Adaptive inexact smoothing Newton method for a nonconforming discretization of a variational inequality

Ibtihel Ben Gharbia, Joëlle Ferzly, Martin Vohralík, Soleiman Yousef

Adaptive inexact smoothing Newton method for a nonconforming discretization of a variational inequality*

Ibtihel Ben Gharbia[a], Joëlle Ferzly[a,b,c,*], Martin Vohralík[b,c], Soleiman Yousef[a]

[a] *IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France*

[b] *Inria, 2 Rue Simone Iff, 75589 Paris, France*

[c] *CERMICS, École des Ponts, 77455 Marne-la-Vallée, France*

Abstract

We develop in this work an adaptive inexact smoothing Newton method for a nonconforming discretization of a variational inequality. As a model problem, we consider the contact problem between two membranes. Discretized with the finite volume method, this leads to a nonlinear algebraic system with complementarity constraints. The non-differentiability of the arising nonlinear discrete problem a priori requests the use of an iterative linearization algorithm in the semismooth class like, e.g., the Newton-min. In this work, we rather approximate the inequality constraints by a smooth nonlinear equality, involving a positive smoothing parameter that should be drawn down to zero. This makes it possible to directly apply any standard linearization like the Newton method. The solution of the ensuing linear system is then approximated by any iterative linear algebraic solver. In our approach, we carry out an a posteriori error analysis where we introduce potential reconstructions in discrete subspaces included in $H^1(\Omega)$, as well as $\mathbf{H}(\mathrm{div},\Omega)$-conforming discrete equilibrated flux reconstructions. With these elements, we design an a posteriori estimate that provides guaranteed upper bound on the energy error between the unavailable exact solution of the continuous level and a postprocessed, discrete, and available approximation, and this at any resolution step. It also offers a separation of the different error components, namely, discretization, smoothing, linearization, and algebraic. Moreover, we propose stopping criteria and design an adaptive algorithm where all the iterative procedures (smoothing, linearization, algebraic) are adaptively stopped; this is in particular our way to fix the smoothing parameter. Finally, we numerically assess the estimate and confirm the performance of the proposed adaptive algorithm, in particular in comparison with the semismooth Newton method.

Keywords: elliptic variational inequality, complementarity constraint, semismooth and smoothing Newton method, equilibrated flux, a posteriori error estimate, stopping criteria

1. Introduction

Variational inequalities have been of great interest to researchers due to their various applications. Possibly expressed as a system of partial differential equations (PDEs) with complementarity constraints, they arise in a variety of fields such as engineering and economics [1], mathematical finance [2], structural mechanics [3], flow processes in porous media [4], and many more. The numerical discretization of such problems yields a finite-dimensional nonlinear algebraic system with complementarity constraints written in the form: find a vector $\boldsymbol{X} \in \mathbb{R}^n, n > 1$, such that

$$\mathbb{E}\boldsymbol{X} = \boldsymbol{F}, \tag{1.1a}$$

$$\boldsymbol{K}(\boldsymbol{X}) \geq \boldsymbol{0},\ \boldsymbol{G}(\boldsymbol{X}) \geq \boldsymbol{0},\ \boldsymbol{K}(\boldsymbol{X}) \cdot \boldsymbol{G}(\boldsymbol{X}) = 0. \tag{1.1b}$$

Let $0 < m < n$ be an integer. The first line (1.1a) derives from the discretization of a linear PDE, where $\mathbb{E} \in \mathbb{R}^{n-m,n}$ is a matrix and $\boldsymbol{F} \in \mathbb{R}^{n-m}$ is a given vector. Denoting by $\boldsymbol{K} : \mathbb{R}^n \to \mathbb{R}^m$ and $\boldsymbol{G} : \mathbb{R}^n \to \mathbb{R}^m$ two (linear) operators,

*This project has partly received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 647134 GATIPOR).

*Corresponding author

Email addresses: ibtihel.ben-gharbia@ifpen.fr (Ibtihel Ben Gharbia), joelle.ferzly@ifpen.fr (Joëlle Ferzly), martin.vohralik@inria.fr (Martin Vohralík), soleiman.yousef@ifpen.fr (Soleiman Yousef)

line (1.1b) expresses the complementarity relationship between the nonnegative vectors $\boldsymbol{K}(\boldsymbol{X})$ and $\boldsymbol{G}(\boldsymbol{X})$, in the sense that if one of them has a positive component, then the corresponding component in the other one must be zero. Countless developments have been made over the years to (approximately) solve problem (1.1). In this regard, we mention the semismooth Newton method [5, 6, 7, 8, 9], the active set-type methods [10], the primal-dual active set strategy which can be interpreted as a semismooth Newton method [11], and projection-type methods [12]. Another class of methods, motivated by the augmented Lagrangian methods, is the one invoking a regularization technique [13, 14]. It can be combined with a path-following strategy to properly update the regularization parameter, see, e.g., [15, 16]. Inspired from the interior-point methods [17, 18], another approach is the non-parametric interior-point method proposed recently in [19]. For an enlightening summary of numerical methods solving problem (1.1), we refer to the books of Ferris et al. [20], Facchinei and Pang [21, 22], Bonnans et al. [23], Ito and Kunisch [24], and Ulbrich [25]. Recently, we have proposed in [26] an adaptive smoothing Newton method for the resolution of nonlinear discrete problems in the form (1.1).

In this work, we consider a system of PDEs with complementarity constraints in an infinite-dimensional framework. Our goal is to estimate the overall error between the unknown PDE solution and a numerical approximation at each resolution step in an adaptive algorithm inspired from [26]. The guiding principle of the considered approach, following [26], is to approximate the complementarity constraints in (1.1b) by a system of smooth (differentiable) nonlinear equations $\boldsymbol{C}_\mu(\boldsymbol{X}) = \boldsymbol{0}$, where $\boldsymbol{C}_\mu : \mathbb{R}^n \to \mathbb{R}^m$ is a smooth (differentiable) approximation of a non-differentiable complementarity function (C-function) $\boldsymbol{C} : \mathbb{R}^n \to \mathbb{R}^m$ with a parameter $\mu > 0$. This reformulation brings us to approximate problem (1.1) at each smoothing step $j \geq 1$, with parameter $\mu^j > 0$, by finding a vector $\boldsymbol{X}^j \in \mathbb{R}^n$ such that

$$\begin{aligned} \mathbb{E}\boldsymbol{X}^j &= \boldsymbol{F}, \\ \boldsymbol{C}_{\mu^j}(\boldsymbol{X}^j) &= \boldsymbol{0}. \end{aligned} \tag{1.2}$$

Hence, any iterative linearization procedure can be directly applied to system (1.2), yielding at each linearization step $k \geq 1$ a linear system

$$\mathbb{A}_{\mu^j}^{j,k-1} \boldsymbol{X}^{j,k} = \boldsymbol{B}_{\mu^j}^{j,k-1}, \tag{1.3}$$

where $\mathbb{A}_{\mu^j}^{j,k-1} \in \mathbb{R}^{n,n}$ is a matrix and $\boldsymbol{B}_{\mu^j}^{j,k-1} \in \mathbb{R}^n$ is a vector. Let us stress, however, that it is impractical to solve (1.3) exactly in applications. Following [27, 28, 29, 30], we solve the latter system only approximately by employing an iterative linear algebraic solver, giving rise, at each smoothing step $j \geq 1$, linearization step $k \geq 1$, and linear algebraic step $i \geq 1$, to a residual vector $\boldsymbol{R}_{\text{alg}}^{j,k,i} \in \mathbb{R}^n$ defined by

$$\boldsymbol{R}_{\text{alg}}^{j,k,i} := \boldsymbol{B}_{\mu^j}^{j,k-1} - \mathbb{A}_{\mu^j}^{j,k-1} \boldsymbol{X}^{j,k,i}. \tag{1.4}$$

In this regard, as we consider numerical approximations, it is crucial to control the error between the unknown PDE solution $\boldsymbol{u}$ and the numerical approximation arising at steps j, k, i, say $\boldsymbol{u}_h^{j,k,i}$, and to approximate systems (1.2) and (1.3) efficiently and accurately while limiting the computational costs. In this respect, we remark that in [26], a posteriori error estimators were only formulated at the discrete level, addressing the error $\boldsymbol{u}_h - \boldsymbol{u}_h^{j,k,i}$ only, and yielding adaptive stopping criteria for the nonlinear and linear solvers but not for the smoothing iterations.

The present paper aims at designing an adaptive algorithm in which the algebraic, linearization, as well as smoothing iterations are adaptively stopped. Our key tool for this is to derive guaranteed a posteriori estimates allowing to obtain a fully computable upper bound on the energy error $e^{j,k,i}$ between the approximate solution $\boldsymbol{u}_h^{j,k,i}$ and the unknown solution $\boldsymbol{u}$, at each step $j \geq 1, k \geq 1$, and $i \geq 1$ of the resolution, in the form

$$e^{j,k,i} \leq \eta_{\text{disc}}^{j,k,i} + \eta_{\text{sm}}^{j,k,i} + \eta_{\text{lin}}^{j,k,i} + \eta_{\text{alg}}^{j,k,i}. \tag{1.5}$$

These computable estimates allow us to identify all sources of error resulting from the numerical simulation, namely the discretization, smoothing, linearization, and linear algebraic solver error. Distinguishing the error components in particular enables to formulate optimal criteria to adaptively stop the various iterative solvers whenever the corresponding error no longer significantly influences the behavior of the overall error, as in [31, 32, 33, 34, 35, 36, 3, 37, 38], and the references therein.

There is a well-developed literature on a posteriori error estimates for PDEs. For a general introduction, we refer for instance to the books of Ainsworth and Oden [39], Repin [40], and Verfürth [41]. For variational inequalities, we can mention the contributions of Repin [42], Belgacem et al. [43], and Bürg and Schröder [44]. In this work, we are interested in the so-called equilibrated fluxes estimates, based on $\mathbf{H}(\text{div}, \Omega)$-conforming and locally conservative flux

reconstructions belonging to the lowest-order Raviart–Thomas–Nédélec space $\mathbf{RT}_0$ (discrete subspace of $\mathbf{H}(\mathrm{div},\Omega)$). We refer the reader to the contributions [45, 46, 33]. As we consider a nonconforming, finite volume, numerical discretization, we will also rely on a potential reconstruction following in particular [47, 48, 49]. This methodology in particular allows us to obtain the unknown constant-free bound in (1.5).

We apply our approach to the following problem that models the contact between two membranes. Let $\Omega \subset \mathbb{R}^2$ be an open polygonal domain. The problem reads: find u_1, u_2, and λ such that

$$\begin{cases} -\beta_1 \Delta u_1 - \lambda = f_1 & \text{in } \Omega, & \text{(1.6a)}\\ -\beta_2 \Delta u_2 + \lambda = f_2 & \text{in } \Omega, & \text{(1.6b)}\\ u_1 - u_2 \geq 0, \quad \lambda \geq 0, \quad (u_1 - u_2)\lambda = 0 & \text{in } \Omega, & \text{(1.6c)}\\ u_1 = g & \text{on } \partial\Omega, & \text{(1.6d)}\\ u_2 = 0 & \text{on } \partial\Omega, & \text{(1.6e)} \end{cases}$$

where the unknowns are the displacements u_1 and u_2 of the two membranes and the Lagrange multiplier λ which characterizes the action, or the reaction $-\lambda$, of one membrane on the other. Equations (1.6a) and (1.6b) describe the kinematic behavior of each membrane under the action of external forces $f_1, f_2 \in L^2(\Omega)$. The constant parameters $\beta_1, \beta_2 > 0$ correspond to the tension of each membrane. Line (1.6c) represents the linear complementarity conditions, $u_1 - u_2 \geq 0$ states that the membranes cannot interpenetrate, $\lambda \geq 0$ stems from the definition of λ, and $(u_1 - u_2)\lambda = 0$ means that where the membranes are not in contact ($u_1 - u_2 > 0$), λ vanishes, and where they are in contact ($u_1 = u_2$), λ is nonnegative. The boundary conditions in (1.6d) and (1.6e) indicate that the first membrane is fixed on the boundary $\partial\Omega$ at $g > 0$, where g is a constant, above the second one, which is fixed at zero.

The contact problem (1.6) has been studied in several works. Existence and uniqueness together with a conforming finite element discretization were studied in [50, 51, 43], see also the references therein. A semismooth Newton method combined with a path-following strategy was introduced and tested in [52]. Recently, in [3], an adaptive inexact Newton method, steered by a posteriori error estimates as in (1.5), was proposed to solve problem (1.6) when discretized by conforming finite elements. In our work, we rather consider the cell-centered finite volume method. We develop an adaptive inexact smoothing Newton method to solve the arising discrete problem, where any of the classical linearization scheme for smooth nonlinearities and any iterative linear algebraic solver can be used.

Let us briefly outline the structure of the paper. In Section 2, we fix notation, present the model problem (1.6) in details, and introduce its finite volume discretization. We recall the semismooth Newton method in Section 3. Then, we introduce a smoothed reformulation of our problem and address its numerical approximation employing an (inexact) smoothing Newton method in Section 4. Next, Sections 5 and 6 are devoted to describe the potential and equilibrated flux reconstructions, enabling to pursue our analysis. In Section 7, we derive an a posteriori error estimate on the error between the exact solution and the approximate solution on any smoothing step $j \geq 1$, any linearization step $k \geq 1$, and any algebraic step $i \geq 1$. We split our guaranteed bound into estimators characterizing the discretization, smoothing, and algebraic errors, and establish a linearization estimator reflecting the linearization error, obtaining an estimate of the form (1.5). This error distinction leads to adaptive stopping criteria that we incorporate in the adaptive inexact smoothing Newton algorithm presented in Section 8. We study numerically the behavior of our a posteriori estimates and the efficiency of the developed algorithm in Section 9. Finally, Section 10 brings forth our conclusions and outlook.

2. Continuous problem and its finite volume discretization

In this section, we first fix notation and present the full and reduced variational formulations of the model problem (1.6). Then, we introduce its finite volume discretization.

2.1. Function spaces, meshes, and notation

We first recall the definition of some functional spaces. For a domain $\Omega \subset \mathbb{R}^2$, let $\mathcal{D}(\Omega)$ be the space of functions $u : \Omega \to \mathbb{R}$ of class $\mathcal{C}^\infty$ with a compact support in Ω. We denote by $L^2(\Omega)$ the space of Lebesgue-measurable functions $u : \Omega \to \mathbb{R}$ such that $\|u\| := (\int_\Omega |u(x)|^2)^{\frac{1}{2}}\,\mathrm{dx} < \infty$. It is a Hilbert space for the scalar product $(u, v) = \int_\Omega u(x)v(x)\mathrm{dx}$. Next, $H^1(\Omega)$ stands for the space of functions in $L^2(\Omega)$ which admit a weak gradient in $[L^2(\Omega)]^2$, and $H^1_0(\Omega)$ stands

for its subspace of functions that vanish on $\partial\Omega$ in the sense of traces. Moreover, $\mathbf{H}(\mathrm{div},\Omega)$ is the space of vector-valued functions $\boldsymbol{u}:\Omega\to\mathbb{R}^2, \boldsymbol{u}\in[L^2(\Omega)]^2$, such that $\nabla\cdot\boldsymbol{u}\in L^2(\Omega)$. The standard notation $\nabla\cdot$ is used for the weak divergence operator. We shall define the sets

$$H^1_g(\Omega) := \left\{u\in H^1(\Omega), u=g \text{ on } \partial\Omega\right\} \text{ and } \Lambda := \left\{\chi\in L^2(\Omega), \chi\geq 0 \text{ a.e. in } \Omega\right\}.$$

We also use in the subsequent sections the notation $\|\cdot\|^2_\omega := (\cdot,\cdot)_\omega$ for the $L^2(\omega)$ norm and scalar product on a subdomain ω of Ω. When $\omega=\Omega$, the subscript is dropped. A similar notation is used for vector-valued functions.

We shall consider a mesh $\mathcal{T}_h$ given by a family of triangles K verifying $\overline{\Omega}=\bigcup_{K\in\mathcal{T}_h}\overline{K}$. We assume that the elements of $\mathcal{T}_h$ are conforming in the sense that the intersection of the closure of two elements is either an empty set, a vertex, or an edge. We also assume that $\mathcal{T}_h$ is admissible, i.e., for all $K\in\mathcal{T}_h$, there is an associated point x_K such that the straight line connecting two points x_K and x_L of two neighboring triangles K and $L\in\mathcal{T}_h$ is orthogonal to $\sigma_{K,L}:=\partial K\cap\partial L$, see [53]; we choose for x_K the circumcenter of K. We denote by $\mathcal{E}_h$ the set of all edges σ of $\mathcal{T}_h$, by $\mathcal{E}_h^{\mathrm{int}}$ the set of interior, and by $\mathcal{E}_h^{\mathrm{ext}}$ the set of boundary edges. To each edge $\sigma\in\mathcal{E}_h$, we associate a unit normal vector $\mathbf{n}_\sigma$. The set of all edges of K is denoted by $\mathcal{E}_K$, which is decomposed into interior edges and boundary edges such that $\mathcal{E}_K=\mathcal{E}_K^{\mathrm{int}}\cup\mathcal{E}_K^{\mathrm{ext}}$. We denote by $\mathbf{n}_{K,\sigma}$ the outward unit normal vector to K on the edge σ.

We then define the broken Sobolev space $H^1(\mathcal{T}_h):=\{u\in L^2(\Omega); u|_K\in H^1(K),\forall K\in\mathcal{T}_h\}$. For a function $u\in H^1(\mathcal{T}_h)$, we denote by $\nabla u\in[L^2(\Omega)]^2$ the broken weak gradient such that $(\nabla u)|_K:=\nabla(u|_K)$.

Next, for a function u and an edge $\sigma\in\mathcal{E}_h^{\mathrm{int}}$ shared by $K,L\in\mathcal{T}_h$ such that $\mathbf{n}_\sigma$ points from K towards L, we define the jump of u on σ as

$$[\![u]\!]_\sigma := (u|_K)|_\sigma - (u|_L)|_\sigma.$$

We set $[\![u]\!]_\sigma = u|_\sigma$ for $\sigma\in\mathcal{E}_h^{\mathrm{ext}}$ in the contact of the second membrane, whereas $[\![u]\!]_\sigma = u|_\sigma - g$ for the first membrane and its approximations. Later, we will simply use the notation $[\![u]\!]$, since there will be no ambiguity, and also extend it componentwise for vector-valued variables.

