Analysis of an Euler implicit-mixed finite element scheme for reactive solute transport in porous media
Radu, A.F.; Pop, I.S.; Attinger, S.

Published: 01/01/2008

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 12. Sep. 2017
ANALYSIS OF AN EULER IMPLICIT - MIXED FINITE ELEMENT SCHEME FOR REACTIVE SOLUTE TRANSPORT IN POROUS MEDIA

FLORIN A. RADU¹,², IULIU SORIN POP³, AND SABINE ATTINGER¹,²

Abstract. In this paper we analyze an Euler implicit-mixed finite element scheme for a porous media solute transport model. The transporting flux is not assumed given, but obtained by solving numerically the Richards equation, a model for sub-surface fluid flow. We prove the convergence of the scheme by estimating the error in terms of the discretization parameters. In doing so we take into account the numerical error occurring in the approximation of the fluid flow. The paper is concluded by numerical experiments, which are in good agreement with the theoretical estimates.

1. Introduction. The need of accurate and efficient numerical schemes to solve reaction-convection-diffusion equations modelling transport in porous media is well recognized. Inappropriate numerical methods can lead to false predictions for the evolution of a contaminated site, as revealed e.g. in [11]. Due to their local mass conservation property, mixed finite element methods are a valuable discretization technique for problems involving flow in porous media [3, 4, 5, 11, 12, 16, 17, 31, 33, 37, 41, 42]. Here, we present and analyze an Euler implicit-mixed finite element scheme (EI-MFE) that is based on the lowest order Raviart-Thomas (RT₀) elements for the equation

\[ \partial_t (\Theta(\psi)c) - \nabla \cdot (D \nabla c - Qc) = \Theta(\psi)r(c) \quad \text{in} \ J \times \Omega, \]

with \( c \) denoting the concentration of the solute, \( D \) the diffusion-dispersion coefficient and \( r(\cdot) \) a reaction term. Here \( J = (0, T] \ (0 < T < \infty) \) is the time interval, whereas \( \Omega \subset \mathbb{R}^d \ (d \geq 1) \) is the computational domain having a Lipschitz continuous boundary \( \Gamma \). The initial and boundary conditions are

\[ c(t = 0) = c_I \quad \text{in} \ \Omega, \quad \text{and} \quad c = 0 \quad \text{on} \ J \times \Gamma. \]

Equation (1.1) models the transport of one component involving non-equilibrium reactions. This situation is considered for the ease of presentation, but the present results can be extended to the case of a multi-component reactive transport, as long as the reactive term \( r(\cdot) \) remains Lipschitz continuous. A general model for the reactive transport of \( M \) mobile and \( N \) immobile species is presented in [34]. To extend the applicability of the present analysis, in Section 4.1 we include also equilibrium sorption effects (see [8, 9, 18]), which complicates the analysis. We mention in particular the case of a Freundlich type isotherm, when the equation becomes degenerate. The case with sorption is considered separately only for an easier understanding of the ideas.

The water flux \( Q \) appearing in (1.1), as well as the water content \( \Theta \), are obtained by solving the Richards equations modeling sub-surface water flow, including unsaturated regions near the surface:

\[ \partial_t \Theta(\psi) - \nabla \cdot (K(\Theta(\psi))\nabla(\psi + z)) = 0 \quad \text{in} \ J \times \Omega. \]
The Richards equation results by using the Darcy law

\[ Q = -K(\Theta(\psi))\nabla(\psi + z), \]

in the mass balance equation for water, which is assumed incompressible

\[ \partial_t \Theta(\psi) + \nabla \cdot Q = 0. \]

Equations (1.4) - (1.5) are completed by initial and boundary conditions

\[ \Psi(t = 0) = \Psi_I \quad \text{in} \quad \Omega, \quad \text{and} \quad \Psi = 0 \quad \text{on} \quad J \times \Gamma. \]

Notice the occurrence of the two functions, the soil-water retention curve \( \Theta(\psi) \), as well as the hydraulic conductivity \( K(\Theta) \), for which several forms are proposed based on laboratory experiments (see e.g. [13]). These functions are strictly increasing and bounded in unsaturated regions, where \( \Theta \) is less than a maximal saturation \( \Theta_S \). Furthermore, the functions are constant in saturated regions, where \( \Theta = \Theta_S \). For the functions that are commonly used, (1.3) is a degenerate elliptic-parabolic equation, and solving it numerically is a challenge in itself. However, in this paper we focus on the solute transport and reaction. The underlying assumption is that these processes do not influence the water flow, therefore \( \Theta \) and \( Q \) are assumed known. In practical computations we first solve (1.3) numerically, and then use the results in (1.1). To be specific, as for the solute transport equation (1.1), the numerical solution of the Richards equation is obtained by an EI-MFE scheme based on \( RT_0 \) elements (see [30] for details). The scheme is briefly presented in Section 3 together with a review of the main results concerning the convergence for saturated/unsaturated flow.

In this paper we analyze the EI-MFE scheme for the transport equation (1.1), by employing techniques that are similar to those used in [5, 31, 37]. The main result shows the convergence of the fully discrete numerical scheme for (1.1). It is obtained in a general framework, by taking into account the low regularity of the solution of the Richards equation. The order of convergence clearly depends on the accuracy of the scheme for water flow. We also show the existence and uniqueness for the solution of the variational problems on which the numerical scheme is constructed. For algorithmic and implementation details we refer to [30, 34].

Several papers are considering numerical schemes for transport equations. We mention [8, 9, 11] for a conformal FEM discretization, [22, 26] for finite volume schemes, and [35, 36] for discontinuous Galerkin methods. Furthermore, a characteristic - mixed method is studied in [4], upwind MFEM are considered in [16, 17], whereas combined finite volume-mixed hybrid finite elements are employed in [20]. Typically either a constant \( \Theta \) is assumed, which corresponds to a saturated flow (see [8, 9, 16, 20]), or a \( \Theta \) that does not depend on time (see [4, 35, 36]). Another possible simplification is to incorporate the term \( c\partial_t \Theta \) in the reactive term \( r(\cdot) \) and to assume that the resulting rate remains Lipschitz continuous [11]. In the general case of a saturated-unsaturated flow this assumption is not satisfied since the factor \( \partial_t \Theta(\psi) \) needs not to be essentially bounded.

An interesting situation is considered in [17], where the term \( c\partial_t \Theta \) is replaced by \( Ac \), with \( A \) a positive constant. Error estimates are obtained without using the Gronwall lemma. A similar situation appears in [20] where the divergence of the water flux is assumed strictly negative: \( \nabla \cdot Q = r \leq 0 \). However, in a general saturated-unsaturated porous media flow such an assumption is not necessary true.
In this paper we consider a MFE discretization of (1.1), requiring a divergence form. Therefore we cannot reduce the terms $c \partial_t \Theta$ and $c \nabla \cdot Q$ by using (1.5), as usually done when dealing with conforming FE. This makes the analysis of a MFE discretization of (1.1) very challenging. Further, we do not assume that the divergence of the water flux is constant or has a constant sign.

The paper is structured as follows. The following section introduces the notations and the working assumptions. In Section 3, the numerical scheme to solve the Richards equation is presented and the main results regarding its convergence are reviewed. Next, Section 4 contains the numerical analysis for the solute transport model, including existence and uniqueness of a solution of the continuous and discrete variational problems. In a brief discussion ending this section we also include equilibrium sorption effects, leading to a possible degeneracy in the solute equation. The analysis is completed by several experiments in Section 5, where the theoretical estimates are confronted with the numerical results. In particular we have investigated in how far the accuracy for the solute is influenced by the errors in approximating the fluid flow.

2. Notations and Assumptions. In what follows we make use of common notations in the functional analysis. By $\langle \cdot, \cdot \rangle$ we mean the inner product on $L^2(\Omega)$, or the duality pairing between $H^1_0(\Omega)$ and $H^{-1}(\Omega)$. Further, $\| \cdot \|$ and $\| \cdot \|_1$ stand for the norms in $L^2(\Omega)$ and $H^1(\Omega)$, respectively. The functions in $H(div; \Omega)$ are vector valued, having a $L^2$ divergence. By $C$ we mean a positive constant, not depending on the unknowns or the discretization parameters.

Furthermore, we let $T_h$ be a regular decomposition of $\Omega \subset \mathbb{R}^d$ into closed $d$-simplices; $h$ stands for the mesh-size (see [15]). Here we assume $\Omega = \bigcup_{T \in T_h} T$, hence $\Omega$ is polygonal. In this way the errors caused by an approximation of a nonpolygonal domain are avoided; we mention [25] for a detailed analysis. We will use the discrete subspaces $W_h \subset L^2(\Omega)$ and $V_h \subset H(div; \Omega)$ defined as

$$W_h := \{ p \in L^2(\Omega) \mid p \text{ is constant on each element } T \in T_h \},$$

$$V_h := \{ q \in H(div; \Omega) \mid q|_T = a + bx \text{ for all } T \in T_h \}.$$  

In other words, $W_h$ denotes the space of piecewise constant functions, while $V_h$ is the $RT_0$ space (see [14]). Notice that $\nabla \cdot q \in W_h$ for any $q \in V_h$.

We will use the following $L^2$ projectors (see [14] and [29], p. 237):

$$P_h : L^2(\Omega) \rightarrow W_h, \quad \langle P_h w - w, w_h \rangle = 0,$$

respectively

$$\Pi_h : H(div; \Omega) \rightarrow V_h, \quad \langle \nabla \cdot (\Pi_h v - v), w_h \rangle = 0,$$

for all $w \in L^2(\Omega), v \in H(div; \Omega)$ and $w_h \in W_h$. For these operators we have

$$\| w - P_h w \| \leq Ch \| w \|_1, \quad \text{respectively} \quad \| v - \Pi_h v \| \leq Ch \| v \|_1$$

for any $w \in H^1(\Omega)$ and $v \in (H^1(\Omega))^d$.

