Study of a chemo-repulsion model with quadratic production.

Part II: Analysis of an unconditional energy-stable fully discrete scheme

F. Guillén-González*, M. A. Rodríguez-Bellido* and D. A. Rueda-Gómez*†

March 8, 2018

Abstract

This work is devoted to study a fully discrete scheme for a repulsive chemotaxis with quadratic production model. By following the ideas presented in [5], we introduce an auxiliary variable (the gradient of the chemical concentration), and then the corresponding Finite Element (FE) backward Euler scheme is mass-conservative and unconditional energy-stable. For this nonlinear scheme, we study some properties like solvability, convergence towards weak solutions, error estimates, and weak, strong and more regular a priori estimates of the scheme. Additionally, we propose two different linear iterative methods to approach the nonlinear scheme: an energy-stable Picard’s method and the Newton’s method. We prove solvability and convergence of both methods to the nonlinear scheme. Finally, we provide some numerical results in agreement with our theoretical analysis about the error estimates.

2010 Mathematics Subject Classification. 35K51, 35Q92, 65M12, 65M15, 65M60, 92C17.
Keywords: Chemorepulsion-production model, fully discrete scheme, finite element method, energy-stability, error estimates.

1 Introduction

The aim of this paper is to study an unconditional energy-stable fully discrete scheme for the following parabolic-parabolic repulsive-productive chemotaxis model (with quadratic production

*Corresponding Author.

†Escuela de Matemáticas, Universidad Industrial de Santander, A.A. 678, Bucaramanga (COLOMBIA). Email: diaruego@uis.edu.co

Dpto. Ecuaciones Diferenciales y Análisis Numérico and IMUS, Universidad de Sevilla, Facultad de Matemáticas, C/ Tarfia, S/N, 41012 Sevilla (SPAIN). Email: guillen@us.es, angeles@us.es

Dpto. Ecuaciones Diferenciales y Análisis Numérico and IMUS, Universidad de Sevilla, Facultad de Matemáticas, C/ Tarfia, S/N, 41012 Sevilla (SPAIN). Email: guillen@us.es, angeles@us.es

Dpto. Ecuaciones Diferenciales y Análisis Numérico and IMUS, Universidad de Sevilla, Facultad de Matemáticas, C/ Tarfia, S/N, 41012 Sevilla (SPAIN). Email: guillen@us.es, angeles@us.es

Dpto. Ecuaciones Diferenciales y Análisis Numérico and IMUS, Universidad de Sevilla, Facultad de Matemáticas, C/ Tarfia, S/N, 41012 Sevilla (SPAIN). Email: guillen@us.es, angeles@us.es
\[ \begin{aligned}
\partial_t u - \Delta u &= \nabla \cdot (u \nabla v) \quad \text{in } \Omega, \ t > 0, \\
\partial_t v - \Delta v + v &= u^2 \quad \text{in } \Omega, \ t > 0, \\
\frac{\partial u}{\partial n} &= \frac{\partial v}{\partial n} = 0 \quad \text{on } \partial \Omega, \ t > 0, \\
u(x,0) &= u_0(x) \geq 0, \quad v(x,0) = v_0(x) \geq 0 \quad \text{in } \Omega,
\end{aligned} \]  

(1)

where \( \Omega \) is a \( n \)-dimensional open bounded domain, \( n = 2, 3 \), with boundary \( \partial \Omega \). The unknowns for this model are \( u(x,t) \geq 0 \), the cell density, and \( v(x,t) \geq 0 \), the chemical concentration.

Problem (1) is conservative in \( u \), because the total mass \( \int_\Omega u(t) \) remains constant in time, as we can check integrating equation (1)1 in \( \Omega \),

\[
\frac{d}{dt} \left( \int_\Omega u \right) = 0, \quad \text{i.e.} \quad \int_\Omega u(t) = \int_\Omega u_0 := m_0, \quad \forall t > 0.
\]

The problem (1) is well-posed ([5]), because there exists global in time “weak-strong” solutions in the following sense: \( u \geq 0 \) and \( v \geq 0 \) a.e. \( (t,x) \in (0, +\infty) \times \Omega \),

\[ (u,v) \in L^\infty(0, +\infty; L^2(\Omega) \times H^1(\Omega)) \cap L^2(0,T; H^1(\Omega) 	imes H^2(\Omega)), \quad \forall T > 0, \]

\[ (\partial_t u, \partial_t v) \in L^{q'}(0,T; H^1(\Omega)' \times L^2(\Omega)), \quad \forall T > 0, \]

\[ (\nabla u, \nabla v) \in L^2(0, +\infty; L^2(\Omega) \times H^1(\Omega)), \]

(2)

where \( q' = 2 \) in the 2-dimensional case (2D) and \( q' = 4/3 \) in the 3-dimensional case (3D) (\( q' \) is the conjugate exponent of \( q = 2 \) in 2D and \( q = 4 \) in 3D), satisfying the \( u \)-equation (1)_1 in a variational sense, the \( v \)-equation (1)_2 a.e. \( (t,x) \in (0, +\infty) \times \Omega \), and the following energy inequality for a.e. \( t_0, t_1 : t_1 \geq t_0 \geq 0 \):

\[
\mathcal{E}(u(t_1), v(t_1)) - \mathcal{E}(u(t_0), v(t_0)) + \int_{t_0}^{t_1} (\|\nabla u(s)\|_{L^2}^2 + \frac{1}{2}\|\Delta v(s)\|_{L^2}^2 + \frac{1}{2}\|\nabla v(s)\|_{L^2}^2) \, ds \leq 0,
\]

(3)

where \( \mathcal{E}(u,v) = \frac{1}{2}\|u\|_{L^2}^2 + \frac{1}{4}\|\nabla v\|_{L^2}^2 \). Moreover, assuming that the following regularity criteria is satisfied:

\[ (u, \nabla v) \in L^\infty(0, +\infty; H^1(\Omega) \times H^1(\Omega)), \]

(4)
(which, at least in 2D domains, is always true), it was proved that there exists a unique global in time strong solution of (1) satisfying

\[
\begin{align*}
(u, v) & \in L^\infty(0, +\infty; H^2(\Omega)^2) \cap L^2(0, T; H^3(\Omega)^2), \\
(\partial_t u, \partial_t v) & \in L^\infty(0, +\infty; L^2(\Omega) \times H^1(\Omega)) \cap L^2(0, +\infty; H^1(\Omega) \times H^2(\Omega)), \\
(\partial_{tt} u, \partial_{tt} v) & \in L^2(0, +\infty; H^1(\Omega)' \times L^2(\Omega)).
\end{align*}
\]

In particular, (5)_1 implies

\[
\| (u, v) \|_{L^\infty(0, +\infty; L^\infty \times L^\infty)} \leq C.
\]

Therefore, it is desired to design numerical methods for the model (1) conserving at the discrete level the main properties of the continuous model, such as mass-conservation, energy-stability, positivity and regularity.

There are only a few works about numerical analysis for chemotaxis models. For instance, for the Keller-Segel system (i.e. with chemo-attraction and linear production), Filbet studied in [4] the existence of discrete solutions and the convergence of a finite volume scheme. Saito, in [9, 10], proved error estimates for a conservative Finite Element (FE) approximation. A mixed FE approximation is studied in [7]. In [3], some error estimates are proved for a fully discrete discontinuous FE method. However, as far as we know, there are not works studying FE schemes satisfying the property of energy-stability related to the energy inequality (3).

In this paper, we propose an unconditional energy-stable fully discrete scheme, which inherit some properties from the continuous model, such as mass-conservation, and weak and strong estimates analogues to (2) and (5)-(6). Moreover, with respect to the nonnegativity of the discrete cell and chemical variables, \( u^n_h \) and \( v^n_h \), we can deduce that \( v^n_h \geq 0 \) (see Remark 3.2), but the cell density nonnegativity \( u^n_h \geq 0 \) can not be assured.

In order to design the scheme, we follow the ideas presented in [5], and we reformulate (1)
introducing a new variable $\sigma = \nabla v$ instead of $v$. Then, model (1) is rewritten as:

$$
\begin{cases}
\partial_t u - \nabla \cdot (\nabla u) = \nabla \cdot (u \sigma) & \text{in } \Omega, \ t > 0, \\
\partial_t \sigma - \nabla (\nabla \cdot \sigma) + \nabla + \text{rot} \text{ (rot } \sigma) = \nabla (u^2) & \text{in } \Omega, \ t > 0, \\
\frac{\partial u}{\partial n} = 0 & \text{on } \partial \Omega, \ t > 0, \\
\sigma \cdot n = 0, \ \ [\text{rot } \sigma \times n]_{\text{tang}} = 0 & \text{on } \partial \Omega, \ t > 0, \\
u(x,0) = u_0(x) \geq 0, \ \sigma(x,0) = \nabla v_0(x) & \text{in } \Omega,
\end{cases} \tag{7}
$$

where (7)$_2$ has been obtained applying the gradient operator to equation (1)$_2$ and adding the term rot(rot $\sigma$) using the fact that rot $\sigma = \text{rot} (\nabla v) = 0$. Once system (7) is solved, we can recover $v$ from $u^2$ by solving

$$
\begin{cases}
\partial_t v - \Delta v + v = u^2 & \text{in } \Omega, \ t > 0, \\
\frac{\partial v}{\partial n} = 0 & \text{on } \partial \Omega, \ t > 0, \\
v(x,0) = v_0(x) > 0 & \text{in } \Omega.
\end{cases} \tag{8}
$$

The outline of this paper is as follows: In Section 2, we give the notation and some preliminary results that will be used along this paper. In Section 3, we study the FE Backward Euler scheme corresponding to formulation (7)-(8), including mass-conservation, unconditional energy-stability, solvability, weak and strong estimates, convergence towards weak solutions, and optimal error estimates of the scheme. In Section 4, we propose two different linear iterative methods in order to approach the nonlinear scheme proposed in Section 3, which are an energy-stable Picard’s method and the Newton’s method. We prove the solvability and the convergence of these methods to the nonlinear scheme. Finally, in Section 5, we present some numerical results in agreement with the theoretical analysis about the error estimates.

## 2 Notations and preliminary results

We recall some functional spaces which will be used throughout this paper. We will consider the usual Sobolev spaces $H^m(\Omega)$ and Lebesgue spaces $L^p(\Omega)$, $1 \leq p \leq \infty$, with norms $\| \cdot \|_m$ and $\| \cdot \|_{L^p}$, respectively. In particular, the $L^2(\Omega)$-norm will be denoted by $\| \cdot \|_0$. We denote by $H^1_0(\Omega) := \{ u \in H^1(\Omega) : u \cdot n = 0 \text{ on } \partial \Omega \}$ and we will use the following equivalent norms in
$H^1(\Omega)$ and $H^1_\sigma(\Omega)$, respectively (see [8] and [1, Corollary 3.5], respectively):

\[
\|u\|_1^2 = \|\nabla u\|_0^2 + \left(\int_\Omega u\right)^2, \quad \forall u \in H^1(\Omega),
\]

\[
\|\sigma\|_1^2 = \|\sigma\|_0^2 + \|
\text{rot } \sigma\|_0^2 + \|
\nabla \cdot \sigma\|_0^2, \quad \forall \sigma \in H^1_\sigma(\Omega).
\]

If $Z$ is a general Banach space, its topological dual will be denoted by $Z'$. Moreover, the letters $C, K$ will denote different positive constants (independent of discrete parameters) which may change from line to line (or even within the same line).

We define the linear elliptic operators

\[
Aw = g \iff \begin{cases} -\Delta w + w = g \text{ in } \Omega, \\ \frac{\partial w}{\partial n} = 0 \text{ on } \partial \Omega, \end{cases}
\]

and

\[
B\sigma = h \iff \begin{cases} -\nabla(\nabla \cdot \sigma) + \text{rot } (\text{rot } \sigma) + \sigma = h \text{ in } \Omega, \\ \sigma \cdot n = 0, \quad [\text{rot } \sigma \times n]_{\text{tang}} = 0 \text{ on } \partial \Omega, \end{cases}
\]

which, in variational form, are given by $A : H^1(\Omega) \to H^1(\Omega)'$ and $B : H^1_\sigma(\Omega) \to H^1_\sigma(\Omega)'$ such that

\[
\langle Aw, \bar{w} \rangle = (\nabla w, \nabla \bar{w}) + (w, \bar{w}), \quad \forall w, \bar{w} \in H^1(\Omega),
\]

\[
\langle B\sigma, \bar{\sigma} \rangle = (\sigma, \bar{\sigma}) + (\nabla \cdot \sigma, \nabla \cdot \bar{\sigma}) + \langle \text{rot } \sigma, \text{rot } \bar{\sigma} \rangle, \quad \forall \sigma, \bar{\sigma} \in H^1_\sigma(\Omega).
\]

We assume the $H^2$-regularity of problems (9) and (10). Consequently, we have the existence of some constants $C > 0$ such that

\[
\|w\|_2 \leq C\|Aw\|_0, \quad \forall w \in H^2(\Omega), \quad \text{and} \quad \|\sigma\|_2 \leq C\|B\sigma\|_0, \quad \forall \sigma \in H^2(\Omega).
\]

Along this paper, we will use repeatedly the classical 3D interpolation inequality

\[
\|u\|_{L^3} \leq C\|u\|_0^{1/2}\|u\|_1^{1/2}, \quad \forall u \in H^1(\Omega).
\]

Finally, we will use the following results (see [6] and [11]):
Lemma 2.1. Assume that \(\delta, \beta, k > 0\) and \(d^n \geq 0\) satisfy
\[
\frac{d^{n+1} - d^n}{k} + \delta d^{n+1} \leq \beta, \quad \forall n \geq 0.
\]
Then, for any \(n_0 \geq 0\),
\[
d^n \leq (1 + \delta k)^{-n-n_0} d^{n_0} + \delta^{-1} \beta, \quad \forall n \geq n_0.
\]