We recall two basic inequalities that will be necessary in order to carry out the analysis in the following sections. Let h_ω denote the diameter of $\omega\subset\Omega$. The Poincaré–Friedrichs and the Poincaré–Wirtinger inequalities state that

$$\|u\|_\omega \leq C_{\mathrm{PF}} h_\omega \|\nabla u\|_\omega \qquad \forall u\in H^1_0(\omega), \tag{2.1a}$$

$$\|u-\overline{u}_\omega\|_\omega \leq C_{\mathrm{PW}} h_\omega \|\nabla u\|_\omega \qquad \forall u\in H^1(\omega), \tag{2.1b}$$

where $\overline{u}_\omega$ is the mean value of the function u over ω given by $\overline{u}_\omega := (u,1)_\omega/|\omega|$ ($|\omega|$ is the measure of ω). The constant C_{PF} can be taken equal to 1, cf. [54, Remark 5.8]. If ω is convex, C_{PW} can be evaluated as $1/\pi$, cf. [55], and it only depends on the geometry of ω if ω is non-convex, cf. [53, Lemma 10.4]. For a function $\boldsymbol{u}=(u_1,u_2)\in\left[H^1_0(\omega)\right]^2$, we introduce the energy semi-norm

$$|||\boldsymbol{u}|||_\omega := \left\{\sum_{\alpha=1}^2 \beta_\alpha \|\nabla u_\alpha\|^2_\omega\right\}^{\frac{1}{2}}. \tag{2.2}$$

We will use the simplified notation $|||\boldsymbol{u}||| := |||\boldsymbol{u}|||_\omega$ when $\omega=\Omega$. We extend this definition in the same way to all $\boldsymbol{u}=(u_1,u_2)\in[H^1(\mathcal{T}_h)]^2$, where it becomes merely a semi-norm. Finally, we define the rescaling of the $H^{-1}(\omega)$ norm

$$|||u|||_{H^{-1}_*(\omega)} := \sup_{\substack{\phi\in H^1_0(\omega)\\ \max(\beta_1^{\frac12},\beta_2^{\frac12})||\nabla\phi||_\omega=1}} \langle u,\phi\rangle, \qquad u\in H^{-1}(\omega). \tag{2.3}$$

2.2. Continuous problem

Setting $\boldsymbol{u}:=(u_1,u_2)$ and $\boldsymbol{v}:=(v_1,v_2)\in\left[H^1(\Omega)\right]^2$, we consider the forms, for $\chi\in L^2(\Omega)$,

$$a(\boldsymbol{u},\boldsymbol{v}) := \sum_{\alpha=1}^2 \beta_\alpha(\nabla u_\alpha,\nabla v_\alpha), \quad b(\boldsymbol{v},\chi) := (\chi, v_1-v_2), \quad l(\boldsymbol{v}) := \sum_{\alpha=1}^2 (f_\alpha, v_\alpha). \tag{2.4}$$

We will also consider in a forthcoming section the extension

$$a(\boldsymbol{u},\boldsymbol{v}) := \sum_{\alpha=1}^{2} \beta_\alpha(\nabla u_\alpha, \nabla v_\alpha) \qquad \boldsymbol{u},\boldsymbol{v} \in [H^1(\mathcal{T}_h)]^2, \tag{2.5}$$

where, recall, ∇ denotes the broken weak gradient on $H^1(\mathcal{T}_h)$.

Given $(f_1, f_2) \in [L^2(\Omega)]^2$ and $g > 0$ a constant, the weak formulation of problem (1.6) is to find $\boldsymbol{u} \in H^1_g(\Omega) \times H^1_0(\Omega)$ and $\lambda \in \Lambda$ such that

$$a(\boldsymbol{u},\boldsymbol{v}) - b(\boldsymbol{v},\lambda) = l(\boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \left[H^1_0(\Omega)\right]^2, \tag{2.6a}$$

$$b(\boldsymbol{u},\chi-\lambda) \geq 0 \qquad \forall \chi \in \Lambda. \tag{2.6b}$$

Problem (2.6) admits a unique weak solution (cf. [51, Proposition 1]).

Define then the convex set $\boldsymbol{\mathcal{K}}_g$ by

$$\boldsymbol{\mathcal{K}}_g := \left\{(v_1,v_2) \in H^1_g(\Omega) \times H^1_0(\Omega),\ v_1 - v_2 \geq 0 \text{ a.e. in } \Omega\right\}. \tag{2.7}$$

We also consider the reduced variational problem: find $\boldsymbol{u} = (u_1,u_2) \in \boldsymbol{\mathcal{K}}_g$ such that

$$a(\boldsymbol{u},\boldsymbol{v}-\boldsymbol{u}) \ \geq\ l(\boldsymbol{v}-\boldsymbol{u}) \qquad \forall \boldsymbol{v} = (v_1,v_2) \in \boldsymbol{\mathcal{K}}_g, \tag{2.8}$$

which is equivalent to (2.6), as proved in [51, Lemma 2]. Note that by the Poincaré–Friedrichs inequality (2.1a), the bilinear form a is coercive on $[H^1_0(\Omega)]^2$. Thus, the well-posedness of (2.8) is a consequence of the Lions–Stampacchia theorem, see [56, Theorem 5.6].

2.3. Finite volume discretization

The finite volume scheme for problem (1.6) reads: find the values $\{u_{1,K}\}_{K\in\mathcal{T}_h}$, $\{u_{2,K}\}_{K\in\mathcal{T}_h}$, and $\{\lambda_K\}_{K\in\mathcal{T}_h}$ such that for all $K \in \mathcal{T}_h$

$$\sum_{\sigma\in\mathcal{E}_K} F_{\alpha,K,\sigma} + (-1)^\alpha |K|\lambda_K = |K| f_{\alpha,K}, \qquad \alpha \in \{1,2\}, \tag{2.9a}$$

$$u_{1,K} - u_{2,K} \geq 0, \quad \lambda_K \geq 0, \quad (u_{1,K} - u_{2,K})\lambda_K = 0, \tag{2.9b}$$

where $f_{\alpha,K} := (f_\alpha, 1)/|K|$. In scheme (2.9), $F_{\alpha,K,\sigma}$ represents the numerical approximation of the flux through the edge σ of the element $K \in \mathcal{T}_h$ and is given by

$$F_{\alpha,K,\sigma} = \begin{cases} -\beta_\alpha |\sigma| \frac{u_{\alpha,L} - u_{\alpha,K}}{d_{K,L}} & \text{if } \sigma \in \mathcal{E}_h^{\text{int}}, \sigma = K \cap L, \\ -\beta_\alpha |\sigma| \frac{u_{\alpha,\sigma} - u_{\alpha,K}}{d_{K,\sigma}} & \text{if } \sigma \in \mathcal{E}_h^{\text{ext}}, \end{cases} \tag{2.10}$$

where for $\sigma \in \mathcal{E}_h^{\text{ext}}$, $u_{1,\sigma} = g$ and $u_{2,\sigma} = 0$, which corresponds to the discretization of the Dirichlet boundary conditions in (1.6). Let for the discretization of problem (1.6), m denotes the number of mesh elements and $n := 3m$. Using that $\mathcal{E}_K = \mathcal{E}_K^{\text{int}} \cup \mathcal{E}_K^{\text{ext}}$, we develop (2.9) and define the stiffness matrix $\mathbb{C}_\alpha \in \mathbb{R}^{m,m}, \alpha \in \{1,2\}$, by

$$\mathbb{C}_{\alpha,K,K} := \sum_{\sigma\in\mathcal{E}_K^{\text{int}}} \frac{|\sigma|}{d_{K,L}} + \sum_{\sigma\in\mathcal{E}_K^{\text{ext}}} \frac{|\sigma|}{d_{K,\sigma}}, \qquad \mathbb{C}_{\alpha,K,L} := -\frac{|\sigma|}{d_{K,L}}, \qquad K,L \in \mathcal{T}_h,\ K \neq L.$$

We also define the diagonal mass matrix $\mathbb{M} \in \mathbb{R}^{m,m}$ by $\mathbb{M}_{K,K} := |K|$, and a vector $\boldsymbol{f}_\alpha \in \mathbb{R}^m$ such that $\boldsymbol{f}_{\alpha,K} := |K| f_{\alpha,K} + \sum_{\sigma\in\mathcal{E}_K^{\text{ext}}} \beta_\alpha \frac{|\sigma|}{d_{K,\sigma}} u_{\alpha,\sigma}, \forall K \in \mathcal{T}_h$. Let $\boldsymbol{X} := [\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{\lambda}]^T \in \mathbb{R}^n$ be the algebraic vector of unknowns of the model such that $\boldsymbol{X}_1 = (u_{1,K})_{K\in\mathcal{T}_h} \in \mathbb{R}^m$, $\boldsymbol{X}_2 = (u_{2,K})_{K\in\mathcal{T}_h} \in \mathbb{R}^m$, and $\boldsymbol{\lambda} = (\lambda_K)_{K\in\mathcal{T}_h} \in \mathbb{R}^m$. Then, the finite volume discretization (2.9a) can be written as: find $\boldsymbol{X} \in \mathbb{R}^n$ such that $\mathbb{E}\boldsymbol{X} = \boldsymbol{F}$, with $\boldsymbol{F} := [\boldsymbol{f}_1, \boldsymbol{f}_2]^T \in \mathbb{R}^{n-m}$ being the right-hand side vector, and $\mathbb{E} \in \mathbb{R}^{n-m,n}$ being a rectangular block matrix defined by

$$\mathbb{E} := \begin{bmatrix} \beta_1\mathbb{C}_1 & \mathbf{0} & -\mathbb{M} \\ \mathbf{0} & \beta_2\mathbb{C}_2 & \mathbb{M} \end{bmatrix}.$$

Overall, (2.9) leads to the following system of algebraic inequalities: find $\boldsymbol{X} \in \mathbb{R}^n$ such that

$$\mathbb{E}\boldsymbol{X} = \boldsymbol{F}, \tag{2.11a}$$

$$\boldsymbol{K}(\boldsymbol{X}) \geq \boldsymbol{0}, \quad \boldsymbol{G}(\boldsymbol{X}) \geq \boldsymbol{0}, \quad \boldsymbol{K}(\boldsymbol{X}) \cdot \boldsymbol{G}(\boldsymbol{X}) = 0, \tag{2.11b}$$

where the linear operators $\boldsymbol{K} : \mathbb{R}^n \to \mathbb{R}^m$ and $\boldsymbol{G} : \mathbb{R}^n \to \mathbb{R}^m$ are defined as

$$\boldsymbol{G}(\boldsymbol{X}) := \boldsymbol{X}_1 - \boldsymbol{X}_2, \quad \text{and} \quad \boldsymbol{K}(\boldsymbol{X}) := \boldsymbol{\lambda}. \tag{2.12}$$

3. Semismooth Newton method

In this section, we consider the semismooth Newton linearization to approximate the solution of the nonlinear system of equations (2.11), see, e.g., [21, 3].

The complementarity constraints (2.11b) written as algebraic inequalities can be expressed as a nonlinear non-differentiable equality by means of C-functions, where C stands for complementarity. We say that a function $\tilde{\boldsymbol{C}} : (\mathbb{R}^m)^2 \to \mathbb{R}^m$, $m \geq 1$, is a C-function if for any pair $(\boldsymbol{x}, \boldsymbol{y}) \in (\mathbb{R}^m)^2$,

$$\tilde{\boldsymbol{C}}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0} \iff \boldsymbol{x} \geq \boldsymbol{0}, \quad \boldsymbol{y} \geq \boldsymbol{0}, \quad \text{and} \quad \boldsymbol{x} \cdot \boldsymbol{y} = 0.$$

As examples, we consider the min and Fischer–Burmeister (F–B) functions

$$\big(\tilde{\boldsymbol{C}}_{\min}(\boldsymbol{x}, \boldsymbol{y})\big)_l := (\min\{\boldsymbol{x}, \boldsymbol{y}\})_l = (\boldsymbol{x}_l + \boldsymbol{y}_l)/2 - |\boldsymbol{x}_l - \boldsymbol{y}_l|/2 \qquad l = 1, \ldots, m, \tag{3.1}$$

$$\big(\tilde{\boldsymbol{C}}_{\mathrm{FB}}(\boldsymbol{x}, \boldsymbol{y})\big)_l := \sqrt{\boldsymbol{x}_l^2 + \boldsymbol{y}_l^2} - (\boldsymbol{x}_l + \boldsymbol{y}_l) \qquad l = 1, \ldots, m. \tag{3.2}$$

For more details on C-functions see [21, 22]. Let us consider a function $\boldsymbol{C} : \mathbb{R}^n \to \mathbb{R}^m$ defined as $\boldsymbol{C}(\boldsymbol{X}) := \tilde{\boldsymbol{C}}\left(\boldsymbol{K}(\boldsymbol{X}), \boldsymbol{G}(\boldsymbol{X})\right)$, where $\tilde{\boldsymbol{C}}$ is any C-function and $\boldsymbol{K}(\cdot), \boldsymbol{G}(\cdot)$ are given in (2.12). This allows to conveniently state constraints (2.11b) in an equality of the form $\boldsymbol{C}(\boldsymbol{X}) = \boldsymbol{0}$. Then, problem (2.11) can be equivalently rewritten as a system of nonlinear algebraic equations: find a vector $\boldsymbol{X} \in \mathbb{R}^n$ such that

$$\mathbb{E}\boldsymbol{X} = \boldsymbol{F}, \tag{3.3a}$$

$$\boldsymbol{C}(\boldsymbol{X}) = \boldsymbol{0}. \tag{3.3b}$$

Note, however, that in general C-functions are not Fréchet-differentiable everywhere.

Next, we detail the semismooth Newton linearization of problem (3.3). Let an initial vector $\boldsymbol{X}^0 \in \mathbb{R}^n$ be given. At the step $k \geq 1$, one looks for $\boldsymbol{X}^k \in \mathbb{R}^n$ such that

$$\mathbb{A}^{k-1}\boldsymbol{X}^k = \boldsymbol{B}^{k-1}, \tag{3.4}$$

where the Jacobian matrix $\mathbb{A}^{k-1} \in \mathbb{R}^{n,n}$ and the right-hand side vector $\boldsymbol{B}^{k-1} \in \mathbb{R}^n$ are given by

$$\mathbb{A}^{k-1} := \begin{bmatrix} \mathbb{E} \\ \mathbb{J}_{\boldsymbol{C}}(\boldsymbol{X}^{k-1}) \end{bmatrix}, \quad \boldsymbol{B}^{k-1} := \begin{bmatrix} \boldsymbol{F} \\ \mathbb{J}_{\boldsymbol{C}}(\boldsymbol{X}^{k-1})\boldsymbol{X}^{k-1} - \boldsymbol{C}(\boldsymbol{X}^{k-1}) \end{bmatrix}. \tag{3.5}$$

We emphasize that equation (3.3a) is linear and a semismooth nonlinearity occurs in the second line (3.3b). In (3.5), $\mathbb{J}_{\boldsymbol{C}}(\boldsymbol{X}^{k-1})$ stands for the Jacobian matrix in the sense of Clarke of the semismooth function $\boldsymbol{C}$ at point $\boldsymbol{X}^{k-1}$, cf. [21, 22]. To give an example, we consider the semismooth min function (3.1) at $\boldsymbol{X}^{k-1}$

$$\boldsymbol{C}(\boldsymbol{X}^{k-1}) = \min\{\boldsymbol{X}_1^{k-1} - \boldsymbol{X}_2^{k-1}, \boldsymbol{\lambda}^{k-1}\} = \min\left\{ \begin{pmatrix} u_{1,K_1}^{k-1} - u_{2,K_1}^{k-1} \\ \vdots \\ u_{1,K_m}^{k-1} - u_{2,K_m}^{k-1} \end{pmatrix}, \begin{pmatrix} \lambda_{K_1}^{k-1} \\ \vdots \\ \lambda_{K_m}^{k-1} \end{pmatrix} \right\}.$$

We define the block matrices $\mathbb{G}$ and $\mathbb{K} \in \mathbb{R}^{m,n}$ by $\mathbb{G} = [\mathbb{I}_{m\times m}, -\mathbb{I}_{m\times m}, \boldsymbol{0}_{m\times m}]$ and $\mathbb{K} = [\boldsymbol{0}_{m\times m}, \boldsymbol{0}_{m\times m}, \mathbb{I}_{m\times m}]$. Then, the l^{th} row of the Jacobian matrix in the sense of Clarke $\mathbb{J}_{\boldsymbol{C}}(\boldsymbol{X}^{k-1})$ is either given by the l^{th} row of $\mathbb{G}$, if $u_{1,K_l}^{k-1} - u_{2,K_l}^{k-1} \leq \lambda_{K_l}^{k-1}$, or by the l^{th} row of $\mathbb{K}$, if $\lambda_{K_l}^{k-1} < u_{1,K_l}^{k-1} - u_{2,K_l}^{k-1}$.

4. Inexact smoothing Newton method

We now address the numerical approximation of the nonsmooth nonlinear problem (3.3) employing a smoothing approach.

4.1. Discrete smoothed problem

We replace $\boldsymbol{C}(\cdot)$ in problem (3.3) by a smoothed C-function $\boldsymbol{C}_\mu(\cdot)$ of class $\mathcal{C}^1$, where $\mu > 0$ is a (small) smoothing parameter. A possible smoothing of the functions (3.1) and (3.2) can be, respectively: for $l = 1, \ldots, m$,

$$\big(\tilde{\boldsymbol{C}}_{\min_\mu}(\boldsymbol{x},\boldsymbol{y})\big)_l = \frac{\boldsymbol{x}_l + \boldsymbol{y}_l}{2} - \frac{\Big(|\boldsymbol{x}-\boldsymbol{y}|_\mu\Big)_l}{2} \qquad \text{with } \big(|\boldsymbol{z}|_\mu\big)_l := \sqrt{\boldsymbol{z}_l^2 + \mu^2}, \tag{4.1}$$

$$\big(\tilde{\boldsymbol{C}}_{\mathrm{FB}_\mu}(\boldsymbol{x},\boldsymbol{y})\big)_l = \sqrt{\mu^2 + \boldsymbol{x}_l^2 + \boldsymbol{y}_l^2} - (\boldsymbol{x}_l + \boldsymbol{y}_l), \tag{4.2}$$

where the μ-smoothed absolute value function $|\cdot|_\mu : \mathbb{R}^m \to \mathbb{R}^m_+$, $m \geq 0$, replaces the absolute value function (not differentiable at $\mathbf{0}$). Note that both functions $\tilde{\boldsymbol{C}}_{\min,\mu}$ and $\tilde{\boldsymbol{C}}_{\mathrm{FB},\mu}$ are of class $\mathcal{C}^\infty$.

We now introduce a smoothing loop with index $j \geq 1$, where $\mu_j > 0$ is a (decreasing) sequence of smoothing parameters. The discrete smoothed problem at each outer smoothing step $j \geq 1$ then reads as follows: find $\boldsymbol{X}^j \in \mathbb{R}^n$ such that

$$\begin{aligned} \mathbb{E}\boldsymbol{X}^j &= \boldsymbol{F}, \\ \boldsymbol{C}_{\mu^j}(\boldsymbol{X}^j) &= \mathbf{0}, \end{aligned} \tag{4.3}$$

with $\boldsymbol{C}_{\mu^j}(\boldsymbol{X}^j) := \tilde{\boldsymbol{C}}_{\mu^j}\left(\boldsymbol{K}(\boldsymbol{X}^j), \boldsymbol{G}(\boldsymbol{X}^j)\right)$. This approach gives rise to the nonlinear algebraic system (4.3) at each smoothing step $j \geq 1$, which is differentiable. Its solution is approximated employing the (inexact) Newton method detailed next.