For the discretization in time we let $N \in \mathbb{N}$ be strictly positive, and define the time step $\tau = T/N$, as well as $t_n = n\tau$ ($n \in \{1, 2, \ldots, N\}$). Given a function $f$ defined on the interval $J$, we define:

$$\overline{f}^n = \frac{1}{\tau} \int_{t_{n-1}}^{t_n} f(t) \, dt, \quad \text{and} \quad f^n = f(t_n),$$

3
whenever $n \in \{1, \ldots, N\}$. For $n = 0$ we take $\mathcal{F}^0 = f^0 = f(0)$.

Throughout this paper we make use of the following assumptions:

(A1) The rate function $r : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous with the constant $L_r$; furthermore, $r(c) = 0$ for all $c \leq 0$.

(A2) The diffusion coefficient $D$ does not depend on $\psi$ or $Q$. For simplicity, let $D = 1$.

(A3) $1 \geq \Theta_S \geq \Theta(x) \geq \Theta_R > 0, \forall x \in \mathbb{R}$.

(A4) The initial $c_I$ is essentially bounded and positive; furthermore, $\Psi_I \in L^2(\Omega)$.

(A5) $Q \in L^\infty(J \times \Omega) \cap L^2(J; H^1(\Omega))$.

Remark 2.1. For the ease of presentation we take $D = 1$. The extension to a positive definite tensor is immediate, requiring a minor change in Lemma 4.35 below.

Remark 2.2. In (A3) we assume that $\Theta(\cdot)$ is uniformly bounded by a strictly positive constant. Since $\Theta$ stands for the water content, the boundedness of $\Theta(\cdot)$ is a reasonable assumption. However, by taking the lower bound strictly positive we disregard the case of a completely dry (fully unsaturated) medium. For commonly used porous media models (see e.g. [13]) such an assumption holds, for example, in the case of a homogeneous medium if the initial and boundary saturation (where prescribed) exceed the lower limit. Furthermore, (A5) also implies that $\partial_t \Theta(\psi) \in L^2(J \times \Omega)$. Since $\Theta(\psi)$ is essentially bounded, we immediately obtain $\Theta(\psi) \in C([0, T]; L^2(\Omega))$.

Remark 2.3. (A5) is also assumed in recent papers referring to the discretization of porous media flow models (see also [11, 20]). Previous results are under stronger assumptions: a constant divergence of the flux ([35, 36, 17]), a constant sign for the water flux ([16]), or a constant water flux [8, 9, 18].

3. Error estimates for Richards' equation. In this section we review some convergence results for the EL-MFE (RT0) discretization of the Richards equation. The particular choice of the finite element space is dictated by the lacking regularity of the solution. We consider three cases: the saturated flow regime (where $\Theta = \Theta_S$), the strictly unsaturated flow regime (where $\Theta < \Theta_S$), and the saturated-unsaturated flow regime where both cases mentioned above are allowed. Denoting by $e_x$ the constant gravitational vector and given the initial pressure $\psi_I \in L^2(\Omega)$, we can define the mixed, time integrated variational form of (1.5)–(1.4).

**Problem 3.1.** Find $(\psi, Q) \in L^2(J; L^2(\Omega)) \times L^2(J; (L^2(\Omega))^d)$ such that $\int_0^t Q(s) \, ds \in L^2(J; H(\text{div}; \Omega))$, and

\begin{align}
\langle c, \Theta(\psi(t)) - \Theta(\psi_I), w \rangle + \langle \nabla \cdot \int_0^t Q(s) \, ds, w \rangle &= 0, \\
\langle \int_0^t K^{-1}(\Theta(\psi(s)))Q(s) \, ds, v \rangle + \langle \int_0^t \psi(s) \, ds, \nabla \cdot v \rangle + \langle \int_0^t e_x \, dt, v \rangle &= 0,
\end{align}

for all $t \in J$, $w \in L^2(\Omega)$ and $v \in H(\text{div}; \Omega)$.

We mention [2, 27] for the existence and uniqueness of a solution for Problem 3.1. In particular, in [2] it is shown that $\Theta(\psi) \in L^\infty(J; L^1(\Omega))$. By (A3) we also have $\Theta(\psi) \in L^\infty(J \times \Omega)$.

With $n \in \{1, 2, \ldots, N\}$ we can define the fully discrete problem at time level $n$:

**Problem 3.2.** Let $h > 0$ be given. Find $(\psi^n_h, Q^n_h) \in W_h \times V_h$ such that

\begin{align}
\langle c, \Theta(\psi^n_h) - \Theta(\psi^{n-1}_h), w_h \rangle + \langle \nabla \cdot Q^n_h, w_h \rangle &= 0, \\
\langle K^{-1}(\Theta(\psi^n_h))Q^n_h, v_h \rangle - \langle \psi^n_h, \nabla \cdot v_h \rangle + \langle e_x, v_h \rangle &= 0,
\end{align}

for all $w_h \in W_h$ and $v_h \in V_h$.
for all \( w_h \in W_h \) and \( v_h \in V_h \).

The discrete initial pressure is taken such that \( \Theta(\psi^0_h) = P_h \Theta(\psi^I) \). Since \( \psi^0_h \in W_h \) is piecewise constant, the same holds for \( \Theta(\psi^0_h) \). Moreover, we have \( \Theta(\psi^0_h) \geq \Theta_R > 0 \).

Assuming that the flow is fully saturated, when \( \Theta = \Theta_S \) for all times and almost everywhere in \( \Omega \), the Richards equation becomes elliptic. In this case one obtains \([14]\)

**Theorem 3.1.** Assume that the flow is fully saturated. Then we have

\[
\|\psi - \psi_h\|^2 + \|Q - Q_h\|^2 \leq Ch^2
\]

If the flow is strictly partially saturated (i.e. \( \Theta_R \leq \Theta < \Theta_S \)), then the Richards equation is nonlinear but regular (non-degenerate). Then we have \([3]\)

**Theorem 3.2.** Let \( (\psi, Q) \) be the solution of Problem 3.1 and \( (\psi^n_h, Q^n_h) \in W_h \times V_h \) solving the discrete problems 3.2. Assuming \((A3)\), there holds

\[
\sum_{n=1}^{N} \tau \|\psi(t_n) - \psi^n_h\|^2 + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\Theta(\psi(t)) - \Theta(\psi^n_h)\|^2 dt \\
+ \sum_{n=1}^{N} \tau \|Q^n - Q^n_h\|^2 + \sum_{n=1}^{N} \tau \|Q(t_n) - Q^n_h\|^2 \leq C(\tau^2 + h^2).
\]

The convergence order of the scheme numerical scheme \((3.3) - (3.4)\) in the partially saturated case has been investigated numerically in several papers (see \([11, 12]\), as well as Section 5). To our knowledge, no rigorous error estimates have been proven up to now in this specific case. This is due to the lacking regularity of the solution, as well as the nonlinearities involved in the coefficient functions. An alternative approach is to combine the two nonlinearities in \((1.3)\) into a single one and use the Kirchhoff transformation \([2, 5, 31, 37]\). For this approach, assuming that the saturation is a Lipschitz continuous function of the pressure, as well as \( \partial_t \Theta \in L^2(J; L^2(\Omega)) \), an \( \tau^2 + h^2 \) convergence order is obtained in \([5, 31, 37]\). The regularity assumption on \( \partial_t \Theta \) is given up in \([31]\), resulting in a reduction of the convergence order to \( \tau + h^2 \). A more general situation is considered in \([33]\), where only Hölder continuity is assumed for \( \Theta(\cdot) \). In this case the convergence result applies to all flow regimes. We also refer to \([41]\) for the analysis of an expanded mixed finite element scheme for equation \((1.3)\).

In the same spirit we mention another well recognized mass conservative method, the multipoint flux approximation method (MPFA) (see e.g. \([6]\)). As proven in \([21]\), this method can be successfully employed for solving the Richards equation.

In this paper we aim at proving the convergence of the numerical scheme for the solute transport. The error estimates proven in the main result here are given in terms of discretization parameters and depend also on the accuracy of the scheme for the water flow. Therefore it is not so relevant which method was used to solve the Richards equation, as long as it produces convergent approximations for the water flux and the saturation.

Having in mind \((A5)\), we assume something similar for the discrete fluxes:

\[
(A5') \quad Q^n_h \in L^\infty(\Omega) \quad \text{for all } n \in \{1, \ldots, N\}.
\]

**4. Error estimates for the EI-MFE scheme for the solute transport equation.** In this section we prove the error estimates for the EI-MFE discretization of the solute transport equation. We first set a continuous mixed variational formulations of \((1.1)\) and analyse the existence and uniqueness of the solution. We continue by giving some stability estimates for the solution of the continuous problem and present the fully discrete scheme. After showing the existence and uniqueness of
a solution also in the discrete case we prove the convergence of the scheme. Following
the discussion in the introduction, we assume that the pressure \( p \) and the water flux
\( \mathbf{Q} \), as well as their approximations \( \psi_n, \mathbf{Q}_n \), are assumed and satisfying the estimates in
Section 3.

**Problem 4.1.** (The continuous problem) Find \( c \in L^2(J; L^2(\Omega)) \) and \( \mathbf{q} \in L^2(J; H(\text{div}; \Omega)) \) such that for almost every \( t \in J \) there holds

\[
\begin{align*}
(\Theta(\psi(t))c(t)) - \Theta(\psi_1)c(t) + \mathbf{q} \cdot \nabla c = \Theta(\psi(t))r(c) & \quad \text{in } J \times \Omega, \\
\langle \mathbf{q}(t), \mathbf{v} \rangle - \langle c(t), \nabla \cdot \mathbf{v} \rangle - \langle c(t)\mathbf{Q}(t), \mathbf{v} \rangle &= 0,
\end{align*}
\]

for all \( \mathbf{v} \in H(\text{div}; \Omega) \) and \( \mathbf{v} \in H(\text{div}; \Omega) \).