Lemma 2.2. (Uniform discrete Gronwall lemma) Let \(k > 0\) and \(d^n, g^n, h^n \geq 0\) such that
\[
\frac{d^{n+1} - d^n}{k} \leq g^n d^n + h^n, \quad \forall n \geq 0.
\]
If for any \(r \in \mathbb{N}\), there exist \(a_1(t_r), a_2(t_r)\) and \(a_3(t_r)\) depending on \(t_r = kr\), such that
\[
k \sum_{n=n_0}^{n_0+r-1} g^n \leq a_1(t_r), \quad k \sum_{n=n_0}^{n_0+r-1} h^n \leq a_2(t_r), \quad k \sum_{n=n_0}^{n_0+r-1} d^n \leq a_3(t_r),
\]
for any integer \(n_0 \geq 0\), then
\[
d^n \leq \left( a_2(t_r) + \frac{a_3(t_r)}{t_r} \right) \exp \{a_1(t_r)\}, \quad \forall n \geq r.
\]

As consequence of Lemma 2.2 and Discrete Gronwall Lemma, we have the following result (see [5, Corollary 2.4.]):

Corollary 2.3. Under hypothesis of Lemma 2.2. Let \(k_0 > 0\) be fixed, then the following estimate holds for all \(k \leq k_0\)
\[
d^n \leq C(d^0, k_0) \quad \forall n \geq 0.
\]

3 Fully Discrete Backward Euler Scheme in variables \((u, \sigma)\)

This section is devoted to design an unconditionally energy-stable scheme for model (1) (for a modified energy in variables \((u, \sigma))\), using a FE discretization in space and the backward Euler discretization in time (considered for simplicity on a uniform partition of \([0, T]\) with time step \(k = T/N : (t_n = nk)_{n=0}^{N}\)). Concerning the space discretization, we consider \(\{T_h\}_{h>0}\) be a family of shape-regular and quasi-uniform triangulations of \(\Omega\) made up of simplexes \(K\) (triangles in two dimensions and tetrahedra in three dimensions), so that \(\Omega = \bigcup_{K \in T_h} K\), where \(h = \max_{K \in T_h} h_K\).
with $h_K$ being the diameter of $K$. Further, let $\mathcal{N}_h = \{a_i\}_{i \in I}$ denote the set of all the nodes of $\mathcal{T}_h$. We choose the following continuous FE spaces for $u$, $\sigma$ and $v$:

$$(U_h, \Sigma_h, V_h) \subset H^1 \times H^1_\sigma \times W^{1,6},$$

generated by $P_k, P_m, P_r$ with $k, m, r \geq 1$.

Now, let $A_h : U_h \rightarrow U_h$, $B_h : \Sigma_h \rightarrow \Sigma_h$ and $\tilde{A}_h : V_h \rightarrow V_h$ be the linear operators defined, respectively, as follows:

$$\begin{align*}
(A_h u_h, \bar{u}_h) &= (\nabla u_h, \nabla \bar{u}_h) + (u_h, \bar{u}_h), \quad \forall \bar{u}_h \in U_h, \\
(B_h \sigma_h, \bar{\sigma}_h) &= (\nabla \cdot \sigma_h, \nabla \cdot \bar{\sigma}_h) + (\text{rot} \sigma_h, \text{rot} \bar{\sigma}_h) + (\sigma_h, \bar{\sigma}_h), \quad \forall \bar{\sigma}_h \in \Sigma_h, \\
(\tilde{A}_h v_h, \bar{v}_h) &= (\nabla v_h, \nabla \bar{v}_h) + (v_h, \bar{v}_h), \quad \forall \bar{v}_h \in V_h.
\end{align*}$$

Moreover, we choose the following interpolation operators:

$$\begin{align*}
\mathcal{R}_h^u : H^1(\Omega) \rightarrow U_h, \quad \mathcal{R}_h^\sigma : H^1_\sigma(\Omega) \rightarrow \Sigma_h, \quad \mathcal{R}_h^v : H^1(\Omega) \rightarrow V_h
\end{align*}$$

such that for all $u \in H^1(\Omega)$, $\sigma \in H^1_\sigma(\Omega)$ and $v \in H^1(\Omega)$, $\mathcal{R}_h^u u \in U_h$, $\mathcal{R}_h^\sigma \sigma \in \Sigma_h$ and $\mathcal{R}_h^v v \in V_h$ satisfy

$$\begin{align*}
(\nabla(\mathcal{R}_h^u u - u), \nabla \bar{u}_h) + (\mathcal{R}_h^u u - u, \bar{u}_h) &= 0, \quad \forall \bar{u}_h \in U_h, \\
(\nabla \cdot (\mathcal{R}_h^\sigma \sigma - \sigma), \nabla \cdot \bar{\sigma}_h) + (\text{rot} (\mathcal{R}_h^\sigma \sigma - \sigma), \text{rot} \bar{\sigma}_h) + (\mathcal{R}_h^\sigma \sigma - \sigma, \bar{\sigma}_h) &= 0, \quad \forall \bar{\sigma}_h \in \Sigma_h, \\
(\nabla(\mathcal{R}_h^v v - v), \nabla \bar{v}_h) + (\mathcal{R}_h^v v - v, \bar{v}_h) &= 0, \quad \forall \bar{v}_h \in V_h,
\end{align*}$$

respectively. Observe that, from Lax-Milgram Theorem, the interpolation operators $\mathcal{R}_h^u$, $\mathcal{R}_h^\sigma$ and $\mathcal{R}_h^v$ are well defined. Moreover, the following interpolation errors hold

$$\begin{align*}
\frac{1}{h}||\mathcal{R}_h^u u - u||_0 + ||\mathcal{R}_h^u u - u||_1 &\leq Ch^k ||u||_{k+1} \quad \forall u \in H^{k+1}(\Omega), \\
\frac{1}{h}||\mathcal{R}_h^\sigma \sigma - \sigma||_0 + ||\mathcal{R}_h^\sigma \sigma - \sigma||_1 &\leq Ch^m ||\sigma||_{m+1} \quad \forall \sigma \in H^{m+1}(\Omega), \\
\frac{1}{h}||\mathcal{R}_h^v v - v||_0 + ||\mathcal{R}_h^v v - v||_1 &\leq Ch^r ||v||_{r+1} \quad \forall v \in H^{r+1}(\Omega).
\end{align*}$$

Also, the following stability properties will be used

$$\| (\mathcal{R}_h^u u, \mathcal{R}_h^\sigma \sigma, \mathcal{R}_h^v v) \|_{W^{1,0}} \leq C \left\| (u, \sigma, v) \right\|_2,$$
which can be obtained from (17)-(19), using the inverse inequality

$$\| (u_h, \sigma_h, v_h) \|_{W^{1,6}} \leq C h^{-1} \| (u_h, \sigma_h, v_h) \|_1 \quad \text{for all } (u_h, \sigma_h, v_h) \in U_h \times \Sigma_h \times V_h,$$

and comparing $R_h^{u, \sigma, v}$ with an average interpolation of Clement or Scott-Zhang type (which is stable in $W^{1,6}$-norm).

**Lemma 3.1.** Assume the $H^2$-regularity for problems (9)-(10) given in (11). Then, the following estimates hold

$$\| u_h \|_{W^{1,6}} \leq C \| A_h u_h \|_0 \quad \forall u_h \in U_h, \quad \| v_h \|_{W^{1,6}} \leq C \| \tilde{A}_h v_h \|_0 \quad \forall v_h \in V_h, \quad (21)$$

$$\| \sigma_h \|_{W^{1,6}} \leq C \| B_h \sigma_h \|_0 \quad \forall \sigma_h \in \Sigma_h. \quad (22)$$

**Proof.** First, we consider regular functions associated to the discrete functions $A_h u_h, \tilde{A}_h v_h$ and $B_h \sigma_h$. We define $u(h), v(h) \in H^2(\Omega)$ and $\sigma(h) \in H^2(\Omega)$ as the solutions of problems

$$\begin{align*}
\left\{ \begin{array}{l}
-\Delta u(h) + u(h) = A_h u_h \quad \text{in } \Omega, \\
\frac{\partial u(h)}{\partial n} = 0 \quad \text{on } \partial \Omega,
\end{array} \right. \quad (23)
\end{align*}$$

$$\begin{align*}
\left\{ \begin{array}{l}
-\Delta v(h) + v(h) = \tilde{A}_h v_h \quad \text{in } \Omega, \\
\frac{\partial v(h)}{\partial n} = 0 \quad \text{on } \partial \Omega,
\end{array} \right. \quad (24)
\end{align*}$$

and

$$\begin{align*}
\left\{ \begin{array}{l}
-\nabla (\nabla \cdot \sigma(h)) + \text{rot} \left( \text{rot } \sigma(h) \right) + \sigma(h) = B_h \sigma_h \quad \text{in } \Omega, \\
\sigma(h) \cdot n = 0, \quad [\text{rot } \sigma(h) \times n]_{\text{tang}} = 0 \quad \text{on } \partial \Omega.
\end{array} \right. \quad (25)
\end{align*}$$

In particular, from (11),

$$\| u(h) \|_2 \leq C \| A_h u_n \|_0, \quad \| v(h) \|_2 \leq C \| \tilde{A}_h v_h \|_0 \quad \text{and} \quad \| \sigma(h) \|_2 \leq C \| B_h \sigma_h \|_0. \quad (26)$$

Now, we decompose the $W^{1,6}$-norm as:

$$\| u_h \|_{W^{1,6}} \leq \| u_h - R_h^u u(h) \|_{W^{1,6}} + \| R_h^u u(h) - u(h) \|_{W^{1,6}} + \| u(h) \|_{W^{1,6}} := I_1 + I_2 + I_3, \quad (27)$$

$$\| v_h \|_{W^{1,6}} \leq \| v_h - R_h^v v(h) \|_{W^{1,6}} + \| R_h^v v(h) - v(h) \|_{W^{1,6}} + \| v(h) \|_{W^{1,6}} := H_1 + H_2 + H_3, \quad (28)$$

8
\[
\|\sigma_h\|_{W^{1,6}} \leq \|\sigma_h - R^n_h \sigma(h)\|_{W^{1,6}} + \|R^n_h \sigma(h) - \sigma(h)\|_{W^{1,6}} + \|\sigma(h)\|_{W^{1,6}} := J_1 + J_2 + J_3. \tag{29}
\]

In order to bound \(J_i\), \(i = 1, 2\), we test \((25)_1\) by \(\sigma_h \in \Sigma_h\) and using \((13)_2\) we have

\[
(\nabla \cdot \sigma_h, \nabla \cdot \sigma_h) + (\text{rot} \sigma_h, \text{rot} \sigma_h) + (\sigma_h, \sigma_h) = (\nabla \cdot \sigma(h), \nabla \cdot \sigma_h) + (\text{rot} \sigma(h), \text{rot} \sigma_h) + (\sigma(h), \sigma_h), \quad \forall \sigma_h \in \Sigma_h. \tag{30}
\]

By subtracting at both sides of equality \((30)\) the terms \((\nabla \cdot R^n_h \sigma(h), \nabla \cdot \sigma_h)\), \((\text{rot} R^n_h \sigma(h), \text{rot} \sigma_h)\) and \((R^n_h \sigma(h), \sigma_h)\), testing by \(\sigma_h = \sigma_h - R^n_h \sigma(h) \in \Sigma_h\) and using the Hölder and Young inequalities, we deduce

\[
\|\sigma_h - R^n_h \sigma(h)\|_1 \leq C\|R^n_h \sigma(h) - \sigma(h)\|_1 \leq C h\|\sigma(h)\|_2, \tag{31}
\]

where in the last inequality interpolation error \((18)\) was used. Then, using in \((31)\) the inverse inequality \(\|\sigma_h\|_{W^{1,6}} \leq C h^{-1}\|\sigma(h)\|_1\) for all \(\sigma_h \in \Sigma_h\), we conclude that for \(i = 1, 2\)

\[
J_i \leq C h^{-1}\|R^n_h \sigma(h) - \sigma(h)\|_1 \leq C\|\sigma(h)\|_2. \tag{32}
\]

Finally,

\[
J_3 = \|\sigma(h)\|_{W^{1,6}} \leq C\|\sigma(h)\|_2. \tag{33}
\]

Therefore, using \((32)-(33)\) in \((29)\), and taking into account \((26)\), we deduce \((22)\). Proceeding analogously for \(I_i\) and \(H_i\), \(i = 1, 2, 3\), we deduce \((21)\). \(\square\)

### 3.1 Definition of the scheme

By taking into account the reformulation \((7)\), we consider the following FE Backward Euler Scheme in variables \((u, \sigma)\) (Scheme US, from now on) which is a first order in time, nonlinear and coupled scheme:

- **Initialization:** We fix \((u^0_h, \sigma^0_h) = (R^n_h u_0, R^n_h \sigma_0) \in U_h \times \Sigma_h\) and \(v^0_h = R^n_h v_0 \in V_h\). Then,
  \[
  \int_\Omega u^0_h = \int_\Omega u_0 = m_0.
  \]

- **Time step:** Given \((u^{n-1}_h, \sigma^{n-1}_h) \in U_h \times \Sigma_h\), compute \((u^n_h, \sigma^n_h) \in U_h \times \Sigma_h\) solving

\[
\begin{align*}
(\delta_t u^n_h, \bar{u}_h) + (\nabla u^n_h, \nabla \bar{u}_h) + (u^n_h \sigma^n_h, \nabla \bar{u}_h) &= 0, \quad \forall \bar{u}_h \in U_h, \\
(\delta_t \sigma^n_h, \bar{\sigma}_h) + (B_h \sigma^n_h, \bar{\sigma}_h) - 2(u^n_h \nabla u^n_h, \bar{\sigma}_h) &= 0, \quad \forall \bar{\sigma}_h \in \Sigma_h,
\end{align*}
\] \(\tag{34}\)

\[
\int_\Omega u^0_h = \int_\Omega u_0 = m_0.
\]
where $\delta_t u^n_h = \frac{u^n_h - u^{n-1}_h}{k}$.

