FINITE ELEMENT APPROXIMATION OF TRANSPORT OF REACTIVE SOLUTES IN POROUS MEDIA

PART 2 ERROR ESTIMATES FOR EQUILIBRIUM ADSORPTION PROCESSES

JOHN W. BARRETT\textsuperscript{1} AND PETER KNABNER\textsuperscript{2}

Abstract. In this paper we analyse a fully practical piecewise linear finite element approximation: involving numerical integration, backward Euler time discretization and possibly regularization and relaxation; of the following degenerate parabolic equation arising in a model of reactive solute transport in porous media: Find $u(x,t)$ such that
\[
\begin{align*}
\partial_t u + \partial_t [\varphi(u)] - \Delta u &= f \quad \text{in} \quad \Omega \times (0, T] \\
u &= 0 \quad \text{on} \quad \partial \Omega \times (0, T] \\
u(\cdot, 0) &= g(\cdot) \quad \text{in} \quad \Omega
\end{align*}
\]
for known data $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, $f$, $g$ and a monotonically increasing $\varphi \in C^0(\mathbb{R}) \cap C^1(-\infty, 0) \cup \{0, \infty\}$ satisfying $\varphi(0) = 0$, which is only locally Hölder continuous, with exponent $p \in (0, 1)$, at the origin; e.g. $\varphi(s) \equiv |s|^p$. This lack of Lipschitz continuity at the origin limits the regularity of the unique solution $u$ and leads to difficulties in the finite element error analysis.

Key words. finite elements, error analysis, degenerate parabolic problem, porous medium

AMS subject classifications. 65M15, 65M60, 35K55, 35K65, 35H35

1. Introduction. This is the second of two papers, in which we study finite element approximations of degenerate parabolic systems and equations, as they arise in the modelling of reactive solute transport in porous media. Here we concentrate on a quasistationary equilibrium adsorption reaction, leading to
\[
\begin{align*}
\partial_t (\Theta u) + \rho \partial_t v - \nabla \cdot (\Theta D \nabla u - qu) &= f \quad \text{in} \quad Q_T \\
v &= \varphi(u) \quad \text{in} \quad Q_T,
\end{align*}
\]
supplemented by initial and boundary conditions for the unknown function $u$, the dissolved concentration. Here $\Omega$ is a bounded domain in $\mathbb{R}^d$, $1 \leq d \leq 3$, $[0, T]$ a fixed time interval and $Q_T \equiv \Omega \times (0, t)$, for $t \in (0, T]$. For more remarks on the model we refer to the introduction of part 1 and for a complete account to the literature cited there. The parameter functions $\Theta$, $\rho$, $D$, $\mu$ are given and fulfill:
\[
\begin{align*}
\partial_t \Theta + \nabla \cdot q = 0, \quad \Theta(x, t) &\geq \Theta_0 > 0, \quad \rho(x) \geq \rho_0 > 0 \quad \text{in} \quad Q_T
\end{align*}
\]
and further conditions such that the linear part of (1.1) defines a uniformly parabolic operator. The nonlinearity $\varphi$ – the adsorption isotherm – is monotone non-decreasing but there are typical examples, which are not Lipschitz continuous at $u = 0$ such as the Freundlich isotherm
\[
\varphi(u) \equiv \alpha u^p \quad \text{for} \quad u \geq 0, \quad \text{where} \quad \alpha \in \mathbb{R}^+ \quad \text{and} \quad p \in (0, 1).
\]

\textsuperscript{1} THIS WORK WAS SUPPORTED BY EC PROJECT "filtration and nonlinear diffusion processes" [contract no. SCI-0018-C(TT)] AND INSTITUT FÜR ANGEWANDTE ANALYSIS UND STOCHASTIK
\textsuperscript{2} DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE, LONDON SW7 2BZ, UK
\textsuperscript{2} INSTITUT FÜR ANGEWANDTE MATHEMATIK, MARTENSSTRASSE 3, 91058 ERLANGEN, GERMANY
Thus in general equation (1.1) is degenerate, exhibiting a finite speed of propagation property, such that a front given by the boundary of the support of $u$, is preserved. In fact, there is a close relation between equation (1.1) and the well-investigated [generalised] "porous medium equation" (see e.g. [1]) which reads

$$\partial_t [\phi(u)] - \Delta u = f \quad \text{in } Q_T$$

with $\phi(u) \equiv \text{sgn}(u)|u|^{1/m}$ for some $m > 1$; i.e. (a model problem of) (1.1) is of the form (1.4a), and (1.1) and (1.4a) are equivalent if we assume that for some $\alpha > 0$

$$\phi'(u) \geq \alpha \text{ for all } u \in \mathbb{R}, u \neq 0.$$  

(1.4b)

A sufficient condition for the finite speed of propagation property is

$$1/\phi \in L^1(0, \delta) \text{ for some } \delta > 0,$$

(1.5) see [16], which has also been proven to be necessary in the one-dimensional case, see [12]. This condition is satisfied by (1.3), see also §2. A common description of chemical non-equilibrium has the form of relaxation kinetics, i.e.

$$\partial_t v = k(\varphi(u) - v)$$

(1.6) with a rate parameter $k > 0$. Equations (1.1a), (1.6) in general form a degenerate system with the aforementioned property. In part 1 we gave a fairly complete order of convergence analysis in energy norms for Galerkin finite element approximations of the system (1.1a), (1.6); based on a technique which is at least applicable for time independent and smooth $\Theta, q, D$. However, the fact that we analysed the Galerkin procedure assumes in addition that the system is not convection-dominated, where we would encounter all the well-known difficulties. This analysis has been presented for a model problem, to which we will restrict ourselves later on.

One may expect that for $k \to \infty$ (P\textsubscript{k}) \equiv (1.1a), (1.6) approximates (P) \equiv (1.1). This may be called a kinetic approximation and will be made rigorous in §2. The aim of this paper is to exploit the kinetic approximation as a proof technique device (and possibly also as an algorithmic device) to study order of convergence estimates for the problem (P) on the basis of the results of part 1 for the relaxed problem (P\textsubscript{k}). There it turned out to be advantageous to introduce a regularized system (P\textsubscript{\varphi, \varepsilon}) obtained by substituting $\varphi$ by a Lipschitz continuous $\varphi\varepsilon$, differing only near $u = 0$. The relaxation is a proof device insofar as the order of convergence estimates, established for the finite element approximation of (P\textsubscript{\varphi, \varepsilon}) for appropriate $\varepsilon = O(h^\gamma), \quad \varepsilon = O(h^\delta), \quad \varepsilon = O(h^\varepsilon)$, where $h$ is the mesh parameter, then carry over to the corresponding finite element approximation of (P\textsubscript{\varepsilon}), the regularized version of (P). We can improve on these convergence estimates by taking into account a non-degeneracy condition, which describes the minimal growth of $u$ away from the front. In the one-dimensional case the following result has been established in [2]. We will assume later on, that $\varphi$ is Hölder continuous near $u = 0$ with exponent $p \in (0, 1)$. If in addition the exponent is sharp, i.e.

$$\varphi(u) \geq \alpha u^p \quad \text{for } u \in [0, \delta_0] \text{ and for some } \alpha, \delta_0 > 0$$

(1.7)

then

$$A\varepsilon(t) \leq C\varepsilon,$$  \hspace{1cm} (N.D.)

(1.8a)
where

\[ A_\varepsilon(t) \equiv \int_0^t m(\Omega_\varepsilon(s)) \, ds. \]

\[ \Omega_\varepsilon(t) \equiv \{ x \in \Omega : u(x, t) \in (0, \varepsilon^{1/(1-p)}) \}, \]

and \( m \) is the Lebesgue measure.

For ease of exposition we will develop our results for the following model problem, which keeps the specific difficulty of the non-Lipschitz nonlinearity, but reduces the handling of standard terms:

(P) Find \( u(x, t) \) such that

\[ \partial_t u + \partial_t |\varphi(u)| - \Delta u = f \quad \text{in } Q_T \]

\[ u = 0 \quad \text{on } \partial \Omega \times (0, T) \quad u(\cdot, 0) = g(\cdot) \quad \text{in } \Omega, \]

where we make the following assumptions on the given data:

[D1]: \( \Omega \subset \mathbb{R}^d, \ 1 \leq d \leq 3 \), with either \( \Omega \) convex polyhedral or \( \partial \Omega \in C^{1,1} \), \( f \in L^\infty(Q_T), g \in L^\infty(\Omega) \cap H_0^1(\Omega) \) and \( \varphi \in C^0(\mathbb{R}) \) is such that

\[ \varphi(0) = 0, \ \varphi(s) > 0 \ \forall s > 0 \ \text{and } \varphi \text{ is monotonically increasing}, \]

\[ \varphi \in C^1(-\infty, 0] \cup (0, \infty), \]

there exist \( L \in \mathbb{R}^+ \) and \( \varepsilon_0, p \in (0, 1) \) such that

\[ |\varphi(a) - \varphi(b)| \leq L |a - b|^p \quad \text{for all } a, b \in [0, \varepsilon_0]. \]

Below we gather together the further assumptions that we require at various stages in the paper.

[D2]: In addition to (D1) we assume that \( f \in H^1(0, T; L^2(\Omega)), g \in H^2(\Omega) \) and that \( k > k_0. \)

[D3]: In addition to (D2) we assume that the constant \( M \) in (2.2c) can be chosen uniformly for all \( s \in \mathbb{R} \) (In view of the bounds (2.5) for \( u \), see Theorem 2.2, this is always achievable by changing \( \varphi(s) \) for \( |s| \geq m = \max \{|-y, u|\} \). Let \( \Omega^h \) be a polyhedral approximation to \( \Omega \) defined by \( \Omega^h \equiv \bigcup_{\kappa \in T^h} \kappa \) with \( \text{dist}(\partial \Omega, \partial \Omega^h) \leq C h^2 \); where \( T^h \) is a partitioning consisting of regular simplices \( \kappa \) with \( h_\kappa \equiv \text{diam}(\kappa) \) and \( h \equiv \max_{\kappa \in T^h} h_\kappa. \) For ease of exposition we assume that \( \Omega^h \subset \Omega. \)

[D4]: In addition to (D3) we assume that \( f \in H^1(0, T; C^0(\Omega)) \cap L^2(0, T; H^2(\Omega)). \)

We prove our basic error bound for a fully practical approximation to (P) under the assumptions (D5), where

[D5]: In addition to (D4) we assume that \( T^h \) is such that

(i) for \( d = 2 \) it is weakly acute; that is, for any pair of adjacent triangles the sum of opposite angles relative to the common side does not exceed \( \pi \); and

(ii) for \( d = 3 \) the angle between any two faces of the same tetrahedron does not exceed \( \pi/2. \)

We improve on these basic error bounds by replacing (D5) by

[D6]: in addition to (D5) we assume that

(i) \( \Omega \subset \mathbb{R}^d, \ d = 1 \) or 2, and \( T^h \) is a quasi-uniform partition if \( d = 2. \)

(ii) \( g \) and \( f \geq 0 \) and
(iii) \( \varphi \in C^2(0, m] \) such that \( \varphi''(s) \leq 0 \) for all \( s \in (0, m] \), where \( m = \max \{ \bar{w} \} \); see (2.5).

We note that in proving the error bounds in this paper, as opposed to the bounds in for example Corollary 3.1 and (4.1), the only place that the acuteness assumption on the partitioning \( T^h \) is required, is in establishing the bounds (3.7) and (3.8).

The layout of this paper is as follows. In the next section we establish the existence and uniqueness of a solution to (P) under assumptions (D1) by firstly establishing these results for a regularized relaxed version \((P_{k,e})\). In addition we recall a number of useful a priori estimates for \((P_{k,e})\) under assumptions (D1) and (D2). In \S3 under assumptions (D5) we prove error bounds for a continuous in time continuous piecewise linear finite element approximation in space, involving numerical integration, of \((P_e)\).

Moreover, we improve on these bounds under assumptions (D6). In addition we note that one can prove superior error bounds for a less practical scheme involving no numerical integration under assumptions (D3). In \S4 we consider a fully practical approximation involving discretization in time using the backward Euler method. Finally, in \S5 we report on a numerical experiment.

The most complete order of convergence analysis until now for the finite element approximation of the porous medium equation, involving time discretization and numerical integration has been given in [11]. In fact many of the proof techniques used in this paper are similar to those used there. However contrary to our approach they consider this approximation directly, taking regularization but not relaxation of the problem into account. A proviso in the comparison lies in the fact that in some places we require the mesh to be (weakly) acute, whereas they do not. Our approach, firstly leads to their resulting error bounds with a less severe time step constraint; that is, \( \tau = C h \) as opposed to their restrictions \( \tau = C h^{1+\eta} \) and \( \tau = C h^{3/2(3-\eta)} \) on not assuming and assuming (N.D.), respectively. Furthermore under additional assumptions, see (D6), we can improve on their error bounds in some cases.

Finally, we note that one could employ alternative forms of relaxation, not considered here. The description of a physically caused non-equilibrium may lead to

\[
(1.10) \quad \partial_t u = k(u - \varphi^{-1}(v)).
\]

For a nonlinearity of the type (1.3), \( \varphi^{-1} \) is Lipschitz continuous, i.e. (1.1a), (1.10) is a regular system. This type of relaxation was used in [15] for the Stefan problem. A semi-explicit time discretization leads to a linear elliptic problem at each time level; whereas for the relaxation (1.6) considered here, a fully implicit time discretization leads to a semilinear elliptic problem of the type studied in [9] and [5]. Therefore from a computational viewpoint it would seem advisable to use a relaxation based on (1.10). However we note that (1.10), as it leads to a regular system, adds a considerable amount of artificial diffusion. This has already been reported in [15].

7] proposes a modification of relaxation to improve on this aspect, which leads to a nonlinear algorithm at each time level. A relaxation based on (1.6) smears much less, as it still leads to a problem with a finite speed of propagation. Therefore for computational purposes, we have to leave it open whether relaxation and of which form is to be preferred. We stress again that in this paper the relaxation based on (1.6) is mainly used as a proof technique to establish order of convergence results for the finite element approximation of (P) (without relaxation), which we are unable to prove otherwise. As an intermediate step, we obtain the same results for an approximation with relaxation. Here the computational complexities of the fully discrete problems are identical.
Throughout the paper we adopt the standard notation for Sobolev spaces. We note that the seminorm $\| \cdot \|_{H^1(\Omega)}$ and the norm $\| \cdot \|_{H^1_0(\Omega)}$ are equivalent on $H^1_0(\Omega)$. The standard $L^2$ inner product over $\Omega$ is denoted by $\langle \cdot , \cdot \rangle$. Throughout $C$ denotes a generic positive constant independent of $\varepsilon$ the regularization parameter, $k$ the relaxation parameter and $h$ the mesh spacing.

2. The continuous problem. In this section we review the existence and uniqueness of a solution to (P). In doing so we will develop various bounds that will be useful in analysing the error in the finite element approximation of (P). Firstly we introduce a regularized version of (P), for $\varepsilon \in (0, \varepsilon_0] \quad (\varepsilon_0$ as in (1.9c)).

