Convergence Analysis of a Symmetric Dual-Wind Discontinuous Galerkin Method for a Parabolic Variational Inequality

Satyajith Bommana Boyana[a], Thomas Lewis[a], Aaron Rapp[b], Yi Zhang[a,*]

[a]Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, NC 27402, United States

[b]Department of Mathematical Sciences, University of the Virgin Islands, St. Thomas, VI 00802, United States

Abstract

This paper investigates a symmetric dual-wind discontinuous Galerkin (DG) method for solving parabolic variational inequalities. By employing a symmetric dual-wind DG discretization in space and a backward Euler discretization in time, we propose a fully discrete scheme to solve a time-dependent obstacle problem. Under reasonable regularity assumptions on the exact solution, we prove the convergence of numerical solutions with rates in the $L^\infty(L^2)$ and $L^2(H^1)$-like energy errors by introducing a new interpolation operator which is a combination of the standard interpolation operator and a positive-preserving interpolation operator. Numerical experiments are provided to validate the effectiveness of the proposed method.

Keywords: parabolic variational inequality, obstacle problem, second order variational inequality, finite element, discontinuous Galerkin methods, *a priori* analysis

2010 MSC: 65K15, 65N30

1. Introduction

Let Ω be a bounded convex polygonal domain in $\mathbb{R}^2$, $T_F > 0$, and set $J = [0, T_F]$. For a given $f \in C(J; L^\infty(\Omega))$ and $\psi \in H^1(\Omega)$ with $\psi \leq 0$ a.e. on $\partial\Omega$, we consider the parabolic variational inequality: For all $t \in (0, T_F]$, find $u(t) \in K \subset H_0^1(\Omega)$ such that

$$(\partial_t u, v - u) + a(u, v - u) \geq (f(t), v - u) \qquad \forall\, v \in K, \tag{1.1a}$$

$$u(0) = u_0, \tag{1.1b}$$

where $u_0 \in K$ is the given initial condition, K is the constrained set

$$K := \{v \in H_0^1(\Omega) : v \geq \psi \text{ a.e. in } \Omega\}, \tag{1.2}$$

and the bilinear forms $(\cdot, \cdot)$ and $a(\cdot, \cdot)$ are defined by

$$(v, w) = \int_\Omega v\, w\, dx, \qquad a(v, w) = \int_\Omega \nabla v \cdot \nabla w\, dx \qquad \forall\, v, w \in H^1(\Omega). \tag{1.3}$$

*Corresponding author

Email addresses: `s_bomman@uncg.edu` (Satyajith Bommana Boyana), `tllewis3@uncg.edu` (Thomas Lewis), `aaron.rapp@uvi.edu` (Aaron Rapp), `y_zhang7@uncg.edu` (Yi Zhang)

Throughout the paper we follow the standard functional space and function notation in [13, 1, 6]. In particular, $W^{k,p}(\Omega)$ for nonnegative integer k and $1 \le p \le \infty$ denotes the Sobolev space, $H^k(\Omega)$ denotes the space when $p = 2$, $C(J;X)$ denotes the space of continuous functions from J to the normed space X, and $L^p(J;X)$ is the space equipped with the norm

$$\|v\|_{L^p(J;X)} = \left(\int_0^{T_F} \|v(t)\|_X^p \, dt \right)^{\frac{1}{p}} \qquad 1 \le p < \infty,$$
$$\|v\|_{L^\infty(J;X)} = \operatorname*{ess\,sup}_{t \in J} \|v(t)\|_X.$$

The problem (1.1) is a generalization of an elliptic obstacle obstacle problem with the obstacle function ψ. It is also a special case of variational inequalities (VIs) that arise from a wide range of applications including mechanics, physics, engineering, economics, finance, mathematical programming, optimal control, and optimization, etc [29, 36, 20]. Due to the presence of the the obstacle function, the variational inequality becomes a nonlinear problem which leads to challenges in both the theoretical and numerical analysis.

The existence and uniqueness of the solution to the problem (1.1) follows from the standard theory for VIs [9, 10, 29, 36, 20]. Under the assumptions on the data

$$f \in C(J; L^\infty(\Omega)), \quad \frac{\partial f}{\partial t} \in L^2(J; L^\infty(\Omega)), \tag{1.4}$$
$$\psi \in W^{2,\infty}(\Omega), \quad u_0 \in W^{2,\infty}(\Omega) \cap K, \tag{1.5}$$

the exact solution u satisfies [10]

$$u \in L^\infty(J; W^{2,p}(\Omega)) \qquad 1 \le p < \infty, \tag{1.6}$$
$$\frac{\partial u}{\partial t} \in L^2(J; H_0^1(\Omega)) \cap L^\infty(J; L^\infty(\Omega)). \tag{1.7}$$

Moreover, denoting $\partial^+ u/\partial t$ to be the right-hand derivative of u with respect to t, the solution to (1.1) satisfies

$$\frac{\partial^+ u}{\partial t} = \Delta u + f \qquad \text{a.e. on } \Omega^+(t), \tag{1.8}$$
$$\frac{\partial^+ u}{\partial t} = \max\{f + \Delta\psi, 0\} \qquad \text{a.e. on } \Omega^0(t), \tag{1.9}$$

where the contact set $\Omega^0(t)$ and the non-contact set $\Omega^+(t)$ are given by

$$\Omega^0(t) := \{x \in \Omega : u(x,t) = \psi(x)\}, \tag{1.10}$$
$$\Omega^+(t) := \{x \in \Omega : u(x,t) > \psi(x)\} \tag{1.11}$$

for all $t \in J$. The boundary between $\Omega^0(t)$ and $\Omega^+(t)$ is called the free boundary which is the main source of singularities for the exact solution. We denote

$$\sigma(t) = \frac{\partial^+ u}{\partial t}(t) - \Delta u(t) + f(t). \tag{1.12}$$

For any $t \in J$, the following complementarity form of (1.1) follows from (1.8)–(1.9)

$$\sigma(t) \ge 0 \quad \text{a.e. in } \Omega, \tag{1.13}$$

$$(\sigma(t), u(t) - \psi) = 0 \quad \text{a.e. in } \Omega. \tag{1.14}$$

The study of VIs is an active area and we refer the reader to [14, 29, 20, 25] for an extensive survey on the numerical analysis of VIs. In the case of elliptic VIs, various finite element methods including conforming, nonconforming, virtual elements, and a family of DG methods have been studied in [16, 11, 40, 8, 7, 5, 4, 38, 39, 24, 31] and the references therein. Optimal *a priori* error estimates were obtained using the established regularity results for the exact solution; the key ingredient in the analysis is the use of the complementarity condition which is valid since the exact solution has the full regularity (i.e., $u \in H^2(\Omega)$). For the parabolic VI (1.1), there is an another difficulty arising in the low regularity of the time derivative of the exact solution (see (1.7)). In [27], a fully discrete scheme with a standard linear conforming finite element method and backward Euler time-stepping was considered, where the error estimate of the form $O(h + \tau^{\frac{3}{4}}(\log(\tau^{-1}))^{\frac{1}{4}})$ was obtained. The sub-optimal rate of τ is due to the low regularity of $\partial u/\partial t$. However, the analysis in [27] was carried out under the assumption of the zero obstacle function. Recently in [21, 22, 23, 33, 34], the analysis has been extended to general obstacle functions for conforming, nonconforming, and a family of DG methods with the help of the positive-preserving interpolation operator [12] with nonhomogeneous Dirichlet boundary conditions. Similar results were achieved in these works by converting the original parabolic VI (1.1) into the problem with a zero obstacle and fixed nonhomogeneous boundary conditions.

In this work, we are interested in applying and analyzing symmetric dual-wind DG methods [30] to approximate solutions of (1.1). Such methods have been applied to elliptic VIs in [31], where sharp error estimates were derived for both linear and quadratic elements. The methods followed the framework of the DG differential calculus [17], where discrete partial derivatives are used to approximate classical partial derivatives. In particular, the dual-wind DG methods utilize both the up-wind discrete gradient operator and the down-wind discrete gradient operator and are stable without explicitly penalizing jump discontinuities as is standard for classical DG methods [2]. The goal of this paper is to extend the work in [31] to the case of parabolic VIs and propose a more simple *a priori* convergence analysis. We will combine the dual-wind DG discretization in space and the backward Euler discretization in time to obtain a fully discrete scheme. Due to the use of the discrete gradient operators, such an analysis is more subtle. To this aim, we assume the data satisfies (1.4)–(1.5) such that the regularity results in (1.6)–(1.7) hold. A new interpolation operator will be introduced to deal with general obstacle functions, and thus we achieve a simplified convergence analysis and prove $O(h + \tau^{\frac{3}{4}}(\log(\tau^{-1}))^{\frac{1}{4}})$ convergence for the proposed method.Furthermore, we demonstrate $O(h+\tau(\log(\tau^{-1}))^{\frac{1}{2}})$ convergence is achievable under stronger assumptions. In addition, the flexibility of a penalty-free method makes this approximation desirable for more complicated problems.

The rest of the paper is organized as follows. In Section 2 we introduce a fully discrete symmetric dual-wind DG method and derive several preliminary results including the interpolation operator that will be useful for the convergence analysis. In Section 3 we present error estimates for the convergence of the proposed method. We report numerical experiments in Section 4. Finally, we include a summary in Section 5.

2. The Fully Discrete Method

The fully discrete dual-wind DG method is formulated in this section. We will first introduce the DG differential operators that are used to define the dual-wind DG method. Next the fully discrete problem is presented and finally some useful lemmas for the discrete differential operators and interpolation error estimates are derived.

2.1. DG Differential Operators

Standard functional space and function notation in [1, 6] will be adopted in this paper. For convenience, we introduce the following notations:

- $\mathcal{T}_h$: a shape-regular simplicial triangulation of $\Omega \subset \mathbb{R}^2$ [13, 6],
- h_T: the diameter of the simplex $T \in \mathcal{T}_h$,
- $h := \max_{T\in\mathcal{T}_h} h_T$: mesh size of the triangulation,
- $\mathcal{E}_h^I$: the set of interior edges in $\mathcal{T}_h$,
- $\mathcal{E}_h^B$: the set of boundary edges in $\mathcal{T}_h$,
- $\mathcal{E}_h := \mathcal{E}_h^I \cup \mathcal{E}_h^B$: the set of edges in $\mathcal{T}_h$,
- $\mathbb{V}_T$: the set of vertices of the simplex $T \in \mathcal{T}_h$,
- $\mathbb{V}_h$: the set of all vertices of $\mathcal{T}_h$,
- $\boldsymbol{W}^{k,p}(\Omega) := [W^{k,p}(\Omega)]^2$: vector-valued Sobolev spaces with each component in $W^{k,p}(\Omega)$,
- $\boldsymbol{H}^k(\Omega) := [H^k(\Omega)]^2$: vector-valued Sobolev spaces with each component in $H^k(\Omega)$,
- $W^{k,p}(\mathcal{T}_h) := \prod_{T\in\mathcal{T}_h} W^{k,p}(T)$: piecewise Sobolev spaces,
- $\boldsymbol{W}^{k,p}(\mathcal{T}_h) := \prod_{T\in\mathcal{T}_h} \boldsymbol{W}^{k,p}(T)$: vector-valued piecewise Sobolev spaces,
- $(v, w)_{\mathcal{T}_h} := \sum_{T\in\mathcal{T}_h} \int_T vw\,dx$: piecewise L^2 inner product over $\mathcal{T}_h$,
- $\langle v, w\rangle_{\mathcal{S}_h} := \sum_{e\in\mathcal{S}_h} \int_e vw\,ds$: piecewise L^2 inner product over a subset $\mathcal{S}_h \subset \mathcal{E}_h$,
- $\|v\|^2_{L^2(\mathcal{T}_h)} := (v, v)_{\mathcal{T}_h}$: square of the L^2 norm over $\mathcal{T}_h$.

We also define $\mathcal{V}_h := W^{1,1}(\mathcal{T}_h) \cap C^0(\mathcal{T}_h)$ and $\boldsymbol{\mathcal{V}}_h := [\mathcal{V}_h]^2$. Let V_h be the space of piecewise linear polynomials with respect to the triangulation $\mathcal{T}_h$, i.e., $V_h = \{v \in L^2(\Omega) : v|_T \in \mathbb{P}_1(T) \quad \forall\, T \in \mathcal{T}_h\}$, where $\mathbb{P}_1(T)$ denotes the space of linear polynomials over T. The corresponding vector-valued DG space is then given by $\boldsymbol{V}_h := [V_h]^2$. Throughout the paper, we will use bold-face notation to indicate vector-valued spaces.