4.2. Newton linearization

Let $j \geq 1$ be fixed and let $\boldsymbol{X}^{j,0}$ be a given initial vector. At each linearization iteration $k \geq 1$, the new approximation $\boldsymbol{X}^{j,k} \in \mathbb{R}^n$ is obtained solving the linear problem written as

$$\mathbb{A}_{\mu^j}^{j,k-1} \boldsymbol{X}^{j,k} = \boldsymbol{B}_{\mu^j}^{j,k-1}, \tag{4.4}$$

where the Jacobian matrix $\mathbb{A}_{\mu^j}^{j,k-1} \in \mathbb{R}^{n,n}$ and the right-hand side vector $\boldsymbol{B}_{\mu^j}^{j,k-1} \in \mathbb{R}^n$ are defined by

$$\mathbb{A}_{\mu^j}^{j,k-1} := \begin{bmatrix} \mathbb{E} \\ \mathbb{J}_{\boldsymbol{C}_{\mu^j}}(\boldsymbol{X}^{j,k-1}) \end{bmatrix}, \quad \boldsymbol{B}_{\mu^j}^{j,k-1} := \begin{bmatrix} \boldsymbol{F} \\ \mathbb{J}_{\boldsymbol{C}_{\mu^j}}(\boldsymbol{X}^{j,k-1})\boldsymbol{X}^{j,k-1} - \boldsymbol{C}_{\mu^j}(\boldsymbol{X}^{j,k-1}) \end{bmatrix}, \tag{4.5}$$

with $\mathbb{J}_{\boldsymbol{C}_{\mu^j}}(\boldsymbol{X}^{j,k-1})$ the standard Jacobian matrix of the smooth function $\boldsymbol{C}_{\mu^j}$ at $\boldsymbol{X}^{j,k-1}$.

4.3. Algebraic resolution

The system of linear algebraic equations (4.4) is typically numerically addressed using an iterative algebraic solver. For a fixed smoothing step $j \geq 1$, a fixed Newton step $k \geq 1$, and a given initial vector $\boldsymbol{X}^{j,k,0}$ (typically, $\boldsymbol{X}^{j,k,0} = \boldsymbol{X}^{j,k-1,\bar{i}}$, the last iterate available from the previous linearization step), the iterative solver generates for $i \geq 1$ (inner loop in k) a sequence $\boldsymbol{X}^{j,k,i}$ approximating $\boldsymbol{X}^{j,k}$ from (4.4) up to the residual given by

$$\boldsymbol{R}_{\mathrm{alg}}^{j,k,i} := \boldsymbol{B}_{\mu^j}^{j,k-1} - \mathbb{A}_{\mu^j}^{j,k-1} \boldsymbol{X}^{j,k,i}. \tag{4.6}$$

Detailing the first two equations of (4.6), we obtain for $\alpha \in \{1,2\}$, at smoothing iteration $j \geq 1$, Newton iteration $k \geq 1$, and linear solver iteration $i \geq 1$, the residual $\boldsymbol{R}_{\mathrm{alg},\alpha,K}^{j,k,i}$ given by

$$\Big(\boldsymbol{R}_{\mathrm{alg},\alpha}^{j,k,i}\Big)_K := |K| f_{\alpha,K} - \sum_{\sigma \in \mathcal{E}_K} F_{\alpha,K,\sigma}^{j,k,i} - (-1)^\alpha |K| \lambda_K^{j,k,i}, \tag{4.7}$$

where $\boldsymbol{R}^{j,k,i}_{\mathrm{alg},\alpha,K}$ is the algebraic residual associated to the element $K \in \mathcal{T}_h, \alpha \in \{1,2\}$, and $F^{j,k,i}_{\alpha,K,\sigma}$ is given by

$$F^{j,k,i}_{\alpha,K,\sigma} := \begin{cases} -\beta_\alpha|\sigma|\frac{u^{j,k,i}_{\alpha,L}-u^{j,k,i}_{\alpha,K}}{d_{K,L}} & \text{if } \sigma \in \mathcal{E}^{\mathrm{int}}_h, \sigma = K \cap L, \\ -\beta_\alpha|\sigma|\frac{u_{\alpha,\sigma}-u^{j,k,i}_{\alpha,K}}{d_{K,\sigma}} & \text{if } \sigma \in \mathcal{E}^{\mathrm{ext}}_h. \end{cases} \tag{4.8}$$

5. Postprocessing of the approximate solution and potential reconstructions

This section introduces $H^1(\Omega)$-conforming reconstructed potentials that will be central in the formulation of our posteriori error estimates.

5.1. Postprocessed potential

The discrete finite volume solution from (2.9) or more precisely from (4.6) is only piecewise constant, see Figure 1, left, for an illustration in one space dimension. Recall that it is defined for all $K \in \mathcal{T}_h$ and $\alpha \in \{1,2\}$ by $u^{j,k,i}_{\alpha h}|_K := u^{j,k,i}_{\alpha,K}$ and $\lambda^{j,k,i}_h|_K := \lambda^{j,k,i}_K$. In particular, setting $\boldsymbol{u}^{j,k,i}_h := (u^{j,k,i}_{1h}, u^{j,k,i}_{2h})$, the discrete solution is such that

$$\begin{aligned} \boldsymbol{u}^{j,k,i}_h &\notin \boldsymbol{\mathcal{K}}_g, \\ -\beta_\alpha \nabla u^{j,k,i}_{\alpha h} &\notin \mathbf{H}(\mathrm{div},\Omega), \qquad \alpha \in \{1,2\}, \\ \nabla\cdot(-\beta_\alpha \nabla u^{j,k,i}_{\alpha h}) &\neq f_\alpha - (-1)^\alpha \lambda^{j,k,i}_h, \qquad \alpha \in \{1,2\}. \end{aligned}$$

In the subsequent sections, we try to mimic the above properties, satisfied by of the weak solution $\boldsymbol{u}$, by building reconstructions from the discrete approximate solution $\boldsymbol{u}^{j,k,i}_h$.

Let $\mathbb{P}_p(K), p \geq 0$, denote the set of polynomials of total degree at most p on the element $K \in \mathcal{T}_h$. First, to be able to evaluate the (broken) gradient of the approximate solution and to measure its distance to the exact solution by the energy (semi-)norm defined in (2.2), it is primordial to transform the piecewise constant solution $\boldsymbol{u}^{j,k,i}_h$ into a higher-order piecewise polynomial. To do so, we locally construct a postprocessed approximation $\tilde{\boldsymbol{u}}^{j,k,i}_h$ that lies in $[\mathbb{P}_2(\mathcal{T}_h)]^2$, the space of piecewise second-order polynomials, following [57, 48].

Definition 5.1 (Postprocessed solution)**.** *We introduce the piecewise quadratic, discontinuous, postprocessed solution* $\tilde{\boldsymbol{u}}^{j,k,i}_h := \left(\tilde{u}^{j,k,i}_{1h}, \tilde{u}^{j,k,i}_{2h}\right) \in [\mathbb{P}_2(\mathcal{T}_h)]^2$ *as follows. Let* $F^{j,k,i}_{\alpha,K,\sigma}$ *be given by* (4.8)*. For* $\alpha \in \{1,2\}$*, let*

$$\frac{\left(\tilde{u}^{j,k,i}_{\alpha h}, 1\right)_K}{|K|} = u^{j,k,i}_{\alpha,K}, \tag{5.1a}$$

$$-\beta_\alpha \nabla \tilde{u}^{j,k,i}_{\alpha h} \in (\mathbb{P}_0(K))^2 + x\mathbb{P}_0(K), \quad -\beta_\alpha \nabla \tilde{u}^{j,k,i}_{\alpha h}|_K \cdot \mathbf{n}_{K,\sigma} = \frac{F^{j,k,i}_{\alpha,K,\sigma}}{|\sigma|} \qquad \forall \sigma \in \mathcal{E}_K. \tag{5.1b}$$

Figure 1, right part, gives an illustration of this postprocessed solution. Condition (5.1a) states that the mean value on each mesh element of the postprocessed solution is given by the original solution, whereas (5.1b) fixes the flux $-\beta_\alpha \nabla \tilde{u}^{j,k,i}_{\alpha h}$ to be in the lowest-order Raviart–Thomas space and its normal component to coincide with the finite volume edge fluxes.

5.2. Non-admissible potential reconstruction

The postprocessed solution $\tilde{\boldsymbol{u}}^{j,k,i}_h$ of Definition 5.1 is not included in the convex space $\boldsymbol{\mathcal{K}}_g$, already by the fact that it does not lie in $H^1_g(\Omega) \times H^1_0(\Omega)$. We will therefore introduce a continuous reconstructed solution $\boldsymbol{s}_h$ that can still be nonphysical, in the sense that it may not satisfy the complementarity constraints, and thus not lie in $\boldsymbol{\mathcal{K}}_g$, but at least it lies in $H^1_g(\Omega) \times H^1_0(\Omega)$.

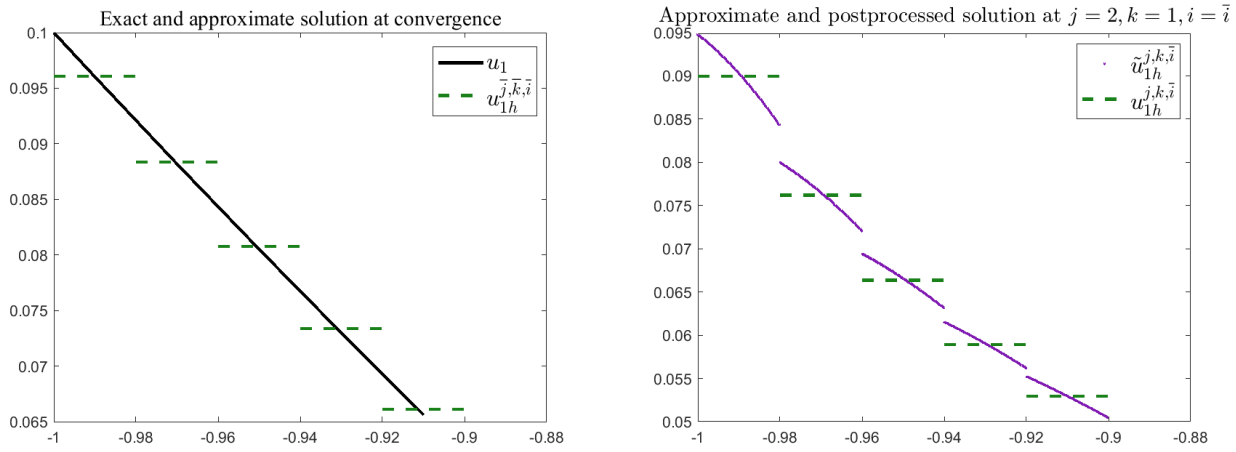

Figure 1: [Adaptive inexact smoothing Newton method, Algorithm 1, one space dimension, zoom on the first 5 elements of the computational mesh $\mathcal{T}_h$] Left: exact solution u_1 and approximate solution $u_{1h}^{\bar{j},\bar{k},\bar{i}}$ at convergence of all solvers. Right: Approximate solution $u_{1h}^{j,k,\bar{i}}$ and postprocessed solution $\tilde{u}_{1h}^{j,k,\bar{i}}$ at steps $(j,k)=(2,1)$ and at convergence of the algebraic solver $(i=\bar{i})$.

Notations. Let X_h^p, $p \geq 1$, stand for the discrete conforming space of piecewise polynomial functions

$$X_h^p := \left\{ v_h \in \mathcal{C}^0(\overline{\Omega});\ v_h|_K \in \mathbb{P}_p(K),\ \forall K \in \mathcal{T}_h \right\} \subset H^1(\Omega). \tag{5.2}$$

We will in the sequel also need the boundary-aware set and space

$$X_{gh}^p := \{ v_h \in X_h^p;\ v_h = g \text{ on } \partial\Omega \} \subset H_g^1(\Omega) \quad \text{and} \quad X_{0h}^p := X_h^p \cap H_0^1(\Omega) \subset H_0^1(\Omega). \tag{5.3}$$

Definition 5.2 (Non-admissible potential reconstruction)**.** *We introduce* $\boldsymbol{s}_h^{j,k,i} := (s_{1h}^{j,k,i}, s_{2h}^{j,k,i})$, *given by, for* $\alpha \in \{1,2\}$,

$$\boldsymbol{s}_h^{j,k,i} := \boldsymbol{\mathcal{I}}_{\mathrm{Os}}(\tilde{\boldsymbol{u}}_h^{j,k,i}) := \left(\mathcal{I}_{\mathrm{Os}}(\tilde{u}_{1h}^{j,k,i}), \mathcal{I}_{\mathrm{Os}}(\tilde{u}_{2h}^{j,k,i}) \right), \tag{5.4}$$

where $\mathcal{I}_{\mathrm{Os}}$ *denotes the Oswald interpolation operator previously considered in, e.g., [48]. This operator associates to the discontinuous piecewise polynomial* $\tilde{u}_{\alpha h}^{j,k,i}, \alpha \in \{1,2\}$, *its conforming interpolant, i.e., continuous and contained in* $H^1(\Omega)$, *by taking averages in all Lagrangian evaluation points and fixing the boundary values to respectively* g *or* 0. *Figure 2 illustrates the postprocessed and the reconstructed solution at a specific smoothing and linearization iterations (left) and at convergence (right). The reconstructed solution is then piecewise second-order polynomial, continuous, and satisfies*

$$\boldsymbol{s}_h^{j,k,i} := (s_{1h}^{j,k,i}, s_{2h}^{j,k,i}) \in X_{gh}^2 \times X_{0h}^2 \subset H_g^1(\Omega) \times H_0^1(\Omega).$$

5.3. Admissible potential reconstruction

It may happen that the potential reconstruction $\boldsymbol{s}_h^{j,k,i}$ defined by (5.4) violates the non-penetration condition $s_{1h}^{j,k,i} - s_{2h}^{j,k,i} \geq 0$, see Figure 3, so that $\boldsymbol{s}_h^{j,k,i} \notin \boldsymbol{\mathcal{K}}_g$, where we recall $\boldsymbol{\mathcal{K}}_g$ is given in (2.7). In order to avoid this, we build from the potential reconstruction $\boldsymbol{s}_h^{j,k,i} \in X_{gh}^2 \times X_{0h}^2 \not\subset \boldsymbol{\mathcal{K}}_g$, a final admissible potential reconstruction $\tilde{\boldsymbol{s}}_h^{j,k,i} \in \boldsymbol{\mathcal{K}}_g$, $\tilde{\boldsymbol{s}}_h^{j,k,i} \in X_{gh}^3 \times X_{0h}^3$. We now provide details on how to build it.

Definition 5.3 (Admissible potential reconstruction)**.** *We employ the following possible procedure, which is composed of two steps:*

Step 1. First, we construct $\hat{\boldsymbol{s}}_h^{j,k,i} \in X_{gh}^2 \times X_{0h}^2 \subset H_g^1(\Omega) \times H_0^1(\Omega)$ *such that for each Lagrangian evaluation node* $\boldsymbol{a}$

$$\hat{\boldsymbol{s}}_h^{j,k,i}(\boldsymbol{a}) := \begin{cases} \left(s_{1h}^{j,k,i}(\boldsymbol{a}), s_{2h}^{j,k,i}(\boldsymbol{a}) \right) & \text{if } s_{1h}^{j,k,i}(\boldsymbol{a}) \geq s_{2h}^{j,k,i}(\boldsymbol{a}), \\ \left(\frac{1}{2}\left(s_{1h}^{j,k,i}(\boldsymbol{a}) + s_{2h}^{j,k,i}(\boldsymbol{a}) \right), \frac{1}{2}\left(s_{1h}^{j,k,i}(\boldsymbol{a}) + s_{2h}^{j,k,i}(\boldsymbol{a}) \right) \right) & \text{if } s_{1h}^{j,k,i}(\boldsymbol{a}) < s_{2h}^{j,k,i}(\boldsymbol{a}). \end{cases} \tag{5.5}$$

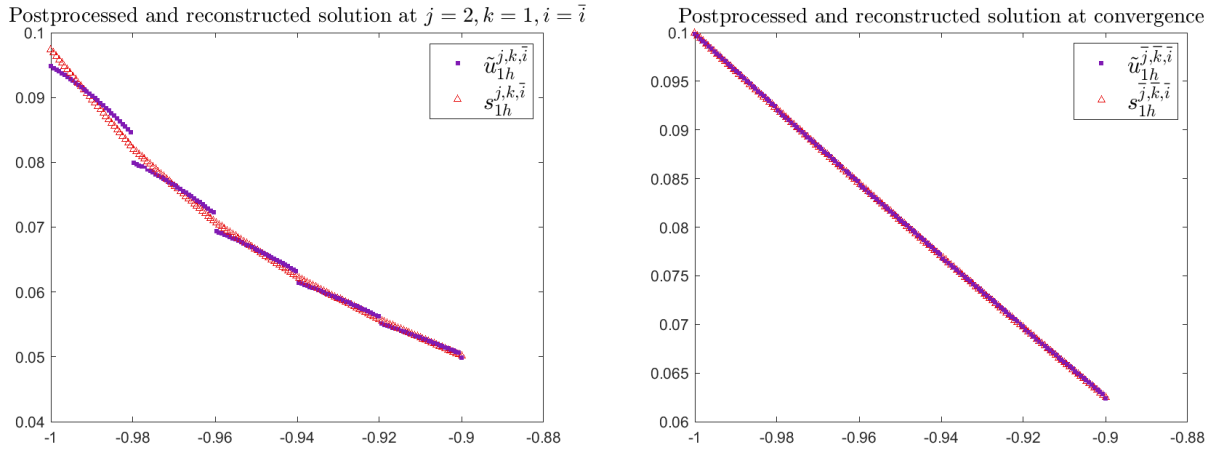

Figure 2: [Adaptive inexact smoothing Newton method, Algorithm 1, one space dimension, zoom on the first 5 elements of the computational mesh $\mathcal{T}_h$] Postprocessed solution $\tilde{u}_{1h}^{j,k,\bar{i}}$ and reconstructed solution $s_{1h}^{j,k,\bar{i}}$ at steps $(j,k) = (2,1)$ and at convergence of the algebraic solver $(i = \bar{i})$, left. Postprocessed solution $\tilde{u}_{1h}^{\bar{j},\bar{k},\bar{i}}$ and reconstructed solution $s_{1h}^{\bar{j},\bar{k},\bar{i}}$ at convergence of all solvers, right.