(P$_\varepsilon$) Find $u_{\varepsilon}(x,t)$ such that

$$\frac{\partial}{\partial t} u_{\varepsilon} + \frac{\partial}{\partial t} [\varphi(\varepsilon u_{\varepsilon})] - \Delta u_{\varepsilon} = f \quad \text{in } Q_T$$

$$u_{\varepsilon} = 0 \quad \text{on } \partial \Omega \times (0, T] \quad u_{\varepsilon}(\cdot, 0) = g_{\varepsilon}(\cdot) \quad \text{in } \Omega,$$

where $g_\varepsilon \in L^\infty(\Omega) \cap H^1_0(\Omega)$ is such that

\begin{align}
(2.1a) \quad & g_\varepsilon \to g \text{ in } L^\infty(\Omega) \cap H^1_0(\Omega) \text{ as } \varepsilon \to 0, \\
(2.1b) \quad & |g_\varepsilon|_{H^2(\Omega)} \leq C \text{ if } g \in H^2(\Omega),
\end{align}

and $\varphi_\varepsilon \in C^{0,1}_{loc}(\mathbb{R})$ is such that

\begin{align}
(2.2a) \quad & \varphi_\varepsilon(s) = \varphi(s) \text{ for } s \notin (0, \varepsilon^{1/(1-p)}), \\
(2.2b) \quad & \varphi_\varepsilon(s) \text{ is strictly monotonically increasing on } [0, \varepsilon^{1/(1-p)}], \\
(2.2c) \quad & \text{for } m \in \mathbb{N} \text{ there exists a } M(m) \in \mathbb{R}^+: \varphi_\varepsilon(b) - \varphi_\varepsilon(a) \leq M(m) \varepsilon^{-1} (b - a) \quad \text{for } -m \leq a \leq b \leq m.
\end{align}

Note that $M$ can be chosen independently of $m$, if $\varphi'$ is bounded in $\mathbb{R} \setminus (0, b)$ for some $b > 0$. In addition we set

$$\Phi_\varepsilon(s) := \int_0^s \varphi_\varepsilon(\sigma) \, d\sigma.$$ 

It is a simple matter to deduce from the conditions (2.2) that for all $|a|, |b| \leq m$

\begin{align}
(2.3a) \quad & |M(m)|^{-1} \varepsilon |\varphi_\varepsilon(a) - \varphi_\varepsilon(b)|^2 \leq |\varphi_\varepsilon(a) - \varphi_\varepsilon(b)|(a-b) \leq M(m) \varepsilon^{-1} |a-b|^2,
\end{align}

and

\begin{align}
(2.3b) \quad & \varphi_\varepsilon(\varepsilon^{1/(1-p)}) = \varphi(\varepsilon^{1/(1-p)}) \leq L \varepsilon^{(1-p)/2},
\end{align}

with $L$ as in (1.9c). The simplest choice for $\varphi_\varepsilon$ is the linear regularization

$$\varphi_\varepsilon(s) \equiv \varepsilon^{-1/(1-p)} \varphi(\varepsilon^{1/(1-p)} s) \quad \text{for } s \in (0, \varepsilon^{1/(1-p)}).$$

In addition it is useful to consider the following problem in which the reaction process is relaxed in time with $k > 0$ being the given relaxation parameter.

(P$_{k,\varepsilon}$) Find $\{u_{k,\varepsilon}(x,t), v_{k,\varepsilon}(x,t)\}$ such that

$$\frac{\partial}{\partial t} u_{k,\varepsilon} + \frac{\partial}{\partial t} v_{k,\varepsilon} - \Delta u_{k,\varepsilon} = f \quad \text{in } Q_T \quad u_{k,\varepsilon} = 0 \quad \text{on } \partial \Omega \times (0, T].$$
\[ \partial_t v_{k,\varepsilon} = k(\varphi_{\varepsilon}(u_{k,\varepsilon}) - v_{k,\varepsilon}) \quad \text{in } Q_T \]

\[ u_{k,\varepsilon}(\cdot, 0) = g_{\varepsilon}(\cdot) \quad v_{k,\varepsilon}(\cdot, 0) = \varphi_{\varepsilon}(g_{\varepsilon}(\cdot)) \quad \text{in } \Omega. \]

The above problem has been studied in part 1. We adopt the notion of weak solution defined there and below we recall some of the results.

**Theorem 2.1.** Let the assumptions (D1) hold. Then for all \( \varepsilon \in (0, \varepsilon_0] \) and \( k > 0 \) there exists a unique weak solution \( \{u_{k,\varepsilon}, v_{k,\varepsilon}\} \) to \((P_{k,\varepsilon})\) such that

\[ u_{k,\varepsilon} \leq u_{k,\varepsilon} \leq \bar{u} \quad \text{and} \quad v \leq v_{k,\varepsilon} \leq \bar{v} \quad \text{in } Q_T, \]

where \( u, \bar{u}, v, \bar{v} \in C(\overline{\Omega}) \) are all independent of \( \varepsilon \) and \( k \). Furthermore, if \( g_{\varepsilon} \) and \( f \geq 0 \) one can take \( \bar{u} = \bar{v} = 0 \).

**Proof.** This result with \( u, \bar{u}, v, \bar{v} \in C[0, T], \) all independent of \( \varepsilon \) and uniformly bounded in \( k \), follows from Theorem 2.1 of part 1. Furthermore noting Remark 2.1 of part 1 yields the above choice of \( u, \bar{u}, v, \bar{v} \). We note for later purposes that \( u \) and \( \bar{u} \) depend only on \( \Omega, \|f\|_{L^\infty(\Omega)} \) and \( |g_{\varepsilon}|_{L^\infty(\Omega)} \).

**Lemma 2.1.** Under assumptions (D1) we have for all \( \varepsilon \in (0, \varepsilon_0], k > 0 \) and \( t \in (0, T] \) that

\[ \varepsilon |\nabla \varphi_{\varepsilon}(u_{k,\varepsilon})|_{L^2(Q_T)}^2 + \int_0^t \langle \nabla u_{k,\varepsilon}(\cdot, s), \nabla \varphi_{\varepsilon}(u_{k,\varepsilon}(\cdot, s)) \rangle \, ds + \langle \Phi_{\varepsilon}(u_{k,\varepsilon}(\cdot, t)), 1 \rangle + \]

\[ k |\varphi_{\varepsilon}(u_{k,\varepsilon}) - v_{k,\varepsilon}|_{L^2(Q_T)}^2 + |v_{k,\varepsilon}(\cdot, t)|_{L^2(\Omega)}^2 + k^{-1} |\partial_t v_{k,\varepsilon}|_{L^2(Q_T)}^2 \leq C. \]  

**Proof.** See Lemma 2.1 of part 1. \( \square \)

**Lemma 2.2.** Let the assumptions (D2) hold. Then we have for all \( \varepsilon \in (0, \varepsilon_0], k > 0 \) and \( t \in (0, T] \) that

\[ |\nabla u_{k,\varepsilon}(\cdot, t)|_{L^2(\Omega)}^2 + |\partial_t u_{k,\varepsilon}|_{L^2(Q_T)}^2 + \varepsilon |\partial_t v_{k,\varepsilon}|_{L^2(Q_T)}^2 + \varepsilon |\partial_t \varphi_{\varepsilon}(u_{k,\varepsilon})|_{L^2(Q_T)}^2 + \]

\[ k^{-1} |\partial_t u_{k,\varepsilon}(\cdot, t)|_{L^2(\Omega)}^2 + \varepsilon |\partial_t v_{k,\varepsilon}(\cdot, t)|_{L^2(\Omega)}^2 + |\nabla(\partial_t u_{k,\varepsilon})|_{L^2(Q_T)}^2 \leq C. \]

**Proof.** Noting (2.1b), see Lemma 2.2 in part 1. \( \square \)

We will prove existence of solutions of problems \((P_{\varepsilon})\) or \((P)\) in the following sense.

**Definition:** \( u_{\varepsilon} \) is a weak solution to \((P_{\varepsilon})\), if \( u_{\varepsilon} \in L^2(0, T; H^1_0(\Omega)) \) is such that \( \varphi_{\varepsilon}(u_{\varepsilon}) \in L^2(Q_T) \) and for all test functions \( \eta \in L^2(0, T; H^1_0(\Omega)) \cap H^1(0, T; L^2(\Omega)) \) with \( \eta(\cdot, T) = 0 \) in \( \Omega \)

\[ \int_{Q_T} \{-[u_{\varepsilon} + \varphi_{\varepsilon}(u_{\varepsilon})] \partial_t \eta + \nabla u_{\varepsilon} \cdot \nabla \eta - f \eta\} \, dx \, dt - \langle [g_{\varepsilon}(\cdot) + \varphi_{\varepsilon}(g_{\varepsilon}(\cdot))], \eta(\cdot, 0) \rangle = 0. \]

A similar definition holds for \((P)\) with \( u_{\varepsilon}, \varphi_{\varepsilon}(u_{\varepsilon}) \) and \( g_{\varepsilon} \) replaced by \( u, \varphi(u) \) and \( g \).

For \( k \in \mathbb{R}^+ \equiv \mathbb{R}^+ \cup \{-\infty\} \) and for sufficiently smooth \( w \) we set

\[ |w|^2_{E_{k,\varepsilon}(Q_t)} \equiv |w|^2_{L^2(Q_t)} + \frac{1}{2} k^{-1} |w(\cdot, t)|_{L^2(\Omega)}^2 \]

and

\[ |w|^2_{E_{k,\varepsilon}(Q_t)} \equiv |w|^2_{E_{k,\varepsilon}(Q_t)} + \frac{1}{2} \int_0^t |\nabla \int_0^s w(\cdot, s) \, ds|^2_{L^2(\Omega)} + k^{-1} |\nabla w|^2_{L^2(Q_t)}. \]
THEOREM 2.2. Let the assumptions (D1) hold. Then for all $\varepsilon \in (0, \varepsilon_0]$ there exists a unique weak solution $u_\varepsilon$ to (P$_\varepsilon$) and

$$u \leq u_\varepsilon \leq \overline{u} \quad \text{in } Q_T,$$

where $u_\varepsilon$, $\overline{u} \in C^0(\overline{\Omega})$ are all independent of $\varepsilon$. Moreover, if $g_\varepsilon$ and $f \geq 0$ then $u_\varepsilon \geq 0$ in $Q_T$. Furthermore, for all $k > 0$ and $t \in (0, T]$ we have that

$$|u_\varepsilon - u_{k,\varepsilon}|^2_{L^2(\Omega)} + \varepsilon|\varphi_\varepsilon(u_\varepsilon) - \varphi_\varepsilon(u_{k,\varepsilon})|^2_{L^2(Q_t)} + \varepsilon|\varphi_\varepsilon(u_\varepsilon) - v_{k,\varepsilon}|^2_{L^2(Q_t)}$$

$$\leq Ck^{-2}|\partial_s v_{k,\varepsilon}|^2_{L^2(Q_T)} \leq \begin{cases} Ck^{-1} & \text{if (D1) holds} \\ C_{\varepsilon}^{-1}k^{-2} & \text{if (D2) holds} \end{cases}$$

(2.8b)

Proof. Existence of a weak solution $u_\varepsilon$ to (P$_\varepsilon$) follows by letting $k \to \infty$ in (P$_{k,\varepsilon}$), from which it is clearly seen from (2.5) that the bounds (2.8a) hold, see [4] for details. Let $\varepsilon \equiv u_\varepsilon^1 - u_\varepsilon^2$, where $u_\varepsilon^1, u_\varepsilon^2$ are two weak solutions of (P$_\varepsilon$). Then subtracting the two defining equations and choosing $\eta(\cdot, t) \equiv \int_{t}^{T} \epsilon\left(\cdot, s\right) ds$ yields $\|e_{\varepsilon,\varepsilon}\|_{L^2(\Omega)} \leq 0$ on noting (2.3a) and hence uniqueness.

Let $e_{\varepsilon}^u \equiv u_\varepsilon - u_{k,\varepsilon}$ and $e_{\varepsilon}^v \equiv \varphi_\varepsilon(u_\varepsilon) - v_{k,\varepsilon}$. Then subtracting the first equation in (P$_{k,\varepsilon}$) from (P$_\varepsilon$), multiplying by $\int_{t}^{T} e_{\varepsilon}^u(\cdot, \sigma) d\sigma$, integrating over $Q_t$, where $s$ is the integration variable in time, and performing integration by parts and noting (2.3a) and the second equation in (P$_{k,\varepsilon}$), yields that

$$\|e_{\varepsilon,\varepsilon}\|_{L^2(\Omega)}^2 + \|e_{\varepsilon,\varepsilon}\|_{L^2(\Omega)}^2 + \int_{t}^{T} \left(\varphi_\varepsilon(u_\varepsilon(\cdot, s)) - \varphi_\varepsilon(u_{k,\varepsilon}(\cdot, s)), e_{\varepsilon,\varepsilon}^u(\cdot, s)\right) ds$$

$$= \int_{t}^{T} \left(\varphi_\varepsilon(u_\varepsilon(\cdot, s)) - \varphi_\varepsilon(u_{k,\varepsilon}(\cdot, s)), e_{\varepsilon,\varepsilon}^v(\cdot, s)\right) ds$$

$$= -k^{-1} \int_{t}^{T} \left(\partial_s v_{k,\varepsilon}(\cdot, s), e_{\varepsilon,\varepsilon}^u(\cdot, s)\right) ds,$$

(2.9)

where $[\inf \underline{u}, \sup \overline{u}] \subseteq [-m, m]$, see Theorem 2.1. Hence the desired result (2.8b) follows from (2.9), (2.6) and (2.7).

THEOREM 2.3. Let the assumptions (D1) hold. Then there exists a unique weak solution $u$ to (P) and for all $\varepsilon \in (0, \varepsilon_0]$ and $t \in (0, T]$ we have that

$$|u - u_\varepsilon|^2_{L^2(\Omega)} + \varepsilon|\varphi(u) - \varphi(u_\varepsilon)|^2_{L^2(Q_t)}$$

$$|u - u_\varepsilon|^2_{L^2(\Omega)} + \varepsilon|\varphi(u) - \varphi(u_\varepsilon)|^2_{L^2(Q_t)} \leq C_{A_{\varepsilon}}(t) \epsilon^{1+p}/(1-p) + C\left[|g - g_\varepsilon| + |\varphi(g) - \varphi_\varepsilon(g_\varepsilon)|^2_{H^{-1}(\Omega)}\right].$$

(2.10)

In addition, the bounds (2.8a) hold for $u_\varepsilon$; and if $g, f \geq 0$ then $u \geq 0$ in $Q_T$.

Proof. Existence of a weak solution $u$ to (P) follows by letting $\varepsilon \to 0$ in (P$_\varepsilon$), from which it is clearly seen that the bounds (2.8a) hold for $u$, see [4] for details. Uniqueness follows as for (P$_\varepsilon$) with $\varphi_\varepsilon$ replaced by $\varphi$, see Theorem 2.2.

