On several numerical strategies to solve Richards' equation in heterogeneous media with Finite Volumes

Sabrina Bassetto, Clément Cancès, Guillaume Enchéry, Quang Huy Tran

On several numerical strategies to solve Richards' equation in heterogeneous media with Finite Volumes

Sabrina Bassetto[*,†] Clément Cancès[†] Guillaume Enchéry[*]
Quang-Huy Tran[*]

June 13, 2021

Abstract

We benchmark several numerical approaches building on upstream mobility two-point flux approximation finite volumes to solve Richards' equation in domains made of several rock-types. Our study encompasses four different different schemes corresponding to different ways to approximate the nonlinear transmission condition systems arising at the interface between different rocks, as well as different resolution strategies based on Newton's method with variable switch. The different methods are compared on filling and drainage test cases with standard nonlinearities of Brooks-Corey and van Genuchten type, as well as with challenging steep nonlinearities.

Keywords

Richards' equation, heterogeneous domains, finite volumes, Newton's method, parametrization based variable switch

Mathematics subject classification

76S05, 65M08, 65H10

1 Introduction

The Richards equation [43] is a popular model for underground water flow in the vadose zone. It consists in a simplification of the incompressible immiscible two-phase Darcy flow model, assuming that the pressure of the gas phase is known and equal to the atmospheric pressure, see for instance [5]. Besides, Richards equation also attracts an important interest from scientists as it provides a relatively simple model that already accounts for many difficulties occurring in complex porous media flows, like degeneracies when one phase (air or water) vanishes, or strong material heterogeneities with severe changes in the physical parameters at the interface between different rocks. We formalize mathematically in Section 1.1 the problem under consideration in this paper, namely Richards equation in heterogeneous domains, before discussing on possible numerical strategies in Section 1.2.

[*]IFP Energies nouvelles, 1 et 4 avenue de Bois Préau, 92852 Rueil-Malmaison Cedex, France. sabrina.bassetto@ifpen.fr, guillaume.enchery@ifpen.fr, quang-huy.tran@ifpen.fr

[†]Inria, Univ. Lille, CNRS, UMR 8524 – Laboratoire Paul Painlevé, 59000 Lille, France. clement.cances@inria.fr

1.1 The Richards equation in heterogeneous domains

Let $\Omega \subset \mathbb{R}^d$ $(1 \leqslant d \leqslant 3)$ be a connected open polyhedral domain, representing the porous matrix in which water flows. The porous matrix is assumed to be heterogeneous, and we particularly focus on severe variations of the rock characteristics at the interface between different rock-types. More precisely, we assume that there exist polyhedral connected and disjointed open subsets $(\Omega_i)_{1\leqslant i\leqslant I}$ such that

$$\overline{\Omega} = \bigcup_{1\leqslant i\leqslant I} \overline{\Omega}_i.$$

Each subdomain Ω_i represents a rock-type, and is assumed to be homogeneous for simplicity. We denote by

$$\Gamma_{i,j} = \overline{\Omega}_i \cap \overline{\Omega}_j, \quad 1 \leqslant i, j \leqslant I,$$

the interface between Ω_i and Ω_j and by

$$\Gamma = \bigcup_{1\leqslant i\neq j\leqslant I} \Gamma_{i,j}$$

the set containing all these interfaces.

Let $T > 0$ be an arbitrary finite time horizon, then Richards' equation in $Q_{i,T} = (0,T) \times \Omega_i$ writes

$$\phi_i\, \partial_t s + \operatorname{div} F = 0, \tag{1.1}$$

$$F + \lambda_i\, \eta_i(s)\, \nabla(p + \varrho g z) = 0, \tag{1.2}$$

$$s - \mathcal{S}_i(p) = 0. \tag{1.3}$$

The unknowns are the water saturation s, the water flux F and the water pressure p. Equation (1.1) encodes the local conservation of the water volume (since water is described as an incompressible fluid). The Darcy-Muskat relation (1.2) relates the water flux to the gradient of the hydraulic head, whereas the last equation (1.3) links the saturation to the pressure. In the above system, ϕ_i stands for the porosity of the i^{th} rock and $\lambda_i > 0$ for its intrinsic permeability (isotropy of the porous medim is assumed here), while ϱ stands for the water density which is assumed to be constant, and g denotes the modulus of the gravity vector. The mobility $\eta_i(s)$ is nonnegative and nondecreasing with respect to the saturation, while the function $\mathcal{S}_i$ relating the water pressure and saturation is nondecreasing and takes its values in $[0,1]$. In accordance with the classical models of the literature — see Section 4.1.3 for the precise description of the models to be used in practice in the numerical simulations — we assume that water is always mobile, i.e., that $\eta_i(\mathcal{S}_i(p)) > 0$ for all $p \in \mathbb{R}$. Water becomes immobile in the dry asymptote, i.e. $\lim_{p\to-\infty} \eta_i(\mathcal{S}_i(p)) = 0$, leading to a degeneracy of hyperbolic type. On the other hand, positive pressures correspond to saturated regimes, i.e. $\mathcal{S}_i(p) = \mathcal{S}_i(0)$ for all $p \geqslant 0$, leading to a degeneracy of elliptic type.

At the interface $\Gamma_{i,j}$, pressure and flux are continuous. More precisely, denote by p_i the trace at $(0,T) \times \Gamma_{i,j}$ of the pressure $p_{|\Omega_i}$ in $Q_{i,T}$, and by F_i the trace at $(0,T) \times \Gamma_{i,j}$ of the flux $F_{|\Omega_i}$ in $Q_{i,T}$, then the transmission conditions across $\Gamma_{i,j}$ write

$$F_i \cdot \nu_i + F_j \cdot \nu_j = 0, \tag{1.4}$$

$$p_i - p_j = 0, \tag{1.5}$$

where ν_i (resp. ν_j) denotes the normal to $\Gamma_{i,j}$ outward w.r.t. Ω_i (resp. Ω_j). Note that since the pressure is continuous, and since $\mathcal{S}_i \neq \mathcal{S}_j$ in general, the saturation is discontinuous across $\Gamma_{i,j}$. The pressure continuity (1.5) has to be relaxed in the case where the water mobility could vanish for finite p. We refer for instance to [7, 17, 21] for formulations with such relaxed pressure continuity conditions at the interfaces. Let us stress that our work can be extended without further difficulties to this more involving setting.

Concerning the boundary conditions, the external boundary $\partial\Omega$, the outward normal of which being denoted by ν, is split into a portion called Γ^D where a constant Dirichlet boundary condition is imposed, and $\Gamma^N = \partial\Omega\backslash\Gamma^D$ where a Neumann boundary condition is fixed:

$$F \cdot \nu = q_N \quad \text{on } (0,T) \times \Gamma^{\mathrm{N}}, \tag{1.6}$$

$$p = p^{\mathrm{D}} \quad \text{on } (0,T) \times \Gamma^{\mathrm{D}}. \tag{1.7}$$

Finally, the initial saturation profile is prescribed,

$$s(0,x) = s^0(x) \text{ in } \Omega. \tag{1.8}$$

1.2 Motivation and positioning of our work

Richards equation is interesting in itself for modeling the infiltration of water in the near subsurface. This motivated the development of many numerical approaches with the aim of being robust while preserving accuracy, especially with respect to mass conservation. For numerical schemes approximating the solutions to Richards equation, we refer for instance to [19] for finite differences, to [29] for control volume finite elements, to [25,27] for two-point flux approximation (TPFA) finite volumes and to [13,34,46] for more advanced finite volume methods, to [6,42,48] for mixed finite elements, or to [36] for discontinuous Galerkin approaches. The above reference list is far from being exhaustive, and we refer to [28] for a review.

The problem being nonlinear and degenerate, an important part of the research effort has been assigned to the design of efficient iterative linearization procedures. Two main approaches then emerge: a first one based on (modified) Picard type fixed point strategies, and second one relying on Newton's method. Suitably designed Picard iteration based methods are known to enjoy robustness at the price of a mere linear convergence speed, see for instance [19,37,41,45]. On the other hand, a crude Newton's algorithm may face severe difficulties to converge, see for instance [35,37] for comparison of different approaches. This motivated the introduction of methods based on variable switch [20,30], nested Newton loops [18], or nonlinear preconditioning technics [8] to increase robustness. Our approach, which is described in Section 2.3 and [3], relies on the so-called parametrization approach introduced in [9,12], which can be interpreted as a generalization of the variable switch approach as well as a (diagonal) nonlinear preconditioning technique.

The second main difficulty to be addressed is the strong heterogeneity of the domain Ω with discontinuous physical characteristics across Γ. Since the pressure is continuous, cf. (1.5), schemes that are based on formulations involving the Kirchhoff transform $\theta_i = \int_0^p \eta_i(\mathcal{S}_i(a))\mathrm{d}a$, which is known to be a powerful tool for the mathematical [1] and numerical [27,42,48] study of Richards equation, will require a specific treatment at the interfaces to maintain the continuity of the pressure. We refer for instance to [10,16,21–23,31] for methods built in this spirit. A more natural approach consists in using discrete fluxes expressed directly in the form (1.2), with degrees of freedom localized on the interface Γ to enforce the continuity of the pressure, as done for instance in [2,14,26,32,40]. Let us also mention [44] where the authors solve the transmission condition (1.4)–(1.5) thanks to an iterative procedure stemming from domain decomposition. In the case of cell centered methods, like for instance TPFA finite volumes, convergence can also be assessed without any specific treatment of the interface, as for example done in our recent contribution [4]. However, the pressure continuity is only imposed at convergence w.r.t. grid refinement, leading to possible loss in the accuracy. Therefore, specific treatments of the interface are needed. As highlighted in [11], the specific treatment of the interface Γ may have a major impact on the Newton's method behavior. The purpose of this paper is to compare several approaches described in Section 3 to deal with the interface transmission condition (1.4)–(1.5) and to depict their pros and cons when confronted to different physical settings described in Section 4.

2 Problem discretization

2.1 Space-time discretization

Let $(\mathscr{T},\mathscr{E})$ be a finite-volume space discretization of Ω satisfying the classical orthogonality condition required for the consistency of the Two Point Flux Approximation (TPFA), see [24, Definition 9.1] for more details. Here $\mathscr{T}$ denotes the set of cells and $\mathscr{E}$ the set of faces. We assume that the mesh is consistent with the geometry in the sense that, for all $K \in \mathscr{T}$, there exists $i \in \{1,\dots,I\}$ such that $K \subset \Omega_i$. We denote by $\mathscr{T}_i = \{K \in \mathscr{T} \mid K \subset \Omega_i\}$. Then for all $f \in \{\mathcal{S},\lambda,\phi,\dots\}$ that depends on the rock-type, we set $f_K = f_i$ if $K \in \mathscr{T}_i$. The set $\mathscr{E}$ is then subdivided into: the set of internal faces shared by cells of the same subdomain $\mathscr{E}_i = \{\sigma = K|L \in \mathscr{E} \mid K,L \in \mathscr{T}_i\}$, the set of the internal faces shared by cells belonging to different subdomains $\mathscr{E}_\Gamma = \{\sigma \in \mathscr{E} \mid \sigma \subset \Gamma\} = \bigcup_{i\neq j}\{\sigma = K|L \in \mathscr{E} \mid K \in \mathscr{T}_i,\ L \in \mathscr{T}_j\}$, the set of Dirichlet faces $\mathscr{E}_{\rm ext}^D = \{\sigma \in \mathscr{E} \,|\, \sigma \subset \Gamma^D\}$ and the set of Neumann faces $\mathscr{E}_{\rm ext}^N = \{\sigma \in \mathscr{E} \,|\, \sigma \subset \Gamma^N\}$. Let us call $\mathscr{E}_{\rm int} = \mathscr{E}_i \cup \mathscr{E}_\Gamma$ the set of all internal faces. We also introduce the local set $\mathscr{E}_K = \{\sigma \in \mathscr{E} \,|\, \sigma \subset \partial K\}$ containing all the faces surrounding a cell K. To each face $\sigma \in \mathscr{E}$ we associate a distance d_σ defined by

$$d_\sigma = \begin{cases} |x_K - x_L| & \text{if } \sigma = K|L \in \mathscr{E}_{\rm int}, \\ d_{K,\sigma} & \text{if } \sigma \in \mathscr{E}_K \cap (\mathscr{E}_{\rm ext}^{\rm D} \cup \mathscr{E}_{\rm ext}^{\rm N}) \end{cases} \tag{2.1}$$

where, for all pair (K,σ) such that $\sigma \in \mathscr{E}_K$, $d_{K,\sigma} = |x_K - x_\sigma|$, with x_K the cell center and x_σ the face center, which is chosen as the intersection of $[x_K,x_L]$ with σ. Moreover, for each cell K, we denote by m_K its Lebesgue measure, and by m_σ the measure of a face σ. The time discretization is given by a vector of values $(t^n)_{0\leqslant 1\leqslant N}$ with $0 = t^0 < t^1 < \cdots < t^N = T$, and we denote by $t^n - t^{n-1} = \Delta t^n$, $1 \leqslant n \leqslant N$, the time steps.

2.2 Upstream TPFA finite-volume scheme

The two-point flux approximation of a diffusive flux, $F_{K\sigma}$, related to the gradient of an unknown u and coming out a cell K through the face σ, is defined by:

$$J_{K\sigma} = a_\sigma(u_{K\sigma} - u_K)$$

where the transmissivity on the face $\sigma \in \mathscr{E}$ is defined by $a_\sigma = \frac{m_\sigma}{d_\sigma}$ and the mirror value $u_{K\sigma}$ by

$$u_{K\sigma} = \begin{cases} u_L & \text{if } \sigma = K|L \in \mathscr{E}_{\rm int}, \\ u_K & \text{if } \sigma \in \mathscr{E}_K \cap \mathscr{E}_{\rm ext}^{\rm N}, \\ u_\sigma & \text{if } \sigma \in \mathscr{E}_K \cap \mathscr{E}_{\rm ext}^{\rm D}. \end{cases}$$

The saturation capillary-pressure relationship (1.3) and the volume balance (1.1) are discretized into

$$s_K^n = \mathcal{S}_K(p_K^n), \qquad K \in \mathscr{T}_i, n \geqslant 1, \tag{2.2}$$

$$m_K\phi_K\frac{s_K^n - s_K^{n-1}}{\Delta t} + \sum_{\sigma\in\mathscr{E}_K} F_{K\sigma}^n = 0, \qquad K \in \mathscr{T}_i,\ n \geqslant 1, \tag{2.3}$$

where the Darcy flux (1.2) is approximated by

$$F_{K\sigma}^n = \begin{cases} a_\sigma\lambda_K\eta_\sigma^n[(p_K^n - p_{K\sigma}^n) + \varrho g\ (z_K - z_{K\sigma})], & \sigma \in \mathscr{E}_K \cap (\mathscr{E}_{\rm int} \cup \mathscr{E}_{\rm ext}^D), \\ m_\sigma q_\sigma^n, & \sigma \in \mathscr{E}_K \cap \mathscr{E}_{\rm ext}^N, \end{cases} \tag{2.4a}$$

with

$$q_\sigma^n = \frac{1}{m_\sigma}\Delta t^n \int_{t^{n-1}}^{t^n}\int_\sigma q_N.$$

In (2.4a), the face mobilities are upwinded in the following way

$$\eta_\sigma^n = \begin{cases} \eta_K(s_K^n) & \text{if } (p_K^n - p_{K\sigma}^n) + \varrho g\,(z_K - z_{K\sigma}) \geqslant 0, \\ \eta_K(s_{K\sigma}^n) & \text{otherwise.} \end{cases} \tag{2.4b}$$

