Convergence analysis for a nonlinear system of parabolic variational inequalities

Yahya Alnashri

Abstract
This work aims to provide a comprehensive and unified numerical analysis for a nonlinear system of parabolic variational inequalities (PVIs) subject to Dirichlet boundary condition. This analysis enables us to establish the existence of an exact solution to the considered model and to prove the convergence for the approximate solution and its approximate gradient. Our results are applicable for several conforming and nonconforming numerical schemes.

MSC: 35J86; 65N12; 65N15; 76S05

Keywords: Nonlinear parabolic variational inequalities; PDEs; Gradient discretization method; Gradient schemes; Obstacle problem; Convergence analysis; Nonconforming P₁ finite element method

1 Introduction
Nonlinear parabolic variational inequalities and PDEs are useful tools to model the coupled biochemical interactions of microbial cells, which are crucial to numerous applications, especially in the medical field and food production [16, 19, 21]. We consider here a nonlinear parabolic system consisting of PDEs and variational inequalities:

\begin{align}
(\partial_t \bar{A} - \text{div}(D_A \nabla \bar{A}) - F(\bar{A}, \bar{B}))(\bar{A} - \chi) &= 0 \quad \text{in } \Omega \times (0, T), \\
\partial_t \bar{A} - \text{div}(D_A \nabla \bar{A}) &\leq F(\bar{A}, \bar{B}) \quad \text{in } \Omega \times (0, T), \\
\bar{A} &\leq \chi \quad \text{in } \Omega \times (0, T), \\
\partial_t \bar{B} - \text{div}(D_B \nabla \bar{B}) &= G(\bar{A}, \bar{B}) \quad \text{in } \Omega \times (0, T), \\
(\bar{A}, \bar{B}) &= (0, 0) \quad \text{on } (\partial \Omega \times (0, T))^2, \\
(\bar{A}(\mathbf{x}, 0), \bar{B}(\mathbf{x}, 0)) &= (A_{\text{ini}}, B_{\text{ini}}) \quad \text{in } (\Omega \times \{0\})^2.
\end{align}

Numerical approximation in parabolic systems of inequalities and generalization of inequalities have received considerable attention in the research literature. Wheeler [23] obtains the error estimate of second order in $L^\infty(L^2)$ for a linear approximation with respect to space and time, with a strong regularity on the solutions, such as $\partial_t \bar{B} \in L^2(0, T; L^2(\Omega))$. 

© The Author(s) 2022. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Johnson [17] analyzes inequality (1.1a), in which \( F = 0 \) and \( D_A \) is constant. The work in [4] considers a model without the barrier and provides \( O(h) \) order of convergence in \( L^\infty(L^2) \)-norm. An \( L^2 \)-error estimate is provided in different studies, such as [5], by using a finite difference in time. Vuik [22] deals with parabolic variational inequalities with a nonlinear source term and derives the convergence rate of the finite element method in space with respect to \( L^\infty \)-norm. He also shows that the general finite difference gives \( O(h) \) in \( L^\infty(L^2) \)-norm under a strong hypothesis on data. Saker et al. [20] discuss the discrete and continuous forms of a Carlson-type inequality, and [18] introduces Minkowski’s inequality by using \( AB \)-fractional integral operators. However, there is a lack of full convergence analyses of numerical schemes for the model (1.1a)–(1.1f) since the coupled nonlinearity of the system and the constraint (the inequality) in the model comprise the primary theoretical challenge. It appears that considerable research is still required, beginning with convergence analysis and testing other varieties of scheme outside conforming methods. Rather than undertaking individual research for every numerical scheme, this work utilizes a gradient discretization method (GDM) to provide a unified and full convergence analysis of numerical methods for (1.1a)–(1.1f) under natural hypotheses on data. The GDM is a generic framework to unify the numerical analysis for diffusion partial differential equations and their corresponding problems. Due to the variety of choice of the discrete elements in the GDM, a series of conforming and nonconforming numerical schemes can be included in the GDM, see [2, 3, 6, 9–13] for more details.

The outline of this paper is as follows. Section 2 is devoted to writing the model (1.1a)–(1.1f) in an equivalent weak sense. Section 3 defines the discrete space and functions followed by the gradient scheme to our model in the weak sense. Section 4 provides the convergence results, Theorem 4.5, which is proved by following the compactness technique under classical hypothesis on continuous model data. Finally, as an example, we present in Sect. 5 the nonconforming \( P_1 \) finite element scheme that has not been applied to the nonlinear model (1.1a)–(1.1f), yet.

2 Continuous setting

Hypothesis 2.1 We assume the following:

1. \( \Omega \subset \mathbb{R}^d (d \geq 1) \) is a bounded connected open set, and \( T > 0 \).
2. \( D_A, D_B : \Omega \to M_d(\mathbb{R}) \) are measurable functions (where \( M_d(\mathbb{R}) \) consists of \( d \times d \) matrices) and there exist \( d_1, d_2 > 0 \) such that for a.e. \( x \in \Omega \), \( D_A(x) \) and \( D_B(x) \) are symmetric with eigenvalues in \([d_1, d_2] \).
3. The constraint function \( \chi \) is in \( H^1(\Omega) \cap C(\overline{\Omega}) \) such that \( \chi \geq 0 \) on the domain boundary \( \partial \Omega \).
4. \( F \) and \( G \) are smooth and Lipschitz functions on \( \mathbb{R}^2 \) with Lipschitz constants \( M_1 \) and \( M_2 \), respectively, and \( M = \max(M_1, M_2) \).
5. \( A_{ini} \in W^{2,\infty}(\Omega) \cap K \), where \( K := \{ \varphi \in H^1_0(\Omega) : \varphi \leq \chi(t) \text{ in } \Omega \} \) and \( B_{ini} \in W^{2,\infty}(\Omega) \).

With the above Hypothesis 2.1, we consider the time-dependent closed convex set

\[ \mathcal{K} := \{ \varphi \in L^2(0, T; H^1_0(\Omega)) : \varphi(t) \in K \text{ for a.e. } t \in (0, T) \} . \]

It is clear that the time-dependent \( \mathcal{K} \) contains at least the constant in time function \( t \mapsto \chi^- := \min(0, \chi) \).
**Definition 2.2** (Weak formulation) Under Hypothesis 2.1, we say that \((\hat{A}, \hat{B})\) is a weak solution of (1.1a)–(1.1f) if the following properties and relations hold:

1. \(A \in \mathbb{K} \cap C^0([0, T]; L^2(\Omega)), \hat{A}(., 0) = A_{ini}, \partial_t \hat{A} \in L^2(0, T; L^2(\Omega)), \)
2. \(B \in C^0([0, T]; L^2(\Omega)), \hat{B}(., 0) = B_{ini}, \partial_t \hat{B} \in L^2(0, T; H^{-1}(\Omega)), \)
3. for all \(\phi \in \mathbb{K}, \) and for all \(\psi \in L^2(0, T; H_0^1(\Omega)), \)

\[
\int_0^T \int_\Omega \partial_t \hat{A}(x, t)(A(x, t) - \psi(x, t)) \, dx \, dt \\
+ \int_0^T \int_\Omega D_A \nabla \hat{A} \cdot \nabla (\hat{A} - \psi)(x, t) \, dx \, dt \\
\leq \int_0^T \int_\Omega F(\hat{A}, \hat{B})(\hat{A}(x, t) - \psi(x, t)) \, dx \, dt, \quad \text{and} \\
\int_0^T \langle \partial_t \hat{B}(x, t), \psi(x, t) \rangle_{H^{-1}, H^1} \, dt + \int_0^T \int_\Omega D_B(x) \nabla \hat{B}(x, t) \cdot \nabla \psi(x, t) \, dx \, dt \\
= \int_0^T \int_\Omega G(\hat{A}(x, t), \hat{B}(x, t)) \psi(x, t) \, dx \, dt,
\]

where \(\langle \cdot, \cdot \rangle_{H^{-1}, H^1} \) is the duality product between \(H^{-1}(\Omega) \) and \(H^1(\Omega). \)

**3 Discrete setting**

We begin with defining the discrete space and operators. These discrete elements are slightly different from those defined in [2, 3], in particular, \(\chi_D, I_D, \) and \(J_D \) are introduced to deal with the nonconstant barrier \(\chi \) and the initial solutions \(A_{ini} \) and \(B_{ini}. \)

**Definition 3.1** (GD for time-dependent obstacle problem) Let \(\Omega \) be an open domain of \(\mathbb{R}^d \) \((d \geq 1)\) and \(T > 0. \) A gradient discretization \(D \) is defined by \(D = (X_{D,0}, \Pi_D, \nabla_D, \chi_D, I_D, J_D, (t^{(m)})_{m=0, \ldots, N}), \) where:

1. The discrete set \(X_{D,0} \) is a finite-dimensional vector space over \(\mathbb{R}, \) taking into account the homogenous Dirichlet boundary condition \((1.1e). \)
2. The linear operator \(\Pi_D : X_{D,0} \rightarrow L^2(\Omega) \) is the reconstruction of the approximate function.
3. The linear operator \(\nabla_D : X_{D,0} \rightarrow L^2(\Omega)^d \) is the reconstruction of the gradient of the function, and must be chosen so that \(\|\nabla_D \cdot \|_{L^2(\Omega)^d} \) is a norm on \(X_{D,0}. \)
4. \(\chi_D \in L^2(\Omega) \) is an approximation of the barrier \(\chi. \)
5. \(I_D : W^{2,\infty}(\Omega) \cap K \rightarrow K_D := \{\phi \in X_{D,0} : \Pi_D \phi \leq \chi_D, \text{ in } \Omega\} \) is a linear and continuous interpolation operator for the initial solution \(A_{ini} \).
6. \(J_D : W^{2,\infty}(\Omega) \rightarrow X_{D,0} \) is a linear and continuous interpolation operator for the solution \(B_{ini}. \)
7. \(t^{(0)} = 0 < t^{(1)} < \cdots < t^{(N)} = T. \)