In the following, we will define the jump and average operator across the edge $e \in \mathcal{E}_h$. For simplicity, we assume the global labeling number of T^+ is larger than that of T^-. For $e = \partial T^+ \cap \partial T^-$ with $T^+, T^- \in \mathcal{T}_h$, we define

$$[\![v]\!]|_e := v^+ - v^-, \qquad \{v\}|_e := \frac{1}{2}\big(v^+ + v^-\big) \qquad \forall\, v \in \mathcal{V}_h,$$

where $v^\pm := v|_{T^\pm}$. For $e = \partial T^+ \cap \partial\Omega$ for a boundary simplex $T^+ \in \mathcal{T}_h$, define

$$[\![v]\!]|_e := v^+, \qquad \{v\}|_e := v^+ \qquad \forall\, v \in \mathcal{V}_h.$$

Set $\boldsymbol{n}_e = (n_e^{(1)}, n_e^{(2)})^t := \boldsymbol{n}_{T^+}|_e = -\boldsymbol{n}_{T^-}|_e$ to be the unit normal on $e \in \mathcal{E}_h^I$. For $i = 1, 2$ and $v \in \mathcal{V}_h$, we define the following two trace operators on $e \in \mathcal{E}_h^I$ in the direction x_i:

$$\mathcal{Q}_i^+(v) := \begin{cases} v|_{T^+}, & \text{if } n_e^{(i)} > 0, \\ v|_{T^-}, & \text{if } n_e^{(i)} < 0, \\ \{v\}, & \text{if } n_e^{(i)} = 0 \end{cases} \quad \text{and} \quad \mathcal{Q}_i^-(v) := \begin{cases} v|_{T^-}, & \text{if } n_e^{(i)} > 0, \\ v|_{T^+}, & \text{if } n_e^{(i)} < 0, \\ \{v\}, & \text{if } n_e^{(i)} = 0. \end{cases}$$

Note that the operators $\mathcal{Q}_i^{\pm}(v)$ can be equivalently defined by $\mathcal{Q}_i^{\pm}(v) := \{v\} \pm \frac{1}{2}\mathrm{sgn}(n_e^{(i)})[\![v]\!]$, which can be understood as the "forward" and "backward" limit of v in the x_i direction on $e \in \mathcal{E}_h^I$. For $e = \partial T^+ \cap \partial\Omega \in \mathcal{E}_h^B$, we simply take $\boldsymbol{n}_e = n_{T^+}$ and $\mathcal{Q}_i^{\pm}(v) = v^+$.

Next, using the "forward" and "backward" trace operators, we are able to define the discrete partial derivative operators $\partial_{h,x_i}^{\pm}, \partial_{h,x_i}^{\pm,g} : \mathcal{V}_h \to V_h$ $(i = 1, 2)$ as

$$\big(\partial_{h,x_i}^{\pm} v, \varphi_h\big)_{\mathcal{T}_h} := \big\langle \mathcal{Q}_i^{\pm}(v) n^{(i)}, [\![\varphi_h]\!] \big\rangle_{\mathcal{E}_h} - \big(v, \partial_{x_i}\varphi_h\big)_{\mathcal{T}_h} \qquad \forall\, \varphi_h \in V_h, \tag{2.1}$$

$$\big(\partial_{h,x_i}^{\pm,g} v, \varphi_h\big)_{\mathcal{T}_h} := \big\langle \mathcal{Q}_i^{\pm}(v) n^{(i)}, [\![\varphi_h]\!] \big\rangle_{\mathcal{E}_h^I} + \big\langle g n^{(i)}, \varphi_h \big\rangle_{\mathcal{E}_h^B} - \big(v, \partial_{x_i}\varphi_h\big)_{\mathcal{T}_h} \qquad \forall\, \varphi_h \in V_h, \tag{2.2}$$

where $g \in L^1(\partial\Omega)$. With these notations, we can define the discrete gradient operators for $v \in \mathcal{V}_h$ as

$$\nabla_h^{\pm} v = (\partial_{h,x_1}^{\pm} v, \partial_{h,x_2}^{\pm} v), \qquad \nabla_{h,g}^{\pm} v = (\partial_{h,x_1}^{\pm,g} v, \partial_{h,x_2}^{\pm,g} v).$$

Furthermore, we define $\overline{\nabla}_h v$, $\overline{\nabla}_{h,g} v$, $\overline{\partial}_{h,x_i} v$, and $\overline{\partial}_{h,x_i}^{g} v$ to be the average of "forward" and "backward" discrete operators, i.e.,

$$\overline{\partial}_{h,x_i} v := \frac{1}{2}(\partial_{h,x_i}^{+} v + \partial_{h,x_i}^{-} v), \qquad \overline{\partial}_{h,x_i}^{g} v := \frac{1}{2}(\partial_{h,x_i}^{+,g} v + \partial_{h,x_i}^{-,g} v),$$
$$\overline{\nabla}_h v := \frac{1}{2}(\nabla_h^{+} v + \nabla_h^{-} v), \qquad \overline{\nabla}_{h,g} v := \frac{1}{2}(\nabla_{h,g}^{+} v + \nabla_{h,g}^{-} v).$$

In particular, whenever $g = 0$, the discrete gradient operators $\nabla_{h,0}^{\pm} v$ and $\overline{\nabla}_{h,0} v$ are defined using $\partial_{h,x_i}^{\pm,0} v$ and $\overline{\partial}_{h,x_i}^{0} v$.

Combining (2.1), (2.2), and integration by parts, the discrete gradient operators also satisfy the following conditions:

$$\begin{aligned}\big(\nabla_h^{\pm} v, \boldsymbol{\varphi}_h)_{\mathcal{T}_h} &= (\nabla v, \boldsymbol{\varphi}_h)_{\mathcal{T}_h} - \big\langle [\![v]\!], \{\boldsymbol{\varphi}_h\} \cdot \boldsymbol{n} \big\rangle_{\mathcal{E}_h^I} \\ &\quad \pm \frac{1}{2} \sum_{i=1}^{2} \big\langle [\![v]\!] |n^{(i)}|, [\![\boldsymbol{\varphi}_h^{(i)}]\!] \big\rangle_{\mathcal{E}_h^I} \qquad \forall\, v \in \mathcal{V}_h,\ \forall\, \boldsymbol{\varphi}_h \in \boldsymbol{V}_h,\end{aligned} \tag{2.3}$$

$$\begin{aligned}(\nabla_{h,g}^{\pm} v, \boldsymbol{\varphi}_h)_{\mathcal{T}_h} &= (\nabla v, \boldsymbol{\varphi}_h)_{\mathcal{T}_h} - \big\langle [\![v]\!], \{\boldsymbol{\varphi}_h\} \cdot \boldsymbol{n} \big\rangle_{\mathcal{E}_h} \\ &\quad + \big\langle g, \boldsymbol{\varphi}_h \cdot \boldsymbol{n} \big\rangle_{\mathcal{E}_h^B} \pm \frac{1}{2} \sum_{i=1}^{2} \big\langle [\![v]\!] |n^{(i)}|, [\![\boldsymbol{\varphi}_h^{(i)}]\!] \big\rangle_{\mathcal{E}_h^I} \qquad \forall\, v \in \mathcal{V}_h,\ \forall\, \boldsymbol{\varphi}_h \in \boldsymbol{V}_h,\end{aligned} \tag{2.4}$$

where $g \in L^1(\partial\Omega)$. Finally, the following properties can be derived immediately by the definition of the discrete gradient operators and integration by parts [17]:

$$(\nabla_{h,g}^{\pm} v, \boldsymbol{\varphi}_h)_{\mathcal{T}_h} = \big(\nabla_{h,0}^{\pm} v, \boldsymbol{\varphi}_h)_{\mathcal{T}_h} + \big\langle g, \boldsymbol{\varphi}_h \cdot \boldsymbol{n} \big\rangle_{\mathcal{E}_h^B} \qquad \forall\, v \in \mathcal{V}_h,\ \forall\, \boldsymbol{\varphi}_h \in \boldsymbol{V}_h, \tag{2.5}$$

$$(\partial_{h,x_i}^{\pm} v_h, \varphi_h)_{\mathcal{T}_h} = -(v_h, \partial_{h,x_i}^{\mp} \varphi_h)_{\mathcal{T}_h} + \big\langle v_h, \varphi_h n^{(i)} \big\rangle_{\mathcal{E}_h^B} \qquad \forall\, v_h,\ \varphi_h \in V_h. \tag{2.6}$$

2.2. The Fully Discrete Dual-Wind DG Method

In this subsection we combine the dual-wind DG method [30, 31] with a backward Euler discretization in time to obtain a fully discrete scheme for (1.1).

Let $t_n = n\tau$ $(n = 0, 1, ..., N)$ be a uniform partition of J with $\tau = \frac{T_F}{N}$ and V_h be the discontinuous piecewise linear space associated with $\mathcal{T}_h$ that is introduced in Section 2.1. We define the approximation of the constrained set K to be

$$K_h := \{v_h \in V_h : v_h(p) \geq \psi(p) \quad \forall\, p \in \mathbb{V}_T,\ \forall\, T \in \mathcal{T}_h\}. \tag{2.7}$$

The fully discrete scheme for approximating of (1.1) seeks $u_h^n \in K_h$ $(n = 1, ..., N)$ such that

$$(\partial u_h^n, v_h - u_h^n) + a_h(u_h^n, v_h - u_h^n) \geq (f(t_n), v_h - u_h^n) \qquad \forall\, v_h \in K_h, \tag{2.8}$$

where

$$\partial u_h^n = \frac{u_h^n - u_h^{n-1}}{\tau} \tag{2.9}$$

and

$$a_h(v_h, w_h) := \frac{1}{2}\Big(\big(\nabla_{h,0}^+ v_h, \nabla_{h,0}^+ w_h\big)_{\mathcal{T}_h} + \big(\nabla_{h,0}^- v_h, \nabla_{h,0}^- w_h\big)_{\mathcal{T}_h}\Big) + \Big\langle \frac{\gamma_e}{h_e} [\![v_h]\!], [\![w_h]\!] \Big\rangle_{\mathcal{E}_h}. \tag{2.10}$$

Here γ_e is a "penalty" parameter on $e \in \mathcal{E}_h$ that will be determined later. We choose u_h^0 to be $I_{h,1}u_0$, where $I_{h,1}$ is the standard nodal interpolation operator that satisfies the following property [13, 6]:

$$\|v - I_{h,1}v\|_{L^2(T)} + h_T|v - I_{h,1}v|_{H^1(T)} \leq Ch_T^{s+1}|v|_{H^{s+1}(T)} \qquad \forall\, T \in \mathcal{T}_h, \tag{2.11}$$

for all $v \in H^{s+1}(\Omega)$ with $s \geq 1$. Utilizing Lemma 2.3 and (2.11), the global estimate in terms of the energy norm was shown in [31] for $s \geq 1$:

$$\|v - I_{h,1}v\|_h^2 \leq Ch^{2s}|v|_{H^{s+1}(\Omega)}^2 \qquad \forall\, v \in H^{s+1}(\Omega). \tag{2.12}$$

Note that the method (2.8) is unconditionally stable with respect to τ similar to the implicit scheme for the heat equation.