**Theorem 4.1.** Assuming (A1)-(A5), the Problem 4.1 has an unique solution.

**Proof.** To prove the existence of a solution we use the Richards equation (1.5)
and obtain

\[
\Theta(\psi)\partial_t c - \Delta c + \nabla c = \Theta(\psi(t))r(c) \quad \text{in } J \times \Omega.
\]

With boundary and initial conditions as in (1.2) and since \( D = 1 \), (1.1) and (4.3) are
formally equivalent. Therefore we start with a solution to (4.3) and then construct a
solution to the mixed formulation of (1.1). The existence and uniqueness of a solution
for the transformed equation (4.3) is provided first in \( H^1(J; H^{-1}(\Omega)) \cap L^2(J; H^1_0(\Omega)) \)
(see [24], Chapter VI.4; alternatively one can prove the result as a limit of solutions
to linearized problems and using the results in Chapter III). Moreover, by (A1) and
(A4) the solution \( c \) remains essentially bounded and positive. Furthermore, using
(A5), the regularity of \( \Omega \) and since the boundary conditions are homogeneous we can
improve the regularity of the solution to \( c \in H^1(J; L^2(\Omega)) \cap L^2(J; H^2(\Omega)) \).

Let now \( \zeta \in H^1_0(\Omega) \) fixed arbitrary. Since \( \partial_\ell c \in L^2(J \times \Omega) \), using the embedding
\( H^1 \hookrightarrow L^4 \), for any \( \varphi \in L^2(J; H^1_0(\Omega)) \) we have

\[
\left| \int_0^T \langle \zeta \partial_\ell c, \varphi \rangle dt \right| \leq \int_0^T \| \partial_\ell c \| \| \zeta \varphi \| dt \leq C \| \varphi \|_{L^2(J; H^1(\Omega))},
\]

showing that \( \zeta \partial_\ell c \in L^2(J; H^{-1}(\Omega)) \). Furthermore, since \( c \) is essentially bounded and
in \( L^2(J; H^2(\Omega)) \), we also have \( \zeta c \in L^2(J; H^1(\Omega)) \). By the regularity of \( \Theta(\Psi) \), for
almost all \( 0 \leq s < t \leq T \) we have

\[
\langle \Theta(\Psi(t)), \zeta c(t) \rangle - \langle \Theta(\Psi(s)), \zeta c(s) \rangle = \int_s^t \langle \partial_\tau \Theta(\Psi(\tau)), \zeta c(\tau) \rangle + \langle \partial_\tau (\zeta c(\tau)), \Theta(\Psi(\tau)) \rangle d\tau.
\]

Using the Richards equation (1.5), this gives

\[
\langle \Theta(\Psi(t)), \zeta c(t) \rangle - \langle \Theta(\Psi(s)), \zeta c(s) \rangle + \int_0^t \langle \nabla \cdot \mathbf{Q}, \zeta c(\tau) \rangle - \langle \partial_\tau (\zeta c(\tau)), \Theta(\Psi(\tau)) \rangle d\tau = 0.
\]

Since \( c \) is a weak solution of (4.3), we immediately obtain

\[
\langle \Theta(\Psi(t))c(t), \zeta \rangle - \langle c(t)\Theta(\Psi(t)), \zeta \rangle + \int_0^t \langle \nabla \cdot (c(t)\mathbf{Q}(\tau)), \zeta \rangle d\tau \\
+ \int_0^t \langle \nabla c(\tau), \zeta \rangle d\tau = \int_0^t \langle \Theta(\Psi(\tau))r(c(\tau)), \zeta \rangle d\tau.
\]

Equation (4.4) holds for any \( \zeta \in H^1_0(\Omega) \). However, since \( c \in L^2(J; H^2(\Omega)) \), we can
transform the inner product in the last term on the left into \( -\langle \Delta c(\tau), \zeta \rangle \). Then by
density arguments, (4.4) holds for any test function \( \zeta \in L^2(\Omega) \). Identifying now \( q = \boldsymbol{r} \cdot \nabla c \in L^2(J; H(\text{div}; \Omega)) \) we have found a solution to Problem 4.1.

For the uniqueness we assume that Problem 4.1 has two solutions \((c_1, q_1), (c_2, q_2)\) in \( L^2(J \times \Omega) \times L^2(J; H(\text{div}; \Omega)) \). With \( c = c_1 - c_2 \) and \( q = q_1 - q_2 \) it follows that for almost all \( t > 0 \) we have

\[
\langle \Theta(\psi(t))c(t), w \rangle + \langle \nabla \cdot \int_0^t q(s) \, ds, w \rangle = \langle \int_0^t \Theta(\psi)(r(c_1) - r(c_2)) \, ds, w \rangle,
\]

\[
\langle q(t), v \rangle - \langle c(t), \nabla \cdot v \rangle - \langle c(t) Q(t), v \rangle = 0,
\]

for all \( w \in L^2(\Omega), v \in H(\text{div}; \Omega) \) and for almost all \( t > 0 \). Taking into above \( w = c(t) \) and \( v = \int_0^t q \, ds \) and adding the resulting gives

\[
\langle \Theta(\psi(t))c(t), c(t) \rangle + \langle q(t), \int_0^t q \, ds \rangle = \langle c(t) Q(t), \int_0^t q \, ds \rangle + \langle \int_0^t \Theta(\psi)(r(c_1) - r(c_2)) \, ds, c(t) \rangle.
\]

(4.5)

Integrating (4.5) in time from 0 to a \( t' \leq T \), since \( \Theta \geq \Theta_R \) we obtain

\[
\Theta_R \int_0^{t'} \|c\|^2 \, dt + \frac{1}{2} \left\| \int_0^{t'} q \, dt \right\|^2 \\
\leq \int_0^{t'} \langle c(t) Q(t), \int_0^t q \, ds \rangle \, dt + \int_0^{t'} \langle \int_0^t \Theta(\psi)(r(c_1) - r(c_2)) \, ds, c(t) \rangle \, dt.
\]

(4.6)

Since \( Q \in L^\infty(J \times \Omega) \) using the Cauchy-Schwarz inequality leads to

\[
\Theta_R \int_0^{t'} \|c\|^2 \, dt + \frac{1}{2} \left\| \int_0^{t'} q \, dt \right\|^2 \\
\leq C^2 \Theta_R \int_0^{t'} \left\| \int_0^t q \, ds \right\|^2 \, dt + \frac{1}{\Theta_R} \int_0^{t'} \left\| \int_0^t \Theta(\psi)(r(c_1) - r(c_2)) \, ds \right\|^2 \, dt.
\]

(4.7)

Recalling that \( \Theta \leq \Theta_S \), using the Lipschitz continuity of \( r(\cdot) \) and the Cauchy-Schwarz inequality, the inequality (4.7) becomes

\[
\Theta_R \int_0^{t'} \|c\|^2 \, dt + \frac{1}{2} \left\| \int_0^{t'} q \, dt \right\|^2 \leq C^2 \Theta_R \int_0^{t'} \left\| \int_0^t q \, ds \right\|^2 \, dt + \frac{L^2_{\Theta} \Theta^2}{\Theta_R} \int_0^{t'} \int_0^t \|c\|^2 \, dt,
\]

(4.8)

for all \( t' \in [0, T] \). Since both solutions satisfy the same initial data, employing the Gronwall lemma gives \( \|c(t)\| = 0 \), implying the uniqueness of \( c \). Furthermore, since \( \int_0^{t'} q \, dt = 0 \) for all \( t' \in [0, T] \), (4.5) also gives \( q = q_1 - q_2 = 0 \), yielding uniqueness for Problem 4.1. \( \square \)

**Remark 4.1.** As follows from the above proof, the solution has more regularity than required in the statement of Problem 4.1. To be specific, we have \( c \in H^1(J; L^2(\Omega)) \cap L^2(J; H^2(\Omega)) \cap L^\infty(J \times \Omega) \), whereas \( q \in L^\infty(J \times \Omega) \). Furthermore, since \( Q \in L^2(J; H^1(\Omega))' \) we immediately obtain \( q \in L^2(J; (H^1(\Omega))') \). Also notice that the regularity of \( c \) and \( \Theta(\psi) \) allows writing (4.1) for all \( t \).
Recalling the notations in (2.5), the following stability estimates will be used later.

**Proposition 4.2.** For the solution of Problem 4.1 we have

\begin{align}
\sum_{n=1}^{N} \tau \|q^n\|_1^2 & \leq C \tag{4.9} \\
\sum_{n=1}^{N} \tau \|\bar{q}^n - q^n\|^2 & \leq C \tau^2 \tag{4.10} \\
\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|c(t) - c^n\|^2 dt & \leq C \tau^2 \tag{4.11} \\
\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|c(t) - \bar{q}^n\|^2 dt & \leq C \tau^2 \tag{4.12} \\
\sum_{n=1}^{N} \tau \|\bar{q}^n\|_1^2 & \leq C \tag{4.13} \\
\end{align}

**Proof.** Since \(q \in L^2(J; (H^1(\Omega))^d)\), the estimate (4.9) follows straightforwardly:

\[
\sum_{n=1}^{N} \tau \|q^n\|_1^2 = \sum_{n=1}^{N} \tau \left\| \frac{1}{\tau} \int_{t_{n-1}}^{t_n} q dt \right\|_1^2 \leq \sum_{n=1}^{N} \tau \left( \int_{t_{n-1}}^{t_n} |q| dt \right)^2 \leq C.
\]

For (4.10) we use \(\partial_t c \in L^2(J; L^2(\Omega))\):

\[
\sum_{n=1}^{N} \tau \|\bar{q}^n - q^n\|^2 = \sum_{n=1}^{N} \frac{1}{\tau^2} \int_{\Omega} \left( \int_{t_{n-1}}^{t_n} c(t_n) - c(t_n) dt \right)^2 dx = \sum_{n=1}^{N} \frac{1}{\tau} \int_{\Omega} \left( \int_{t_{n-1}}^{t_n} \int_{t}^{t_n} \partial_t c(s) ds dt \right)^2 dx \\
\leq \sum_{n=1}^{N} \int_{\Omega} \int_{t_{n-1}}^{t_n} \int_{t_{n-1}}^{t_n} (\partial_t c)^2 ds dt dx \leq \sum_{n=1}^{N} \tau^2 \|\partial_t c\|_{L^2(t_{n-1}, t_n; L^2(\Omega))}^2 \leq C \tau^2.
\]

The proof for (4.11) and (4.12) follows similarly, whereas (4.13) is a consequence of \(c \in L^2(J; H^1(\Omega))\).