**Remark 3.2** For a general obstacle \(\chi, \) most of numerical methods fail to approximate the solution \(\hat{A} \) by elements inside the set \(K_D. \) For an example, in the \(p \) finite element method, we consider only the values of \(\hat{A} \) at the vertices of the mesh, which only guarantee that these values satisfy the barrier condition \((1.1c) \) only at these vertices, not necessarily at any point in \(\Omega. \) We define here the set \(K_D \) based on the approximate barrier \(\psi_D \) to be able to construct an interpolant that belongs to \(K_D. \) However, there is no need to use an approximate barrier, if the barrier \(\chi \) is assumed to be constant.
For any \( \varphi = (\varphi^{(n)})_{n=0,...,N} \in X^{N+1}_{D,0} \), we define space–time functions as follows: the reconstructed function \( \Pi_D \varphi : \Omega \times [0, T] \to \mathbb{R} \) and the reconstructed gradient \( \nabla_D \varphi : \Omega \times [0, T] \to \mathbb{R}^d \) are given by:

\[
\Pi_D \varphi(t), 0) = \Pi_D \varphi^{(0)}(0) \quad \text{and} \quad \forall n = 0, \ldots, N - 1, \forall t \in (t^{(n)}, t^{(n+1)}], \forall \mathbf{x} \in \Omega,
\]

\[
\Pi_D \varphi(t, t) = \Pi_D \varphi^{(n+1)}(\mathbf{x}) \quad \text{and} \quad \nabla_D \varphi(t, t) = \nabla_D \varphi^{(n+1)}(\mathbf{x}).
\]

Setting \( \delta t^{(n+\frac{1}{2})} = t^{(n+1)} - t^{(n)} \), for \( n = 0, \ldots, N - 1 \), and \( \delta t_D = \max_{n=0,\ldots,N-1} \delta t^{(n+\frac{1}{2})} \), the discrete derivative \( \delta_D \varphi \in L^\infty(0, T; L^2(\Omega)) \) of \( \varphi \in X^{N+1}_{D,1} \) is defined by

\[
\delta_D \varphi(t) = \delta_D^{(n+\frac{1}{2})} \varphi := \frac{\Pi_D \varphi^{(n+1)} - \Pi_D \varphi^{(n)}}{\delta t^{(n+\frac{1}{2})}}, \quad \forall n = 0, \ldots, N - 1 \quad \text{and} \quad t \in (t^{(n)}, t^{(n+1)}].
\]

In order to construct a good approximate scheme, we require four properties: coercivity, consistency, limit-conformity, and compactness. The first three respectively connect to the Poincaré inequality, the interpolation error, and the Stokes formula. The compactness property enables us to deal with the nonlinearity caused by the reaction terms \( F \) and \( G \).

**Definition 3.3** (Coercivity) If \( D \) is a gradient discretization, set

\[
C_D = \max_{\varphi \in X^D, \varphi \neq 0} \frac{\| \Pi_D \varphi \|_{L^2(\Omega)}}{\| \nabla_D \varphi \|_{L^2(\Omega)^d}}.
\]

A sequence \( (D_m)_{m \in \mathbb{N}} \) of gradient discretizations is coercive if \( (C_{D_m})_{m \in \mathbb{N}} \) remains bounded.

**Definition 3.4** (Consistency) If \( D \) is a gradient discretization, let \( S_D : \mathcal{K} \to [0, \infty) \) and \( \tilde{S}_D : H^1_0(\Omega) \to [0, \infty) \) be defined by

\[
\forall w \in \mathcal{K}, \quad S_D(w) = \min_{\varphi \in \mathcal{K}_D} \left( \| \Pi_D \varphi - w \|_{L^2(\Omega)} + \| \nabla_D \varphi - \nabla w \|_{L^2(\Omega)^d} \right), \quad (3.1)
\]

\[
\forall w \in H^1_0(\Omega), \quad \tilde{S}_D(w) = \min_{\varphi \in X^D_D} \left( \| \Pi_D \varphi - w \|_{L^2(\Omega)} + \| \nabla_D w - \nabla \varphi \|_{L^2(\Omega)^d} \right). \quad (3.2)
\]

A sequence \( (D_m)_{m \in \mathbb{N}} \) of gradient discretizations is consistent if, as \( m \to \infty \),

- for all \( w \in \mathcal{K} \), \( S_{D_m}(w) \to 0 \),
- for all \( w \in H^1_0(\Omega) \), \( \tilde{S}_{D_m}(w) \to 0 \),
- for all \( w \in W^{2,\infty}(\Omega) \cap \mathcal{K} \), \( \Pi_{D_m} I_{D_m} w \to w \) strongly in \( L^2(\Omega) \),
- for all \( w \in W^{2,\infty}(\Omega) \), \( \Pi_{D_m} I_{D_m} w \to w \) strongly in \( L^2(\Omega) \),
- \( (\| \nabla_{D_m} I_{D_m} \psi \|_{L^2(\Omega)^d})_{m \in \mathbb{N}} \) is bounded,
- \( \delta_{D_m} \to 0 \).

**Definition 3.5** (Limit-conformity) If \( D \) is a gradient discretization, let \( W_D : H_{\text{div}}(\Omega) := \{ \psi \in L^2(\Omega)^d : \text{div} \psi \in L^2(\Omega) \} \to [0, +\infty) \) be defined by

\[
W_D(\psi) = \sup_{\varphi \in X^D, \varphi \neq 0} \frac{| \int_\Omega (\nabla_D \varphi \cdot \psi + \Pi_D \varphi \text{div}(\psi)) \, d\mathbf{x} |}{\| \nabla_D \varphi \|_{L^2(\Omega)^d}}. \quad (3.3)
\]

A sequence \( (D_m)_{m \in \mathbb{N}} \) of gradient discretizations is limit-conforming if for all \( \psi \in H_{\text{div}}(\Omega) \), \( W_{D_m}(\psi) \to 0 \), as \( m \to \infty \).
**Definition 3.6** (Compactness) A sequence of gradient discretizations \((\mathcal{D}_m)_{m \in \mathbb{N}}\) is compact if, for any sequence \((\phi_m)_{m \in \mathbb{N}} \in X_{\mathcal{D}_m,0}\) such that \(\|\nabla_{\mathcal{D}_m} \phi_m\|_{L^2(\Omega)}\) is bounded, the sequence \((\Pi_{\mathcal{D}_m} \phi_m)_{m \in \mathbb{N}}\) is relatively compact in \(L^2(\Omega)\).

**Definition 3.7** (Gradient scheme problem) Find sequences \(\{\kappa_n\} = \{\kappa^0, ...\infty, N\}\) of piecewise-constant in time functions by \(T\) sequence \((\kappa_n)\), if, for any sequence \((\kappa_n)\),

\[
\int_\Omega \delta_D^{(n_1+\frac{1}{2})} A(\kappa) \Pi_D (A^{(n_1)}(\kappa) - \phi(\kappa)) \, d\kappa + \int_\Omega D \nabla_D A^{(n_1)}(\kappa) \cdot \nabla_D (A^{(n_1)}(\kappa) - \phi(\kappa)) \, d\kappa \]

\[
\leq \int_\Omega F (\Pi_D A^{(n_1)} + \Pi_D B^{(n_1)} + 1) \Pi_D (A^{(n_1)}(\kappa) - \phi(\kappa)) \, d\kappa \, dt, \quad \text{and}
\]

\[
\int_\Omega \delta_D^{(n_1+\frac{1}{2})} B(\kappa) \Pi_D \psi(\kappa) \, d\kappa + \int_\Omega D \nabla_D B^{(n_1)}(\kappa) \cdot \nabla_D \psi(\kappa) \, d\kappa = \int_\Omega G (\Pi_D A^{(n_1)} + \Pi_D B^{(n_1)} + 1) \Pi_D \psi(\kappa) \, d\kappa \, dt. \tag{3.4b}
\]

### 4 Main Results

Let the time interval \([0,T]\) be divided into \(\ell_k\) intervals of length \(\kappa\), where \(\kappa\) tends to zero as \(\ell_k \to \infty\). Let \(I_i\) be the characteristic function of \(I_i = [i\kappa, (i+1)\kappa], i = 0, ... \ell_k\). We define a set of piecewise-constant in time functions by

\[
\mathbb{I}_k = \left\{ w_k(\kappa,t) = \sum_{i=1}^{\ell_k} I_i(t) \phi_i(\kappa) : \phi_i \in C^0_0(\mathbb{R}) \text{ and } \phi_i \leq \chi \text{ in } \Omega \text{ a.e.} \right\}. \tag{4.1}
\]

