MONASH University

Faculty of Science
School of Mathematical Sciences

DOCTORAL THESIS

Presented by

Yahya ALNASHRI

The Gradient Discretisation Method For Variational Inequalities

A thesis submitted for the degree of Doctor of Philosophy

MONASH University
Discipline: **MATHEMATICS**

May 2017

Abstract

The purpose of this thesis is to develop a gradient discretisation method for elliptic and parabolic, linear and nonlinear, variational inequalities. The gradient discretisation method is a framework which enables a unified convergence analysis of many different methods – such as finite elements (*conforming, non-conforming* and mixed) and finite volumes methods – for 2^{nd} order diffusion equations.

Using the gradient discretisation method framework, we perform the numerical analysis of variational inequalities. We first establish error estimates for numerical approximations of linear elliptic variational inequalities. Using compactness techniques, we prove the convergence of numerical schemes for nonlinear elliptic variational inequalities based on Leray-Lions operators. We also show the uniform-in-time convergence for linear parabolic variational inequalities.

As numerical applications of this framework, we design, analyse and test the hybrid mimetic mixed (HMM) method for variational inequalities. Several numerical experiments are presented and demonstrate the accuracy of the proposed method, and confirm our theoretical rates of convergence, on grids with various cell geometries

Keywords: Elliptic variational inequalities, parabolic variational inequalities, obstacle problem, Signorini boundary conditions, nonlinear variational inequalities, nonlinear operators, Leray-Lions operator, seepage model, gradient discretisation method, gradient schemes, gradient discretisation, hybrid mimetic mixed methods, error estimates, convergence, monotonicity algorithm.

List of publications

The following is a list of publications/unpublished manuscripts resulting from this thesis.

- Y. Alnashri and J. Droniou. Gradient schemes for the Signorini and the obstacle problems, and application to hybrid mimetic mixed methods. Computers and Mathematics with Applications, 72, pp. 2788–2807, 2016.
- Y. Alnashri and J. Droniou. Convergence analysis of numerical schemes for non–linear variational inequalities. Submitted.
- Y. Alnashri and J. Droniou. Gradient schemes for an obstacle problem. Springer Proc. Math. Stat. Vol. 77. Finite volumes for complex applications. VII. Methods and theoretical aspects (Berlin, 2014). Springer, Cham, 2014, pp. 67–75.
- Y. Alnashri and J. Droniou. Gradient discretisation method for parabolic variational inequalities. In preparation.

Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer reviewed journals and two unpublished manuscripts. The core theme of the thesis is mathematical and numerical analysis. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the School of Mathematical Sciences, Monash University, under the supervision of Dr. Jérôme Droniou.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research. As is standard in mathematics, each author contributes equally to each aspect of the work.

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: [redacted] **Date:** May 05, 2017

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student and co-authors' contributions to this work.

Main Supervisor signature: [redacted] **Date:** May 05, 2017

This is dedicated to

my parent Ahmed Alnashri and Safiah Alnashri

my wife Eishah Alnashri

my sons, Ahmed, Abdulmajeed and Abdulkareem

my brothers

Acknowledgements

All praise and glory to Allah who provided me the blessings and strength to complete this work.

My sincere appreciation goes to my main supervisor, Dr. Jérôme Droniou, for his friendship, patient guidance and encouragement, in responding to my questions and queries, and for his valuable advice on presenting my thesis in this form. I find myself extremely lucky to have a supervisor like him. My appreciation also goes to my second supervisor, Dr. Todd Oliynyk.

I would like to extend my thanks to Monash University, both staff and organisation for assisting me throughout the completion of this research. My special appreciation goes to the School of Mathematical Sciences at Monash University for providing the facilities to enable me to undertake this research and for giving me the opportunities to attend several local and international conferences. Many thanks to Umm-Alqura University for providing me with a scholarship to complete my studies. I extend my thanks to the Saudi Arabia Cultural Mission (SACM) in Australia for their continuous support. Special thanks and appreciation to my friends, Ali Najmi and Abdullelah Al Thobaity for their help, support and encouragement during my PhD journey. I also thank Dr. Lisa Lines, who provided proofreading services, according to the guidelines laid out in the university endorsed national "Guidelines for Editing Research Theses".

My father and mother, Ahmed Alnashri and Safiah Alnashri are deserving thanks and appreciation for their prayers, patience and their encouragement of me to complete this study. I also thank my brothers for their support. Finally, I thank with love, my wife, Eishah Alnashri for her support and for the sacrifices she made for me to achieve this thesis.

Content

Chapter 1

Introduction

Variational inequalities (VIs) are problems involving partial differential equations, in which the solution must satisfy inequality conditions imposed either inside the domain or on a part of the boundary. Numerous problems that arise in fluid dynamics, elasticity, biomathematics, mathematical economics and control theory are modelled by VIs [57, 68, 53].

The topic of this thesis is the numerical analysis of variational inequalities. It aims to provide the generic convergence results of several numerical schemes for VIs, from error estimates in the case of linear models, to convergence analysis by compactness techniques in the case of nonlinear models. This generic analysis is done by applying the gradient discretisation method to such models.

The next section describes the seepage model, an application of VIs. Section 1.2 presents the principles of the gradient discretisation method. The main results are detailed in the last section of this introduction.

1.1 A motivation: the seepage problem

Free boundary problems, in which partial differential equations (PDEs) are written on a domain whose boundary is not given but has to be located as a part of the solution, are a common kind of variational inequalities. The main drivers of this research are unconfined seepage models, which describe the flow of water in nonhomogeneous dam as a free boundary problem, whose unknown surface is located between the wet and the dry regions. These models involve nonlinear quasi-variational inequalities obtained through the Baiocchi transform [7], or quasi-linear classical variational inequalities through an extension of the Darcy velocity inside the dry domain [96, 28, 71, 95].

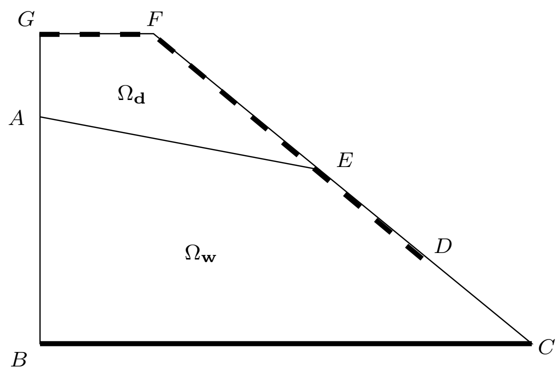

Figure 1.1: Geometry of the dam.

As in [96], Figure 1.1 shows the two-dimensional seepage flow through a nonhomogeneous dam. Let p denote to the pore water pressure and $\bar{u}$ denote to the total head at point $\boldsymbol{x} = (x, y)$. Neglecting the gravity, the

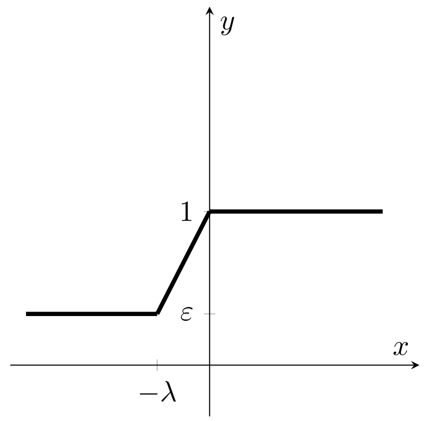

Figure 1.2: Heaviside function.

relation linking the pore pressure and the total head is

$$\bar{u} = y + \frac{p}{\mu} \quad \text{in } \Omega_{\mathbf{w}},$$

where μ is the constant water density and $\Omega_{\mathbf{w}}$ is the region occupied by the water. Using an approximated Heaviside function $\mathcal{H}_\varepsilon^\lambda$, the Darcy velocity of the flow, denoted by $\mathbf{V}$, can be written on the whole domain as

$$\mathbf{V} = \mathcal{H}_\varepsilon^\lambda(\bar{u} - y)\mathbf{K}\nabla\bar{u} \quad \text{in } \Omega, \tag{1.1.1}$$

where $\mathbf{K}$ denotes to the absolute permeability of the dam. This Heaviside function is chosen such that it is equal to 1 in the saturated part $\Omega_{\mathbf{w}}$, in which $\bar{u} - y \geq 0$, and the discharge velocity is approximately zero in the unsaturated part $\Omega_{\mathbf{d}} = \Omega \setminus \Omega_{\mathbf{w}}$ (Ω being the dam). As shown in Figure 1.2, one possible choice of the Heaviside function is

$$\mathcal{H}_\varepsilon^\lambda(\rho) = \begin{cases} 1 & \text{if } \rho \geq 0\,, \\ \frac{1-\varepsilon}{\lambda}\rho + 1 & \text{if } -\lambda < \rho < 0\,, \\ \varepsilon & \text{if } \rho \leq -\lambda. \end{cases} \tag{1.1.2}$$

The flow must satisfy the continuity equation,

$$\operatorname{div}\mathbf{V} = 0 \quad \text{in } \Omega. \tag{1.1.3}$$

The flow on the boundaries is governed by a mixture of Dirichlet, Neumann and inequalities boundary conditions. On the upstream surface AB and downstream CD, the total head of water is assumed to satisfy a nonhomogeneous Dirichlet boundary condition,

$$\bar{u} = g \quad \text{on } AB \cup CD, \tag{1.1.4}$$

where g is equal to the upstream flow depth on AB and it is equal to the downstream flow depth on CD.

The following flux boundary condition is imposed on the surface BC to reflects the fact that there is no flow at the bottom of the dam:

$$\mathbf{V} \cdot \mathbf{n} = 0 \text{ on } BC, \tag{1.1.5}$$

where $\mathbf{n}$ is the unit outward vector on the boundary.

The unknown point E, which needs to be located as a part of a solution, splits the free boundary surface DF into two parts. The total head of water $\bar{u}$ is assumed to be less than its ordinate y at the same point on the unsaturated part EF, and thus water may not enter or exit the dry domain on this part. This means that

$$\bar{u} < y \Rightarrow \mathbf{V} \cdot \mathbf{n} = 0 \quad \text{on } EF.$$

When $\bar{u}$ is higher than or equals to y (the pore pressure is nonnegative), the water tends to travel along the lower part DE. Due to the continuity of $\bar{u}$, the pore pressure cannot be strictly greater than zero on this saturated part, therefore, the flow occurs on this region only when the total head of water reaches the ordinate y (pressure vanishes). This means

$$\bar{u} = y \Rightarrow \mathbf{V} \cdot \mathbf{n} \leq 0 \quad \text{on } ED.$$

Finally, there is no flow and the pressure vanishes at seepage point E. On the whole surface, either the pore pressure is equal to zero or the flux is equal to zero. The flow and the pore pressure on the free boundary can be mathematically controlled by the Signorini boundary conditions,

$$\left.\begin{aligned} \bar{u} &\leq y \\ \mathbf{V} \cdot \mathbf{n} &\leq 0 \\ \mathbf{V} \cdot \mathbf{n}(\bar{u} - y) &= 0 \end{aligned}\right\} \quad \text{on } DFGA. \tag{1.1.6}$$

The seepage model can be described by a free boundary problem, whose solution satisfies Equations (1.1.1), (1.1.3), (1.1.4), (1.1.5) and the Signorini boundary condition (1.1.6).

1.2 Gradient discretisation method

The gradient discretisation method (GDM) is a framework for the analysis of numerical schemes for diffusion partial differential equations (PDEs) problems. This framework consists in discretising the weak variational formulations of PDEs using a small number of discrete elements, called a gradient discretisation. The scheme thus attained is called a gradient scheme. Under a few assumptions on the gradient discretisation, the corresponding gradient scheme can be shown to converge.

Previous studies [41] showed that this framework includes many well-known numerical schemes: *conforming, non–conforming* and mixed finite elements methods (including the *non-conforming* Crouzeix–Raviart method and the Raviart–Thomas method), hybrid mimetic mixed methods (which contain hybrid mimetic finite differences, hybrid finite volumes/SUSHI scheme and mixed finite volumes), *nodal* mimetic finite differences, and finite volumes methods (such as some multi-points flux approximation and discrete duality finite volume methods). Various boundary conditions can be considered within the gradient discretisation framework. This framework has been analysed for several linear and nonlinear elliptic and parabolic problems, including the Leray-Lions, Stokes, Richards, Stefan's and elasticity equations. We refer the reader to [40, 48, 42, 39, 49, 46] for more details, and the monograph [37] for a complete presentation.

1.2.1 Construction of the scheme

Consider the following simple linear elliptic PDE model:

$$\begin{aligned} -\mathrm{div}(\nabla \bar{u}) &= f \quad \text{in } \Omega, \\ \bar{u} &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.2.1}$$

where Ω is an open bounded connected subset of $\mathbb{R}^d$ $(d = 1, 2, 3)$ with a boundary $\partial\Omega$ and f is in $L^2(\Omega)$. The weak formulation of this problem is given by

$$\begin{aligned} &\text{Find } u \in H_0^1(\Omega), \text{ such that, } \forall v \in H_0^1(\Omega), \\ &\int_\Omega \nabla \bar{u}(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} = \int_\Omega f(\boldsymbol{x}) v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{1.2.2}$$

The principle of the GDM for such a model relies on defining a discrete space and reconstruction operators, which together are called a gradient discretisation (GD). Replacing the continuous spaces and operators in the weak formulation (1.2.2) by these discrete elements yields a numerical scheme for this problem, called a gradient scheme (GS).

Selecting the gradient discretisation mostly depends on the boundary conditions (BCs). For homogenous Dirichlet BCs, the gradient discretisation is a triplet $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$, where the space of degrees of freedom $X_{\mathcal{D},0}$ is a discrete version of the continuous space $H^1_0(\Omega)$, $\Pi_{\mathcal{D}}$ is a function reconstruction operator that relates an element of $X_{\mathcal{D},0}$ to a function in $L^2(\Omega)$, and $\nabla_{\mathcal{D}}$ is a gradient reconstruction in $L^2(\Omega)^d$ from the degrees of freedom. It must be chosen such that $||\nabla_{\mathcal{D}} \cdot ||_{L^2(\Omega)^d}$ is a norm on $X_{\mathcal{D},0}$. Substituting these discrete elements in place of the continuous space and operators in (1.2.2) gives the gradient scheme for (1.2.2),

$$\begin{aligned}&\text{Find } u \in X_{\mathcal{D},0} \text{ such that, } \forall v \in X_{\mathcal{D},0},\\ &\int_\Omega \nabla_{\mathcal{D}} u(\boldsymbol{x}) \cdot \nabla_{\mathcal{D}} v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} = \int_\Omega f(\boldsymbol{x}) \Pi_{\mathcal{D}} v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{1.2.3}$$

Owing to the versatile choice of the GD, many numerical schemes can be written in the setting of this gradient scheme. We provide here two simple examples of numerical schemes. In each of these schemes, we define $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$ such that (1.2.3) is the corresponding scheme. In both of following examples, let $\mathcal{T}$ be a conforming triangulation of Ω.

Example 1.2.1 (*Conforming* $\mathbb{P}1$ **finite element method).** The discrete space $X_{\mathcal{D},0}$ is made of vectors of values at the *nodes* of the mesh, the operator $\Pi_{\mathcal{D}} v$ is the piecewise linear continuous function that takes these values at the *nodes*, and $\nabla_{\mathcal{D}} v = \nabla(\Pi_{\mathcal{D}} v)$. Using this gradient discretisation $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$ in (1.2.3) gives the *conforming* $\mathbb{P}1$ finite element method to (1.2.2).

Example 1.2.2 (*Non-conforming* $\mathbb{P}1$ **finite element method).** We take $X_{\mathcal{D},0}$ as the space of piecewise linear functions on $\mathcal{T}$, which are continuous at the edge mid-points and take zero values at mid-points of all boundary edges. The GD is completed by setting $\Pi_{\mathcal{D}} = \mathbf{Id}$, and $\nabla_{\mathcal{D}} = \nabla_{\mathrm{B}}$, which is a broken gradient defined by

$$\text{for all } v \in X_{\mathcal{D},0}, \text{ for all triangle } T \in \mathcal{T}, \forall \boldsymbol{x} \in T, \quad \nabla_{\mathrm{B}} v(\boldsymbol{x}) = \nabla(v_{|T}).$$

Another example, the HMM method, is covered in details in Chapter 3.

1.2.2 Properties for the convergence analysis of GS

The quality of the discrete elements $(X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$ is measure through the constant $C_{\mathcal{D}}$, and the functions $S_{\mathcal{D}} : H^1_0(\Omega) \to [0, +\infty)$ and $W_{\mathcal{D}} : H_{\mathrm{div}}(\Omega) \to [0, +\infty)$ respectively defined by

$$C_{\mathcal{D}} = \max_{v \in X_{\mathcal{D},0} \setminus \{0\}} \frac{\|\Pi_{\mathcal{D}} v\|_{L^2(\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}}, \tag{1.2.4}$$

$$\forall \varphi \in H^1_0(\Omega),\ S_{\mathcal{D}}(\varphi) = \min_{v \in X_{\mathcal{D},0}} \big(\|\Pi_{\mathcal{D}} v - \varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla\varphi\|_{L^2(\Omega)^d}\big) \tag{1.2.5}$$

and

$$\forall \boldsymbol{\psi} \in H_{\mathrm{div}}(\Omega),\ W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v \in X_{\mathcal{D},0} \setminus \{0\}} \frac{1}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}} \Big| \int_\Omega (\nabla_{\mathcal{D}} v \cdot \boldsymbol{\psi} + \mathit{\Pi}_{\mathcal{D}} v \,\mathrm{div}(\boldsymbol{\psi}))\,\mathrm{d}\boldsymbol{x} \Big|, \tag{1.2.6}$$

where $H_{\mathrm{div}}(\Omega) = \{\boldsymbol{\psi} \in L^2(\Omega)^d : \mathrm{div}\boldsymbol{\psi} \in L^2(\Omega)\}$.

The constant $C_{\mathcal{D}}$ is used to measure the coercivity (it yields a discrete Poincaré inequality). The function $S_{\mathcal{D}}$ measures the accuracy of approximating smooth functions and their gradients by elements defined from the discrete space. In some contexts, it is called the interpolation error. The following Stokes' formula is an important property satisfied by the usual gradient:

$$\int_\Omega (\nabla u \cdot \varphi + u \,\mathrm{div}\varphi)\,\mathrm{d}\boldsymbol{x} = 0,\ \forall u \in H_0^1(\Omega), \varphi \in H_{\mathrm{div}}(\Omega).$$

The discrete version of this formula is usually not valid, especially in the case of *non-conforming* methods. The function $W_\mathcal{D}$ assesses how well the function reconstruction $\Pi_\mathcal{D}$ and the gradient reconstruction $\nabla_\mathcal{D}$ satisfy the Stokes' formula. If the method is *conforming*, that is $X_{\mathcal{D},0} \subset H_0^1(\Omega)$, $\Pi_\mathcal{D} = \mathbf{Id}$ and $\nabla_\mathcal{D} = \nabla$, then $W_\mathcal{D} \equiv 0$.

Based on these constant and functions, error estimates can be obtained, in the GDM, between the solution $\bar{u}$ to (1.2.2) and the solution u to (1.2.3):

$$||\nabla_\mathcal{D} u - \nabla \bar{u}||_{L^2(\Omega)^d} \le 2S_\mathcal{D}(\bar{u}) + W_\mathcal{D}(\nabla \bar{u}) \tag{1.2.7}$$

and

$$||\Pi_\mathcal{D} u - \bar{u}||_{L^2(\Omega)} \le (C_\mathcal{D} + 1)S_\mathcal{D}(\bar{u}) + C_\mathcal{D} W_\mathcal{D}(\nabla \bar{u}). \tag{1.2.8}$$

Let us now explain the roles of the three indicators $C_\mathcal{D}$, $S_\mathcal{D}$ and $W_\mathcal{D}$ in establishing the above error estimates. In general, obtaining error estimates for *conforming* methods starts by taking a generic $v = v_h \in V_h \subset H_0^1(\Omega)$ as a test function in the continuous problem (1.2.2). Subtracting this problem, written for v_h, from the approximate scheme yields the following relation between the exact and the approximate solutions,

$$\int_\Omega \nabla(\bar{u} - u_h) \cdot \nabla v_h \,\mathrm{d}\boldsymbol{x} = 0. \tag{1.2.9}$$

For the GDM, which includes *non–conforming* methods, this formula does not necessarily make sense since there is no hope for the continuous space $H_0^1(\Omega)$ and the discrete space $X_{\mathcal{D},0}$ to share a common test function as it does in the *conforming* case. The function $W_\mathcal{D}$ is required to deduce an approximate form of (1.2.9); applying the definition of $W_\mathcal{D}$ to $\boldsymbol{\psi} = \nabla \bar{u} \in H_{\mathrm{div}}(\Omega)$ gives

$$\left| \int_\Omega (\nabla \bar{u} \cdot \nabla_\mathcal{D} v + \Pi_\mathcal{D} v \Delta \bar{u})\,\mathrm{d}\boldsymbol{x} \right| \le ||\nabla_\mathcal{D} v||_{L^2(\Omega)^d} W_\mathcal{D}(\nabla \bar{u}),$$

which leads to, since $-\Delta \bar{u} = f$ and u is a solution to the gradient scheme (1.2.3)

$$\int_\Omega (\nabla \bar{u} - \nabla_\mathcal{D} u) \cdot \nabla_\mathcal{D} v \,\mathrm{d}\boldsymbol{x} \le ||\nabla_\mathcal{D} v||_{L^2(\Omega)^d} W_\mathcal{D}(\nabla \bar{u}). \tag{1.2.10}$$

Let $v = w - u$ where $w \in X_{\mathcal{D},0}$, to get

$$\begin{aligned}\int_\Omega (\nabla_\mathcal{D} w - \nabla_\mathcal{D} u) \cdot (\nabla_\mathcal{D} w - \nabla_\mathcal{D} u)\,\mathrm{d}\boldsymbol{x} \le &\int_\Omega (\nabla_\mathcal{D} w(x) - \nabla \bar{u}(x)) \cdot (\nabla_\mathcal{D} w(x) - \nabla_\mathcal{D} u(x))\,\mathrm{d}\boldsymbol{x} \\ &+ ||\nabla_\mathcal{D}(w - u)||_{L^2(\Omega)^d} W_\mathcal{D}(\nabla \bar{u}).\end{aligned}$$

This inequality becomes, thanks to the Cauchy–Schwarz inequality,

$$||\nabla_\mathcal{D} u - \nabla_\mathcal{D} w||_{L^2(\Omega)^d} \le ||\nabla \bar{u} - \nabla_\mathcal{D} w||_{L^2(\Omega)^d} + W_\mathcal{D}(\nabla \bar{u}). \tag{1.2.11}$$

The role of the function $S_\mathcal{D}$ rises in providing the best interpolant choice, given by

$$w = \arg \min_{w \in X_{\mathcal{D},0}} S_\mathcal{D}(\bar{u}).$$

This interpolant leads to the following bounds, thanks to the definition of $S_\mathcal{D}$:

$$||\nabla_\mathcal{D} w - \nabla \bar{u}||_{L^2(\Omega)^d} \le S_\mathcal{D}(\bar{u}) \quad \text{and} \quad ||\Pi_\mathcal{D} w - \bar{u}||_{L^2(\Omega)} \le S_\mathcal{D}(\bar{u}), \tag{1.2.12}$$

which gives in turn

$$\begin{aligned}||\nabla_{\mathcal{D}}u - \nabla\bar{u}||_{L^2(\Omega)^d} &\leq ||\nabla_{\mathcal{D}}u - \nabla_{\mathcal{D}}w||_{L^2(\Omega)^d} + ||\nabla_{\mathcal{D}}w - \nabla\bar{u}||_{L^2(\Omega)^d}\\ &\leq ||\nabla_{\mathcal{D}}u - \nabla_{\mathcal{D}}w||_{L^2(\Omega)^d} + S_{\mathcal{D}}(\bar{u}).\end{aligned}$$

From this inequality and (1.2.11), the estimate (1.2.7) is concluded.

The constant $C_{\mathcal{D}}$ is to ensure the following discrete Poincaré inequality

$$||\Pi_{\mathcal{D}}v||_{L^2(\Omega)} \leq C_{\mathcal{D}}||\nabla_{\mathcal{D}}v||_{L^2(\Omega)^d}, \quad \text{for all } v \in X_{\mathcal{D},0}.$$

The application of this inequality to $u - w$, together with (1.2.11) and (1.2.12), leads to

$$||\Pi_{\mathcal{D}}(u - w)||_{L^2(\Omega)} \leq C_{\mathcal{D}}(S_{\mathcal{D}}(\bar{u}) + W_{\mathcal{D}}(\nabla\bar{u})).$$

Using the triangular inequality, one can write $||\Pi_{\mathcal{D}}u - \bar{u}||_{L^2(\Omega)} \leq ||\Pi_{\mathcal{D}}u - \Pi_{\mathcal{D}}w||_{L^2(\Omega)} + ||\Pi_{\mathcal{D}}w - \bar{u}||_{L^2(\Omega)}$. The above inequality and (1.2.12) show that Estimate (1.2.8) holds.

Choosing a sequence of gradient discretisation $(\mathcal{D}_m)_{m\in\mathbb{N}}$ such that $(C_{\mathcal{D}_m})_{m\in\mathbb{N}}$ remains bounded and that $(S_{\mathcal{D}_m}(\bar{u}))_{m\in\mathbb{N}}$ and $(W_{\mathcal{D}_m}(\nabla\bar{u}))_{m\in\mathbb{N}}$ converge to 0, Estimates (1.2.7) and (1.2.8) show that $(\Pi_{\mathcal{D}_m}u_m)_{m\in\mathbb{N}}$ converges strongly to $\bar{u}$ in $L^2(\Omega)$ and $(\nabla_{\mathcal{D}_m}u_m)_{m\in\mathbb{N}}$ converges strongly to $\nabla\bar{u}$ in $L^2(\Omega)^d$.

1.2.3 GDM and nonlinear problems

The GDM can also be applied to nonlinear problem, such as stationary diffusion problems:

$$\begin{aligned}-\mathrm{div}(\Lambda(\bar{u})\nabla\bar{u}) &= f \quad \text{in } \Omega,\\ \bar{u} &= 0 \quad \text{on } \partial\Omega,\end{aligned} \tag{1.2.13}$$

with the same assumptions as in Section 1.2.1, and Λ is a Carathéodory function from $\Omega \times \mathbb{R}$ to the set of $d \times d$ symmetric matrices $S_d(\mathbb{R})$ such that it has eigenvalues in $(\underline{\lambda}, \overline{\lambda}) \subset (0, +\infty)$. The weak formulation of this problem is

$$\begin{aligned}&\text{Find } u \in H_0^1(\Omega), \text{ such that}, \forall v \in H_0^1(\Omega),\\ &\int_\Omega \Lambda(\bar{u})\nabla\bar{u}(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} = \int_\Omega f(\boldsymbol{x})v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{1.2.14}$$

Using the same gradient discretisation $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$ as for the linear case, the related gradient scheme for (1.2.14) is

$$\begin{aligned}&\text{Find } u \in X_{\mathcal{D},0} \text{ such that}, \forall v \in X_{\mathcal{D},0},\\ &\int_\Omega \Lambda(\Pi_{\mathcal{D}}u)\nabla_{\mathcal{D}}u(\boldsymbol{x}) \cdot \nabla_{\mathcal{D}}v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} = \int_\Omega f(\boldsymbol{x})\Pi_{\mathcal{D}}v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{1.2.15}$$

In order for a sequence of gradient discretisation $(\mathcal{D}_m)_{m\in\mathbb{N}}$ to provide converging gradient schemes for the nonlinear problem (1.2.14), it is expected that:

(P1) $(C_{\mathcal{D}_m})_{m\in\mathbb{N}}$ remains bounded,

(P2) $\lim_{m\to\infty} S_{\mathcal{D}_m}(\varphi) = 0,\ \forall\varphi \in H_0^1(\Omega)$,

(P3) $\lim_{m\to\infty} W_{\mathcal{D}_m}(\boldsymbol{\psi}) = 0,\ \forall\boldsymbol{\psi} \in H_{\mathrm{div}}$,

(P4) for all $(u_m)_{m\in\mathbb{N}}$ such that $u_m \in X_{\mathcal{D}_m,0}$ for all $m \in \mathbb{N}$ and $(\nabla_{\mathcal{D}_m}u_m)_{m\in\mathbb{N}}$ is bounded in $L^2(\Omega)^d$, then the sequence $(\Pi_{\mathcal{D}_m}u_m)_{m\in\mathbb{N}}$ is relatively compact in $L^2(\Omega)$.

Here $C_{\mathcal{D}}$, $S_{\mathcal{D}}$ and $W_{\mathcal{D}}$ are respectively defined by (1.2.4), (1.2.5) and (1.2.6). These properties are precisely called the *coercivity*, the *consistency*, the *limit-conformity* and *compactness* of the sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$. The convergence of the gradient scheme (1.2.15) towards an exact solution of (1.2.14) is usually proved by a compactness technique, whose key ideas are:

1. establish an energy estimate on the approximate solution to the scheme (1.2.15). The coercivity property is essential for the existence of this estimate,
2. if such a bound is obtained, the limit-conformity property shows the weak convergence of $\Pi_{\mathcal{D}_m}u_m$ (resp. $\nabla_{\mathcal{D}_m}u_m$) in $L^2(\Omega)$ (resp. $L^2(\Omega)^d$) to a function $\bar{u}\in H_0^1(\Omega)$ (resp. $\nabla\bar{u}$), as stated in the lemma below, proved in [37]. Due to the compactness property, the convergence of $\Pi_{\mathcal{D}_m}u_m$ is indeed strong.

   Lemma 1.2.3. *Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretisations for homogeneous Dirichlet BCs, which is coercive in the sense of* **(P1)***, and limit-conforming in the sense of* **(P3)***. Let $u_m\in X_{\mathcal{D}_m,0}$ be such that $(||\nabla_{\mathcal{D}_m}u_m||_{L^2(\Omega)^d})_{m\in\mathbb{N}}$ remains bounded. Then there exists a sequence of $(\mathcal{D}_m,u_m)$, denoted in the same way, and $\bar{u}\in H_0^1(\Omega)$ such that $\Pi_{\mathcal{D}_m}u_m$ converges weakly in $L^2(\Omega)$ to $\bar{u}$, and $\nabla_{\mathcal{D}_m}u_m$ converges weakly in $L^2(\Omega)^d$ to $\nabla\bar{u}$.*

3. show that the limit $\bar{u}\in H_0^1(\Omega)$ is the solution to (1.2.2). During this stage, the consistency is required to construct an interpolation w of a given test function $\varphi\in H_0^1(\Omega)$ such that $\Pi_{\mathcal{D}_m}w_m\to\varphi$ in $L^2(\Omega)$ and $\nabla_{\mathcal{D}_m}w_m\to\nabla\varphi$ in $L^2(\Omega)^d$ as $m\to\infty$. The strong convergence of the sequence of an approximate gradient $\nabla_{\mathcal{D}_m}u_m$ usually depends on the assumptions on the continuous model.

1.3 Summary of the thesis

This thesis studies the application of the gradient discretisation method to different kinds of elliptic and parabolic variational inequalities modelling the flow of water in porous media.

Chapter 2

The aim of this chapter is to study the gradient discretisation method for linear variational inequalities. The main VI considered in this chapter involves a mixture of Dirichlet, Neumann and Signorini boundary conditions, with each one set on a different part of the boundary:

$$\left\{\begin{array}{rl} -\operatorname{div}(\Lambda\nabla\bar{u})=f & \text{in }\Omega,\\ \bar{u}=0 & \text{on }\Gamma_1,\\ \Lambda\nabla\bar{u}\cdot\mathbf{n}=0 & \text{on }\Gamma_2,\\ \left.\begin{array}{r}\bar{u}\le a\\ \Lambda\nabla\bar{u}\cdot\mathbf{n}\le 0\\ (a-\bar{u})\Lambda\nabla\bar{u}\cdot\mathbf{n}=0\end{array}\right\} & \text{on }\Gamma_3,\end{array}\right. \tag{1.3.1}$$

where Λ is a function from Ω to $S_d(\mathbb{R})$, the set of $d\times d$ symmetric matrices, that it has eigenvalues in $(\underline{\lambda},\overline{\lambda})\subset(0,+\infty)$, Ω is a bounded open set of $\mathbb{R}^d$ $(d\ge 1)$, $\mathbf{n}$ is the unit outer normal to $\partial\Omega$ and $(\Gamma_1,\Gamma_2,\Gamma_3)$ is a partition of $\partial\Omega$.

Several linear models involve possibly heterogeneous and anisotropic tensors. This happens to contact problems in elasticity models [62] involving composite materials (for which the stiffness tensor depends on the position), and in lubrication problems [21] (in which the tensor is a function of the first fundamental form of the film).

We focus on the weak formulation of (1.3.1), given by,

$$\left\{\begin{array}{l}\text{Find }\bar{u}\in\mathcal{K}:=\{v\in H^1(\Omega)\ :\ \gamma(v)=0\text{ on }\Gamma_1,\ \gamma(v)\le a\text{ on }\Gamma_3\text{ a.e.}\}\ \text{ s.t.,}\\ \forall v\in\mathcal{K}\,,\quad\displaystyle\int_\Omega\Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x})\cdot\nabla(\bar{u}-v)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\le\int_\Omega f(x)(\bar{u}(\boldsymbol{x})-v(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}.\end{array}\right. \tag{1.3.2}$$

Here γ is the trace operator.

The first step to discretise this problem is to define a relevant gradient discretisation. For the Signorini problem, the gradient discretisation consists of a discrete space $X_{\mathcal{D},\Gamma_{2,3}}$, a linear function reconstruction $\Pi_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to L^2(\Omega)$, a linear trace reconstruction $\mathbb{T}_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to H^{1/2}(\partial\Omega)$, and a linear gradient reconstruction $\nabla_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to L^2(\Omega)^d$, the latter known to induce a norm on $X_{\mathcal{D},\Gamma_{2,3}}$. Substituting these discrete elements in the continuous weak formulation (1.3.2) gives the gradient scheme for the Signorini problem,

$$\left\{\begin{array}{l} \text{Find } u \in \mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},\Gamma_{2,3}} \ : \ \mathbb{T}_{\mathcal{D}} v \le a \text{ on } \Gamma_3\} \ \text{ s.t.}, \forall v \in \mathcal{K}_{\mathcal{D}}\,, \\ \displaystyle\int_\Omega \Lambda(\boldsymbol{x})\nabla_{\mathcal{D}} u(\boldsymbol{x}) \cdot \nabla_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \le \int_\Omega f(\boldsymbol{x})\Pi_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}. \end{array}\right. \tag{1.3.3}$$

As mentioned in Section 1.2.2, the quality of these discrete elements $(X_{\mathcal{D},\Gamma_{2,3}}, \Pi_{\mathcal{D}}, \mathbb{T}_{\mathcal{D}}, \nabla_{\mathcal{D}})$ is usually measured by three indicators $C_{\mathcal{D}}$, $S_{\mathcal{D}}$ and $W_{\mathcal{D}}$. The constant $C_{\mathcal{D}}$ is defined by

$$C_{\mathcal{D}} = \max_{v \in X_{\mathcal{D},\Gamma_{2,3}} \setminus \{0\}} \Big(\frac{\|\Pi_{\mathcal{D}} v\|_{L^2(\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}} + \frac{\|\mathbb{T}_{\mathcal{D}} v\|_{H^{1/2}(\partial\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}} \Big). \tag{1.3.4}$$

This definition does not only depend on the norm of $\Pi_{\mathcal{D}}$ as in the PDEs models, but it also includes the norm of the reconstruction trace $\mathbb{T}_{\mathcal{D}}$. We define the function $S_{\mathcal{D}} : \mathcal{K} \times \mathcal{K}_{\mathcal{D}} \to [0, +\infty)$ by

$$\forall (\varphi, v) \in \mathcal{K} \times \mathcal{K}_{\mathcal{D}},\ S_{\mathcal{D}}(\varphi, v) = \|\Pi_{\mathcal{D}} v - \varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla\varphi\|_{L^2(\Omega)^d}.$$

To build an interpolant inside $\mathcal{K}_{\mathcal{D}}$ from the elements of $\mathcal{K}$, the function $S_{\mathcal{D}}$ is defined in such a way as to act on the discrete set where the approximate solution and test functions are sought, and on the continuous set in which the exact solution lies. Finally, we define the function $W_{\mathcal{D}} : H_{\mathrm{div}}(\Omega) \to [0, +\infty)$ by for all $\boldsymbol{\psi} \in H_{\mathrm{div}}(\Omega)$,

$$W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v \in X_{\mathcal{D},\Gamma_{2,3}} \setminus \{0\}} \frac{1}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}} \Big| \int_\Omega (\nabla_{\mathcal{D}} v \cdot \boldsymbol{\psi} + \Pi_{\mathcal{D}} v \,\mathrm{div}(\boldsymbol{\psi})) \,\mathrm{d}\boldsymbol{x} - \langle \gamma_{\mathbf{n}}(\boldsymbol{\psi}), \mathbb{T}_{\mathcal{D}} v \rangle \Big|, \tag{1.3.5}$$

where $\langle \cdot, \cdot \rangle$ denotes to the duality product between $H^{1/2}(\partial\Omega)$ and $H^{-1/2}(\partial\Omega)$. This definition contains a major change with respect to (1.2.6). This is due to insufficient regularities on the solution on the boundary Γ_3, preventing the reconstruction operator $\mathbb{T}_{\mathcal{D}}$ from constructing functions in $L^2(\partial\Omega)$, see Remark 2.2.1 for details.

The main contribution of Chapter 2 is to carry out, thanks to GDM, a unified convergence analysis of several numerical methods for linear variational inequalities, and to find a generic formula to obtain the convergence rates of these numerical schemes. The main result of this chapter is Theorem 2.2.7, which gives general error estimates between the solution to (1.3.2) and the solution to (1.3.3):

$$\|\nabla_{\mathcal{D}} u - \nabla\bar{u}\|_{L^2(\Omega)^d} \le \sqrt{\frac{2}{\underline{\lambda}} G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+} + \frac{1}{\underline{\lambda}}[W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + (\overline{\lambda} + \underline{\lambda}) S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})], \quad \forall v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}} \tag{1.3.6}$$

and

$$\|\Pi_{\mathcal{D}} u - \bar{u}\|_{L^2(\Omega)} \le C_{\mathcal{D}} \sqrt{\frac{2}{\underline{\lambda}} G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+} + \frac{1}{\underline{\lambda}}[C_{\mathcal{D}} W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + (C_{\mathcal{D}}\overline{\lambda} + \underline{\lambda}) S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})], \quad \forall v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}} \tag{1.3.7}$$

where

$$G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = \langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), \mathbb{T}_{\mathcal{D}} v_{\mathcal{D}} - \gamma\bar{u} \rangle \quad \text{and} \quad G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+ = \max(0, G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})).$$

Dealing with variational inequalities however requires us to establish new estimates with an additional term $G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$. This additional term apparently leads to a degraded rate of convergence, but we show that the optimal rates can be easily recovered for many classical methods.

The first general error estimate for VIs is established by Falk [50]. The novelty of our estimates is that they are applicable to both *conforming* and *non-conforming* numerical schemes. A key idea in Falk's proof requires the use of a common test function, that is admissible in both the continuous and the discrete problems. Since we are dealing here with possibly *nonconforming* schemes, the technique in [50] cannot be used to obtain error estimates. Instead, we develop a similar technique that used in Section 1.2.2 for PDEs, that starts by applying the definition of $W_{\mathcal{D}}$ in (1.3.5) to $\boldsymbol{\psi} = \Lambda\nabla\bar{u}$ and $v = w - u$, such that $w \in \mathcal{K}_{\mathcal{D}}$. In order to obtain the equivalent for VIs of (1.2.10), we need to develop a new technique to deal with the duality product appearing in the definition of $W_{\mathcal{D}}$.

The general error estimates formulas stemming from our convergence analysis are simpler than the ones provided in the literature such as in [50], since they only depend upon the choice of an interpolant in $\mathcal{K}_{\mathcal{D}}$. With an interpolant constructed from the values of $\bar{u}$ at the vertices, Estimates (1.3.6) and (1.3.7) show that a first order *conforming* numerical method has an order one error estimate on the H^1 norm. We also obtain the error estimates for *non-conforming* methods, see Theorem 2.2.10. This theorem can be used to obtain convergence rates for new methods that can be written in the manner of gradient scheme (1.3.3).

This chapter also considers an obstacle model, where inequalities are imposed inside the domain, and the main corresponding results are given in Theorems 2.2.8 and 2.2.11.

The results of Chapter 2 have been published in [3].

Chapter 3

The focus of this chapter is the discretisation of linear variational inequalities by the hybrid mimetic mixed (HMM) method, which contains the hybrid finite volume methods [47], the (mixed/hybrid) mimetic finite differences methods [18] and the mixed finite volume methods [36]. These methods were developed for anisotropic heterogeneous diffusion equations on generic grids, as often encountered in engineering problems related to flows in porous media.

Using the generic setting of the GDM, we define the HMM methods for the linear Signorini and the obstacle problems. Theorems 2.2.10 and 2.2.11 enable us to establish the convergence rates of the HMM schemes applied on generic mesh.

The second part of this chapter concerns the computation of the HMM solutions in practice. For the Signorini model, we first show that the gradient scheme (1.3.3), written for a particular HMM gradient discretisation, can be represented in the manner of finite volume methods, which are based on balance and conservation equations. Let $\mathcal{M}$ denote to the set of cells K, $\mathcal{E}$ denote to the edges of cells, and $\mathcal{E}_K$ denotes to the edges of a given cell K. Also, let u_K (resp. u_σ) represents an approximation value of the unknown $\bar{u}$ at $\overline{\boldsymbol{x}}_K$, the centre of mass of K (resp. $\overline{\boldsymbol{x}}_\sigma$, the centre of mass of σ). The numerical flux $F_{K,\sigma}(u)$ is defined as a linear function of the unknowns $(u_K)_{K\in\mathcal{M}}, (u_\sigma)_{\sigma\in\mathcal{E}}$ (the precise definition can be found in the formula (3.3.5)). The finite volume presentation of HMM for the Signorini problem is:

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) = \int_K f \,\mathrm{d}\boldsymbol{x}, \quad \forall K \in \mathcal{M}, \tag{1.3.8}$$

$$F_{K,\sigma}(u) + F_{L,\sigma}(u) = 0, \quad \text{if } \sigma \in \mathcal{E}_K \cap \mathcal{E}_L, K \neq L, \tag{1.3.9}$$

$$u_\sigma = 0, \quad \forall \sigma \in \mathcal{E}_{\text{ext}}(\text{external edges}) \text{ such that } \sigma \subset \Gamma_1, \tag{1.3.10}$$

$$F_{K,\sigma}(u) = 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \subset \Gamma_2, \tag{1.3.11}$$

$$F_{K,\sigma}(u)(u_\sigma - a_\sigma) = 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \subset \Gamma_3, \tag{1.3.12}$$

$$-F_{K,\sigma}(u) \le 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \subset \Gamma_3, \tag{1.3.13}$$

$$u_\sigma \le a_\sigma, \quad \forall \sigma \in \mathcal{E}_{\text{ext}} \text{ such that } \sigma \subset \Gamma_3, \tag{1.3.14}$$

where a_σ is a constant approximation of a on σ.

An iterative monotonicity algorithm proposed in [59] is used to calculate the solution of this scheme. It begins with setting $\mathbb{G}^{(0)} = \mathcal{E}_{\Gamma_3} := \{\sigma \in \mathcal{E}_{\rm ext} : \ \sigma \subset \Gamma_3\}$ and $\mathbb{H}^{(0)} = \emptyset$. At any step n, we assume that there are two known disjointed sets such that $\mathbb{G}^{(n)} \cup \mathbb{H}^{(n)} = \mathcal{E}_{\Gamma_3}$, and we solve the linear system of Equations (1.3.8), (1.3.9), (1.3.10), (1.3.11), together with

$$\begin{aligned} F_{K,\sigma}(u^{(n)}) = 0, &\quad \forall K \in \mathcal{M}\,, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \in \mathbb{G}^{(n)},\\ u_\sigma^{(n)} = a_\sigma, &\quad \forall \sigma \in \mathcal{E}_{\rm ext} \text{ such that } \sigma \in \mathbb{H}^{(n)}. \end{aligned}$$

We then consider the edges included in Γ_3 to define the new two sets $\mathbb{G}^{(n+1)}$ and $\mathbb{H}^{(n+1)}$ for the next step $n+1$. We seek edges in $\mathbb{G}^{(n)}$ that break the condition (1.3.14) and we transfer them to the new set $\mathbb{H}^{(n+1)}$, which also gathers all edges $\sigma \in \mathbb{H}^{(n)}$ such that $-F_{K,\sigma}(u^{(n)}) < 0$. The set $\mathbb{G}^{(n+1)}$ consists of edges in the old set $\mathbb{G}^{(n)}$, that meet the constraint (1.3.14), and edges σ in $\mathbb{H}^{(n)}$ where $-F_{K,\sigma}(u^{(n)}) \geq 0$. The algorithm stops at $N \in \mathbb{N}$ such that $\mathbb{G}^{(N)} = \mathbb{G}^{(N+1)}$ and $\mathbb{H}^{(N)} = \mathbb{H}^{(N+1)}$. The solution to (1.3.8)–(1.3.14) is then $u = u^{(N)}$.

Different numerical tests are provided in Chapter 3 to evaluate the behaviour of the HMM methods on general meshes. We develop a numerical test on a Signorini problem, for which a non-trivial exact solution is known in order to confirm the theoretical rates of convergence.

Most of materials in Chapter 3 have been published in [3].

Chapter 4

The main objective of Chapter 4 is to use the GDM to provide a complete numerical analysis of nonlinear variational inequalities based on Leray-Lions operators. Firstly, we are interested in the nonlinear Signorini model,

$$\left\{\begin{array}{rl} -\operatorname{div} \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) = f & \text{in } \Omega,\\ \bar{u} = g & \text{on } \Gamma_1,\\ \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \mathbf{n} = 0 & \text{on } \Gamma_2,\\ \left.\begin{array}{r} \bar{u} \leq a\\ \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \mathbf{n} \leq 0\\ (a-\bar{u})\mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \mathbf{n} = 0 \end{array}\right\} & \text{on } \Gamma_3. \end{array}\right. \tag{1.3.16}$$

The weak formulation of this model is: for a given $p \in (1,\infty)$ associated with $\mathbf{a}$,

$$\left\{\begin{array}{l} \text{Find } \bar{u} \in \mathcal{K} := \{v \in W^{1,p}(\Omega) \ : \ \gamma(v) = g \text{ on } \Gamma_1, \ \gamma(v) \leq a \text{ on } \Gamma_3 \text{ a.e.}\} \text{ such that,}\\ \forall v \in \mathcal{K}, \quad \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}(\boldsymbol{x}), \nabla \bar{u}(\boldsymbol{x})) \cdot \nabla(\bar{u}-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \leq \int_\Omega f(\boldsymbol{x})(\bar{u}-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}. \end{array}\right. \tag{1.3.17}$$

The analysis provided in this chapter covers different possible choices of the Leray-Lions operators $\mathbf{a}$. By taking the p-Laplacian operator, $\mathbf{a}(\boldsymbol{x}, \bar{u}, \xi) = |\xi|^{p-2}\xi$, we can study the Bulkley fluid model, which describes blood flows [85], food processing [54] and Bingham fluid flows [76]. The seepage model described in Section 1.1 is realised by taking $\mathbf{a}(\boldsymbol{x}, \bar{u}, \xi) = \Lambda(\boldsymbol{x}, \bar{u})\xi$, where $\Lambda : \Omega \times S_d(\mathbb{R})$ is a Carathéodory function.

For problems involving the Signorini and nonhomogeneous Dirichlet BCs, a gradient discretisation is made of a set of discrete unknowns $X_{\mathcal{D}} = X_{\mathcal{D},\Gamma_{2,3}} \oplus X_{\mathcal{D},\Gamma_1}$ (a direct sum of two finite dimensional spaces on $\mathbb{R}$), an interpolant operator for the trace $\mathcal{I}_{\mathcal{D},\Gamma_1} : W^{1-\frac{1}{p},p}(\partial\Omega) \to X_{\mathcal{D},\Gamma_1}$, a function reconstruction $\Pi_{\mathcal{D}} : X_{\mathcal{D}} \to W^{1,p}(\Omega)$, a trace reconstruction $\mathbb{T}_{\mathcal{D}} : X_{\mathcal{D}} \to L^p(\partial\Omega)$, and a gradient reconstruction $\nabla_{\mathcal{D}} : X_{\mathcal{D}} \to W^{1,p}(\Omega)^d$, which must be chosen to define a norm on $X_{\mathcal{D},\Gamma_{2,3}}$. Using this gradient discretisation in (1.3.17) yields the gradient scheme

$$\left\{\begin{array}{l} \text{Find } u \in \mathcal{K}_{\mathcal{D}} := \{v \in \mathcal{I}_{\mathcal{D},\Gamma_1} g + X_{\mathcal{D},\Gamma_{2,3}} \ : \ \mathbb{T}_{\mathcal{D}} v \leq a \text{ on } \Gamma_3\} \text{ such that } \forall v \in \mathcal{K}_{\mathcal{D}},\\ \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}} u(\boldsymbol{x}), \nabla_{\mathcal{D}} u(\boldsymbol{x})) \cdot \nabla_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \leq \int_\Omega f(\boldsymbol{x}) \Pi_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}. \end{array}\right. \tag{1.3.18}$$

The main results of this chapter is Theorem 4.2.9, which shows that, up to a subsequence of a gradient discretisation and approximate solutions $(\mathcal{D}_m, u_m)$, $\Pi_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)$ to an exact solution $\bar{u}$ to (1.3.17), and $\nabla_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)^d$ to $\nabla\bar{u}$. Proving this theorem via error estimate, as in linear problems, would be unsuccessful here since this technique requires the uniqueness of the solution and strong assumptions on the data; such assumptions are not valid in applications. In order to prove the convergence results, we follow the compactness technique, the idea of which is given in Section 1.2.3. The coercivity property (Definition 4.2.5) is required to obtain a bound on the approximate solution. This coercivity is still measured through the constant $C_{\mathcal{D}}$ defined by (1.3.4) where the $L^2(\Omega)$, $L^2(\Omega)^d$, $H^{1/2}(\partial\Omega)$ norms are respectively replaced with $L^p(\Omega)$ $L^p(\Omega)^d$, $L^p(\partial\Omega)$. Using such a bound and the limit-conformity of $(\mathcal{D}_m)_{m\in\mathbb{N}}$ (Definition 4.2.7) shows that there is a such sequence of $(\mathcal{D}_m)_{m\in\mathbb{N}}$ and $\bar{u} \in W^{1,p}(\Omega)$, such that along this subsequence $\Pi_{\mathcal{D}_m} u_m$ converges weakly to $\bar{u}$ in $L^p(\Omega)$, $\nabla_{\mathcal{D}_m} u_m$ converges weakly to $\nabla\bar{u}$ in $L^2(\Omega)^d$, and $\mathbb{T}_{\mathcal{D}_m} u_m$ converges weakly to $\gamma\bar{u}$ in $L^p(\partial\Omega)$, see [37, Lemma 2.57]. To obtain these weak limits, we note that it is sufficient to apply the limit-conformity indicator $W_{\mathcal{D}}$ to smooth functions. In our case, therefore, we restrict the definition of $W_{\mathcal{D}}$ to the functions in $C^2(\overline{\Omega})$: for all $\boldsymbol{\psi} \in C^2(\overline{\Omega})^d$, such that $\boldsymbol{\psi}\cdot\mathbf{n} = 0$ on Γ_2,

$$W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v\in X_{\mathcal{D},\Gamma_{2,3}}\setminus\{0\}} \frac{\left|\displaystyle\int_\Omega (\nabla_{\mathcal{D}} v\cdot\boldsymbol{\psi} + \Pi_{\mathcal{D}} v \,\mathrm{div}(\boldsymbol{\psi}))\,\mathrm{d}\boldsymbol{x} - \int_{\Gamma_3} \boldsymbol{\psi}\cdot\mathbf{n}\mathbb{T}_{\mathcal{D}} v\,\mathrm{d}\boldsymbol{x}\right|}{\|\nabla_{\mathcal{D}} v\|_{L^p(\Omega)^d}}.$$

This new definition is also the reason we can consider $\mathbb{T}_{\mathcal{D}}$ with values in $L^p(\Omega)$ but not $W^{1-\frac{1}{p},p}(\partial\Omega)$. We use the compactness property (Definition 4.2.8) to deal with the nonlinearity coming from the operator $\mathbf{a}$. This property shows that $\Pi_{\mathcal{D}_m} u_m$ actually converges strongly to $\bar{u}$ in $L^p(\Omega)$, up to a subsequence. The consistency property (Definition 4.2.6) enables us, for any $\varphi \in \mathcal{K}$, to find an interpolant $(v_m)_{m\in\mathbb{N}} \in \mathcal{K}_{\mathcal{D}_m}$ such that $\Pi_{\mathcal{D}_m} v_m \to \varphi$ in $L^p(\Omega)$ and $\nabla_{\mathcal{D}_m} v_m \to \nabla\varphi$ in $L^p(\Omega)^d$. These convergences, together with the Minty trick, allow us to pass to the limit in (1.3.18), written for $\Pi_{\mathcal{D}_m} u_m$, $\nabla_{\mathcal{D}_m} u_m$ and $\mathbb{T}_{\mathcal{D}_m} u_m$, to deduce that $\bar{u}$ is an exact solution to (1.3.1). If the operator $\mathbf{a}$ is assumed to be strictly monotone, the strong converge of $\nabla_{\mathcal{D}_m} u_m$ can be proved by following the same idea as in [40].

We also extend our analysis to the nonlinear obstacle problem and the Bulkley model. As an application of generic notions of GDM, we develop the HMM methods to three nonlinear models. The convergence of the resulting schemes are the consequences of Theorems 4.2.9, 4.3.9 and 4.3.10.

We present numerical results to illustrate the power of the HMM scheme in finding the location of the seepage point, even though the meshes are distorted. For the numerical solution, an iterative fixed point and the monotonicity algorithms are used at the same time to deal with the nonlinearity caused by the inequalities in the model, and by the nonlinear operator in the diffusion equation.

The results of Chapter 4 have been submitted for publication [2].

Chapter 5

The final chapter concerns linear parabolic variational inequalities (PVIs). First, we consider the following parabolic Signorini model:

$$\left\{\begin{aligned} \partial_t\bar{u} - \mathrm{div}(\Lambda\nabla\bar{u}) &= f && \text{in } \Omega\times(0,T),\\ \bar{u} &= 0 && \text{on } \Gamma_1\times(0,T),\\ \left.\begin{aligned}\bar{u} &\le a\\ \Lambda\nabla\bar{u}\cdot\mathbf{n} &\le 0\\ (a-\bar{u})\Lambda\nabla\bar{u}\cdot\mathbf{n} &= 0\end{aligned}\right\} & && \text{on } \Gamma_2\times(0,T),\\ \bar{u}(\boldsymbol{x},0) &= \bar{u}_{\mathrm{ini}} && \text{in } \Omega\times\{0\}.\end{aligned}\right. \tag{1.3.19}$$

Chapter 5 uses the GDM to provide a unified convergence analysis of numerical schemes for the PVIs. The time-dependent gradient discretisation $\mathcal{D}^T$ consists of the spatial gradient discretisation $(X_{\mathcal{D},\Gamma_2}, \Pi_{\mathcal{D}}, \mathbb{T}_{\mathcal{D}}, \nabla_{\mathcal{D}})$ as in the elliptic problems; an interpolation operator $J_{\mathcal{D}}$ to deal with the initial condition; and a discretisation of the time interval $(0,T)$. The gradient scheme for the weak formulation of (1.3.19) is to find a sequence $(u^{(n)})_{n=0,\dots,N} \subset \mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},\Gamma_2} : \mathbb{T}_{\mathcal{D}} v \le a \text{ on } \Gamma_2\}$ such that

$$\begin{cases} u^{(0)} = J_{\mathcal{D}}\bar{u}_{\rm ini}, \text{ and for all } v = (v^n)_{n=0,\dots,N} \subset \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\int_0^T\int_\Omega \delta_{\mathcal{D}}u(t)\,\Pi_{\mathcal{D}}(u-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \int_0^T\int_\Omega \Lambda(x)\nabla_{\mathcal{D}}u(\boldsymbol{x},t)\cdot\nabla_{\mathcal{D}}(u-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\ \displaystyle\le \int_0^T\int_\Omega f(\boldsymbol{x},t)\Pi_{\mathcal{D}}(u-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \end{cases} \tag{1.3.20}$$

Here $\delta_{\mathcal{D}}u$ denotes to the discrete time derivative function defined from $(0,T)$ to $L^2(\Omega)$, and $\Pi_{\mathcal{D}}$ and $\nabla_{\mathcal{D}}$ are extended as a space–time reconstruction; $\Pi_{\mathcal{D}}v : \Omega\times(0,T)\to\mathbb{R}$ and $\nabla_{\mathcal{D}}v : \Omega\times(0,T)\to\mathbb{R}^d$.

The main results are stated in Theorem 5.2.7, which gives the uniform-in-time convergence of $\Pi_{\mathcal{D}_m}u_m$, the strong convergence of $\nabla_{\mathcal{D}_m}u_m$, and the weak convergence of $\delta_{\mathcal{D}_m}u_m$ in time. We establish in Lemma 5.4.3 priori energy estimates on approximate gradients and discrete time derivatives of solutions. These estimates allow us to invoke the compactness technique as in Chapter 4. Passing to the limit in Scheme (1.3.20), written for sequences u_m and v_m in $\mathcal{K}_{\mathcal{D}_m}$, encounters a challenge in finding for a given $\varphi \in L^2(\Omega\times(0,T))$ such that $\varphi\le a$ on Γ_2, an interpolant $v_m \in \mathcal{K}_{\mathcal{D}_m}$ such that $\Pi_{\mathcal{D}_m}v_m$ converges to a φ and $\nabla_{\mathcal{D}_m}v_m \to \nabla\varphi$ in $L^2(\Omega\times(0,T))^d$. In parabolic PDE models, finding such an interpolant hinges on the fact that the set of tensorial functions in $C^\infty(\Omega\times(0,T))$ is dense in the space $L^2(0,T;H_0^1(\Omega))$. Using such a result would not however enable us to satisfy the condition $\mathbb{T}_{\mathcal{D}}\le a$ on Γ_2. To overcome this difficulty, we define a set of piecewise-constant in time functions that satisfy the barrier condition. For any $\varphi\in L^2(\Omega\times(0,T))$ such that $\varphi\le a$ on Γ_2, we can find a piecewise-constant in time function $(\bar{w}_\kappa)_{\kappa>0}$ such that $w_\kappa\le a$ on Γ_2 for all κ, and that converges to φ in $L^2(0,T;H^1(\Omega))$, as $\kappa\to 0$ (see Remark 5.2.2). We show in Lemma 5.4.2 that for any such piecewise-constant in time function $\bar{w}_\kappa$, there exists a sequence $(w_m)_{m\in\mathbb{N}}\in\mathcal{K}_{\mathcal{D}_m}$ such that for all $m\in\mathbb{N}$, $\Pi_{\mathcal{D}_m}w_m$ converges strongly to $\bar{w}_\kappa$ in $L^2(\Omega\times(0,T))$, $\nabla_{\mathcal{D}_m}w_m$ converges strongly to $\nabla\bar{w}_\kappa$ in $L^2(\Omega\times(0,T))$ and $\mathbb{T}_{\mathcal{D}_m}w_m$ converges strongly to $\gamma\bar{w}_\kappa$ in $L^2(\partial\Omega\times(0,T))$.