Step 2. We point out that even if the inequality $(\hat{s}_{1h}^{j,k,i} - \hat{s}_{2h}^{j,k,i})(\boldsymbol{a}) \geq 0$ *is satisfied by the above first construction step for all Lagrangian nodes* $\boldsymbol{a}$*, this does not necessarily imply that* $\hat{s}_{1h}^{j,k,i} \geq \hat{s}_{2h}^{j,k,i}$ *everywhere, see the left part of Figure 4. To guarantee the requested property, we proceed as follows:*

a) First, go through all internal edges $\sigma \in \mathcal{E}^{\rm int}$ *of the mesh* $\mathcal{T}_h$*. Consider the second-degree polynomial* $\hat{s}_\sigma := (\hat{s}_{1h}^{j,k,i} - \hat{s}_{2h}^{j,k,i})|_\sigma$ *on the edge* σ*. If* $\hat{s}_\sigma \geq 0$*, i.e.* $\hat{s}_\sigma$ *is nonnegative over* σ*, set* $c_\sigma := 0$*. Otherwise,* $\hat{s}_\sigma$ *takes negative values inside* σ*. Let* ω_σ *be the subdomain formed by the two triangles that share the edge* σ*. Consider the edge bubble function* ψ_σ*, a non-negative piecewise second-order polynomial defined over* ω_σ*, continuous over* σ*, zero on* $\partial\omega_\sigma$*, with* $\|\psi_\sigma\|_{\infty,\omega_\sigma} = 1$*. Let* c_σ *be the smallest positive constant such that* $(\hat{s}_\sigma + c_\sigma\psi_\sigma|_\sigma) \geq 0$ *on* σ*.*

b) Second, go through all elements K *of* $\mathcal{T}_h$*. Consider the second-degree polynomial* $\hat{s}_K := (\hat{s}_{1h}^{j,k,i} - \hat{s}_{2h}^{j,k,i})|_K + (\sum_{\sigma\in\mathcal{E}_K^{\rm int}} c_\sigma\psi_\sigma)|_K$ *on the triangle* K*. If* $\hat{s}_K \geq 0$*, set* $c_K := 0$*. Otherwise, consider the element bubble function* ψ_K*, a non-negative third-order polynomial defined over* K*, zero on* ∂K*, with* $\|\psi_K\|_{\infty,K} = 1$*. Let* c_K *be the smallest positive constant such that* $\hat{s}_K + c_K\psi_K \geq 0$ *on the element* K*.*

c) The last step of our construction is to define $\tilde{\boldsymbol{s}}_h^{j,k,i}$*, for* $\alpha \in \{1,2\}$*, by*

$$\tilde{s}_{\alpha h}^{j,k,i} := \hat{s}_{\alpha h}^{j,k,i} - (-1)^\alpha \frac{1}{2} \sum_{\sigma\in\mathcal{E}_h^{\rm int}} c_\sigma\psi_\sigma - (-1)^\alpha \frac{1}{2} \sum_{K\in\mathcal{T}_h} c_K\psi_K. \tag{5.6}$$

This yields

$$\tilde{\boldsymbol{s}}_h^{j,k,i} \in X_{gh}^3 \times X_{0h}^3 \subset H_g^1(\Omega) \times H_0^1(\Omega), \quad \text{with } \tilde{s}_{1h}^{j,k,i} \geq \tilde{s}_{2h}^{j,k,i},$$

so that

$$\tilde{\boldsymbol{s}}_h^{j,k,i} \in \boldsymbol{\mathcal{K}}_g.$$

An illustration of the two steps described above is given in Figure 4.

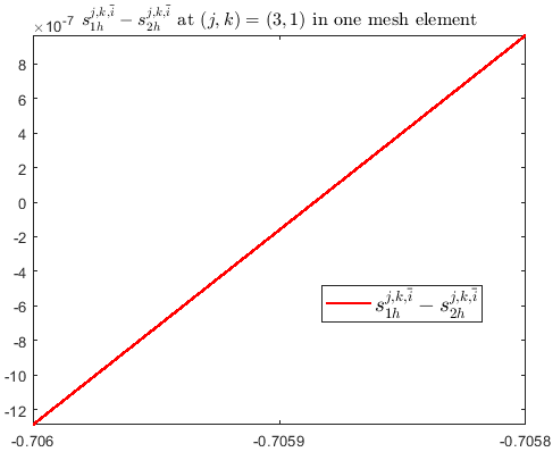

Figure 3: [Adaptive inexact smoothing Newton method, Algorithm 1, one space dimension, zoom on one element of the computational mesh $\mathcal{T}_h$] $s_{1h}^{j,k,\bar{i}} - s_{2h}^{j,k,\bar{i}}$ at steps $(j,k) = (3,1)$, at convergence of the algebraic solver $(i = \bar{i})$.

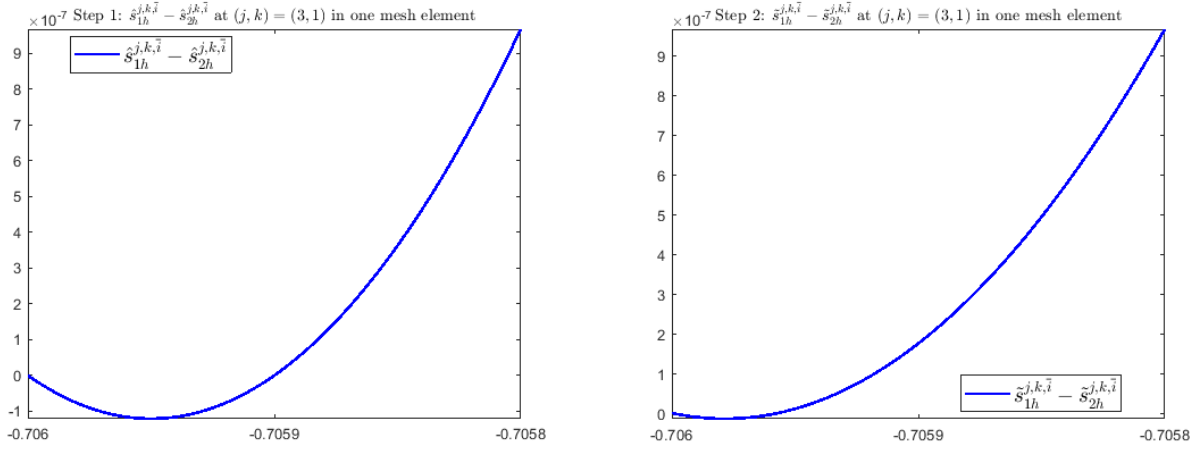

Figure 4: [Adaptive inexact smoothing Newton method, Algorithm 1, one space dimension, zoom on one element of the computational mesh $\mathcal{T}_h$] $\hat{s}_{1h}^{j,k,\bar{i}} - \hat{s}_{2h}^{j,k,\bar{i}}$ after the reconstruction step 1, left, and $\tilde{s}_{1h}^{j,k,\bar{i}} - \tilde{s}_{2h}^{j,k,\bar{i}}$ after the reconstruction step 2, right, at steps $(j,k) = (3,1)$ and at convergence of the algebraic solver $(i = \bar{i})$.

6. Flux reconstructions

We present in this section a construction of an equilibrated flux $\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}$ providing a discrete approximation of the exact flux $-\beta_\alpha \nabla u_\alpha$, cf. [48]. For this purpose, we will need the lowest-order Raviart–Thomas finite-dimensional subspace of $\mathbf{H}(\mathrm{div}, \Omega)$, defined by

$$\mathbf{RT}_0(\Omega) := \{\boldsymbol{v}_h \in \mathbf{H}(\mathrm{div}, \Omega)\,;\ \boldsymbol{v}_h|_K \in [\mathbb{P}_0(K)]^2 + \boldsymbol{x}\mathbb{P}_0(K)\}, \quad \forall K \in \mathcal{T}_h.$$

In particular, $\boldsymbol{v}_h \in \mathbf{RT}_0(\Omega)$ is such that $(\nabla \cdot \boldsymbol{v}_h)|_K \in \mathbb{P}_0(K), \forall K \in \mathcal{T}_h$, and $(\boldsymbol{v}_h \cdot \mathbf{n})|_\sigma \in \mathbb{P}_0(\sigma),\ \forall \sigma \in \mathcal{E}_K$. For more details, we refer to [58].

Let $\Pi_{\mathbb{P}_0}$ denote the $L_2(\Omega)$-orthogonal projection onto $\mathbb{P}_0(\mathcal{T}_h)$, the space of piecewise constants. An equilibrated flux reconstruction $\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}$ is a piecewise vector-valued polynomial function, designed to approximate $\boldsymbol{\sigma}_\alpha = -\beta_\alpha \nabla u_\alpha$,

and satisfying

$$\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i} \in \mathbf{RT}_0(\Omega), \tag{6.1a}$$

$$\nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i} = \Pi_{\mathbb{P}_0}(f_\alpha) - (-1)^\alpha \lambda_h^{j,k,i} \ \in \mathbb{P}_0(\mathcal{T}_h). \tag{6.1b}$$

The remaining difference between f_α and $\Pi_{\mathbb{P}_0}(f_\alpha)$ will be considered in the next section, giving rise to the so-called data oscillation. Note that the reconstructed flux mimics the properties of the weak flux. Indeed, (6.1b) is a discrete form of the condition $\nabla\cdot\boldsymbol{\sigma}_\alpha = f_\alpha - (-1)^\alpha\lambda$, where only the mean values of the divergence of $\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}$ need to coincide with the mean values of f_α on each mesh element. This can equivalently be written as

$$\left(\nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i} + (-1)^\alpha\lambda_h^{j,k,i}, 1\right)_K = (f_\alpha, 1)_K, \qquad \forall K \in \mathcal{T}_h.$$

We would like to emphasize that since the construction of the fluxes is based on the first two diffusion equations in (1.6) that are linear, there is no need to construct any linearization error flux as in [33]. To cope with inexact algebraic solver, though, we define the algebraic error flux reconstruction as follows.

Definition 6.1 (Algebraic error flux reconstruction)**.** *Let the smoothing step $j \geq 1$, the step of the nonlinear solver $k \geq 1$, and the step of the linear solver $i \geq 1$ be fixed. Given $\boldsymbol{R}_{\mathrm{alg},\alpha,K}^{j,k,i}$ defined in (4.7), and following [59, Concept 4.1], we can define the algebraic error flux reconstruction $\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}$ in $\mathbf{RT}_0(\mathcal{T}_h)$ for $\alpha \in \{1,2\}$ as follows*

$$\nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}|_K = \frac{\boldsymbol{R}_{\mathrm{alg},\alpha,K}^{j,k,i}}{|K|}, \qquad \forall K \in \mathcal{T}_h. \tag{6.2}$$

Definition 6.2 (Total flux reconstruction)**.** *The total flux reconstruction $\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i} \in \mathbf{RT}_0(\mathcal{T}_h)$ is defined by*

$$\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i} := -\beta_\alpha\nabla\tilde{\boldsymbol{u}}_{\alpha h}^{j,k,i} + \tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}. \tag{6.3}$$

Lemma 6.3 (Total flux reconstruction)**.** *There holds (6.1).*

Proof. First, condition (6.1a) follows from Definition 5.1 of the postprocessed solution together with Definition 6.1. To show (6.1b), we apply the Green formula and then employ (5.1b) and (4.7) which shows

$$\begin{aligned}
\left(\nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}, 1\right)_K &= \left(\nabla\cdot(-\beta_\alpha\nabla\tilde{\boldsymbol{u}}_{\alpha h}^{j,k,i}) + \nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}, 1\right)_K \\
&= \sum_{\sigma\in\mathcal{E}_K}\left(-\beta_\alpha\nabla\tilde{\boldsymbol{u}}_{\alpha h}^{j,k,i}\cdot\mathbf{n}_{K,\sigma}, 1\right)_\sigma + \left(\nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}, 1\right)_K \\
&\overset{(5.1b),(6.2)}{=} \sum_{\sigma\in\mathcal{E}_K} F_{\alpha,K,\sigma}^{j,k,i} + \boldsymbol{R}_{\mathrm{alg},\alpha,K}^{j,k,i} \\
&\overset{(4.7)}{=} \left(f_{\alpha,K} - (-1)^\alpha\lambda_K^{j,k,i}, 1\right)_K.
\end{aligned}$$

□

Remark 6.4 (Practical approximate algebraic error flux reconstruction)**.** *We use below a simple and practical approach to approximate the algebraic error flux reconstruction $\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}$, following [33, Section 4]. Let $\nu > 0$ be a user-given fixed parameter. Performing ν additional steps of the linear solver, then computing $-\beta_\alpha\nabla\tilde{\boldsymbol{u}}_{\alpha h}^{j,k,i+\nu}$ as in (5.1b) with $i+\nu$ in place of i, an algebraic error flux reconstruction can be defined as*

$$\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i} := -\beta_\alpha\nabla\tilde{\boldsymbol{u}}_{\alpha h}^{j,k,i+\nu} - \left(-\beta_\alpha\nabla\tilde{\boldsymbol{u}}_{\alpha h}^{j,k,i}\right),$$

satisfying (6.2) approximately.

7. A posteriori error estimates

Equipped with the key ingredients of the a posteriori analysis, namely the postprocessing and reconstructions of Sections 5 and 6, we are now in a position to rigorously derive an a posteriori estimate for the displacements. This allows to obtain a fully computable error upper bound at any smoothing step $j \geq 1$, any linearization step $k \geq 1$, and any step of the algebraic solver $i \geq 1$ of the inexact smoothing Newton method of Section 4. Let us stress that, for $j \geq 1, k \geq 1$, and $i \geq 1$, the conditions $(u_{1h}^{j,k,i} - u_{2h}^{j,k,i}) \geq 0$, $\lambda_h^{j,k,i} \geq 0$, and $\lambda_h^{j,k,i}(u_{1h}^{j,k,i} - u_{2h}^{j,k,i}) = 0$ are not necessarily satisfied, see Figure 5 for an illustration. In addition to the developments of Section 5, to deal with the possible violation of condition $\lambda_h^{j,k,i} \geq 0$, we define the negative and positive parts of $\lambda_h^{j,k,i}$ by

$$\lambda_h^{j,k,i} = \lambda_h^{j,k,i,\mathrm{pos}} + \lambda_h^{j,k,i,\mathrm{neg}}, \qquad \lambda_h^{j,k,i,\mathrm{pos}} := \max\{\lambda_h^{j,k,i}, 0\}, \qquad \lambda_h^{j,k,i,\mathrm{neg}} := \min\{\lambda_h^{j,k,i}, 0\}.$$

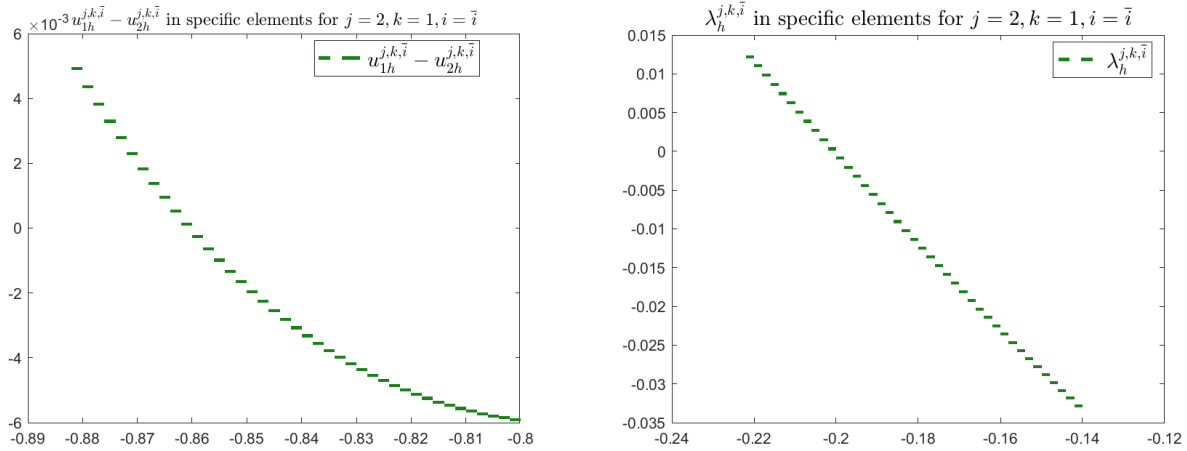

Figure 5: [Adaptive inexact smoothing Newton method, Algorithm 1, one space dimension, zoom on some elements of the computational mesh $\mathcal{T}_h$] $u_{1h}^{j,k,\bar{i}} - u_{2h}^{j,k,\bar{i}}$, left, and $\lambda_h^{j,k,\bar{i}}$, right, in specific elements, at steps $(j,k) = (2,1)$ and at convergence of the algebraic solver $(i = \bar{i})$.

7.1. A posteriori error estimate for the displacements

Recall that C_{PF} and C_{PW} are the Poincaré constants from (2.1). Let $C_{\beta,\Omega} := C_{\mathrm{PF}} h_\Omega (\frac{1}{\beta_1} + \frac{1}{\beta_2})^{\frac{1}{2}}$. We introduce for each element different estimators $\eta_{\cdot,K}^{j,k,i}, K \in \mathcal{T}_h$ together with their global counterparts $\eta_{\cdot}^{j,k,i} := \{\sum_{K\in\mathcal{T}_h} (\eta_{\cdot,K}^{j,k,i})^2\}^{\frac{1}{2}}$. We then have the following theorem.