**Lemma 4.1** For \(T > 0\), let \((\mathcal{D}_m)_{m \in \mathbb{N}}\) be a sequence of gradient discretizations that is consistent. Let \(\tilde{w}_k \in L^k\) be a piecewise constant in time function, where \(\mathbb{I}_k\) is the set defined by (4.1). Then there exists a sequence \((w_m)_{m \in \mathbb{N}}\) such that \(w_m = (w_m^{(n)})_{n=0,1,...N}\) for \(m \in \mathbb{N}\) and \(n \to \infty\),

\[
\Pi_{\mathcal{D}_m} w_m \to \tilde{w}_k \quad \text{strongly in } L^2(\Omega) \times (0,T), \tag{4.2a}
\]

\[
\nabla_{\mathcal{D}_m} w_m \to \nabla \tilde{w}_k \quad \text{strongly in } L^2(\Omega) \times (0,T)^d. \tag{4.2b}
\]

**Proof** Write \(\tilde{w}_k(\kappa,t) = \sum_{i=1}^{\ell_k} I_i(t) \phi_i(\kappa)\), where \(\phi_i \in C^0_0(\mathbb{R}) \cap \mathcal{K}\). Let \(s \in (0,T)\) and choose \(n := n(s)\) such that \(s \in (t_i^{(m)}, t_i^{(n+1)}]\). Let \(w_m \in X_{\mathcal{D}_m,0}\) be defined by \(w_m = \sum_{i=1}^{t_i^{(m+1)}} I_i(t^{(n+1)}) P_{\mathcal{D}_m} \phi_i\), where

\[
P_{\mathcal{D}_m}(\phi) = \arg\min_{\omega \in \mathcal{K}_{\mathcal{D}_m}} (\|\Pi_{\mathcal{D}_m} \omega - \phi\|_{L^2(\Omega)}^2 + \|\nabla_{\mathcal{D}_m} \omega - \nabla \phi\|_{L^2(\Omega)^d})^2. \tag{4.3}
\]
For \( i = 1, \ldots, \ell_e \), we define \( \xi^i_m : (0, T) \to \mathbb{R} \) by \( \xi^i_m(s) = 1_{I_i}(t^{(i+1)}) \) for \( s \in (0, T) \). Using the relation \( ab - cd = (a - c)b + c(b - d) \), we obtain, for all \( s \in (0, T) \) and a.e. \( x \in \Omega \),

\[
(\Pi_{D_m} w_m - \bar{w}_e)(x, s) = \sum_{i=1}^{\ell_e} (\xi^i_m(s) - 1_{I_i}(s)) \Pi_{D_m} P_{D_m} \phi_i(x) + \sum_{i=1}^{\ell_e} 1_{I_i}(s) (\Pi_{D_m} P_{D_m} \phi_i - \phi_i)(x).
\]

An application of the definition of \( S_{D_m} \) yields

\[
\| \Pi_{D_m} w_m - \bar{w}_e \|_{L^2(\Omega \times (0, T))} \leq \sum_{i=1}^{\ell_e} \| \xi^i_m(s) - 1_{I_i}(s) \|_{L^2(0, T)} \| \Pi_{D_m} P_{D_m} \phi_i \|_{L^2(\Omega)}
\]

\[
+ \sum_{i=1}^{\ell_e} \| 1_{I_i}(s) \|_{L^2(0, T)} \| \Pi_{D_m} P_{D_m} \phi_i - \phi_i \|_{L^2(\Omega)}
\]

\[
\leq \sum_{i=1}^{\ell_e} \| \xi^i_m(s) - 1_{I_i}(s) \|_{L^2(0, T)} \left( S_{D_m}(\phi_i) + \| \phi_i \|_{L^2(\Omega)} \right)
\]

\[
+ C_1 \sum_{i=1}^{\ell_e} S_{D_m}(\phi_i),
\]

where \( C_1 = \sum_{i=1}^{\ell_e} \| 1_{I_i} \|_{L^2(0, T)} \). Using consistency, one obtains \( S_{D_m}(\phi_i) \to 0 \) as \( m \to \infty \), for any \( i = 0, \ldots, \ell_e \), which implies that the second term on the right-hand side vanishes. In the case in which both \( s, t^{(i+1)} \in I_i \) or both \( s, t^{(i+1)} \not\in I_i \), the quantity \( \xi^i_m(s) - 1_{I_i}(s) \) equals zero. In the case in which \( s \in I_i \) and \( t^{(i+1)} \not\in I_i \) or \( s \not\in I_i \) and \( t^{(i+1)} \in I_i \), one can deduce, writing \( I_i = [a_i, b_i] \) and because \( s \) is chosen such that \( |s - t^{(i+1)}| \leq \delta t_{D_m} \),

\[
\| \xi^i_m(s) - 1_{I_i}(s) \|_{L^2(0, T)} \leq \text{measure}([a_i - \delta t_{D_m}, a_i + \delta t_{D_m}] \cup [b_i - \delta t_{D_m}, b_i + \delta t_{D_m}])
\]

\[
\leq 4\delta t_{D_m}.
\]

This shows that the first term on the right-hand side of \( (4.4) \) tends to zero when \( m \to \infty \). Hence, \( (4.2a) \) is concluded. The proof of \( (4.2b) \) is obtained by the same reasoning, replacing \( \bar{w}_e \) by \( \nabla \bar{w}_e \) and \( \Pi_{D_m} w_m \) by \( \nabla D_m w_m \).

\[\square\]

**Lemma 4.2** (Energy estimates) Let Hypothesis 2.1 hold. If \( D \) is a gradient discretization such that \( \delta_D < \frac{1}{2M} \), \( K_D \) is a nonempty set, and \( (A, B) \in K_D \times X_{D,0} \) is a solution of the approximate scheme \( (3.4a) - (3.4b) \), then there exists a constant \( C_2 \geq 0 \) only depending on \( \Omega, \alpha, T, M, C_0 := \max(F(0), G(0)) \), \( \| \Pi_D I_D A_{m1} \|_{L^2(\Omega)}, \| \nabla_D I_D A_{m1} \|_{L^2(\Omega)^d}, \) and \( \| \Pi_D I_D B_{m1} \|_{L^2(\Omega)} \), such that

\[
\| \delta_D A \|_{L^2(\Omega \times (0, T))} + \| \nabla_D A \|_{L^\infty(0, T; L^2(\Omega)^d)}
\]

\[
+ \| \Pi_D B \|_{L^\infty(0, T; L^2(\Omega))} + \| \nabla_D B \|_{L^2(\Omega \times (0, T))} \leq C_2.
\]
Proof. We start by taking \( \varphi : = A^{(n)} \) (it belongs to \( \mathcal{K}_D \)) and the function \( \psi : = \delta t^{(n+\frac{1}{2})}B^{(n+1)} \) in (3.4a)-(3.4b) to get

\[
\delta t^{(n+\frac{1}{2})} \int_{\Omega} \left| \delta_D^{(n+\frac{1}{2})} A \right|^2 \ dx + \int_{\Omega} D_A \nabla_D A^{(n+1)} \cdot \nabla_D (A^{(n+1)} - A^{(n)}) \ dx \\
\leq \delta t^{(n+\frac{1}{2})} \int_{\Omega} F(\Pi_D A^{(n+1)}, \Pi_D B^{(n+1)}) \delta_D^{(n+\frac{1}{2})} A \ dx,
\]

(4.6)

and

\[
\int_{\Omega} \left( \Pi_D B^{(n+1)}(x) - \Pi_D B^{(n)}(x) \right) \Pi_D B^{(n+1)}(x) \ dx \\
+ \int_{\Omega} (\Pi_D B^{(n+1)} + D_B \left| \nabla_D B^{(n+1)} \right|^2 \ dx + \int_{\Omega} D_B \left| \nabla_D B^{(n+1)} \right|^2 \ dx \ dx + \int_{\Omega} \Pi_D B^{(n+1)}(x) \ dx \\
= \delta t^{(n+\frac{1}{2})} \int_{\Omega} G(\Pi_D A^{(n+1)}, \Pi_D B^{(n+1)}) \delta_D^{(n+\frac{1}{2})} A \ dx.
\]

(4.7)

Applying the fact that \( (r - s) \cdot \mathbf{r} \geq \frac{1}{2} |r|^2 - \frac{1}{2} |s|^2 \) to the second term on the left-hand side of (4.6) and to the first term on the left-hand side of (4.7), it follows that

\[
\delta t^{(n+\frac{1}{2})} \int_{\Omega} \left( \left| \delta_D^{(n+\frac{1}{2})} A \right|^2 - \left| \nabla_D A^{(n)} \right|^2 \ dx + \frac{d_1}{2} \int_{\Omega} \left( \left| \nabla_D A^{(n)} \right|^2 - \left| \nabla_D A^{(n+1)} \right|^2 \ dx \right) \ dx \\
\leq \delta t^{(n+\frac{1}{2})} \int_{\Omega} F(\Pi_D A^{(n+1)}, \Pi_D B^{(n+1)}) \delta_D^{(n+\frac{1}{2})} A \ dx,
\]

and

\[
\frac{1}{2} \int_{\Omega} \left( \left| \Pi_D B^{(n+1)}(x) \right|^2 - \left| \Pi_D B^{(n)}(x) \right|^2 \ dx + \frac{d_1}{2} \int_{\Omega} \left( \left| \nabla_D B^{(n+1)} \right|^2 \ dx + \int_{\Omega} \left| \nabla_D B^{(n+1)} \right|^2 \ dx \right) \ dx + \int_{\Omega} \Pi_D B^{(n+1)}(x) \ dx \\
\leq \delta t^{(n+\frac{1}{2})} \int_{\Omega} G(\Pi_D A^{(n+1)}, \Pi_D B^{(n+1)}) \delta_D^{(n+\frac{1}{2})} A \ dx.
\]

