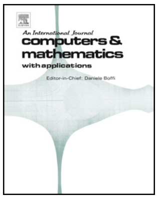
Computers and Mathematics with Applications 152 (2023) 155–167

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa



An adaptive solution strategy for Richards' equation

Jakob S. Stokke [a], Koondanibha Mitra [b], Erlend Storvik [a,c], Jakub W. Both [a], Florin A. Radu [a,*]

[a] Center for Modeling of Coupled Subsurface Dynamics, Department of Mathematics, University of Bergen, Bergen, Norway
[b] Computational Mathematics group, Hasselt University, Hasselt, Belgium
[c] Department of Computer science, Electrical engineering and Mathematical sciences, Western Norway University of Applied Sciences, Førde, Norway

ARTICLE INFO

Keywords:
Richards equation
Linearisation schemes
Newton method
L-scheme
Flow in porous media
Finite elements

ABSTRACT

Flow in variably saturated porous media is typically modeled by the Richards equation, a nonlinear elliptic-parabolic equation which is notoriously challenging to solve numerically. In this paper, we propose a robust and fast iterative solver for Richards' equation. The solver relies on an adaptive switching algorithm, based on rigorously derived *a posteriori* indicators, between two linearization methods: L-scheme and Newton. Although a combined L-scheme/Newton strategy was introduced previously in [1], here, for the first time we propose a reliable and robust criteria for switching between these schemes. The performance of the solver, which can be in principle applied to any spatial discretization and linearization methods, is illustrated through several numerical examples.

1. Introduction

In this paper, we consider the pressure head ψ based formulation of the Richards equation

$$\partial_t \theta(\boldsymbol{x}, \psi) - \nabla \cdot [K(\boldsymbol{x}, \theta)(\psi))\nabla(\psi + z)] = f, \tag{1}$$

where $\theta : \Omega \times \mathbb{R} \to [0,1]$ is the water content, K is the rank 2 permeability tensor of the porous medium, z is the height against the gravitational direction, and f is a source/sink term. Richards' equation is used to model the flow of water in saturated/unsaturated porous media. It is a highly nonlinear and degenerate elliptic-parabolic equation which makes solving it a very challenging task, see e.g. the review work of [2]. We refer to [3] for the existence and uniqueness of a weak solution of Richards' equation.

There are plenty of works regarding discretization of Richards' equation. Due to the low regularity of solutions of (1), see [4], generally, a backward Euler (implicit) scheme (3) is employed to discretize it in time, see e.g. [1,5]. Regarding spatial discretization we mention continuous Galerkin finite elements [6,7], mixed or expanded mixed finite elements [8–12], finite volumes [13,14] (see also the recent review [15]), or multipoint flux approximation (MPFA) [16]. Regardless of the choice of the spatial discretization method, one has to solve at each time step a nonlinear, finite-dimensional problem. In this paper, we will focus on how to efficiently solve these problems using iterative linearization techniques.

The main iterative linearization methods used for this type of nonlinear problem are the Newton method, Picard or modified Picard, L-scheme, the Jaeger-Kacur method, or combinations of them. Perhaps the most common choice is the Newton method [17,18] which converges quadratically provided the initial guess is close enough to the final solution. For a r-Hölder continuous θ' function ($r \in (0,1]$) and the initial guess equal to the solution of the previous time step, it was shown in [10] that the Newton scheme is $(1+r)^{\text{th}}$ order convergent if

$$\tau \leq C\theta_m^{\frac{2+r}{r}} h^d, \tag{2}$$

where $\tau > 0$ is the time step size, $h > 0$ the mesh size, $d \in \mathbb{N}$ the spatial dimension, $C > 0$ a constant which depends on the domain and the nonlinearities, and $\theta_m := \inf \theta' \geq 0$. However, for simulations in 2 or 3 dimensions, condition (2) is quite restrictive particularly if the mesh size h is small, or if the problem is degenerate ($\theta_m = 0$). This fact is corroborated by numerical simulations in [1,19] which show that the Newton method fails to converge in many such cases. One can improve the robustness of Newton method by using a damped version of it. Line search, variable switching [20] or trust-regions techniques [21] are examples of such. Alternatively, one can increase the robustness of Newton's method by performing first a few fixed-point iterations. This was proposed in [17,18] by using the Picard method and in [1] by using the L-scheme. Nevertheless, the switching between the schemes was not based on an *a posteriori* indicator, but done in a heuristic manner.

[*] Corresponding author.
E-mail address: florin.radu@uib.no (F.A. Radu).

https://doi.org/10.1016/j.camwa.2023.10.020

Available online 25 October 2023

The other linearization schemes are fixed-point type schemes, typically more robust, however only linearly convergent. It has been shown in [22,13] that the Picard method does not perform well for Richards' equation. A modified Picard method was proposed in [22]. The modified Picard coincides with Newton's method for the case of a constant permeability, therefore it inherits robustness problems. The L-scheme, first proposed in [23,24,1], is a stabilized Picard method and it was designed to be unconditionally converging irrespective of the choice of the initial guess even in degenerate settings and for larger time steps. The L-scheme (see Definition 2.3) uses a global constant as a stabilization coefficient, does not involve the computation of any derivatives, and thus, is not only more stable but also consumes less computational time per iteration due to easier assembly of the stiffness matrices which are better conditioned. Numerical results in [1,19] clearly demonstrate this. However, they also reveal that the L-scheme converges considerably slower in terms of number of iterations compared to the Newton scheme and at a linear rate. Furthermore, its overall performance strongly depends on the careful choice of a tuning parameter; despite theoretical stability, an improper choice may effectively result in stagnation. The sensitivity of the performance of the L-scheme with respect to the stabilization can be significantly relaxed when combining the L-scheme with Anderson acceleration [25]. Indeed, for Richards equation extended to deformable porous media and solved by an L-scheme, it has been demonstrated that, first, the stabilization parameter can be chosen outside the theoretical range, and second, the non-degenerate convergence can be retained in case of previous divergence or accelerated, as also discussed from a theoretical perspective [26]. Similar stabilizing properties of the Anderson acceleration have been also discussed for general fixed-point methods [27,28]. Other fixed point iterations schemes include Jäger-Kacŭr scheme [29] which converges unconditionally albeit slowly, and is more computationally expensive than the L-scheme per iteration, see Table 1. The modified L-scheme, proposed in [19], shows stability similar to the L-scheme while having much faster convergence rates (scaling with τ); yet, the convergence is still linear.

In this paper, we investigate a hybrid strategy, dynamically switching between the L-scheme and Newton's method. This utilizes the advantages of both methods: the unconditional stability of the L-scheme, and the quadratic convergence of Newton's method when close to the exact solution. The crucial difference to previous works on hybrid approaches, e.g. [1,17], is the adaptive nature of the switch between both linearization methods. A switch from the L-scheme to Newton's method is performed when the iterate is sufficiently close to the solution. This finally allows us to balance robustness and speed.

The main challenge in implementing this strategy originates from deriving a rigorous switching criteria between the schemes. Since, the *a priori* estimates, such as the ones provided in [10], involve unknown constants and assume the worst-case scenario, we pursue an *a posteriori* estimate-based approach here instead. A rigorous and efficient *a posteriori* estimator for the fully degenerate Richards equation involving linearization errors was derived in [30] in the continuous space-time setting. For the time-discrete problem (3), a robust, efficient, and reliable estimator was derived in [31] using an orthogonal decomposition result dividing the total error into a discretization and a linearization component. Furthermore, its effectiveness was demonstrated numerically. These papers serve as the main inspirations in deriving the *a posteriori* based switching criteria in Section 3 and an adaptive L-scheme algorithm in Appendix A. Nevertheless, since we are only interested in computing the linearization error component, the computation of equilibrated flux will be avoided wherever possible.

The paper is organized as follows. In Section 2, we introduce the mathematical notation, state the assumptions, define the fully-discrete solution, and elaborate on different linearization methods. In Section 3, the adaptive switching algorithm is developed. Firstly, a concept of linearization error is introduced along with the derivation of a predictive indicator for linearization error of the next iteration. The adaptive algorithm compares the linearization error with the estimator to determine the exact switching points. In Section 4, four numerical test cases (partially saturated, degenerate, recharge of a drainage trench and a heterogeneous medium) are presented which illustrate the robustness and computational efficiency of the adaptive scheme compared to the standard Newton's method or the L-scheme. Section 5 contains the conclusions of this work. The paper ends with two appendices, one concerning an adaptive L-scheme and the other on the details of the computation of the equilibrated flux.

2. Mathematical and numerical formulation

We consider Richards' equation in the space-time domain $\mathcal{G} = \Omega \times [0, T]$, where Ω is a bounded domain in $\mathbb{R}^d$ with a Lipschitz continuous boundary $\partial\Omega$, and $T > 0$. Let $(\cdot,\cdot)$ and $\|\cdot\|$ be the inner product and norm of the square-integrable functions in Ω, i.e. $L^2(\Omega)$, respectively. Moreover, using common notation from functional analysis, $H^1(\Omega)$ represents the Sobolev space of functions with first-order weak derivatives in $L^2(\Omega)$, and $H^1_0(\Omega)$ its subspace containing functions with vanishing trace at the boundary.

Assumption 1. For the material properties θ and K, and source term f in (1), the following assumptions are made:

(a) The saturation function $\theta(\boldsymbol{x},\psi)$ (for $\boldsymbol{x} \in \Omega$, $\psi \in \mathbb{R}$) is Lipschitz continuous and monotonically increasing with respect to ψ with L_θ and $\theta_m \geq 0$ denoting the global Lipschitz constant and the lower bound for the derivative respectively.

(b) The permeability tensor $K : \Omega \times [0,1] \to \mathbb{R}^{d\times d}$ satisfies the uniform (pseudo) ellipticity condition, i.e., for constants $\kappa_M \geq \kappa_m \geq 0$,

$$\kappa_m |\boldsymbol{z}|^2 \leq \boldsymbol{z}^{\mathrm{T}} K\, \boldsymbol{z} \leq \kappa_M |\boldsymbol{z}|^2, \quad \forall\, \boldsymbol{z} \in \mathbb{R}^d.$$

Moreover, $K(\boldsymbol{x}, \theta(\boldsymbol{x},\psi))$ (denoted later by $K \circ \theta$) is Lipschitz continuous with respect to ψ for all $\boldsymbol{x} \in \Omega$, with the global Lipschitz constant being L_κ.

(c) The source function satisfies $f \in C(0,T; L^2(\Omega))$.

Note that these assumptions are consistent with the commonly used Brooks-Corey [32] and van Genuchten [33] parametrizations of the functions θ and K. To simplify notation we write $K(\boldsymbol{x},\theta) = K(\theta)$ and $\theta(\boldsymbol{x},\psi) = \theta(\psi)$ throughout the paper, although they can be treated as spatially heterogeneous everywhere in our analysis.

2.1. Time-discretization: backward Euler

To discretize the Richards equation in time we consider the backward-Euler time discretization of (1). For this implicit scheme, no CFL conditions need to be satisfied for stability (thus avoiding restrictions on the time step size). Moreover, it does not require higher-order time regularity (unlike the Crank-Nicholson scheme) to converge to the time-continuous solutions. We subdivide the time-interval $[0,T]$ uniformly N times with time step size $\tau = T/N$ and discrete time steps $t_n = \tau n$, where $n \in \{1,\ldots,N\}$. Then, we look for a sequence $\{\psi^n\}_{n=1}^N$ of functions in Ω, satisfying the time-discrete system

$$\frac{\theta(\psi^n) - \theta(\psi^{n-1})}{\tau} - \nabla \cdot \left[K(\theta(\psi^n))\nabla(\psi^n + z)\right] = f(t_n). \tag{3}$$

Denoting $f(t_n)$ by f^n subsequently, a more precise and general definition of the weak solutions of (3) is given below. For simplicity, we assume homogeneous Dirichlet boundary condition although our results are valid for Dirichlet and Neumann boundary conditions in general.