We now proceed with the EI-MFE scheme for Problem 4.1.

**Problem 4.2.** (The discrete problem) Let \(n \in \{1, \ldots, N\}\), and \(\Theta(\psi_h^n), \Theta(\psi_h^{n-1})\), \(Q_h^n\), as well as \(c_h^{n-1}\) be given. Find \((c_h^n, q_h^n) \in W_h \times V_h\) such that

\begin{align}
\Theta(\psi_h^n) c_h^n - \Theta(\psi_h^{n-1}) c_h^{n-1}, w_h & + \tau \langle \nabla \cdot q_h^n, w_h \rangle = \tau (\Theta(\psi_h^n) r(c_h^n), w_h), \tag{4.14} \\
\langle c_h^n, v_h \rangle - \langle c_h^n, \nabla \cdot v_h \rangle & - \langle c_h^n Q_h^n, v_h \rangle = 0, \tag{4.15}
\end{align}

for all \(w_h \in W_h\) and \(v_h \in V_h\).

Initially we take \(c_h^0 = \frac{P_h(\Theta(\psi_1))}{P_h(\Theta(\psi_1))} \). The particular form of the initial data is allowed by the lower bound on \(\Theta\) and will be used when proving Theorem 4.5 below.

**Theorem 4.3.** Assuming (A1)-(A3) and that the time step \(\tau\) is small enough, the Problem 4.2 has an unique solution.

**Proof.** We first show uniqueness. Let \((c_h^{n,1}, q_h^{n,1})\) and \((c_h^{n,2}, q_h^{n,2})\) be two solutions of Problem 4.2. With \(c_h^n := c_h^{n,1} - c_h^{n,2} \in W_h\) and \(q_h^n := q_h^{n,1} - q_h^{n,2} \in V_h\) we have

\[
\langle \Theta(\psi_h^n) c_h^n, w_h \rangle + \tau \langle \nabla \cdot q_h^n, w_h \rangle = \tau (\Theta(\psi_h^n) (r(c_h^{n,1}) - r(c_h^{n,2})), w_h),
\]

8
for all \( w_h \in W_h \) and \( v_h \in V_h \). Taking \( w_h = c^n_h \) and \( v_h = \tau q^n_h \), adding the above gives

\[
(\Theta(\psi^n_h)c^n_h, c^n_h) + \tau \|q^n_h\|^2
\]

(4.18)

\[
= \tau(c^n_h Q^n_h, q^n_h) + \tau(\Theta(\psi^n_h)(r(c^n_{h,1}) - r(c^n_{h,2})), c^n_h).
\]

Using now (A1), (A3), (A5') and the Cauchy-Schwarz inequality we immediately get

\[
||c^n_h||^2 + \tau ||q^n_h||^2 \leq C||q^n_h||^2 + C\tau^2 ||q^n_h||^2.
\]

(4.19)

If \( \tau \) is sufficiently small, this gives gives \( c^n_h = 0 \) and \( q^n_h = 0 \), so the solution is unique.

To show the existence of a solution for the fully discrete Problem 4.2 we let \( \{w_1, \ldots, w_{n_1}\} \cup \{v_1, \ldots, v_{n_2}\} \) be a basis for \( W_h \times V_h \) and introduce the mapping \( \mathcal{P} : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^{n_1+n_2} \) that will be defined below. To do so we first notice that given \( \hat{\alpha}, \alpha \in \mathbb{R}^{n_1} \) and \( \hat{\beta}, \beta \in \mathbb{R}^{n_2} \), with \( \hat{\xi} = (\hat{\alpha}, \hat{\beta}) \), \( \xi = (\alpha, \beta) \) one can consider

\[
((\hat{\xi}, \xi)) := (\hat{\alpha}, \alpha)_{n_1} + \tau(\hat{\beta}, \beta)_{n_2}, \text{ and } ||\xi|| := ((\xi, \xi))^{1/2}
\]

for defining an inner product, respectively a norm on \( \mathbb{R}^{n_1+n_2} \). Here \((\cdot, \cdot)_p\) stands for the euclidian inner product in \( \mathbb{R}^p \).

Further, any \( \xi = (\alpha, \beta) \in \mathbb{R}^{n_1+n_2} \) determines uniquely a pair \( (\hat{w}, \hat{v}) \in W_h \times V_h \) by \( \hat{w} = \sum_{k=1}^{n_1} \alpha_k w_k \), respectively \( \hat{v} = \sum_{k=1}^{n_2} \beta_k v_k \). Then we let \( \hat{\xi} = (\hat{\alpha}, \hat{\beta}) \in \mathbb{R}^{n_1+n_2} \) be given by

\[
\hat{\alpha}_k = (\Theta(\psi^n_h)\hat{w} - \Theta(\psi^{n-1}_h)\hat{c}^{n-1}_h, w_k) + \tau(\nabla \cdot \hat{v}, w_k) - \tau(\Theta(\psi^n_h)\tau(\hat{w}), w_k)
\]

for all \( k = 1, \ldots, n_1 \), respectively

\[
\hat{\beta}_k = (\hat{v}, v_k) - (\hat{w}, \nabla \cdot v_k) - (\hat{w} Q^n_h, v_k)
\]

for all \( k = 1, \ldots, n_2 \). Having determined the above we define \( \mathcal{P}(\xi) = \hat{\xi} \). Notice that finding a \( \xi \) such that \( \mathcal{P}(\xi) = 0 \) immediately gives a solution to Problem 4.2.

Clearly, \( \mathcal{P} \) is continuous. Moreover, for any \( \xi = (\alpha, \beta) \in \mathbb{R}^{n_1+n_2} \), with \( (\hat{w}, \hat{v}) \in W_h \times V_h \) introduced above we have

\[
((\mathcal{P}(\xi), \xi)) = (\Theta(\psi^n_h)\hat{w} - \Theta(\psi^{n-1}_h)\hat{c}^{n-1}_h, \hat{w}) + \tau\|\hat{v}\|^2 - \tau(\hat{w} Q^n_h, v_k) - \tau(\Theta(\psi^n_h)\tau(\hat{w}), w_k).
\]

Recalling (A1) - (A5), using the Cauchy inequality, as well as the inequality of means we obtain

\[
((\mathcal{P}(\xi), \xi)) \geq \Theta_R \|\hat{w}\|^2 + \frac{1}{2}\|\hat{v}\|^2 - \frac{\tau C_Q}{2}\|\hat{w}\|^2
\]

\[
- (\tau \Theta_S L_\tau + \frac{\Theta_S}{2}) \|\hat{w}\|^2 - \frac{1}{2\Theta_S} \|\Theta(\psi^{n-1}_h)\hat{c}^{n-1}_h\|^2,
\]

where \( C_Q = \|Q\|_{L^\infty} \) and \( L_\tau \) is the Lipschitz constant of \( r \). For \( \tau \) small enough we have \( \Theta_R - \tau (C_Q + 2L_\tau \Theta_S) \geq m > 0 \), yielding

\[
((\mathcal{P}(\xi), \xi)) \geq \frac{1}{2} \min\{m, 1\} (\|\xi\|^2 - K),
\]

9
where $K = \|\Theta(\psi_{h}^{n-1})c_{h}^{n-1}\|^{2}/(2\Theta R)$. This gives \((P(\xi), \xi) \geq \min\{m, 1\}K > 0\) for all $\xi \in \mathbb{R}^{n+1}$ satisfying $\|\xi\|^2 = 2K$, yielding the existence of a solution (see Lemma 1.4, p. 140 in [39]).

In the fully saturated case, since $\Theta(\psi) = \Theta_S$ is constant, standard techniques for parabolic equations (see, e.g., [23]) can be used to prove the convergence of the scheme. In this case we obtain:

**Theorem 4.4.** (Fully saturated flow) Let $(c, q)$ solve Problem 4.1, and $(c_n^h, q_n^h)$ solve the problems 4.2 for all $n \in \{1, \ldots, N\}$. Assuming (A1)-(A5), there holds

\[
\max_{n=1,\ldots,N} \|c(t_n) - c_n^h\|^2 + \sum_{n=1}^{N} \|q(t_n) - q_n^h\|^2 \leq C(\tau^2 + h^2).
\]

In the saturated-unsaturated flow regime we start with the following

**Theorem 4.5.** (Saturated-unsaturated flow) Let $(c, q)$ solve Problem 4.1 and $(c_n^h, q_n^h)$ solve the problems 4.2 for all $n \in \{1, \ldots, N\}$. Assuming (A1)-(A5) we have

\[
\sum_{n=1}^{N} \tau \|c_{n}^h\|^2 + \sum_{n=1}^{N} \|\Pi_h q_n^h - q_n^h\|^2 \leq C \left\{ \tau^2 + h^2 + \sum_{n=1}^{N} \tau \|Q_n^h - Q_n^h\|^2 \right\}
\]

\[
+ \sum_{n=1}^{N} \tau \|\Theta(\psi(t_n)) - \Theta(\psi_n^h)\|^2 + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\Theta(\psi) - \Theta(\psi_n^h)\|^2 dt.
\]

where the constant $C$ does not depend on the discretization parameters.