Summing the above inequalities over \( n \in [0, m - 1] \), where \( m = 0, \ldots, N \) gives

\[
\| \delta_D A \|^2_{L^2(\Omega \times (0,t^{(m)})]} + \frac{d_1}{2} \left( \| \nabla_D A^{(n)} \|^2_{L^2(\Omega \times (0,t^{(m)})]} - \| \nabla_D A^{(n+1)} \|^2_{L^2(\Omega \times (0,t^{(m)})]} \right) \\
\leq \sum_{n=0}^{m-1} \delta t^{(n+\frac{1}{2})} \int_{\Omega} F(\Pi_D A^{(n+1)}, \Pi_D B^{(n+1)}) \delta_D^{(n+\frac{1}{2})} A \ dx,
\]

(4.8)

and

\[
\frac{1}{2} \left( \| \Pi_D B^{(n)} \|^2_{L^2(\Omega \times (0,t^{(m)})]} - \| \Pi_D B^{(n)} \|^2_{L^2(\Omega \times (0,t^{(m)})]} \right) + \frac{d_1}{2} \sum_{n=0}^{m-1} \delta t^{(n+\frac{1}{2})} \| \nabla_D B^{(n)} \|^2_{L^2(\Omega \times (0,t^{(m)})]} \\
\leq \sum_{n=0}^{m-1} \delta t^{(n+\frac{1}{2})} \int_{\Omega} G(\Pi_D A^{(n+1)}, \Pi_D B^{(n+1)}) \delta_D^{(n+\frac{1}{2})} A \ dx.
\]

(4.9)
This, together with the Cauchy–Schwarz inequality, implies that
\[
\| \delta^D A \|_{L^2(\Omega \times (0,t) ; dx)}^2 + \frac{d_1}{2} \left( \| \nabla^D A^{(0)} \|_{L^2(\Omega ; dx)}^2 - \| \nabla^D A^{(0)} \|_{L^2(\Omega ; dt)}^2 \right) \\
\leq \sum_{n=0}^m \delta t^{(n+\frac{1}{2})} \left\| F(\Pi^D A^{(n+1)}, \Pi^D B^{(n+1)}) \right\|_{L^2(\Omega \times (0,T))} \| \delta^D (n+\frac{1}{2}) A \|_{L^2(\Omega \times (0,t))}
\]
and
\[
\frac{1}{2} \left( \| \Pi^D B^{(m)} \|_{L^2(\Omega)}^2 - \| \Pi^D B^{(0)} \|_{L^2(\Omega)}^2 \right) + d_2 \sum_{n=0}^{m-1} \delta t^{(n+\frac{1}{2})} \| \nabla^D B^{(n)} \|_{L^2(\Omega ; dt)} \\
\leq \sum_{n=0}^m \delta t^{(n+\frac{1}{2})} \left\| G(\Pi^D A^{(n+1)}, \Pi^D B^{(n+1)}) \right\|_{L^2(\Omega \times (0,T))} \| \Pi^D B^{(n+1)} \|_{L^2(\Omega \times (0,T))}.
\]
Using the Lipschitz condition, we arrive at
\[
\| \delta^D A \|_{L^2(\Omega \times (0,t) ; dx)}^2 + \frac{d_1}{2} \left( \| \nabla^D A^{(0)} \|_{L^2(\Omega ; dx)}^2 - \| \nabla^D A^{(0)} \|_{L^2(\Omega ; dt)}^2 \right) \\
\leq \sum_{n=0}^m \delta t^{(n+\frac{1}{2})} \| \delta^D (n+\frac{1}{2}) A \|_{L^2(\Omega \times (0,t))} \\
\times (M \| \Pi^D A^{(n+1)} \|_{L^2(\Omega)} + M \| \Pi^D B^{(n+1)} \|_{L^2(\Omega)} + C_0),
\]
and
\[
\frac{1}{2} \left( \| \Pi^D B^{(m)} \|_{L^2(\Omega)}^2 - \| \Pi^D B^{(0)} \|_{L^2(\Omega)}^2 \right) + d_2 \sum_{n=0}^{m-1} \delta t^{(n+\frac{1}{2})} \| \nabla^D B^{(n)} \|_{L^2(\Omega ; dt)} \\
\leq \sum_{n=0}^m \delta t^{(n+\frac{1}{2})} \left( M \| \Pi^D B^{(n+1)} \|_{L^2(\Omega)}^2 + M \| \Pi^D B^{(n+1)} \|_{L^2(\Omega)}^2 + C_0 \| \Pi^D B^{(n+1)} \|_{L^2(\Omega)}^2 \right).
\]
This, together with Young’s inequality, gives, whenever \( 1 - \sum_{i=1}^3 \epsilon_i > 0, \)
\[
\| \delta^D A \|_{L^2(\Omega \times (0,t) ; dx)}^2 + \frac{d_1}{2} \left( \| \nabla^D A^{(0)} \|_{L^2(\Omega ; dx)}^2 - \| \nabla^D A^{(0)} \|_{L^2(\Omega ; dt)}^2 \right) \\
\leq \sum_{i=1}^3 \frac{\epsilon_i}{2} \| \delta^D (n+\frac{1}{2}) A \|_{L^2(\Omega \times (0,t))}^2 \\
+ \sum_{n=0}^m \frac{\epsilon_i^{(n+\frac{1}{2})}}{2} \left( \frac{M}{2 \epsilon_{i1}} \| \Pi^D A^{(n+1)} \|_{L^2(\Omega)}^2 + \frac{M}{2 \epsilon_{i2}} \| \Pi^D B^{(n+1)} \|_{L^2(\Omega)}^2 + \frac{1}{2 \epsilon_{i3}} C_0 \right), (4.10)
\]
and
\[
\frac{1}{2} \left( \| \Pi^D B^{(m)} \|_{L^2(\Omega ; dt)}^2 - \| \Pi^D B^{(0)} \|_{L^2(\Omega ; dt)}^2 \right) + d_2 \sum_{n=0}^{m-1} \delta t^{(n+\frac{1}{2})} \| \nabla^D B^{(n)} \|_{L^2(\Omega ; dt)}^2
\]
\[
\leq \sum_{n=0}^{m} \delta t^{n+\frac{1}{2}} \left( M \left( 1 + \frac{1}{2\varepsilon_4} \right) \| \Pi_D B^{(n+1)} \|_{L^2(\Omega)}^2 \\
+ \frac{\varepsilon_4}{2} \| \Pi_D A^{(n+1)} \|_{L^2(\Omega)}^2 + \frac{\varepsilon_5}{2} c_0^2 \right). 
\]

(4.11)

Thanks to Gronwall inequality [15, Lemma 5.1], inequality (4.11) can be rewritten as

\[
\frac{1}{2} \| \Pi_D B^{(m)} \|_{L^2(\Omega)^d}^2 + d_2 \sum_{n=0}^{m-1} \delta t^{n+\frac{1}{2}} \| \nabla_D B^{(n)} \|_{L^2(\Omega)^d}^2 \\
\leq e^{C_3} \left( \frac{T \varepsilon_5}{2} c_0^2 + \frac{\varepsilon_4}{2} \sum_{n=0}^{m-1} \delta t^{n+\frac{1}{2}} \| \Pi_D A^{(n+1)} \|_{L^2(\Omega \times (0, T))}^2 + \frac{1}{2} \| \Pi_D B^{(0)} \|_{L^2(\Omega)}^2 \right),
\]

where \( C_3 \) depends on \( T, M, \) and \( \varepsilon_4 \). Combining this inequality with (4.10) yields

\[
\left( 2 - \sum_{i=1}^{3} \frac{\varepsilon_i}{2} \right) \| \nabla_D A \|^2_{L^2(\Omega \times (0, T))} + \left( \frac{d_1}{2} + \frac{M}{2\varepsilon_1} - \frac{\varepsilon_4}{2} e^{C_3} \right) \| \nabla_D A^{(m_0)} \|^2_{L^2(\Omega)^d} + \| \Pi_D B^{(m_0)}(x) \|^2_{L^2(\Omega)} + \| \Pi_D B^{(m_0)}(x, t) \|^2_{L^2(\Omega \times (0, T))} \\
\leq \left( \frac{T}{2\varepsilon_3} + \frac{T \varepsilon_5}{2} e^{C_3} \right) c_0^2 + \frac{d_1}{2} \| \nabla_D A \|^2_{L^2(\Omega)^d} + e^{C_3} \| \Pi_D B^{(0)} \|^2_{L^2(\Omega)}.
\]

Taking the supremum over \( m \in [0, N] \) and using the real inequality \( \sup_n (r_n + s_n) \leq \sup_n (r_n) + \sup_n (s_n) \), we obtain the desired estimates.

\[ \square \]

In the following definition, we introduce a dual norm [8], which is defined on the space \( \Pi_D(X_{D,0}) \subset L^2(\Omega) \), to ensure the required compactness results.