Theorem 7.1 (A posteriori estimate for the displacements)**.** *Let $\boldsymbol{u} \in \boldsymbol{\mathcal{K}}_g$ be the weak solution of (2.8). Consider the finite volume discretization (4.7)–(4.8) on smoothing step $j \geq 1$, linearization step $k \geq 1$, and algebraic step $i \geq 1$. Let the postprocessed solution $\tilde{\boldsymbol{u}}_h^{j,k,i}$ be given following Definition 5.1, and the admissible potential reconstruction $\tilde{\boldsymbol{s}}_h^{j,k,i}$ following Definition 5.3. Next, let the algebraic error flux reconstruction be given following Definition 6.1, and the total flux reconstruction following Definition 6.2. Let $\boldsymbol{\Pi}_0^\sigma$ be the $L^2(\sigma)$-orthogonal projection onto constants. For $\alpha \in \{1,2\}$, define the local elementwise estimators*

$$\eta_{\mathrm{nonc},K}^{j,k,i} := \left|\!\left|\!\left| \tilde{\boldsymbol{s}}_h^{j,k,i} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\!\right|\!\right|_K, \tag{7.1a}$$

$$\eta_{\mathrm{osc},K,\alpha} := C_{\mathrm{PW}} h_K \beta_\alpha^{-\frac{1}{2}} \left\| f_\alpha - \Pi_{\mathbb{P}_0}(f_\alpha) \right\|_K, \qquad \eta_{\mathrm{osc}} := \left(\sum_{K\in\mathcal{T}_h} \sum_{\alpha=1}^{2} (\eta_{\mathrm{osc},K,\alpha})^2 \right)^{\frac{1}{2}}, \tag{7.1b}$$

$$\eta_{\mathrm{alg},K,\alpha}^{j,k,i} := \beta_\alpha^{-\frac{1}{2}} \left\| \tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i} \right\|_K, \qquad \eta_{\mathrm{alg}}^{j,k,i} := \left(\sum_{K\in\mathcal{T}_h} \sum_{\alpha=1}^{2} (\eta_{\mathrm{alg},K,\alpha}^{j,k,i})^2 \right)^{\frac{1}{2}}, \tag{7.1c}$$

$$\eta_{\mathrm{sm,lin,alg},1,K}^{j,k,i} := C_{\beta,\Omega} \left\| \lambda_h^{j,k,i,\mathrm{neg}} \right\|_K, \qquad \eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i} := 2 \left(\lambda_h^{j,k,i,\mathrm{pos}}, \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i} \right)_K. \tag{7.1d}$$

Then, defining the total estimator by

$$\eta^{j,k,i} := \left\{ \left(\eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta_{\mathrm{nonc}}^{j,k,i} + \eta_{\mathrm{sm,lin,alg},1}^{j,k,i} \right)^2 + \sum_{K \in \mathcal{T}_h} \eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i} \right\}^{\frac{1}{2}}, \tag{7.2}$$

the following a posteriori error estimate holds for the energy semi-norm, as well as for the energy semi-norm augmented by the jump term for the the postprocessed solution

$$\left|\!\left|\!\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\!\right|\!\right| \leq \eta^{j,k,i}, \tag{7.3a}$$

$$\left|\!\left|\!\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\!\right|\!\right| + \left\{ \sum_{\sigma \in \mathcal{E}_h} |\sigma|^{-1} \left\| \left[\!\left[\boldsymbol{u} - \boldsymbol{\Pi}_0^{\sigma}(\tilde{\boldsymbol{u}}_h^{j,k,i}) \right]\!\right] \right\|_{\sigma}^2 \right\}^{\frac{1}{2}} \leq \eta^{j,k,i} + \left\{ \sum_{\sigma \in \mathcal{E}_h} |\sigma|^{-1} \left\| \left[\!\left[\boldsymbol{\Pi}_0^{\sigma}(\tilde{\boldsymbol{u}}_h^{j,k,i}) \right]\!\right] \right\|_{\sigma}^2 \right\}^{\frac{1}{2}}. \tag{7.3b}$$

Remark 7.2 (Estimates (7.3))**.** *The estimate (7.3a) gives a fully computable upper bound on the energy semi-norm of the error between the exact solution $\boldsymbol{u}$ and its approximation $\tilde{\boldsymbol{u}}_h^{j,k,i}$ at each smoothing, linearization, and algebraic iterations j, k, and $i \geq 1$. The data oscillation estimators $\eta_{\mathrm{osc},K,\alpha}$ come from the fact that the source term is not necessarily piecewise constant, whereas $\eta_{\mathrm{alg},K,\alpha}^{j,k,i}$ reflect the algebraic error. The estimators $\eta_{\mathrm{sm,lin,alg},1,K}^{j,k,i}$ and $\eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i}$ reflect inconsistencies in the contact conditions at the discrete level, whereas $\eta_{\mathrm{nonc},K}^{j,k,i}$ evaluates the nonconformity of the postprocessed solution $\tilde{\boldsymbol{u}}_h^{j,k,i}$, i.e. the fact that it does not lie in $\boldsymbol{\mathcal{K}}_g$. Finally, (7.3b) adds an error jump term to the left which equals the jump estimator on the right since $[\![u_\alpha]\!] = 0$, $\alpha \in \{1, 2\}$. This transforms the energy semi-norm into a norm.*

Proof. We first remark that (7.3b) follows from (7.3a) by adding to both sides of the inequality the same term, since $[\![\boldsymbol{u}]\!] = 0$. To prove (7.3a), we distinguish the following two cases.

Case 1. If $\left|\!\left|\!\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\!\right|\!\right| \leq \left|\!\left|\!\left| \boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i} \right|\!\right|\!\right|$, we just have to estimate $\left|\!\left|\!\left| \boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i} \right|\!\right|\!\right|$.

The reduced problem (2.8) for the test function $\boldsymbol{v} = \tilde{\boldsymbol{s}}_h^{j,k,i} \in \boldsymbol{\mathcal{K}}_g$ gives

$$a(\boldsymbol{u}, \boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}) \leq l(\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}). \tag{7.4}$$

Denoting $\boldsymbol{w} := \boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}$, we use (7.4) and add and subtract $a(\tilde{\boldsymbol{u}}_h^{j,k,i}, \boldsymbol{w})$ and $b(\boldsymbol{w}, \lambda_h^{j,k,i})$ to get, also employing the notations (2.4),

$$\begin{aligned} a(\boldsymbol{w}, \boldsymbol{w}) \leq\ & l(\boldsymbol{w}) + b(\boldsymbol{w}, \lambda_h^{j,k,i}) - a(\tilde{\boldsymbol{u}}_h^{j,k,i}, \boldsymbol{w}) + a(\tilde{\boldsymbol{u}}_h^{j,k,i} - \tilde{\boldsymbol{s}}_h^{j,k,i}, \boldsymbol{w}) - b(\boldsymbol{w}, \lambda_h^{j,k,i}) \\ =\ & \sum_{\alpha=1}^{2} \left(f_\alpha - (-1)^\alpha \lambda_h^{j,k,i}, w_\alpha \right) - \sum_{\alpha=1}^{2} \beta_\alpha \left(\nabla \tilde{u}_{\alpha h}^{j,k,i}, \nabla w_\alpha \right) + a(\tilde{\boldsymbol{u}}_h^{j,k,i} - \tilde{\boldsymbol{s}}_h^{j,k,i}, \boldsymbol{w}) - b(\boldsymbol{w}, \lambda_h^{j,k,i}). \end{aligned} \tag{7.5}$$

As $\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i} \in \mathbf{H}(\mathrm{div}, \Omega)$ by (6.1a), and as, relying on Definition 5.3, $w_\alpha \in H_0^1(\Omega)$, the Green formula gives

$$\left(\nabla \cdot \tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}, w_\alpha \right) = - \left(\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}, \nabla w_\alpha \right) \qquad \forall \alpha \in \{1, 2\}. \tag{7.6}$$

Then, from (6.3) and (7.6), we have

$$\begin{aligned} a(\boldsymbol{w}, \boldsymbol{w}) \leq & \sum_{\alpha=1}^{2} \sum_{K \in \mathcal{T}_h} \left(f_\alpha - (-1)^\alpha \lambda_h^{j,k,i} - \nabla \cdot \tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}, w_\alpha \right)_K - \sum_{\alpha=1}^{2} \sum_{K \in \mathcal{T}_h} \left(\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}, \nabla w_\alpha \right)_K \\ & + a(\tilde{\boldsymbol{u}}_h^{j,k,i} - \tilde{\boldsymbol{s}}_h^{j,k,i}, \boldsymbol{w}) - b(\boldsymbol{w}, \lambda_h^{j,k,i}). \end{aligned} \tag{7.7}$$

It remains to bound each of the four terms in (7.7).

Using for the first term the flux property (6.1b) and the Cauchy–Schwarz and Poincaré–Wirtinger inequalities

(2.1b) as $w_\alpha|_K \in H^1(K)$, we have

$$\left(f_\alpha - (-1)^\alpha \lambda_h^{j,k,i} - \nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}, w_\alpha\right)_K = (f_\alpha - \Pi_{\mathbb{P}_0}(f_\alpha), w_\alpha - \overline{w}_{\alpha,K})_K \le \eta_{\mathrm{osc},K,\alpha} \left\|\beta_\alpha^{\frac{1}{2}}\nabla w_\alpha\right\|_K,$$

$$\left(\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}, \nabla w_\alpha\right)_K \le \eta_{\mathrm{alg},K,\alpha}^{j,k,i} \left\|\beta_\alpha^{\frac{1}{2}}\nabla w_\alpha\right\|_K,$$

where $\overline{w}_{\alpha,K}$ denotes the mean value of w_α on K. By applying the Cauchy–Schwarz inequality and using the definition of the energy semi-norm (2.2), we obtain

$$\sum_{\alpha=1}^{2}\sum_{K\in\mathcal{T}_h}\left(f_\alpha - (-1)^\alpha \lambda_h^{j,k,i} - \nabla\cdot\tilde{\boldsymbol{\sigma}}_{\alpha h}^{j,k,i}, w_\alpha\right)_K \le \eta_{\mathrm{osc}} \left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right|, \tag{7.8}$$

$$-\sum_{\alpha=1}^{2}\sum_{K\in\mathcal{T}_h}\left(\tilde{\boldsymbol{\sigma}}_{\alpha h,\mathrm{alg}}^{j,k,i}, \nabla w_\alpha\right)_K \le \eta_{\mathrm{alg}}^{j,k,i} \left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right|. \tag{7.9}$$

For the third term of (7.7), applying the Cauchy–Schwarz inequality, we get

$$a(\tilde{\boldsymbol{u}}_h^{j,k,i} - \tilde{\boldsymbol{s}}_h^{j,k,i}, \boldsymbol{w}) \le \underbrace{\left|\!\left|\!\left|\tilde{\boldsymbol{u}}_h^{j,k,i} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right|}_{\eta_{\mathrm{nonc}}^{j,k,i}} \left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right|. \tag{7.10}$$

Next, as $\boldsymbol{u} \in \boldsymbol{\mathcal{K}}_g$, $-b(\boldsymbol{u}, \lambda_h^{j,k,i,\mathrm{pos}}) \le 0$, and since $\boldsymbol{w} = \boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}$, we have

$$-b(\boldsymbol{w}, \lambda_h^{j,k,i,\mathrm{pos}}) \le\ b(\tilde{\boldsymbol{s}}_h^{j,k,i}, \lambda_h^{j,k,i,\mathrm{pos}}).$$

Using the fact that $\lambda_h^{j,k,i} = \lambda_h^{j,k,i,\mathrm{pos}} + \lambda_h^{j,k,i,\mathrm{neg}}$, the last term of (7.7) will be estimated as

$$-b(\boldsymbol{w}, \lambda_h^{j,k,i}) \le -\, b(\boldsymbol{w}, \lambda_h^{j,k,i,\mathrm{neg}}) + b(\tilde{\boldsymbol{s}}_h^{j,k,i}, \lambda_h^{j,k,i,\mathrm{pos}}) \tag{7.11a}$$

$$= -\,(\lambda_h^{j,k,i,\mathrm{neg}}, w_1 - w_2) + \left(\lambda_h^{j,k,i,\mathrm{pos}}, \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i}\right). \tag{7.11b}$$

The Cauchy–Schwarz inequality and the definition of the energy norm (2.2) lead to

$$\|\nabla(w_1 - w_2)\| \le \sum_{\alpha=1}^{2} \beta_\alpha^{-\frac{1}{2}} \left\|\beta_\alpha^{\frac{1}{2}}\nabla w_\alpha\right\| \le \left(\sum_{\alpha=1}^{2}\beta_\alpha^{-1}\right)^{\frac{1}{2}} \left(\sum_{\alpha=1}^{2}\left\|\beta_\alpha^{\frac{1}{2}}\nabla w_\alpha\right\|^2\right)^{\frac{1}{2}} \le \left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right)^{\frac{1}{2}} |||\boldsymbol{w}|||. \tag{7.12}$$

The Poincaré–Friedrichs inequality (2.1a) together with (7.12) give

$$-b(\boldsymbol{w}, \lambda_h^{j,k,i}) \ \le\ \eta_{\mathrm{sm,lin,alg},1}^{j,k,i} \left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right| + \frac{1}{2}\sum_{K\in\mathcal{T}_h} \underbrace{2\left(\lambda_h^{j,k,i,\mathrm{pos}}, \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i}\right)_K}_{\eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i}}. \tag{7.13}$$

Finally, due to the results (7.8), (7.9), (7.10), and (7.13) we have

$$\left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right|^2 \le \left(\eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta_{\mathrm{nonc}}^{j,k,i} + \eta_{\mathrm{sm,lin,alg},1}^{j,k,i}\right) \left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right| + \frac{1}{2}\sum_{K\in\mathcal{T}_h} \eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i}. \tag{7.14}$$

The Young inequality $ab \le \frac{1}{2}\left(a^2 + b^2\right),\ (a,b) \ge 0$, applied to the first term of (7.14) finally gives

$$\left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right| \le \eta^{j,k,i} := \left\{\left(\eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta_{\mathrm{nonc}}^{j,k,i} + \eta_{\mathrm{sm,lin,alg},1}^{j,k,i}\right)^2 + \sum_{K\in\mathcal{T}_h} \eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i}\right\}^{\frac{1}{2}}.$$

Case 2. If $\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right| \le \left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right|$, we have

$$\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right|^2 = a(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}) = a(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{u}-\tilde{\boldsymbol{s}}_h^{j,k,i}) + a(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i},\tilde{\boldsymbol{s}}_h^{j,k,i}-\tilde{\boldsymbol{u}}_h^{j,k,i}). \tag{7.15}$$

We start by estimating the first term of (7.15), while still denoting $\boldsymbol{w} = \boldsymbol{u}-\tilde{\boldsymbol{s}}_h^{j,k,i}$, as

$$\begin{aligned} a(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{w}) \le\ & l(\boldsymbol{w}) - a(\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{w}) + b(\boldsymbol{w},\lambda_h^{j,k,i}) - b(\boldsymbol{w},\lambda_h^{j,k,i}) \\ & \le \sum_{\alpha=1}^{2}\left(f_\alpha - (-1)^\alpha \lambda_h^{j,k,i}, w_\alpha\right) - \sum_{\alpha=1}^{2}\beta_\alpha\left(\nabla\tilde{u}_{\alpha h}^{j,k,i},\nabla w_\alpha\right) - b(\boldsymbol{w},\lambda_h^{j,k,i}), \end{aligned} \tag{7.16}$$

using again (2.4) and (2.8), as in (7.5). The three terms in (7.16) are identical to the terms in (7.5), estimated in (7.8), (7.9), and (7.13), respectively. Invoking the hypothesis of this case, we can thus write

$$\begin{aligned} a(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{w}) &\le \left(\eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta_{\mathrm{sm,lin,alg},1}^{j,k,i}\right)\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{s}}_h^{j,k,i}\right|\!\right|\!\right| + \frac{1}{2}\sum_{K\in\mathcal{T}_h}\eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i} \\ &\le \left(\eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta_{\mathrm{sm,lin,alg},1}^{j,k,i}\right)\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right| + \frac{1}{2}\sum_{K\in\mathcal{T}_h}\eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i}. \end{aligned}$$

The Cauchy–Schwarz inequality yields for the second term of (7.15)

$$a(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i},\tilde{\boldsymbol{s}}_h^{j,k,i}-\tilde{\boldsymbol{u}}_h^{j,k,i}) \le \left|\!\left|\!\left|\tilde{\boldsymbol{s}}_h^{j,k,i}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right|\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right| = \eta_{\mathrm{nonc}}^{j,k,i}\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right|.$$

By combining the previous results, we then obtain

$$\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right|^2 \le \left(\eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta_{\mathrm{nonc}}^{j,k,i} + \eta_{\mathrm{sm,lin,alg},1}^{j,k,i}\right)\left|\!\left|\!\left|\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right| + \frac{1}{2}\sum_{K\in\mathcal{T}_h}\eta_{\mathrm{sm,lin,alg},2,K}^{j,k,i}. \tag{7.17}$$

The Young inequality $ab \le \frac{1}{2}\left(a^2+b^2\right)$, $(a,b)\ge 0$, applied to the first term of (7.17) provides now again immediately the desired result. □

7.2. A posteriori error estimate for the actions

We present here an a posteriori estimate for the actions $\lambda_h^{j,k,i}$, extending [43, Corollary 3.5] to the nonconforming and inexact solvers setting.

Theorem 7.3 (A posteriori estimate for the actions)**.** *Let the assumptions and notations of Theorem 7.1 hold. The following a posteriori error estimate holds between the solution $\lambda\in\Lambda$ of problem (2.6) and the approximation $\lambda_h^{j,k,i}$ given by (4.7)–(4.8)*

$$\left|\!\left|\!\left|\lambda-\lambda_h^{j,k,i}\right|\!\right|\!\right|_{H_*^{-1}(\Omega)} \le \eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta^{j,k,i}. \tag{7.18}$$

Proof. Let $\beta_{\mathrm{m}} := \max(\beta_1,\beta_2)$. From the definition (2.3) of the norm of $H_*^{-1}(\Omega)$ and of the form b in (2.4) we have

$$\left|\!\left|\!\left|\lambda-\lambda_h^{j,k,i}\right|\!\right|\!\right|_{H_*^{-1}(\Omega)} = \sup_{\substack{v\in H_0^1(\Omega)\\ \beta_{\mathrm{m}}\|\nabla v\|^2=1}} (\lambda-\lambda_h^{j,k,i},v) = \sup_{\substack{\boldsymbol{\varphi}\in[H_0^1(\Omega)]^2\\ \beta_{\mathrm{m}}\sum_{\alpha=1}^2\|\nabla\varphi_\alpha\|^2=1}} b(\boldsymbol{\varphi},\lambda-\lambda_h^{j,k,i}).$$

Fix $\boldsymbol{\varphi}\in[H_0^1(\Omega)]^2$ such that $\beta_{\mathrm{m}}\sum_{\alpha=1}^2\|\nabla\varphi_\alpha\|^2 = 1$. It follows from (2.6a) that $-b(\boldsymbol{\varphi},\lambda-\lambda_h^{j,k,i}) = l(\boldsymbol{\varphi}) - a(\boldsymbol{u},\boldsymbol{\varphi}) + b(\boldsymbol{\varphi},\lambda_h^{j,k,i})$. By simply adding and subtracting $a(\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{\varphi})$, where the action of the form a on $\tilde{\boldsymbol{u}}_h^{j,k,i}$ is defined in (2.5), we obtain

$$-b(\boldsymbol{\varphi},\lambda-\lambda_h^{j,k,i}) = l(\boldsymbol{\varphi}) + b(\boldsymbol{\varphi},\lambda_h^{j,k,i}) - a(\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{\varphi}) - a(\boldsymbol{u}-\tilde{\boldsymbol{u}}_h^{j,k,i},\boldsymbol{\varphi}).$$

The first three terms are identical to the first three terms in (7.5) but with $\boldsymbol{\varphi}$ instead of $\boldsymbol{w}$. They are estimated in

(7.8) and (7.9), leading to

$$l(\boldsymbol{\varphi}) + b(\boldsymbol{\varphi}, \lambda_h^{j,k,i}) - a(\tilde{\boldsymbol{u}}_h^{j,k,i}, \boldsymbol{\varphi}) \leq (\eta_{\text{osc}} + \eta_{\text{alg}}^{j,k,i}) \, |||\boldsymbol{\varphi}||| \, .$$

The last term is estimated as $-a(\boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i}, \boldsymbol{\varphi}) \leq \left|\left|\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\right|\right|$, since $|||\boldsymbol{\varphi}||| \leq 1$. Through these estimations we get

$$-b(\boldsymbol{\varphi}, \lambda - \lambda_h^{j,k,i}) \leq \eta_{\text{osc}} + \eta_{\text{alg}}^{j,k,i} + \left|\left|\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\right|\right| . \tag{7.19}$$

We obtain the desired result by combining (7.19) to (7.3a). □

7.3. Distinguishing the different error components

The aim of this section is to identify the various error components in the a posteriori estimators from Theorem 7.1, which will lead to a posteriori stopping criteria.