Definition 2.1 *(Backward Euler time-discretization of* (1)*).* Let $\psi^0 \in L^2(\Omega)$ be given. Then the sequence $\{\psi^n\}_{n=1}^N \subset H^1_0(\Omega)$ is the backward Euler solution of (1) if for all $n \in \{1,\ldots,N\}$, and $v \in H^1_0(\Omega)$,

$$\frac{1}{\tau}(\theta(\psi^n) - \theta(\psi^{n-1}), v) + (K(\theta(\psi^n))\nabla(\psi^n + z), \nabla v) = (f^n, v). \tag{4}$$

2.2. Space-discretization: continuous Galerkin finite elements

We consider the finite element method to discretize (4) further in space. Let $\mathcal{T}_h$ be a triangulation of Ω into closed d-simplices, where $h := \max_{E\in\mathcal{T}_h}(\mathrm{diam}(E))$ denotes the mesh size. Assuming Ω is a polygon, the Galerkin finite element space is

$$V_h = \left\{ v_h \in H^1_0(\Omega) |\ v_{h|E} \in \mathcal{P}_p(E),\ T \in \mathcal{T}_h \right\}, \tag{5}$$

where $\mathcal{P}_p(E)$ denotes the space of p-order polynomials on E, $p \in \mathbb{N}$. Then, the fully discrete Galerkin formulation of Richards' equation reads

Definition 2.2 *(Fully discrete solution of* (1)*).* Let $\psi^0_h := \psi^0 \in L^2(\Omega)$. Then the sequence $\{\psi^n_h\}_{n=1}^N \subset V_h$ is the fully discrete solution of (1) if for all $n \in \{1, \dots, N\}$, and $v_h \in V_h$,

$$(\theta(\psi^n_h) - \theta(\psi^{n-1}_h), v_h) + \tau(K(\theta(\psi^n_h))\nabla(\psi^n_h + z), \nabla v_h) = \tau(f^n, v_h). \tag{6}$$

2.3. Iterative linearization schemes

To obtain the solution of the nonlinear problem (6) an iterative linearization scheme is generally employed. To investigate the trade-off between the stability and speed of such schemes, we focus on two linearization strategies that will be representatives of linearly and quadratically convergent methods with convergence meant in the L^2 sense.

2.3.1. Linearly convergent schemes: the L-scheme

Where the quadratically convergent Newton method utilizes a proper first-order Taylor expansion of the nonlinear terms in (6), the linearly convergent methods that we consider here, only exploit an expansion of the monotone components, i.e. the nonlinear saturation function. Moreover, the expansion does not need to be exact. Consider the following scheme: Given $\psi^{n-1}_h, \psi^{n,j-1}_h \in V_h$, find $\psi^{n,j}_h \in V_h$ such that

$$\begin{aligned}&(\mathcal{L}(\psi^{n,j-1}_h)(\psi^{n,j}_h - \psi^{n,j-1}_h), v_h) + \tau(K(\theta(\psi^{n,j-1}_h))\nabla(\psi^{n,j}_h + z), \nabla v_h)\\ &\quad = \tau(f^n, v_h) - (\theta(\psi^{n,j-1}_h) - \theta(\psi^{n-1}_h), v_h),\end{aligned} \tag{7}$$

for all $v_h \in V_h$, where $\mathcal{L} : \mathbb{R} \to [0, \infty)$ is a predetermined positive weight function, and $j \in \mathbb{N}$ is the iteration index. Observe that, provided $\kappa_m > 0$ in Assumption 1, the problem above is linear, monotone, and Lipschitz with respect to $\psi^{n,j}_h$, and hence a unique weak solution of (7) exists. Moreover, if the iteration converges, i.e. if $\psi^{n,j}_h \to \psi^n_h$ strongly in $H^1_0(\Omega)$, then ψ^n_h indeed solves (6). There can be many different choices of the function $\mathcal{L}$ which leads to different linearization schemes, see Table 1. For the rest of this paper, we mainly focus on the case when $\mathcal{L}$ is constant which leads to the widely studied L-scheme.

Definition 2.3 *(L-scheme).* Let $\psi^{n-1}_h, \psi^{n,0}_h \in L^2(\Omega)$ and $L > 0$ be given. Then the L-scheme solves for the sequence $\{\psi^{n,j}_h\}_{j\in\mathbb{N}} \subset V_h$ which satisfies for all iteration indices $j \in \mathbb{N}$, and $v_h \in V_h$

$$\begin{aligned}&L((\psi^{n,j}_h - \psi^{n,j-1}_h), v_h) + \tau(K(\theta(\psi^{n,j-1}_h))\nabla(\psi^{n,j}_h + z), \nabla v_h)\\ &= \tau(f^n, v_h) - (\theta(\psi^{n,j-1}_h) - \theta(\psi^{n-1}_h), v_h).\end{aligned} \tag{8}$$

Different choices of $\mathcal{L}$ and the resulting schemes are listed below

Remark 1 *(Non-constant L for heterogeneous media).* For the L-scheme, spatially varying constitutive laws, e.g., the water content, are ideally handled by using spatially varying linearization L. The proofs can be adapted by utilizing weighted inner products, with the weights varying in space accordingly; similar ideas have been successfully applied in the context of iterative splitting schemes close to the L-scheme but applied for poroelasticity modeled by Biot's equations [34] and can be transferred to the case of Richards' equation. Likewise, the practical performance of the linearization is expected to be effectively dominated by the global supremum of the locally evaluated convergence rates (and not by the more pessimistic convergence rate evaluated in respective global upper and lower bounds of the single parameters). For more details on the techniques, we refer the reader to [34].

Table 1
Different linearly convergent schemes (7) defined along with their linearization weight function $\mathcal{L}$.

Scheme	$\mathcal{L}(\psi)$
Picard	0
Modified Picard [22]	$\theta'(\psi)$
Jäger-Kacŭr [29]	$\sup_{\xi\in\mathbb{R}} \frac{\theta(\xi)-\theta(\psi)}{\xi-\psi}$
L-scheme [23,24,1]	$L > 0$ constant
Modified L-scheme [19]	$\theta'(\psi) + M\tau,\ M > 0$ constant

It has been shown in [1, Theorem 1] that if $L \geq \frac{1}{2}\sup_{\xi\in\mathbb{R}} \theta'(\xi)$, then the L-scheme iterations converge irrespective of the initial guess under minor restrictions on the time step size τ and independent of the mesh size. However, numerical results in [1,19] reveal that the convergence of the L-scheme can be relatively slow, depending on the choice of the stabilization parameter L, see please the Appendix A for an adaptive L-scheme. One can enhance the convergence speed by computing L using the previous iterates and derivatives. In general, taking L as the Jacobian matrix, would lead to the modified Picard method [22]. This is exploited in the modified Picard scheme, first proposed in [22], uses $\mathcal{L}(\psi^{n,j-1}) = \theta'(\psi^{n,j-1})$, complying with the first-order Taylor series expansion $\theta(\psi^{n,j}) \approx \theta(\psi^{n,j-1}) + \theta'(\psi^{n,j-1})(\psi^{n,j} - \psi^{n,j-1})$. As a result, if converging it requires fewer iterations compared to the L-scheme although the convergence is still linear. Nevertheless, this choice of the $\mathcal{L}$ function may lead to divergence of the scheme for larger time step sizes, as predicted in [10] and observed numerically in [1,19]. In an attempt to resolve this issue, a modified L-scheme was proposed in [19] that inherits the characteristics of both the L-scheme (except that it is using derivatives and the linear systems are not necessarily well conditioned) and the Picard scheme. The modified L-scheme exhibits increased stability compared to the Picard scheme while retaining its speed. However, the modified L-scheme converges unconditionally under the additional restriction that $\psi^{n,0}_h = \psi^{n-1}_h$ and the discrete time-derivative $(\psi^n_h - \psi^{n-1}_h)/\tau$ is in $L^\infty(\Omega)$. Since the objective of this paper is to start the linearization iterations with a stable scheme, and then switch to a quadratically converging scheme when its convergence can be guaranteed, the rest of the study will be with respect to the L-scheme which is arguably the most stable among the schemes presented in Table 1 and the cheapest in terms of computing time per iteration (due to well-conditioned linear systems and not involving derivatives). Nonetheless, we remark that our methodology generalizes to all other linearly converging iterative methods.

Remark 2 *(Generality of the results).* Although the analysis of Section 3 primarily focuses on the switching between L-scheme and the Newton method, the same techniques can be directly extended to cover switching between the schemes in Table 1 and Newton. Moreover, the L-adaptive strategy in Appendix A can be extended to the modified L-scheme (see Table 1) to select the parameter $M > 0$ adaptively.

2.3.2. Quadratically convergent scheme: the Newton method

The Newton method uses the first order Taylor series expansions of all the nonlinear functions in (1) to ensure quadratic rates of convergence.

Definition 2.4 *(The Newton method).* Let $\psi^{n-1}_h, \psi^{n,0}_h \in L^2(\Omega)$ be given. Then the Newton method solves for the sequence $\{\psi^{n,j}_h\}_{j\in\mathbb{N}} \subset V_h$ which satisfies for all iteration indices $j \in \mathbb{N}$, and $v_h \in V_h$

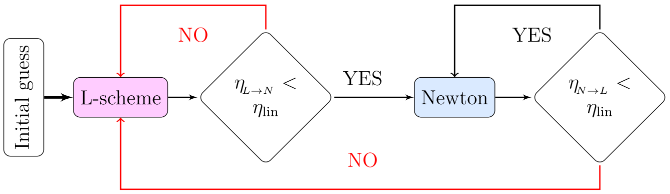

Initial guess
L-scheme
NO
$\eta_{L\to N} < \eta_{\rm lin}$
YES
Newton
YES
$\eta_{N\to L} < \eta_{\rm lin}$
NO

Fig. 1. Flowchart of Adaptive switching algorithm between L-scheme and Newton's method.

$$\begin{aligned}&(\theta'(\psi_h^{n,j-1})(\psi_h^{n,j}-\psi_h^{n,j-1}),v_h)+\tau(K(\theta(\psi_h^{n,j-1}))\nabla(\psi_h^{n,j-1}+z),\nabla v_h)\\&\quad+\tau\left((K\circ\theta)'(\psi_h^{n,j-1})\nabla(\psi_h^{n,j-1}+z)(\psi_h^{n,j}-\psi_h^{n,j-1}),\nabla v_h\right)\\&\quad=\tau(f^n,v_h)-(\theta(\psi_h^{n,j-1})-\theta(\psi_h^{n-1}),v_h).\end{aligned}\tag{9}$$

However, this comes at the cost of decreased numerical stability as discussed in Section 1. In the next section we combine the L-scheme and the Newton method in a consistent manner in order to obtain a linearization strategy that is both stable and fast.

3. A posteriori estimate based adaptive switching between L-scheme and Newton

In this section, we develop the switching algorithm between L-scheme and the Newton method using *a posteriori* error analysis (see Fig. 1). For comparing the errors between different linearization schemes we introduce a uniform notion of linearization errors $\eta_{\rm lin}$ in Section 3.1 based on arguments in [31]. The idea behind the adaptive algorithm is to start with the L-scheme and derive an estimator $\eta_{L\to N}$ in Section 3.2 that predicts from the $j^{\rm th}$ and $(j-1)^{\rm th}$ iterate the linearization error for the next iteration if done using the Newton scheme. If the error is predicted to decrease, then the iteration switches to Newton. Then another estimator $\eta_{N\to L}$ is derived in Section 3.3 which predicts the linearization error of the next step of the Newton iteration. The algorithm switches back to the L-scheme in case the error is predicted to increase. In fact, we go one step further in Appendix A and derive an estimator $\eta_{L\to L}$ to predict if the L-scheme itself will converge and to tune the value of L accordingly. Finally, the full algorithm is laid out in Section 3.4 based on these estimators.