To measure the error, we introduce the notation

$$\|v\|_{1,h}^2 := \frac{1}{2}\Big(\|\nabla_{h,0}^+ v\|_{L^2(\Omega)}^2 + \|\nabla_{h,0}^- v\|_{L^2(\Omega)}^2\Big) \qquad \forall\, v \in \mathcal{V}_h, \tag{2.13}$$

$$\|v\|_h^2 := \|v\|_{1,h}^2 + \sum_{e \in \mathcal{E}_h} \frac{\gamma_e}{h_e}\big\|[\![v]\!]\big\|_{L^2(e)}^2 \qquad \forall\, v \in \mathcal{V}_h. \tag{2.14}$$

It was shown in [30] that when the triangulation $\mathcal{T}_h$ is quasi-uniform and each simplex in the triangulation has at most one boundary edge, then there exists a constant $C_* > 0$ independent of h such that

$$C_* \sum_{e \in \mathcal{E}_h} h_e^{-1}\|[\![v_h]\!]\|_{L^2(e)}^2 \leq \|v_h\|_{1,h}^2 \qquad \forall\, v_h \in V_h. \tag{2.15}$$

Let $\gamma_{\min} := \min_{e \in \mathcal{E}_h} \gamma_e$. In the case of $\gamma_{min} > 0$, we have

$$\gamma_{\min} \sum_{e \in \mathcal{E}_h} h_e^{-1}\big\|[\![v_h]\!]\big\|_{L^2(e)}^2 \leq \|v_h\|_h^2 \qquad \forall\, v_h \in V_h. \tag{2.16}$$

In the case of $-C_* < \gamma_{\min} \leq 0$, we assume $\mathcal{T}_h$ is quasi-uniform and each simplex in the triangulation has at most one boundary edge such that (2.15) holds and we have [31]

$$\big(C_* + \gamma_{\min}\big) \sum_{e \in \mathcal{E}_h} h_e^{-1}\big\|[\![v_h]\!]\big\|_{L^2(e)}^2 \leq \|v_h\|_h^2 \qquad \forall\, v_h \in V_h. \tag{2.17}$$

From (2.16) and (2.17), it is clear that $\|\cdot\|_h$ is non-negative and defines a mesh-dependent norm. The bilinear form $a_h(\cdot,\cdot)$ is naturally coercive with respect to $\|\cdot\|_h$ for each case. Furthermore, we have the following boundedness results for $a_h(\cdot,\cdot)$ with respect to $\|\cdot\|_h$:

$$\gamma_{\min} \geq 0: \qquad a_h(v, w) \leq \|v\|_h\|w\|_h \qquad \forall\, v,\, w \in V_h + H^2(\Omega), \tag{2.18}$$

$$-C_* < \gamma_{\min} < 0: \qquad a_h(v_h, w_h) \le \|v_h\|_h \|w_h\|_h \qquad \forall\, v_h,\, w_h \in V_h. \tag{2.19}$$

Let $I_{h,1} : H^2(\Omega) \longrightarrow V_h \cap H^1(\Omega)$ be the standard nodal interpolation operator. It is clear $I_{h,1}$ preserves the constraints, i.e., $I_{h,1}\psi \in K_h$. Since K_h is nonempty and convex, we can apply the coercivity and boundedness of the bilinear form to conclude that the fully discrete problem (2.8) is well-posed (see [20, 36, 29]).

2.3. Preliminary Results

In this subsection, we establish several technical lemmas that will be crucial in the error analysis in Section 3.

First, we immediately obtain from (2.14), (2.16), and (2.17) that there exists a positive constant C independent of h such that

$$\|v_h\|_{1,h} \le C\|v_h\|_h \qquad \forall\, v_h \in V_h. \tag{2.20}$$

The following lemma establishes the relation between the classical gradient and the discrete gradient operator whose proof can be found in [31].

Lemma 2.1. *If $\gamma_{\min} > 0$, there holds*

$$\|\nabla v_h\|^2_{L^2(\mathcal{T}_h)} \le C\left(1 + \frac{1}{\gamma_{\min}}\right)\|v_h\|_h^2 \qquad \forall v_h \in V_h. \tag{2.21}$$

If $-C_* \le \gamma_{\min} \le 0$, there holds

$$\|\nabla v_h\|^2_{L^2(\mathcal{T}_h)} \le C\left(1 + \frac{1 + |\gamma_{\min}|}{C_* + \gamma_{\min}}\right)\|v_h\|_h^2 \qquad \forall v_h \in V_h. \tag{2.22}$$

It is well-known that the discrete Poincaré inequality holds [3, 19]:

$$\|v\|^2_{L^2(\Omega)} \le C\left(\|\nabla v\|^2_{L^2(\mathcal{T}_h)} + \sum_{e \in \mathcal{E}_h} \frac{1}{h_e}\|[\![v]\!]\|^2_{L^2(e)}\right) \qquad \forall\, v \in H^1(\mathcal{T}_h). \tag{2.23}$$

However the piecewise H^1-seminorm is evaluated using the classical gradient operator. With the help of Lemma 2.1, we are able to derive a similar discrete Poincaré inequality in the DG space using the energy norm $\|\cdot\|_h$.

Lemma 2.2. *There exists a positive constant C independent of h such that*

$$\|v\|^2_{L^2(\Omega)} \le C\|v\|_h^2 \qquad \forall v \in V_h. \tag{2.24}$$

Proof. We first assume $\gamma_{\min} > 0$. Then the estimate (2.24) immediately follows from (2.23), (2.21), and (2.16). In the case $-C_* \le \gamma_{\min} \le 0$, we can apply (2.22) and (2.17) instead of (2.21) and (2.16) to obtain the estimate (2.24). □

One important property of the discrete derivatives of a function v is that they reduce to the L^2 projection of the derivative of v provided $v \in H^1(\Omega)$. Indeed, the following lemma can be proved using (2.3)–(2.4) (see [17]).

Lemma 2.3. *For any $v \in H^1(\Omega)$, both $\partial^{\pm}_{h,x_i} v$ and $\overline{\partial}_{h,x_i} v$ are the L^2 projections of $\partial_{x_i} v$ onto V_h, i.e.,*

$$(\partial^{\pm}_{h,x_i} v, w_h)_{\mathcal{T}_h} = (\overline{\partial}_{h,x_i} v, w_h)_{\mathcal{T}_h} = (\partial_{x_i} v, w_h)_{\mathcal{T}_h} \qquad \forall\, w_h \in V_h. \tag{2.25}$$

Moreover, if $v \in H^1(\Omega)$ satisfies $v = g$ on $\partial\Omega$ for some $g \in L^1(\partial\Omega)$, then both $\partial^{\pm,g}_{h,x_i} v$ and $\overline{\partial}^{g}_{h,x_i} v$ are the L^2 projection of $\partial_{x_i} v$ onto V_h.

From Lemma 2.3, we obtain a useful formula connecting the Laplacian operator and the discrete bilinear form.

Lemma 2.4. *Let $v \in H^2(\Omega) \cap H^1_0(\Omega)$ and $v_h \in V_h$. There holds*

$$a_h(v, v_h) = (-\Delta v, v_h)_{\mathcal{T}_h} - \langle \{\overline{\nabla}_{h,0} v - \nabla v\} \cdot \mathbf{n}, [\![v_h]\!] \rangle_{\mathcal{E}_h}. \tag{2.26}$$

Proof. Since $v_h \in V_h$ and $v = 0$ on $\partial\Omega$, we have

$$\begin{aligned}
a_h(v, v_h) &= \frac{1}{2}\big((\nabla^+_{h,0} v, \nabla^+_{h,0} v_h)_{\mathcal{T}_h} + (\nabla^-_{h,0} v, \nabla^-_{h,0} v_h)_{\mathcal{T}_h}\big) && (2.27)\\
&= \frac{1}{2}\big((\nabla^+_{h,0} v, \nabla v_h)_{\mathcal{T}_h} + (\nabla^-_{h,0} v, \nabla v_h)_{\mathcal{T}_h}\big) - \langle \{\overline{\nabla}_{h,0} v\} \cdot \mathbf{n}, [\![v_h]\!] \rangle_{\mathcal{E}_h}\\
&= \frac{1}{2}\big((\nabla v, \nabla v_h)_{\mathcal{T}_h} + (\nabla v, \nabla v_h)_{\mathcal{T}_h}\big) - \langle \{\overline{\nabla}_{h,0} v)\} \cdot \mathbf{n}, [\![v_h]\!] \rangle_{\mathcal{E}_h}\\
&= (\nabla v, \nabla v_h)_{\mathcal{T}_h} - \langle \{\overline{\nabla}_{h,0} v\} \cdot \mathbf{n}, [\![v_h]\!] \rangle_{\mathcal{E}_h}\\
&= (-\Delta v, v_h)_{\mathcal{T}_h} - \langle \{\overline{\nabla}_{h,0} v - \nabla v_h\} \cdot \mathbf{n}, [\![v_h]\!] \rangle_{\mathcal{E}_h}, && (2.28)
\end{aligned}$$

where in the second equality we applied (2.4), in the third equality we used the L^2 projection properties of $\nabla^{\pm}_{h,0}$ in Lemma 2.3, and in the last equality we applied integration by parts. □

In the analysis, we need to consider the interpolation of the time derivative $\frac{\partial u}{\partial t}$. Due to its low regularity, the standard nodal interpolation $I_{h,1} \frac{\partial u}{\partial t}$ is not well-defined. Therefore, a new operator is needed to deal with the low regularity and the obstacle constraints simultaneously. In [12, 21], a positive-preserving operator $I_{h,2}$ from $H^1(\Omega) \cap C(\overline{\partial\Omega})$ to $V_h \cap H^1(\Omega)$ was constructed by

$$I_{h,2} v(x) := \sum_{p \in \mathbb{V}_h} \alpha_p \phi_p(x), \tag{2.29}$$

where ϕ_p is the nodal basis function associated with the vertex $p \in \mathbb{V}_h$ and

$$\alpha_p = \begin{cases} \frac{1}{|B_p|} \int_{B_p} v(x)\, dx & \forall\, p \in \mathbb{V}_h \cap \Omega, \\ v(p) & \forall\, p \in \mathbb{V}_h \cap \partial\Omega. \end{cases} \tag{2.30}$$

Here B_p denotes the largest inscribed disk for the local patch $\omega_p := \bigcup\{T \in \mathcal{T}_h : p \in \mathbb{V}_T\}$. The following approximation properties can be found in [21]:

$$\begin{aligned}
\|v - I_{h,2} v\|_{L^2(T)} &\le C h_T \|v\|_{H^1(S_T)} && \forall\, v \in H^1_0(\Omega), && (2.31)\\
\|v - I_{h,2} v\|_{L^2(T)} &\le C h_T^2 \|v\|_{H^2(S_T)} && \forall\, v \in H^2(\Omega), && (2.32)\\
\|\nabla(v - I_{h,2} v)\|_{L^2(T)} &\le C h_T \|v\|_{H^2(S_T)} && \forall\, v \in H^2(\Omega), && (2.33)
\end{aligned}$$

where $S_T := \cup_{p \in \mathbb{V}_T} \omega_p$. The definition of $I_{h,2}$ implies it is a positive-preserving operator, i.e., $v \ge 0$ in Ω implies that $I_{h,2} v \ge 0$ in Ω.

Based on $I_{h,1}$ and $I_{h,2}$, we construct a new operator $\Pi_h : H_0^1(\Omega) \longrightarrow V_h$ as

$$\Pi_h v = I_{h,1}\psi + I_{h,2}(v - \psi). \tag{2.34}$$

Since $\psi \in H^2(\Omega)$, the operator is well-defined. Note that if $v \in K$, we have $\Pi_h v \in K_h$. Indeed, for $p \in \mathbb{V}_h \cap \Omega$, since $v \geq \psi$ in Ω, we have

$$\Pi_h v(p) = I_{h,1}\psi(p) + I_{h,2}(v - \psi)(p) \geq \psi(p) + 0 = \psi(p).$$

For $p \in \mathbb{V}_h \cap \partial\Omega$, we have

$$\Pi_h v(p) = I_{h,1}\psi(p) + I_{h,2}(v - \psi)(p) = \psi(p) - \psi(p) = 0 \geq \psi(p)$$

due to the fact that $v = 0$ on $\partial\Omega$, $\psi \leq 0$ on $\partial\Omega$, and the property of $I_{h,2}$.