Once the scheme US is solved, given $v^{n-1}_h \in V_h$, we can recover $v^n_h = v^n_h((u^n_h)^2) \in V_h$ solving:

$$
\left( \delta_t v^n_h, \bar{v}_h \right) + \left( \tilde{A}_h v^n_h, \bar{v}_h \right) = \left( (u^n_h)^2, \bar{v}_h \right), \quad \forall \bar{v}_h \in V_h.
$$

(35)

Given $u^n_h \in U_h$ and $v^{n-1}_h \in V_h$, Lax-Milgram theorem implies that there exists a unique $v^n_h \in V_h$ solution of (35).

Remark 3.2. By using the mass-lumping technique in all terms of (35) excepting the self-diffusion term $(\nabla v^n_h, \nabla \bar{v}_h)$, approximating by $P_1$-continuous FE and imposing a condition based on a geometrical property of the triangulation, related to the fact that the interior angles of the triangles or tetrahedra must be at most $\pi/2$, we can prove that if $v^{n-1}_h \geq 0$ then $v^n_h \geq 0$. However, in all numerical simulations that we have made without using mass-lumping, we have not found any example in which, beginning with $v^0_h \geq 0$ we obtain $v^n_h(a_i) < 0$, for some $n > 0$ and $a_i$.

3.2 Solvability, Energy-Stability and Convergence

Assuming that the functions $\bar{u}_h = 1 \in U_h$ and $\bar{v}_h = 1 \in V_h$, we deduce that the scheme US conserves in time the total mass $\int_\Omega u^n_h$, that is,

$$
\int_\Omega u^n_h = \int_\Omega u^{n-1}_h = \cdots = \int_\Omega u^0_h,
$$

and we have the following behavior for $\int_\Omega v^n_h$:

$$
\delta_t \left( \int_\Omega v^n_h \right) = \int_\Omega (u^n_h)^2 - \int_\Omega v^n_h.
$$

Now, we establish some results concerning to the solvability and energy-stability of scheme US, but we will omit their proofs because those follow the same ideas given in [5] (Theorem 4.3, Lemma 4.6 and Theorem 4.8, respectively).

**Theorem 3.3. (Unconditional existence and conditional uniqueness) There exists $(u^n_h, \sigma^n_h) \in U_h \times \Sigma_h$ solution of the scheme US. Moreover, if**

$$
k ||(u^n_h, \sigma^n_h)||_1^4 \text{ is small enough,}
$$

(36)

**then the solution is unique.**
Remark 3.4. In the case of 2D domains, from estimate (54) below, the uniqueness restriction (36) can be relaxed to $kK_0^2$ small enough, where $K_0$ is a constant depending on data $(\Omega, u_0, \sigma_0)$, but independent of $(k, h)$ and $n$.

Remark 3.5. In 3D domains, using the inverse inequality $\|u_h\|_1 \leq \frac{C}{h}\|u_h\|_0$ (see Lemma 4.5.3 in [2], p. 111) and estimate (41) below, we have that

$$\|(u_h^n, \sigma_h^n)\|_1^4 \leq \frac{C}{h^4}\|(u_h^n, \sigma_h^n)\|_0^4 \leq \frac{C}{h^4}C_0^2$$

and therefore, the uniqueness restriction (36) can be rewritten as

$$\frac{kC_1}{h^4} \text{ small enough,} \quad (37)$$

where $C_1$ is a positive constant depending on data $(\Omega, u_0, \sigma_0)$, but independent of $n$.

Definition 3.6. A numerical scheme with solution $(u_n, \sigma_n)$ is called energy-stable with respect to the energy

$$E(u, \sigma) = \frac{1}{2}\|u\|_0^2 + \frac{1}{4}\|\sigma\|_0^2, \quad (38)$$

if this energy is time decreasing, that is

$$E(u_h^n, \sigma_h^n) \leq E(u_h^{n-1}, \sigma_h^{n-1}), \quad \forall n. \quad (39)$$

Lemma 3.7. (Unconditional stability) The scheme $US$ is unconditionally energy-stable with respect to $E(u, \sigma)$. In fact, if $(u_h^n, \sigma_h^n)$ is a solution of the scheme $US$, then the following discrete energy law holds

$$\delta_t E(u_h^n, \sigma_h^n) + \frac{k}{2}\|\delta_t u_h^n\|_0^2 + \frac{k}{4}\|\delta_t \sigma_h^n\|_0^2 + \|\nabla u_h^n\|_0^2 + \frac{1}{2}\|\sigma_h^n\|_1^2 = 0. \quad (40)$$

Remark 3.8. Looking at (40), we can say that the scheme $US$ introduces the following two first order “numerical dissipation terms”:

$$\frac{k}{2}\|\delta_t u_h^n\|_0^2 \quad \text{and} \quad \frac{k}{4}\|\delta_t \sigma_h^n\|_0^2.$$

From the (local in time) discrete energy law (40), we deduce the following global in time estimates for $(u_h^n, \sigma_h^n)$ solution of scheme $US$:
Theorem 3.9. (Uniform Weak estimates of scheme US) Let \((u^n_h, \sigma^n_h)\) be a solution of scheme US. Then, the following estimates hold

\[
\|(u^n_h, \sigma^n_h)\|_0^2 + k^2 \sum_{m=1}^{n} \|((\delta_t u^n_m, \delta_t \sigma^n_m))\|_0^2 + k \sum_{m=1}^{n} \|((\nabla u^n_m, \sigma^n_m))\|_{L^2 \times H^1}^2 \leq C_0, \quad \forall n \geq 1,
\] (41)

\[
k \sum_{m=n_0+1}^{n+n_0} \|(u^n_m, \sigma^n_m)\|_1^2 \leq C_0 + C_1(nk), \quad \forall n \geq 1,
\] (42)

where the integer \(n_0 \geq 0\) is arbitrary, with positive constants \(C_0, C_1\) depending on the data \((\Omega, u_0, \sigma_0)\), but independent of \((k, h)\) and \((n, n_0)\).

3.2.1 Weak estimates of \(v^n_h\) in 3D domains

In this fully discrete scheme US it is not clear how to quantify the relation \(\sigma^n_h \simeq \nabla v^n_h\), and therefore, the uniform estimates for \(v_n\) cannot be obtained directly from estimates for \(\sigma^n_h\). Thus, in this subsection we will obtain directly uniform weak estimates for \(v^n_h\).

Lemma 3.10. (Estimate of \(\left|\int_{\Omega} v^n_h\right|\)) Let \(v^n_h\) be the solution of (35). Then, it holds

\[
\left|\int_{\Omega} v^n_h\right| \leq K_0, \quad \forall n \geq 0,
\] (43)

where \(K_0\) is a positive constant depending on the data \(u_0, \sigma_0, v_0\), but independent of \(k, h\) and \(n\).

Proof. The proof follows as in Corollary 4.10 of [5]. \(\square\)

Lemma 3.11. (Discrete duality estimates for \(v^n_h\)) Let \(v^n_h\) be the solution of (35). Then, the following estimates hold

\[
\|\tilde{A}_h^{-1} v^n_h\|_1^2 \leq K_0, \quad \forall n \geq 0,
\] (44)

\[
k \sum_{m=n_0+1}^{n+n_0} \|v^n_m\|_0^2 \leq K_0 + K_1(nk), \quad \forall n \geq 1,
\] (45)

with positive constants \(K_0, K_1\) depending on the data \(\Omega, u_0, \sigma_0, v_0\), but independent of \((k, h)\) and \((n, n_0)\).

Proof. Testing (35) by \(\tilde{v} = \tilde{A}_h^{-1} v^n_h\), and using (21)2 and (41), it is not difficult to deduce

\[
\delta_t \left(\frac{1}{2} \|\tilde{A}_h^{-1} v^n_h\|_1^2\right) + \|v^n_h\|_0^2 \leq \|u^n_h\|_0^2 \|\tilde{A}_h^{-1} v^n_h\|_{L^\infty} \leq C\|\tilde{A}_h^{-1} v^n_h\|_{W^{1,6}} \leq C\|v^n_h\|_0 \leq \frac{1}{2} \|v^n_h\|_0^2 + C,
\]
which implies that
\[ \delta_t \left( \| \tilde{A}_h^{-1} v_h^n \|_1^2 \right) + \| v_h^n \|_0^2 \leq C, \]  
(46)
where \( C \) is a constant independent of \((k, h)\) and \( n \). Then, using that \( \| v_h^n \|_0^2 \geq C \| \tilde{A}_h^{-1} v_h^n \|_1^2 \) (owing to \((21)_2\)) in (46), we deduce
\[ (1 + Ck)\| \tilde{A}_h^{-1} v_h^n \|_1^2 - \| \tilde{A}_h^{-1} v_h^{n-1} \|_1^2 \leq Ck, \]  
(47)
and therefore, using Lemma 2.1 in (47), we obtain (44). Finally, multiplying (46) by \( k \) and adding from \( m = n_0 + 1 \) to \( m = n + n_0 \), using (44), we conclude (45).

\[ \square \]

**Lemma 3.12.** (Weak estimates for \( v_h^n \)) Under hypothesis of Lemma 3.11, the following estimates hold
\[ \| v_h^n \|_0^2 \leq C_2, \quad \forall n \geq 0, \]  
(48)
\[ k \sum_{m=n_0+1}^{n+n_0} \| v_h^m \|_1^2 \leq C_2 + C_3(nk), \quad \forall n \geq 1, \]  
(49)
with positive constants \( C_2, C_3 \) depending on the data \( \Omega, u_0, \sigma_0, v_0 \), but independent of \((k, h)\) and \((n, n_0)\).

**Proof.** Testing (35) by \( \bar{v} = v_h^n \) we obtain
\[ \delta_t \left( \frac{1}{2} \| v_h^n \|_0^2 \right) + \frac{1}{2k} \| v_h^n - v_h^{n-1} \|_0^2 + \| v_h^n \|_1^2 = ((u_h^n)^2, v_h^n - v_h^{n-1}) + ((u_h^n)^2, v_h^{n-1}) \]  
\[ \leq \frac{1}{4k} \| v_h^n - v_h^{n-1} \|_0^2 + Ck \| u_h^n \|_{L^4}^4 + \frac{1}{2} \| u_h^n \|_{L^4}^2 \| v_h^{n-1} \|_0^2 + \frac{1}{2} \| u_h^n \|_{L^4}^2, \]
which implies that
\[ \| v_h^n \|_0^2 - \| v_h^{n-1} \|_0^2 + k \| v_h^n \|_1^2 \leq Ck^2 \| u_h^n \|_{L^4}^4 + k \| u_h^n \|_{L^4}^2 \| v_h^{n-1} \|_0^2 + k \| u_h^n \|_{L^4}^2. \]  
(50)
Moreover, taking into account that \( k \| u_h^n \|_{L^4}^2 \leq kC \| u_h^n \|_{L^4}^2 \), from estimate (42) we deduce
\[ k \| u_h^n \|_{L^4}^2 \leq C_0 + C_1 k. \]  
(51)
Then, from (50) and (51), we have
\[ \| v_h^n \|_0^2 - \| v_h^{n-1} \|_0^2 + k \| v_h^n \|_1^2 \leq (CC_0 + CC_1 k + 1)k \| u_h^n \|_{L^4}^2 + k \| u_h^n \|_{L^4}^2 \| v_h^{n-1} \|_0^2, \]  
(52)
which, in particular implies
\[
\|v^n_h\|_0^2 - \|v^{n-1}_h\|_0^2 \leq C k\|u^n_h\|_{L^2}^2 + k\|u^n_h\|_{L^4}^2 \|v^{n-1}_h\|_0^2.
\] (53)

Therefore, taking into account estimates (42) and (45), applying Corollary 2.3 in (53), we conclude (48). Finally, summing for \(m\) from \(n_0 + 1\) to \(n + n_0\) in (52), and using (42) and (48), we deduce (49).

\[
\square
\]

3.2.2 Convergence

Starting from the previous stability estimates, proceeding as in Theorem 4.11 of [5] we can prove the convergence towards weak solutions as \((k, h) \to 0\). Concretely, by introducing the functions:

- \((\tilde{u}_{h,k}, \tilde{\sigma}_{h,k})\) are continuous functions on \([0, +\infty)\), linear on each interval \((t_{n-1}, t_n)\) and equal to \((u^n_h, \sigma^n_h)\) at \(t = t_n, n \geq 0\);

- \((u^n_{h,k}, \sigma^n_{h,k})\) as the piecewise constant functions taking values \((u^n_h, \sigma^n_h)\) on \((t_{n-1}, t_n], n \geq 1\), then, we have the following result:

**Theorem 3.13. (Convergence)** There exists a subsequence \((k', h')\) of \((k, h)\), with \(k', h' \downarrow 0\), and a weak solution \((u, \sigma)\) of (7) in \((0, +\infty)\), such that \((\tilde{u}_{h',k'}, \tilde{\sigma}_{h',k'})\) and \((u^n_{h',k'}, \sigma^n_{h',k'})\) converge to \((u, \sigma)\) weakly-* in \(L^\infty(0, +\infty; L^2(\Omega) \times L^2(\Omega))\), weakly in \(L^2(0, T; H^1(\Omega) \times H^1(\Omega))\) and strongly in \(L^2(0, T; L^2(\Omega) \times L^2(\Omega))\), for any \(T > 0\).

Note that, since the positivity of \(u^n_h\) cannot be assured, then the positivity of the limit function \(u\) cannot be proven. Moreover, if we introduce the functions:

- \(\tilde{v}_{h,k}\) are continuous functions on \([0, +\infty)\), linear on each interval \((t_{n-1}, t_n)\) and equal to \(v^n_h\), at \(t = t_n, n \geq 0\);

- \(v^n_{h,k}\) as the piecewise constant functions taking values \(v^n_h\) on \((t_{n-1}, t_n], n \geq 1\), proceeding as in Lemma 4.12 of [5] and taking into account the estimates (48)-(49), the following result can be proved:

**Corollary 3.14.** There exists a subsequence \((k', h')\) of \((k, h)\), with \(k', h' \downarrow 0\), and a weak solution \(v\) of (8) in \((0, +\infty)\), such that \(\tilde{v}_{h',k'}\) and \(v^n_{h',k'}\) converge to \(v\) weakly-* in \(L^\infty(0, +\infty; L^2(\Omega))\), weakly in \(L^2(0, T; H^1(\Omega))\) and strongly in \(L^2(0, T; L^2(\Omega))\), for any \(T > 0\).
**Remark 3.15.** From the equivalence of problems (1) and (7)-(8) established in [5], and taking into account Theorem 3.13 and Corollary 3.14, we deduce that the limit pair \((u, v)\) is a weak-strong solution of problem (1).