Finally to show (2.10), let $e^u \equiv u - u_\varepsilon$ and $e^v \equiv \varphi(u) - \varphi_\varepsilon(u_\varepsilon)$. Subtracting the first equation in (P$_\varepsilon$) from (P), multiplying by $\int_{t}^{T} e^u(\cdot, \sigma) d\sigma$, integrating over $Q_t$, yields

$$\|e^u\|^2_{L^2(\Omega)} = -\int_{t}^{T} \left(e^u(\cdot, s), e^u(\cdot, s)\right) ds + \left(e^u(\cdot, 0) + e^v(\cdot, 0), \int_{t}^{T} e^u(\cdot, s) ds\right).$$
and therefore, noting (2.3a),
\[
\|e^u\|_{E^2(\infty, t)}^2 + [M(m)]^{-1}\|e^e\|_{L^2(\Omega)}^2 \\
\leq \int_0^t \langle e^u(s, \cdot), (\zeta - u)(s, \cdot) \rangle \, ds + \int_0^t \langle e^u(s, \cdot), e^e(s, \cdot) \rangle \, ds \\
\leq C A_e(t) e^{(1 + p)/(1 - p)} + C \|g - g_e\|_{H^{-1}(\Omega)} + [\varphi(g) - \varphi_e(g_e)]^2_{H^{-1}(\Omega)},
\]
where \(\zeta \equiv \varphi_e^{-1}(\varphi(u))\) if \(\varphi(u) \in (0, \varphi(\varepsilon^{1/(1 - p)})]\) and \(\zeta \equiv u\) otherwise, and \([\inf u, \sup \bar{u}] \subseteq [-m, m] \). 

Next we discuss possible choices of \(g_e\). On choosing
(a) \(g_e = g\) it follows from Sobolev embedding, (2.2a), (2.2b) and (2.3b) that
\[
\|g - g_e\|_{H^{-1}(\Omega)}^2 + [\varphi(g) - \varphi_e(g_e)]^2_{H^{-1}(\Omega)} \leq C \|\varphi(g) - \varphi_e(g_e)\|_{L^p(\Omega)}^2
\]
(2.11a)
\[
\leq C [m(\Omega_{g,e})]^{2/\mu} e^{2p/(1 - p)},
\]
where \(\Omega_{g,e} \equiv \{x \in \Omega : g(x) \in (0, \varepsilon^{1/(1 - p)})\}\), \(\mu = 1\) if \(d = 1, \mu > 1\) if \(d = 2\) and \(\mu = 6/5\) if \(d = 3\). Hence if \(m(\Omega_{g,e})^{2/\mu} e^{2p/(1 - p)} \leq C A_e(T) e^{(1 + p)/(1 - p)}\), that is,
(2.11b)
\[
m(\Omega_{g,e}) \leq C [A_e(T) e]^{\mu/2}
\]
then it is acceptable to choose \(g_e = g\). However, for general \(g \in H^2(\Omega)\) this choice leads to an inferior rate of convergence as \(\varepsilon \to 0\) compared to the first term on the right hand side of (2.10).

(b) \(g_e\) such that \(\varphi_e(g_e) = \varphi(g)\) yields that \(g_e = g\) if \(g \in (0, \varepsilon^{1/(1 - p)})\) and hence it follows that
\[
\|g - g_e\|_{H^{-1}(\Omega)}^2 + [\varphi(g) - \varphi_e(g_e)]^2_{H^{-1}(\Omega)} \leq C m(\Omega_{g,e}) e^{(1 + p)/(1 - p)}.
\]
(2.12)

However, \(g_e \notin H^2(\Omega)\) in general and therefore many of the a priori estimates do not hold, e.g. (2.7). Below we seek an alternative choice for \(g_e\).

**Corollary 2.3.** Let the assumptions (D2) hold. Then for all \(\varepsilon \in (0, \varepsilon_0]\) on choosing \(g_e \in H^1_0(\Omega)\) to be the unique solution of
\[
(\nabla g_e, \nabla \eta) + \langle \varphi_e(g_e), \eta \rangle = (\nabla g, \nabla \eta) + \langle \varphi(g), \eta \rangle \quad \forall \eta \in H^1_0(\Omega)
\]
yields that \(\|g - g_e\|_{L^\infty(\Omega)} \leq \varepsilon^{1/(1 - p)}\) and
\[
\varepsilon \|g - g_e\|_{H^1(\Omega)}^2 + \|g - g_e\|_{H^1(\Omega)}^2 + \varepsilon [\varphi(g) - \varphi_e(g_e)]^2_{L^2(\Omega)} + [\varphi(g) - \varphi_e(g_e)]^2_{H^{-1}(\Omega)}
\]
(2.13a)
\[
\leq C m(\Omega_{g,e}) e^{(1 + p)/(1 - p)}.
\]
Furthermore, if \(\varphi(s) \geq \varphi_e(s)\) for all \(s \in (0, \varepsilon^{1/(1 - p)})\) then
(2.13b)
\[
g_e \geq g \quad \text{in } \overline{\Omega}
\]
and hence if \(g, f \geq 0\) then \(u_{k,\varepsilon}, v_{k,\varepsilon}, u_{\varepsilon} \geq 0\) in \(Q_T\).

**Proof.** From Theorem 2.1 in [9] it follows that \(\|g - g_e\|_{L^\infty(\Omega)} \leq \varepsilon^{1/(1 - p)}\). Furthermore we have from (2.3a) that
\[
C [\varphi(g) - \varphi_e(g_e)]^2_{H^{-1}(\Omega)} \leq \|g - g_e\|_{H^1(\Omega)}^2 + C \varepsilon [\varphi(g) - \varphi_e(g_e)]^2_{L^2(\Omega)}
\]
\[
\leq \|g - g_e\|_{H^1(\Omega)}^2 + [\varphi(g) - \varphi_e(g_e)]^2_{H^{-1}(\Omega)}
\]
\[
= [\varphi(g) - \varphi_e(g_e)]^2_{H^{-1}(\Omega)}
\]
\[
\leq C e^{-1} \|g - g_e\|_{L^2(\Omega)}^2 \leq C m(\Omega_{g,e}) e^{(1 + p)/(1 - p)}.
\]
where $\zeta \equiv \varphi_{e}^{-1}(\varphi(g))$ if $\varphi(g) \in (0, \varphi(\varepsilon^{1/(1-p)}))$ and $\zeta \equiv g$ otherwise. By elliptic regularity $|g - g_{e}|_{L^{2}(\Omega)}^{2} \leq C|\varphi_{e}(g) - \varphi_{e}(g_{e})|_{L^{2}(\Omega)}^{2}$. Hence the desired result (2.14a).

(2.14b) follows from noting that

$$
|\nabla [g - g_{e}]_{+}|_{L^{2}(\Omega)}^{2} = \langle \nabla (g - g_{e}), \nabla [g - g_{e}]_{+} \rangle = \langle \varphi_{e}(g_{e}) - \varphi(g), [g - g_{e}]_{+} \rangle
$$

$$
\leq \langle \varphi_{e}(g - \varphi(g), [g - g_{e}]_{+} \rangle \leq 0.
$$

Finally (2.14b) and Theorems 2.1 and 2.2 yield that if $g, f \geq 0$ then $u_{k}, v_{k, e}, u_{e} \geq 0$ in $Q_{T}$. □

Throughout the rest of this paper $g_{e}$ is as defined by (2.13). Note that we will recover the use of $g_{e} = g$, and its interpolant in the finite element approximation, if $g$ satisfies (2.11b); see Lemma 3.5 and Remark 3.1.

Problem (P) is strongly related to a degenerate problem, which has been investigated intensively, the (generalised) porous medium equation

$$
\begin{equation}
\partial_{t} w - \Delta |\beta(w)| = f \text{ in } Q_{T}.
\end{equation}
$$

where $\beta : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and strictly increasing, and without loss of generality $\beta(0) = 0$. The (classical) porous medium equation is given by $\beta(w) \equiv sgn(w)|w|^m$ for some $m > 1$. A change of variables yields (1.4a) with $\phi = \beta^{-1}$. Obviously (P) is of the form (1.4a). On the other hand, (1.4a) can be written in the form of (P), if we assume that $\phi$ satisfies (1.9) and, as in [11], for some $\alpha > 0$

$$
\begin{equation}
\phi'(s) \geq \alpha \text{ for all } s \in \mathbb{R}.
\end{equation}
$$

where we allow for $\phi'(0) = \infty$; as we can substitute $\partial_{t} u$ by $\alpha \partial_{t} u$ in the definition of (P), which amounts to substituting $\varphi$ by $\varphi/\alpha$ and scaling $t$ by $1/\alpha$. Actually, we can even cast problem (1.4a) in the form of (P) if we only assume for every $m > 0$

$$
\begin{equation}
\phi'(s) \geq \alpha(m) > 0 \text{ for all } s \in [-m, m].
\end{equation}
$$

This condition is satisfied, if e.g. $\beta \in C^{1}(\mathbb{R})$ and $\beta'(s) > 0$ for $s \neq 0$. This can be seen as follows. As already noted, we can substitute $\partial_{t} u$ by $\alpha \partial_{t} u$ in the definition of (P), (P) and (P) without affecting the developed theory. In particular the bounds $\bar{u}, \bar{u}$ for the u-component are independent of $\alpha$ and $\varphi$. Choose $m = \max\{|u|_{L^{\infty}(\Omega)}, \bar{u}|_{L^{\infty}(\Omega)}\}$ and $\alpha = \alpha(m)$ according to (2.17). This $\alpha$ we take in the definition of (P) and $\varphi \equiv \varphi^{(\alpha)}$ defined by

$$
\begin{equation}
\varphi^{(\alpha)}(s) \equiv \begin{cases} 
\phi(-m) + \alpha m & s \leq -m \\
\phi(s) & |s| \leq m \\
\phi(m) & s \geq m
\end{cases}
\end{equation}
$$

Then $\varphi^{(\alpha)}$ satisfies (1.9), if $\phi$ does so, and as the solution of (P) fulfills $|u|_{L^{\infty}(Q_{T})} \leq m$, we have that $\alpha u + \varphi^{(\alpha)}(u) = \phi(u)$, i.e. the solution of (1.4a) is the solution (P).

The existence result for (P) in Theorem 2.3 is not new. It is quite comparable to the basic results for the generalised porous medium equation (compare e.g. [13]). What is of importance for the following, is the precise information about its approximation by (P).
3. A continuous in time finite element approximation. We now consider the continuous piecewise linear finite element approximation to $(P_{k,c})$. Assuming (D3), we introduce

\[ S_h \equiv \{ \chi \in C(\overline{\Omega}^h) : \chi|_{\kappa} \text{ is linear for all } \kappa \in T^h \} \]

and

\[ S_h^0 \equiv \{ \chi \in S_h : \chi = 0 \text{ on } \partial \Omega^h \}. \]

In the analysis that follows we extend $\chi \in S_h$ from $\overline{\Omega}^h$ to $\overline{\Omega}^h \setminus \overline{\Omega}$ by zero. Let $\pi_h : C^0(\Omega) \to S_h^0$ denote the interpolation operator such that for any $w \in C^0(\Omega)$, $\pi_h w \in S_h$ satisfies

\[ (\pi_h w)(x_i) = w(x_i) \quad \text{for all nodes } x_i \text{ of the partition } T^h. \]

Let $P_h^0 : L^2(\Omega) \to S_h$ denote the $L^2$ projection such that for any $w \in L^2(\Omega)$, $P_h^0 w \in S_h$ satisfies

\[ \langle w - P_h^0 w, \chi \rangle = 0 \quad \forall \chi \in S_h^0. \]

Let $P_h^1 : H^1(\Omega) \to S_h^0$ denote the $H^1$ semi-norm projection such that for any $w \in H^1(\Omega)$, $P_h^1 w \in S_h^0$ satisfies

\[ \langle \nabla (w - P_h^1 w), \nabla \chi \rangle = 0 \quad \forall \chi \in S_h^0. \]

We recall the standard approximation results, for all $\kappa \in T^h$

\[ |w - \pi_h w|_{W^{m,q}(\kappa)} \leq Ch^{q-m} |w|_{W^{2,q}(\kappa)} \quad \text{for } m = 0 \text{ and } 1 \quad \text{and} \]

\[ \forall q \in [1, \infty] \text{ if } d \leq 2 \text{ and } \forall q \in (3/2, \infty] \text{ if } d = 3, \]

\[ |w - P_h^0 w|_{L^2(\Omega)} \leq Ch^m |w|_{H^m(\Omega)} \quad \text{for } m = 0, 1 \text{ and } 2 \]

and

\[ |w - P_h^1 w|_{L^2(\Omega)} + h |w - P_h^1 w|_{H^2(\Omega)} \leq Ch^m |w|_{H^m(\Omega)} \quad \text{for } m = 1 \text{ and } 2. \]

where in (3.1a) we note the embedding $W^{2,1}(\kappa) \subset C^0(\k)$ in the case $d=2$, see for example p. 300 in [8]. Another result that will be useful later is that

\[ |(I - \pi_h) \varphi_c(\chi)|_{L^2(\Omega)} \leq Ch |\nabla \pi_h[\varphi_c(\chi)]|_{L^2(\Omega)} \quad \forall \chi \in S_h^0; \]


The standard Galerkin approximation to $(P_{k,c})$ is then:

\[ \left( P_{k,c}^h \right) \] Find $u_{k,c}^h \in H^1(0,T;S_h^0)$ and $v_{k,c}^h \in H^1(0,T;S_h^0)$ such that

\[ \langle \partial_t u_{k,c}^h + \partial_x v_{k,c}^h, \chi \rangle + \langle \nabla u_{k,c}^h, \nabla \chi \rangle = \langle f, \chi \rangle \quad \forall \chi \in S_h^0 \]

\[ \langle \partial_t v_{k,c}^h, \chi \rangle = k \langle \varphi_c(u_{k,c}^h) - v_{k,c}^h, \chi \rangle \quad \forall \chi \in S_h \]

\[ u_{k,c}^h(\cdot,0) = P_h^0 g_c(\cdot) \quad v_{k,c}^h(\cdot,0) = P_h^0 [\varphi_c(g_c(\cdot))]. \]

The above approximation is not practical as it requires the term $\langle \varphi_c(u_{k,c}^h), \chi \rangle$ to be integrated exactly. This is obviously difficult in practice and it is computationally more convenient to consider a scheme where numerical integration is applied to all the terms and the initial data is interpolated as opposed to being projected. Below we introduce and analyse such a scheme.
For all \( w_1, w_2 \in C^0(\Omega^h) \) we set
\[
\langle w_1, w_2 \rangle_h = \int_{\Omega^h} \pi_h(w_1 w_2)
\]
as an approximation to \( \langle w_1, w_2 \rangle \). On setting
\[
|w|_h \equiv [\langle w, w \rangle_h]^{1/2} \quad \text{for } w \in C^0(\Omega^h),
\]
we recall the well-known results
\[
(3.3a) \quad |\chi|_{L^2(\Omega^h)} \leq |\chi|_h \leq C_1 |\chi|_{L^2(\Omega^h)} \quad \forall \chi \in S^h.
\]
and for \( m = 0 \) or 1
\[
(3.3b) \quad \left| \int_{\Omega^h} \chi_1 \chi_2 - (\chi_1, \chi_2)^h \right| \leq C_2 h^{1+m} \|\chi_1\|_{H^1(\Omega^h)} \|\chi_2\|_{H^m(\Omega^h)} \quad \forall \chi_1, \chi_2 \in S^h.
\]

Assuming (D4), a more practical approximation to \( (P_{k,\varepsilon}) \) than \( (P^h_{k,\varepsilon}) \) is then:
\[
(P^\varepsilon_{k,\varepsilon}) \quad \text{Find } u^\varepsilon_{k,\varepsilon} \in H^1(0, T; S^h_0) \text{ and } v^\varepsilon_{k,\varepsilon} \in H^1(0, T; S^h) \text{ such that}
\]
\[
\begin{align*}
&\langle \partial_t u^\varepsilon_{k,\varepsilon} + \partial_x v^\varepsilon_{k,\varepsilon}, \chi \rangle^h + \langle \nabla u^\varepsilon_{k,\varepsilon}, \nabla \chi \rangle^h = \langle f, \chi \rangle^h \quad \forall \chi \in S^h_0 \\
&\langle \partial_t v^\varepsilon_{k,\varepsilon}, \chi \rangle^h = k \left< \varphi_\varepsilon(u^\varepsilon_{k,\varepsilon}) - v^\varepsilon_{k,\varepsilon}, \chi \right>^h \quad \forall \chi \in S^h
\end{align*}
\]
where \( g^\varepsilon \in S^h_0 \) is a suitable approximation to \( g_\varepsilon \), the unique solution of (2.13), satisfying \( \|\nabla g^\varepsilon, \nabla \chi \| \leq C |\chi|_{L^2(\Omega)} \forall \chi \in S^h_0 \). Hence it follows that \( |g^\varepsilon|_{H^1(\Omega)} \| \varphi_\varepsilon(g^\varepsilon)\|_{h} \leq C \).