The initial condition (1.8) is discretized into

$$s_K^0 = \frac{1}{m_K}\int_K s^0, \quad \forall K \in \mathscr{T}_i, \tag{2.5}$$

and the Dirichlet boundary condition (1.7) into

$$p_\sigma^D = \frac{1}{m_\sigma}\int_\sigma p^D, \quad \forall \sigma \in \mathscr{E}_{\text{ext}}^D. \tag{2.6}$$

2.3 Switch of variable and parametrization technique

Let us now detail the resolution strategy for problem (2.2)–(2.6). A natural approach to solve this nonlinear system is to choose the pressure $(p_K)_{K\in\mathscr{T}}$ as primary unknown and to solve it via an iterative method such as Newton's one. However, the pressure variable is known to be an inefficient choice for dry soils $s \ll 1$, because of the degeneracy of Richards' equation, where schemes in which saturation is the primary variable outperform. On the other hand, the knowledge of the saturation is not sufficient to describe the pressure curve in saturated regions where the pressure-saturation relation cannot be inverted. This motivated the design of schemes which introduce a switch of variable [20, 30]. Our approach is based on the technique proposed by Brenner and Cancès [9], in which a third generic variable τ is introduced to become the primary unknown of the system. Then, removing the subscript i related to the rock-type for convenience, the idea is to choose a parametrization of the graph $\{p, \mathcal{S}(p)\}$, i.e., to construct two monotone functions $\mathfrak{s} : I \to [s_{\text{rw}}, 1 - s_{\text{rn}}]$ and $\mathfrak{p} : I \to \mathbb{R}$ such that $\mathfrak{s}(\tau) = \mathcal{S}(\mathfrak{p}(\tau))$ and $\mathfrak{s}'(\tau) + \mathfrak{p}'(\tau) > 0$ for all $\tau \in I \subset \mathbb{R}$. The latter non-degeneracy assumption ensures that for all $p \in \mathbb{R}$, there exists a unique $\tau \in \mathbb{R}$ such that $(p, \mathcal{S}(p)) = (\mathfrak{p}(\tau), \mathfrak{s}(\tau))$. Such a parametrization is not unique, for instance we can choose $I = \mathbb{R}$, $\mathfrak{p} = \text{Id}$ which is equivalent to solving the system always in pressure, but this is not recommended as seen before. Here, we set $I = \mathbb{R}$ and

$$\mathfrak{s}(\tau) = \begin{cases} \mathcal{S}(\kappa(\tau - \tau_*) + p_*) & \text{if } \tau \leqslant \tau_*, \\ s_{\text{rw}} + \tau(1 - s_{\text{rn}} - s_{\text{rw}}) & \text{if } \tau_* \leqslant \tau \leqslant \tau_{\text{s}}, \\ \mathcal{S}(p_{\text{s}} + \varsigma(\tau - \tau_{\text{s}})) & \text{if } \tau \geqslant \tau_{\text{s}}, \end{cases} \tag{2.7a}$$

$$\mathfrak{p}(\tau) = \begin{cases} \kappa(\tau - \tau_*) + p_* & \text{if } \tau \leqslant \tau_*, \\ \mathcal{S}^{-1}(s_{\text{rw}} + \tau(1 - s_{\text{rn}} - s_{\text{rw}})) & \text{if } \tau_* \leqslant \tau \leqslant \tau_{\text{s}}, \\ p_{\text{s}} + \varsigma(\tau - \tau_{\text{s}}) & \text{if } \tau \geqslant \tau_{\text{s}}. \end{cases} \tag{2.7b}$$

In the above formulas, $(p_{\text{s}}, s_{\text{s}}) = (\mathfrak{p}(\tau_{\text{s}}), \mathfrak{s}(\tau_{\text{s}}))$ is referred later on as the switching point, at which one passes from τ behaving as the saturation to τ behaving as the pressure (recall that Newton's iterations are not sensitive to linear changes of variables). Another switch is incorporated at $(p_*, s_*) = (\mathfrak{p}(\tau_*), \mathfrak{s}(\tau_*))$ to improve Newton's robustness in presence of heterogeneities. The parameter τ_* is chosen so small that the solution $(p_K^n, s_K^n)_{K\in\mathscr{T}}$ to the scheme is always larger that (p_*, s_*). The parameters κ and ς are chosen so that $\mathfrak{p}$ is C^1, leading to the expressions

$$\kappa = \frac{1 - s_{\text{rn}} - s_{\text{rw}}}{\mathcal{S}'(p_*^+)}, \quad \text{and} \quad \varsigma = \frac{1 - s_{\text{rn}} - s_{\text{rw}}}{\mathcal{S}'(p_{\text{s}}^-)}, \tag{2.8}$$

where $\mathcal{S}'(p_*^+)$ and $\mathcal{S}'(p_{\text{s}}^-)$ respectively denote the limits of $\mathcal{S}'(p)$ as p tends to p_* and p_{s} from above and below. Then if $\mathcal{S}$ is C^1, so is $\mathfrak{s} = \mathcal{S} \circ \mathfrak{p}$. When $\mathcal{S}$ is convex then concave, as in the Brooks-Corey and van Genuchten settings detailed in Section 4.1.3, then choosing τ_{s} such that $(p_{\text{s}}, s_{\text{s}})$

is the inflexion point of the graph of $\mathcal{S}$ ensures that both $\mathfrak{p}$ and the restriction of $\mathfrak{s}$ to $[\tau_*, +\infty)$ are concave. Moreover, if $\mathcal{S}$ belongs to $C^2(\mathbb{R})$ as in the van Genuchten setting, then so do the restrictions of $\mathfrak{p}$ and $\mathfrak{s}$ to $(\tau_*, +\infty)$. An example of parametrized curves $\mathfrak{p}, \mathfrak{s}$ corresponding to van Genuchten pressure-saturation laws is shown in Figure 1.

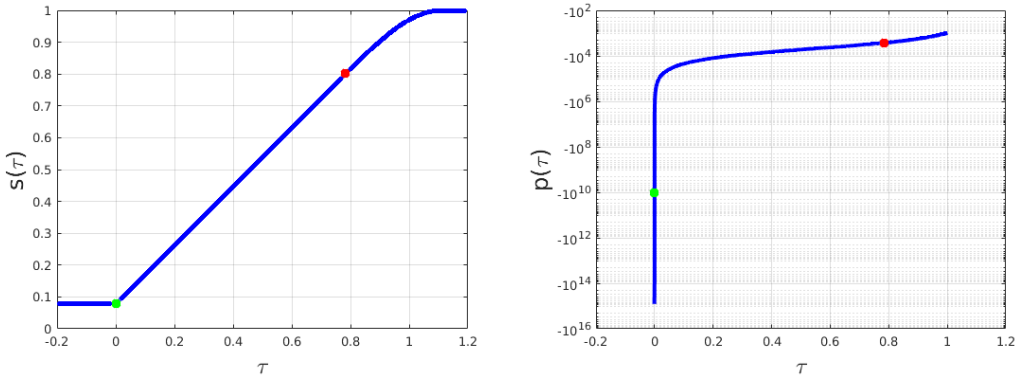

Figure 1: Plot of saturation and pressure parametrized van Genuchten-Mualem curves, using values of rock type 1 reported in Table 2 . The green dot indicate the value for $\tau = \tau_*$ and the magenta one $\tau = \tau_s$.

Applying the parametrization to our equations, we obtain the parametrized system:

$$m_K \phi_K \frac{\mathfrak{s}_K(\tau_K^n) - \mathfrak{s}_K(\tau_K^{n-1})}{\Delta t} + \sum_{\sigma \in \mathscr{E}_K} F_{K\sigma}^n = 0, \qquad K \in \mathscr{T}_i,\ n \geqslant 1, \tag{2.9}$$

where the fluxes (2.4a) become

$$F_{K\sigma}^n = \begin{cases} a_\sigma \lambda_K \eta_\sigma^n [(\mathfrak{p}_K(\tau_K^n) - \mathfrak{p}_K(\tau_{K\sigma}^n)) + \rho g\ (z_K - z_{K\sigma})], & \sigma \in \mathscr{E}_K \cap (\mathscr{E}_{\rm int} \cup \mathscr{E}_{\rm ext}^D), \\ m_\sigma q_N, & \sigma \in \mathscr{E}_K \cap \mathscr{E}_{\rm ext}^N, \end{cases} \tag{2.10}$$

and the upwinded face mobilities turn into

$$\eta_\sigma^n = \begin{cases} \eta_K(\mathfrak{s}_K(\tau_K^n)) & \text{if } (\mathfrak{p}_K(\tau_K^n) - \mathfrak{p}_K(\tau_{K\sigma}^n)) + \varrho g\ (z_K - z_{K\sigma}) \geqslant 0, \\ \eta_K(\mathfrak{s}_K(\tau_{K\sigma}^n)) & \text{otherwise.} \end{cases} \tag{2.11}$$

Finally, we rewrite the initial condition as

$$\tau_K^0 = \mathfrak{s}_K^{-1}\left(\frac{1}{m_K}\int_K s^0\right), \quad \forall K \in \mathscr{T}_i, \tag{2.12}$$

and the Dirichlet boundary condition as

$$\tau_\sigma^D = \mathfrak{p}_K^{-1}\left(\frac{1}{m_\sigma}\int_\sigma p^D\right), \quad \forall \sigma \in \mathscr{E}_{\rm ext}^D. \tag{2.13}$$

We have not specified yet how the interface fluxes $F_{K\sigma}^n$ for $\sigma = \mathscr{E}_\Gamma$ are treated. This specification is the purpose of Section 3. In the case of a homogeneous domain where $\Gamma = \varnothing$, the resulting system $\mathcal{F}_n(\boldsymbol{\tau}^n) = \mathbf{0}$ which is fully equivalent to (2.2)–(2.6), admits a unique solution $\boldsymbol{\tau}^n$ (for details see [4, Proposition 3.6-3.7]).

Remark 2.1. *The practical resolution of the nonlinear system relies on Newton's method. In the homogeneous setting, the method we use is the one that is presented in [3, Section 2] with some*

differences. The first one concerns the approximation of the k_r law by the van Genuchten Mualem model, that we explain in Section 4.1.3. Another one is related to the values of τ: here $I = \mathbb{R}$, so no projection of τ after each Newton iteration to avoid $\tau < s_{\rm rw}$ is required. Then, when we treat dry zones, we risk to manage a singular Jacobian matrix. In order to avoid this we impose that $k_r(s_{\rm rw}) = 10^{-33}$ when evaluating $\mathbb{J}$ to allow the pressure-gravity motor not to be zero. Finally, to help Newton's algorithm recover the good direction when the ℓ_∞ norm of the residual exceeds 10^2, a relaxation is activated with 0.3 as relaxing constant.

3 Numerical treatment of the interface

This section is devoted to the presentation of different strategies to approximate the transmission conditions (1.4)–(1.5) across the faces $\sigma \in \mathscr{E}_\Gamma$ located at an interface between to different rock-types. We propose four schemes, referred as methods A to D. For the last one, two different iterative Newton-based solvers are proposed.

3.1 Method A

This method basically consists in treating the interfaces as standard bulk faces, leading to the formula

$$F^n_{K\sigma} = a_\sigma \lambda_\sigma \eta^n_\sigma [(\mathfrak{p}_K(\tau^n_K) - \mathfrak{p}_L(\tau^n_L)) + \varrho g\ (z_K - z_L)], \quad \sigma = K|L \in \mathscr{E}_\Gamma, \tag{3.1}$$

where the face permeabilities $(\lambda_\sigma)_{\sigma\in\mathscr{E}_\Gamma}$ are given by

$$\lambda_\sigma = \frac{\lambda_K \lambda_L d_\sigma}{\lambda_K d_{L,\sigma} + \lambda_L d_{K,\sigma}}, \quad \sigma = K|L \in \mathscr{E}_\Gamma, \tag{3.2}$$

and the upwind face mobilities turn into

$$\eta^n_\sigma = \begin{cases} \eta_K(\mathfrak{s}_K(\tau^n_K)) & \text{if } (\mathfrak{p}_K(\tau^n_K) - \mathfrak{p}_L(\tau^n_L)) + \varrho g\ (z_K - z_L) \geqslant 0, \\ \eta_L(\mathfrak{s}_L(\tau^n_L)) & \text{otherwise.} \end{cases} \tag{3.3}$$

Therefore, the continuity of the normal flux (1.4) is exactly transposed into the local conservation condition

$$F^n_{K\sigma} + F^n_{L\sigma} = 0, \qquad \forall\, \sigma = K|L \in \mathscr{E}_\Gamma,\ n \geqslant 1. \tag{3.4}$$

On the other hand, $p^n_K \neq p^n_L$ in general. The pressure continuity (1.5) is recovered asymptotically as d_σ tends to 0 from (3.1). More precisely, assuming that $|F^n_{K\sigma}| \leqslant C m_\sigma$, then we deduce from (3.1) that $|\mathfrak{p}_K(\tau^n_K) - \mathfrak{p}_L(\tau^n_L)| \leqslant C d_\sigma$, where the constant C has been updated and further depends on $\lambda_\sigma, \max_i \|\eta_i \circ \mathcal{S}_i\|_\infty$ and ϱg.

The scheme (2.9)–(2.11), complemented by the interface fluxes (3.1)–(3.3), has been shown in [4] to be well-posed in the sense that the corresponding nonlinear system admits a unique solution $(\tau^n_K)_{K\in\mathscr{T}}$. Further, the rigorous convergence of the scheme as the mesh size and the time steps tend to 0 is also established. However, the numerical results presented in [4] (as well as those presented in what follows) show that the expected first order convergence can be lost in presence of heterogeneities. Methods B, C, and D have been designed as remedies to this loss of accuracy, which takes its origin in the poor approximation of the pressure continuity (1.5) by Method A.