Following the same idea of the Signorini case, we develop the GDM to a time-dependent version of the obstacle problem, and state the convergence results in Theorem 5.3.4.

Based on notions of the GDM, we define the HMM method for both parabolic problems. Theorem 5.2.7 proves the convergence of the HMM scheme if the corresponding gradient discretisation $\mathcal{D}$ satisfies the three properties (coercivity, consistency and limit-conformity) and $||\nabla_{\mathcal{D}_m}J_{\mathcal{D}_m}\bar{u}_{\rm ini}||_{L^2(\Omega)^d}$ is bounded. Under classical assumptions on the mesh, we verify the three properties, and we construct an interpolant $J_{\mathcal{D}_m}$ such that it is bounded in the discrete norm $||\nabla_{\mathcal{D}}\cdot||_{L^2(\Omega)^d}$ and $\Pi_{\mathcal{D}_m}J_{\mathcal{D}_m}\bar{u}_{\rm ini}$ converges strongly to $\bar{u}_{\rm ini}$ in $L^2(\Omega)$.

We conduct numerical tests to assess the efficiency of the HMM method in solving parabolic variational inequalities. Numerically, at each time step $t^{(n)}$, we apply the monotonicity algorithm given in Chapter 3 to solve a number of systems of elliptic equations. The numerical results highlight that the number of iterations of the monotonicity algorithm is reduced along the time steps if we use the final sets at time $t^{(n)}$ as initial guesses for the monotonicity algorithm at the next time step $t^{(n+1)}$. This is due to the expected slow movement of the solution to the PVI between the time steps $t^{(n)}$ and $t^{(n+1)}$. Therefore, solving PVI could be less expensive than solving several disconnected elliptic VIs.

Chapter 2

Linear elliptic variational inequalities

Abstract. We show in this chapter that the gradient discretisation method can be extended to liner elliptic variational inequalities involving mixed Dirichlet, Neumann and Signorini boundary conditions. This extension allows us to provide error estimates for numerical approximations of such models, recovering known convergence rates for some methods, and establishing new convergence rates for schemes not previously studied for variational inequalities.

2.1 Introduction

The aim of this chapter is to develop a gradient discretisation method for elliptic linear variational inequalities with different types of boundary conditions.

Linear variational inequalities are used in the study of triple points in electrochemical reactions [69], of contact in linear elasticity [62], and of lubrication phenomena [21]. The study of linear models is also a first step towards a complete analysis of nonlinear models.

In this work, we consider two types of linear elliptic variational inequalities: the Signorini problem and the obstacle problem. The Signorini problem is formulated as

$$-\mathrm{div}(\Lambda\nabla\bar{u}) = f \quad \text{in } \Omega, \tag{2.1.1a}$$

$$\bar{u} = 0 \quad \text{on } \Gamma_1, \tag{2.1.1b}$$

$$\Lambda\nabla\bar{u}\cdot\mathbf{n} = 0 \quad \text{on } \Gamma_2, \tag{2.1.1c}$$

$$\left.\begin{array}{r} \bar{u} \leq a \\ \Lambda\nabla\bar{u}\cdot\mathbf{n} \leq 0 \\ \Lambda\nabla\bar{u}\cdot\mathbf{n}(a-\bar{u}) = 0 \end{array}\right\} \quad \text{on } \Gamma_3. \tag{2.1.1d}$$

The obstacle problem is

$$(\mathrm{div}(\Lambda\nabla\bar{u}) + f)(\psi - \bar{u}) = 0 \quad \text{in } \Omega, \tag{2.1.2a}$$

$$-\mathrm{div}(\Lambda\nabla\bar{u}) \leq f \quad \text{in } \Omega, \tag{2.1.2b}$$

$$\bar{u} \leq \psi \quad \text{in } \Omega, \tag{2.1.2c}$$

$$\bar{u} = 0 \quad \text{on } \partial\Omega. \tag{2.1.2d}$$

Here Ω is a bounded open set of $\mathbb{R}^d$ $(d \geq 1)$, $\mathbf{n}$ is the unit outer normal to $\partial\Omega$ and $(\Gamma_1, \Gamma_2, \Gamma_3)$ is a partition of $\partial\Omega$ (precise assumptions are stated in the next section).

Mathematical theories associated with the existence, uniqueness and stability of the solutions to variational inequalities have been extensively developed [45, 68, 57]. From the numerical point of view, different methods have been considered to approximate variational inequalities. For example, Bardet and Tobita [10] applied a finite difference scheme to the unconfined seepage problem. Extensions of the discontinuous Galerkin method to solve the obstacle problem can be found in [91, 31]. Although this method is still applicable when the functions are discontinuous along the elements boundaries, the exact solution must be in the space H^2 to ensure the consistency and the convergence. Numerical verification methods, which aim at finding a set in which a solution exists, have been developed for a few variational inequality problems. In particular we cite the obstacle problems, the Signorini problem and elasto-plastic problems (see [84] and references therein).

Falk [50] was the first to provide a general error estimate for the approximation by *conforming* methods of the obstacle problem. This estimate showed that the $\mathbb{P}1$ finite element error is $\mathcal{O}(h)$. Yang and Chao [88] showed that the convergence rate of *non-conforming* finite elements method for the Signorini problem has order one, under an $H^{5/2}$-regularity assumption on the solution.

Herbin and Marchand [61] showed that if $\Lambda \equiv \mathbf{Id}$ and if the grid satisfies the orthogonality condition required by the two-point flux approximation (TPFA) finite volume method, then for both problems the solutions provided by this scheme converge in $L^2(\Omega)$ to the unique solution as the mesh size tends to zero.

The gradient discretisation method framework enables us to design a unified convergence analysis of many numerical methods for linear variational inequalities; using the theorems stemming from this analysis, we recover known estimates for some schemes, and we establish new estimates for methods that have not, to our best knowledge, been previously studied for variational inequalities.

This chapter is organised as follows. In Section 2.2, we present the gradient discretisation method framework for variational inequalities, and we state our main error estimates. We then show in Section 2.3 an example of a method that is contained into this framework, therefore satisfying the error estimates established in Section

2.3. Section 2.4 is devoted to the proof of the main results. We present in Section 2.5 an extension of our results to the case of approximate barriers, that is natural in many schemes in which the exact barrier is replaced with an interpolant depending on the approximation space. The chapter is completed by an appendix recalling the proof of equivalence between the strong and the weak formulations, and the definition of the normal trace of H_{div} functions and the.

2.2 Assumptions and main results

2.2.1 Weak formulations

Let us start by stating our assumptions and weak formulations for the Signorini and the obstacle problems.

Hypothesis 2.2.1 (Signorini problem). *We make the following assumptions on the data in* (2.1.1)*:*

1. Ω *is an open bounded connected subset of* $\mathbb{R}^d$*,* $d \geq 1$ *and* Ω *has a Lipschitz boundary,*
2. Λ *is a measurable function from* Ω *to* $\mathrm{M}_d(\mathbb{R})$ *(where* $\mathrm{M}_d(\mathbb{R})$ *is the set of* $d \times d$ *matrices) and there exists* $\underline{\lambda}$, $\overline{\lambda} > 0$ *such that, for a.e.* $x \in \Omega$, $\Lambda(\boldsymbol{x})$ *is symmetric with eigenvalues in* $[\underline{\lambda}, \overline{\lambda}]$*,*
3. Γ_1, Γ_2 *and* Γ_3 *are measurable pairwise disjoint subsets of* $\partial\Omega$ *such that* $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3 = \partial\Omega$ *and* Γ_1 *has a positive* $(d-1)$*-dimensional measure,*
4. $f \in L^2(\Omega)$, $a \in L^2(\partial\Omega)$.

Under Hypothesis 2.2.1, the weak formulation of Problem (2.1.1) is

$$\left\{\begin{array}{l} \text{Find } \bar{u} \in \mathcal{K} := \{v \in H^1(\Omega) \ : \ \gamma(v) = 0 \text{ on } \Gamma_1,\ \gamma(v) \leq a \text{ on } \Gamma_3 \text{ a.e.}\} \ \text{ s.t.,} \\ \forall v \in \mathcal{K}\,, \quad \displaystyle\int_\Omega \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x}) \cdot \nabla(\bar{u} - v)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \int_\Omega f(\boldsymbol{x})(\bar{u}(\boldsymbol{x}) - v(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}\,, \end{array}\right. \tag{2.2.1}$$

where $\gamma : H^1(\Omega) \mapsto H^{1/2}(\partial\Omega)$ is the trace operator. We refer the reader to Proposition 2.A.1 for the proof of equivalence between the strong and the weak formulations. It was shown in [45] that, if $\mathcal{K}$ is not empty (which is the case, for example, if $a \geq 0$ on $\partial\Omega$) then there exists a unique solution to Problem (2.2.1). In the sequel, we will assume that the barrier a is such that $\mathcal{K}$ is not empty.

Hypothesis 2.2.2 (Obstacle problem). *Our assumptions on the data in* (2.1.2) *are:*

1. Ω *and* Λ *satisfy (1) and (2) in Hypothesis 2.2.1,*
2. $f \in L^2(\Omega)$, $\psi \in L^2(\Omega)$.

Under Hypotheses 2.2.2, we consider the obstacle problem (2.1.2) under the following weak form:

$$\left\{\begin{array}{l} \text{Find } \bar{u} \in \mathcal{K} := \{v \in H_0^1(\Omega) \ : \ v \leq \psi \text{ in } \Omega\} \ \text{ such that, for all } v \in \mathcal{K}, \\ \displaystyle\int_\Omega \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x}) \cdot \nabla(\bar{u} - v)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \int_\Omega f(\boldsymbol{x})(\bar{u}(\boldsymbol{x}) - v(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}. \end{array}\right. \tag{2.2.2}$$

It can be seen [93] that if $\bar{u} \in C^2(\overline{\Omega})$ and Λ is Lipschitz continuous, then (2.1.2) and (2.2.2) are indeed equivalent, see Proposition 2.A.2 for the details. The proof of this equivalence can be easily adapted to the case where the solution belongs to $H^2(\Omega)$. If ψ is such that $\mathcal{K}$ is not empty, which we assume from here on, then (2.2.2) has a unique solution.

2.2.2 Construction of gradient discretisation method

Gradient schemes provide a general formulation of different numerical methods. Each gradient scheme is based on a gradient discretisation, which is a set of discrete space and operators used to discretise the weak formulation of the problem under study. Actually, a gradient scheme consists of replacing the continuous space and operators used in the weak formulations by the discrete counterparts provided by a gradient discretisation. In this part, we define gradient discretisations for the Signorini and the obstacle problems, and we list the properties that are required of a gradient discretisation to give rise to a converging gradient scheme.

2.2.2.1 Signorini problem

Definition 2.2.3 (Gradient discretisation for the Signorini problem). A gradient discretisation $\mathcal{D}$ for Problem (2.2.1) is defined by $\mathcal{D} = (X_{\mathcal{D},\Gamma_{2,3}}, \Pi_{\mathcal{D}}, \mathbb{T}_{\mathcal{D}}, \nabla_{\mathcal{D}})$, where

1. the set of discrete unknowns $X_{\mathcal{D},\Gamma_{2,3}}$ is a finite dimensional vector space on $\mathbb{R}$, taking into account the homogeneous boundary conditions on Γ_1,
2. the linear mapping $\Pi_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to L^2(\Omega)$ is the reconstructed function,
3. the linear mapping $\mathbb{T}_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to H^{1/2}_{\Gamma_1}(\partial\Omega)$ is the reconstructed trace, where $H^{1/2}_{\Gamma_1}(\partial\Omega) = \{v \in H^{1/2}(\partial\Omega) : v = 0 \text{ on } \Gamma_1\}$,
4. the linear mapping $\nabla_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to L^2(\Omega)^d$ is a reconstructed gradient, which must be defined such that $\|\nabla_{\mathcal{D}} \cdot \|_{L^2(\Omega)^d}$ is a norm on $X_{\mathcal{D},\Gamma_{2,3}}$.

As explained above, a gradient scheme for (2.2.1) is obtained by simply replacing in the weak formulation of the problem the continuous space and operators by the discrete space and operators coming from a gradient discretisation.

Definition 2.2.4 (Gradient scheme for the Signorini problem). Let $\mathcal{D}$ be the gradient discretisation in the sense of Definition 2.2.3. The gradient scheme for Problem (2.2.1) is

$$\left\{ \begin{array}{l} \text{Find } u \in \mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},\Gamma_{2,3}} \; : \; \mathbb{T}_{\mathcal{D}} v \le a \text{ on } \Gamma_3\} \text{ such that, } \forall v \in \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\int_\Omega \Lambda(\boldsymbol{x}) \nabla_{\mathcal{D}} u(\boldsymbol{x}) \cdot \nabla_{\mathcal{D}}(u-v)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \le \int_\Omega f(\boldsymbol{x}) \Pi_{\mathcal{D}}(u-v)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \end{array} \right. \tag{2.2.3}$$

The accuracy of a gradient discretisation is measured through three indicators, related to its *coercivity*, *consistency* and *limit-conformity*. The good behaviour of these indicators along a sequence of gradient discretisations ensures that the solutions to the corresponding gradient schemes converge towards the solution to the continuous problem.

To measure the coercivity of a gradient discretisation $\mathcal{D}$ in the sense of Definition 2.2.3, we define the norm $C_{\mathcal{D}}$ of the linear mapping $\Pi_{\mathcal{D}}$ by

$$C_{\mathcal{D}} = \max_{v \in X_{\mathcal{D},\Gamma_{2,3}} \setminus \{0\}} \Big(\frac{\|\Pi_{\mathcal{D}} v\|_{L^2(\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}} + \frac{\|\mathbb{T}_{\mathcal{D}} v\|_{H^{1/2}(\partial\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}} \Big). \tag{2.2.4}$$

The consistency of a gradient discretisation $\mathcal{D}$ in the sense of Definition 2.2.3 is measured by $S_{\mathcal{D}} : \mathcal{K} \times \mathcal{K}_{\mathcal{D}} \to [0, +\infty)$ defined by

$$\forall (\varphi, v) \in \mathcal{K} \times \mathcal{K}_{\mathcal{D}}, \; S_{\mathcal{D}}(\varphi, v) = \|\Pi_{\mathcal{D}} v - \varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla\varphi\|_{L^2(\Omega)^d}. \tag{2.2.5}$$

To measure the limit-conformity of the gradient discretisation $\mathcal{D}$ in the sense of Definition 2.2.3, we introduce $W_{\mathcal{D}} : H_{\rm div}(\Omega) \to [0,+\infty)$ defined by

$$\forall \boldsymbol{\psi} \in H_{\rm div}(\Omega),$$
$$W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v\in X_{\mathcal{D},\Gamma_{2,3}}\setminus\{0\}} \frac{1}{\|\nabla_{\mathcal{D}}v\|_{L^2(\Omega)^d}} \Big| \int_\Omega (\nabla_{\mathcal{D}}v\cdot\boldsymbol{\psi} + \Pi_{\mathcal{D}}v\,{\rm div}(\boldsymbol{\psi}))\,{\rm d}\boldsymbol{x} - \langle \gamma_{\mathbf{n}}(\boldsymbol{\psi}), \mathbb{T}_{\mathcal{D}}v\rangle \Big|, \tag{2.2.6}$$

where $H_{\rm div}(\Omega) = \{\boldsymbol{\psi} \in L^2(\Omega)^d : {\rm div}\boldsymbol{\psi} \in L^2(\Omega)\}$. We refer the reader to Appendix 2.B for the definitions of the normal trace $\gamma_{\mathbf{n}}$ on $H_{\rm div}(\Omega)$, and of the duality product $\langle\cdot,\cdot\rangle$ between $(H^{1/2}(\partial\Omega))'$ and $H^{1/2}(\partial\Omega)$.

Remark 2.2.1. The definition (2.2.4) of $C_{\mathcal{D}}$ does not only include the norm of $\Pi_{\mathcal{D}}$, as in the obstacle problem case detailed below, but also quite naturally the norm of the other reconstruction operator $\mathbb{T}_{\mathcal{D}}$.

The definition (2.2.6) takes into account the nonzero boundary conditions on a part of $\partial\Omega$. This was already noted in the case of gradient discretisations adapted for PDEs with mixed boundary conditions; however, a major difference must be raised here. $W_{\mathcal{D}}$ will be applied to $\boldsymbol{\psi} = \Lambda\nabla\bar{u}$. For PDEs [40, 48, 37] the boundary conditions ensure that $\gamma_{\mathbf{n}}(\boldsymbol{\psi}) \in L^2(\partial\Omega)$ and thus that $\langle\gamma_{\mathbf{n}}(\boldsymbol{\psi}), \mathbb{T}_{\mathcal{D}}v\rangle$ can be replaced with $\int_{\Gamma_3}\gamma_{\mathbf{n}}(\boldsymbol{\psi})\mathbb{T}_{\mathcal{D}}v\,{\rm d}\boldsymbol{x}$ in $W_{\mathcal{D}}$. Hence, in the study of gradient schemes for PDEs with mixed boundary conditions, $\mathbb{T}_{\mathcal{D}}v$ only needs to be in $L^2(\partial\Omega)$. For the Signorini problem we cannot ensure that $\gamma_{\mathbf{n}}(\boldsymbol{\psi}) \in L^2(\partial\Omega)$; we only know that $\gamma_{\mathbf{n}}(\boldsymbol{\psi}) \in H^{1/2}(\partial\Omega)$. The definition of $\mathbb{T}_{\mathcal{D}}$ therefore needs to be changed to ensure that this reconstructed trace takes values in $H^{1/2}(\partial\Omega)$ instead of $L^2(\partial\Omega)$.

2.2.2.2 Obstacle problem

The definition of a gradient discretisation for the obstacle problem is not different from the definition of a gradient discretisation for elliptic PDEs with homogeneous Dirichlet conditions [40, 48].

Definition 2.2.5 (Gradient discretisation for the obstacle problem). A gradient discretisation $\mathcal{D}$ for homogeneous Dirichlet boundary conditions is defined by a triplet $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$, where

1. the set of discrete unknowns $X_{\mathcal{D},0}$ is a finite dimensional vector space over $\mathbb{R}$, taking into account the boundary condition (2.1.2d).
2. the linear mapping $\Pi_{\mathcal{D}} : X_{\mathcal{D},0} \to L^2(\Omega)$ gives the reconstructed function,
3. the linear mapping $\nabla_{\mathcal{D}} : X_{\mathcal{D},0} \to L^2(\Omega)^d$ gives a reconstructed gradient, which must be defined such that $\|\nabla_{\mathcal{D}}\cdot\|_{L^2(\Omega)^d}$ is a norm on $X_{\mathcal{D},0}$.

Definition 2.2.6 (Gradient scheme for the obstacle problem). Let $\mathcal{D}$ be a gradient discretisation in the sense of Definition 2.2.5. The corresponding gradient scheme for (2.2.2) is given by

$$\left\{ \begin{array}{l} \text{Find } u \in \mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},0} \ : \ \Pi_{\mathcal{D}}v \le \psi \text{ in } \Omega\} \ \text{ such that, } \forall v \in \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\int_\Omega \Lambda(\boldsymbol{x})\nabla_{\mathcal{D}}u(\boldsymbol{x})\cdot\nabla_{\mathcal{D}}(u-v)(\boldsymbol{x})\,{\rm d}\boldsymbol{x} \le \int_\Omega f(\boldsymbol{x})\Pi_{\mathcal{D}}(u-v)(\boldsymbol{x})\,{\rm d}\boldsymbol{x}. \end{array}\right. \tag{2.2.7}$$

The coercivity, consistency and limit-conformity of a gradient discretisation in the sense of Definition 2.2.5 are defined through the following constant and functions:

$$C_{\mathcal{D}} = \max_{v\in X_{\mathcal{D},0}\setminus\{0\}} \frac{\|\Pi_{\mathcal{D}}v\|_{L^2(\Omega)}}{\|\nabla_{\mathcal{D}}v\|_{L^2(\Omega)^d}}, \tag{2.2.8}$$

$$\forall(\varphi, v) \in \mathcal{K}\times\mathcal{K}_{\mathcal{D}},\ S_{\mathcal{D}}(\varphi, v) = \|\Pi_{\mathcal{D}}v - \varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}}v - \nabla\varphi\|_{L^2(\Omega)^d}, \tag{2.2.9}$$

and

$$\forall \boldsymbol{\psi} \in H_{\mathrm{div}}(\Omega),\ W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v\in X_{\mathcal{D},0}\setminus\{0\}} \frac{1}{\|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d}} \Big| \int_\Omega (\nabla_{\mathcal{D}} v \cdot \boldsymbol{\psi} + \Pi_{\mathcal{D}} v \,\mathrm{div}(\boldsymbol{\psi}))\,\mathrm{d}\boldsymbol{x}\Big|. \tag{2.2.10}$$

The only indicator that changes with respect to [40, 48] is $S_{\mathcal{D}}$. For PDEs, the consistency requires to consider $S_{\mathcal{D}}$ on $H_0^1(\Omega) \times X_{\mathcal{D},0}$ and to ensure that $\inf_{v\in X_{\mathcal{D}_m}} S_{\mathcal{D}_m}(\varphi, v) \to 0$ as $m \to \infty$ for any $\varphi \in H_0^1(\Omega)$. Here, the domain of $S_{\mathcal{D}}$ is adjusted to the set to which the solution of the variational inequality belongs (namely $\mathcal{K}$), and to the set in which we can pick the test functions of the gradient scheme (namely $\mathcal{K}_{\mathcal{D}}$).

We note that this double reduction does not necessarily facilitate, with respect to the case of PDEs, the proof of the consistency of the sequence of gradient discretisations. In practice, however, the proof developed for the PDE case also provides the consistency of gradient discretisations for variational inequalities.

2.2.3 Error estimates

We present here our main error estimates for the gradient schemes approximations of Problems (2.2.1) and (2.2.2).

2.2.3.1 Signorini problem

Theorem 2.2.7 (Error estimate for the Signorini problem). *Under Hypothesis 2.2.1, let $\bar{u} \in \mathcal{K}$ be the solution to Problem (2.2.1). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 2.2.3 and $\mathcal{K}_{\mathcal{D}}$ is nonempty, then there exists a unique solution $u \in \mathcal{K}_{\mathcal{D}}$ to the gradient scheme (2.2.3). Furthermore, this solution satisfies the following inequalities, for any $v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}}$:*

$$\|\nabla_{\mathcal{D}} u - \nabla \bar{u}\|_{L^2(\Omega)^d} \le \sqrt{\frac{2}{\underline{\lambda}} G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+} + \frac{1}{\underline{\lambda}}[W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + (\overline{\lambda} + \underline{\lambda})S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})], \tag{2.2.11}$$

and

$$\|\Pi_{\mathcal{D}} u - \bar{u}\|_{L^2(\Omega)} \le C_{\mathcal{D}}\sqrt{\frac{2}{\underline{\lambda}} G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+} + \frac{1}{\underline{\lambda}}[C_{\mathcal{D}} W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + (C_{\mathcal{D}}\overline{\lambda} + \underline{\lambda})S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})], \tag{2.2.12}$$

where

$$G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = \langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), \mathbb{T}_{\mathcal{D}} v_{\mathcal{D}} - \gamma\bar{u}\rangle \quad \text{and} \quad G^+_{\mathcal{D}} = \max(0, G_{\mathcal{D}}). \tag{2.2.13}$$

Here, the quantities $C_{\mathcal{D}}$, $S_{\mathcal{D}}(\bar{u},\, v_{\mathcal{D}})$ and $W_{\mathcal{D}}$ are defined by (2.2.4), (2.2.5) *and* (2.2.6)*, respectively.*

2.2.3.2 Obstacle problem

We give in [1, Theorem 1] an error estimate for gradient schemes for the obstacle problem. For low-order methods with piecewise constant approximations, such as the HMM schemes, this theorem provides an $\mathcal{O}(\sqrt{h})$ rate of convergence for the function and the gradient (h is the mesh size). The following theorem improves [1, Theorem 1] by introducing the free choice of interpolant $v_{\mathcal{D}}$. This enables us, in Section 2.2.4.2, to establish much better rates of convergence –namely $\mathcal{O}(h)$ for HMM, for example.

Theorem 2.2.8 (Error estimate for the obstacle problem). *Under Hypothesis 2.2.2, let $\bar{u} \in \mathcal{K}$ be the solution to Problem (2.2.2). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 2.2.5 and if $\mathcal{K}_{\mathcal{D}}$ is nonempty, then there exists a unique solution $u \in \mathcal{K}_{\mathcal{D}}$ to the gradient scheme (2.2.7). Moreover, if* $\mathrm{div}(\Lambda\nabla\bar{u}) \in L^2(\Omega)$ *then this solution satisfies the following estimates, for any $v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}}$:*

$$\|\nabla_{\mathcal{D}} u - \nabla \bar{u}\|_{L^2(\Omega)^d} \le \sqrt{\frac{2}{\underline{\lambda}} E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+} + \frac{1}{\underline{\lambda}}[W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + (\overline{\lambda} + \underline{\lambda})S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})], \tag{2.2.14}$$

$$\|\Pi_{\mathcal{D}} u - \bar{u}\|_{L^2(\Omega)} \le C_{\mathcal{D}} \sqrt{\frac{2}{\underline{\lambda}} E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+} + \frac{1}{\underline{\lambda}} [C_{\mathcal{D}} W_{\mathcal{D}}(\Lambda \nabla \bar{u}) + (C_{\mathcal{D}} \overline{\lambda} + \underline{\lambda}) S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})], \tag{2.2.15}$$

where

$$E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = \int_\Omega (\mathrm{div}(\Lambda \nabla \bar{u}) + f)(\bar{u} - \Pi_{\mathcal{D}} v_{\mathcal{D}}) \,\mathrm{d}x \quad \text{and} \quad E_{\mathcal{D}}^+ = \max(0, E_{\mathcal{D}}). \tag{2.2.16}$$

Here, the quantities $C_{\mathcal{D}}$, $S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$ and $W_{\mathcal{D}}$ are defined by (2.2.8)*,* (2.2.9) *and* (2.2.10)*, respectively.*

Remark 2.2.2. We note that, if Λ is Lipschitz-continuous, the assumption $\mathrm{div}(\Lambda \nabla \bar{u}) \in L^2(\Omega)$ is reasonable given the H^2-regularity result on $\bar{u}$ of [16].

Compared with the previous general estimates, such as Falk's Theorem [50], our estimates seem to be simpler since they only involve one choice of interpolant $v_{\mathcal{D}}$ whereas Falk's estimate depends on choices of $v_h \in K_h$ and $v \in K$. Yet, our estimates provide the same final orders of convergence as Falk's estimate. We also note that Estimates (2.2.11), (2.2.12), (2.2.14) and (2.2.15) are applicable to *conforming* and *non-conforming* methods, whereas the estimates in [50] seem to be applicable only to *conforming* methods.

2.2.4 Orders of convergence

In what follows, we consider polytopal open sets Ω and gradient discretisations based on polytopal meshes of Ω, as defined in [41, 37]. The following definition is a simplification (it does not include the vertices) of the definition in [41], that will however be sufficient to our purpose.

Definition 2.2.9 (Polytopal mesh). Let Ω be a bounded polytopal open subset of $\mathbb{R}^d$ ($d \ge 1$). A polytopal mesh of Ω is given by $\mathcal{T} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:

1. $\mathcal{M}$ is a finite family of non empty connected polytopal open disjoint subsets of Ω (the cells) such that $\overline{\Omega} = \cup_{K \in \mathcal{M}} \overline{K}$. For any $K \in \mathcal{M}$, $|K| > 0$ is the measure of K and h_K denotes the diameter of K.
2. $\mathcal{E}$ is a finite family of disjoint subsets of $\overline{\Omega}$ (the edges of the mesh in 2D, the faces in 3D), such that any $\sigma \in \mathcal{E}$ is a non empty open subset of a hyperplane of $\mathbb{R}^d$ and $\sigma \subset \overline{\Omega}$. We assume that for all $K \in \mathcal{M}$ there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \cup_{\sigma \in \mathcal{E}_K} \overline{\sigma}$. We then denote by $\mathcal{M}_\sigma = \{K \in \mathcal{M} : \sigma \in \mathcal{E}_K\}$. We then assume that, for all $\sigma \in \mathcal{E}$, $\mathcal{M}_\sigma$ has exactly one element and $\sigma \subset \partial\Omega$, or $\mathcal{M}_\sigma$ has two elements and $\sigma \subset \Omega$. We let $\mathcal{E}_{\mathrm{int}}$ be the set of all interior faces, i.e. $\sigma \in \mathcal{E}$ such that $\sigma \subset \Omega$, and $\mathcal{E}_{\mathrm{ext}}$ the set of boundary faces, i.e. $\sigma \in \mathcal{E}$ such that $\sigma \subset \partial\Omega$. For $\sigma \in \mathcal{E}$, the $(d-1)$-dimensional measure of σ is $|\sigma|$, the centre of gravity of σ is $\overline{\boldsymbol{x}}_\sigma$, and the diameter of σ is h_σ.
3. $\mathcal{P} = (\boldsymbol{x}_K)_{K \in \mathcal{M}}$ is a family of points of Ω indexed by $\mathcal{M}$ and such that, for all $K \in \mathcal{M}$, $\boldsymbol{x}_K \in K$ ($\boldsymbol{x}_K$ is sometimes called the "centre" of K). We then assume that all cells $K \in \mathcal{M}$ are strictly $\boldsymbol{x}_K$-star-shaped, meaning that if $\boldsymbol{x}$ is in the closure of K then the line segment $[\boldsymbol{x}_K, \boldsymbol{x})$ is included in K.

For a given $K \in \mathcal{M}$, let $\mathbf{n}_{K,\sigma}$ be the unit vector normal to σ outward to K and denote by $d_{K,\sigma}$ the orthogonal distance between $\boldsymbol{x}_K$ and $\sigma \in \mathcal{E}_K$. The size of the discretisation is $h_{\mathcal{M}} = \sup\{h_K : K \in \mathcal{M}\}$.

For most gradient discretisations for PDEs based on first order methods, including HMM methods and *conforming* or *non-conforming* finite elements methods, explicit estimates on $S_{\mathcal{D}}$ and $W_{\mathcal{D}}$ can be established [37]. The proofs of these estimates are easily transferable to the above setting of gradient discretisations for variational inequalities, and give

$$W_{\mathcal{D}}(\boldsymbol{\psi}) \le C h_{\mathcal{M}} \|\boldsymbol{\psi}\|_{H^1(\Omega)^d}, \quad \forall \boldsymbol{\psi} \in H^1(\Omega)^d, \tag{2.2.17}$$

$$\inf_{v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}}} S_{\mathcal{D}}(\varphi, v_{\mathcal{D}}) \le C h_{\mathcal{M}} \|\varphi\|_{H^2(\Omega)}, \quad \forall \varphi \in H^2(\Omega) \cap \mathcal{K}. \tag{2.2.18}$$

Hence, if Λ is Lipschitz-continuous and $\bar{u} \in H^2(\Omega)$, then the terms involving $W_\mathcal{D}$ and $S_\mathcal{D}$ in (2.2.11), (2.2.12), (2.2.14) and (2.2.15) are of order $\mathcal{O}(h_\mathcal{M})$, provided that $v_\mathcal{D}$ is chosen to optimise $S_\mathcal{D}$. Since the terms $G_\mathcal{D}$ and $E_\mathcal{D}$ can be bounded above by $C||\gamma(\bar{u}) - \mathbb{T}_\mathcal{D} v_\mathcal{D}||_{L^2(\partial\Omega)}$ and $C||\bar{u} - \Pi_\mathcal{D} v_\mathcal{D}||_{L^2(\Omega)}$, respectively, we expect these to be of order $h_\mathcal{M}$. Hence, the dominating terms in the error estimates are $\sqrt{G_\mathcal{D}^+}$ and $\sqrt{E_\mathcal{D}^+}$. In first approximation, these terms seem to behave as $\sqrt{h_\mathcal{M}}$ for a first order *conforming* or *non-conforming* method. This initial brute estimate can however usually be improved, as we will show in Theorems 2.2.10 and 2.2.11, even for *non-conforming* methods based on piecewise constant reconstructed functions, and lead to the expected $\mathcal{O}(h_\mathcal{M})$ global convergence rate.

2.2.4.1 Signorini problem

For a first order *conforming* numerical method, the $\mathbb{P}1$ finite element method for example, if $\bar{u}$ is in $H^2(\Omega)$, then the classical interpolant $v_\mathcal{D} \in X_{\mathcal{D},\Gamma_{2,3}}$ constructed from the values of $\bar{u}$ at the vertices satisfies [17]

$$\begin{aligned}
\|\Pi_\mathcal{D} v_\mathcal{D} - \bar{u}\|_{L^2(\Omega)} &\le C h_\mathcal{M}^2 \|D^2 \bar{u}\|_{L^2(\Omega)^{d\times d}},\\
\|\nabla_\mathcal{D} v_\mathcal{D} - \nabla \bar{u}\|_{L^2(\Omega)^d} &\le C h_\mathcal{M} \|D^2 \bar{u}\|_{L^2(\Omega)^{d\times d}},\\
||\mathbb{T}_\mathcal{D} v_\mathcal{D} - \gamma(\bar{u})||_{L^2(\partial\Omega)} &\le C h_\mathcal{M}^2 \|D^2 \gamma(\bar{u})\|_{L^2(\partial\Omega)^{d\times d}}.
\end{aligned}$$

If a is piecewise affine on the mesh then this interpolant $v_\mathcal{D}$ lies in $\mathcal{K}_\mathcal{D}$ and can therefore be used in Theorem 2.2.7. We then have $S_\mathcal{D}(\bar{u}, v_\mathcal{D}) = \mathcal{O}(h_\mathcal{M})$ and $G_\mathcal{D}(\bar{u}, v_\mathcal{D}) = \mathcal{O}(h_\mathcal{M}^2)$, and (2.2.11)–(2.2.12) give an order one error estimate on the H^1 norm. If a is not linear, the definition of $\mathcal{K}_\mathcal{D}$ is usually relaxed, see Section 2.5.

A number of low-order methods have piecewise constant approximations of the solution, e.g. finite volume methods or finite element methods with mass lumping. For those, there is no hope of finding an interpolant which gives an order $h_\mathcal{M}^2$ approximation of $\bar{u}$ in $L^2(\Omega)$ norm. We can however prove, for such methods, that $\sqrt{G_\mathcal{D}^+}$ behaves better than the expected $\sqrt{h_\mathcal{M}}$ order. In the following theorem, we denote by $W^{2,\infty}(\partial\Omega)$ the functions v on $\partial\Omega$ such that, for any face F of $\partial\Omega$, $v_{|F} \in W^{2,\infty}(F)$.

Theorem 2.2.10 (Signorini problem: order of convergence for non-conforming reconstructions). *Under Hypothesis 2.2.1, let $\mathcal{D}$ be a gradient discretisation in the sense of Definition 2.2.3, such that $\mathcal{K}_\mathcal{D} \neq \emptyset$, and let $\mathcal{T} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be a polytopal mesh of Ω. Let $\bar{u}$ and u be the respective solutions to Problems* (2.2.1) *and* (2.2.3). *We assume that the barrier a is constant, that $\gamma_\mathbf{n}(\Lambda\nabla\bar{u}) \in L^2(\partial\Omega)$ and that $\gamma(\bar{u}) \in W^{2,\infty}(\partial\Omega)$. We also assume that there exists an interpolant $v_\mathcal{D} \in \mathcal{K}_\mathcal{D}$ such that $S_\mathcal{D}(\bar{u}, v_\mathcal{D}) \le C_1 h_\mathcal{M}$ and that, for any $\sigma \in \mathcal{E}_{\rm ext}$, $||\mathbb{T}_\mathcal{D} v_\mathcal{D} - \gamma(\bar{u})(\boldsymbol{x}_\sigma)||_{L^2(\sigma)} \le C_2 h_\sigma^2 |\sigma|^{1/2}$, where $\boldsymbol{x}_\sigma \in \sigma$ (here, the constants C_i do not depend on the edge or the mesh). Then, there exists C depending only on Ω, Λ, $\bar{u}$, a, C_1, C_2 and an upper bound of $C_\mathcal{D}$ such that*

$$\|\nabla_\mathcal{D} u - \nabla\bar{u}\|_{L^2(\Omega)^d} + \|\Pi_\mathcal{D} u - \bar{u}\|_{L^2(\Omega)} \le C h_\mathcal{M} + C W_\mathcal{D}(\Lambda\nabla\bar{u}). \tag{2.2.19}$$

Remark 2.2.3. Since $\gamma_\mathbf{n}(\Lambda\nabla\bar{u}) \in L^2(\partial\Omega)$ the reconstructed trace $\mathbb{T}_\mathcal{D}$ can be taken with values in $L^2(\partial\Omega)$ (see the discussion at the end of Section 2.2.2.1). In particular, $\mathbb{T}_\mathcal{D} v$ can be piecewise constant. The $H^{1/2}(\partial\Omega)$ norm in the definition (2.2.4) of $C_\mathcal{D}$ is then replaced with the $L^2(\partial\Omega)$ norm, and the duality product in (2.2.6) is replaced with a plain integral on $\partial\Omega$.

For the two-point finite volume method, an error estimate similar to (2.2.19) is stated in [61], under the assumption that the solution is in $H^2(\Omega)$. It however seems to us that the proof in [61] uses an estimate which requires $\gamma(\bar{u}) \in W^{2,\infty}(\partial\Omega)$, as in Theorem 2.2.10.

We notice that the assumption that a is constant is compatible with some models, such as an electrochemical reaction and friction mechanics [55]. This condition on a can be relaxed or, as detailed in Section 2.5, a can be approximated by a simpler barrier – the definition of which depends on the considered scheme.

2.2.4.2 Obstacle problem

As explained in the introduction to this section, to estimate the order of convergence for the obstacle problem we only need to estimate $E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$. This can be readily done for $\mathbb{P}_1$ finite elements, for example. If ψ is linear or constant, letting $v_{\mathcal{D}} = \bar{u}$ at all vertices (as in [25]) shows that $v_{\mathcal{D}}$ is an element of $\mathcal{K}_{\mathcal{D}}$, and that $S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = \mathcal{O}(h_{\mathcal{M}})$ and $E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = \mathcal{O}(h_{\mathcal{M}}^2)$. Therefore, Theorem 2.2.8 provides an order one error estimate.

The following theorem shows that, as for the Signorini problem, the error estimate for the obstacle problem is of order one even for methods with piecewise constant reconstructed functions.

Theorem 2.2.11 (**Obstacle problem: order of convergence for non-conforming reconstructions).** *Under Hypothesis 2.2.2, let $\mathcal{D}$ be a gradient discretisation in the sense of Definition 2.2.5, such that $\mathcal{K}_{\mathcal{D}} \neq \emptyset$, and let $\mathcal{T} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be a polytopal mesh of Ω. Let $\bar{u}$ and u be the respective solutions to Problems* (2.2.2) *and* (2.2.7)*. We assume that ψ is constant, that* $\mathrm{div}(\Lambda\nabla\bar{u}) \in L^2(\Omega)$*, and that* $\bar{u} \in W^{2,\infty}(\Omega)$*. We also assume that there exists an interpolant $v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}}$ such that $S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) \leq C_1 h_{\mathcal{M}}$ and, for any $K \in \mathcal{M}$, $||\Pi_{\mathcal{D}} v_{\mathcal{D}} - \bar{u}(\boldsymbol{x}_K)||_{L^2(K)} \leq C_2 h_K^2 |K|^{1/2}$ (here, the various constants C_i do not depend on the cell or the mesh). Then, there exists C depending only on Ω, $\underline{\lambda}$, $\overline{\lambda}$, C_1, C_2 and an upper bound of $C_{\mathcal{D}}$ such that*

$$\|\nabla_{\mathcal{D}} u - \nabla\bar{u}\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}} u - \bar{u}\|_{L^2(\Omega)} \leq C h_{\mathcal{M}} + C W_{\mathcal{D}}(\Lambda\nabla\bar{u}). \tag{2.2.20}$$

As for the Signorini problem, under regularity assumptions the condition that ψ is constant can be relaxed, see Remark 2.4.1. However, if ψ is not constant it might be more practical to use an approximation of this barrier in the numerical discretisation – see Section 2.5.

Remark 2.2.4 (**Convergence without regularity assumptions).** The various regularity assumptions made on the data or the solutions in all previous theorems are used to state optimal orders of convergence. The convergence of gradient schemes can however be established under the sole regularity assumptions stated in Hypothesis 2.2.1 and 2.2.2, see [1].

2.3 Example of gradient schemes

Many numerical schemes can be seen as gradient discretisation method. We show here that the family of *conforming* Galerkin methods can be recast as gradient schemes when applied to variational inequalities.

Example 2.3.1 (Galerkin methods)**.** Any Galerkin method, including *conforming* finite elements and spectral methods of any order, fits into the gradient scheme framework.

For the Signorini problem, if V is a finite dimensional subspace of $\{v \in H^1(\Omega) : \gamma(v) = 0 \text{ on } \Gamma_1\}$, we let $X_{\mathcal{D},\Gamma_{2,3}} = V$, $\Pi_{\mathcal{D}} = \mathbf{Id}$, $\nabla_{\mathcal{D}} = \nabla$ and $\mathbb{T}_{\mathcal{D}} = \gamma$. Then $(X_{\mathcal{D},\Gamma_{2,3}}, \Pi_{\mathcal{D}}, \mathbb{T}_{\mathcal{D}}, \nabla_{\mathcal{D}})$ is a gradient discretisation and the corresponding gradient scheme corresponds to the Galerkin approximation of (2.2.2). Then, $C_{\mathcal{D}}$ defined by (2.2.4) is bounded above by the maximum between the Poincaré constant and the norm of the trace operator $\gamma : H^1(\Omega) \to H^{1/2}(\partial\Omega)$.

Conforming Galerkin methods for the obstacle problem consist in taking $X_{\mathcal{D},0} = V$ a finite dimensional subspace of $H_0^1(\Omega)$, $\Pi_{\mathcal{D}} = \mathbf{Id}$ and $\nabla_{\mathcal{D}} = \nabla$. For this gradient discretisation, $C_{\mathcal{D}}$ (defined by (2.2.8)) is bounded above by the Poincaré constant in $H_0^1(\Omega)$.

For both the Signorini and the obstacle problems, $W_{\mathcal{D}}$ is identically zero and the error estimate is solely dictated by the interpolation error $S_{\mathcal{D}}(\bar{u}, v)$ and by $G_{\mathcal{D}}(\bar{u}, v)$ or $E_{\mathcal{D}}(\bar{u}, v)$, as expected. For $\mathbb{P}1$ finite elements, if the barrier (a or ψ) is piecewise affine on the mesh then the classical interpolant v constructed from the nodal values of $\bar{u}$ belongs to $\mathcal{K}$. As we saw, using this interpolant in Estimates (2.2.14) and (2.2.15) leads to the expected order 1 convergence.

If the barrier is more complex, then it is usual to consider some piecewise affine approximation of it in the definition of the scheme, and the gradient scheme is then modified to take into account this approximate barrier (see Section 2.5); the previous natural interpolant of $\bar u$ then belongs to the (approximate) discrete set $\widetilde{\mathcal{K}}_{\mathcal{D}}$ used to define the scheme, and gives a proper estimate of $S_{\mathcal{D}}(\bar u, v)$ and $G_{\mathcal{D}}(\bar u, v)$ or $E_{\mathcal{D}}(\bar u, v)$.

2.4 Proofs of the theorems

Proof of Theorem 2.2.7

Since $\mathcal{K}_{\mathcal{D}}$ is a nonempty convex closed set, the existence and uniqueness of a solution to Problem (2.2.3) follows from Stampacchia's theorem.

We note that $\Lambda\nabla\bar u \in H_{\rm div}(\Omega)$ since

$$-{\rm div}(\Lambda\nabla\bar u) = f \text{ in the sense of distributions} \tag{2.4.1}$$

(use $v = \bar u \pm \varphi$ in (2.2.1), with $\varphi \in C_c^\infty(\Omega)$). Taking $\boldsymbol{\psi} = \Lambda\nabla\bar u$ in the definition (2.2.6) of $W_{\mathcal{D}}$, for any $z \in X_{\mathcal{D},\Gamma_{2,3}}$ we get

$$\int_\Omega \nabla_{\mathcal{D}} z\cdot\Lambda\nabla\bar u \,{\rm d}\boldsymbol{x} + \int_\Omega \Pi_{\mathcal{D}} z\,{\rm div}(\Lambda\nabla\bar u)\,{\rm d}\boldsymbol{x} \le \|\nabla_{\mathcal{D}} z\|_{L^2(\Omega)^d} W_{\mathcal{D}}(\Lambda\nabla\bar u) + \langle\gamma_{\mathbf{n}}(\Lambda\nabla\bar u), \mathbb{T}_{\mathcal{D}} z\rangle. \tag{2.4.2}$$

Let us focus on the last term on the right-hand side of the above inequality. For any $v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}}$, we have

$$\langle\gamma_{\mathbf{n}}(\Lambda\nabla\bar u), \mathbb{T}_{\mathcal{D}}(v_{\mathcal{D}} - u)\rangle = \langle\gamma_{\mathbf{n}}(\Lambda\nabla\bar u), \mathbb{T}_{\mathcal{D}} v_{\mathcal{D}} - \gamma\bar u\rangle + \langle\gamma_{\mathbf{n}}(\Lambda\nabla\bar u), \gamma\bar u - \mathbb{T}_{\mathcal{D}} u\rangle. \tag{2.4.3}$$

The definition of the space $H^{1/2}_{\Gamma_1}(\partial\Omega)$ shows that there exists $w \in H^1(\Omega)$ such that $\gamma w = \mathbb{T}_{\mathcal{D}} u$ (this implies that $w \in \mathcal{K}$). According to the definition of the duality product (2.B), we have

$$\langle\gamma_{\mathbf{n}}(\Lambda\nabla\bar u), \gamma\bar u - \mathbb{T}_{\mathcal{D}} u\rangle = \int_\Omega \Lambda\nabla\bar u\cdot\nabla(\bar u - w)\,{\rm d}\boldsymbol{x} + \int_\Omega {\rm div}(\Lambda\nabla\bar u)(\bar u - w)\,{\rm d}\boldsymbol{x}.$$

Since $\bar u$ satisfies (2.4.1), it follows that

$$\langle\gamma_{\mathbf{n}}(\Lambda\nabla\bar u), \gamma\bar u - \mathbb{T}_{\mathcal{D}} u\rangle = \int_\Omega \Lambda\nabla\bar u\cdot\nabla(\bar u - w)\,{\rm d}\boldsymbol{x} - \int_\Omega f(\bar u - w)\,{\rm d}\boldsymbol{x}, \tag{2.4.4}$$

which leads to, taking w as a test function in the weak formulation (2.2.1),

$$\langle\gamma_{\mathbf{n}}(\Lambda\nabla\bar u), \gamma\bar u - \mathbb{T}_{\mathcal{D}} u\rangle \le 0. \tag{2.4.5}$$

From this inequality, recalling (2.4.1) and (2.4.3) and setting $z = v_{\mathcal{D}} - u \in X_{\mathcal{D},\Gamma_{2,3}}$ in (2.4.2), we deduce that

$$\int_\Omega \nabla_{\mathcal{D}}(v_{\mathcal{D}} - u)\cdot\Lambda\nabla\bar u\,{\rm d}\boldsymbol{x} + \int_\Omega f\Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\,{\rm d}\boldsymbol{x} \le \|\nabla_{\mathcal{D}}(v_{\mathcal{D}} - u)\|_{L^2(\Omega)^d} W_{\mathcal{D}}(\Lambda\nabla\bar u) + G_{\mathcal{D}}(\bar u, v_{\mathcal{D}}). \tag{2.4.6}$$

We use the fact that u is the solution to Problem (2.2.3) to bound from below the term involving f and we get

$$\int_\Omega \Lambda\nabla_{\mathcal{D}}(v_{\mathcal{D}} - u)\cdot(\nabla\bar u - \nabla_{\mathcal{D}} u)\,{\rm d}\boldsymbol{x} \le \|\nabla_{\mathcal{D}}(v_{\mathcal{D}} - u)\|_{L^2(\Omega)^d} W_{\mathcal{D}}(\Lambda\nabla\bar u) + G_{\mathcal{D}}(\bar u, v_{\mathcal{D}})^+$$

(note that we also used the bound $G_{\mathcal{D}} \le G_{\mathcal{D}}^+$). Adding and subtracting $\nabla_{\mathcal{D}} v_{\mathcal{D}}$ in $\nabla\bar u - \nabla_{\mathcal{D}} u$, and using Cauchy-Schwarz's inequality, this leads to

$$\underline{\lambda}\|\nabla_{\mathcal{D}} v_{\mathcal{D}} - \nabla_{\mathcal{D}} u\|^2_{L^2(\Omega)^d} \le \|\nabla_{\mathcal{D}} v_{\mathcal{D}} - \nabla_{\mathcal{D}} u\|_{L^2(\Omega)^d}\Big(W_{\mathcal{D}}(\Lambda\nabla\bar u) + \overline{\lambda} S_{\mathcal{D}}(\bar u, v_{\mathcal{D}})\Big) + G_{\mathcal{D}}(\bar u, v_{\mathcal{D}})^+.$$

Applying Young's inequality gives

$$\|\nabla_{\mathcal{D}} v_{\mathcal{D}} - \nabla_{\mathcal{D}} u\|_{L^2(\Omega)^d} \le \sqrt{\frac{2}{\underline{\lambda}} G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+ + \frac{1}{\underline{\lambda}^2}\left[W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + \overline{\lambda} S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})\right]^2}. \tag{2.4.7}$$

Estimate (2.2.11) follows from $\|\nabla_{\mathcal{D}} u - \nabla\bar{u}\|_{L^2(\Omega)^d} \le \|\nabla_{\mathcal{D}} u - \nabla_{\mathcal{D}} v_{\mathcal{D}}\|_{L^2(\Omega)^d} + \|\nabla_{\mathcal{D}} v_{\mathcal{D}} - \nabla\bar{u}\|_{L^2(\Omega)^d}$ and $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$. Using (2.2.4) and (2.4.7), we obtain

$$\|\Pi_{\mathcal{D}} v_{\mathcal{D}} - \Pi_{\mathcal{D}} u\|_{L^2(\Omega)} \le C_{\mathcal{D}} \sqrt{\frac{2}{\underline{\lambda}} G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+ + \frac{1}{\underline{\lambda}^2}\left[W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + \overline{\lambda} S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})\right]^2}.$$

By writing $\|\Pi_{\mathcal{D}} u - \bar{u}\|_{L^2(\Omega)} \le \|\Pi_{\mathcal{D}} u - \Pi_{\mathcal{D}} v_{\mathcal{D}}\|_{L^2(\Omega)} + \|\Pi_{\mathcal{D}} v_{\mathcal{D}} - \bar{u}\|_{L^2(\Omega)}$, the above inequality shows that

$$\|\Pi_{\mathcal{D}} u - \bar{u}\|_{L^2(\Omega)} \le C_{\mathcal{D}} \sqrt{\frac{2}{\underline{\lambda}} G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+ + \frac{1}{\underline{\lambda}^2}\Big(W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + \overline{\lambda} S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})\Big)^2} + S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}).$$

Applying $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ again, Estimate (2.2.12) is obtained and the proof is complete.

Proof of Theorem 2.2.8

The proof of this theorem is very similar to the proof of [1, Theorem 1]. We however give some details for the sake of completeness.

As for the Signorini problem, the existence and uniqueness of the solution to Problem (2.2.7) follows from Stampacchia's theorem. Let us now establish the error estimates. Under the assumption that $\mathrm{div}(\Lambda\nabla\bar{u}) \in L^2(\Omega)$, we note that $\Lambda\nabla\bar{u} \in H_{\mathrm{div}}(\Omega)$. For any $w \in X_{\mathcal{D},0}$, using $\boldsymbol{\psi} = \Lambda\nabla\bar{u}$ in the definition (2.2.10) of $W_{\mathcal{D}}$ therefore implies

$$\int_\Omega \nabla_{\mathcal{D}} w \cdot \Lambda\nabla\bar{u}\,\mathrm{d}\boldsymbol{x} + \int_\Omega \Pi_{\mathcal{D}} w\,\mathrm{div}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} \le \|\nabla_{\mathcal{D}} w\|_{L^2(\Omega)^d} W_{\mathcal{D}}(\Lambda\nabla\bar{u}). \tag{2.4.8}$$

For any $v_{\mathcal{D}} \in \mathcal{K}_{\mathcal{D}}$, one has

$$\begin{aligned}\int_\Omega \Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\mathrm{div}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} &= \int_\Omega (\Pi_{\mathcal{D}} u - \psi)(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} \\ &\quad + \int_\Omega (\psi - \Pi_{\mathcal{D}} v_{\mathcal{D}})(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} - \int_\Omega f\Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{2.4.9}$$

It is well known that the solution to the weak formulation (2.2.2) satisfies (2.1.2b) in the sense of distributions (use test functions $v = \bar{u} - \varphi$ in (2.2.2), with $\varphi \in C_c^\infty(\Omega)$ nonnegative). Hence, under our regularity assumptions, (2.1.2b) holds a.e. and, since $u \in \mathcal{K}_{\mathcal{D}}$, we obtain $\int_\Omega (\Pi_{\mathcal{D}} u - \psi)(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} \le 0$. Hence,

$$\begin{aligned}&\int_\Omega \Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\mathrm{div}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} \\ &\qquad \le \int_\Omega (\psi - \Pi_{\mathcal{D}} v_{\mathcal{D}})(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} - \int_\Omega f\Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x} \\ &\qquad = \int_\Omega (\psi - \bar{u})(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} + \int_\Omega (\bar{u} - \Pi_{\mathcal{D}} v_{\mathcal{D}})(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} - \int_\Omega f\Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Our regularity assumptions ensure that $\bar{u}$ satisfies (2.1.2a). Therefore, by definition of $E_{\mathcal{D}}$,

$$\int_\Omega \mathit{\Pi}_{\mathcal{D}}(u - v_{\mathcal{D}})\mathrm{div}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} \le E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) - \int_\Omega f\Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x}. \tag{2.4.10}$$

From this inequality and setting $w = v_{\mathcal{D}} - u \in X_{\mathcal{D},0}$ in (2.4.8), we obtain

$$\int_\Omega \nabla_{\mathcal{D}}(v_{\mathcal{D}} - u)\cdot\Lambda\nabla\bar{u}\,\mathrm{d}\boldsymbol{x} + \int_\Omega f\Pi_{\mathcal{D}}(u - v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x} \le \|\nabla_{\mathcal{D}}(v_{\mathcal{D}} - u)\|_{L^2(\Omega)^d} W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}).$$

The rest of the proof can be handled in much the same way as the proof of Theorem 2.2.7, taking over the reasoning from (2.4.6).

Proof of Theorem 2.2.10

We follow the same technique used in [61]. According to Remark 2.2.3 and due to the regularity on the solution, since $\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} = \gamma(\bar{u}) = 0$ on Γ_1 and $\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}) = 0$ on Γ_2, we can write

$$G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = \int_{\Gamma_3} (\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} - \gamma(\bar{u}))\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x}.$$

We notice first that the assumptions on $\bar{u}$ ensure that it is a solution of the Signorini problem in the strong sense. By applying Theorem 2.2.7, we have

$$\|\nabla_{\mathcal{D}}u - \nabla\bar{u}\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}}u - \bar{u}\|_{L^2(\Omega)} \le C\left(W_{\mathcal{D}}(\Lambda\nabla\bar{u}) + S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) + \sqrt{G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+}\right) \tag{2.4.11}$$

with C depending only on $\underline{\lambda}$, $\overline{\lambda}$ and an upper bound of $C_{\mathcal{D}}$. Since $S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) \le C_1 h_{\mathcal{M}}$ by assumption, it remains to estimate the last term $G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+$. We start by writing

$$\begin{aligned}
G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) &= \int_{\Gamma_3} (\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} - a)\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} + \int_{\Gamma_3} (a - \gamma(\bar{u}))\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x}\\
&= \int_{\Gamma_3} (\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} - a)\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x}\\
&= \sum_{\sigma\in\mathcal{E}\,,\,\sigma\subset\Gamma_3} \int_\sigma (\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} - a)\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x}\\
&=: \sum_{\sigma\in\mathcal{E}\,,\,\sigma\subset\Gamma_3} G_\sigma(\bar{u}, v_{\mathcal{D}}),
\end{aligned}$$

where the term involving $(a - \gamma(\bar{u}))\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})$ has been eliminated by using (2.1.1d). We then split the study on each σ depending on the cases: either (i) $\gamma\bar{u} < a$ a.e. on σ, or (ii) $\mathrm{meas}(\{\boldsymbol{y}\in\sigma : \gamma\bar{u}(\boldsymbol{y}) - a = 0\}) > 0$ (where meas is the $(d-1)$-dimensional measure).

In Case (i), we have $\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}) = 0$ in σ since $\gamma\bar{u}$ satisfies (2.1.1d). Hence, $G_\sigma(\bar{u}, v_{\mathcal{D}}) = 0$.

In Case (ii), let us denote ∇_σ the tangential gradient to σ, and let us recall that, as a consequence of Stampacchia's lemma, if $w \in W^{1,1}(\sigma)$ then $\nabla_\sigma w = 0$ a.e. on $\{\boldsymbol{y}\in\sigma : w(\boldsymbol{y}) = 0\}$ (here, "a.e." is for the measure meas). Hence, with $w = \gamma\bar{u} - a$ we obtain at least one $\boldsymbol{y}_0 \in \sigma$ such that $(\gamma\bar{u} - a)(\boldsymbol{y}_0) = 0$ and $\nabla_\sigma(\gamma\bar{u} - a)(\boldsymbol{y}_0) = 0$. Let F be the face of $\partial\Omega$ that contains σ. Using a Taylor's expansion along a path on F between $\boldsymbol{y}_0$ and $\boldsymbol{x}_\sigma$, we deduce

$$|(\gamma\bar{u} - a)(\boldsymbol{x}_\sigma)| \le L_F h_\sigma^2, \tag{2.4.12}$$

where L_F depends on F and on the Lipschitz constant of the tangential derivative $\nabla_{\partial\Omega}(\gamma\bar{u} - a)$ on F (the Lipschitz-continuity of $\nabla_{\partial\Omega}(\gamma\bar{u} - a)$ on this face follows from our assumption that $\gamma\bar{u} - a \in W^{2,\infty}(\partial\Omega)$). Recalling that $||\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} - \gamma\bar{u}(\boldsymbol{x}_\sigma)||_{L^2(\sigma)} \le C_2 h_\sigma^2|\sigma|^{1/2}$, we infer that $||\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} - a||_{L^2(\sigma)} \le Ch_{\mathcal{M}}^2|\sigma|^{1/2}$, and therefore, that $|G_\sigma(\bar{u}, v_{\mathcal{D}})| \le Ch_{\mathcal{M}}^2||\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})||_{L^2(\sigma)}|\sigma|^{1/2}$. Here, C depends only on Ω, $\bar{u}$ and a. Gathering the upper bounds on each G_σ and using the Cauchy–Schwarz inequality gives $G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) \le Ch_{\mathcal{M}}^2||\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})||_{L^2(\partial\Omega)}$, with C depending only on Ω, Λ, $\bar{u}$ and a. This implies $G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+ \le Ch_{\mathcal{M}}^2$, and the proof is complete by plugging this estimate into (2.4.11).