Corollary 7.4 (A posteriori error estimate distinguishing the error components)**.** *We define for $\alpha \in \{1, 2\}$ and $K \in \mathcal{T}_h$ the smoothing, discretization, linearization, and algebraic estimators as follows:*

$$\eta_{\text{disc}}^{j,k,i} := \eta_{\text{osc}} + \eta_{\text{nonc}}^{j,k,i} + \left(\left| \sum_{K \in \mathcal{T}_h} 2 \left(\lambda_h^{j,k,i,\text{pos}}, \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i} - \tilde{u}_{1h}^{j,k,i} + \tilde{u}_{2h}^{j,k,i} \right)_K \right| \right)^{\frac{1}{2}} , \tag{7.20a}$$

$$\eta_{\text{sm,lin,alg}}^{j,k,i} := \eta_{\text{sm,lin,alg},1}^{j,k,i} + \left(\left| \sum_{K \in \mathcal{T}_h} 2 \left(\lambda_h^{j,k,i,\text{pos}}, u_{1h}^{j,k,i} - u_{2h}^{j,k,i} \right)_K \right| \right)^{\frac{1}{2}} , \tag{7.20b}$$

$$\eta_{\text{lin,alg}}^{j,k,i} := \eta_{\text{lin,alg},1}^{j,k,i} + \left(\left| \sum_{K \in \mathcal{T}_h} 2 \left(\lambda_h^{j,k,i,\text{pos}} - \lambda_h^{j,k-1,\bar{i},\text{pos}}, u_{1h}^{j,k,i} - u_{2h}^{j,k,i} - u_{1h}^{j,k-1,\bar{i}} + u_{2h}^{j,k-1,\bar{i}} \right)_K \right| \right)^{\frac{1}{2}} , \tag{7.20c}$$

$$\eta_{\text{alg}}^{j,k,i} := \left(\sum_{K \in \mathcal{T}_h} \sum_{\alpha=1}^{2} \left(\eta_{\text{alg},K,\alpha}^{j,k,i} \right)^2 \right)^{\frac{1}{2}} , \tag{7.20d}$$

with

$$\eta_{\text{lin,alg},1,K}^{j,k,i} := C_{\beta,\Omega} \left\| \lambda_h^{j,k,i,\text{neg}} - \lambda_h^{j,k-1,\bar{i},\text{neg}} \right\|_K , \quad \text{and} \quad \eta_{\text{lin,alg},1}^{j,k,i} := \left(\sum_{K \in \mathcal{T}_h} \left(\eta_{\text{lin,alg},1,K}^{j,k,i} \right)^2 \right)^{\frac{1}{2}} .$$

Then,

$$\left|\left|\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\right|\right| \leq \eta^{j,k,i} \leq \eta_{\text{disc}}^{j,k,i} + \eta_{\text{sm,lin,alg}}^{j,k,i} + \eta_{\text{lin,alg}}^{j,k,i} + \eta_{\text{alg}}^{j,k,i} . \tag{7.21}$$

Proof. From (7.3a), employing the inequality $(a+b)^{\frac{1}{2}} \leq a^{\frac{1}{2}} + b^{\frac{1}{2}}$, for $a, b \geq 0$, we have

$$\left|\left|\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\right|\right| \leq \eta^{j,k,i} \leq \eta_{\text{osc}} + \eta_{\text{alg}}^{j,k,i} + \eta_{\text{nonc}}^{j,k,i} + \eta_{\text{sm,lin,alg},1}^{j,k,i} + \left(\sum_{K \in \mathcal{T}_h} \eta_{\text{sm,lin,alg},2,K}^{j,k,i} \right)^{\frac{1}{2}} . \tag{7.22}$$

We then decompose $\eta_{\text{sm,lin,alg},2,K}^{j,k,i}$ by adding and subtracting the components of $\tilde{\boldsymbol{u}}_h^{j,k,i}$ as follows

$$\begin{aligned} \eta_{\text{sm,lin,alg},2,K}^{j,k,i} &= 2 \left(\lambda_h^{j,k,i,\text{pos}}, \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i} \right)_K \\ &= 2 \left(\lambda_h^{j,k,i,\text{pos}}, \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i} - \tilde{u}_{1h}^{j,k,i} + \tilde{u}_{2h}^{j,k,i} \right)_K + 2 \left(\lambda_h^{j,k,i,\text{pos}}, \tilde{u}_{1h}^{j,k,i} - \tilde{u}_{2h}^{j,k,i} \right)_K \\ &\overset{(5.1a)}{=} 2 \left(\lambda_h^{j,k,i,\text{pos}}, \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i} - \tilde{u}_{1h}^{j,k,i} + \tilde{u}_{2h}^{j,k,i} \right)_K + 2 \left(\lambda_h^{j,k,i,\text{pos}}, u_{1h}^{j,k,i} - u_{2h}^{j,k,i} \right)_K . \end{aligned} \tag{7.23}$$

We now combine (7.23) together with (7.22) inserting the absolute values. This leads to

$$\left|\left|\left| \boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i} \right|\right|\right| \leq \eta_{\text{disc}}^{j,k,i} + \eta_{\text{sm,lin,alg}}^{j,k,i} + \eta_{\text{alg}}^{j,k,i} .$$

Finally, we define the linearization estimator $\eta_{\mathrm{lin,alg}}^{j,k,i}$ analogously to the smoothing estimator $\eta_{\mathrm{sm,lin,alg}}^{j,k,i}$, considering the terms $\lambda_h^{j,k,i} - \lambda_h^{j,k-1,\bar{i}}$ and $\boldsymbol{u}_h^{j,k,i} - \boldsymbol{u}_h^{j,k-1,\bar{i}}$ estimating the linearization error. □

Remark 7.5 (Nature of the estimators)**.** *The nonconformity and oscillation estimators $\eta_{\mathrm{nonc}}^{j,k,i}$ and η_{osc} considered as discretization estimators vanish when the computational effort grows, i.e. when the number of mesh elements goes to infinity. The smoothing estimator $\eta_{\mathrm{sm,lin,alg}}^{j,k,i}$ stems from the error in the algebraic system, linearization, and smoothing. It goes to zero at convergence of all the solvers, since when j, k, and $i \to \infty$, we have $\lambda_h^{j,k,i} \geq 0$ and $\lambda_h^{j,k,i}(u_{1h}^{j,k,i} - u_{2h}^{j,k,i}) = 0$. The linearization estimator $\eta_{\mathrm{lin,alg}}^{j,k,i}$ reflects the error stemming from both linearization and algebraic resolution and vanishes when k and $i \to \infty$. Finally, the algebraic estimator $\eta_{\mathrm{alg}}^{j,k,i}$ evaluating the error in the algebraic iterative resolution of the linear system (4.4) vanishes when $i \to \infty$.*

8. Stopping criteria and adaptive inexact smoothing algorithm

We derive in this section adaptive stopping criteria for the linear, the nonlinear solver, and the smoothing iterations, based on the estimators of Corollary 7.4.

Let three user-specified parameters ζ_{sm}, ζ_{lin}, and ζ_{alg} be given in $]0,1]$, representing the desired relative size (percentage) of the smoothing, linearization, and algebraic errors, respectively. Below, we denote by $\overline{j}, \overline{k}$, and $\overline{i}$ the last (stopping) smoothing, linearization and algebraic step, respectively. The stopping criterion for the algebraic step i at each linearization step k and smoothing step j is chosen as

$$\eta_{\mathrm{alg}}^{j,k,\overline{i}} < \zeta_{\mathrm{alg}} \eta_{\mathrm{lin,alg}}^{j,k,\overline{i}}. \tag{8.1}$$

This criterion expresses that there is no need to continue with the algebraic iterations once the linearization error component starts to dominate. Similarly, to stop the Newton iterations at each smoothing step j, we apply

$$\eta_{\mathrm{lin,alg}}^{j,\overline{k},\overline{i}} < \zeta_{\mathrm{lin}} \eta_{\mathrm{sm,lin,alg}}^{j,\overline{k},\overline{i}}, \tag{8.2}$$

which requires the linearization estimator to be sufficiently small with respect to the smoothing estimator. Finally, we stop the outer smoothing loop whenever

$$\eta_{\mathrm{sm,lin,alg}}^{\overline{j},\overline{k},\overline{i}} < \zeta_{\mathrm{sm}} \eta_{\mathrm{disc}}^{\overline{j},\overline{k},\overline{i}}, \tag{8.3}$$

i.e. when the smoothing estimator is ζ_{sm}-times smaller than the discretization estimator. As for the amount of smoothing, we will proceed following [26] and diminish it by a fixed factor $\zeta \in]0,1[$ on each smoothing step. We are now ready to present our adaptive inexact smoothing Newton algorithm that includes the above adaptive criteria for stopping the iterative solvers.

Algorithm 1 Adaptive inexact smoothing Newton algorithm

1. Initialization
Choose parameters $\zeta \in]0,1[$ and $\zeta_{\rm sm}, \zeta_{\rm lin}, \zeta_{\rm alg} \in]0,1]$.
Choose an initial smoothing parameter $\mu^1 > 0$, a number of additional algebraic solver steps $\nu \geq 1$, and an initial approximation $\boldsymbol{X}^0 \in \mathbb{R}^n$. Set $j := 1$ and $\overline{j} = 0$.

2. Smoothing j-loop

2.1 Set $\boldsymbol{X}^{j,0} := \boldsymbol{X}^0, k := 1$, and $\overline{k} = 0$.

2.2 Newton linearization k-loop

2.2.1 From $\boldsymbol{X}^{j,k-1}$ define $\mathbb{A}_{\mu^j}^{j,k-1} \in \mathbb{R}^{n,n}$ and $\boldsymbol{B}_{\mu^j}^{j,k-1} \in \mathbb{R}^n$ by the Newton linearization (4.5).

2.2.2 Consider the problem of finding a solution $\boldsymbol{X}^{j,k}$ to

$$\mathbb{A}_{\mu^j}^{j,k-1} \boldsymbol{X}^{j,k} = \boldsymbol{B}_{\mu^j}^{j,k-1}. \tag{8.4}$$

2.2.3 Set $\boldsymbol{X}^{j,k,0} := \boldsymbol{X}^{j,k-1}$ as initial guess for the iterative algebraic solver.
Set $i := 1$, and if $j = 1$ and $k = 1$, set $\overline{i} = 0$.

2.2.4 Algebraic solver i-loop

i) Perform ν steps of the iterative algebraic solver for the solution of (8.4), yielding, on step $i + \nu$, an approximation $\boldsymbol{X}^{j,k,i+\nu}$ to $\boldsymbol{X}^{j,k}$ satisfying

$$\mathbb{A}_{\mu^j}^{j,k-1} \boldsymbol{X}^{j,k,i+\nu} = \boldsymbol{B}_{\mu^j}^{j,k-1} - \boldsymbol{R}_{\rm alg}^{j,k,i+\nu}.$$

ii) Set $i := i + \nu$. Compute the estimators given in (7.20).

iii) If $\eta_{\rm alg}^{j,k,i} < \zeta_{\rm alg} \eta_{\rm lin,alg}^{j,k,i}$, set $\overline{i} := i$ and stop. If not, go to **i)**.

2.2.5 If $\eta_{\rm lin,alg}^{j,k,\overline{i}} < \zeta_{\rm lin} \eta_{\rm sm,lin,alg}^{j,k,\overline{i}}$, set $\overline{k} := k$ and stop. If not, set $k := k + 1$ and go to **2.2.1**.

2.3 If $\eta_{\rm sm,lin,alg}^{j,\overline{k},\overline{i}} < \zeta_{\rm sm} \eta_{\rm disc}^{j,\overline{k},\overline{i}}$, set $\overline{j} := j$ and stop.
If not, set $j := j + 1$, $\boldsymbol{X}^{j,0} := \boldsymbol{X}^{j-1,\overline{k},\overline{i}}$, and $\mu^j := \zeta \mu^{j-1}$. Then set $k := 1$ and go to **2.2.1**.

9. Numerical results

In this section, we numerically illustrate the efficiency of our theoretical developments considering problem (1.6). Our main goals are to assess the sharpness of the guaranteed bound (7.3) and to show that Algorithm 1 performs well and leads to smaller number of iterations in comparison with usual stopping criteria as well as the classical semismooth Newton method.

We carry out computations fixing the tensions in (1.6a) and (1.6b) as $\beta_1 = 1$ and $\beta_2 = 1$. The boundary condition g of the first membrane in (1.6d) is taken equal to 0.1. We consider the one-dimensional domain $\Omega = (-1, 1)$, (all the theoretical developments apply here), and use the following analytical solution for $x \in \Omega$, following [43],

$$u_1(x) := g(2x^2 - 1), \quad u_2(x) := \begin{cases} 2g(1-x^2)(2x^2-1) & \text{if } x < \frac{-1}{\sqrt{2}} \text{ or } x > \frac{1}{\sqrt{2}}, \\ g(2x^2-1) & \text{otherwise,} \end{cases} \quad \lambda(x) := \begin{cases} 0 & \text{if } x < \frac{-1}{\sqrt{2}} \text{ or } x > \frac{1}{\sqrt{2}}, \\ 2g & \text{otherwise.} \end{cases}$$

This triple is the solution of (1.6) for the data f_1 and f_2 given by

$$f_1(x) := \begin{cases} -4g & \text{if } x < \frac{-1}{\sqrt{2}} \text{ or } x > \frac{1}{\sqrt{2}}, \\ -6g & \text{otherwise,} \end{cases} \qquad f_2(x) := \begin{cases} -12g(1-4x^2) & \text{if } x < \frac{-1}{\sqrt{2}} \text{ or } x > \frac{1}{\sqrt{2}}, \\ -2g & \text{otherwise.} \end{cases}$$

For all the tests, the number of mesh elements is $m = 10000$, leading to the overall number of unknowns $n = 30000$. We choose the initial guess as $\boldsymbol{X}^0 = [\mathbf{1}g, \mathbf{0}, \mathbf{0}] \in \mathbb{R}^n$, where $\mathbf{1} = [1, \ldots, 1]^T \in \mathbb{R}^m$. The implementation was done

in the MATLAB software. The value of the coefficients $\zeta_{\rm sm}, \zeta_{\rm lin}$, and $\zeta_{\rm alg}$ from the adaptive stopping criteria in Section 8 is 0.1. The parameters in Algorithm 1 are set as: $\mu^1 = 1,\ \zeta = 0.1$, and $\nu = 4$.

9.1. Semismooth Newton-min

First, for comparison, to find an approximate solution to the algebraic system (2.11), we employ the semismooth Newton-min method described in Section 3 in which the stopping criterion for the linearization requests the relative total residual of problem (3.3) $\boldsymbol{R}^k_{\rm rel} := \|\boldsymbol{R}(\boldsymbol{X}^k)\|/\|\boldsymbol{R}(\boldsymbol{X}^0)\|$ to be below 10^{-8}, where

$$\boldsymbol{R}(\boldsymbol{V}) := \begin{bmatrix} \boldsymbol{F} - \mathbb{E}\boldsymbol{V} \\ -\boldsymbol{C}(\boldsymbol{V}) \end{bmatrix}, \quad \boldsymbol{V} \in \mathbb{R}^n. \tag{9.1}$$

The evolution of the relative total residual is shown in Figure 6. In its right part, we zoom on the last 10 Newton-min iterations. We observe that the curve goes down slowly until step 893, where the convergence gets extremely fast.

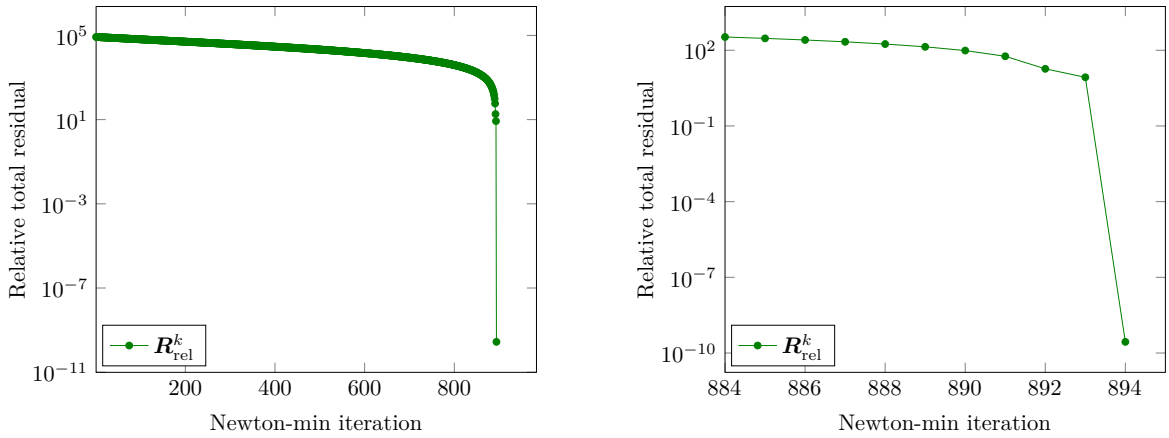

Figure 6: [Semismooth Newton-min method of Section 3] Relative total residual as a function of Newton-min iterations, left, and as a function of the last 10 Newton-min iterations, right.

9.2. Adaptive smoothing Newton-min

In this section, we employ Algorithm 1 with an "exact" resolution of the system of algebraic equations (8.4), i.e., we skip steps 2.2.3 and 2.2.4. We drop the notation "alg" from estimators (7.20b) and (7.20c), whereas $\eta^{j,k,i}_{\rm alg}$ of (7.20d) vanishes. First, we want to emphasize the performance of the adaptive smoothing method employing the adaptive stopping criterion to stop the nonlinear solver. To this end, we use two linearization stopping criteria: the adaptive criterion (8.2) $\eta^{j,k}_{\rm lin} < \zeta_{\rm lin}\eta^{j,k}_{\rm sm,lin}$ and the classical one on the relative linearization residual of problem (4.3) $\boldsymbol{R}^{j,k}_{\rm lin,rel} := \|\boldsymbol{R}_{\rm lin}(\boldsymbol{X}^{j,k})\|/\|\boldsymbol{R}_{\rm lin}(\boldsymbol{X}^{1,0})\|$ lying below 10^{-8}, with $\boldsymbol{R}_{\rm lin}(\cdot)$ given by

$$\boldsymbol{R}_{\rm lin}(\boldsymbol{V}) := \begin{bmatrix} \boldsymbol{F} - \mathbb{E}\boldsymbol{V} \\ -\boldsymbol{C}_{\mu^j}(\boldsymbol{V}) \end{bmatrix}, \quad \boldsymbol{V} \in \mathbb{R}^n. \tag{9.2}$$

We show in Figure 7 the number of performed Newton iterations employing the smoothing Newton method with exact algebraic resolution, during the fourth smoothing step ($j = 4$). It can be noticed that the use of the adaptive stopping criterion brings down the number of iterations from 20 to 12.