3.2 Method B

This method, introduced in [4] consists in adding two thin cells, denoted by $I_{\sigma,K}$ and $I_{\sigma,L}$, of thickness $\delta_B \ll d_\sigma$ on both sides of each face $\sigma = K|L \in \mathscr{E}_\Gamma$ located at a rock-type interface, as depicted in Figure 2. This leads to the adjunction of two additional unknowns $\tau^n_{\sigma,K}$ and $\tau^n_{\sigma,L}$ per interface $\sigma = K|L \in \mathscr{E}_\Gamma$, that will allow for a more precise approximation of the pressure continuity condition (1.5). Define $(F^n_{K\sigma})_{\sigma\in\mathscr{E}_\Gamma}$ to be used in (2.9) by setting for $\sigma = K|L \in \mathscr{E}_\Gamma$ and $n \geqslant 1$:

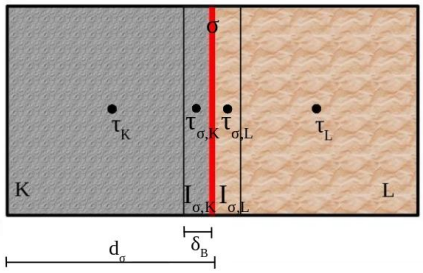

Figure 2: Method B: introduction of two thin cells on both sides of a face located between two rock types

$$F^n_{K\sigma} = a_{\sigma,K}\lambda_K\eta^n_{\sigma,K}\left(\mathfrak{p}_K(\tau^n_K) - \mathfrak{p}_K(\tau^n_{\sigma,K}) + \varrho g(z_K - z_{\sigma,K})\right), \tag{3.5}$$

$$F^n_{L\sigma} = a_{\sigma,L}\lambda_L\eta^n_{\sigma,L}\left(\mathfrak{p}_L(\tau^n_L) - \mathfrak{p}_L(\tau^n_{\sigma,L}) + \varrho g(z_L - z_{\sigma,L})\right), \tag{3.6}$$

where we have set

$$a_{\sigma,K} = \frac{m_\sigma}{d_{K,\sigma} - \delta_{\rm B}/2}, \quad a_{\sigma,L} = \frac{m_\sigma}{d_{L,\sigma} - \delta_{\rm B}/2},$$

where $z_{\sigma,K}$ and $z_{\sigma,L}$ are the respective vertical coordinates of

$$x_{\sigma,K} = x_K + \frac{d_{K,\sigma} - \delta_{\rm B}/2}{d_\sigma}(x_L - x_K) \quad \text{and} \quad x_{\sigma,L} = x_L + \frac{d_{L,\sigma} - \delta_{\rm B}/2}{d_\sigma}(x_K - x_L)$$

and

$$\eta^n_{\sigma,K} = \begin{cases} \eta_K(\mathfrak{s}_K(\tau^n_K)) & \text{if } \mathfrak{p}_K(\tau^n_K) + \varrho g z_K \geqslant \mathfrak{p}_K(\tau^n_{K\sigma}) + \varrho g z_{\sigma,K}, \\ \eta_K(\mathfrak{s}_K(\tau^n_{\sigma,K})) & \text{otherwise.} \end{cases}$$

Two equations are required to determine $\tau^n_{\sigma,K}$ and $\tau^n_{\sigma,L}$. These equations are local conservation laws in the thin cells $I_{\sigma,K}$ and $I_{\sigma,L}$. Denote by $m_{\sigma,K}$ and $m_{\sigma,L}$ the Lebesgue measure of the thin cells $I_{\sigma,K}$ and $I_{\sigma,L}$ respectively ($m_{K\sigma} = m_\sigma\delta_{\rm B}$ for Cartesian grids as depicted in Figure 2), then $(\tau^n_{K\sigma}, \tau^n_{L\sigma})$ are determined by

$$\phi_K\frac{\mathfrak{s}_K(\tau^n_{\sigma,K}) - \mathfrak{s}_K(\tau^{n-1}_{\sigma,K})}{\Delta t}m_{\sigma,K} + F^n_\sigma - F^n_{K\sigma} = 0, \tag{3.7}$$

$$\phi_L\frac{\mathfrak{s}_L(\tau^n_{\sigma,L}) - \mathfrak{s}_L(\tau^{n-1}_{\sigma,L})}{\Delta t}m_{\sigma,L} - F^n_\sigma - F^n_{L\sigma} = 0, \tag{3.8}$$

where F^n_σ is the flux from $I_{\sigma,K}$ to $I_{\sigma,L}$ defined by

$$F^n_\sigma = \frac{m_\sigma}{\delta_{\rm B}}\lambda_\sigma\eta^n_\sigma\left(\mathfrak{p}_K(\tau^n_{\sigma,K}) - \mathfrak{p}_L(\tau^n_{\sigma,L}) + \varrho g(z_{\sigma,K} - z_{\sigma,L})\right), \tag{3.9}$$

with λ_σ given by (3.2) and

$$\eta^n_\sigma = \begin{cases} \eta_K(\mathfrak{s}_K(\tau^n_{\sigma,K})) & \text{if } \mathfrak{p}_K(\tau^n_{\sigma,K}) + \varrho g z_{\sigma,K} \geqslant \mathfrak{p}_L(\tau^n_{\sigma,L}) + \varrho g z_{\sigma,L}, \\ \eta_L(\mathfrak{s}_L(\tau^n_{\sigma,L})) & \text{otherwise.} \end{cases} \tag{3.10}$$

Assuming that $|F^n_\sigma| \leqslant Cm_\sigma$, then we deduce from (3.9) that $|\mathfrak{p}_K(\tau^n_{\sigma,K}) - \mathfrak{p}_L(\tau^n_{\sigma,L})| \leqslant C\delta_{\rm B}$, improving the pressure continuity with respect to Method A since $\delta_{\rm B} \ll d_\sigma$. On the other hand, summing (3.7) and (3.8) yields

$$|F^n_{K\sigma} + F^n_{L\sigma}| \leqslant C\frac{m_\sigma\delta_{\rm B}}{\Delta t} \underset{\delta_{\rm B}\to 0}{\longrightarrow} 0. \tag{3.11}$$

Note that even if (3.11) can be interpreted as a defect in the approximation of (1.4), Method B is still conservative since we keep track of this defect thanks to the discrete conservation laws (3.7)–(3.8). The volume m_K of the cell K is updated into

$$m_K \leftarrow m_K - \sum_{\sigma \in \mathscr{E}_K \cap \mathscr{E}_\Gamma} m_{\sigma,K} \tag{3.12}$$

in (2.9) for all K having interface edges.

3.3 Method C

This method takes inspiration from [11,12] and consists in adding only one thin cell, I_σ, of thickness $\delta_C \ll d_\sigma$, which overlaps the rock-type interface as shown in Figure 3. For $\sigma = K|L \in \mathscr{E}_\Gamma$, we denote by $m_{\sigma,K}$ and $m_{\sigma,L}$ the Lebesgue measures of $I_{\sigma,K} := I_\sigma \cap K$ and $I_{\sigma,L} := I_\sigma \cap L$ respectively. The

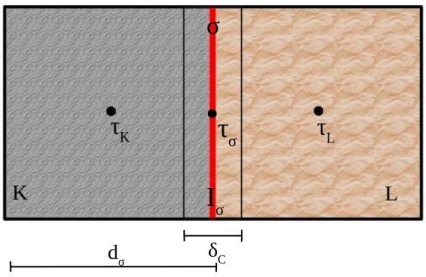

Figure 3: Method C: one extra thin cell I_σ overlaps the interface located between two rock types.

system is enriched with only one extra unknown τ_σ^n per face $\sigma \in \mathscr{E}_\Gamma$, in opposition to Method B where two additional unknowns where needed. The new cell I_σ is shared by two subcells $I_{\sigma,K}$ and $I_{\sigma,L}$ corresponding to different lithologies. To enforce one single pressure in the cell, we introduce a second parametrization and define monotone functions $\omega_{\sigma,K}, \omega_{\sigma,L}$ with $\omega'_{\sigma,K} + \omega'_{\sigma,L} > 0$ such that

$$\mathfrak{p}_K(\omega_{\sigma,K}(\tau)) = \mathfrak{p}_L(\omega_{\sigma,L}(\tau)), \qquad \forall \tau. \tag{3.13}$$

As for the parametrization $(\mathfrak{p}, \mathfrak{s})$ of the graph of $\mathcal{S}$, an infinite number of admissible $(\omega_{\sigma,K}, \omega_{\sigma,L})$ satisfying (3.13) can be built. We further investigate two choices.

The first possibility, named with exponent 1, consists in setting

$$\omega^1_{\sigma,K}(\tau) = \tau, \quad \omega^1_{\sigma,L}(\tau) = \mathfrak{p}_L^{-1} \circ \mathfrak{p}_K(\tau), \tag{3.14}$$

the orientation of the cell being such that $\omega^1_{\sigma,L}$ is concave (choose K and L such that $p_{b,K} > p_{b,L}$ or $\alpha_K > \alpha_L$ in the Brooks-Corey and van Genuchten settings described in Section 4.1.3 respectively), hence so does $\mathfrak{p}_L \circ \omega^1_{\sigma,L}$. As it appears on figures 4 and 5, the derivative of $\omega^1_{\sigma,L}$ might blow up (the value of κ_K defined by (2.8) is very large in practice).

Our second proposition, named with exponent 2, is tailored to maintain control on the derivatives of $\omega^2_{\sigma,K}$ and $\omega^2_{\sigma,L}$. To this end, keeping the same orientation $K|L$ of the interface σ, we set

$$\omega^2_{\sigma,K}(\tau) = \begin{cases} \mathfrak{p}_K^{-1} \circ \mathfrak{p}_L(\tau) & \text{if } \tau \leqslant \beta_L, \\ \tau + \beta_K - \beta_L & \text{if } \tau \geqslant \beta_L. \end{cases} \qquad \omega^2_{\sigma,L}(\tau) = \begin{cases} \tau & \text{if } \tau \leqslant \beta_L, \\ \mathfrak{p}_L^{-1} \circ \mathfrak{p}_K(\tau + \beta_K - \beta_L) & \text{if } \tau \geqslant \beta_L. \end{cases} \tag{3.15}$$

the parameters β_K and β_L are uniquely determined by the conditions

$$\mathfrak{p}_K(\beta_K) = \mathfrak{p}_L(\beta_L) \quad \text{and} \quad \mathfrak{p}'_K(\beta_K) = \mathfrak{p}'_L(\beta_L)$$

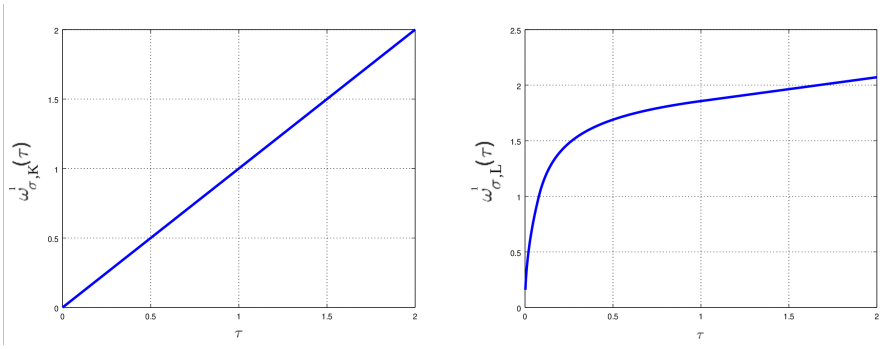

Figure 4: Behaviour of $\omega^1_{\sigma,K}(\cdot)$ and $\omega^1_{\sigma,L}(\cdot)$ functions using the Brooks and Corey model.

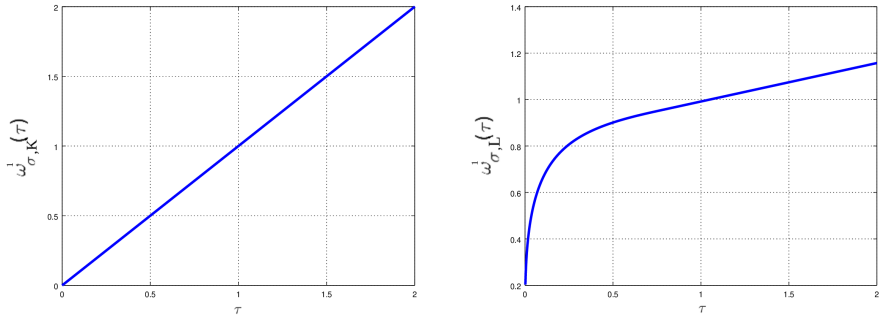

Figure 5: Behaviour of $\omega^1_{\sigma,K}(\cdot)$ and $\omega^1_{\sigma,L}(\cdot)$ functions using the van Genuchten Mualem model.

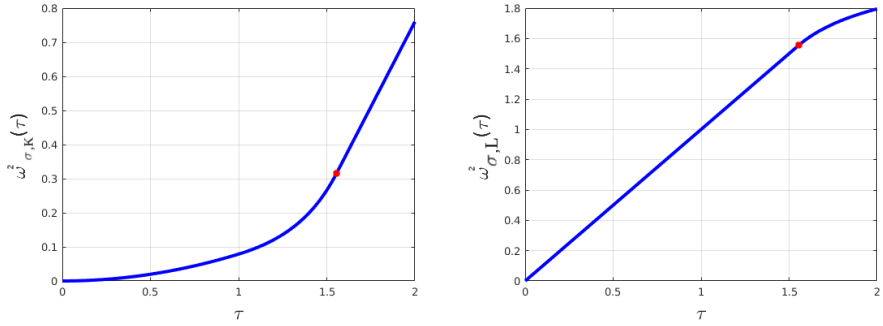

Figure 6: Behaviour of $\omega^2_{\sigma,K}(\cdot)$ and $\omega^2_{\sigma,L}(\cdot)$ functions using the Brooks and Corey model.

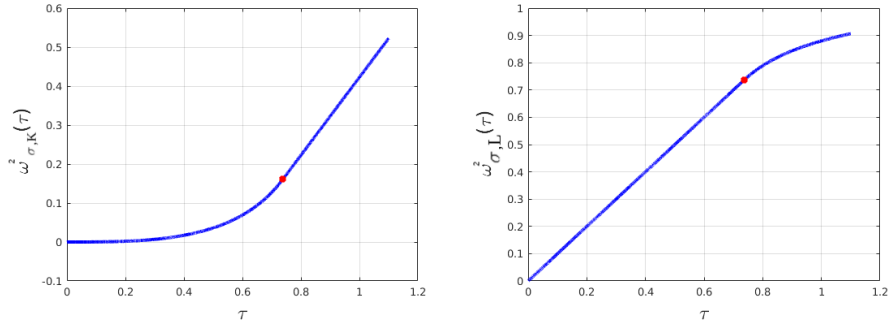

Figure 7: Behaviour of $\omega^2_{\sigma,K}(\cdot)$ and $\omega^2_{\sigma,L}(\cdot)$ functions using the van Genuchten Mualem model.

since $\mathfrak{p}_K, \mathfrak{p}_L$ are increasing and concave. This yields 1-Lipschitz continuous functions $\omega^2_{\sigma,K}$ and $\omega^2_{\sigma,L}$ as depicted on figures 6 and 7.

With a double parametrization $(\omega_{\sigma,K}, \omega_{\sigma,L})_{\sigma\in\mathscr{E}_\Gamma}$ at hand, Method C then consists in writing a discrete conservation law in I_σ. While in Method B, the sub-cells $I_{\sigma,K}$ and $I_{\sigma,L}$ had different pressures generating an in-between flux F^n_σ (3.9), here the two sub-cells share the same pressure

$$p^n_\sigma = \mathfrak{p}_K(\omega_{\sigma,K}(\tau^n_\sigma)) = \mathfrak{p}_L(\omega_{\sigma,L}(\tau^n_\sigma)), \qquad \sigma = K|L \in \mathscr{E}_\Gamma, \tag{3.16}$$

thanks to (3.13). The discrete volume conservation on I_σ then reads

$$\begin{aligned}\phi_K \frac{\mathfrak{s}_K(\omega_{\sigma,K}(\tau^n_\sigma)) - \mathfrak{s}_K(\omega_{\sigma,K}(\tau^{n-1}_\sigma))}{\Delta t} m_{\sigma,K} \\ + \phi_L \frac{\mathfrak{s}_L(\omega_{\sigma,L}(\tau^n_\sigma)) - \mathfrak{s}_L(\omega_{\sigma,L}(\tau^{n-1}_\sigma))}{\Delta t} m_{\sigma,L} - F^n_{K\sigma} - F^n_{L\sigma} = 0,\end{aligned} \tag{3.17}$$

where the fluxes $F^n_{K\sigma}$ and $F^n_{L\sigma}$ from K to I_σ and from L to I_σ are given by

$$F^n_{K\sigma} = \frac{m_\sigma}{d_{K,\sigma}} \lambda_K \eta^n_{\sigma,K} \left(\mathfrak{p}_K(\tau^n_K) - p^n_\sigma + \varrho g(z_K - z_\sigma)\right), \tag{3.18}$$

$$F^n_{L\sigma} = \frac{m_\sigma}{d_{L,\sigma}} \lambda_L \eta^n_{\sigma,L} \left(\mathfrak{p}_L(\tau^n_L) - p^n_\sigma + \varrho g(z_L - z_\sigma)\right), \tag{3.19}$$

with

$$\eta^n_{\sigma,K} = \begin{cases} \eta_K(\mathfrak{s}_K(\tau^n_K)) & \text{if } \mathfrak{p}_K(\tau^n_K) + \varrho g z_K \geqslant p^n_\sigma + \varrho g z_\sigma, \\ \eta_K(\mathfrak{s}_K(\omega_{\sigma,K}(\tau^n_\sigma))) & \text{otherwise.} \end{cases} \tag{3.20}$$

In (3.18)–(3.20), p^n_σ is given by (3.16), which should be thought as the discrete counterpart to the pressure continuity (1.5) across the interface. Concerning the continuity of the fluxes (1.4), it follows from (3.17) that

$$|F^n_{K\sigma} + F^n_{L\sigma}| \leqslant C \frac{m_\sigma \delta_{\mathrm{C}}}{\Delta t} \underset{\delta_{\mathrm{C}}\to 0}{\longrightarrow} 0, \tag{3.21}$$

meaning that (1.4) is recovered only asymptotically. Nevertheless, with δ_{C} small, flux continuity is captured in an accurate way. Moreover, as for Method B, Method C is locally conservative if one corrects the cell size m_K as prescribed by (3.12).