**Proof.** Taking $t = t_n$ in (4.1) and recalling the notation in (2.5) we have

\[
\langle \Theta(\psi^n)c^n, w \rangle + \sum_{k=1}^{n} \tau \nabla \cdot (\Pi_h q^n), w \rangle = \langle \Theta(\psi^n)c_I, w \rangle + \sum_{k=1}^{n} \tau \Theta(\psi^n)r(c_{k}^h), w \rangle,
\]

for all $w \in L^2(\Omega)$.

Before dealing with the flux equation (4.2) we mention that the analysis in [5, 37] carried out for the Richards equation leaves the flux equation unchanged, and considers the time integrated variant of the balance equation (4.1). Here we proceed as in [33] and integrate also (4.2) in time to obtain

\[
\langle \Pi_h q^n, v \rangle - \langle r^n, \nabla \cdot v \rangle - \langle c \Pi_h q^n, v \rangle = 0,
\]

for all $v \in H(\text{div}; \Omega)$. Summing up (4.14) written for the time steps $t_k, k = 1, \ldots, n$, subtracting the result from (4.22), as well as (4.15) from (4.23), and employing the projectors defined in (2.2), (2.3) gives

\[
\langle \Theta(\psi^n)c^n - \Theta(\psi_n^h)c_n^h, w_h \rangle + \sum_{k=1}^{n} \tau \langle \nabla \cdot \Pi_h q^n, q_k^h, w_h \rangle = \sum_{k=1}^{n} \tau \langle \Theta(\psi^n)r(c_{k}^h) - \Theta(\psi_n^h)r(c_{k}^h), w_h \rangle,
\]

for all $w_h \in W_h$ and

\[
\langle q^n - q_n^h, v_h \rangle - \langle \Pi_h q^n - q_n^h, \nabla \cdot v_h \rangle - \langle c \Pi_h q^n - Q_n^h c_n^h, v_h \rangle = 0,
\]

for all $v_h \in V_h$. The specific choice of the initial data gives no error at $t = 0$ in (4.24).
Defining

\[
(4.26) \quad e_n^\theta = \sum_{k=1}^{n} (P_k q_k^\theta - q_k^\theta), \quad \forall n \in \{1, \ldots, N\},
\]

we take \( w_h = P_h \bar{\tau}^n - c_h^n \in \text{W}_h \) in (4.24) and \( v_h = \tau e_n^\theta \in \text{V}_h \) in (4.25), add the resulting and obtain

\[
\langle \Theta(\psi^n) c^n - \Theta(\psi_h^n) c_h^n, P_h \bar{\tau}^n - c_h^n \rangle + \tau \langle \bar{\tau}^n - q_h^n, e_n^\theta \rangle - \tau \langle \bar{c}Q^n_h - Q_h^n \psi_h^n, e_n^\theta \rangle
\]
\[
= \sum_{k=1}^{n} \tau (\Theta(\psi) r(c))^k - \Theta(\psi_h^k) r(c_h^k), P_h \bar{\tau}^n - c_h^n).
\]

Summing the above for \( n = 1 \) to \( N \) yields

\[
(4.27) \quad \sum_{n=1}^{N} (\Theta(\psi^n) c^n - \Theta(\psi_h^n) c_h^n, P_h \bar{\tau}^n - c_h^n) + \sum_{n=1}^{N} \tau (\bar{\tau}^n - q_h^n, e_n^\theta)
\]
\[
= \sum_{n=1}^{N} \tau (\bar{Q}^n_h c_h^n, e_n^\theta) + \sum_{n=1}^{N} \sum_{k=1}^{n} \tau (\Theta(\psi) r(c))^k - \Theta(\psi_h^k) r(c_h^k), P_h \bar{\tau}^n - c_h^n).
\]

Further we estimate the terms in (4.27), which are denoted by \( T_1, \ldots, T_4 \). \( T_1 \) gives

\[
T_1 = \sum_{n=1}^{N} (\Theta(\psi^n) c^n - \Theta(\psi_h^n) c_h^n, P_h \bar{\tau}^n - c_h^n)
\]
\[
= \sum_{n=1}^{N} (\Theta(\psi_h^n) c^n - c_h^n, P_h \bar{\tau}^n - c_h^n) + \sum_{n=1}^{N} ((\Theta(\psi^n) - \Theta(\psi_h^n)) c^n, P_h \bar{\tau}^n - c_h^n)
\]
\[
= \sum_{n=1}^{N} (\Theta(\psi_h^n) (c^n - \bar{\tau}^n), \bar{\tau}^n - c_h^n) + \sum_{n=1}^{N} (\Theta(\psi_h^n) (c^n - \bar{\tau}^n), P_h \bar{\tau}^n - \bar{\tau}^n)
\]
\[
+ \sum_{n=1}^{N} ((\Theta(\psi^n) - \Theta(\psi_h^n)) c^n, P_h \bar{\tau}^n - \bar{\tau}^n) + \sum_{n=1}^{N} ((\Theta(\psi^n) - \Theta(\psi_h^n)) c^n, \bar{\tau}^n - c_h^n)
\]
\[
(4.28) =: T_{11} + T_{12} + T_{13} + T_{14} + T_{15} + T_{16}.
\]

Using (A3) and the Cauchy-Schwarz inequality, for any \( \delta_{11} > 0 \) we have

\[
(4.29) \quad T_{11} \leq \frac{\Theta_S}{2\delta_{11}} \sum_{n=1}^{N} \|c^n - \bar{\tau}^n\|^2 + \frac{\Theta_S \delta_{11}}{2} \sum_{n=1}^{N} \|\bar{\tau}^n - c_h^n\|^2.
\]

The estimate for \( T_{12} \) is similar:

\[
(4.30) \quad T_{12} \leq \frac{\Theta_S}{2} \sum_{n=1}^{N} \|c^n - \bar{\tau}^n\|^2 + \frac{\Theta_S}{2} \sum_{n=1}^{N} \|P_h \bar{\tau}^n - \bar{\tau}^n\|^2.
\]

The term \( T_{13} \) in (4.28) is positive. There holds

\[
(4.31) \quad T_{13} \geq \Theta_n \sum_{n=1}^{N} \|\bar{\tau}^n - c_h^n\|^2.
\]
Furthermore, for any $\delta_{14} > 0$ we have

$$ T_{14} \leq \frac{\Theta_S \delta_{14}}{2} \sum_{n=1}^{N} \|e_n\|^2 + \frac{\Theta_S}{2 \delta_{14}} \sum_{n=1}^{N} \|P_n e_n - e_n\|^2. $$

Similarly, using the boundedness of $c$ (see Remark 4.1) we have

$$ T_{15} \leq C \sum_{n=1}^{N} \|\Theta(\psi^n) - \Theta(\psi^n_h)\|^2 + C \sum_{n=1}^{N} \|P_n e_n - e_n\|^2, $$

whereas for any $\delta_{16} > 0$, $T_{16}$ gives

$$ T_{16} \leq \frac{C}{2 \delta_{16}} \sum_{n=1}^{N} \|\Theta(\psi^n) - \Theta(\psi^n_h)\|^2 + \frac{C \delta_{16}}{2} \sum_{n=1}^{N} \|e_n - c^n_h\|^2. $$

To estimate $T_2$ we use the following elementary equality:

$$ 2 \sum_{n=1}^{N} (a_n, \sum_{k=1}^{n} a_k) = \sum_{n=1}^{N} \|a_n\|^2 + \sum_{n=1}^{N} \|a_n\|^2, $$

for any set of $d$-dimensional real vectors $a_k \in \mathbb{R}^d$ ($k \in \{1, \ldots, N\}, d \geq 1$). This gives

$$ T_2 = \sum_{n=1}^{N} \tau(q^n - \Pi_h q_n^n, e_q^n) + \sum_{n=1}^{N} \tau(\Pi_h q^n - q_h^n, e_q^n) $$

$$ \leq \sum_{n=1}^{N} \tau(q^n - \Pi_h q_n^n, e_q^n) + \frac{\tau}{2} \|e_q^n\|^2 + \frac{\tau}{2} \sum_{n=1}^{N} \|\Pi_h q^n - q_h^n\|^2. $$

To conclude with $T_2$, we estimate the first term on the right, which is denoted by $T_{21}$:

$$ T_{21} \leq \frac{1}{2} \sum_{n=1}^{N} \|q^n - \Pi_h q_n^n\|^2 + \frac{1}{2} \sum_{n=1}^{N} \tau^2 \|e_q^n\|^2. $$

The convective term $T_3$ is split into four terms, denoted $T_{31}, \ldots, T_{34}$:

$$ T_3 = \sum_{n=1}^{N} \tau(q^n - q_n^n, e_q^n) + \sum_{n=1}^{N} \tau(q^n - q_h^n, e_q^n) $$

$$ + \sum_{n=1}^{N} \tau(q_h^n (c^n - e^n), e_q^n) + \sum_{n=1}^{N} \tau(q_h^n (e^n - c_h^n), e_q^n). $$