3.1. Linearization errors and iteration-dependent energy norms

In [31] it is shown that the total numerical error corresponding to a finite element-based linearization scheme can be orthogonally decomposed into a discretization component and a linearization component if the errors are computed using an iteration-dependent energy norm (for linearly convergent schemes in Table 1 this is just the energy norm invoked by the symmetric bilinear form associated with the unknown $\psi_h^{n,j}$ in (7)). Here, we are only interested in the linearization component which is defined as the difference between successive iterates in the aforementioned energy norm, i.e.,

$$\eta_{\rm lin}^j := \left|\!\left|\!\left|\psi_h^{n,j}-\psi_h^{n,j-1}\right|\!\right|\!\right|_{\mathcal{L},\psi_h^{n,j-1}},\tag{10}$$

where $|\!|\!|\cdot|\!|\!|_{\mathcal{L},\psi_h^{n,j-1}}$ represents the particular H^1 equivalent-norm defined using the iterate $\psi_h^{n,j-1}$ and associated with the linearization scheme denoted by $\mathcal{L}$. The fully computable estimator $\eta_{\rm lin}^j$ encapsulates the entirety of the linearization error, as shown in Section 5 of [31], and hence, will be used as its sole measure in the subsequent sections. We mention explicitly the energy norms of the two schemes that are discussed: With reference to Definition 2.3, the energy norm for L-scheme is defined as

$$|\!|\!|\xi|\!|\!|_{L,\psi_h^{n,j-1}} := \left(\int_\Omega L\xi^2+\tau\left|K(\theta(\psi_h^{n,j-1}))^{\frac{1}{2}}\nabla\xi\right|^2\right)^{\frac{1}{2}}\tag{11}$$

for all $\xi\in H_0^1(\Omega)$, and with reference to Definition 2.4 the norm for the Newton method is

$$|\!|\!|\xi|\!|\!|_{N,\psi_h^{n,j-1}} := \left(\int_\Omega \theta'(\psi_h^{n,j-1})\,\xi^2+\tau|K(\theta(\psi_h^{n,j-1}))^{\frac{1}{2}}\nabla\xi|^2\right)^{\frac{1}{2}}.\tag{12}$$

3.2. L-scheme to Newton switching estimate

For some $i\in\mathbb{N}$, let the sequence $\{\psi_h^{n,j}\}_{j=1}^i\subset V_h$ be obtained using the L-scheme (8), and in the $(i+1)^{\rm th}$-iteration we want to test for switching to the Newton scheme. Let $\tilde{\psi}_h^{n,i+1}\in V_h$ be the solution of the Newton scheme (9) having $\psi_h^{n,i}$ as the previous iterate. In this section, we will assume the following:

Assumption 2 *(Convection term is not dominant).* For a given $i\in\mathbb{N}$, there exists a constant $C_N^i\in[0,2)$ such that

$$\tau|K(\theta(\psi_h^{n,i}))^{-\frac{1}{2}}(K\circ\theta)'(\psi_h^{n,i})\nabla(\psi_h^{n,i}+z)|^2\le(C_N^i)^2\theta'(\psi_h^{n,i}),\tag{13}$$

a.e. in Ω.

The assumption above is also required to show the coercivity of the linear problem (9) for $j=i+1$, and hence, to show the existence of solution $\tilde{\psi}_h^{n,i+1}$. Observe that, since $\psi_h^{n,i}$ is known, the constant C_N^i is fully computable. Additionally, it is smaller than 2 if the numerical flux is bounded, and τ is small. Notably, the estimate holds even in the degenerate case when $\theta'(\psi_h^{n,i})=0$, since the left-hand side has $(\theta'(\psi_h^{n,i}))^2$. To cover the degenerate case, we also introduce the concept of an equilibrated flux.

Definition 3.1 *(Equilibrated flux $\boldsymbol{\sigma}_L^i$ for degenerate regions).* For a predetermined $\epsilon>0$, let $\mathcal{T}_{\rm deg}^{i,\epsilon}:=\{K\in\mathcal{T}_h : \inf\theta'(\psi_h^{n,i})<\epsilon \text{ in } K\}$. Let $\Pi_h : L^2(\Omega)\to\mathcal{P}_p(\mathcal{T}_h)$ be the $\mathcal{P}_p$ projection operator, i.e. $(\Pi_h u,v_h)=(u,v_h)$ for all $u\in L^2(\Omega)$ and $v_h\in\mathcal{P}_p(\mathcal{T}_h)$. Moreover, let $\mathbf{RT}_p(\mathcal{T}_h)$ be the $p^{\rm th}$-order Raviart-Thomas space on $\mathcal{T}_h$, i.e., $\boldsymbol{\sigma}\in\mathbf{RT}_p(\mathcal{T}_h)$ implies $\boldsymbol{\sigma}|_K\in(\mathcal{P}_p(K))^d+\boldsymbol{x}\mathcal{P}_p(K)$ for all $K\in\mathcal{T}_h$. Then, we define $\boldsymbol{\sigma}_L^i\in\mathbf{RT}_p(\mathcal{T}_h)\cap\boldsymbol{H}(\mathrm{div},\Omega)$ as

$$\nabla\cdot\boldsymbol{\sigma}_L^i=\begin{cases}\frac{1}{\tau}\Pi_h(L(\psi_h^{n,i}-\psi_h^{n,i-1})-(\theta(\psi_h^{n,i})-\theta(\psi_h^{n,i-1}))) & \text{in } \mathcal{T}_{\rm deg}^{i,\epsilon},\\ 0 & \text{otherwise}.\end{cases}\tag{14}$$

We defer to Appendix B for discussions on how to compute $\boldsymbol{\sigma}_L^i$ in practice. Then, we have the following result.

Proposition 1 *(Error control of L-scheme to Newton switching step). For a given $\psi_h^{n,0},\psi_h^{n-1}\in V_h$, let $\{\psi_h^{n,j}\}_{j=1}^i\subset V_h$ solve (8) for some $i\in\mathbb{N}$. Let $\tilde{\psi}_h^{n,i+1}\in V_h$ be the solution of (9) with the previous iterate $\psi_h^{n,i}$. Recall Definition 3.1. Then, under the Assumptions 1–2, one has*

$$\left|\!\left|\!\left|\tilde{\psi}_h^{n,i+1}-\psi_h^{n,i}\right|\!\right|\!\right|_{N,\psi_h^{n,i}}\le\eta_{L\to N}^i,$$

where,

$$\eta_{L\to N}^i := \frac{2}{2-C_N^i}\left(\left[\eta_{L\to N}^{i,\rm poten}\right]^2+\tau\left[\eta_{L\to N}^{i,\rm flux}\right]^2\right)^{\frac{1}{2}}$$

with

$$\eta_{L\to N}^{i,\rm poten} := \left\|\theta'(\psi_h^{n,i})^{-\frac{1}{2}}\left(L\left(\psi_h^{n,i}-\psi_h^{n,i-1}\right)-\left(\theta(\psi_h^{n,i})-\theta(\psi_h^{n,i-1})\right)\right)\right\|_{\mathcal{T}_h\setminus\mathcal{T}_{\rm deg}^{i,\epsilon}},$$

$$\eta_{L\to N}^{i,\rm flux} := \left\|K(\theta(\psi_h^{n,i}))^{-\frac{1}{2}}\left[\left(K(\theta(\psi_h^{n,i}))-K(\theta(\psi_h^{n,i-1}))\right)\nabla\left(\psi_h^{n,i}+z\right)+\boldsymbol{\sigma}_L^i\right]\right\|.$$

Proof. Observe from (9) that $\delta\psi_h^{i+1} := \tilde{\psi}_h^{n,i+1} - \psi_h^{n,i} \in V_h$ satisfies

$$
\begin{aligned}
&(\theta'(\psi_h^{n,i})\delta\psi_h^{i+1}, v_h) + \tau(K(\theta(\psi_h^{n,i}))\nabla\delta\psi_h^{i+1}, \nabla v_h)\\
&\quad + \tau\left((K\circ\theta)'(\psi_h^{n,i})\nabla(\psi_h^{n,i}+z)\,\delta\psi_h^{i+1}, \nabla v_h\right)\\
&\quad = \tau(f^n, v_h) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n-1}), v_h) - \tau(K(\theta(\psi_h^{n,i}))\nabla(\psi_h^{n,i}+z), \nabla v_h),
\end{aligned} \tag{15}
$$

for all $v_h \in V_h$. Inserting the test function $v_h = \delta\psi_h^{i+1}$ in (15), one has

$$
\begin{aligned}
\left|\!\left|\!\left|\delta\psi_h^{i+1}\right|\!\right|\!\right|^2_{N,\psi_h^{n,i}} &\overset{(12)}{=} \int_\Omega \left(\theta'(\psi_h^{n,i})|\delta\psi_h^{i+1}|^2 + \tau|K(\theta(\psi_h^{n,i}))^{\frac{1}{2}}\nabla\delta\psi_h^{i+1}|^2\right)\\
&\overset{(15)}{=} \underbrace{-\tau\left((K\circ\theta)'(\psi_h^{n,i})\nabla(\psi_h^{n,i}+z)\,\delta\psi_h^{i+1}, \nabla\delta\psi_h^{i+1}\right)}_{=:T_1}\\
&\quad + \underbrace{\tau(f^n, \delta\psi_h^{i+1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n-1}), \delta\psi_h^{i+1}) - \tau(K(\theta(\psi_h^{n,i}))\nabla(\psi_h^{n,i}+z), \nabla\delta\psi_h^{i+1})}_{=:T_2}.
\end{aligned} \tag{16a}
$$

Calling $\boldsymbol{\sigma}^i = (K\circ\theta)'(\psi_h^{n,i})\nabla(\psi_h^{n,i}+z)$ for brevity, we estimate that

$$
\begin{aligned}
T_1 &:= -\tau(\boldsymbol{\sigma}^i\delta\psi_h^{i+1}, \nabla\delta\psi_h^{i+1})\\
&\leq \left(\tau\int_\Omega |K(\theta(\psi_h^{n,i}))^{-\frac{1}{2}}\boldsymbol{\sigma}^i|^2(\delta\psi_h^{i+1})^2\right)^{\frac{1}{2}}\left(\tau\int_\Omega |K(\theta(\psi_h^{n,i}))^{\frac{1}{2}}\nabla\delta\psi_h^{i+1}|^2\right)^{\frac{1}{2}}\\
&\overset{(13)}{\leq} C_N^i\left(\int_\Omega \theta'(\psi_h^{n,i})(\delta\psi_h^{i+1})^2\right)^{\frac{1}{2}}\left(\tau\int_\Omega |K(\theta(\psi_h^{n,i}))^{\frac{1}{2}}\nabla\delta\psi_h^{i+1}|^2\right)^{\frac{1}{2}}\\
&\leq \frac{C_N^i}{2}\int_\Omega\left(\theta'(\psi_h^{n,i})|\delta\psi_h^{i+1}|^2 + \tau|K(\theta(\psi_h^{n,i}))^{\frac{1}{2}}\nabla\delta\psi_h^{i+1}|^2\right)\\
&= \frac{C_N^i}{2}\left|\!\left|\!\left|\delta\psi_h^{i+1}\right|\!\right|\!\right|^2_{N,\psi_h^{n,i}}.
\end{aligned} \tag{16b}
$$

For estimating the last term, we observe from the divergence theorem that

$$
\begin{aligned}
-(\boldsymbol{\sigma}_L^i, \nabla\delta\psi_h^{i+1}) &= (\nabla\cdot\boldsymbol{\sigma}_L^i, \delta\psi_h^{i+1})\\
&\overset{(14)}{=} \tfrac{1}{\tau}(\Pi_h(L(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1}))), \delta\psi_h^{i+1})_{\mathcal{T}_{\rm deg}^{i,\epsilon}}\\
&= \tfrac{1}{\tau}(L(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1})), \delta\psi_h^{i+1})_{\mathcal{T}_{\rm deg}^{i,\epsilon}}
\end{aligned}
$$

The last equality follows from the definition of the projection operator Π_h and $\delta\psi_h^{i+1} \in V_h \subset \mathcal{P}_p(\mathcal{T}_h)$. Using this result, along with (8) and $\delta\psi_h^{i+1} \in V_h$, one has

$$
\begin{aligned}
T_2 &:= \tau(f^n, \delta\psi_h^{i+1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n-1}), \delta\psi_h^{i+1}) - \tau(K(\theta(\psi_h^{n,i}))\nabla\psi_h^i, \nabla\delta\psi_h^{i+1})\\
&\overset{(8)}{=} (L(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1})), \delta\psi_h^{i+1})\\
&\qquad - \tau((K(\theta(\psi_h^{n,i})) - K(\theta(\psi_h^{n,i-1})))\nabla(\psi_h^{n,i}+z), \nabla\delta\psi_h^{i+1})\\
&= (L(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1})), \delta\psi_h^{i+1}) + \tau(\boldsymbol{\sigma}_L^i, \nabla\delta\psi_h^{i+1})\\
&\qquad - \tau((K(\theta(\psi_h^{n,i})) - K(\theta(\psi_h^{n,i-1})))\nabla(\psi_h^{n,i}+z) + \boldsymbol{\sigma}_L^i, \nabla\delta\psi_h^{i+1})\\
&= (L(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1})), \delta\psi_h^{i+1})_{\mathcal{T}_h\setminus\mathcal{T}_{\rm deg}^{i,\epsilon}}\\
&\qquad - \tau((K(\theta(\psi_h^{n,i})) - K(\theta(\psi_h^{n,i-1})))\nabla(\psi_h^{n,i}+z) + \boldsymbol{\sigma}_L^i, \nabla\delta\psi_h^{i+1})\\
&\overset{(14)}{\leq} (\theta'(\psi_h^{n,i})^{-\frac{1}{2}}(L(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1}))),\\
&\qquad \theta'(\psi_h^{n,i})^{\frac{1}{2}}\delta\psi_h^{i+1})_{\mathcal{T}_h\setminus\mathcal{T}_{\rm deg}^{i,\epsilon}} + \tau[\eta_{L\to N}^{i,\rm flux}]\,\|K(\psi_h^{n,i})^{\frac{1}{2}}\nabla\delta\psi_h^{i+1}\|\\
&\leq [\eta_{L\to N}^{i,\rm poten}]\cdot\|\theta'(\psi_h^{n,i})^{\frac{1}{2}}\delta\psi_h^{i+1}\| + \sqrt{\tau}\,[\eta_{L\to N}^{i,\rm flux}]\cdot\sqrt{\tau}\|K(\psi_h^{n,i})^{\frac{1}{2}}\nabla\delta\psi_h^{i+1}\|.
\end{aligned} \tag{16c}
$$

Combining (16), using the Cauchy-Schwarz inequality along with the definition of $\eta_{L\to N}^i$, one has the estimate. □

3.3. Newton to L-scheme switching estimate

Assuming that the L-scheme converges unconditionally, after switching to Newton we would want to switch back to the L-scheme only if linearization error of the Newton scheme increases with iterations. Similar to before, we can estimate if this is going to happen in the $(i+1)^{\rm th}$-step, purely from the iterates up to the $i^{\rm th}$-step. For this purpose, we introduce another equilibrated flux.