3.4 Method D

The last method we propose, referred as Method D, consists in enforcing both the pressure continuity and the flux continuity across the interface, at the price of one edge unknown τ^n_σ on each

$\sigma \in \mathscr{E}_\Gamma$ on the interface between different rocks. Such an approach has already been proposed for instance in [10, 11, 16, 21]. Letting δ_C tend to 0 in Method C (cf. Figure 8, and more precisely in (3.17)), one recovers the flux continuity

$$F^n_{K\sigma} + F^n_{L\sigma} = 0, \qquad \sigma = K|L \in \mathscr{E}_\Gamma, \tag{3.22}$$

with $F^n_{K\sigma}$ and $F^n_{L\sigma}$ respectively defined by (3.18) and (3.19), whereas pressure continuity is still ensured by (3.16). We propose then two numerical strategies, later referred as Methods D_1 and D_2 to solve the resulting nonlinear system.

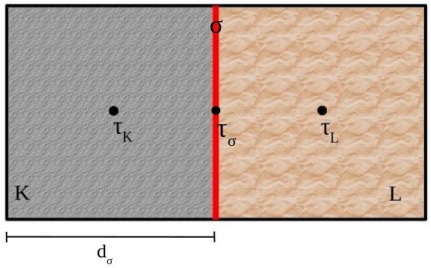

Figure 8: Method D: introduction of a face unknown τ^n_σ with no associated volume.

3.4.1 Method D_1: Schur complement based elimination of the face unknowns

With the rock-type face unknowns the obtained system is made of $\#\mathscr{T} + \#\mathcal{E}_\Gamma$ equations

$$\mathcal{F}(\boldsymbol{\tau}^n_{\mathscr{T}}, \boldsymbol{\tau}^n_{\mathcal{E}_\Gamma}) = \begin{pmatrix} \mathcal{F}_{\mathscr{T}}(\boldsymbol{\tau}^n_{\mathscr{T}}, \boldsymbol{\tau}^n_{\mathcal{E}_\Gamma}) \\ \mathcal{F}_{\mathcal{E}_\Gamma}(\boldsymbol{\tau}^n_{\mathscr{T}}, \boldsymbol{\tau}^n_{\mathcal{E}_\Gamma}) \end{pmatrix} = \mathbf{0} \tag{3.23}$$

where $\boldsymbol{\tau}^n_{\mathscr{T}} = (\tau^n_K)_{K\in\mathscr{T}}$, $\boldsymbol{\tau}^n_{\mathcal{E}_\Gamma} = (\tau^n_\sigma)_{\sigma\in\mathscr{E}_\Gamma}$, and where $\mathcal{F}_{\mathscr{T}}$ corresponds to the volume conservation laws (2.9) and $\mathcal{F}_{\mathcal{E}_\Gamma}$ to the flux conservation across the interfaces (3.22).

In what follows, we are interested in the resolution of the system (3.23) at a prescribed time step n. For notation convenience, the superscript n is dropped in this section. Denote by $\left(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}\right)_{\ell\geqslant 0}$ a sequence of approximation of $(\boldsymbol{\tau}_{\mathscr{T}}, \boldsymbol{\tau}_{\mathcal{E}_\Gamma})$ given by iterations of Newton's method. The Jacobian matrix of $\mathcal{F}$ at $\left(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}\right)$, $\ell \geqslant 0$ can be split into four blocks:

$$\mathbb{J}\mathcal{F}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}) = \left[\begin{array}{c|c} A^\ell & B^\ell \\ \hline C^\ell & D^\ell \end{array}\right]$$

where

$$A^\ell = \frac{\partial \mathcal{F}_{\mathscr{T}}}{\partial \boldsymbol{\tau}_{\mathscr{T}}}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}), \quad B^\ell = \frac{\partial \mathcal{F}_{\mathscr{T}}}{\partial \boldsymbol{\tau}^n_{\mathcal{E}_\Gamma}}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}), \quad C^\ell = \frac{\partial \mathcal{F}_{\mathcal{E}_\Gamma}}{\partial \boldsymbol{\tau}^n_{\mathscr{T}}}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}), \quad D^\ell = \frac{\partial \mathcal{F}_{\mathcal{E}_\Gamma}}{\partial \boldsymbol{\tau}^n_{\mathcal{E}_\Gamma}}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}).$$

Then the matrix D^ℓ is diagonal with negative diagonal entries because of the monotonicity of $F^n_{K\sigma}$, $F^n_{L\sigma}$ with respect to τ^n_σ that can be deduced from the monotonicity of $\mathfrak{p}_K$, $\mathfrak{p}_L$ and $\omega_{\sigma,K}$, $\omega_{\sigma,L}$. Therefore, D^ℓ can be inverted for free.

A Newton iteration to solve (3.23) then consists in computing an increment $\boldsymbol{\Delta}^\ell = \left(\boldsymbol{\Delta}^\ell_{\mathscr{T}}, \boldsymbol{\Delta}^\ell_{\mathcal{E}_\Gamma}\right)^T$ which is solution to

$$\mathbb{J}\mathcal{F}(\boldsymbol{\tau}^\ell_{\mathscr{T}})\boldsymbol{\Delta}^\ell = -\mathcal{F}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}) = -\begin{pmatrix} \mathcal{F}_{\mathscr{T}}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}) \\ \mathcal{F}_{\mathscr{E}_\Gamma}(\boldsymbol{\tau}^\ell_{\mathscr{T}}, \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}) \end{pmatrix} = -\begin{pmatrix} \mathcal{F}^\ell_{\mathscr{T}} \\ \mathcal{F}^\ell_{\mathscr{E}_\Gamma} \end{pmatrix},$$

or equivalently

$$(A^\ell - B^\ell \ \left(D^\ell\right)^{-1} \ C^\ell)\boldsymbol{\Delta}^\ell_{\mathscr{T}} = -\mathcal{F}^\ell_{\mathscr{T}} + B^\ell \ \left(D^\ell\right)^{-1} \mathcal{F}^\ell_{\mathcal{E}_\Gamma}, \tag{3.24}$$

and

$$\boldsymbol{\Delta}^\ell_{\mathcal{E}_\Gamma} = -(D^\ell)^{-1}\mathcal{F}^\ell_{\mathcal{E}_\Gamma} + (D^\ell)^{-1} \ C^\ell \ \boldsymbol{\Delta}^\ell_{\mathscr{T}}.$$

Then the unknowns are updated by $\boldsymbol{\tau}^{\ell+1} = \boldsymbol{\tau}^\ell + \boldsymbol{\Delta}^\ell$. The linear system (3.24) then consists in $\#\mathscr{T}$ equations.

3.4.2 Method D_2: face unknowns elimination thanks to a bisection method

We present here an alternative approach to solve the nonlinear system (3.23). The strategy consists here in computing increments of the cell unknowns via Newton's method and updating the face unknowns by solving exactly the flux conservation (3.22) on each interface and at each Newton iteration ℓ via the bisection method. The unknowns $\{\tau^\ell_\sigma\}_{\sigma\in\mathscr{E}_\Gamma}$ are here considered as functions of the neighbouring cell unknowns τ^ℓ_K and τ^ℓ_L for $\sigma = K|L \in \mathscr{E}_\Gamma$ (we still drop the time index n). Therefore, the dependency of the interface fluxes w.r.t. the cell unknowns expresses as

$$\frac{\partial F_{K,\sigma}(\tau^\ell_K, \tau_\sigma(\tau^\ell_K, \tau^\ell_L))}{\partial \tau_K} = \frac{\partial F_{K,\sigma}}{\partial \tau_K} + \frac{\partial F_{K,\sigma}}{\partial \tau_\sigma}\frac{\partial \tau_\sigma}{\partial \tau_K}.$$

The Jacobian matrix associated to the reduced system then writes

$$\mathbb{J}\mathcal{F}(\boldsymbol{\tau}^\ell_{\mathscr{T}}) = A^\ell - B^\ell E^\ell \quad \text{where} \quad E^\ell = \frac{\partial \boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}}{\partial \boldsymbol{\tau}_{\mathscr{T}}} = \left(D^{-1}\right)^\ell C^\ell.$$

The increment of the ℓ^{th} Newton iteration thus solves (3.24) with

$$\mathcal{F}_{\mathcal{E}_\Gamma}(\boldsymbol{\tau}^\ell_{\mathcal{E}_\Gamma}) = \mathbf{0}. \tag{3.25}$$

We detail now how we solve the flux-conservation subsystem (3.25) knowing $\left(\boldsymbol{\tau}^\ell_K\right)_{K\in\mathscr{T}}$. For each $\sigma \in \mathscr{E}_\Gamma$ and for all outer Newton loop iteration ℓ, we build a sequence $\left(\vartheta^{\ell,k}_\sigma\right)_{k\geqslant 0} = \left(p^{\ell,k}_\sigma + \rho g z_\sigma\right)_{k\geqslant 0}$ approximating the interface hydraulic head at the interface. More precisely, define

$$\vartheta^\ell_K = \mathfrak{p}_K(\tau^\ell_K) + \rho g z_K, \qquad \vartheta^\ell_L = \mathfrak{p}_L(\tau^\ell_L) + \rho g z_L, \qquad \sigma = K|L \in \mathscr{E}_\Gamma,$$

and

$$F^\ell_{K,\sigma}(\vartheta_\sigma) = \frac{m_\sigma}{d_{K,\sigma}}\lambda_K \eta^\ell_{\sigma,K}(\vartheta_\sigma)(\vartheta^\ell_K - \vartheta_\sigma), \quad F^\ell_{L,\sigma}(\vartheta_\sigma) = \frac{m_\sigma}{d_{L,\sigma}}\lambda_L \eta^\ell_{\sigma,L}(\vartheta_\sigma)(\vartheta^\ell_L - \vartheta_\sigma),$$

with

$$\eta^\ell_{\sigma,K}(\vartheta_\sigma) = \begin{cases} \eta_K \circ \mathfrak{s}_K \circ \mathfrak{p}_K^{-1}(\vartheta^\ell_K - \rho g z_K) & \text{if } \vartheta_\sigma \leqslant \vartheta^\ell_K, \\ \eta_K \circ \mathfrak{s}_K \circ \mathfrak{p}_K^{-1}(\vartheta_\sigma - \rho g z_\sigma) & \text{otherwise} \end{cases}$$

and a similar definition for $\eta^\ell_{\sigma,L}$, then one readily checks that $F^\ell_{K,\sigma}$ is decreasing w.r.t. ϑ_σ, and that

$$F^\ell_{K,\sigma}\left(\min\left(\vartheta^\ell_K, \vartheta^\ell_L\right)\right) \geqslant 0 \geqslant F^\ell_{K,\sigma}\left(\max\left(\vartheta^\ell_K, \vartheta^\ell_L\right)\right).$$

Therefore, the continuous and decreasing function $G^\ell_\sigma : \mathbb{R} \to \mathbb{R}$ defined by

$$G^\ell_\sigma(\vartheta_\sigma) = \frac{F^\ell_{K,\sigma}(\vartheta_\sigma) + F^\ell_{L,\sigma}(\vartheta_\sigma)}{a_\sigma \mu^{-1}(\lambda_K + \lambda_L)(|\vartheta^\ell_K| + |\vartheta^\ell_L|)}$$

vanishes at some $\vartheta^\ell_\sigma \in \left[\min\left(\vartheta^\ell_K, \vartheta^\ell_L\right), \max\left(\vartheta^\ell_K, \vartheta^\ell_L\right)\right]$, from which one deduces

$$\tau^\ell_\sigma = \omega^{-1}_{\sigma,K} \circ \mathfrak{p}_K^{-1}(\vartheta^\ell_\sigma - \rho g z_\sigma) = \omega^{-1}_{\sigma,L} \circ \mathfrak{p}_L^{-1}(\vartheta^\ell_\sigma - \rho g z_\sigma)$$

with $(\omega_{\sigma,K}, \omega_{\sigma,L})$ being a double parametrization as introduced in Section 3.3. Then we solve the nonlinear equation $G_\sigma^\ell(\vartheta_\sigma) = 0$ thanks to the classical bisection method, stopping the iterations over k when either $\left|G_\sigma^\ell(\vartheta_\sigma^k)\right| < \epsilon_{\text{bis}}$ or $\frac{|\vartheta_0^k - \vartheta_1^k|}{\min(|\vartheta_0^k|, |\vartheta_1^k|)} < \gamma$ with $\epsilon_{\text{bis}} = 10^{-16}$ and $\gamma = 10^{-15}$ in our simulations. We can finally remark that Method D_2 does not depend on the choice of the double parametrization.

4 Numerical results

We now present numerical results obtained for different test cases. In all cases, we consider a two-dimensional layered domain $\Omega = [0\text{m}, 5\text{m}] \times [-3\text{m}, 0\text{m}]$ made up of two rock types denoted by RT0 and RT1 respectively, RT0 being less permeable than RT1. The domain Ω is partitioned into three connected subdomains: $\Omega_1 = [1\text{m}, 4\text{m}] \times [-1\text{m}, 0\text{m}]$, $\Omega_2 = [0\text{m}, 5\text{m}] \times [-3\text{m}, -2\text{m}]$ and $\Omega_3 = \Omega \setminus (\Omega_1 \cup \Omega_2)$, as depicted in Figure 9.

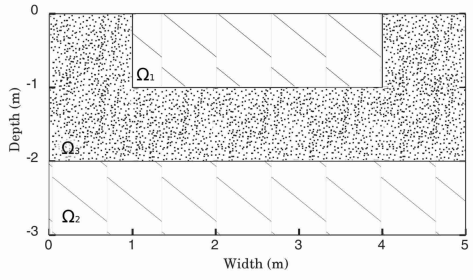

Figure 9: Simulation domain $\Omega = [0\text{m}, 5\text{m}] \times [-3\text{m}, 0\text{m}]$.

4.1 Description of the test cases

Both filling and drainage configurations are considered along with the two classical Brooks-Corey and van Genuchten-Mualem hydraulic models. These analytical models are first used in a setting where the pressure-saturation relationship and its inverse have moderate derivatives (non-steep cases). We then only consider the Brooks-Corey model and coefficients where the pressure-saturation dependence has sharp variations (steep cases).

4.1.1 Filling case

This test case has already been considered in [18, 30, 33, 39]. The rock-type repartition is reported in Figure 10. Starting from an initially dry domain Ω, where the initial capillary pressure is set to $-47.088 \cdot 10^5$Pa, water flows from a portion $\Gamma_N = \{(x, y) \,|\, x \in [1\text{m}, 4\text{m}], y = 0\text{m}\}$ of the top boundary at a constant rate of 0.5m/day. A no-flow boundary condition is applied elsewhere. The simulation stops after 1 day.