**Theorem 3.1.** Let the assumptions (D4) hold. Then for all \( \varepsilon \in (0, \varepsilon_0] \), \( h > 0 \) there exists a unique solution \( \{u^\varepsilon_{k,\varepsilon}, v^\varepsilon_{k,\varepsilon}\} \) to \( (P^\varepsilon_{k,\varepsilon}) \) and \( |u^\varepsilon_{k,\varepsilon}|_{L^\infty(Q_T)} \| v^\varepsilon_{k,\varepsilon}\|_{L^\infty(Q_T)} \leq C(h, k) \).

**Proof.** See Theorem 3.1 in part 1. \( \square \)

**Lemma 3.1.** Let the assumptions (D4) hold. Then we have for all \( \varepsilon \in (0, \varepsilon_0] \), \( h > 0 \) and \( t \in (0, T] \) that
\[
\begin{align*}
\nabla u^\varepsilon_{k,\varepsilon}(\cdot, t) &\in L^2(\Omega) + |\partial_t u^\varepsilon_{k,\varepsilon}(\cdot, t)|^2_{L^2(Q_T)} + \varepsilon |\partial_x v^\varepsilon_{k,\varepsilon}(\cdot, t)|^2_{L^2(Q_T)} + \varepsilon |\partial_t \pi_h[\varphi_\varepsilon(u^\varepsilon_{k,\varepsilon})]|^2_{L^2(Q_T)} + k^{-1} \\
&+ k^{-1} \left[ |\partial_t u^\varepsilon_{k,\varepsilon}(\cdot, t)|^2_{L^2(\Omega)} + \varepsilon |\partial_t v^\varepsilon_{k,\varepsilon}(\cdot, t)|^2_{L^2(\Omega)} + |\nabla \partial_t u^\varepsilon_{k,\varepsilon}(\cdot, t)|^2_{L^2(Q_T)} \right] \leq C.
\end{align*}
\]

**Proof.** Noting the bounds on \( g^\varepsilon \) above, see Lemma 3.1 in part 1. \( \square \)

In order to analyse the approximation \( (P^h_{k,\varepsilon}) \) it is convenient to introduce an associated linear problem of \( (P^h_{k,\varepsilon}) \):
\[
(P^\pm_{k,\varepsilon}) \quad \text{Find } u^{\pm}_{k,\varepsilon} \in H^1(0, T; S^h_0) \text{ and } v^{\pm}_{k,\varepsilon} \in H^1(0, T; S^h) \text{ such that}
\]
\[
\begin{align*}
&\langle \partial_t u^{\pm}_{k,\varepsilon} + \partial_x v^{\pm}_{k,\varepsilon}, \chi \rangle + \langle \nabla u^{\pm}_{k,\varepsilon}, \nabla \chi \rangle = \langle f, \chi \rangle \quad \forall \chi \in S^h_0 \\
&\langle \partial_t v^{\pm}_{k,\varepsilon}, \chi \rangle = k \left< \varphi_\varepsilon(u^{\pm}_{k,\varepsilon}) - v^{\pm}_{k,\varepsilon}, \chi \right> \quad \forall \chi \in S^h
\end{align*}
\]
\[
\begin{align*}
u^{\pm}_{k,\varepsilon}(\cdot, 0) &= P^h_0 g_\varepsilon(\cdot) \\
v^{\pm}_{k,\varepsilon}(\cdot, 0) &= P^h_0[\varphi_\varepsilon(g_\varepsilon(\cdot))].
\end{align*}
\]
The existence and uniqueness of \( \{v_{h,\bar{v}}^{h,\bar{v}}(\cdot), h_{k,\bar{k}}^{h,\bar{h}}(\cdot)\} \) solving \((P_{k,\bar{k}}^{h,\bar{h}})\) for all \( \varepsilon \in (0, \varepsilon_0) \) and \( h > 0 \) is easily established under assumptions (D3) and we have the following result.

**Lemma 3.2.** Let the assumptions (D3) hold. Then we have for all \( \varepsilon \in (0, \varepsilon_0) \), \( h > 0 \) and \( t \in (0, T) \) that

\[
\begin{align*}
&\|u_{k,\bar{v}}^{h,\bar{h}}(\cdot) - u_{k,\bar{v}}^{h,\bar{h}}(\cdot)\|_{L^2(\Omega)}^2 + h^2 \|\nabla \int_0^t (u_{k,\bar{v}}^{h,\bar{h}}(\cdot), s) ds\|_{L^2(\Omega)}^2 \\
&\quad \leq C h^4 \left[ \|u_{k,\bar{v}}^{h,\bar{h}}(\cdot)\|_{L^2(0,T;H^2(\Omega))}^2 + \|u_{k,\bar{v}}^{h,\bar{h}}(\cdot)\|_{H^2(\Omega)}^2 \right] \leq C \varepsilon^{-1} h^4.
\end{align*}
\]  

**Proof.** The first inequality in (3.5) is proved in Lemma 3.2 of part 1. Under the stated assumptions on \( \Omega \) we have that

\[
\|u_{k,\bar{v}}^{h,\bar{h}}(\cdot)\|_{L^2(0,T;H^2(\Omega))} \leq C \left[ \|\partial_t u_{k,\bar{v}}^{h,\bar{h}}(\cdot)\|_{L^2(\Omega)} + \|\partial_t v_{k,\bar{v}}^{h,\bar{h}}(\cdot)\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} \right] \leq C \varepsilon^{-1/2},
\]

where we have noted (2.7). Hence the desired result (3.5). \( \Box \)

Assuming (D5), it is easy to deduce that the stiffness matrix \( \{(\nabla \chi_i, \nabla \chi_j)\}_{i,j=1}^n \) where \( \{\chi_i\}_{i=1}^n \) are the internal nodes of the partitioning and \( \chi_i \in S^h_0 \) is such that \( \chi_i(x_i) = \delta_{ij}, \ 1 \leq i, j \leq n \); is an M-matrix. From this property one can deduce that

\[
M^{-1} \varepsilon \|\nabla \pi_h[\varphi_\varepsilon(\chi)]\|_{L^2(\Omega)}^2 \leq \langle \nabla \chi, \nabla \pi_h[\varphi_\varepsilon(\chi)] \rangle \quad \forall \chi \in S^h_0,
\]

see e.g. §2.4.2 of [10]. Furthermore it follows from (3.7), (3.1a), (2.3a) and (3.2) that for all \( w \in H^2(\Omega) \cap H_0^2(\Omega) \)

\[
M^{-1} \varepsilon \|\nabla \pi_h[\varphi_\varepsilon(w)]\|_{L^2(\Omega)}^2 \leq \langle \nabla w, \nabla \varphi_\varepsilon(w) \rangle + C \varepsilon^{-1} h^2 \|w\|_{H^2(\Omega)}^2,
\]

see the derivation of (3.12) in part 1 for details.

**Corollary 3.1.** Let the assumptions (D5) hold. Then the unique solution \( \{v_{k,\bar{v}}^{h,\bar{h}}(\cdot), \ v_{k,\bar{v}}^{h,\bar{h}}(\cdot)\} \) to \((P_{k,\bar{k}}^{h,\bar{h}})\), \( \varepsilon \in (0, \varepsilon_0) \) and \( h > 0 \), satisfies the bounds (2.5) with \( y, \overline{u}, \overline{v} \) and \( \overline{v} \in C(\Omega) \) all independent of \( \varepsilon, k \) and \( h \). In particular, if \( g_\varepsilon \) and \( f \geq 0 \) then \( \tilde{v}_{k,\bar{v}}^{h,\bar{h}} \geq 0 \) in \( \Omega \).

**Proof.** See Corollary 3.1 in part 1 and compare the proof of Theorem 1.1 for a justification of the time independence of the bounds. \( \Box \)

**Lemma 3.3.** Under assumptions (D5) we have for all \( \varepsilon \in (0, \varepsilon_0) \) and \( h > 0 \), provided \( \varepsilon^{-1} kh^2 \leq C \), and for all \( t \in (0, T) \) that

\[
\begin{align*}
&\|u_{k,\bar{v}}^{h,\bar{h}}(\cdot) - u_{k,\bar{v}}^{h,\bar{h}}(\cdot)\|_{E_{k,\bar{k}}(\varepsilon)}^2 + \varepsilon \|\pi_h[\varphi_\varepsilon(u_{k,\bar{v}}^{h,\bar{h}}(\cdot)) - \varphi_\varepsilon(u_{k,\bar{v}}^{h,\bar{h}}(\cdot))]|_{L^2(\Omega)}^2 \\
&\quad \leq C \varepsilon^{-1} \|g_\varepsilon\|_{H^1(\Omega)}^2 + C \|P_{k,\bar{k}}^{h,\bar{h}} g_\varepsilon - g_\varepsilon\|_{L^2(\Omega)}^2 + C \|P_{k,\bar{k}}^{h,\bar{h}} g_\varepsilon - g_\varepsilon\|_{L^2(\Omega)}^2 + C \|\pi_h[\varphi_\varepsilon(u_{k,\bar{v}}^{h,\bar{h}}(\cdot)) - \varphi_\varepsilon(u_{k,\bar{v}}^{h,\bar{h}}(\cdot))]|_{L^2(\Omega)}^2 \\
&\quad \leq \sup_{\chi \in S^h_0} \|\varphi_\varepsilon(u_{k,\bar{v}}^{h,\bar{h}}(\cdot), \chi) - \langle \varphi_\varepsilon(u_{k,\bar{v}}^{h,\bar{h}}(\cdot), \chi) \rangle \|_{L^2(\Omega)}^2.
\end{align*}
\]

**Proof.** See Lemma 3.3 of part 1. \( \Box \)

We now improve on the bound (3.9) in the physically interesting case of given data \( g \) and \( f \geq 0 \) yielding \( u \geq 0 \) in \( Q_T \). Assuming (D6), we set \( \varphi_\varepsilon \) to be the following quadratic regularization of \( \varphi \)

\[
\varphi_\varepsilon(s) = \begin{cases} \ s^2 + bs & \text{for } s \in [0, \delta] \\ \varphi(s) & \text{otherwise} \end{cases}
\]

\[
\text{for } s \in \mathbb{R}.
\]
where \( a \equiv \delta^{-1} \varphi'(\delta) - \delta^{-2} \varphi(\delta) \), \( b \equiv -\varphi'(\delta) + 2\delta^{-1} \varphi(\delta) \) and \( \delta \equiv \varepsilon^{1/(1-p)} \) so that \( \varphi_\varepsilon \in C^1[0,\infty) \). From (D6)(iii) it follows that \( \varphi(\delta) \geq \delta \varphi'(\delta) \), which in turn yields that \( \beta \leq \delta \leq \varepsilon^{1/(1-p)} \) so that \( \varepsilon \leq \alpha \leq (C_1 \varepsilon^s)^{1/(1-p)} \leq \alpha \leq 0 \), see (2.3b), and hence \( \varphi_\varepsilon \) satisfies the conditions (2.2). Therefore all the results proved so far in this paper hold under the assumptions (D6). We note for example that \( \varphi(s) \equiv \alpha s^p \) for \( s \geq 0 \) with \( p \in (0,1) \) and \( x \in \mathbb{R}^+ \) satisfies (1.9) and (D6)(iii).

Assuming (D6), it follows for all \( w \in H^1(\Omega) \cap W^{2,1}(\Omega) \) with \( w(x) \in [0,m] \) for \( x \in \Omega \) that

\[
|\varphi_\varepsilon(w)|_{W^{2,1}(\Omega)} \leq C\varepsilon^{-1}||w||_{W^{2,1}(\Omega)},
\]

see (3.25) in part 1. From (D6)(i) we have the discrete Sobolev imbedding

\[
|\chi|_{L^{\infty}(\Omega)} \leq C[\ln(1/h)]^r ||\chi||_{H^1(\Omega)} \leq C[\ln(1/h)]^r ||\nabla \chi||_{L^2(\Omega)}, \quad \forall \chi \in S^h \theta^r,
\]

where \( r = 0 \) if \( d = 1 \) and \( r = 1/2 \) if \( d = 2 \); see for example p. 67 in [14]. As noted in part 1, the quasi-uniformity restriction is not really restrictive in practice.

**Lemma 3.4.** Under assumptions (D6) we have for all \( \varepsilon \in (0,\varepsilon_0] \) and \( h > 0 \), provided \( \varepsilon^{-1} kh^2 \leq C \), and for all \( t \in (0,T] \) that

\[
\begin{align*}
|u_{\varepsilon,k}^h - \hat{u}_k^h|_{L^2(\Omega)}^2 & + \varepsilon |\pi_h[\varphi_\varepsilon(u_{\varepsilon,k}^h) - \varphi_\varepsilon(\hat{u}_k^h)]|_{L^2(\Omega)}^2 \\
& \leq C kh^2[\varepsilon^{-2}[\ln(1/h)]^r]|u_{\varepsilon,k}^h|_{L^2(0,T;W^{2,1}(\Omega))}^2 + ||g^h||_{H^1(\Omega)}^2 + \\
& + C \sup_{\chi \in S^h \theta^r} \left\{ |\langle \varphi_\varepsilon(g_t), \chi \rangle - \langle \varphi_\varepsilon(\hat{u}_k^h), \chi \rangle| / ||\chi||_{H^1(\Omega)} \right\}^2.
\end{align*}
\]

**Proof.** See Lemma 3.4 of part 1.

One can also consider the corresponding approximation without relaxing the reaction

**Theorem 3.2.** Let the assumptions (D4) hold. Then for all \( \varepsilon \in (0,\varepsilon_0] \) and \( h > 0 \) there exists a unique solution \( \hat{u}_k^h \) to (\( P_{\varepsilon,k}^h \)). Moreover, for all \( t \in (0,T] \) we have that

\[
|\hat{u}_k^h - \hat{u}_k^h|_{L^2(\Omega)}^2 + \varepsilon |\pi_h[\varphi_\varepsilon(\hat{u}_k^h) - \varphi_\varepsilon(\hat{u}_k^h)]|_{L^2(\Omega)}^2 + \varepsilon |\pi_h[\varphi_\varepsilon(\hat{u}_k^h)] - \hat{u}_k^h|_{L^2(\Omega)}^2 \\
\leq C \varepsilon^{-2} |\hat{u}_k^h|_{L^2(Q_T)}^2 \leq C \varepsilon^{-1} k^{-2}.
\]

In addition under the assumptions (D5) \( \hat{u}_k^h \) satisfies the bounds (2.8a). In particular, if \( g^h \) and \( f \geq 0 \) then \( \hat{u}_k^h \geq 0 \) in \( Q_T \).

