CONVERGENCE ANALYSIS OF MIXED NUMERICAL SCHEMES FOR REACTIVE FLOW IN A POROUS MEDIUM

K. KUMAR†, I. S. POP‡, AND F. A. RADU§

Abstract. This paper deals with the numerical analysis of an upscaled model describing the reactive flow in a porous medium. The solutes are transported by advection and diffusion and undergo precipitation and dissolution. The reaction term and, in particular, the dissolution term have a particular, multivalued character, which leads to stiff dissolution fronts. We consider the Euler implicit method for the temporal discretization and the mixed finite element for the discretization in space. More precisely, we use the lowest order Raviart–Thomas elements. As an intermediate step we consider also a semidiscrete mixed variational formulation (continuous in space). We analyze the numerical schemes and prove the convergence to the continuous formulation. Apart from the proof for the convergence, this also yields an existence proof for the solution of the model in mixed variational formulation. Numerical experiments are performed to study the convergence behavior.

Key words. numerical analysis, reactive flows, weak formulation, implicit scheme, mixed finite element discretization

AMS subject classifications. 35A35, 65L60, 65J20

DOI. 10.1137/120880938

1. Introduction. Reactive flows in a porous medium have a wide range of applications ranging from spreading of polluting chemicals leading to ground water contamination (see [40] and references therein) to biological applications such as tissue and bone formation, or pharmaceutical applications [27], or technological applications such as operation of solid batteries. A common feature of the above applications is the transport and reactions of ions/solutes. In this work, we deal with the transport of ions/solutes taking place through the combined process of convection and diffusion. For reactions, we focus on a specific class, namely, the precipitation and dissolution processes, where the ions undergo combination (precipitation) to form a crystal. The reverse process of dissolution takes place where the crystal gets dissolved.

We consider here an upscaled model defined on a Darcy scale. This implies that the solid grains and the pore space are not distinguished and the equations are defined everywhere. Consequently, the crystals formed as a result of reactions among ions and the ions themselves are defined everywhere in the domain. Such models fall in the general category of reactive porous media flow models. For Darcy-scale models related directly to precipitation and dissolution processes we refer to [7, 29, 33, 34] (see also the references therein). Here we adopt the ideas proposed first in [23], and extended in a series of papers [15, 16, 17]. These papers are referring to Darcy-scale models; the pore-scale counterpart is considered in [18], where distinction is made for the domains...
delineating the pore space and the solid grains. The transition from the pore-scale model to the upscaled model is obtained, for instance, via homogenization arguments. For a simplified situation of a two-dimensional (2D) strip, the rigorous arguments are provided in [18]; see also [28, 1] for the upscaling procedure in transport dominated flow regimes. For a similar situation, but tracking the geometry changes due to the reactions leading to the free boundary problems, the formal arguments are presented in [24] and [30].

We are motivated by analyzing appropriate numerical methods for solving the reactive flows for an upscaled model. Considering the mixed variational formulation is an attractive proposition as it preserves the mass locally. Our main goal here is to provide the convergence of a mixed finite element discretization for such a model for dissolution and precipitation in porous media, involving a multivalued dissolution rate. Before discussing the details and specifics, we briefly review some of the relevant numerical works. For continuously differentiable rates the convergence of (adaptive) finite volume discretizations is studied in [22, 32]; see also [10] for the convergence of a finite volume discretization of a copper-leaching model. In a similar framework, discontinuous Galerkin methods are discussed in [39] and upwind mixed finite element methods (MFEMs) are considered in [11, 12]; combined finite volume–mixed hybrid finite elements are employed in [20, 21]. Non-Lipschitz, but Hölder continuous rates are considered using conformal finite element method (FEM) schemes in [4, 5]. Similarly, for Hölder continuous rates (including equilibrium and nonequilibrium cases), MFEMs are analyzed rigorously in [36, 38], whereas [37] provides error estimates for the coupled system describing unsaturated flow and reactive transport. In all these cases, the continuity of the reaction rates allows error estimates to be obtained. A characteristic MFEM for the advection dominated transport has been treated in [3] and characteristic FEM schemes for contaminant transport giving rise to possibly non-Lipschitz reaction rates are treated in [13] where the convergence and the error estimates have been provided. A parabolic problem coupled with linear ODEs at the boundary have been treated in [2] using a characteristic MFEM. Conformal schemes both for the semidiscrete and fully discrete (FEM) cases for the upscaled model under consideration have been treated in [25].

The main difficulty here is due to the particular description of the dissolution rate, which may become discontinuous. To deal with this, we consider a regularization of this term and the corresponding sequence of regularized equations. The regularization parameter \( \delta \) is dependent on the time discretization parameter \( \tau \) in such a way that as \( \tau \downarrow 0 \), it is ensured that \( \delta \downarrow 0 \). Thus, obtaining the limit of the discretized scheme automatically yields, by virtue of the regularization parameter also vanishing, the original equation. In proving the convergence results, compactness arguments are employed. These arguments rely on a priori estimates providing weak convergence. However, strong convergence is needed to deal with the nonlinear terms in the reaction rates. Translation estimates are used to achieve this.

We consider both the semidiscrete and the fully discrete cases with the proof for the latter case following closely the ideas of the semidiscrete case. However, there are important differences, particularly in the way the translation estimates are obtained. Whereas in the semidiscrete case, we use the dual problem for obtaining the translation estimates, in the fully discrete case, we use the properties of the discrete \( H^1_0 \) norm following the finite volume framework [19]. The convergence analysis of appropriate numerical schemes for the problem considered here is a stepping stone for coupled flow and reactive transport problems (for example, Richards’ equation coupled with precipitation-dissolution reaction models).
The paper is structured as follows. We begin with a brief description of the model in section 2. We define the mixed variational formulation in section 2.2 where we prove the uniqueness of the solution with the existence coming from the convergence proof. Next, in sections 3 and 4 the time-discrete, respectively, fully discrete numerical schemes are considered and the proofs for the convergence are provided. The numerical experiments are shown in section 5 followed by the conclusions and discussions in section 6.

2. The mathematical model. We consider a Darcy-scale model that describes the reactive transport of the ions/solutes in a porous medium. The solutes are subjected to convective transport and in addition they undergo diffusion and reactions in the bulk. Below we provide a brief description and the assumptions of the model; we refer to [15], or [16] for more details.

Let \( \Omega \subset \mathbb{R}^2 \) be the domain occupied by the porous medium, and assume \( \Omega \) open, connected, bounded, and with Lipschitz boundary \( \Gamma \). Further, let \( T > 0 \) be a fixed but arbitrarily chosen time, and define \( \Omega^T = (0, T] \times \Omega \), and \( \Gamma^T = (0, T] \times \Gamma \). At the outset, we assume that the fluid velocity \( q \) is known, divergence free, and essentially bounded.

Usually, two or more different types of ions react to produce precipitate (an immobile species). A simplified model will be considered here where we include only one mobile species. This makes sense if the boundary and initial data are compatible (see [15], or [16]). Then, denoting by \( v \) the concentration of the (immobile) precipitate, and by \( u \) the cation concentration, the model reduces to

\[
\begin{cases}
\partial_t(u + v) + \nabla \cdot (qu - \nabla u) = 0 & \text{in } \Omega^T, \\
u = 0 & \text{on } \Gamma^T, \\
u = u_I & \text{in } \Omega, \text{ for } t = 0,
\end{cases}
\]

for the ion transport, and

\[
\begin{cases}
\partial_t v = (r(u) - w) & \text{on } \Omega^T, \\
w \in H(v) & \text{on } \Omega^T, \\
v = v_I & \text{on } \Omega, \text{ for } t = 0,
\end{cases}
\]

for the precipitate. For the ease of presentation we restrict ourselves to homogeneous Dirichlet boundary conditions. The assumptions for the initial conditions will be given below.

In the system considered above, we assume all the quantities and variables as dimensionless. To simplify the exposition, the diffusion is assumed to be 1, the extension to a positive definite diffusion tensor being straightforward. Further, we assume that the Damköhler number is scaled to 1, as well as an eventual factor in the time derivative of \( v \) in (2.2), appearing in the transition form the pore scale to the core scale.

The assumptions on the precipitation rate \( r \) are the following:

\( (A_{r1}) \) \( r(\cdot) : \mathbb{R} \to [0, \infty) \) is Lipschitz continuous in \( \mathbb{R} \) with Lipschitz constant \( L_r \);

\( (A_{r2}) \) there exists a unique \( u_* \geq 0 \), such that

\[
r(u) = \begin{cases} 0 & \text{for } u \leq u_*, \\
\text{strictly increasing for } u \geq u_* & \text{with } r(\infty) = \infty.
\end{cases}
\]

The interesting part is the structure of the dissolution rate. We interpret it as a process encountered strictly at the surface of the precipitate layer, so the rate is
In the setting above, a unique precipitation gain, or positive. This can be summarized as
dissolution) is either zero, if the solute present there is insufficient to produce a net precipitation gain, or positive.

\[ w \in H(v), \quad \text{where } H(v) = \begin{cases} \{0\} & \text{if } v < 0, \\ [0,1] & \text{if } v = 0, \\ \{1\} & \text{if } v > 0. \end{cases} \]

In the setting above, a unique \( u^* \) exists for which \( r(u^*) = 1 \). If \( u = u^* \) for all \( t \)
and \( x \), then the system is in equilibrium: no precipitation or dissolution occurs, since
the precipitation rate is balanced by the dissolution rate regardless of the presence of absence of crystals (see [25, section 5] for some illustrations). Then, as follows from
\[ \text{for a.e. } (t,x) \in \Omega^T \text{ where } v = 0, \text{ the dissolution rate satisfies} \]

\[ w = \begin{cases} r(u) & \text{if } u < u^*, \\ 1 & \text{if } u \geq u^*. \end{cases} \]

Since we will work with the model in the mixed formulation, we define the flux as

\[ Q = -\nabla u + qu. \]

2.1. Notation. We adopt the standard notation from functional analysis. In
particular, by \( H^1(\Omega) \) we mean the space of functions in \( H^1(\Omega) \), having a vanishing trace on \( \Gamma \), and \( H^{-1} \) is its dual. By \( (\cdot,\cdot) \) we mean the \( L^2 \) inner product or the duality pairing between \( H^1_0 \) and \( H^{-1} \). Further, \( \|\cdot\| \) stands for the norms induced by the \( L^2 \) inner product. For other norms, we explicitly state it. The functions in \( H(\text{div};\Omega) \) are vector valued having an \( L^2 \) divergence. Furthermore, \( C \) denotes a generic constant
the value of which might change from line to line and is independent of unknown variables or the discretization parameters.

Having introduced this notation we can state the assumptions on the initial conditions:

\( (A_{I1}) \) The initial data \( u_I \) and \( v_I \) are nonnegative and essentially bounded;

\( (A_{I2}) \) \( u_I, v_I \in H^1_0(\Omega) \).

We have taken the initial conditions in \( H^1_0 \) to avoid technicalities. Alternatively,
one can approximate the initial conditions by taking the convolutions with smooth functions. The \( H^1_0 \) regularity for \( v_I \) is used for obtaining strong convergence results,
for which \( L^2 \) regularity is not sufficient.