We now employ the adaptive smoothing Newton-min method, with an exact algebraic resolution, including the adaptive stopping criteria (8.2) and (8.3) to stop the linearization and smoothing steps, respectively. In terms of numbers, 6 smoothing iterations and 41 cumulated linearization iterations are needed to reach the end of the simulation, as seen from Figure 8 left, compared to 894 linearization iterations employing the semismooth Newton-min method above.

The various estimators given in (7.20) are presented in the left part of Figure 8. Each set of curves represents one smoothing step (fixed value j). From each set one can see that the linearization estimator is dominant and

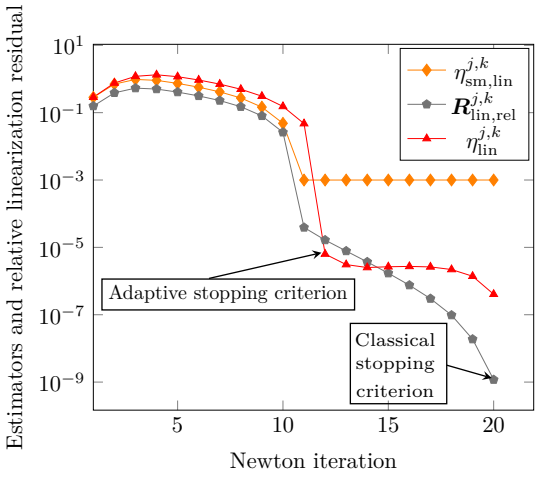

Figure 7: [Adaptive smoothing Newton-min method, Algorithm 1, exact resolution of the algebraic system (8.4)] Estimators and relative linearization residual as a function of the Newton iterations at a specific smoothing step ($j = 4$, k varies).

close to the total estimator, until becoming smaller than the smoothing estimator, when the adaptive stopping criterion $\eta^{j,k}_{\mathrm{lin}} < \zeta_{\mathrm{lin}}\eta^{j,k}_{\mathrm{sm,lin}}$ is satisfied. The smoothing estimator satisfies (8.3) from the cumulated Newton-min iteration $k = 40$. Computational savings in terms of linearization iterations can be evaluated considering the results in Figure 8, right. A comparison of the number of performed Newton iterations employing the semismooth Newton-min method of Section 9.1 and the adaptive smoothing Newton-min method of the present section shows a significant gain reaching a factor of roughly 22.

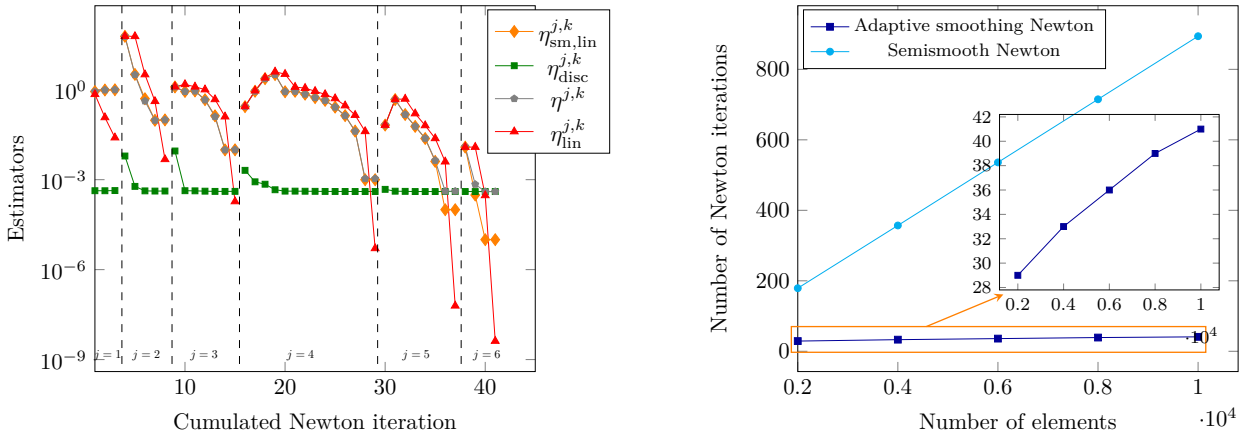

Figure 8: [Adaptive smoothing Newton-min method, Algorithm 1, exact resolution of the algebraic system (8.4)] Estimators as a function of the cumulated Newton iterations, left. Comparison between the number of performed Newton iterations employing the Newton-min method of Section 9.1 and the adaptive smoothing Newton-min method of Section 9.2, right.

9.3. Adaptive inexact smoothing Newton-min

This section is devoted to present the results obtained employing the adaptive inexact smoothing Newton-min algorithm of Algorithm 1 in Section 8. We consider at each Newton step $k \geq 1$ the GMRES iterative algebraic solver for the system (8.4), see [60], with an ILU preconditioner. To shed more light on the importance of the adaptive stopping criterion for stopping the linear solver, we compare the adaptive resolution where the stopping criterion for the GMRES is given by (8.1) with the classical resolution where the algebraic iterations are stopped using the relative algebraic residual, i.e.,

$$\boldsymbol{R}^{j,k,i}_{\mathrm{alg,rel}} := \frac{\left\|\mathbb{M}_2\backslash(\mathbb{M}_1\backslash(\boldsymbol{B}^{j,k-1}_{\mu^j} - \mathbb{A}^{j,k-1}_{\mu^j}\boldsymbol{X}^{j,k,i}))\right\|}{\left\|\mathbb{M}_2\backslash(\mathbb{M}_1\backslash(\boldsymbol{B}^{j,k-1}_{\mu^j} - \mathbb{A}^{j,k-1}_{\mu^j}\boldsymbol{X}^{j,k-1}))\right\|} \leq 10^{-10}, \tag{9.3}$$

where $\mathbb{M}_1$ and $\mathbb{M}_2$ denote the preconditioner matrices. Figure 9 shows the algebraic estimator, linearization estimator, and relative algebraic residual, computed every $\nu = 4$ algebraic steps, at specific smoothing and linearization steps $(j,k) = (4,1)$ for the classical and adaptive resolutions. We observe that 188 algebraic iterations are needed to meet the classical criterion (9.3), whereas only 20 iterations are required if we terminate the algebraic solver according to criterion (8.1).

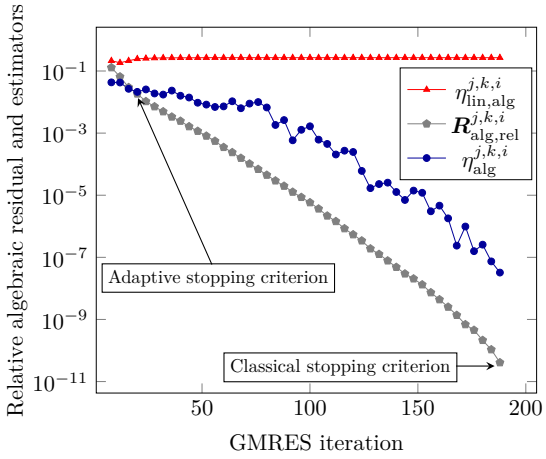

Figure 9: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Estimators and relative algebraic residual as a function of the GMRES iterations at smoothing and linearization steps $(j,k) = (4,1)$ using the adaptive stopping criterion (8.1) and the classical one (9.3).

We now employ the entire Algorithm 1 featuring also the adaptive stopping criterion for the algebraic solver. To satisfy adaptive criteria (8.1), (8.2), and (8.3), 6 smoothing iterations, 41 cumulated Newton-min iterations, and 2552 cumulated GMRES iterations are needed. We also assess the quality of the a posteriori error estimates of Theorems 7.1 and 7.3 by means of the effectivity indices resulting from estimates (7.3a), (7.3b), and (7.18) defined as

$$\mathrm{I}_{\text{eff}}^{j,k,i} := \frac{\eta^{j,k,i}}{|||\boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i}|||}, \tag{9.4a}$$

$$\bar{\mathrm{I}}_{\text{eff}}^{j,k,i} := \frac{\eta^{j,k,i} + \left\{\sum_{\sigma\in\mathcal{E}_h} |\sigma|^{-1} \|[\![\boldsymbol{\Pi}_0^\sigma(\tilde{\boldsymbol{u}}_h^{j,k,i})]\!]\|_\sigma^2\right\}^{\frac12}}{|||\boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i}||| + \left\{\sum_{\sigma\in\mathcal{E}_h} |\sigma|^{-1} \|[\![\boldsymbol{u} - \boldsymbol{\Pi}_0^\sigma(\tilde{\boldsymbol{u}}_h^{j,k,i})]\!]\|_\sigma^2\right\}^{\frac12}}, \tag{9.4b}$$

$$\tilde{\mathrm{I}}_{\text{eff}}^{j,k,i} := \frac{\eta_{\text{osc}} + \eta_{\text{alg}}^{j,k,i} + 2\eta^{j,k,i} + \left\{\sum_{\sigma\in\mathcal{E}_h} |\sigma|^{-1} \|[\![\boldsymbol{u} - \boldsymbol{\Pi}_0^\sigma(\tilde{\boldsymbol{u}}_h^{j,k,i})]\!]\|_\sigma^2\right\}^{\frac12}}{|||\boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i}||| + |||\lambda - \lambda_h^{j,k,i}|||_{H_*^{-1}(\Omega)} + \left\{\sum_{\sigma\in\mathcal{E}_h} |\sigma|^{-1} \|[\![\boldsymbol{u} - \boldsymbol{\Pi}_0^\sigma(\tilde{\boldsymbol{u}}_h^{j,k,i})]\!]\|_\sigma^2\right\}^{\frac12}}. \tag{9.4c}$$

See Remark 9.1 for details on approximately computing the dual norm.

Remark 9.1 (Computing approximately the dual norm)**.** *In practice, the dual norm* $|||\lambda - \lambda_h^{j,k,i}|||_{H^{-1}(\Omega)}$ *with* $\lambda - \lambda_h^{j,k,i} \in \Lambda$*, is not easily computable. We provide here a practical way to approximate this norm and evaluate it numerically following [61]. We consider the following elliptic problem that consists in finding, for a given* $f \in L^2(\Omega)$*, the function* $\phi : \Omega \to \mathbb{R}$ *such that*

$$\begin{aligned} -\Delta\phi &= f \qquad \text{in } \Omega, \\ \phi &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{9.5}$$

The weak formulation of problem (9.5) *consists in finding* $\phi \in H_0^1(\Omega)$ *such that*

$$(\nabla\phi, \nabla v) = (f, v) \qquad \forall v \in H_0^1(\Omega). \tag{9.6}$$

Then the definition of the $H^{-1}(\Omega)$ norm together with (9.6) give

$$\|f\|_{H^{-1}(\Omega)} = \sup_{v\in H_0^1(\Omega);\ ||\nabla v||_\Omega=1} (f,v) \overset{(9.6)}{=} \sup_{v\in H_0^1(\Omega);\ ||\nabla v||_\Omega=1} (\nabla\phi,\nabla v) = \|\nabla\phi\|_\Omega.$$

We consider the cell-centered finite volume method to find an approximate solution to problem (9.5) on a refined mesh. Assuming that the discretization error is negligible, we employ $\|\nabla\tilde{\phi}_h\|_\Omega$, where $\tilde{\phi}_h$ is obtained by a postprocessing as in Definition 5.1, to approximate $\|f\|_{H^{-1}(\Omega)}$.

The results are reported in Table 1 where we show at each smoothing step j and linearization step k: the last algebraic step $\bar{i}$, the estimators, and the effectivity indices (9.4) at convergence of the algebraic solver ($i=\bar{i}$). We observe that we indeed have a guaranteed upper bound on all steps $j\geq 1, k\geq 1$, and $\bar{i}$, and that all the effectivity indices take excellent values when all the three stopping criteria (8.1)–(8.3) are satisfied on the last line of Table 1.

j	k	$\bar{i}$	$\eta_{\rm disc}^{j,k,\bar{i}}$	$\eta_{\rm sm,lin,alg}^{j,k,\bar{i}}$	$\eta_{\rm lin,alg}^{j,k,\bar{i}}$	$\eta_{\rm alg}^{j,k,\bar{i}}$	$\eta^{j,k,\bar{i}}$	$\mathrm{I}_{\rm eff}^{j,k,\bar{i}}$	$\bar{\mathrm{I}}_{\rm eff}^{j,k,\bar{i}}$	$\tilde{\mathrm{I}}_{\rm eff}^{j,k,\bar{i}}$
1	1	12	4.25e-04	8.72e-01	7.06e-01	5.71e-02	8.74e-01	1.70	1.64	2.01
1	2	12	4.27e-04	9.78e-01	1.19e-01	6.83e-04	9.78e-01	1.68	1.63	1.93
1	3	12	4.33e-04	9.97e-01	2.52e-02	4.18e-04	9.97e-01	1.68	1.64	1.93
2	1	20	2.16e-01	5.89e+01	5.94e+01	1.49e+00	6.03e+01	140.12	46.87	75.11
2	2	16	8.45e-03	3.33e+00	6.03e+01	2.78e+00	6.06e+00	65.93	31.89	60.43
2	3	12	4.43e-04	9.40e-01	3.65e+00	2.63e-01	1.12e+00	47.56	14.21	22.96
2	4	12	4.19e-04	9.50e-02	7.66e-01	1.81e-02	9.67e-02	4.91	2.02	2.89
2	5	12	4.18e-04	9.84e-02	6.23e-03	6.04e-04	9.84e-02	4.27	1.97	2.68
3	1	12	1.83e-02	1.16e+00	1.16e+00	8.51e-02	1.23e+00	37.75	13.62	16.84
3	2	12	1.64e-02	2.01e-01	1.15e+00	1.10e-01	3.17e-01	10.29	4.07	5.88
3	3	28	1.29e-02	2.34e-01	3.37e-01	1.47e-02	2.50e-01	34.75	4.77	8.18
3	4	24	8.35e-03	4.43e-02	2.51e-01	2.11e-02	6.19e-02	16.86	1.97	3.10
3	5	28	1.86e-03	9.33e-03	3.66e-02	1.90e-03	9.73e-03	10.51	1.15	1.32
3	6	48	4.08e-04	9.91e-03	8.06e-04	2.89e-05	9.92e-03	13.22	1.16	1.32
4	1	16	3.45e-03	2.42e-01	2.40e-01	1.56e-02	2.56e-01	73.81	5.16	8.53
4	2	12	2.73e-03	2.80e-02	2.53e-01	6.15e-03	3.37e-02	14.96	1.53	1.86
4	3	44	6.58e-04	2.31e-01	2.32e-01	1.60e-02	2.46e-01	258.31	5.24	9.68
4	4	8	1.13e-03	2.83e-03	2.47e-01	1.65e-02	1.84e-02	23.26	1.31	1.80
4	5	60	4.75e-04	1.04e-01	1.03e-01	7.75e-03	1.11e-01	231.17	2.94	5.01
4	6	8	1.12e-03	1.73e-03	1.25e-01	1.18e-02	1.28e-02	26.74	1.22	1.60
4	7	60	4.04e-04	2.48e-02	2.40e-02	2.04e-03	2.62e-02	63.62	1.45	1.95
4	8	8	4.50e-04	1.10e-03	2.77e-02	2.15e-03	2.78e-03	6.78	1.04	1.12
4	9	156	4.03e-04	9.97e-04	3.51e-05	3.47e-06	1.08e-03	2.63	1.01	1.03
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6	1	20	5.07e-04	2.87e-03	2.84e-03	2.81e-04	3.52e-03	8.61	1.05	1.12
6	2	16	4.79e-04	4.73e-05	3.38e-03	1.36e-04	5.76e-04	1.41	1.00	1.01
6	3	60	4.14e-04	3.12e-03	3.12e-03	5.37e-05	3.57e-03	8.73	1.06	1.12
6	4	12	4.80e-04	2.73e-05	3.18e-03	1.69e-04	5.75e-04	1.41	1.00	1.02
6	5	96	4.06e-04	1.39e-03	1.39e-03	1.33e-04	1.92e-03	4.69	1.03	1.06
6	6	08	4.34e-04	2.02e-05	1.66e-03	4.36e-05	4.46e-04	1.09	1.00	1.01
6	7	316	4.02e-04	2.87e-03	2.86e-03	2.61e-04	3.52e-03	8.62	1.05	1.12
6	8	08	4.15e-04	1.21e-05	2.88e-03	1.66e-05	4.18e-04	1.02	1.00	1.01
6	9	680	4.01e-04	1.00e-05	1.73e-07	1.30e-08	4.01e-04	1.00	1.00	1.01

Table 1: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Last algebraic step $\bar{i}$, estimators (7.20) and effectivity indices (9.4) at each smoothing step j and each Newton-min step k, at convergence of the algebraic solver ($i=\bar{i}$).

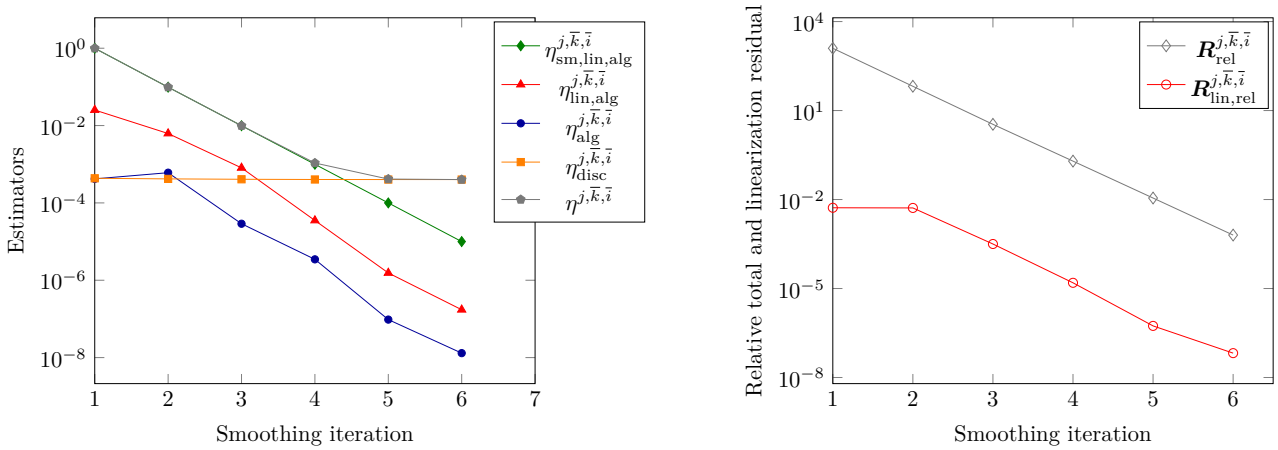

Figure 10: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Estimators of Section 7.3, left, and relative linearization and total residuals, right, as a function of the smoothing iterations j at convergence of the algebraic and linearization solvers (j varies, $k = \overline{k}, i = \overline{i}$).