**Proof.** Existence and uniqueness of a solution to (\( P_{\varepsilon,k}^h \)) follows from a discrete analogue of Theorem 2.2. The first inequality in (3.14) is a discrete analogue of the first inequality in (2.8b). The second inequality in (3.14) follows from (3.4). The bounds (2.8a) follow from (3.14), the equivalence of norms on \( S^h \) and Corollary 3.1. \( \square \)
Lemma 3.5. Let the assumptions (D4) hold. Then for all \( \varepsilon \in (0, \varepsilon_0) \) and \( h > 0 \)
let \( \bar{\varepsilon} \in (0, \varepsilon] \) and \( \bar{h} \in (0, h] \) be such that

\[
(3.15a) \quad \left[ m(\Omega_{e,\bar{\varepsilon}})^2 \right] / (1-p) \leq C A_{\varepsilon}(T) \varepsilon^{(1+\mu)(1-p)},
\]
\[
(3.15b) \quad h \leq C (\bar{\varepsilon}/\varepsilon)^{1/2} \bar{h}.
\]

where \( \mu = 1 \) if \( d = 1, \mu > 1 \) if \( d = 2 \) and \( \mu = 6/5 \) if \( d = 3 \). Let \( T^h \) be a subpartitioning
of \( T \) consisting of regular simplices \( \kappa \) with \( h_\kappa \equiv \text{diam}(\kappa) \) and \( h \equiv \max_{\kappa \in T_h} h_\kappa \).
such that \( \Omega^h \equiv \cup_{\kappa \in T_h} \kappa \equiv \cup_{\kappa \in T_h} \overline{\kappa} \) and \( \overline{\kappa} = \cup \overline{\kappa} \) for all \( \kappa \in T^h \). Let \( S^h \) be the
associated continuous piecewise linear space, \( \pi_\kappa : C^0(\overline{\Omega^h}) \to S^h \) be the corresponding
interpolation operator and \( \Omega_{e,\bar{\varepsilon}} \equiv \cup_{\kappa \in T^h} \kappa \), where \( T^h = \{ \kappa \in T^h : \kappa \cap \Omega_{e,\bar{\varepsilon}} \neq \emptyset \} \).

Then we have on choosing \( g^h \in S^h \) to be the unique solution of

\[
(3.16a) \quad \langle \nabla g^h, \nabla \chi \rangle + \langle \varphi_\varepsilon(g^h), \chi \rangle^h = \langle \nabla \pi_\kappa g, \nabla \chi \rangle + \langle \varphi_\varepsilon(g), \chi \rangle^h + \int_{\Omega_{e,\bar{\varepsilon}}} \pi_\kappa[\varphi_\varepsilon(g) - \varphi_\varepsilon(g)] \chi \quad \forall \chi \in S^h,
\]

that \( \langle \nabla g^h, \nabla \chi \rangle \leq C |\chi|_{L^2(\Omega)} \forall \chi \in S^h \) as required in (P_{e,\varepsilon}). In addition if assumption
(D5) holds and \( \varphi_\varepsilon(s) \geq \varphi_\varepsilon(s) \) for all \( s \in (0, e^{1/(1-p)}) \) then

\[
(3.16b) \quad g^h \geq \pi_\kappa g \quad \text{in} \quad \overline{\Omega}
\]

and hence if \( g, f \geq 0 \) then \( \varphi_\varepsilon^h, \psi_\varepsilon^h, \psi_\varepsilon^h \geq 0 \) in \( Q_T \).

Furthermore, provided \( \varepsilon^{-1} h^2 < C \), we have that

\[
|P^l_{g^h} g^h - g^h_{L^2(\Omega)}| + \sup_{\chi \in S^h} \left\{ \left| \langle \varphi_\varepsilon(g^h), \chi \rangle - \langle \varphi_\varepsilon(g^h), \chi \rangle \right|^h \right\} / \| \chi \|_{H^1(\Omega)}^2 \leq C A_{\varepsilon}(T) \varepsilon^{(1+\mu)(1-p)} + \frac{C \varepsilon^{-1} h^2}{C \varepsilon^{-2} h^4 \ln(1/h)} \| \chi \|_{H^1(\Omega)}^2 \quad \text{if (D5)}
\]

\[
\quad \text{or (D6) holds for both} \ T^h \text{and} \ T^h.
\]

Proof. Choosing \( \chi \equiv g^h \) in (3.16a) yields that \( \| g^h \|_{H^1(\Omega)} \leq C, g \in H^2(\Omega) \), and hence \( \| \varphi_\varepsilon(g^h) \|_{L^2(\Omega)} \leq C \). By noting that \( |\nabla \pi_\kappa g, \nabla \chi| \leq |g|_{H^2(\Omega)} |\chi|_{H^1(\Omega)} \), see e.g. (3.4c)
in part 1, it follows that \( |\nabla g^h, \nabla \chi| \leq |\chi|_{L^2(\Omega)} \forall \chi \in S^h \). The result (3.16b) is a discrete analogue of (2.14b) and follows from choosing \( \chi \equiv \sum_{j=1}^l |g(x_j) - g^h(x_j)|_1 x_j \) in (3.16a) and noting that the stiffness matrix is an M-matrix. (3.16b) and Corollary
3.1 yield that if \( g, f \geq 0 \) then \( d_\varepsilon, \varepsilon_\varepsilon, \psi_\varepsilon \geq 0 \) in \( Q_T \).

From (2.13) and (3.16a) we have that

\[
\langle \nabla (P^l_{g^h} g - g^h), \nabla \chi \rangle + \frac{1}{2} \left\{ \langle \varphi_\varepsilon(g), \chi \rangle - \langle \varphi_\varepsilon(g), \chi \rangle^h \right\} = \frac{1}{2} \left\{ \langle \varphi_\varepsilon(g), \chi \rangle - \langle \varphi_\varepsilon(g), \chi \rangle^h \right\} + \int_{\Omega_{e,\bar{\varepsilon}}} (I - \pi_\kappa)[\varphi_\varepsilon(g) - \varphi_\varepsilon(g)] \chi \quad \forall \chi \in S^h,
\]

that is

\[
\langle \nabla (P^l_{g^h} g - g^h), \nabla \chi \rangle + \langle \varphi_\varepsilon(g) - \varphi_\varepsilon(g^h), \chi \rangle^h = \langle \varphi_\varepsilon(g) - \varphi_\varepsilon(g), \chi \rangle + \langle \varphi_\varepsilon(g), \chi \rangle^h
\]
\[ + \left[ \varphi_\varepsilon(g), \chi - \langle \varphi_\varepsilon(g), \chi \rangle^h \right] - \left[ \langle \varphi_\varepsilon(g_\varepsilon), \chi \rangle - \langle \varphi_\varepsilon(g\varepsilon), \chi \rangle^h \right] + \]

\[ \int_{\Omega} (I - \pi_h)[\varphi_\varepsilon(g) - \varphi_\varepsilon(g)] \chi \quad \forall \chi \in S_h^0; \]

where \( d_\varepsilon \equiv g_\varepsilon - g \in H^2(\Omega) \), noting Corollary 2.3, and \( d_h^0 \equiv g_h^0 - \pi_h g \in S_h^0 \).

Similarly to (2.11a) it follows that

\[ |\langle \varphi_\varepsilon(g) - \varphi_\varepsilon(g), \chi \rangle| \leq C|\varphi_\varepsilon(g) - \varphi_\varepsilon(g)|_{H^{-1}(\Omega)}||\chi||_{H^1(\Omega)}; \]

\[ \leq C\left( \frac{\left| \varphi_\varepsilon(g) \right|_{H^{-1}(\Omega)}}{m(\Omega, \varepsilon)} \right)^{1/\varepsilon} \left| \varepsilon \right|_{H^1(\Omega)}. \]

Assuming (D5) it follows from (3.2), (3.1a) and (2.2c) that for all \( w \in H^2(\Omega) \cap H_h^1(\Omega) \) and for \( \varepsilon^{-1}h^2 \leq C \)

\[ (I - \pi_h)\varphi_\varepsilon(w)_{L^2(\Omega)} \leq \| (I - \pi_h)\varphi_\varepsilon(\pi_h w) \|_{L^2(\Omega)} + |\varphi_\varepsilon(w) - \varphi_\varepsilon(\pi_h w)|_{L^2(\Omega)} \]

\[ \leq Ch\left| \nabla \varphi_\varepsilon(\pi_h w) \right|_{L^2(\Omega)} + C\varepsilon^{-1}h^2 |w|_{H^2(\Omega)}; \]

\[ \leq C\varepsilon^{-1/2}h[1 + ||w||_{H^2(\Omega)}], \]

since (3.8) implies that

\[ \varepsilon \left| \nabla \varphi_\varepsilon(\pi_h w) \right|_{L^2(\Omega)}^2 \leq -C(\Delta w, \varphi_\varepsilon(w)) + C\varepsilon^{-1}h^2 |w|_{H^2(\Omega)}; \]

\[ \leq C[1 + ||w||^2_{H^2(\Omega)}]. \]

Hence (3.20a), (3.20b) and (3.3b) yield that

\[ |\langle \varphi_\varepsilon(w), \chi \rangle - \langle \varphi_\varepsilon(w), \chi \rangle^h| \leq \| (I - \pi_h)\varphi_\varepsilon(w), \chi \| + |\pi_h \varphi_\varepsilon(w), \chi \rangle - \langle \varphi_\varepsilon(w), \chi \rangle^h | \]

\[ \leq C\varepsilon^{-1/2}h[1 + ||w||_{H^2(\Omega)}] ||\chi||_{H^1(\Omega)}; \]

As (D5) holds also for \( T^h \) and (3.15) yields that \( \varepsilon^{-1}h^2 \leq \varepsilon^{-1}h^2 \leq C\varepsilon^{-1}h^2 \), we have similarly to (3.20) that

\[ \int_{\Omega} (I - \pi_h)[\varphi_\varepsilon(g) - \varphi_\varepsilon(g)] \chi \leq C\varepsilon^{-1/2}h[1 + ||g||_{H^2(\Omega)}] ||\chi||_{L^2(\Omega)}; \]

Choosing \( \chi \equiv P_h^1 d_\varepsilon - d_h^0 \) in (3.18b) and noting (2.3a), (3.19), (3.21), (3.22), (3.15a), (3.15b) and (3.1c) yields that

\[ |P_h^1 d_\varepsilon - d_h^0|_{H^1(\Omega)} + C|\varphi_\varepsilon(g_\varepsilon) - \varphi_\varepsilon(g_h^0)|^2 \]

\[ \leq |P_h^1 d_\varepsilon - d_h^0|_{H^1(\Omega)} + |\varphi_\varepsilon(g_\varepsilon) - \varphi_\varepsilon(g_h^0), g_\varepsilon - g_h^0|^2 \]

\[ \leq |\varphi_\varepsilon(g_\varepsilon) - \varphi_\varepsilon(g_h^0), (I - P_h^1) d_\varepsilon|^2 + CA_h(T)\varepsilon(1+p)(1-p) + C\varepsilon^{-1}h^2 \]

\[ \leq CA_h(T)\varepsilon(1+p)(1-p) + C\varepsilon^{-1}h^2. \]

Under assumptions (D6) we have in place of (3.20a), after noting (3.1a) and (3.11), that

\[ |(I - \pi_h)\varphi_\varepsilon(w)|_{L^2(\Omega)} \leq Ch^2|\varphi_\varepsilon(w)|_{W^{2,1}(\Omega)} \leq C\varepsilon^{-1}h^2 ||w||_{W^{2,1}(\Omega)}. \]

Hence in place of (3.21), we have on noting (3.12), (3.3b) and (3.20b) that

\[ |\langle \varphi_\varepsilon(w), \chi \rangle - \langle \varphi_\varepsilon(w), \chi \rangle^h| \leq |(I - \pi_h)\varphi_\varepsilon(w), \chi \| + |\pi_h \varphi_\varepsilon(w), \chi \rangle - \langle \varphi_\varepsilon(w), \chi \rangle^h | \]

\[ \leq C\varepsilon^{-1}h^2 ||(1/h)\right|| ||w||_{H^2(\Omega)} ||\chi||_{H^1(\Omega)}. \]
Similarly, in place of (3.22) we have that

\[
\int_{\Omega_{\epsilon,h}} (I - \pi_{h})[\varphi_{\epsilon}(g) - \varphi_{\epsilon}(g)] \chi \leq C\varepsilon^{-1}h^{2}[\ln(1/h)^{2}] ||g||_{H^{2}(\Omega)} ||\chi||_{H^{1}(\Omega)}. \tag{3.26}
\]

Therefore, similarly to (3.23), it follows that

\[
P_{h}^{\epsilon}d_{\epsilon} - d_{\epsilon}^{h} \leq C\varepsilon[\varphi_{\epsilon}(g_{\epsilon}) - \varphi_{\epsilon}(g_{\epsilon}^{h})]_{h}^{2} \leq CA_{\epsilon}(T)\varepsilon^{(1+p)/(1-p)} + C\varepsilon^{-2}h^{4}||\ln(1/h)||^{2r}. \tag{3.27}
\]

Combining (3.18a), (3.19), (3.15), (3.21), (3.22), (3.23), (3.25), (3.26) and (3.27) yields the desired result (3.17).

\[\Box\]

**Remark 3.1.** Firstly we note the construction of \(g_{h}^{\epsilon}\) as in (3.16a) is fully practical. Furthermore in line with (2.11b) if \(m(\Omega_{\epsilon,h})^{2/\mu} \leq CA_{\epsilon}(T)\varepsilon\), then one can choose \(\varepsilon = \varepsilon_{0}\) and hence \(g_{h}^{\epsilon} \equiv \pi_{h}g\) in this case.

Secondly we note that (a) the linear regularization, (2.4), and (b) the quadratic regularization, (3.10), are such that \(\varphi(s) \geq \varphi_{\epsilon}(s) \geq \varphi_{\epsilon}(s)\) for all \(s \in (0, \varepsilon^{(1/(1-p))})\) and \(\varepsilon \in (0, \varepsilon_{0})\) if (a) \(\varphi''(s) < 0\) and (b) \(\varphi''(s) > 0\) for all \(s \in (0, \varepsilon^{(1/(1-p))})\), respectively. These hold for the model case \(\varphi(s) \equiv \alpha s^{p}\) for \(s \in (0, \varepsilon^{(1/(1-p))})\).