Definition 3.2 *(Equilibrated flux $\boldsymbol{\sigma}_N^i$ for degenerate regions (Newton scheme)).* Recalling Definition 3.1, we define $\boldsymbol{\sigma}_N^i \in \mathbf{RT}_p(\mathcal{T}_h) \cap \boldsymbol{H}(\mathrm{div},\Omega)$ as

$$
\nabla\cdot\boldsymbol{\sigma}_N^i = \begin{cases} \frac{1}{\tau}\Pi_h(\theta'(\psi_h^{n,i})(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1}))) & \text{in } \mathcal{T}_{\rm deg}^{i,\epsilon},\\ 0 & \text{otherwise}. \end{cases} \tag{17}
$$

The corresponding result mirroring Proposition 1 is

Proposition 2 *(Error control of Newton to Newton step). For a given $\psi_h^{n,0}, \psi_h^{n-1} \in V_h$, let $\{\psi_h^{n,j}\}_{j=1}^{i+1} \subset V_h$ solve (9) for some $i \in \mathbb{N}$. Then, under Assumptions 1–2, one has*

$$
\left|\!\left|\!\left|\psi_h^{n,i+1} - \psi_h^{n,i}\right|\!\right|\!\right|_{N,\psi_h^{n,i}} \leq \eta_{N\to L}^i,
$$

where

$$
\eta_{N\to L}^i := \tfrac{2}{(2-C_N^i)}\left([\eta_{N\to L}^{i,\rm poten}]^2 + \tau[\eta_{N\to L}^{i,\rm flux}]^2\right)^{\frac{1}{2}}
$$

with

$$
\begin{aligned}
\eta_{N\to L}^{i,\rm poten} &:= \|\theta'(\psi_h^{n,i})^{-\frac{1}{2}}(\theta'(\psi_h^{n,i-1})(\psi_h^{n,i} - \psi_h^{n,i-1})\\
&\qquad - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1})))\|_{\mathcal{T}_h\setminus\mathcal{T}_{\rm deg}^{i,\epsilon}},\\
\eta_{N\to L}^{i,\rm flux} &:= \left\| \begin{matrix} \Big[(K(\theta(\psi_h^{n,i})) - K(\theta(\psi_h^{n,i-1})))\nabla(\psi_h^{n,i}+z) \\ -(K\circ\theta)'(\psi_h^{n,i-1})(\psi_h^{n,i} - \psi_h^{n,i-1})\nabla(\psi_h^{n,i-1}+z))\Big]\,K(\theta(\psi_h^{n,i}))^{-\frac{1}{2}} \\ +K(\theta(\psi_h^{n,i}))^{-\frac{1}{2}}\boldsymbol{\sigma}_N^i \end{matrix} \right\|.
\end{aligned}
$$

The proof is identical to the proof of Proposition 1 and hence is left for the avid reader.

Remark 3 *(Effectivity of the estimators $\eta_{L\to N}^i$ and $\eta_{N\to L}^i$).* The estimators $\eta_{L\to N}^i$ and $\eta_{N\to L}^i$ predict the linearization error $\eta_{\rm lin}^{i+1}$ of the $(i+1)^{\rm th}$ iteration if done using the Newton scheme (9). In the cases where the iteration is done indeed using the Newton scheme, the sharpness of the estimate can be measured using the **effectivity index**, i.e., if $(i+1)^{\rm th}$ iteration is Newton then

$$
(\text{Eff. Ind.})_i := \begin{cases} \eta_{L\to N}^i/\eta_{\rm lin}^{i+1} & \text{if } i^{\rm th} \text{ iteration is L-scheme},\\ \eta_{N\to L}^i/\eta_{\rm lin}^{i+1} & \text{if } i^{\rm th} \text{ iteration is Newton}. \end{cases} \tag{18}
$$

Observe that it is always greater than 1 due to Propositions 1 and 2 and an effectivity index close to 1 implies a sharp estimate. The estimators are expected to be quite accurate since mainly the Cauchy-Schwarz inequality is used to derive them, except for estimate (16b) where the term T_1 is bounded above using the global approximation in Assumption 2. This expected sharpness is shown to be the case through the numerical experiments of Section 4, see in particular Figs. 5 and 8.

3.4. A-posteriori estimate based adaptive linearization algorithm

With the above estimates in mind, we propose a switching algorithm between the L-scheme and the Newton method. The linearization scheme used at iteration $j = i + 1$ should be Newton if the linearization error, predicted by the estimators $\eta^i_{L\to N}$ and $\eta^i_{N\to L}$, is smaller than the linearization error η^i_{lin} of the i^{th} step, see (10). However, to optimize the algorithm we take a few numerical considerations into account first.

3.4.1. *Computational considerations*

To speed up the computations of this switching criteria, we make a few more reductions

- **[Equilibrated flux]** If the saturated domain is much smaller than the unsaturated domain, then we take $\boldsymbol{\sigma}^i_L = \boldsymbol{\sigma}^i_N = 0$. Through a numerical example it is shown in Section 4.4 that this does not change the number of iterations required by the algorithm.
- **[Switching condition]** The condition $\eta^i_{L\to N} \le \eta^i_{\text{lin}}$ might be difficult to satisfy if the estimators are not sharp (see Remark 3), and even when it is satisfied it might require large values of i. Hence, to expedite the switching between L-scheme and Newton, we will use the criteria $\eta^i_{L\to N} < C_{\text{tol}}\,\eta^i_{\text{lin}}$ for a constant $C_{\text{tol}} > 1$.

3.4.2. *Adaptive linearization algorithm*

Under these considerations we propose the following adaptive algorithm:

Algorithm 1 L-scheme/Newton *a-posteriori switching.*

```
Require: $\psi^{n,0} \in L^2(\Omega)$ as initial guess.
Ensure: Scheme = $\boxed{\text{L-scheme}}$, $C_{\text{tol}} = 1.5$
  for i = 1,2,.. do
    if Scheme = $\boxed{\text{L-scheme}}$ then
      Compute iterate using $\boxed{\text{L-scheme}}$, i.e., (8)
      if $C^i_N \ge 2 - \textit{tolerance}$ then continue.
      else if $\eta^i_{L\to N} \le C_{\text{tol}}\eta^i_{\text{lin}}$ then
        Set Scheme = $\boxed{\text{Newton}}$
    else
      Compute iterate using $\boxed{\text{Newton}}$, i.e., (9)
      if $\eta^i_{N\to L} > \eta^i_{\text{lin}}$ then
        Set Scheme = $\boxed{\text{L-scheme}}$
```

Remark 4 *(Combining L-scheme adaptivity).* In Appendix A, we further propose an algorithm to adaptively select L in order to expedite the convergence of the L-scheme. This can directly be implemented in conjunction to Algorithm 1 to improve the convergence speed of the composite scheme. Nevertheless, we have refrained from combining these schemes for the ease of presentation.

Remark 5 *(Computational cost of the estimators).* In the non-degenerate case, the quantities C^i_N, $\eta^i_{L\to N}$ and $\eta^i_{N\to L}$, can be directly computed from the iterates $\psi^{n,i}_h$ and $\psi^{n,i-1}_h$ by inserting $\boldsymbol{\sigma}^i_L = \boldsymbol{\sigma}^i_N = 0$, see Propositions 1 and 2. Hence, the cost of computing the estimators is small in comparison to the cost of the iterations. Since the L-scheme iterations are less expensive than the Newton iterations, the L/N scheme generally performs better or similarly to the Newton scheme time-wise. This is evident from the numerical experiments, e.g. see Fig. 3b. In the degenerate case, global computation are required for computing $\boldsymbol{\sigma}^i_L$ and $\boldsymbol{\sigma}^i_N$ if they are used. We discuss the computation of these equilibrated fluxes in Appendix B and their computation can be made relatively inexpensive by precomputing the associated stiffness matrices. The computational cost for the estimators can be reduced even further by evaluating them only for selected iterations. Nevertheless, we do not pursue this option for the sake of simplicity.

4. Numerical results

In this section, we perform several numerical examples that demonstrate the robustness and efficiency of the proposed algorithm for switching between Newton's method and the L-scheme. This is done through careful comparison between the switching algorithm, hereafter called the L/N-scheme, the standard Newton method and the L-scheme. It is important to note that the L-scheme includes a tuning parameter that significantly affects the performance of the method. As a remedy, we choose two different values, L_1 and L_2 in the performance comparison. Here, L_1 is a quasi-optimal choice of tuning parameter and will be defined for each specific subproblem, see Table 2, and $L_2 = \sup\{\theta'(\psi)\}$. For the L/N-scheme, L_1 is always chosen for the L-scheme iterations. The linear systems arising from the linearization schemes are solved using a direct solver.

Table 2
Parameter values for all test cases. The parameters are presented in column format, where each column corresponds to the parameters for the specified test case.

Parameters	Test case 1	Test case 2	Test case 3	Test case 4
van Genuchten-Mualem				
θ_R	0.026	0.026	0.131	
θ_S	0.42	0.42	0.396	
K_S	0.12	0.12	$4.96 \cdot 10^{-2}$	
α	0.551	0.95	0.423	
n	2.9	2.9	2.06	
L-scheme				
L_1	0.1	0.15	$3.501 \cdot 10^{-3}$	0.25
$L_2 = L_\theta$	0.136	0.2341	$4.501 \cdot 10^{-3}$	0.33

To measure the performance of each separate method, we examine both the number of iterations and computational time that they require to satisfy the stopping criterion

$$\left|\left|\left|\psi^{n,j}_h - \psi^{n,j-1}_h\right|\right|\right|_{\mathcal{L},\psi^{n,j-1}_h} < 10^{-7},$$

where $|||\cdot|||_{\mathcal{L},\psi^{n,j-1}_h}$ is the iteration and linearization-dependent energy norm for the pressure head, with $\mathcal{L} \in \{L, N\}$. Here, the computational time covers the entire simulations and all experiments were performed on an Acer Swift 3, with an Intel core i7-1165G7-processor.

In total, four different test cases for the numerical experiments are considered:

- Test case 1: The first test case is taken from [35], although it is modified in the sense that we disregard surfactant transport. Here, the flow is always partially saturated.
- Test case 2: The second test case can be found in [1], and it considers extraction/injection above the water table.
- Test case 3: The third test case is a known problem that is studied in [1,36–38]. Here, a time-dependent Dirichlet boundary condition is used to describe the recharge of a groundwater reservoir from a drainage trench.
- Test case 4: The final test case considers a heterogeneous and anisotropic medium, it is also found in [30].