We furthermore assume that \( \Omega \) is polygonal. Therefore it admits regular decompositions into simplices and the errors due to nonpolygonal domains are avoided. The spatial discretization will be defined on such a regular decomposition \( \mathcal{T}_h \) into 2D simplices (triangles); \( h \) stands for the mesh size. We provide the exposition for two dimensions but extending the results to three dimensions is similar. We define the following sets:

\[ \mathcal{V} := H^1((0,T);L^2(\Omega)), \]

\[ \mathcal{S} := L^2((0,T);H(\text{div};\Omega)), \]

\[ \mathcal{W} := \{w \in L^\infty(\Omega^T); 0 \leq w \leq 1\}. \]

In addition, for the fully discrete situation, we use the following discrete subspaces
\( \mathcal{V}_h \subset L^2(\Omega) \) and \( \mathcal{S}_h \subset H(\text{div};\Omega) \) defined as follows:

\[ \mathcal{V}_h := \{u \in L^2(\Omega) \mid \text{u is constant on each element } T \in \mathcal{T}_h\}, \]

\[ \mathcal{S}_h := \{Q \in H(\text{div};\Omega) \mid Q|_T = a + bx \text{ for all } T \in \mathcal{T}_h\}. \]
In other words, \( \mathcal{V}_h \) denotes the space of piecewise constant functions, while \( \mathcal{S}_h \) is the \( RT_h \) space. Clearly from the above definitions, \( \nabla \cdot Q \in \mathcal{V}_h \) for any \( Q \in \mathcal{S}_h \).

We also define the following usual projections:

\[
P_h : L^2(\Omega) \to \mathcal{V}_h, \quad (P_h v - v, v_h)_h = 0
\]

for all \( v_h \in \mathcal{V}_h \). Similarly, the projection \( \Pi_h \) is defined on \( (H^1(\Omega))^2 \) such that

\[
\Pi_h : (H^1(\Omega))^2 \to \mathcal{S}_h, \quad (\nabla \cdot (\Pi_h Q - Q), v)_h = 0
\]

for all \( v_h \in \mathcal{V}_h \). Following [35, p. 237] (also see [9]), this operator can be extended to \( H(\text{div}; \Omega) \) and also for the above operators there holds

\[
\|v - P_h v\| \leq C h \|v\|_{H^1(\Omega)} \quad \text{for all } v \in H^1(\Omega),
\]

\[
\|Q - \Pi_h Q\| \leq C h \|Q\|_{H^1(\Omega)} \quad \text{for all } Q \in (H^1(\Omega))^2,
\]

\[
\|\nabla \cdot Q - \nabla \cdot (\Pi_h Q)\| \leq C h \|Q\|_{H^2(\Omega)} \quad \text{for all } Q \in (H^2(\Omega))^2.
\]

For the spatial discretization we will work with the approximation \( q_h \) of the Darcy velocity \( q \), defined on the given mesh \( T_h \). For this approximation we assume that there exists an \( M_q > 0 \) such that \( s.t. \|q_h\|_{L^\infty} \leq M_q \) (the same estimate being valid for \( q \)), and as \( h \to 0 \)

\[
\|q_h - q\| \to 0.
\]

Having stated the assumptions, we proceed by introducing the mixed variational formulation and analyzing the convergence of its discretization.

### 2.2. Continuous mixed variational formulation.

Except for some particular situations, one cannot expect the existence of classical solutions to (2.1)–(2.2) and we work with the weak formulation. A weak solution of (2.1)–(2.2) written in mixed form is defined as follows.

**Definition 2.1.** A quadruple \((u, Q, v, w) \in \mathcal{V} \times \mathcal{S} \times \mathcal{V} \times \mathcal{W}\) with \( u|_{t=0} = u_l, v|_{t=0} = v_t \) is a mixed weak solution of (2.1)–(2.2) if \( w \in H(\cdot) \) a.e. and for all \( t \in (0, T) \) and \((\phi, \theta, \psi) \in L^2(\Omega) \times L^2(\Omega) \times H(\text{div}; \Omega)\) we have

\[
(\partial_t u, \phi) + (\nabla \cdot Q, \phi) + (\partial_t v, \phi) = 0,
\]

\[
(\partial_t v, \theta) - (r(u) - w, \theta) = 0,
\]

\[
(Q, \psi) - (u, \nabla \cdot \psi) - (qu, \psi) = 0.
\]

The proof for the existence of a solution for (2.9) is obtained by the convergence of the numerical schemes considered below. Therefore, we give the proof for the uniqueness of the solution. The following lemma shows the uniqueness without further details on \( w \). As mentioned in (2.5) the inclusion \( w \in H(\cdot) \) can be made more precise.

**Lemma 2.2.** The mixed weak formulation (2.9) has at most one solution.

**Proof.** Assume there exist two solution quadruples \((u_1, Q_1, v_1, w_1)\) and \((u_2, Q_2, v_2, w_2)\), and define \( u := u_1 - u_2, \quad Q := Q_1 - Q_2, \quad v := v_1 - v_2, \quad w := w_1 - w_2 \). Clearly, at \( t = 0 \) we have \( u(0, x) = 0 \) and \( v(0, x) = 0 \) for all \( x \).

Subtracting (2.9)2 for \( u_2, v_2, \) and \( w_2 \) from the equations for \( u_1, v_1, \) and \( w_1 \) and taking (for \( t \leq T \) arbitrary) \( \theta = \chi_{(0,t)} v \), using monotonicity of \( H \) and the Lipschitz continuity of \( r(\cdot) \) leads to

\[
\|v(t, \cdot)\|^2 = \int_0^t \int_\Omega (r(u_1) - r(u_2))v(s, x)dxds - \int_0^t \int_\Omega (H(v_1) - H(v_2))v(s, x)dxds
\leq \frac{1}{2} \int_0^t L_1^2 \|u(s, \cdot)\|^2ds + \frac{1}{2} \int_0^t \|v(s, \cdot)\|^2ds.
\]
Then Gronwall’s lemma gives
\begin{equation}
\|v(t,\cdot)\|^2 \leq C \int_0^t \|u(s,\cdot)\|^2 ds.
\end{equation}

Next, we choose \( \phi = \chi_{(0,t)}(u(t,x)) \) in the difference between the two equalities (2.9) to get
\[
\|u(t,\cdot)\|^2 + \left( \int_0^t \nabla \cdot Q(s,\cdot) ds, u(t,\cdot) \right) + (v(t,\cdot), u(t,\cdot)) = 0.
\]

Similarly, choosing \( \psi = \int_0^t Q(s)ds \) in (2.9) yields
\[
\int_\Omega \left( \int_0^t Q(s,x) ds \right) dx - \int_\Omega u(t,x) \left( \int_0^t \nabla \cdot Q(s,x) ds \right) dx
= \int_\Omega q u(t,x) \left( \int_0^t Q(s,x) ds \right) dx.
\]

Combining the above gives
\[
\|u(t,\cdot)\|^2 + \int_\Omega v(t,x)u(t,x)dx + \int_\Omega Q(t,x)\int_0^t Q(s,x)dsdx
\leq \frac{1}{4}\|u(t,\cdot)\|^2 + M_q^2 \left\| \int_0^t Q(s,\cdot)ds \right\|^2,
\]
which implies
\[
\|u(t,\cdot)\|^2 + (Q(t,\cdot), \int_0^t Q(s,\cdot)ds) \leq \frac{1}{2}\|u(t,\cdot)\|^2 + M_q^2 \left\| \int_0^t Qds \right\|^2 + \|v(t,\cdot)\|^2.
\]

Using (2.10) we obtain
\begin{equation}
\frac{1}{2}\|u(t,\cdot)\|^2 + (Q(t,\cdot), \int_0^t Q(s,\cdot)ds) \leq C \int_0^t \|u(s,\cdot)\|^2 ds + M_q^2 \left\| \int_0^t Q(s,\cdot)ds \right\|^2.
\end{equation}

The uniqueness follows now by applying Gronwall’s lemma.

3. **Semidiscrete mixed variational formulation.** As announced, to avoid dealing with inclusion in the description of dissolution rate, the numerical scheme relies on the regularization of the Heaviside graph. With this aim, with \( \delta > 0 \) we define
\begin{equation}
H_\delta(z) := \begin{cases} 
1, & z > \delta, \\
\frac{z}{\delta}, & 0 \leq z \leq \delta, \\
0, & z < 0.
\end{cases}
\end{equation}

Next, with \( N \in \mathbb{N}, \tau = \frac{T}{N} \), and \( t_n = n\tau, n = 1, \ldots, N \), we consider a first order time discretization with uniform time stepping, which is implicit in \( u \) and explicit in \( v \). At each time step \( t_n \) we use \( (u_n^{n-1}, v_n^{n-1}) \in (L^2(\Omega), L^2(\Omega)) \) determined at


\[ t_{n-1} \] to find the next approximation \((u^n_0, Q^n_0, v^n_0, w^n_0)\). The procedure is initiated with \(u^0_0 = u_I, v^0_0 = v_L\). Specifically, we look for \((u^n_0, v^n_0, Q^n_0) \in (L^2(\Omega), L^2(\Omega), H(\text{div}; \Omega))\) satisfying the following time-discrete problem.

Problem \(P^\text{mef,n}_\delta\). Given \((u^{n-1}_0, v^{n-1}_0) \in (L^2(\Omega), L^2(\Omega))\), find \((u^n_0, Q^n_0, v^n_0, w^n_0) \in (L^2(\Omega), H(\text{div}; \Omega), L^2(\Omega), L^\infty(\Omega))\) such that

\[
\begin{align*}
(u^n_\delta - u^{n-1}_\delta, \phi) + \tau(\nabla \cdot Q^n_\delta, \phi) + (v^n_\delta - v^{n-1}_\delta, \phi) &= 0, \\
(v^n_\delta - v^{n-1}_\delta, \theta) - \tau(r(u^n_\delta), \theta) - \tau(H_\delta(v^n_\delta), \theta) &= 0, \\
(Q^n_\delta, \psi) - (u^n_\delta, \nabla \cdot \psi) - (q u^n_\delta, \psi) &= 0
\end{align*}
\]

for all \((\phi, \theta, \psi) \in (L^2(\Omega), L^2(\Omega), H(\text{div}; \Omega))\). For completeness we define \(w^n_\delta = H_\delta(v^n_\delta)\). This is a system of elliptic equations for \(u^n_\delta, Q^n_\delta, v^n_\delta\) given \(u^{n-1}_\delta, v^{n-1}_\delta \in H^1_0(\Omega), v^{n-1}_\delta \in L^2(\Omega)\). For stability reasons, we choose \(\delta = O(\tau^{\frac{1}{2}})\) (see [14, 25] for detailed arguments) which implies that \(\frac{\tau}{\delta^2}\) goes to zero as \(\tau \searrow 0\). This in turn allows us to consider the solutions along the sequence of regularized Heaviside function with the regularization parameter \(\delta\) automatically vanishing in the limit of \(\tau \searrow 0\).

The existence of a solution for Problem \(P^\text{mef,n}_\delta\) will result from the convergence of the fully discrete scheme, which is proved in the appendix by keeping \(\tau\) and \(\delta\) fixed, and passing to the limit \(h \searrow 0\). For now, we prove the uniqueness of the solution.

\textbf{Lemma 3.1.} Problem \(P^\text{mef,n}_\delta\) has at most one solution triple \((u^n_\delta, Q^n_\delta, v^n_\delta)\).