Proof of Theorem 2.2.11

The proof is very similar to the proof of Theorem 2.2.10. As in this previous proof, we only have to estimate $E_{\mathcal{D}}(\bar u, v_{\mathcal{D}})$, which can be re-written as

$$\begin{aligned}E_{\mathcal{D}}(\bar u, v_{\mathcal{D}}) &= \int_\Omega (\mathrm{div}(\Lambda\nabla\bar u)+f)(\bar u-\psi)\,\mathrm{d}\boldsymbol{x} + \int_\Omega (\mathrm{div}(\Lambda\nabla\bar u)+f)(\psi-\Pi_{\mathcal{D}}v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x}\\ &= \int_\Omega (\mathrm{div}(\Lambda\nabla\bar u)+f)(\psi-\Pi_{\mathcal{D}}v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x}\\ &= \sum_{K\in\mathcal{M}}\int_K (\mathrm{div}(\Lambda\nabla\bar u)+f)(\psi-\Pi_{\mathcal{D}}v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x} =: \sum_{K\in\mathcal{M}} E_K(\bar u, v_{\mathcal{D}}),\end{aligned}$$

where the term involving $(\mathrm{div}(\Lambda\nabla\bar u)+f)(\bar u-\psi)$ has been eliminated using (2.1.2a). Each term $E_K(\bar u, v_{\mathcal{D}})$ is then estimated by considering two cases, namely: (i) either $\bar u<\psi$ on K, in which case $E_K(\bar u, v_{\mathcal{D}})=0$ since $f+\mathrm{div}(\Lambda\bar u)=0$ on K, or (ii) $|\{\boldsymbol{y}\in K\;:\;\bar u(\boldsymbol{y})=\psi\}|>0$, in which case $E_K(\bar u, v_{\mathcal{D}})$ is estimated by using a Taylor expansion and the assumption $||\Pi_{\mathcal{D}}v_{\mathcal{D}}-\bar u(\boldsymbol{x}_K)||_{L^2(K)}\le C_2h_K^2|K|^{1/2}$.

Remark 2.4.1. If $(\boldsymbol{x}_K)_{K\in\mathcal{M}}$ are the centres of gravity of the cells, the assumption that ψ is constant can be relaxed under additional regularity hypotheses. Namely, if $F:=\mathrm{div}(\Lambda\nabla\bar u)+f\in H^1(K)$ and $\psi\in H^2(K)$ for any cell K, then by letting F_K and ψ_K be the mean values on K of F and ψ, respectively, we can write

$$\begin{aligned}E_K(\bar u, v_{\mathcal{D}}) &= \int_K F(\psi-\Pi_{\mathcal{D}}v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x}\\ &= \int_K F(\psi(\boldsymbol{x}_K)-\Pi_{\mathcal{D}}v_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x} + \int_K (F-F_K)(\psi-\psi_K)\,\mathrm{d}\boldsymbol{x} + \int_K F(\psi_K-\psi(\boldsymbol{x}_K))\,\mathrm{d}\boldsymbol{x}\\ &=: T_{1,K}+T_{2,K}+T_{3,K}.\end{aligned}$$

The term $T_{1,K}$ is estimated as in the proof (since $\bar u(\boldsymbol{x}_K)-\psi(\boldsymbol{x}_K)$ can be estimated using a Taylor expansion about $\boldsymbol{y}_0$). We estimate $T_{2,K}$ by using the Cauchy-Schwarz inequality and classical estimates between an H^1 function and its average:

$$|T_{2,K}|\le ||F-F_K||_{L^2(K)}||\psi-\psi_K||_{L^2(K)}\le Ch_K^2||F||_{H^1(K)}||\psi||_{H^1(K)}.$$

As for $T_{3,K}$, we use the fact that $\psi\in H^2(\Omega)$ and that $\boldsymbol{x}_K$ is the center of gravity of K to write $|\psi(\boldsymbol{x}_K)-\psi_K|\le Ch_K^2||\psi||_{H^2(K)}$. Combining all these estimates to bound $E_K(\bar u, v_D)$ and using Cauchy-Schwarz inequalities leads to an upper bound in $\mathcal{O}(h_{\mathcal{M}}^2)$ for $E_{\mathcal{D}}(\bar u, v_D)$.

2.5 The case of approximate barriers

For general barrier functions (a in the Signorini problem, ψ in the obstacle problem), it might be challenging to find a proper approximation of $\bar u$ inside $\mathcal{K}_{\mathcal{D}}$. Consider, for example, the $\mathbb{P}1$ finite element method; the approximation is usually constructed using the values of $\bar u$ at the nodes of the mesh, which only ensures that this approximation is bounded above by the barrier at these nodes, not necessarily everywhere else in the domain. It is therefore usual to relax the upper bound imposed on the solution to the numerical scheme, and to consider only approximate barriers in the schemes.

2.5.1 The Signorini problem

We consider the gradient scheme (2.2.3) with, instead of $\mathcal{K}_{\mathcal{D}}$, the following convex set:

$$\widetilde{\mathcal{K}}_{\mathcal{D}} = \{v\in X_{\mathcal{D},\Gamma_1}\;:\;\mathbb{T}_{\mathcal{D}}v\le a_{\mathcal{D}}\text{ on }\Gamma_3\}, \tag{2.5.1}$$

where $a_{\mathcal{D}} \in L^2(\partial\Omega)$ is an approximation of a. The following theorems state error estimates for this modified gradient scheme.

Theorem 2.5.1 (Error estimates for the Signorini problem with approximate barrier). *Under the assumptions of Theorem 2.2.7, if $\widetilde{\mathcal{K}}_{\mathcal{D}}$ is not empty then there exists a unique solution u to the gradient scheme* (2.2.3) *in which $\mathcal{K}_{\mathcal{D}}$ has been replaced with $\widetilde{\mathcal{K}}_{\mathcal{D}}$. Moreover, if $a - a_{\mathcal{D}} \in H^{1/2}_{\Gamma_1}(\partial\Omega)$, Estimates* (2.2.11) *and* (2.2.12) *hold for any $v_{\mathcal{D}} \in \widetilde{\mathcal{K}}_{\mathcal{D}}$, provided that $G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$ is replaced with*

$$\widetilde{G}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) + \langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), a - a_{\mathcal{D}}\rangle$$

Proof. The proof is identical to that of Theorem 2.2.7, provided we can control the left-hand side of (2.4.3). For any $v_{\mathcal{D}} \in \widetilde{\mathcal{K}}_{\mathcal{D}}$, we write

$$\langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), \mathbb{T}_{\mathcal{D}}(v_{\mathcal{D}} - u)\rangle = \langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), \mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} - \gamma\bar{u}\rangle + \langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), \gamma\bar{u} - (\mathbb{T}_{\mathcal{D}}u + a - a_{\mathcal{D}})\rangle + \langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), a - a_{\mathcal{D}}\rangle.$$

We note that the first term in the right-hand side is $G_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$; hence, the first and last terms in this right-hand side add up to $\widetilde{G}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$ and we only need to prove that the second term is nonpositive. We take $w \in H^1(\Omega)$ such that $\gamma(\bar{w}) = \mathbb{T}_{\mathcal{D}}u + a - a_{\mathcal{D}}$, and we notice that w belongs to $\mathcal{K}$ since $\mathbb{T}_{\mathcal{D}}u$ and $a - a_{\mathcal{D}}$ both belong to $H^{1/2}_{\Gamma_1}(\partial\Omega)$, and since $\mathbb{T}_{\mathcal{D}}u + a - a_{\mathcal{D}} \le a$ on Γ_3. Similarly to (2.4.4), the definition of w shows that

$$\langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), \gamma\bar{u} - (\mathbb{T}_{\mathcal{D}}u + a - a_{\mathcal{D}})\rangle = \int_\Omega \Lambda\nabla\bar{u} \cdot \nabla(\bar{u} - w)\,\mathrm{d}\boldsymbol{x} - \int_\Omega f(\bar{u} - w)\,\mathrm{d}\boldsymbol{x}.$$

Hence, by (2.2.1), $\langle \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}), \gamma\bar{u} - (\mathbb{T}_{\mathcal{D}}u + a - a_{\mathcal{D}})\rangle \le 0$ and, as required, the left-hand side of (2.4.3) is bounded above by $\widetilde{G}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$. □

In the case of the $\mathbb{P}1$ finite element method, $a_{\mathcal{D}}$ is the $\mathbb{P}_1$ approximation of a on $\partial\Omega$ constructed from its values at the nodes. The interpolant $v_{\mathcal{D}}$ of $\bar{u}$ mentioned in Section 2.2.4.1 is bounded above by $a = a_{\mathcal{D}}$ at the nodes, and thus everywhere by $a_{\mathcal{D}}$; $v_{\mathcal{D}}$ therefore, belongs to $\widetilde{\mathcal{K}}_{\mathcal{D}}$ and can be used in Theorem 2.5.1. Under H^2 regularity of a, we have $||a_{\mathcal{D}} - a||_{L^2(\partial\Omega)} = \mathcal{O}(h^2_{\mathcal{M}})$. Hence, if $\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}) \in L^2(\partial\Omega)$, we see that $\widetilde{G}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = \mathcal{O}(h^2_{\mathcal{M}})$ and Theorem 2.5.1 thus gives an order one estimate.

For several low-order methods (e.g. HMM, see [40]), the interpolant $v_{\mathcal{D}}$ of the exact solution $\bar{u}$ is constructed such that $\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} = \gamma(\bar{u})(\overline{\boldsymbol{x}}_\sigma)$ on each $\sigma \in \mathcal{E}_{\mathrm{ext}}$. The natural approximate barrier is then piecewise constant, and an order one error estimate can be obtained as shown in the next result.

Theorem 2.5.2 (Signorini problem: order of convergence for piecewise constant reconstructions). *Under Hypothesis 2.2.1, let $\mathcal{D}$ be a gradient discretisation, in the sense of Definition 2.2.3, and let $\mathcal{T} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be a polytopal mesh of Ω. Let $a_{\mathcal{D}} \in L^2(\partial\Omega)$ be such that $a_{\mathcal{D}} - a = 0$ on Γ_1 and, for any $\sigma \in \mathcal{E}_{\mathrm{ext}}$ such that $\sigma \subset \Gamma_3$, $a_{\mathcal{D}} = a(\boldsymbol{x}_\sigma)$ on σ, where $\boldsymbol{x}_\sigma \in \sigma$. Let $\bar{u}$ be the solution to* (2.2.1)*, and let us assume that $\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}) \in L^2(\partial\Omega)$ and that $\gamma(\bar{u}) - a \in W^{2,\infty}(\partial\Omega)$. We also assume that there exists an interpolant $v_{\mathcal{D}} \in X_{\mathcal{D},\Gamma_{2,3}}$ such that $S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) \le C_1 h_{\mathcal{M}}$ with C_1 not depending on $\mathcal{D}$ or $\mathcal{T}$, and, for any $\sigma \subset \Gamma_3$, $\mathbb{T}_{\mathcal{D}}v_{\mathcal{D}} = \gamma(\bar{u})(\boldsymbol{x}_\sigma)$ on σ.*

Then $\widetilde{\mathcal{K}}_{\mathcal{D}} \neq \emptyset$ and, if u is the solution to (2.2.3) *with $\mathcal{K}_{\mathcal{D}}$ replaced with $\widetilde{\mathcal{K}}_{\mathcal{D}}$, it holds*

$$\|\nabla_{\mathcal{D}}u - \nabla\bar{u}\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}}u - \bar{u}\|_{L^2(\Omega)} \le Ch_{\mathcal{M}} + CW_{\mathcal{D}}(\Lambda\nabla\bar{u}) \tag{2.5.2}$$

where C depends only on Ω, Λ, $\bar{u}$, a, C_1 and an upper bound of $C_{\mathcal{D}}$.

Proof. Clearly, $v_{\mathcal{D}} \in \widetilde{\mathcal{K}}_{\mathcal{D}}$. The conclusion follows from Theorem 2.5.1 if we prove that $\widetilde{G}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})^+ = \mathcal{O}(h^2_{\mathcal{M}})$. Note that, following Remark 2.2.3, it makes sense to consider a piecewise constant reconstructed trace $\mathbb{T}_{\mathcal{D}}v$.

With our choices of $a_{\mathcal{D}}$ and $\Pi_{\mathcal{D}} v_{\mathcal{D}}$, and using the fact that $\bar{u}$ is the solution to the Signorini problem in the strong sense (and satisfies therefore (2.1.1d)), we have

$$
\begin{aligned}
\widetilde{G}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) &= \int_{\Gamma_3} (\mathbb{T}_{\mathcal{D}} v_{\mathcal{D}} - \gamma\bar{u})\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} + \int_{\Gamma_3} (a - a_{\mathcal{D}})\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\Gamma_3} \gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})(\mathbb{T}_{\mathcal{D}} v_{\mathcal{D}} - a_{\mathcal{D}})\,\mathrm{d}\boldsymbol{x} \\
&= \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}},\,\sigma\subset\Gamma_3} \int_{\sigma} (\gamma\bar{u}(\boldsymbol{x}_\sigma) - a(\boldsymbol{x}_\sigma))\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u})\,\mathrm{d}\boldsymbol{x} \\
&=: \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}},\,\sigma\subset\Gamma_3} \widetilde{G}_{\sigma}(\bar{u}, v_{\mathcal{D}}).
\end{aligned}
$$

We then deal edge by edge, considering two cases as in the proof of Theorem 2.2.10. In Case (i), where $\gamma\bar{u} < a$ a.e. on σ, we have $\widetilde{G}_{\sigma}(\bar{u}, v_{\mathcal{D}}) = 0$ since $\gamma_{\mathbf{n}}(\Lambda\nabla\bar{u}) = 0$ in σ. In Case (ii), we estimate $\widetilde{G}_{\sigma}(\bar{u}, v_{\mathcal{D}})$ using the Taylor expansion (2.4.12). □

2.5.2 The obstacle problem

With $\psi_{\mathcal{D}} \in L^2(\Omega)$ an approximation of ψ, we consider the new convex set

$$
\widetilde{\mathcal{K}}_{\mathcal{D}} = \{v \in X_{\mathcal{D},0} \;:\; \Pi_{\mathcal{D}} v \le \psi_{\mathcal{D}} \text{ in } \Omega\} \tag{2.5.3}
$$

and we write the gradient scheme (2.2.7) with $\widetilde{\mathcal{K}}_{\mathcal{D}}$ instead of $\mathcal{K}_{\mathcal{D}}$. The following theorems are the equivalent for the obstacle problem of Theorems 2.5.1 and 2.5.2.

Theorem 2.5.3 (Error estimates for the obstacle problem with approximate barrier). *Under the assumptions of Theorem 2.2.8, if $\widetilde{\mathcal{K}}_{\mathcal{D}}$ is not empty then there exists a unique solution u to the gradient scheme* (2.2.7) *in which $\mathcal{K}_{\mathcal{D}}$ has been replaced with $\widetilde{\mathcal{K}}_{\mathcal{D}}$. Moreover, Estimates* (2.2.14) *and* (2.2.15) *hold for any $v_{\mathcal{D}} \in \widetilde{\mathcal{K}}_{\mathcal{D}}$, provided that $E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$ is replaced with*

$$
\widetilde{E}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) = E_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) + \int_{\Omega} (\mathrm{div}(\Lambda\nabla\bar{u}) + f)(\psi_{\mathcal{D}} - \psi)\,\mathrm{d}\boldsymbol{x}.
$$

Proof. We follow exactly the proof of Theorem 2.2.8, except that we introduce $\psi_{\mathcal{D}}$ instead of ψ in (2.4.9). The first term in the right-hand side of this equation is then still bounded above by 0, and the second term is written

$$
\int_{\Omega} (\psi_{\mathcal{D}} - \Pi_{\mathcal{D}} v_{\mathcal{D}})(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} = \int_{\Omega} (\psi_{\mathcal{D}} - \psi)(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x} + \int_{\Omega} (\psi - \Pi_{\mathcal{D}} v_{\mathcal{D}})(\mathrm{div}(\Lambda\nabla\bar{u}) + f)\,\mathrm{d}\boldsymbol{x}.
$$

The first term in this right-hand side corresponds to the additional term in $\widetilde{E}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$, whereas the second term is the one handled in the proof of Theorem 2.2.8. □

Theorem 2.5.4 (Obstacle problem: order of convergence for piecewise constant reconstructions). *Let $\mathcal{D}$ be a gradient discretisation, in the sense of Definition 2.2.5, and let $\mathcal{T} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be a polytopal mesh of Ω. Let $\psi_{\mathcal{D}} \in L^2(\Omega)$ be defined by*

$$
\forall K \in \mathcal{M}\,,\; \psi_{\mathcal{D}} = \psi(\boldsymbol{x}_K) \text{ on } K.
$$

Let $\bar{u}$ be the solution to (2.2.2)*, and let us assume that* $\mathrm{div}(\Lambda\nabla\bar{u}) \in L^2(\Omega)$ *and that* $\bar{u} - \psi \in W^{2,\infty}(\Omega)$*. We also assume that there exists an interpolant $v_{\mathcal{D}} \in X_{\mathcal{D},0}$ such that $S_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}}) \le C_1 h_{\mathcal{M}}$, with C_1 not depending on $\mathcal{D}$ or $\mathcal{T}$, and that $\Pi_{\mathcal{D}} v_{\mathcal{D}} = \bar{u}(\boldsymbol{x}_K)$ on K, for any $K \in \mathcal{M}$.*

Then $\widetilde{\mathcal{K}}_{\mathcal{D}} \neq \emptyset$ and, if u is the solution to (2.2.7) *with $\mathcal{K}_{\mathcal{D}}$ replaced with $\widetilde{\mathcal{K}}_{\mathcal{D}}$,*

$$\|\nabla_{\mathcal{D}} u - \nabla \bar{u}\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}} u - \bar{u}\|_{L^2(\Omega)} \leq C h_{\mathcal{M}} + C W_{\mathcal{D}}(\Lambda \nabla \bar{u}) \tag{2.5.4}$$

where C depends only on Ω, Λ, $\bar{u}$, ψ, C_1 and an upper bound of $C_{\mathcal{D}}$.

Proof. The proof can be conducted by following the same ideas as in the proof of Theorem 2.5.2. □

Let us now compare our results with previous studies. In [17, 50], $\mathcal{O}(h_{\mathcal{M}})$ error estimates are established for $\mathbb{P}_1$ and mixed finite elements applied to the Signorini problem and the obstacle problem. This order was obtained under the assumptions that $\bar{u} \in W^{1,\infty}(\Omega) \cap H^2(\Omega)$ and a is constant for the Signorini problem, and that $\Lambda \equiv \mathbf{Id}$ and $\bar{u}$ and ψ are in $H^2(\Omega)$ for the obstacle problem. Our results generalise these orders of convergence to the case of a Lipschitz-continuous Λ and a nonconstant a (note that mixed finite elements are also part of the gradient schemes framework, see [41, 37]).

Studies of *non-conforming* methods for variational inequalities are scarcer. We cite [87], which applies Crouzeix–Raviart methods to the obstacle problem and obtains an order $\mathcal{O}(h_{\mathcal{M}})$ under strong regularity assumptions, namely $f \in L^\infty(\Omega)$, $\bar{u} - \psi \in W^{2,\infty}(\Omega)$, $\psi \in H^2(\Omega)$, $\Lambda \equiv \mathbf{Id}$ and the free boundary has a finite length. For the Signorini problem, under the assumptions that a is constant, $\bar{u} \in W^{2,\infty}(\Omega)$ and $\Lambda \equiv \mathbf{Id}$, Wang and Hua [92] give a proof of an order $\mathcal{O}(h_{\mathcal{M}})$ for Crouzeix–Raviart methods. The natural Crouzeix–Raviart interpolants are piecewise linear on the edges and the cells, and if we therefore use it as $a_{\mathcal{D}}$ for the Signorini problem, we see that under an H^2 regularity assumption on a, we have $||a - a_{\mathcal{D}}||_{L^2(\partial\Omega)} \leq C h_{\mathcal{M}}^2$. Hence, the additional term in $\widetilde{G}_{\mathcal{D}}(\bar{u}, v_{\mathcal{D}})$ in Theorem 2.5.1 is of order $h_{\mathcal{M}}^2$ and this theorem therefore gives back the known $\mathcal{O}(h_{\mathcal{M}})$ order of convergence for the Crouzeix–Raviart scheme applied to the Signorini problem. This is obtained under slightly more general assumptions, since Λ does not need to be $\mathbf{Id}$ here. The same applies to the obstacle problem through Theorem 2.5.3.

We finally notice that most previous research investigates the Signorini problem under the assumption that the barrier a is a constant. On the contrary, Theorems 2.2.7, 2.5.1 and 2.5.2 presented here are valid for nonconstant barriers.

Appendix

2.A Equivalence between the weak and strong formulations

There is no guarantee to find a smooth classical solution in $C^2(\overline{\Omega})$ to the Signorini and the obstacle models, (2.1.1) and (2.1.2), even under strong assumptions on the data, such as taking f in $C^1(\Omega)$, see [73], for example. However, the existence of these solutions can be proved in the weak sense. The following propositions show that, for smooth solutions, the strong and weak formulations of the Signorini (resp. the obstacle) problem are equivalent.

Proposition 2.A.1. *Let Hypothesis 2.2.1 hold. We assume that $\bar{u} \in C^2(\overline{\Omega})$, then $\bar{u}$ is a weak solution to* (2.2.1) *if and only if $\bar{u}$ satisfies* (2.1.1).

Proof. The following proof is introduced in [55]. We temporarily assume that the solution $\bar{u} \in C^2(\overline{\Omega}) \cap \mathcal{K}$ and satisfies (2.1.1). We multiply Equation (2.1.1a) by $\bar{u} - v$, for a generic $v \in \mathcal{K}$, and apply the divergence theorem to obtain

$$\int_\Omega \nabla\bar{u} \cdot \nabla(\bar{u} - v)\,\mathrm{d}\boldsymbol{x} = \int_{\partial\Omega} \nabla\bar{u} \cdot \mathbf{n}(\bar{u} - v)\,\mathrm{d}\boldsymbol{x} + \int_\Omega f(\bar{u} - v)\,\mathrm{d}\boldsymbol{x}, \quad \forall v \in \mathcal{K}.$$

Introducing the barrier a, this equality becomes

$$\int_\Omega \nabla\bar{u} \cdot \nabla(\bar{u} - v)\,\mathrm{d}\boldsymbol{x} = \int_{\Gamma_3} \nabla\bar{u} \cdot \mathbf{n}(\bar{u} - a)\,\mathrm{d}\boldsymbol{x} + \int_{\Gamma_3} \nabla\bar{u} \cdot \mathbf{n}(a - v)\,\mathrm{d}\boldsymbol{x} + \int_\Omega f(\bar{u} - v)\,\mathrm{d}\boldsymbol{x}, \quad \forall v \in \mathcal{K}.$$

Due to (2.1.1b), the first term of the right-hand side of this equality is zero, and the second term is non–positive (because $v \in \mathcal{K}$). Therefore, $\bar{u}$ satisfies the inequality of Problem (2.2.1). Since this formulation still makes sense if $\bar{u} \in \mathcal{K}$ and $f \in L^2(\Omega)$, then $\bar{u}$ is a solution to the weak formulation problem (2.2.1).

Conversely, letting $u \in C^2(\overline{\Omega}) \cap \mathcal{K}$ be a solution to (2.2.1), we show that $\bar{u}$ also satisfies the strong formulation (2.1.1). Take $v = \bar{u} \pm \varphi$ such that $\varphi \in C_0^\infty(\overline{\Omega})$ as a test function in (2.2.1) to deduce

$$\int_\Omega (f - \nabla\bar{u})\varphi\,\mathrm{d}\boldsymbol{x} = 0, \qquad \text{for all } \varphi \in C_0^\infty(\overline{\Omega}).$$

From the definition of distribution, it follows that

$$\int_\Omega (f + \operatorname{div}(\nabla\bar{u}))\varphi\,\mathrm{d}\boldsymbol{x} = 0, \qquad \text{for all } \varphi \in C_0^\infty(\overline{\Omega}),$$

which yields (2.1.1a).

Let g be a smooth function with compact support in Γ_2. Let $\varphi \in C^\infty(\overline{\Omega})$ such that $\gamma\varphi = g$ on Γ_2 and $\gamma\varphi = 0$ on $\Gamma_1 \cup \Gamma_3$. Insert $v = \bar{u} \pm \varphi$ respectively in (2.2.1) to get

$$\int_\Omega (\nabla\bar{u} - f)\varphi\,\mathrm{d}\boldsymbol{x} = 0, \quad \text{for all } \varphi \in C^\infty(\overline{\Omega}), \text{ such that } \gamma\varphi = 0 \text{ on } \Gamma_1 \cup \Gamma_3.$$

By utilising (2.1.1a) proved previously and the divergence theorem, we obtain

$$\int_{\Gamma_2} \nabla\bar{u} \cdot \mathbf{n}\, g\,\mathrm{d}\boldsymbol{x} = 0, \quad \text{for all } g \in C_c^\infty(\Gamma_2),$$

that proves (2.1.1c).

Take $v = \bar{u} - \varphi$ in (2.2.1), such that $\varphi \in C^\infty(\overline{\Omega})$, $\gamma\varphi = 0$ on Γ_1 and $\gamma\varphi > 0$ on Γ_3. Applying the previous reasoning and using (2.1.1c), we see that the second inequality of (2.1.1d) holds.

It remains to verify the last equality of the Signorini boundary conditions (2.1.1d). To do so, it is sufficient to show that $\nabla\bar{u} \cdot \mathbf{n} = 0$ on the set $A := \{\boldsymbol{x} \in \Gamma_3 : \ \bar{u}(\boldsymbol{x}) < a(\boldsymbol{x})\}$. Let g be a smooth function with compact support in A and $\widetilde{g}$ be a zero extension of g on Γ_3. Let φ be a function of $H^1(\Omega)$ such that $\gamma(\varphi) = \widetilde{g}$ on Γ_3 and $\gamma\varphi = 0$ on Γ_1. Introduce $\delta \geq 0$ by

$$\delta = \operatorname{ess\,inf}(a - \gamma\bar{u}(\boldsymbol{x}),\ \boldsymbol{x} \in \operatorname{supp} g).$$

There exists $\mu = \frac{\delta}{||g||_{L^\infty(A)}}$, such that $v = \bar{u} \pm \mu\varphi$ belong to $\mathcal{K}$. Setting v as a tests function in (2.2.1), and reasoning as previously, we infer

$$\int_A \nabla\bar{u} \cdot \mathbf{n}\, g \,\mathrm{d}\boldsymbol{x} = 0, \quad \text{for all } g \in C_c^\infty(A),$$

which concludes the proof. □

Proposition 2.A.2. *Let Hypothesis 2.2.2 hold. We assume that $\bar{u} \in C^2(\overline{\Omega})$, then $\bar{u}$ is a solution to the variational problem* (2.2.2) *if and only if $\bar{u}$ satisfies* (2.1.2)*.*

Proof. The proof is the same as in [93]. Temporarily, assume that solution $\bar{u}$ of (2.1.2) is in $C_0^2(\overline{\Omega}) \cap \mathcal{K}$. For any $v \in \mathcal{K}$, multiplying both sides of Equation (2.1.2a) by $\bar{u} - v$ and integrating over Ω, we obtain

$$\begin{aligned}\int_\Omega (-\operatorname{div}(\nabla\bar{u}) - f)(\bar{u} - v) &= \int_\Omega (-\operatorname{div}(\nabla\bar{u}) - f)(\bar{u} - \psi)\,\mathrm{d}\boldsymbol{x} + \int_\Omega (-\operatorname{div}(\nabla\bar{u}) - f)(\psi - v)\,\mathrm{d}\boldsymbol{x} \\ &\leq 0, \quad \text{for all } v \in \mathcal{K},\end{aligned}$$

due to (2.1.2a), leading to the first term on the right-hand side is zero, and to $v \leq \psi$ in Ω, that shows, together with (2.1.2b), that the second term is non–positive. Integrating by part this relation, the divergence theorem proves that $\bar{u}$ must satisfy the inequality formula in the problem (2.2.2). This inequality also makes sense if $\bar{u} \in \mathcal{K}$.

Conversely, suppose that $\bar{u} \in C^2(\overline{\Omega}) \cap \mathcal{K}$ is a weak solution of (2.2.2) and let $v = \bar{u} - w$ as a test function such that $w \in C_0^\infty(\Omega)$ and $w \geq 0$ in Ω to get

$$\int_\Omega (\nabla\bar{u} \cdot \nabla w - fw)\,\mathrm{d}\boldsymbol{x} \leq 0, \qquad \text{for all } w \in C_0^\infty(\Omega) \text{ and } w \geq 0.$$

Integration by part of this inequality gives again,

$$\int_\Omega (-\operatorname{div}(\nabla\bar{u}) - f) w \,\mathrm{d}\boldsymbol{x} \leq 0, \qquad \text{for all } w \in C_0^\infty(\Omega) \text{ and } w \geq 0,$$

which shows that differential inequality (2.1.2b) is fulfilled.

We assume that $\bar{u} < \psi$ for some points $\boldsymbol{x}_0 \in \Omega$. Then $\boldsymbol{x}_0$ has a neighbourhood $\mathcal{N}(\boldsymbol{x}_0) \subset \Omega$ such that $\bar{u} \leq \psi - C$ in $\mathcal{N}(\boldsymbol{x}_0)$, where $C > 0$. Let $w \in C_0^\infty(\mathcal{N}(\boldsymbol{x}_0))$, then there exists $\mu = \frac{C}{||w||_{L^\infty(\Omega)}}$ such that $v := \bar{u} \pm \mu w \leq \psi$. Choosing v as a test function in (2.2.2), the integration by part produces

$$\int_\Omega (-\operatorname{div}(\nabla\bar{u}) - f) w \,\mathrm{d}\boldsymbol{x} = 0, \quad \text{for all } w \in C_0^\infty(\mathcal{N}(\boldsymbol{x}_0)).$$

Then $-\operatorname{div}(\nabla\bar{u}) = f$ in $\mathcal{N}(\boldsymbol{x}_0)$ if $\bar{u}(\boldsymbol{x}_0) < \psi(\boldsymbol{x}_0)$ for $\boldsymbol{x}_0 \in \Omega$, which implies that (2.1.2a) is verified. Other boundary conditions (2.1.2c) and (2.1.2d) are easily derived from the definition of the closed convex set $\mathcal{K}$. □

2.B Normal trace

We recall here some classical notions on the normal trace of a function in H_{div}. Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain. Then there exists a surjective trace mapping $\gamma : H^1(\Omega) \to H^{1/2}(\partial\Omega)$. For $\boldsymbol{\psi} \in H_{\text{div}}$, the normal trace $\gamma_{\mathbf{n}}(\boldsymbol{\psi}) \in (H^{1/2}(\partial\Omega))'$ of $\boldsymbol{\psi}$ is defined by

$$\langle \gamma_{\mathbf{n}}(\boldsymbol{\psi}), \gamma(w) \rangle = \int_\Omega \boldsymbol{\psi} \cdot \nabla w \, \mathrm{d}\boldsymbol{x} + \int_\Omega \mathrm{div}\boldsymbol{\psi} \cdot w \, \mathrm{d}\boldsymbol{x}, \quad \text{for any } w \in H^1(\Omega),$$

where $\langle \cdot, \cdot \rangle$ denotes the duality product between $H^{1/2}(\partial\Omega)$ and $(H^{1/2}(\partial\Omega))'$. Stoke's formula in $H_0^1(\Omega) \times H_{\text{div}}(\Omega)$ shows that this definition makes sense, i.e. the right-hand side is unchanged if we apply to $\widetilde{w} \in H^1(\Omega)$ such that $\gamma(w) = \gamma(\widetilde{w})$.

Chapter 3

Application to hybrid mimetic mixed (HMM) methods

Abstract. As an application of the gradient discretisation method framework, we design and analyse the hybrid mimetic mixed (HMM) method to linear elliptic variational inequalities. We also propose an implementation methodology to solve the HMM method for linear variational inequalities in practice. For this purpose, we re-write the methods in the sense of classical finite volume discretisations, resulting in a linear system of equations. We provide different numerical results that demonstrate the accuracy of these schemes, and confirm our theoretical rates of convergence obtained under general grids.

3.1 Introduction

Using the gradient discretisation method framework, this chapter aims to develop the hybrid mimetic mixed (HMM) method, containing the mixed/hybrid mimetic finite difference methods, to linear elliptic variational inequalities. It also aims to compute the solution of the HMM scheme in practice and to provide some numerical experiments to illustrate the validity of theoretical results obtained in the previous chapter.

The notions of $\mathcal{T}$ (the polytopal mesh of Ω) defined in Definition 2.2.9 are still valid throughout this chapter. We also take $d = 2$ here. The HMM scheme is a family of the finite volume (FV) methods, which are based on cells unknowns $(u_K)_{K\in\mathcal{M}}$ and edges unknowns $(u_\sigma)_{\sigma\in\mathcal{E}_K}$ respectively approximating the values $(\bar{u}(\overline{\boldsymbol{x}}_K))_{K\in\mathcal{M}}$ (recall that $\overline{\boldsymbol{x}}_K$ is centre of mass of K) and $(\bar{u}(\overline{\boldsymbol{x}}_\sigma))_{\sigma\in\mathcal{E}_K}$ (recall that $\overline{\boldsymbol{x}}_\sigma$ is centre of mass of σ). Let us summarise the principle of the FV methods in the setting of simple PDEs,

$$\begin{aligned} -\mathrm{div}(\Lambda\nabla\bar{u}) &= f \quad \text{in } \Omega, \\ \bar{u} &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{3.1.1}$$

with the same assumptions on data given in Section 1.2. Throughout this chapter, let $\overline{F}_{K,\sigma}(\bar{u})$ denote to the exact flux on edge σ of a cell K given by $-\int_\sigma \Lambda\nabla\bar{u}\cdot\mathbf{n}_{K,\sigma}\,\mathrm{d}\boldsymbol{x}$. We integrate (3.1.1) on cell K to get

$$\sum_{\sigma\in\mathcal{E}_K} \overline{F}_{K,\sigma}(\bar{u}) = \int_K f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.1.2}$$

This equality is called the balance of fluxes equation and it can also directly be derived from the original physical model, see [34] for more details. The FV methods also maintain the physical conservation laws of the model,

$$\overline{F}_{K,\sigma}(\bar{u}) + \overline{F}_{L,\sigma}(\bar{u}) = 0, \qquad \text{for any } \sigma\in\mathcal{E}_K\cap\mathcal{E}_L. \tag{3.1.3}$$

The main idea of the classical FV schemes is to find $F_{K,\sigma}(u)$, which is expected to approximate the exact flux $\overline{F}_{K,\sigma}(\bar{u})$ in such a way that it satisfies the following discrete versions of the balance and conservativity laws equations:

$$\text{for all } K\in\mathcal{M} \ : \ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) = \int_K f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \tag{3.1.4}$$

$$\text{for all } \sigma\in\mathcal{E}_K\cap\mathcal{E}_L, \quad F_{K,\sigma}(u) + F_{L,\sigma}(u) = 0, \tag{3.1.5}$$

where the flux $F_{K,\sigma} = F_{K,\sigma}(u)$ is a linear function of the unknowns $(u_K)_{K\in\mathcal{M}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$.

In this chapter, we show that the HMM method fits into the gradient discretisation method framework and therefore its approximate solution provided by the HMM method satisfies the error estimates established in Section 2.2.

This chapter is organised as follows: in the next section we recall the idea of the HMM method in the setting of the above PDEs model. In Section 3.4 we design the HMM method for the linear Signorini problem, and establish a convergence rate for this scheme, as an application of the gradient discretisation method. We also show that this method can be reformulated in the sense of the balance and conservativity equations, we describe an implementation procedure and we discuss its convergence. Section 3.5.2 is concerned with studying the HMM method for the obstacle problem and a computation of this scheme's solution. We provide in Section 3.6 numerical results that demonstrate the excellent behaviour of this new scheme on test cases from the literature. We develop a test case for the Signorini problem with analytic solutions, and we use it to assess the practical convergence rates of the HMM methods. Further, to numerically confirm our theoretical results, we also perform numerical tests on different types of meshes. Finally, an appendix explains how to build a local matrix that is used to assemble the HMM matrix.

3.2 Three forms of the HMM method for (3.1.1)

As previously mentioned, the HMM method is a framework containing three methods, the hybrid finite volume (HFV) methods, the mixed finite volume (MFV) methods and the (mixed-hybrid) mimetic finite differences (MFD) methods, see [38]. Let us summarise the principle of these three methods. In this section, let the discrete unknowns space be defined as

$$X_{\mathcal{D},0} = \{v = ((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}) \ : \ v_K, v_\sigma \in \mathbb{R},\ v_\sigma = 0 \text{ for all } \sigma \in \mathcal{E}_{\rm ext}\}. \tag{3.2.1}$$

Let $\mathcal{F}$ be the set of discrete fluxes around all cells. The notation $\mathcal{F}_K$ denotes to the set of fluxes $F_K(u) = (F_{K,\sigma}(u))_{\sigma\in\mathcal{E}_K}$ restricted on the boundary of a given cell K, such that $F_{K,\sigma}(u)$ is still approximating $-\int_\sigma \Lambda\nabla\bar{u}\cdot\mathbf{n}_{K,\sigma}\,\mathrm{d}\boldsymbol{x}$. The central idea of all three methods is to express discrete equations of (3.1.1) in terms of discrete unknowns $(u_K)_{K\in\mathcal{M}}$, $(u_\sigma)_{\sigma\in\mathcal{E}}$ and $(F_{K,\sigma}(u))_{K,\sigma}$ by using approximation of the exact solution $\bar{u}$ and its exact gradient $\nabla\bar{u}$.

- **HFV method:** It is the form that is closest to the gradient scheme idea. A cell-wise constant gradient is defined by
$$\nabla_K v = \frac{1}{|K|}\sum_{\sigma\in\mathcal{E}_K}|\sigma|(v_\sigma - v_K)\mathbf{n}_{K,\sigma}. \tag{3.2.2}$$
It is proved in [37, Lemma B.6] that this gradient is linearly exact, that is, if φ is affine and $v_K = \varphi(\boldsymbol{x}_K)$ and $v_\sigma = \varphi(\overline{\boldsymbol{x}}_\sigma)$, then $\nabla_K v = \nabla\varphi$. Although this gradient is consistent, it does not satisfy a Poincaré inequality and therefore cannot be directly used in the weak formulation to approximate the bilinear term $\int_\Omega \Lambda\nabla u\cdot\nabla v\,\mathrm{d}\boldsymbol{x}$, that appears in the weak formulation of Problem (3.1.1). For this purpose, a stabilisation term is required,
$$R_K(v) = (v_\sigma - v_K - \nabla_K v\cdot(\boldsymbol{x}_\sigma - \boldsymbol{x}_K))_{\sigma\in\mathcal{E}_K} \in \mathbb{R}^{\mathrm{Card}(\mathcal{E}_K)}. \tag{3.2.3}$$
The HFV method for the weak formulation of (3.1.1) reads: seek $u = ((u_K)_{K\in\mathcal{M}}, (u_\sigma)_{\sigma\in\mathcal{E}}) \in X_{\mathcal{D},0}$, such that for any $v \in X_{\mathcal{D},0}$,
$$\sum_{K\in\mathcal{M}}|K|\Lambda_K\nabla_K u\cdot\nabla_K v + \sum_{K\in\mathcal{M}} R_K(v)^T\mathbb{B}_K R_K(u) = \sum_{K\in\mathcal{M}} v_K\int_K f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.2.4}$$
Here Λ_K is the value of Λ in the cell K (it is usual – although not mandatory – to assume that Λ is piecewise constant on the mesh) and, for $K\in\mathcal{M}$, $\mathbb{B}_K$ is a symmetric positive definite matrix of size $\mathrm{Card}(\mathcal{E}_K)$.
- **MFV method:** It is a standard format of the finite volume schemes and is based on the same discrete unknowns as in the HFV scheme. If $F_K \in \mathcal{F}$, then $\mathbf{v}_K(F_K)$ is defined by
$$\mathbf{v}_K(F_K) = -\frac{1}{|K|}\Lambda_K^{-1}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(\overline{\boldsymbol{x}}_\sigma - \overline{\boldsymbol{x}}_K), \quad \forall K\in\mathcal{M}. \tag{3.2.5}$$
This vector can be considered as the approximation of the gradient $\nabla\bar{u}$ in K. Taking the quantity
$$T_{K,\sigma}(F_K) = F_{K,\sigma} + \Lambda_K\mathbf{v}_K(F_K)\cdot\mathbf{n}_{K,\sigma}, \tag{3.2.6}$$
the space $\mathcal{F}_K$ can be equipped with an inner product defined by
$$\langle F_K, G_K\rangle_K = |K|\mathbf{v}(F_K)\cdot\Lambda_K\mathbf{v}(G_K) + T_K(G_K)^T\mathbb{E}_K T_K(F_K), \tag{3.2.7}$$

where $\mathbb{E}_K$ is a symmetric positive definite matrix. The formula controlling all discrete unknowns $(u_K)_{K\in\mathcal{M}}$, $(u_\sigma)_{\sigma\in\mathcal{E}}$ and $(F_{K,\sigma}(u))_{\sigma\in\mathcal{E}_K}$ is

$$\forall K \in \mathcal{M},\ \forall G_K \in \mathcal{F}_K : \quad \langle F_K(u), G_K\rangle_K = \sum_{\sigma\in\mathcal{E}_K} (u_K - u_\sigma) G_{K,\sigma}. \tag{3.2.8}$$

The MFV presentation for Problem (3.1.1) is defined by Equations [(3.1.4), (3.1.5), (3.2.5), (3.2.6), (3.2.7) and (3.2.8)].

- **MFD method:** Let $\widehat{X}_{\mathcal{D},0}$ be the restriction of the space $X_{\mathcal{D},0}$ on the values at cells; $\widehat{X}_{\mathcal{D},0} = \{\widehat{v} = (v_K)_{K\in\mathcal{M}} : v_K \in \mathbb{R}\}$. This space can be equipped with a standard L^2–inner product defined by

$$[\widehat{u}, \widehat{v}]_{\widehat{X}_{\mathcal{D},0}} = \sum_{K\in\mathcal{M}} |K| u_K v_K.$$

Given a symmetric definite positive matrix $\mathbb{M}_K$, the space $\mathcal{F}_K$ can be also endowed with the local inner product,

$$[F, G]_K = F_K^T \mathbb{M}_K G_K. \tag{3.2.9}$$

A discrete divergence operator $\mathcal{DIV} : \mathcal{F} \to \widehat{X}_{\mathcal{D},0}$ is defined by

$$\mathcal{DIV}^h G = \frac{1}{|K|} \sum_{\sigma\in\mathcal{E}_K} G_{K,\sigma}, \quad \forall K \in \mathcal{M}.$$

The definition of the MFD method is based on finding a matrix $\mathbb{M}_K$ that must satisfy the stability and the discrete Stoke formula conditions, respectively called **(S1)** and **(S2)** in [18],

(S1) There exists $s_\star > 0$, $S_\star > 0$ independent of the mesh such that

$$\forall K \in \mathcal{M}, \forall G \in \mathcal{F}\ : s_\star \sum_{\sigma\in\mathcal{E}_K} |K|(G_{K,\sigma})^2 \le [G, G]_K \le S_\star \sum_{\sigma\in\mathcal{E}_K} |K|(G_{K,\sigma})^2.$$

(S2) For all affine function v and for all $G \in \mathcal{F}$,

$$[(\Lambda\nabla v)^I, G]_K + \int_K v(\mathcal{DIV}^h G)_K \,\mathrm{d}\boldsymbol{x} = \sum_{\sigma\in\mathcal{E}_K} \frac{1}{|\sigma|} G_{K,\sigma} \int_\sigma v \ \mathrm{d}S,$$

where $((\Lambda\nabla v)^I)_\sigma^K = \int_\sigma -\Lambda_K \nabla v \cdot \mathbf{n}_{K,\sigma} \,\mathrm{d}S$.

The discrete flux operator $\mathcal{G}^h : \widehat{X}_{\mathcal{D},0} \to \mathcal{F}$ is defined as the adjoint operator of $\mathcal{DIV}^h$,

$$[F, \mathcal{G}^h u]_{\mathcal{F}} = [u, \mathcal{DIV}^h F]_{\widehat{X}_{\mathcal{D},0}}.$$

The MFD method to (3.1.1) is to find $(\widehat{u}, F) \in \widehat{X}_{\mathcal{D},0} \times \mathcal{F}$ such that

$$\mathcal{DIV}^h F = f_h, \quad F = \mathcal{G}^h u,$$

where f_h is the mean values of the source function f.

3.3 All HMM methods are gradient schemes

The operator $\Pi_{\mathcal{D}}$ and the piecewise constant gradient $\nabla_{\mathcal{D}}$ for the HMM methods are defined as follows:

$$\forall v \in X_{\mathcal{D},0}\,,\ \forall K \in \mathcal{M}\,:\,\Pi_{\mathcal{D}} v = v_K \text{ on } K. \tag{3.3.1}$$

$$\forall v \in X_{\mathcal{D},0},\ \forall K \in \mathcal{M}\,:\,\nabla_{\mathcal{D}} v = \nabla_K v + \frac{\sqrt{d}}{d_{K,\sigma}}(A_K R_K(v))_\sigma \mathbf{n}_{K,\sigma} \text{ on } \mathrm{co}(\{\sigma, \boldsymbol{x}_K\}), \tag{3.3.2}$$

where $\mathrm{co}(S)$ is the convex hull of the set S and

- $R_K(v)$ is defined by (3.2.3),
- A_K is an isomorphism of the vector space $\mathrm{Im}(R_K)$.

The right-hand side of (3.2.4) is equal to the right-hand side of the gradient scheme for the weak formulation of (3.1.1), (see (1.2.3) in Chapter 1). It is proved in [37, Proposition 12.5] that for any symmetric positive definite matrix $\mathbb{B}_K$, there exists an isomorphism A_K, such that for all $(u, v) \in X^2_{\mathcal{D},0}$,

$$|K|\Lambda_K \nabla_K u \cdot \nabla_K v + R_K(v)^T \mathbb{B}_K R_K(u) = \int_K \Lambda_K \nabla_{\mathcal{D}} u \cdot \nabla_{\mathcal{D}} v \,\mathrm{d}\boldsymbol{x}. \tag{3.3.3}$$

Moreover, it shows that this relation is satisfied, provided that, for all $(\xi, \eta) \in (\mathrm{Im}(R_K))^2$,

$$\xi^T \mathbb{B}_K \eta = (A_K(\xi))^T \mathbb{D}_K (A_K(\eta)), \tag{3.3.4}$$

where $\mathbb{D}_K = \mathrm{diag}\left(\frac{|\sigma|}{d_{K,\sigma}} \Lambda_K \mathbf{n}_{K,\sigma} \cdot \mathbf{n}_{K,\sigma}\right)$ is a diagonal definite positive matrix.

Summing (3.3.3) over $K \in \mathcal{M}$, the left-hand side of (3.2.4) is identical to the left-hand side of the gradient schemes (1.2.3). Therefore, using this gradient discretisation $(X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$ in the gradient scheme (1.2.3) gives the HMM method to (3.1.1).

The above scheme can also be expressed in terms of fluxes as a classical finite volume method. The fluxes on the edges of a cell K that correspond to $u \in X_{\mathcal{D},0}$ are defined by:

$$\begin{aligned}\forall v \in X_{\mathcal{D},0}, \quad \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma) &= |K|\Lambda_K \nabla_K u_K \cdot \nabla_K v_K + R_K^T(u) \mathbb{B}_K R_K(v) \\ &= \int_K \Lambda_K \nabla_{\mathcal{D}} u \cdot \nabla_{\mathcal{D}} v \,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{3.3.5}$$

Note that the flux $F_{K,\sigma}(u)$ thus constructed is an approximation of $-\int_\sigma \Lambda \nabla \bar{u} \cdot \mathbf{n}_{K,\sigma} \,\mathrm{d}\boldsymbol{x}$. Based on this relation, the HMM scheme (3.2.4) can be seen as: find u in $X_{\mathcal{D},0}$, such that for any v in $X_{\mathcal{D},0}$,

$$\sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} F(u)(v_K - v_\sigma) = \sum_{K \in \mathcal{M}} v_K \int_K f \,\mathrm{d}\boldsymbol{x}. \tag{3.3.6}$$

This formulation enables the equivalence between the HMM scheme (3.2.4) and the finite volumes scheme (3.1.4)–(3.1.5). Setting $v = 1$ only on a given cell K_0 and $v = 0$ on all edges and other cells yields (3.1.4) for $K = K_0$. The conservation law is obtained by taking $v = 1$ on a common edge σ between cells K and L and $v = 0$ on the other degrees of freedom.

Conversely, let u satisfy Problem (3.1.4)–(3.1.5); for $v \in X_{\mathcal{D},0}$, we have

$$\sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} F(u) v_K = \sum_{K \in \mathcal{M}} v_K \int_K f \,\mathrm{d}\boldsymbol{x}. \tag{3.3.7}$$

Thanks to the conservation equation and the boundary condition, we deduce

$$\begin{aligned}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)v_\sigma &= \sum_{\sigma\in\mathcal{E}_{\rm int}} F_{K,\sigma}(u)u_\sigma - \sum_{\sigma\in\mathcal{E}_{\rm ext}} F_{K,\sigma}(u)u_\sigma\\ &= 0.\end{aligned}$$

Combining this equation with (3.3.7) leads to u satisfies (3.3.6).

3.4 The HMM method for the Signorini problem

3.4.1 Construction of the HMM method

Let $\mathcal{T}$ be a polytopal mesh of Ω. The discrete space to consider is

$$X_{\mathcal{D},\Gamma_{2,3}} = \{v = ((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}) \ : \ v_K \in \mathbb{R}, v_\sigma \in \mathbb{R}, v_\sigma = 0 \text{ for all } \sigma\in\mathcal{E}_{\rm ext} \text{ such that } \sigma\subset\Gamma_1\}. \tag{3.4.1}$$

We assume here that the mesh is compatible with $(\Gamma_i)_{i=1,2,3}$, in the sense that for any $i = 1,2,3$, each boundary edge is either fully included in Γ_i or disjoint from this set.

The operators $\Pi_{\mathcal{D}}$ and $\nabla_{\mathcal{D}}$ are defined by (3.3.1) and (3.3.2). It is natural to consider a piecewise constant trace reconstruction:

$$\forall\sigma\in\mathcal{E}_{\rm ext} \ : \ \mathbb{T}_{\mathcal{D}}v = v_\sigma \text{ on } \sigma. \tag{3.4.2}$$

This reconstructed trace operator does not take values in $H^{1/2}_{\Gamma_1}(\partial\Omega)$, and the corresponding gradient discretisation $\mathcal{D}$ is therefore not admissible in the sense of Definition 2.2.3. However, for sufficiently regular $\bar{u}$, we can consider reconstructed traces in $L^2(\partial\Omega)$ only, see Remark 2.2.3.

With $\mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},\Gamma_{2,3}} \ : \ v_\sigma \le a$ on σ, for all $\sigma\in\mathcal{E}_{\rm ext}$ such that $\sigma\subset\Gamma_3\}$, the gradient scheme (2.2.3) corresponding to this gradient discretisation can be recast as

$$\left\{\begin{array}{l}\text{Find } u\in\mathcal{K}_{\mathcal{D}} \text{ such that, for all } v\in\mathcal{K}_{\mathcal{D}}\,,\\ \displaystyle\sum_{K\in\mathcal{M}}|K|\Lambda_K\nabla_K u\cdot\nabla_K(u-v) + \sum_{K\in\mathcal{M}} R_K(u-v)^T\mathbb{B}_K R_K(u) \le \sum_{K\in\mathcal{M}}(u_K - v_K)\int_K f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},\end{array}\right. \tag{3.4.3}$$

where Λ_K is the value of Λ in the cell K and, for $K\in\mathcal{M}$, $\mathbb{B}_K$ is a symmetric positive definite matrix of size $\mathrm{Card}(\mathcal{E}_K)$. This matrix is associated to A_K, see Section 3.3.

Assume the existence of $\theta > 0$ such that,

$$\begin{aligned}\max_{K\in\mathcal{M}}\left(\max_{\sigma\in\mathcal{E}_K}\frac{h_K}{d_{K,\sigma}} + \mathrm{Card}(\mathcal{E}_K)\right) + \max_{\sigma\in\mathcal{E}_{\rm int},\mathcal{M}_\sigma=K,L}\left(\frac{d_{K,\sigma}}{d_{L,\sigma}} + \frac{d_{L,\sigma}}{d_{K,\sigma}}\right)&\\ + \max\left\{\frac{|K|}{h_K|\sigma|} : K\in\mathcal{M}, \sigma\in\mathcal{E}_K\right\} \le \theta&\end{aligned} \tag{3.4.4}$$

and, for all $K\in\mathcal{M}$ and $\mu\in\mathbb{R}^{\mathcal{E}_K}$,

$$\begin{aligned}\frac{1}{\theta}\sum_{\sigma\in\mathcal{E}_K}|D_{K,\sigma}|\left|\frac{R_{K,\sigma}(\mu)}{d_{K,\sigma}}\right|^2 &\le \sum_{\sigma\in\mathcal{E}_K}|D_{K,\sigma}|\left|\frac{(A_K R_K(\mu))_\sigma}{d_{K,\sigma}}\right|^2\\ &\le \theta\sum_{\sigma\in\mathcal{E}_K}|D_{K,\sigma}|\left|\frac{R_{K,\sigma}(\mu)}{d_{K,\sigma}}\right|^2.\end{aligned} \tag{3.4.5}$$

Under the above assumptions on the mesh and on the matrices $\mathbb{B}_K$, it is shown in [37] that the constant $C_{\mathcal{D}}$ can be bounded above by quantities depending on the θ, but not on the mesh size, and that $W_{\mathcal{D}}(\boldsymbol{\psi})$ is of order

$\mathcal{O}(h_\mathcal{M})$ if $\boldsymbol{\psi}$ is in $(H^1)^d$. If $d \le 3$ and $\bar{u} \in H^2(\Omega)$, we can construct the interpolant $v_\mathcal{D} = ((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{M}})$ with $v_K = \bar{u}(\overline{\boldsymbol{x}}_K)$ and $v_\sigma = \bar{u}(\overline{\boldsymbol{x}}_\sigma)$ and, by [37, Lemma 12.8 and Proposition A.10], we have

$$S_\mathcal{D}(\bar{u}, v_\mathcal{D}) = ||\Pi_\mathcal{D} v_\mathcal{D} - \bar{u}||_{L^2(\Omega)} + ||\nabla_\mathcal{D} v_\mathcal{D} - \nabla u||_{L^2(\Omega)^d} \le C h_\mathcal{M} ||\bar{u}||_{H^2(\Omega)}.$$

Under the assumption that $\bar{u} \in W^{2,\infty}(\Omega)$, this estimate is also proved in [41] (with $||\bar{u}||_{H^2(\Omega)}$ replaced with $||\bar{u}||_{W^{2,\infty}(\Omega)}$). If $\bar{u} \in \mathcal{K}$ and a is constant, this interpolant $v_\mathcal{D}$ belongs to $\mathcal{K}_\mathcal{D}$. Moreover, given the definition (3.4.2) of $\mathbb{T}_\mathcal{D}$, we have $(\mathbb{T}_\mathcal{D} v_\mathcal{D})_{|\sigma} - \gamma(\bar{u})(\overline{\boldsymbol{x}}_\sigma) = v_\sigma - \bar{u}(\overline{\boldsymbol{x}}_\sigma) = 0$. Hence, using this $v_\mathcal{D}$ in Theorem 2.2.10, we see that the HMM scheme for the Signorini problem enjoys an order 1 rate of convergence. A nonconstant barrier a can be approximated by a piecewise constant barrier on $\mathcal{E}_{\rm ext}$ and Theorem 2.5.2 shows that, with this approximation, we still have an order 1 rate of convergence. Note that since this convergence also involves the gradients, a first order convergence is optimal for a low-order method such as the HMM method.

3.4.2 Recast as a finite volumes scheme

We show here that the HMM method given in the previous section can then be re-written in terms of the balance and conservativity of the fluxes.

With $a_\sigma \in L^2(\partial\Omega)$ a constant approximation of a, we consider the HMM scheme (3.4.3), replacing $\mathcal{K}_\mathcal{D}$ with

$$\widetilde{\mathcal{K}}_\mathcal{D} := \{v \in X_{\mathcal{D},\Gamma_{2,3}} \ : \ v_\sigma \le a_\sigma \text{ on } \sigma, \text{ for all } \sigma \in \mathcal{E}_{\rm ext} \text{ such that } \sigma \subset \Gamma_3\}.$$

Lemma 3.4.1. *Let $\mathcal{T}$ be admissible mesh in the sense of Definition 2.2.9. Assume that for any $\sigma \in \mathcal{E}$, there $i = 1, 2, 3$ such that $\sigma \subset \Gamma_i$. Under Hypothesis 2.2.1, u is a solution to Problem (3.4.3), in which $\mathcal{K}_\mathcal{D}$ has been replaced with $\widetilde{\mathcal{K}}_\mathcal{D}$ if and only if u is a solution to the following problem:*

$$\begin{aligned}
\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) &= m(K) f_K, \quad \forall K \in \mathcal{M}, && \text{(3.4.6a)}\\
F_{K,\sigma}(u) + F_{L,\sigma}(u) &= 0, \quad \textit{if } \sigma \in \mathcal{E}_K \cap \mathcal{E}_L, K \ne L, && \text{(3.4.6b)}\\
u_\sigma &= 0, \quad \forall \sigma \in \mathcal{E}_{\rm ext} \textit{ such that } \sigma \subset \Gamma_1, && \text{(3.4.6c)}\\
F_{K,\sigma}(u) &= 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \textit{ such that } \sigma \subset \Gamma_2, && \text{(3.4.6d)}\\
F_{K,\sigma}(u)(u_\sigma - a_\sigma) &= 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \textit{ such that } \sigma \subset \Gamma_3, && \text{(3.4.6e)}\\
-F_{K,\sigma}(u) &\le 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \textit{ such that } \sigma \subset \Gamma_3, && \text{(3.4.6f)}\\
u_\sigma &\le a_\sigma, \quad \forall \sigma \in \mathcal{E}_{\rm ext} \textit{ such that } \sigma \subset \Gamma_3. && \text{(3.4.6g)}
\end{aligned}$$

Here, the flux $F_{K,\sigma}(u)$ is given by (3.3.5) and $f_K = \frac{1}{\mathrm{m}(K)} \int_K f \,\mathrm{d}\boldsymbol{x}$, for any $K \in \mathcal{M}$.

Remark 3.4.1. The above problem (3.4.6) is the HMM discretisation of the strong Signorini problem (2.1.1) based on the physical characteristics of the model. Integrating (2.1.1a) over a cell K and using Stokes formula give the fluxes balance equation (3.4.6a). The boundary data on Γ_1 and on the impermeable surface Γ_2 are respectively approximated by (3.4.6c) and (3.4.6d). Finally, the Signorini boundary conditions, controlling the flow and the pressure of water on the free boundary Γ_3, are naturally expressed by (3.4.6e)–(3.4.6g).

