199–1215, 2004.
- TN Narasimhan and PA Witherspoon. An integrated finite difference method for analyzing fluid flow in porous media. *Water Resources Research*, 12(1):57–64, 1976.
- J. Simmons, BL Landrum, JM Pinson, and PB Crawford. Swept areas after breakthrough in vertically fractured five-spot patterns. *Trans. AIME*, 216:73, 1959.
- P. Sochala, A. Ern, and S. Piperno. Mass conservative bdf-discontinuous galerkin/explicit finite volume schemes for coupling subsurface and overland flows. *Computer Methods in Applied Mechanics and Engineering*, 198(27):2122–2136, 2009.
- M.T. Van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44(5):892–898, 1980.

The date of receipt and acceptance  
will be inserted by the editor.

Copyright © 2010 IFPEN Energies nouvelles

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than IFP Energies nouvelles must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee: Request permission from Documentation, IFP Energies nouvelles, fax. +33 1 47 52 70 78, or [revueogst@ifpen.fr](mailto:revueogst@ifpen.fr).