Lemma 2.5. *For any $v \in H^2(\Omega)$, there exists a positive constant C independent of h such that*

$$\|v - \Pi_h v\|_{L^2(\Omega)} + h|v - \Pi_h v|_{H^1(\Omega)} + h\|v - \Pi_h v\|_h \leq Ch^2(\|v\|_{H^2(\Omega)} + \|\psi\|_{H^2(\Omega)}). \tag{2.35}$$

Proof. By the definition (2.34), we have

$$\begin{aligned} v - \Pi_h v &= \psi + (v - \psi) - I_{h,1}\psi - I_{h,2}(v - \psi) \\ &= (\psi - I_{h,1}\psi) + ((v - \psi) - I_{h,2}(v - \psi)). \end{aligned}$$

Using the local approximation properties for $I_{h,2}$ in (2.32)–(2.33), we have

$$\|(v - \psi) - I_{h,2}(v - \psi)\|_{L^2(\Omega)} + h|(v - \psi) - I_{h,2}(v - \psi)|_{H^1(\Omega)} \leq Ch^2(\|v\|_{H^2(\Omega)} + \|\psi\|_{H^2(\Omega)}). \tag{2.36}$$

Furthermore, we use (2.33) and Lemma 2.3 to obtain

$$\|(v - \psi) - I_{h,2}(v - \psi)\|_h \leq Ch(\|v\|_{H^2(\Omega)} + \|\psi\|_{H^2(\Omega)}). \tag{2.37}$$

Indeed, this can be established in the same way as those in $\|\psi - I_{h,1}\psi\|_h$ (see [31]). Now combining (2.36) with the approximation properties for $I_{h,1}$ in (2.11), we have

$$\begin{aligned} \|v - \Pi_h v\|_{L^2(\Omega)} + h|v - \Pi_h v|_{H^1(\Omega)} &\leq \|\psi - I_{h,1}\psi\|_{L^2(\Omega)} + \|(v - \psi) - I_{h,2}(v - \psi)\|_{L^2(\Omega)} \\ &\quad + |\psi - I_{h,1}\psi|_{H^1(\Omega)} + |(v - \psi) - I_{h,2}(v - \psi)|_{H^1(\Omega)} \\ &\leq Ch^2(\|v\|_{H^2(\Omega)} + \|\psi\|_{H^2(\Omega)}). \end{aligned} \tag{2.38}$$

Similarly, using (2.12) and (2.37), we obtain the estimate

$$\|v - \Pi_h v\|_h \leq Ch(\|v\|_{H^2(\Omega)} + \|\psi\|_{H^2(\Omega)}). \tag{2.39}$$

The final estimate (2.35) follows from (2.38) and (2.39). □

3. An *a Priori* Error Analysis

In this section we establish an error estimate for the fully discrete method proposed in Section 2. We denote $e^n = u(t_n) - u_h^n$ to be error between the exact solution and the discrete solution at the time point $t_n = n\tau$ $(n = 0, 1, 2, ..., N)$. The goal is to estimate the error in the $L^\infty(L^2)$ and $L^2(H^1)$-like norms, i.e.,

$$\max_{1\leq n\leq N} \|e^n\|_{L^2(\Omega)} + \Big(\sum_{n=1}^{N} \tau\|e^n\|_h^2\Big)^{\frac{1}{2}}.$$

To this aim, we decompose the error as $e^n = \eta^n + \xi^n$, where

$$\eta^n = u(t_n) - \Pi_h u(t_n),$$
$$\xi^n = \Pi_h u(t_n) - u_h^n,$$

and Π_h is the interpolation operator that was introduced in Section 2.3. By Lemma 2.5, we have

$$\max_{1\le n\le N} \|\eta^n\|_{L^2(\Omega)} + \Big(\sum_{n=1}^{N} \tau\|\eta^n\|_h^2\Big)^{\frac{1}{2}} \le Ch(\|u\|_{L^\infty(J;H^2(\Omega))} + \|\psi\|_{H^2(\Omega)}). \tag{3.1}$$

Applying the triangle inequality, it suffices to estimate the term $\max_{1\le n\le N} \Big\{\|\xi^n\|_{L^2(\Omega)} + \Big(\sum_{n=1}^{N} \tau\|\xi^n\|_h^2\Big)^{\frac{1}{2}}\Big\}$. We begin with an abstract lemma concerning a single step estimate.

Lemma 3.1. *Let u and u_h^n $(n = 1, 2, ..., N)$ be the solution to* (1.1) *and* (2.8)*, respectively. There holds*

$$\frac{1}{2}\|\xi^n\|_{L^2(\Omega)}^2 + \tau\|\xi^n\|_h^2 \le \frac{1}{2}\|\xi^{n-1}\|_{L^2(\Omega)}^2 + S_{1,n} + S_{2,n} + S_{3,n} + S_{4,n}, \tag{3.2}$$

where

$$\begin{aligned}
S_{1,n} &= -(\eta^n - \eta^{n-1}, \xi^n)_{\mathcal{T}_h},\\
S_{2,n} &= -\tau a_h(\eta^n, \xi^n),\\
S_{3,n} &= \int_{t_{n-1}}^{t_n} \Big[\big(\frac{\partial^+ u}{\partial t}(t_n), \xi^n\big)_{\mathcal{T}_h} + a_h(u(t_n), \xi^n) - (f(t_n), \xi^n)_{\mathcal{T}_h}\Big]\, dt,\\
S_{4,n} &= \int_{t_{n-1}}^{t_n} \Big(\frac{\partial u}{\partial t}(t) - \frac{\partial^+ u}{\partial t}(t_n), \xi^n\Big)_{\mathcal{T}_h}\, dt.
\end{aligned}$$

Proof. Using the definition of the bilinear form $a_h(\cdot,\cdot)$ (see (2.10) and (2.14)) and the discrete formulation (2.8), we have

$$\begin{aligned}
(\partial\xi^n, \xi^n)_{\mathcal{T}_h} + \|\xi^n\|_h^2 &= (\partial\xi^n, \xi^n)_{\mathcal{T}_h} + a_h(\xi^n, \xi^n)\\
&= [(\partial\Pi_h u(t_n), \xi^n)_{\mathcal{T}_h} + a_h(\Pi_h u(t_n), \xi^n)]\\
&\quad - [(\partial u_h^n, \xi^n)_{\mathcal{T}_h} + a_h(u_h^n, \xi^n)]\\
&\le [(\partial\Pi_h u(t_n), \xi^n)_{\mathcal{T}_h} + a_h(\Pi_h u(t_n), \xi^n)] - (f(t_n), \xi^n)_{\mathcal{T}_h}\\
&= [(\partial u(t_n), \xi^n)_{\mathcal{T}_h} + a_h(u(t_n), \xi^n)] - (f(t_n), \xi^n)_{\mathcal{T}_h}\\
&\quad - [(\partial\eta^n, \xi^n)_{\mathcal{T}_h} + a_h(\eta^n, \xi^n)].
\end{aligned} \tag{3.3}$$

We then multiply τ on both sides of (3.3) to get

$$\begin{aligned}
\tau(\partial\xi^n, \xi^n)_{\mathcal{T}_h} + \tau\|\xi^n\|_h^2 \le \tau[(\partial u(t_n), \xi^n)_{\mathcal{T}_h} + a_h(u(t_n), \xi^n)] - \tau(f(t_n), \xi^n)_{\mathcal{T}_h}\\
- \tau[(\partial\eta^n, \xi^n)_{\mathcal{T}_h} + a_h(\eta^n, \xi^n)].
\end{aligned} \tag{3.4}$$

Note that

$$\begin{aligned}
\tau(\partial\xi^n, \xi^n)_{\mathcal{T}_h} &= (\xi^n - \xi^{n-1}, \xi^n)_{\mathcal{T}_h}\\
&= \frac{1}{2}\|\xi^n\|_{L^2(\Omega)}^2 - \frac{1}{2}\|\xi^{n-1}\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\xi^n - \xi^{n-1}\|_{L^2(\Omega)}^2,
\end{aligned} \tag{3.5}$$

and

$$\tau(\partial u(t_n), \xi^n)_{\mathcal{T}_h} = \int_{t_{n-1}}^{t_n} \Big(\frac{\partial u}{\partial t}(t), \xi^n\Big)_{\mathcal{T}_h} dt, \tag{3.6}$$

which is a direct consequence of (2.9). Finally, the estimate (3.2) is obtained by dropping $\frac{1}{2}\|\xi^n - \xi^{n-1}\|^2_{L^2(\Omega)}$ and rearranging the terms. □

In the following lemma, we estimate the remaining terms $S_{1,n}$, $S_{2,n}$, $S_{3,n}$, and $S_{4,n}$ on the right-hand side of (3.2). In particular, in order to estimate $S_{4,n}$, we set $J_n = (t_{n-1}, t_n]$ $(n = 1, 2, ..., N)$ and introduce

$$D_n := \bigcup_{t \in J_n} (\Omega^+(t) \cup \Omega^+(t_n)) \setminus \overline{\Omega^+(t) \cap \Omega^+(t_n)}, \tag{3.7}$$

which measures the change of the non-contact regions in J_n [27].

Lemma 3.2. *Let u and u_h^n $(n = 1, 2, ..., N)$ be the solutions to* (1.1a)–(1.1b) *and* (2.8)*, respectively. There exists $\epsilon > 0$ and a positive constant C independent of h and τ such that*

$$\begin{aligned}
\frac{1}{2}\|\xi^n\|^2_{L^2(\Omega)} + \tau\|\xi^n\|^2_h \le{}& \frac{1}{2}\|\xi^{n-1}\|^2_{L^2(\Omega)} + \frac{C}{\epsilon}h^2\Big\|\frac{\partial u}{\partial t}\Big\|^2_{L^2(J_n;H^1(\Omega))} + C\epsilon\tau\|\xi^n\|^2_h \\
&+ \frac{C}{\epsilon}h^2\tau(\|u(t_n)\|^2_{H^2(\Omega)} + \|\psi\|^2_{H^2(\Omega)}) \\
&+ \frac{C}{\epsilon}\tau(h^2\|\sigma(t_n)\|^2_{L^2(\Omega)} + h^2\|u(t_n)\|^2_{H^2(\Omega)} + h^2\|\psi\|^2_{H^2(\Omega)}) \\
&+ \frac{C}{\epsilon}\Big(\tau^2\Big\|\frac{\partial u}{\partial t}\Big\|^2_{L^2(J_n;H^1(\Omega))} + h^2\tau\|u\|_{L^\infty(J_n;H^2(\Omega))}\Big) \\
&+ \frac{C}{\epsilon}\tau^2\Big\|\frac{\partial f}{\partial t}\Big\|^2_{L^2(J_n;L^2(\Omega))} + CQ_n(p),
\end{aligned} \tag{3.8}$$

where

$$Q_n(p) = \tau\|f + \Delta\psi\|_{L^\infty(J_n;L^\infty(\Omega))}\|\xi^n\|_{L^p(\Omega)}\mathrm{meas}(D_n)^{\frac{1}{q}}, \tag{3.9}$$

with $\frac{1}{p} + \frac{1}{q} = 1$ $(p \ge 1)$, and $\mathrm{meas}(D_n)$ denotes the Lebesgue measure of D_n.

Proof. First, we estimate $S_{1,n}$ on the right-hand side of (3.2). By Young's inequality, we have

$$|S_{1,n}| \le \frac{C}{\epsilon\tau}\|\eta^n - \eta^{n-1}\|^2_{L^2(\Omega)} + \epsilon\tau\|\xi^n\|^2_{L^2(\Omega)} \tag{3.10}$$

for some $\epsilon > 0$. Since ψ is independent of t, $\frac{\partial u}{\partial t} = 0$ on $\partial\Omega$, and the interpolation estimate $I_{h,2}$ in (2.31), we have

$$\begin{aligned}
\|\eta^n - \eta^{n-1}\|^2_{L^2(\Omega)} &= \Big\|\int_{t_{n-1}}^{t_n} \frac{\partial \eta}{\partial t}\, dt\Big\|^2_{L^2(\Omega)} \\
&\le \Big(\int_{t_{n-1}}^{t_n} \Big\|\frac{\partial(u - \Pi_h u)}{\partial t}\Big\|_{L^2(\Omega)} dt\Big)^2 \\
&= \Big(\int_{t_{n-1}}^{t_n} \Big\|\frac{\partial u}{\partial t} - \frac{\partial}{\partial t}(I_{h,1}(\psi) + I_{h,2}(u - \psi))\Big\|_{L^2(\Omega)} dt\Big)^2
\end{aligned} \tag{3.11}$$

$$
\begin{aligned}
&= \Big(\int_{t_{n-1}}^{t_n} \|\frac{\partial u}{\partial t} - I_{h,2}(\frac{\partial u}{\partial t})\|_{L^2(\Omega)}\, dt\Big)^2 \\
&= \Big(\Big(\int_{t_{n-1}}^{t_n} \|\frac{\partial u}{\partial t} - I_{h,2}(\frac{\partial u}{\partial t})\|^2_{L^2(\Omega)}\, dt\Big)^{\frac{1}{2}} \Big(\int_{t_{n-1}}^{t_n} 1^2\, dt\Big)^{\frac{1}{2}} \Big)^2 \\
&\le h^2 \tau \|\frac{\partial u}{\partial t}\|^2_{L^2(J_n;H^1(\Omega))},
\end{aligned}
$$

where we also used the fact that $I_{h,2}$ and $\frac{\partial}{\partial t}$ are interchangeable. Combining (3.10) and (3.11), we obtain

$$
|S_{1,n}| \le \frac{C}{\epsilon} h^2 \|\frac{\partial u}{\partial t}\|^2_{L^2(J_n;H^1(\Omega))} + \epsilon\tau \|\xi^n\|^2_{L^2(\Omega)}. \tag{3.12}
$$

To estimate $S_{2,n}$, we will consider two cases. Whenever $\gamma_{\min} \ge 0$, we apply (2.18) and Young's inequality to obtain

$$
\begin{aligned}
|S_{2,n}| &\le \tau \|\eta^n\|_h \|\xi^n\|_h \\
&\le \frac{C}{\epsilon} \tau \|\eta^n\|_h^2 + \epsilon\tau \|\xi^n\|_h^2.
\end{aligned} \tag{3.13}
$$