**Definition 4.3** If \( D \) is a gradient discretization, then the dual norm \( \| \cdot \|_{\star, D} \) on \( \Pi_D(X_{D,0}) \) is given by

\[
\forall U \in \Pi_D(X_{D,0}),
\| U \|_{\star, D} = \sup \left\{ \int_{\Omega} U(x) \Pi_D \psi(x) \, dx : \psi \in X_{D,0}, \| \nabla_D \psi \|_{L^2(\Omega)^d} = 1 \right\}. 
\]

(4.12)

**Lemma 4.4** Under Hypothesis 2.1, let \( D \) be a gradient discretization, which is coercive. If \( B \in X_{D,0} \) satisfies (3.4b), then there exists a constant \( C_4 \) depending only on \( C_1, M, \Omega, T, \) and \( \| \Pi_D B^{(0)} \|_{L^2(\Omega)} \), such that

\[
\int_0^T \| \delta_D B(t) \|_{\star, D}^2 \, dt \leq C_4.
\]

(4.13)
Proof Putting \( \psi = \phi \) in (3.4b), together with the Cauchy–Schwarz inequality and the coercivity property, implies

\[
\int_{\Omega} \delta_{n}^{(m+1)} B(x) \Pi D \phi(x) \, dx \\
\leq d_{2} \| \nabla D B^{(m+1)} \|_{L^{2}(\Omega \times (0,T))^{d}} \| \nabla D \phi \|_{L^{2}(\Omega \times (0,T))^{d}} \\
+ (M \| \Pi D B^{(m+1)} \|_{L^{2}(\Omega \times (0,T))} + M \| \Pi D A^{(m+1)} \|_{L^{2}(\Omega \times (0,T))} + C_{0}) \| \Pi D \phi \|_{L^{2}(\Omega)} \\
\leq \| \nabla D \phi \|_{L^{2}(\Omega \times (0,T))^{d}} \left[ d_{2} \| \nabla D B^{(m+1)} \|_{L^{2}(\Omega \times (0,T))^{d}} \\
+ C_{D}(2 \Pi D B^{(m+1)} \|_{L^{2}(\Omega \times (0,T))} + M \| \Pi D A^{(m+1)} \|_{L^{2}(\Omega \times (0,T))} + C_{0}) \right].
\]

Taking the supremum over \( \phi \in \mathcal{X}_{D,0} \) with \( \| \nabla D \phi \|_{L^{2}(\Omega)^{d}} = 1 \), multiplying by \( \delta d^{(m)} \), summing over \( n \in [0, N - 1] \), and using (4.5) yield the desired estimate. \( \square \)

Theorem 4.5 Under Hypothesis (2.1), let \( (\mathcal{D}_{m})_{m \in \mathbb{N}} \) be a sequence of gradient discretizations, that is coercive, limit-conforming, consistent, compact, and such that \( \mathcal{K}_{D_{m}} \) is a nonempty set for any \( m \in \mathbb{N} \). For \( m \in \mathbb{N} \), let \( (A_{m}, B_{m}) \in \mathcal{K}_{D_{m}}^{N_{m+1}} \times X_{D_{m},0}^{N_{m+1}} \) be solutions to the scheme (3.4a)–(3.4b) with \( D = \mathcal{D}_{m} \). Then there exists a solution \((\bar{A}, \bar{B})\) for the discrete problem (2.1a)–(2.1b), and a subsequent of gradient discretizations, indexed by \( (\mathcal{D}_{m})_{m \in \mathbb{N}} \), such that, as \( m \to \infty \),

1. \( \Pi D_{m} A_{m} \to A \) and \( \Pi D_{m} B_{m} \to B \) strongly in \( L^{\infty}(0,T;L^{2}(\Omega)) \),
2. \( \nabla D_{m} A_{m} \to \nabla A \) and \( \nabla D_{m} B_{m} \to \nabla B \) strongly in \( L^{2}(\Omega \times (0,T))^{d} \),
3. \( \delta D_{m} A_{m} \) converges weakly to \( \partial_{A} \) in \( L^{2}(\Omega \times (0,T)) \).

Proof The proof is divided into four stages and its idea is inspired by [1].

Step 1: Existence of approximate solutions. At \((n+1)\), we see that (3.4a) and (3.4b) respectively express a system of nonlinear elliptic variational inequality on \( A^{(m+1)} \) and nonlinear equations on \( B^{(m+1)} \). For \( w = (w_{1}, w_{2}) \in \mathcal{K}_{D} \times X_{D,0} \), we see that \((A, B) \in \mathcal{K}_{D} \times X_{D,0} \) satisfies

\[
\begin{align*}
& a^{(m+1)}(A^{(m+1)} - \psi) \leq L(A^{(m+1)} - \psi), \quad \forall \psi \in \mathcal{K}_{D}, \quad \text{and} \\
& \begin{align*}
& \int_{\Omega} \Pi D B^{(m+1)}(x) \Pi D \psi(x) + \int_{\Omega} \Pi D B^{(m+1)}(x) \cdot \nabla D \psi(x) \\
& = \int_{\Omega} G(\Pi D w_{1}, \Pi D w_{2}) \Pi D \psi(x) \\
& \quad \forall \psi \in X_{D,0}. 
\end{align*}
\end{align*}
\]

where \( \alpha := \frac{1}{\delta d^{(m+1)}}, \) and the bilinear and linear forms are defined by

\[
\begin{align*}
& a^{(m+1)}(\phi, z) = \alpha \int_{\Omega} \Pi D \phi \Pi D z \, dx + D_{A} \int_{\Omega} \nabla D \phi \cdot \nabla D z \, dx, \quad \forall \phi, z \in \mathcal{K}_{D} \quad \text{and} \\
& L(z) = \int_{\Omega} F(\Pi D w_{1}, \Pi D w_{2}) \Pi D z \, dx + \alpha \int_{\Omega} \Pi D A^{(m)} \Pi D z \, dx, \quad \forall z \in \mathcal{K}_{D}. 
\end{align*}
\]

Stampacchia’s theorem implies that there exists \( \bar{A} \in \mathcal{K}_{D} \) satisfying inequality (4.14a). The second equation (4.14b) describes a linear square system. Taking \( \psi = B^{(m+1)} \) in
(4.14b), using the similar reasoning as in the proof of Lemma 4.2, and setting \( G = 0 \) yield \( \| \nabla B^{(m+1)} \|_{L^2(\Omega)^d} = 0 \). This shows that the matrix corresponding to the linear system is invertible. Consider the continuous mapping \( T : K_D \times X_{\mathcal{D},0} \rightarrow K_D \times X_{\mathcal{D},0} \), where \( T(w) = (A, B) \) with \((A, B)\) being the solution to (4.14a)–(4.14b). The existence of a solution \((A^{(m+1)}, B^{(m+1)})\) to the nonlinear system is a consequence of Brouwer’s fixed point theorem.

**Step 2: Strong convergence of** \( \Pi_D A_m \) **and** \( \Pi_D B_m \) **in** \( L^\infty(0, T; L^2(\Omega)) \) **and the weak convergence of** \( \delta_D A_m \) **in** \( L^2(\Omega \times (0, T)) \). Applying estimate (4.5) to the sequence of solutions \((A_m)_{m \in \mathbb{N}}, (B_m)_{m \in \mathbb{N}}\) of the scheme (3.4a)–(3.4b) shows that both \( \| \nabla_D A_m \|_{L^2(\Omega \times (0, T))^d} \) and \( \| \nabla_D B_m \|_{L^2(\Omega \times (0, T))^d} \) are bounded. Using [8, Lemma 4.8], there exists a sequence, still denoted by \((D_m^i)_{m \in \mathbb{N}}, \hat{A}, \hat{B} \in L^2(0, T; H^2_0(\Omega))\) such that, as \( m \to \infty \), \( \Pi_D A_m \) converges weakly to \( \hat{A} \) in \( L^2(\Omega \times (0, T)) \), \( \nabla_D A_m \) converges weakly to \( \nabla \hat{A} \) in \( L^2(\Omega \times (0, T))^d \), \( \Pi_D B_m \) converges weakly to \( \tilde{B} \) in \( L^2(\Omega \times (0, T)) \), and \( \nabla_D B_m \) converges weakly to \( \nabla \tilde{B} \) in \( L^2(\Omega \times (0, T))^d \). Since \( A_m \in K_D \), passing to the limit in \( \Pi_D A_m \leq \chi \) in \( \Omega \) shows that \( \hat{A} \leq \chi \) in \( \Omega \). Thanks to [8, Theorem 4.31], estimate (4.5) shows that \( \hat{A} \in C([0, T]; L^2(\Omega)) \), \( \Pi_D A_m \) converges strongly to \( \hat{A} \) in \( L^\infty(0, T; L^2(\Omega)) \), and \( \delta_D A_m \) converges weakly to \( \partial_t \hat{A} \) in \( L^2(0, T; L^2(\Omega)) \).

Let us show the strong convergence of \( \Pi_D B_m \) to \( \tilde{B} \) in \( L^\infty(0, T; L^2(\Omega)) \). Indeed, \( \Pi_D B_m \) converges strongly to \( \tilde{B} \) in \( L^2(\Omega \times (0, T)) \), thanks to estimate (4.13), consistency, limit-conformity, and compactness, as well as [8, Theorem 4.14]. We can apply the dominated convergence theorem to show that \( G(\Pi_D A_m, \Pi_D B_m) \to G(\hat{A}, \hat{B}) \) in \( L^2(\Omega \times (0, T)) \), thanks to the assumptions on \( G \) given in Hypothesis 2.1.