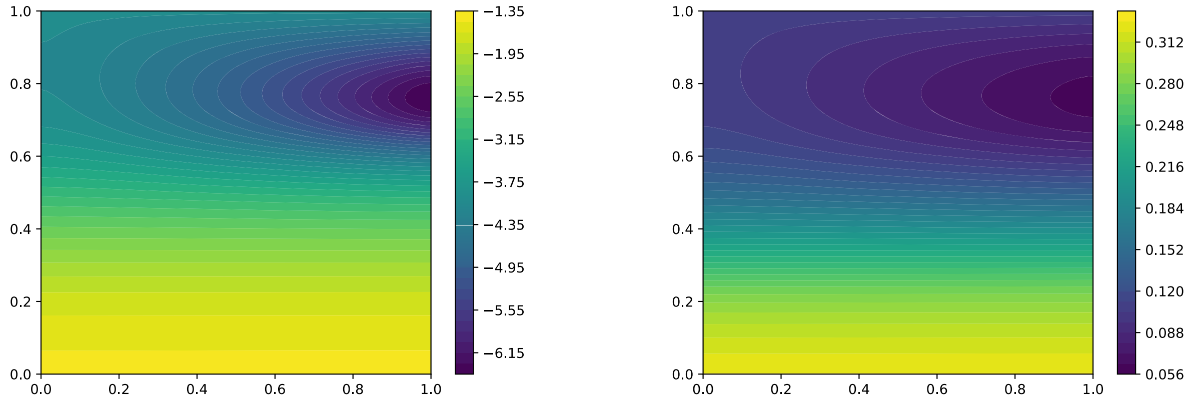
For the first three test cases, the van Genuchten-Mualem parametrization [33] is used to describe the relation between the saturation, the pressure head and the permeability,

$$\theta(\psi) = \begin{cases} \theta_R + (\theta_S - \theta_R)\left[\frac{1}{1+(-\alpha\psi)^n}\right]^{\frac{n-1}{n}}, & \psi \le 0, \\ \theta_S, & \psi > 0, \end{cases}$$
$$K(\Theta(\psi)) = \begin{cases} K_s\,(\Theta(\psi))^{\frac{1}{2}}\left[1 - \left(1 - \Theta(\psi)^{\frac{n}{n-1}}\right)^{\frac{n-1}{n}}\right]^2, & \psi \le 0, \\ K_s, & \psi > 0. \end{cases} \tag{19}$$

Here,

$$\Theta(\psi) = \frac{\theta(\psi) - \theta_R}{\theta_S - \theta_R},$$

with θ_S and θ_R being the water volume and the residual water content respectively, K_s the hydraulic conductivity of the fully saturated porous



(a) Pressure head profile ψ at final time $T = 1$.

(b) Saturation profile $\theta(\psi)$ at final time $T = 1$.

Fig. 2. Test case 1 - Strictly unsaturated medium.

medium, and α and n soil related parameters. In the last test case we use a parameterization similar to the Brooks-Corey model [32],

$$K(\boldsymbol{x},\theta)=\bar{\mathbf{K}}(\boldsymbol{x})\theta^3,\quad \theta(\psi)=\begin{cases}(2-\psi)^{-\frac{1}{3}}, & \text{if } \psi<1,\\ 1, & \text{if } \psi\geq 1,\end{cases} \tag{20}$$

where $\bar{\mathbf{K}}(\boldsymbol{x})$ is defined in (21). Assumption 2, i.e. the convection term being non-dominant, holds in all the examples. For the last three test cases the degenerate domain is nonempty, $\mathcal{T}^{i,\epsilon}_{deg} \neq \emptyset$. In practice the switching between linearization techniques is successful even without the computation of the equilibrated fluxes. Therefore we only compute the equilibriated flux in example 4, despite the switching mechanism working equally both with and without the flux.

In all of the test-cases, triangular linear conforming finite elements with mesh diameter h are applied together with the implicit Euler time-discretization with time step size τ, as described in Sections 2.1 and 2.2. The mesh diameter h and time step size τ vary between the different experiments and will be specified for each individual experiment. We note that the numerical experiments are expected to perform equivalently for other spatial discretization methods such as the Raviart-Thomas mixed finite elements or discontinuous Galerkin finite elements.

The finite element implementation is Python based and uses the simulation toolbox PorePy [39] for grid management and pyFreeFem [40] for the computation of the equilibriated fluxes. Unless it is stated, the default number of CPUs used is 8. It is available for download at https://github.com/MrShuffle/RichardsEquation/releases/tag/v1.0.1.

4.1. Test case 1 - strictly unsaturated medium

In this test case, we consider a strictly unsaturated porous medium, and use the van Genuchten-Mualem parametrization that is described by parameters from Table 2. The test case is heavily inspired by [35], and the domain is given by $\Omega = \Omega_1 \cup \Omega_2$, where $\Omega_1 = [0,1]\times[0,1/4]$ and $\Omega_2 = [0,1]\times(1/4,1]$. We consider the time interval $[0,T]$, where $T=\tau$ varies with choice of time step size τ, as we only take one time step. As initial condition, we choose the pressure head

$$\boldsymbol{\psi}^0(x,z)=\begin{cases}-z-1/4 & (x,z)\in\Omega_1\\ -4 & (x,z)\in\Omega_2,\end{cases}$$

where x represents the positional variable in the horizontal direction and z in the vertical direction. A Dirichlet boundary condition is imposed at the top boundary that complies with the initial condition. For the rest of the boundary no-flow boundary conditions are used, and the following source term is applied

$$f(x,z)=\begin{cases}0 & (x,z)\in\Omega_1\\ 0.06\cos\left(\frac{4}{3}\pi(z)\right)\sin(x) & (x,z)\in\Omega_2.\end{cases}$$

The solution after one time step with time step size $\tau = 1$, is given in Fig. 2a.

4.1.1. Comparison of convergence properties

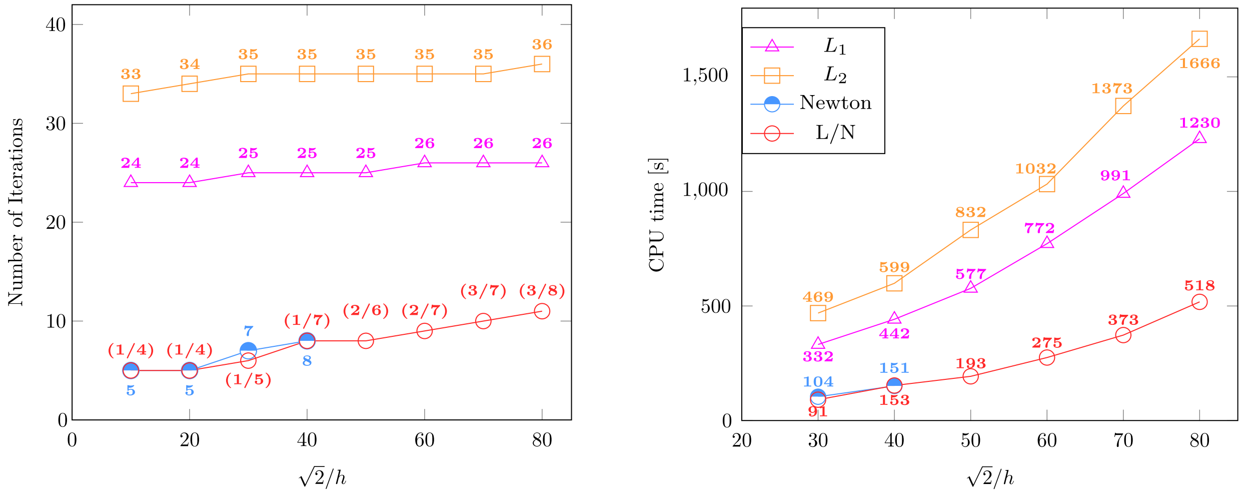
Here, we discuss the performance and convergence properties of the newly proposed L/N-scheme and compare it to the Newton method and the L-scheme. In Fig. 3a, the number of iterations for different choices of the mesh size parameters, with time step size $\tau = 0.01$ are presented. As expected the L-scheme is robust and converges in each scenario, for both L_1 and L_2. Newton's method, however, only converges for sufficiently coarse meshes. Yet, when converging, it converges in fewer iterations than the L-scheme. Finally, the hybrid L/N method converges in as few if not fewer iterations as the Newton method (when it converges) and converges robustly, and in far fewer iterations than the L-scheme for the other mesh sizes.

Furthermore, a similar experiment is performed for a fixed mesh size $h = \sqrt{2}/40$, and varying time step sizes, see Fig. 4a. For larger time step sizes the Newton method diverges, while the other methods converge robustly. Again the L/N-scheme converges with the performance expected of Newton's method, in addition to being as robust as the L-scheme. We highlight the enormous difference in the number of iterations for the largest time step size $\tau = 1$ in Fig. 4a.

Then, the performance of the linearization schemes is compared in terms of computational time, cf. Fig. 3b and Fig. 4b. One can observe virtually the same performance for the hybrid method as for Newton's method when the latter converges. The former in fact is sometimes slightly faster, due to each L-scheme iteration being slightly less expensive than a Newton iteration, see Remark 6. In addition, the hybrid method continues to show the same performance for the cases in which Newton's method does not converge. Finally, Fig. 3b shows that, for all meshes, the computational time of the L-schemes is consistent with the reported numbers of iterations in Fig. 3a with L_1 being the fastest. Although it uses more than double the computational time of the L/N-scheme.

Overall, the newly proposed L/N-scheme shows the best performance. It is as fast as Newton's method when it converges, and is significantly more robust.

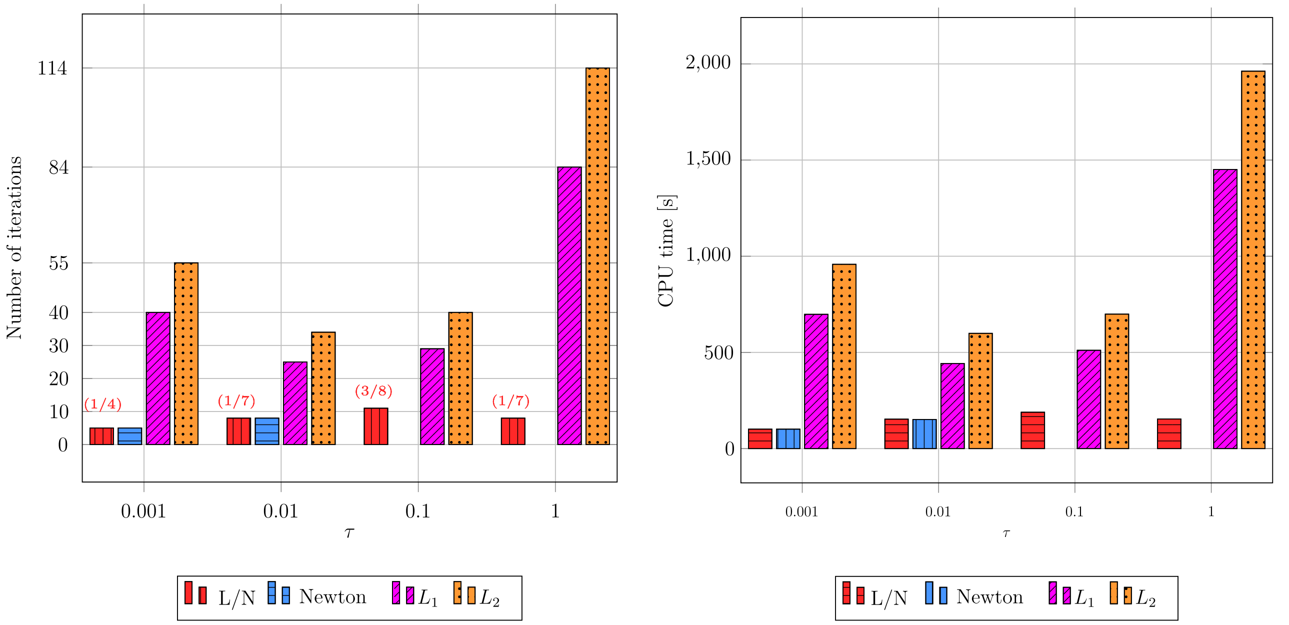
Remark 6 *(Computational time per iteration).* It is known that condition numbers for matrices coming from systems linearized by Newton's method are higher than for those linearized by the L-scheme [1]. Therefore, each iteration of Newton's method, when implemented without preconditioning, takes more time than each L-scheme iteration.



(a) Total number of iterations. The numbers in the red parentheses correspond to (number of L-scheme iterations/number of Newton iterations).

(b) Computational time in seconds.

Fig. 3. Test case 1 - Strictly unsaturated medium: Performance metrics for all linearization schemes for fixed $\tau = 0.01$ and varying mesh size.



(a) Number of iterations for different time step sizes.

(b) Total computational time in seconds for different time step sizes.

Fig. 4. Test case 1 - Strictly unsaturated medium: Performance comparison for all of the linearization schemes for different time step sizes and fixed mesh size $h = \sqrt{2}/40$.