### 3.3 Uniform Strong Estimates

In this subsection, we are going to establish a priori estimates in strong norms for any solution \((u_n, \sigma_n)\) of the scheme US and \(v^n_h\) of (35). We will assume the estimate

\[
\|(u^n_h, \sigma^n_h)\|_1^2 \leq K_0, \quad \forall n \geq 0,
\]

with \(K_0 > 0\) a constant depending on the initial data, but independent of \((k, h)\) and \(n\). Note that estimate (54) can be proven in 2D domains, following line to line the proof of Theorem 4.20 in [5].

#### 3.3.1 Uniform Strong Estimates of the scheme US

**Theorem 3.16. (Strong estimates)** Let \((u^n_h, \sigma^n_h)\) be a solution of the scheme US satisfying the assumption (54). Then, the following estimate holds

\[
k^n \sum_{m=n_0+1}^{n+n_0} \left( \|(\delta_t u^n_m, \delta_t \sigma^n_m)\|_0^2 + \|(u^n_m, \sigma^n_m)\|_{W^{1,6}}^2 \right) \leq K_1 + K_2(nk), \quad \forall n \geq 1,
\]

for any integer \(n_0 \geq 0\), with positive constants \(K_1, K_2\) depending on \((\Omega, u_0, \sigma_0)\), but independent of \((k, h)\) and \((n, n_0)\).

**Proof.** The proof follows as in Theorem 4.14 of [5], but in this case it is necessary to use the estimate

\[
\|(u^n_h, \sigma^n_h)\|_{W^{1,6}} \leq C(\|(\delta_t u^n_h, \delta_t \sigma^n_h)\|_0 + \|(u^n_h, \sigma^n_h)\|_1^3 + \|u^n_h\|_0),
\]

which is deduced from (21) and (22). \(\square\)

**Theorem 3.17. (More regular estimates)** Assume that \((u_0, \sigma_0) \in H^2(\Omega) \times H^2(\Omega)\). Under the hypothesis of Theorem 3.16, the following estimates hold

\[
\|(\delta_t u^n_h, \delta_t \sigma^n_h)\|_0^2 \leq K_3, \quad \forall n \geq 1,
\]

15
\[ k \sum_{m=n_0+1}^{n+n_0} \| (\delta_t u_h^m, \delta_t \sigma_h^m) \|^2_1 \leq K_4 + K_5(nk), \quad \forall n \geq 1, \tag{58} \]
\[ \|(u_h^n, \sigma_h^n)\|^2_{W^{1,6}} \leq K_6, \quad \forall n \geq 0, \tag{59} \]

for any integer \( n_0 \geq 0 \), with positive constants \( K_3, K_4, K_5, K_6 \) depending on data \((\Omega, u_0, \sigma_0)\), but independent of \((k, h)\) and \((n, n_0)\).

**Proof.** The proof follows as in Theorem 4.16 of [5], but in this case, in order to obtain (59) it is necessary to use (56). \( \square \)

**Remark 3.18.** In particular, from (59) one has \( \|(u_h^n, \sigma_h^n)\|_{L^{\infty}} \leq K_7 \) for all \( n \geq 0 \), with \( K_7 > 0 \) a constant independent of \((k, h)\) and \( n \).

### 3.3.2 Uniform Strong estimates of \( v_h^n \)

**Theorem 3.19.** (Strong estimates for \( v_h^n \)) Assume (54) and let \( v_h^n \) be the solution of (35). Then, the following estimates hold

\[ \|v_h^n\|^2_1 \leq C_1, \quad \forall n \geq 0, \tag{60} \]
\[ k \sum_{m=n_0+1}^{n+n_0} (\|\delta_t v_h^m\|^2_0 + \|\tilde{A}_h v_h^m\|^2_0) \leq C_1 + C_2(nk), \quad \forall n \geq 1, \tag{61} \]

for any integer \( n_0 \geq 0 \), with positive constants \( C_1, C_2 \) depending on \( \Omega, u_0, \sigma_0, v_0 \), but independent of \((k, h)\) and \((n, n_0)\).

**Proof.** Testing (35) by \( \tilde{A}_h v_h^n \) and \( \delta_t v_h^n \), and using the Hölder and Young inequalities, we obtain

\[ \delta_t (\|v_h^n\|^2_1) + \frac{1}{2} \|\tilde{A}_h v_h^n\|^2_0 + \frac{1}{2} \|v_h^n\|^2_0 \leq \|u_h^n\|^4_{L^4}, \tag{62} \]

which, taking into account (21)2 and (54), in particular implies

\[ (1 + Ck)\|v_h^n\|^2_1 - \|v_h^{n-1}\|^2_1 \leq kK_0^2. \]

Thus, from Lemma 2.1, we deduce

\[ \|v_h^n\|^2_1 \leq (1 + Ck)^{-n}\|v_h^0\|^2_1 + C K_0^2 \leq \|v_h^0\|^2_1 + C K_0^2, \quad \forall n \geq 0, \]

16
which implies (60). Moreover, multiplying (62) by \( k \) and adding from \( m = n_0 + 1 \) to \( m = n + n_0 \), using (54) and (60), we deduce (61).

\[ \square \]

**Theorem 3.20.** (More regular estimates for \( v^n_h \)) Assume that \( v_0 \in H^2(\Omega) \). Under hypothesis of Theorems 3.17 and 3.19, the following estimates hold

\[
\|\delta_t v^n_h\|_0^2 \leq C_3, \quad \forall n \geq 1,
\]

\[
k \sum_{m=n_0+1}^{n+n_0} \|\delta_t v^m_h\|_1^2 \leq C_4 + C_5(nk), \quad \forall n \geq 1,
\]

\[
\|v^n_h\|_{W^{1,6}} \leq C_6, \quad \forall n \geq 0,
\]

for any integer \( n_0 \geq 0 \), with positive constants \( C_3, C_4, C_5, C_6 \) depending on data \( \Omega, u_0, \sigma_0, v_0 \), but independent of \((k, h)\) and \((n, n_0)\).

**Proof.** Denote by \( \bar{v}^n_h = \delta_t v^n_h \). Then, making the time discrete derivative of (35) (using \( \delta_t (u^n_h)^2 = (u^n_h + u^{n-1}_h)\delta_t u^n_h \)), testing by \( \bar{v}^n_h \) and using (54) and (57), we obtain

\[
\frac{1}{2} \frac{d}{dt}(\|\bar{v}^n_h\|_0^2) + \frac{1}{2} \|\bar{v}^n_h\|_1^2 \leq C\|u^n_h + u^{n-1}_h\|_{L^2}^2 \|\delta_t u^n_h\|_0^2 \leq C.
\]

In particular,

\[
(1 + k)\|\bar{v}^n_h\|_0^2 - \|\bar{v}^n_h\|_0^2 \leq kC.
\]

Then, from Lemma 2.1, we deduce

\[
\|\bar{v}^n_h\|_0^2 \leq (1 + k)^{-(n-1)}\|\bar{v}^1_h\|_0^2 + C, \quad \forall n \geq 1.
\]

Observe that from (35) we have

\[
(\delta_t v^1_h, \bar{v}_h) + (\bar{A}_h(v^1_h - v^0_h), \bar{v}_h) + (\bar{A}_h v^0_h, \bar{v}_h) = ((u^1_h)^2, \bar{v}_h), \quad \forall \bar{v}_h \in V_h.
\]

Then, testing (68) by \( \bar{v}_h = \delta_t v^1_h \) and using the Hölder and Young inequalities and (54), we can obtain

\[
\|\delta_t v^1_h\|_0^2 \leq C\|\bar{A}_h v^0_h\|_0^2 + C\|u^1_h\|_{L^4}^4.
\]
Moreover, considering the linear and continuous operator  $\tilde{A}_h : H^1(\Omega) \to V_h$ defined as

\[(\tilde{A}_h v, \tilde{v}_h) = (\nabla v, \nabla \tilde{v}_h) + (v, \tilde{v}_h), \quad \forall \tilde{v}_h \in V_h,\]

(which is an extension of $\tilde{A}_h$ to $H^1(\Omega)$), using the inverse inequality $\|v_h\|_1 \leq \frac{1}{h} \|v_h\|_0$ for all $v_h \in V_h$, and the interpolation error (19), we have

\[
\|\tilde{A}_h v_h^0\|_0 \leq \|\tilde{A}_h^* (R_h v_0 - v_0)\|_0 + \|\tilde{A}_h^* v_0\|_0
\]

\[
\leq C \frac{1}{h} \|\nabla (R_h v_0 - v_0)\|_0 + C \|R_h v_0 - v_0\|_0 + \|v_0\|_2 \leq C \|v_0\|_2. \tag{70}
\]

Thus, using (54) and (70) in (69), we conclude that $\|\tilde{v}_h^1\|_0^2 \leq C$, where the constant $C$ is independent of $(k, h)$. Therefore, using this fact in (67), we conclude (63). Moreover, multiplying (66) by $k$ and adding from $m = n_0 + 1$ to $m = n + n_0$, using (63), we deduce (64). Finally, taking into account (21), we have

\[
\|v_h^n\|_{W^{1,6}} \leq \|\tilde{A}_h v_h^n\|_0 \leq \|\delta_t v_h^n\|_0 + \|u_h^n\|_{L^4}^2,
\]

which, taking into account that from (20) we have $\|v_h^0\|_{W^{1,6}} = \|R_h^* v_0\|_{W^{1,6}} \leq C \|v_0\|_2$, and using (54) and (63), implies (65). □

### 3.4 Error estimates at finite time

In this subsection, we will obtain error estimates for any solution $(u_h^n, \sigma_h^n)$ of the scheme US and $v_h^n$ of (35), with respect to sufficiently regular solutions $(u, \sigma)$ of (7) and $v$ of (8) respectively. In our analysis, in order to obtain optimal error estimates we need to assume that both spaces $U_h, \Sigma_h$ are generated by $P_m$-continuous FE and $V_h$ is generated by $P_{m+1}$-continuous FE, with $m \geq 1$. This is a natural assumption taking into account that the energy norm for $v$ in the continuous model has one order greater than the energy norms for $u, \sigma$.

#### 3.4.1 Error estimates for scheme US

We start introducing the following notations for the errors at $t = t_n$: $e_u^n = u(t_n) - u_h^n$ and $e_{\sigma}^n = \sigma(t_n) - \sigma_h^n$, and for the discrete in time derivative of these errors: $\delta_t e_u^n = e_u^n - e_u^{n-1} - \frac{k}{h}$ and $\delta_t e_{\sigma}^n = e_{\sigma}^n - e_{\sigma}^{n-1} - \frac{k}{h}$. Then, subtracting (7) at $t = t_n$ and the scheme US, we have that $(e_u^n, e_{\sigma}^n)$
satisfies

\[(\delta_t e^n_u, \bar{u}_h) + (\nabla e^n_u, \nabla \bar{u}_h) + (e^n_u \sigma(t_n) + u^n_h e^n_{\sigma}, \nabla \bar{u}_h) = (\xi^n_1, \bar{u}_h), \quad \forall \bar{u}_h \in U_h, \tag{71}\]

\[(\delta_t e^n_{\sigma}, \bar{\sigma}_h) + (B e^n_{\sigma}, \bar{\sigma}_h) = 2(e^n_u \nabla u(t_n) + u^n_h \nabla e^n_{\sigma}, \bar{\sigma}_h) + (\xi^n_2, \bar{\sigma}_h), \quad \forall \bar{\sigma}_h \in \Sigma_h, \tag{72}\]

where \(\xi^n_1, \xi^n_2\) are the consistency errors associated to the scheme US, that is, \(\xi^n_1 = \delta_t(u(t_n)) - u(t_n)\) and \(\xi^n_2 = \delta_t(\sigma(t_n)) - \sigma(t_n)\). Now, considering the interpolation operators \(R^n_u\) and \(R^n_{\sigma}\) defined in (14)-(15), we decompose \(e^n_u\) and \(e^n_{\sigma}\) as follows

\[e^n_u = (I - R^n_u)u(t_n) + R^n_u u(t_n) - u^n_h = e^n_{u,i} + e^n_{u,h}, \tag{73}\]

\[e^n_{\sigma} = (I - R^n_{\sigma})\sigma(t_n) + R^n_{\sigma} \sigma(t_n) - \sigma^n_h = e^n_{\sigma,i} + e^n_{\sigma,h}, \tag{74}\]

where \(e^n_{u,i}\) is the interpolation error and \(e^n_{u,h}\) is the discrete error of \(u\). Then, taking into account (14)-(15), from (71)-(74) we have

\[
\begin{align*}
(\delta_t e^n_{u,h}, \bar{u}_h) + (\nabla e^n_{u,h}, \nabla \bar{u}_h) + (e^n_{u,h} \sigma(t_n) + u^n_h e^n_{\sigma,h}, \nabla \bar{u}_h) &= (\xi^n_1, \bar{u}_h) \quad - (\delta_t e^n_{u,i}, \bar{u}_h) - (e^n_{u,i} \sigma(t_n) + u^n_h e^n_{\sigma,i}, \nabla \bar{u}_h) + (e^n_{u,i}, \bar{u}_h), \quad \forall \bar{u}_h \in U_h, \tag{75} \\
(\delta_t e^n_{\sigma,h}, \bar{\sigma}_h) + (B e^n_{\sigma,h}, \bar{\sigma}_h) &= (\xi^n_2, \bar{\sigma}_h) + 2(e^n_{u,h} \nabla u(t_n) + u^n_h \nabla e^n_{\sigma,h}, \bar{\sigma}_h) \\
&+ 2(e^n_{u,i} \nabla u(t_n) + u^n_h \nabla e^n_{\sigma,i}, \bar{\sigma}_h) - (\delta_t e^n_{\sigma,i}, \bar{\sigma}_h), \quad \forall \bar{\sigma}_h \in \Sigma_h. \tag{76}
\end{align*}
\]

Notice that \(\int_{\Omega} e^n_{u,h} = 0\) (since \(u^n_h = R^n_{\sigma} u_0\) and from (14) \(\int_{\Omega} R^n_{\sigma} u(t_n) = \int_{\Omega} u(t_n) = m_0\)), hence the following norms are equivalents: \(\|\nabla e^n_{u,h}\|_0 \simeq \|e^n_{u,h}\|_1\).