Proof of Lemma 3.4.1. The proof is inspired from [61]. Using the definition of flux (3.3.5), the equivalent formulation to Problem (3.4.3), written with $\widetilde{\mathcal{K}}_\mathcal{D}$, is:

$$\left\{\begin{array}{l}
\text{Find } u \in \widetilde{\mathcal{K}}_\mathcal{D} \text{ such that, for all } v \in \widetilde{\mathcal{K}}_\mathcal{D}\,,\\
\displaystyle\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_K - u_\sigma) - \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma) \le \sum_{K\in\mathcal{M}} (u_K - v_K) \int_K f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.
\end{array}\right. \quad (3.4.7)$$

First, assume that u is a solution to (3.4.6). With $v \in \widetilde{\mathcal{K}}_{\mathcal{D}}$, multiplying (3.4.6a) by $(u_K - v_K)$ and summing over control volumes yield

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_K - v_K) = \sum_{K\in\mathcal{M}} (u_K - v_K)\int_K f\,\mathrm{d}\boldsymbol{x}.$$

We introduce u_σ and v_σ to obtain

$$\begin{aligned}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_K - u_\sigma) - \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma) + \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_\sigma - v_\sigma)\\ = \sum_{K\in\mathcal{M}} (u_K - v_K)\int_K f\,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{3.4.8}$$

If we prove that the third term on the left-hand side is nonnegative, we deduce that u satisfies (3.4.7). Since $u_\sigma = v_\sigma = 0$ whenever $\sigma \in \mathcal{E}_{\mathrm{ext}}$ such that $\sigma \subset \Gamma_1$ and $F_{K,\sigma}(u) = 0$ whenever $\sigma \in \mathcal{E}_{\mathrm{ext}}$ such that $\sigma \in \Gamma_2$, we have

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_\sigma - v_\sigma) = \sum_{\sigma\in\mathcal{E}|\sigma\in\mathcal{E}_K\cap\mathcal{E}_L} (F_{K,\sigma}(u) + F_{L,\sigma}(u))(u_\sigma - v_\sigma) + \sum_{K\in\mathcal{M}}\sum_{\substack{\sigma\in\mathcal{E}_K\\ \sigma\subset\Gamma_3}} F_{K,\sigma}(u)(u_\sigma - v_\sigma).$$

Thanks to the conservativity condition (3.4.6b), the first term in the right-hand side also vanishes. Introduce a_σ in this relation to get

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_\sigma - v_\sigma) = \sum_{\substack{\sigma\in\mathcal{E}_K\\ \sigma\subset\Gamma_3}} F_{K,\sigma}(u)(u_\sigma - a_\sigma) + \sum_{\substack{\sigma\in\mathcal{E}_K\\ \sigma\subset\Gamma_3}} F_{K,\sigma}(u)(a_\sigma - v_\sigma).$$

Owing to (3.4.6e), the first term on the right-hand side is equal to zero. From (3.4.6f) and the fact that $v \in \widetilde{\mathcal{K}}_{\mathcal{D}}$, the last term on the right-hand side is nonnegative and thus the proof of the first part is complete.

Let us now prove the converse. Assume that $u \in \widetilde{\mathcal{K}}_{\mathcal{D}}$ is a solution to Problem (3.4.7). Let $L \in \mathcal{M}$ and $v = u \pm w$ with $w_L = 1$, $w_K = 0$ for any $K \in \mathcal{M}$ and $L \neq K$, and $w_\sigma = 0$ for any $\sigma \in \mathcal{E}$. v is an element of $\widetilde{\mathcal{K}}_{\mathcal{D}}$. Thus, inserting this v in inequality of Problem (3.4.7) leads to, respectively

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)w_K - \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)w_\sigma \geq \sum_{K\in\mathcal{M}} w_K \int_K f\,\mathrm{d}\boldsymbol{x}.$$

and

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)w_K - \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)w_\sigma \leq \sum_{K\in\mathcal{M}} w_K \int_K f\,\mathrm{d}\boldsymbol{x}.$$

Since $w_\sigma = 0$ for any $\sigma \in \mathcal{E}$ and $w_K = 0$ for all $L \neq K \in \mathcal{M}$, splitting the sum over edges and cells in the above two inequalities yields

$$\sum_{\sigma\in\mathcal{E}_L} F_{L,\sigma}(u) = \int_L f\,\mathrm{d}\boldsymbol{x}.$$

In a similar way, let $s \in \mathcal{E}_{\mathrm{int}}$ and take $v = u \pm w$ in (3.4.7) with $w \in X_{\mathcal{D},\Gamma_{2,3}}$ such that $w_K = 0$ for all $K \in \mathcal{M}$ and $w_s = 1$ for $s \in \mathcal{E}_{\mathrm{int}}$ and $w_\sigma = 0$ for all $\sigma \in \mathcal{E}$, $s \neq \sigma$. (3.4.7) implies $F_{K,s}(u) + F_{L,s}(u) = 0$ for K, L such that $s \in \mathcal{E}_K \cap \mathcal{E}_L$.

Pick $s \in \mathcal{E}_{\mathrm{ext}}$ such that $s \subset \Gamma_2$ and choose $v = u \pm w$ with $w \in X_{\mathcal{D},\Gamma_{2,3}}$ such that $w_K = 0$ for all $K \in \mathcal{M}$ and $w_s = 1$ for $s \in \mathcal{E}_{\mathrm{ext}}$ such that $s \subset \Gamma_2$ and $w_\sigma = 0$ for all $\sigma \in \mathcal{E}$, $s \neq \sigma$. Inserting this v in (3.4.7) gives $F_{K,s}(u) = 0$, which proves (3.4.6d). Since the solution $u \in \widetilde{\mathcal{K}}_{\mathcal{D}}$, (3.4.6c) and (3.4.6g) are obviously fulfilled.

To prove (3.4.6f), let $s \in \mathcal{E}_{\text{ext}}$ such that $s \subset \Gamma_3$ and choose $w \in X_{\mathcal{D},\Gamma_{2,3}}$ such that $w_K = 0$ for all $K \in \mathcal{M}$ and $w_s = -1$ and $w_\sigma = 0$ for all $\sigma \in \mathcal{E}$, $s \neq \sigma$. Setting $v = u + w$ in (3.4.7) ($v \in \widetilde{\mathcal{K}}_{\mathcal{D}}$), we infer

$$-F_{K,s}(u) \leq 0, \quad s \in \mathcal{E}_K, \text{ such that } s \subset \Gamma_3.$$

To verify (3.4.6e), let $E = \{\sigma \in \mathcal{E}_{\text{ext}} : \sigma \subset \Gamma_3 \text{ and } u_\sigma < a_\sigma\}$. Choose $s \in E$, we only need to show that $F_{K,s}(u)(a_s - u_s) = 0$. Letting $t = a_s - u_s > 0$, define $v_1 = u + tw$ and $v_2 = u - tw$ such that $w_K = 0$ for any $K \in \mathcal{M}$ and $w_s = -1$, and $w_\sigma = 0$ for all $\sigma \in \mathcal{E}_{\text{ext}}$, $\sigma \subset \Gamma_3$, and $s \neq \sigma$. Since $v_{2s} = a_s$ and $v_{1s} = u_s - t \leq a_s$, we see that $v_1, v_2 \in \widetilde{\mathcal{K}}_{\mathcal{D}}$. Setting $v = v_1$ (respectively $v = v_2$) as a test function in (3.4.7), and reasoning as previously, we obtain

$$F_{K,s}(a_s - u_s) \geq 0 \quad \text{and} \quad F_{K,s}(a_s - u_s) \leq 0,$$

which concludes the proof. □

3.4.3 Resolution of the HMM for the Signorini problem

This section explains how to compute the solution of the HMM scheme (3.4.6) in practice. To deal with the non–linearity caused by the Signorini boundary conditions, the iteration monotonicity algorithm proposed in [59] can be naturally applied to this flux approximation formulation . This algorithm has been used to calculate the solution of $\mathbb{P}1$ finite elements and two-point flux approximation finite volumes to variational inequality in [59]. Its convergence and efficiency relies on the discrete maximum–minimum principles. We note that by [77, Section 4.2] the HMM method is monotone when applied to isotropic diffusion on meshes made of acute triangles. In this case, the convergence of the monotonicity algorithm can be established as in [59] for finite element and two-point finite volume methods. In our numerical tests, we noticed that, even when applied on meshes for which the monotonicity of HMM is unknown or fails, the monotonicity algorithm actually still converges.

The monotonicity algorithm approximates the solution to the discrete variational inequalities by sequences of solutions to linear problems. If we assume that there are two known sets $\mathbb{G}$ and $\mathbb{H}$ such that $\mathbb{G} \cup \mathbb{H} = \Gamma_3$, $\mathbb{G} \cap \mathbb{H} = \emptyset$ and

$$\begin{aligned} u_\sigma &\leq a_\sigma, \text{ for all } \sigma \in \mathbb{G}, \\ -F_{K,\sigma}(u) &< 0 \text{ for all } K \in \mathcal{M}, \text{ s.t. } \sigma \in \mathcal{E}_K \cap \mathbb{H}. \end{aligned} \tag{3.4.9}$$

The set $\mathbb{H}$ is not any set that satisfies (3.4.9) but the largest one that gathers all $\sigma \subset \Gamma_3$ such that $-F_{K,\sigma}(u) < 0$. Therefore, the discrete Signorini boundary conditions (3.4.6e)–(3.4.6g) would be equivalent to

$$\begin{aligned} F_{K,\sigma}(u) &= 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathbb{G}, \\ u_\sigma &= a_\sigma, \quad \forall \sigma \in \mathbb{H}. \end{aligned} \tag{3.4.10}$$

The monotonicity algorithm (explained in Algorithm 1) is an iteration process; at each iteration we solve (3.4.6a)–(3.4.6d) together with (3.4.10), which is a square linear system on the unknown $(u_K)_{K\in\mathcal{M}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$. At each step, we also iterate on all edges included in Γ_3 to switch edges that break the constraints (3.4.9).

Given that the gradients $(\nabla_{K,\sigma} u)_{\sigma\in\mathcal{E}_K}$ are constructed based on the discrete unknowns $(u_\sigma)_{\sigma\in\mathcal{E}_K}$ and u_K, the fluxes can be expressed in terms of these unknowns,

$$F_{K,\sigma}(u) = \sum_{\sigma'\in\mathcal{E}_K} W^K_{\sigma,\sigma'}(u_K - u_\sigma), \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K, \tag{3.4.12}$$

where $(W^K_{\sigma,\sigma'})_{\sigma,\sigma'\in\mathcal{E}_K^2}$ is a symmetric positive definite matrix, whose construction is given in Appendix 3.A. There are two possible ways to solve the linear system (3.4.11a)–(3.4.11f) in Algorithm 1. For simplicity, we drop the indexes (n).

Algorithm 1 Monotonicity algorithm for the Signorini model

Set $\mathbb{G}^{(0)} = \{\sigma \in \mathcal{E} \ : \ \sigma \subset \Gamma_3\}$ and $\mathbb{H}^{(0)} = \emptyset$
Set $N = \mathrm{Card}(\mathbb{G}^{(0)})$ ▷ Theoretical bound on the iterations
while $n \le N$ **do**
  $\mathbb{G}^{(n)}$ and $\mathbb{H}^{(n)}$ being known, find the solution $u^{(n)}$ to the system

$$\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}(u^{(n)}) = m(K) f_K, \quad \forall K \in \mathcal{M}, \tag{3.4.11a}$$

$$F_{K,\sigma}(u^{(n)}) + F_{L,\sigma}(u^{(n)}) = 0, \quad \forall \sigma \in \mathcal{E}_K \cap \mathcal{E}_L, K \neq L \tag{3.4.11b}$$

$$u_\sigma^{(n)} = 0, \quad \forall \sigma \in \mathcal{E}_{\mathrm{ext}} \text{ such that } \sigma \subset \Gamma_1, \tag{3.4.11c}$$

$$F_{K,\sigma}(u^{(n)}) = 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \subset \Gamma_2, \tag{3.4.11d}$$

$$F_{K,\sigma}(u^{(n)}) = 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \in \mathbb{G}^{(n)}, \tag{3.4.11e}$$

$$u_\sigma^{(n)} = a_\sigma, \quad \forall \sigma \in \mathcal{E}_{\mathrm{ext}} \text{ such that } \sigma \in \mathbb{H}^{(n)}. \tag{3.4.11f}$$

  Set $\mathbb{G}^{(n+1)} = \{\sigma \in \mathbb{G}^{(n)} \ : \ u_\sigma^{(n)} \le a_\sigma\} \cup \{\sigma \in \mathbb{H}^{(n)} \ : \ -F_{K,\sigma}(u^{(n)}) \ge 0\}$
  Set $\mathbb{H}^{(n+1)} = \{\sigma \in \mathbb{H}^{(n)} \ : \ -F_{K,\sigma}(u^{(n)}) < 0\} \cup \{\sigma \in \mathbb{G}^{(n)} \ : \ u_\sigma^{(n)} > a_\sigma\}$
  if $\mathbb{G}^{(n+1)} = \mathbb{G}^{(n)}$ and $\mathbb{H}^{(n+1)} = \mathbb{H}^{(n)}$ **then**
    Exit “while” loop
  end if
end while
Set $u = u^{(n)}$, $\mathbb{G} = \mathbb{G}^{(n+1)}$ and $\mathbb{H} = \mathbb{H}^{(n+1)}$

- **Without elimination of cells unknowns:** We produce a linear system, whose matrix corresponds to the cells and to the edges. This system, therefore, is of size $\mathrm{Card}(\mathcal{E}) + \mathrm{Card}(\mathcal{M})$. Plugging (3.4.12) in the linear system (3.4.11a), (3.4.11b) and (3.4.11d) directly provides the following linear equations:

$$\forall K \in \mathcal{M}, \sum_{\sigma \in \mathcal{E}_K} \sum_{\sigma' \in \mathcal{E}_K} W^K_{\sigma,\sigma'}(u_K - u_{\sigma'}) = f_K, \tag{3.4.13}$$

$$\forall \sigma \in \mathcal{E}_K \cap \mathcal{E}_L, \sum_{\sigma' \in \mathcal{E}_K} W^K_{\sigma,\sigma'}(u_K - u_{\sigma'}) + \sum_{\sigma' \in \mathcal{E}_L} W^L_{\sigma,\sigma'}(u_L - u_{\sigma'}) = 0, \tag{3.4.14}$$

$$\forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ s.t. } \sigma \subset \Gamma_2 \text{ or } \sigma \in \mathbb{G}, \sum_{\sigma' \in \mathcal{E}_K} W^K_{\sigma,\sigma'}(u_K - u_{\sigma'}) = 0. \tag{3.4.15}$$

The entries of the system matrix corresponding to the edges which are of Dirichlet boundary condition are directly provided by Equations (3.4.11c) and (3.4.11f).

In the assembly of the system, all elements on the right-hand side are zeros except the ones corresponding to degrees of freedom associated to cells K and to edges $\sigma \in \mathbb{H}$. Only the right-hand side of Equation (3.4.15) needs to be modified in the case of nonhomogeneous Neumann and Dirichlet BCs.

- **The hybrid resolution (elimination of cells unknowns):** The size of the system given in the previous way can be reduced. As in [24, 36], algebraic processes are used to eliminate the discrete unknowns $(u_K)_{K\in\mathcal{M}}$ in order to produce a sparse linear system with $\mathrm{Card}(\mathcal{E})$ unknowns. Let $b_{K,\sigma'} = \sum_{\sigma\in\mathcal{E}_K} -W^K_{\sigma,\sigma'}$ and $b_K = \sum_{\sigma'\in\mathcal{E}_K} b_{K,\sigma'}$. From (3.4.13), it follows that

$$u_K = \frac{1}{b_K}\left(\mathrm{m}(K) f_K + \sum_{\sigma' \in \mathcal{E}_K} b_{K,\sigma'} u_{\sigma'}\right), \quad \forall K \in \mathcal{M}. \tag{3.4.16}$$

This equation is used to compute the discrete unknowns $(u_K)_{K\in\mathcal{M}}$ if $(u_\sigma)_{\sigma\in\mathcal{E}}$ are known. Reporting (3.4.16) in (3.4.12) and using the conservation equation (3.4.11b), we have

$$\begin{aligned}\sum_{\sigma' \in \mathcal{E}_K}\left(-W^K_{\sigma,\sigma'} - \frac{b_{K,\sigma} b_{K,\sigma'}}{b_K}\right) u_{\sigma'} + \sum_{\sigma' \in \mathcal{E}_L}\left(-W^L_{\sigma,\sigma'} - \frac{b_{L,\sigma} b_{L,\sigma'}}{b_L}\right) u_{\sigma'} \\ = \frac{b_{K,\sigma}}{b_K}\mathrm{m}(K) f_K + \frac{b_{L,\sigma}}{b_L}\mathrm{m}(K) f_L, \quad \forall \sigma \in \mathcal{E}_K \cap \mathcal{E}_L,\end{aligned} \tag{3.4.17}$$

that controls the relation between the interior edges. For the exterior edges involving the Neumann boundary condition, again, we report (3.4.12) in (3.4.16) and substitute the flux into the equation $F_{K,\sigma} = 0$ to obtain for all $K \in \mathcal{M}$ and for all $\sigma \in \mathcal{E}_K$ such that $\sigma \subset \Gamma_2$ or $\sigma \in \mathbb{G}$,

$$\sum_{\sigma' \in \mathcal{E}_K}\left(-W^K_{\sigma,\sigma'} - \frac{b_{K,\sigma} b_{K,\sigma'}}{b_K}\right) u_{\sigma'} = \frac{b_{K,\sigma}}{b_K}\mathrm{m}(K) f_K. \tag{3.4.18}$$

Finally, other edges $\sigma \in \mathcal{E}_{\mathrm{ext}}$ such that $\sigma \subset \Gamma_2$ or $\sigma \in \mathbb{H}$ can be easily handled by Equations (3.4.11c) and (3.4.11f). Equations (3.4.17), (3.4.18), (3.4.11c) and (3.4.11f) provide a symmetric linear system, whose unknowns are $(u_\sigma)_{\sigma\in\mathcal{E}}$.

nonhomogeneous Neumann boundary condition can be easily included in the previous computation by adding the boundary data to the right-hand side of Equations (3.4.18).

Based on the discrete maximum and minimum principles, it is proved in [59] that the iterations number of the monotonicity algorithm is theoretically bounded by the number of edges included in Γ_3. As previously stated, satisfying the maximum principle is a sufficient condition, but not a necessary one, to ensure the convergence of the algorithm as the number of iterations might be still bounded, even though the monotonicity property is not satisfied. The following results investigate the convergence of the algorithm when such a bound exists. Their proofs are inspired from [59].

Lemma 3.4.2. *Under Hypothesis 2.2.1, for each iteration n, there exists a unique solution to Problem* (3.4.11) *introduced in Algorithm 1.*

Proof. Let $g = ((g_K)_{K\in\mathcal{M}}, (g_\sigma)_{\sigma\in\mathcal{E}})$, where $g_K \in \mathbb{R}$ and $g_\sigma \in \mathbb{R}$, be defined by $g_K = 0$ for all $K \in \mathcal{M}$, $g_\sigma = 0$ for all $\sigma \in \mathcal{E}_{\mathrm{ext}}$ such that $\sigma \subset \Gamma_1$ and $g_\sigma = a_\sigma$ for all $\sigma \in \mathbb{H}^{(n)}$. Consider

$$X_{\mathcal{D},\Gamma_2} = \{v = ((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}) \;:\; v_K \in \mathbb{R}, v_\sigma \in \mathbb{R}, v_\sigma = 0 \text{ for all } \sigma \in \mathcal{E}_{\mathrm{ext}} \text{ such that } \sigma \subset \Gamma_1 \text{ or } \sigma \in \mathbb{H}^{(n)}\}.$$

The weak formulation of Problem (3.4.11) is:

$$\begin{aligned}
&\text{Find } u^{(n)} \in X_{\mathcal{D},\Gamma_2} + g \text{ such that for all } v \in X_{\mathcal{D},\Gamma_2},\\
&\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u^{(n)})(v_K - v_\sigma) = \sum_{K\in\mathcal{M}} v_K \int_K f \,\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

This problem is the linear system coming from the HMM method for mixed BCs, for which we know there is a solution. □

Lemma 3.4.3. *Let Hypothesis 2.2.1 hold and $(\mathbb{G}^{(n)})_{n\in\mathbb{N}}$ and $(\mathbb{H}^{(n)})_{n\in\mathbb{N}}$ be the sets determined by applying the monotonicity Algorithm 1. If there exists $N \in \mathbb{N}$ such that $\mathbb{G}^{(N)} = \mathbb{G}^{(N+1)}$ and $\mathbb{H}^{(N)} = \mathbb{H}^{(N+1)}$, then the solution $u^{(N)}$ to* (3.4.11) *is also the unique solution to Problem* (3.4.6)*.*

Proof. Set $\mathbb{G} = \mathbb{G}^{(N)}$, $\mathbb{H} = \mathbb{H}^{(N)}$ and $u = u^{(N)}$. Thus u is a solution to the problem stated in the algorithm. We only need to verify the discrete Signorini conditions (3.4.6e)–(3.4.6g). Due to $\mathbb{H}^{(N)} = \mathbb{H}^{(N+1)}$, we have

$$F_{K,\sigma}(u) < 0, \quad \forall\sigma \in \mathbb{H},\ \forall K \in \mathcal{M}.$$

From (3.4.11e), one obtains

$$F_{K,\sigma}(u) = 0, \quad \forall\sigma \in \mathbb{G},\ \forall K \in \mathcal{M}.$$

Hence, (3.4.6f) holds since $\mathbb{G}\cup\mathbb{H} = \Gamma_3$. Likewise, from (3.4.11f) and the fact that $\mathbb{G}^{(N)} = \mathbb{G}^{(N+1)}$, it is deduced that

$$\begin{aligned}
u_\sigma &\le a_\sigma, \quad \forall\sigma \in \mathbb{G},\\
u_\sigma &= a_\sigma, \quad \forall\sigma \in \mathbb{H},
\end{aligned}$$

which verifies (3.4.6g). It is clear that (3.4.6e) follows from the property $\Gamma_3 = \mathbb{G}\cup\mathbb{H}$ and $\mathbb{G}\cap\mathbb{H} = \emptyset$. □

3.5 The HMM method for the obstacle problem

The idea in the obstacle problem is very similar to the Signorini problem. We nevertheless provide the details here, for the sake of completeness.

3.5.1 Construction of the HMM method

We still take $\mathcal{T}$ a polytopal mesh of Ω, and we consider $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$, where the elements $X_{\mathcal{D},0}$, $\Pi_{\mathcal{D}}$ and $\nabla_{\mathcal{D}}$ are defined by (3.2.1), (3.3.1) and (3.3.2). Using $\mathcal{D}$ in (2.2.7), we obtain the HMM method for the obstacle problem.

Setting $\mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},0} \ : \ v_K \le \psi \text{ on } K, \text{ for all } K \in \mathcal{M}\}$, this method reads

$$\left\{ \begin{array}{l} \text{Find } u \in \mathcal{K}_{\mathcal{D}} \text{ such that, for all } v \in \mathcal{K}_{\mathcal{D}}\,, \\ \displaystyle\sum_{K\in\mathcal{M}} |K|\Lambda_K \nabla_K u \cdot \nabla_K(u-v) + \sum_{K\in\mathcal{M}} R_K(u-v)^T \mathbb{B}_K R_K(u) \le \sum_{K\in\mathcal{M}} (u_K - v_K) \int_K f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \end{array} \right. \tag{3.5.1}$$

Using the same interpolant as in Section 3.4.1, we see that Theorem 2.2.11 (for a constant barrier ψ) and Theorem 2.5.4 (for a piecewise constant approximation of ψ) provide an order 1 convergence rate, under Assumptions (3.4.4) and (3.4.5).

3.5.2 Recast as a finite volumes scheme

The following lemma shows that the above HMM scheme can then be reformulated by the means of the balance and conservativity of the fluxes, a useful formulation for computational purposes. To do this, we replace the set $\mathcal{K}_{\mathcal{D}}$ in the scheme (3.5.1) with the set

$$\widetilde{\mathcal{K}}_{\mathcal{D}} := \{v \in X_{\mathcal{D},0} \ : \ v_K \le \psi_K \text{ on } K, \text{ for all } K \in \mathcal{M}\},$$

where $\psi_K \in L^2(\Omega)$ is an approximation of ψ on K, see Section 2.5.

Lemma 3.5.1. *Let $\mathcal{D}$ be admissible mesh. Under Hypothesis 2.2.2, u is a solution to the HMM scheme* (3.5.1)*, in which $\mathcal{K}_{\mathcal{D}}$ has been replaced with $\widetilde{\mathcal{K}}_{\mathcal{D}}$ if and only if u is a solution to the following problem:*

$$\left(-\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) + m(K) f_K\right)(\psi_K - u_K) = 0, \quad \forall K \in \mathcal{M}, \tag{3.5.2a}$$

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) \le m(K) f_K, \quad \forall K \in \mathcal{M}, \tag{3.5.2b}$$

$$u_K \le \psi_K, \quad \forall K \in \mathcal{M}, \tag{3.5.2c}$$

$$F_{K,\sigma}(u) + F_{L,\sigma}(u) = 0, \quad \textit{if } \sigma \in \mathcal{E}_K \cap \mathcal{E}_L, K \ne L, \tag{3.5.2d}$$

$$u_\sigma = 0, \quad \forall \sigma \in \mathcal{E}_{\mathrm{ext}}, \tag{3.5.2e}$$

where $f_K = \frac{1}{\mathrm{m}(K)} \int_K f \,\mathrm{d}\boldsymbol{x}$, for any $K \in \mathcal{M}$.

Remark 3.5.1. The above problem describes the HMM formulation to the obstacle model (2.1.2). By integrating $-\mathrm{div}(\Lambda\nabla\bar{u}) \le f$ on a control volume K leads to Inequality (3.5.2b). Equation (3.5.2a) is naturally the discretisation form of (2.1.2b). Inequality (3.5.2c) and Equation (3.5.2e) respectively translate the continuous barrier inequality (2.1.2c) and the homogeneous Dirichlet boundary condition (2.1.2d).

Proof of Lemma 3.5.1. The idea of the proof is taken from [61]. Based on the definition of flux (3.3.5), the discrete problem (3.5.1), with $\widetilde{\mathcal{K}}_{\mathcal{D}}$ can be formulated as:

$$\left\{ \begin{array}{l} \text{Find } u \in \widetilde{\mathcal{K}}_{\mathcal{D}} \text{ such that, for all } v \in \widetilde{\mathcal{K}}_{\mathcal{D}}\,, \\ \displaystyle\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_K - u_\sigma) - \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma) \le \sum_{K\in\mathcal{M}} (u_K - v_K) \int_K f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \end{array} \right. \tag{3.5.3}$$

Assume that u satisfies (3.5.2) and let us show that u is a solution to (3.5.3). It is clear that

$$\sum_{K\in\mathcal{M}}\left(-\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)+\int_K f\,\mathrm{d}\boldsymbol{x}\right)(u_K-v_K)=\sum_{K\in\mathcal{M}}\left(-\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)+\int_K f\,\mathrm{d}\boldsymbol{x}\right)(u_K-\psi_K)$$
$$+\sum_{K\in\mathcal{M}}\left(-\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)+\int_K f\,\mathrm{d}\boldsymbol{x}\right)(\psi_K-v_K).$$

Thanks to (3.5.2a), the first term in the right-hand side is zero. We use (3.5.2b) and the fact that $v\in\widetilde{\mathcal{K}}_{\mathcal{D}}$ (i.e., $v_K\le\psi_K,\forall K\in\mathcal{M}$), to see that the last term is nonnegative. Hence, it follows that

$$\sum_{K\in\mathcal{M}}\left(-\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)+\int_K f\,\mathrm{d}\boldsymbol{x}\right)(u_K-v_K)\ge 0,$$

which gives

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)(u_K-v_K)\le\sum_{K\in\mathcal{M}}(u_K-v_K)\int_K f\,\mathrm{d}\boldsymbol{x}.$$

Adding and subtracting u_σ and v_σ in this relation lead to

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)(u_K-u_\sigma)-\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)(v_K-v_\sigma)\quad+\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)(u_\sigma-v_\sigma)$$
$$\le\sum_{K\in\mathcal{M}}(u_K-v_K)\int_K f\,\mathrm{d}\boldsymbol{x}.$$

By writing

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)(u_\sigma-v_\sigma)=\sum_{\sigma\in\mathcal{E}|\sigma\in\mathcal{E}_K\cap\mathcal{E}_L}(F_{K,\sigma}(u)+F_{L,\sigma}(u))(u_\sigma-v_\sigma)+\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\rm ext}}F_{K,\sigma}(u)(u_\sigma-v_\sigma),$$

the left-hand side vanishes due to the conservation of fluxes (3.5.2d) and the zero boundary condition (3.5.2e). Therefore, we deduce that u is a solution to (3.5.3).

Let now u be a solution to Problem (3.5.3). We note that (3.5.2c) and (3.5.2e) immediately follow from the fact that u belongs to the set $\widetilde{\mathcal{K}}_{\mathcal{D}}$. Pick $L\in\mathcal{M}$ and set $v=u-w$ as a test function in (3.5.3) with $w\in X_{\mathcal{D},0}$ such that $w_L=1$ and $w_K=0$ for any $K\in\mathcal{M}$ and $K\neq L$, and $w_\sigma=0$ for any $\sigma\in\mathcal{E}$. We deduce that

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)(w_K-w_\sigma)\le\sum_{K\in\mathcal{M}}w_K\int_K f\,\mathrm{d}\boldsymbol{x},$$

which becomes, since $w_\sigma=0$ for any $\sigma\in\mathcal{E}$,

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u)w_K\le\sum_{K\in\mathcal{M}}w_K\int_K f\,\mathrm{d}\boldsymbol{x}.$$

By splitting control volumes, it follows that

$$\sum_{\sigma\in\mathcal{E}_L}F_{L,\sigma}(u)-\int_L f\,\mathrm{d}\boldsymbol{x}\le 0.$$

It remains to verify (3.5.2a). Let $S=\{K\in\mathcal{M}\ :\ u_K<\psi_K\}$. Choose $L\in S$ and let us prove that $\sum_{\sigma\in\mathcal{E}_L}F_{L,\sigma}(u)=\int_L f\,\mathrm{d}\boldsymbol{x}$. Consider $w\in X_{\mathcal{D},0}$ such that $w_L=-1$ and $w_K=0$ for any $K\in\mathcal{M}$ and $K\neq L$, and $w_\sigma = 0$ for any $\sigma \in \mathcal{E}$. Letting $v_1 = u + \mu w$ and $v_2 = u - \mu w$ where $\mu_K = \psi_K - u_K > 0$, we see that both v_1 and v_2 are in $\widetilde{\mathcal{K}}_{\mathcal{D}}$. Setting $v := v_1$ (respectively $v := v_2$) in (3.5.3) and reasoning as previously (splitting the sum over the cells and making use of the characteristic of w), we obtain

$$\sum_{\sigma\in\mathcal{E}_L} F_{L,\sigma}(u) \leq \int_L f \,\mathrm{d}\boldsymbol{x} \quad \text{and} \quad \sum_{\sigma\in\mathcal{E}_L} F_{L,\sigma}(u) \geq \int_L f \,\mathrm{d}\boldsymbol{x},$$

which completes the proof. □

3.5.3 Resolution of the HMM for the obstacle problem

The monotonicity algorithm is also used to compute the approximate solution of the HMM scheme to the obstacle problem. Its idea here is based on dividing the domain into two disjoint groups: the first part gathers all the cells in which the fluxes balance equation holds, and the second part is made of the remaining cells, in which the solution equals the barrier:

$$\mathbb{I} = \{K \in \mathcal{M} \ : \ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) = \mathrm{m}(K) f_K \text{ and } \psi_K \leq u_K\} \ \text{ and}$$

$$\mathbb{J} = \{K \in \mathcal{M} \ : \ \psi_K = u_K \text{ and } \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) < \mathrm{m}(K) f_K\}.$$

If there are two known sets $\mathbb{I}$ and $\mathbb{J}$ such that $\mathbb{I} \cup \mathbb{J} = \mathcal{M}$, $\mathbb{I} \cap \mathbb{J} = \emptyset$ and

$$u_K \leq \psi_K, \text{ for all } K \in \mathbb{I},$$

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) < \mathrm{m}(K) f_K \text{ for all } K \in \mathbb{J}.$$

Therefore, the discrete version (3.5.2a)–(3.5.2c) can be expressed as

$$\begin{aligned} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) &= \mathrm{m}(K) f_K, \quad \text{for all } K \in \mathbb{I}, \\ u_K &= \psi_K, \quad \text{for all } K \in \mathbb{J}. \end{aligned} \tag{3.5.4}$$

The set $\mathbb{J}$ contains all cells $K \in \mathcal{M}$, where $\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) < \mathrm{m}(K) f_K$. The monotonicity algorithm detailed in Algorithm 2 aims to find a solution to the above problem and identify the two sets of cells $\mathbb{I}$ and $\mathbb{J}$ as a part of the solution.

Thanks to the monotonicity algorithm, the non–linearity caused by the inequalities in the model is eliminated and thus we solve at each step n a linear system of Equations (3.5.5) with unknowns $((u_K^{(n)})_{K\in\mathcal{M}}, (u_\sigma^{(n)})_{\sigma\in\mathcal{E}})$. Let us describe the processes for calculating a solution for this system. For simplicity, the iteration indicator n is dropped.

- **Without elimination of cells unknowns:** Problem (3.5.5) is a linear system, whose matrix is of size $\mathrm{Card}(\mathcal{E}) + \mathrm{Card}(\mathcal{M})$. The equations to assemble this system are:

$$\sum_{\sigma\in\mathcal{E}_K} \sum_{\sigma'\in\mathcal{E}_K} W^K_{\sigma,\sigma'}(u_K - u_{\sigma'}) = f_K, \quad \forall K \in \mathbb{I}, \tag{3.5.6}$$

$$\sum_{\sigma'\in\mathcal{E}_K} W^K_{\sigma,\sigma'}(u_K - u_{\sigma'}) + \sum_{\sigma'\in\mathcal{E}_L} W^L_{\sigma,\sigma'}(u_L - u_{\sigma'}) = 0, \quad \forall \sigma \in \mathcal{E}_K \cap \mathcal{E}_L, \tag{3.5.7}$$

  together with Equation (3.5.5b) and (3.5.5d).

Algorithm 2 Monotonicity algorithm for the obstacle model

1: Set $\mathbb{I}^{(0)} = \mathcal{M}$ and $\mathbb{J} = \emptyset$
2: Set $N = \mathrm{Card}(\mathbb{I}^{(0)})$ ▷ Theoretical bound on the iterations
3: **while** $n \leq N$ **do**
4: $\quad$ $\mathbb{I}^{(n)}$ and $\mathbb{J}^{(n)}$ being known, find the solution $u^{(n)}$ to the system

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u^{(n)}) = \mathrm{m}(K) f_K, \quad \forall K \in \mathbb{I}^{(n)}, \tag{3.5.5a}$$

$$u_K^{(n)} = \psi_K, \quad \forall K \in \mathbb{J}^{(n)}, \tag{3.5.5b}$$

$$F_{K,\sigma}(u^{(n)}) + F_{L,\sigma}(u^{(n)}) = 0, \quad \forall \sigma \in \mathcal{E}_K \cap \mathcal{E}_L, K \neq L \tag{3.5.5c}$$

$$u_\sigma^{(n)} = 0, \quad \forall \sigma \in \mathcal{E}_{\mathrm{ext}}. \tag{3.5.5d}$$

5: $\quad$ Set $\mathbb{I}^{(n+1)} = \{K \in \mathbb{I}^{(n)} \ : \ u_K^{(n)} \leq \psi_K\} \cup \{K \in \mathbb{J}^{(n)} \ : \ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u^{(n)}) \geq \mathrm{m}(K) f_K\}$
6: $\quad$ Set $\mathbb{J}^{(n+1)} = \{K \in \mathbb{J}^{(n)} \ : \ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u^{(n)}) < \mathrm{m}(K) f_K\} \cup \{K \in \mathbb{I}^{(n)} \ : \ u_K^{(n)} > \psi_K\}$
7: $\quad$ **if** $\mathbb{I}^{(n+1)} = \mathbb{I}^{(n)}$ and $\mathbb{J}^{(n+1)} = \mathbb{J}^{(n)}$ **then**
8: $\quad\quad$ Exit "while" loop
9: $\quad$ **end if**
10: **end while**
11: Set $u = u^{(n)}$, $\mathbb{I} = \mathbb{I}^{(n+1)}$ and $\mathbb{J} = \mathbb{J}^{(n+1)}$

- **The hybrid resolution (elimination of cells unknowns):** Reporting (3.5.6) in (3.5.5a) leads to

$$u_K = \frac{1}{b_K}(\mathrm{m}(K) f_K + \sum_{\sigma'\in\mathcal{E}_K} b_{K,\sigma'} u_{\sigma'}), \quad \forall K \in \mathbb{I}. \tag{3.5.8}$$

Equation (3.5.5b) governs the unknowns $(u_K)_{K\in\mathbb{J}}$. Again $b_{K,\sigma'} = \sum_{\sigma\in\mathcal{E}_K} -W^K_{\sigma,\sigma'}$ and $b_K = \sum_{\sigma'\in\mathcal{E}_K} b_{K,\sigma'}$. It remains to build a linear system to determine the edges unknowns $(u_\sigma)_{\sigma\in\mathcal{E}}$. The degrees of freedoms corresponding to the boundary edges can be directly governed by (3.5.5d). For any interior edge $\sigma \in \mathcal{E}_K \cap \mathcal{E}_L$, four cases are considered:

1. If both cells K and L are in $\mathbb{I}$, reporting (3.5.8) in (3.5.6) and using the conservation equation (3.5.5c), we have

$$\sum_{\sigma'\in\mathcal{E}_K}\left(-W^K_{\sigma,\sigma'} - \frac{b_{K,\sigma} b_{K,\sigma'}}{b_K}\right) u_{\sigma'} + \sum_{\sigma'\in\mathcal{E}_L}\left(-W^L_{\sigma,\sigma'} - \frac{b_{L,\sigma} b_{L,\sigma'}}{b_L}\right) u_{\sigma'} = \frac{b_{K,\sigma}}{b_K}\mathrm{m}(K) f_K + \frac{b_{L,\sigma}}{b_L}\mathrm{m}(K) f_L.$$

2. In the case where both K and L are in $\mathbb{J}$, using (3.5.5b), (3.5.5c) and (3.4.12), it follows directly that

$$\sum_{\sigma'\in\mathcal{E}_K} -W^K_{\sigma,\sigma'} u_{\sigma'} + \sum_{\sigma'\in\mathcal{E}_L} -W^L_{\sigma,\sigma'} u_{\sigma'} = \sum_{\sigma'\in\mathcal{E}_K} -W^K_{\sigma,\sigma'} \psi_K + \sum_{\sigma'\in\mathcal{E}_L} -W^L_{\sigma,\sigma'} \psi_L.$$

3. If $K \in \mathbb{I}$ and $L \in \mathbb{J}$, by similar argument, it can be obtained

$$\sum_{\sigma'\in\mathcal{E}_K}\left(-W^K_{\sigma,\sigma'} - \frac{b_{K,\sigma} b_{K,\sigma'}}{b_K}\right) u_{\sigma'} + \sum_{\sigma'\in\mathcal{E}_L} -W^L_{\sigma,\sigma'} u_{\sigma'} = \frac{b_{K,\sigma}}{b_K}\mathrm{m}(K) f_K + \sum_{\sigma'\in\mathcal{E}_L} -W^L_{\sigma,\sigma'} \psi_L.$$

4. If $K \in \mathbb{J}$ and $L \in \mathbb{I}$, we have

$$\sum_{\sigma' \in \mathcal{E}_K} -W^K_{\sigma,\sigma'} u_{\sigma'} + \sum_{\sigma' \in \mathcal{E}_L} \left(-W^L_{\sigma,\sigma'} - \frac{b_{L,\sigma} b_{L,\sigma'}}{b_L} \right) u_{\sigma'} = \sum_{\sigma' \in \mathcal{E}_K} -W^K_{\sigma,\sigma'} \psi_K + \frac{b_{L,\sigma}}{b_L} \mathrm{m}(K) f_L.$$

It has also been proved in [59] that the number of cells Card($\mathcal{M}$) is the theoretical upper bound of the iterations number of the monotonicity Algorithm 2 when the discrete maximum principle is satisfied by the $\mathbb{P}1$ finite element and the two-point finite volume methods. The following lemmas discuss some theoretical findings associated with the convergence of our resolution.

Lemma 3.5.2. *Under Hypothesis 2.2.2, for each iteration n, there exists a unique solution to Problem* (3.5.5) *in Algorithm 2.*

Proof. Under Hypothesis 2.2.2, the equivalent weak formulation of Problem (3.5.5) is to find $u \in X_{\mathcal{D},0}$ such that for all $v \in X_{\mathcal{D},0}$,

$$\begin{aligned} \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma) = \sum_{K \in \mathcal{M}} v_K \int_K f \, \mathrm{d}\boldsymbol{x}, \quad \text{such that } v_K = 0 \text{ if } K \in \mathbb{J}, \\ u_K = \psi_K, \quad \text{for all } K \in \mathbb{J}. \end{aligned} \tag{3.5.9}$$

This scheme is equivalent to solving the linear system $Mu = Z$ with unknowns $u = ((u_K)_{K\in\mathcal{M}}, (u_\sigma)_{\sigma\in\mathcal{E}})$ and the right-hand side is $Z_K = \mathrm{m}(K) f_K$ for all $K \in \mathbb{I}$, $Z_K = \psi_K$, for all $K \in \mathbb{J}$, and $Z_\sigma = 0$, for all $\sigma \in \mathcal{E}$. To show that this system is invertible, we prove $Mu = 0$ implies $u = 0$. It is obvious that $u_K = 0$ for all $K \in \mathbb{J}$. Therefore, it is admissible to take $v = u$ in (3.5.9) to see, thanks to the definition of fluxes (3.3.5)

$$\sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}(u)(u_K - u_\sigma) = \int_\Omega \Lambda_K \nabla_{\mathcal{D}} u \cdot \nabla_{\mathcal{D}} u \, \mathrm{d}\boldsymbol{x} = 0,$$

which leads to $u_K = 0$ for all $K \in \mathcal{M}$ and $u_\sigma = 0$ for all $\sigma \in \mathcal{E}$, owing to the construction of the discrete gradient ($||\nabla_{\mathcal{D}} \cdot ||_{L^2(\Omega)^d}$ is a norm on $X_{\mathcal{D},0}$). □

Lemma 3.5.3. *Let Hypothesis 2.2.2 hold and the two sets $(\mathbb{I}^{(n)})_{n\in\mathbb{N}}$ and $(\mathbb{J}^{(n)})_{n\in\mathbb{N}}$ be the ones determined by applying the monotonicity Algorithm 2. If there exists $N \in \mathbb{N}$ such that $\mathbb{I}^{(N)} = \mathbb{I}^{(N+1)}$ and $\mathbb{J}^{(N)} = \mathbb{J}^{(N+1)}$, then the solution $u^{(N)}$ to the scheme* (3.5.5) *in Algorithm 2 is also the unique solution to Problem* (3.5.2).

Proof. Set $\mathbb{I} = \mathbb{I}^{(N)}$, $\mathbb{J} = \mathbb{J}^{(N)}$ and $u = u^{(N)}$. Due to $\mathbb{J}^{(N)} = \mathbb{J}^{(N+1)}$, one has

$$\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}(u) \leq m(K) f_K, \quad \forall K \in \mathbb{J}.$$

From (3.5.5a), it follows that

$$\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}(u) = m(K) f_K, \quad \forall K \in \mathbb{I},$$

which shows that u satisfies (3.5.2b) since $\mathcal{M} = \mathbb{I} \cup \mathbb{J}$. Likewise, from $\mathbb{I}^{(N)} = \mathbb{I}^{(N+1)}$ and (3.5.5b), we conclude

$$\begin{aligned} u_K &\leq \psi_K, \quad \forall K \in \mathbb{I}, \\ u_K &= \psi_K, \quad \forall K \in \mathbb{J}. \end{aligned}$$

Hence, (3.5.2c) holds. Finally, (3.5.2a) is a consequence of the fact that $\mathcal{M} = \mathbb{I} \cup \mathbb{J}$ and $\mathbb{I} \cap \mathbb{J} = \emptyset$.

Table 3.1: Number of iterations of the monotonicity algorithm in Test 3.6.1

Mesh size h	0.0625	0.0500	0.0250	0.0156
$\#\{\sigma \subset \Gamma_3\}$	16	20	40	46
NITER	4	5	5	6

3.6 Numerical results

We present numerical experiments to highlight the efficiency of the HMM methods for variational inequalities, and to verify our theoretical results given in the previous chapter. All tests below are performed by MATLAB code using the "PDE Toolbox".

The choice of $A_K = \beta_K \mathbf{Id}$, with $\beta_K > 0$, in the schemes (3.4.3) and (3.5.1) satisfies Assumption (3.4.5). The solutions of the HMM schemes in the following tests are computed based on this choice, therefore the matrix $\mathbb{B}_K$ to be considered here is

$$\mathbb{B}_K = \mathbb{D}_K = \text{diag}\left(\frac{|\sigma|}{d_{K,\sigma}}\Lambda_K \mathbf{n}_{K,\sigma} \cdot \mathbf{n}_{K,\sigma}\right). \tag{3.6.1}$$

3.6.1 The Signorini problem

Test 3.6.1. We investigate the Signorini problem from [88]:

$$\begin{aligned} -\Delta \bar{u} &= 2\pi \sin(2\pi x) && \text{in } \Omega = (0,1)^2, \\ \bar{u} &= 0 && \text{on } \Gamma_1, \\ \nabla \bar{u} \cdot \mathbf{n} &= 0 && \text{on } \Gamma_2, \\ \left.\begin{aligned} \bar{u} &\geq 0 \\ \nabla \bar{u} \cdot \mathbf{n} &\geq 0 \\ \bar{u}\nabla \bar{u} \cdot \mathbf{n} &= 0 \end{aligned}\right\} & && \text{on } \Gamma_3, \end{aligned}$$

with $\Gamma_1 = [0,1] \times \{1\}$, $\Gamma_2 = (\{0\} \times [0,1]) \cup (\{1\} \times [0,1])$ and $\Gamma_3 = [0,1] \times \{0\}$. Here, the domain is meshed with triangles produced by "INITMESH".

Figure 3.1 presents the graph of the approximate solution obtained on a mesh of size 0.05. The solution compares very well with linear finite elements solution from [92]; the graph seems to perfectly capture the point where the condition on Γ_3 changes from Dirichlet to Neumann at about 0.7 on the $x-axis$. Table 3.1 shows the number of iterations (NITER) of the monotonicity algorithm, required to obtain the HMM solution for various mesh sizes. [59] proves that this number of iterations is theoretically bounded by the number of edges included in Γ_3. We observe that NITER is much less than this worst-case bound. We also notice the robustness of this monotonicity algorithm: reducing the mesh size does not significantly affect the number of iterations.

Test 3.6.2. Setting $\Omega = (0,1)$, we consider the Signorini model with nonhomogeneous Dirichlet boundary condition presented in [61]:

$$\begin{aligned} -\text{div}(\Lambda \nabla \bar{u}) &= f && \text{in } \Omega, \\ \bar{u} &= g && \text{on } \Gamma_1, \\ \Lambda \nabla \bar{u} \cdot \mathbf{n} &= 0 && \text{on } \Gamma_2, \\ \left.\begin{aligned} \bar{u} &\geq -1 \\ \Lambda \nabla \bar{u} \cdot \mathbf{n} &\geq -2 \\ (\Lambda \nabla \bar{u} \cdot \mathbf{n} + 2)(\bar{u} + 1) &= 0 \end{aligned}\right\} & && \text{on } \Gamma_3, \end{aligned}$$

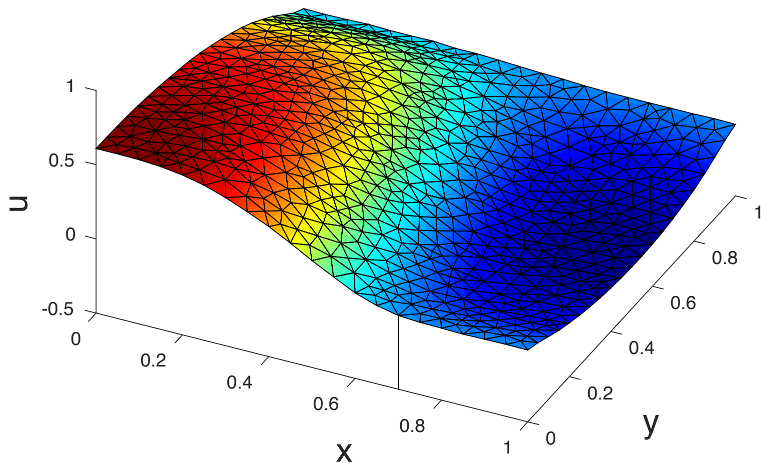

Figure 3.1: The HMM solution for Test 3.6.1 with $h = 0.05$. The line represents where the boundary condition on Γ_3 changes from Neumann to Dirichlet.

with $\Gamma_1 = \{0\} \times [0,1]$, $\Gamma_2 = [0,1] \times \{1\} \cup (\{1\} \times [0,1]$ and $\Gamma_3 = [0,1] \times \{0\}$.

The source term $f \in C^2(\overline{\Omega})$ and the boundary data $g \in C^2(\overline{\Omega})$ are defined such that the solution $\bar{u}$ to the above problem satisfies the following properties:

1. $\bar{u} \in C^2(\Omega)$,
2. $\bar{u}(x,0) \geq -1$, $\nabla\bar{u}(x,0)\cdot\mathbf{n} = -2$ for all $x \in [0, \frac{1}{2}]$,
3. $\bar{u}(x,0) = -1$, $\nabla\bar{u}(x,0)\cdot\mathbf{n} \geq -2$ for all $x \in [\frac{1}{2}, 1]$,
4. $\nabla\bar{u}(\boldsymbol{x})\cdot\mathbf{n} = 0$, for all $\boldsymbol{x} = (x,y) \in \Gamma_2$.

Let us explain the construction of these functions f and g. First, one possible choice of g such that $g(0,0) = g(0,1) = 0$ is

$$g(y) = 2(y^3 - 2y^2 + y).$$

With this function g, we define a function $V \in C^2(\overline{\Omega})$ such that $V(\frac{1}{2}, 0) = -1$ in the following way:

$$V(x,y) = \begin{cases} g(y) + 16x^3 - 12x^2, & \text{for } x \in (0, \frac{1}{2}) \text{ and } y \in (0,1), \\ a_3(y)x^3 + a_2(y)x^2 + a_1(y)x + a_0(y), & \text{for } x \in (\frac{1}{2}, 1) \text{ and } y \in (0,1). \end{cases} \tag{3.6.2}$$

The task now is to determine the coefficients a_3, a_2, a_1 and a_0 depending on y. To ensure $\nabla\bar{u}\cdot\mathbf{n} = 0$ on Γ_2, impose $\partial_x V(\frac{1}{2}, y) = \partial_x V(1,y) = 0$, which gives

$$\begin{aligned} \tfrac{3}{4}a_3 + a_2 + a_1 &= 0 \\ 3a_3 + 2a_2 + a_1 &= 0. \end{aligned}$$

With an algebraic elimination processes, one deduces

$$a_3 = -\frac{2}{3}a_1 \quad \text{and} \quad a_2 = \frac{1}{2}a_1. \tag{3.6.3}$$

Substituting this relation into (3.6.2) implies, for $x \in (\frac{1}{2}, 1)$ and for $y \in (0,1)$,

$$V(x,y) = -\frac{2}{3}a_1 x^3 + \frac{1}{2}a_1 x^2 + a_1 x + a_0.$$

From the continuity of the function V at $\boldsymbol{x} = (\frac{1}{2}, y)$, one has

$$\frac{13}{24}a_1 + a_0 = g(y) - 1,$$

which is simplified as

$$a_1 = \frac{24}{13}(g - 1 - a_0), \quad \text{for } y \in (0,1). \tag{3.6.4}$$

Using this equation and (3.6.3), it follows that, for all $(x,y) \in (\frac{1}{2}, 1) \times (0,1)$,

$$V(x,y) = -\frac{48}{39}(g(y)-1)x^3 + \frac{12}{13}(g(y)-1)x^2 + \frac{24}{13}(g(y)-1)x + \left(\frac{48}{39}x^3 - \frac{12}{13}x^2 - \frac{24}{13}x + 1\right)a_0(y).$$

Setting $F(x,y) = (-\frac{48}{39}x^3 + \frac{12}{13}x^2 + \frac{24}{13}x)(g(y)-1)$ and $G(x) = \frac{48}{39}x^3 - \frac{12}{13}x^2 - \frac{24}{13}x + 1$, we get

$$V(x,y) = F(x,y) + G(x)a_0(y). \tag{3.6.5}$$

To satisfy the condition "$\bar{u}(x,0) \geq -1$ on the left part of Γ_3", take $V(x,0) = -1$, thus

$$\frac{48}{39}x^3 - \frac{12}{13}x^2 - \frac{24}{13}x + \left(\frac{48}{39}x^3 - \frac{12}{13}x^2 - \frac{24}{13}x + 1\right)a_0(0) = -1,$$

which yields $a_0(0) = -1$.

To guarantee that $\nabla V \cdot \mathbf{n} = 0$ on a part of Γ_2, in which $y = 1$, impose $\partial_y V = 0$. Then

$$\left(-\frac{48}{39}x^3 + \frac{12}{13}x^2 + \frac{24}{13}x\right)g'(1) + G(x)a_0'(1) = 0.$$

Since $g'(1) = 0$, this relation leads to $a_0'(1) = 0$.

To meet the condition that $-\nabla V \cdot \mathbf{n} \leq 2$ on $[\frac{1}{2}, 1] \times \{0\}$, the inequality $\partial_y V \leq 2$ must be satisfied on this half of Γ_3. It follows that

$$\left(\frac{48}{39}x^3 - \frac{12}{13}x^2 - \frac{24}{13}x\right)(2 - a_0'(0)) \leq 2.$$

Choosing $a_0'(0) = 2$, this inequality holds. Hence a_0 can be defined by

$$a_0(y) = -(1-y)^2.$$

The remaining unknowns a_1, a_2 and a_3 can respectively be determined by (3.6.4) and (3.6.3). Reporting these coefficients in (3.6.2) gives

$$V(x,y) = \frac{1}{13}(-16x^3 + 12x^2 + 24x)(2y^3 - 3y^2) - (1-y)^2, \quad \text{for } x \in (\frac{1}{2}, 1) \text{ and } y \in (0,1).$$

Finally, take $f = -\operatorname{div}(V)$ to deduce: for $x \leq \frac{1}{2}$,

$$f(x,y) = 2(y^3 - 2y^2 + y) + 16x^3 - 12x^2,$$

Table 3.2: Number of iterations of the monotonicity algorithm in Test 3.6.2

Mesh size h	0.0625	0.0500	0.0250	0.0156
$\#\{\sigma \subset \Gamma_3\}$	16	20	40	46
NITER	4	5	5	6

and, for $x > \frac{1}{2}$,

$$f(x,y) = \frac{1}{13}(12y-6)(-16x^3+12x^2+24x) + \frac{1}{13}(96x-24)(3y^2-2y^3) - 2.$$

Figure 3.2 depicts the computed HMM solution to the Signorini model described in Test 3.6.2 on a mesh which is of size $h = 0.05$, while Table 3.2 reports the iterations number of the monotonicity algorithm. As in the previous test, the solution is computed on a triangle mesh generated by "INITMESH", but the final outputs are presented based on "XYGRID". The graph shows that the transition between the Dirichlet and the Neumann BCs happens around $(\frac{1}{2}, 0)$, which matches the results obtained by the two–points finite volume (TPFV) method [61]. Table 3.2 also confirms that the required iterations number of the monotony algorithm is still far from the theoretical bounded.

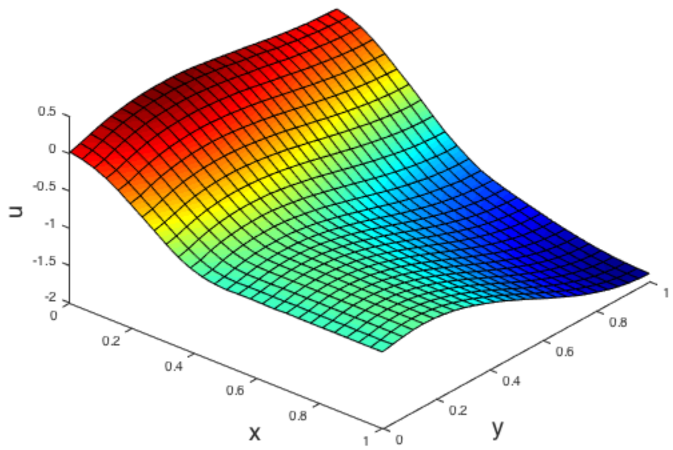

Figure 3.2: The HMM solution for Test 3.6.2 with $h = 0.05$.

To more rigorously assess the rates of convergence, we develop here a new heterogeneous test case for the Signorini boundary conditions, which has an analytical solution with non-trivial one-sided conditions on Γ_3 (the analytical solution switches from homogeneous Dirichlet to homogeneous Neumann at the mid-point of this boundary).

Test 3.6.3. In this test case, we consider (2.1.1a)–(2.1.1d) with the geometry of the domain presented in Figure 3.3, left. The exact solution is

$$\bar{u}(x,y) = \begin{cases} P(y)h(x) & \text{for } y \in (0, \frac{1}{2}) \text{ and } x \in (0,1), \\ xg(x)G(y) & \text{for } y \in (\frac{1}{2}, 1) \text{ and } x \in (0,1). \end{cases}$$

To ensure that (2.1.1b) holds on the lower part of Γ_1, and that the normal derivatives $\Lambda\nabla\bar{u}\cdot\mathbf{n}$ at the interface $y=\frac{1}{2}$ match, we take P such that $P(0)=P(\frac{1}{2})=P'(\frac{1}{2})=0$. Assuming that $P<0$ on $(0,\frac{1}{2})$, the conditions $\partial_n\bar{u}=-\partial_x\bar{u}=0$ and $\bar{u}<0$ on the lower half of Γ_3, as well as the homogeneous condition $\partial_n\bar{u}=0$ on the lower half of Γ_2, will be satisfied if h is such that $h'(0)=h'(1)=0$ and $h(0)=1$. We choose

$$P(y)=-y\left(y-\frac{1}{2}\right)^2 \quad \text{and } h(x)=\cos(\pi x).$$

Given our choice of P, ensuring that $\mathrm{div}(\Lambda\nabla\bar{u})\in L^2(\Omega)$ (i.e. that $\bar{u}$ and the normal derivatives match at $y=\frac{1}{2}$) can be done by taking G such that $G(\frac{1}{2})=G'(\frac{1}{2})=0$. The boundary condition (2.1.1b) on the upper part of Γ_1 is satisfied if $G(1)=0$. The boundary conditions $\partial_n\bar{u}=0$ on the upper part of Γ_2, and $\bar{u}=0$ and $\partial_n\bar{u}<0$ on the upper part of Γ_3, are enforced by taking g such that $g(1)+g'(1)=0$ and $g(0)>0$ (with $G>0$ on $(\frac{1}{2},1)$). We select here

$$g(x)=\cos^2\left(\frac{\pi}{2}x\right)=\frac{1+\cos(\pi x)}{2} \quad \text{and} \quad G(y)=(1-y)\left(y-\frac{1}{2}\right)^2.$$

This $\bar{u}$ is then the analytical solution to the following problem

$$\begin{aligned}
-\mathrm{div}(\Lambda\nabla\bar{u})=f &\quad \text{in } \Omega,\\
\bar{u}=0 &\quad \text{on } \Gamma_1,\\
\Lambda\nabla\bar{u}\cdot\mathbf{n}=0 &\quad \text{on } \Gamma_2,\\
\left.\begin{aligned}\bar{u}\le 0\\ \Lambda\nabla\bar{u}\cdot\mathbf{n}\le 0\\ \bar{u}\Lambda\nabla\bar{u}\cdot\mathbf{n}=0\end{aligned}\right\} &\quad \text{on } \Gamma_3,
\end{aligned}$$

with, for $y\le\frac{1}{2}$,

$$f(x,y)=100\left[\pi^2 y\cos(\pi x)\left(y-\frac{1}{2}\right)^2-2y\cos(\pi x)-2\cos(\pi x)(2y-1)\right],$$

and, for $y>\frac{1}{2}$,

$$\begin{aligned}
f(x,y)=&\frac{\pi^2 x}{2}\cos(\pi x)(y-1)\left(y-\frac{1}{2}\right)^2-2x\cos^2\left(\frac{\pi x}{2}\right)(3y-2)\\
&+\pi\sin(\pi x)(y-1)\left(y-\frac{1}{2}\right)^2.
\end{aligned}$$

We test the scheme on a sequence of meshes that are (mostly) made of hexagonal cells. The third mesh in this sequence (with mesh size $h=0.07$) is represented in Figure 3.3, right.