\textbf{Proof.} Since \(u^n_\delta = H_\delta(v^n_\delta)\), it has no influence on the existence or uniqueness of the solution. Therefore, we consider only the triples \((u^n_\delta, Q^n_\delta, v^n_\delta)\). Assume that for the same \((u^{n-1}_\delta, v^{n-1}_\delta)\) there are two solution triples \((u^n_{\delta,i}, Q^n_{\delta,i}, v^n_{\delta,i})\), \(i = 1, 2\) providing a solution to Problem \(P^\text{mef,n}_\delta\). Define

\[
\begin{align*}
&u^n_{\delta} := u^n_{\delta,1} - u^n_{\delta,2}, \\
&Q^n_{\delta} := Q^n_{\delta,1} - Q^n_{\delta,2}, \\
&v^n_{\delta} := v^n_{\delta,1} - v^n_{\delta,2}.
\end{align*}
\]

We now consider the equations for the differences above. Taking \(\theta = v^n_\delta\) in (3.2)_2 gives

\[
\|v^n_\delta\|^2 = \tau r(u^n_{\delta,1}) - r(u^n_{\delta,2}), \|v^n_\delta\| \leq \tau L\|u^n_\delta\| \|v^n_\delta\|
\]

as the \(H^1\) terms cancel because of explicit discretization. This gives, \(\|v^n_\delta\| \leq C\tau\|u^n_\delta\|\).

Further, with \(\phi = u^n_{\delta,1}, \theta = u^n_{\delta,2}, \psi = \tau Q^n_\delta\), from (3.2) we obtain

\[
\|u^n_{\delta,1}\|^2 + \tau\|Q^n_\delta\|^2 + \tau r(u^n_{\delta,1}) - r(u^n_{\delta,2}) = \tau(q u^n_{\delta,1}, Q^n_\delta).
\]

Since \(\tau\) is monotone, the Cauchy–Schwarz inequality and boundedness of \(q\) give

\[
\|u^n_{\delta}\|^2 + \frac{1}{\tau} \|Q^n_\delta\|^2 \leq \frac{1}{\tau} M_q^2 \|u^n_{\delta}\|^2.
\]

For \(\tau < \frac{2}{M_q^2}\), we obtain \(\|u^n_{\delta}\| = 0\) and thereby \(\|Q^n_\delta\| = 0\). Together with the above bound on \(\|v^n_{\delta}\|\), we conclude \(u^n_{\delta,1} = u^n_{\delta,2} = 0\) and \(Q^n_{\delta,1} = 0\).

\textbf{Lemma 3.2.} It holds that

\[
\begin{align*}
&\sup_{k=1,\ldots,N} \left( \|u^n_k\| + \frac{1}{\tau} \|v^n_k - v^{k-1}_k\| + \|v^n_k\| + \|Q^n_k\| \right) \leq C, \\
&\sum_{n=1}^{N} \left( \frac{1}{\tau} \|u^n_k - u^{n-1}_k\|^2 + \|Q^n_k - Q^{n-1}_k\|^2 + \tau \|\nabla \cdot Q^n_k\|^2 \right) \leq C.
\end{align*}
\]
Proof. We start by showing (3.3). To this aim we choose \( \phi = u^n_0, \ \psi = \tau Q^n_0, \ \theta = u^n_0 \) as test functions in (3.2), and add the resulting to obtain

\[
(u^n_0 - u^{n-1}_0, u^n_0) + \tau \|Q^n_0\|^2 - \tau (q u^n_0, Q^n_0) + \tau (r(u^n_0), u^n_0) = \tau (H_{\delta}(v^{n-1}_0), u^n_0).
\]

Since \( q \) and \( H_{\delta} \) are bounded and \( r(u^n_0)u^n_0 \geq 0 \), by Young’s inequality we get

\[
\|u^n_0\|^2 - \|u^{n-1}_0\|^2 + \|u^n_0 - u^{n-1}_0\|^2 + 2\tau \|Q^n_0\|^2 + 2\tau (r(u^n_0), u^n_0)
\]

\[
= 2\tau (q u^n_0, Q^n_0) + 2\tau (H_{\delta}(v^{n-1}_0), u^n_0) \leq \tau \|Q^n_0\|^2 + C\tau \|u^n_0\|^2 + C\tau + C\tau \|u^n_0\|^2.
\]

Summing over \( n = 1, \ldots, k \) (where \( k \in \{1, \ldots, N\} \) is arbitrary) gives

\[
\|u^n_0\|^2 + \sum_{n=1}^k \|u^n_0 - u^{n-1}_0\|^2 + \tau \sum_{n=1}^k \|Q^n_0\|^2 \leq \|u_I\|^2 + C + C\tau \sum_{n=1}^k \|u^n_0\|^2,
\]

and the first term in (3.3) follows from the discrete Gronwall lemma.

For the second term in (3.3), consider (3.2)\(_2\), choose \( \theta = v^n_0 - v^{n-1}_0 \), and apply the Cauchy–Schwarz inequality for the right-hand side,

\[
\|v^n_0 - v^{n-1}_0\|^2 \leq \tau \|r(u^n_0)\| \|v^n_0 - v^{n-1}_0\| + \tau \|H_{\delta}(v^{n-1}_0)\| \|v^n_0 - v^{n-1}_0\|.
\]

Using the previous bound, the boundedness of \( H_{\delta} \) and the Lipschitz continuity of \( r \) imply the conclusion. To prove the third term, choose \( \theta = v^n_0 \) in (3.2)\(_2\), rewrite the right-hand side, and using the monotonicity of \( H_{\delta} \) and the Cauchy–Schwarz inequality

\[
\|v^n_0\|^2 - \|v^{n-1}_0\|^2 + \|v^n_0 - v^{n-1}_0\|^2 \leq 2\tau C \|u^n_0\| \|v^n_0\| + 2\tau (H_{\delta}(v^{n-1}_0), v^n_0 - v^{n-1}_0).
\]

By Young’s inequality this leads to

\[
\|v^n_0\|^2 - \|v^{n-1}_0\|^2 + \frac{1}{2} \|v^n_0 - v^{n-1}_0\|^2 \leq \tau \|v^n_0\|^2 + C\tau \|u^n_0\|^2 + 2\tau^2 \|H_{\delta}\|^2.
\]

Summing over \( n = 1, \ldots, k \) (with \( k \in \{1, \ldots, N\} \) arbitrary), this gives

\[
\|v^n_0\|^2 + \sum_{n=1}^k \|v^n_0 - v^{n-1}_0\|^2 \leq \|v_I\|^2 + \tau \sum_{n=1}^k \|v^n_0\|^2 + C\tau \sum_{n=1}^k \|u^n_0\|^2 + \sum_{n=1}^k 4\tau^2 \|H_{\delta}\|^2
\]

\[
\leq \tau \sum_{n=1}^k \|v^n_0\|^2 + C + C\tau,
\]

where we have used the estimates proved before and the bounds on initial data. Now the inequality follows from the discrete Gronwall lemma. We proceed with the last term in (3.3). To this aim, we need to specify the initial flux \( Q^n_0 = -\nabla u_I + q u_I \in (L^2(\Omega))^d \). With \( \phi = u^n_0 - u^{n-1}_0 \), (3.2)\(_1\) gives

\[
(u^n_0 - u^{n-1}_0, u^n_0) + \tau (\nabla \cdot Q^n_0, u^n_0 - u^{n-1}_0) + (v^n_0 - v^{n-1}_0, u^n_0 - u^{n-1}_0) = 0.
\]

Now take \( \psi = \tau Q^n_0 \) to obtain

\[
\tau (Q^n_0 - Q^{n-1}_0, Q^n_0) - \tau (u^n_0 - u^{n-1}_0, \nabla \cdot Q^n_0) - \tau (q(u^n_0 - u^{n-1}_0), Q^n_0) = 0.
\]
Further, use (3.7) and rewrite the above left-hand side to obtain
\[
2 \|u^n_\delta - u^{n-1}_\delta\|^2 + \tau \|Q^n_\delta\|^2 - \tau \|Q^{n-1}_\delta\|^2 + \tau \|Q^n_\delta - Q^{n-1}_\delta\|^2 = 2\tau(q(u^n_\delta - u^{n-1}_\delta), Q^n_\delta) - 2(v^n_\delta - v^{n-1}_\delta, u^n_\delta - u^{n-1}_\delta).
\]

The right-hand side can be estimated using Young’s inequality to get
\[
(3.8) \quad \sum_{n=1}^k \|u^n_\delta - u^{n-1}_\delta\|^2 + \tau \|Q^n_\delta\|^2 + \tau \sum_{n=1}^k \|Q^n_\delta - Q^{n-1}_\delta\|^2 \\
\leq C\tau + C\tau^2 \sum_{n=1}^k \|Q^n_\delta\|^2 + \tau \|Q^n_\delta\|^2.
\]

The estimate follows now by the discrete Gronwall lemma. Moreover, from (3.8) we also get the first and second terms in (3.4). Finally, we take \(\phi = \nabla \cdot Q^n_\delta\) in (3.2) to use Young’s inequality, and previous estimates to bound the third term in (3.4).

**3.1. Enhanced compactness.** As will be seen below, the above estimates are not sufficient to retrieve the desired limiting equations. To complete the proof of convergence, stronger compactness properties are needed. These are obtained by translation estimates. To this aim, we define the translation in space
\[
\Delta_x f(\cdot) := f(\cdot + \xi), \quad \xi \in \mathbb{R}^2.
\]

With \(\xi \in \mathbb{R}^2\), we consider \(\Omega_\xi \subset \Omega\) such that \(\Omega_\xi := \{ x \in \Omega : \text{dist}(x, \Gamma) > \xi \}\). In this way, the translations \(\Delta_x f(x)\) with \(x \in \Omega\) are well defined.

For reasons of brevity, the norms and the inner products for the translations should be understood with respect to \(\Omega_\xi\) unless explicitly stated otherwise. Let us, first consider the translation for \(u^n_\delta\).

**Lemma 3.3.** It holds that \(\sum_{n=1}^N \tau \|\Delta_x u^n_\delta\|^2 \leq C|\xi|\).

**Proof.** For (3.2) we have after translation in space
\[
(\Delta_x Q^n_\delta, \psi) - (\Delta_x u^n_\delta, \nabla \cdot \psi) - (\Delta_x (Qu^n_\delta), \psi) = 0.
\]

We construct an appropriate test function to obtain the estimate above. Take \(\eta^n\) such that
\[
\begin{cases}
-\Delta \eta^n = \Delta_x u^n_\delta & \text{in } \Omega \\
\eta^n = 0 & \text{on } \Gamma
\end{cases}
\]

and choose \(\psi = \nabla \eta^n\) (note that \(\psi \in H(\text{div}; \Omega)\)) to obtain
\[
(3.9) \quad (\Delta_x Q^n_\delta, \nabla \eta^n) + (\Delta_x u^n_\delta, \Delta_x \eta^n) - (\Delta_x (Qu^n_\delta), \nabla \eta^n) = 0.
\]

Note that \(\eta^n\) satisfies \(\|\Delta \eta^n\| = \|\Delta_x u^n_\delta\|\), and therefore \(\|\eta^n\|_{H^2(\Omega)} \leq C(\Omega) \|\Delta_x u^n_\delta\|\).