Water flows according to the following dynamics. It starts invading the dry porous space in Ω_1. When it reaches the interface with Ω_3, capillary forces creates a suction force on water from Ω_1 to Ω_3. But, on the other hand, the low permeability value in RT1 is set against this water flow through Ω_3. The simulation ends before water reaches the bottom part corresponding to Ω_2.

4.1.2 Drainage case

This test case is designed as a two-dimensional extension of a one-dimensional test case proposed by [38] and addressed in [18, 39]. We simulate a vertical drainage starting from saturated initial and

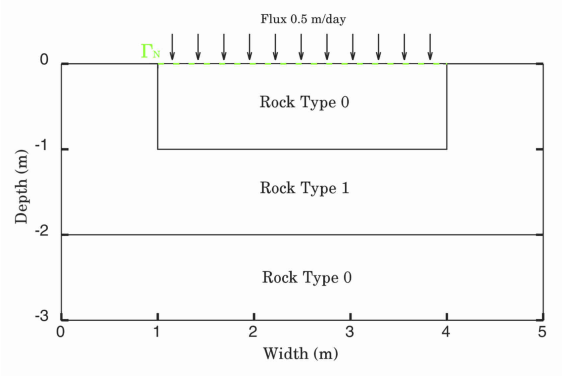

Figure 10: Boundary condition for the filling case

boundary conditions during $105 \cdot 10^4$ s. The initial pressure is hydro-static, that is $p^0(z) = -\rho g z$. A Dirichlet boundary condition $p_D = 0$ Pa is imposed on the bottom boundary, $\Gamma_D = \{(x, y) \,|\, x \in [0\text{m}, 5\text{m}], y = -3\text{m}\}$. The rock-type distribution of Ω is shown in Figure 11 along with the bottom boundary condition.

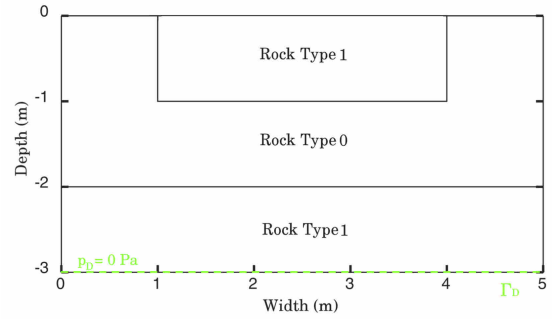

Figure 11: Boundary condition for the drainage test

Note that rock types RT0 and RT1 are here reversed compared to the previous case. Thus, at the top interface between Ω_1 and Ω_3, capillarity acts in opposition to gravity and to the evolution of the system towards a dryer configuration. The interface between Ω_2 and Ω_3 acts in the opposite way: both capillarity and gravity contribute to the drainage of the RT0 subdomain.

4.1.3 Hydraulic models

For two-phase problems, water saturation and capillary pressure are linked through the relation $s = \mathcal{S}(p)$. Here $\mathcal{S} : \mathbb{R} \to [0, 1]$ is nondecreasing. It satisfies $\mathcal{S}(p) = 1 - s_{\text{rn}}$ if $p \geqslant p_b$ and $\mathcal{S}(p) \to s_{\text{rw}}$ as $p \to -\infty$, with s_{rw} (resp. s_{rn}) the residual wetting (resp. non-wetting) saturation. In the following, to model the two-phase flow characteristics for both rock types,

- either the Brooks-Corey [15] model:

$$s = \mathcal{S}(p) = \begin{cases} s_{\rm rw} + (1 - s_{\rm rn} - s_{\rm rw}) \left(\frac{p}{p_b}\right)^{-n} & \text{if } p \leqslant p_b, \\ 1 - s_{\rm rn} & \text{if } p > p_b, \end{cases}$$
$$k_r(s) = s_{\rm eff}^{3+\frac{2}{n}}, \qquad s_{\rm eff} = \frac{s - s_{\rm rw}}{1 - s_{\rm rn} - s_{\rm rw}}; \tag{4.1}$$

- or the van Genuchten-Mualem [47] model:

$$s = \mathcal{S}(p) = \begin{cases} s_{\rm rw} + (1 - s_{\rm rn} - s_{\rm rw}) \left[1 + \left|\frac{\alpha}{\rho g} p\right|^n\right]^{-m} & \text{if } p \leqslant 0, \\ 1 - s_{\rm rn} & \text{if } p > 0, \end{cases}$$
$$k_r(s) = s_{\rm eff}^{\frac{1}{2}} \{1 - [1 - s_{\rm eff}^{\frac{1}{m}}]^m\}^2, \qquad s_{\rm eff} = \frac{s - s_{\rm rw}}{1 - s_{\rm rn} - s_{\rm rw}}, \qquad m = 1 - \frac{1}{n} \tag{4.2}$$

are used. In both models, we have denoted by $k_r(\cdot)$ the relative permeability which, with the water viscosity $\mu = 10^{-3}$ Pa·s, defines the water mobility thanks to $\eta(\cdot) = k_r(\cdot)/\mu$. The parameters used for both rock types are given in Table 1 for cases using the Brooks-Corey model and in Table 2 for the other ones. These parameters have been chosen in such a way that water is more likely to be in RT1 than in RT0: indeed, at a fixed pressure, the water saturation is higher in RT1 than in RT0. This can be observed on the plots of the capillary-pressure functions depicted in Figures 12–14. On these figures, the relative permeability functions are also shown. Let us, in particular, remark the non-Lipschitz character of the relative permeability in the van Genuchten-Mualem case. Thus, in order to avoid infinite values for the derivative of $k_r(s)$ when $s \to 1 - s_{\rm rn}$, we approximate it for $s \in [s_{lim}, 1 - s_{\rm rn}]$ using a second degree polynomial $\widetilde{k_r}(s)$. This polynomial satisfies the following conditions: $k_r(s_{lim}) = \widetilde{k_r}(s_{lim})$, $\widetilde{k_r}'(s_{lim}) = k_r'(s_{lim})$ and $\widetilde{k_r}(1 - s_{\rm rn}) = 1$ where s_{lim} is chosen so that $s_{\rm eff} = 0.998$.

	$1 - s_{rn}$	s_{rw}	p_b[Pa]	n	λ[m^2]	ϕ
RT0	1.0	0.1	$-1.4708 \cdot 10^3$	3.0	10^{-11}	0.35
RT1	1.0	0.2	$-3.4301 \cdot 10^3$	1.5	10^{-13}	0.35

Table 1: Parameters used for the Brooks-Corey model

	$1 - s_{\rm rn}$	$s_{\rm rw}$	n	λ [m^2]	α [m^{-1}]	ϕ
RT0 (Sand)	1.0	0.0782	2.239	$6.3812 \cdot 10^{-12}$	2.8	0.3658
RT1 (Clay)	1.0	0.2262	1.3954	$1.5461 \cdot 10^{-13}$	1.04	0.4686

Table 2: Parameters used for the van Genuchten-Mualem model

4.2 Comparison of the results in non-steep cases

We now analyze the results obtained on the test cases which were previously introduced. We use uniform time discretizations. The time step Δt depends on the test case and is reported in Table 5 together with the others numerical parameters used for these simulations. The detailed results

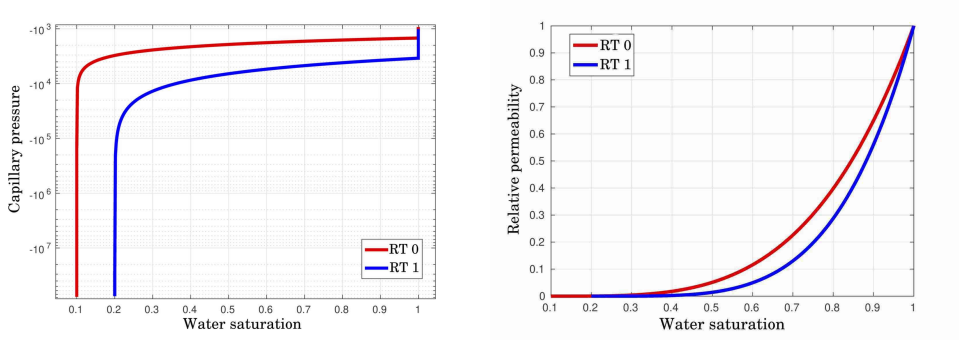

Figure 12: Capillary pressure and relative permeability curves for the Brooks-Corey model

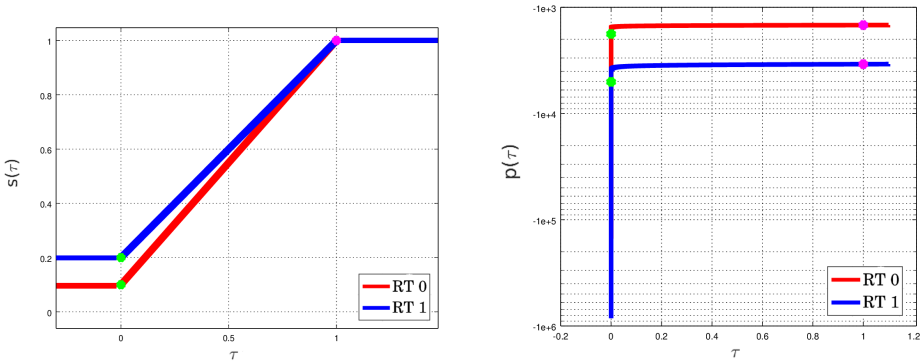

Figure 13: Parametrized saturation and pressure functions using the Brooks and Corey model and parameters of Table 1. The green dot indicates the value for $\tau = \tau_*$ and the magenta one $\tau = \tau_s$.

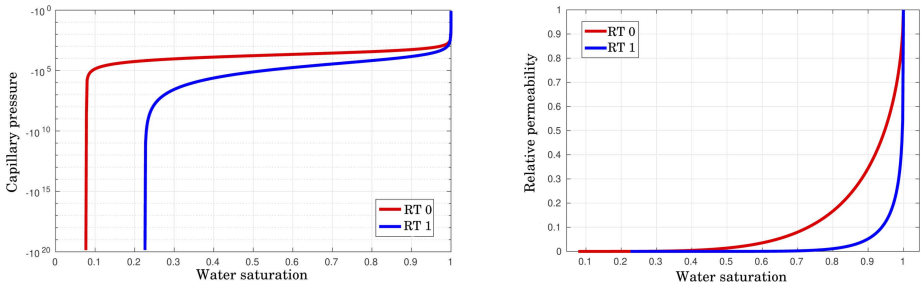

Figure 14: Capillary pressure and relative permeability curves for the van Genuchten-Mualem model

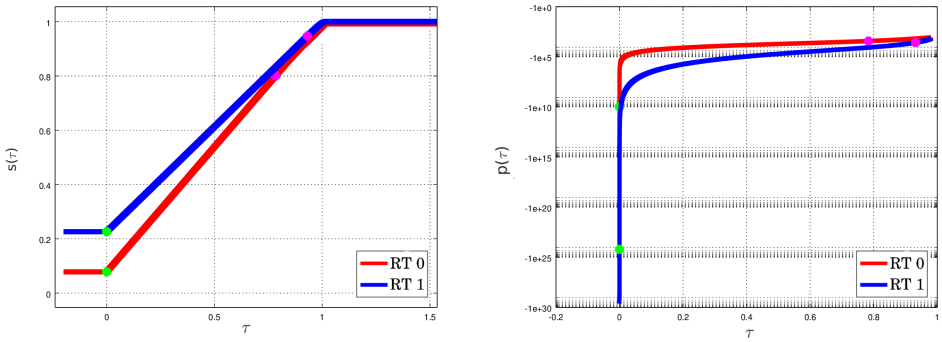

Figure 15: Parametrized saturation and pressure functions using the van Genuchten Mualem model using parameters of Table 2. The green dot indicate the value for $\tau = \tau_*$ and the magenta one $\tau = \tau_s$.

for the different cases, methods and meshes along with figures of the solutions are reported in Appendix A.

In Table 3 we present a brief classification of the proposed methods based on both robustness (R) and accuracy (A) criteria. For each criterion, the colour choice corresponds to the following glossary: green for good, orange for average, red for bad. Regarding the robustness, a non-convergent method is classified as red; orange is used if a method faces many times difficulties during Newton's resolution (maximal number of iterations reached, a much larger number of total iterations in comparison to other methods...); the green label is used in other cases. Thus, a method, having a relative error in the same order as the best performing one, is tagged as green; if an error has one (resp. several) order(s) of magnitude more than the best performing one, the label of the corresponding method is taken as orange (resp. red).

Let us discuss each test case in details, starting with the filling test case simulated with the Brooks and Corey model. Table 6 shows that Method A has the smallest saturation relative error with the coarsest mesh. Subsequent refinements then enable to reduce the error related to methods B,C,D at a higher convergence rate, leading to errors on the finest grid that are smaller than the one obtained with the classical scheme A. All the methods face difficulties in the Brooks-Corey filling case for the 3^{rd} mesh and the first time step. This is due to our non-optimal choice of a uniform time discretization. A simple time step adaptation strategy similar to the one used in the steep case would fix this issue.

Keeping the Brooks and Corey model, if we now analyze the results reported in Table 7 for the drainage case, we notice that methods B, C and D always have a smaller error than method A that converges again at a slower rate. Concerning Method C, it behaves as Method A in terms of accuracy and is fairly cheaper in terms of iterations with respect to this one.

We now consider the results obtained with the van Genuchten Mualem model. Table 8 summarizes the results obtained with the filling case. We can notice that all methods have approximately the same errors and convergence rates. Regarding Newton's cost, the conclusions are similar to the ones made for the Brooks and Corey tests.

In the drainage case (see Table 9), methods B and C turn out to be more precise and to converge faster that Method A. On the other hand, Methods B and C require more Newton iterations than Method A and D that almost have the same iterations' cost. Moreover we observe that all methods require an important maximum number of iterations to converge which is greater than 50 (reaching the number of 100 iteration for the finer meshes.)

Throughout all these non-steep tests we can also notice that the number of Newton iterations to reach convergence with method B is larger than for the other ones.

Let us now make one last comment on the results obtained using the two proposed double

	Method A		Method B		Method C		Method D_1		Method D_2	
	R	A	R	A	R	A	R	A	R	A
Brooks and Corey Filling case										
Brooks and Corey Drainage case										
van Genuchten Mualem Filling case										
van Genuchten Mualem Drainage case										

Table 3: Summary of methods' robustness (R) and accuracy (A) classification. Method D_1 and Method D_2 denote Method D with Schur complement and bisection method respectively. Color legend: green= good, orange= passing, red=bad.

parametrizations (see Eq. (3.14)-(3.15)) in Method D: they provide the same solutions with the same accuracy in all tests. They only differ in terms of Newton iterations which slightly vary from one parametrization to the other one according to the test case.