Whenever $-C_* \le \gamma_{\min} < 0$, we have by (2.10), (2.19), the Cauchy-Schwarz inequality, Young's inequality, and the trace theorem with scaling that

$$
\begin{aligned}
|S_{2,n}| &\le C\tau (\|\nabla^+_{h,0}\eta^n\|_{L^2(\Omega)} + \|\nabla^-_{h,0}\eta^n\|_{L^2(\Omega)})(\|\nabla^+_{h,0}\xi^n\|_{L^2(\Omega)} + \|\nabla^-_{h,0}\xi^n\|_{L^2(\Omega)}) \\
&\quad + \tau \sum_{e\in\mathcal{E}_h} \frac{|\gamma_e|}{h_e} \|[\![\eta^n]\!]\|_{L^2(e)} \|[\![\xi^n]\!]\|_{L^2(e)} \\
&\le \frac{C}{\epsilon}\tau \Big(\|\eta^n\|_h^2 + \sum_{T\in\mathcal{T}_h} h_T^{-2} \|\eta^n\|^2_{L^2(T)} + \sum_{T\in\mathcal{T}_h} \|\nabla \eta^n\|^2_{L^2(T)} \Big) + \epsilon\tau \|\xi^n\|_h^2.
\end{aligned} \tag{3.14}
$$

By Lemma 2.5, (3.13), and (3.14), we have

$$
|S_{2,n}| \le \frac{C}{\epsilon} h^2 \tau (\|u(t_n)\|^2_{H^2(\Omega)} + \|\psi\|^2_{H^2(\Omega)}) + \epsilon\tau \|\xi^n\|_h^2. \tag{3.15}
$$

Next, we estimate $S_{3,n}$. By Lemma 2.4, we have

$$
S_{3,n} = \tau(\sigma(t_n), \xi^n)_{\mathcal{T}_h} - \tau \langle \{\overline{\nabla}_{h,0} u(t_n) - \nabla u(t_n)\} \cdot \mathbf{n}, [\![\xi^n]\!] \rangle_{\mathcal{E}_h}, \tag{3.16}
$$

where $\sigma(t_n) := \frac{\partial^+ u}{\partial t}(t_n) - \Delta u(t_n) - f(t_n)$. The first term on the right-hand side of (3.16) can be estimated by

$$
\begin{aligned}
\tau(\sigma(t_n), \xi^n)_{\mathcal{T}_h} &= \tau(\sigma(t_n), \Pi_h u(t_n) - u(t_n))_{\mathcal{T}_h} + \tau(\sigma(t_n), u(t_n) - \psi)_{\mathcal{T}_h} \\
&\quad + \tau(\sigma(t_n), \psi - I_{h,1}\psi)_{\mathcal{T}_h} + \tau(\sigma(t_n), I_{h,1}\psi - u_h^n)_{\mathcal{T}_h} \\
&\le \tau(\sigma(t_n), \Pi_h u(t_n) - u(t_n))_{\mathcal{T}_h} + \tau(\sigma(t_n), \psi - I_{h,1}\psi)_{\mathcal{T}_h}, \\
&\le \tau \|\sigma(t_n)\|_{L^2(\Omega)} (\|\Pi_h u(t_n) - u(t_n)\|_{L^2(\Omega)} + \|\psi - I_{h,1}\psi\|_{L^2(\Omega)}) \\
&\le Ch^2\tau \|\sigma(t_n)\|_{L^2(\Omega)} (\|u(t_n)\|_{H^2(\Omega)} + \|\psi\|_{H^2(\Omega)}) \\
&\le Ch^2\tau (\|\sigma(t_n)\|^2_{L^2(\Omega)} + \|u(t_n)\|^2_{H^2(\Omega)} + \|\psi\|^2_{H^2(\Omega)}),
\end{aligned} \tag{3.17}
$$

where we used the complementarity form (1.13)–(1.14) at $t = t_n$, (2.11), and Lemma 2.5. In the case of $\gamma_{\min} > 0$, we have

$$\begin{aligned}
&\tau\langle\{\overline{\nabla}_{h,0}u(t_n) - \nabla u(t_n)\}\cdot\mathbf{n}, [\![\xi^n]\!]\rangle_{\mathcal{E}_h} \qquad (3.18)\\
&\le \tau\Big(\sum_{e\in\mathcal{E}_h}\frac{h_e}{\gamma_e}\|\overline{\nabla}_{h,0}u(t_n) - \nabla u(t_n)\|^2_{L^2(e)}\Big)^{\frac{1}{2}}\Big(\sum_{e\in\mathcal{E}_h}\frac{\gamma_e}{h_e}\|[\![\xi^n]\!]\|^2_{L^2(e)}\Big)^{\frac{1}{2}}\\
&\le \frac{C}{\epsilon}\tau\Big(\sum_{T\in\mathcal{T}_h}\|\overline{\nabla}_{h,0}u(t_n) - \nabla u(t_n)\|^2_{L^2(T)} + h_T^2\|\nabla(\overline{\nabla}_{h,0}u(t_n) - \nabla u(t_n))\|^2_{L^2(T)}\Big) + \epsilon\tau\|\xi^n\|_h^2\\
&\le \frac{C}{\epsilon}h^2\tau|u(t_n)|^2_{H^2(\Omega)} + \epsilon\tau\|\xi^n\|_h^2
\end{aligned}$$

by Lemma 2.3 and Young's inequality. A similar result can be obtained whenever $-C_* \le \gamma_{\min} \le 0$. We combine (3.16)–(3.18) to conclude that

$$|S_{3,n}| \le \frac{C}{\epsilon}h^2\tau(\|\sigma(t_n)\|^2_{L^2(\Omega)} + \|u(t_n)\|^2_{H^2(\Omega)} + \|\psi\|^2_{H^2(\Omega)}) + \epsilon\tau\|\xi^n\|_h^2. \qquad (3.19)$$

For the last term $S_{4,n}$, we rewrite it as

$$\begin{aligned}
S_{4,n} = &\int_{t_{n-1}}^{t_n}\int_{\Omega^+(t)}\frac{\partial^+ u}{\partial t}(t)\,\xi^n\,dx\,dt + \int_{t_{n-1}}^{t_n}\int_{\Omega^0(t)}\frac{\partial^+ u}{\partial t}(t)\,\xi^n\,dx\,dt \qquad (3.20)\\
&- \int_{t_{n-1}}^{t_n}\int_{\Omega^+(t_n)}\frac{\partial^+ u}{\partial t}(t_n)\,\xi^n\,dx\,dt - \int_{t_{n-1}}^{t_n}\int_{\Omega^0(t_n)}\frac{\partial^+ u}{\partial t}(t_n)\,\xi^n\,dx\,dt,
\end{aligned}$$

based on the decomposition $\Omega = \Omega^+(t)\cup\Omega^0(t)$ for any $t\in J$. By (1.8)–(1.9), we can further rewrite $S_{4,n}$ as

$$\begin{aligned}
S_{4,n} = &\int_{t_{n-1}}^{t_n}\int_{\Omega^+(t)}\Delta(u(t)-\psi)\xi^n\,dxdt + \int_{t_{n-1}}^{t_n}\int_{\Omega^+(t)}(f(t)+\Delta\psi)\xi^n\,dxdt \qquad (3.21)\\
&+ \int_{t_{n-1}}^{t_n}\int_{\Omega^0(t)}\max\{f(t)+\Delta\psi, 0\}\xi^n\,dxdt - \int_{t_{n-1}}^{t_n}\int_{\Omega^+(t_n)}\Delta(u(t_n)-\psi)\xi^n\,dxdt\\
&- \int_{t_{n-1}}^{t_n}\int_{\Omega^+(t_n)}(f(t_n)+\Delta\psi)\xi^n\,dxdt - \int_{t_{n-1}}^{t_n}\int_{\Omega^0(t)}\max\{f(t_n)+\Delta\psi, 0\}\xi^n\,dxdt\\
:= &\,S_{4,n}^{(1)} + S_{4,n}^{(2)} + S_{4,n}^{(3)},
\end{aligned}$$

where

$$\begin{aligned}
S_{4,n}^{(1)} &= \int_{t_{n-1}}^{t_n}\int_{\Omega^+(t)}\Delta(u(t)-\psi)\xi^n\,dxdt - \int_{t_{n-1}}^{t_n}\int_{\Omega^+(t_n)}\Delta(u(t_n)-\psi)\xi^n\,dxdt,\\
S_{4,n}^{(2)} &= \int_{t_{n-1}}^{t_n}\int_{\Omega\setminus D_n}(\tilde f(x,t) - \tilde f(x,t_n))\xi^n\,dxdt,\\
S_{4,n}^{(3)} &= \int_{t_{n-1}}^{t_n}\int_{D_n}(\tilde f(x,t) - \tilde f(x,t_n))\xi^n\,dxdt,
\end{aligned}$$

and $\tilde f(\cdot,\cdot)$ is given by

$$\tilde f(x,t) = \begin{cases} f(x,t)+\Delta\psi(x), & x\in\Omega^+(t),\\ \max(f(x,t)+\Delta\psi(x), 0), & x\in\Omega^0(t). \end{cases} \qquad (3.22)$$

In the following, we will estimate $S_{4,n}^{(1)}$, $S_{4,n}^{(2)}$, and $S_{4,n}^{(3)}$ separately. Since $u(t)=\psi$ on $\Omega^0(t)$, we rewrite $S_{4,n}^{(1)}$ as

$$\begin{aligned} S_{4,n}^{(1)} &= \int_{t_{n-1}}^{t_n}\int_\Omega \Delta(u(t)-\psi)\xi^n\,dxdt - \int_{t_{n-1}}^{t_n}\int_\Omega \Delta(u(t_n)-\psi)\xi^n\,dxdt \\ &= \int_{t_{n-1}}^{t_n}\int_\Omega \Delta(u(t)-u(t_n))\xi^n\,dxdt \\ &= -\int_{t_{n-1}}^{t_n}(\nabla(u(t)-u(t_n)),\nabla\xi^n)_{\mathcal{T}_h}dt + \int_{t_{n-1}}^{t_n}\langle \partial_n u(t)-\partial_n u(t_n), [\![\xi^n]\!]\rangle_{\mathcal{E}_h}\,dt, \end{aligned} \tag{3.23}$$

where we applied integration by parts. By the Cauchy-Schwarz inequality, Young's inequality, and Lemma 2.2, we have

$$\begin{aligned} -\int_{t_{n-1}}^{t_n}(\nabla(u(t)-u(t_n)),\nabla\xi^n)_{\mathcal{T}_h}dt &= \int_{t_{n-1}}^{t_n}\Big(\int_t^{t_n}\nabla\frac{\partial u}{\partial t}(z)\,dz,\nabla\xi^n\Big)_{\mathcal{T}_h}dt \\ &\le \int_{t_{n-1}}^{t_n}\Big\|\int_t^{t_n}\nabla\frac{\partial u}{\partial t}(z)\,dz\Big\|_{L^2(\mathcal{T}_h)}\|\nabla\xi^n\|_{L^2(\mathcal{T}_h)}\,dt \\ &\le \int_{t_{n-1}}^{t_n}\int_{t_{n-1}}^{t_n}\big\|\nabla\frac{\partial u}{\partial t}(z)\big\|_{L^2(\mathcal{T}_h)}\,dz\|\nabla\xi^n\|_{L^2(\mathcal{T}_h)}\,dt \\ &\le \int_{t_{n-1}}^{t_n}\big\|\frac{\partial u}{\partial t}\big\|_{L^2(J_n;H^1(\Omega))}\tau^{\frac{1}{2}}\|\nabla\xi^n\|_{L^2(\mathcal{T}_h)}\,dt \\ &\le \frac{C}{\epsilon}\tau^2\big\|\frac{\partial u}{\partial t}\big\|^2_{L^2(J_n;H^1(\Omega))}+\epsilon\tau\|\nabla\xi^n\|_{L^2(\mathcal{T}_h)} \\ &\le \frac{C}{\epsilon}\tau^2\big\|\frac{\partial u}{\partial t}\big\|^2_{L^2(J_n;H^1(\Omega))}+C\epsilon\tau\|\xi^n\|_h. \end{aligned} \tag{3.24}$$