Let \( t_0 \in [0, T] \) and define the sequence \( t_m \in [0, T] \) such that \( t_m \to t_0 \), as \( m \to \infty \). Consider \( k(m) \in [0, N_m - 1] \) such that \( k_m \in (t^{(m)}, t^{(m+1)}) \). Following the technique used in Lemma 4.2, one can obtain

\[
\frac{1}{2} \int_{\Omega} (\Pi_D B(x, t_m))^2 \, dx \\
\leq \frac{1}{2} \int_{\Omega} (\Pi_D f_D B_m(x))^2 \, dx - \int_{\Omega} D_B(\nabla_D B(x, t)) \, dx \, dt \\
\leq \int_{\Omega} G(\Pi_D A(x, t), \Pi_D B(x, t)) \Pi_D B(x, t) \, dx \, dt. \tag{4.15}
\]

Using the characteristic function, it is obvious that, as \( m \to \infty \),

\[
\Pi_D B_m \to \tilde{B} \quad \text{strongly in } L^2(\Omega \times (0, T)) \quad \text{and} \\
1_{[0, t^{(m)}]} \nabla \tilde{B} \to 1_{[0, t_0]} \nabla \tilde{B} \quad \text{strongly in } L^2(\Omega \times (0, T))^d.
\]

These convergence results imply that

\[
\int_0^{t_0} \int_{\Omega} D_B(\nabla B(x, t)) \, dx \, dt \\
= \int_0^{t_{(m)}} \int_{\Omega} 1_{[0, t_0]} D_B(\nabla B(x, t)) \, dx \, dt \\
= \lim_{m \to \infty} \int_0^{t_{(m+1)}} \int_{\Omega} 1_{[0, t_0]} D_B \nabla \tilde{B}(x, t) \cdot \nabla_D B_m(x, t) \, dx \, dt.
\]
Letting \( \bar{\psi} = \bar{B}1_{[0,t_0)}(t) \) in (3.4b) and integrating by parts, we obtain

\[
\frac{1}{2} \int_{\Omega} \left( \tilde{B}(\mathbf{x}, t_0) \right)^2 d\mathbf{x} + \int_{t_0}^{t_0} \int_{\Omega} \mathbf{D}_B(\nabla \tilde{B}(\mathbf{x}, t))^2 d\mathbf{x} dt
\]

(4.18)

From (4.17) and (4.18), we obtain

\[
\limsup_{m \to \infty} \int_{\Omega} \left( \Pi_{D_m} B_m(\mathbf{x}, t_m) \right)^2 d\mathbf{x} \leq \int_{\Omega} \tilde{B}(\mathbf{x}, t_0)^2 d\mathbf{x}. 
\]

(4.19)

Estimates (4.5) and (4.13), together with [8, Theorem 4.19], imply the weak convergence of \( (\Pi_{D_m} B_m)_{m \in \mathbb{N}} \) to \( \tilde{B} \) in \( L^2(\Omega) \); it is indeed uniform in \( [0, T] \). This yields the weak convergence of \( \Pi_{D_m} B_m(\cdot, s_m) \) to \( \tilde{B}(\cdot, t_0) \) in \( L^2(\Omega) \). As a consequence of estimate (4.19), this convergence of \( \Pi_{D_m} B_m(\cdot, s_m) \) holds in the strong sense in \( L^2(\Omega) \). With the continuity of \( \tilde{B} : [0, T] \to L^2(\Omega) \), we can apply [8, Lemma C.13] to conclude the strong convergence of \( \Pi_{D_m} B_m \) in \( L^\infty(0, T; L^2(\Omega)) \).

**Step 3: Convergence towards the solution of the continuous model.** Recall that \( A_{(0)}^{(0)} = I_{D_m} A_{ini} \), therefore the consistency shows that \( \Pi_{D_m} A_{(0)}^{(0)} \) converges strongly to \( A_{ini} \) in \( L^2(\Omega) \), as \( m \to \infty \). Hence, \( \tilde{A} \in C([0,T]; L^2(\Omega)) \cap \mathbb{K} \) and \( \tilde{A} \) satisfies all the conditions except for the integral inequality imposed on the exact solution of problem (2.1a). Let us now show that this integral relation holds. With Hypothesis 2.1, the dominated convergence theorem leads to \( F(\Pi_{D_m} A_m, \Pi_{D_m} B_m) \to F(\tilde{A}, \tilde{B}) \) in \( L^2(\Omega \times (0, T)) \). The \( L^2 \)-weak convergence of \( \nabla_{D_m} A_m \) yields

\[
\int_0^T \int_\Omega \mathbf{D}_A \nabla \tilde{A} \cdot \nabla \tilde{A} d\mathbf{x} dt \leq \liminf_{m \to \infty} \int_0^T \int_\Omega \mathbf{D}_A \nabla_{D_m} A_m \cdot \nabla_{D_m} A_m d\mathbf{x} dt. 
\]

(4.20)

Fix \( \kappa > 0 \) and let \( \tilde{w}_k \in L_\kappa \), where \( L_\kappa \) is defined by (4.1). Thanks to Lemma 4.1, there exists a sequence \( (w_m)_{m \in \mathbb{N}} \) such that \( w_m \in L_{D_m}^{\kappa N_m+1} \) and \( \Pi_{D_m} w_m \to \tilde{w}_k \) strongly in \( L^2(\Omega \times (0, T)) \).
and $\nabla D_n w_m \rightarrow \nabla w_\kappa$ strongly in $L^2(\Omega \times (0, T))^d$. Setting $\psi := w_m$ as a generic function in (3.4a), inequality (4.20) implies that

$$
\int_0^T \int_\Omega D_A \nabla \tilde{\phi} \cdot \nabla \tilde{\phi} \, dx \, dt \\
\leq \liminf_{m \to \infty} \left[ \int_0^T \int_\Omega F(\Pi_{D_m} A_m, \Pi_{D_m} B_m) \Pi_{D_m} (A_m - w_m) \, dx \, dt \\
+ \int_0^T \int_\Omega D_A \nabla D_m A_m \cdot \nabla D_m w_m \, dx \, dt - \int_0^T \int_\Omega \delta_{D_m} A_m \Pi_{D_m} (A_m - w_m) \, dx \, dt \right].
$$

With the weak–strong convergences, we pass to the limit in this relation to obtain, for all $w_\kappa \in L_\kappa$ and all $\kappa > 0$,

$$
\int_0^T \int_\Omega D_A \nabla \tilde{\phi} \cdot \nabla \tilde{\phi} \, dx \, dt + \int_0^T \int_\Omega D_A \nabla \tilde{\phi} \cdot \nabla (\tilde{\phi} - w_\kappa) \, dx \, dt \\
\leq \int_0^T \int_\Omega F(\tilde{\phi}(\textbf{x}, t), B(\textbf{x}, t))(\tilde{\phi} - w_\kappa)(\textbf{x}, t) \, dx \, dt.
$$

By the density of the set $C_0^{\infty}(\Omega) \cap K$ in $K$ proved in [14], every $\psi \in K$ can be approximated by a piecewise constant function in time $w_\kappa \in L_\kappa$ such that $w_\kappa \to \psi$ strongly in $L^2(0, T; H_0^1(\Omega))$ as $\kappa \to 0$ (note that $w_\kappa \leq \chi$ in $\Omega \times (0, T)$). Hence, (2.1a) is verified.

Let us verify the integral equality (2.1b). Let $\psi$ be a generic function in the space $L^2(0, T; L^2(\Omega))$ which satisfies $\delta_1 \psi \in L^2(\Omega \times (0, T))$ and $\psi(T, \cdot) = 0$. Using the technique in [8, Lemma 4.10], we can construct $w_m = (w_m^{(n)})_{n=0}^{N_m} \in X_{D_m}^{N_m+1}$, such that $\Pi_{D_m} w_m \to \psi$ in $L^2(0, T; L^2(\Omega))$ and $\delta_{D_m} w_m \to \delta_1 \psi$ strongly in $L^2(\Omega \times (0, T))$. Take $\psi = \delta_{D_m}^{(\frac{1}{2})} w_m^{(n)}$ as a generic function in (3.4b) and sum over $n \in [0, N_m - 1]$ to get

$$
\sum_{n=0}^{N_m-1} \int_\Omega \left[ \Pi_{D_m} B_m^{(n+1)}(\textbf{x}) - \Pi_{D_m} B_m^{(n)}(\textbf{x}) \right] \Pi_{D_m} w_m^{(n)}(\textbf{x}) \, dx \\
+ \int_0^T \int_\Omega D_B(\textbf{x}) \nabla D_m B_m(\textbf{x}, t) \cdot \nabla D_m w_m(\textbf{x}, t) \, dx \, dt \tag{4.21}
$$

Applying [8, Eq. (D.15)] to the right-hand side yields, thanks to $w_0 = 0$,

$$
\int_0^T \int_\Omega \Pi_{D_m} B_m(\textbf{x}, t) \delta_{D_m} w_m(\textbf{x}, t) \, dx \, dt - \int_\Omega \Pi_{D_m} B_m^{(0)}(\textbf{x}) \Pi_{D_m} w_m^{(0)}(\textbf{x}) \, dx \\
+ \int_0^T \int_\Omega D_B(\textbf{x}) \nabla D_m B_m(\textbf{x}, t) \cdot \nabla D_m w_m(\textbf{x}, t) \, dx \, dt
$$

\[= \int_0^T \int_\Omega G(\Pi_{D_m} A(\textbf{x}, t), \Pi_{D_m} B(\textbf{x}, t)) \Pi_{D_m} w_m(\textbf{x}, t) \, dx \, dt.\]
Using the consistency, we see that $\Pi_{D_m} B_m^{(0)} = \Pi_{D_m} B_{m\infty} \to B_{m\infty}$ in $L^2(\Omega)$. This, when passing to the limit $m \to \infty$, implies, for all $\psi$ in $L^2(0,T; H^1(\Omega))$,

$$
- \int_0^T \int_{\Omega} \partial_t \psi(\mathbf{x},t) \tilde{B}(\mathbf{x},t) \, d\mathbf{x} \, dt + \int_0^T \int_{\Omega} \mathbf{D}_B \nabla \tilde{B} \cdot \nabla \psi(\mathbf{x},t) \, d\mathbf{x} \, dt \\
= \int_0^T \int_{\Omega} G(\tilde{A}, \tilde{B}) \psi(\mathbf{x},t) \, d\mathbf{x} \, dt.
$$

Since $C^\infty([0,T]; H^1(\Omega))$ is dense in $L^2(0,T; H^1(\Omega))$, integrating by parts shows that the above equality can be expressed in the sense of distributions, which is equivalent to (2.1b).