4.3 Tests with Brooks-Corey model and steep capillary-pressure curves

The aim of this section is to evaluate the robustness of Newton's algorithm when used with the four previous methods and steep capillary-pressure curves. We use the same filling and drainage cases with Brooks-Corey model as in the previous section. We here only change the value of the parameter n which is now equal to 120 for rock-type RT0 and 60 for RT1, making the problem (1.1)–(1.8) close to a strongly-degenerate parabolic case. The corresponding capillary pressure curves are represented in Figure 16. The time evolution for these tests is adaptive:

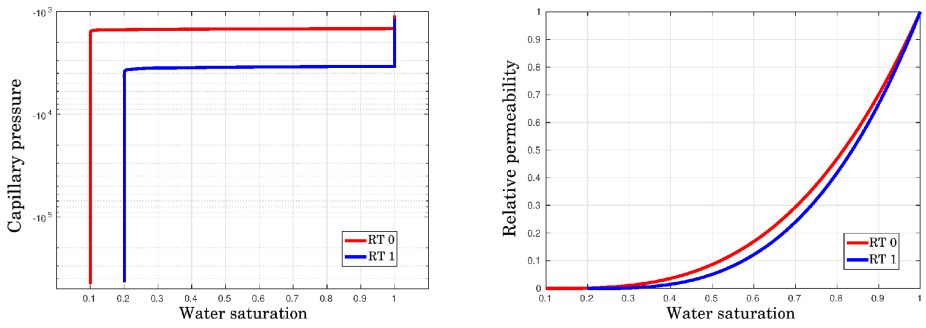

Figure 16: Steep cases: capillary-pressure and relative permeability curves for the Brooks-Corey model

- **Filling case**
  Minimal, maximal and initial time steps are such that $\Delta t_{min} = 10^{-6}s$, $\Delta t_{max} = 10^{4}s$, $\Delta t^0 = 10^{-6}s$ and, for $n \geqslant 0$,

$$\Delta t^{n+1} = \min(\Delta t_{max}, 1.2\Delta t^n)$$

in case of Newton's convergence or

$$\Delta t^{n+1} = \max(\Delta t_{min}, \Delta t^n/2)$$

in the absence of convergence. In the latter case, for $\Delta t = \Delta t_{min}$, the simulation stops. $N_{max} = 30$ is taken as maximal number of Newton's iterations.

- **Drainage case**
  Minimal, maximal and initial time steps are such that $\Delta t_{min} = 1s$, $\Delta t_{max} = 10^5 s$, $\Delta t^0 = 1s$ and, for $n \geqslant 0$,

$$\Delta t^{n+1} = \begin{cases} 2.5\Delta t^n & \text{if } \Delta t < 500s, \\ \min(\Delta t_{max}, 1.2\Delta t^n) & \text{otherwise,} \end{cases}$$

  in case of a successful time step, or

$$\Delta t^{n+1} = \begin{cases} \max(\Delta t_{min}, \Delta t^n/5) & \text{if } \Delta t < 500s, \\ \Delta t^n/2 & \text{otherwise,} \end{cases}$$

  in the absence of convergence with $N_{max} = 30$ iterations. If Newton does not converge for $\Delta t = \Delta t_{min}$ the simulation stops.

4.3.1 Comparison of the results

Table 4 shows that only methods A, B, C^2, D_1^2 and D_2 converge for all test cases. Here and hereafter, the exponents 1 or 2 refer to the choice of the double parametrization presented in Section 3.3. Figure 23 reports on the evolution of the cumulated number of Newton iterations for the filling case with the 50×30 cells mesh. Apart from method D which faces difficulties at the beginning, all curves evolve in the same way. These conclusions remain valid for the 400×240 cells mesh with an exception for method B whose number of iterations increased as it can be observed in Figure 24.
Figures 26 and 27 show the results obtained on the drainage case with meshes of resolutions 50×30 and 400×240 respectively. In both cases, the methods C^1 and D_1^1 face more difficulties to converge than the other ones around time $t = 348500s$. It corresponds to the moment at which the cells line in Ω_2 below the interface between Ω_3 and Ω_2 starts to empty. Note that the number of Newton iterations also increases for method B at that particular time on the finer mesh too. Method C^2 also encounters difficulties on the coarser mesh but at a earlier time (when the cells line in Ω_1 below the interface between Ω_1 and Ω_3 starts to empty) and to a lesser extent. On the whole, the results on this last case show a higher degree of robustness for methods A and the second proposition of the double parameterisation which has been designed with the aim of controlling and bounding its derivatives, as it can be seen in Figures 17.

4.4 Overall methods' evaluation

Using this glossary and the results obtained in the steep and non-steep cases, we proceed, in the following of this section, to a general evaluation of the five studied methods.

Let us start with Method A. In this approach, rock-type interface faces are treated like classical inner faces and the pressure continuity on these interfaces is not enforced. Nevertheless, if the simulation is performed on a sufficiently refined mesh, a good approximation of this condition can be obtained. In the previous tests, and in particular in the steep ones (see Section 4.3), this method turns out to be very robust. On fine meshes its accuracy is, in general, close to the ones of other methods. In the filling case with the Brooks and Corey model (see Table 6), this method is even

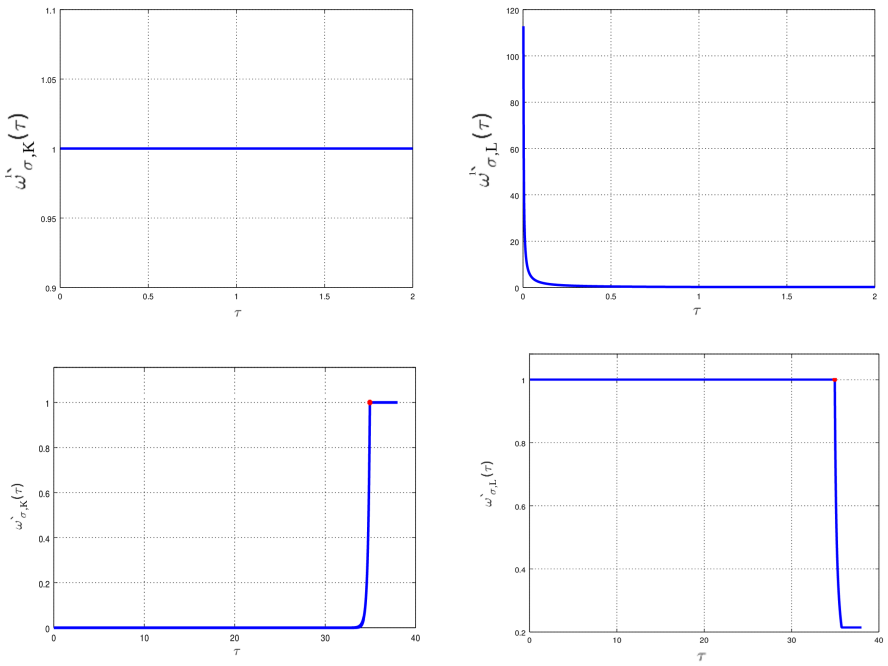

Figure 17: Comparison between $\omega'^1_{\sigma,K}(\cdot), \omega'^1_{\sigma,L}(\cdot)$ (above) and $\omega'^2_{\sigma,K}(\cdot), \omega'^2_{\sigma,L}(\cdot)$ (below) when using steep capillary pressure curves.

		Method A	Method B	Method C		Method D_1		Method D_2	
	Mesh			dp_1	dp_2	dp_1	dp_2	dp_1	dp_2
Filling case	50×30								
	400×240								
Drainage case	50×30								
	400×240								

Table 4: Summary of methods' robustness classification for steep tests. Method D_1 and Method D_2 denote Method D with Schur complement and bisection method respectively. Color legend: green= good, orange= passing, red=not converge.

the most accurate one on coarse meshes. A noticeable drawback of this method is the loss of the linear convergence rate when used with the Brooks and Corey model (see Tables 6–7).

Method B is the first approach we propose with a specific treatment for the rock-type interfaces, which only entails moderate changes in terms of implementation compared to method A. We here just add two thin cells around the rock-type interfaces and neglect, for these new cells, the fluxes through the faces with small measures. It features a rather good robustness since it also converges in the steep cases (see Section 4.3). For non-steep cases, it always provides a rather accurate solution and good robustness. Compared to method A, the linear convergence rate is recovered at the price of about 10% extra Newton iterations.

Method C is the first method which strongly enforces the pressure continuity on the rock-type interfaces. Here, the interfaces are thickened in thin cells and the pressure continuity is ensured by introducing a second parametrization (3.13) for which we propose two different forms detailed in Equations (3.14)-(3.15). This parametrization should be calculated beforehand and depends on the chosen petro-physical model. Thus, this method involves more changes for its implementation with respect to the previous one. In non-steep simulations, the two proposed double parametrizations provide the same solutions with the same accuracy with just a slight difference in the required number of iterations. Moreover it behaves as Method B in all non-steep simulations in terms of error. In the steep tests a remarkable difference of performance between the use of the two proposed parametrizations for the pressure continuity at interfaces arises: the first proposition of parametrization converges only in drainage case while the second one always converges showing, generally, a competitive robustness.

The last studied method Method D guarantees the flux conservation between all cells of the initial meshes and pressure continuity at rock-type interfaces. As for Method C, it also uses a double parametrization (3.13). In the non-steep cases, the application of the two proposed second parametrizations, as in Method C, provides the same solutions with the same accuracy with just a slight difference in the required number of iterations, as already remarked in Section 4.2. Moreover, Methods D_1 and D_2 show in all tests a good robustness and the same rather good accuracy: in drainage cases their accuracy is fairly better than Method A when using the Brooks and Corey model and, when employing the van Genuchten Mualem model, they show a relative error almost halved with respect to the one of Method A. In filling cases they have almost the same accuracy as Method A. Methods D_1 and D_2 always recover a first-order convergence except for the drainage test case with the van Genuchten Mualem model in which the convergence rate is slightly degraded. The fact that Methods D_1 and D_2 show the same accuracy is not surprising: the only difference between the two methods is how we solve the system that, for its part, does not change. In steep tests both double parametrizations employed in Method D_1 and D_2 make the simulation converge, apart for the filling test case in which Method D_1^1 fails, just showing in some cases a difference of behaviour in terms of robustness as detailed in Section 4.3.1.

Finally we can conclude that if we want to perform simulation for test cases with steep pressure curves, we can choose between Method A or B or, if one does not mind making larger code changes, methods C^2, D_1^2 or D_2 can also be used. If it is not the case, for coarse meshes, Method A ensures a good robustness and an accurate solution without any particular treatment for interfaces. For more refined meshes, even if its accuracy is slightly lower or comparable -it depends on the specific test case- to that of Method D, Method B is easier to implement and the least intrusive with respect to methods introducing a treatment for interfaces. So, if the choice is based on an accuracy criterion actually Method B, C and D are almost equivalent; in terms of ease of implementation the best choice is Method B.

A Figures and data related to the non-steep cases

	Δt	τ_*	ϵ	ϵ_{bis}	γ	δ_B	δ_C
Filling - Brooks and Corey	500	10^{-10}	10^{-12}	10^{-16}	10^{-15}	10^{-6}	$2 \cdot 10^{-6}$
Filling - van Genuchten Mualem	500	10^{-8}	10^{-12}	10^{-16}	10^{-15}	10^{-6}	$2 \cdot 10^{-6}$
Drainage - Brooks and Corey	1000	10^{-10}	10^{-12}	10^{-16}	10^{-15}	10^{-6}	$2 \cdot 10^{-6}$
Drainage - van Genuchten Mualem	1000	10^{-8}	10^{-12}	10^{-16}	10^{-15}	10^{-6}	$2 \cdot 10^{-6}$

Table 5: Numerical parameters used in the examples

A.1 Filling case using Brooks and Corey model

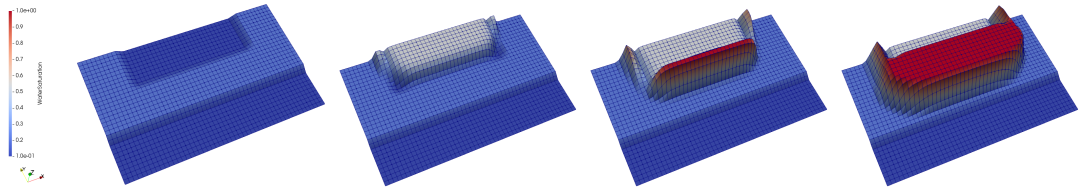

Figure 18: Evolution of the saturation profile for $t \in \{0s, 21.5 \cdot 10^3 s, 41.5 \cdot 10^3 s, 86.4 \cdot 10^3 s\}$ for the non-steep filling case, using Brooks and Corey model, Method B and the 50×30 cells mesh.

Method A	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	9.60719e-2	8.08028e-2	6.41616e-2	5.18869e-2
Rate of convergence	–	0.25	0.333	0.306
♯ total iterations	647	777	1074	1236
♯ avg iterations	3	4	6	7
♯ max iterations	18	21	168	32
Method B	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.43731e-1	1.0421e-1	6.3325e-2	2.76736e-2
Rate of convergence	–	0.464	0.719	1.194
♯ total iterations	835	959	1279	1428
♯ avg iterations	4	5	7	8
♯ max iterations	19	21	168	38
Method C	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.46706e-1	1.06227e-1	6.45985e-2	2.84733e-2
Rate of convergence	–	0.465	0.7156	1.182
♯ total iterations	690	796(794)	1106(1102)	1253(1247)
♯ avg iterations	3	4	6	7
♯ max iterations	20(18)	21	168	29
Method D_1	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.7701e-1	1.26129e-1	7.75469e-2	3.66345e-2
Rate of convergence	–	0.489	0.702	1.082
♯ total iterations	620(634)	721(734)	1001(1018)	1140(1146)
♯ avg iterations	3	4	5	6
♯ max iterations	17	20	155	32
Method D_2	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.7701e-1	1.26129e-1	7.75469e-2	3.66345e-2
Rate of convergence	–	0.489	0.702	1.082
♯ total iterations	590	714	999	1140
♯ avg iterations	3	4	5	6
♯ max iterations	17	20	155	32
♯ avg it. bisection per face	16	15	15	14

Table 6: Results for the non-steep filling case using Brooks and Corey model. For methods C and D_1, we specify within parentheses the number of iterations corresponding to the second choice of double parametrization when it differs from the one obtained with the first choice, which is reported without parentheses.

A.2 Drainage case using Brooks and Corey model

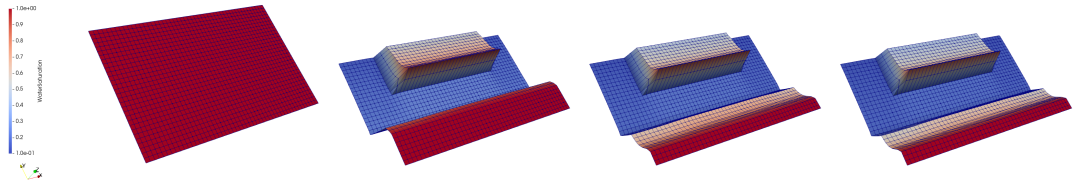

Figure 19: Evolution of the saturation profile for $t \in \{0s, 35 \cdot 10^4 s, 70 \cdot 10^4 s, 105 \cdot 10^4 s\}$ for the non-steep drainage case, using Brooks and Corey model, Method B and the 50×30 cells mesh.