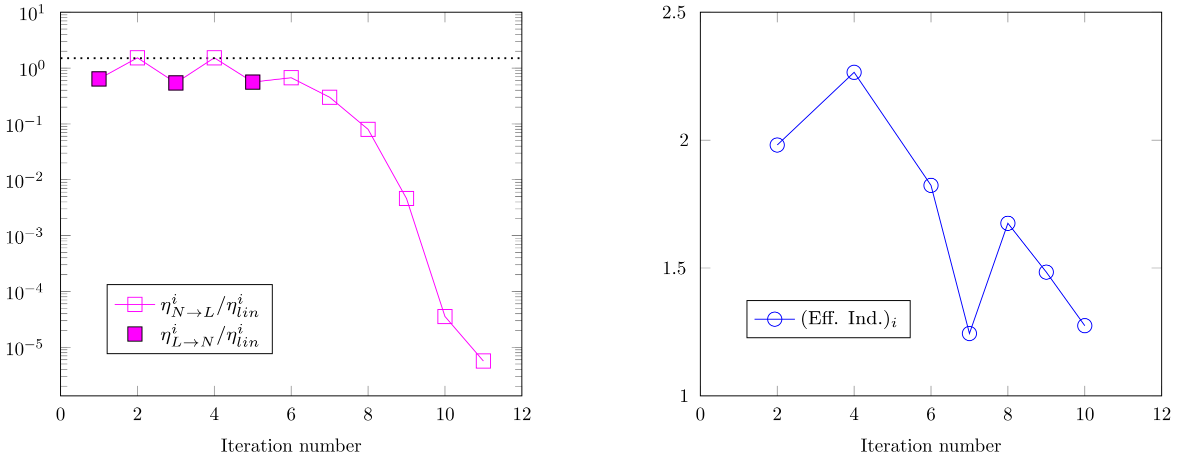
Remark 7 *(Computational time for the coarsest mesh).* The computational times of the coarsest meshes are omitted due to the use of multiprocessing in the implementations. This causes the most time consuming part to be the spawn process of the local assembly on each element. As a result, the computational times for the coarsest meshes are very similar for all the linearization methods.

4.1.2. Switching characteristics

Finally, the dynamic switch between the L-scheme and Newton's method is inspected in further detail. In Fig. 5, the evolution of the indicators for the switch is displayed for a fixed mesh and time step size. The example particularly demonstrates the ability of the hybrid method to switch back and forth between both linearizations before switching fully to Newton. In addition, the final number of L-scheme iterations is kept at its minimum. The plot also shows the effectivity indices introduced in (18) and discussed in Remark 3. The effectivity index is greater than 1 in all cases, which validates Propositions 1 and 2 and it stays between 1.27 to 2.3, implying that the estimators $\eta^i_{L\to N}$ and $\eta^i_{N\to L}$ are sharp.

4.2. Test case 2 - variably saturated medium

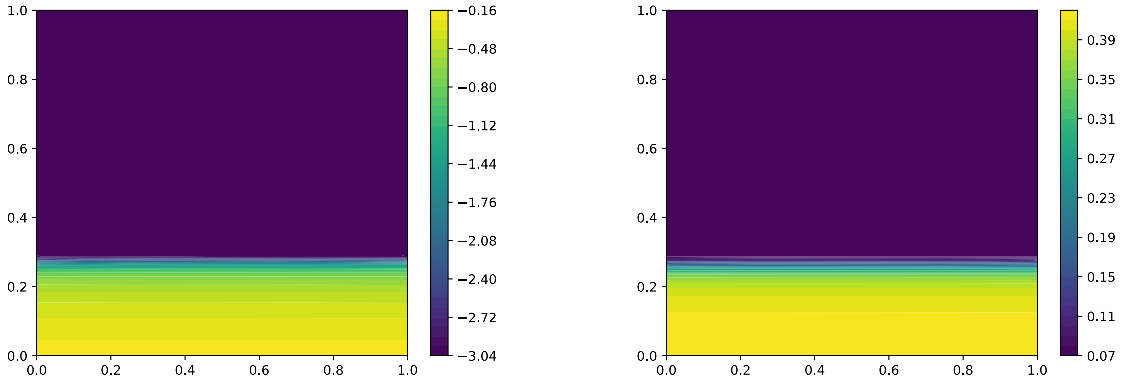
The example parameters are as in Table 2, Test case 2. We consider a variably saturated medium, $\Omega = \Omega_{gw} \cup \Omega_{vad}$, where the groundwater zone is $\Omega_{gw} = [0,1] \times [0,1/4)$ and a vadose zone is $\Omega_{vad} = [0,1] \times [1/4,1]$. Here, we consider the time interval $[0,T]$, where $T = 0.01$ and we only take one time step with $\tau = 0.01$. As initial condition, we choose the pressure head



(a) Evolution of switching indicators for L/N-scheme where the dashed line is $C_{\text{tol}} = 1.5$. The L/N-scheme oscillates between the linearization strategies, but eventually recovers.

(b) Efficency index. Notice that the iterations correspond to the ones in Figure 5(a), and that only the ones where the Newton method is performed are counted, i.e., iteration 1,3 and 5 are removed.

Fig. 5. Test case 1 - Strictly unsaturated medium: Evolution of switching indicators for the L/N-scheme and efficiency indices (18) for the Newton iterations (see Remark 3). Here, the mesh size is $h = \sqrt{2}/80$ and time step size $\tau = 0.01$.



(a) Pressure head profile ψ at $T = 0.01$.

(b) Saturation profile $\theta(\psi)$ at final time $T = 0.01$.

Fig. 6. Test case 2 - Variably saturated medium.

$$\boldsymbol{\psi}^0(x,z) = \begin{cases} -z + 1/4 & (x,z) \in \Omega_{gw} \\ -3 & (x,z) \in \Omega_{vad}, \end{cases}$$

where x represents the positional variable in the horizontal direction and z in the vertical direction. On the surface a constant Dirichlet boundary condition is imposed, being equal to the initial condition at all times. For the rest of the boundary no-flow boundary conditions are used. We apply the following source term

$$f(x,z) = \begin{cases} 0 & (x,z) \in \Omega_{gw} \\ 0.006\cos\left(\frac{4}{3}\pi(z-1)\right)\sin(2\pi x) & (x,z) \in \Omega_{vad}. \end{cases}$$

After one time step the pressure head profile is given in Fig. 6.

4.2.1. Comparison of convergence properties

The iteration count for the second test case for different mesh sizes and fixed time step for all linearization schemes is illustrated in Fig. 7a. Again the L-scheme converges in every case. However, Newton's method does not converge for any mesh size. The hybrid method needs the fewest number of iterations, which shows that the dynamic switch is successful.

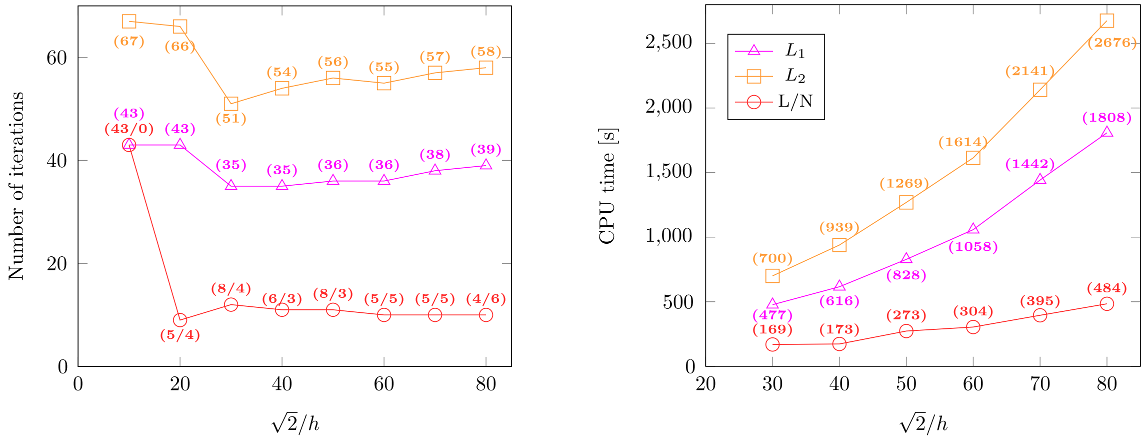
The CPU time performance of the linearization schemes is compared in Fig. 7b. Both versions of the L-scheme take computational times consistent with the number of iterations, with the simulations with the parameter L_1 being less expensive. However, the L-scheme (using L_1) requires approximately 373% of the computational time of the hybrid method including the computation of the switching indicators. In addition, the benefit of a few additional L-scheme iterations further decreases the computational time of the hybrid method.

4.2.2. Switching characteristics

We also give a more in-depth look to the dynamic switch between the Newton's method and the L-scheme. In Fig. 8, the evolution of the switching indicators is shown for a fixed time step and a fixed mesh size. After 8 L-scheme iterations the switching indicator $\eta_{L\to N}$ becomes lower than C_{tol} and then Newton's method converges. From Fig. 7a the number of L-scheme iterations required before the switching indicator becomes small enough to switch to Newton's method varies with the mesh size. Note that for the coarsest mesh no switch to Newton's method happens. This is due to $\eta_{L\to N}/\eta^i_{lin}$ approaching C_{tol}, but never becoming smaller.

4.3. Test case 3 - recharge of a groundwater reservoir

Here, we consider a known problem [38], also used e.g. in [1], which models the recharge of a groundwater reservoir from a drainage



(a) Total number of iterations. The numbers in the red parentheses correspond to (number of L-scheme iterations/number of Newton iterations).

(b) Computational time in seconds.

Fig. 7. Test case 2 - Variably saturated medium: Performance metrics for all linearization schemes for fixed $\tau = 0.01$ and varying mesh size.

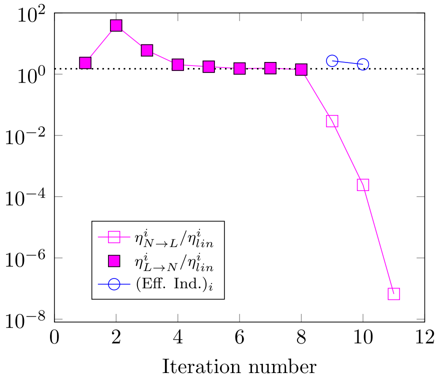


Fig. 8. Test case 2 - Variably saturated medium: Evolution of switching indicators for L/N-scheme for fixed $h = \sqrt{2}/50$ and $\tau = 0.01$. The dashed line is $C_{\text{tol}} = 1.5$, the switching criterion from L-scheme to Newton's method. The effectivity indices (18) corresponding to the Newton iterations are also plotted and they remain below 2.8.

trench in two spatial dimensions. The domain $\Omega \subset \mathbb{R}^2$ represents a vertical segment of the subsurface. One portion of the right side of the domain is fixed by a constant Dirichlet boundary condition. A time-dependent Dirichlet boundary condition on parts of the upper boundary is used to mimic the drainage trench. No-flow conditions are utilized on the remaining parts of the boundary. The used parameters are given in Table 2 Test case 3, corresponding to silt loam. The geometry is given by

$$\Omega = [0,2] \times [0,3],$$

$$\Gamma_{D_1} = [0,1] \times (3),$$

$$\Gamma_{D_2} = (2) \times [0,1],$$

$$\Gamma_N = \Omega \backslash \left\{ \Gamma_{D_1} \cup \Gamma_{D_2} \right\},$$

and the initial pressure head distribution and boundary conditions are

$$\psi(0,x,z) = 1 - z$$

$$\psi(t,x,z) = \begin{cases} -2 + 35.2t, & \text{if } t \le \frac{1}{16}, \quad \text{on } \Gamma_{D_1}, \\ 0.2, & \text{if } t > \frac{1}{16}, \quad \text{on } \Gamma_{D_1}, \\ 1 - z, & \text{on } \Gamma_{D_2}, \end{cases}$$

$$-K(\theta(\psi(t,x,z)))\nabla(\psi(t,x,z) + z) \cdot \nu = 0, \quad \text{on } \Gamma_N,$$

Table 3
Test case 3 - Recharge of a groundwater reservoir: Average number of iterations per time step, total number of iterations and computational time for 2501 nodes.

	No. Itr	CPU time [s]
L_1	274	6136
L_2	330	7356
Newton	39	980
L/N	(10/30)	1021

where ν is the outward normal vector. The solution is computed over 9 timesteps, where the time unit is in days, with time step size $\tau = 1/48$ and with a regular mesh consisting of 2501 nodes. The pressure head and saturation profile at the final time for the L/N-scheme is shown in Fig. 9.