**Step 4: Proof of the strong convergence of $\nabla_{D_m} A_m$ and $\nabla_{D_m} B_m$.** From the weak--strong convergences, we have, for all $\tilde{w}_\kappa \in \mathbb{K}$,

$$
\limsup_{m \to \infty} \int_0^T \int_{\Omega} \mathbf{D}_A \nabla_{D_m} A_m \cdot \nabla_{D_m} A_m \, d\mathbf{x} \, dt \\
\leq \int_0^T \int_{\Omega} F(\tilde{A}, \tilde{B}) (\tilde{A} - \tilde{w}_\kappa) \, d\mathbf{x} \, dt \\
+ \int_0^T \int_{\Omega} \mathbf{D}_A \nabla \tilde{A} \cdot \nabla \tilde{w}_\kappa \, d\mathbf{x} \, dt - \int_0^T \int_{\Omega} \partial_t \tilde{A} (\tilde{A} - \tilde{w}_\kappa) \, d\mathbf{x} \, dt.
$$

Thanks to the density results, for any $\varphi \in \mathbb{K}$, we can find $(\tilde{w}_\kappa)_{\kappa > 0}$ that converges to $\varphi$ in $L^2(0,T; H^1(\Omega))$, as $\kappa \to 0$. Therefore, we infer, for all $\varphi \in \mathbb{K}$,

$$
\limsup_{m \to \infty} \int_0^T \int_{\Omega} \mathbf{D}_A \nabla_{D_m} A_m \cdot \nabla_{D_m} A_m \, d\mathbf{x} \, dt \\
\leq \int_0^T \int_{\Omega} F(\tilde{A}, \tilde{B}) (\tilde{A} - \varphi) \, d\mathbf{x} \, dt \\
+ \int_0^T \int_{\Omega} \mathbf{D}_A \nabla \tilde{A} \cdot \nabla \varphi \, d\mathbf{x} \, dt - \int_0^T \int_{\Omega} \partial_t \tilde{A} (\tilde{A} - \varphi) \, d\mathbf{x} \, dt.
$$

Taking $\varphi = \tilde{A}$, the above relation yields

$$
\limsup_{m \to \infty} \int_0^T \int_{\Omega} \mathbf{D}_A \nabla_{D_m} A_m \cdot \nabla_{D_m} A_m \, d\mathbf{x} \, dt \leq \int_0^T \int_{\Omega} \mathbf{D}_A \nabla \tilde{A} \cdot \nabla \tilde{A} \, d\mathbf{x} \, dt.
$$

Together with (4.20), we obtain

$$
\lim_{m \to \infty} \int_0^T \int_{\Omega} \mathbf{D}_A \nabla_{D_m} A_m \cdot \nabla_{D_m} A_m \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} \mathbf{D}_A \nabla \tilde{A} \cdot \nabla \tilde{A} \, d\mathbf{x} \, dt,
$$

which implies

$$
0 \leq d_1 \limsup_{m \to \infty} \int_0^T \int_{\Omega} | \nabla \tilde{A} - \nabla_{D_m} A_m |^2 \, d\mathbf{x} \, dt \\
\leq \limsup_{m \to \infty} \left[ \int_0^T \int_{\Omega} \mathbf{D}_A \nabla \tilde{A} \cdot \nabla \tilde{A} + \int_0^T \int_{\Omega} \mathbf{D}_A \nabla_{D_m} A_m \cdot \nabla_{D_m} A_m \, d\mathbf{x} \, dt \\
- 2 \int_0^T \int_{\Omega} \mathbf{D}_A \nabla \tilde{A} \cdot \nabla_{D_m} A_m \, d\mathbf{x} \, dt \right] = 0,
$$
showing that $\nabla D_m A_m \rightarrow \nabla \bar{A}$ strongly in $L^2(\Omega \times (0, T))^d$. To show the strong convergence of $\nabla D_m B_m$, we begin by writing

$$\int_0^T \int_\Omega \left( \nabla D_m B_m(\mathbf{x}, t) - \nabla \bar{B}(\mathbf{x}, t) \right) \cdot \left( \nabla D_m B_m(\mathbf{x}, t) - \nabla \bar{B}(\mathbf{x}, t) \right) d\mathbf{x} dt$$

$$= \int_0^T \int_\Omega \nabla D_m B_m(\mathbf{x}, t) \cdot \nabla D_m B_m(\mathbf{x}, t) d\mathbf{x} dt$$

$$- \int_0^T \int_\Omega \nabla D_m B_m(\mathbf{x}, t) \cdot \nabla \bar{B}(\mathbf{x}, t) d\mathbf{x} dt$$

$$- \int_0^T \int \nabla \bar{B}(\mathbf{x}, t) \cdot \left( \nabla D_m B_m(\mathbf{x}, t) - \nabla \bar{B}(\mathbf{x}, t) \right) d\mathbf{x} dt. \quad (4.22)$$

Setting $\psi := B_m$ in (3.4b) and $\psi = \bar{B}$ in (2.1b), and taking $D_B(\mathbf{x}) = \text{Id}$, when passing to the limit superior, yields

$$\limsup_{m \to \infty} \int_0^T \int_\Omega \nabla D_m B_m(\mathbf{x}, t) \cdot \nabla D_m B_m(\mathbf{x}, t) d\mathbf{x} dt$$

$$= \int_0^T \int_\Omega G(\bar{A}, \bar{B}) \bar{B}(\mathbf{x}, t) d\mathbf{x} dt - \int_0^T \int_\Omega \partial_t \bar{B}(\mathbf{x}, t) \bar{B}(\mathbf{x}, t) d\mathbf{x} dt$$

$$= \int_0^T \int \nabla \bar{B}(\mathbf{x}, t) \cdot \nabla \bar{B}(\mathbf{x}, t) d\mathbf{x} dt.$$

Passing to the limit in (4.22) and using the above inequality, we reach the desired convergence result. \hfill \Box

5 An example of schemes covered by the analysis

Many numerical schemes fit into the analysis provided in this work. We construct here the nonconforming $P1$ finite element scheme for our model. Let $\mathcal{T} = (\mathcal{M}, \mathcal{E})$ be the polytopal mesh of $\Omega$ defined in [8, Definition 7.2], in which $\mathcal{M}$ and $\mathcal{E}$ consist of cells $K$ and edges $\sigma$, respectively. The elements of gradient discretization $D$ associated with the nonconforming $P1$ finite element scheme are:

- $X_{D,0} = \{ w = (w_\sigma)_{\sigma \in \mathcal{E}} : w_\sigma \in \mathbb{R} \text{ and } w_\sigma = 0 \text{ for all } \sigma \in \mathcal{E}_{\text{ext}} \}$.
- For all $w \in X_{D,0}$ and for all $K \in \mathcal{M}$, for a.e. $\mathbf{x} \in K$,

$$\Pi_D w(\mathbf{x}) = \sum_{\sigma \in \mathcal{E}_K} w_\sigma e^\sigma_K(\mathbf{x}),$$

where $e^\sigma_K$ is a basis function.

- For all $w \in X_{D,0}$ and all $K \in \mathcal{M}$, for a.e. $\mathbf{x} \in K$,

$$(\nabla_D w)_K = \nabla \left[ (\Pi_D w)_K \right] = \sum_{\sigma \in \mathcal{E}_K} w_\sigma \nabla e^\sigma_K.$$

- The approximate obstacle $\chi_D$ is defined by

$$\chi_D := \int_\sigma \chi(\mathbf{x}) d\mathbf{x}.$$
• For all $\omega \in W^{2,\infty}(\Omega)$, we can construct the interpolants $I_D \omega = I_D^\omega = (x_\sigma)_{\sigma \in E}$ with $\chi_\sigma = \omega(\bar{x}_\sigma)$.

Substituting these elements into scheme (3.4a)–(3.4b) yields the nonconforming $P_1$ finite element scheme for problem (2.1a)–(2.1b) and its convergence is therefore obtained from Theorem 4.5. Droniou et al. [7] show that $D$ given here satisfies the three properties, namely coercivity, limit-conformity, and compactness. Let us discuss the consistency property in the sense of Definition 3.4. It is shown in [7] that $\tilde{S}_{D_m}(\psi) \to 0$, for all $\psi \in H^1_0(\Omega)$, which verifies the second item. Similarly, we can prove that $S_{D_m}(\psi) \to 0$ for all $\psi \in C^2(\Omega) \cap C$. For any $\psi$, let $\omega = (\omega_\sigma)_{\sigma \in E} \in X_{D,0}$ be the interpolant such that $\omega_\sigma = \psi(\bar{x}_\sigma)$, for all $\sigma \in E$. We clearly deduce $\Pi_D \psi \leq \chi_D$ in $\Omega$. By the density results established in [14], we see that the first item is fulfilled.