**Theorem 3.21.** We assume that there exists \((u, \sigma)\) an exact solution of (7) with the following regularity:

\[
\begin{align*}
(u, \sigma) &\in L^\infty(0, T; H^{m+1}(\Omega) \times H^{m+1}(\Omega)), \quad (u_t, \sigma_t) \in L^2(0, T; H^{m+1}(\Omega) \times H^{m+1}(\Omega)), \\
(u_{tt}, \sigma_{tt}) &\in L^2(0, T; H^1(\Omega)' \times H^1_0(\Omega)').
\end{align*}
\]

Let \((u_h^n, \sigma_h^n)\) be a solution of the scheme US. Then, if

\[k(\|\tilde{u}(u, \sigma)\|_{L^\infty(H^1)} + \|\tilde{u}(u, \sigma)\|_{L^\infty(H^2)}) \quad \text{is small enough}, \tag{78}\]
the following a priori error estimate holds

$$\|(e_{u,h}^n, e_{\sigma,h}^n)\|_{L^2 \cap H^1} \leq C(T)(k + h^{m+1})$$  \hfill (79)

where \(C(T) = K_1 T \exp(K_2 T)\), with \(K_1, K_2 > 0\) independent of \((k, h)\).

Recall that \(u, \sigma\) are approximated by \(\mathbb{P}_m\)-continuous FE.

**Proof.** Taking \(\bar{u}_h = e_{u,h}^n\) in (75), \(\bar{\sigma}_h = \frac{1}{2}e_{u,h}^n\) in (76) and adding, the terms \((u_h^n \nabla e_{u,h}^n, e_{\sigma,h}^n)\) cancel, and we obtain

$$\delta_t \left( \frac{1}{2} \|e_{u,h}^n\|_0^2 + \frac{1}{4} \|e_{\sigma,h}^n\|_0^2 \right) + \frac{1}{2} \|e_{u,h}^n, e_{\sigma,h}^n\|_1^2 = (\xi_{1}^n, e_{u,h}^n) + \frac{1}{2} (\xi_{2}^n, e_{\sigma,h}^n) - (\delta_t e_{u,i}^n, e_{u,h}^n)$$

$$- \frac{1}{2} \left( (\delta_t e_{\sigma,i}^n, e_{\sigma,h}^n) - (e_{u,h}^n, \sigma(t_n) \cdot \nabla e_{u,h}^n - \nabla u(t_n) \cdot e_{\sigma,h}^n) - (e_{u,i}^n, \sigma(t_n) \cdot \nabla e_{u,h}^n - \nabla u(t_n) \cdot e_{\sigma,h}^n) \right)$$

$$- (u_h^n, e_{\sigma,i}^n \cdot \nabla e_{u,h}^n - \nabla e_{u,i}^n \cdot e_{\sigma,h}^n) + (e_{u,i}^n, e_{u,h}^n) = \sum_{m=1}^8 I_m.$$  \hfill (80)

Then, using the Hölder and Young inequalities, the 3D interpolation inequality (12), the interpolation errors (17)-(18), the stability property (20) and the hypothesis (77), we control the terms on the right hand side of (80) as follows

$$I_1 + I_2 \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C_\varepsilon \|\xi_{1}^n, \xi_{2}^n\|_{H^1 \times (H^1)^2}^2$$

$$\quad \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C \int_{t_{n-1}}^{t_n} \|(u(t), \sigma(t))\|_{H^1 \times (H^1)^2}^2 dt,$$  \hfill (81)

$$I_5 \leq \|e_{u,h}^n\|_{L^2}(\|\nabla u(t_n)\|_0 \|e_{\sigma,h}^n\|_{L^0} + \|\nabla \cdot \sigma(t_n)\|_0 \|e_{u,h}^n\|_{L^0})$$

$$\quad \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C_\varepsilon \|\nabla u(t_n), \nabla \cdot \sigma(t_n)\|_{L^2}^2 \|e_{u,h}^n\|_{0}^2,$$  \hfill (82)

$$I_6 \leq \|e_{u,i}^n\|_0(\|\nabla e_{u,h}^n\|_0 \|\sigma(t_n)\|_{L^\infty} + \|\nabla u(t_n)\|_{L^2} \|e_{\sigma,h}^n\|_{L^0})$$

$$\quad \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C_\varepsilon \|\nabla u(t_n), \sigma(t_n)\|_{L^2}^2 \|e_{u,i}^n\|_0^2$$

$$\quad \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C \varepsilon (2^{m+1}) \|u(t_n)\|_{m+1}^2 \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C \varepsilon (2^{m+1}),$$  \hfill (83)

$$I_7 \leq \|(e_{u,i}^n, e_{\sigma,i}^n \cdot \nabla e_{u,h}^n - \nabla e_{u,i}^n \cdot e_{\sigma,h}^n)\| + \|(R_h^n u(t_n), e_{\sigma,i}^n \cdot \nabla e_{u,h}^n - \nabla e_{u,i}^n \cdot e_{\sigma,h}^n)\|$$

$$\quad \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C_\varepsilon \|e_{u,h}^n\|_0^2 \|(e_{u,i}^n, e_{\sigma,i}^n)\|_{W^{1,3} \times L^\infty}^2 + C_\varepsilon \|R_h^n u(t_n)\|_{H^{1,3} \cap L^\infty}^2 \|e_{u,i}^n, e_{\sigma,i}^n\|_0^2$$

$$\quad \leq \varepsilon \|(e_{u,h}^n, e_{\sigma,h}^n)\|_1^2 + C \|(u(t_n), \sigma(t_n))\|_{L^2}^2 \|e_{u,h}^n\|_0^2 \varepsilon (2^{m+1}) \|u(t_n), \sigma(t_n)\|_{m+1}^2 \|u(t_n)\|_2^2$$

20
Then, multiplying (77) into account (76)

**Remark 3.22.** Under hypothesis of Theorem 3.21, one has in particular

\[ \| (u_n^1, \sigma_n^1) \|^2_1 \leq C + C(T) \left( k + \frac{h^{2(m+1)}}{k} \right), \]

Therefore, under the hypothesis

\[ \frac{h^{2(m+1)}}{k} \leq C, \tag{88} \]

we have the estimate

\[ \| (u_n^1, \sigma_n^1) \|^2_1 \leq C, \tag{89} \]
hence the hypothesis (36) providing uniqueness of the scheme is equivalent to \( k \) small enough.

Finally, since for any choice of \((k, h)\) either (37) (see Remark 3.5) or (88) holds, one has the uniqueness of \((u^n_h, \sigma^n_h)\) solution of (34) only imposing \( k \) small enough.

### 3.4.2 Error estimates for \( v^n_h \) solution of (35)

We introduce the following notation for the errors in \( t = t_n: e^n_v = v(t_n) - v^n_h, \) and for the discrete in time derivative of this error: \( \delta_t e^n_v = \frac{e^n_v - e^{n-1}_v}{k}. \) Then, subtracting (8) at \( t = t_n \) and (35), we have that \( e^n_v \) satisfies

\[
(\delta_t e^n_v, \bar{v}_h) + \langle A e^n_v, \bar{v}_h \rangle = \langle (u(t_n) + u^n_h)e^n_u, \bar{v}_h \rangle + \langle \xi^n_3, \bar{v}_h \rangle, \quad \forall \bar{v}_h \in V_h,
\]

(90)

where \( \xi^n_3 \) is the consistency error associated to (35), that is, \( \xi^n_3 = \delta_t(v(t_n)) - v_t(t_n). \) Now, considering the interpolation operator \( \mathcal{R}^n_h \) defined in (16), we decompose \( e^n_v \) as follows

\[
e^n_v = (\mathcal{I} - \mathcal{R}^n_h)v(t_n) + \mathcal{R}^n_h v(t_n) - v^n_h = e^n_{v,i} + e^n_{v,h},
\]

(91)

where \( e^n_{v,i} \) is the interpolation error and \( e^n_{v,h} \) is the discrete error of \( v. \) Then, taking into account (16), from (90)-(91) we have

\[
\left( \delta_t e^n_{v,h}, \bar{v}_h \right) + \left( A e^n_{v,h}, \bar{v}_h \right) = \left( \xi^n_3, \bar{v}_h \right) + \langle (u(t_n) + u^n_h)(e^n_{u,h} + e^n_{u,i}), \bar{v}_h \rangle - \left( \delta_t e^n_{v,i}, \bar{v}_h \right), \quad \forall \bar{v}_h \in V_h\]

(92)

**Theorem 3.23.** Under hypothesis of Theorem 3.21. Let \( v^n_h \) be the solution of (35), and assume the following regularity for \( v \) exact solution of (8):

\[
(v_t, v_{tt}) \in L^2(0; T; H^{m+2}(\Omega) \times H^1(\Omega')).
\]

(93)

Then, the a priori error estimate holds

\[
\|e^n_{v,h}\|_{L^\infty L^2 \cap L^2 H^1} \leq C(T)(k + h^{m+1}),
\]

(94)

where \( C(T) = K_1 T \exp(K_2 T), \) with \( K_1, K_2 > 0 \) independent of \((k, h)\).

**Proof.** Taking \( \bar{v}_h = e^n_{v,h} \) in (92) and using the Hölder and Young inequalities, we obtain

\[
\delta_t \left( \frac{1}{2} \|e^n_{v,h}\|_0^2 \right) + \frac{k}{2} \|\delta_t e^n_{v,h}\|_0^2 + \frac{1}{2} \|e^n_{v,h}\|_1^2 \leq C \|\xi^n_3\|_{(H^1)'},
\]

22
+C\|u(t_n) + u^n_h\|^2_{L^2}\left(\|e^n_{u,h}\|^2_0 + \|e^n_{u,i}\|^2_0\right) + C\|(I - R^n_{h})\delta v(t_n)\|^2_0. \tag{95}

Using (17), (19) and proceeding as in (81) and (86), we bound the terms on the right hand side in (95) and we deduce

\delta_t \left(\|e^n_{v,h}\|^2_0 + \|e^n_{u,h}\|^2_1\right) \leq Ck \int_{t_{n-1}}^{t_n} \|v_{tt}(t)\|^2_{(H^1)} dt + C\|u(t_n) + u^n_h\|^2_{L^3}\|e^n_{u,h}\|^2_0

+C\|u(t_n) + u^n_h\|^2_{L^3} + C\int_{t_{n-1}}^{t_n} \|v_t\|^2_{m+2} dt. \tag{96}

Then, multiplying (96) by k, adding from n = 1 to n = r, we obtain (recall $e^0_{v,h} = 0$):

\|e^n_{v,h}\|^2_0 + \sum_{n=1}^{r} \|e^n_{u,h}\|^2_1 \leq Ck^2 \int_0^{t_r} \|v_{tt}(t)\|^2_{(H^1)} dt + C\int_0^{t_r} \|u(t_n) + u^n_h\|^2_{L^3} dt

+C\int_0^{t_r} \|v_t\|^2_{m+2} dt.

Then, using (42), (77), (93) and (79), we conclude (94). \qed

**Theorem 3.24.** Under hypothesis of Theorem 3.23, but assuming the regularity:

\[ v_{tt} \in L^2(0; T; L^2(\Omega)), \tag{97} \]

the a priori error estimate

\[ \|e^n_{v,h}\|_{H^1 \cap W^{1,6}} \leq C(T)(k + h^{m+1}) \tag{98} \]

holds, where $C(T) = K_1 T \exp(K_2 T)$, with $K_1, K_2 > 0$ independent of $(k, h)$.