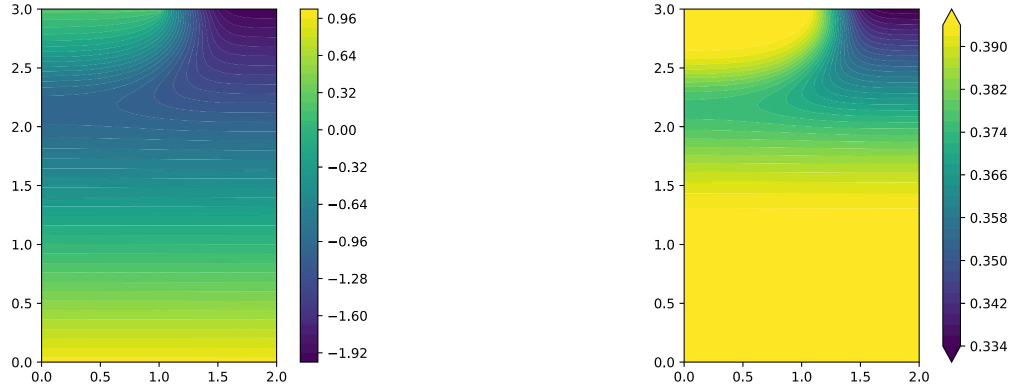
4.3.1. *Comparison of convergence properties*

The performance of all schemes for test case 3 is displayed in Table 3. All schemes converge for this example. The Newton method requires the least amount of iterations. However, the hybrid method only needs one more iteration. Both uses significantly less iterations than the L-schemes. For all time steps except one, only one L-scheme iteration is needed per time step, which indicates a successful dynamic switch for almost all time steps. The evolution in time of the schemes can be seen in Fig. 10, where the hybrid scheme and Newton's method become slightly better with time due to better initial guesses. However, the L-schemes use significantly more iterations as the problem becomes more nonlinear with time.

The computational time for the L-schemes is much higher than both Newton's method and the hybrid method, which is consistent with the expense per iteration discussed in Remark 6. More significantly, the L/N-scheme performs almost the same as Newton's method.

4.4. *Test case 4 - heterogeneous and anisotropic medium*

For this test case we consider a heterogeneous and anisotropic medium, also used in [30]. Here, the CPUs have been load balanced to optimize computational time for all schemes on the given mesh size. We consider permeability and saturation functions which are similar to the Brooks-Corey model (20) and a zero source term, i.e. $f = 0$. The domain is the unit square and the medium is made heterogeneous and anisotropic by



(a) Pressure head profile ψ at 4.5 hours.

(b) Saturation profile $\theta(\psi)$ at 4.5 hours.

Fig. 9. Test case 3 - Recharge of a groundwater reservoir.

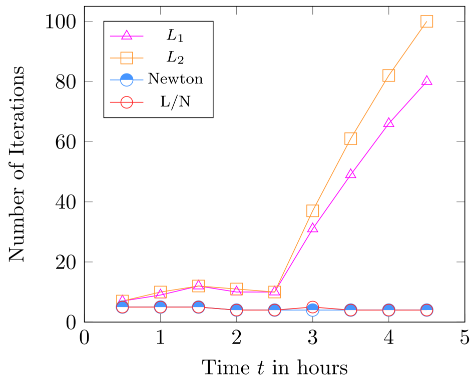


Fig. 10. Test case 3 - Recharge of a groundwater reservoir: Number of iterations per time step.

$$\bar{\mathbf{K}} = \begin{cases} \bar{\mathbf{K}}_1, & \text{if } z > 0.5, \\ K_\phi \boldsymbol{Q} \bar{\mathbf{K}}_1 \boldsymbol{Q}^T, & \text{if } z <= 0.5, \end{cases} \qquad \text{where } \bar{\mathbf{K}}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \end{pmatrix}, \quad \boldsymbol{Q} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}, \tag{21}$$

where α reflects a slanted alignment of the principle axes of $\bar{\mathbf{K}}$ and K_ϕ signifies a change in porosity. We chose $\alpha = \pi/3$ and $K_\phi = 0.1$ to be fixed. The initial condition is discontinuous,

$$\psi(0, x, z) = \begin{cases} 0.9, & \text{if } z > 0.5, \\ -3, & \text{if } z \leq 0.5, \end{cases}$$

and we take a constant Dirichlet boundary condition on $\Gamma_{D_1} = (0, 0.5) \times (1)$ and $\Gamma_{D_2} = (1) \times (0, 0.5)$ being equal to the initial condition. For the remainder of the boundary no-flow conditions are used. We compute the solution on a uniform mesh with diameter $h = \sqrt{2}/80$ over 20 time steps using a time step size of $\tau = 0.1$. The pressure and saturation profile at time $T = 1$ is visualized in Fig. 11. At the interface $z = 0.5$, close to $x = 0$, degeneracy occurs due to the jump in $\bar{\mathbf{K}}$ and as a result of the no-flow boundary condition.

4.4.1. Comparison of performance properties

In this example, all of the schemes converge. The performance metrics are displayed in Table 4. The L-scheme uses considerably more iterations than Newton's method and the hybrid scheme, where Newton's method uses one less iteration than the hybrid. As the hybrid scheme only uses one L-scheme iteration per time step the switching is successful. In Fig. 12 the evolution of the schemes' performance with time is visualized, with number of iterations per time step. The hybrid scheme and Newton's method uses fewer iterations with time. However, similar to the previous test case, the L-scheme's convergence rate becomes slower after $t = 1.5$ due to increased nonlinearity of the problem.

Table 4
Test case 4 - Heterogeneous and anisotropic medium: Total number of iterations and total computational time.

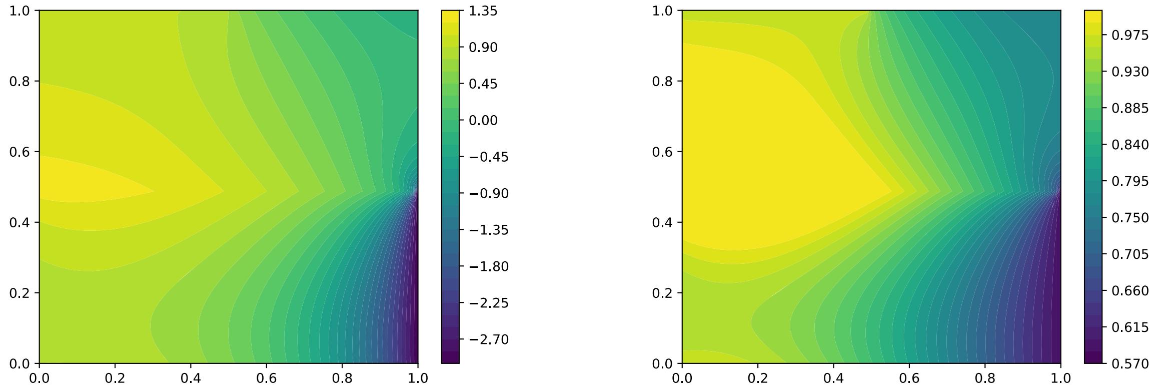
	No. Itr	CPU time [s]
L_1	393	8189
L_2	508	10686
Newton	137	3401
L/N	(20/118)	3097

The hybrid scheme performs the best for this example in terms of computational time. Newton's method is slower, but still significantly faster than the L-schemes. The computational cost of the equilibrated flux is small in comparison with assembly and solution of the linear system for this example. However, it is worth noting that the hybrid scheme uses the same number of iterations if the equilibrated flux is not computed, i.e., if $\boldsymbol{\sigma}^i_L = \boldsymbol{\sigma}^i_N = 0$ is inserted. Hence, this choice could further decrease the computational time.

5. Conclusions

In this paper, we considered solving Richards' equation, which models the flow of water through saturated/unsaturated porous media (soil). After applying backward Euler time-discretization and continuous Galerkin finite element space-discretization to Richards' equation, to solve the resulting nonlinear finite-dimensional problem we developed a hybrid iterative linearization strategy that combines the L-scheme with the Newton method. The idea behind this is to use the robust, but only first-order convergent L-scheme to stabilize the quadratically convergent Newton method. The switching between the two schemes is done in an adaptive manner using *a posteriori* indicators which predict the linearization error of the next iteration using a concept of iteration-dependent energy norms. After each iteration, it is checked whether the Newton method is predicted to decrease the linearization error of the next iteration. If so, then the Newton method is used, otherwise, the iteration is done using the L-scheme. The hybrid scheme is now robust, but still quadratically convergent after switching to the Newton scheme.

The performance of the hybrid scheme is tested on illustrative, realistic numerical examples which reveal that the scheme is as robust as the L-scheme and it converges in cases where Newton fails. Moreover, in cases when Newton converges, the hybrid scheme takes roughly the same amount of iterations and computational time and is considerably faster than even the optimized L-scheme. Lastly, we comment that the



(a) Pressure head profile ψ at $T = 1$.

(b) Saturation profile $\theta(\psi)$ at $T = 1$.

Fig. 11. Test case 4 - Heterogeneous and anisotropic medium.

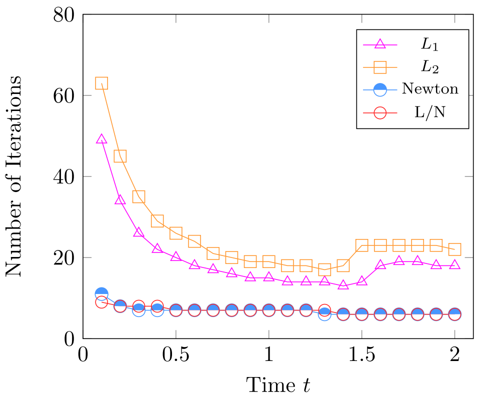


Fig. 12. Test case 4 - Heterogeneous and anisotropic medium: Number of iterations per time step.

scheme is quite general as it can, in principle, be extended to other spatial discretization and linearization methods.

Data availability

No data was used for the research described in the article.

Acknowledgements

JSS and FAR acknowledge the support of the VISTA program, The Norwegian Academy of Science and Letters and Equinor, JWB is funded in part through the Center of Sustainable Subsurface Resources (Norwegian Research Council project 331841) and the 'FracFlow' project funded by Equinor, Norway through Akademiaavtalen. KM acknowledges the support of FWO (Fonds Wetenschappelijk Onderzoek) for funding him through the 'Junior Postdoctoral Fellowship' (project code 1209322N) and to Akademiaavtalen for funding his visit to the University of Bergen.

Appendix A. An adaptive L-scheme

As discussed in Sections 1 and 2.3.1, the L-scheme converges unconditionally provided that $L \geq \frac{1}{2}\sup_{\xi\in\mathbb{R}} \theta'(\xi)$ and the time step size τ is smaller than a constant independent of the mesh size. However, numerical results in [1] suggest that the optimal rate of convergence of the L-scheme is obtained for a considerably smaller L although convergence cannot always be guaranteed for such values. Hence, to speed up the computations, it is possible to start the iterations with a smaller value of L and then use the *a posteriori* estimates to decide if L is to be increased or not. Analogous to Propositions 1 and 2 we state a result that allows us to do this rigorously.

Proposition 3 *(Error control of L-scheme). For a given $\psi_h^{n,0}, \psi_h^{n-1} \in V_h$, let $\{\psi_h^{n,j}\}_{j=1}^{i+1} \subset V_h$ solve (8) for some $i \in \mathbb{N}$. Then under Assumption 1,*

$$\left\|\left|\psi_h^{n,i+1} - \psi_h^{n,i}\right|\right\|_{L,\psi_h^{n,i}} \leq \eta^i_{L\to L},$$

where

$$\eta^i_{L\to L} := \left([\eta^{i,\text{poten}}_{L\to L}]^2 + \tau[\eta^{i,\text{flux}}_{L\to L}]^2\right)^{\frac{1}{2}}$$

with

$$\eta^{i,\text{poten}}_{L\to L} := \|L^{-\frac{1}{2}}(L(\psi_h^{n,i} - \psi_h^{n,i-1}) - (\theta(\psi_h^{n,i}) - \theta(\psi_h^{n,i-1})))\|,$$

$$\eta^{i,\text{flux}}_{L\to L} := \left\|(K(\theta(\psi_h^{n,i})) - K(\theta(\psi_h^{n,i-1})))K(\theta(\psi_h^{n,i}))^{-\frac{1}{2}}\nabla(\psi_h^{n,i} + z)\right\|.$$

The detailed proof is again omitted. Observe that for the estimate above, neither Assumption 2 nor any separate treatment of the degenerate domains is required.