Method A	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	4.48867e-2	2.60531e-2	1.64213e-2	1.110698e-2
Rate of convergence	–	0.785	0.666	0.564
♯ total iterations	2598	2848	3258	3819
♯ avg iterations	2	2	3	3
♯ max iterations	21	24	29	32
Method B	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.80469e-2	9.9613e-3	4.83626e-3	1.77811e-3
Rate of convergence	–	0.857	1.042	1.443
♯ total iterations	2845	3056	3448	3918
♯ avg iterations	2	2	3	3
♯ max iterations	20	24	20	32
Method C	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.81634e-2	1.00638e-2	4.92295e-3	1.84945e-3
Rate of convergence	–	0.851	1.032	1.412
♯ total iterations	2659(2653)	2893(2887)	3304(3298)	3804(3798)
♯ avg iterations	2	2	3	3
♯ max iterations	20(21)	24	28(29)	32
Method D_1	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	3.03634e-2	1.62917e-2	8.64114e-3	4.18359e-3
Rate of convergence	–	0.898	0.915	1.046
♯ total iterations	2659(2665)	2919(2905)	3329(3324)	3863(3861)
♯ avg iterations	2	2	3	3
♯ max iterations	21(20)	24	29(30)	32
Method D_2	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	3.03634e-2	1.62917e-2	8.64114e-3	4.18359e-3
Rate of convergence	–	0.898	0.915	1.046
♯ total iterations	2614	2894	3320	3856
♯ avg iterations	2	2	3	3
♯ max iterations	21	24	29	32
♯ avg it. bisection per face	35	34	33	32

Table 7: Results for the non-steep drainage case using Brooks and Corey model. For methods C and D_1, we specify within parentheses the number of iterations corresponding to the second choice of double parametrization when it differs from the one obtained with the first choice, which is reported without parentheses.

A.3 Filling case using van Genuchten Mualem model

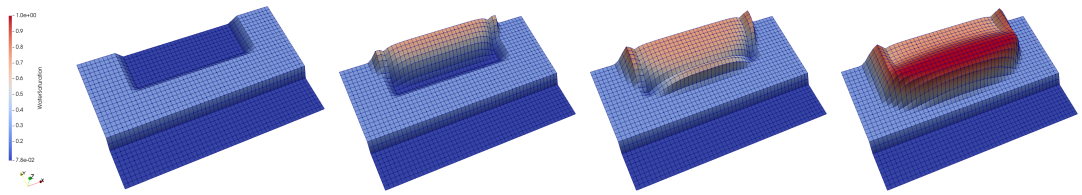

Figure 20: Evolution of the saturation profile for $t \in \{0s, 21.5 \cdot 10^3 s, 41.5 \cdot 10^3 s, 86.4 \cdot 10^3 s\}$ for the non-steep filling case using Van Genuchten model, Method B and the 50×30 cells mesh.

Method A	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.05534e-1	7.48124e-2	4.55216e-2	2.1125e-2
Rate of convergence	–	0.496	0.717	1.108
♯ total iterations	575	667	782	930
♯ avg iterations	3	3	4	5
♯ max iterations	9	12	15	18
Method B	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.23187e-1	8.67715e-2	5.30592e-2	2.40712e-2
Rate of convergence	–	0.506	0.71	1.14
♯ total iterations	836	900	959	1076
♯ avg iterations	4	5	5	6
♯ max iterations	9	12	15	18
Method C	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.23321e-1	8.68681e-2	5.31222e-2	2.41106e-2
Rate of convergence	–	0.509	0.706	1.14
♯ total iterations	571(573)	678(675)	779(800)	934(960)
♯ avg iterations	3	3	4	5
♯ max iterations	9	12	15	18
Method D_1	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.50856e-1	1.0295e-1	6.20923e-2	2.91945e-2
Rate of convergence	–	0.551	0.729	1.089
♯ total iterations	579(581)	677(676)	785(814)	933(985)
♯ avg iterations	3	3	4	5
♯ max iterations	9	12	15	18
Method D_2	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.50856e-1	1.0295e-1	6.20923e-2	2.91945e-2
Rate of convergence	–	0.551	0.729	1.089
♯ total iterations	579	674	783	933
♯ avg iterations	3	3	4	5
♯ max iterations	9	12	15	18
♯ avg it. bisection per face	15	13	12	11

Table 8: Results for the non-steep filling case using van Genuchten Mualem model.For methods C and D_1, we specify within parentheses the number of iterations corresponding to the second choice of double parametrization when it differs from the one obtained with the first choice, which is reported without parentheses.

A.4 Drainage case using van Genuchten Mualem model

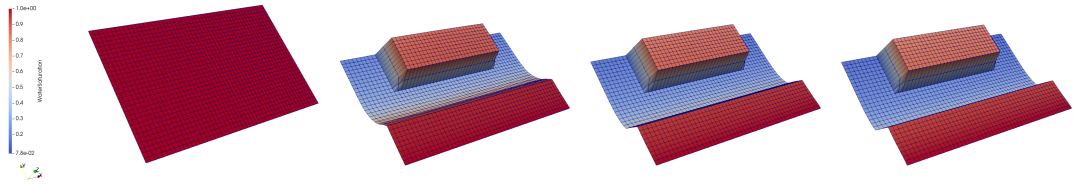

Figure 21: Evolution of the saturation profile for $t \in \{0s, 35 \cdot 10^4 s, 70 \cdot 10^4 s, 105 \cdot 10^4 s\}$ for the non-steep drainage case using Van Genuchten model, Method B and the 50×30 cells mesh.

Method A	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	1.50494e-2	7.3434e-3	3.57016e-3	1.66693e-3
Rate of convergence	–	1.035	1.04	1.099
♯ total iterations	2333	2330	2325	2326
♯ avg iterations	2	2	2	2
♯ max iterations	65	67	67	72
Method B	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	6.52099e-3	3.11282e-3	1.35196e-3	4.53001e-4
Rate of convergence	–	1.064	1.203	1.577
♯ total iterations	2949	3006	3028	3236
♯ avg iterations	2	2	2	3
♯ max iterations	97	96	96	100
Method C	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	6.52104e-3	3.11287e-3	1.35201e-3	4.53048e-4
Rate of convergence	–	1.067	1.203	1.577
♯ total iterations	2855	2818(2817)	2904(2902)	2970(2962)
♯ avg iterations	2	2	2	2
♯ max iterations	99	98(97)	97(95)	101(93)
Method D_1	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	6.58861e-3	3.56638e-3	1.87745e-3	1.01069e-3
Rate of convergence	–	0.886	0.926	0.893
♯ total iterations	2349	2351(2350)	2350	2367
♯ avg iterations	2	2	2	2
♯ max iterations	76	77(76)	80	84
Method D_2	50×30	100×60	200×120	400×240
$\frac{\|\|s-s_{ref}\|\|_{L^2([0,T],\Omega)}}{\|\|s_{ref}\|\|_{L^2([0,T],\Omega)}}$	6.58861e-3	3.56638e-3	1.87745e-3	1.01069e-3
Rate of convergence	–	0.886	0.923	0.893
♯ total iterations	2348	2345	2350	2367
♯ avg iterations	2	2	2	2
♯ max iterations	76	76	80	84
♯ avg it. bisection per face	34	33	32	31

Table 9: Results for the non-steep drainage case using van Genuchten Mualem model. For methods C and D_1, we specify within parentheses the number of iterations corresponding to the second choice of double parametrization when it differs from the one obtained with the first choice, which is reported without parentheses.

B Figures and data related to the steep cases

B.1 Filling case

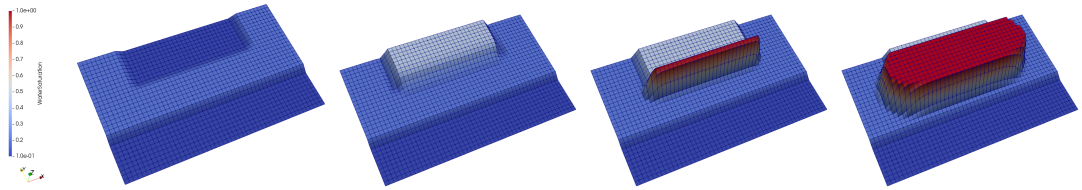

Figure 22: Evolution of the saturation profile for $t \in \{0s, 5422.843s, 37844.5s, 86.4 \cdot 10^3 s\}$ for the steep filling case using Brooks and Corey model, Method B and the 50×30 cells mesh

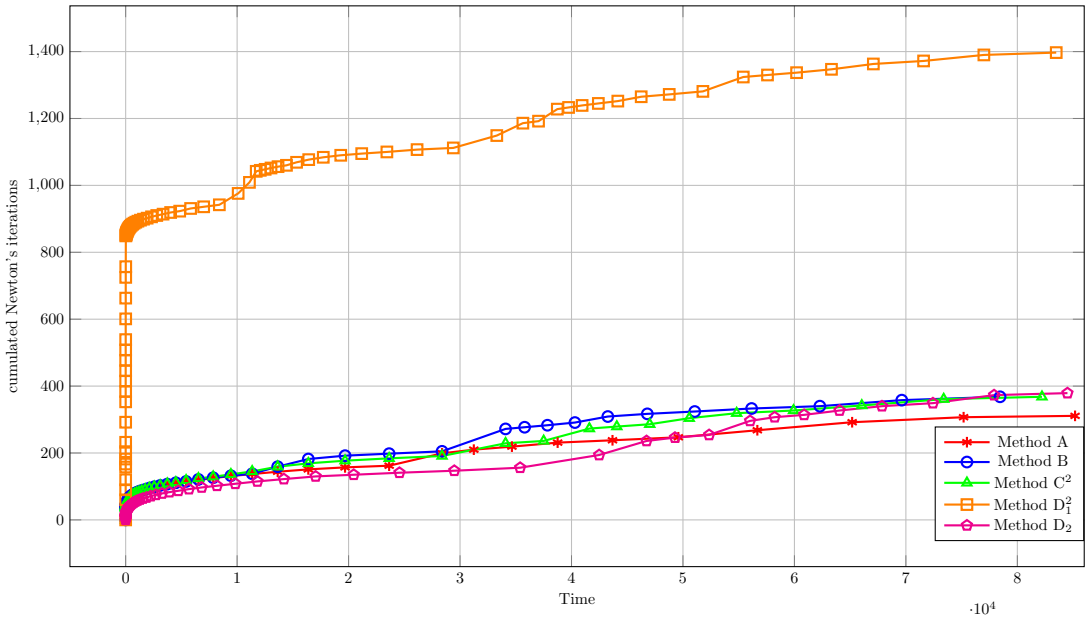

Figure 23: Steep filling case: Evolution of the cumulated number of Newton's iterations for the 50×30 cells mesh

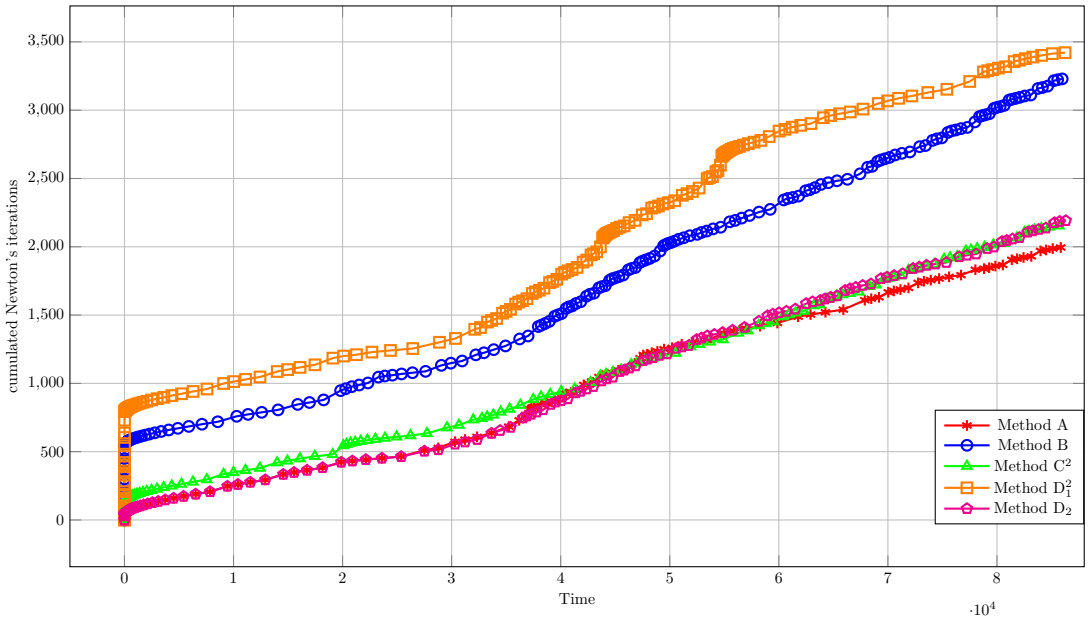

Figure 24: Steep filling case: Evolution of the cumulated number of Newton's iterations for the 400×240 cells mesh

B.2 Drainage case

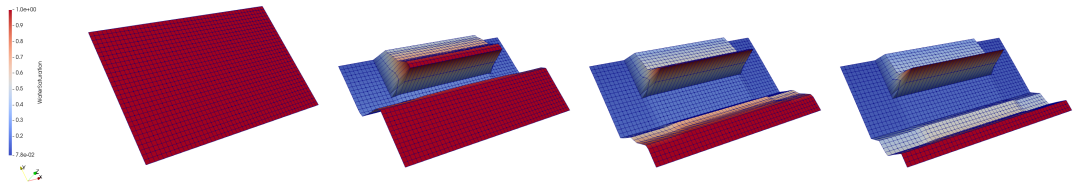

Figure 25: Evolution of the saturation profile for $t \in \{0s, 81593.8s, 308776s, 105 \cdot 10^4 s\}$ for the steep drainage case using Brooks and Corey model, Method B and the 50×30 cells mesh

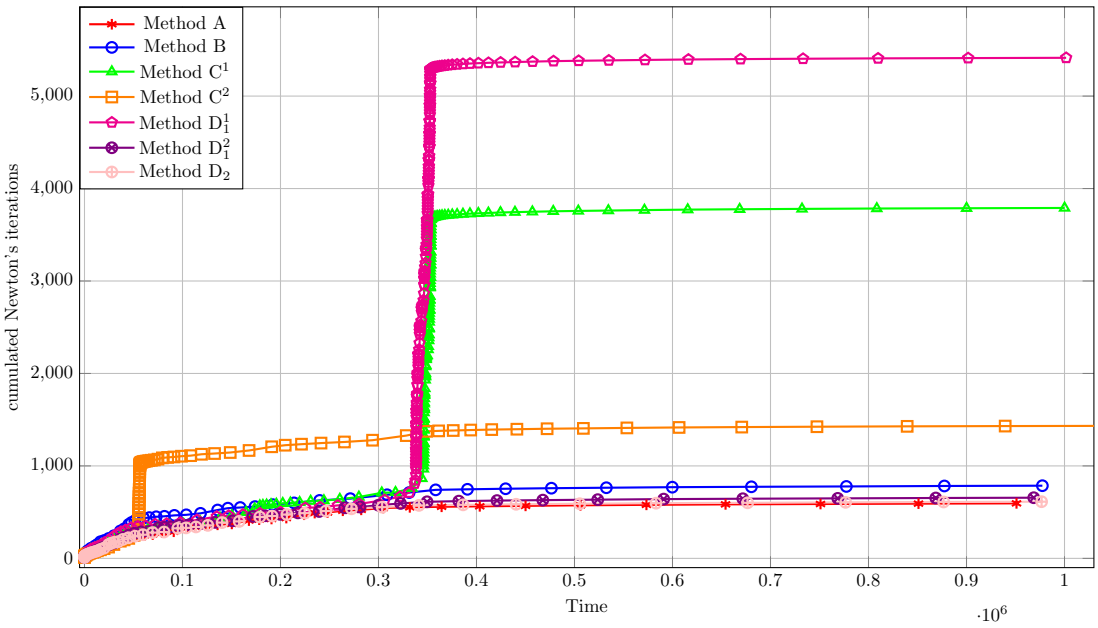

Figure 26: Steep drainage case: Evolution of the cumulated number of Newton's iterations for the 50×30 cells mesh