**Theorem 3.3.** On choosing \(g_{h}^{\epsilon}\) in \(P_{h}^{\epsilon}(T)\) defined by (3.16a); we have for all \(\varepsilon \in (0, \varepsilon_{0})\) and \(h > 0\), provided \(\varepsilon^{-1}kh^{2} \leq C\), and for all \(t \in (0, T)\):

(i) Under assumptions \((D5)\)

\[
|u - \bar{u}_{k,h}^{\epsilon}|_{L^{2}(\Omega)}^{2} + \varepsilon|\varphi(u) - \pi_{h}[\varphi_{\epsilon}(\bar{u}_{k,h}^{\epsilon})]|_{L^{2}(Q_{t})}^{2} \leq C[\varepsilon^{(1+p)/(1-p)}\{A_{\epsilon}(T) + m(\Omega_{\epsilon,h})\} + \varepsilon^{-1}(k^{-2} + h^{2})]. \tag{3.28}
\]

(ii) Under assumptions \((D6)\)

\[
|u - \bar{u}_{k,h}^{\epsilon}|_{L^{2}(Q_{t})}^{2} \leq E, \quad \int_{0}^{t} \left( u - \bar{u}_{k,h}^{\epsilon} \right)(\cdot, s) \left( u - \bar{u}_{k,h}^{\epsilon} \right)(\cdot, s) ds \leq \min\{E, \varepsilon^{-1}h^{2}\} \quad \text{and} \quad \varepsilon|\varphi(u) - \pi_{h}[\varphi_{\epsilon}(\bar{u}_{k,h}^{\epsilon})]|_{L^{2}(Q_{t})}^{2} \leq \min\{E, h^{2}\}; \tag{3.29a}
\]

where

\[
|u - \bar{u}_{k,h}^{\epsilon}|_{L^{2}(Q_{t})}^{2} \leq E, \quad \int_{0}^{t} \left( u - \bar{u}_{k,h}^{\epsilon} \right)(\cdot, s) \left( u - \bar{u}_{k,h}^{\epsilon} \right)(\cdot, s) ds \leq \min\{E, \varepsilon^{-1}h^{2}\} \quad \text{and} \quad \varepsilon|\varphi(u) - \pi_{h}[\varphi_{\epsilon}(\bar{u}_{k,h}^{\epsilon})]|_{L^{2}(Q_{t})}^{2} \leq \min\{E, h^{2}\}; \tag{3.29b}
\]

where

\[
E \equiv C[\varepsilon^{(1+p)/(1-p)}\{A_{\epsilon}(T) + m(\Omega_{\epsilon,h})\} + \varepsilon^{-1}k^{-2} + \varepsilon^{-3}kh^{4}||\ln(1/h)||^{2r}]. \tag{3.29c}
\]

\(r = 0\) if \(d = 1\) and \(r = 1/2\) if \(d = 2\).

**Proof.** It follows similarly to (3.20) that

\[
|\varphi(u_{k,h}) - \pi_{h}[\varphi_{\epsilon}(u_{k,h})]|_{L^{2}(Q_{t})}^{2} \leq Ck^{2}|\nabla[\pi_{h}\varphi_{\epsilon}(u_{k,h})]|_{L^{2}(Q_{t})}^{2} + C\varepsilon^{-2}h^{4}||u_{k,h}||_{L^{2}(0,T;H^{2}(\Omega))}^{2} \leq C\varepsilon^{-1}; \tag{3.30a}
\]

since (3.6), (2.7) and (3.8) yield that

\[
|u_{k,h}||_{L^{2}(0,T;H^{2}(\Omega))}^{2} \leq Ck ; \quad \text{and} \quad |\nabla[\pi_{h}\varphi_{\epsilon}(u_{k,h})]|_{L^{2}(Q_{t})}^{2} \leq C\varepsilon^{-1}. \tag{3.30b}
\]

The results (3.28) and (3.29) then follow immediately from combining (2.10), (2.14a), (2.8b), (3.5), (3.6), (3.17), (3.30a) with (3.9) and (3.13), respectively.

**Corollary 3.3a.** Let the assumptions \((D5)\) hold. Then for all \(h > 0\) and \(t \in (0, T)\):

\[\Box\]
(i) Under no assumptions on non-degeneracy and on choosing \( \varepsilon = C h^{1-p} \leq \varepsilon_0 \) and \( k = C h^{-1} \), we have that

\[
\|u - \bar{u}_{k,\varepsilon}^h\|_{L^2(\Omega)} + \left\| \int_0^t (u - \bar{u}_{k,\varepsilon}^h)(\cdot, s) \, ds \right\|_{H^1(\Omega)} \leq C h^{(1+p)/2}
\]

and

\[
\|\varphi (u) - \pi_h[\varphi_x(\bar{u}_{k,\varepsilon}^h)]\|_{L^2(\Omega)} \leq C h^p.
\]

(ii) Assuming (N.D.) and that \( m(\Omega_{q,\varepsilon}) \leq C\varepsilon \), then on choosing \( \varepsilon = C h^{2(1-p)/(3-p)} \leq \varepsilon_0 \) and \( k = C h^{-1} \), we have that

\[
\|u - \bar{u}_{k,\varepsilon}^h\|_{L^2(\Omega)} + \left\| \int_0^t (u - \bar{u}_{k,\varepsilon}^h)(\cdot, s) \, ds \right\|_{H^1(\Omega)} \leq C h^{2/(3-p)}
\]

and

\[
\|\varphi (u) - \pi_h[\varphi_x(\bar{u}_{k,\varepsilon}^h)]\|_{L^2(\Omega)} \leq C h^{(1+p)/(3-p)}.
\]

Proof. The results follow directly from (3.28).

**Corollary 3.3b.** Let the assumptions (D6). Then for all \( h > 0 \) and \( t \in (0, T) \):

(i) Under no assumptions on non-degeneracy and on choosing \( \varepsilon = C\{h^{2}[\ln(1/h)]^r\}^{(1-p)/(5-2p)} \leq \varepsilon_0 \) and \( k = C\{h^{2}[\ln(1/h)]^r\}^{-2/(5-2p)} \) we have for all \( p \in (1/2, 1) \) that

\[
\|u - \bar{u}_{k,\varepsilon}^h\|_{L^2(\Omega)} \leq C\{h^{2}[\ln(1/h)]^r\}^{(1+p)/(5-2p)},
\]

\[
\left\| \int_0^t (u - \bar{u}_{k,\varepsilon}^h)(\cdot, s) \, ds \right\|_{H^1(\Omega)} \leq C\{h^{2}[\ln(1/h)]^r\}^{3/(5-2p)}
\]

and

\[
\|\varphi (u) - \pi_h[\varphi_x(\bar{u}_{k,\varepsilon}^h)]\|_{L^2(\Omega)} \leq C\{h^{2}[\ln(1/h)]^r\}^{q/(5-2p)}.
\]

where \( q = \min\{2p, 3/2\} \).

(ii) Assuming (N.D.) and that \( m(\Omega_{q,\varepsilon}) \leq C\varepsilon \), then on choosing \( \varepsilon = C\{h^{2}[\ln(1/h)]^r\}^{4(1-p)/(13-7p)} \leq \varepsilon_0 \) and \( k = C\{h^{2}[\ln(1/h)]^r\}^{-2(3-p)/(13-7p)} \), we have for all \( p \in (1/3, 1) \) that

\[
\|u - \bar{u}_{k,\varepsilon}^h\|_{L^2(\Omega)} \leq C\{h^{2}[\ln(1/h)]^r\}^{4/(13-7p)},
\]

\[
\left\| \int_0^t (u - \bar{u}_{k,\varepsilon}^h)(\cdot, s) \, ds \right\|_{H^1(\Omega)} \leq C\{h^{2}[\ln(1/h)]^r\}^{3(3-p)/(2(13-7p))}
\]

and

\[
\|\varphi (u) - \pi_h[\varphi_x(\bar{u}_{k,\varepsilon}^h)]\|_{L^2(\Omega)} \leq C\{h^{2}[\ln(1/h)]^r\}^{q/(13-7p)},
\]

where \( q = \min\{2(p+1), 3(3-p)/2\} \).

Proof. The results follow directly from (3.29).

We note that (3.33a) & (3.33c) improve on (3.31a) & (3.31b), and (3.34a) & (3.34c) improve on (3.32a) & (3.32b).
Theorem 3.4. For all \( h > 0 \) and \( t \in (0, T) \) the error bounds (3.31)–(3.34) hold under the stated assumptions and choices of \( \varepsilon \) with \( \hat{u}_{k,\varepsilon}^h \) replaced by \( \hat{u}_{k,\varepsilon}^h \).

Proof. The above error bounds follow by combining (3.14) with (3.31)–(3.34). \( \Box \)

Remark 3.2. Of course the above analysis simplifies for \( \{u_{k,\varepsilon}^n, v_{k,\varepsilon}^n\} \) and \( \bar{u}_k^h \), the unique solutions of the less practical schemes (P\(_{k,\varepsilon}^h\)) and (P\(_{k,\varepsilon}^h\)) respectively. (P\(_{k,\varepsilon}^h\)) being the same as (\( \bar{P}_{k,\varepsilon}^h \)), but with all the required integrals performed exactly. In addition one can improve on the error bounds above, see Remark 3.1 in part 1.

4. A fully discrete and practical finite element approximation. In this section we analyse the fully discrete practical approximation to (P\(_{k,\varepsilon}^h\)) with timestep \( \tau = T/N \):

\[
(P_{k,\varepsilon}^{h,\tau}) \text{ For } n = 1 \rightarrow N \text{ find } u_{k,\varepsilon}^{h,n} \in S_0^h \text{ and } v_{k,\varepsilon}^{h,n} \in S^h \text{ such that }
\]

\[
\tau^{-1} \left( \langle \hat{u}_{k,\varepsilon}^{h,n} - u_{k,\varepsilon}^{h,n-1}, \chi \rangle + \langle \nabla \hat{u}_{k,\varepsilon}^{h,n}, \nabla \chi \rangle \right) = \langle f^n, \chi \rangle^h \quad \forall \chi \in S_0^h
\]

\[
\tau^{-1} \left( \langle \hat{v}_{k,\varepsilon}^{h,n} - v_{k,\varepsilon}^{h,n-1}, \chi \rangle + k \langle \varphi_e(\hat{v}_{k,\varepsilon}^{h,n}) - \hat{v}_{k,\varepsilon}^{h,n}, \chi \rangle^h \right) = \langle \bar{f}_e, \chi \rangle^h \quad \forall \chi \in S^h
\]

\[
\hat{u}_{k,\varepsilon}^{h,0}(\cdot) = g_e^h(\cdot) \quad \hat{v}_{k,\varepsilon}^{h,0}(\cdot) = \pi_h(\varphi_e(g_e^h(\cdot)))
\]

where \( f^n(\cdot) \equiv f(\cdot, n\tau) \).

Let \( \hat{U}_{k,\varepsilon} \in L^\infty(0, T; S_0^h) \) and \( \hat{V}_{k,\varepsilon} \in L^\infty(0, T; S^h) \) be such that for \( n = 1 \rightarrow N \)

\[
\hat{u}_{k,\varepsilon}(\cdot, t) \equiv u_{k,\varepsilon}^{h,n}(\cdot) \quad \text{ and } \quad \hat{v}_{k,\varepsilon}(\cdot, t) \equiv v_{k,\varepsilon}^{h,n}(\cdot) \quad \text{ if } t \in ([n-1]T, n\tau];
\]

and \( U_{k,\varepsilon}^L \in C^0([0, T]; S_0^h) \) and \( V_{k,\varepsilon}^L \in C^0([0, T]; S^h) \) be such that for \( n = 1 \rightarrow N \)

\[
U_{k,\varepsilon}^L(\cdot, t) \equiv [(t - (n-1)\tau)u_{k,\varepsilon}^{h,n}(\cdot) + (n\tau - t)\hat{u}_{k,\varepsilon}^{h,n-1}(\cdot)]/\tau \quad \text{ if } t \in ([n-1]T, n\tau]
\]

and

\[
V_{k,\varepsilon}^L(\cdot, t) \equiv [(t - (n-1)\tau)v_{k,\varepsilon}^{h,n}(\cdot) + (n\tau - t)\hat{v}_{k,\varepsilon}^{h,n-1}(\cdot)]/\tau \quad \text{ if } t \in ([n-1]T, n\tau]
\]

Then \( \hat{P}_{k,\varepsilon}^{h,\tau} \) can be restated: for almost every \( t \in (0, T) \):

\[
\left( \langle \partial_t \hat{U}_{k,\varepsilon} + \partial_x \hat{V}_{k,\varepsilon}, \chi \rangle^h + \langle \nabla \hat{U}_{k,\varepsilon}, \nabla \chi \rangle \right) = \langle \hat{f}, \chi \rangle^h \quad \forall \chi \in S_0^h
\]

\[
\left( \langle \partial_t \hat{V}_{k,\varepsilon}, \chi \rangle^h + k \langle \varphi_e(\hat{V}_{k,\varepsilon}) - \hat{V}_{k,\varepsilon}, \chi \rangle^h \right) = \langle \bar{f}_e, \chi \rangle^h \quad \forall \chi \in S^h
\]

\[
\hat{U}_{k,\varepsilon}^L(\cdot, 0) = g_e^h(\cdot) \quad \hat{V}_{k,\varepsilon}^L(\cdot, 0) = \pi_h(\varphi_e(g_e^h(\cdot)))
\]

where \( \hat{f}(\cdot, t) \equiv f^n(\cdot) \equiv f(\cdot, n\tau) \) if \( t \in ([n-1]T, n\tau], n = 1 \rightarrow N \).

Theorem 4.1. Let the assumptions (D4) hold. Then for all \( \varepsilon \in (0, \varepsilon_0], h \) and \( \tau > 0 \) there exists a unique solution \( \{\hat{U}_{k,\varepsilon}, \hat{V}_{k,\varepsilon}\} \) to \( \hat{P}_{k,\varepsilon}^{h,\tau} \). Moreover, if the assumptions (D5) hold then

\[
(4.1) \quad \underline{U} \leq \hat{U}_{k,\varepsilon} \leq U \quad \text{ and } \quad \underline{V} \leq \hat{V}_{k,\varepsilon} \leq V \quad \text{ in } Q_T,
\]

where \( \underline{U}, U, \underline{V} \) and \( V \in \mathbb{R} \) are independent of \( h, \tau, \varepsilon, \) and \( k \). In particular, if \( g_e^h \) and \( f \geq 0 \) then one can take \( \underline{U} = \underline{V} = 0 \).

Proof. See Theorem 4.1 in part 1. \( \Box \)
Lemma 4.1. Under the assumptions (D4) we have for all $\varepsilon \in (0, \varepsilon_0]$, $h$, $\tau > 0$ and $m = 0 \to N$ that

$$
|\hat{u}^h_{k,\varepsilon} - \bar{U}_{k,\varepsilon}|^2_{E_2(k,m\tau)} + \varepsilon \|\pi_h[\varphi(\hat{u}^h_{k,\varepsilon}) - \varphi(\bar{U}_{k,\varepsilon})]|^2_{L^2(Q_m\tau)} + \varepsilon \|\hat{v}^h_{k,\varepsilon} - \bar{V}_{k,\varepsilon}|^2_{E_1(k,m\tau)}
\leq C \tau^2 \left( |\partial_t \hat{u}^h_{k,\varepsilon}|^2_{L^2(Q\tau)} + (\tau + k^{-1})^{-1} \|
abla(\partial_t \hat{u}^h_{k,\varepsilon})|^2_{L^2(Q\tau)} + |\partial_t \pi_h[\varphi(\hat{u}^h_{k,\varepsilon})]|^2_{L^2(Q\tau)} + |\partial_t \hat{v}^h_{k,\varepsilon}|^2_{L^2(Q\tau)} + |\partial_t [\pi_h f]|^2_{L^2(Q\tau)} \right).
$$

Proof. See Lemma 4.1 in part 1.