**Proof.** Taking $\vec{\delta}_h = \delta_v e^n_{v,h}$ in (92) and using the Hölder and Young inequalities, we obtain

\[ \delta_t \left(\frac{1}{2}\|e^n_{v,h}\|^2_1 + \frac{k}{2}\|\delta e^n_{v,h}\|^2_1 + \frac{k}{2}\|\delta e^n_{u,h}\|^2_0\right) \leq \frac{k}{2}\|\delta e^n_{v,h}\|^2_1 + \frac{k}{2}\|\delta e^n_{u,h}\|^2_0 + C\|u(t_n) + u^n_h\|^2_{L^3}\|e^n_{u,h}\|^2_0

+C\|u(t_n) + u^n_h\|e^n_{u,i}\|^2_0 + C\|(I - R^n_{h})\delta v(t_n)\|^2_0. \tag{99} \]

Using the Hölder inequality, the interpolation error (17), the stability property (20) and the hypothesis (77), we have

\[ \|u(t_n) + u^n_h\|e^n_{u,i}\|^2_0 \leq C\|u(t_n) + R^n_{h}u(t_n)\|^2_{L^\infty}\|e^n_{u,i}\|^2_0 + C\|e^n_{u,h}\|^2_{L^6}\|e^n_{u,i}\|^2_0 \]


\[ \leq C h^{2(m+1)} + C \| e_{u,h}^n \|_{L^6}^2. \]  

(100)

Therefore, from (99), proceeding as in (81) and (86) and using (100), we deduce

\[
\begin{align*}
\delta t \left( \| e_{u,h}^n \|_1^2 \right) + \| \tilde{A}_h e_{u,h}^n \|_0^2 & \leq C k \int_{t_{n-1}}^{t_n} \| v(t) \|_0^2 dt \\
+ (C \| u(t_n) + u_h^n \|_{L^3}^2 + C') \| e_{u,h}^n \|_{L^6}^2 + C h^{2(m+1)} & \leq \frac{C h^{2(m+2)}}{k} \int_{t_{n-1}}^{t_n} \| v(t) \|_{L^{m+2}}^2 dt.
\end{align*}
\]

Now, in order to bound the term \( \| u(t_n) + u_h^n \|_{L^3}^2 \), we split the argument into two cases:

1. **Estimates assuming \( h << f(k) \) (\( h \) small enough with respect to \( k \))**:

From (79) we have that

\[
k \sum_{n=1}^{r} \| e_{u,h}^n \|_1^2 \leq C(T)(k^2 + h^{2(m+1)}),
\]

which implies that

\[
\| e_{u,h}^n \|_1 \leq C(T)(k^{1/2} + \frac{h^{m+1}}{k^{1/2}}).
\]

(101)

Moreover, using the interpolation inequality (12), (79), (20), (77) and (101), we obtain

\[
\begin{align*}
\| u(t_n) + u_h^n \|_{L^3}^2 & \leq C \| u(t_n) \|_{L^3}^2 + C \| \mathcal{R}_h^2 u(t_n) \|_{L^3}^2 + C \| e_{u,h}^n \|_{L^3}^2 \\
& \leq C + C(T)(k + h^{m+1})(k^{1/2} + \frac{h^{m+1}}{k^{1/2}}) \leq C
\end{align*}
\]

(102)

under the hypothesis

\[
\frac{h^{2(m+1)}}{k^{1/2}} \leq C.
\]

(103)

2. **Estimates assuming \( k << g(k) \) (\( k \) small enough with respect to \( h \))**:

Using the inverse inequality \( \| u_h \|_{L^3} \leq \frac{C}{k^{1/2}} \| u_h \|_0 \) for all \( u_h \in U_h \), (20), (77) and (79), we have that

\[
\begin{align*}
\| u(t_n) + u_h^n \|_{L^3}^2 & \leq C \| u(t_n) \|_{L^3}^2 + C \| \mathcal{R}_h^2 u(t_n) \|_{L^3}^2 + C \| e_{u,h}^n \|_{L^3}^2 \\
& \leq \frac{C}{h} \| e_{u,h}^n \|_0^2 + C \leq \frac{C(T)}{h}(k^2 + h^{2(m+1)}) + C \leq C
\end{align*}
\]

under the hypothesis

\[
\frac{k^2}{h} < C.
\]

(104)
Therefore, since for any choice of \((k,h)\) either (103) or (104) holds, we arrive at

\[
\delta_t \left( \|e_{v,h}^n\|_1^2 \right) + \|A_h e_{v,h}^n\|_0^2 \leq C k \int_{t_{n-1}}^{t_n} \|v_{tt}(t)\|_0^2 dt \\
+C \|e_{u,h}^n\|_0^2 + C h^{2(m+1)} + \frac{C h^{2(m+2)}}{k} \int_{t_{n-1}}^{t_n} \|v_t\|_{m+2}^2 dt.
\] (105)

Multiplying (105) by \(k\), adding from \(n = 1\) to \(n = r\), recalling that \(e_{v,h}^0 = 0\) and using (77), (93), (97) and (79), we conclude (98).

\[\square\]

4 Linear iterative methods to approach the Backward Euler scheme

In this section, we propose two different linear iterative methods to approach the Backward Euler scheme \(US\), which are an energy-stable Picard’s method and the Newton’s method. We prove the solvability and the convergence of these methods to the nonlinear scheme.

4.1 Picard Method

In order to approximate the solution \((u_h^n, \sigma_h^n)\) of the nonlinear scheme \(US\), we consider the following Picard method: Let \((u_h^{n-1}, \sigma_h^{n-1}) \in U_h \times \Sigma_h\) be fixed. Given \(u_h^{l-1} \in U_h\) (assuming \(u_h^0 = u_h^{n-1}\) at the first iteration step), find \((u_h^l, \sigma_h^l) \in U_h \times \Sigma_h\) solving the linear coupled problem:

\[
\begin{cases}
\frac{1}{k}(u_h^l, \bar{u}_h) + (\nabla u_h^l, \nabla \bar{u}_h) + (u_h^{l-1} \sigma_h^l, \nabla \bar{u}_h) = \frac{1}{k}(u_h^{n-1}, \bar{u}_h), \quad \forall \bar{u}_h \in U_h, \\
\frac{1}{k}(\sigma_h^l, \bar{\sigma}_h) + (B_h \sigma_h^l, \bar{\sigma}_h) - 2(u_h^{l-1} \nabla u_h^l, \bar{\sigma}_h) = \frac{1}{k}(\sigma_h^{n-1}, \bar{\sigma}_h), \quad \forall \bar{\sigma}_h \in \Sigma_h,
\end{cases}
\] (106)

until that the stopping criteria \(\max \left\{ \frac{\|u_h^l - u_h^{l-1}\|_0}{\|u_h^{l-1}\|_0}, \frac{\|\sigma_h^l - \sigma_h^{l-1}\|_0}{\|\sigma_h^{l-1}\|_0} \right\} \leq tol\) (with \(tol > 0\) being a tolerance parameter) be satisfied.

**Theorem 4.1. (Unconditional Unique Solvability)** There exists a unique \((u_h^l, \sigma_h^l)\) solution of (106).

**Proof.** Since (106) can be rewritten as a square linear algebraic system, it is sufficient to prove uniqueness. Suppose that there exist \((u_{h,1}^l, \sigma_{h,1}^l), (u_{h,2}^l, \sigma_{h,2}^l) \in U_h \times \Sigma_h\) two possible solutions of (106). Then defining \(u_h^l = u_{h,1}^l - u_{h,2}^l\) and \(\sigma_h^l = \sigma_{h,1}^l - \sigma_{h,2}^l\), we have that \((u_h^l, \sigma_h^l) \in U_h \times \Sigma_h\) satisfies

\[
\frac{1}{k}(u_h^l, \bar{u}_h) + (\nabla u_h^l, \nabla \bar{u}_h) + (u_h^{l-1} \sigma_h^l, \nabla \bar{u}_h) = 0, \quad \forall \bar{u}_h \in U_h,
\] (107)
\[
\frac{1}{k}(\sigma_h^l, \sigma_h) + (B_h \sigma_h^l, \sigma_h) - 2(u_h^{l-1} \nabla u_h^l, \sigma_h) = 0, \quad \forall \sigma_h \in \Sigma_h. \quad (108)
\]

Taking \( \bar{u}_h = u_h^l \) and \( \bar{\sigma}_h = \frac{1}{2} \sigma_h^l \) in (107)-(108) and adding, the terms \( (u_h^{l-1} \nabla u_h^l, \sigma_h) \) cancel, and we obtain
\[
\frac{1}{2k} \|u_h^l, \sigma_h^l\|_0^2 + \frac{1}{2} \|\nabla u_h^l, \sigma_h^l\|_{L^2 \times H^1}^2 \leq 0,
\]
and thus we conclude that \( \|u_h^l, \sigma_h^l\|_1 = 0 \), which implies \( u_{h,1}^l = u_{h,2}^l \) and \( \sigma_{h,1}^l = \sigma_{h,2}^l \).

**Theorem 4.2. (Local uniqueness of solution of scheme US and Convergence of Picard’s method)** Given \( (u_h^{n-1}, \sigma_h^{n-1}) \), there exists \( r > 0 \) (large enough) such that if
\[
k \|u_h^{n-1}, \sigma_h^{n-1}\|_1^4 \quad \text{and} \quad kr^4 \quad \text{are small enough},
\]
then the scheme US has a unique solution \( (u_h^n, \sigma_h^n) \) in \( \overline{B}_r((u_h^{n-1}, \sigma_h^{n-1})) := \{(u, \sigma) \in U_h \times \Sigma_h : \|u - u_h^{n-1}, \sigma - \sigma_h^{n-1}\|_1 \leq r\} \). Moreover, the sequence of solutions \( \{u_h^0, \sigma_h^0\}_{i \geq 0} \) of the iterative algorithm (106) (assuming \( (u_h^0, \sigma_h^0) = (u_h^{n-1}, \sigma_h^{n-1}) \) at the first iteration step), converges to \( (u_h^n, \sigma_h^n) \) strongly in \( H^1(\Omega) \).

**Proof.** We consider the operator \( R : U_h \to U_h \), given by \( R(\bar{u}) = u \), where \( (u, \sigma) \) satisfies (106) with \( u_h^{l-1} = \bar{u} \) and \( (u_h^l, \sigma_h^l) = (u, \sigma) \), that is,
\[
\frac{1}{k}(u, \bar{u}_h) + (\nabla u, \nabla \bar{u}_h) + (\bar{u} \sigma, \nabla \bar{u}_h) = \frac{1}{k}(u_h^{n-1}, \bar{u}_h), \quad \forall \bar{u}_h \in U_h,
\]
\[
\frac{1}{k}(\sigma, \bar{\sigma}_h) + (B_h \sigma, \bar{\sigma}_h) - 2(\bar{u} \nabla u, \bar{\sigma}_h) = \frac{1}{k}(\sigma_h^{n-1}, \bar{\sigma}_h), \quad \forall \bar{\sigma}_h \in \Sigma_h.
\]

Observe that from Theorem 4.1, we have that for any \( \bar{u} \in U_h \) there exists a unique \( (u, \sigma) \in U_h \times \Sigma_h \) solution of (110)-(111). Thus, \( R \) is well defined. Now, before to prove that \( R \) is contractive, we will construct a ball \( \overline{B}_r(u_h^{n-1}) = \{u \in U_h : \|u - u_h^{n-1}\|_1 \leq r\} \subset U_h \) such that \( R(\overline{B}_r(u_h^{n-1})) \subset \overline{B}_r(u_h^{n-1}) \). In order to define \( r \), we consider \( w = u - u_h^{n-1} \) and \( \tau = \sigma - \sigma_h^{n-1} \).

Then, from (110)-(111) we have that \( (w, \tau) \) verifies
\[
\frac{1}{k}(w, \bar{u}_h) + (\nabla w, \nabla \bar{u}_h) = -(\bar{u} \tau, \nabla \bar{u}_h) - (\bar{u} \sigma_h^{n-1}, \nabla \bar{u}_h), \quad \forall \bar{u}_h \in U_h,
\]
\[
\frac{1}{k}(\tau, \bar{\sigma}_h) + (B_h \tau, \bar{\sigma}_h) = 2(\bar{u} \nabla w, \bar{\sigma}_h) - (B_h \sigma_h^{n-1}, \bar{\sigma}_h) + 2(\bar{u} \nabla u_h^{n-1}, \bar{\sigma}_h), \quad \forall \bar{\sigma}_h \in \Sigma_h.
\]

Testing by \( \bar{u}_h = w \) and \( \bar{\sigma}_h = \frac{1}{2} \tau \) in (112)-(113) and adding, the terms \( (\bar{u} \nabla w, \tau) \) cancel, and
using the fact that \( \int_{\Omega} w = 0 \) as well as the 3D interpolation inequality (12), we obtain

\[
\frac{1}{2k} \| (w, \tau) \|^2 + \frac{1}{2} \| (w, \tau) \|^2 \leq \frac{1}{8} \| (w, \tau) \|^2 + C \| (u_{h}^{n-1}, \sigma_{h}^{n-1}) \|^2
\]

\[
+ \frac{1}{8} \| \tilde{u} - u_{h}^{n-1} \| + \frac{1}{8} \| u_{h}^{n-1} \| + \frac{1}{8} \| (w, \tau) \|^2 + C \| (u_{h}^{n-1}, \sigma_{h}^{n-1}) \|^2 \| (w, \tau) \|^2.
\]

Therefore, from (114) we deduce

\[
\left[ \frac{1}{2k} - C \| (u_{h}^{n-1}, \sigma_{h}^{n-1}) \|^2 \right] \| (w, \tau) \|^2 + \frac{1}{4} \| (w, \tau) \|^2 \leq C \| (u_{h}^{n-1}, \sigma_{h}^{n-1}) \|^2 + \frac{1}{8} \| \tilde{u} - u_{h}^{n-1} \|^2.
\]

Thus, if \( k < \frac{1}{2C \| (u_{h}^{n-1}, \sigma_{h}^{n-1}) \|^2} \), from (115) we conclude

\[
\| (w, \tau) \|^2 \leq C \| (u_{h}^{n-1}, \sigma_{h}^{n-1}) \|^2 + \frac{1}{2} \| \tilde{u} - u_{h}^{n-1} \|^2.
\]

Then, choosing \( r > 0 \) large enough such that

\[
C \| (u_{h}^{n-1}, \sigma_{h}^{n-1}) \|^2 \leq \frac{1}{2} r^2,
\]

from (116) we deduce that \( R(\overline{B}_{r}(u_{h}^{n-1})) \subseteq \overline{B}_{r}(u_{h}^{n-1}) \). Then, we take the restriction of \( R \) to \( \overline{B}_{r}(u_{h}^{n-1}) \), that is, \( R_{r} : \overline{B}_{r}(u_{h}^{n-1}) \rightarrow \overline{B}_{r}(u_{h}^{n-1}) \). Let’s prove that \( R_{r} \) is contractive. Let \( \tilde{u}_{1}, \tilde{u}_{2} \in \overline{B}_{r}(u_{h}^{n-1}) \), and \( (u_{1}, \sigma_{1}) \) and \( (u_{2}, \sigma_{2}) \) solutions of (110)-(111) corresponding to \( \tilde{u}_{1} \) and \( \tilde{u}_{2} \) respectively (i.e., \( R_{r}(\tilde{u}_{1}) = u_{1} \) and \( R_{r}(\tilde{u}_{2}) = u_{2} \)). Then, from (110)-(111) we have that \( (u_{1} - u_{2}, \sigma_{1} - \sigma_{2}) \in U_{h} \times \Sigma_{h} \) satisfies

\[
\frac{1}{k} (u_{1} - u_{2}, \tilde{u}_{h}) + (\nabla (u_{1} - u_{2}), \nabla \tilde{u}_{h}) + (\tilde{u}_{1}(\sigma_{1} - \sigma_{2}), \nabla \tilde{u}_{h}) + ((\tilde{u}_{1} - \tilde{u}_{2})\sigma_{2}, \nabla \tilde{u}_{h}) = 0, \ \forall \tilde{u}_{h} \in U_{h},
\]

\[
\frac{1}{k} (\sigma_{1} - \sigma_{2}, \tilde{\sigma}_{h}) + (B_{h}(\sigma_{1} - \sigma_{2}), \tilde{\sigma}_{h}) - 2(\tilde{u}_{1} \nabla (u_{1} - u_{2}), \tilde{\sigma}_{h}) - 2((\tilde{u}_{1} - \tilde{u}_{2}) \nabla u_{2}, \tilde{\sigma}_{h}) = 0, \ \forall \tilde{\sigma}_{h} \in \Sigma_{h}.
\]