Next, we estimate

$$\begin{aligned} &\int_{t_{n-1}}^{t_n}\langle \partial_n u(t)-\partial_n u(t_n), [\![\xi^n]\!]\rangle_{\mathcal{E}_h}\,dt \\ \le&\frac{C}{\epsilon}\int_{t_{n-1}}^{t_n}\|h_e^{\frac{1}{2}}\partial_n(u(t)-u(t_n))\|^2_{L^2(\mathcal{E}_h)}+\epsilon\sum_{e\in\mathcal{E}_h}\frac{1}{h_e}\|[\![\xi^n]\!]\|^2_{L^2(e)}\,dt. \end{aligned} \tag{3.25}$$

Using the trace theorem with scaling, we have

$$\begin{aligned} \|h_e^{\frac{1}{2}}\partial_n(u(t)-u(t_n))\|^2_{L^2(\mathcal{E}_h)} &\le C\sum_{T\in\mathcal{T}_h}\|\nabla(u(t)-u(t_n))\|^2_{L^2(T)} \\ &\quad + C\sum_{T\in\mathcal{T}_h}h_T^2(\|u(t)\|^2_{H^2(T)}+\|u(t_n)\|^2_{H^2(T)}). \end{aligned} \tag{3.26}$$

Moreover, we have

$$\sum_{T\in\mathcal{T}_h}\|\nabla(u(t)-u(t_n))\|^2_{L^2(T)} \le \sum_{T\in\mathcal{T}_h}\int_T\Big(\int_{t_{n-1}}^{t_n}\big|\nabla\frac{\partial u}{\partial t}(\sigma)\big|\,d\sigma\Big)^2dx \tag{3.27}$$

$$
\begin{aligned}
&\le \tau \sum_{T\in\mathcal{T}_h} \int_{t_{n-1}}^{t_n} \int_T \big|\nabla \frac{\partial u}{\partial t}(\sigma)\big|^2 \, dx\, d\sigma \\
&\le \tau \int_{t_{n-1}}^{t_n} \sum_{T\in\mathcal{T}_h} \big|\frac{\partial u}{\partial t}\big|_{H^1(T)}^2 \, d\sigma \\
&\le \tau \big\|\frac{\partial u}{\partial t}\big\|_{L^2(J_n;H^1(\Omega))}^2.
\end{aligned}
$$

From (3.26)–(3.27), we have

$$
\|h_e^{\frac{1}{2}} \partial_n (u(t) - u(t_n))\|_{L^2(\mathcal{E}_h)}^2 \le C(\tau \big\|\frac{\partial u}{\partial t}\big\|_{L^2(J_n;H^1(\Omega))}^2 + h^2 \|u\|_{L^\infty(J_n;H^2(\Omega))}). \tag{3.28}
$$

Then it follows from (3.25) and (3.28) that

$$
\begin{aligned}
\int_{t_{n-1}}^{t_n} \langle \partial_n u(t) - \partial_n u(t_n), [\![\xi^n]\!]\rangle_{\mathcal{E}_h} \, dt \le \frac{C}{\epsilon}\Big(\tau^2 \big\|\frac{\partial u}{\partial t}\big\|_{L^2(J_n;H^1(\Omega))}^2 + h^2\tau \|u\|_{L^\infty(J_n;H^2(\Omega))}\Big) \\
+ \epsilon\tau \sum_{e\in\mathcal{E}_h} \frac{1}{h_e} \|[\![\xi^n]\!]\|_{L^2(e)}^2.
\end{aligned} \tag{3.29}
$$

Finally, we combine (3.23), (3.24), (3.29), and (2.16)–(2.17) to get

$$
S_{4,n}^{(1)} \le \frac{C}{\epsilon}\Big(\tau^2 \big\|\frac{\partial u}{\partial t}\big\|_{L^2(J_n;H^1(\Omega))}^2 + h^2\tau \|u\|_{L^\infty(J_n;H^2(\Omega))}\Big) + C\epsilon\tau \|\xi^n\|_h^2. \tag{3.30}
$$

Now we focus on the estimation of $S_{4,n}^{(2)}$. For any $x \in \Omega \setminus D_n$, it follows from the definition of D_n that either $x \in \Omega^+(t_n) \cap \Omega^+(t)$ or $x \in \Omega^0(t_n) \cap \Omega^0(t)$ for $t \in J_n$. Whenever $x \in \Omega^+(t_n) \cap \Omega^+(t)$, we have

$$
\begin{aligned}
|\tilde{f}(x,t) - \tilde{f}(x,t_n)| &= |f(x,t) + \Delta\psi(x) - f(x,t_n) - \Delta\psi(x)| \\
&= |f(x,t) - f(x,t_n)|.
\end{aligned} \tag{3.31}
$$

Whenever $x \in \Omega^0(t_n) \cap \Omega^0(t)$, we have

$$
\begin{aligned}
|\tilde{f}(x,t) - \tilde{f}(x,t_n)| &= |\max\{f(x,t) + \Delta\psi(x), 0\} - \max\{f(x,t_n) + \Delta(x)\psi, 0\}| \\
&\le |f(x,t) - f(x,t_n)|.
\end{aligned} \tag{3.32}
$$

Combining these two estimates, we conclude that

$$
|\tilde{f}(x,t) - \tilde{f}(x,t_n)| \le |f(x,t) - f(x,t_n)| \qquad \forall\, x \in \Omega \setminus D_n,\ t \in J_n, \tag{3.33}
$$

and thus

$$
\begin{aligned}
S_{4,n}^{(2)} &\le \int_{t_{n-1}}^{t_n} \int_{\Omega\setminus D_n} |f(x,t) - f(x,t_n)|\, |\xi^n| \, dxdt \\
&\le \int_{t_{n-1}}^{t_n} \frac{C}{\epsilon}\tau \big\|\frac{\partial f}{\partial t}\big\|_{L^2(J_n;L^2(\Omega))}^2 + \epsilon \|\xi^n\|_{L^2(\Omega)}^2 \, dt \\
&\le \frac{C}{\epsilon}\tau^2 \big\|\frac{\partial f}{\partial t}\big\|_{L^2(J_n;L^2(\Omega))}^2 + \epsilon\tau \|\xi^n\|_{L^2(\Omega)}^2 \\
&\le \frac{C}{\epsilon}\tau^2 \big\|\frac{\partial f}{\partial t}\big\|_{L^2(J_n;L^2(\Omega))}^2 + C\epsilon\tau \|\xi^n\|_h^2,
\end{aligned} \tag{3.34}
$$

where in the second inequality we applied the similar technique as used in (3.27) and in the last equality we used Lemma 2.2.

To estimate $S_{4,n}^{(3)}$, we apply the definitions of D_n and $\tilde{f}(\cdot,\cdot)$ and the Hölder inequality with $\frac{1}{p}+\frac{1}{q}=1$ for $p\in[1,\infty)$ to obtain

$$\begin{aligned} S_{4,n}^{(3)} &= \int_{t_{n-1}}^{t_n}\int_{D_n}(\tilde{f}(x,t)-\tilde{f}(x,t_n))\xi^n\,dxdt \\ &\le C\tau\|f+\Delta\psi\|_{L^\infty(J_n;L^\infty(\Omega))}\int_{D_n}|\xi^n|\,dx \\ &\le C\tau\|f+\Delta\psi\|_{L^\infty(J_n;L^\infty(\Omega))}\|\xi^n\|_{L^p(\Omega)}\text{meas}(D_n)^{1/q}. \end{aligned} \tag{3.35}$$

From (3.21), (3.30), (3.34), and (3.35), we then have

$$\begin{aligned} |S_{4,n}| \le{}& \frac{C}{\epsilon}\Big(\tau^2\|\frac{\partial u}{\partial t}\|^2_{L^2(J_n;H^1(\Omega))}+h^2\tau\|u\|_{L^\infty(J_n;H^2(\Omega))}\Big)+C\epsilon\tau\|\xi^n\|_h^2 \\ &+\frac{C}{\epsilon}\tau^2\|\frac{\partial f}{\partial t}\|^2_{L^2(J_n;L^2(\Omega))}+Q_n(p), \end{aligned} \tag{3.36}$$

where $Q_n(p)$ is given in (3.9).

The final estimate (3.8) follows from Lemma 3.1, (3.12), (3.15), (3.19), and (3.36). □

From Lemma 3.2, it is key to estimate $Q_n(p)$. To this aim, we assume there exists a positive constant C independent of N such that

$$\sum_{n=1}^{N}\text{meas}(D_n)\le C, \tag{3.37}$$

which indicates the contact and non-contact sets do not change too frequently [27]. By rearranging the terms we could assume $\text{meas}(D_n)\le\frac{C}{n}$ $(n=1,2,...,N)$ based on (3.37). In particular, such a condition holds whenever

$$\text{meas}(D_n)\le\frac{C}{N}\qquad\forall\, n=1,2,...,N. \tag{3.38}$$

Theorem 3.3. *Let $e^n=u(t_n)-u_h^n$ $(n=0,1,2,...,N)$, where u (resp., u_h^n) denotes the solution to* (1.1) *(resp.,* (2.8)*). Under the assumption* (3.37)*, there exists a positive constant C independent of h and τ such that*

$$\max_{1\le n\le N}\|e^n\|_{L^2(\Omega)}+\Big(\sum_{n=1}^{N}\tau\|e^n\|_h^2\Big)^{\frac{1}{2}}\le C\big[h+\tau^{\frac{3}{4}}(\log\tau^{-1})^{\frac{1}{4}}\big]. \tag{3.39}$$

In addition, under the assumption (3.38)*, there exists a positive constant C independent of h and τ such that*

$$\max_{1\le n\le N}\|e^n\|_{L^2(\Omega)}+\Big(\sum_{n=1}^{N}\tau\|e^n\|_h^2\Big)^{\frac{1}{2}}\le C\big[h+\tau(\log\tau^{-1})^{\frac{1}{2}}\big]. \tag{3.40}$$

Proof. Summing over $n=1,2,...,N$ in (3.8), we have

$$\max_{1\le n\le N}\|\xi^n\|^2_{L^2(\Omega)}+\sum_{n=1}^{N}\tau\|\xi^n\|_h^2\le C\|\xi^0\|^2_{L^2(\Omega)}+C\epsilon\sum_{n=1}^{N}\tau\|\xi^n\|_h^2 \tag{3.41}$$

$$+ C_\epsilon h^2\Big(\|\frac{\partial u}{\partial t}\|^2_{L^2(J;H^1(\Omega))} + \|u\|_{L^\infty(J;H^2(\Omega))} + \|\psi\|^2_{H^2(\Omega)} + \|f\|_{L^\infty(J;H^1(\Omega))}\Big)$$
$$+ C_\epsilon \tau^2\Big(\|\frac{\partial u}{\partial t}\|^2_{L^2(J;H^1(\Omega))} + \|\frac{\partial f}{\partial t}\|^2_{L^2(J;L^2(\Omega))}\Big)$$
$$+ C\sum_{n=1}^N Q_n(p_n),$$

where $C_\epsilon := \frac{C}{\epsilon}$. Choosing $\epsilon > 0$ sufficiently small, the second term $C\epsilon\sum_{n=1}^N \|\xi^n\|_h^2\tau$ on the right-hand side of (3.41) can be absorbed to the left-hand side. Note that in Lemma 3.2, the choice of p in (3.8) can vary for different n. Therefore we used the notation p_n in (3.41).

Next, we will focus on the estimation of $\sum_{n=1}^N Q_n(p_n)$. We first assume (3.37) holds. By the Sobolev Imbedding Theorem [28]

$$\frac{1}{\sqrt{p}}\|v\|_{L^p(\Omega)} \le C(\|\nabla v\|_{L^2(\mathcal{T}_h)} + \|v\|_{L^2(\Omega)}) \qquad \forall\, v \in W^{1,p}(\mathcal{T}_h), \tag{3.42}$$

Lemma 2.1 and Lemma 2.2, we have

$$\|\xi^n\|_{L^p(\Omega)} \le C\sqrt{p}\|\xi^n\|_h \qquad \forall\, p \ge 1. \tag{3.43}$$

Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be a disjoint partition of $\{1, 2, ..., N\}$. We choose $p_n = 2$ for $n \in \mathcal{N}_1$ and $p_n = p$ with $p \in [1, \infty)$ for $n \in \mathcal{N}_2$, and we apply (3.43), the Cauchy-Schwarz inequality, and Young's inequality to obtain

$$\begin{aligned}
\sum_{n=1}^N Q_n(p_n) &= \sum_{n\in\mathcal{N}_1} Q_n(2) + \sum_{n\in\mathcal{N}_2} Q_n(p) \\
&\le C\|f+\Delta\psi\|_{L^\infty(J;L^\infty(\Omega))}\Big(\tau\sum_{n\in\mathcal{N}_1}\|\xi^n\|_{L^2(\Omega)}\mathrm{meas}(D_n)^{\frac12} + \tau\sum_{n\in\mathcal{N}_2}\|\xi^n\|_{L^p(\Omega)}\mathrm{meas}(D_n)^{\frac1q}\Big) \\
&\le C\|f+\Delta\psi\|_{L^\infty(J;L^\infty(\Omega))}\Big(\tau\sum_{n\in\mathcal{N}_1}\|\xi^n\|_{L^2(\Omega)}\mathrm{meas}(D_n)^{\frac12} + \tau\sum_{n\in\mathcal{N}_2}\sqrt{p}\|\xi^n\|_h\mathrm{meas}(D_n)^{\frac1q}\Big) \\
&\le C\|f+\Delta\psi\|_{L^\infty(J;L^\infty(\Omega))}\Big(\epsilon\max_{1\le n\le N}\|\xi^n\|^2_{L^2(\Omega)} + C_\epsilon\tau^2\big(\sum_{n\in\mathcal{N}_1}\mathrm{meas}(D_n)^{\frac12}\big)^2 \\
&\quad + \epsilon\tau\sum_{n=1}^N\|\xi^n\|_h^2 + C_\epsilon p\tau\sum_{n\in\mathcal{N}_2}\mathrm{meas}(D_n)^{\frac2q}\Big),
\end{aligned} \tag{3.44}$$

where q satisfies $1/p + 1/q = 1$.