Corollary 4.1. Under the assumptions (D4) we have for all $\varepsilon \in (0, \varepsilon_0]$, $h$, $\tau > 0$ and $m = 0 \to N$ that

$$
|\hat{u}^h_{k,\varepsilon} - \bar{U}_{k,\varepsilon}|^2_{E_2(k,m\tau)} + \varepsilon \|\pi_h[\varphi(\hat{u}^h_{k,\varepsilon}) - \varphi(\bar{U}_{k,\varepsilon})]|^2_{L^2(Q_m\tau)} + \varepsilon \|\hat{v}^h_{k,\varepsilon} - \bar{V}_{k,\varepsilon}|^2_{E_1(k,m\tau)}
\leq C \varepsilon^{-1} + (\tau + k^{-1})^{-1} k \tau^2.
$$

(4.2)

Proof. The result follows immediately from Lemma 4.1 and (3.4). \qed

Below we will present an alternative bound to (4.2). Firstly, we prove an analogue of Lemma 3.1.

Lemma 4.2. Under assumptions (D4) we have for all $\varepsilon \in (0, \varepsilon_0]$, $h$, $\tau > 0$ and $t \in (0, T]$ that

$$
\nabla \bar{U}_{k,\varepsilon}(\cdot, t)|^2_{L^2(\Omega)} + |\partial_t \bar{U}_{k,\varepsilon}|^2_{L^2(Q\tau)} + \varepsilon |\partial_t \hat{V}_{k,\varepsilon}|^2_{L^2(Q\tau)} + \varepsilon \tau \sum_{n=1}^\infty |\pi_h[\varphi(\bar{U}_{k,\varepsilon}(\cdot, n\tau)) - \varphi(\bar{U}_{k,\varepsilon}(\cdot, (n-1)\tau))]|^2_{L^2(\Omega)}
\leq C(1 + (k\tau)^{-1}).
$$

(4.3)

Proof. We adapt the proof given for Lemma 2.2 in part 1. We adopt the difference notation $D_t^+ \hat{u}^{h,n}_{k,\varepsilon} = (\hat{u}^{h,n+1}_{k,\varepsilon} - \hat{u}^{h,n}_{k,\varepsilon})/\tau$, $D_t^- \hat{u}^{h,n}_{k,\varepsilon} = (\hat{u}^{h,n}_{k,\varepsilon} - \hat{u}^{h,n-1}_{k,\varepsilon})/\tau$ and $\delta_t^\pm D_t^\pm$ and note that

$$
\sum_{n=1}^m [(a^n - a^{n-1}) a^n] = \frac{1}{2}[(a^n)² - (a^n)^2] + \sum_{n=1}^m (a^n - a^{n-1})².
$$

(4.4)

Subtracting successive equations in $(\hat{P}_{k,\varepsilon}^h)$ yields for $n = 1 \to N - 1$

$$
\langle \delta_t^2 \hat{u}^{h,n}_{k,\varepsilon} + \hat{v}^{h,n}_{k,\varepsilon}, \lambda \rangle^h + \langle \nabla D_t^+ \hat{u}^{h,n}_{k,\varepsilon}, \nabla \lambda \rangle^h = \langle D_t^+ f^n, \lambda \rangle^h \quad \forall \lambda \in S^h
$$

(4.5a)

$$
\langle \delta_t^2 \hat{u}^{h,n}_{k,\varepsilon}, \lambda \rangle^h = k \langle D_t^+ [\varphi(\hat{u}^{h,n}_{k,\varepsilon}) - \hat{v}^{h,n}_{k,\varepsilon}], \lambda \rangle^h \quad \forall \lambda \in S^h
$$

(4.5b)

and hence

$$
\langle k^{-1} \delta_t^2 \hat{u}^{h,n}_{k,\varepsilon} + D_t^+ [\hat{u}^{h,n}_{k,\varepsilon} + \varphi(\hat{u}^{h,n}_{k,\varepsilon})], \lambda \rangle^h + \langle k^{-1} \nabla D_t^+ \hat{u}^{h,n}_{k,\varepsilon} + \nabla \hat{u}^{h,n+1}_{k,\varepsilon}, \nabla \lambda \rangle^h
= \langle f^{n+1} + k^{-1} D_t^- f^n, \lambda \rangle^h \quad \forall \lambda \in S^h.
$$

(4.6)
Choosing $\chi \equiv D_t^+ \hat{u}^{h,n}_{k,\varepsilon}$ in (4.6) and summing from $n = 1 \rightarrow m$ and noting (4.4) yields for $m \equiv 1 \rightarrow N - 1$ that

$$
\begin{align*}
\frac{1}{2} k^{-1} \left[ |D_t^+ \hat{u}^{h,m}_{k,\varepsilon}|^2_{h} + \tau \sum_{n=1}^{m} |\delta_t^0 \hat{u}^{h,n}_{k,\varepsilon}|^2_{h} - |D_t^+ \hat{u}^{h,0}_{k,\varepsilon}|^2_{h} \right] + \\
+ \frac{1}{2} |\nabla \hat{u}^{h,m+1}_{k,\varepsilon}|^2_{L^2(\Omega)} + (\tau + 2k^{-1}) \tau \sum_{n=1}^{m} |\nabla (D_t^+ \hat{u}^{h,n}_{k,\varepsilon})|^2_{L^2(\Omega)} - |\nabla \hat{u}^{h,1}_{k,\varepsilon}|^2_{L^2(\Omega)} \\
+ \tau \sum_{n=1}^{m} \langle D_t^+ \hat{u}^{h,n}_{k,\varepsilon} + \varphi \tau (u^{h,n}_{k,\varepsilon}), D_t^+ \hat{u}^{h,n}_{k,\varepsilon} \rangle = \tau \sum_{n=1}^{m} (f^{n+1} + k^{-1} D_t^+ f^n, D_t^+ \hat{u}^{h,n}_{k,\varepsilon})
\end{align*}
$$

and hence

$$
\begin{align*}
k^{-1} |D_t^+ \hat{u}^{h,m}_{k,\varepsilon}|^2_{h} + \tau \sum_{n=0}^{m} |D_t^+ \hat{u}^{h,n}_{k,\varepsilon}|^2_{h} + |\nabla \hat{u}^{h,m+1}_{k,\varepsilon}|^2_{L^2(\Omega)} + 2\tau \sum_{n=0}^{m} \langle D_t^+ [\varphi \tau (u^{h,n}_{k,\varepsilon})] D_t^+ \hat{u}^{h,n}_{k,\varepsilon} \rangle \\
\leq (\tau + k^{-1}) |D_t^+ \hat{u}^{h,0}_{k,\varepsilon}|^2_{h} + 2\tau \langle D_t^+ [\varphi \tau (u^{h,0}_{k,\varepsilon})] D_t^+ \hat{u}^{h,0}_{k,\varepsilon} \rangle \\
+ |\nabla \hat{u}^{h,1}_{k,\varepsilon}|^2_{L^2(\Omega)} + \tau \sum_{n=1}^{m} |f^{n+1} + k^{-1} D_t^+ f^n|^2_{h}.
\end{align*}
$$

Next we note from the first equations in (P^{h,\varepsilon}_{1}) and the initial conditions that

$$
(4.8a) \quad (1 + k\tau) \langle D_t^+ \hat{u}^{h,0}_{k,\varepsilon}, \chi \rangle^h = k\tau \langle D_t^+ [\varphi \tau (u^{h,0}_{k,\varepsilon})], \chi \rangle^h \quad \forall \chi \in S^h
$$

and hence

$$
(4.8b) \quad \langle D_t^+ \hat{u}^{h,0}_{k,\varepsilon}, \chi \rangle^h + \frac{k\tau}{1 + k\tau} \langle D_t^+ [\varphi \tau (u^{h,0}_{k,\varepsilon})], \chi \rangle^h + \langle \nabla \hat{u}^{h,1}_{k,\varepsilon}, \nabla \chi \rangle = \langle f^1, \chi \rangle^h \quad \forall \chi \in S^h
$$

In addition it follows from (2.3a) that for $n = 0 \rightarrow N - 1$

$$
(4.9) \quad M^{-1} \varepsilon |D_t^+ [\varphi \tau (u^{h,n}_{k,\varepsilon})]|^2_{h} \leq \langle D_t^+ [\varphi \tau (u^{h,n}_{k,\varepsilon})], D_t^+ \hat{u}^{h,n}_{k,\varepsilon} \rangle.
$$

Choosing $\chi \equiv D_t^+ \hat{u}^{h,0}_{k,\varepsilon}$ in (4.8b) and noting (4.9) and Lemma 3.5 yields that

$$
\tau |D_t^+ \hat{u}^{h,0}_{k,\varepsilon}|^2_{h} + 2\tau \frac{k\tau}{1 + k\tau} \langle D_t^+ [\varphi \tau (u^{h,0}_{k,\varepsilon})], D_t^+ \hat{u}^{h,0}_{k,\varepsilon} \rangle + |\nabla \hat{u}^{h,1}_{k,\varepsilon}|^2_{L^2(\Omega)}
$$

$$
\leq |\nabla \hat{u}^{h,1}_{k,\varepsilon}|^2_{L^2(\Omega)} + \tau |f^1|^2_{h} \leq C.
$$

From (4.7), (4.9), (4.10) and the assumptions on $f$ it follows that

$$
\begin{align*}
k^{-1} |D_t^+ \hat{u}^{h,m}_{k,\varepsilon}|^2_{h} + \tau \sum_{n=0}^{m} |D_t^+ \hat{u}^{h,n}_{k,\varepsilon}|^2_{h} + |\nabla \hat{u}^{h,m+1}_{k,\varepsilon}|^2_{L^2(\Omega)} + \varepsilon \tau \sum_{n=0}^{m} |D_t^+ [\varphi \tau (u^{h,n}_{k,\varepsilon})]|^2_{h}
\end{align*}
$$

$$
\leq C \left[ 1 + (k\tau)^{-1} + \tau \sum_{n=1}^{N} |f^{n+1} + k^{-1} D_t^+ f^n|^2_{h} \right] \leq C \left[ 1 + (k\tau)^{-1} \right].
$$

Choosing $\chi \equiv D_t^+ \hat{u}^{h,n}_{k,\varepsilon}$ in (4.5b) and summing from $n = 1 \rightarrow m$ and noting (4.4) yields for $m = 1 \rightarrow N - 1$ that

$$
\begin{align*}
k^{-1} |D_t^+ \hat{u}^{h,m}_{k,\varepsilon}|^2_{h} + \tau \sum_{n=1}^{m} |D_t^+ \hat{u}^{h,n}_{k,\varepsilon}|^2_{h} \leq k^{-1} |D_t^+ \hat{u}^{h,0}_{k,\varepsilon}|^2_{h} + \tau \sum_{n=1}^{m} |D_t^+ [\varphi \tau (u^{h,n}_{k,\varepsilon})]|^2_{h}.
\end{align*}
$$
Hence noting (4.8a) we obtain for \( m = 1 \to N - 1 \) that

\[
(4.12) \quad k^{-1} |D_t^+ \tilde{v}_{k, \tau}^h|^2_{h} + \tau \sum_{n=0}^{m} |D_t^+ \tilde{v}_{k, \tau}^h|^2_{h} \leq \tau \sum_{n=0}^{m} |D_t^+ [\varphi_{\varepsilon}(U_{k, \tau}^h)]|^2_{h}.
\]

Combining (4.11) and (4.12) and noting (3.3a) yields the desired result (4.3).

**Lemma 4.3.** Under the assumptions (D5) we have for all \( \varepsilon \in (0, \varepsilon_0) \), \( h, \tau > 0 \) and for \( m = 0 \to N \) that

\[
|u_{k, \tau}^h - U_{k, \tau}^{h}|^2_{E_{z,(k \cdot m)}} + \varepsilon |\pi_h [\varphi_{\varepsilon}(u_{k, \tau}^h) - \varphi_{\varepsilon}(U_{k, \tau}^h)]|^2_{L^2(Q_{m \tau})} + \varepsilon \|v_{k, \tau}^h - V_{k, \tau}^{h}\|^2_{E_{1,(k \cdot m)}} \leq C(\varepsilon^{-1} \tau^2 + \varepsilon^{-1} k^{-1} \tau + k^{-2}).
\]

**Proof.** Let \( E_u \equiv \tilde{u}_{k, \tau}^h - U_{k, \tau}^{h}, E_{U} \equiv \tilde{U}_{k, \tau}^h - U_{k, \tau}^{h}, E_{V} \equiv \tilde{V}_{k, \tau}^h - V_{k, \tau}^{h} \) and \( E_f \equiv f - f_r \). First, we note that

\[
(4.14a) \quad |\tilde{U}_{k, \tau}^h - U_{k, \tau}^{h}|^2_{L^2(Q_{\tau})} = \sum_{n=1}^{N} \int_{(n-1)\tau}^{n\tau} (\tau - t)^2 |\partial_t \tilde{U}_{k, \tau}^h|^2 dt \leq \tau^2 |\partial_t \tilde{U}_{k, \tau}^h|^2_{L^2(Q_{\tau})}
\]

and the equivalent result with \( U \) replaced by \( V \). Hence it follows from (4.3) that

\[
(4.14b) \quad |\tilde{U}_{k, \tau}^h - U_{k, \tau}^{h}|^2_{L^2(Q_{\tau})} + \varepsilon |\tilde{V}_{k, \tau}^h - V_{k, \tau}^{h}|^2_{L^2(Q_{\tau})} \leq C(\tau^2 + k^{-1} \tau).
\]

Similarly we have \( |E_f|^2_{L^2(Q_{\tau})} \leq C \tau^2 \).

It follows from \( (\tilde{P}_{h, \tau}) \) and \( (\tilde{P}_{h, \tau}^r) \) that \( E_{u}^{h}(\cdot, 0) = 0 \), \( E_{u}^{L}(\cdot, 0) = 0 \) and for almost every \( t \in (0, T) \)

\[
(4.15a) \quad \langle \partial_t E_{u}^L + \partial_t E_{v}^L, \chi \rangle^h + \langle \nabla E_{u}, \nabla \chi \rangle^h = \langle E_{f}, \chi \rangle^h \quad \forall \chi \in S_{0}^h
\]

\[
(4.15b) \quad \langle \partial_t E_{v}^L, \chi \rangle^h = k \langle [\varphi_{\varepsilon}(u_{k, \tau}^h) - \varphi_{\varepsilon}(U_{k, \tau}^h)] - E_{v}, \chi \rangle^h \quad \forall \chi \in S_{0}^h.
\]

Choosing \( \chi \equiv \int_{0}^{t} E_{u}(\cdot, s) d\sigma \) in (4.15a), integrating over \((0, t)\) in time, where \( s \) is the integration variable in time yields that

\[
(4.16) \quad \int_{0}^{t} \langle \tilde{U}_{k, \tau}^h - U_{k, \tau}^{h}, \chi \rangle^h + \int_{0}^{t} E_{f}(\cdot, \sigma) d\sigma, E_{u}(\cdot, s) \rangle^h ds.
\]

Similarly choosing \( \chi \equiv E_{u} \) in (4.15a) yields that

\[
(4.17) \quad \int_{0}^{t} \left[ \langle [\tilde{U}_{k, \tau}^h - U_{k, \tau}^{h}], \partial_t E_{u}^L, \rangle^h + \langle E_{f},\partial_t E_{u}^L, \rangle^h \right] ds.
\]

From (4.16), (4.17), (3.4), (3.3) and (4.14b) it follows that

\[
\int_{0}^{t} |E_{u}(\cdot, s)|^2_{h} ds + \frac{1}{2} \int_{0}^{t} E_{u}(\cdot, s) d\sigma_{L^2(\Omega)} +
\]
\[+k^{-1} \left[ \frac{1}{2} |E_u^L(\cdot, t)|_h^2 + \int_0^t \| \nabla E_u(\cdot, s)|^2_{L^2(\Omega)} \, ds \right] +\]
\[+ \int_0^t \left\langle \varphi_\varepsilon(\hat{u}_{k,\varepsilon}(\cdot, s)) - \varphi_\varepsilon(\hat{U}_{k,\varepsilon}(\cdot, s)), E_u(\cdot, s) \right\rangle^h ds +\]
\[= \int_0^t \left\langle [\hat{U}_{k,\varepsilon}^L - \hat{U}_{k,\varepsilon}](\cdot, s), \hat{V}_{k,\varepsilon}(\cdot, s) \right\rangle^h ds +\]
\[+k^{-1} \int_0^t \left\langle [\hat{U}_{k,\varepsilon} - \hat{U}_{k,\varepsilon}^L](\cdot, s), \delta_s E_u^L(\cdot, s) \right\rangle^h ds +\]
\[+ \int_0^t \left\langle k^{-1} E_f(\cdot, s) + \int_0^s \! E_f(\cdot, \sigma) \, d\sigma, E_u(\cdot, s) \right\rangle^h ds\]
\[(4.18)\]
\[\leq C[\varepsilon^{-1} + (k\tau)^{-1}][\tau^2 + k^{-1} \tau] \leq C(\varepsilon^{-1} \tau^2 + \varepsilon^{-1} k^{-1} \tau + k^{-2}).\]

The desired result for \(u\) in (4.13) then follows from (4.18), (3.3a) and (2.3a).