Using (A5) and the Cauchy-Schwarz inequality, for $T_{31}$ we obtain

$$ T_{31} \leq \frac{1}{2} \sum_{n=1}^{N} \|\frac{1}{\tau} \int_{t_{n-1}}^{t_n} Q(t)(c(t) - c^n)dt\|^2 + \frac{\tau^2}{2} \sum_{n=1}^{N} \|e_q^n\|^2 $$

$$ \leq \frac{C}{2 \tau} \sum_{n=1}^{N} \|c(t) - c^n\|^2 dt + \frac{\tau^2}{2} \sum_{n=1}^{N} \|e_q^n\|^2. $$
Similarly, using (A5'), for $T_{32}$ and $T_{33}$ we get

\begin{equation}
T_{32} \leq \frac{C}{2} \sum_{n=1}^{N} \|Q - Q_n\|^2 + \frac{\tau^2}{2} \sum_{n=1}^{N} \|e_q^n\|^2,
\end{equation}

and

\begin{equation}
T_{33} \leq \frac{C}{2} \sum_{n=1}^{N} \|c^n - \tau r\|^2 + \frac{\tau^2}{2} \sum_{n=1}^{N} \|e_q^n\|^2.
\end{equation}

Furthermore, using again (A5'), for any $\delta > 0$ we obtain

\begin{equation}
T_{34} \leq \frac{C\delta_{34}}{2} \sum_{n=1}^{N} \|c^n - c_k^n\|^2 + \frac{\tau^2}{2\delta_{34}} \sum_{n=1}^{N} \|e_q^n\|^2.
\end{equation}

In estimating the reaction term, for any $\delta > 0$ we have

\begin{equation}
T_4 \leq \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{\tau}{\delta_4} \|\Theta(\psi) r(c) - \Theta(\psi_k^n) r(c_k^n)\|^2 + \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{\tau\delta_4}{2} \|P_h\| \cdot c_k^n - c_h^n\|^2
\end{equation}

\begin{equation}
\leq \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{\tau}{\delta_4} \|\Theta(\psi) r(c) - \Theta(\psi_k^n) r(c_k^n)\|^2 + \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{\tau\delta_4}{2} \|P_h\| \cdot c_k^n - c_h^n\|^2.
\end{equation}

We denote the terms on the right by $T_{41}$ and $T_{42}$, and use the inequality $\|f + g\|^2 \leq 2(\|f\|^2 + \|g\|^2)$ to obtain

\begin{equation}
T_{42} \leq \sum_{n=1}^{N} T\delta_4 \|c^n - c_k^n\|^2 + \sum_{n=1}^{N} T\delta_4 \|P_h\| \cdot c_k^n - c_h^n\|^2.
\end{equation}

For $T_{41}$ we obtain in a similar manner

\begin{equation}
T_{41} \leq \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{\tau}{\delta_4} \|\Theta(\psi) r(c) - \Theta(\psi_k^n) r(c)\|^2
\end{equation}

\begin{equation}
+ \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{2\tau}{\delta_4} \|\Theta(\psi_k^n) (r(c) - r(c_k^n))\|^2 + \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{2\tau}{\delta_4} \|\Theta(\psi_k^n) r(c) - r(c_k^n)\|^2.
\end{equation}

We denote the terms in the right hand side above by $T_{411}, T_{412}$ and $T_{413}$. Using the boundedness of $c$ (see Remark 4.1), the Lipschitz continuity of $r(\cdot)$, as well as the Cauchy-Schwarz inequality, $T_{411}$ gives

\begin{equation}
T_{411} = \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{\tau}{\delta_4} \int_{x_{k-1}}^{x_k} \left( \frac{1}{\tau} \int_{t_{k-1}}^{t_k} (\Theta(\psi) - \Theta(\psi_k^n)) r(t) \, dt \right)^2 \, dx.
\end{equation}

\begin{equation}
\leq \frac{1}{\delta_4} \sum_{n=1}^{N} \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|\Theta(\psi) - \Theta(\psi_k^n)\|^2 \, dt \leq \frac{\tau}{\delta_4} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\Theta(\psi) - \Theta(\psi_k^n)\|^2 \, dt.
\end{equation}

Using (A3) and the Lipschitz continuity of $r(\cdot)$, for the second term in (4.45) we have

\begin{equation}
T_{412} = \sum_{n=1}^{N} \sum_{k=1}^{n} \frac{2}{\delta_4} \int_{t_{k-1}}^{t_k} (r(c(t)) - r(c_k^n)) \Theta(\psi_k^n) \, dt
\end{equation}

\begin{equation}
\leq \frac{2C^2 \tau L^2}{\delta_4} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|c(t) - \psi_k^n\|^2 \, dt \leq \frac{C}{\delta_4} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|c(t) - \psi_k^n\|^2 \, dt.
\end{equation}
Finally, we use (A1) and (A3) to estimate $T_{413}$:

$$T_{413} \leq C \frac{1}{\tau} \sum_{n=1}^{N} \left( \sum_{k=1}^{n} \| \bar{e}^n - c_h^n \|^2 + \tau \| v^N_q \|^2 \right).$$

Choosing $\delta_{11}, \delta_{14}, \delta_{16}, \delta_{34}$ and $\delta_1$ properly and using (4.27)–(4.48), a $C > 0$ not depending on $h$ or $\tau$ exists such that

$$\frac{1}{C} \left\{ \sum_{n=1}^{N} \| \bar{e}^n - c_h^n \|^2 + \tau \| v^N_q \|^2 \right\} \leq \sum_{n=1}^{N} \left( \| \Theta(\psi(t_n)) - \Theta(\psi_h^n) \|^2 + \sum_{n=1}^{N} \frac{1}{\tau} \int_{t_{n-1}}^{t_n} \| \Theta(\psi) - \Theta(\psi_h^n) \|^2 dt \right. + \sum_{n=1}^{N} \| \bar{q}^n - \bar{q}_h^n \|^2 \left. + \frac{1}{2} \sum_{n=1}^{N} \| c(t) - c_h^n \|^2 \right) + \frac{1}{2} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \| (c(t) - c_h^n) \|^2 + \sum_{n=1}^{N} \| \bar{e}^n - c_h^n \|^2 + \sum_{n=1}^{N} \tau \sum_{k=1}^{n} \| \bar{e}^k - c_h^k \|^2.$$

Using the projector estimates in (2.4), the regularity of $c$ and $q$ stated in Remark 4.1, as well as Proposition 4.2, applying the discrete Gronwall lemma to (4.49) gives

$$\sum_{n=1}^{N} \| \bar{e}^n - c_h^n \|^2 + \tau \| v^N_q \|^2 \leq C \left\{ \tau + \frac{h^2}{\tau} + \sum_{n=1}^{N} \left( \| \bar{q}^n - \bar{q}_h^n \|^2 + \| \psi(t_n) - \Theta(\psi_h^n) \|^2 \right) \right\} .$$

Now (4.21) follows straightforwardly. 

**Remark 4.2.** Using the stability estimates in Proposition 4.2, (4.21) also implies:

$$\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \| c(t) - c_h^n \|^2 dt + \tau \sum_{n=1}^{N} \left( \| \bar{q}^n - \bar{q}_h^n \|^2 \right) \leq C \left\{ \tau^2 + \frac{h^2}{\tau} + \sum_{n=1}^{N} \left( \frac{1}{\tau} \int_{t_{n-1}}^{t_n} \| \Theta(\psi) - \Theta(\psi_h^n) \|^2 dt \right) \right\} .$$

As follows from (4.21) and (4.51), the estimates for the MFEM scheme applied to the reactive transport equation depend on the errors in the approximation of the Richards equation. In what follows we consider the explicit convergence order, depending only on the discretization parameters. As presented in Theorem 4.4 optimal estimates can be obtained in the fully saturated case. Furthermore, in the strictly unsaturated flow regime we can use Theorem 3.2 to obtain the following.

**Theorem 4.6.** (strictly unsaturated flow) Let $(c, q)$ solve Problem 4.1 and $(c_h^n, q_h^n)$ solve the problems 4.2 for $n \in \{1, \ldots, N\}$. Assuming (A1)–(A5), there holds

$$\sum_{n=1}^{N} \tau \| c^n - c_h^n \|^2 + \left\| \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} q(t) - q_h^n dt \right\| \leq C \left( \tau^2 + h^2 \right) ,$$
where the constant $C$ does not depend on the discretization parameters.

4.1. Error estimates for the case with sorption. In this section we add equilibrium sorption effects to the reactive flow model (1.1). The results are given for the general, unsaturated-saturated flow case. In the case of a fully saturated flow, sharper results can be obtained.

With $s$ denoting the concentration of the adsorbed solute, the transport equation becomes (see, e.g. [8, 9, 18] and the references therein)

$$	ag{4.53} \partial_t (\Theta(\psi)c) + \rho_0 \partial_t s - \nabla \cdot (D \nabla c - Qc) = \Theta(\psi) r(c) \quad \text{in } J \times \Omega,$$

with $\rho_0$ denoting the density of the soil, which is assumed constant. Depending on the fastness of the adsorption process, we can speak about equilibrium process (if adsorption is much faster than diffusion or transport), or non-equilibrium (when all time scales are in balance). In this paper we restrict to equilibrium kinetics, which leads to a degenerate parabolic model. In this case we have

$$
(4.54) \quad s = \phi(c),
$$

with $\phi$ denoting a sorption isotherm. Typical examples of $\phi$ are discussed in [30]: linear, Freundlich, Langmuir or Freundlich-Langmuir. From mathematical point of view, the most interesting is the Freundlich isotherm

$$
(4.55) \quad \phi(c) = c^{\alpha}, \quad \text{for } c \geq 0, \quad \text{and with } \alpha \in (0, 1].
$$

For completeness we extend the function $\phi$ by 0 for all negative arguments. Notice the singularity of the derivative at $c = 0$, therefore $\phi$ is not Lipschitz continuous. In the general setting we assume

$$(A6) \quad \text{The sorption isotherm } \phi \text{ is nondecreasing and Hölder continuous with an exponent } \alpha \in (0, 1], \text{ i.e. } |\phi(a) - \phi(b)| \leq C|a - b|^\alpha \quad \forall a, b \in \mathbb{R}.$$ 

As for Problem 4.1, the continuous mixed variational formulation of (4.53) reads

**Problem 4.3.** (The continuous problem) Find $c \in L^2(J; L^2(\Omega))$ and $q \in L^2(J; H(\text{div}; \Omega))$ such that for almost all $t \in J$ there holds

$$
(\Theta(\psi(t))c(t) - \Theta(\psi_1)c_t, w) + \rho_0 (\phi(c(t)) - \phi(c_t), w) + \langle \nabla \cdot \int_0^t q \, ds, w \rangle = \langle \int_0^t \Theta(\psi)r(c) \, ds, w \rangle,
$$

$$
(4.56) \quad \langle q, v \rangle - \langle c, \nabla \cdot v \rangle + \langle cQ, v \rangle = 0,
$$

for all $w \in L^2(\Omega)$ and $v \in H(\text{div}; \Omega)$.