The relative errors on $\bar{u}$ and $\nabla\bar{u}$, and the corresponding orders of convergence (computed from one mesh to the next one), are presented in Table 3.3. For the HMM method, Theorem 2.2.10 predicts a convergence of order 1 on the gradient if the solution belongs to H^2 and $\Lambda\nabla\bar{u}\in H^1$; the observed numerical rates are slightly below these values, probably due to the fact that Λ is not Lipschitz-continuous here (and thus the regularity $\Lambda\nabla\bar{u}\in H^1$ is not satisfied). As in the previous test, the number of iterations of the monotonicity algorithm remains well below the theoretical bound.

The solution to the HMM scheme for the third mesh in the family used in Table 3.3 is plotted in Figure 3.4. On Γ_3, the saturated constraint for the exact solution changes from Neumann $\nabla\bar{u}\cdot\mathbf{n}=0$ to Dirichlet $\bar{u}=0$ at $y=0.5$. It is clear in Figure 3.4 that the HMM scheme captures well this change of constraint. The slight bump visible at $y=1/2$ is most probably due to the mesh not being aligned with the heterogeneity of Λ (hence, in

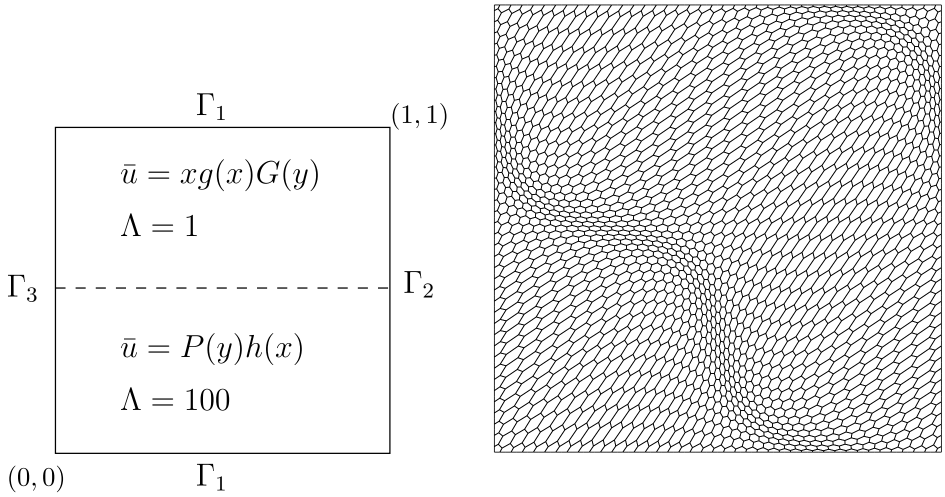

Figure 3.3: Test 3.6.3: geometry and diffusion (left), typical grid (right).

Table 3.3: Error estimate and number of iterations for Test 3.6.3

mesh size h	0.24	0.13	0.07	0.03
$\frac{\Vert\bar{u}-\Pi_{\mathcal{D}}u\Vert_{L^2(\Omega)}}{\Vert\bar{u}\Vert_{L^2(\Omega)}}$	0.6858	0.2531	0.1355	0.0758
Order of convergence	–	1.60	0.92	0.84
$\frac{\Vert\nabla\bar{u}-\nabla_{\mathcal{D}}u\Vert_{L^2(\Omega)}}{\Vert\nabla\bar{u}\Vert_{L^2(\Omega)}}$	0.4360	0.2038	0.1041	0.0542
Order of convergence	–	1.22	0.99	0.95
$\#\{\sigma \subset \Gamma_3\}$	20	40	80	160
NITER	5	5	6	7

some cells the diffusion tensor takes two – very distinct – values, and its approximation by one constant value smears the solution).

When meshes are aligned with data heterogeneities, such bumps do not appear. This is illustrated in Figure 3.5, in which we represent the solution obtained when using a "Kershaw" mesh as in the FVCA5 benchmark [60]. This mesh has size $h = 0.16$ and 34 edges on Γ_3. The monotonicity algorithm converges in 7 iterations. The relative L^2 error on $\bar{u}$ and $\nabla\bar{u}$ are respectively 0.017 and 0.019. As expected on these kinds of extremely distorted meshes, the solution has internal oscillations, but otherwise is qualitatively good. In particular, despite distorted cells near the boundary Γ_3, the transition between the Dirichlet and the Neumann boundary conditions is well captured.

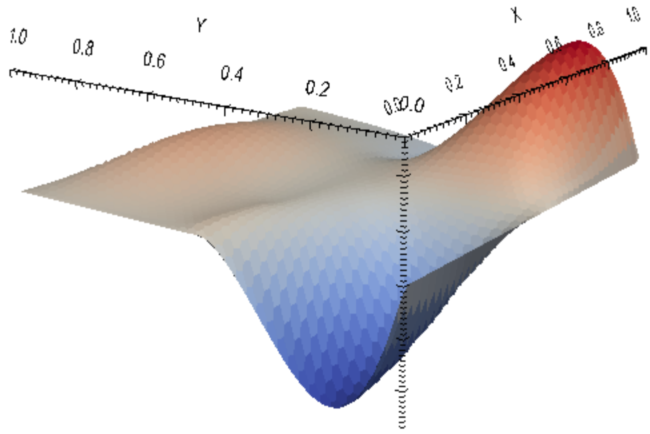

Figure 3.4: The HMM solution for Test 3.6.3 on an hexagonal mesh with $h = 0.07$.

3.6.2 The obstacle problem

Test 3.6.4. In this test case taken from [59], we apply the HMM method to the homogeneous obstacle problem:

$$\begin{aligned} (\Delta\bar{u} + C)(\psi - \bar{u}) = 0, &\quad \text{in } \Omega = (0,1)^2, \\ -\Delta\bar{u} \geq C, &\quad \text{in } \Omega, \\ \bar{u} \geq \psi, &\quad \text{in } \Omega, \\ \bar{u} = 0, &\quad \text{on } \partial\Omega. \end{aligned}$$

The constant C is negative, and the obstacle function $\psi(x, y) = -\min(x, 1-x, y, 1-y)$ satisfies $\psi = 0$ on $\partial\Omega$.

Figure 3.6 shows the graph of the HMM solution to the above obstacle problem on a triangular mesh of size $h = 0.05$. Table 3.4 illustrates the performance of the method and the algorithm. Here again, the number of iterations required to obtain the solution is far less than the number of cells, which is a theoretical bound in the case of obstacle problem [59]. Our results compare well with the results obtained by semi-iterative Newton-type methods in [20], which indicate that decreasing $|C|$ contributes to the difficulty of the problem (leading to an increased number of iterations). We note that, for a mesh of nearly 14,000 cells, we only require 29 iterations if $|C| = 5$ and 14 iterations if $|C| = 20$. On a mesh of 10,000 cells, the semi-iterative Newton-type method of [20] requires 32 iterations if $|C| = 5$ and 9 iterations if $|C| = 20$. Figure 3.9 presents the contact regions based on a cartesian mesh. The black dots represent the set of cell centres where the approximate solution u is equal to the obstacle ψ. The figure also illustrates the fact that decreasing $|C|$ results in maximising the number of meshes where the diffusion equation holds; this might be a reason for the increase in the number of iterations. In this situation, starting from the initial $\mathbb{I}^{(0)} = \mathcal{M}$, that is the diffusion equation is everywhere satisfied, is an appropriate choice to achieve the solution in a small number of iterations. For instance, for $|C| = 5$ and using triangular mesh (with $h = 0.016$), 42 iterations are needed to obtain the solution if we begin with $\mathbb{J}^{(0)} = \mathcal{M}$, the case where the initial solution is everywhere equal to the barrier. With the initial guess $\mathbb{I}^{(0)} = \mathcal{M}$, 29 iterations are enough.

To achieve a better interpolation at the boundary, the "MATLAB SURF function", which takes into account the data at the edges midpoints, is used to plot the HMM solution to the problem in Test 4. This is demonstrated by the graph in Figure 3.7 in which the solution exactly takes zero values at the boundary, which is unlike the one in Figure 3.6, where the solution seems to be not as well approximated at some points of the boundary (due

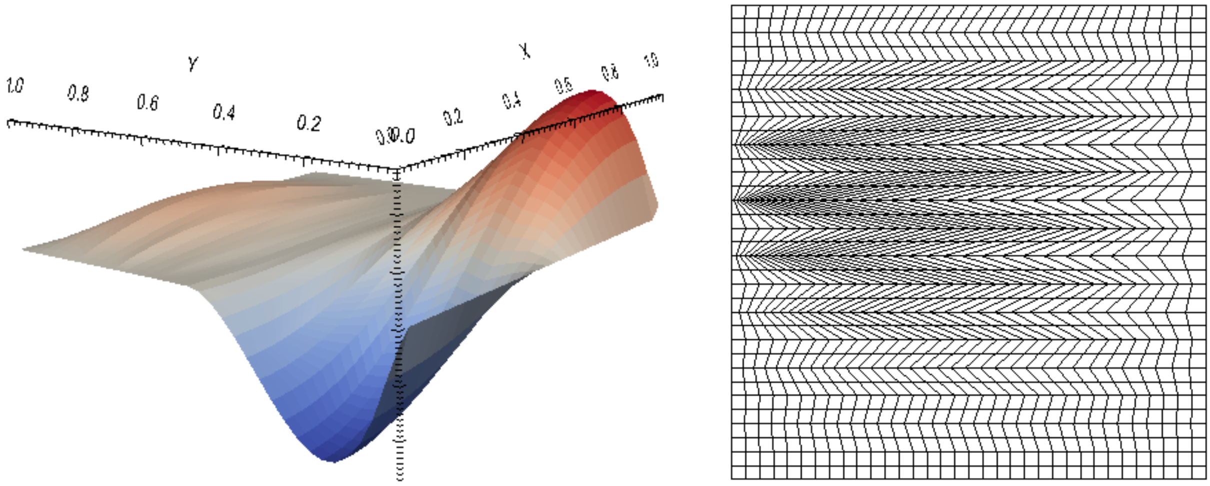

Figure 3.5: The HMM solution (left) for Test 3.6.3 on a Kershaw mesh (right).

to the use of the "MATLAB PDEPRTNI function" that provides linearly interpolated values at node points).

We conduct the same test on a Kershaw mesh (Figure 3.5, right) with size $h = 0.5$ to examine the performance of applying the HMM method to the obstacle model on a general grid. Figure 3.8 shows that the results obtained by means of this type of mesh are in agreement with those achieved by using a triangle mesh. It requires the same iterations number to determine the solution and decreasing $|C|$ still results in raising the iterations number.

Table 3.4: Test 3.6.4: number of iterations for various C (with a triangle mesh)

mesh size h	0.016	0.025	0.050	0.062
Card($\mathcal{M}$)	14006	5422	1342	872
NITER ($C = -5$)	29	20	12	11
NITER ($C = -10$)	23	18	10	9
NITER ($C = -15$)	19	13	9	8
NITER ($C = -20$)	14	12	7	8

Test 3.6.5. To evaluate the validity of the convergence rate of HMM for the obstacle problem, we implement the method to the model with an available solution [6],

$$\begin{aligned} -\bar{u}(\Delta\bar{u} + f) &= 0, \quad \text{in } \Omega = (-1,1)^2, \\ -\Delta\bar{u} &\geq f, \quad \text{in } \Omega, \\ \bar{u} &\geq 0, \quad \text{in } \Omega, \\ \bar{u} &= 0, \quad \text{on } \partial\Omega, \end{aligned}$$

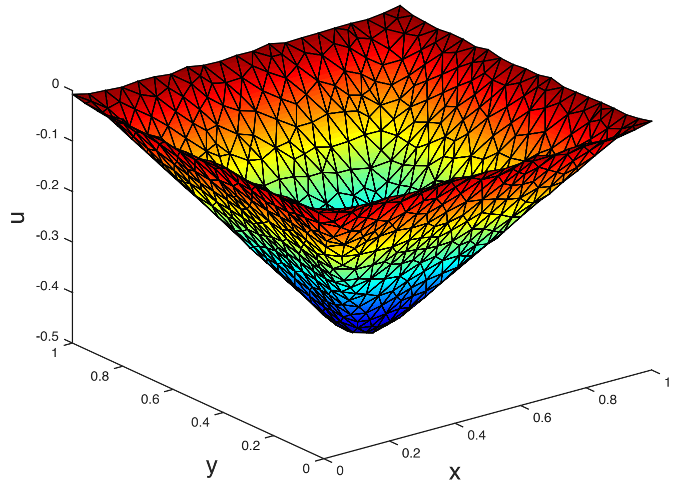

Figure 3.6: The HMM solution for Test 3.6.4 with $h = 0.05$, $C = -20$.

where the function f is defined by:

$$f(x,y) = \begin{cases} -8(2x^2 + 2y^2 - r^2), & \text{if } \sqrt{x^2+y^2} > r \\ -8r^2(1 - x^2 - y^2 + r^2), & \text{if } \sqrt{x^2+y^2} \le r. \end{cases}$$

The exact solution to this problem is given by:

$$\bar{u}(x,y) = (\max\{x^2 + y^2 - r^2\})^2.$$

The test is performed on three different sequences of meshes, triangular, hexahedral (Figure 3.3, right) and Kershaw type (Figure. 3.5, right). Figure 3.10 shows the approximate HMM solution with the case $r = 0.7$ on the hexahedral mesh (with size $h = 0.05$).

For the three types of meshes, the relative error on the solution and on its gradient and the corresponding convergence rate is given in Tables 3.5, 3.6 and 3.7. Theorem 2.2.11 shows that the approximate HMM solution to the obstacle problem enjoys an order 1 rate of convergence; we see that the hexahedral and the “Kershaw” meshes present slightly better rates on this test case than the triangular mesh. Indeed, the HMM method is based on the number of unknowns edges, that controls the number of degrees of freedom. For instance, a triangular mesh (with $h = 0.125$) includes 1337 edges while a hexahedral (with $h = 0.13$) and a Kershaw mesh (with $h = 0.17$) respectively produce 5200 and 9384 edges. We also see a super-convergence in L^2 norm (a case where the solution behaves better than the order 1, initially expected here since $\Pi_{\mathcal{D}}$ is piecewise constant reconstruction). The tables also state that the monotony algorithm converges in a small number of iterations even on meshes of small size.

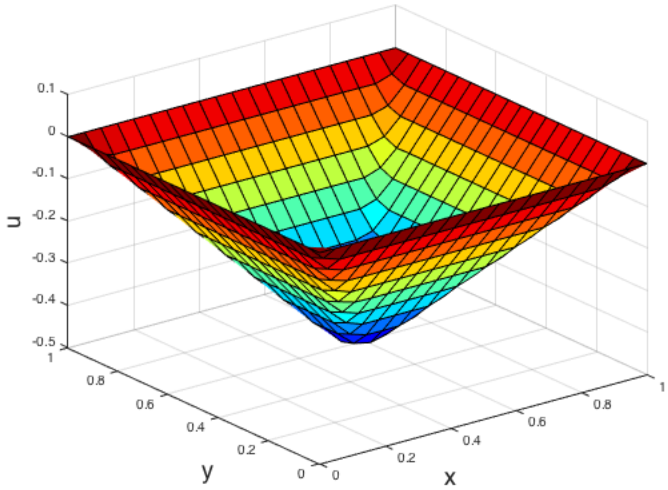

Figure 3.7: The HMM solution for Test 3.6.4 with $h = 0.05$, $C = -20$.

Table 3.5: Error estimate and number of iterations for Test 3.6.5 with a triangular mesh

mesh size h	0.125	0.063	0.031	0.016
$\frac{\|\bar{u}-\Pi_{\mathcal{D}}u\|_{L^2(\Omega)}}{\|\bar{u}\|_{L^2(\Omega)}}$	0.0822	0.0199	0.0047	0.0012
Order of convergence	–	2.05	2.08	1.97
$\frac{\|\nabla\bar{u}-\nabla_{\mathcal{D}}u\|_{L^2(\Omega)}}{\|\nabla\bar{u}\|_{L^2(\Omega)}}$	0.0530	0.0259	0.0126	0.0063
Order of convergence	–	1.03	1.04	1.00
Card($\mathcal{M}$)	870	3474	14120	56980
NITER	10	15	28	49

Table 3.6: Error estimate and number of iterations for Test 3.6.5 with a hexahedral mesh

mesh size h	0.48	0.26	0.13	0.07
$\frac{\|\bar{u}-\Pi_{\mathcal{D}}u\|_{L^2(\Omega)}}{\|\bar{u}\|_{L^2(\Omega)}}$	0.0314	0.0089	0.0027	0.0007
Order of convergence	–	2.06	1.72	2.18
$\frac{\|\nabla\bar{u}-\nabla_{\mathcal{D}}u\|_{L^2(\Omega)}}{\|\nabla\bar{u}\|_{L^2(\Omega)}}$	0.0593	0.0230	0.0081	0.0029
Order of convergence	–	1.54	1.51	1.35
Card($\mathcal{M}$)	121	441	1681	6561
NITER	5	7	13	23

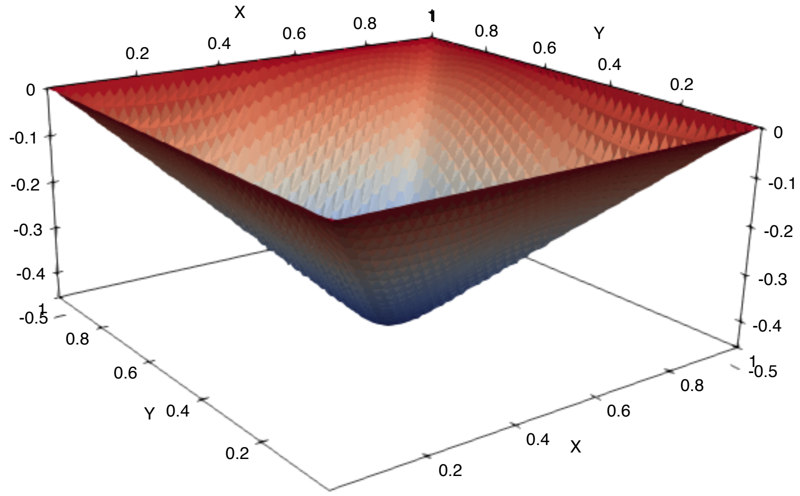

Figure 3.8: The HMM solution for Test 3.6.4 on an hexagonal mesh with $h = 0.06$ and $C = -20$.

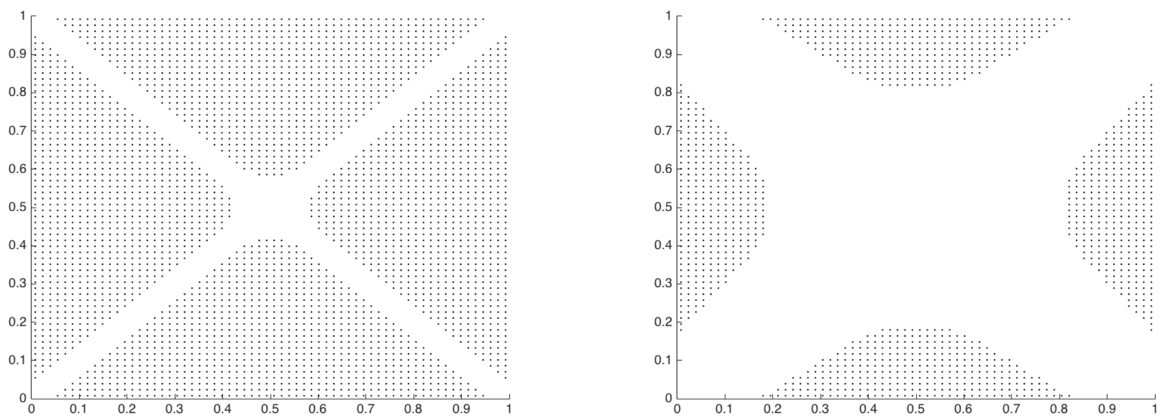

Figure 3.9: Coincidence set corresponding to Test 3.6.4, $|C| = 20$ (left) and $|C| = 5$ (right).

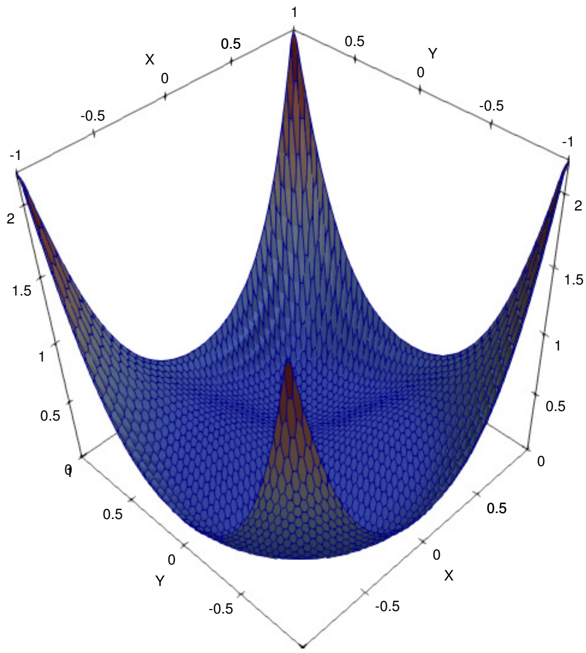

Figure 3.10: The HMM solution for Test 3.6.5 on an hexagonal mesh with $h = 0.06$ and r=0.7

Table 3.7: Error estimate and number of iterations for Test 3.6.5 with a Kershaw mesh

mesh size h	0.66	0.33	0.22	0.17
$\frac{\Vert\bar{u}-\Pi_{\mathcal{D}}u\Vert_{L^2(\Omega)}}{\Vert\bar{u}\Vert_{L^2(\Omega)}}$	0.0255	0.0071	0.0033	0.0019
Order of convergence	–	1.84	1.89	2.14
$\frac{\Vert\nabla\bar{u}-\nabla_{\mathcal{D}}u\Vert_{L^2(\Omega)}}{\Vert\nabla\bar{u}\Vert_{L^2(\Omega)}}$	0.0329	0.0110	0.0061	0.0041
Order of convergence	–	1.58	1.45	1.54
Card($\mathcal{M}$)	289	1156	2601	4624
NITER	8	15	23	32

Appendix

3.A Computation of the local HMM matrix

We present here a way to compute the local matrix $(W^K_{\sigma,\sigma'})_{\sigma,\sigma'\in\mathcal{E}_K^2}$ used in the implementation procedures, precisely in Equation (3.4.12).

As previously mentioned, all the numerical tests are performed by taking the isomorphism $A_K = \beta_K\mathbf{Id}$. Letting $F_K(u) = (F_{K,\sigma}(u))_{\sigma\in\mathcal{E}_K}$, $U_K = (u_K - u_\sigma)_{\sigma\in\mathcal{E}_K}$ and $\mathbb{W}_K = (W^K_{\sigma,\sigma'})_{\sigma,\sigma'\in\mathcal{E}_K}$, the relation (3.4.12) can be rewritten $F_K(u) = \mathbb{W}_K U_K$. Equation (3.3.5) can be written, thanks to Equation (3.3.3) and the particular choice of A_K

$$V_K^T\mathbb{W}_K U_K = |K|\Lambda_K\nabla_K u\cdot\nabla_K v + \beta^2(R_K(v))^T\mathbb{B}_K(R_K(u)), \tag{3.A.1}$$

where $\mathbb{B}_K$ is a diagonal positive definite matrix defined by (3.6.1) and $V_K = (v_K - v_\sigma)_{\sigma\in\mathcal{E}_K}$.

One can write $\nabla_K u = \mathbb{L}_K U_K$, with

$$\mathbb{L}_K = -\frac{|\sigma|}{|K|}\mathbf{n}_{K,\sigma}.$$

Then, $R_K(u) = (\mathbb{I}_K - \mathbb{X}_K\mathbb{L}_K)U_K$, where

- $\mathbb{I}_K$ is the $\mathrm{Card}(\mathcal{E}_K)$ identity matrix,
- $\mathbb{X}_K$ is the $\mathrm{Card}(\mathcal{E}_K)\times d$ matrix with rows $((\overline{\boldsymbol{x}}_\sigma - \boldsymbol{x}_K)^T)_{\sigma\in\mathcal{E}}$.

Equation (3.A.1) can be also reduced to

$$V_K^T\mathbb{W}_K U_K = V_K^T(|K|\mathbb{L}_K^T\Lambda_K\mathbb{L}_K)U_K + V_K^T(\beta_K^2(\mathbb{I}_K - \mathbb{X}_K\mathbb{L}_K)^T\mathbb{B}_K(\mathbb{I}_K - \mathbb{X}_K\mathbb{L}_K))U_K,$$

which gives

$$\mathbb{W}_K = |K|\mathbb{L}_K^T\Lambda_K\mathbb{L}_K + \beta_K^2(\mathbb{I}_K - \mathbb{X}_K\mathbb{L}_K)^T\mathbb{B}_K(\mathbb{I}_K - \mathbb{X}_K\mathbb{L}_K).$$

Chapter 4

Nonlinear elliptic variational inequalities

Abstract. Using the gradient discretisation method, we provide a complete and unified numerical analysis for nonlinear variational inequalities based on Leray-Lions operators and subject to nonhomogeneous Dirichlet and Signorini boundary conditions. This analysis is proved to be easily extended to the obstacle and Bulkley models, which can be formulated as nonlinear VIs. It also allows us to obtain the convergence results for many *conforming* and *nonconforming* numerical schemes included in the GDM, and not previously studied for these models. Our theoretical results are applied to the HMM method. Numerical results are provided for HMM on the seepage model, and demonstrate that, even on distorted meshes, this method provides accurate results.

4.1 Introduction

The purpose of this chapter is to provide a complete and unified convergence analysis of numerical schemes for nonlinear variational inequalities. Nonlinear variational inequalities are related to a wide range of applications. In particular, unconfined seepage models can be used to study the construction of earth dams, embankments and hydraulic design. With nonlinear variational inequalities, one can also study the Bulkley fluid model, which is applicable to different phenomena and processes, such as blood flow [85], food processing [54] and Bingham fluid flows [76].

We consider here variational inequalities related to elliptic equations of the type

$$-\operatorname{div}\mathbf{a}(\boldsymbol{x},u,\nabla u) = f \quad \text{in } \Omega, \tag{4.1.1a}$$

$$u = g \quad \text{on } \partial\Omega, \tag{4.1.1b}$$

where Ω is an open bounded connected subset of $\mathbb{R}^d$, $d \geq 1$, with boundary $\partial\Omega$. Precise assumptions on data will be stated in the next sections.

The theory on PDEs of the kind (4.1.1) has been covered in several works, see [89, 33, 51, 70] and references therein. A number of numerical analyses on these models has also been carried out, starting from the approximation of the p-Laplace equation, with proved rates of convergences, by $\mathbb{P}1$ finite elements in [11]. Subsequent works consider more general Leray-Lions models, possibly transient, and establish either error estimates (under regularity assumptions on the solution to the PDE), or prove the convergence towards a solution with minimal regularity. We refer the reader to [40, 33, 30, 29, 58, 78, 4, 5] for a few examples. Several algorithms can be used to compute the solution to the corresponding nonlinear numerical schemes, from basic fixed-point iterations (which corresponds to the Kaçanov method [70]) to Newton methods, to multigrid techniques [12], to augmented Lagrangian algorithms [65].

The mathematical theory of Variational inequalities (VI) based on equations of the kind (4.1.1) is well understood, see e.g. [74, 72, 86, 23]. We note that [86] considers an obstacle problem with measure source terms rather than $W^{1,p}(\Omega)'$ source terms (the theory for the corresponding PDEs is developed in [14]). [63] studies nonlinear quasi-variational inequalities and proposes a semismooth Newton iteration to obtain a solution.

Apparently, the numerical analysis for VI based on (4.1.1) is much more limited in scope. Under strong monotonicity assumptions on the operator, [82] develops a convergence analysis of *conforming* numerical schemes for nonlinear VI. [56] develops the analysis of *conforming* finite elements method for VI involving a nonlinear proper function. The $\mathbb{P}1$ finite element method is applied to the obstacle problem, restricted to a p-Laplacian operator, homogenous Dirichlet boundary condition and zero barrier inside the domain, and an a priori error estimate is obtained under $W^{2,p}$ regularity on the solution [66, 79]. The Bulkley model is also approximated by $\mathbb{P}_1$ finite elements [26, 27] and the Lagrange methods [26]. In [96], a seepage model is approximated by a finite elements method, but no convergence analysis is carried out. The authors utilise a fixed point method (Kaçanov) to treat the nonlinearity and to compute the solution to the scheme.

This chapter is organised as follows. Section 4.2 details the nonlinear Signorini problem, its approximation by the gradient discretisation method, and the corresponding convergence results. Section 4.3 shows that the GDM can successfully be adapted to the obstacle problem and the Bulkley fluid model. A short section, Section 4.4, describes the case where the barriers of the Signorini and obstacle problems are approximated as part of the discretisation process. In Section 4.5 we apply our results to the HMM scheme, and we provide numerical tests to highlight the efficiency of this method for solving the seepage model on a range of different meshes. Section 4.A contains basic materials corresponding to the Leray-Lions operators.

4.2 Nonlinear Signorini problem

4.2.1 Continous problem

We first consider the following nonlinear Signorini problem:

$$-\operatorname{div} \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) = f \quad \text{in } \Omega, \tag{4.2.1a}$$

$$\bar{u} = g \quad \text{on } \Gamma_1, \tag{4.2.1b}$$

$$\mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_2, \tag{4.2.1c}$$

$$\left.\begin{aligned} \bar{u} &\le a \\ \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \mathbf{n} &\le 0 \\ (a - \bar{u})\mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \mathbf{n} &= 0 \end{aligned}\right\} \quad \text{on } \Gamma_3. \tag{4.2.1d}$$

Here $\mathbf{n}$ denotes to the unit outer normal to $\partial\Omega$ that consists of three parts $(\Gamma_1, \Gamma_2, \Gamma_3)$. The assumptions on the Leray-Lions operator $\mathbf{a}$ are standard:

$$\mathbf{a} : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d \text{ is a Carathéodory function,} \tag{4.2.2}$$

(i.e., for a.e. $\boldsymbol{x} \in \Omega$, $(u, \xi) \mapsto \mathbf{a}(\boldsymbol{x}, u, \xi)$ is continuous and, for all $(u, \xi) \in \mathbb{R} \times \mathbb{R}^d$, $\boldsymbol{x} \to \mathbf{a}(\boldsymbol{x}, u, \xi)$ is measurable) and, for some $p \in (1, \infty)$ and $p' = \frac{p}{p-1}$,

$$\begin{aligned} &\exists \overline{a} \in L^{p'}(\Omega), \exists \mu > 0 : \\ &|\mathbf{a}(\boldsymbol{x}, u, \xi)| \le \overline{a}(\boldsymbol{x}) + \mu |\xi|^{p-1}, \text{ for a.e. } \boldsymbol{x} \in \Omega,\ \forall u \in \mathbb{R},\ \forall \xi \in \mathbb{R}^d, \end{aligned} \tag{4.2.3}$$

$$\exists \underline{a} > 0 : \mathbf{a}(\boldsymbol{x}, u, \xi) \cdot \xi \ge \underline{a} |\xi|^p, \text{ for a.e. } \boldsymbol{x} \in \Omega,\ \forall u \in \mathbb{R},\ \forall \xi \in \mathbb{R}^d, \tag{4.2.4}$$

$$(\mathbf{a}(\boldsymbol{x}, u, \xi) - \mathbf{a}(\boldsymbol{x}, u, \chi)) \cdot (\xi - \chi) \ge 0 \text{ for a.e. } \boldsymbol{x} \in \Omega,\ \forall u \in \mathbb{R},\ \forall \xi, \chi \in \mathbb{R}^d. \tag{4.2.5}$$

Assumptions (4.2.3), (4.2.4) and (4.2.5) are called the growth, coercivity and the monotonicity conditions, respectively. Setting $\mathbf{a}(\boldsymbol{x}, u, \nabla u) = |\nabla u|^{p-2} \nabla u$ in (4.2.1a) gives in particular the p-Laplacian operator.

Remark 4.2.1. With $p = 2$ and $\mathbf{a}(\boldsymbol{x}, u, \nabla u) = \Lambda(\boldsymbol{x}, u) \nabla u$, Problem (4.2.1) covers the seepage models, where $\Lambda(\boldsymbol{x}, u)$ is defined based on a permeability tensor $\mathbf{K}$ and a penalised Heaviside function depending on a fixed function. Using the penalised Heaviside function, the Darcy law, initially only valid in the wet domain, can be extended to the dry domain. We refer the reader to [96] and references therein for more details.

Assumptions 4.2.1. *The assumptions on the data in Problem* (4.2.1) *are the following:*

1. *the operator* $\mathbf{a}$ *satisfies* (4.2.2)–(4.2.5) *and the domain* Ω *has a Lipschitz boundary,*
2. *the parts of the boundary,* Γ_1, Γ_2 *and* Γ_3*, are assumed to be measurable and pairwise disjoint subsets of* $\partial\Omega$ *such that* $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3 = \partial\Omega$ *and the* $(d-1)$*-dimensional measure of* Γ_1 *is non zero,*
3. *the source term* f *belongs to* $L^{p'}(\Omega)$*, the barrier* a *belongs to* $L^p(\partial\Omega)$ *and the boundary data* g *belongs to* $W^{1-\frac{1}{p},p}(\partial\Omega)$*,*
4. *the closed convex set* $\mathcal{K} := \{v \in W^{1,p}(\Omega) \ :\ \gamma(v) = g \text{ on } \Gamma_1,\ \gamma(v) \le a \text{ on } \Gamma_3 \text{ a.e.}\}$ *is nonempty.*

Based on Assumptions 4.2.1, Problem (4.2.1) can be written in the following weak sense:

$$\left\{ \begin{aligned} &\text{Find } \bar{u} \in \mathcal{K} \text{ such that for all } v \in \mathcal{K}, \\ &\quad \int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}(\boldsymbol{x}), \nabla \bar{u}(\boldsymbol{x})) \cdot \nabla(\bar{u} - v)(\boldsymbol{x}) \, \mathrm{d}x \le \int_\Omega f(\boldsymbol{x})(\bar{u} - v)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \end{aligned} \right. \tag{4.2.6}$$

The existence of a solution to this problem is ensured by the classical results in [74] as follows.

Lemma 4.2.2. *Under Assumptions 4.2.1, Problem* (4.2.6) *has at least one weak solution.*

Proof. Letting $\widetilde{g} \in \mathcal{K}$ be a lifting of g such that $\gamma(\widetilde{g}) = g$, we define the nonlinear operator $\mathcal{A} : V = W^{1,p}_{\Gamma_1}(\Omega) \to V' = W^{-1,p'}_{\Gamma_1}(\Omega)$ by

$$\langle \mathcal{A}(\widetilde{u}), \widetilde{w}\rangle = \int_\Omega \mathbf{a}(\boldsymbol{x}, \widetilde{u}+\widetilde{g}, \nabla(\widetilde{u}+\widetilde{g}))\cdot\nabla\widetilde{w}\,\mathrm{d}\boldsymbol{x}, \quad \forall \widetilde{u}, \widetilde{w} \in W^{1,p}_{\Gamma_1}(\Omega),$$

where $\langle\cdot,\cdot\rangle$ denotes the duality product between the space $W^{1,p}_{\Gamma_1}(\Omega) := \{v \in W^{1,p}(\Omega) : \ \gamma(v) = 0 \text{ on } \Gamma_1\}$ and its dual space $W^{-1,p'}_{\Gamma_1}(\Omega)$. Letting $\widetilde{u} = \bar{u} - \widetilde{g}$ and $\widetilde{w} = w - \widetilde{g}$, Problem (4.2.1) can be recast as

$$\begin{cases} \text{Find } \widetilde{u} \in \mathcal{K} - \widetilde{g}, \text{ such that for all } \widetilde{w} \in \mathcal{K} - \widetilde{g}, \\ \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \widetilde{u}+\widetilde{g}, \nabla(\widetilde{u}+\widetilde{g}))\cdot\nabla(\widetilde{u}-\widetilde{v})\,\mathrm{d}\boldsymbol{x} \le \int_\Omega f(\widetilde{u}-\widetilde{v})\,\mathrm{d}\boldsymbol{x}. \end{cases}$$

With $\mathbf{a}$ satisfying the standard assumptions of a Leray-Lions operator (4.2.2)–(4.2.5), it is proved in [74, Section 2.6] that $\mathcal{A}$ is of the calculus of variations on $W^{1,p}_{\Gamma_1}(\Omega)$ and it is therefore pseudo-monotone. Let us verify that, for a fixed $\widetilde{\varphi} \in W^{1,p}_{\Gamma_1}(\Omega)$ such that $\gamma(\widetilde{\varphi}) \le a$ on Γ_3,

$$\lim_{\|\widetilde{u}\|_V\to\infty} \frac{\langle \mathcal{A}(\widetilde{u}), \widetilde{u}-\widetilde{\varphi}\rangle}{\|\widetilde{u}\|_V} = +\infty. \tag{4.2.7}$$

To do this, we start with

$$\begin{aligned} \langle \mathcal{A}(\widetilde{u}), \widetilde{u}-\widetilde{\varphi}\rangle &= \int_\Omega \mathbf{a}(\boldsymbol{x}, \widetilde{u}+\widetilde{g}, \nabla(\widetilde{u}+\widetilde{g}))\cdot\nabla(\widetilde{u}+\widetilde{g})\,\mathrm{d}\boldsymbol{x} - \int_\Omega \mathbf{a}(\boldsymbol{x}, \widetilde{u}+\widetilde{g}, \nabla(\widetilde{u}+\widetilde{g}))\cdot\nabla(\widetilde{\varphi}+\widetilde{g})\,\mathrm{d}\boldsymbol{x} \\ &\ge \int_\Omega \underline{a}|\nabla(\widetilde{u}+\widetilde{g})|^p\,\mathrm{d}\boldsymbol{x} - \int_\Omega (\overline{a} + \mu|\nabla(\widetilde{u}+\widetilde{g})|^{p-1})\cdot\nabla(\widetilde{\varphi}+\widetilde{g})\,\mathrm{d}\boldsymbol{x}. \end{aligned}$$

Using Holder's inequality, we get

$$\begin{aligned} \langle \mathcal{A}(\widetilde{u}), \widetilde{u}-\widetilde{\varphi}\rangle &\ge \int_\Omega \underline{a}|\nabla(\widetilde{u}+\widetilde{g})|^p\,\mathrm{d}\boldsymbol{x} - C_1 + \left(\int_\Omega \mu^{p'}|\nabla(\widetilde{u}+\widetilde{g})|^p\,\mathrm{d}\boldsymbol{x}\right)^{(p-1)/p}\left(\int_\Omega |\nabla(\widetilde{\varphi}+\widetilde{g})|^p\,\mathrm{d}\boldsymbol{x}\right)^{1/p} \\ &\ge \underline{a}\|\nabla(\widetilde{u}+\widetilde{g})\|^p_{L^p(\Omega)^d} - C_1 - \mu\|\nabla(\widetilde{\varphi}+\widetilde{g})\|_{L^p(\Omega)^d}\|\nabla(\widetilde{u}+\widetilde{g})\|^{p-1}_{L^p(\Omega)^d}, \end{aligned} \tag{4.2.8}$$

where $C_1 = \int_\Omega \overline{a}\ \nabla(\widetilde{\varphi}+\widetilde{g})\,\mathrm{d}\boldsymbol{x}$ (note that C_1 does not depend on $\widetilde{u}$). Applying Young's inequality to the third term in the right-hand side yields

$$\|\nabla(\widetilde{u}+\widetilde{g})\|^{p-1}_{L^p(\Omega)^d}\|\nabla(\widetilde{\varphi}+\widetilde{g})\|_{L^p(\Omega)^d} \le \frac{\varepsilon}{p'}\|\nabla(\widetilde{u}+\widetilde{g})\|^p_{L^p(\Omega)^d} + \frac{1}{p\varepsilon^{p/p'}}\|\nabla(\widetilde{\varphi}+\widetilde{g})\|^p_{L^p(\Omega)^d},$$

where $\varepsilon > 0$ is chosen to satisfy $\underline{a} - \mu\frac{\varepsilon}{p'} > 0$. Together with (4.2.8), we obtain

$$\langle \mathcal{A}(\widetilde{u}), \widetilde{u}-\widetilde{\varphi}\rangle \ge (\underline{a} - \mu\frac{\varepsilon}{p'})\|\nabla(\widetilde{u}+\widetilde{g})\|^p_{L^p(\Omega)^d} - C_2, \tag{4.2.9}$$

where $C_2 = C_1 + \mu\frac{1}{p\varepsilon^{p/p'}}\|\nabla(\widetilde{\varphi}+\widetilde{g})\|^p_{L^p(\Omega)^d}$. We always have $||x+y||^p_V \le 2^{(p-1)}(||x||^p_V + ||y||^p_V)$, so $||x||^p_V \ge 2^{(p-1)}||x+y||^p_V - ||y||^p_V$. We apply this inequality to $x = \nabla(\widetilde{u}+\widetilde{g})$ and $y = -\nabla\widetilde{g}$, to deduce

$$||\nabla(\widetilde{u}+\widetilde{g})||^p_{L^p(\Omega)^d} \ge 2^{(1-p)}||\nabla\widetilde{u}||^p_{L^p(\Omega)^d} - ||\nabla\widetilde{g}||_{L^p(\Omega)^d}.$$

Then Inequality (4.2.9) becomes, with $C_3 > 0$ and C_4 not depending on $\widetilde{u}$,

$$\langle \mathcal{A}(\widetilde{u}), \widetilde{u}-\widetilde{\varphi}\rangle \ge C_3||\nabla\widetilde{u}||^p_{L^p(\Omega)^d} + C_4.$$

Using this inequality, we can easily obtain (4.2.7). Hence the assumptions for [74, Theorem 8.2, Chapter 2] are satisfied, and therefore Problem (4.2.6) has at least one solution. □

4.2.2 The gradient discretisation method

Definition 4.2.3 (Gradient discretisation for the Signorini BCs). A gradient discretisation $\mathcal{D}$ for Signorini boundary conditions and nonhomogeneous Dirichlet BCs is $\mathcal{D} = (X_{\mathcal{D}}, \Pi_{\mathcal{D}}, \mathcal{I}_{\mathcal{D},\Gamma_1}, \mathbb{T}_{\mathcal{D}}, \nabla_{\mathcal{D}})$, where:

1. the set of discrete unknowns $X_{\mathcal{D}} = X_{\mathcal{D},\Gamma_{2,3}} \oplus X_{\mathcal{D},\Gamma_1}$ is a direct sum of two finite dimensional spaces on $\mathbb{R}$. The first space corresponds to the interior degrees of freedom and to the boundaries degrees of freedom on $\Gamma_2 \cup \Gamma_3$. The second space corresponds to the boundary degrees of freedom on Γ_1,
2. the linear mapping $\Pi_{\mathcal{D}} : X_{\mathcal{D}} \to L^p(\Omega)$ reconstructs functions from the degrees of freedom,
3. the linear mapping $\mathcal{I}_{\mathcal{D},\Gamma_1} : W^{1-\frac{1}{p},p}(\partial\Omega) \to X_{\mathcal{D},\Gamma_1}$ interpolates the traces of functions in $W^{1,p}(\Omega)$ on the degrees of freedom,
4. the linear mapping $\mathbb{T}_{\mathcal{D}} : X_{\mathcal{D}} \to L^p(\partial\Omega)$ reconstructs traces from the degrees of freedom,
5. the linear mapping $\nabla_{\mathcal{D}} : X_{\mathcal{D}} \to L^p(\Omega)^d$ reconstructs gradients from the degrees of freedom. It must be such that $\|\nabla_{\mathcal{D}} \cdot \|_{L^p(\Omega)^d}$ is a norm on $X_{\mathcal{D},\Gamma_{2,3}}$.

As already explained, the gradient scheme is obtained by taking the weak formulation (4.2.6) of the model, and replacing the continuous elements (space, function, gradient, trace...) by the discrete elements provided by the chosen gradient discretisation.

Definition 4.2.4 (Gradient schemes for Signorini problem). Let $\mathcal{D}$ be a gradient discretisation in the sense of Definition 4.2.3. The corresponding gradient scheme for Problem (4.2.6) is

$$\left\{\begin{array}{l} \text{Find } u \in \mathcal{K}_{\mathcal{D}} := \mathcal{K}_{\mathcal{D}} := \{v \in \mathcal{I}_{\mathcal{D},\Gamma_1} g + X_{\mathcal{D},\Gamma_{2,3}} \ : \ \mathbb{T}_{\mathcal{D}} v \le a \text{ on } \Gamma_3\} \text{ such that } \forall v \in \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}} u(\boldsymbol{x}), \nabla_{\mathcal{D}} u(\boldsymbol{x})) \cdot \nabla_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \le \int_\Omega f(\boldsymbol{x}) \Pi_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}. \end{array}\right. \tag{4.2.10}$$

We presented in Chapter 2 three properties called coercivity, GD-consistency and limit-conformity to assess the accuracy of gradient schemes for linear VI. These properties are sufficient to establish error estimates and prove the convergence of the GDM for VI based on *linear* differential operator. For nonlinear problems, an additional property called compactness is required to ensure the convergence of the GDM. Let us describe these four properties in the context of Signorini boundary conditions.

Definition 4.2.5 (Coercivity). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 4.2.3, set

$$C_{\mathcal{D}} = \max_{v \in X_{\mathcal{D},\Gamma_{2,3}} \setminus \{0\}} \Big(\frac{\|\Pi_{\mathcal{D}} v\|_{L^p(\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^p(\Omega)^d}} + \frac{\|\mathbb{T}_{\mathcal{D}} v\|_{L^p(\partial\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^p(\Omega)^d}} \Big). \tag{4.2.11}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretisations is **coercive** if $(C_{\mathcal{D}_m})_{m\in\mathbb{N}}$ remains bounded.

Definition 4.2.6 (GD-Consistency). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 4.2.3, define $S_{\mathcal{D}} : \mathcal{K} \to [0, +\infty)$ by

$$\forall \varphi \in \mathcal{K}, \ S_{\mathcal{D}}(\varphi) = \min_{v \in \mathcal{K}_{\mathcal{D}}} \left(\|\Pi_{\mathcal{D}} v - \varphi\|_{L^p(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla\varphi\|_{L^p(\Omega)^d} \right). \tag{4.2.12}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretisations is **GD-consistent** (or simply *consistent*, for short) if for all $\varphi \in \mathcal{K}$, $\lim_{m\to\infty} S_{\mathcal{D}_m}(\varphi) = 0$.

Definition 4.2.7 (Limit-conformity). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 4.2.3, define $W_{\mathcal{D}} : C^2(\overline{\Omega})^d \to [0,+\infty)$ by

$$\begin{aligned} &\forall \boldsymbol{\psi} \in C^2(\overline{\Omega})^d, \text{ such that } \boldsymbol{\psi}\cdot\mathbf{n} = 0 \text{ on } \Gamma_2,\\ &W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v\in X_{\mathcal{D},\Gamma_{2,3}}\setminus\{0\}} \frac{\left|\displaystyle\int_\Omega (\nabla_{\mathcal{D}} v\cdot\boldsymbol{\psi} + \Pi_{\mathcal{D}} v \,\mathrm{div}(\boldsymbol{\psi}))\,\mathrm{d}\boldsymbol{x} - \int_{\Gamma_3} \boldsymbol{\psi}\cdot\mathbf{n}\mathbb{T}_{\mathcal{D}} v\,\mathrm{d}\boldsymbol{x}\right|}{\|\nabla_{\mathcal{D}} v\|_{L^p(\Omega)^d}}. \end{aligned} \tag{4.2.13}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretisations is **limit-conforming** if, for all $\psi \in C^2(\overline{\Omega})^d$ such that $\boldsymbol{\psi}\cdot\mathbf{n} = 0$ on Γ_2, $\lim_{m\to\infty} W_{\mathcal{D}_m}(\psi) = 0$.

Remark 4.2.2. The convergence of sequence of approximate solution is obtained by following the compactness technique, which requires finding a weak limit to the reconstructed function, its trace and its gradient. In order to prove the existence of such a limit, which is the result of [37, Lemma 2.57], we need to apply the above definition only to smooth functions $\boldsymbol{\psi}$. Therefore the function $W_{\mathcal{D}}$ in the above definition only needs to be defined on the space $C^2(\overline{\Omega})^d$.

Definition 4.2.8 (Compactness). A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretisations is **compact** if, for any sequence $(u_m)_{m\in\mathbb{N}}$ with $u_m \in \mathcal{K}_{\mathcal{D}_m}$ and such that $(\|\nabla_{\mathcal{D}_m} u_m\|_{L^p(\Omega)^d})_{m\in\mathbb{N}}$ is bounded, the sequence $(\Pi_{\mathcal{D}_m} u_m)_{m\in\mathbb{N}}$ is relatively compact in $L^p(\Omega)$.

4.2.3 Convergence results

We can now state and prove our main convergence theorem for the gradient discretisation method applied to the nonlinear Signorini problem.

Theorem 4.2.9 (Convergence of the GDM for the nonlinear Signorini problem). *Under Assumptions 4.2.1, let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretisations in the sense of Definition 4.2.3, which is coercive, GD-consistent, limit-conforming and compact, and such that $\mathcal{K}_{\mathcal{D}_m}$ is nonempty for any m. Then, for any $m\in\mathbb{N}$, the gradient scheme* (4.2.10) *has at least one solution $u_m \in \mathcal{K}_{\mathcal{D}_m}$.*

Assume furthermore that

$$\begin{aligned} &\exists \varphi_g \in W^{1,p}(\Omega) \text{ s.t. } \gamma(\varphi_g) = g \text{ and}\\ &\lim_{m\to\infty} \min\{\|\Pi_{\mathcal{D}_m} v_m - \varphi_g\|_{L^p(\Omega)} + \|\mathbb{T}_{\mathcal{D}_m} v_m - \gamma(\varphi_g)\|_{L^p(\Gamma_3)}\\ &\qquad + \|\nabla_{\mathcal{D}_m} v_m - \nabla\varphi_g\|_{L^p(\Omega)^d} \,:\, v - \mathcal{I}_{\mathcal{D}_m,\Gamma_1}\gamma(\varphi_g) \in X_{\mathcal{D}_m,\Gamma_{2,3}}\} = 0. \end{aligned} \tag{4.2.14}$$

Then, up to a subsequence, $\Pi_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)$ to a weak solution $\bar{u}$ of Problem (4.2.6)*, and $\nabla_{\mathcal{D}_m} u_m$ converges weakly to $\nabla\bar{u}$ in $L^p(\Omega)^d$.*

Moreover, if $\mathbf{a}$ *is strictly monotonic in the sense*

$$(\mathbf{a}(\boldsymbol{x},s,\xi) - \mathbf{a}(\boldsymbol{x},s,\chi))\cdot(\xi-\chi) > 0, \text{ for a.e. } \boldsymbol{x}\in\Omega,\ \forall s\in\mathbb{R}, \forall \xi,\chi\in\mathbb{R}^d \text{ with } \xi\neq\chi, \tag{4.2.15}$$

then $\nabla_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)^d$ to $\nabla\bar{u}$.

Remark 4.2.3. Assumption (4.2.14) is obviously always satisfied if $g = 0$ (take $\varphi_g = 0$). For most sequences of gradient discretisations, the convergence stated in (4.2.14) actually holds for any φ with $g = \gamma(\varphi)$, and corresponds to the GD-consistency of the method for nonhomogeneous Fourier BCs (see [37, Remark 2.58 and Definition 2.49]).

Proof. The proof is inspired from [40], and follows the general path described in [34, Section 1.2] and [35, Section 2.2].

Step 1: existence of a solution to the GS.
Let $\widetilde{g} \in \mathcal{K}$ be a lifting of g, such that $\gamma(\widetilde{g}) = g$. Introduce

$$g_{\mathcal{D}} = \operatorname*{argmin}_{v \in \mathcal{K}_{\mathcal{D}}} \left(\|\Pi_{\mathcal{D}} v - \widetilde{g}\|_{L^p(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla \widetilde{g}\|_{L^p(\Omega)^d} \right). \tag{4.2.16}$$

Let $\langle \cdot, \cdot \rangle$ be the duality product between the finite dimensional space $X_{\mathcal{D},\Gamma_{2,3}}$ and its dual $X'_{\mathcal{D},\Gamma_{2,3}}$. Define the operator $\mathcal{A}_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to X'_{\mathcal{D},\Gamma_{2,3}}$ by, for $\widehat{u}, \widehat{w} \in X_{\mathcal{D},\Gamma_{2,3}}$,

$$\langle \mathcal{A}_{\mathcal{D}}(\widehat{u}), \widehat{w} \rangle = \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}}(\widehat{u} + g_{\mathcal{D}})(\boldsymbol{x}), \nabla_{\mathcal{D}}(\widehat{u} + g_{\mathcal{D}})(\boldsymbol{x})) \cdot \nabla_{\mathcal{D}}(\widehat{w} + g_{\mathcal{D}})(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

Applying the same reasoning as in [74], we check that $\mathcal{A}_{\mathcal{D}}$ is an operator of the calculus of variations (this is made extremely easy here, due to the finite dimension of $X_{\mathcal{D},\Gamma_{2,3}}$). See Proposition 4.A.4 for details.

The existence of a solution to the scheme (4.2.10) is then a consequence of [74, Theorem 8.2, Chapter 2] since, setting $\widehat{u} = u - g_{\mathcal{D}}$ and $\widehat{w} = w - g_{\mathcal{D}}$, this scheme can be re-written

$$\begin{gathered} \text{find } \widehat{u} \in \mathcal{K}_{\mathcal{D}} - g_{\mathcal{D}}, \text{ such that for all } \widehat{w} \in \mathcal{K}_{\mathcal{D}} - g_{\mathcal{D}}, \\ \langle \mathcal{A}_{\mathcal{D}}(\widehat{u}), \widehat{u} - \widehat{w} \rangle \le \ell(\widehat{u} - \widehat{w}), \end{gathered}$$

where $\ell \in X'_{\mathcal{D},\Gamma_{2,3}}$ is defined by $\ell(\widehat{w}) = \int_\Omega f \Pi_{\mathcal{D}} \widehat{w} \, \mathrm{d}x$.

Step 2: convergence towards the solution to the continuous model.
Let us start by estimating $\|\nabla_{\mathcal{D}_m} u_m\|_{L^p(\Omega)^d}$. In (4.2.10), set $u := u_m$, and $v := v_m$ a generic element in $\mathcal{K}_{\mathcal{D}_m}$. By using the Holder's inequality and due to the coercivity assumption (4.2.4), it follows that

$$\begin{aligned} \underline{a} \|\nabla_{\mathcal{D}_m} u_m\|^p_{L^p(\Omega)^d} &\le \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \, \mathrm{d}\boldsymbol{x} \\ &\le \|f\|_{L^{p'}(\Omega)} \|\Pi_{\mathcal{D}_m}(u_m - v_m)\|_{L^p(\Omega)} + \|\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m)\|_{L^{p'}(\Omega)^d} \|\nabla_{\mathcal{D}_m} v_m\|_{L^p(\Omega)^d}. \end{aligned}$$

Since $u_m - v_m$ is an element in $X_{\mathcal{D}_m,\Gamma_{2,3}}$, applying the coercivity property (see Definition 4.2.5) gives C_p not depending on m such that $\|\Pi_{\mathcal{D}_m}(u_m - v_m)\|_{L^p(\Omega)} \le C_p \|\nabla_{\mathcal{D}_m}(u_m - v_m)\|_{L^p(\Omega)^d}$, and thus, using the growth assumption (4.2.3),

$$\begin{aligned} \underline{a} \|\nabla_{\mathcal{D}_m} u_m\|^p_{L^p(\Omega)^d} \le{}& C_p \|f\|_{L^{p'}(\Omega)} \left(\|\nabla_{\mathcal{D}_m} u_m\|_{L^p(\Omega)^d} + \|\nabla_{\mathcal{D}_m} v_m\|_{L^p(\Omega)^d} \right) \\ &+ \left(\|\overline{a}\|_{L^{p'}(\Omega)^d} + \mu \|\nabla_{\mathcal{D}_m} u_m\|^{p-1}_{L^p(\Omega)^d} \right) \|\nabla_{\mathcal{D}_m} v_m\|_{L^p(\Omega)^d}. \end{aligned}$$

Applying Young's inequality to this relation shows that

$$\|\nabla_{\mathcal{D}_m} u_m\|^p_{L^p(\Omega)^d} \le C_1 \left(\|\nabla_{\mathcal{D}_m} v_m\|^p_{L^p(\Omega)^d} + \|f\|^{p'}_{L^{p'}(\Omega)} + \|\overline{a}\|^{p'}_{L^{p'}(\Omega)^d} \right), \tag{4.2.17}$$

where C_1 does not depend on m. Let us now define, for $\varphi \in \mathcal{K}$, an element $P_{\mathcal{D}_m}\varphi$ of $\mathcal{K}_{\mathcal{D}_m}$ by

$$P_{\mathcal{D}_m}(\varphi) = \operatorname*{argmin}_{v \in \mathcal{K}_{\mathcal{D}_m}} (\|\Pi_{\mathcal{D}_m} v - \varphi\|_{L^p(\Omega)} + \|\nabla_{\mathcal{D}_m} v - \nabla \varphi\|_{L^p(\Omega)^d}). \tag{4.2.18}$$

Take $\varphi \in \mathcal{K}$ and let $v_m := P_{\mathcal{D}_m}\varphi$ in (4.2.17). By the triangle inequality

$$\|\nabla_{\mathcal{D}_m} v_m\|_{L^p(\Omega)^d} \le S_{\mathcal{D}_m}(\varphi) + \|\nabla \varphi\|_{L^p(\Omega)^d},$$

and the GS-consistency of $\mathcal{D}_m$ shows that $\|\nabla_{\mathcal{D}_m} v_m\|_{L^p(\Omega)^d}$ is bounded. This shows that $\|\nabla_{\mathcal{D}_m} u_m\|_{L^p(\Omega)^d}$ remains bounded.

Now, using (4.2.14), [37, Lemma 2.57] (slightly adjusted to the fact that the limit-conformity involves here functions such that $\boldsymbol{\psi}\cdot\mathbf{n}=0$ on Γ_2, see (4.2.13)) asserts the existence of $\bar u\in W^{1,p}(\Omega)$ and a subsequence, still denoted by $(\mathcal{D}_m)_{m\in\mathbb{N}}$, such that $\gamma\bar u = g$ on Γ_1, $\Pi_{\mathcal{D}_m}u_m$ converges weakly to $\bar u$ in $L^p(\Omega)$, $\nabla_{\mathcal{D}_m}u_m$ converges weakly to $\nabla\bar u$ in $L^p(\Omega)^d$ and $\mathbb{T}_{\mathcal{D}_m}u_m$ converges weakly to $\gamma\bar u$ in $L^p(\Gamma_3)$. Since $u_m\in\mathcal{K}_{\mathcal{D}_m}$, we have $\mathbb{T}_{\mathcal{D}_m}u_m\le a$ on Γ_3 and thus $\gamma\bar u\le a$ on Γ_3. In other words, $\bar u$ belongs to $\mathcal{K}$. By the compactness hypothesis, the convergence of $\Pi_{\mathcal{D}_m}u_m$ to $\bar u$ is actually strong in $L^p(\Omega)$. Up to another subsequence, we can therefore assume that this convergence holds almost everywhere on Ω.