This implies that translations of \(\nabla \eta^n\) are controlled,
\[
(3.10) \quad \|\nabla (\Delta_x \eta^n)\|_{L^2(\Omega)} \leq C|\xi| \|\nabla \cdot (\nabla \eta^n)\| \leq C(\Omega)|\xi| \|\Delta_x u^n_\delta\|.
\]

Recalling (3.3), this gives
\[
(3.11) \quad \|\nabla (\Delta_x \eta^n)\| \leq C|\xi|.
\]
Thus we have the following estimate

\begin{equation}
\tau \sum_{n=1}^{N} \|\Delta\xi v_n^\theta\|^2 = \tau \sum_{n=1}^{N} (q v_n^\theta, \nabla (\Delta\xi \eta^n)) + \tau \sum_{n=1}^{N} (Q_n^\theta, \nabla (\Delta\xi \eta^n))
\end{equation}

\begin{equation}
\leq \tau \sum_{n=1}^{N} \|Q_n^\theta\| \|\nabla (\Delta\xi \eta^n)\| + \tau \sum_{n=1}^{N} \|q\|_{L^\infty(\Omega)} \|u_n^\theta\| \|\nabla (\Delta\xi \eta^n)\|.
\end{equation}

The conclusion follows by (3.3), the Young inequality and (3.11). □

The estimate (3.13) follows from above and from (3.15), whereas (3.14) is a direct consequence of Lemma 3.3 and the assumptions on \(v_l\). □
3.2. Convergence. For proving the convergence of the time discretization scheme, we consider the sequence of time-discrete quadruples \( \{ (u^n_\delta, Q^n_\delta, v^n_\delta, w^n_\delta) \}_{n=0}^{N} \) solving Problem \( \mathbf{P}^{nef,n}_\delta \), and construct a time-continuous approximation by linear interpolation. In this sense, for \( t \in (t_{n-1}, t_n) \) \((n = 1, \ldots, N)\) we define

\[
Z^\tau(t) := z^n_\delta \frac{(t - t_{n-1})}{\tau} + z^{n-1}_\delta \frac{(t_n - t)}{\tau},
\]

where \( Z^\tau \) may refer to any of \( (U^\tau, Q^\tau, V^\tau) \) and \( z^n_\delta \) the discrete counterpart. Further for completeness, \( W^\tau(t) := H_\delta(V^\tau(t)) \). The estimates in Lemma 3.2 can be translated directly to \( (U^\tau, Q^\tau, V^\tau, W^\tau) \).

**Lemma 3.5.** A constant \( C > 0 \) exists s.t. for any \( \tau \) and \( \delta = O(\sqrt{\tau}) \) the following \( L^2(0,T; L^2(\Omega)) \) estimates hold:

\[
\begin{align}
\|U^\tau\| + \|V^\tau\| + \|Q^\tau\| &\leq C, \\
\|\partial_t U^\tau\| + \|\partial_t V^\tau\| + \|
abla \cdot Q^\tau\| &\leq C.
\end{align}
\]

**Proof.** Equation (3.18) follows easily from (3.3). For instance,

\[
\|U^\tau\|^2 \leq 2 \|u^n_\delta\|^2 + 2 \|u^{n-1}_\delta\|^2 \leq C
\]

and other estimates follow similarly. To estimate \( \|\partial_t V^\tau\| \) we note that, whenever \( t \in (t_{n-1}, t_n) \), \( \partial_t V^\tau = \frac{v^n_\delta - v^{n-1}_\delta}{\tau} \) implying

\[
\int_0^T \|\partial_t V^\tau\|^2 dt = \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \frac{1}{\tau} \|v^n_\delta - v^{n-1}_\delta\|^2 dt \leq \sum_{n=1}^N \frac{1}{\tau} \|v^n_\delta - v^{n-1}_\delta\|^2 \leq C \tau N \leq C,
\]

where we have used the estimate (3.3). The proof for \( \partial_t U^\tau \) is the same as above and uses the estimate (3.4). The only remaining part in (3.19) is to show that \( \nabla \cdot Q^\tau \in L^2(0,T; L^2(\Omega)) \). To see this note that

\[
\nabla \cdot Q^\tau = \nabla \cdot Q^{n-1}_\delta + \frac{t - t_{n-1}}{\tau} \nabla \cdot (Q^n_\delta - Q^{n-1}_\delta).
\]

Obtaining the bounds is now a simple exercise involving the a priori estimates already obtained for \( \nabla \cdot Q^n_\delta \) above in Lemma 3.2. We spare the details. \( \Box \)

Note that the estimates above are uniform in \( \tau \), if \( \delta = O(\sqrt{\tau}) \) and we have \( (U^\tau, Q^\tau, V^\tau, W^\tau) \in \mathcal{V} \times \mathcal{S} \times \mathcal{V} \times L^\infty(\Omega^T) \). Moreover, we have the following lemma.

**Lemma 3.6.** A quadruple \( (u, Q, v, w) \in \mathcal{V} \times \mathcal{S} \times \mathcal{V} \times L^\infty(\Omega^T) \) exists s.t. along a sequence \( \tau \searrow 0 \) (and with \( \delta = O(\sqrt{\tau}) \)) we have by compactness arguments

1. \( U^\tau \to u \) weakly in \( L^2((0,T); L^2(\Omega)) \),
2. \( \partial_t U^\tau \to \partial_t u \) weakly in \( L^2((0,T); L^2(\Omega)) \),
3. \( Q^\tau \to Q \) weakly in \( L^2((0,T); L^2(\Omega)^d) \),
4. \( \nabla \cdot Q^\tau \to \nabla \cdot Q \) weakly in \( L^2((0,T); L^2(\Omega)) \),
5. \( V^\tau \to v \) weakly in \( L^2((0,T); L^2(\Omega)) \),
6. \( \partial_t V^\tau \to \partial_t v \) weakly in \( L^2((0,T); L^2(\Omega)) \),
7. \( W^\tau \to w \) weakly-star in \( L^\infty(\Omega) \).

In the above only weak convergence of \( U^\tau \) in \( L^2(0,T; L^2(\Omega)) \) is obtained, which is not sufficient for passing to the limit for nonlinear term \( r(U^\tau) \). To obtain strong convergence, we use translation estimates as derived in Lemma 3.3.

**Lemma 3.7.** It holds that \( U^\tau \to u \) strongly in \( L^2((0,T); L^2(\Omega)) \).
Proof. The proof relies on applying the Riesz–Fréchet–Kolmogorov theorem (see, e.g., Theorem 4.26 in [8]). In view of \( \partial_t U^\tau \in L^2(0, T; L^2(\Omega)) \), the translation in time is already controlled. What we need is to control the translation in space:

\[
I_{\xi} := \int_0^T \int_{\Omega_{\xi}} |\Delta_\xi U^\tau|^2 \, dx \, dt \to 0 \text{ as } |\xi| \searrow 0.
\]

The definition of \( U^\tau \) immediately implies that \( |I_{\xi}| \leq \sum_{n=1}^N \tau (2\|\Delta_\xi u^\tau_n\|^2 + 2\|\Delta_\xi u^{n-1}_\delta\|^2) \).

Using Lemma 3.3 we find that \( |I_{\xi}| \leq C|\xi| \), where \( C \) is independent of \( \tau \) and \( \delta \), implying the strong convergence.

To identify \( w \) with \( H(v) \) we further need the strong convergence of \( V^\tau \).

**Lemma 3.8.** For \( V^\tau \), it holds that \( V^\tau \to v \) strongly in \( L^2((0, T); L^2(\Omega)) \).

**Proof.** The proof is similar to the above proof of Lemma 3.7 and uses the estimate in Lemma 3.4. The details are spared.

### 3.3. The Limit Equations

Once the strong convergence is obtained, the following theorem provides the existence of the weak solution in the mixed variational formulation.

**Theorem 3.9.** The limit quadruple \((u, Q, v, w)\) is a solution in the sense of Definition 2.1.

**Proof.** By the weak convergence, the estimates in Lemma 3.5 carry over for the limit quadruple \((u, Q, v, w)\). Moreover, the time-continuous approximation in (3.17) satisfies

\[
(\partial_t U^\tau, \phi) + (\nabla \cdot Q^\tau, \phi) + (\partial_t V^\tau, \phi) = (\nabla \cdot (Q^\tau - Q^\delta), \phi), \tag{3.20}
\]

\[
(\partial_t V^\tau, \theta) - (r(U^\tau) - W^\tau, \theta) = (H_\delta(V^\tau) - H_\delta(v^{n-1}_\delta), \theta) + (r(u^\delta_n) - r(U^\tau), \theta), \tag{3.21}
\]

\[
(Q^\tau, \psi) - (U^\tau, \nabla \cdot \psi) - (qU^\tau, \psi) = (Q^\tau - Q^\delta, \psi) - (U^\tau - u^\delta_n, \nabla \cdot \psi) - (q(U^\tau - u^\delta_n), \psi), \tag{3.22}
\]

for all \((\phi, \theta, \psi) \in (L^2(0, T; H^{1}_\delta(\Omega)), \mathcal{V}, \mathcal{S})\). Note that, in fact, (3.20) also holds for \( \phi \in \mathcal{V} \). Here we choose a better space to identify the limit, where we prove that the term on the right is vanishing along a sequence \( \tau \searrow 0 \). By density arguments, the limit will hold for \( \phi \in \mathcal{V} \).

Consider first (3.20) and note that by Lemma 3.6, the left-hand side converges to the desired limit. It only remains to show that the right-hand side, denoted by \( I_1 \), vanishes as \( \tau \searrow 0 \). Integrating by parts, which is allowed due to the choice of \( \phi \in L^2(0, T; H^{1}_\delta(\Omega)) \), one has

\[
|I_1| \leq \left( \sum_{n=1}^N \tau C\|Q^\delta_n - Q^{n-1}_\delta\|^2 \right)^{\frac{1}{2}} \left( \int_0^T \|\nabla \phi\|^2 \, dt \right)^{\frac{1}{2}} \to 0
\]

due to the estimate (3.4).

Next, we consider (3.21). First we prove that the last two integrals on the right-hand side, denoted by \( I_2 \) and \( I_3 \), vanish. For \( I_2 \) we use the Lipschitz continuity of \( H_\delta \) and the definition of \( V^\tau \) to obtain

\[
|I_2| \leq \left( \sum_{n=1}^N \frac{\tau}{\theta^2} \|v^\delta_n - v^{n-1}_\delta\|^2 \right)^{\frac{1}{2}} \left( \int_0^T \|\theta\|^2 \, dt \right)^{\frac{1}{2}}.
\]
Using (3.3) we have $|I_2| \leq C \frac{1}{\tau} \left( \int_0^T \|Q(t)\|^2 dt \right)^{\frac{1}{2}}$. By the choice of $\delta$, $\frac{3}{2} \gamma_1 > 0$ as $\tau \searrow 0$, implying that $I_2$ vanishes in the limit. For $I_3$ we use the Lipschitz continuity of $r$ and (3.4) to conclude the same.