Testing by \( \tilde{u}_{h} = u_{1} - u_{2}, \ \tilde{\sigma}_{h} = \frac{1}{2}(\sigma_{1} - \sigma_{2}) \) and adding, the terms \((\tilde{u}_{1}(\sigma_{1} - \sigma_{2}), \nabla (u_{1} - u_{2}))\) cancel, and using the Hölder and Young inequalities, the 3D interpolation inequality (12) and taking into account that \( \int_{\Omega} u_{1} - u_{2} = 0 \), we obtain

\[
\frac{1}{2k} \| (u_{1} - u_{2}, \sigma_{1} - \sigma_{2}) \|_{0}^{2} + \| u_{1} - u_{2} \|_{0}^{2} + \frac{1}{2} \| \sigma_{1} - \sigma_{2} \|_{1}^{2}
\]

\[
\leq C \| \tilde{u}_{1} - \tilde{u}_{2} \|_{1} (\| \sigma_{2} \| \| u_{1} - u_{2} \|_{L^3} + \| u_{2} \|_{1} \| \sigma_{1} - \sigma_{2} \|_{L^3})
\]

27
\[ \leq \frac{1}{4} \| \bar{u}_1 - \bar{u}_2 \|^2 + \frac{1}{2} \| u_1 - u_2 \|^2 + \frac{1}{4} \| \sigma_1 - \sigma_2 \|^2 + C \| (u_1 - u_2, \sigma_1 - \sigma_2) \|_0^2 \| (u_2, \sigma_2) \|_1^4 \]

and thus, we deduce that

\[
\frac{1}{k} \| (u_1 - u_2, \sigma_1 - \sigma_2) \|_0^2 + \| u_1 - u_2 \|_1^2 + \frac{1}{2} \| \sigma_1 - \sigma_2 \|_1^2 \\
\leq \frac{1}{2} \| \bar{u}_1 - \bar{u}_2 \|_1^2 + C \| (u_1 - u_2, \sigma_1 - \sigma_2) \|_0^2 \| (u_2, \sigma_2) \|_1^4. \tag{118}
\]

Therefore, since from (116) and (117) we have \( \| (u_2, \sigma_2) \|_1^4 \leq C (r^4 + \| (u_{h-1}^n, \sigma_{h-1}^n) \|_1^4) \), if \( \frac{1}{2k} > C r^4 \) and \( \frac{1}{2k} > C \| (u_{h-1}^n, \sigma_{h-1}^n) \|_1^4 \), from (118) we have

\[ \| R_r(\bar{u}_1) - R_r(\bar{u}_2) \|_1^2 \leq \frac{1}{2} \| \bar{u}_1 - \bar{u}_2 \|_1^2, \]

which implies that \( R_r \) is contractive. Then, as a consequence of the Banach fixed point theorem, we conclude that there exists a unique fixed point of \( R_r, R_r(u) = u \). Thus, \( (u, \sigma) \) is the unique solution of the scheme \( \mathbf{U}S \) in \( B_r(u_{h-1}) \). Additionally, we conclude that the sequence of solutions \( \{ u^l_h, \sigma^l_h \}_{l \geq 0} \) of the iterative algorithm (106), where \( (u^0_h, \sigma^0_h) = (u_{h-1}^n, \sigma_{h-1}^n) \), converges to the solution \( (u^n_h, \sigma^n_h) \).

**Remark 4.3.** In the case of 2D Domains, from estimate (54), the restriction (109) can be relaxed to \( k \leq K_0 \), where \( K_0 \) is a constant depending on data \((\Omega, u_0, \sigma_0)\), but independent of \((k, h)\) and \( n \).

**Remark 4.4.** We have that the restriction (109) is equivalent to (36). Therefore, under hypothesis of Theorem 3.21 and arguing as in Remark 3.22, the conclusion of Theorem 4.2 remains true only assuming \( k \) small enough.

### 4.2 Newton’s Method

In this subsection, in order to approximate the solution \( (u^n_h, \sigma^n_h) \) of the nonlinear scheme \( \mathbf{U}S \), we consider Newton’s algorithm: Let \( (u_{h-1}^n, \sigma_{h-1}^n) \in U_h \times \Sigma_h \) be fixed. Given \( (u_{h-1}^l, \sigma_{h-1}^l) \in U_h \times \Sigma_h \),
In the following theorem, we will use this lemma to prove the convergence of Newton’s method. Let

$$\text{Theorem 4.6. (Conditional convergence of Newton’s Method)}$$

be a Banach space and consider a sequence \( \{e_l\}_{l \geq 0} \subseteq X \), such that

$$\|e_l\|_X^2 \leq C (\|e_{l-1}\|_X^2)^2, \quad \forall l \geq 1 \quad \text{and} \quad \|e_0\|_X^2 \text{ is small enough.}$$

Then, \( e_l \) converges to 0 as \( l \to +\infty \) in the \( X \)-norm.

In the following theorem, we will use this lemma to prove the convergence \((u_h^0, \sigma_h^0) \to (u_h^n, \sigma_h^n)\) in the \( H^1(\Omega) \)-norm.

**Theorem 4.6. (Conditional convergence of Newton’s Method)** Let \((u_h^n, \sigma_h^n)\) be a fixed solution of the scheme \( U^S \) and let \((u_h^l, \sigma_h^l)\) be any solution of (119). There exists \( \delta_0 > 0 \) small enough such that if

$$\|e_u^0, e_\sigma^0\|_1^2 \leq \delta_0, \quad k \|(u_h^n, \sigma_h^n)\|_1^4 \quad \text{and} \quad k(\delta_0)^2 \text{ are small enough,}$$

then \( \{u_h^l, \sigma_h^l\}_{l \geq 0} \) converges to \((u_h^n, \sigma_h^n)\) in the \( H^1(\Omega) \)-norm as \( l \to +\infty \).

**Proof.** We can define problem (34) in a vectorial way,

$$\begin{align*}
(0, 0) &= \langle F(u_h^n, \sigma_h^n), (u_h, \sigma_h) \rangle = \langle F_1(u_h^n, \sigma_h^n), \bar{u}_h \rangle + \langle F_2(u_h^n, \sigma_h^n), \bar{\sigma}_h \rangle,
\end{align*}$$

where each \( F_i(u_h^n, \sigma_h^n) \) corresponds with the equation (34), \( i = 1, 2 \). Therefore, Newton’s method (119) reads

$$\langle F'(u_h^{i-1}, \sigma_h^{i-1})(u_h^i - u_h^{i-1}, \sigma_h^i - \sigma_h^{i-1}), (\bar{u}_h, \bar{\sigma}_h) \rangle = -\langle F(u_h^{i-1}, \sigma_h^{i-1}), (\bar{u}_h, \bar{\sigma}_h) \rangle,$$
which can be rewritten as

\[
(0, 0) = \left(\langle F_1(u_h^{l-1}, \sigma_h^{l-1}), \tilde{u}_h \rangle, \langle F_2(u_h^{l-1}, \sigma_h^{l-1}), \tilde{\sigma}_h \rangle\right)
+ \left(\langle F'_1(u_h^{l-1}, \sigma_h^{l-1})(u_h^l - u_h^{l-1}, \sigma_h^{l-1} - \sigma_h^{l-1}), \tilde{u}_h \rangle, \langle F'_2(u_h^{l-1}, \sigma_h^{l-1})(u_h^l - u_h^{l-1}, \sigma_h^{l-1} - \sigma_h^{l-1}), \tilde{\sigma}_h \rangle\right) + \frac{1}{2} \left(\langle (u_h^l - u_h^{l-1}, \sigma_h^{l-1} - \sigma_h^{l-1})^t F''_1(u_h^{n+\varepsilon}, \sigma_h^{n+\varepsilon})(u_h^l - u_h^{l-1}, \sigma_h^{l-1} - \sigma_h^{l-1}), \tilde{u}_h \rangle, \langle (u_h^l - u_h^{l-1}, \sigma_h^{l-1} - \sigma_h^{l-1})^t F''_2(u_h^{n+\varepsilon}, \sigma_h^{n+\varepsilon})(u_h^l - u_h^{l-1}, \sigma_h^{l-1} - \sigma_h^{l-1}), \tilde{\sigma}_h \rangle\right),
\]

(122)

Moreover, from a vectorial Taylor’s formula of \( F(u_h^n, \sigma_h^n) \) with center at \((u_h^{l-1}, \sigma_h^{l-1})\), and using (121), we have that

\[
(0, 0) = \left(\langle F_1(u_h^n, \sigma_h^n), \tilde{u}_h \rangle, \langle F_2(u_h^n, \sigma_h^n), \tilde{\sigma}_h \rangle\right)
+ \left(\langle F'_1(u_h^n, \sigma_h^n)(u_h^n - u_h^{l-1}, \sigma_h^n - \sigma_h^{l-1}), \tilde{u}_h \rangle, \langle F'_2(u_h^n, \sigma_h^n)(u_h^n - u_h^{l-1}, \sigma_h^n - \sigma_h^{l-1}), \tilde{\sigma}_h \rangle\right) + \frac{1}{2} \left(\langle (u_h^n - u_h^{l-1}, \sigma_h^n - \sigma_h^{l-1})^t F''_1(u_h^{n+\varepsilon}, \sigma_h^{n+\varepsilon})(u_h^n - u_h^{l-1}, \sigma_h^n - \sigma_h^{l-1}), \tilde{u}_h \rangle, \langle (u_h^n - u_h^{l-1}, \sigma_h^n - \sigma_h^{l-1})^t F''_2(u_h^{n+\varepsilon}, \sigma_h^{n+\varepsilon})(u_h^n - u_h^{l-1}, \sigma_h^n - \sigma_h^{l-1}), \tilde{\sigma}_h \rangle\right),
\]

(123)

where \( u^{n+\varepsilon} = \varepsilon u_h^n + (1 - \varepsilon)u_h^{l-1}, \sigma^{n+\varepsilon} = \varepsilon \sigma_h^n + (1 - \varepsilon)\sigma_h^{l-1} \), and \( F'_i \) and \( F''_i \) denote the Jacobian and the Hessian of \( F_i \) \((i = 1, 2)\), respectively. Therefore, denoting by \( e^l_u = u_h^n - u_h^l \) and \( e^l_{\sigma} = \sigma_h^n - \sigma_h^l \), from (122)-(123), we deduce

\[
\left\langle \frac{\partial F_1}{\partial u}(u_h^{l-1}, \sigma_h^{l-1})(e^l_u) + \frac{\partial F_1}{\partial \sigma}(u_h^{l-1}, \sigma_h^{l-1})(e^l_{\sigma}), \tilde{u}_h \right\rangle = -\frac{1}{2} \left(\langle (e^l_u, e^l_{\sigma})^t F''_1(u_h^{n+\varepsilon}, \sigma_h^{n+\varepsilon})(e^l_u, e^l_{\sigma}), \tilde{u}_h \rangle, \right)
\]

(124)

\[
\left\langle \frac{\partial F_2}{\partial u}(u_h^{l-1}, \sigma_h^{l-1})(e^l_u) + \frac{\partial F_2}{\partial \sigma}(u_h^{l-1}, \sigma_h^{l-1})(e^l_{\sigma}), \tilde{\sigma}_h \right\rangle = -\frac{1}{2} \left(\langle (e^l_u, e^l_{\sigma})^t F''_2(u_h^{n+\varepsilon}, \sigma_h^{n+\varepsilon})(e^l_u, e^l_{\sigma}), \tilde{\sigma}_h \rangle, \right)
\]

(125)

Thus, from (124)-(125) and taking into account that \( F''_i \) are constant matrices, we arrive at

\[
\frac{1}{k}(e^l_u, \tilde{u}_h) + \langle \nabla e^l_u, \nabla \tilde{u}_h \rangle + (e^l_u \sigma_h^{l-1}, \nabla \tilde{u}_h) + (u_h^{l-1} e^l_{\sigma}, \nabla \tilde{u}_h) = -(e^l_u e^l_{\sigma}, \nabla \tilde{u}_h), \quad \forall \tilde{u}_h \in U_h,
\]

(126)

\[
\frac{1}{k}(e^l_{\sigma}, \tilde{\sigma}_h) + (B_h e^l_{\sigma}, \tilde{\sigma}_h) + 2(u_h^{l-1} e^l_u, \nabla \tilde{\sigma}_h) = -(|e^l_u|^2, \nabla \cdot \tilde{\sigma}_h), \quad \forall \tilde{\sigma}_h \in \Sigma_h.
\]

(127)