By Lemma 3 in [27], there exist a positive constant C independent of τ and the choices of $\mathcal{N}_1$, $\mathcal{N}_2$, p such that

$$\tau^2\Big(\sum_{n\in\mathcal{N}_1}\mathrm{meas}(D_n)^{\frac12}\Big)^2 + p\tau\sum_{n\in\mathcal{N}_2}\mathrm{meas}(D_n)^{\frac2q} \le C\tau^{\frac32}(\log\tau^{-1})^{\frac12}. \tag{3.45}$$

With sufficiently small $\epsilon > 0$, we then obtain the error estimate

$$\max_{1\le n\le N}\|\xi^n\|^2_{L^2(\Omega)} + \sum_{n=1}^N\tau\|\xi^n\|_h^2 \le C\|\xi^0\|^2_{L^2(\Omega)} \tag{3.46}$$

$$+ C_\epsilon h^2\Big(\|\frac{\partial u}{\partial t}\|^2_{L^2(J;H^1(\Omega))} + \|u\|_{L^\infty(J;H^2(\Omega)} + \|\psi\|^2_{H^2(\Omega)} + \|f\|_{L^\infty(J;H^1(\Omega)}\Big)$$
$$+ C_\epsilon \tau^2\Big(\|\frac{\partial u}{\partial t}\|^2_{L^2(J;H^1(\Omega))} + \|\frac{\partial f}{\partial t}\|^2_{L^2(J;L^2(\Omega))}\Big)$$
$$+ C\tau^{\frac{3}{2}}(\log \tau^{-1})^{\frac{1}{2}}(\|f\|_{L^\infty(J;L^\infty(\Omega))} + \|\psi\|_{W^{2,\infty}(\Omega)}).$$

The final estimate (3.39) follows from (3.46), Lemma 2.5, (2.11), and (3.1).

Now we assume (3.38) holds. In (3.44), we choose $\mathcal{N}_1 = \{1, 2, ..., M\}$ and $\mathcal{N}_2 = \{M+1, ..., N\}$, where M equals to the integer part of $(N \log N)^{\frac{1}{2}}$, and $p = \log N$. Note that $\frac{C}{N} = O(\tau)$. A direct calculation shows that

$$\sum_{n=1}^{N} Q_n(p_n) = \sum_{n\in\mathcal{N}_1} Q_n(2) + \sum_{n\in\mathcal{N}_2} Q_n(p) \tag{3.47}$$
$$\leq C\|f + \Delta\psi\|_{L^\infty(J;L^\infty(\Omega))}\Big(\epsilon \max_{n=1}^{N} \|\xi^n\|^2_{L^2(\Omega)} + \epsilon\sum_{n=1}^{N} \tau\|\xi^n\|^2_h + \tau^2 \log \tau^{-1}\Big).$$

We then obtain the estimate (3.40) under the assumption (3.38). □

Remark 3.4. The rate of convergence in Theorem 3.3 is sharp with respect to the spatial mesh-size h. However due to the low regularity of $\frac{\partial u}{\partial t}$, the optimal rate with respect to the temporal mesh-size τ is not directly available. Under the assumption (3.37), we showed a convergence rate close to $\frac{3}{4}$, which agrees with the results in [27, 34]. Indeed, under the assumption

$$\sum_{n=1}^{N} (\text{meas}(D_n))^{\frac{1}{2}} \leq C, \tag{3.48}$$

we can derive the optimal convergence $O(h + \tau)$ by slightly revising (3.44) in the proof of Theorem 3.3. However, such an assumption may not hold in practice. Under a more reasonable assumption (3.38), we recover a convergence rate close to 1 in τ.

4. Numerical Experiments

In this section, we report several numerical tests to illustrate the performance of the fully discrete methods (2.8) with $\gamma_e := \gamma$ for all $e \in \mathcal{E}_h$ where $\gamma \in \{-1, 0, 1\}$. At each time step, we solve the discrete problem by the primal-dual active set strategy [26]. To verify the rates in Theorem 3.3 for Example 1 and Example 2, we measure the relative total error (RTE)

$$\text{RTE} = \frac{\max\limits_{1\leq n\leq N} \|u(t_n) - u_h^n\|_{L^2(\Omega)} + \left(\sum\limits_{n=1}^{N} \tau\|u(t_n) - u_h^n\|^2_h\right)^{\frac{1}{2}}}{\max\limits_{1\leq n\leq N} \|u(t_n)\|_{L^2(\Omega)} + \left(\sum\limits_{n=1}^{N} \tau\|\nabla u(t_n)\|^2_{L^2(\Omega)}\right)^{\frac{1}{2}}}, \tag{4.1}$$

where N is the number of time steps on the interval $J = [0, T_F]$.

4.1. Example 1

In this example, we consider a benchmark where the contact set is the oscillating moving circle [35, 34]. Let $\Omega = [-1,1]^2$, $J = [0, 0.25]$, and $\psi = 0$. With $r_1 = \frac{1}{3}$ and $\omega = 4$, define

$$r_0(t) = \frac{1}{3} + 0.3\sin(4\omega\pi t),$$
$$c(t) = r_1(\cos(\omega\pi t), \sin(\omega\pi t)),$$

and the sets

$$\Omega^0(t) = \{\|x - c(t)\|_2 \le r_0(t)\},$$
$$\Omega^+(t) = \{\|x - c(t)\|_2 > r_0(t)\}.$$

Define the load function as

$$f(x,t) = \begin{cases} 4\left[r_0^2(t) - 2\|x-c(t)\|_2^2 - \frac{1}{2}(\|x-c(t)\|_2^2 - r_0^2(t))((x-c(t))c'(t) + r_0(t)r_0'(t))\right], & x \in \Omega^+(t), \\ -4r_0^2\left(1 - \|x-c(t)\|_2^2 + r_0^2(t)\right), & x \in \Omega^0(t). \end{cases}$$

The exact solution is

$$u(x,t) = \begin{cases} \frac{1}{2}\left(\|x-c(t)\|_2^2 - r_0^2(t)\right)^2, & x \in \Omega^+(t), \\ 0, & x \in \Omega^0(t). \end{cases} \tag{4.2}$$

Note that this example has nonhomogeneous boundary data $g(t)$ determined by (4.2). As such, the fully discrete scheme (2.8) is modified as

$$(\partial u_h^n, v_h - u_h^n) + a_h(u_h^n, v_h - u_h^n) \ge F_h^n(v_h - u_h^n) \qquad \forall\, v_h \in K_h, \tag{4.3}$$

where

$$F_h^n(v) := (f(t_n), v)_{\mathcal{T}_h} + \left\langle \frac{\gamma_e}{h_e} g(t_n), v \right\rangle_{\mathcal{E}_h^B} - \left\langle g(t_n), \overline{\nabla}_{h,0} v \cdot \boldsymbol{n} \right\rangle_{\mathcal{E}_h^B} \qquad \forall\, v \in V_h.$$

In Table 4.1, we display the relative errors (RTE) and rates of convergence. To determine a rate for h, we fix $\tau = 5 \times 10^{-5}$ and vary $h = \{1/2, 1/4, ..., 1/32\}$. From Table 4.1, we observe the $O(h)$ convergence for all three penalty values as predicted in Theorem 3.3. To determine a rate for τ, we reduce the mesh-size h by a factor of $\frac{1}{2}$ and we doubled the number of time steps. Furthermore, the denominator in (4.1) is calculated on the finest temporal mesh. Numerical results indicate order 1 convergence with respect to τ. Indeed, from the definition of the exact solution (4.2), we can verify that the set D_n satisfies the assumption (3.38). Therefore our numerical results agree with the error estimate in (3.40) of Theorem 3.3.

4.2. Example 2

In this numerical experiment we consider a homogeneous problem with a non-zero obstacle, which is a slight modification from [37]. Let $\Omega = [0,1]^2$, $J = [0,1]$, $\alpha(t) = \frac{1}{2} + \frac{1}{4}\sin(2\pi t)$, and define the obstacle function to be $\psi(x) = x_1(1-x_1)x_2(1-x_2)$. We define the exact solution for the problem (1.1a)–(1.1b)

$$u(x,t) = \begin{cases} 100x_1(x_1 - \alpha(t))^2 x_2(1-x_2) + 2x_1(1-x_1) + x_2(1-x_2), & x_1 < \alpha(t), \\ 2x_1(1-x_1) + x_2(1-x_2), & x_1 \ge \alpha(t) \end{cases}$$

Table 4.1: Errors and convergence rates with respect to h and τ for Example 1

	h rate			τ rate		
	h	RTE	Rate	N	RTE	Rate
$\gamma=-1$	1	2.6157e-01	—	5	1.4707e+00	—
	1/2	1.2528e-01	1.0621	10	7.3719e-01	0.9964
	1/4	6.1189e-02	1.0338	20	3.7174e-01	0.9877
	1/8	3.0775e-02	0.9915	40	1.9030e-01	0.9660
	1/16	1.5542e-02	0.9855	80	9.6538e-02	0.9791
	1/32	7.8714e-03	0.9815	160	4.8663e-02	0.9883
$\gamma=0$	1	2.6754e-01	—	5	1.4781e+00	—
	1/2	1.2948e-01	1.0471	10	7.4718e-01	0.9842
	1/4	6.3395e-02	1.0303	20	3.7762e-01	0.9845
	1/8	3.1755e-02	0.9974	40	1.9307e-01	0.9678
	1/16	1.5981e-02	0.9907	80	9.7812e-02	0.9811
	1/32	8.0731e-03	0.9851	160	4.9259e-02	0.9896
$\gamma=1$	1	2.7322e-01	—	5	1.4842e+00	—
	1/2	1.3281e-01	1.0407	10	7.5462e-01	0.9759
	1/4	6.5049e-02	1.0298	20	3.8189e-01	0.9826
	1/8	3.2479e-02	1.0020	40	1.9511e-01	0.9689
	1/16	1.6304e-02	0.9942	80	9.8756e-02	0.9823
	1/32	8.2231e-03	0.9875	160	4.9705e-02	0.9905

such that the load function is given by

$$f(x,t)=\begin{cases}\partial_t u-\Delta u, & x_1<\alpha(t),\\ 0, & x_2\geq\alpha(t).\end{cases}$$

In Table 4.2, we display the relative errors (RTE) and rates of convergence. We fix $\tau=5\times10^{-5}$ and vary $h=\{1/2,1/4,...,1/64\}$ to calculate the rate in h. We observe the $O(h)$ convergence for all three penalty values, which agrees with Theorem 3.3. A direct calculation demonstrated the assumption (3.38) holds. The error estimate in (3.40) of Theorem 3.3 predicted the order 1 convergence with respect to τ, which is also confirmed from Table 4.2.