The sorption model (4.53) may degenerate for $c = 0$. Here we are interested in the convergence of the numerical scheme, and we do not focus on questions concerning the existence, uniqueness and the regularity of a solution. Therefore we assume that Problem 4.3 has a unique solution, having the regularity mentioned in Remark 4.1. This allows us maintaining the working framework of the previous section.

Below we define the fully discrete scheme for Problem 4.3:

**Problem 4.4.** Let $n \in \{1, \ldots, N\}$ and $\Theta(\psi^n_h), \Theta(\psi^{n-1}_h), Q^n_h, c^{n-1}_h$ be given. Find $(c^n_h, q^n_h) \in W_h \times V_h$ such that

$$
(\Theta(\psi^n_h)c^n_h - \Theta(\psi^{n-1}_h)c^{n-1}_h, w_h) + \rho_0 (\phi(c^n_h) - \phi(c^{n-1}_h), w_h) + \tau \langle \nabla \cdot q^n_h, w_h \rangle = \tau (\Theta(\psi^n_h) r(c^n_h), w_h),
$$

$$
(4.58) \quad \langle q^n_h, v_h \rangle - \langle c^n_h, \nabla \cdot v_h \rangle + \langle c^n_h Q^n_h, v \rangle = 0,
$$

$$
(4.59) \quad \langle q^n_h, v_h \rangle - \langle c^n_h, \nabla \cdot v_h \rangle + \langle c^n_h Q^n_h, v \rangle = 0,
$$

15
for all \( w_h \in W_h \) and \( v_h \in V_h \).

Initially we take a \( c_0^h \in W_h \) such that \( \Theta(q^0_h)c_0^h + \phi(c_0^h) = P_h(\Theta(q_1)c_1 + \phi(c_1)) \). This choice, allowed since \( \Theta(q^0_h) \geq \Theta_R > 0 \) while \( \phi \) is nondecreasing, gives for any \( w_h \in W_h \)

\[
(4.60) \quad \langle \Theta(q^0_h)c_0^h + \phi(c_0^h), w_h \rangle = \langle \Theta(q_1)c_1 + \phi(c_1), w_h \rangle.
\]

For the discretization of Problem 4.3 we have the following convergence result:

**Theorem 4.7.** (saturated-unsaturated flow) Let \((c, q)\) solve Problem 4.3 and \((c_n^h, q_n^h)\) the solution of Problem 4.4 for all \( n \in \{1, \ldots, N\} \). Assuming \((A1)-(A6)\), a \( C > 0 \) not depending on the discretization parameters exists such that

\[
\sum_{n=1}^{N} \tau \| v^n - c_n^h \|^2 + \tau^2 \left\| \sum_{n=1}^{N} (\Pi_h v^n - q_n^h) \right\|^2 \leq C \left\{ \tau \frac{4\alpha}{1+\alpha} + h^{1+\alpha} + \sum_{n=1}^{N} \tau \| q^n - Q_n^h \|^2 \right. \]

\[
+ \left. \sum_{n=1}^{N} \tau \| \Theta(q(t_n)) - \Theta(q_n^h) \|^2 + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \| \Theta(q(t)) - \Theta(q_n^h) \|^2 dt \right\}.
\]

**Proof.** We follow the ideas in the proof of (4.21). When compared to the scheme for Problem 4.1, to discretize the case with sorption one adds the term \( \rho_0(\phi(c^h) - \phi(c^h_{n-1}), w_h) \) in (4.58). In the present proof we focus on the differences brought by this additional term. Specifically, in the left hand side of (4.27) we have to add

\[
T_S = \rho_0 \sum_{n=1}^{N} (\phi(c^n) - \phi(c_n^h), P_h v^n - c_n^h).
\]

To estimate \( T_S \) we split it into

\[
(4.62) \quad T_S = \rho_0 \sum_{n=1}^{N} (\phi(c^n) - \phi(c_n^h), c^n - c_n^h) + \rho_0 \sum_{n=1}^{N} (\phi(c^n) - \phi(c_n^h), P_h v^n - c^n)
\]

and denote the terms on the right by \( T_{S1} \) and \( T_{S2} \). By the monotonicity of the sorption isotherm \( \phi \), \( T_{S1} \) is positive; moreover, (A6) gives

\[
(4.63) \quad T_{S1} \geq \rho_0 \| \phi(c^n) - \phi(c_n^h) \|_{L^{1+\alpha}(\Omega)}^{1+\alpha}.
\]

For \( T_{S2} \) we use the Young inequality

\[
ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad \text{for any } a, b > 0 \text{ and } p, q \in (1, \infty) \text{ such that } \frac{1}{p} + \frac{1}{q} = 1.
\]

For any \( \delta > 0 \), with \( a = \delta|\phi(c^n) - \phi(c_n^h)|, b = \frac{|P_h v^n - c^n|}{\delta} \) and \( p = \frac{1+\alpha}{\alpha} \) this gives

\[
(4.64) \quad T_{S2} \leq \frac{\alpha \rho_0}{1+\alpha} \delta \frac{1+\alpha}{\delta} \| \phi(c^n) - \phi(c_n^h) \|_{L^{1+\alpha}(\Omega)}^{1+\alpha}
\]

\[
+ \rho_0 \delta^{-1+\alpha} \sum_{n=1}^{N} \| P_h v^n - c^n \|_{L^{1+\alpha}(\Omega)}^{1+\alpha}.
\]

Denoting the two terms above by \( T_{S21} \) and \( T_{S22} \), we notice that choosing \( \delta \) properly allows compensating the first one by (4.63). Since \( \alpha \in (0, 1] \), to estimate \( T_{S22} \) we use
(2.4), the continuous embedding $L^2(\Omega) \subseteq L^{1+\alpha}(\Omega)$, the stability estimates (4.10) and (4.13), as well as the Young inequality
\[
x^{1+\alpha} \leq \frac{1 - \alpha}{2} \tau^2 + \frac{1 + \alpha}{2} \tau^{\frac{2(\alpha-1)}{1+\alpha}}x^2,
\]
(valid for all $x > 0$) and obtain
\[
T_{S22} \leq \frac{\rho_b}{1 + \alpha} \sum_{n=1}^{N} \left\| \frac{\partial (\bar{P} \bar{c} - c^n)}{\partial \tau} \right\|_{1+\alpha} \left\| \bar{c} - c^n \right\|_{1+\alpha}
\leq Ch^{1+\alpha} \sum_{n=1}^{N} \left( \frac{1 - \alpha}{2} \tau^2 + \frac{1 + \alpha}{2} \tau^{\frac{2(\alpha-1)}{1+\alpha}} \right)
+C \sum_{n=1}^{N} \left( \frac{1 - \alpha}{2} \tau^2 + \frac{1 + \alpha}{2} \tau^{\frac{2(\alpha-1)}{1+\alpha}} \right)
\leq C \left( \frac{h^{1+\alpha}}{\tau} + \tau^{1+\frac{2(\alpha-1)}{1+\alpha}} \right).
\]
(4.65)

Now we can use (4.62)-(4.65) and proceed as in the proof of (4.21) to obtain the estimates in (4.61).

5. Numerical Results. This section contains some numerical simulations for the flow and contaminant transport. To be specific, the water content as well as the flux are not considered as given, but obtained by solving the Richards equation (1.5). Once these being determined, we determine the solute concentration by solving (1.1). To do so, we apply the El-MFE schemes presented in Sections 3 and 4.

With $\Omega = (0,1)^2$ and $T = 1$ (hence $J = (0,1)$), for determining the saturation and the flux we solve
\[
\partial_t \Theta(\psi) - \Delta \psi = f \quad \text{in } J \times \Omega
\]
in two situations, which are called later as Example 1 and Example 2. First we consider a linear problem by taking $\Theta(\psi) = \psi$, whereas $\Theta(\psi) = \psi^2$ in the second case. The right hand side $f$ is chosen such that the equation (5.1) admits the analytical solution
\[
\psi_{ex}(t,x,y) = 4 - 2x - 4tx(1-x)y^2(1-y)^2,
\]
yielding a saturation $\Theta$ and a flux $Q$ satisfying (A3) and (A5). The initial and boundary conditions are also chosen accordingly: $\psi(0,x,y) = 4 - 2x$, as well as $\psi = \psi_{ex}$ along $J \times \partial \Omega$.

Next we consider the solute transport equation
\[
\partial_t (\Theta(\psi)c) - \nabla \cdot (D \nabla c - Qc) = \Theta(\psi)r(c) + g \quad \text{in } J \times \Omega
\]
with $\Theta(\psi)$ and $Q$ provided by (5.1). We choose $D = 10.0$, whereas $r(c) = c^2$. Furthermore, $g$ and the initial and boundary conditions are chosen such that the equation (5.3) has the exact solution
\[
c_{ex}(t,x,y) = 100(1 - \exp(-t))y^4(1-y)^4(1-x) + 1.
\]
Table 5.1
Numerical results for the Richards equation, Example 1.