To complete this step, it remains to show that $\bar u$ is a solution to (4.2.6). We use the Minty trick. From assumption (4.2.3), the sequence $\mathcal{A}_{\mathcal{D}_m}=\mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\nabla_{\mathcal{D}_m}u_m)$ is bounded in $L^{p'}(\Omega)^d$ and converges weakly up to a subsequence to some $\mathcal{A}$ in $L^{p'}(\Omega)^d$. Owing to the GD-consistency of the gradient discretisations, for all $\varphi\in\mathcal{K}$ we have $\Pi_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi)\to\varphi$ strongly in $L^p(\Omega)$ and $\nabla_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi)\to\nabla\varphi$ strongly in $L^p(\Omega)^d$. Taking $v:=P_{\mathcal{D}_m}\varphi$ as a test function in the gradient scheme (4.2.10) and passing to the superior limit gives

$$
\begin{aligned}
&\limsup_{m\to\infty}\int_\Omega \mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\nabla_{\mathcal{D}_m}u_m)\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}\\
&\quad\le \limsup_{m\to\infty}\left(\int_\Omega f(\Pi_{\mathcal{D}_m}u_m-\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\varphi)\,\mathrm{d}\boldsymbol{x}+\int_\Omega \mathcal{A}_{\mathcal{D}_m}\cdot\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\varphi\,\mathrm{d}\boldsymbol{x}\right)\\
&\quad\le \int_\Omega f(\bar u-\varphi)\,\mathrm{d}\boldsymbol{x}+\int_\Omega\mathcal{A}\cdot\nabla\varphi\,\mathrm{d}\boldsymbol{x},\quad \text{for all }\varphi\text{ in }\mathcal{K}.
\end{aligned}
$$

Choosing $\varphi=\bar u$, yields

$$
\limsup_{m\to\infty}\int_\Omega \mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\nabla_{\mathcal{D}_m}u_m)\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}\le\int_\Omega\mathcal{A}\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}. \tag{4.2.19}
$$

Using the monotonicity assumption (4.2.5), one writes, for $\mathbf{G}\in L^p(\Omega)^d$,

$$
\begin{aligned}
&\liminf_{m\to\infty}\Big[\int_\Omega \mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\nabla_{\mathcal{D}_m}u_m)\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}-\int_\Omega \mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\nabla_{\mathcal{D}_m}u_m)\cdot\mathbf{G}\,\mathrm{d}\boldsymbol{x}\\
&\qquad-\int_\Omega \mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\mathbf{G})\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}+\int_\Omega \mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\mathbf{G})\cdot\mathbf{G}\,\mathrm{d}\boldsymbol{x}\Big]\\
&\quad=\liminf_{m\to\infty}\int_\Omega\Big[\mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\nabla_{\mathcal{D}_m}u_m)-\mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\mathbf{G})\Big]\cdot\Big[\nabla_{\mathcal{D}_m}u_m-\mathbf{G}\Big]\,\mathrm{d}\boldsymbol{x}\\
&\quad\ge 0.
\end{aligned}\tag{4.2.20}
$$

The strong convergence of $\Pi_{\mathcal{D}_m}u_m$ shows that $\mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\mathbf{G})\to\mathbf{a}(\boldsymbol{x},\bar u,\mathbf{G})$ in $L^{p'}(\Omega)^d$. Hence, passing to the limit in (4.2.20),

$$
\begin{aligned}
&\liminf_{m\to\infty}\int_\Omega \mathbf{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}u_m,\nabla_{\mathcal{D}_m}u_m)\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}-\int_\Omega\mathcal{A}\cdot\nabla\mathbf{G}\,\mathrm{d}\boldsymbol{x}\\
&\qquad-\int_\Omega\mathbf{a}(\boldsymbol{x},\bar u,\mathbf{G})\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}+\int_\Omega\mathbf{a}(\boldsymbol{x},\bar u,\mathbf{G})\cdot\mathbf{G}\,\mathrm{d}\boldsymbol{x}\ge 0.
\end{aligned}\tag{4.2.21}
$$

Combining this inequality with (4.2.19) yields the following inequality

$$
\int_\Omega\mathcal{A}\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}-\int_\Omega\mathcal{A}\cdot\mathbf{G}\,\mathrm{d}\boldsymbol{x}-\int_\Omega\mathbf{a}(\boldsymbol{x},\bar u,\mathbf{G})\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}+\int_\Omega\mathbf{a}(\boldsymbol{x},\bar u,\mathbf{G})\cdot\mathbf{G}\,\mathrm{d}\boldsymbol{x}\ge 0.
$$

Assume that $\boldsymbol{\phi} \in C_c^\infty(\Omega)^d$ and $\alpha > 0$. Putting $\mathbf{G} = \nabla\bar{u} + \alpha\boldsymbol{\phi}$ and dividing by α, one obtains

$$-\int_\Omega (\mathcal{A} - \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u} + \alpha\boldsymbol{\phi})) \cdot \boldsymbol{\phi} \, \mathrm{d}\boldsymbol{x} \geq 0, \ \forall \boldsymbol{\phi} \in C_c^\infty(\Omega)^d, \ \forall \alpha > 0.$$

Letting $\alpha \to 0$ and applying the dominated convergence theorem yields

$$-\int_\Omega (\mathcal{A} - \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u})) \cdot \boldsymbol{\phi} \, \mathrm{d}\boldsymbol{x} \geq 0, \ \forall \boldsymbol{\phi} \in C_c^\infty(\Omega)^d.$$

Applied to $-\boldsymbol{\phi}$ instead of $\boldsymbol{\phi}$, this leads to

$$-\int_\Omega (\mathcal{A} - \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u})) \cdot \boldsymbol{\phi} \, \mathrm{d}\boldsymbol{x} = 0, \ \forall \boldsymbol{\phi} \in C_c^\infty(\Omega)^d,$$

which implies that

$$\mathcal{A} = \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u}), \text{ a.e. on } \Omega. \tag{4.2.22}$$

Setting $\mathbf{G} = \nabla\bar{u}$ in (4.2.21), it follows that

$$\int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u}) \cdot \nabla\bar{u} \, \mathrm{d}\boldsymbol{x} \leq \liminf_{m\to\infty} \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \, \mathrm{d}\boldsymbol{x}, \tag{4.2.23}$$

which gives, since u_m is a solution to the gradient scheme (4.2.10), for all $\varphi \in \mathcal{K}$,

$$\int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u}) \cdot \nabla\bar{u} \, \mathrm{d}\boldsymbol{x} \leq \liminf_{m\longrightarrow\infty} \left[\int_\Omega f \Pi_{\mathcal{D}_m}(u_m - P_{\mathcal{D}_m}\varphi) \, \mathrm{d}\boldsymbol{x} + \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi) \, \mathrm{d}\boldsymbol{x} \right].$$

Using (4.2.22) and the strong convergence of $\nabla_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi)$ to $\nabla\varphi$ yields

$$\int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u}) \cdot \nabla\bar{u} \, \mathrm{d}\boldsymbol{x} \leq \int_\Omega f(\bar{u} - \varphi) \, \mathrm{d}\boldsymbol{x} + \int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u}) \cdot \nabla\varphi \, \mathrm{d}\boldsymbol{x},$$

and therefore shows that $\bar{u}$ is a solution to (4.2.6).

Step 3: strong convergence of the gradients, if a is strictly monotonic.
Owing to (4.2.19) and (4.2.22),

$$\limsup_{m\to\infty} \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \, \mathrm{d}\boldsymbol{x} \leq \int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u}) \cdot \nabla\bar{u} \, \mathrm{d}\boldsymbol{x}. \tag{4.2.24}$$

Together with (4.2.23), we conclude that

$$\lim_{m\to\infty} \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \, \mathrm{d}\boldsymbol{x} = \int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u}) \cdot \nabla\bar{u} \, \mathrm{d}\boldsymbol{x}. \tag{4.2.25}$$

The remaining reasoning to obtain the strong convergence of $\nabla_{\mathcal{D}_m} u_m$ is exactly like in [40]. For the sake of completeness, we recall it. Equality (4.2.25) leads to

$$\lim_{m\to\infty} \int_\Omega (\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) - \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u})) \cdot (\nabla_{\mathcal{D}_m} u_m - \nabla\bar{u}) \, \mathrm{d}\boldsymbol{x} = 0.$$

Making use of the fact that $(\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) - \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u})) \cdot (\nabla_{\mathcal{D}_m} u_m - \nabla\bar{u}) \geq 0$ for a.e. $\boldsymbol{x} \in \Omega$, we deduce that

$$(\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) - \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u})) \cdot (\nabla_{\mathcal{D}_m} u_m - \nabla\bar{u}) \to 0 \text{ in } L^1(\Omega).$$

Up a subsequence, the convergence holds almost everywhere. The strict monotonicity assumption (4.2.15) and [40, Lemma 3.2] yield $\nabla_{\mathcal{D}_m} u_m \to \nabla \bar{u}$ a.e. as $m \to \infty$. Furthermore, as a consequence, $\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \to \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \nabla \bar{u}$ a.e. Since $\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \geq 0$, and taking into account (4.2.25), [40, Lemma 3.3] gives the strong convergence of $\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m$ to $\mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \cdot \nabla \bar{u}$ in $L^1(\Omega)$ as $m \to \infty$. As a consequence of this L^1-convergence, we obtain the equi-integrability of the sequence of functions $\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m$. This provides, with (4.2.4), the equi-integrability of $(|\nabla_{\mathcal{D}_m} u_m|^p)_{m\in\mathbb{N}}$. The strong convergence of $\nabla_{\mathcal{D}_m} u_m$ to $\nabla \bar{u}$ in $L^p(\Omega)$ is then directly implied by the Vitali theorem.

□

4.3 Nonlinear obstacle problem and generalised Bulkley fluid model

4.3.1 Continuous problems

4.3.1.1 Nonlinear obstacle problem

We are concerned here with other kinds of variational inequalities problems. The first one is an obstacle model, in which the inequalities are imposed inside the domain Ω. It is formulated as

$$(\operatorname{div} \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) + f)(\psi - \bar{u}) = 0 \quad \text{in } \Omega, \tag{4.3.1a}$$
$$-\operatorname{div} \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) \leq f \quad \text{in } \Omega, \tag{4.3.1b}$$
$$\bar{u} \leq \psi \quad \text{in } \Omega, \tag{4.3.1c}$$
$$\bar{u} = h \quad \text{on } \partial\Omega. \tag{4.3.1d}$$

Assumptions 4.3.1. *The assumptions on the data of the obstacle model* (4.3.1) *are:*

1. *the operator* $\mathbf{a}$ *and the domain* Ω *satisfy the same properties as in Assumption 4.2.1,*
2. *the function* f *belongs to* $L^{p'}(\Omega)$*, the boundary function* h *is in* $W^{1-\frac{1}{p},p}(\partial\Omega)$ *and the obstacle function* ψ *belongs to* $L^p(\Omega)$*,*
3. *the closed convex set* $\mathcal{K} := \{v \in W^{1,p}(\Omega) \ : \ v \leq \psi \text{ in } \Omega,\ \gamma(v) = h \text{ on } \partial\Omega\}$ *is nonempty.*

The weak formulation of the obstacle problem (4.3.1) is

$$\left\{ \begin{array}{l} \text{Find } \bar{u} \in \mathcal{K} \text{ such that, for all } v \in \mathcal{K}, \\ \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}(\boldsymbol{x}), \nabla \bar{u}(\boldsymbol{x})) \cdot \nabla(\bar{u} - v)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \leq \int_\Omega f(\boldsymbol{x})(\bar{u} - v)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \end{array} \right. \tag{4.3.2}$$

4.3.1.2 Generalised Bulkley model

The second problem is called Bulkley model, whose weak formulation is given by

$$\left\{ \begin{array}{l} \text{Find } \bar{u} \in W_0^{1,p}(\Omega) \text{ such that, for all } v \in W_0^{1,p}(\Omega), \\ \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}(\boldsymbol{x}), \nabla \bar{u}(\boldsymbol{x})) \cdot \nabla(\bar{u} - v)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} + \int_\Omega |\nabla \bar{u}(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} - \int_\Omega |\nabla v(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \\ \displaystyle\leq \int_\Omega f(\boldsymbol{x})(\bar{u} - v)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \end{array} \right. \tag{4.3.3}$$

Here the operator $\mathbf{a}$ is assumed to satisfies (4.2.2)–(4.2.5) and the domain Ω has a Lipschitz boundary. Models considered in the removal of materials from a duct by using fluids [54] are included in (4.3.3) by setting $\mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla \bar{u}) = |\nabla \bar{u}|^{p-2} \nabla \bar{u}$.

As for the Signorini problem, [74, Theorems 8.2 and 8.5, Chap. 2] and respectively yield the existence of a solution to problems (4.3.2) and (4.3.3).

4.3.2 Discrete problems

4.3.2.1 Obstacle problem

Let us recall the definition of a gradient discretisation for nonhomogeneous Dirichlet boundary conditions [37].

Definition 4.3.2 (GD for nonhomogeneous Dirichlet boundary conditions). A gradient discretisation $\mathcal{D}$ for nonhomogeneous Dirichlet boundary conditions is defined by $\mathcal{D} = (X_{\mathcal{D}}, \Pi_{\mathcal{D}}, \mathcal{I}_{\mathcal{D},\partial\Omega}, \nabla_{\mathcal{D}})$, where:

1. the set of discrete unknowns $X_{\mathcal{D}} = X_{\mathcal{D},0} \oplus X_{\mathcal{D},\partial\Omega}$ is a direct sum of two finite dimensional spaces on $\mathbb{R}$, representing respectively the interior degrees of freedom and the boundary degrees of freedom,
2. the linear mapping $\Pi_{\mathcal{D}} : X_{\mathcal{D}} \to L^p(\Omega)$ provides the reconstructed function,
3. the linear mapping $\mathcal{I}_{\mathcal{D},\partial\Omega} : W^{1-\frac{1}{p},p}(\partial\Omega) \to X_{\mathcal{D},\partial\Omega}$ provides an interpolation operator for the trace of functions in $W^{1,p}(\Omega)$,
4. the linear mapping $\nabla_{\mathcal{D}} : X_{\mathcal{D}} \to L^p(\Omega)^d$ gives a reconstructed gradient, which must be defined such that $\|\nabla_{\mathcal{D}} \cdot \|_{L^p(\Omega)^d}$ is a norm on $X_{\mathcal{D},0}$.

Definition 4.3.3 (GS for the nonlinear obstacle problem). Let $\mathcal{D}$ be a gradient discretisation in the sense of Definition 4.3.2. The corresponding gradient scheme for (4.3.2) is given by

$$\left\{\begin{array}{l}\text{Find } u \in \mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},0} + \mathcal{I}_{\mathcal{D},\partial\Omega} h \ : \ \Pi_{\mathcal{D}} v \le \psi \text{ in } \Omega\} \ \text{ s.t., } \forall v \in \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}} u(\boldsymbol{x}), \nabla_{\mathcal{D}} u(\boldsymbol{x})) \cdot \nabla_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \le \int_\Omega f(\boldsymbol{x}) \Pi_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.\end{array}\right. \tag{4.3.4}$$

4.3.2.2 Generalised Bulkley model

Definition 4.3.4 (GD for homogeneous Dirichlet boundary conditions). A gradient discretisation $\mathcal{D}$ for homogeneous Dirichlet boundary conditions $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$, where $X_{\mathcal{D},0}$ is a finite dimensional vector space over $\mathbb{R}$, taking into account the zero boundary condition in the space $W_0^{1,p}(\Omega)$, and $\Pi_{\mathcal{D}}$ and $\nabla_{\mathcal{D}}$ are as in Definition 4.3.2 but defined on $X_{\mathcal{D},0}$.

Note that the gradient discretisation $\mathcal{D}$ defined in Definition 2.2.5 is a particular case of this definition when $p = 2$.

Definition 4.3.5 (GS for the Bulkley model). Let $\mathcal{D}$ be a gradient discretisation in the sense of Definition 4.3.4. The corresponding gradient scheme for (4.3.3) is given by

$$\left\{\begin{array}{l}\text{Find } \bar{u} \in X_{\mathcal{D},0}, \text{ such that for all } v \in X_{\mathcal{D},0}, \\ \displaystyle\int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}} u(\boldsymbol{x}), \nabla_{\mathcal{D}} u(\boldsymbol{x})) \cdot \nabla_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} + \int_\Omega |\nabla_{\mathcal{D}} u(\boldsymbol{x})| \,\mathrm{d}\boldsymbol{x} - \int_\Omega |\nabla_{\mathcal{D}} v(\boldsymbol{x})| \,\mathrm{d}\boldsymbol{x} \\ \displaystyle\le \int_\Omega f(\boldsymbol{x}) \Pi_{\mathcal{D}}(u-v)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.\end{array}\right. \tag{4.3.5}$$

4.3.2.3 Properties of GD

Except for the restriction to the convex sets $\mathcal{K}$ and $\mathcal{K}_{\mathcal{D}}$ in the GD-consistency, all the properties of GD required for the convergence analysis of the GDM on the nonlinear obstacle and Bulkley models are similar to the corresponding ones for GD adapted to PDEs as in [40, 37].

Definition 4.3.6 (Coercivity). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 4.3.2 or Definition 4.3.4, define

$$C_{\mathcal{D}} = \max_{v \in X_{\mathcal{D},0} \setminus \{0\}} \frac{\|\Pi_{\mathcal{D}} v\|_{L^p(\Omega)}}{\|\nabla_{\mathcal{D}} v\|_{L^p(\Omega)^d}}. \tag{4.3.6}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of such gradient discretisations is **coercive** if $(C_{\mathcal{D}_m})_{m\in\mathbb{N}}$ remains bounded.

Definition 4.3.7 (GD-Consistency). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 4.3.2, let $S_{\mathcal{D}} : \mathcal{K} \to [0, +\infty)$ be defined by

$$\forall \varphi \in \mathcal{K},\ S_{\mathcal{D}}(\varphi) = \min_{v \in \mathcal{K}_{\mathcal{D}}} \left(\|\Pi_{\mathcal{D}} v - \varphi\|_{L^p(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla \varphi\|_{L^p(\Omega)^d} \right). \tag{4.3.7}$$

If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 4.3.4, $S_{\mathcal{D}}$ is defined the same way with $(\mathcal{K}, \mathcal{K}_{\mathcal{D}})$ replaced by $(W_0^{1,p}(\Omega), X_{\mathcal{D},0})$. A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of such gradient discretisations is **GD-consistent** if for all $\varphi \in \mathcal{K}$, $\lim_{m\to\infty} S_{\mathcal{D}_m}(\varphi) = 0$.

Definition 4.3.8 (Limit-conformity). If $\mathcal{D}$ is a gradient discretisation in the sense of Definition 4.3.2 or Definition 4.3.4, define $W_{\mathcal{D}} : W^{\mathrm{div},p'}(\Omega) \to [0, +\infty)$ by

$$\forall \boldsymbol{\psi} \in C^2(\overline{\Omega})^d, \quad W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v \in X_{\mathcal{D},0} \setminus \{0\}} \frac{1}{\|\nabla_{\mathcal{D}} v\|_{L^p(\Omega)^d}} \left| \int_\Omega (\nabla_{\mathcal{D}} v \cdot \boldsymbol{\psi} + \Pi_{\mathcal{D}} v \,\mathrm{div}(\boldsymbol{\psi}))\,\mathrm{d}\boldsymbol{x} \right|. \tag{4.3.8}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of such gradient discretisations is **limit-conforming** if $\lim_{m\to\infty} W_{\mathcal{D}_m}(\psi) = 0$ for all $\psi \in C^2(\overline{\Omega})^d$.

Finally, Definition 4.2.8 (compactness) remains the same as the case of gradient discretisation $\mathcal{D}$ in the sense of Definition 4.3.2 or Definition 4.3.4, with $\mathcal{K}_{\mathcal{D}_m}$ replaced by $X_{\mathcal{D}_m,0}$ in the latter case.

4.3.3 Convergence results

The following two theorems state the convergence properties of the GDM for the nonlinear obstacle problem and the Bulkley model.

Theorem 4.3.9 (Convergence of the GDM for the nonlinear obstacle problem). *Under Assumptions 4.3.1, let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretisations in the sense of Definition 4.3.2, which is coercive, GD-consistent, limit-conforming and compact, and such that $\mathcal{K}_{\mathcal{D}_m}$ is a nonempty set for any m.*

Then, for any $m \in \mathbb{N}$, the gradient scheme (4.3.4) has at least one solution $u_m \in \mathcal{K}_{\mathcal{D}_m}$ and, up to a subsequence, $\Pi_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)$ to a weak solution $\bar{u}$ of Problem (4.3.2) and $\nabla_{\mathcal{D}_m} u_m$ converges weakly in $L^p(\Omega)^d$ to $\nabla\bar{u}$.

If the strict monotonicity (4.2.15) is assumed, then $\nabla_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)^d$ to $\nabla\bar{u}$.

Theorem 4.3.10 (Convergence of the GDM for the Bulkley model). *Under Assumptions (4.2.2)–(4.2.5) and $f \in L^{p'}(\Omega)$, let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretisations in the sense of Definition 4.3.4, which is coercive, GD-consistent, limit-conforming and compact. Then, for any $m \in \mathbb{N}$, the gradient scheme (4.3.5) has at least one solution $u_m \in X_{\mathcal{D},0}$ and, up to a subsequence, $\Pi_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)$ to a weak solution $\bar{u}$ of Problem (4.3.3) and $\nabla_{\mathcal{D}_m} u_m$ converges weakly to $\nabla\bar{u}$ in $L^p(\Omega)^d$.*

If we also assume that **a** *is strictly monotonic in the sense of (4.2.15), then $\nabla_{\mathcal{D}_m} u_m$ converges strongly in $L^p(\Omega)^d$ to $\nabla\bar{u}$.*

The proof of Theorem 4.3.9 is extremely similar to the proof of Theorem 4.2.9. We therefore only provide the proof of Theorem 4.3.10.

Proof of Theorem 4.3.10. Let us define the operator $\mathcal{A}_{\mathcal{D}} : X_{\mathcal{D},0} \to X'_{\mathcal{D},0}$ and the functional $J_{\mathcal{D}} : X_{\mathcal{D},0} \to \mathbb{R}^+$ as follows:

$$\langle \mathcal{A}_{\mathcal{D}}(u), v\rangle = \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}} u(\boldsymbol{x}), \nabla_{\mathcal{D}} u(\boldsymbol{x})) \cdot \nabla_{\mathcal{D}} v(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \quad \text{and}$$
$$J_{\mathcal{D}}(u) = \int_\Omega |\nabla_{\mathcal{D}} u(\boldsymbol{x})| \,\mathrm{d}\boldsymbol{x},$$

where $\langle \cdot, \cdot \rangle$ is the duality product between $X'_{\mathcal{D},0}$ and $X_{\mathcal{D},0}$. Applying the same arguments as in [74], one can easily prove that the operator $\mathcal{A}_{\mathcal{D}}$ is pseudo-monotone and obtain, since $J_{\mathcal{D}} \geq 0$,

$$\frac{\langle \mathcal{A}_{\mathcal{D}}(u), u - \phi\rangle + J_{\mathcal{D}}(u)}{\|\nabla_{\mathcal{D}} u\|_{L^p(\Omega)^d}} \to +\infty \quad \text{as} \quad \|\nabla_{\mathcal{D}} u\|_{L^p(\Omega)^d} \to \infty.$$

A direct application of [74, Theorem 8.5, Chapter 2] gives the existence of a solution to (4.3.5).

We now show that $\|\nabla_{\mathcal{D}_m} u_m\|_{L^p(\Omega)^d}$ is bounded. Choose $u := u_m$, and $v := 0 \in X_{\mathcal{D}_m,0}$ in (4.3.5). Due to the coercivity assumption (4.2.4), the Hölder inequality and the coercivity, one has

$$\begin{aligned}
\underline{a}\|\nabla_{\mathcal{D}_m} u_m\|^p_{L^p(\Omega)^d} &\leq \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \,\mathrm{d}\boldsymbol{x}\\
&\leq \int_\Omega f \Pi_{\mathcal{D}_m} u_m \,\mathrm{d}\boldsymbol{x}\\
&\leq C_p \|f\|_{L^{p'}(\Omega)} \|\nabla_{\mathcal{D}_m} u_m\|_{L^p(\Omega)^d}.
\end{aligned}$$

This shows that $\|\nabla_{\mathcal{D}_m} u_m\|_{L^p(\Omega)^d}$ is bounded. According to [37, Lemma 2.12], there exists $\bar{u} \in W_0^{1,p}(\Omega)$ and a subsequence, denoted by the same way $(\mathcal{D}_m)_{m\in\mathbb{N}}$, such that $\Pi_{\mathcal{D}_m} u_m$ converges weakly to $\bar{u}$ in $L^p(\Omega)$ and $\nabla_{\mathcal{D}_m} u_m$ converges weakly to $\nabla\bar{u}$ in $L^p(\Omega)^d$. In fact, the strong convergence of the sequence $\Pi_{\mathcal{D}_m} u_m$ to $\bar{u}$ in $L^p(\Omega)$ is ensured by the compactness property. The growth assumption (4.2.3) shows that the sequence $\mathcal{A}_{\mathcal{D}_m} = \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m)$ is bounded in $L^{p'}(\Omega)^d$ and thus, up to a subsequence, that it converges weekly to some $\mathcal{A}$ in this space.

Defining $P_{\mathcal{D}_m}$ as in (4.2.18) with $\mathcal{K}$ and $\mathcal{K}_{\mathcal{D}_m}$ replaced with $W_0^{1,p}(\Omega)$ and $X_{\mathcal{D},0}$, respectively, the consistency guarantees that $\Pi_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi) \to \varphi$ strongly in $L^p(\Omega)$ and $\nabla_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi) \to \nabla\varphi$ strongly in $L^p(\Omega)^d$, for all $\varphi \in W_0^{1,p}$. Inserting $v := P_{\mathcal{D}_m}\varphi$ into the gradient scheme (4.3.4), we obtain

$$\begin{aligned}
&\int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} u_m \,\mathrm{d}\boldsymbol{x}\\
&\quad\leq \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m) \cdot \nabla_{\mathcal{D}_m} P_{\mathcal{D}_m}\varphi \,\mathrm{d}\boldsymbol{x} + \int_\Omega f \Pi_{\mathcal{D}_m}(u_m - P_{\mathcal{D}_m}\varphi) \,\mathrm{d}\boldsymbol{x}\\
&\quad\quad - \int_\Omega |\nabla_{\mathcal{D}_m} u_m| \,\mathrm{d}\boldsymbol{x} + \int_\Omega |\nabla_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi)| \,\mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{4.3.9}$$

All the terms except the last two can be handled as in Theorem 4.2.9. From the strong convergence of the sequence of $P_{\mathcal{D}_m}\varphi$, letting $m \to \infty$ in the last term implies

$$\lim_{m\to\infty} \int_\Omega |\nabla_{\mathcal{D}_m}(P_{\mathcal{D}_m}\varphi)| \,\mathrm{d}\boldsymbol{x} = \int_\Omega |\nabla\varphi| \,\mathrm{d}\boldsymbol{x}. \tag{4.3.10}$$

Estimating $\liminf_{m\to\infty} \int_\Omega |\nabla_{\mathcal{D}_m} u_m| \,\mathrm{d}\boldsymbol{x}$ is rather standard. For any $\mathbf{w} \in L^\infty(\Omega)^d$ such that $|\mathbf{w}| \leq 1$, write $\int_\Omega \mathbf{w} \cdot \nabla_{\mathcal{D}_m} u_m \,\mathrm{d}\boldsymbol{x} \leq \int_\Omega |\nabla_{\mathcal{D}_m} u_m| \,\mathrm{d}\boldsymbol{x}$. The weak convergence in $L^p(\Omega)^d$ of $\nabla_{\mathcal{D}_m} u_m$ then yields

$$\int_\Omega \mathbf{w} \cdot \nabla\bar{u} \,\mathrm{d}\boldsymbol{x} = \lim_{m\to\infty} \int_\Omega \mathbf{w} \cdot \nabla_{\mathcal{D}_m} u_m \,\mathrm{d}\boldsymbol{x} \leq \liminf_{m\to\infty} \int_\Omega |\nabla_{\mathcal{D}_m} u_m| \,\mathrm{d}\boldsymbol{x}.$$

Taking the supremum over $\mathbf{w}$ leads to

$$\int_\Omega |\nabla \bar{u}|\,\mathrm{d}\boldsymbol{x} \le \liminf_{m\to\infty} \int_\Omega |\nabla_{\mathcal{D}_m} u_m|\,\mathrm{d}\boldsymbol{x}.$$

From this estimation and (4.3.10), passing to the superior limit in Inequality (4.3.9) gives

$$\limsup_{m\to\infty} \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m)\cdot\nabla_{\mathcal{D}_m} u_m\,\mathrm{d}\boldsymbol{x} \\ \le \int_\Omega \mathcal{A}\cdot\nabla\varphi\,\mathrm{d}\boldsymbol{x} + \int_\Omega f(\bar{u}-\varphi)\,\mathrm{d}\boldsymbol{x} - \int_\Omega |\nabla\bar{u}|\,\mathrm{d}\boldsymbol{x} + \int_\Omega |\nabla\varphi|\,\mathrm{d}\boldsymbol{x}. \quad (4.3.11)$$

Since this inequality holds for any $\varphi \in W_0^{1,p}(\Omega)$, making $\varphi = \bar{u}$ gives

$$\limsup_{m\to\infty} \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m)\cdot\nabla_{\mathcal{D}_m} u_m\,\mathrm{d}\boldsymbol{x} \le \int_\Omega \mathcal{A}\cdot\nabla\varphi\,\mathrm{d}\boldsymbol{x}. \quad (4.3.12)$$

Exactly as Theorem 4.2.9, it is then shown that $\mathcal{A} = \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u})$ and

$$\int_\Omega \mathbf{a}(\boldsymbol{x}, \bar{u}, \nabla\bar{u})\cdot\nabla\bar{u}\,\mathrm{d}\boldsymbol{x} \le \liminf_{m\to\infty} \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m)\cdot\nabla_{\mathcal{D}_m} u_m\,\mathrm{d}\boldsymbol{x}.$$

Substituting $\mathcal{A}$ and using this relation in (4.3.11) show that $\bar{u}$ is a solution to Problem (4.3.3). The rest of proof follows along the same steps of Theorem 4.2.9. □

4.4 Approximate barriers

Let us now discuss the case of approximate barriers. In most numerical methods, as the $\mathbb{P}1$ finite elements for instance, the standard interpolant of smooth function v is constructed by taking the value of v at interpolation nodes. When v is bounded by the barrier (a for the Signorini problem, ψ for the obstacle problem), this interpolation may not satisfy the barriers conditions at any point on the boundary/in the domain, especially for the case of the nonconstant barriers. It is therefore classical to modify these barriers conditions when discretising the model. This modification can often be written in the following way.

Using $a_{\mathcal{D}} \in L^p(\partial\Omega)$ (for the Signorini problem) or $\psi_{\mathcal{D}} \in L^p(\Omega)$ (for the obstacle problem), which are respectively approximations of a or ψ, we introduce the convex sets as

$$\mathcal{K}_{a_{\mathcal{D}}} := \{v \in \mathcal{I}_{\mathcal{D},\Gamma_1} g + X_{\mathcal{D},\Gamma_{2,3}} \;:\; \mathbb{T}_{\mathcal{D}} v \le a_{\mathcal{D}} \text{ on } \Gamma_3\}$$

or

$$\mathcal{K}_{\psi_{\mathcal{D}}} := \{v \in \mathcal{I}_{\mathcal{D},\Gamma_1} h + X_{\mathcal{D},0} \;:\; \Pi_{\mathcal{D}} v \le \psi_{\mathcal{D}}\}.$$

The schemes (4.2.10) or (4.3.4) are then modified by replacing the set $\mathcal{K}_{\mathcal{D}}$ by $\mathcal{K}_{a_{\mathcal{D}}}$ in the Signorini case, or by $\mathcal{K}_{\psi_{\mathcal{D}}}$ in the obstacle case. The convergence results for this case of approximate barriers are given in the following theorems, whose proofs are identical to that of Theorem 4.2.9 (see [3, Section 6] for the case of approximate barriers in gradient schemes for linear VI).

Theorem 4.4.1. *Under the assumptions of Theorem 4.2.9, let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretisations in the sense of Definition 4.2.3, which is coercive, limit-conforming, compact, and GD-consistent, in which $S_{\mathcal{D}}$ is defined using $\mathcal{K}_{a_{\mathcal{D}}}$ instead of $\mathcal{K}_{\mathcal{D}}$. Assume that each $\mathcal{K}_{a_{\mathcal{D}_m}}$ is nonempty.*

Then, for any $m \in \mathbb{N}$, there exists at least one solution $u_m \in \mathcal{K}_{a_{\mathcal{D}_m}}$ to the gradient scheme (4.2.10) *in which $\mathcal{K}_{\mathcal{D}_m}$ has been replaced with $\mathcal{K}_{a_{\mathcal{D}_m}}$. If $a_{\mathcal{D}_m} \to a$ in $L^p(\partial\Omega)$ as $m \to \infty$, then the convergences of $\Pi_{\mathcal{D}_m} u_m$ and $\nabla_{\mathcal{D}_m} u_m$ stated in Theorem 4.2.9 still holds.*

Theorem 4.4.2. *Under the assumptions of Theorem 4.3.9, let* $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *be a sequence of gradient discretisations in the sense of Definition 4.3.2, which is coercive, limit-conforming, compact, and GD-consistent in which* $S_{\mathcal{D}}$ *is defined using* $\mathcal{K}_{\psi_{\mathcal{D}}}$ *instead of* $\mathcal{K}_{\mathcal{D}}$*. Assume that* $\mathcal{K}_{\psi_{\mathcal{D}_m}}$ *is nonempty for any* m*.*

Then, for any $m\in\mathbb{N}$*, there exists at least one solution* $u_m\in\mathcal{K}_{\psi_{\mathcal{D}_m}}$ *to the gradient scheme* (4.3.4) *in which* $\mathcal{K}_{\mathcal{D}_m}$ *has been replaced with* $\mathcal{K}_{\psi_{\mathcal{D}_m}}$*. Furthermore, if* $\psi_{\mathcal{D}_m}\to\psi$ *in* $L^p(\Omega)$ *as* $m\to\infty$*, then the convergence of* $\Pi_{\mathcal{D}_m}u_m$ *and* $\nabla_{\mathcal{D}_m}u_m$ *given in Theorem 4.3.9 still holds.*

4.5 Application to the hybrid mixed mimetic methods

The gradient discretisation method is used here to design a hybrid mimetic mixed (HMM) scheme for nonlinear variational inequalities. In Chapter 3 we established the HMM method for the linear Signorini and obstacle problems (i.e., $\mathbf{a}(\boldsymbol{x},\bar{u},\nabla\bar{u})=\Lambda(\boldsymbol{x})\nabla\bar{u}$). The only other application, that we are aware of, mimetic method to variational inequalities only concerns linear variational inequalities and the nodal mimetic finite difference method [6]. Our application of an HMM scheme for nonlinear variational inequalities seems to be the first one of a scheme for these models on generic meshes.

The notion of polytopal mesh described in Definition 2.2.9 still valid here to construct the HMM method for the nonlinear variational inequalities.

4.5.1 HMM for the nonlinear Signorini problem

Let $\mathcal{T}$ be a polytopal mesh that is aligned with the boundaries $(\Gamma_i)_{i=1,2,3}$, that is, for any $i=1,2,3$, each boundary edge is either fully included in Γ_i or disjoint from this set. We describe here a gradient discretisation that corresponds, for linear diffusion problems and standard boundary conditions, to the HMM method [40, 37].

Define two discrete spaces as follows:

$$X_{\mathcal{D},\Gamma_{2,3}}=\{v=((v_K)_{K\in\mathcal{M}},(v_\sigma)_{\sigma\in\mathcal{E}})\ :\ v_K\in\mathbb{R},\ v_\sigma\in\mathbb{R},v_\sigma=0\text{ for all }\sigma\in\mathcal{E}_{\rm ext}\text{ such that }\sigma\subset\Gamma_1\}\tag{4.5.1}$$

and

$$\begin{aligned}X_{\mathcal{D},\Gamma_1}=\{v=((v_K)_{K\in\mathcal{M}},(v_\sigma)_{\sigma\in\mathcal{E}})\ :\ &v_K\in\mathbb{R},\ v_\sigma\in\mathbb{R},\\ &v_K=0\text{ for all }K\in\mathcal{M},\ v_\sigma=0\text{ for all }\sigma\in\mathcal{E}_{\rm int}\text{ and}\\ &v_\sigma=0\text{ for all }\sigma\in\mathcal{E}_{\rm ext}\text{ such that }\sigma\subset\Gamma_2\cup\Gamma_3\}.\end{aligned}\tag{4.5.2}$$

The space $X_{\mathcal{D}}$ is the direct sum of these two spaces, and the function reconstruction $\Pi_{\mathcal{D}}$, the piecewise constant trace reconstruction $\mathbb{T}_{\mathcal{D}}$ and the gradient reconstruction $\nabla_{\mathcal{D}}$ are given by (3.3.1), (3.4.2) and (3.3.2).

Although $\mathcal{I}_{\mathcal{D},\Gamma_1}$ is formally defined on the whole space $W^{1-\frac{1}{p},p}(\partial\Omega)$, to define and analyse the gradient scheme (4.2.10) we only need to consider $\mathcal{I}_{\mathcal{D},\Gamma_1}g$ where g is the specific boundary condition (4.2.1b). When this boundary condition is known to be more regular than $W^{1-\frac{1}{p},p}(\partial\Omega)$, the definition of $\mathcal{I}_{\mathcal{D},\Gamma_1}$ can take advantage of this regularity. Given the assumption in Proposition 4.5.1 below, we therefore set

$$\forall g\in C^2(\overline{\Omega})\ :\ \mathcal{I}_{\mathcal{D},\Gamma_1}g=g(\overline{\boldsymbol{x}}_\sigma),\ \text{for all }\sigma\in\mathcal{E}_{\rm ext}\text{ such that }\sigma\subset\Gamma_1.\tag{4.5.3}$$

The convex set is then $\mathcal{K}_{\mathcal{D}}:=\{v\in X_{\mathcal{D},\Gamma_{2,3}}+\mathcal{I}_{\mathcal{D},\Gamma_1}g\ :\ v_\sigma\le a\text{ on }\sigma,\forall\sigma\in\mathcal{E}_{\rm ext}\text{ s.t. }\sigma\subset\Gamma_3\}$ and, translating the gradient scheme (4.2.10), the HMM discretisation of Problem (4.2.6) is identical to the gradient scheme (4.2.10) corresponding to these gradient discretisation.

The convergence of the HMM scheme is a consequence of Theorem 4.2.9 and the four properties proved in the following proposition. The condition (4.2.14) follows from the GD-consistency of HMM for Fourier boundary conditions (see [37, Definition 2.49 and Section 12.2.2]).

Proposition 4.5.1. *Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of HMM GD given by* (4.5.1)*,* (4.5.2)*,* (4.5.3)*,* (3.3.1)*,* (3.4.2) *and* (3.3.2)*, for certain polytopal meshes $(\mathcal{T}_m)_{m\in\mathbb{N}}$. Assume the existence of $\theta > 0$ such that, for any $m \in \mathbb{N}$,*

$$\begin{aligned}\max_{K\in\mathcal{M}_m}\left(\max_{\sigma\in\mathcal{E}_K}\frac{h_K}{d_{K,\sigma}}+\mathrm{Card}(\mathcal{E}_K)\right)+\max_{\sigma\in\mathcal{E}_{m,\mathrm{int}},\mathcal{M}_\sigma=K,L}\left(\frac{d_{K,\sigma}}{d_{L,\sigma}}+\frac{d_{L,\sigma}}{d_{K,\sigma}}\right)\\+\max\left\{\frac{|K|}{h_K|\sigma|}:K\in\mathcal{M}_m,\sigma\in\mathcal{E}_K\right\}\le\theta\end{aligned}\tag{4.5.4}$$

and, for all $K \in \mathcal{M}_m$ and $\mu \in \mathbb{R}^{\mathcal{E}_K}$,

$$\begin{aligned}\frac{1}{\theta}\sum_{\sigma\in\mathcal{E}_K}|D_{K,\sigma}|\Big|\frac{R_{K,\sigma}(\mu)}{d_{K,\sigma}}\Big|^p\le\sum_{\sigma\in\mathcal{E}_K}|D_{K,\sigma}|\Big|\frac{(A_KR_K(\mu))_\sigma}{d_{K,\sigma}}\Big|^p\\\le\theta\sum_{\sigma\in\mathcal{E}_K}|D_{K,\sigma}|\Big|\frac{R_{K,\sigma}(\mu)}{d_{K,\sigma}}\Big|^p.\end{aligned}\tag{4.5.5}$$

Then the sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is coercive, limit-conforming and compact in the sense of Definitions 4.2.5, 4.2.6 and 4.2.7. If moreover $C^2(\overline{\Omega})\cap\mathcal{K}$ is dense in $\mathcal{K}$ and the function a is piecewise constant on $\mathcal{E}_{\mathrm{ext}}$, the sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is consistent.

Note that the assumption that $C^2(\overline{\Omega})\cap\mathcal{K}$ is dense in $\mathcal{K}$ already appears in [56], to ensure the convergence of numerical schemes for variational inequalities with the case of the homogeneous Dirichlet boundary condition and a constant barriers.

Proof. The coercivity, limit-conformity and compactness follow as in the case of the HMM for PDEs, see [37, Theorem 12.12]. We therefore only discuss the consistency property. Due to the density of $C^2(\overline{\Omega})\cap\mathcal{K}$ in $\mathcal{K}$, as in [37, Lemma 2.13] we see that the GD-consistency follows if we prove that $S_{\mathcal{D}_m}(\varphi)\to 0$ for all $\varphi\in C^2(\overline{\Omega})\cap\mathcal{K}$. For such a φ, Let $v_m = ((v_K)_{K\in\mathcal{M}_m},(v_\sigma)_{\sigma\in\mathcal{E}_m})\in X_{\mathcal{D}_m}$ be the interpolant such that, $v_K=\varphi(\overline{\boldsymbol{x}}_K)$ for all $K\in\mathcal{M}$ and $v_\sigma=\varphi(\overline{\boldsymbol{x}}_\sigma)$ for all $\sigma\in\mathcal{E}$. From the definition (4.5.3) of $\mathcal{I}_{\mathcal{D},\Gamma_1}g$, we see that $v_m-\mathcal{I}_{\mathcal{D},\Gamma_1}g\in X_{\mathcal{D}_m,\Gamma_{2,3}}$. Moreover, since a is piecewise constant on $\mathcal{E}_{\mathrm{ext}}$, we clearly have $\mathbb{T}_{\mathcal{D}_m}v_m\le a$ on Γ_3 since $\varphi\in\mathcal{K}$. Hence, the interpolant v_m belongs to $\mathcal{K}_{\mathcal{D}_m}$. From the proof of [37, Theorem 12.12 and Proposition 7.36],

$$\|\Pi_{\mathcal{D}_m}v_m-\varphi\|_{L^p(\Omega)}+\|\nabla_{\mathcal{D}_m}v_m-\nabla\varphi\|_{L^p(\Omega)^d}\le Ch_{\mathcal{M}_m}$$

with C not depending on m. This shows that that $\lim_{m\to\infty}S_{\mathcal{D}_m}(\varphi)=0$ and concludes the proof. □

4.5.2 HMM methods for the nonlinear obstacle problem and Bulkley model

We use the notations introduced in Section 4.5.1 to reconstruct the HMM schemes corresponding to the two gradient schemes problems defined in Section 4.3.2. The elements of gradient discretisation $\mathcal{D}$ to consider here are given by

$$X_{\mathcal{D},0}=\{v=((v_K)_{K\in\mathcal{M}},(v_\sigma)_{\sigma\in\mathcal{E}})\ :\ v_K\in\mathbb{R},v_\sigma\in\mathbb{R},v_\sigma=0\text{ for all }\sigma\in\mathcal{E}_{\mathrm{ext}}\text{ such that }\sigma\subset\partial\Omega\},$$

$$X_{\mathcal{D},\partial\Omega}=\{v=((v_K)_{K\in\mathcal{M}},(v_\sigma)_{\sigma\in\mathcal{E}})\ :\ v_K\in\mathbb{R},v_\sigma\in\mathbb{R},v_K=0\text{ for all }K\in\mathcal{M},\ v_\sigma=0\text{ for all }\sigma\in\mathcal{E}_{\mathrm{int}}\}.$$

The discrete mappings $\mathcal{I}_{\mathcal{D},\partial\Omega}$, $\Pi_{\mathcal{D}}$ and $\nabla_{\mathcal{D}}$ are as in Section 4.5.1.

Setting $\mathcal{K}_{\mathcal{D}}:=\{v\in X_{\mathcal{D},0}+\mathcal{I}_{\mathcal{D},\partial\Omega}h\ :\ v_K\le\psi_K$ on K, for all $K\in\mathcal{M}\}$, the HMM methods for (4.3.2) and (4.3.3) are respectively the gradient schemes (4.3.4) and (4.3.5) coming from the above gradient discretisation.

Recall that the coercivity, limit-conformity and compactness for sequences of GD adapted to the obstacle problem are the same properties as for sequences of GD for PDEs with nonhomogeneous Dirichlet boundary conditions. The proof of these properties follow therefore from [37], under the regularity assumptions (4.5.4) and (4.5.5). The GD-consistency follows, if the barrier ψ is a piecewise constant and $C^2(\overline{\Omega}) \cap \mathcal{K}$ is dense in $\mathcal{K}$, as in Proposition 4.5.1.

For the Bulkley model, all the properties of GD are identical to those for PDEs with homogeneous Dirichlet boundary conditions, and therefore follow (still under the assumptions (4.5.4) and (4.5.5)) from [40].

Using these properties, the convergence of the HMM method for each of the problems is a straightforward consequence of Theorems 4.3.9 and 4.3.10.

Remark 4.5.1. If the barriers a and ψ are non piecewise constants, for the HMM method they would normally be approximated by piecewise constants. The convex sets $\mathcal{K}_{\mathcal{D}}$ would be modified as described in Section 4.4, and Theorems 4.4.1 and 4.4.2 would insure the convergence of the HMM schemes.

4.6 Numerical results

We demonstrate here the efficiency of the HMM method for solving nonlinear Signorini problems by considering the meaningful example of the seepage model. Due to the double nonlinearity in the model, two iterative algorithms are used in conjunction to compute a numerical solution: fixed point iterations to deal with the nonlinear operator, and a monotonicity algorithm for the inequalities coming from the imposed Signorini boundary conditions.

We consider a test case from [96]. The geometry of the domain Ω representing the dam is illustrated in Fig 4.1. Letting $\boldsymbol{x} = (x, y)$, the model reads

$$
\begin{aligned}
-\operatorname{div}(\Lambda(\boldsymbol{x}, \bar{u})\nabla \bar{u}) = 0 &\quad \text{in } \Omega,\\
\bar{u} = g &\quad \text{on } \Gamma_1,\\
\Lambda(\boldsymbol{x}, \bar{u})\nabla \bar{u} \cdot \mathbf{n} = 0 &\quad \text{on } \Gamma_2,\\
\left.\begin{aligned}
\bar{u} \le y\\
\Lambda(\boldsymbol{x}, \bar{u})\nabla \bar{u} \cdot \mathbf{n} \le 0\\
\Lambda(\boldsymbol{x}, \bar{u})\nabla \bar{u} \cdot \mathbf{n}(y - \bar{u}) = 0
\end{aligned}\right\} &\quad \text{on } \Gamma_3,
\end{aligned}
$$

with

$$
\begin{aligned}
&\Gamma_1 = \{(x, y) \in \mathbb{R}^2 : \ x = 0 \text{ and } y \in [0, 5]\} \cup \{(x, y) \in \mathbb{R}^2 \ \ x + y = 7 \text{ and } y \in [0, 1]\},\\
&\Gamma_2 = \{(x, y) \in \mathbb{R}^2 : \ y = 0\},\\
&\Gamma_3 = \{(x, y) \in \mathbb{R}^2 : \ y = 5 \text{ and } x \in [0, 2]\} \cup \{(x, y) \in \mathbb{R}^2 \ \ x + y = 7 \text{ and } y \in (1, 5]\}.
\end{aligned}
$$

The boundary condition g is defined by $g(0, y) = 5$ for all $y \in [0, 5]$ and $g(x, y) = 1$, for all $x \in (0, 7)$. We set $\mathbf{a}(\boldsymbol{x}, s, \xi) = \mathcal{H}_{\varepsilon}^{\lambda}(s - y)\xi$ in which the regularised Heaviside function $\mathcal{H}_{\varepsilon}^{\lambda}$ is given by (1.1.2). Here, both λ and ε are taken as 10^{-3}.

As stated above, to obtain the solution to this problem, we first apply a simple fixed point iterations (Algorithm 3), whose idea is to generate a sequence $(u^{(n)})_{n\in\mathbb{N}} \in \mathcal{K}_{\mathcal{D}}$ by solving linear VI problems, in which the nonlinearity in the operator has been fixed to the previous element in the sequence.

At any iteration n in Algorithm 3, we need to solve a linear VI. To compute its solution, introduce the fluxes $(F^w_{K,\sigma}(u))_{K\in\mathcal{M},\sigma\in\mathcal{E}_K}$ defined by: for all $K \in \mathcal{M}$ and all $u, v, w \in X_{\mathcal{D}}$,

$$
\sum_{\sigma\in\mathcal{E}_K} |\sigma| F^w_{K,\sigma}(u)(v_K - v_\sigma) = \int_K \mathcal{H}_{\varepsilon}^{\lambda}(w_K)\nabla_{\mathcal{D}} u \cdot \nabla_{\mathcal{D}} v \,\mathrm{d}\boldsymbol{x}.
$$

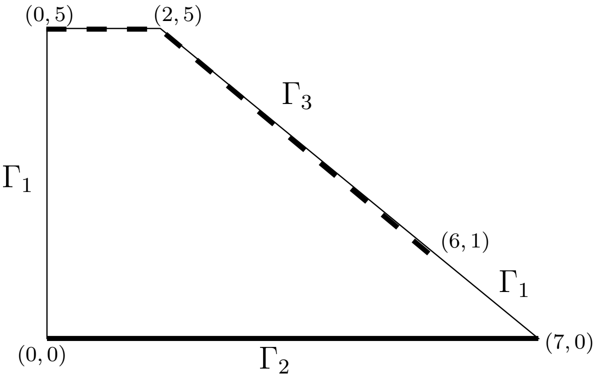

Figure 4.1: Geometry for the numerical tests.

Algorithm 3 Fixed point algorithm

Let δ be a small number (stopping criteria) and $u^{(0)} = 0$ ▷ For us, $\delta = 10^{-3}$.
for $n = 1, 2, 3, \ldots$ **do**
    Solve the following linear problem ▷ $u^{(n)}$ is known

$$\left\{\begin{array}{l} \text{Find } u^{(n+1)} \in \mathcal{K}_{\mathcal{D}} \text{ such that, for all } v \in \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\sum_{K\in\mathcal{M}} \int_K \Lambda(\boldsymbol{x}, u^{(n)}) \nabla_{\mathcal{D}} u^{(n+1)} \cdot \nabla_{\mathcal{D}}(u^{(n+1)} - v)\,\mathrm{d}\boldsymbol{x} \le \sum_{K\in\mathcal{M}} (u_K^{(n+1)} - v_K) \int_K f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \end{array}\right. \tag{4.6.1}$$

    if $\|u^{(n+1)} - u^{(n)}\|_{L^2(\Omega)} \le \delta \|u^{(n)}\|_{L^2(\Omega)}$ **then**
        Exit "for" loop
    end if
end for
Set $u = u^{(n+1)}$

Choosing $w = u^{(n)}$ in this relation, Problem (4.6.1) can be recast as [3]

$$\sum_{\sigma\in\mathcal{E}_K} |\sigma| F_{K,\sigma}(u^{(n+1)}) = m(K) f_K, \quad \forall K \in \mathcal{M} \tag{4.6.2}$$

$$F_{K,\sigma}(u^{(n+1)}) + F_{L,\sigma}(u^{(n+1)}) = 0, \quad \forall \sigma \in \mathcal{E}_{\text{int}} \text{ with } \mathcal{M}_\sigma = \{K, L\}, \tag{4.6.3}$$

$$u_\sigma^{(n+1)} = g, \quad \forall \sigma \in \mathcal{E}_{\text{ext}} \text{ such that } \sigma \subset \Gamma_1, \tag{4.6.4}$$

$$F_{K,\sigma}(u^{(n+1)}) = 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \subset \Gamma_2, \tag{4.6.5}$$

$$F_{K,\sigma}(u^{(n+1)})(u_\sigma^{(n+1)} - \overline{y}_\sigma) = 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \subset \Gamma_3, \tag{4.6.6}$$

$$-F_{K,\sigma}(u^{(n+1)}) \le 0, \quad \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K \text{ such that } \sigma \subset \Gamma_3, \tag{4.6.7}$$

$$u_\sigma^{(n+1)} \le \overline{y}_\sigma, \quad \forall \sigma \in \mathcal{E}_{\text{ext}} \text{ such that } \sigma \subset \Gamma_3. \tag{4.6.8}$$

Here $\overline{y}_\sigma$ denotes to y-coordinate of the centre of gravity of edge σ. The monotonicity algorithm performed in Chapter 3 is used to solve this nonlinear system at each iteration n (see Algorithm 4 for completeness). This algorithm only requires, at each of its steps, to solve a square linear system on the unknowns $(u_K)_{K\in\mathcal{M}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$.

Algorithm 4 Monotonicity algorithm

```
1: Set 𝔸^(0) = {σ ∈ ℰ : σ ⊂ Γ_3} and 𝔹 = ∅
2: Set I = Card(𝔸^(0))                          ▷ Theoretical bound on the iterations
3: while i ≤ I do
4:     𝔸^(i) and 𝔹^(i) being known, find the solution u^(n+1) to the system
```

$$
\begin{aligned}
\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u^{(i)}) &= m(K)f_K, && \forall K\in\mathcal{M},\\
F_{K,\sigma}(u^{(i)}) + F_{L,\sigma}(u^{(i)}) &= 0, && \forall\sigma\in\mathcal{E}_K\cap\mathcal{E}_L, K\neq L\\
u_\sigma^{(i)} &= 0, && \forall\sigma\in\mathcal{E}_{\mathrm{ext}} \text{ such that } \sigma\subset\Gamma_1,\\
F_{K,\sigma}(u^{(i)}) &= 0, && \forall K\in\mathcal{M}, \forall\sigma\in\mathcal{E}_K \text{ such that } \sigma\subset\Gamma_2,\\
F_{K,\sigma}(u^{i}) &= 0, && \forall K\in\mathcal{M}, \forall\sigma\in\mathcal{E}_K \text{ such that } \sigma\in\mathbb{A}^{(i)},\\
u_\sigma^{(i)} &= \overline{y}_\sigma, && \forall\sigma\in\mathcal{E}_{\mathrm{ext}} \text{ such that } \sigma\in\mathbb{B}^{(i)}.
\end{aligned}
\tag{4.6.9}
$$

```
5:     Set 𝔸^(i+1) = {σ ∈ 𝔸^(i) : u_σ^(i) ≤ a_σ} ∪ {σ ∈ 𝔹^(i) : −F_{K,σ}(u^(i)) ≥ 0}
6:     Set 𝔹^(i+1) = {σ ∈ 𝔹^(i) : −F_{K,σ}(u^(i)) < 0} ∪ {σ ∈ 𝔸^(i) : u_σ^(i) > a_σ}
7:     if 𝔸^(i+1) = 𝔸^(i) and 𝔹^(i+1) = 𝔹^(i) then
8:         Exit "while" loop
9:     end if
10: end while
11: Set 𝔸 = 𝔸^(i+1) and 𝔹 = 𝔹^(i+1)
```

Our numerical tests are conducted on two different mesh types given in Figure 4.2. The first type (left) is build on hexagonal cells with maximum size $h_{\mathcal{M}} = 0.69$ and a number of edges including in Γ_3 equal to $N = 72$. For the second type (right) of mesh inspired by the "Kershaw mesh" in [60], the cells still have the same maximum size as the first one and $N = 92$. The fixed point algorithm (Algorithm 3) converges respectively in 5 and 6 iterations, and the maximum number of iterations of the monotonicity algorithm (Algorithm 4) is also still far from the bound, with 6 for the first mesh and 5 for the second mesh.

The monotonicity algorithm offers a useful way to determine the location of the seepage point. Following the interpretation of the model in [96], the seepage point should split the free boundary Γ_3 into upper and lower parts in the following way: (1) there is no flow on the upper part (so $F_{K,\sigma} = 0$ for every edge σ in this part); (2) the pore pressure vanishes on the lower part (so $\bar{u} = y$ on this part); (3) both conditions are met at the seepage point. Note that the first and second conditions are naturally expressed by the last two equations in (4.6.9). Since any edge in the set $\mathbb{B}$ cannot satisfy the last property (due to the strict inequality $u_\sigma < y_\sigma$), the seepage point does not lie on those edges. This point can thus be located at the edge σ in the set $\mathbb{A}$ whose midpoint has the largest ordinate $\overline{y}_\sigma$. Considering the mesh size and the fact that the HMM solution is computed at the mid-point of edges, our numerical results point out the seepage occurs at point with ordinate $\overline{y}_\sigma \in [3.31, 3.65]$ for the hexahedral mesh, and $\overline{y}_\sigma \in [3.28, 3.63]$ for the Kershaw mesh. This location is in perfect agreement with the numerical tests in [96].

Testing the scheme on a mesh with a large number of cells, the seepage location does not change much. We observe that the seepage position moves up only by 1% for a 1681-cell refinement of the hexahedral mesh, and

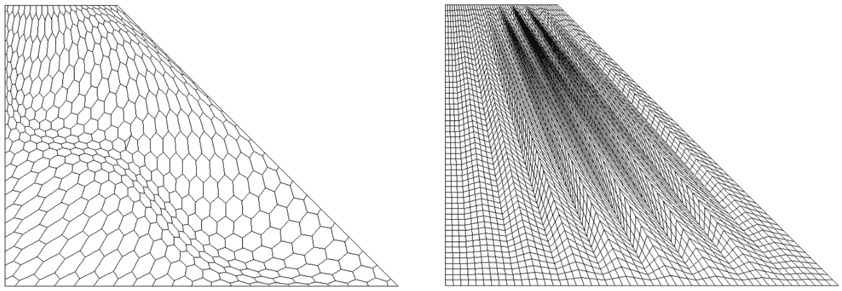

Figure 4.2: The first mesh type (hexahedral, left) and the second mesh type (Kershaw, right).

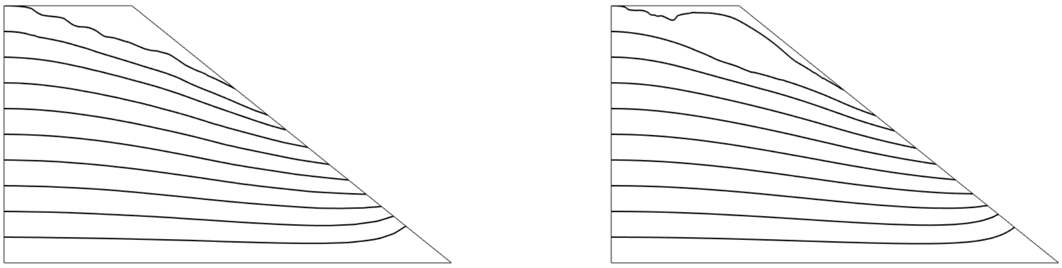

Figure 4.3: Isolines obtained on the hexahedral mesh (left) and on the Kershaw mesh (right).

by 2% for a 4642-cell refinement of the Kershaw mesh.