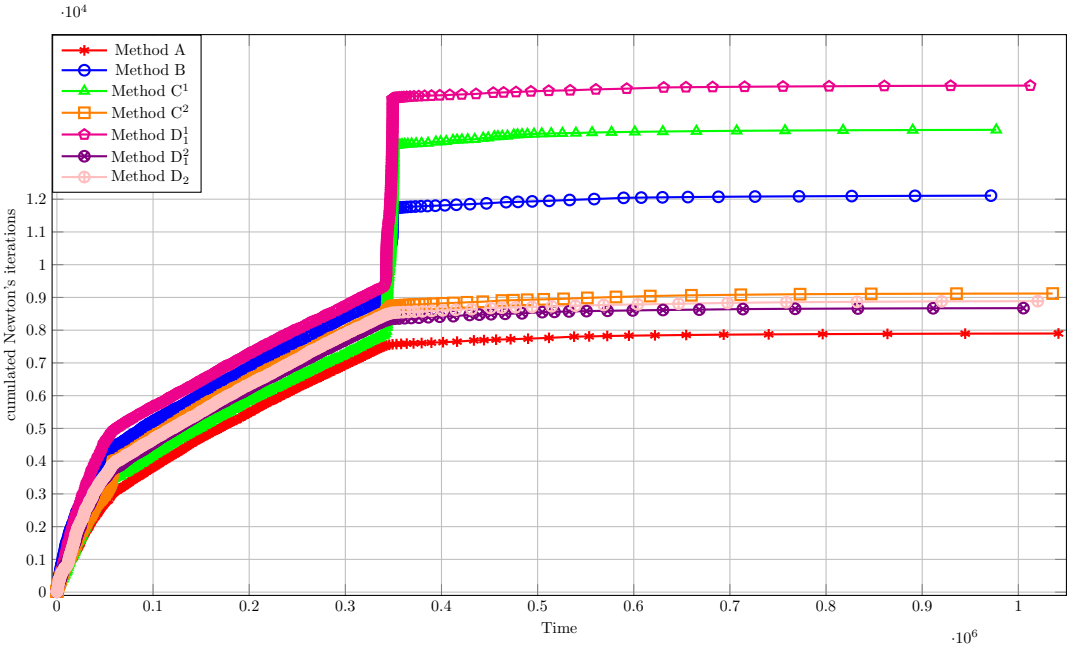

Figure 27: Steep drainage case: Evolution of the cumulated number of Newton's iterations for the 400×240 cells mesh

References

[1] H. W. Alt, S. Luckhaus, and A. Visintin, *On nonstationary flow through porous media*, Ann. Mat. Pura Appl. (4), 136 (1984), pp. 303–316, https://doi.org/10.1007/BF01773387.

[2] T. Arbogast, M. Juntunen, J. Pool, and M. F. Wheeler, *A discontinuous Galerkin method for two-phase flowin a porous medium enforcing $H(div)$ velocity and continuous capillary pressure*, Comput. Geosci., 17 (2013), pp. 1055–1078, https://doi.org/10.1007/s10596-013-9374-y.

[3] S. Bassetto, C. Cancès, G. Enchéry, and Q. H. Tran, *Robust Newton solver based on variable switch for a finite volume discretization of Richards equation*, in Finite Volumes for Complex Applications IX - Methods, Theoretical Aspects, Examples, R. Klöfkorn, E. Keilegavlen, F. A. Radu, and J. Fuhrmann, eds., vol. 323 of Springer Proceedings in Mathematics & Statistics, 2020, pp. 385–394, https://doi.org/10.1007/978-3-030-43651-3_35.

[4] S. Bassetto, C. Cancès, G. Enchéry, and Q. H. Tran, *Upstream mobility Finite Volumes for the Richards equation in heterogenous domains*. preprint, Jan. 2021, https://hal.archives-ouvertes.fr/hal-03109483.

[5] J. Bear and Y. Bachmat, *Introduction to modeling of transport phenomena in porous media*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990.

[6] L. Bergamaschi and M. Putti, *Mixed finite elements and Newton-type linearizations for the solution of Richards' equation*, Int. J. Numer. Meth. Eng., 45 (1999), pp. 1025–1046, https://doi.org/10.1002/(SICI)1097-0207(19990720)45:8<1025::AID-NME615>3.0.CO;2-G.

[7] M. Bertsch, R. Dal Passo, and C. J. van Duijn, *Analysis of oil trapping in porous media flow*, SIAM J. Math. Anal., 35 (2003), pp. 245–267, https://doi.org/10.1137/S0036141002407375.

[8] K. Brenner, *Acceleration of newton's method using nonlinear jacobi preconditioning*, in Finite Volumes for Complex Applications IX - Methods, Theoretical Aspects, Examples, R. Klöfkorn, E. Keilegavlen, F. A. Radu, and J. Fuhrmann, eds., Cham, 2020, Springer International Publishing, pp. 395–403, https://doi.org/10.1007/978-3-030-43651-3_36.

[9] K. Brenner and C. Cancès, *Improving Newton's method performance by parametrization: The case of the Richards equation*, SIAM J. Numer. Anal., 55 (2017), pp. 1760–1785, https://doi.org/https://doi.org/10.1137/16M1083414.

[10] K. Brenner, C. Cancès, and D. Hilhorst, *Finite volume approximation for an immiscible two-phase flow in porous media with discontinuous capillary pressure*, Comput. Geosci., 17 (2013), pp. 573–597, https://doi.org/10.1007/s10596-013-9345-3.

[11] K. Brenner, J. Droniou, R. Masson, and E. H. Quenjel, *Total-velocity-based finite volume discretization of two-phase Darcy flow in highly heterogeneous media with discontinuous capillary pressure*. hal-02957054, 2020, https://hal.archives-ouvertes.fr/hal-02957054.

[12] K. Brenner, M. Groza, L. Jeannin, R. Masson, and J. Pellerin, *Immiscible two-phase Darcy flow model accounting for vanishing and discontinuous capillary pressures: application to the flow in fractured porous media*, Comput. Geosci., 21 (2017), pp. 1075–1094, https://doi.org/10.1007/s10596-017-9675-7.

[13] K. Brenner, D. Hilhorst, and V. D. Huy Cuong, *A gradient scheme for the discretization of the Richards equation*, in Finite volumes for complex applications. VII. Elliptic, Parabolic and Hyperbolic problems, J. Fuhrmann, M. Ohlberger, and C. Rohde, eds., vol. 78 of Springer Proc. Math. Stat., Springer, Cham, 2014, pp. 537–545, https://doi.org/10.1007/978-3-319-05591-6_53.

[14] K. Brenner, R. Masson, and E. H. Quenjel, *Vertex approximate gradient discretization preserving positivity for two-phase Darcy flows in heterogeneous porous media*, J. Comput. Phys., 409 (2020), p. 109357, https://doi.org/10.1016/j.jcp.2020.109357.

[15] R. Brooks and A. Corey, *Hydraulic properties of porous media*, Hydrology Paper, 7 (1964), pp. 26–28.

[16] C. Cancès, *Finite volume scheme for two-phase flow in heterogeneous porous media involving capillary pressure discontinuities*, ESAIM: M2AN, 43 (2009), pp. 973–1001, https://doi.org/10.1051/m2an/2009032.

[17] C. Cancès, T. Gallouët, and A. Porretta, *Two-phase flows involving capillary barriers in heterogeneous porous media*, Interfaces Free Bound., 11 (2009), pp. 239–258, https://doi.org/10.4171/IFB/210.

[18] V. Casulli and P. Zanolli, *A nested Newton-type algorithm for finite volume methods solving Richards' equation in mixed form*, SIAM J. Sci. Comp., 32 (2010), pp. 2255–2273, https://doi.org/https://doi.org/10.1137/100786320.

[19] 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 (1990), pp. 1483–1496, https://doi.org/https://doi.org/10.1029/WR026i007p01483.

[20] H.-J. G. Diersch and P. Perrochet, *On the primary variable switching technique for simulating unsaturated–saturated flows*, Adv. Water Resour., 23 (1999), pp. 271–301, https://doi.org/https://doi.org/10.1016/S0309-1708(98)00057-8.

[21] G. Enchéry, R. Eymard, and A. Michel, *Numerical approximation of a two-phase flow in a porous medium with discontinuous capillary forces*, SIAM J. Numer. Anal., 43 (2006), pp. 2402–2422, https://doi.org/10.1137/040602936.

[22] A. Ern, I. Mozolevski, and L. Schuh, *Discontinuous Galerkin approximation of two-phase flows in heterogeneous porous media with discontinuous capillary pressures*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 1491–1501, https://doi.org/10.1016/j.cma.2009.12.014, http://dx.doi.org/10.1016/j.cma.2009.12.014.

[23] B. G. Ersland, M. S. Espedal, and R. Nybø, *Numerical methods for flow in a porous medium with internal boundaries*, Comput. Geosci., 2 (1998), pp. 217–240, https://doi.org/https://doi.org/10.1023/A:1011554320427.

[24] R. Eymard, T. Gallouët, and R. Herbin, *Finite volume methods*. Ciarlet, P. G. (ed.) et al., in Handbook of numerical analysis. North-Holland, Amsterdam, pp. 713–1020, 2000.

[25] R. Eymard, T. Gallouët, R. Herbin, M. Gutnic, and D. Hilhorst, *Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies*, M3AS: Math. Models Meth. Appl. Sci., 11 (2001), pp. 1505–1528, https://doi.org/10.1142/S0218202501001446.

[26] R. Eymard, C. Guichard, R. Herbin, and R. Masson, *Gradient schemes for two-phase flow in heterogeneous porous media and Richards equation*, Z. Angew. Math. Mech., 94 (2014), pp. 560–585, https://doi.org/10.1002/zamm.201200206.

[27] R. Eymard, M. Gutnic, and D. Hilhorst, *The finite volume method for Richards equation*, Comput. Geosci., 3 (1999), pp. 259–294, https://doi.org/10.1023/A:1011547513583.

[28] M. W. Farthing and F. L. Ogden, *Numerical solution of Richards' equation: A review of advances and challenges*, Soil Science Society of America Journal, 81 (2017), pp. 1257–1269, https://doi.org/https://doi.org/10.2136/sssaj2017.02.0058.

[29] P. A. Forsyth, *A control volume finite element approach to NAPL groundwater contamination.*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 1029–1057, https://doi.org/10.1137/0912055.

[30] P. A. Forsyth, Y. S. Wu, and K. Pruess, *Robust numerical methods for saturated-unsaturated flow with dry initial conditions in heterogeneous media*, Adv. Water Resour., 18 (1995), pp. 25–38, https://doi.org/https://doi.org/10.1016/0309-1708(95)00020-J.

[31] F. P. Hamon, B. T. Mallison, and H. A. Tchelepi, *Implicit hybrid upwinding for two-phase flow in heterogeneous porous media with buoyancy and capillarity*, Comput. Methods Appl. Mech. Engrg., 331 (2018), pp. 701–727, https://doi.org/10.1016/j.cma.2017.10.008.

[32] H. Hoteit and A. Firoozabadi, *Numerical modeling of two-phase flow in heterogeneous permeable media with different capillarity pressures*, Advances in Water Resources, 31 (2008), pp. 56–73, https://doi.org/https://doi.org/10.1016/j.advwatres.2007.06.006.

[33] M. R. Kirkland, R. G. Hills, and P. J. Wierenga, *Algorithms for solving Richards equation for variably saturated soils*, Water Resour. Res., 28 (1992), pp. 2049–2058, https://doi.org/https://doi.org/10.1029/92WR00802.

[34] R. A. Klausen, F. A. Radu, and G. T. Eigestad, *Convergence of MPFA on triangulations and for Richards' equation*, Internat. J. Numer. Methods Fluids, 58 (2008), pp. 1327–1351, https://doi.org/10.1002/fld.1787, http://dx.doi.org/10.1002/fld.1787.

[35] F. Lehmann and P. Ackerer, *Comparison of iterative methods for improved solutions of the fluid flow equation in partially saturated porous media*, Transport in Porous Media, 31 (1998), pp. 275–292, https://doi.org/https://doi.org/10.1023/A:1006555107450.

[36] H. Li, M. W. Farthing, C. N. Dawson, and M. C. T., *Local discontinuous Galerkin approximations to Richards' equation*, Adv. Water. Resour., 30 (2007), pp. 555–575, https://doi.org/10.1016/j.advwatres.2006.04.011.

[37] F. List and F. Radu, *A study on iterative methods for solving richards' equation*, Computational Geosciences, 20 (2016), https://doi.org/https://doi.org/10.1007/s10596-016-9566-3.

[38] F. Marinelli and D. S. Dunford, *Semianalytical solution to Richards equation for layered porous media*, J. Irrig. Drain. Eng., 124 (1998), pp. 290–299, https://doi.org/https://doi.org/10.1061/(ASCE)0733-9437(1998)124:6(290).

[39] D. McBride, M. Cross, N. Croft, C. Bennett, and J. Gebhardt, *Computational modelling of variably saturated flow in porous media with complex three-dimensional geometries*, Int. J. Numer. Meth. Fluids, 50 (2006), pp. 1085–1117, https://doi.org/https://doi.org/10.1002/fld.1087.

[40] J. Niessner, R. Helmig, H. Jakobs, and J. E. Roberts, *Interface condition and linearization schemes in the Newton iterations for two-phase flow in heterogeneous porous media*, Adv. Water Resour., 28 (2005), pp. 671–687, https://doi.org/10.1016/j.advwatres.2005.01.006.

[41] I. Pop, F. Radu, and P. Knabner, *Mixed finite elements for the richards' equation: linearization procedure*, Journal of Computational and Applied Mathematics, 168 (2004), pp. 365–373, https://doi.org/https://doi.org/10.1016/j.cam.2003.04.008. Selected Papers from the Second International Conference on Advanced Computational Methods in Engineering (ACOMEN 2002).

[42] F. A. Radu, I. S. Pop, and P. Knabner, *Order of convergence estimates for an Euler implicit, mixed finite element discretization of Richards' equation*, SIAM J. Numer. Anal., 42 (2004), pp. 1452–1478, https://doi.org/10.1137/S0036142902405229.

[43] L. A. Richards, *Capillary conduction of liquids through porous mediums*, Physics, 1 (1931), pp. 318–333, https://doi.org/10.1063/1.1745010.

[44] D. Seus, K. Mitra, I. S. Pop, F. A. Radu, and C. Rohde, *A linear domain decomposition method for partially saturated flow in porous media*, Computer Methods in Applied Mechanics and Engineering, 333 (2018), p. 331,Äì355, https://doi.org/10.1016/j.cma.2018.01.029, http://dx.doi.org/10.1016/j.cma.2018.01.029.

[45] M. Slodicka, *A robust and efficient linearization scheme for doubly nonlinear and degenerate parabolic problems arising in flow in porous media*, SIAM Journal on Scientific Computing, 23 (2002), pp. 1593–1614, https://doi.org/10.1137/S1064827500381860.

[46] D. Svyatskiy and K. Lipnikov, *Second-order accurate finite volume schemes with the discrete maximum principle for solving Richards' equation on unstructured meshes*, Adv. Water Resour., 104 (2017), pp. 114–126, https://doi.org/10.1016/j.advwatres.2017.03.015.

[47] M. T. van Genuchten, *A closed-form equation for predicting the hydraulic conductivity of unsaturated soils*, Soil Sci. Soc. Amer. J., 44 (1980), pp. 892–898, https://doi.org/10.2136/sssaj1980.03615995004400050002x.

[48] C. S. Woodward and C. N. Dawson, *Analysis of expanded mixed finite element methods for a nonlinear parabolic equation modeling flow into variably saturated porous media*, SIAM J. Numer. Anal., 37 (2000), pp. 701–724, https://doi.org/10.1137/S0036142996311040, https://doi.org/10.1137/S0036142996311040.