<table>
<thead>
<tr>
<th>N</th>
<th>τ = h</th>
<th>E_R</th>
<th>γ_R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>7.099274e-05</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>1.559765e-05</td>
<td>2.19</td>
</tr>
<tr>
<td>3</td>
<td>0.025</td>
<td>-3.472974e-06</td>
<td>2.17</td>
</tr>
<tr>
<td>4</td>
<td>0.0125</td>
<td>8.717040e-07</td>
<td>1.99</td>
</tr>
<tr>
<td>5</td>
<td>0.00625</td>
<td>2.167524e-07</td>
<td>2.01</td>
</tr>
</tbody>
</table>

All simulations are carried out with τ = h. We work on a uniform mesh, starting with h = 0.1 and refine it successively by halving the grid size and the time step. For the Richards equation we compute the error:

\[ E_R := \sum_{n=1}^{N} \tau \| \psi_{ex}(t_n) - \psi_h^n \|_2^2 + \sum_{n=1}^{N} \tau \| Q_{ex}(t_n) - Q_h^n \|_2^2, \]

whereas for the transport equation we calculate the error for the concentration:

\[ E_T := \sum_{n=1}^{N} \tau \| c_{ex}^n - c_h^n \|_2^2. \]

To compare the numerical results with the theoretical estimates we have estimated the reduction order of two two successive errors, computed for two discretization parameters - say \( h_1 \) and \( h_2 \). If the estimate for the error E is \( h^\gamma \) for some \( \gamma > 0 \), then the ratio \( E_1/E_2 \) should behave asymptotically as \((h_1/h_2)^\gamma\). The tables presenting the numerical results have a generic structure: a column for the discretization parameters, another one for the errors, and finally a column for the exponent \( \gamma \) estimated from two successive calculations.

We start with Examples 1 and 2, which are regular parabolic problems. These cases can be assimilated to a strictly unsaturated flow. Following Theorem 3.2, \( E_R \) should behave at least as \( \tau^2 + h^2 \). Since \( \tau = h \), asymptotically we should obtain \( \gamma_R = 2 \). The numerical results for the first two examples are included in Tables 5.1 and 5.3.

Using the saturation and the flux calculated above, we solve (5.3). The numerical results are presented in Tables 5.2 and 5.4. Following Theorem 4.6, the asymptotic order for \( E_T \) should be again \( \gamma_T = 2 \). Notice the good agreement of the numerical results with the theoretical estimates in Theorems 3.2 and 4.6.

Table 5.2
Numerical results for the solute transport, Example 1.

<table>
<thead>
<tr>
<th>N</th>
<th>τ = h</th>
<th>E_T</th>
<th>γ_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>2.844167e-05</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>6.539999e-06</td>
<td>2.12</td>
</tr>
<tr>
<td>3</td>
<td>0.025</td>
<td>1.561840e-06</td>
<td>2.06</td>
</tr>
<tr>
<td>4</td>
<td>0.0125</td>
<td>4.067868e-07</td>
<td>1.94</td>
</tr>
<tr>
<td>5</td>
<td>0.00625</td>
<td>1.029391e-07</td>
<td>1.98</td>
</tr>
</tbody>
</table>

In the next example the Richards equation (5.1) is degenerate. This is due to the
nonlinearity
\[
\Theta(\psi) := \begin{cases}
\frac{19}{6}, & \text{if } \psi < 1, \\
\frac{\psi^3}{3} + \frac{3\psi^2}{2} - 2\psi + 4, & \text{if } 1 \leq \psi \leq 2, \\
\frac{10}{3}, & \text{if } 2 < \psi.
\end{cases}
\]

It is to see that \(\Theta(\cdot)\) is \(C^1\) and Lipschitz continuous, whereas \(\Theta' = 0\) on \(\mathbb{R}\setminus (1, 2)\). Therefore (5.1) is parabolic whenever \(\psi \in (1, 2)\), and elliptic for the other values of \(\psi\). The initial and boundary conditions, as well as the right hand side \(f(\cdot)\) are chosen such that (5.1) is solved by

\[
\psi_{ex}(t, x, y) = 64tx(1-x)y(1-y).
\]

With the pressure above we calculate the water content \(\Theta\) and the water flux \(Q\) and consider the solute transport equation (5.3) with homogeneous Dirichlet boundary conditions and zero initial condition. As before we take \(r(c) = c^2\), and the source \(g\) such that

\[
c_{ex}(t, x, y) = tx(1-x)y(1-y)
\]

is a solution of (5.3). Tables 5.5 and 5.6 are presenting the numerical results for the Richards equation, respectively the solute transport. Notice that the convergence order is still at least 2. This is in spite of the theoretical estimates of order \(\tau + h^2\), which are obtained even in a weaker norm, but for the Richards equation in the saturated/unsaturated flow regime (see [5, 31, 33, 37]). This improved convergence for the flow leads to better results for the solute transport, similar to the theoretical estimates in Theorem 4.5.

Having in mind the above results, the following question arises naturally: in how far the estimates in (4.21) are reflecting a correct dependence of the solute error on
Table 5.5
Numerical results for the Richards equation, Example 3.

<table>
<thead>
<tr>
<th>N</th>
<th>τ = h</th>
<th>( E_R )</th>
<th>γ_R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>1.448942e-01</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>1.123385e-02</td>
<td>3.69</td>
</tr>
<tr>
<td>3</td>
<td>0.025</td>
<td>1.750283e-03</td>
<td>2.68</td>
</tr>
<tr>
<td>4</td>
<td>0.0125</td>
<td>4.331334e-04</td>
<td>2.01</td>
</tr>
<tr>
<td>5</td>
<td>0.00625</td>
<td>1.080408e-04</td>
<td>2.00</td>
</tr>
</tbody>
</table>

the errors in the flux and saturation? In this sense we mention the following situation: given a stationary flow and its numerical approximation satisfying

\[ \sum_{n=1}^{N} \tau \| Q_{ex}^n - Q_h^n \|_2^2 \leq Ch^{\beta_1}, \]

(5.10)
is it possible to obtain a more accurate approximation of the solute, say an error of order \( h^{\beta_2} \) with \( \beta_2 > \beta_1 \)?

This situation appears in many practical situations, when one wants to know if it makes sense to use higher order finite elements for the solute transport, in combination with a lower order finite element approximation of the flow (see also [11, 12]). Having in mind this very important question, we have carried out the last numerical test, by considering a constant saturation \( \Theta = 1 \) and flux \( Q = (1,0) \) in the solute transport equation (5.3). In discretizing (5.3) we use an approximation of the flux \( Q \) given by \( Q_h = (1 + 10\sqrt{h}, 0) \), yielding for the given \( T \) and \( \Omega \)

\[ E_R = \sum_{n=1}^{N} \tau \| Q_{ex}^n - Q_h^n \|_2^2 = 100h. \]

(5.11)

Further we maintain the framework of Example 3. Table 5.7 presents the error \( E_T \) given in (5.6), as well as the estimated order of convergence \( \gamma_T \). It is worth noticing that the present computations suggest a decrease of the accuracy down to the order 1, identical to the order of \( E_R \). This suggests an answer to the question above, in the sense that the approximation of the solute should be in balance with the one for the flow. Nevertheless, as revealed in several numerical experiments carried out for problems that are not dominated by convection, or at least where the errors in the approximation of the convection are not dominating, the reduction in the convergence order only appears on very fine grids.

Table 5.6
Numerical results for the solute transport, Example 3.

<table>
<thead>
<tr>
<th>N</th>
<th>τ = h</th>
<th>( E_T )</th>
<th>( \gamma_T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>1.862225e-05</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>1.912738e-06</td>
<td>3.28</td>
</tr>
<tr>
<td>3</td>
<td>0.025</td>
<td>2.730756e-07</td>
<td>2.81</td>
</tr>
<tr>
<td>4</td>
<td>0.0125</td>
<td>5.605202e-08</td>
<td>2.28</td>
</tr>
<tr>
<td>5</td>
<td>0.00625</td>
<td>1.318858e-08</td>
<td>2.09</td>
</tr>
</tbody>
</table>

6. Conclusions. We have analyzed a numerical scheme for a porous media solute transport model, where the transporting flux is obtained by solving numerically
Table 5.7

Numerical results for the solute transport, Example 4.

<table>
<thead>
<tr>
<th>N</th>
<th>$\tau = h$</th>
<th>$E_T$</th>
<th>$\gamma_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>1.204681e-05</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>5.097287e-06</td>
<td>1.24</td>
</tr>
<tr>
<td>3</td>
<td>0.025</td>
<td>2.299894e-06</td>
<td>1.15</td>
</tr>
<tr>
<td>4</td>
<td>0.0125</td>
<td>1.086988e-06</td>
<td>1.08</td>
</tr>
<tr>
<td>5</td>
<td>0.00625</td>
<td>5.278606e-07</td>
<td>1.04</td>
</tr>
</tbody>
</table>

The Richards equation. The spatial discretization is mixed and based on the lowest order Raviart - Thomas finite elements, whereas the time stepping is performed by the Euler implicit method. We have proven the convergence of the scheme by estimating the error in terms of the discretization parameters. In doing so we have taken into account the numerical error occurring in the approximation of the fluid flow. Furthermore, we avoid some of the commonly made assumptions concerning the boundedness of the time derivative of the saturation, or a strict sign of the water flux or of its divergence. The numerical experiments agree with the estimates derived theoretically.

Acknowledgements. Part of the work of the first author was done during his stay at the Max Planck Institute for Mathematics in the Sciences in Leipzig. The work of the second author was supported by the Dutch government through the national program BSIK: knowledge and research capacity, in the ICT project BRICKS (http://www-bsik-bricks.nl), theme MSV1.

REFERENCES