For the first term on the left in (3.21), the limit is straightforward. For the limit of the second term, with strong convergence of $U^\tau$ and weak-* convergence of $W^\tau$ we get

$$
\lim_{\tau \searrow 0} \frac{r(U^\tau) - W^\tau, \theta}{r(u) - w, \theta},
$$

leading to the limiting equation

$$
(\partial_t v, \theta) = (r(u) - w, \theta) \text{ for all } \theta \in \mathcal{V}.
$$

Now we consider (3.22) and denote the corresponding integrals on the right-hand side, respectively, by $I_4, I_5$, and $I_6$. By the definition of $Q^\tau$ and (3.4), as $\tau \searrow 0$ we obtain

$$
|I_4| \leq \left( \sum_{n=1}^N \tau \|Q_n^\tau - Q_n^{\tau-1}\|^2 \right)^{\frac{1}{2}} \left( \int_0^T \|\psi\|^2 dt \right)^{\frac{1}{2}} \rightarrow 0.
$$

Similarly, $I_5$ and $I_6$ vanish in the limit using an a priori estimate for $u_n^\tau - u_n^{\tau-1}$ as in Lemma 3.2. With this the limit equation takes the form

$$
(Q, \psi) - (u, \nabla \cdot \psi) - (q u, \psi) = 0.
$$

To conclude the proof what remains is to show that $w = H(v)$. Since we have $V^\tau$ strongly converging, we also obtain $V^\tau \rightarrow v$ pointwise a.e. and further, as $\tau \searrow 0$, by construction $\delta \searrow 0$. For the set $R_+ := \{(t, x) : v(t, x) > 0\}$, let us assume $\mu := \{(t, x, z) : v(t, x) < 0\}$. Then the pointwise convergence implies the existence of an $\varepsilon_\mu > 0$ such that $V^\tau > \mu$ for all $\varepsilon \leq \varepsilon_\mu$. Then for any $\varepsilon \leq \varepsilon_\mu$ we have $W^\tau = 0$ implying $\varepsilon = 1$. A similar conclusion also holds for $R_-$, where $R_- := \{(t, x) : v(t, x) < 0\}$.

For the case when $v = 0$; consider the set $R_0 := \{(t, x, z) : v(t, x, z) = 0\}$. Now in the interior of the set $R_0$, $\partial_t v = 0$. Next, from the weak convergence of $\partial_t V^\tau, W^\tau, r(U^\tau)$, we have the following limit equation

$$
(\partial_t v, \theta) = (r(u) - w, \theta).
$$

Hence, for the interior of the set $R_0$, we obtain $w = r(u)$. Furthermore, the bounds $0 \leq W^\tau \leq 1$ with weak-* convergence of $W^\tau$ to $w$ imply the same bounds on $w$ and hence, $w = r(u)$ with $0 \leq r(u) \leq 1$.

4. The fully discrete formulation. Following the semidiscrete scheme, we now consider the fully discrete system (discretized in both space and time) and show its convergence. The steps for the proof of convergence are similar to the semidiscrete situation and wherever the proof is similar, we suppress the details. Further, to simplify notation, henceforth, we suppress the subscript $\delta$.

Starting with $u_h^0 = u_I, v_h^0 = v_I$, with $n \in \{1, \ldots, N\}$, the approximation $(u_h^n, v_h^n, Q^n_h, w_h^n)$ of $(u(t_n), v(t_n), Q(t_n), w(t_n))$ at $t = t_n$ solves the following problem.

Problem $P_h^n$. Given $(u_h^{n-1}, v_h^{n-1}) \in (\mathcal{V}_h, \mathcal{Q}_h)$ find $(u_h^n, v_h^n, Q_h^n, w_h^n) \in (\mathcal{V}_h, \mathcal{Q}_h, \mathcal{S}_h, L^\infty(\Omega))$ satisfying

$$
\begin{align*}
(u_h^n - u_h^{n-1}, \phi) &+ \tau (\nabla \cdot Q_h^n, \phi) - (v_h^n - v_h^{n-1}, \phi) = 0, \\
(v_h^n - v_h^{n-1}, \theta) &- \tau (r(u_h^n), \theta) - \tau (H_\delta(v_h^{n-1}), \theta) = 0, \\
(Q_h^n, \psi) - (u_h^n, \nabla \cdot \psi) - (q_h u_h^n, \psi) = 0
\end{align*}
$$

(4.1)
for all \((\phi, \theta, \psi) \in \mathcal{V}_h \times \mathcal{V}_h \times \mathcal{S}_h\). For completion, we define \(w_h^n = H_\delta(v_h^n)\). For stability reasons, as before, we choose \(\delta = O(\tau^{\frac{1}{2}})\) (see [14, 25] for detailed arguments).

From (4.1)_1 and (4.1)_2, we eliminate \(v_h^n\), which is computed after having obtained \((u_h^n, Q_h^n)\) satisfying for all \((\phi, \psi) \in (\mathcal{V}_h, \mathcal{S}_h)\)

\[
(u_h^n - u_h^{n-1}, \phi) + \tau(\nabla \cdot Q_h^n, \phi) + \tau(r(u_h^n) - H_\delta(u_h^{n-1}), \phi) = 0,
\]

\[
(Q_h^n, \psi) - (u_h^n, \nabla \cdot \psi) - (q_h u_h^n, \psi) = 0.
\]

In the above formulation, the nonlinearities only involve \(u_h^n\); the \(H_\delta\) is known from the previous time step and is in \(L^\infty\).

The existence follows from [37, Theorem 4.3], which treats a more general case. Its proof is based on [41, Lemma 1.4, p. 140]. Following the ideas in section 3, one can prove that (4.2) has a unique solution pair \((u_h^n, Q_h^n)\). This also determines \(v_h^n\) and \(w_h^n\) uniquely. We summarize the above result.

**Lemma 4.1.** Problem \(P_h^n\) has a unique solution quadruple \((u_h^n, Q_h^n, v_h^n, w_h^n)\).

Similar to the semidiscrete case, one can prove the following stability estimates.

**Lemma 4.2.** The following estimates hold:

\[
\sup_{k=1,\ldots,N} \left( \left\| \psi_h^k \right\| + \frac{1}{\tau} \left\| e_h^k - e_h^{k-1} \right\| + \left\| e_h^k \right\| + \left\| Q_h^k \right\| \right) \leq C,
\]

\[
\sum_{n=1}^N \left( \frac{1}{\tau} \left\| u_h^n - u_h^{n-1} \right\|^2 + \left\| Q_h^n - Q_h^{n-1} \right\|^2 \right) + \tau \left\| \nabla \cdot Q_h^n \right\|^2 + \tau \left\| \nabla \cdot (Q_h^n - Q_h^{n-1}) \right\|^2 \leq C.
\]

We continue with the steps analogous to the semidiscrete situation. As in (3.17) we consider the time-continuous approximation by the piecewise linear interpolations of the time-discrete solutions. Let \(Z_h^n\) denote the interpolated solution with \(Z\) standing for \(U, V\), or \(Q\) and \(W_h^n\) defined analogously. As before, the estimates in Lemma 4.2 carry over for the time-continuous approximation (the proof is omitted).

**Lemma 4.3.** The time-continuous approximations satisfy the following estimates:

\[
\| \partial_t U_h^n \|^2 + \| \nabla \cdot Q_h^n \|^2 + \| U_h^n \|^2 + \| V_h^n \|^2 + \| \partial_t V_h^n \|^2 + \| Q_h^n \|^2 \leq C,
\]

\[
0 \leq W_h^n \leq 1.
\]

Here the norms are taken with respect to \(L^2(0, T; L^2(\Omega))\). The estimates are uniform in \(\tau\) and \(\delta\) and furthermore we have \((U_h^n, Q_h^n, V_h^n, W_h^n) \in V \times S \times V \times L^\infty(\Omega)\).

Clearly, if \(\tau \searrow 0\) with \(\delta = O(\tau^{\frac{1}{2}})\) implies that both \(\delta, \frac{\tau}{\delta} \searrow 0\). By compactness arguments, we get from Lemma 4.3 the following convergence result.

**Lemma 4.4.** Along a sequence \(\tau \searrow 0\), it holds that

1. \(U_h^n \rightharpoonup u\) weakly in \(L^2((0, T); L^2(\Omega))\),
2. \(\partial_t U_h^n \rightharpoonup \partial_t u\) weakly in \(L^2((0, T); H^{-1}(\Omega))\),
3. \(Q_h^n \rightharpoonup Q\) weakly in \(L^2((0, T); L^2(\Omega)^d)\),
4. \(\nabla \cdot Q_h^n \rightharpoonup \chi\) weakly in \(L^2((0, T); L^2(\Omega))\),
5. \(V_h^n \rightharpoonup v\) weakly in \(L^2((0, T); L^2(\Omega))\),
6. \(\partial_t V_h^n \rightharpoonup \partial_t v\) weakly in \(L^2((0, T); L^2(\Omega))\),
7. \(W_h^n \rightharpoonup w\) weakly-star in \(L^\infty(\Omega)\).

As in the semidiscrete case, identification of the above limit \(\chi\) with \(\nabla \cdot Q\) is obtained via smooth test functions. Note that the above lemma only provides weak
convergence for \( U_h^\tau, V_h^\tau \); in the wake of nonlinearities, the strong convergence is needed. However, the techniques from the semidiscrete case cannot be applied directly. This is because the translation of a function that is piecewise constant on the given mesh need not be piecewise constant on that mesh. We therefore adopt the finite volume framework in [19] in order to overcome this difficulty.

### 4.1. Strong convergence

In what follows, we establish the required strong convergence of \( U_h^\tau \) followed by that of \( V_h^\tau \). We provide the notation used below in the framework of finite volumes. Let \( E \) denote the set of edges of the simplices \( T_h \). Also, we have that \( E = E_{\text{int}} \cup E_{\text{ext}} \) with \( E_{\text{ext}} = E \cap \partial \Omega \) and \( E_{\text{int}} = E \setminus E_{\text{ext}} \). We adopt the following notation:

\[
|T| = \text{the area of } T \in T_h, \quad x_i = \text{the center of the circumcircle of } T,
\]

\[
l_{ij} = \text{the edge between } T_i \text{ and } T_j, \quad \sigma_{ij} = \frac{|l_{ij}|}{d_{ij}}.
\]

In analogy with the spatially continuous case, we define the following discrete inner product for any \( u_h^n, v_h^n \in V_h \):

\[
(u_h^n, v_h^n)_h := \sum_{T_i \in T_h} |T_i| u_h^n(i) v_h^n(i), \quad (u_h^n, v_h^n)_{1,h} := \sum_{l_{ij} \in E} |\sigma_{ij}|(u_h^n(i) - u_h^n(j))(v_h^n(i) - v_h^n(j)).
\]

The discrete inner product gives rise to the discrete \( H^1_0 \) norm, which is

\[
\| u_h^n \|^2_{1,h} = \sum_{l_{ij} \in E} |\sigma_{ij}|(u_h^n(i) - u_h^n(j))^2.
\]

In [19], the following discrete Poincaré inequality is proved: \( \| u_h^n \| \leq C\| u_h^n \|_{1,h} \), with \( C \) independent of \( h \) or \( u_h^n \). Based on Lemma 4 in [19], below we show that the translations are controlled by the discrete \( \| \cdot \|_{1,h} \) norm.

**Lemma 4.5.** Let \( \Omega \) be an open bounded set of \( \mathbb{R}^2 \) and let \( T_h \) be an admissible mesh. For a given \( u \) defined in \( \Omega \) and extended to 0 outside \( \Omega \) we have

\[
\| \nabla \xi u \|^2_{L^2(\mathbb{R}^2)} \leq \| u \|^2_{1,h} |\xi|(|\xi| + C \text{size}(T_h)) \text{ for all } \xi \in \mathbb{R}^2.
\]

This shows that for a sequence \( \{ u_h^n \} \) having the discrete \( H^1_0 \) norm uniformly bounded, the \( L^2 \)-norm of the translations \( \nabla \xi u_h^n \) vanishes uniformly with respect to \( h \) as \( \eta \searrow 0 \). This is an essential step in proving the strong \( L^2 \)-convergence for \( u_h^n \). Here we only need to show that \( u_h^n \) has bounded discrete \( H^1_0 \) norm.