Figure 4.3 provides the isolines presentation of the Darcy's velocity field of the solution which are generated by MATLAB via ISOLINES function. As expected, the distorted cells at the top of the domain provoke perturbations of the isolines there, but quite remarkably do not impact the location of the seepage point. Otherwise, the isolines are very similar to the ones in [96].

Appendix

4.A Basic results on nonlinear operator

Let us begin by recalling the concept of calculus of variations as defined in [74, Chapter 2].

Definition 4.A.1 (hemi-continuous mapping). Let V be a reflexive vector space and $(\cdot,\cdot)$ be the duality product between V and its dual space V'. The mapping $T : V \to V'$ is said to be *hemi-continuous* if $\lambda \in \mathbb{R} \mapsto (T(u+\lambda v), w) \in \mathbb{R}$ is continuous for all $u, v, w \in V$.

Definition 4.A.2 (Calculus of variations). Let V be a reflexive vector space and $(\cdot,\cdot)$ be the duality product between V and its dual space V'. The mapping $T : V \to V'$ is of the "*calculus of variations*" if $T(u) = B(u,u)$ where $B : V \times V \to V'$ satisfies:

(H1) $\forall u$, $v \to B(u,v)$ is *hemi-continuous bounded* $V \to V'$ and

$$(B(u,u) - B(u,v), u - v) \geq 0.$$

(H2) If $(w_m)_{m\in\mathbb{N}} \to u$ weakly in V and if $(B(w_m,w_m) - B(w_m,u), w_m - u) \to 0$, as $m \to \infty$ then, for all v, $B(w_m,v) \to B(u,v)$ weakly in V', as $m \to \infty$.

(H3) If $w_n \to u$ weakly in V, as $m \to \infty$ and if $B(w_m,u) \to F$ weakly in V', as $m \to \infty$, then $(B(w_m,u), w_m) \to (F,u)$, as $m \to \infty$.

The following lemma taken from [43] enables us to pass to the limit in the nonlinear form $\mathbf{a}(\boldsymbol{x}, u_m, \nabla\phi)$.

Lemma 4.A.3. *Let Ω be a bounded subset of $\mathbb{R}^N$. Suppose $G : \Omega \times \mathbb{R} \to \mathbb{R}^N$ is a Carathéodory function such that there exists positive constants C_1, γ such that, for a.e. $\boldsymbol{x} \in \Omega$,*

$$|G(\boldsymbol{x},\xi)| \leq C_1(1 + |\xi|^\gamma),\ \forall \xi \in \mathbb{R}, \forall n \in \mathbb{N}. \tag{4.A.1}$$

If $p \in [\gamma,\infty]$ and $(u_m)_{m\in\mathbb{N}} \subset L^p(\Omega)$ satisfies that $u_m \to u$ in $L^p(\Omega)$ as $m \to \infty$, then

$$G(\boldsymbol{x}, u_m) \to G(\boldsymbol{x}, u) \ \text{in}\ L^{p/\gamma}(\Omega) \ \text{as}\ m \to \infty.$$

Proof. For $v \in L^p(\Omega)$, then $G(\cdot,v) \in L^{p/\gamma}(\Omega)$ and condition (4.A.1) asserts that there exists a constant C_1 such that

$$\|G(\cdot,v)\|_{L^{p/\gamma}(\Omega)} \leq C_1(\text{meas}(\Omega)^{\gamma/p} + \|v\|^\gamma_{L^p(\Omega)}). \tag{4.A.2}$$

Let $u_m \to u$ in $L^p(\Omega)$ and we assume that $G(\cdot,u_m)$ does not converges to $G(\cdot,u)$ in $L^{p/\gamma}(\Omega)$. This means that there exists $\varepsilon > 0$ such that

$$\|G_m(\cdot,u_m) - G(\cdot,u)\|_{L^{p/\gamma}(\Omega)} > \varepsilon. \tag{4.A.3}$$

Since $u_m \to u$ in $L^p(\Omega)$ as $m \to \infty$, the partial converse dominated convergence theorem shows that there exists a subsequence of $(u_m)_{m\in\mathbb{N}}$ and a function $v \in L^p(\Omega)$ such that $u_m(\boldsymbol{x}) \to u(\boldsymbol{x})$ for a.e. $\boldsymbol{x} \in \Omega$, and $|u_m(\boldsymbol{x})| \leq v(\boldsymbol{x})$ for a.e. $\boldsymbol{x} \in \Omega$. Using the fact that G is a Carathéodory function, we obtain, $G(\boldsymbol{x}, u_m(x)) \to G(\boldsymbol{x}, u(x))$ for a.e. $\boldsymbol{x} \in \Omega$. In addition, for a.e. $\boldsymbol{x} \in \Omega$,

$$|G(\boldsymbol{x}, u_m(x))| \leq C_1(1 + (v(\boldsymbol{x}))^\gamma),$$

therefore, by the dominated convergence theorem, we have

$$G(\cdot, u_m) \to G(\cdot, u) \quad \text{in } L^{p/\gamma}(\Omega),$$

which is a contradiction to (4.A.3). $\square$

Proposition 4.A.4. *Let $X_{\mathcal{D},\Gamma_{2,3}}$ be the finite dimensional space described in Definition 4.2.3 and $\langle \cdot, \cdot \rangle$ be the duality product between $X_{\mathcal{D},\Gamma_{2,3}}$ and its dual $X'_{\mathcal{D},\Gamma_{2,3}}$. Let the operator $\mathcal{A}_{\mathcal{D}} : X_{\mathcal{D},\Gamma_{2,3}} \to X'_{\mathcal{D},\Gamma_{2,3}}$ be defined by, for $\widetilde{u}, \widetilde{w} \in X_{\mathcal{D},\Gamma_{2,3}}$,*

$$\langle \mathcal{A}_{\mathcal{D}}(\widetilde{u}), \widetilde{w} \rangle = \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}}(\widetilde{u} + g_{\mathcal{D}}), \nabla_{\mathcal{D}}(\widetilde{u} + g_{\mathcal{D}})) \cdot \nabla_{\mathcal{D}}(\widetilde{w} + g_{\mathcal{D}}) \,\mathrm{d}\boldsymbol{x},$$

where $g_{\mathcal{D}}$ is defined by (4.2.16) and $\mathbf{a}$ is the Leray-Lions operator satisfying the assumptions (4.2.2)–(4.2.5). Then the mapping $\mathcal{A}_{\mathcal{D}}$ is of the "calculus of variations".

Proof. For $\widetilde{u}, \widetilde{v} \in X_{\mathcal{D},\Gamma_{2,3}}$, let us introduce $B(\widetilde{u}, \widetilde{v}) \in X'_{\mathcal{D},\Gamma_{2,3}}$ by, for all $\widetilde{w} \in X_{\mathcal{D},\Gamma_{2,3}}$,

$$\langle B(\widetilde{u}, \widetilde{v}), \widetilde{w} \rangle = \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}}(\widetilde{u} + g_{\mathcal{D}}), \nabla_{\mathcal{D}}(\widetilde{v} + g_{\mathcal{D}})) \cdot \nabla_{\mathcal{D}}(\widetilde{w} + g_{\mathcal{D}}) \,\mathrm{d}\boldsymbol{x}.$$

Then, we see that, for $\widetilde{u} \in X_{\mathcal{D},\Gamma_{2,3}}$, $B(\widetilde{u}, \widetilde{u}) = \mathcal{A}(\widetilde{u})$.

Let $(\widetilde{u}_m)_{m\in\mathbb{N}} \in X_{\mathcal{D},\Gamma_{2,3}}$ such that $\widetilde{u}_m \to \widetilde{u}$ weakly in $X_{\mathcal{D},\Gamma_{2,3}}$ and let us prove that $B(\widetilde{u}_m, \widetilde{v}) \to B(\widetilde{u}, \widetilde{v})$ weakly in $X'_{\mathcal{D},\Gamma_{2,3}}$. Since weak convergence implies to strong convergence in a finite dimensional space, $\widetilde{u}_m \to \widetilde{u}$ in $X_{\mathcal{D},\Gamma_{2,3}}$, which, in turn, implies that $\Pi_{\mathcal{D}}\widetilde{u}_m \to \Pi_{\mathcal{D}}\widetilde{u}$ in $L^p(\Omega)$. Lemma 4.A.3 shows that $\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}}(\widetilde{u}_m + g_{\mathcal{D}}), \nabla_{\mathcal{D}}(\widetilde{v} + g_{\mathcal{D}})) \to \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}}(\widetilde{u} + g_{\mathcal{D}}), \nabla_{\mathcal{D}}(\widetilde{v} + g_{\mathcal{D}}))$ in $L^{p'}(\Omega)$. Hence, $B(\widetilde{u}_m, \widetilde{v}) \to B(\widetilde{u}, \widetilde{v})$ in $X'_{\mathcal{D},\Gamma_{2,3}}$, that verifies **(H2)** of Definition 4.A.2.

From the above argument (with $\widetilde{u}_m \to \widetilde{u}$ in $X_{\mathcal{D},\Gamma_{2,3}}$), we conclude that B is hemi-continous. By the monotonicity assumption (4.2.5), we know that

$$\langle B(\widetilde{u}, \widetilde{u}) - B(\widetilde{u}, \widetilde{v}), \widetilde{u} - \widetilde{v} \rangle \geq 0, \quad \text{for all } \widetilde{u}, \widetilde{v} \in X_{\mathcal{D},\Gamma_{2,3}}.$$

To check **(H1)** of Definition 4.A.2, we only need to prove that B is bounded. By the growth assumption (4.2.3), we get

$$\begin{aligned}
\Big| \langle B(\widetilde{u}, \widetilde{v}), \widetilde{w} \rangle \Big| &= \Big| \int_\Omega \mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}}(\widetilde{u} + g_{\mathcal{D}}), \nabla_{\mathcal{D}}(\widetilde{v} + g_{\mathcal{D}})) \cdot \nabla_{\mathcal{D}}\widetilde{w} \,\mathrm{d}\boldsymbol{x} \Big| \\
&\leq \left(\int_\Omega |\mathbf{a}(\boldsymbol{x}, \Pi_{\mathcal{D}}(\widetilde{u} + g_{\mathcal{D}}), \nabla_{\mathcal{D}}(\widetilde{v} + g_{\mathcal{D}}))|^{p'} \,\mathrm{d}\boldsymbol{x} \right)^{1/p'} \left(\int_\Omega |\nabla_{\mathcal{D}}\widetilde{w}|^p \,\mathrm{d}\boldsymbol{x} \right)^{1/p} \\
&\leq \int_\Omega \left(\overline{a} + \mu |\nabla_{\mathcal{D}}(\widetilde{v} + g_{\mathcal{D}})|^{p-1} \right) \,\mathrm{d}\boldsymbol{x} \; \|\nabla_{\mathcal{D}}\widetilde{w}\|_{L^p(\Omega)^d} \\
&\leq \left(C_2 \|\nabla_{\mathcal{D}}\widetilde{v}\|^{p-1}_{L^p(\Omega)^d} + C_3 \right) \|\nabla_{\mathcal{D}}\widetilde{w}\|_{L^p(\Omega)^d},
\end{aligned}$$

where $C_2 = \mu$ and $C_3 = \|\overline{a}\|_{L^{p'}(\Omega)} + \mu \|\nabla_{\mathcal{D}} g_{\mathcal{D}}\|^{p-1}_{L^p(\Omega)^d}$. From the above inequality, we deduce

$$\|B(\widetilde{u}, \widetilde{v})\|_{X'_{\mathcal{D},\Gamma_{2,3}}} \leq C_2 \|\nabla_{\mathcal{D}}\widetilde{v}\|^{p-1}_{L^p(\Omega)^d} + C_3,$$

and then, since $\|\nabla \cdot\|_{L^p(\Omega)^d}$ is a norm on $X_{\mathcal{D},\Gamma_{2,3}}$, this proves that B is bounded.

Finally, let $(\widetilde{u}_m)_{m\in\mathbb{N}} \in X_{\mathcal{D},\Gamma_{2,3}}$ such that $\widetilde{u}_m \to \widetilde{u}$ weakly in $X_{\mathcal{D},\Gamma_{2,3}}$ and $B(\widetilde{u}_m, \widetilde{v}) \to F$ weakly in $X'_{\mathcal{D},\Gamma_{2,3}}$. It is clear to see that passing to the limit in $\langle B(\widetilde{u}_m, \widetilde{v}), \widetilde{u}_m \rangle$ gives $\langle F, \widetilde{u} \rangle$, which verifies **(H3)** of Definition 4.A.2. $\square$

Chapter 5

Linear parabolic variational inequalities

Abstract. In this chapter, we adapt the gradient discretisation method to two linear parabolic variational inequalities (PVIs), the time-dependent Signorini and obstacle problems. We prove bounds on the approximate solutions in discrete norms. Together with the classical set of properties (coercivity, space–time consistency and limit-conformity), these estimates enable us to obtain the convergence of the gradient schemes to the weak solutions of the parabolic variational inequalities. We introduce a numerical scheme based on the HMM discretisation and provide numerical experiments.

5.1 Introduction

The main objective of this chapter is the convergence analysis of numerical methods for linear parabolic variational inequalities modelling the diffusion in porous media. Semipermeable membrane including osmosis phenomenon, and problems concerning the control of temperature at thermal boundaries are standard applications of the PVIs in porous media and physic [76].

In what follows, let $[0,T] \subset \mathbb{R}$ and $\Omega \subset \mathbb{R}^d$ ($d = 1,2,3$) be a bounded connected open set. One model considered here involves homogeneous Dirichlet and Signorini boundary conditions, each one set on a different part of the boundary. The Signorini BCs are of the form

$$\bar{u} \leq a,\ \nabla\bar{u}\cdot\mathbf{n} \leq 0,\ \text{ and } (a-\bar{u})\nabla\bar{u}\cdot\mathbf{n} = 0 \quad \text{on } \Gamma_2 \times (0,T). \tag{5.1.1}$$

Here $\mathbf{n}$ still denotes the unit outer normal to $\partial\Omega$, which is split into two parts, Γ_1 where the Dirichlet boundary condition is imposed and Γ_2. The Signorini BCs (5.1.1) typically express the flow of fluid among a domain with a semi-permeable membrane boundary [76], and unsaturated flow in porous media [64].

We also consider a linear model appearing in financial mathematics, a parabolic obstacle problem where unilateral conditions are imposed inside a domain,

$$\partial_t\bar{u} - \operatorname{div}(\Lambda\nabla\bar{u}) \geq f,\ \bar{u} \geq \psi,\ \text{ and } (\partial_t\bar{u} - \operatorname{div}(\Lambda\nabla\bar{u}) - f)(\bar{u}-\psi) = 0 \quad \text{in } \Omega\times(0,T).$$

This chapter aims to provide preliminary results on the gradient discretisation method for parabolic VIs. Obtaining a general error estimate will be the topic of future work.

The theoretical results concerning the existence and uniqueness of the solution, as well as its regularity, to the linear PVIs can be found in foundational studies such as [75, 8, 15, 90]. Discretisation of linear PVIs is undertaken in various works. In [22], a discretisation framework, using the backwards Euler and Galerkin methods, is developed for the parabolic Signorini problem.

Regarding $\mathbb{P}1$ finite element methods for parabolic obstacle problem, we refer the reader to [32, 94, 83, 81, 52, 80, 67]. The L^∞–convergence and the error estimate for $\mathbb{P}1$ finite element method, applied on triangular meshes with acute angles, are obtained in [52] under regularity assumptions on the solution ($\partial_t\bar{u}$ and $\Delta\bar{u}$ in $L^\infty(\Omega\times(0,T))$). [80] provides a posteriori error estimate of order $\mathcal{O}(h+\tau)$, where τ is the time step, for linear finite element method for the parabolic obstacle problem, provided that the initial solution $\bar{u}_0$ is smooth.

[13] designs a two-points finite volume scheme for the parabolic obstacle model involving a convection term, and establishes energy estimates on the approximate solution in a discrete $L^2(0,T;H_0^1(\Omega))$ norm, and on its discrete time derivative in the $L^2(\Omega\times(0,T))$ norm. Using such estimates, [13] establishes the strong convergence of the scheme's solution to the exact solution in $L^2(\Omega\times(0,T))$, as well as the weak convergence of the discrete time derivative of the scheme's solution to the time derivative of the continuous solution in $L^2(\Omega\times(0,T))$. [9] studies a posteriori and a priori error estimates for a discontinuous Galerkin method for parabolic obstacle problems.

In this chapter, we adapt the gradient discretisation method to the linear parabolic Signorini and obstacle problems. This adaptation provides a unified convergence analysis of numerical methods for these problems. It also enables us to obtain convergence theorems of numerical schemes for PVIs. As an application of the gradient discretisation framework, we extend the HMM method to PVIs.

The outline of this chapter is as follows. Section 5.2 states the continuous parabolic Signorini model and its weak formulation, and studies the gradient discretisation method for this model. Section 5.3 is devoted to the parabolic obstacle model. In Section 5.4 we prove some estimates on the approximate solutions for both models and prove the convergence of the gradient discretisation method. In Section 5.5, we introduce the HMM method to the parabolic Signorini and obstacle problems. Finally, numerical tests are given in Section 5.6 to evaluate the behaviour of the method.

5.2 Parabolic Signorini problem

5.2.1 Notations

Let us start with recalling basic notations of vector spaces to be used throughout the current chapter. If V be a Banach space equipped with the norm $||\cdot||_V$, we denote by $L^p(0,T;V)$ the space of Lebesgue measurable functions $w:\Omega\times[0,T]\to V$ such that $||w||_{L^p(0,T;V)}<\infty$, where

$$||w||_{L^p(0,T;V)} := \left(\int_0^T ||w||_V^p \,\mathrm{d}t\right)^{1/p}, \quad \text{if } p\neq\infty,$$
$$||w||_{L^\infty(0,T;V)} := \operatorname{esssup}_{t\in(0,T)} ||w||_V.$$

The spaces $L^p(0,T;V)$ are Banach spaces w.r.t. the above norms. If V is endowed with a Hilbert product $\langle\cdot,\cdot\rangle_V$, the space $L^2(0,T;V)$ is also a Hilbert space w.r.t the inner product

$$\langle f,g\rangle_{L^2(0,T;V)} = \int_0^T \langle f,g\rangle_V \,\mathrm{d}t.$$

In what follows, we use these notations with $V=H^1_{\Gamma_1}(\Omega)$ for the Signorini problem, and $V=H^1_0(\Omega)$ for the obstacle problem. Here $H^1_{\Gamma_1}(\Omega)$ denotes to the space of all functions in $H^1(\Omega)$ with zero values on the boundary Γ_1. Let $v(t)$ denote the function $\boldsymbol{x}\mapsto v(\boldsymbol{x},t)$.

5.2.2 Continuous model and weak formulation

The first model to be considered is

$$\partial_t\bar{u} - \operatorname{div}(\Lambda\nabla\bar{u}) = f \quad \text{in } \Omega\times(0,T), \tag{5.2.1a}$$
$$\bar{u} = 0 \quad \text{on } \Gamma_1\times(0,T), \tag{5.2.1b}$$
$$\left.\begin{aligned} \bar{u} &\le a \\ \Lambda\nabla\bar{u}\cdot\mathbf{n} &\le 0 \\ (a-\bar{u})\Lambda\nabla\bar{u}\cdot\mathbf{n} &= 0 \end{aligned}\right\} \quad \text{on } \Gamma_2\times(0,T), \tag{5.2.1c}$$
$$\bar{u}(\boldsymbol{x},0) = \bar{u}_{\text{ini}} \quad \text{in } \Omega\times\{0\}. \tag{5.2.1d}$$

Assumptions 5.2.1. *The data in* (5.2.1) *are assumed to satisfy the following:*

1. Ω *is an open bounded connected subset of* $\mathbb{R}^d$ $(d=1,2,3)$ *with a Lipschitz boundary, and* $T>0$,
2. Λ *is a measurable function from* Ω *to* $\mathrm{M}_d(\mathbb{R})$ *(where* $\mathrm{M}_d(\mathbb{R})$ *is the set of* $d\times d$ *matrices) and there exists* $\underline{\lambda}$, $\overline{\lambda}>0$ *such that, for a.e.* $\boldsymbol{x}\in\Omega$, $\Lambda(\boldsymbol{x})$ *is symmetric with eigenvalues in* $[\underline{\lambda},\overline{\lambda}]$,
3. $\partial\Omega=\Gamma_1\cup\Gamma_2$, *where* Γ_1 *and* Γ_2 *are measurable pairwise disjoint subsets of* $\partial\Omega$ *such that* Γ_2 *is open and the measure of* Γ_1 *is strictly positive,*
4. $0<a\in\mathbb{R}$ *and* $f\in L^2(\Omega\times(0,T))$,
5. $\bar{u}_{\text{ini}}\in\mathcal{K}:=\{v\in H^1_{\Gamma_1}(\Omega):\ \gamma v\le a$ *on* Γ_2 *a.e.*$\}$.

Based on the set $\mathcal{K}$ given in the above assumptions, we define the following closed convex subset of the space $L^2(0,T;H^1_{\Gamma_1}(\Omega))$:

$$\mathbb{K} = \{v\in L^2(0,T;H^1_{\Gamma_1}(\Omega)):\ v(t)\in\mathcal{K} \text{ for a.e. } t\in(0,T)\}.$$

Let t be an arbitrary point in $(0,T)$ and $\bar{u}$ be a classical solution to the above model (5.2.1). Then $\bar{u}$ belongs to the set $\mathbb{K}$, since it satisfies (5.2.1b) and (5.2.1c). Multiplying Equation (5.2.1a) by $\bar{u}(\boldsymbol{x},t)-v(\boldsymbol{x},t)$, with

$v \in \mathbb{K}$, and following the same idea as in the elliptic model (Appendix 2.A), we deduce that $\bar{u} \in \mathbb{K}$ satisfies the inequality, for all $v \in \mathbb{K}$,

$$\int_\Omega \partial_t \bar{u}(\boldsymbol{x},t)(u(\boldsymbol{x},t) - v(\boldsymbol{x},t))\,\mathrm{d}\boldsymbol{x} + \int_\Omega \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x},t)\cdot\nabla(\bar{u}-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}$$
$$\le \int_\Omega f(\boldsymbol{x},t)(\bar{u}(\boldsymbol{x},t) - v(\boldsymbol{x},t))\,\mathrm{d}\boldsymbol{x}.$$

Under Assumptions 5.2.1, integrating this inequality over the time interval $[0,T]$ generates the weak formulation of Problem (5.2.1),

$$\left\{\begin{array}{l} \text{Find } \bar{u} \in \mathbb{K}\cap C^0([0,T];L^2(\Omega)), \text{ such that } \partial_t\bar{u} \in L^2(0,T;L^2(\Omega)),\ \bar{u}(\boldsymbol{x},0) = \bar{u}_{\mathrm{ini}}, \text{ and} \\ \displaystyle\int_0^T\int_\Omega \partial_t\bar{u}(\boldsymbol{x},t)(u(\boldsymbol{x},t)-v(\boldsymbol{x},t))\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \int_0^T\int_\Omega \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x},t)\cdot\nabla(\bar{u}-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\ \displaystyle\le \int_0^T\int_\Omega f(\boldsymbol{x},t)(\bar{u}(\boldsymbol{x},t)-v(\boldsymbol{x},t))\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \quad \text{for all } v\in\mathbb{K}. \end{array}\right. \tag{5.2.2}$$

In [74, 44] it is shown that there exists a unique weak solution to (5.2.2). Conversely, let $\bar{u} \in \mathbb{K}$ be a solution to (5.2.2) and for $t \in (0,T)$, take

$$v(\boldsymbol{x},t) = \bar{u}(\boldsymbol{x},t) \pm \varphi(\boldsymbol{x},t),$$

where φ is an arbitrary function in $C_0^\infty(\Omega\times(0,T))$. Taking the function v as a test function in (5.2.2) (because it belongs to $\mathbb{K}$) gives

$$\int_0^T\int_\Omega (\partial_t\bar{u}(\boldsymbol{x},t)\varphi(\boldsymbol{x},t) + \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x},t)\cdot\nabla\varphi(\boldsymbol{x},t) - f(\boldsymbol{x},t)\varphi(\boldsymbol{x},t))\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t = 0.$$

By integration by part, we see that $\bar{u}$ satisfies the evolution equation (5.2.1a) in $\Omega\times(0,T)$. The boundary conditions in (5.2.1) can be verified as in the case of the elliptic model in Appendix 2.A.

Remark 5.2.1. Let the time interval $[0,T]$ be divided into ℓ_κ intervals of length κ, where κ tends to zero as $\ell_\kappa \to \infty$. Let $\mathbf{1}_{I_i}$ be the characteristic function of $I_i = [i\kappa,(i+1)\kappa)$, $i = 0,...,\ell_\kappa$. We define a set of piecewise-constant in time functions by

$$\mathbb{L}_\kappa = \left\{ w_\kappa(\boldsymbol{x},t) = \sum_{i=1}^{\ell_\kappa} \mathbf{1}_{I_i}(t)\varphi_i(\boldsymbol{x}) \ :\ \varphi \in C^2(\overline{\Omega}) \text{ and } \varphi \le a \text{ on } \Gamma_2 \text{ a.e.} \right\}. \tag{5.2.3}$$

By the density of the set $C^2(\overline{\Omega})\cap\mathcal{K}$ in $\mathcal{K}$ established in [56], every $v \in \mathbb{K}$ can be approximated by a piecewise constant function in time $w_\kappa \in \mathbb{L}_\kappa$ such that w_κ converges strongly to v in $L^2(0,T;H^1(\Omega))$ as $\kappa\to 0$ (note that $\gamma(w_\kappa) \le a$ on $\Gamma_2\times(0,T)$). Hence, Problem (5.2.2) is equivalent to

$$\left\{\begin{array}{l} \text{Find } \bar{u} \in \mathbb{K}\cap C^0([0,T];L^2(\Omega)), \text{ such that } \partial_t\bar{u} \in L^2(0,T;L^2(\Omega)),\ \bar{u}(x,0) = \bar{u}_{\mathrm{ini}}, \text{ and} \\ \displaystyle\int_0^T\int_\Omega \partial_t\bar{u}(\boldsymbol{x},t)(u-w_\kappa)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \int_0^T\int_\Omega \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x},t)\cdot\nabla(\bar{u}-w_\kappa)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\ \displaystyle\le \int_0^T\int_\Omega f(\boldsymbol{x},t)(\bar{u}-w_\kappa)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \quad \text{for all } w_\kappa\in\mathbb{L}_\kappa, \text{ for all } \kappa > 0. \end{array}\right. \tag{5.2.4}$$

5.2.3 Discrete problem and main results

In this section, we present the approximate problems for the parabolic Signorini model. As in the elliptic case, the gradient scheme for such a model is obtained by replacing the continuous spaces and operators in the weak formulations (5.2.2) by a set of discrete elements called a space–time gradient discretisation. Besides the notations of previous gradient discretisation (Definition 2.2.3), time steps and interplants are introduced to deal with the time dependency and initial solutions.

Definition 5.2.2 (GD for time-dependent Signorini problem). Let Ω be an open subset of $\mathbb{R}^d$ (with $d = 1,2,3$) and $T > 0$. A space–time gradient discretisation $\mathcal{D}^T$ for problems with the Signorini and homogeneous Dirichlet boundary conditions is a family $\mathcal{D}^T = (\mathcal{D}, J_{\mathcal{D}}, (t^{(n)})_{n=0,...,N})$, where:

1. $\mathcal{D} = (X_{\mathcal{D},\Gamma_2}, \Pi_{\mathcal{D}}, \mathbb{T}_{\mathcal{D}}, \nabla_{\mathcal{D}})$ is a (time–independent) gradient discretisation, whose elements are defined by:
   - $X_{\mathcal{D},\Gamma_2}$, a finite dimensional vector space on $\mathbb{R}$, taking into account the zero boundary conditions on Γ_1,
   - the linear mapping $\Pi_{\mathcal{D}} : X_{\mathcal{D},\Gamma_2} \to L^2(\Omega)$ is the function reconstruction,
   - the linear mapping $\mathbb{T}_{\mathcal{D}} : X_{\mathcal{D},\Gamma_2} \to L^2(\partial\Omega)$ is the trace reconstruction,
   - the linear mapping $\nabla_{\mathcal{D}} : X_{\mathcal{D},\Gamma_2} \to L^2(\Omega)^d$ is a gradient reconstruction, which must be defined such that $\|\nabla_{\mathcal{D}} \cdot \|_{L^2(\Omega)^d}$ is a norm on $X_{\mathcal{D},\Gamma_2}$.
2. $J_{\mathcal{D}} : \mathcal{K} \to \mathcal{K}_{\mathcal{D}}$ is an interpolation operator, where the discrete set $\mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},\Gamma_2} : \ \mathbb{T}_{\mathcal{D}} v \le a \text{ on } \Gamma_2\}$,
3. $t^{(0)} = 0 < t^{(1)} < .... < t^{(N)} = T$.

Remark 5.2.2. It is classical, for any $v = (v^{(n)})_{n=0,...,N} \in X_{\mathcal{D},\Gamma_2}^{N+1}$, to define the three space–time functions, the reconstructed function $\Pi_{\mathcal{D}} v : \Omega \times [0,T] \to \mathbb{R}$, the reconstructed gradient $\nabla_{\mathcal{D}} v : \Omega \times [0,T] \to \mathbb{R}^d$ and the reconstructed trace $\mathbb{T}_{\mathcal{D}} v : \Gamma_2 \times [0,T] \to \mathbb{R}$, given by:

$$
\begin{aligned}
&\Pi_{\mathcal{D}} v(\cdot,0) = \Pi_{\mathcal{D}} v^{(0)} \text{ and } \forall n = 0,...,N-1, \ \forall t \in (t^{(n)}, t^{(n+1)}], \forall \boldsymbol{x} \in \Omega,\\
&\Pi_{\mathcal{D}} v(\boldsymbol{x},t) = \Pi_{\mathcal{D}} v^{(n+1)}(\boldsymbol{x}), \ \nabla_{\mathcal{D}} v(\boldsymbol{x},t) = \nabla_{\mathcal{D}} v^{(n+1)}(\boldsymbol{x}),\\
&\text{and } \mathbb{T}_{\mathcal{D}} v(\boldsymbol{x},t) = \mathbb{T}_{\mathcal{D}} v^{(n+1)}(\boldsymbol{x}).
\end{aligned}
$$

Setting $\delta t^{(n+\frac{1}{2})} = t^{(n+1)} - t^{(n)}$, for $n = 0,...,N-1$, and $\delta t_{\mathcal{D}} = \max_{n=0,...,N-1} \delta t^{(n+\frac{1}{2})}$, the discrete derivative $\delta_{\mathcal{D}} v \in L^\infty(0,T;L^2(\Omega))$ of $v \in X_{\mathcal{D},\Gamma_2}^{N+1}$ is defined by

$$
\delta_{\mathcal{D}} v(t) = \delta_{\mathcal{D}}^{(n+\frac{1}{2})} v := \frac{\Pi_{\mathcal{D}} v^{(n+1)} - \Pi_{\mathcal{D}} v^{(n)}}{\delta t^{(n+\frac{1}{2})}}, \text{ for all } n = 0,...,N-1 \text{ and } t \in (t^{(n)}, t^{(n+1)}]. \tag{5.2.5}
$$

Based on the above gradient discretisation, the gradient scheme (GS) for the Signorini problem is given in the following definition.

Definition 5.2.3 (Gradient scheme). The gradient scheme for Problem (5.2.2) consists in a sequence $u = (u^{(n)})_{n=0,...,N} \subset \mathcal{K}_{\mathcal{D}}$, such that $u^{(0)} = J_{\mathcal{D}} \bar{u}_{\rm ini}$, and for all $n = 0,...,N-1$,

$$
\left\{
\begin{aligned}
&\int_\Omega \delta_{\mathcal{D}}^{(n+\frac{1}{2})} u(\boldsymbol{x})\, \Pi_{\mathcal{D}}(u^{(n+1)}(\boldsymbol{x}) - v(\boldsymbol{x}))\, \mathrm{d}\boldsymbol{x} + \int_\Omega \Lambda(\boldsymbol{x}) \nabla_{\mathcal{D}} u^{(n+1)}(\boldsymbol{x}) \cdot \nabla_{\mathcal{D}}(u^{(n+1)} - v)(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}\\
&\le \frac{1}{\delta t^{(n+\frac{1}{2})}} \int_{t^{(n)}}^{t^{(n+1)}} \int_\Omega f(\boldsymbol{x},t) \Pi_{\mathcal{D}}(u^{(n+1)}(\boldsymbol{x}) - v(\boldsymbol{x}))\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t, \text{ for all } v \in \mathcal{K}_{\mathcal{D}}.
\end{aligned}
\right. \tag{5.2.6}
$$

Multiplying this scheme by $\delta t^{(n+\frac{1}{2})}$, putting $v = v^{(n)}$, summing over n and using the notations in Remark 5.2.2, the above gradient scheme is equivalent to finding a sequence $(u^{(n)})_{n=0,\ldots,N} \subset \mathcal{K}_{\mathcal{D}}$ such that

$$\left\{\begin{array}{l} u^{(0)} = J_{\mathcal{D}}\bar{u}_{\rm ini}, \text{ and for all } v = (v^n)_{n=0,\ldots,N} \subset \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\int_0^T\int_\Omega \delta_{\mathcal{D}}u(t)\,\Pi_{\mathcal{D}}(u-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \int_0^T\int_\Omega \Lambda(\boldsymbol{x})\nabla_{\mathcal{D}}u(\boldsymbol{x},t)\cdot\nabla_{\mathcal{D}}(u-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\ \displaystyle\leq \int_0^T\int_\Omega f(\boldsymbol{x},t)\Pi_{\mathcal{D}}(u-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \end{array}\right. \tag{5.2.7}$$

The convergence of this GS is guaranteed by the usual three properties. The coercivity and the limit-conformity are similar to those in the steady problems, whereas the consistency needs to be modified to include information on the time steps and initial interpolant. For the sake of completeness, we recall them all.

Definition 5.2.4 (Coercivity). If $T > 0$ and $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is a sequence of space–time gradient discretisation in the sense of Definition 5.2.2, then $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is **coercive** if the sequence of time-independent $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is coercive. This means that $(C_{\mathcal{D}_m})_{m\in\mathbb{N}}$ remains bounded, where

$$C_{\mathcal{D}} = \max_{v\in X_{\mathcal{D},\Gamma_2}\setminus\{0\}} \Big(\frac{\|\Pi_{\mathcal{D}}v\|_{L^2(\Omega)}}{\|\nabla_{\mathcal{D}}v\|_{L^2(\Omega)^d}} + \frac{\|\mathbb{T}_{\mathcal{D}}v\|_{L^2(\partial\Omega)}}{\|\nabla_{\mathcal{D}}v\|_{L^2(\Omega)^d}}\Big). \tag{5.2.8}$$

Definition 5.2.5 (Space–time consistency). If $T > 0$ and $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is a sequence of space–time gradient discretisation in the sense of Definition 5.2.2, then $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is **consistent** if:

1. $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is consistent, that is, for all $\varphi \in \mathcal{K}$, $\lim_{m\to\infty} S_{\mathcal{D}_m}(\varphi) = 0$, where $S_{\mathcal{D}} : \mathcal{K} \to [0,+\infty)$ is defined by
$$\forall\varphi\in\mathcal{K},\ S_{\mathcal{D}}(\varphi) = \min_{v\in\mathcal{K}_{\mathcal{D}}} \left(\|\Pi_{\mathcal{D}}v-\varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}}v - \nabla\varphi\|_{L^2(\Omega)^d}\right), \tag{5.2.9}$$
2. $\Pi_{\mathcal{D}_m}J_{\mathcal{D}_m}\bar{u}_{\rm ini} \to \bar{u}_{\rm ini}$ in $L^2(\Omega)$, as $m\to\infty$,
3. $\delta t_{\mathcal{D}_m} \to 0$, as $m\to\infty$.

Definition 5.2.6 (Limit-conformity). If $T > 0$ and $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is a sequence of space–time gradient discretisation in the sense of Definition 5.2.2, then $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is **limit-conforming** if the sequence of time-independent $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is limit-conforming. This means that for all $\boldsymbol{\psi} \in C^2(\overline{\Omega})^d$, $\lim_{m\to\infty} W_{\mathcal{D}_m}(\boldsymbol{\psi}) = 0$, where $W_{\mathcal{D}} : C^2(\overline{\Omega}) \to [0,+\infty)$ is defined by

$$\forall\boldsymbol{\psi}\in C^2(\overline{\Omega})^d,\quad W_{\mathcal{D}}(\boldsymbol{\psi}) = \sup_{v\in X_{\mathcal{D},\Gamma_2}\setminus\{0\}} \frac{\Big|\displaystyle\int_\Omega (\nabla_{\mathcal{D}}v\cdot\boldsymbol{\psi} + \Pi_{\mathcal{D}}v\,\mathrm{div}(\boldsymbol{\psi}))\,\mathrm{d}\boldsymbol{x} - \int_{\Gamma_2}\boldsymbol{\psi}\cdot\mathbf{n}\mathbb{T}_{\mathcal{D}}v\,\mathrm{d}\boldsymbol{x}\Big|}{\|\nabla_{\mathcal{D}}v\|_{L^2(\Omega)^d}}.$$

Remark 5.2.3. As in the nonlinear elliptic Signorini problem (see Remark 4.2.2), we need to define $W_{\mathcal{D}}$ in the above definition only on smooth functions $\boldsymbol{\psi} \in C^2(\overline{\Omega})$. Since the convergence of the GS is proved based on the compactness technique, this definition is sufficient to use Lemma 5.4.1 throughout the proof of convergence.

The convergence results of the gradient scheme for the Signorini problem is stated in the following theorem, whose proof is detailed in Section 5.4.2.

Theorem 5.2.7 (Convergence of the GS for the Signorini problem). *Under Assumptions 5.2.1, let $\bar{u}$ be the unique solution to* (5.2.2)*. Let $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ be a sequence of space–time gradient discretisations in the sense of Definition 5.2.2, that is coercive, space–time consistent and limit-conforming in the sense of Definitions 5.2.4, 5.2.5 and 5.2.6, and such that $\mathcal{K}_{\mathcal{D}_m}$ is a nonempty set for any $m\in\mathbb{N}$. Also, assume that $(||\nabla_{\mathcal{D}_m}J_{\mathcal{D}_m}\bar{u}_{\rm ini}||_{L^2(\Omega)^d})_{m\in\mathbb{N}}$ is bounded. Then for each $m\in\mathbb{N}$ there is a unique solution $u_m \in \mathcal{K}_{\mathcal{D}_m}^{N_m+1}$ to the scheme* (5.2.6) *with $\mathcal{D}^T = \mathcal{D}_m^T$ and, as $m\to\infty$,*

1. $\Pi_{\mathcal{D}_m} u_m$ *converges strongly to* $\bar{u}$ *in* $L^\infty(0,T;L^2(\Omega))$,
2. $\nabla_{\mathcal{D}_m} u_m$ *converges strongly to* $\nabla\bar{u}$ *in* $L^2(\Omega\times(0,T))^d$,
3. $\delta_{\mathcal{D}_m} u_m$ *converges weakly to* $\partial_t\bar{u}$ *in* $L^2(\Omega\times(0,T))$.

5.3 Parabolic obstacle problem

5.3.1 Continuous model and weak formulation

This section is concerned with a linear parabolic obstacle model given by

$$\begin{aligned}
(\partial_t\bar{u}-\operatorname{div}(\Lambda\nabla\bar{u})-f)(\bar{u}-\psi)&=0 \quad \text{in } \Omega\times(0,T), && (5.3.1a)\\
\partial_t\bar{u}-\operatorname{div}(\Lambda\nabla\bar{u})&\geq f \quad \text{in } \Omega\times(0,T), && (5.3.1b)\\
\bar{u}&\geq\psi \quad \text{in } \Omega\times(0,T), && (5.3.1c)\\
\bar{u}&=0 \quad \text{on } \partial\Omega\times(0,T), && (5.3.1d)\\
\bar{u}(\boldsymbol{x},0)&=\bar{u}_{\rm ini} \quad \text{in } \Omega\times\{0\}, && (5.3.1e)
\end{aligned}$$

with the following assumptions on the data:

$$\begin{aligned}
&T>0, \text{ and } \Omega \text{ and } \Lambda \text{ satisfies Items 1 and 2 in Assumptions 5.2.1},\\
&f\in L^2(\Omega\times(0,T)) \text{ and } \psi\in H^1(\Omega)\cap C(\overline{\Omega}), \text{ such that } \psi\leq 0 \text{ on } \partial\Omega, \qquad (5.3.2)\\
&\bar{u}_{\rm ini}\in\mathcal{K}:=\{v\in H_0^1(\Omega):\ v\geq\psi \text{ in } \Omega\}.
\end{aligned}$$

The time-dependent closed convex subset of the space $L^2(0,T;H_0^1(\Omega))$ is defined by

$$\mathbb{K}=\{v\in L^2(0,T;H_0^1(\Omega)):\ v(t)\in\mathcal{K} \text{ for a.e. } t\in(0,T)\}.$$

If $\bar{u}$ is a solution to (5.3.3), then $\bar{u}$ belongs to the set $\mathbb{K}$. Assume (5.3.2) hold. Following standard variational techniques (as in the case of the elliptic model, Proposition 2.A.2), the weak formulation of Problem (5.3.3) is:

$$\left\{\begin{aligned}
&\text{Find } \bar{u}\in\mathbb{K}\cap C^0([0,T];L^2(\Omega)), \text{ such that } \partial_t\bar{u}\in L^2(0,T;L^2(\Omega)),\ \bar{u}(\boldsymbol{x},0)=\bar{u}_{\rm ini}, \text{ and}\\
&\int_0^T\int_\Omega \partial_t\bar{u}(\boldsymbol{x},t)(u(\boldsymbol{x},t)-v(\boldsymbol{x},t))\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t+\int_0^T\int_\Omega \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x},t)\cdot\nabla(\bar{u}-v)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&\leq\int_0^T\int_\Omega f(\boldsymbol{x},t)(\bar{u}(\boldsymbol{x},t)-v(\boldsymbol{x},t))\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \qquad \text{for all } v\in\mathbb{K}.
\end{aligned}\right. \qquad (5.3.3)$$

It has been proved in [74, 44] that there exists a unique weak solution to Problem (5.3.3). Indeed the set $\mathcal{K}$ contains ψ^+ (it belongs to $H_0^1(\Omega)$) and thus the set $\mathbb{K}$ contains the constant in time function $t\mapsto\psi^+$.

Thanks to the fact that the set $C^2(\Omega)\cap\mathcal{K}$ is dense in the set $\mathcal{K}$ (proved in [56]), Remark 5.2.1 is still applicable to the obstacle problem, and the weak formulation (5.3.3) can be recast as

$$\left\{\begin{aligned}
&\text{Find } \bar{u}\in\mathbb{K}\cap C^0([0,T];L^2(\Omega)), \text{ such that } \partial_t\bar{u}\in L^2(0,T;L^2(\Omega)),\ \bar{u}(\boldsymbol{x},0)=\bar{u}_{\rm ini}, \text{ and}\\
&\int_0^T\int_\Omega \partial_t\bar{u}(\boldsymbol{x},t)(u-w_\kappa)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t+\int_0^T\int_\Omega \Lambda(\boldsymbol{x})\nabla\bar{u}(\boldsymbol{x},t)\cdot\nabla(\bar{u}-w_\kappa)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&\leq\int_0^T\int_\Omega f(\boldsymbol{x},t)(\bar{u}-w_\kappa)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \qquad \text{for all } w_\kappa\in\mathbb{L}_\kappa, \text{ and for all } \kappa>0,
\end{aligned}\right. \qquad (5.3.4)$$

where

$$\mathbb{L}_\kappa=\left\{w_\kappa(\boldsymbol{x},t)=\sum_{i=1}^{\ell_\kappa}\mathbf{1}_{I_i}(t)\varphi_i(\boldsymbol{x})\ :\ \varphi\in C^2(\overline{\Omega}) \text{ and } \varphi\geq\psi \text{ in } \Omega\right\}. \qquad (5.3.5)$$

5.3.2 Discrete problem and main results

Definition 5.3.1 (GD for time-dependent obstacle problem). Let Ω be an open subset of $\mathbb{R}^d$ (with $d = 1, 2, 3$) and $T > 0$. A space–time gradient discretisation $\mathcal{D}^T$ for the obstacle problem with homogeneous Dirichlet boundary conditions is a family $\mathcal{D}^T = (\mathcal{D}, \psi_{\mathcal{D}}, J_{\mathcal{D}}, (t^{(n)})_{n=0,...,N})$, where:

1. $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$ is a gradient discretisation in the sense of Definition 2.2.5,
2. $\psi_{\mathcal{D}} \in L^2(\Omega)$ is an approximation of the barrier ψ,
3. $J_{\mathcal{D}} : \mathcal{K} \to \mathcal{K}_{\mathcal{D}}$ is an interpolation operator, where $\mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},\Gamma_2} : \Pi_{\mathcal{D}} v \geq \psi_{\mathcal{D}}, \text{ in } \Omega\}$,
4. $t^{(0)} = 0 < t^{(1)} < .... < t^{(N)} = T$.

Remark 5.3.1. We define here the set $\mathcal{K}_{\mathcal{D}}$ based on the approximate barrier $\psi_{\mathcal{D}} \in L^2(\Omega)$ to be able to construct an interpolant that belongs to $\mathcal{K}_{\mathcal{D}}$. This is unlike the Signorini problem where the barrier a is assumed to be constant to use the density results, and thus there is no need to use an approximate barrier.

With the notations recalled in Remark 5.2.2 and using the discrete elements given in the above definition, the gradient scheme for (5.3.3) is: seek $u = (u^{(n)})_{n=0,...,N} \subset \mathcal{K}_{\mathcal{D}}$, such that $u^{(0)} = J_{\mathcal{D}} \bar{u}_{\rm ini} \in \mathcal{K}_{\mathcal{D}}$ and for all $n = 0, ..., N-1$,

$$\left\{\begin{array}{l} \displaystyle\int_\Omega \delta_{\mathcal{D}}^{(n+\frac{1}{2})} u(\boldsymbol{x})\, \Pi_{\mathcal{D}}(u^{(n+1)}(\boldsymbol{x}) - v(\boldsymbol{x}))\, \mathrm{d}\boldsymbol{x} + \int_\Omega \Lambda(\boldsymbol{x}) \nabla_{\mathcal{D}} u^{(n+1)}(\boldsymbol{x}) \cdot \nabla_{\mathcal{D}}(u^{(n+1)} - v)(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\ \displaystyle\leq \frac{1}{\delta t^{(n+\frac{1}{2})}} \int_{t^{(n)}}^{t^{(n+1)}} \int_\Omega f(\boldsymbol{x}, t) \Pi_{\mathcal{D}}(u^{(n+1)}(\boldsymbol{x}) - v(\boldsymbol{x}))\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t, \text{ for all } v \in \mathcal{K}_{\mathcal{D}}. \end{array}\right. \tag{5.3.6}$$

Applying this scheme to $v = v^{(n)}$, multiplying by $\delta t^{(n+\frac{1}{2})}$ and summing over $n = 0, ..., N$, we obtain an equivalent problem consisting in finding a sequence $(u^{(n)})_{n=0,...,N} \subset \mathcal{K}_{\mathcal{D}}$ such that

$$\left\{\begin{array}{l} u^{(0)} = J_{\mathcal{D}} \bar{u}_{\rm ini} \text{ and for all } v = (v^n)_{n=1,...N} \subset \mathcal{K}_{\mathcal{D}}, \\ \displaystyle\int_0^T \int_\Omega \delta_{\mathcal{D}} u(t)\, \Pi_{\mathcal{D}}(u - v)(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t + \int_0^T \int_\Omega \Lambda(\boldsymbol{x}) \nabla_{\mathcal{D}} u(\boldsymbol{x}, t) \cdot \nabla_{\mathcal{D}}(u - v)(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t \\ \displaystyle\leq \int_0^T \int_\Omega f(\boldsymbol{x}, t) \Pi_{\mathcal{D}}(u - v)(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t. \end{array}\right. \tag{5.3.7}$$

The three properties used to assess the convergence of the above gradient scheme are listed in the following definitions.

Definition 5.3.2 (Coercivity and limit-conformity). Assume that $T > 0$ and $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is a sequence of space–time gradient discretisation in the sense of Definition 5.3.1. Then the sequence $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is:

1. **coercive** if the sequence of the gradient discretisation $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is coercive; that is $(C_{\mathcal{D}_m})_{m\in\mathbb{N}}$ remains bounded, where $C_{\mathcal{D}}$ is given by (2.2.8),
2. **limit-conforming** if the sequence of the gradient discretisation $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is limit-conforming; for all $\boldsymbol{\psi} \in H_{\rm div}(\Omega)$, $\lim_{m\to\infty} W_{\mathcal{D}_m}(\boldsymbol{\psi}) = 0$, where $W_{\mathcal{D}}$ is defined by (2.2.10).

Definition 5.3.3 (Space–time consistency). If $T > 0$. A sequence $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ of space–time gradient discretisation in the sense of Definition 5.3.1 is **consistent** if:

1. $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is consistent; that is, for all $\varphi \in \mathcal{K}$, $\lim_{m\to\infty} S_{\mathcal{D}_m}(\varphi) = 0$, where $S_{\mathcal{D}} : \mathcal{K} \to [0, +\infty)$ is defined by
$$\forall \varphi \in \mathcal{K},\ S_{\mathcal{D}}(\varphi) = \min_{v\in\mathcal{K}_{\mathcal{D}}} \left(\|\Pi_{\mathcal{D}} v - \varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla\varphi\|_{L^2(\Omega)^d} \right), \tag{5.3.8}$$

2. $\psi_{\mathcal{D}_m} \to \psi$ in $L^2(\Omega)$, as $m \to \infty$,
3. $\Pi_{\mathcal{D}_m} J_{\mathcal{D}_m} \bar{u}_{\rm ini} \to \bar{u}_{\rm ini}$ in $L^2(\Omega)$, as $m \to \infty$,
4. $\delta t_{\mathcal{D}_m} \to 0$, as $m \to \infty$.

The following theorem, whose proof is identical to that of Theorem 5.2.7 and therefore is omitted, states the convergence of the gradient scheme to a weak solution of the obstacle problem. The inclusion of the approximate barrier $\psi_{\mathcal{D}}$ in the discrete set $\mathcal{K}_{\mathcal{D}}$ does not entail any major change in the convergence analysis.

Theorem 5.3.4 (Convergence of the GS for the obstacle problem). *We assume* (5.3.2) *and consider a sequence $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ of space–time gradient discretisations, in the sense of Definition 5.3.1, that is coercive, limit-conforming and consistent in the sense of Definitions 5.3.2 and 5.3.3, and such that $\mathcal{K}_{\mathcal{D}_m}$ is a nonempty set for any $m \in \mathbb{N}$. Also, we assume that $(||\nabla_{\mathcal{D}_m} J_{\mathcal{D}_m} \bar{u}_{\rm ini}||_{L^2(\Omega)^d})_{m\in\mathbb{N}}$ is bounded. Let $\bar{u}$ be the unique solution to* (5.3.3). *Then for any $m \in \mathbb{N}$ there is a unique solution $u_m \in \mathcal{K}_{\mathcal{D}_m}^{N_m+1}$ to the GS* (5.2.6) *with $\mathcal{D}^T = \mathcal{D}_m^T$ and the following convergence occur, as $m \to \infty$:*

1. *$\Pi_{\mathcal{D}_m} u_m \to \bar{u}$ strongly in $L^\infty(0,T;L^2(\Omega))$,*
2. *$\nabla_{\mathcal{D}_m} u_m \to \nabla\bar{u}$ strongly in $L^2(\Omega\times(0,T))^d$,*
3. *$\delta_{\mathcal{D}_m} u_m \to \partial_t \bar{u}$ weakly in $L^2(\Omega\times(0,T))$.*

5.4 Proof of the convergence of GS for the Signorini problem (Theorem 5.2.7)

5.4.1 Technical lemmas

To prove the convergence result, we establish here some estimates on the scheme's solution and its gradient. Let us first recall the results of [37, Lemma 4.7] concerning the regularity of the limit in the space–time setting.

Lemma 5.4.1. *Let $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ be a sequence of space–time gradient discretisations in the sense of Definition 5.2.2, which is coercive and limit-conforming in the sense of Definitions 5.2.4 and 5.2.6. Let $u_m \in X_{\mathcal{D}_m,\Gamma_2}^{N_m+1}$ be such that $(||\nabla_{\mathcal{D}_m} u_m||_{L^2(0,T;L^2(\Omega)^d)})_{m\in\mathbb{N}}$ is bounded. Then, there exists $\bar{u} \in L^2(0,T;H^1_{\Gamma_1}(\Omega))$ such that, up to a subsequence, as $m \to \infty$,*

$$\begin{aligned}
&\Pi_{\mathcal{D}_m} u_m \to \bar{u} \textit{ weakly in } L^2(\Omega\times(0,T)),\\
&\nabla_{\mathcal{D}_m} u_m \to \nabla\bar{u} \textit{ weakly in } L^2(\Omega\times(0,T))^d,\\
&\mathbb{T}_{\mathcal{D}_m} u_m \to \gamma\bar{u} \textit{ weakly in } L^2(\partial\Omega\times(0,T)).
\end{aligned}$$

Lemma 5.4.2 (Interpolation of space–time functions). *For $T>0$, let $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ be a sequence of space–time gradient discretisations in the sense of Definition 5.2.2, that is space–time-consistent in the sense of Definition 5.2.5. Let $\bar{w}_\kappa \in \mathbb{L}_\kappa$ be a piecewise constant in time function, where $\mathbb{L}_\kappa$ is the set defined by* (5.2.3). *Then the following statements hold.*

1. *There exists a sequence $(w_m)_{m\in\mathbb{N}}$ such that $w_m = (w_m^{(n)})_{n=0,\ldots,N_m} \in \mathcal{K}_{\mathcal{D}_m}^{N_m+1}$ for all $m \in \mathbb{N}$, and, as $m \to \infty$,*

$$\Pi_{\mathcal{D}_m} w_m \to \bar{w}_\kappa \textit{ strongly in } L^2(\Omega\times(0,T)), \tag{5.4.2a}$$

$$\nabla_{\mathcal{D}_m} w_m \to \nabla\bar{w}_\kappa \textit{ strongly in } L^2(\Omega\times(0,T))^d. \tag{5.4.2b}$$

2. *If* $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ *is coercive and limit-conforming in the sense of Definitions 5.2.4 and 5.2.6, then the sequence* $(w_m)_{m\in\mathbb{N}}$ *in Item 1 also satisfies*

$$\mathbb{T}_{\mathcal{D}_m} w_m \to \gamma(\bar{w}_\kappa) \text{ weakly in } L^2(\partial\Omega\times(0,T)), \text{ as } m\to\infty. \tag{5.4.3}$$

Proof. Since the set $\mathbb{L}_\kappa$ is defined in Remark 5.2.1, write $\bar{w}_\kappa(\boldsymbol{x},t) = \sum_{i=1}^{\ell_\kappa} \mathbf{1}_{I_i}(t)\varphi_i(\boldsymbol{x})$ such that $\varphi_i \in C^2(\overline{\Omega})\cap\mathcal{K}$. Let $s\in(0,T)$ and choose $n := n(s)$ such that $s\in(t^{(n(s))},t^{(n(s)+1)}]$. Let $w_m \in X_{\mathcal{D}_m,\Gamma_2}$ be defined by $w_m = \sum_{i=1}^{\ell_\kappa} \mathbf{1}_{I_i}(t^{(n(s)+1)})P_{\mathcal{D}_m}\varphi_i$ and

$$P_{\mathcal{D}_m}(\varphi) = \underset{v\in\mathcal{K}_{\mathcal{D}_m}}{\operatorname{argmin}} \big(\|\Pi_{\mathcal{D}_m} v - \varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}_m} v - \nabla\varphi\|_{L^2(\Omega)^d}\big). \tag{5.4.4}$$

For $i=1,...,\ell_\kappa$, we define $\chi_m^i : (0,T)\to\mathbb{R}$ by $\chi_m^i(s) = \mathbf{1}_{I_i}(t^{(n(s)+1)})$ 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. $\boldsymbol{x}\in\Omega$,

$$(\Pi_{\mathcal{D}_m} w_m - \bar{w}_\kappa)(\boldsymbol{x},s) = \sum_{i=1}^{\ell_\kappa} \big(\chi_m^i(s) - \mathbf{1}_{I_i}(s)\big)\Pi_{\mathcal{D}_m} P_{\mathcal{D}_m}\varphi_i(\boldsymbol{x}) + \sum_{i=1}^{\ell_\kappa} \mathbf{1}_{I_i}(s)\left(\Pi_{\mathcal{D}_m} P_{\mathcal{D}_m}\varphi_i - \varphi_i\right)(\boldsymbol{x}).$$

An application of the definition of $S_{\mathcal{D}_m}$ yields

$$\begin{aligned}
||\Pi_{\mathcal{D}_m} w_m - \bar{w}_\kappa||_{L^2(\Omega\times(0,T))} &\le \sum_{i=1}^{\ell_\kappa} ||\chi_m^i(s) - \mathbf{1}_{I_i}(s)||_{L^2(0,T)} ||\Pi_{\mathcal{D}_m} P_{\mathcal{D}_m}\varphi_i||_{L^2(\Omega)} \\
&\quad + \sum_{i=1}^{\ell_\kappa} ||\mathbf{1}_{I_i}(s)||_{L^2(0,T)} ||\Pi_{\mathcal{D}_m} P_{\mathcal{D}_m}\varphi_i - \varphi_i||_{L^2(\Omega)} \\
&\le \sum_{i=1}^{\ell_\kappa} ||\chi_m^i(s) - \mathbf{1}_{I_i}(s)||_{L^2(0,T)} \left(S_{\mathcal{D}_m}(\varphi_i) + ||\varphi_i||_{L^2(\Omega)}\right) + C_1 \sum_{i=1}^{\ell_\kappa} S_{\mathcal{D}_m}(\varphi_i),
\end{aligned} \tag{5.4.5}$$

where $C_1 = \sum_{i=1}^{\ell_\kappa} ||\mathbf{1}_{I_i}||_{L^2(0,T)}$. From the consistency property, it follows that $S_{\mathcal{D}_m}(\varphi_i)\to 0$ as $m\to\infty$, for any $i=0,...,\ell_\kappa$, implying that the second term on the right-hand side vanishes. In the case in which both s, $t^{(n(s)+1)}\in I_i$ or both s, $t^{(n(s)+1)}\notin I_i$, the quantity $\chi_m^i(s)-\mathbf{1}_{I_i}(s)$ equals zero. In the case in which $s\in I_i$ and $t^{(n(s)+1)}\notin I_i$ or $s\notin I_i$ and $t^{(n(s)+1)}\in I_i$, one can deduce (writing $I_i=[a_i,b_i]$ and because s is chosen such that $|s-t^{(n(s)+1)}|\le\delta t_{\mathcal{D}_m}$

$$\begin{aligned}
||\chi_m^i(s) - \mathbf{1}_{I_i}(s)||^2_{L^2(0,T)} &\le \text{measure}([a_i-\delta t_{\mathcal{D}_m}, a_i+\delta t_{\mathcal{D}_m}]\cup[b_i-\delta t_{\mathcal{D}_m}, b_i+\delta t_{\mathcal{D}_m}]) \\
&\le 4\delta t_{\mathcal{D}_m}.
\end{aligned}$$

This shows that the first term on the right-hand side of (5.4.5) tends to zero when $m\to\infty$. Hence, (5.4.2a) is concluded. The proof of (5.4.2b) is obtained by the same reasoning, replacing $\bar{w}_\kappa$ by $\nabla\bar{w}_\kappa$ and $\Pi_{\mathcal{D}_m} w_m$ by $\nabla_{\mathcal{D}_m} w_m$. Item 2 follows by applying Lemma 5.4.1 to v_m. □

Lemma 5.4.3 (Estimate on the discrete gradient and time derivative). *Under Assumptions 5.2.1, let* $\mathcal{D}^T$ *be a space–time gradient discretisation in the sense of Definition 5.2.2, such that* $\mathcal{K}_\mathcal{D}$ *is a nonempty set, and let* $u\in\mathcal{K}_\mathcal{D}$ *be a solution of the gradient scheme* (5.2.6). *If* $||\nabla_\mathcal{D} J_\mathcal{D}\bar{u}_{\rm ini}||_{L^2(\Omega)^d}$ *is bounded by* C_2, *then there exists a constant* $C_3\ge 0$ *only depending on* C_2 *and* f, *such that*

$$||\delta_\mathcal{D} u||_{L^2(\Omega\times(0,T))} + ||\nabla_\mathcal{D} u||_{L^\infty(0,T;L^2(\Omega)^d)} \le C_3. \tag{5.4.6}$$

Proof. Setting, as a test function in Scheme (5.2.6), the function $v := u^{(n)}$ (it belongs to $\mathcal{K}_{\mathcal{D}}$) leads to

$$\delta t^{(n+\frac{1}{2})}\int_\Omega |\delta_{\mathcal{D}}^{(n+\frac{1}{2})}u|^2 \,\mathrm{d}\boldsymbol{x} + \int_\Omega \Lambda\nabla_{\mathcal{D}}u^{(n+1)}\cdot\nabla_{\mathcal{D}}(u^{(n+1)}-u^{(n)})\,\mathrm{d}\boldsymbol{x} \le \int_{t^{(n)}}^{t^{(n+1)}}\int_\Omega f\delta_{\mathcal{D}}^{(n+\frac{1}{2})}u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{5.4.7}$$

Using the fact that $(r-s)\cdot r \ge \frac{1}{2}|r|^2 - \frac{1}{2}|s|^2$ with $r = \Lambda^{1/2}\nabla_{\mathcal{D}}u^{(n+1)}$ and $s = \Lambda^{1/2}\nabla_{\mathcal{D}}u^{(n)}$, it follows

$$\int_\Omega \Lambda\nabla_{\mathcal{D}}u^{(n+1)}\cdot\nabla_{\mathcal{D}}(u^{(n+1)}-u^{(n)})\,\mathrm{d}\boldsymbol{x} \ge \frac{\underline{\lambda}}{2}\int_\Omega\left(|\nabla_{\mathcal{D}}u^{(n+1)}|^2 - |\nabla_{\mathcal{D}}u^{(n)}|^2\right)\,\mathrm{d}\boldsymbol{x}.$$

Plugging this inequality into (5.4.7) gives

$$\delta t^{(n+\frac{1}{2})}\int_\Omega |\delta_{\mathcal{D}}^{(n+\frac{1}{2})}u|^2 \,\mathrm{d}\boldsymbol{x} + \frac{\underline{\lambda}}{2}\int_\Omega\left(|\nabla_{\mathcal{D}}u^{(n+1)}|^2 - |\nabla_{\mathcal{D}}u^{(n)}|^2\right)\,\mathrm{d}\boldsymbol{x} \le \int_{t^{(n)}}^{t^{(n+1)}}\int_\Omega f\delta_{\mathcal{D}}^{(n+\frac{1}{2})}u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

We sum this inequality on $n = 0,...,m-1$, for some $m = 0,...,N$, to obtain

$$\int_0^{t^{(m)}}\int_\Omega |\delta_{\mathcal{D}}u|^2\,\mathrm{d}\boldsymbol{x} + \frac{\underline{\lambda}}{2}\int_\Omega\left(|\nabla_{\mathcal{D}}u^{(m)}|^2 - |\nabla_{\mathcal{D}}u^{(0)}|^2\right)\,\mathrm{d}\boldsymbol{x} \le \int_0^{t^{(m)}}\int_\Omega f\delta_{\mathcal{D}}u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{5.4.8}$$

Applying the Cauchy-Schwarz' and Young's inequalities to the term on the right-hand side implies that, for all $m = 0,...,N$,

$$\int_0^{t^{(m)}}\int_\Omega f\delta_{\mathcal{D}}u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \le \frac{1}{2}||f||^2_{L^2(\Omega\times(0,t^{(m)}))} + \frac{1}{2}||\delta_{\mathcal{D}}u||^2_{L^2(\Omega\times(0,t^{(m)}))}.$$

Substituting this relation into the inequality (5.4.8) leads to, for all $m = 0,...,N$,

$$\frac{1}{2}||\delta_{\mathcal{D}}u||^2_{L^2(\Omega\times(0,t^{(m)}))} + \frac{\underline{\lambda}}{2}\int_\Omega |\nabla_{\mathcal{D}}u^{(m)}|^2\,\mathrm{d}\boldsymbol{x} \le \frac{1}{2}||f||^2_{L^2(\Omega\times(0,t^{(m)}))} + \frac{\underline{\lambda}}{2}||\nabla_{\mathcal{D}}u^{(0)}||^2_{L^2(\Omega)^d}.$$

Thanks to the fact that $\sup_n(a_n) + \sup_n(b_n) \le 2\sup_n(a_n+b_n)$, taking the supremum on $m = 0,...,N$ gives

$$\frac{1}{2}||\delta_{\mathcal{D}}u||_{L^2(\Omega\times(0,T))} + \frac{\underline{\lambda}}{2}\sup_{m=0,...,N}\int_\Omega |\nabla_{\mathcal{D}}u^{(m)}|^2\,\mathrm{d}\boldsymbol{x} \le ||f||^2_{L^2(\Omega\times(0,T))} + \underline{\lambda}||\nabla_{\mathcal{D}}u^{(0)}||^2_{L^2(\Omega)^d}.$$

This leads to Estimate (5.4.6). □

5.4.2 Proof of Theorem 5.2.7

The proof follows the same compactness technique used in Chapter 4. It is divided into three steps.