4.3. Example 3

In this experiment, we consider a problem without a known exact solution. Let $\Omega=[0,1]^2$, $J=[0,1]$, $\psi(x)=\cos(3\pi x_1)+\cos(3\pi x_2)$, and $f=-5$. We also take $u(x,t)=\psi(x)$ for $t\in J$ on $\partial\Omega$. To estimate the rate of convergence with respect to h, we fix $\tau=5\times10^{-4}$ and vary $h=\{1/2,1/4,...,1/64\}$. Since the exact solution is unknown, we replace $u(t_n)$ with $u^n_{h/2}$ in (4.1) to estimate the relative total error:

$$\mathrm{RTE}=\frac{\max\limits_{1\leq n\leq N}\|u^n_{h/2}-u^n_h\|_{L^2(\Omega)}+\left(\sum\limits_{n=1}^N\tau\|u^n_{h/2}-u^n_h\|^2_{h/2}\right)^{\frac12}}{\max\limits_{1\leq n\leq N}\|u^n_{h/2}\|_{L^2(\Omega)}+\left(\sum\limits_{n=1}^N\tau\|\nabla u^n_{h/2}\|^2_{L^2(\mathcal{T}_h)}\right)^{\frac12}}. \tag{4.4}$$

Table 4.2: Errors and convergence rates with respect to h and τ for Example 2

	h rate			τ rate		
	h	RTE	Rate	N	RTE	Rate
$\gamma=-1$	1/2	3.1968e-01	—	5	3.3185e-01	—
	1/4	2.0704e-01	0.6267	10	2.2153e-01	0.5830
	1/8	1.0875e-01	0.9289	20	1.1520e-01	0.9434
	1/16	5.4894e-02	0.9862	40	5.9319e-02	0.9576
	1/32	2.7533e-02	0.9955	80	2.9988e-02	0.9841
	1/64	1.3777e-02	0.9989	160	1.5077e-02	0.9920
$\gamma=0$	1/2	3.5288e-01	—	5	3.6516e-01	—
	1/4	2.2616e-01	0.6418	10	2.4108e-01	0.5990
	1/8	1.1824e-01	0.9356	20	1.2481e-01	0.9498
	1/16	5.9684e-02	0.9864	40	6.4082e-02	0.9618
	1/32	2.9937e-02	0.9954	80	3.2362e-02	0.9856
	1/64	1.4981e-02	0.9988	160	1.6262e-02	0.9928
$\gamma=1$	1/2	3.7934e-01	—	5	3.9179e-01	—
	1/4	2.4017e-01	0.6594	10	2.5542e-01	0.6172
	1/8	1.2508e-01	0.9412	20	1.3173e-01	0.9553
	1/16	6.3104e-02	0.9871	40	6.7485e-02	0.9649
	1/32	3.1646e-02	0.9957	80	3.4054e-02	0.9867
	1/64	1.5837e-02	0.9987	160	1.7106e-02	0.9933

In order to estimate the rate of convergence with respect to τ, we reduce the mesh-size h by a factor of $\frac{1}{2}$ and doubled the number of time steps. For convenience, we denote $u^n_{\tau,h}$ by the numerical solution at time point t_n with the time step size τ and the spatial mesh size h. Then we use $u_{\tau/2,h/2}(t_n)$ to approximate $u(t_n)$ and calculate the relative total error in the following:

$$\mathrm{RTE} = \frac{\max\limits_{1\le n\le N} \|u_{\tau/2,h/2}(t_n) - u^n_{\tau,h}\|_{L^2(\Omega)} + \left(\sum\limits_{n=1}^{N} \tau \|u_{\tau/2,h/2}(t_n) - u^n_h\|^2_{h/2}\right)^{\frac{1}{2}}}{\max\limits_{1\le n\le N} \|u_{\tau/2,h/2}(t_n)\|_{L^2(\Omega)} + \left(\sum\limits_{n=1}^{N} \tau \|\nabla u_{\tau/2,h/2}(t_n)\|^2_{L^2(\mathcal{T}_h)}\right)^{\frac{1}{2}}}. \tag{4.5}$$

The corresponding simulated errors and rates are displayed in Table 4.3, which agrees with the results in Theorem 3.3.

5. Summary

In this paper, we have formulated a fully discrete dual-wind DG method for a parabolic variational inequality. Under the reasonable assumptions on the contact sets, error estimates are obtained in the energy norm. Several numerical examples are presented which demonstrate the theoretical rates of convergence established in Theorem 3.3. It is possible to extend the analysis to the time-dependent obstacle if the time derivative of the obstacle is sufficiently smooth. It is also of great interest to extend the analysis to quadratic dual-wind DG methods and investigate the rate of convergence. However such an analysis requires the construction of a positive-preserving interpolation error estimator that also satisfies

Table 4.3: Errors and convergence rates with respect to h and τ for Example 3

	h rate			τ rate		
	h	RTE	Rate	N	RTE	Rate
$\gamma = -1$	1/2	8.7405e-02	—	5	1.0172e-01	—
	1/4	4.9549e-02	0.8189	10	5.3537e-02	0.9260
	1/8	2.3074e-02	1.1026	20	2.4176e-02	1.1469
	1/16	1.1057e-02	1.0612	40	1.1640e-02	1.0545
	1/32	5.4271e-03	1.0268	80	5.7795e-03	1.0101
	1/64	2.6971e-03	1.0088	160	2.9071e-03	0.9913
$\gamma = 0$	1/2	8.8387e-02	—	5	1.0306e-01	—
	1/4	4.9660e-02	0.8318	10	5.3774e-02	0.9384
	1/8	2.2845e-02	1.1202	20	2.3977e-02	1.1652
	1/16	1.0786e-02	1.0827	40	1.1376e-02	1.0757
	1/32	5.2329e-03	1.0435	80	5.5879e-03	1.0256
	1/64	2.5841e-03	1.0180	160	2.7955e-03	0.9992
$\gamma = 1$	1/2	8.9865e-02	—	5	1.0306e-01	—
	1/4	5.0125e-02	0.8422	10	5.3774e-02	0.9385
	1/8	2.2841e-02	1.1339	20	2.3977e-02	1.1652
	1/16	1.0655e-02	1.1001	40	1.1376e-02	1.0757
	1/32	5.1215e-03	1.0570	80	5.5879e-03	1.0256
	1/64	2.5146e-03	1.0262	160	2.7955e-03	0.9992

the constraints on the midpoint of the triangulation. This together with the *a posteriori* error analysis for the adaptive algorithm of the proposed method will be investigated in the future.

Acknowledgements The work of the second author was partially supported by the National Science Foundation under Grant DMS-2111059. The work of the fourth author was partially supported by the National Science Foundation under Grant DMS-2111004.

References

[1] R.A. Adams and J.J.F. Fournier. *Sobolev Spaces (Second Edition).* Academic Press, Amsterdam, 2003.

[2] D. Arnold, F. Brezzi, B. Cockburn, and D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2002.

[3] S.C. Brenner. Poincaré-Friedrichs inequality for piecewise H^1 functions. *SIAM J. Numer. Anal.*, 41:306–324, 2003.

[4] S.C. Brenner, C.B. Davis, and L.-Y. Sung. A partition of unity method for a class of fourth order elliptic variational inequalities. Comput. Methods Appl. Mech Engrg., 276:612–626, 2014.

[5] S.C. Brenner, C.B. Davis, and L.-Y. Sung. A partition of unity method for the displacement obstacle problem of clamped Kirchhoff plates. J. Comput. Appl. Math., 265:3–16, 2014.

[6] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods (Third Edition).* Springer-Verlag, New York, 2008.

[7] S.C. Brenner, L.-Y. Sung, H. Zhang and Y. Zhang. A Morley finite element method for the displacement obstacle problem of clamped Kirchhoff plates. *J. Comput. Appl. Math.*, 254:31–42, 2013.

[8] S.C. Brenner, L.-Y. Sung, and Y. Zhang. Finite element methods for the displacement obstacle problem of clamped plates. *Math. Comp.*, 81:1247–1262, 2012.

[9] H. Brézis, *Seuil de régularité pour certains problèmes unilatéraux*, C. R. Ac. Sci. Paris, 273:35–37, 1971.

[10] H. Brézis, *Problèmes unilatéraux*, J. Math. Pures Appl., 57:1–168, 1972.

[11] F. Brezzi, W. W. Hager, and P. A. Raviart, *Error estimates for the finite element solution of variational inequalities I. Primal theory*, Numer. Math., 28:431–443, 1977.

[12] Z. Chen and R. Nochetto. Residual type a posteriori error estimates for elliptic obstacle problems. *Numer. Math.*, 84:527–548, 2000.

[13] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam, 1978.

[14] G. Duvaut and J.-L. Lions, *Inequalities in Mechanics and Physics*, Springer-Verlag, Berlin Heidelberg, 1976.

[15] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34:441–463, 1980.

[16] R. Falk, *Error estimates for the approximation of a class of variational inequalities*, Math. Comp., 28:963–971, 1974.

[17] X. Feng, T. Lewis, and M. Neilan. Discontinuous Galerkin finite element differential calculus and applications to numerical solutions of linear and nonlinear partial differential equations. *J. Comput. Appl. Math.*, 299:68–91, 2016.

[18] A. Friedman. *Variational Principles and Free-Boundary Problems.* Robert E. Krieger Publishing Co. Inc., Malabar, FL, second edition, 1988.

[19] V. Girault, B. Rivière, and M. Wheeler. A discontinuous Galerkin method with non-overlapping domain decomposition for the Stokes and Navier-Stokes problems. *Math. Comp.*, 74:53–84, 2005.

[20] R. Glowinski, J. L. Lions, and I. Tremolieres, *Numerical Analysis of Variational Inequalities*, North-Holland, Amsterdam, 1981.

[21] T. Gudi and P. Majumder. Convergence analysis of finite element method for a parabolic obstacle problem. *J. Comp., Appl. Math.*, 357:85–102, 2019.

[22] T. Gudi and P. Majumder. Conforming and discontinuous Galerkin FEM in space for solving parabolic obstacle problem. *Comput. Math. Appl.*, 78(12):3896–3915, 2019.

[23] T. Gudi and P. Majumder. Crouzeix-Raviart finite element approximation for the parabolic obstacle problem. *Comput. Meth. Appl. Math.*, 20(2):273–292, 2020.

[24] T. Gudi and K. Porwal, A reliable residual based a posteriori error estimator for a quadratic finite element method for the elliptic obstacle problem. *Comput. Methods Appl. Math.*, 15:145–160, 2015.

[25] J. Haslinger, I. Hlaváček, and J. Nečas, *Numerical methods for Unilateral Problems in Solid Mechanics.* Handbook of Numerical Analysis, ISSN: 1570-8659, Vol: 4, Page: 313-485, 1996.

[26] M. Hintermüller, K. Ito and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13:865–888, 2003.

[27] C. Johnson. A convergence estimate for an approximation of a parabolic variational inequality. *SIAM J. Numer. Anal.*, 13:599-606, 1976.

[28] S. Kesavan. *Topics in Functional Analysis and Applications.* New Age International Ltd, Publishers, NewDelhi, 1989.

[29] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications.* SIAM, Philadelphia, 2000.

[30] T. Lewis and M. Neilan. Convergence analysis of a symmetric dual-wind discontinuous Galerkin method. *J. Sci. Comput.*, 59:602-625, 2014.

[31] T. Lewis, A. Rapp, and Y. Zhang. Convergence analysis of symmetric dual-wind discontinuous Galerkin approximation methods for the obstacle problem. *J. Math. Anal. Appl.*, 485, Article 23840, 2020.

[32] J.-L. Lions and G. Stampacchia. Variational inequalities. *Comm. Pure Appl. Math.*, 20:493–519, 1967.

[33] P. Majumder. A convergence analysis of semi-discrete and fully-discrete nonconforming FEM for the parabolic obstacle problem. *Int. J. Comput. Math*, DOI: 10.1080/00207160.2020.1858285.

[34] P. Majumder. A priori error analysis of semi-discrete and fully-discrete DG finite element methods for a parabolic variational inequality. preprent, 2021.

[35] K.S. Moon, R.H. Nochetto, T. von Petersdorff, and C.S. Zhang. A posteriori error analysis for parabolic variational inequalities. *M2AN Math. Model. Numer. Anal.*, 41(3):485-511, 2007.

[36] J.-F. Rodrigues. *Obstacle Problems in Mathematical Physics.* North-Holland Publishing Co., Amsterdam, 1987.

[37] R. Trémolières, J.-L. Lions, and R. Glowinski. *Numerical Analysis of Variational Inequalities.* Studies in Mathematics and Its Applications, North-Holland, 1981.

[38] F. Wang and H. Wei. Virtual element method for the obstacle problem. *IMA J. Numer. Anal.*, 2018.

[39] F. Wang, W. Han, and X. Cheng. Discontinuous Galerkin methods for solving elliptic variational inequalities. *SIAM J. Numer. Anal.*, 48(2):708–733, 2010.

[40] L. Wang. On the quadratic finite element approximation to the obstacle problem. *Numer. Math.*, 92:771–778, 2002.