**Lemma 4.6.** For the sequence \( u_h^n \), the following inequality holds with \( C \) independent of \( h \) and \( n \):

\[
\| u_h^n \|_{1,h} \leq C(\| Q_h^n \|_{1,h} + \| u_h^n \|_{1,h}).
\]

**Proof.** The approach is inspired from the semidiscrete situation and is adapted to the present context by defining an appropriate test function. Define

\[
|T_i| f_h^n(T_i) := \sum_{l_{ij} \in E} |\sigma_{ij}|(u_h^n(i) - u_h^n(j))
\]

and note that by the definition of \( \| \cdot \|^2_{1,h} \),

\[
(f_h^n, u_h^n) = \sum_i |T_i| f_h^n(T_i) u_h^n(T_i) = \sum_{l_{ij} \in E} |\sigma_{ij}|(u_h^n(i) - u_h^n(j))^2 = \| u_h^n \|^2_{1,h}.
\]
Further, by using the Cauchy–Schwarz we obtain
\[ \| f_h^n \|_{L^2(\Omega)}^2 = \sum_i |T_i| |f_h^n(T_i)|^2 \leq \sum_{i,j} |l_{ij}| \frac{(u^n_i - u^n_j)^2}{|T_i|} \frac{1}{|T_i|} \sum_{i,j} |l_{ij}| \frac{|l_{ij}|}{d_{ij}} \]
which implies that \( \| f_h^n \| \leq \| u^n_h \|_{1,h} \). Note that \( f_h^n \in L^2(\Omega) \) and hence, there exists \( \psi_h \in S_h \) which satisfies
\begin{align}
\nabla \cdot \psi_h &= f_h^n \quad \text{in } \Omega, \\
\psi_h &= 0 \quad \text{on } \Gamma.
\end{align}
By the bounds on \( f_h^n \) above, it also holds that \( \| \psi_h \|_{L^2(\Omega)} \leq C \| f_h^n \|_{L^2(\Omega)} \leq C \| u^n_h \|_{1,h} \).
By (4.13), \( (u^n_h, \nabla \cdot \psi_h) = (u^n_h, f^n_h) = \| u^n_h \|^2 \). Now choose for the test function \( \psi = \psi_h \) in (4.13) to obtain
\[ \| u^n_h \|^2_{1,h} = (u^n_h, \nabla \cdot \psi_h) = (Q^n_h, \psi_h) - (q_h u^n_h, \psi_h) \leq \| Q^n_h \| \| \psi_h \| + M_q \| u^n_h \| \| \psi_h \| \leq C \| Q^n_h \| \| u^n_h \|_{1,h} + C M_q \| u^n_h \| \| u^n_h \|_{1,h} \]
and the conclusion follows. \( \square \)

In view of the above lemma, obtaining the relative compactness in \( L^2 \) is straightforward.

**Lemma 4.7.** Along a sequence \((\tau, h)\) converging to \((0, 0)\) (and with \( \delta = O(\sqrt{T}) \)), \( U^\tau_h \) converges strongly in \( L^2(0, T; L^2(\Omega)) \).

**Proof.** Since \( \partial_t U^\tau_h \) is in \( L^2 \), the translation with respect to time is already controlled. What remains is to consider the translation with respect to space. Take (4.11) and sum over \( n = 1, \ldots, N \) to obtain
\[ \tau \sum_{n=1}^N \| u^n_h \|^2_{1,h} \leq C \tau \sum_{n=1}^N (\| Q^n_h \|^2 + \| u^n_h \|^2) \leq C \]
using (4.3). Now use Lemma 4.5 to control the translations by the \( \| \cdot \|_{1,h} \) norm (after extending \( u^n_h \) by 0 outside \( \Omega \); for simplicity retain the same notation):
\[ \tau \sum_{n=1}^N \| u^n_h(\cdot + \xi) - u^n_h \|^2_{L^2(\mathbb{R}^2)} \leq C \| \xi \| \| \text{size}(T_h). \]
which, in turn, provides a similar estimate for \( U^\tau_h \)
\[ \tau \sum_{n=1}^N \| U^\tau_h(\cdot + \xi) - U^\tau_h \|^2_{L^2(\mathbb{R}^2)} \leq C \| \xi \| \| \text{size}(T_h). \]
The Riesz–Fréchet–Kolmogorov compactness theorem proves the assertion. \( \square \)

The strong convergence of \( U^\tau_h \) leads to the strong convergence of \( V^\tau_h \).

**Lemma 4.8.** Along a sequence \((\tau, h)\) converging to \((0, 0)\), \( V^\tau_h \) converges strongly to \( v \) in \( L^2(0, T; L^2(\Omega)) \).

**Proof.** As before, the translations with respect to time are already controlled by virtue of \( \partial_t V^\tau_h \in L^2 \). We now consider the case for the translation with respect to space. Since both \( u^n_h, v^n_h \) are piecewise constants in each simplex \( T \), we have for every \( x \in T \)
\[ v^n_h(x) = v^{n-1}_h(x) + \tau (r(u^n_h(x)) - \tau H_\delta(v^{n-1}_h(x))) , \]
\[ v^n_h(x + \xi) = v^{n-1}_h(x + \xi) + \tau (r(u^n_h(x + \xi)) - \tau H_\delta(v^{n-1}_h(x + \xi))) , \]

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
so that for any \( x \in \Omega \), we have \( \Delta_\xi (v_h^n - v_h^{n-1}) = \tau \Delta_\xi r(u_h^n) - \tau \Delta_\xi H_h(v_h^{n-1}) \). Multiplying by \( \Delta_\xi v_h^n \) and rewriting the left-hand side, we have

\[
(4.17) \quad \frac{1}{2} \{ |\Delta_\xi v_h^n|^2 - |\Delta_\xi v_h^{n-1}|^2 + |\Delta_\xi (v_h^n - v_h^{n-1})|^2 \} \leq \tau L_r |\Delta_\xi u_h^n||\Delta_\xi v_h^n| - \tau (\Delta_\xi H_h(v_h^{n-1})) \Delta_\xi v_h^n.
\]

Rewriting the last term to exploit monotonicity of \( H_h \),

\[
\frac{1}{2} \{ |\Delta_\xi v_h^n|^2 - |\Delta_\xi v_h^{n-1}|^2 + |\Delta_\xi (v_h^n - v_h^{n-1})|^2 \} \leq \tau L_r^2 |\Delta_\xi u_h^n|^2 + \frac{1}{4} |\Delta_\xi v_h^n|^2 + \frac{1}{4} |\Delta_\xi (v_h^n - v_h^{n-1})|^2 + \frac{\tau^2}{\delta^2} |\Delta_\xi v_h^{n-1}|^2.
\]

Integrating over \( \Omega \) and summing over \( n = 1, \ldots, k \) for any \( k \in \{1, \ldots, N\} \) gives

\[
\frac{1}{2} \|\Delta_\xi v_h^k\|^2 + \frac{1}{4} \sum_{n=1}^k \|\Delta_\xi (v_h^n - v_h^{n-1})\|^2 \\
\leq \|\Delta_\xi v_{\ell,k}\|^2 + \tau \sum_{n=1}^k L_r^2 \|\Delta_\xi u_h^n\|^2 + \frac{1}{4} \tau \sum_{n=1}^k \|\Delta_\xi v_h^n\|^2 + \frac{\tau^2}{\delta^2} \sum_{n=1}^k \|\Delta_\xi v_h^{n-1}\|^2,
\]

where the norms are taken with respect to \( \Omega \). Choosing \( \delta = O(\tau^{\frac{1}{2}}) \) leads to

\[
\frac{1}{2} \|\Delta_\xi v_h^k\|^2 + \frac{1}{4} \sum_{n=1}^k \|\Delta_\xi (v_h^n - v_h^{n-1})\|^2 \\
\leq \|\Delta_\xi v_{\ell,h}\|^2 + \tau \sum_{n=1}^k L_r^2 \|\Delta_\xi u_h^n\|^2 + C \tau \sum_{n=1}^k \|\Delta_\xi v_h^n\|^2.
\]

Applying Gronwall’s lemma, \( \sup_{k=1, \ldots, N} \|\Delta_\xi v_h^k\|^2 \leq C \|\Delta_\xi v_{\ell,h}\|^2 + \tau \sum_{n=1}^N \|\Delta_\xi u_h^n\|^2 \).

The strong convergence of \( U_h^1 \) in \( L^2(0, T; L^2(\Omega)) \) implies that the last term vanishes in the limit of \( |\xi| \searrow 0 \) (see the proof of Lemma 4.7). To estimate the translations for the initial condition we consider \( v_{\ell,h} \) as the finite volume approximation of \( v_\ell \) defined (formally) by \( -\Delta v_{\ell,h} = -\Delta v_\ell \) in \( \Omega \), with homogenous Dirichlet boundary conditions. This implies \( \|v\|_{1,h} \leq C \|\nabla v_\ell\| \leq C \). Since the translations are approaching 0 if \( \|v_{\ell,h}\|_{1,h} \leq C \) (uniformly in \( h \)), \( \|\Delta_\xi v_{\ell,h}\| \to 0 \) as \( |\xi| \searrow 0 \). From the above we conclude that \( \|\Delta_\xi v_h^n\|^2 \to 0 \) as \( |\xi| \searrow 0 \). Finally, note that the definition of \( V_h^\tau \) implies the rough estimate

\[
\int_0^T \|\Delta_\xi V_h^\tau\|^2 dt \leq 2\tau \sum_{n=1}^N \|\Delta_\xi v_h^n\|^2 + 2\tau \sum_{n=1}^N \|\Delta_\xi v_h^{n-1}\|^2,
\]

and the right-hand side vanishes uniformly in \( h \) as \( |\xi| \searrow 0 \); hence \( V_h^\tau \) converges strongly.

\[\square\]

### 4.2. The limit equations.

Up to now we obtained the convergence of the fully discrete triples \((U_h^\tau, V_h^\tau, Q_h^\tau)\) along a sequence \((\tau, h)\) approaching \((0, 0)\) with \( \delta = O(\sqrt{\tau}) \). Clearly, the \((L^\infty\) weakly-star\) convergence extends to the sequence \( W_h^\tau = H_h(V_h^\tau) \). In what follows, we identify the limit discussed in the preceding section as the weak formulation (2.9).
Theorem 4.9. The limit quadruple \((u, Q, v, w)\) is a weak solution in the sense of Definition 2.1.

Proof. By the weak convergence, the estimates in Lemma 4.3 carry over for the limit triple \((u, Q, v)\). By (4.1) we have

\[
\int_0^T (\partial_t U_h^T, \phi) dt + \int_0^T (\nabla \cdot Q_h^T, \phi) dt + \int_0^T (\partial_t V_h^T, \phi) dt
\]

\[
= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\partial_t U_h^T, \phi - \phi_h) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\nabla \cdot Q_h^T - \nabla \cdot Q_h^T, \phi) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\nabla \cdot Q_h^T - \nabla \cdot Q_h^T, \phi_h - \phi) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\partial_t V_h^T, \phi - \phi_h) dt
\]

for all \(\phi \in L^2(0, T; H^1_0(\Omega))\), and where \(\phi_h\) is the projection \(\phi_h = P\phi\) introduced in section 2.1. Note that we assume again an \(H^1\) regularity in space for the test function \(\phi\). We use this to control the terms involving \(\|\phi - \phi_h\|\) by using the property (2.7). A usual density argument lets the result hold for all \(\phi \in \mathcal{V}\).