Let $\varphi_m = (\omega_\sigma)_{\sigma \in E_m} \in K_{D_m}$ and $\psi_m = (\omega_\sigma)_{\sigma \in E_m} \in X_{D_m,0}$ be the interpolants such that $\varphi_m = I_D A_m$ and $\psi_m = I_D B_m$. Now [8, (B.11) in Lemma B.7] with $p = 2$ shows that there exist $C_5, C_6 > 0$ not depending on $m$ such that

$$\|\tilde{A}_{m} - \Pi_D I_D A_m\|^2_{L^2(\Omega)} \leq C_5^2 \|\nabla A_m\|^2_{L^2(\Omega)} \quad \text{and}$$

$$\|\tilde{B}_{m} - \Pi_D I_D B_m\|^2_{L^2(\Omega)} \leq C_5^2 \|\nabla B_m\|^2_{L^2(\Omega)}.$$

Passing to the limit, we see the right-hand sides tend to 0 (thanks to the classical regularity hypothesis on $A_m$ and $B_m$), and therefore the third and fourth items of the consistency property are verified. Finally, it is established in [8, proof of Theorem 12.12] that, for $\varphi \in W^{1,p}(\Omega)$, we can construct a function $\omega_m = (\omega_\sigma)_{\sigma \in E_m} \in K_{D_m}$ and find a $C_7 > 0$ not depending on $m$ such that

$$\|\nabla D_m \omega_m\|_{L^p(\Omega)} \leq C_7 \|\nabla \varphi\|_{L^p(\Omega)}.$$  

Applying this estimate (with $p = 2$) to $\varphi = A_{m}$ and $\omega_m = I_D A_m$, we deduce that

$$\|\nabla D_m I_D A_m\|^2_{L^2(\Omega)}$$

is bounded. The nonconforming $P_1$ finite element method for problem (1.1a)–(1.1f) is such that, for all $n = 0, \ldots, N - 1$, the following holds:

\[
\left(\frac{|\sigma|}{\delta t^{(n+1)}} (A^{(n+1)}_\sigma - A^{(n)}_\sigma) + \sum_{\sigma \in E_K} |\sigma| A^{(n+1)}_\sigma n_{K,\sigma} - |\sigma| F(A^{(n+1)}_\sigma, B^{(n+1)}_\sigma)\right) \\
\times (A^{(n+1)}_\sigma - \chi_\sigma) = 0, \quad \text{for all } K \in M \text{ and for all } \sigma \in E_K,
\]

\[
\frac{|\sigma|}{\delta t^{(n+1)}} (A^{(n+1)}_\sigma - A^{(n)}_\sigma) + \sum_{\sigma \in E_K} |\sigma| A^{(n+1)}_\sigma n_{K,\sigma} \geq |\sigma| F(A^{(n+1)}_\sigma, B^{(n+1)}_\sigma) \quad \text{for all } \sigma \in E,
\]

\[
A^{(n+1)}_\sigma \geq \chi_\sigma \quad \text{for all } \sigma \in E,
\]

\[
\frac{|\sigma|}{\delta t^{(n+1)}} (B^{(n+1)}_\sigma - B^{(n)}_\sigma) + \sum_{\sigma \in E_K} |\sigma| B^{(n+1)}_\sigma n_{K,\sigma} = |\sigma| G(A^{(n+1)}_\sigma, B^{(n+1)}_\sigma),
\]

for all $K \in M$ and for all $\sigma \in E_K$,

\[
A^{(n+1)}_\sigma = B^{(n+1)}_\sigma = 0 \quad \text{for all } \sigma \in \mathcal{E}_{\text{ext}},
\]

\[
(A^{(0)}_\sigma, B^{(0)}_\sigma) = (A_{\text{ini}}(x_\sigma, 0), B_{\text{ini}}(x_\sigma, 0)) \quad \text{for all } \sigma \in \mathcal{E}.
\]
6 Conclusion

We developed a gradient discretization for nonlinear system of parabolic variational inequalities. We established the existence of a continuous solution and convergence results without nonphysical hypothesis on the model data. We designed a nonconforming $P_1$ finite element method for the system of a parabolic obstacle problem and showed its convergence.

Acknowledgements

Not applicable.

Funding

Not applicable.

Availability of data and materials

Not applicable.

Declarations

Competing interests

The author declares that he has no competing interests.

Authors’ contributions

This article has one author. The author read and approved the final manuscript.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 31 May 2021 Accepted: 22 December 2021 Published online: 21 January 2022

References

1. Alnashri, Y., Alzubaidi, H.: A gradient discretisation method for anisotropic reaction-diffusion models with applications to the dynamics of brain tumors. Comput. Methods Appl. Math. 21, 753–775 (2021)
2. Alnashri, Y., Droniou, J.: Gradient schemes for the Signorini and the obstacle problems, and application to hybrid mimetic mixed methods. Comput. Math. Appl. 72, 2788–2807 (2016)
3. Alnashri, Y., Droniou, J.: A gradient discretization method to analyze numerical schemes for nonlinear variational inequalities, application to the seepage problem. SIAM J. Numer. Anal. 56, 2375–2405 (2018)
4. Barrett, J.W., Deckelnick, K.: Existence, uniqueness and approximation of a doubly-degenerate nonlinear parabolic system modelling bacterial evolution. Math. Models Methods Appl. Sci. 17, 1095–1127 (2007)
5. Boulaaras, S.M., Haouir, M.: The finite element approximation of evolutionary Hamilton–Jacobi–Bellman equations with nonlinear source terms. Indag. Math. 24, 161–173 (2013)
6. Droniou, J., Eymard, R.: Uniform-in-time convergence of numerical methods for non–linear degenerate parabolic equations. Numer. Math. 132, 721–766 (2016)
7. Droniou, J., Eymard, R., Feron, P.: Gradient schemes for Stokes problem. IMA J. Numer. Anal. 36, 1636–1669 (2015)
8. Droniou, J., Eymard, R., Gallouët, T., Guichard, C., Herbin, R.: The Gradient Discretisation Method, Mathematics & Applications. Springer, Heidelberg (2018)
9. Droniou, J., Eymard, R., Herbin, R.: Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic problems. Math. Models Methods Appl. Sci. 23, 2395–2432 (2013)
10. Droniou, J., Eymard, R., Herbin, R.: Gradient schemes: generic tools for the numerical analysis of diffusion equations. M2AN Math. Model. Numer. Anal. 50, 749–781 (2016)
11. Eymard, R., Feron, P., Gallouët, T., Herbin, R., Guichard, C.: Gradient schemes for the Stefan problem. Int. J. Finite Vol. 10s (2013)
12. Eymard, R., Guichard, C., Herbin, R.: Small-stencil 3D schemes for diffusive flows in porous media. ESAIM: Math. Model. Numer. Anal. 46, 265–290 (2012)
13. Eymard, R., Guichard, C., Herbin, R., Masson, R.: Gradient schemes for two-phase flow in heterogeneous porous media and Richards equation. Z. Angew. Math. Mech. 94, 560–585 (2014)
14. Glowinski, R.: Numerical Methods for Non Linear Variational Problems. Tata Institute of Fundamental Research, Bombay (1980)
15. Heywood, J.G., Rannacher, R.: Finite-element approximation of the nonstationary Navier–Stokes problem part IV: error analysis for second-order time discretization. SIAM J. Numer. Anal. 27, 353–384 (1990)
16. Islam, S., Alam, M.N., Al-Asad, M.F., Tunc, C.: An analytical technique for solving new computational solutions of the modified Zakharov–Kuznetsov equation arising in electrical engineering. J. Appl. Comput. Mech. 7, 715–726 (2021)
17. Johnson, C.: A convergence estimate for an approximation of a parabolic variational inequality. SIAM J. Numer. Anal. 13, 599–606 (1976)
18. Khan, H., Abdeljawad, T., Tunc, C., Alkhazzan, A., Khan, A.: Minkowski’s inequality for the AB-fractional integral operator. J. Inequal. Appl. (2019)
19. Peszynska, M., Trykozko, A., Il'tis, G., Schlüeter, S., Wildenschindler, D.: Biofilm growth in porous media: experiments, computational modeling at the porescale, and upscaling. Adv. Water Resour. 95, 288–301 (2016)
20. Saker, S.H., Tunc, C., Mahmoud, R.R.: New Carlson–Bellman and Hardy–Littlewood dynamic inequalities. Math. Inequal. Appl., 967–983 (2018)
21. Verba, C., Thurber, A., Alleau, Y., Koley, D., Colwell, F., Torres, M.: Mineral changes in cement-sandstone matrices induced by biocementation. Int. J. Greenh. Gas Control 49, 312–322 (2016)
22. Vuik, C.: An $L^r$-error estimate for an approximation of the solution of a parabolic variational inequality. Numer. Math. 57, 453–472 (1990)
23. Wheeler, M.F.: A priori $L^2$ error estimates for Galerkin approximations to parabolic partial differential equations. SIAM J. Numer. Anal. 10, 723–759 (1973)