Testing by \( \tilde{u}_h = e^l_u \) and \( \tilde{\sigma}_h = e^l_{\sigma} \) in (126) and (127) respectively, taking into account that
\[ \int_{\Omega} e_u^l = 0 \] and using the Hölder and Young inequalities as well as the 3D interpolation inequality (12), we obtain

\[
\frac{1}{k} \|(e_u^l, e^l_{\sigma})\|_0^2 + \|(e_u^l, e^l_{\sigma})\|_1^2 \leq \frac{1}{2} \|(e_u^l, e^l_{\sigma})\|_1^2 + C \|(e_u^l, e^l_{\sigma})\|_0^2 \|(u_h^{l-1}, \sigma_h^{l-1})\|_1^2 + C \|(e_u^l, e^l_{\sigma})\|_1^4 \tag{128}
\]

In order to use an induction strategy, we can assume the hypothesis

\[ \|(e_u^l, e^l_{\sigma})\|_1^2 \leq \delta_0, \]

which implies that

\[ \|(u_h^{l-1}, \sigma_h^{l-1})\|_1 \leq \|(u_h^n, \sigma_h^n)\|_1 + \sqrt{\delta_0}, \tag{129} \]

where \( \delta_0 > 0 \) is a small enough constant. Therefore, from (128)-(129) we have

\[
\left( \frac{1}{k} - C(\|(u_h^n, \sigma_h^n)\|_1^2 + (\delta_0)^2) \right) \|(e_u^l, e^l_{\sigma})\|_0^2 + \frac{1}{2} \|(e_u^l, e^l_{\sigma})\|_1^2 \leq C \left( \|(e_u^{l-1}, e^l_{\sigma}^{l-1})\|_1^2 \right)^2. \tag{130}
\]

Thus, if \( \frac{1}{2k} > C(\|(u_h^n, \sigma_h^n)\|_1^2 \) and \( \frac{1}{2k} > C(\delta_0)^2 \) (which is possible owing to (120)\textsubscript{2} and (120)\textsubscript{3}), from (130) we obtain

\[ \|(e_u^l, e^l_{\sigma})\|_1^2 \leq C \left( \|(e_u^{l-1}, e^l_{\sigma}^{l-1})\|_1^2 \right)^2. \tag{131} \]

Therefore, choosing \( \delta_0 \) small enough such that \( \delta_0 C \leq 1 \), the inequality \( \|(e_u^l, e^l_{\sigma})\|_1^2 \leq \delta_0 \) holds. Indeed, if we assume \( \|(e_u^0, e^0_{\sigma})\|_1^2 \leq \delta_0 \), we obtain the following recurrence expression

\[ \|(e_u^l, e^l_{\sigma})\|_1^2 \leq \|(e_u^{l-1}, e^l_{\sigma}^{l-1})\|_1^2 \leq \cdots \leq \|(e_u^0, e^0_{\sigma})\|_1^2 \leq \delta_0. \tag{132} \]

Hence, from (131) the hypothesis of Lemma 4.5 are satisfied, and we conclude the convergence of \((u_h^l, \sigma_h^l)\) to \((u_h^n, \sigma_h^n)\) in the \(H^1(\Omega)\)-norm.

\[ \Box \]

\textbf{Remark 4.7.} If (54) is satisfied (recall that this estimate holds, at least, in 2D Domains), we can determine \( \delta_0 \) in terms of \( k \). Indeed, from (58), we have that

\[ \|(e_u^0, e^0_{\sigma})\|_1^2 = \|(u_h^n - u_h^{n-1}, \sigma_h^n - \sigma_h^{n-1})\|_1^2 \leq k(K_4 + K_5 k), \]

and thus, we consider \( \delta_0 := k(K_4 + K_5 k) \). Then, hypothesis (120) in Theorem 4.6 are only imposed on \( k \), and (120)\textsubscript{2} is reduced to \( k \leq K_0 \), where \( K_0 \) is a constant depending on data.
(Ω, u₀, σ₀), but independent of (k, h) and n.

**Remark 4.8.** Since restriction (120)₂ is equivalent to (36), analogously as in Remark 3.5, under the hypothesis of Theorem 3.21, we have that the conclusion of Theorem 4.6 remains true assuming k small enough, (120)₁ and (120)₃.

Now, observe that from (132), we have the following uniform estimate for \((uₗ, σₗⁿ)\) solution of (119):

\[
\|(uₗ, σₗⁿ)\|₁ ≤ \|(uₗⁿ, σₗⁿ)\|₁ + \sqrt{δ₀}, \quad ∀ l ≥ 0.
\]

(133)

Then, using the above estimate, we will prove the conditional unique solvability of (119).

**Theorem 4.9. (Conditional Unique Solvability)** Assume (120). Then there exists a unique \((uₗ, σₗ)\) solution of (119).

**Proof.** By linearity, it suffices to prove uniqueness of solution of (119). Suppose that there exist \((uₗ₁, σₗ₁), (uₗ₂, σₗ₂)\) ∈ \(U_h \times Σ_h\) two solutions of (119). Then, denoting \(uₗ = uₗ₁ - uₗ₂\) and \(σₗ = σₗ₁ - σₗ₂\), we arrive at

\[
\begin{align*}
\frac{1}{k}(uₗ, uₗ) + (\nabla uₗ, \nabla uₗ) + (uₗ⁻¹σₗ, \nabla uₗ) + (uₗ⁻¹σₗ⁻¹, \nabla uₗ) & = 0, \quad ∀ uₗ ∈ U_h, \quad (134) \\
\frac{1}{k}(σₗ, σₗ) + (Bₗσₗ, σₗ) - 2(uₗ⁻¹∇ uₗ, σₗ) - 2(uₗ⁻¹∇ uₗ, σₗ) & = 0, \quad ∀ σₗ ∈ Σ_h. \quad (135)
\end{align*}
\]

Taking \(uₗ = uₗ₁\) and \(σₗ = \frac{1}{2}σₗ₁\) in (134)-(135), taking into account that \(\int uₗ₁ = 0\) and using the Hölder and Young inequalities as well as the interpolation inequality (12), we obtain

\[
\frac{1}{2k}(uₗ₁, σₗ₁)₁² + \frac{1}{2}(uₗ₁, σₗ₁)₂² ≤ \frac{1}{4}(uₗ₁, σₗ₁)₁² + C(uₗ⁻¹, σₗ⁻¹)₂² \cdot (uₗ₁, σₗ₁)₁² \leq 0,
\]

which, using (133) (recall that (133) holds assuming (120)), implies that

\[
\left[\frac{1}{k} - C(\|(uₗ₁, σₗ₁)\|₁² + (δ₀)²)\right] \|(uₗ₁, σₗ₁)\|₁² + \frac{1}{2}(uₗ₁, σₗ₁)₂² ≤ 0.
\]

(136)

Therefore, assuming (120)₂–₃, from (136) we conclude that \(\|(uₗ₁, σₗ₁)\|₁ = 0\), and therefore, \(uₗ₁ = uₗ₂\) and \(σₗ₁ = σₗ₂\). Thus, there exists a unique \((uₗ, σₗ)\) solution of (119).
5 Numerical results

In this section, we consider the nonlinear scheme US with right hand sides $f(x,t)$, $g(x,t)$ and $h(x,t)$ in (34) and (35) respectively, where these right hand sides are chosen corresponding to the exact solutions $u = e^{-t}(\cos(2\pi x)\cos(2\pi y) + 2)$, $v = (1 + \sin(t))(\cos(2\pi x)\cos(2\pi y) + 2)$ and $\sigma = \nabla v = (1 + \sin(t))(-2\pi \sin(2\pi x)\cos(2\pi y), -2\pi \sin(2\pi y)\cos(2\pi x))$. In our computation, we take $\Omega = [0,1] \times [0,1]$, and we use a uniform partition with $m + 1$ nodes in each direction. We choose the spaces for $u$, $\sigma$ and $v$, generated by $\mathcal{P}_1, \mathcal{P}_1, \mathcal{P}_2$-continuous FE, respectively. The linear iterative method used to approach the nonlinear scheme US is the Newton Method, and in all the cases, the iterative method stops when the relative error in $L^2$-norm is less than $\varepsilon = 10^{-6}$.

In order to check numerically the error estimates obtained in our theoretical analysis, we choose $k = 10^{-5}$ and the numerical results with respect to time $T = 0.001$ are listed in Tables 1-3. We can see that when $h \to 0$, $\|u(t_n) - u^h_n\|_{L^2 H^1}$ is convergent in optimal rate $O(h)$, and $\|u^h_n - R^u_h u^h_n\|_{L^2 H^1}$, $\|u(t_n) - u^h_n\|_{L^\infty L^2}$, $\|u^h_n - R^u_h u^h_n\|_{L^\infty L^2}$, $\|v(t_n) - v^h_n\|_{L^\infty H^1}$ and $\|v^h_n - R^v_h v^h_n\|_{L^\infty H^1}$ are convergent in optimal rate $O(h^2)$.

| $m \times m$ | $\|u(t_n) - u^h_n\|_{L^\infty L^2}$ Order | $\|u^h_n - R^u_h u^h_n\|_{L^2 H^1}$ Order |
|---------------|--------------------------------|----------------------------------|
| 40 $\times$ 40 | $2.5 \times 10^{-3}$ - | $1.5 \times 10^{-4}$ - |
| 50 $\times$ 50 | $1.6 \times 10^{-4}$ 1.9970 | $9 \times 10^{-5}$ 1.9846 |
| 60 $\times$ 60 | $1.1 \times 10^{-4}$ 1.9980 | $7 \times 10^{-5}$ 1.9896 |
| 70 $\times$ 70 | $8 \times 10^{-5}$ 1.9985 | $5 \times 10^{-5}$ 1.9923 |
| 80 $\times$ 80 | $6 \times 10^{-5}$ 1.9989 | $4 \times 10^{-5}$ 1.9938 |

Table 1 – Error orders for $\|u(t_n) - u^h_n\|_{L^\infty L^2}$ and $\|u^h_n - R^u_h u^h_n\|_{L^2 H^1}$.

| $m \times m$ | $\|u(t_n) - u^h_n\|_{L^\infty H^1}$ Order | $\|u^h_n - R^u_h u^h_n\|_{L^2 H^1}$ Order |
|---------------|--------------------------------|----------------------------------|
| 40 $\times$ 40 | $1.1 \times 10^{-2}$ - | $5.219 \times 10^{-2}$ - |
| 50 $\times$ 50 | $8.9 \times 10^{-3}$ 0.9978 | $3.348 \times 10^{-3}$ 1.9896 |
| 60 $\times$ 60 | $7.4 \times 10^{-3}$ 0.9995 | $2.328 \times 10^{-3}$ 1.9937 |
| 70 $\times$ 70 | $6.3 \times 10^{-3}$ 0.9989 | $1.711 \times 10^{-3}$ 1.9966 |
| 80 $\times$ 80 | $5.5 \times 10^{-3}$ 0.9992 | $1.310 \times 10^{-3}$ 1.9988 |

Table 2 – Error orders for $\|u(t_n) - u^h_n\|_{L^\infty H^1}$ and $\|u^h_n - R^u_h u^h_n\|_{L^2 H^1}$.

References

[1] C. Amrouche and N.E.H. Seloula, $L^p$-theory for vector potentials and Sobolev’s inequalities for vector fields: application to the Stokes equations with pressure boundary conditions.
\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
$m \times m$ & $\|v(t_n) - v^h_n\|_{L^\infty H^1}$ & Order & $\|v^h_n - R^h v^h_n\|_{L^\infty H^1}$ & Order \\
\hline
40 \times 40 & $1.08 \times 10^{-2}$ & - & $9.875 \times 10^{-3}$ & - \\
50 \times 50 & $6.9 \times 10^{-3}$ & 1.9985 & $5.526 \times 10^{-4}$ & 2.6014 \\
60 \times 60 & $4.8 \times 10^{-3}$ & 1.9990 & $3.448 \times 10^{-4}$ & 2.5874 \\
70 \times 70 & $3.5 \times 10^{-3}$ & 1.9993 & $2.318 \times 10^{-4}$ & 2.5768 \\
80 \times 80 & $2.7 \times 10^{-3}$ & 1.9995 & $1.645 \times 10^{-4}$ & 2.5684 \\
\hline
\end{tabular}
\caption{Error orders for $\|v(t_n) - v^h_n\|_{L^\infty H^1}$ and $\|v^h_n - R^h v^h_n\|_{L^\infty H^1}$.}
\end{table}

Math. Models Methods Appl. Sci. 23 (2013), no. 1, 37–92.

[2] S.C. Brenner and L.R. Scott, The mathematical theory of finite element methods, Third edition, Texts in Applied Mathematics, 15. Springer, New York (2008).

[3] Y. Epshteyn and A. Izmirlioglu, Fully discrete analysis of a discontinuous finite element method for the Keller-Segel chemotaxis model. J. Sci. Comput. 40 (2009), no. 1-3, 211–256.

[4] F. Filbet, A finite volume scheme for the Patlak-Keller-Segel chemotaxis model. Numer. Math. 104 (2006), no. 4, 457–488.

[5] F. Guillén-González, M.A. Rodríguez-Bellido and D.A. Rueda-Gómez, Study of a chemorepulsion model with quadratic production. Part I: Analysis of the continuous problem and time-discrete numerical schemes. (Submitted).

[6] Y. He and K. Li, Asymptotic behavior and time discretization analysis for the non-stationary Navier-Stokes problem. Numer. Math. 98 (2004), no. 4, 647–673.

[7] A. Marrocco, Numerical simulation of chemotactic bacteria aggregation via mixed finite elements. M2AN Math. Model. Numer. Anal. 37 (2003), no. 4, 617–630.

[8] J. Necas, Les Méthodes Directes en Théorie des Equations Elliptiques. Editeurs Academia, Prague (1967).

[9] N. Saito, Conservative upwind finite-element method for a simplified Keller-Segel system modelling chemotaxis. IMA J. Numer. Anal. 27 (2007), no. 2, 332–365.

[10] N. Saito, Error analysis of a conservative finite-element approximation for the Keller-Segel system of chemotaxis. Commun. Pure Appl. Anal. 11 (2012), no. 1, 339–364.

[11] J. Shen, Long time stability and convergence for fully discrete nonlinear Galerkin methods. Appl. Anal. 38 (1990), 201–229.