The left-hand side gives the desired limit terms; it only remains to show that the right-hand side vanishes in the limit. Denote the successive integrals on the right by \(I_i, i = 1, \ldots, 5\). We deal with each term separately.

For \(I_1\) we use (4.5) to obtain that as \(h \downarrow 0\)

\[
|I_1| \leq \|\partial_t U_h^T\|_{L^2(0, T; L^2(\Omega))} \left( \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\phi - \phi_h\|_{L^2(\Omega)}^2 dt \right)^{1/2} \leq C h \|\nabla \phi\|_{L^2(0, T; L^2(\Omega))} \to 0.
\]

Similarly, by (4.4), for \(I_2\) one gets

\[
|I_2| \leq \left( \sum_{n=1}^N \tau \|Q_h^n - Q_h^{n-1}\|^2 \right)^{1/2} \left( \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\nabla \phi\|_{L^2(\Omega)}^2 dt \right)^{1/2} \leq C \tau^{1/2}.
\]

Clearly \(I_2\) vanishes in the limit of \(\tau \downarrow 0\). The estimates for \(I_3\) are analogous and use the estimate (4.5). The treatment of \(I_4\) and \(I_5\) is similar and relies on (4.4) and estimate for \(\partial_t V_h^T\), respectively.

Next we consider (4.1)\(_2\), which we rewrite as

\[
\int_0^T (\partial_t V_h^T, \theta) dt - \int_0^T (r(U_h^T) - W_h^T, \theta) dt
\]

\[
= \int_0^T (\partial_t V_h^T, \theta - \theta_h) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (H_\delta(V_h^n) - H_\delta(v_h^{n-1}), \theta) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (H_\delta(v_h^{n-1}) - \theta, \theta_h) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (r(u_h^n) - r(U_h^n), \theta) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (r(u_h^n), \theta_h - \theta) dt
\]
for \( \theta \in L^2(0,T;H^1_0(\Omega)) \) and \( \theta_h \) is the \( P_h \) projection of \( \theta \). A better regularity of \( \theta \) is again chosen for identifying the limits and controlling the errors due to the projections. We would retrieve the desired limiting equations once we prove that the integrals on the right-hand side vanish. Let us denote the successive integrals by \( J_i \), \( i = 1, \ldots, 5 \), respectively. For \( J_1 \) we get, by using (4.5) and recalling the projection estimate (2.7),

\[
|J_1| \leq \left( \int_0^T \| \partial_t V_h^2 \| dt \right)^{1/2} \left( \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \| (\theta - \theta_h) \|^2 dt \right)^{1/2} \leq Ch\|\theta\|_{L^2(0,T;H^1_0(\Omega))}
\]

which vanishes in the limit as \( h \searrow 0 \). For \( J_2 \), we use the definition of \( W^\tau \) and Lipschitz continuity of \( H_\delta \) to obtain

\[
|J_2| \leq \sum_{n=1}^N \frac{1}{\tau} \| v_h^n - v_h^{n-1} \| \| \theta \| \leq \sum_{n=1}^N \frac{\tau C_{\delta}}{\delta} \| \theta \|
\]

by using (4.3); and further, using \( \tau/\delta \searrow 0 \) by the construction of \( \delta \) we obtain \( J_2 \to 0 \). Next, we consider \( J_3 \):

\[
|J_3| \leq C \left( \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \| \theta - \theta_h \|^2 dt \right)^{1/2} \leq Ch\|\nabla \theta\|_{L^2(0,T;L^2(\Omega))} \to 0 \text{ as } h \searrow 0
\]

because of (2.7). To continue,

\[
|J_4| \leq \left( \sum_{n=1}^N \tau L_r^2 \| u_h^n - u_h^{n-1} \|^2 \right)^{1/2} \left( \sum_{n=1}^N \tau \| \theta_h \|^2 \right)^{1/2} \to 0,
\]

\[
|J_5| \leq L_r \left( \sum_{n=1}^N \tau \| u_h^n \|^2 dt \right)^{1/2} \left( \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \| (\theta - \theta_h) \|^2 dt \right)^{1/2} \leq Ch\|\theta\|_{L^2(0,T;H^1_0(\Omega))} \to 0,
\]

where we use the estimate (4.4) for \( J_4 \).

Let us consider the next equation, that is, (4.1)_3. We have, by realigning the terms,

\[
\int_0^T (Q_h^n, \psi) dt - \int_0^T (U_h^n, \nabla \cdot \psi) dt - \int_0^T (q_h U_h^n, \psi) dt
\]

\[
= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (Q_h^n - Q_h^n, \psi) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (Q_h^n, \psi - \psi_h) dt
\]

\[
+ \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (u_h^n - U_h^n, \nabla \cdot \psi) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (U_h^n, \nabla \cdot (\psi - \psi_h)) dt
\]

\[
+ \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (q_h (u_h^n - U_h^n), \psi) dt + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (q_h u_h^n, \psi - \psi_h) dt
\]

for all \( \psi \in L^2(0,T;H^2(\Omega)) \) and where \( \psi_h \) is chosen as the \( P_h \) projection of \( \psi \). As before the left-hand side converges to the desired limits. This is obvious except for the third term where we use the \( L^\infty \) and strong convergence of \( q_h \). Indeed,

\[
\int_0^T (q_h U_h^n, \psi) dt = \int_0^T (q U_h^n, \psi) dt + \int_0^T ((q_h - q) U_h^n, \psi) dt
\]
and the first term on the right-hand side passes to the desired limit. We show that the second term vanishes in the limit. Note that \( q_h - q \in (C^\infty(\Omega))^2 \) and hence, \((q_h - q)U_h^T\) has a weak limit. Now choose \( \psi \in L^2(0,T; (C^\infty(\Omega))^2) \) so that \( \psi \in (L^\infty(\Omega))^2 \). Now

\[
\|(q_h - q)U_h^T\|_{(L^1(\Omega))^2} \leq \|q_h - q\|_{(L^2(\Omega))^2} \|U_h^T\|_{(L^2(\Omega))}
\]

and use the strong convergence of \( q_h \) in \( L^2 \) to conclude that the weak limit is indeed 0.

Now we show that the right-hand side of (4.19) vanishes in the limit. Let us denote the integrals by \( K_i, i = 1, \ldots, 6 \). The successive terms will be treated as before. We begin with \( K_1 \):

\[
|K_1| \leq \|\psi\|_{L^2(0,T;H^1(\Omega))} \left( \sum_{n=1}^N \tau \|Q_h^n - Q_h^{n-1}\|^2 \right)^{1/2} \leq C\tau^{1/2}
\]

using bounds given in (4.4). Thus, \( K_1 \) goes to 0 in the limit. For \( K_2 \), recalling the bound (4.3) and the projection estimate (2.7),

\[
|K_2| \leq \left( \sum_{n=1}^N \tau\|Q_h^n\|^2 \right)^{1/2} \left( \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\psi - \psi_h\|_{L^2(\Omega)} dt \right)^{1/2} \leq Ch.
\]

Similarly, we have

\[
|K_3| \leq \left( \sum_{n=1}^N \|u_h^n - u_h^{n-1}\|^2 \right)^{1/2} \left( \int_0^T \|\nabla \cdot \psi\|^2 dt \right)^{1/2} \leq C\tau,
\]

\[
|K_4| \leq \left( \sum_{n=1}^N \|u_h^n\|^2 \right)^{1/2} \left( \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\nabla \cdot (\psi - \psi_h)\|^2 dt \right)^{1/2} \leq Ch,
\]

\[
|K_5| \leq M_q \left( \sum_{n=1}^N \tau\|u_h^n - u_h^{n-1}\|^2 \right)^{1/2} \|\psi\|_{L^2(0,T;L^2(\Omega))} \leq C\tau,
\]

\[
|K_6| \leq M_q \left( \sum_{n=1}^N \tau\|u_h^n\|^2 \right)^{1/2} \left( \sum_{n=1}^N \tau\|\psi - \psi_h\|^2 \right)^{1/2} \leq Ch,
\]

all of which vanish in the limit.

The identification of \( w \) with \( H(v) \) is identical to the semidiscrete case. Note that the limit quadruple \((u, Q, v, w)\) indeed satisfies (2.9), but for test functions having a better regularity in space, \( \phi \in L^2(0,T;H^3(\Omega)) \), \( \theta \in L^2(0,T;H^1(\Omega)) \), and \( \psi \in L^2(0,T;H^2(\Omega)) \). In view of the regularity of \( u, v, Q \), density arguments can be employed to show that the limit equations also hold for \( \phi \in L^2(0,T;L^2(\Omega)), \theta \in L^2(0,T;L^2(\Omega)) \), \( \psi \in L^2(0,T;H(div;\Omega)) \), which completes the proof.

5. **Numerical computations.** We will study the numerical computations in two parts: the first part deals with the illustration of the physical characteristics of the model and the second, the convergence studies. For the former, we will consider the reactive processes taking place depending upon the appropriate choice of boundary and initial conditions. For the convergence studies, we will consider a test problem for which we have constructed an exact solution. We begin with the illustration of the physical properties of the model.
5.1. Illustrative computations. Let us take \( \Omega := (0,1) \times (0,1) \) and we make the following choices:

\[
T = 1, \quad q = (0.5, 0), \quad D = 1, \quad r(u) = 2u_+ (u - 0.5)_+; \quad h = 0.0125, \quad \tau = 0.005.
\]

For the boundary condition, \(-\nu \cdot \nabla u = 0\) on \( \Gamma \setminus \{x = 0\}\), with the boundary condition at \( x = 0 \) specified in the different cases considered below. Note that for this choice of precipitation rate, \( r(u) = 0 \) for \( u \leq 0.5 \) and \( r(1) = 1 \) so that \( u = 1 \) is the equilibrium solution, that is, for which the dissolution rates balance the precipitation rate. To specify the initial condition, let us consider \( \Omega_v := (0.2, 0.8) \times (0.2, 0.8) \) with clearly \( \Omega_v (\subset \Omega) \) being a smaller square inside the original domain. We study the following cases:

**Case (a). Equilibrium situation.** For the initial conditions, we choose \( u_I = 1, \quad v_I = 0.1 \chi_{\Omega_v} \) with \( \chi_{\Omega_v} \) denoting the characteristic function for the set \( \Omega_v \). We impose \( u = 1 \) at \( x = 0 \) as the boundary condition. With \( u = 1 \) being the equilibrium solution, the initial and the boundary conditions ensure that no changes take place in the solution as the initial conditions satisfy the equations (2.1)–(2.2). This is easily confirmed numerically by computing

\[
\|u_I - u_h^n\| = 1.04e - 11, \quad \|v_I - v_h^n\| = 7.98e - 7
\]

in the \( L^2(0,T;L^2(\Omega)) \) norm.