Step 1: existence and uniqueness of an approximate solution.

Note that at any time step $(n+1)$, we need to solve a gradient scheme for a linear elliptic variational inequality: setting $\alpha = \frac{1}{\delta t^{(n+\frac{1}{2})}}$, find $u^{(n+1)} \in \mathcal{K}_{\mathcal{D}}$, such that for all $v \in \mathcal{K}_{\mathcal{D}}$,

$$b(u^{(n+1)}, u^{(n+1)} - v) \le L(u^{(n+1)} - v), \tag{5.4.9}$$

with the bilinear form $b(v,w)$ and the linear form $L(w)$ respectively defined by

$$b(v,w) = \alpha\int_\Omega \Pi_{\mathcal{D}}v\Pi_{\mathcal{D}}w\,\mathrm{d}\boldsymbol{x} + \int_\Omega \nabla_{\mathcal{D}}v\cdot\nabla_{\mathcal{D}}w\,\mathrm{d}\boldsymbol{x},\quad \text{for all } v,w\in\mathcal{K}_{\mathcal{D}} \quad \text{and}$$

$$L(w)=\int_\Omega f\Pi_{\mathcal{D}}w\,\mathrm{d}\boldsymbol{x}+\alpha\int_\Omega \Pi_{\mathcal{D}}u^{(n)}\Pi_{\mathcal{D}}w\,\mathrm{d}\boldsymbol{x},\quad \text{for all } w\in\mathcal{K}_{\mathcal{D}}.$$

The assumptions for Stampacchia's theorem can easily be verified, and therefore there exists a unique weak solution to (5.4.9). This leads to the existence and uniqueness of the solution to (5.2.6).

Step 2: convergence towards the solution to the continuous model.

Applying Estimate (5.4.6) to the sequence of solutions $(u_m)_{m\in\mathbb{N}}$ of the scheme (5.2.7) shows that the norm $||\nabla_{\mathcal{D}_m}u_m||_{L^2(\Omega\times(0,T))^d}$ is bounded. Using Lemma 5.4.1, there exists a sequence, still denoted by $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$, and $\bar u\in L^2(0,T;H^1_{\Gamma_1}(\Omega))$, such that, as $m\to\infty$, $\Pi_{\mathcal{D}_m}u_m$ converges weakly to $\bar u$ in $L^2(\Omega\times(0,T))$ and $\nabla_{\mathcal{D}_m}u_m$ converges weakly to $\nabla\bar u$ in $L^2(\Omega\times(0,T))^d$, and $\mathbb{T}_{\mathcal{D}_m}u_m$ converges weakly to $\gamma(\bar u)$ in $L^2(\partial\Omega\times(0,T))$. Since $u_m\in\mathcal{K}_{\mathcal{D}_m}$, passing to the limit in $\mathbb{T}_{\mathcal{D}_m}u_m\le a$ on Γ_2 shows that $\gamma(\bar u)\le a$ on $\Gamma_2\times(0,T)$, because Γ_2 is an open set in $\partial\Omega$. Thanks to [37, Theorem 4.31], Estimate (5.4.6) shows that $\bar u\in C([0,T];L^2(\Omega))$, $\Pi_{\mathcal{D}_m}u_m$ converges strongly to $\bar u$ in $L^\infty(0,T;L^2(\Omega))$ and $\delta_{\mathcal{D}_m}u_m$ converges weakly to $\partial_t\bar u$ in $L^2(0,T;L^2(\Omega))$. Recall that $u_m^{(0)}=J_{\mathcal{D}_m}\bar u_{\rm ini}$, therefore the space–time consistency shows that $\Pi_{\mathcal{D}_m}u_m^{(0)}$ converges strongly to $\bar u_{\rm ini}$ in $L^2(\Omega)$, as $m\to\infty$. Hence, $\bar u\in C([0,T];L^2(\Omega))\cap\mathbb{K}$ and $\bar u$ satisfies all conditions except the integral inequality imposed on the exact solution of Problem (5.2.2). Let us now show that this integral relation holds. According to Remark 5.2.1, it is sufficient to show that $\bar u$ is a weak solution to (5.2.4). The L^2–weak convergence of $\nabla_{\mathcal{D}_m}u_m$ yields

$$\int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\le\liminf_{m\to\infty}\int_0^T\int_\Omega \Lambda\nabla_{\mathcal{D}_m}u_m\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{5.4.10}$$

Fix $\kappa>0$ and let $\bar w_\kappa\in\mathbb{L}_\kappa$. Thanks to Lemma 5.4.2, there exists a sequence $(w_m)_{m\in\mathbb{N}}$ such that $w_m\in\mathcal{K}_{\mathcal{D}_m}^{N_m+1}$ and $\Pi_{\mathcal{D}_m}w_m$ converges to $\bar w_\kappa$ strongly in $L^2(\Omega\times(0,T))$ and $\nabla_{\mathcal{D}_m}w_m$ converges to $\nabla\bar w_h$ strongly in $L^2(\Omega\times(0,T))^d$. Setting $v:=w_m$ as a test function in the scheme (5.2.7), Inequality (5.4.10) implies that

$$\begin{aligned}\int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\le\liminf_{m\to\infty}\Bigg[\int_0^T\int_\Omega f\Pi_{\mathcal{D}_m}(u_m-w_m)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t+\int_0^T\int_\Omega \Lambda\nabla_{\mathcal{D}_m}u_m\cdot\nabla_{\mathcal{D}_m}w_m\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ -\int_0^T\int_\Omega \delta_{\mathcal{D}_m}u_m\Pi_{\mathcal{D}_m}(u_m-w_m)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\Bigg].\end{aligned}$$

Using the convergences of u_m and w_m leads to

$$\begin{aligned}\int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\le\int_0^T\int_\Omega f(\bar u-\bar w_\kappa)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t+\int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla(\bar u-\bar w_\kappa)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ -\int_0^T\int_\Omega \partial_t\bar u(\bar u-\bar w_\kappa)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t,\end{aligned}$$

which shows that $\bar u$ is a weak solution to (5.2.4).

Step 3: proof of the strong convergence of the discrete gradients $\nabla_{\mathcal{D}_m}u_m$.

In view of the discrete inequality (5.2.7) and the previous convergences, we have

$$\begin{aligned}\limsup_{m\to\infty}\int_0^T\int_\Omega \Lambda\nabla_{\mathcal{D}_m}u_m\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\le\int_0^T\int_\Omega f(\bar u-\bar w_\kappa)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t+\int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla\bar w_\kappa\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ -\int_0^T\int_\Omega \partial_t\bar u(\bar u-\bar w_\kappa)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t,\qquad\text{for all } \bar w_\kappa\in\mathbb{L}_\kappa,\end{aligned}$$

According to Remark 5.2.1, for any $\bar v\in\mathbb{K}$, we can find $(\bar w_\kappa)_{\kappa>0}$ that converges, as $\kappa\to0$, to v in $L^2(0,T;H^1(\Omega))$.

Therefore, we infer

$$\limsup_{m\to\infty}\int_0^T\int_\Omega \Lambda\nabla_{\mathcal{D}_m}u_m\cdot\nabla_{\mathcal{D}_m}u_m \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \le \int_0^T\int_\Omega f(\bar u - v)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla v\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\ - \int_0^T\int_\Omega \partial_t\bar u(\bar u - v)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \quad \text{for all } \bar v\in\mathbb{K}.$$

Taking $v=\bar u$ in this inequality yields

$$\limsup_{m\to\infty}\int_0^T\int_\Omega \Lambda\nabla_{\mathcal{D}_m}u_m\cdot\nabla_{\mathcal{D}_m}u_m \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \le \int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

Together with (5.4.10), we deduce

$$\lim_{m\to\infty}\int_0^T\int_\Omega \Lambda\nabla_{\mathcal{D}_m}u_m\cdot\nabla_{\mathcal{D}_m}u_m \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t = \int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla\bar u\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

Thanks to this relation and the weak convergence of $\nabla_{\mathcal{D}_m}u_m$, we see that

$$\begin{aligned} 0 &\le \limsup_{m\to\infty}\underline{\lambda}\int_0^T\int_\Omega |\nabla\bar u - \nabla_{\mathcal{D}_m}u_m|^2\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \\ &\le \limsup_{m\to\infty}\left[\int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla\bar u + \int_0^T\int_\Omega \Lambda\nabla_{\mathcal{D}_m}u_m\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t - 2\int_0^T\int_\Omega \Lambda\nabla\bar u\cdot\nabla_{\mathcal{D}_m}u_m\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\right] = 0, \end{aligned}$$

which completes the proof.

5.5 Application to the HMM method

5.5.1 HMM method for the parabolic Signorini problem

We still use here the notions of polytopal mesh $\mathcal{T}$ of Ω, given in Definition 2.2.9. The discrete space to consider is

$$X_{\mathcal{D},\Gamma_2} = \{v = ((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}) \;:\; v_K\in\mathbb{R}, v_\sigma\in\mathbb{R}, v_\sigma = 0 \text{ for all } \sigma\in\mathcal{E}_{\rm ext} \text{ such that } \sigma\subset\Gamma_1\} \tag{5.5.1}$$

The mesh is assumed to be compatible with the boundary: for any $\sigma\in\mathcal{E}$, there is $i=1,2$ such that $\sigma\subset\Gamma_i$. The operators $\Pi_\mathcal{D}$, $\mathbb{T}_\mathcal{D}$ and $\nabla_\mathcal{D}$ are still defined by (3.3.1), (3.4.2) and (3.3.2). The discrete set $\mathcal{K}_\mathcal{D}$ is defined by

$$\mathcal{K}_\mathcal{D} := \{v\in X_{\mathcal{D},\Gamma_2} \;:\; v_\sigma \le a \text{ on } \sigma, \text{ for all } \sigma\in\mathcal{E}_{\rm ext} \text{ such that } \sigma\subset\Gamma_2\}.$$

The interpolation operator $J_\mathcal{D}$ is defined by: for all $\varphi\in\mathcal{K}$, $J_\mathcal{D}\varphi = v\in\mathcal{K}_\mathcal{D}$ is defined by

$$\begin{aligned} v_K &= \frac{1}{|K|}\int_K \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \quad \forall K\in\mathcal{M} \\ v_\sigma &= \frac{1}{|\sigma|}\int_\sigma \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \quad \forall\sigma\in\mathcal{E}. \end{aligned} \tag{5.5.2}$$

The gradient scheme (5.2.3) stemming from such a space–time GD can be written as

$$\left\{\begin{aligned} &\text{Find } (u^{(n)})_{n=0,\ldots,N}\subset\mathcal{K}_\mathcal{D} \text{ s.t., } u^{(0)} = J_\mathcal{D}\bar u_{\rm ini}, \text{ and for all } n=0,...,N-1, \text{ for all } v\in\mathcal{K}_\mathcal{D} \\ &\sum_{K\in\mathcal{M}}\frac{|K|}{\delta t^{(n+\frac12)}}\left(u_K^{(n+1)} - u_K^{(n)}\right)(u_K^{(n)} - v) + \sum_{K\in\mathcal{M}}|K|\Lambda_K\nabla_K u^{(n+1)}\cdot\nabla_K(u^{(n+1)} - v) \\ &+\sum_{K\in\mathcal{M}} R_K(u^{(n+1)} - v)^T\mathbb{B}_K R_K(u^{(n+1)}) \le \sum_{K\in\mathcal{M}}(u_K^{(n+1)} - v_K)\int_K f^{(n+1)}\,\mathrm{d}\boldsymbol{x}. \end{aligned}\right. \tag{5.5.3}$$

Here, for all $n = 0, ..., N-1$, the function $f^{(n+1)}$ is the average over $(t^{(n)}, t^{(n+1)})$ of $f(\boldsymbol{x}, \cdot)$. The matrix $\mathbb{B}_K$, Λ_K, ∇_K and R_K are defined in Section 3.2.

Proposition 5.5.1. *Let $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ be a sequence of HMM space–time gradient discretisation given by* (5.5.1), (3.3.1), (3.4.2), (3.3.2) *and* (5.5.2) *for certain polytopal meshes $(\mathcal{T}_m)_{m\in\mathbb{N}}$. Assume that there exists $\theta > 0$, not depending on m, such that* (4.5.4) *and* (4.5.5) *hold.*

Then the sequence $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ is coercive and limit-conforming in the sense of Definitions 5.2.4 and 5.2.6. If $\delta t_{\mathcal{D}_m} \to 0$ as $m \to \infty$, then $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is space–time consistent in the sense of Definition 5.2.5.

Proof. Recall that a sequence of space–time gradient discretisation $(\mathcal{D}_m, J_{\mathcal{D}_m}, (t^{(n)})_{n=0,1,...,N_m})_{m\in\mathbb{N}}$ is coercive (resp. limit-conforming) if its space-independent time $(\mathcal{D}_m)_{m\in\mathbb{N}}$ is coercive (resp. limit-conforming). Therefore, the coercivity and the limit-conformity for $(\mathcal{D}_m^T)_{m\in\mathbb{N}}$ considered here follow from Proposition 4.5.1.

It remains to prove the space–time consistency. Thanks to the density of $C^2(\overline{\Omega})\cap\mathcal{K}$ in $\mathcal{K}$, the consistency of the spatial gradient discretisation $\mathcal{D}$ also follows from Proposition 4.5.1. Let $v_m = ((v_K)_{K\in\mathcal{M}_m}, (v_\sigma)_{\sigma\in\mathcal{E}_m}) \in \mathcal{K}_{\mathcal{D}_m}$ be the interpolant such that $v_m = J_{\mathcal{D}_m}\bar{u}_{\rm ini}$. Applying [37, Estimate (B.11), in Lemma B.7] with $p = 2$, we can find C_4 not depending on m such that

$$||\bar{u}_{\rm ini}(\boldsymbol{x}) - v_K||_{L^2(K)} \le C_4 h_K ||\nabla\bar{u}_{\rm ini}(\boldsymbol{x})||_{L^2(K)}, \quad \forall K \in \mathcal{M}_m.$$

Squaring this estimate and summing this relation over $K \in \mathcal{M}_m$ gives

$$||\bar{u}_{\rm ini} - \Pi_{\mathcal{D}} J_{\mathcal{D}_m}\bar{u}_{\rm ini}||^2_{L^2(\Omega)} \le C_4^2 h^2_{\mathcal{M}_m} ||\nabla\bar{u}_{\rm ini}||^2_{L^2(\Omega)}.$$

This shows that $\lim_{m\to\infty} ||\bar{u}_{\rm ini} - \Pi_{\mathcal{D}_m} J_{\mathcal{D}_m}\bar{u}_{\rm ini}||_{L^2(\Omega)} = 0$ and concludes the proof. □

It is shown in [37, Proof of Theorem 12.12] that, for $\varphi \in W^{1,p}(\Omega)$, there exists an interpolant $v_m = ((v_K)_{K\in\mathcal{M}_m}, (v_\sigma)_{\sigma\in\mathcal{E}_m}) \in X_{\mathcal{D},0}$ defined by (5.5.2), and there is $0 < C_5$ not depending on m such that

$$||\nabla_{\mathcal{D}_m} v_m||_{L^p(\Omega)^d} \le C_5 ||\nabla\varphi||_{L^p(\Omega)^d}.$$

Applying this estimate (with $p = 2$) to $\varphi = \bar{u}_{\rm ini}$ and $v_m = J_{\mathcal{D}_m}\bar{u}_{\rm ini}$ shows that $||\nabla_{\mathcal{D}_m} J_{\mathcal{D}_m}\bar{u}_{\rm ini}||_{L^2(\Omega)^d}$ is bounded. With these properties, the convergence of the above HMM scheme is a consequence of Theorem (5.2.7).

5.5.2 HMM method for the obstacle problem

We consider the space–time gradient discretisation $\mathcal{D}^T$, whose elements $X_{\mathcal{D},0}$, $\Pi_{\mathcal{D}}$, $\nabla_{\mathcal{D}}$, $J_{\mathcal{D}}$ are respectively defined by (4.5.1), (3.3.1), (3.3.2) and (5.5.2). We define the element $\psi_{\mathcal{D}}$ as the function equal to ψ_K on K, for each $K \in \mathcal{M}$. Using the scheme (5.3.7) with this particular gradient discretisation $\mathcal{D}^T$ yields the HMM method for the parabolic obstacle problem. Setting $\mathcal{K}_{\mathcal{D}} := \{v \in X_{\mathcal{D},0} \;:\; v_K \ge \psi_K \text{ in } K, \text{ for all } K \in \mathcal{M}\}$, this method is given by

$$\left\{\begin{array}{l}
\text{Find } (u^{(n)})_{n=0,...,N} \subset \mathcal{K}_{\mathcal{D}} \text{ s.t., } J_{\mathcal{D}}\bar{u}_{\rm ini} = u^{(0)}, \text{ and for all } n = 0, ..., N-1, \text{ for all } v \in \mathcal{K}_{\mathcal{D}} \\[2mm]
\displaystyle\sum_{K\in\mathcal{M}} \frac{|K|}{\delta t^{(n+\frac{1}{2})}} \left(u_K^{(n+1)} - u_K^{(n)}\right)(u_K^{(n)} - v) + \sum_{K\in\mathcal{M}} |K|\Lambda_K \nabla_K u^{(n+1)} \cdot \nabla_K (u^{(n+1)} - v) \\[2mm]
\displaystyle + \sum_{K\in\mathcal{M}} R_K(u^{(n+1)} - v)^T \mathbb{B}_K R_K(u^{(n+1)}) \le \sum_{K\in\mathcal{M}} (u_K^{(n+1)} - v_K) \int_K f^{(n+1)}\,\mathrm{d}\boldsymbol{x},
\end{array}\right. \tag{5.5.4}$$

where $\psi_K = ⨍_K \psi\,\mathrm{d}x$ (the mean value of ψ over K).

In Section 4.5.2, we show that the sequence of HMM space-independent time $(\mathcal{D}_m)_{m\in\mathbb{N}}$ considered here is coercive, consistent and limit-conforming. The space–time GD consistency and the boundedness of the norm $||\nabla_{\mathcal{D}_m} J_{\mathcal{D}_m}\bar{u}_{\rm ini}||_{L^2(\Omega)^d}$ can exactly be verified as in Section 5.5.1. With the above definition of $\psi_{\mathcal{D}}$, we can obtain the convergence of $(\psi_{\mathcal{D}_m})_{m\in\mathbb{N}}$ to ψ in $L^2(\Omega)$ as in the proof of the convergence of $J_{\mathcal{D}_m}\bar{u}_{\rm ini}$ in the previous section. Theorem 5.3.4 provides, therefore, the convergence of the scheme (5.5.4) under the assumptions (4.5.4) and (4.5.4).

5.6 Numerical results

We present here two numerical tests to illustrate the numerical behaviour of the HMM schemes for the parabolic Signorini and obstacle problems.

In Section 3.2, we introduce the formula of the fluxes $(F_{K,\sigma}(u))_{K\in\mathcal{M},\sigma\in\mathcal{E}}$ defined for $u \in X_{\mathcal{D},\Gamma_2}$ for the Signorini problem and $u \in X_{\mathcal{D},0}$ for the obstacle problem. The HMM methods (5.5.3) and (5.5.4) can be re-written in terms of the balance and conservativity of the fluxes in the following ways:

- Parabolic Signorini problem:

$$\frac{\mathrm{m}(K)}{\delta t^{(n+\frac{1}{2})}}\left(u^{(n+1)}-u^{(n)}\right)+\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u^{(n+1)})=\mathrm{m}(K)f_K^{(n+1)},\quad \forall K\in\mathcal{M},\ \forall n=0,...,N-1,$$

$$F_{K,\sigma}(u^{(n+1)})+F_{L,\sigma}(u^{(n+1)})=0,\quad \forall\sigma\in\mathcal{E}_K\cap\mathcal{E}_L, L\neq K,\ \forall n=0,...,N-1,$$
$$u_\sigma^{(n+1)}=0,\quad \forall\sigma\in\mathcal{E}_{\rm ext}\text{ such that }\sigma\subset\Gamma_1,\ \forall n=0,...,N-1,$$
$$F_{K,\sigma}(u^{(n+1)})(a_\sigma-u_\sigma^{(n+1)})=0,\quad \forall K\in\mathcal{M},\forall\sigma\in\mathcal{E}_K,\text{ such that }\sigma\subset\Gamma_2,\ \forall n=0,...,N-1,$$
$$-F_{K,\sigma}(u^{(n+1)})\le 0,\quad \forall K\in\mathcal{M},\forall\sigma\in\mathcal{E}_K\text{ such that }\sigma\subset\Gamma_2,\ \forall n=0,...,N-1,$$
$$u_\sigma^{(n+1)}\le a_\sigma,\quad \forall\sigma\in\mathcal{E}_{\rm ext}\text{ such that }\sigma\subset\Gamma_2,\ \forall n=0,...,N-1.$$

- Parabolic obstacle problem:

$$\left(\frac{\mathrm{m}(K)}{\delta t^{(n+\frac{1}{2})}}\left(u^{(n+1)}-u^{(n)}\right)+\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u^{(n+1)})-m(K)f_K^{(n+1)}\right)(u_K^{(n+1)}-\psi_K)=0,$$
$$\forall K\in\mathcal{M},\ \forall n=0,...,N-1,$$

$$\frac{\mathrm{m}(K)}{\delta t^{(n+\frac{1}{2})}}\left(u^{(n+1)}-u^{(n)}\right)+\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}(u^{(n)})\ge m(K)f_K^{(n+1)},\quad \forall K\in\mathcal{M},\ \forall n=0,...,N-1,$$
$$u_K^{(n+1)}\ge\psi_K,\quad \forall K\in\mathcal{M},\ \forall n=0,...,N-1,$$
$$F_{K,\sigma}(u^{(n+1)})+F_{L,\sigma}(u^{(n+1)})=0,\quad \forall\sigma\in\mathcal{E}_K\cap\mathcal{E}_L, K\neq L,\ \forall n=0,...,N-1,$$
$$u_\sigma^{(n+1)}=0,\quad \forall\sigma\in\mathcal{E}_{\rm ext},\ \forall n=0,...,N-1.$$

For both problems, at any time step n, we need to solve a nonlinear system of equations. We apply the monotonicity Algorithm 1 to solve the nonlinear system for the Signorini problem and Algorithm 2 for the obstacle problem.

Test 5.6.1. We consider the Signorini problem (5.2.1), in which the domain is square $\Omega=(0,1)\times(0,1)$ with $\Gamma_2=\{0\}\times[0,1]$ and Γ_1 the remaining boundary of Ω. The time interval is $[0,T]=[0,0.5]$, the initial condition and the barrier are $\bar{u}_{\rm ini}=a=0$. The source term f is given by

$$f(t)=2(y-t)(1+\cos\pi x).$$

Figure 5.1 shows the HMM solution to this parabolic Signorini problem at $t=0.1$ (left) and at the final time $t=0.5$ (right) on a hexahedral mesh, which is of size $h=0.07$ and has 80 edges on Γ_2. In this experiment, we take as time step $\delta t=0.01$. The graph indicates that the approximate solution at $t=0.5$ switches from homogeneous Dirichlet to homogeneous Neumann around the mid-point of the boundary Γ_2 whereas there is no change in the constraints at $t=0.1$. We conduct also the test on a “Kershaw” mesh as in Figure 3.5 (right). This mesh has size $h=0.11$ and 51 edges on Γ_3. The results are shown in Figure 5.2. Although this type of mesh presents extreme distortions, the HMM scheme still captures the shift in the constraints, that happens in the case where $t=0.5$.

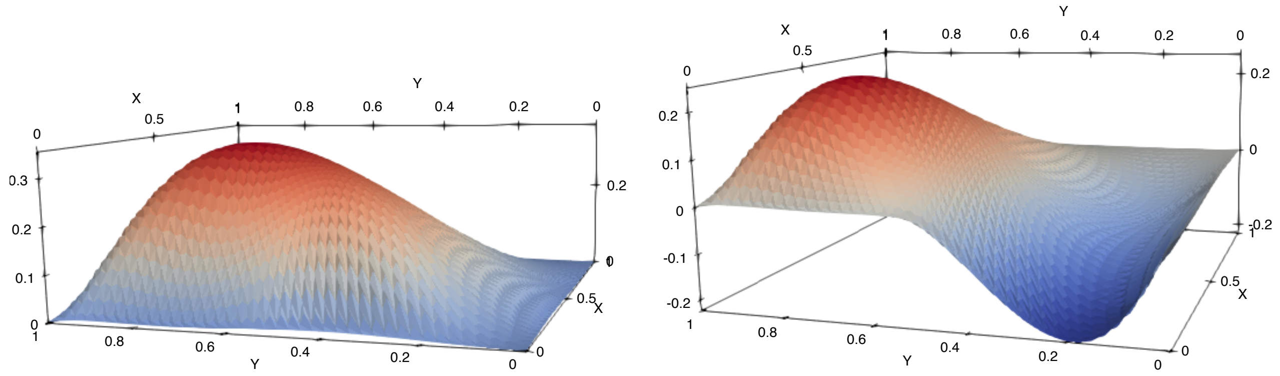

Figure 5.1: Test 5.6.1: the HMM solutions on a hexahedral mesh at $t = 0.1$ (left) and $t = 0.5$ (right).

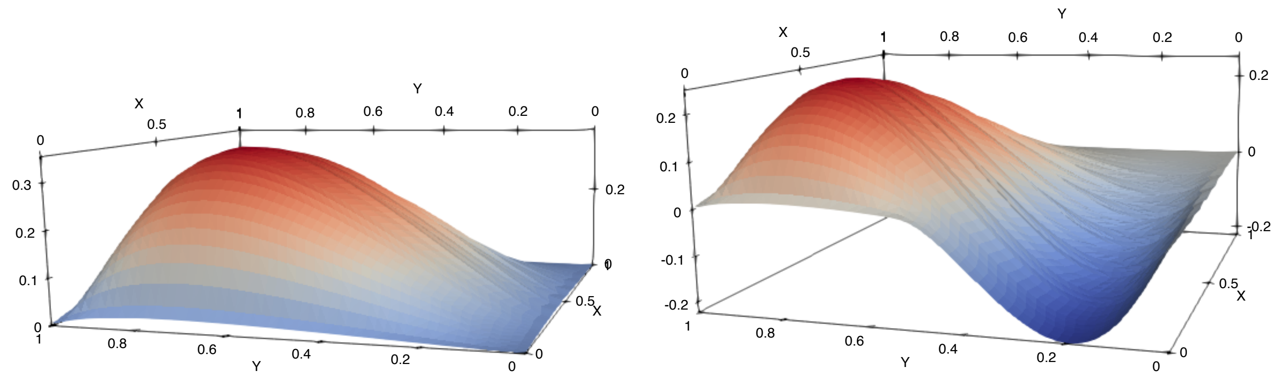

Figure 5.2: Test 5.6.1: the HMM solutions on a “Kershaw“ mesh at $t = 0.1$ (left) and $t = 0.5$ (right)

Test 5.6.2. In this experiment, we consider the obstacle problem (5.3.3) with particular data as in [19]: $\Omega = (-1, 1)^2$, $T = 0.1$, $f = -4$ and $\bar{u}_{\text{ini}} = \psi$, where the barrier function ψ is chosen such that

$$\psi(x, y) = \max\{0, -0.1 + 0.6 \exp(-10r^2), 0.5 - r\}, \quad \text{with} \quad r = \sqrt{x^2 + y^2}.$$

The test is performed on two different mesh types; a mesh consisting of 6561 hexagonal cells and a cartesian mesh with 4096 cells. The time step is $\delta t = 0.005$. Figure 5.3 presents the HMM solution to the above obstacle problem computed at the final time $T = 0.1$. The difference function $u - \psi$ at $t = 0.1$ is plotted on Figure 5.4 with the same types of meshes.

Figure 5.5 provides a presentation of the coincidence set based on a hexahedral mesh (left) and a cartesian mesh (right). The black area presents the set of cell centers where the approximate solution u reaches the barrier ψ. For the cartesian mesh, the contact regions are very similar to the ones obtained by the finite difference method in [19]. For instance, the maximum y ordinate of points $\mathbf{x} \in \Omega$, where the solution is strictly larger than the obstacle, is located around $y = 0.6$.

Solving parabolic variational inequalities in practice is more expensive than linear parabolic partial differential equations models. At each time step, we iterate to solve a number of systems of elliptic equations (see Algorithm 2). To determine the initial two sets $\mathbb{I}$ and $\mathbb{J}$ introduced in Algorithm 2, we assume that $\mathbb{I}^{(0)} = \mathcal{M}$, that is the solution is everywhere equal to the barrier at the initial step. After determining the final $\mathbb{I}$ and $\mathbb{J}$ at time $t^{(n)}$, we use these sets as initial guesses for the monotonicity algorithm at time $t^{(n+1)}$. Given that the solution to the PVI is not expected to move a lot between $t^{(n)}$ and $t^{(n+1)}$, these initial guesses are not far from the correct regions at time $t^{(n+1)}$. As a consequence, the number of iterations is reduced as the time step increases. This is illustrated in Table 5.1; from 8 iteration (starting from the guess $I^{(0)} = \mathcal{M}$) at $t^{(1)}$ to 2 iterations at time $t^{(4)}$ and after.

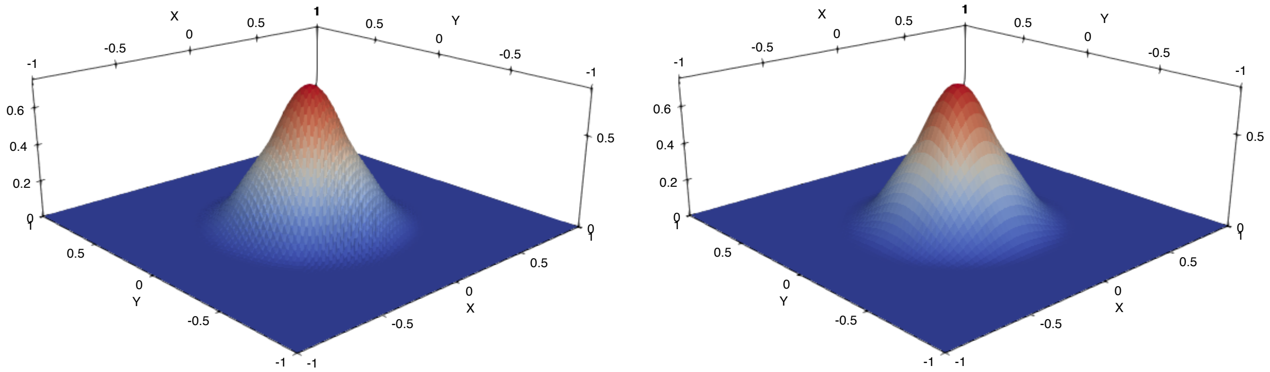

Figure 5.3: The HMM solutions for Test 5.6.2 on a hexahedral mesh (left) and on a cartesian mesh (right).

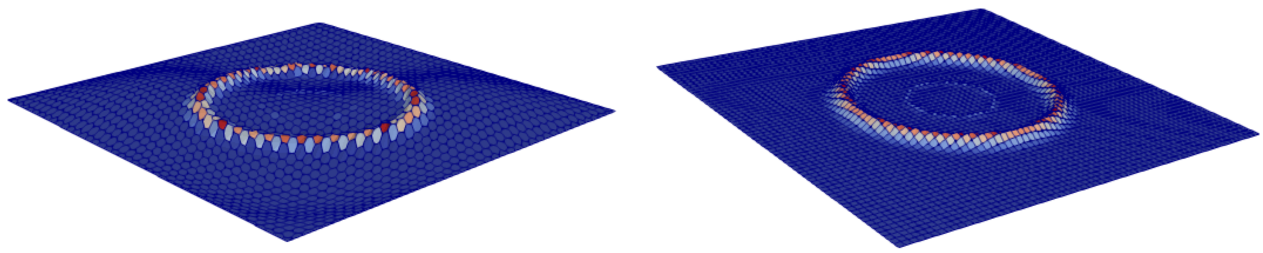

Figure 5.4: Test 5.6.2: the difference function $u - \psi$ on a hexahedral mesh (left) and on a cartesian mesh (right).

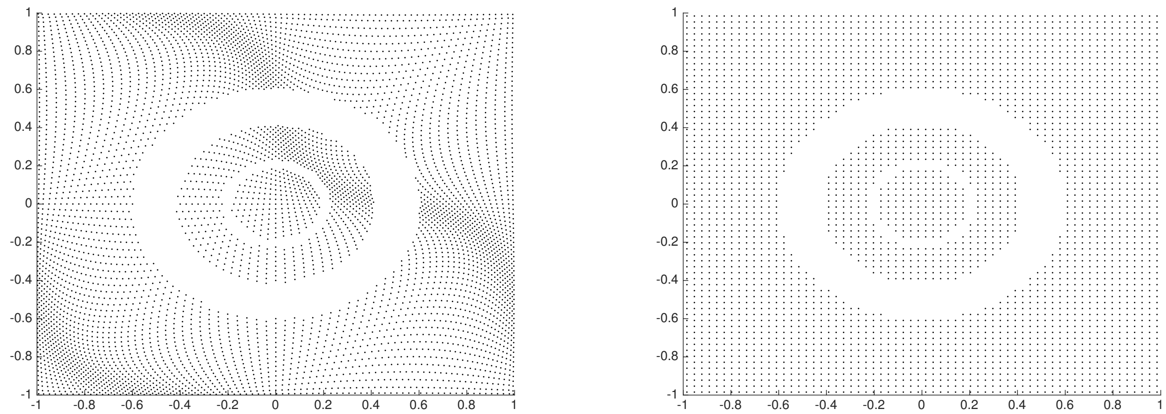

Figure 5.5: Test 5.6.2: coincidence set to the HMM solutions, a hexahedral mesh (left) and a cartesian mesh (right).

Table 5.1: Test 5.6.2: relation between time steps (n) and number of iterations of the monotonicity algorithm (NITER).

Hexahedral mesh						
Time step n	1	2	3	4	5	N
NITER	8	3	3	2	2	1
Cartesian mesh						
Time step n	1	2	3	4	5	N
NITER	6	2	2	2	2	1

Conclusion

We extended a gradient discretisation method framework to elliptic and parabolic variational inequalities. For linear elliptic problems, we established, based on the three quantities constant and functions, general error estimates and orders of convergence. These results present new, and simpler, proofs of optimal orders of convergence for some methods previously studied for variational inequalities. They also allowed us to establish new orders of convergence for recent methods, designed to deal with anisotropic heterogeneous diffusion PDEs on generic grids but not yet studied for the variational inequalities. Using the GDM framework, we provided a unified and complete convergence covering all numerical schemes contained in this framework, for nonlinear elliptic and linear parabolic variational inequalities. Based on a limit number of properties, we presented a proof that does not require uniqueness of a solution or strong assumptions on the solution.

As an application of the GDM framework, we designed an HMM method for elliptic and parabolic Signorini and obstacle problems. We also proved the convergence of the method towards the exact solution in the nonlinear elliptic and linear parabolic variational inequalities, as well as, we obtained the rate of convergence in the linear elliptic case. We introduced an easy implementation processes including the monotonicity algorithm to calculate these solutions in practice.

Through various numerical tests, we showed the efficiency of the HMM method in solving variational inequalities on generic meshes, especially, in determining the location of the seepage point in the seepage model. Tests with analytical solutions were considered to confirm the validity of the general error estimates.

Bibliography

[1] Y. Alnashri and J. Droniou, *Gradient schemes for an obstacle problem*, in Finite Volumes for Complex Applications VII-Methods and Theoretical Aspects, J. Fuhrmann, M. Ohlberger, and C. Rohde, eds., vol. 77, Springer International Publishing, 2014, pp. 67–75.

[2] Y. Alnashri and J. Droniou, *Convergence analysis of numerical schemes for non-linear variational inequalities, application to the seepage problem*, (2016). submitted.

[3] ——, *Gradient schemes for the Signorini and the obstacle problems, and application to hybrid mimetic mixed methods*, Computers and Mathematics with Applications, (2016), pp. 2788–2807.

[4] B. Andreianov, F. Boyer, and F. Hubert, *Besov regularity and new error estimates for finite volume approximations of the p-Laplacian*, Numer. Math., 100 (2005), pp. 565–592.

[5] ——, *Discrete Duality Finite Volume schemes for Leray–Lions-type elliptic problems on general 2D meshes*, Num. Meth. PDEs, 23 (2007), pp. 145–195.

[6] P. Antonietti, L. Beirão da Veiga, N. Bigoni, and M. Verani, *Mimetic finite differences for non-linear and control problems*, Mathematical Models and Methods in Applied Sciences, 24 (2014), pp. 1457–1493.

[7] C. Baiocchi, *Sur un problème à frontière libre traduisant le filtrage de liquides à travers des milieux poreux*, CR Acad. Sci. Paris, 273 (1971), pp. 1215–1217.

[8] ——, *Discretization of evolution variational inequalities*, in Partial differential equations and the calculus of variations, Springer, 1989, pp. 59–92.

[9] L. Banz and E. P. Stephan, *hp–adaptive IPDG/TDG–FEM for parabolic obstacle problems*, Computers & Mathematics with Applications, 67 (2014), pp. 712–731.

[10] J. Bardet and T. obita, *A practical method for solving free-surface seepage problems*, Computers and Geotechnics, 29 (2002), pp. 451 – 475.

[11] J. W. Barrett and W. Liu, *Finite element approximation of the p-laplacian*, mathematics of computation, 61 (1993), pp. 523–537.

[12] R. Bermejo and J.-A. Infante, *A multigrid algorithm for the p-Laplacian*, SIAM journal on scientific computing, 21 (2000), pp. 1774–1789.

[13] J. Berton and R. Eymard, *Finite volume methods for the valuation of american options*, ESAIM: Mathematical Modelling and Numerical Analysis, 40 (2006), pp. 311–330.

[14] I. Boccardo and T. Gallouët, *Nonlinear elliptic equations with right hand side measures*, Communications in Partial Differential Equations, 17 (1992), pp. 189–258.

[15] H. Brezis, *Opeérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, vol. 5, Elsevier, 1973.

[16] H. Brezis and G. Stampacchia, *Sur la régularité de la solution dínéquations elliptiques*, Bull. Soc. Math., 96 (1968), pp. pp. 153–180.

[17] F. Brezzi, W. Hager, and P. Raviart, *Error estimates for the finite element solution of variational inequalities*, Numerische Mathematik, 31 (1978), pp. 1–16.

[18] F. Brezzi, K. Lipnikov, and V. Simoncini, *A family of mimetic finite difference methods on polygonal and polyhedral meshes*, Math. Models Methods Appl. Sci., 15 (2005), pp. 1533–1551.

[19] L. Brugnano and A. Sestini, *Numerical solution of obstacle and parabolic obstacle problems based on piecewise linear systems*, in AIP Conference Proceedings, vol. 1168, AIP, 2009, pp. 746–749.

[20] ——, *Iterative solution of piecewise linear systems for the numerical solution of obstacle problems*, Journal of Numerical Analysis, Industrial and Applied Mathematics, 6 (2011), pp. 67–82.

[21] G. Capriz, *Variational techniques for the analysis of a lubrication problem*, in Mathematical Aspects of Finite Element Methods, Springer, 1977, pp. 47–55.

[22] C. Carstensen and J. Gwinner, *A theory of discretization for nonlinear evolution inequalities applied to parabolic signorini problems*, Annali di Matematica Pura ed Applicata, 177 (1999), pp. 363–394.

[23] J. Chabrowski, *On an obstacle problem for degenerate elliptic operators involving the critical sobolev exponent*, Journal of Fixed Point Theory and Applications, 4 (2008), pp. 137–150.

[24] C. Chainais-Hillairet and J. Droniou, *Convergence analysis of a mixed finite volume scheme for an elliptic-parabolic system modeling miscible fluid flows in porous media*, SIAM Journal on Numerical Analysis, 45(5) (2007), pp. 2228–2258.

[25] P. G. Ciarlet, *The finite element method for elliptic problems*, vol. 40 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, pp. xxviii+530. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[26] J. C. De los Reyes and S. G. Andrade, *Numerical simulation of two-dimensional bingham fluid flow by semismooth newton methods*, Journal of computational and applied mathematics, 235 (2010), pp. 11–32.

[27] J. C. De los Reyes and S. González Andrade, *Numerical simulation of two-dimensional Bingham fluid flow by semismooth Newton methods*, J. Comput. Appl. Math., 235 (2010), pp. 11–32.

[28] C. Desai and G. Li, *A residual flow procedure and application for free surface flow in porous media*, Advances in Water Resources, 6 (1983), pp. 27–35.

[29] D. A. Di Pietro and J. Droniou, *A hybrid high-order method for Leray–Lions elliptic equations on general meshes*, Math. Comp., (2016), p. 32p. To appear. `http://arxiv.org/abs/1508.01918`.

[30] ——, *$W^{s,p}$-approximation properties of elliptic projectors on polynomial spaces, with application to the error analysis of a hybrid high-order discretisation of Leray–Lions problems*, (2016). Submitted.

[31] J. Djoko, *Discontinuous galerkin finite element methods for variational inequalities of first and second kinds*, Numerical Methods for Partial Differential Equations, 24 (2008), pp. 296–311.

[32] S. Dongyang, G. Hongbo, et al., *A class of crouzeix-raviart type nonconforming finite element methods for parabolic variational inequality problem with moving grid on anisotropic meshes*, Hokkaido Mathematical Journal, 36 (2007), pp. 687–709.

[33] J. Droniou, *Finite volume schemes for fully non-linear elliptic equations in divergence form*, ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, 40 (2006), pp. 1069–1100.

[34] J. Droniou, *Finite volume schemes for diffusion equations: introduction to and review of modern methods*, Math. Models Methods Appl. Sci., 24 (2014), pp. 1575–1619.

[35] ———, *Introduction to discrete functional analysis techniques for the numerical study of diffusion equations with irregular data*, in Proceedings of the 17th Biennial Computational Techniques and Applications Conference (CTAC-2014, Canberra), J. Sharples and J. Bunder, eds., vol. 56 of ANZIAM J., 2015, pp. C101–C127.

[36] J. Droniou and R. Eymard, *A mixed finite volume scheme for anisotropic diffusion problems on any grid*, Numer. Math., 105 (2006), pp. 35–71.

[37] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin, *The gradient discretisation method: A framework for the discretisation and numerical analysis of linear and nonlinear elliptic and parabolic problems*, 2016. 425p, https://hal.archives-ouvertes.fr/hal-01382358, version 3.

[38] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin, *A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methodss*, Mathematical Models and Methods in Applied Sciences, 20 (2010), pp. 265–295.

[39] J. Droniou, R. Eymard, and C. Guichard, *Uniform-in-time convergence of numerical schemes for Richards' and Stefan's models*, in Finite Volumes for Complex Applications VII - Methods and Theoretical Aspects (Berlin), M. O. J. Fuhrmann and C. R. Eds., eds., vol. 77, Springer Proceedings in Mathematics and Statistics, 2014, pp. 247–254.

[40] J. Droniou, R. Eymard, and R. Herbin, *Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic problems*, Mathematical Models and Methods in Applied Sciences, 23 (2013), pp. 2395–2432.

[41] J. Droniou, R. Eymard, and R. Herbin, *Gradient schemes: generic tools for the numerical analysis of diffusion equations*, M2AN Math. Model. Numer. Anal., 50 (2016), pp. 749–781. Special issue – Polyhedral discretization for PDE.

[42] J. Droniou and B. P. Lamichhane, *Gradient schemes for linear and non-linear elasticity equations*, Numerische Mathematik, (2014), pp. 1–27.

[43] J. Droniou and K. S. Talbot, *On a miscible displacement model in porous media flow with measure data*, SIAM Journal on Mathematical Analysis, 46 (2014), pp. 3158–3175.

[44] G. Duvaut and J. Lions, *Les inéquations variationnelles en mécanique et en physique*, Dunod, Paris, (1972).

[45] G. Duvaut and J. Lions, *Inequalities in Mechanics and Physics*, Springer Berlin Heidelberg, 1976.

[46] R. Eymard, P. Feron, T. Gallouët, R. Herbin, and C. Guichard, *Gradient schemes for the Stefan problem*, International Journal On Finite Volumes, 10s (2013).

[47] R. Eymard, T. Gallouët, and R. Herbin, *Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces*, IMA J. Numer. Anal., 30 (2010), pp. 1009–1043.

[48] R. Eymard, C. Guichard, and R. Herbin, *Small-stencil 3D schemes for diffusive flows in porous media*, ESAIM. Mathematical Modelling and Numerical Analysis, 46 (2012), pp. 265–290. Copyright - EDP Sciences, SMAI, 2011; Last updated - 2013-03-05.

[49] R. Eymard, C. Guichard, R. Herbin, and R. Masson, *Gradient schemes for two-phase flow in heterogeneous porous media and Richards equation*, ZAMM Z. Angew. Math. Mech., 94 (2014), pp. 560–585.

[50] R. Falk, *Error estimates for the approximation of a class of variational inequalities*, Mathematics of Computation, 28 (1974), pp. 963–971.

[51] M. Feistauer and V. Sobotíková, *Finite element approximation of nonlinear elliptic problems with discontinuous coefficients*, RAIRO-Modélisation mathématique et analyse numérique, 24 (1990), pp. 457–500.

[52] A. Fetter, *L^∞-error estimate for an approximation of a parabolic variational inequality*, Numerische Mathematik, 50 (1987), pp. 557–565.

[53] A. Friedman, *Variational principles and free-boundary problems*, New York : Wiley, 1982.

[54] I. A. Frigaard, S. Leimgruber, and O. Scherzer, *Variational methods and maximal residual wall layers*, Journal of Fluid Mechanics, 483 (2003), pp. 37–65.

[55] S. Gerbi, R. Herbin, and E. Marchand, *Existence of a solution to a coupled elliptic system with a signorini condition*, Advances in Differential Equations, 4 (1999), pp. 225–250.

[56] R. Glowinski, *Numerical methods for non-linear variational problems*, Tata Institute of fundamental research, Bombay, 1980.

[57] R. Glowinski, J. Lions, and R. Tremolieres, *Numerical analysis of variational inequalities*, North-Holland Publishing Company, 8 ed., 1981.

[58] R. Glowinski and J. Rappaz, *Approximation of a nonlinear elliptic problem arising in a non-Newtonian fluid flow model in glaciology*, ESAIM: Math. Model Numer. Anal. (M2AN), 37 (2003), pp. 175–186.

[59] R. Herbin, *A monotonic method for the numerical solution of some free boundary value problems*, SIAM journal on numerical analysis, 40 (2002), pp. 2292–2310.

[60] R. Herbin and F. Hubert, *Benchmark on discretization schemes for anisotropic diffusion problems on general grids*, in Finite volumes for complex applications V, ISTE, London, 2008, pp. 659–692.

[61] R. Herbin and E. Marchand, *Finite volume approximation of a class of variational inequalities*, IMA Journal of Numerical Analysis, 21 (2001), pp. 553–585.

[62] P. Hild and S. Nicaise, *Residual a posteriori error estimators for contact problems in elasticity*, ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, 41 (2007), pp. 897–923.

[63] M. Hintermller and C. N. Rautenberg, *A sequential minimisation technique for elliptic quasi-variational inequalities with gradient constraints*, SIAM Journal on Optimization, 22 (2012), pp. 1224–1257.

[64] R. Hu, Y.-F. Chen, C.-B. Zhou, and H.-H. Liu, *A numerical formulation with unified unilateral boundary condition for unsaturated flow problems in porous media*, Acta Geotechnica, pp. 1–15.

[65] R. R. Huilgol and Z. You, *Application of the augmented Lagrangian method to steady pipe flows of Bingham, Casson and Herschel–Bulkley fluids*, Journal of non-newtonian fluid mechanics, 128 (2005), pp. 126–143.

[66] G. Jouvet and E. Bueler, *Steady, shallow ice sheets as obstacle problems: well-posedness and finite element approximation*, SIAM Journal on Applied Mathematics, 72 (2012), pp. 1292–1314.

[67] J. Kacur and R. Van Keer, *Solution of degenerate parabolic variational inequalities with convection*, ESAIM: Mathematical Modelling and Numerical Analysis, 37 (2003), pp. 417–431.

[68] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Society for Industrial and Applied Mathematics, 2000.

[69] M. Kleitz, L. Dessemond, R. Jimenez, F. Petitbon, R. Herbin, and E. Marchand, *Micro-modelling of the cathode and experimental approaches*, in Proc. of the second European Solid Oxide Fuel Cell Forum, Oslo, Norway, 1996.

[70] M. Křížek and P. Goittaanmäki, *Finite element approximation of variational problems and applications*, vol. 50, Longman Scientific & Technical, 1990.

[71] S. J. Lacy and J. H. Prevost, *Flow through porous media: A procedure for locating the free surface*, International Journal for Numerical and Analytical Methods in Geomechanics, 11 (1987), pp. 585–601.

[72] V. K. Le and K. Schmitt, *On boundary value problems for degenerate quasilinear elliptic equations and inequalities*, Journal of Differential Equations, 144 (1998), pp. 170 – 218.

[73] H. Lewy and G. Stampacchia, *On the regularity of the solution of a variational inequality*, Communications on Pure and Applied Mathematics, 22 (1969), pp. 153–188.

[74] J. Lions, *Quelques méthodes de résolution des problemes aux limites non linéaires*, vol. 31, Dunod Paris, 1969.

[75] J. Lions and G. Stampacchia, *Variational inequalities*, Communications on pure and applied mathematics, 20 (1967), pp. 493–519.

[76] J.-L. Lions, *Inequalities in mechanics and physics*, Springer, 1976.

[77] K. Lipnikov, G. Manzini, and D. Svyatskiy, *Analysis of the monotonicity conditions in the mimetic finite difference method for elliptic problems*, Journal of Computational Physics, 230 (2011), pp. 2620 – 2642.

[78] W. Liu and N. Yan, *Quasi-norm a priori and a posteriori error estimates for the nonconforming approximation of p-Laplacian*, Numer. Math., 89 (2001), pp. 341–378.

[79] W. B. Liu and J. W. Barrett, *Numerical methods for non-linear variational problems*, ESAIM: Mathematical Modelling and Numerical Analysis, 28 (1994), pp. 725–744.

[80] K.-S. Moon, R. H. Nochetto, T. Von Petersdorff, and C.-s. Zhang, *A posteriori error analysis for parabolic variational inequalities*, ESAIM: Mathematical Modelling and Numerical Analysis, 41 (2007), pp. 485–511.

[81] R. H. Nochetto, T. von Petersdorff, and C.-S. Zhang, *A posteriori error analysis for a class of integral equations and variational inequalities*, Numerische Mathematik, 116 (2010), pp. 519–552.

[82] J. Oden and N. Kikuchi, *Theory of variational inequalities with applications to problems of flow through porous media*, International Journal of Engineering Science, 18 (1980), pp. 1173 – 1284.

[83] F. Pérez, L. Ferragut, and J. M. Cascón, *An adaptive method for the stefan problem and its application to endoglacial conduits*, Advances in Engineering Software, 38 (2007), pp. 423–428.

[84] C. S. Ryoo and R. P. Agarwal, *Numerical inclusion methods of solutions for variational inequalities*, International Journal for Numerical Methods in Engineering, 54 (2002).

[85] D. Sankar and U. Lee, *Two-fluid Herschel–Bulkley model for blood flow in catheterized arteries*, Journal of Mechanical Science and Technology, 22 (2008), pp. 1008–1018.

[86] C. Scheven, *Gradient potential estimates in non-linear elliptic obstacle problems with measure data*, Journal of Functional Analysis, 262 (2012), pp. 2777–2832.

[87] D. Shi, C. Wang, and Q. Tang, *Anisotropic crouzeix-raviart type nonconforming finite element methods to variational inequality problem with displacement obstacle*, Journal of Computational Mathematics, 33 (2015), pp. 86–99.

[88] D. Shi and C. Xu, *Eq 1 rot nonconforming finite element approximation to signorini problem*, Science China Mathematics, 56 (2013), pp. 1301–1311.

[89] M. Struwe and J. Jost, *Variational methods, applications to nonlinear partial differential equations and hamiltonian systems*, Jahresbericht der Deutschen Mathematiker Vereinigung, 96 (1994), pp. 24–25.

[90] N. N. Ural'tseva, *Hölder continuity of gradients of solutions of parabolic equations with boundary conditions of signorini type*, in Dokl. Akad. Nauk SSSR, vol. 280, 1985, pp. 563–565.

[91] F. Wang, W. Han, and X. Cheng, *Discontinuous galerkin methods for solving elliptic variational inequalities*, SIAM Journal on Numerical Analysis, 48 (2010), pp. 708–733.

[92] L. Wang and D. Hua, *The nonconforming finite element method for signorini problem*, Journal of Computational Mathematics, 25 (2007), pp. 67–80.

[93] H. Weimin and A. Kendall, *Theoretical Numerical Analysis*, Springer New York, 3 ed., 2009.

[94] X. Yang, G. Wang, and X. Gu, *Numerical solution for a parabolic obstacle problem with nonsmooth initial data*, Numerical Methods for Partial Differential Equations, 30 (2014), pp. 1740–1754.

[95] S. Yinghe, S. Ying, and G. Xiaoming, *Free boundary problem of the 2d seepage flow*, Applied Mathematics and Mechanics, 17 (1996), pp. 549–554.

[96] H. Zheng, H. Chao Dai, and D. F. Liu, *A variational inequality formulation for unconfined seepage problems in porous media*, Applied Mathematical Modelling, 33 (2009), pp. 437 – 450.