A.1. L-adaptive algorithm

Based on Proposition 3, we propose an algorithm that selects optimal L-values adaptively.

Algorithm 2 The L-adaptive scheme.

```
Require: $\psi^{n,0} \in L^2(\Omega)$ as initial guess, $L_M := \sup_{\psi\in\mathbb{R}} \theta'(\psi)$, and $L_m := L_M/8$
Ensure: $C_{L\to L} = \sqrt{2}$, $L = L_m$
  for i = 1,2,.. do
      Compute iterate using L-scheme, i.e., (8)
      if $\eta^i_{L\to L} > \eta^i_{\text{lin}}$ then
          Replace $L_m = L$, $L = \min(C_{L\to L} L, L_M)$, and continue.
      else if $\eta^j_{L\to L} > 0.8\,\eta^j_{\text{lin}}$ for $j \in \{i, i-1, i-2\}$ then
          Replace $L = \max(0.9L, 1.1L_m)$ and continue.
```

A.2. Numerical result

In Fig. 13 we show a result where the L-adaptive scheme is superior to a fixed L-approach. In this case, $L_\theta/2$ is too small for convergence due to a large time step size. Compared with fixed L_1 with the same mesh size and time step size, see Fig. 4, the number of iterations is improved by 20. For smaller time steps, the numerical results reveal that Algorithm 2 results in roughly the same number of iterations compared to a fixed and optimized $L = L_1$ lesser than L_θ. But in all examples considered, it uses fewer iterations than simply choosing $L = L_2 = L_\theta$.

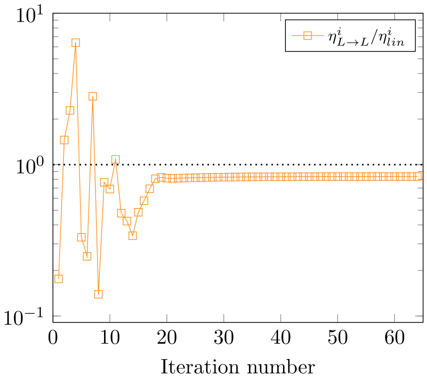


Fig. 13. Test case 1 - Strictly unsaturated medium: L-scheme with L-adaptivity and initial stabilization parameter $L_0 = L_2/8$, $h = \sqrt{2}/40$ and $\tau = 1$.

The advantage of such an adaptive technique is that an optimization study of L does not need to be conducted prior to the simulation. However, since the L-adaptive strategy does not significantly improve the behavior of the L-scheme over the optimized $L = L_1$, we refrained from including it in Algorithm 1 for the sake of simplicity.

Appendix B. Computation of equilibrated flux

Recalling Definitions 3.1 and 3.2, let us propose a simple algorithm to compute an equilibrated flux $\boldsymbol{\sigma}_h \in \mathbf{RT}_p(\mathcal{T}_h) \cap \boldsymbol{H}(\mathrm{div}, \Omega)$ satisfying $\nabla \cdot \boldsymbol{\sigma}_h = \Pi_h f$ in $\mathcal{T}^{i,\epsilon}_{\mathrm{deg}}$, and $\nabla \cdot \boldsymbol{\sigma}_h = 0$ otherwise, where $f \in L^2(\Omega)$. Defining $\boldsymbol{Q}_h := \mathbf{RT}_p(\mathcal{T}_h) \cap \boldsymbol{H}(\mathrm{div}, \Omega)$ and $\tilde{V}_h := \{v_h \in \mathcal{P}_p(\mathcal{T}_h) | \operatorname{Tr}_{\partial\Omega}(v_h) = 0\}$, we seek a pair $(\boldsymbol{\sigma}_h, r_h) \in \boldsymbol{Q}_h \times \tilde{V}_h$ that satisfies the mixed finite element problem,

$$(K(1)^{-1}\boldsymbol{\sigma}_h, \boldsymbol{q}_h) = (r_h, \nabla \cdot \boldsymbol{q}_h), \qquad \forall \boldsymbol{q}_h \in \boldsymbol{Q}_h, \tag{22a}$$

$$(\nabla \cdot \boldsymbol{\sigma}_h, v_h) = (f, v_h), \qquad \forall v_h \in \tilde{V}_h. \tag{22b}$$

The advantage of this flux is that it minimizes $\|K(1)^{-\frac{1}{2}}\boldsymbol{\sigma}_h\|$ which appears in the estimates in Propositions 1 and 2. For practical purposes, a much coarser mesh can be used outside of $\mathcal{T}^{i,\epsilon}_{\mathrm{deg}}$ to compute it, and the stiffness matrix can be precomputed to accelerate the computation.

References

[1] F. List, F.A. Radu, A study on iterative methods for solving Richards' equation, Comput. Geosci. 20 (2) (2016) 341–353.
[2] M.W. Farthing, F.L. Ogden, Numerical solution of Richards' equation: a review of advances and challenges, Soil Sci. Soc. Am. J. 81 (6) (2017) 1257–1269.
[3] H.W. Alt, S. Luckhaus, Quasilinear elliptic-parabolic differential equations, Math. Z. 183 (3) (1983) 311–341.
[4] H.W. Alt, S. Luckhaus, A. Visintin, On nonstationary flow through porous media, Ann. Mat. Pura Appl. 136 (1) (1984) 303–316.
[5] F.A. Radu, I.S. Pop, P. Knabner, Error estimates for a mixed finite element discretization of some degenerate parabolic equations, Numer. Math. 109 (2008) 285–311.
[6] T. Arbogast, An error analysis for Galerkin approximations to an equation of mixed elliptic-parabolic type, Technical Report TR90-33, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 1990.
[7] T. Arbogast, M. Obeyesekere, M.F. Wheeler, Numerical methods for the simulation of flow in root-soil systems, SIAM J. Numer. Anal. 30 (1993) 1677–1702.
[8] T. Arbogast, M.F. Wheeler, N.Y. Zhang, A non-linear mixed finite element method for a degenerate parabolic equation arising in flow in porous media, SIAM J. Numer. Anal. 33 (1996) 1669–1687.
[9] C. Woodward, C. Dawson, Analysis of expanded mixed finite element methods for a non-linear parabolic equation modeling flow into variably saturated porous media, SIAM J. Numer. Anal. 37 (2000) 701–724.
[10] F.A. Radu, I.S. Pop, P. Knabner, On the convergence of the Newton method for the mixed finite element discretization of a class of degenerate parabolic equation, Numer. Math. Adv. Appl. 42 (2006) 1194–1200.
[11] F.A. Radu, W. Wang, Error estimates for a mixed finite element discretization of some degenerate parabolic equations, Nonlinear Anal., Real World Appl. 15 (2014) 266–275.
[12] M. Bause, P. Knabner, Computation of variably saturated subsurface flow by adaptive mixed hybrid finite element methods, Adv. Water Resour. 27 (2004) 565–581.
[13] R. Eymard, M. Gutnic, D. Hilhorst, The finite volume method for Richards equation, Comput. Geosci. 3 (3–4) (1999) 259–294.
[14] R. Eymard, D. Hilhorst, M. Vohralik, A combined finite volume-nonconforming/mixed-hybrid finite element scheme for degenerate parabolic problems, Numer. Math. 105 (1) (2006) 73–131.
[15] S. Bassetto, C. Cancès, G. Enchéry, Q.-H. Tran, On several numerical strategies to solve Richards' equation in heterogeneous media with finite volumes, Comput. Geosci. 26 (5) (2022) 1297–1322.
[16] R.A. Klausen, F.A. Radu, G.T. Eigestad, Convergence of MPFA on triangulations and for Richards' equation, Int. J. Numer. Methods Fluids 58 (2008) 1327–1351.
[17] L. Bergamaschi, M. Putti, Mixed finite elements and Newton-type linearizations for the solution of Richards' equation, Int. J. Numer. Methods Eng. 45 (8) (1999) 1025–1046.
[18] F. Lehmann, P. Ackerer, Comparison of iterative methods for improved solutions of the fluid flow equation in partially saturated porous media, Transp. Porous Media 31 (3) (1998) 275–292.
[19] K. Mitra, I.S. Pop, A modified L-scheme to solve nonlinear diffusion problems, Comput. Math. Appl. (1987) 77 (6) (2019) 1722–1738.
[20] K. Brenner, C. Cances, Improving Newton's method performance by parametrization: the case of the Richards equation, SIAM J. Numer. Anal. 55 (4) (2017) 1760–1785.
[21] X. Wang, H.A. Tchelepi, Trust-region based solver for nonlinear transport in heterogeneous porous media, J. Comput. Phys. 253 (2013) 114–137.
[22] M. Celia, E. Bouloutas, R. Zarba, General mass-conservative numerical solution for the unsaturated flow equation, Water Resour. Res. 26 (7) (1990) 1483–1496.
[23] I.S. Pop, F.A. Radu, P. Knabner, Mixed finite elements for the Richards' equation: linearization procedure, J. Comput. Appl. Math. 168 (1–2) (2004) 365–373.
[24] M. Slodicka, A robust and efficient linearization scheme for doubly nonlinear and degenerate parabolic problems arising in flow in porous media, SIAM J. Sci. Comput. 23 (5) (2002) 1593–1614.
[25] D.G. Anderson, Iterative procedures for nonlinear integral equations, J. ACM (JACM) 12 (4) (1965) 547–560.
[26] J.W. Both, K. Kumar, J.M. Nordbotten, F.A. Radu, Anderson accelerated fixed-stress splitting schemes for consolidation of unsaturated porous media, Comput. Math. Appl. 77 (6) (2019) 1479–1502.
[27] C. Evans, S. Pollock, L.G. Rebholz, M. Xiao, A proof that Anderson acceleration improves the convergence rate in linearly converging fixed-point methods (but not in those converging quadratically), SIAM J. Numer. Anal. 58 (1) (2020) 788–810.
[28] S. Pollock, L.G. Rebholz, Anderson acceleration for contractive and noncontractive operators, IMA J. Numer. Anal. 41 (4) (2021) 2841–2872.
[29] W. Jäger, J. Kačur, Solution of doubly nonlinear and degenerate parabolic problems by relaxation schemes, ESAIM, Math. Model. Numer. Anal. 29 (5) (1995) 605–627.
[30] K. Mitra, M. Vohralík, A posteriori error estimates for the Richards equation, working paper or preprint, Aug. 2021, https://hal.inria.fr/hal-03328944.
[31] K. Mitra, M. Vohralík, Guaranteed, locally efficient, and robust a posteriori estimates for nonlinear elliptic problems in iteration-dependent norms: an orthogonal decomposition result based on iterative linearization, https://inria.hal.science/hal-04156711, 2023.
[32] R. Brooks, A. Corey, Properties of porous media affecting fluid flow, J. Irrig. Drain. Div. 92 (2) (1966) 61–90.
[33] M.T. van Genuchten, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Sci. Soc. Am. J. 44 (5) (1980) 892–898.
[34] J.W. Both, M. Borregales, J.M. Nordbotten, K. Kumar, F.A. Radu, Robust fixed stress splitting for Biot's equations in heterogeneous media, Appl. Math. Lett. 68 (2017) 101–108.
[35] D. Illiano, I.S. Pop, F.A. Radu, Iterative schemes for surfactant transport in porous media, Comput. Geosci. 25 (2) (2021) 805–822.
[36] P. Knabner, Finite Element Simulation of Saturated-Unsaturated Flow Through Porous Media, Birkhäuser, Boston, 1987, pp. 83–93, Ch. 6.
[37] R. Haverkamp, M. Vauclin, J. Touma, P.J. Wierenga, G. Vachaud, A comparison of numerical simulation models for one-dimensional infiltration, Soil Sci. Soc. Am. J. 41 (2) (1977) 285–294.
[38] E. Schneid, Hybrid-Gemischte Finite-Elemente-Diskretisierung der Richards-Gleichung, Naturwissenschaftliche Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg, 2000.
[39] E. Keilegavlen, R. Berge, A. Fumagalli, M. Starnoni, I. Stefansson, J. Varela, I. Berre, Porepy: an open-source software for simulation of multiphysics processes in fractured porous media, Comput. Geosci. 25 (1) (2021) 243–265.
[40] A. Abramian, O. Devauchelle, G. Seizilles, E. Lajeunesse, Boltzmann distribution of sediment transport, Phys. Rev. Lett. 123 (2019), https://doi.org/10.1103/PhysRevLett.123.014501.