**Case (b). Dissolution fronts.** To initiate the dissolution front, we choose the boundary condition at \( x = 0 \); hence we choose

\[
u_I = 1, \quad v_I = 0.1 \chi_{\Omega_v}, \quad u = 0 \text{ at } x = 0.
\]

Clearly, at the boundary one has \( r(u) - H(v) < 0 \), and this is propagated inside the entire domain, initiating a dissolution process. The numerical results in Figure 5.1 show a depletion of the precipitate and the occurrence of dissolution, with the support of \( v \) shrinking as time proceeds.

**Fig. 5.1.** Concentration profiles for the solute \( u \) (top) and the precipitate \( v \) (bottom) at different times: \( t = 0.05, 0.1, 0.2 \). Note that initially \( u = 1; \quad v = 0.2 \) in \( \Omega_v \) and 0 elsewhere. As \( t \) increases, the dissolution front moves rightwards.
5.2. Convergence studies. We consider a test problem similar to (2.1)–(2.2), but including a right-hand side in the first equation (see [26] where we first announced part of these results). This is chosen in such a way that the problem has an exact solution, which is used then to test the convergence of the mixed finite element scheme. Specifically, for $T = 1$ and $\Omega = (0, 5) \times (0, 1)$, and with $r(u) = [u]^2_+$ (where $[u]_+ := \max\{0, u\}$), we consider the problem

$$
\begin{align*}
\partial_t(u + v) + \nabla \cdot (q u - \nabla u) &= f & \text{in } \Omega^T, \\
\partial_t v &= (r(u) - w) & \text{on } \Omega^T, \\
w &\in H(v) & \text{on } \Omega^T.
\end{align*}
$$

Here $q = (1, 0)$ is a constant velocity, whereas

$$
f(t, x, y) = \frac{1}{2} e^{x-t-5} \left(1 - e^{x-t-5}\right) - \frac{1}{2} \left(1 - \frac{1}{2} e^{x-t-5}\right) - \begin{cases} 0 & \text{if } x < t, \\
\frac{e^{x-t-5} - 1}{e^5 - 1} & \text{if } x \geq t,
\end{cases}
$$

and the boundary and initial conditions are such that

$$
u(t, x, y) = \left(1 - e^{x-t-5}\right)^{\frac{1}{2}}, \quad \text{and} \quad v(t, x, y) = \begin{cases} 0 & \text{if } x < t, \\
\frac{e^{x-t-5} - 1}{e^5 - 1} & \text{if } x \geq t,
\end{cases}
$$

providing $w(t, x, y) = \begin{cases} 1 & \text{if } x < t, \\
1 - e^{x-t-5} & \text{if } x \geq t
\end{cases}$ form a solution triple.

We consider the mixed finite element discretization of the problem above, based on the time stepping in section 4 and the lowest order Raviart–Thomas elements $RT_0$. The numerical scheme was implemented in the software package ug [6]. The simulations are carried out for a constant mesh diameter $h$ and time step $\tau$, satisfying $\tau = h$. We start with $h = 0.2$, and refine the mesh (and, correspondingly, $\tau$ and $\delta$) four times successively by halving $h$ up to $h = 0.0125$. We compute the errors for $u$ and $v$ in the $L^2$ norms,

$$
E^h_u = \|u - U^\tau\|_{L^2(\Omega^T)}, \quad \text{respectively,} \quad E^h_v = \|v - V^\tau\|_{L^2(\Omega^T)}.
$$

These are presented in Tables 5.1 and 5.2 for two different choices of $\delta$. Although theoretically no error estimates could be given due to the particular character of the dissolution rate, the tabulated results also include an estimate of the convergence order, based on the reduction factor between two successive calculations:

$$
\alpha = \log_2 (E^h_u / E^{\frac{h}{2}}_u) \quad \text{and} \quad \beta = \log_2 (E^h_v / E^{\frac{h}{2}}_v).
$$

**Table 5.1**

<table>
<thead>
<tr>
<th>$h$</th>
<th>$|u - U^\tau|^2$</th>
<th>$\alpha$</th>
<th>$|v - V^\tau|^2$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>1.1700e-01</td>
<td>1.8409e-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>6.414e-02</td>
<td>0.43</td>
<td>9.927e-02</td>
<td>0.45</td>
</tr>
<tr>
<td>0.05</td>
<td>3.396e-02</td>
<td>0.46</td>
<td>5.317e-02</td>
<td>0.45</td>
</tr>
<tr>
<td>0.025</td>
<td>1.726e-02</td>
<td>0.49</td>
<td>2.785e-02</td>
<td>0.47</td>
</tr>
<tr>
<td>0.0125</td>
<td>8.42e-03</td>
<td>0.52</td>
<td>1.420e-02</td>
<td>0.49</td>
</tr>
</tbody>
</table>
The tests are carried out with two choices of $\delta$: $\delta = \sqrt{\tau}$ (which is supported by the theory) and $\delta = 5\tau$ (still providing stability, but without having any rigorous convergence proof). As resulting from Tables 5.1 and 5.2, for this test case it appears that the method converges sublinearly, respectively, linearly. This suggests that practically $\delta = O(\tau)$ (by maintaining, however, the stability of the scheme) is better.

Similar results are observed for an implicit discretization for $v$ (although this scheme is not analyzed here; this can be obtained with minor modifications of the proofs here). The implicit scheme provides a set of coupled nonlinear equation for the triple $(u^n h, Q^n h, v^n h)$. A Newton iteration is used to solve the resulting system (see [36, 38], where the Newton method is applied to similar problems). The tests are applied to the case described before, and the results are presented in Tables 5.3 and 5.4. As in the semis implicit scheme, we see that for the test problem the convergence rate is sublinear for the case $\delta = \sqrt{\tau}$ and linear for $\delta = 5\tau$.

### Table 5.2

Convergence results for the mixed scheme, with the explicit discretization for $v$; $h = \tau$ and $\delta = 5\tau$.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$|u - U^\tau|^2$</th>
<th>$\alpha$</th>
<th>$|v - V^\tau|^2$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>2.774958e-01</td>
<td>4.300839e-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>1.263555e-01</td>
<td>0.57</td>
<td>1.881785e-01</td>
<td>0.59</td>
</tr>
<tr>
<td>0.05</td>
<td>2.353196e-02</td>
<td>1.21</td>
<td>8.905975e-02</td>
<td>0.54</td>
</tr>
<tr>
<td>0.025</td>
<td>1.072494e-02</td>
<td>0.57</td>
<td>1.792503e-02</td>
<td>1.16</td>
</tr>
<tr>
<td>0.0125</td>
<td>2.308967e-03</td>
<td>1.11</td>
<td>4.368191e-03</td>
<td>1.02</td>
</tr>
</tbody>
</table>

### Table 5.3

Convergence results for the fully implicit mixed discretization; here $h = \tau$ and $\delta = \sqrt{\tau}$.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$|u - U^\tau|^2$</th>
<th>$\alpha$</th>
<th>$|v - V^\tau|^2$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>1.031e-01</td>
<td>1.593e-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>5.925e-02</td>
<td>0.40</td>
<td>9.023e-02</td>
<td>0.41</td>
</tr>
<tr>
<td>0.05</td>
<td>3.247e-02</td>
<td>0.43</td>
<td>5.031e-02</td>
<td>0.42</td>
</tr>
<tr>
<td>0.025</td>
<td>1.686e-02</td>
<td>0.47</td>
<td>2.703e-02</td>
<td>0.45</td>
</tr>
<tr>
<td>0.0125</td>
<td>8.313e-03</td>
<td>0.51</td>
<td>1.3980e-02</td>
<td>0.47</td>
</tr>
</tbody>
</table>

### Table 5.4

Convergence results for the fully implicit mixed discretization; here $h = \tau$ and $\delta = 5\tau$.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$|u - U^\tau|^2$</th>
<th>$\alpha$</th>
<th>$|v - V^\tau|^2$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>2.522029e-01</td>
<td>3.905949e-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>1.172746e-01</td>
<td>0.55</td>
<td>1.724693e-01</td>
<td>0.59</td>
</tr>
<tr>
<td>0.05</td>
<td>3.964776e-02</td>
<td>0.78</td>
<td>6.064739e-02</td>
<td>0.75</td>
</tr>
<tr>
<td>0.025</td>
<td>1.047443e-02</td>
<td>0.96</td>
<td>1.739749e-02</td>
<td>0.90</td>
</tr>
<tr>
<td>0.0125</td>
<td>2.282752e-03</td>
<td>1.10</td>
<td>4.302130e-03</td>
<td>1.01</td>
</tr>
</tbody>
</table>

### 6. Conclusions

We have considered the semidiscrete and fully discrete numerical methods for the upscaled equations. These equations describe the transport and reactions of the solutes. The numerical methods are based on a mixed variational formulation where we have a separate equation for the flux. These numerical methods retain the local mass conservation property. The reaction terms are nonlinear and the dissolution term is multivalued described by a Heaviside graph. To avoid dealing with the inclusions, we use the regularized Heaviside function with the regularization parameter $\delta$ dependent on the time step $\tau$. This implies that in the limit of vanishing discretization parameters automatically yields $\delta \searrow 0$. For the fully discrete situation, we have used lowest order Raviart–Thomas elements. The convergence analysis of
both formulations have been proved using compactness arguments, based on translation estimates. In particular, a discrete $H^1$ norm is used in the proof for the fully discrete scheme.

The work is complemented by the numerical experiments where we study some illustrative examples exhibiting the physical properties of the model. Further, a test case is considered where we construct an exact solution and compare the numerical solution. This numerical study provides us convergence rates for the problem under consideration.

Appendix A. Existence of solution for $P^m_{\delta f,n}$. In this appendix, we prove the existence of a solution for Problem $P^m_{\delta f,n}$. We keep $\tau$ and $\delta$ fixed and let $h \searrow 0$ in the fully discrete problem $P^m_h$. The limit will solve Problem $P^m_{\delta f,n}$. All the steps are similar to the fully discrete case discussed before; therefore, we only give the outline of the proof. Along a sequence $h \searrow 0$, Lemma 4.2 provides the following convergence results:

1. $u^n_h \rightharpoonup u^n_\delta$ weakly in $L^2(\Omega)$,
2. $Q^n_h \rightharpoonup Q^n_\delta$ weakly in $L^2(\Omega)^d$,
3. $\nabla \cdot Q^n_h \rightharpoonup \chi$ weakly in $L^2(\Omega)$,
4. $v^n_h \rightharpoonup v^n_\delta$ weakly in $L^2(\Omega)$.

As before, identification of $\chi$ with $\nabla \cdot Q^n_\delta$ takes place via standard arguments. Further, Lemma 4.6 with the estimate (4.3) gives $\|u^n_h\|_{1,\delta} \leq C$ and after extending $u^n_h$ by 0, Lemma 4.5 implies

$$\|\Delta_{\xi} u^n_h\|_{L^2(\mathbb{R}^2)} \leq C|\xi|(|\xi| + \text{size}(T_h)).$$

Since the right-hand side vanishes uniformly as $|\xi| \searrow 0$, the use of the Riesz–Fréchet–Kolmogorov compactness theorem yields strong convergence of $u^n_h$ to $u^n_\delta$. Now one can use the projection properties and pass $h \searrow 0$ to show that the limit solves $P^m_{\delta f,n}$. Note that having $v$ discretized explicitly in (4.1)$_2$, no nonlinearities in $v^n_h$ are involved and therefore there is no need for strong convergence for $v^n_h$. 

Acknowledgment. Part of the work was completed when K. Kumar visited the Institute of Mathematics, University of Bergen.

REFERENCES