Similarly, choosing \(\chi \equiv E_u\) in (4.15b) yields that
\[\frac{1}{2} k^{-1} |E_v^L(\cdot, t)|_h^2 + \int_0^t |E_v(\cdot, s)|_h^2 \, ds =\]
\[\int_0^t \left\langle \varphi_\varepsilon(\hat{u}_{k,\varepsilon}(\cdot, s)) - \varphi_\varepsilon(\hat{U}_{k,\varepsilon}(\cdot, s)), E_v(\cdot, s) \right\rangle^h ds +\]
\[+k^{-1} \int_0^t \left\langle [\hat{V}_{k,\varepsilon} - \hat{V}_{k,\varepsilon}^L](\cdot, s), \delta_s E_v^L(\cdot, s) \right\rangle^h ds.\]

The desired result for \(v\) in (4.13) then follows from (4.19), the result for \(u\) in (4.13), (4.14b), (4.3), (3.4) and (3.3a).

\(\square\)

**Theorem 4.2**

(a) Let the assumptions (D5) hold. Then for the stated choices of \(\varepsilon\) and \(k\), we have that the error bounds (3.31) and (3.32) hold for \(t = mL\), \(m = 0 \to N\), with \(\hat{u}_{k,\varepsilon}^h\) replaced by \(\hat{U}_{k,\varepsilon}\) with \(\tau \leq C k^{-1} = Ch\).

(b) Let assumptions (D6) hold. Then for the stated choices of \(\varepsilon\) and \(k\), we have that the error bounds (3.33) with \(\tau \leq C k^{-1} = C[\varepsilon h^2 [\ln(1/h)]^{1/2]}^{(3-2p)}\) and \(p \in (1/2, 1]\) and (3.34) with \(\tau \leq C k^{-1} = C[h^2 \{\ln(1/h)\}^{2(3-p)}]^{(3-2p)}\) and \(p \in (1/3, 1]\) hold for \(t = mL\), \(m = 0 \to N\), with \(\hat{u}_{k,\varepsilon}^h\) replaced by \(\hat{U}_{k,\varepsilon}\).

Proof. These results follow from (4.13), (3.22) and (2.33). We note that using (4.2) in place of (4.13) leads to a more restrictive bound on \(\tau\).  \(\square\)

Finally we extend the above results to the problem:

\((\hat{P}_{\varepsilon}^{h,\tau})\) For \(n = 1 \to N\) find \(\hat{u}_{k,\varepsilon}^{h,n} \in S_0^n\) such that
\[\tau^{-1} \left( (\hat{u}_{k,\varepsilon}^{h,n} - \hat{u}_{k,\varepsilon}^{h,n-1}) + [\varphi_\varepsilon(\hat{u}_{k,\varepsilon}^{h,n}) - \varphi_\varepsilon(\hat{u}_{k,\varepsilon}^{h,n-1})], \chi \right)^h + \langle \nabla \hat{u}_{k,\varepsilon}^{h,n}, \nabla \chi \rangle = \langle f^h, \chi \rangle^h \forall \chi \in S_0^n\]
\[\hat{u}_{k,\varepsilon}^{h,0}(\cdot) = g^h(\cdot).\]

**Theorem 4.3.** Let the assumptions (D4) hold. Then for all \(\varepsilon \in (0, \varepsilon_0]\), \(h\) and \(\tau > 0\) there exists a unique solution \(\hat{U}_{k,\varepsilon}\) to \((\hat{P}_{\varepsilon}^{h,\tau})\). In addition for \(m = 0 \to N\) we have that
\[\left\| \hat{U}_{k,\varepsilon} - \hat{U}_{k,\varepsilon} \right\|^2_{L^2(t, t + \tau)} + \varepsilon \left\| \pi_\varepsilon [\varphi_\varepsilon(\hat{U}_{k,\varepsilon}) - \varphi_\varepsilon(\hat{U}_{k,\varepsilon})] \right\|^2_{L^2(Q(t, t + \tau))} +\]
\[ + \varepsilon \left| \pi_h \left[ \varphi (\tilde{U}_t) \right]_{x=0} - \tilde{V}_{k,\varepsilon} \right|_{L^2(I, \Omega)}^2 \leq C \varepsilon^{-1} k^{-2} [1 + (k/\tau)^{-1}] \]

Moreover, under the assumptions (D5) we have

(i) the first bound in (4.1) holds true for \( \tilde{U}_t \). In particular, if \( y^h \) and \( f \geq 0 \) then \( \tilde{U}_t \geq 0 \) in \( Q_T \).

(ii) on choosing \( \tau = C h \), the error bounds (3.31) and (3.32) hold for \( t = m \tau \), \( m = 0 \rightarrow N \), under the stated choices of \( \varepsilon \) with \( \tilde{u}^h_{k,\varepsilon} \) replaced by \( \tilde{U}_t \).

Furthermore, under the assumptions (D6) we have that the following error bounds hold with \( \tilde{u}^h_{k,\varepsilon} \) replaced by \( \tilde{U}_t \) for \( t = m \tau \), \( m = 0 \rightarrow N \), under the stated choices of \( \varepsilon \)

(i) (3.33) for \( p \in (1/2, 1) \) and \( \tau = C \left\{ h^2 \left[ \ln (1/h) \right] \right\}^{2/(5-2p)} \)

(ii) (3.34) for \( p \in (1/3, 1) \) and \( \tau = C \left\{ h^2 \left[ \ln (1/h) \right] \right\}^{2(3-p)/(13-7p)} \)

Proof. Existence and uniqueness of a solution to \( (P_{h,\tau}^{h,\varepsilon}) \) follows as in the proof of Theorem 4.1 of part 1. The first inequality in (4.20) is a discrete analogue of the first inequality in (2.8b), and is proved in a similar way. The second inequality in (4.20) follows from (4.3). The first bound in (4.1) follows from (4.20) and the equivalence of norms on \( S^h \). The above error bounds follow by combining (4.20) with Theorem 4.2.

As stated in sections 1 and 2, problem (P) is equivalent to the generalised porous medium equation, whose finite element approximation by \( (P_{h,\tau}^{h,\varepsilon}) \) is analysed in [11]. There the error bounds (3.31a) and (3.32a) for \( \tilde{u}^h_{k,\varepsilon} \) replaced by \( \tilde{U}_t \) are proved under the same choices of \( \varepsilon \), but with \( \tau = Ch^{1+p} \) and \( \tau = Ch^4/(3-p) \), respectively. Therefore Theorem 4.3 above improves on these results as we require only \( \tau = Ch \). As stated previously we have assumed that the mesh is (weakly) acute, whereas they do not. Furthermore, under additional assumptions we have the improved error bounds (3.33) and (3.34).

Remark 4.1. For the numerical computations it seems to be advantageous to use the unrelaxed form \( (P_{h,\tau}^{h,\varepsilon}) \) instead of the relaxed form \( (P_{h,\tau}^{h,\varepsilon}) \), as the unknown \( v \) does not appear. However, note that for a fixed time level \( n \) the set of equations in \( (P_{h,\tau}^{h,\varepsilon}) \) is equivalent to

\[ \tau^{-1} \left( \tilde{u}^{h,n}_{k,\varepsilon} - \tilde{u}^{h,n-1}_{k,\varepsilon} + \frac{k\tau}{1 + k\tau} \left( \varphi (\tilde{u}^{h,n}_{k,\varepsilon}) - \tilde{u}^{h,n-1}_{k,\varepsilon} \right) \right) + \left( \nabla \tilde{u}^{h,n}_{k,\varepsilon}, \nabla \chi \right) = \left( f^n, \chi \right) \forall \chi \in S^h_0 \]

This means that the computational complexity of both variants is nearly identical.

5. A numerical experiment. Finally we discuss some numerical experiments. As an example we take a transformation of the well-known Barenblatt solution

\[ w(x,t) = (t + 1)^{-m/\alpha} \left[ 1 - \frac{m-1}{2m(m+1)} \left( \frac{x}{(t+1)^{1/\alpha}} \right)^{1+\alpha} \right]^{m-1} \]

of the classical porous medium equation (2.15) with \( \beta(w) \equiv sgn(w)|w|^m \), \( m > 1 \), and \( f \equiv 0 \) in one space dimension; see e.g. [1]. Then \( u \equiv w^m \) can be interpreted to be a
solution of problem (P) with
\[ \varphi(s) \equiv \begin{cases} 
[ms^{1/m} - s]_+ & \text{if } s \leq 1 \\
m - 1 & \text{if } s \geq 1 
\end{cases} \]

and for the appropriate choice of \( Q_T \) and initial condition \( q \), if \(-\Delta u\) is replaced by \(-m\Delta u\) or equivalently a time scaling is performed. As one can take \( \tau = 0 \) and \( \bar{u} = 1 \) in (2.5), it follows that \( \varphi \) satisfies all the properties required in (D1), (D3) and (D6) for \( p = 1/m \). Furthermore, \( u \) satisfies the non-degeneracy condition (N.D.) and \( m(\Omega_{
u, \epsilon}) \leq C\epsilon \), such that according to Remark 3.1 we are allowed to use \( \pi_h g \) as the discrete initial data.

As data we take \( \Omega \equiv (-10, 10) \) and \( T = 1 \) so that the support of the solution \( u \) is compactly contained in \( Q_T \) for \( p = 0.1, 0.3, 0.5, 0.7, 0.9 \). We consider a uniform space discretization and hence all the assumptions (D1)-(D6) are satisfied. We compute the solutions of the fully discrete problem with regularization, i.e. \( (p_{h, \epsilon}^r) \), for \( h = 20/J, J = 320(10)^{1/2} \). As (N.D.) holds and \( m(\Omega_{
u, \epsilon}) \leq C\epsilon \) we choose \( \tau = 0.1h, \epsilon = 0.1h^{2(1-p)/(3-p)} \) in accordance with (D5) and \( \tau = 0.1h^{2(3-p)/(13-7p)} \), \( \epsilon = 0.1h^{2(1-p)/(13-7p)} \) in accordance with (D6) for \( p \geq 0.5 \). The constant 0.1 was chosen so that \( \tau \) and \( \epsilon \) were reasonably small on the coarsest mesh. The resulting systems of nonlinear algebraic equations at each time step are solved by a modified nonlinear SOR method, see [5], to an accuracy well below the expected discretization error.

<table>
<thead>
<tr>
<th>( h^{-1} )</th>
<th>( p = 0.1 )</th>
<th>( p = 0.3 )</th>
<th>( p = 0.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (D5) )</td>
<td>( 0.69 )</td>
<td>( 0.74 )</td>
<td>( 0.80 )</td>
</tr>
<tr>
<td>( E^h \times 10^4 )</td>
<td>( \alpha^h )</td>
<td>( E^h \times 10^4 )</td>
<td>( \alpha^h )</td>
</tr>
<tr>
<td>4</td>
<td>102.10</td>
<td>114.09</td>
<td>40.91</td>
</tr>
<tr>
<td>8</td>
<td>59.69</td>
<td>0.77</td>
<td>42.65</td>
</tr>
<tr>
<td>16</td>
<td>41.00</td>
<td>0.54</td>
<td>17.12</td>
</tr>
<tr>
<td>32</td>
<td>21.13</td>
<td>0.96</td>
<td>7.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( h^{-1} )</th>
<th>( p = 0.7 )</th>
<th>( p = 0.9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (D5) )</td>
<td>( 0.87 )</td>
<td>( 0.99 )</td>
</tr>
<tr>
<td>( E^h \times 10^4 )</td>
<td>( \alpha^h )</td>
<td>( E^h \times 10^4 )</td>
</tr>
<tr>
<td>4</td>
<td>35.34</td>
<td>30.85</td>
</tr>
<tr>
<td>8</td>
<td>15.42</td>
<td>1.20</td>
</tr>
<tr>
<td>16</td>
<td>7.15</td>
<td>1.11</td>
</tr>
<tr>
<td>32</td>
<td>3.44</td>
<td>1.06</td>
</tr>
</tbody>
</table>

The error \( \| u - \bar{U}_t \|_{L^2(Q_T)} \) is approximated for practical purposes by
\[ E^h \equiv \left[ \frac{1}{N} \sum_{n=1}^{N} |u(\cdot, n\tau) - \bar{U}_t(\cdot)|^2 h^2 \right]^{1/2}, \]

where \( N\tau \leq T = 1 < (N + 1)\tau \). It follows from (3.3a), a bound similar to (4.14a), (4.3) and (4.14a) for \( k = \infty \) and noting that \( u \in C^\nu(Q_T) \), where \( \nu = \min\{\frac{1}{1-p}, 2\} \),
that

\[
[E^h]^2 \leq C[|\pi_h u - U^L|_{L^2(Q_T)}^2 + C\tau^2]_{L^2(Q_T)}^2 + C|u - U^L|_{L^2(Q_T)}^2 + C|U - U^L|_{L^2(Q_T)}^2 + C\tau^2
\]

\[
\leq C|u - \bar{U}|_{L^2(Q_T)}^2 + C\bar{U}^2 + C\tau^2.
\]

Noting Theorem 4.3, we see that the approximation $E^h$ is of sufficient accuracy.

We estimate the rate of convergence of $E^h$ by setting

\[
\alpha^h = \frac{\ln[E^{2h}/E^h]}{\ln 2}.
\]

Inspecting Tables 5.1–5.2 we see that the actual convergence rates for the approximations $\bar{U}_e$ are better than that predicted by the theory; which appear in the tables next to the corresponding assumption, i.e. $\alpha^h = 2/(3 - \mu)$ and $8/(13 - 7\mu)$ under assumptions (D5) and (D6), respectively; (the latter being restricted to $\mu > 1/3$). It should be noted though that the actual convergence rates have not yet settled down, especially for smaller values of $\mu$.

**Acknowledgments.** We would like to thank John Aspden, Department of Mathematics, Imperial College for the numerical computations in §5.

**REFERENCES**