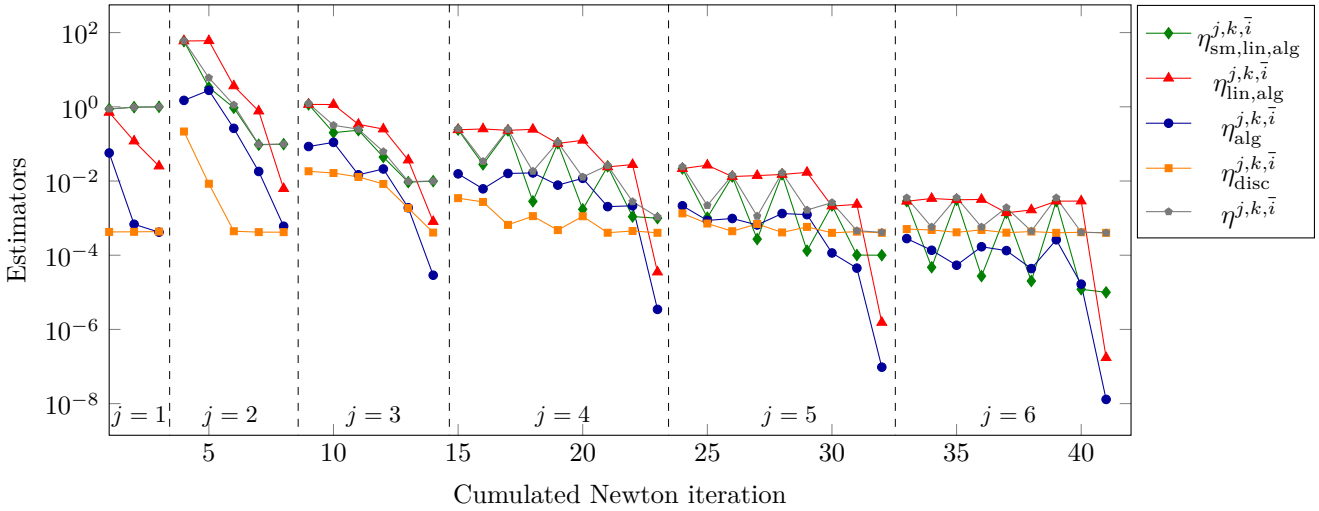

Figure 11: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Estimators of Section 7.3 as a function of the cumulated Newton-min iterations at convergence of the algebraic solver (j and k vary, $i = \overline{i}$).

We next plot in Figure 10, left, the evolution of the various estimators as a function of the smoothing iterations in j when the stopping criteria (8.2) and (8.1) have been satisfied. The curve of the smoothing estimator goes down at each smoothing step while the discretization estimator stagnates. In the right part, we show the relative total residual $\boldsymbol{R}_{\mathrm{rel}}^{j,\overline{k},\overline{i}} := \|\boldsymbol{R}(\boldsymbol{X}^{j,\overline{k},\overline{i}})\|/\|\boldsymbol{R}(\boldsymbol{X}^0)\|$ with $\boldsymbol{R}(\cdot)$ given in (9.1) and the relative linearization residual $\boldsymbol{R}_{\mathrm{lin,rel}}^{j,\overline{k},\overline{i}} := \|\boldsymbol{R}_{\mathrm{lin}}(\boldsymbol{X}^{j,\overline{k},\overline{i}})\|/\|\boldsymbol{R}_{\mathrm{lin}}(\boldsymbol{X}^{1,0})\|$ with $\boldsymbol{R}_{\mathrm{lin}}(\cdot)$ given in (9.2) during the smoothing iterations. Let us point out that $\boldsymbol{R}_{\mathrm{rel}}^{j,\overline{k},\overline{i}}$ steadily decreases as we tighten the smoothing. The residual $\boldsymbol{R}_{\mathrm{lin,rel}}^{j,\overline{k},\overline{i}}$ in turn systematically takes smaller values. The estimators as a function of the cumulated Newton-min iterations are then illustrated in Figure 11. We remark that at each smoothing step the linearization estimator and the algebraic estimator (blue) steadily decrease, while the discretization estimator roughly stagnates. The oscillating behavior of $\eta_{\mathrm{sm,lin,alg}}^{j,k,\overline{i}}$ is explained by the fact that it involves $\eta_{\mathrm{sm,lin,alg},1}^{j,k,\overline{i}}$ given in (7.1d) that takes values varying between 0 and 6.91e+01 depending on whether the constraint $\lambda_h^{j,k,i} \geq 0$ is satisfied or not. Moreover, Figure 12 shows the evolution of the estimators during the cumulated algebraic steps for $j = \{1, 2\}$. The two sets of curves separated by the dashed line represent two smoothing steps whereas the inner sets separated by the dotted lines represent the linearization steps. As expected, the discretization and smoothing estimators typically stagnate while the algebraic estimator decreases until step $\overline{i}$, at which criterion (8.1) is satisfied.

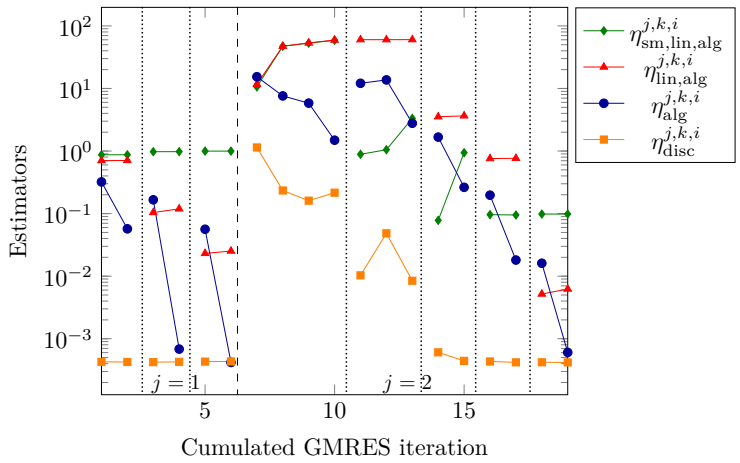

Figure 12: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Estimators of Section 7.3 as a function of the GMRES iterations during the first 2 smoothing iterations ($j = \{1, 2\}$, k and i vary).

Next, Figure 13 shows the effectivity indices (9.4) during the cumulated Newton-min iterations. It can be seen that the index $\mathrm{I}_{\text{eff}}^{j,k,\bar{i}}$ defined as the ratio of the total error estimator and the actual energy error takes bigger values than the indices $\bar{\mathrm{I}}_{\text{eff}}^{j,k,\bar{i}}$ featuring the jump term and the estimate $\tilde{\mathrm{I}}_{\text{eff}}^{j,k,\bar{i}}$ featuring the jump term and the action. When the stopping criteria (8.1)–(8.3) are reached, all the indices approach the optimal value of one.

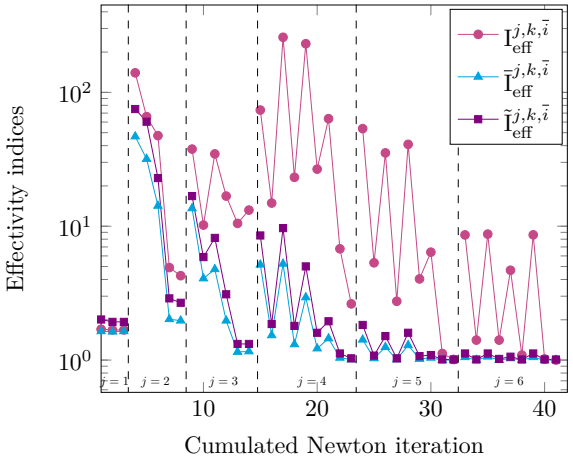

Figure 13: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Effectivity indices given in (9.4) using the total estimator $\eta^{j,k,i}$ given in (7.2), as a function of the cumulated Newton-min iterations, at convergence of the algebraic solver ($i = \bar{i}$).

The estimators and the effectivity indices at convergence of all solvers, i.e., when the criteria (8.1), (8.2), and (8.3) have been satisfied, are plotted in Figure 14 as a function of the number of mesh elements. Notice that the discretization estimator essentially coincides with the total estimator. We observe that the accuracy of our estimators increases in function of the computational effort.

We are also interested in the comparison between the adaptive GMRES (adaptive stopping criterion (8.1)) and the classical GMRES (standard stopping criteria (9.3)) with regard to the number of performed iterations. As seen from Figure 15, the adaptive algebraic resolution does not impact the number of smoothing steps. It slightly affects the number of cumulated Newton steps but leads to an important decrease of the number of GMRES iterations compared with the classical resolution. In this regard, we numerically explore the influence of the coefficients $\zeta_{\text{sm}}, \zeta_{\text{lin}}$, and ζ_{alg} in the adaptive stopping criteria of Section 8 on the smoothing algorithm. We summarize the results obtained in Table 2. We observe that choosing ζ_{sm} or ζ_{lin} small does not considerably affect the overall number of iterations. However, setting ζ_{alg} small increases notably the number of algebraic and linearization iterations.

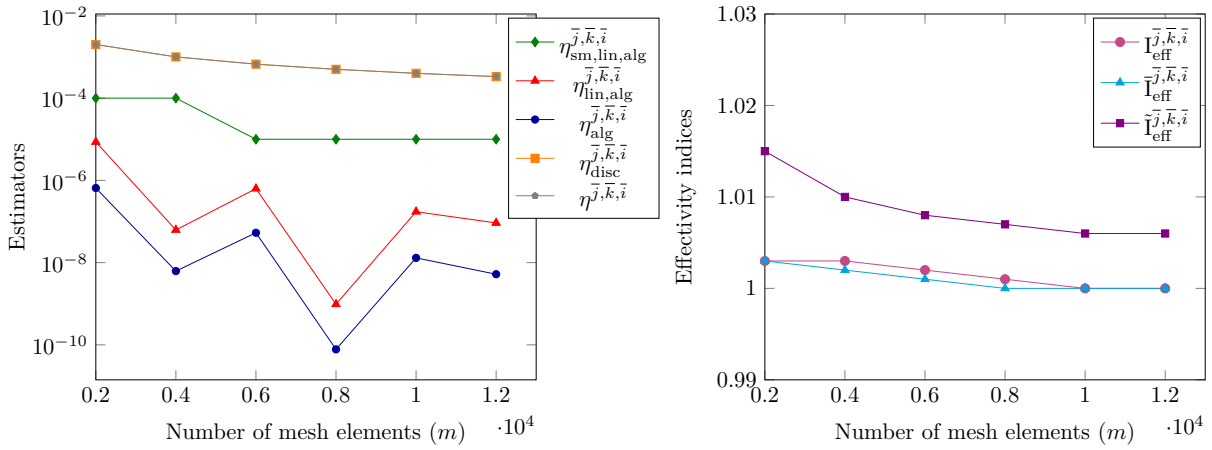

Figure 14: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Estimators, left, and effectivity indices, right, as a function of the number of mesh elements m at convergence of all the solvers.

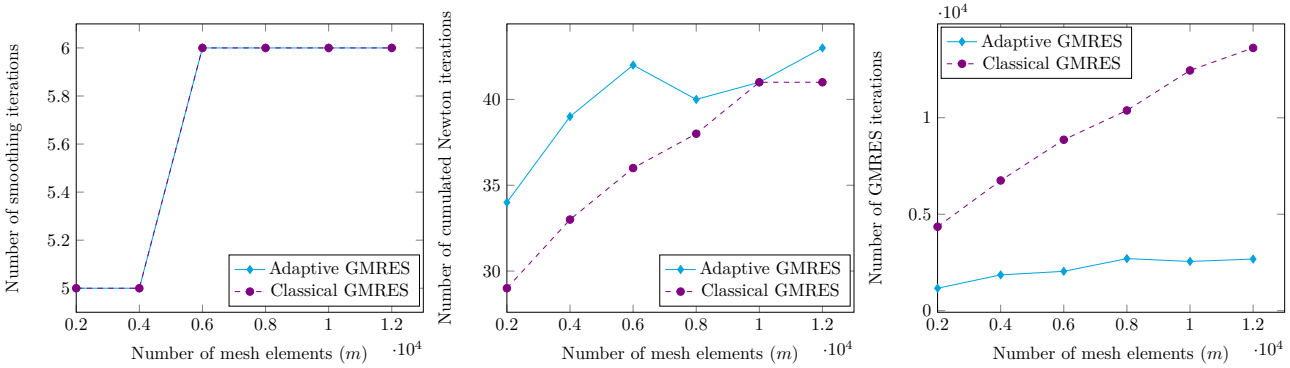

Figure 15: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Number of smoothing iterations (left), cumulated Newton-min iterations (center), and of cumulated algebraic iterations (right) as a function of the number of mesh elements, employing the adaptive stopping criterion (8.1) and the classical one (9.3) for stopping the GMRES solver.

ζ_{sm}	ζ_{lin}	ζ_{alg}	# Smoothing iter.	# Cumul. Newton iter.	# Cumul. GMRES iter.	$\boldsymbol{R}^{\bar{j},\bar{k},\bar{i}}_{\mathrm{rel}}$
10^{-1}	10^{-1}	10^{-1}	6	41	2552	6.33e-04
10^{-2}	10^{-2}	10^{-2}	7	63	9108	3.67e-05
10^{-2}	10^{-1}	10^{-1}	7	45	3652	3.66e-05
10^{-1}	10^{-2}	10^{-1}	6	51	3944	6.33e-04
10^{-1}	10^{-1}	10^{-2}	6	57	6996	6.33e-04

Table 2: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Number of smoothing, cumulated Newton, and cumulated GMRES iterations as well as the relative norm of the total residual vector defined in (9.1) for various parameters $\zeta_{\mathrm{sm}}, \zeta_{\mathrm{lin}}$, and ζ_{alg} in the adaptive stopping criteria of Section 8.

10. Conclusions and outlook

The motivation of the present work was to propose an adaptive inexact smoothing Newton method based on rigorous a posteriori error estimates for solving nonlinear algebraic systems with complementarity constraints arising from finite volume discretizations. We considered in particular the problem modeling the contact between two membranes. We treated the non-differentiable nonlinearity in the constraints by means of a smoothed C-function, which allowed a direct application of the standard Newton method. We designed a posteriori error estimates between the exact and approximate solution, enabling to identify the error components (discretization, smoothing, linearization, algebraic) and yielding adaptive stopping criteria. These criteria together with a simple way of tightening the smoothing became the cornerstones of the developed adaptive algorithm. We finally provided numerical tests employing our adaptive method and the existing semismooth Newton method. The results agree

with theoretical developments and confirm that the adaptivity allows for important computational savings in terms of number of iterations. Future work will consist in applying this method to several synthetic cases of petroleum reservoir simulation, see [62].

Appendix A.

The a posteriori estimate (7.3) of Section 7.1 involves the L_2-norm of $\lambda_h^{j,k,i,\mathrm{neg}}$ and the global domain diameter h_Ω. This gives a guaranteed upper bound, but is not very sharp. We present here an alternative upper bound on the energy error that is typically sharper, but not guaranteed anymore.

Remark Appendix A.1 (Alternative bound)**.** *From* (7.11a)*,* $-b(\boldsymbol{w}, \lambda_h^{j,k,i})$ *can be decomposed as follows*

$$\begin{aligned} -b(\boldsymbol{w}, \lambda_h^{j,k,i}) &= -\left(\lambda_h^{j,k,i}, (u_1 - \tilde{s}_{1h}^{j,k,i}) - (u_2 - \tilde{s}_{2h}^{j,k,i})\right) \\ &\approx \frac{1}{2}\left\{\sum_{K\in\mathcal{T}_h} 2\left(\lambda_h^{j,k,i}, u_{2h}^{j,k,i} - u_{1h}^{j,k,i} + \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i}\right)_K\right\}. \end{aligned}$$

This will give us the following result

$$\left|\!\left|\!\left|\boldsymbol{u} - \tilde{\boldsymbol{u}}_h^{j,k,i}\right|\!\right|\!\right| \lesssim \eta_{\mathrm{alt}}^{j,k,i} := \left\{\left(\eta_{\mathrm{osc}} + \eta_{\mathrm{alg}}^{j,k,i} + \eta_{\mathrm{nonc}}^{j,k,i}\right)^2 + \sum_{K\in\mathcal{T}_h} 2\left(\lambda_h^{j,k,i}, u_{2h}^{j,k,i} - u_{1h}^{j,k,i} + \tilde{s}_{1h}^{j,k,i} - \tilde{s}_{2h}^{j,k,i}\right)_K\right\}^{\frac{1}{2}}. \quad \text{(A.1)}$$

The corresponding effectivity indices are illustrated during the cumulated linearization iterations in Figure A.16. We indeed observe a general improvement at the effectivity indices, though they become (importantly) below one at the initial iterations.

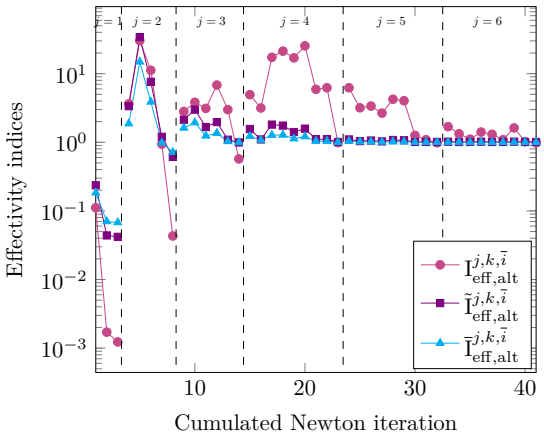

Figure A.16: [Adaptive inexact smoothing Newton-min method, Algorithm 1] Effectivity indices using the alternative total estimator $\eta_{\mathrm{alt}}^{j,k,i}$ as a function of the cumulated Newton-min iterations, at convergence of the algebraic solver ($i = \bar{i}$).

References

[1] M. C. Ferris, J. S. Pang, Engineering and economic applications of complementarity problems, SIAM Rev. 39 (4) (1997) 669–713. doi:10.1137/S0036144595285963.

[2] M. C. Ferris, K. Sinapiromsaran, Formulating and solving nonlinear programs as mixed complementarity problems, Vol. 481 of Lecture Notes in Econom. and Math. Systems, Springer, Berlin, 2000. doi:10.1007/978-3-642-57014-8_10.

[3] J. Dabaghi, V. Martin, M. Vohralík, Adaptive inexact semismooth Newton methods for the contact problem between two membranes, J. Sci. Comput. 84, 28 (2). doi:10.1007/s10915-020-01264-3.

[4] I. Ben Gharbia, J. Jaffré, Gas phase appearance and disappearance as a problem with complementarity constraints, Math. Comput. Simul. 99 (2014) 28–36. doi:10.1016/j.matcom.2013.04.021.

[5] T. De Luca, F. Facchinei, C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, Math. Programming 75 (3, Ser. A) (1996) 407–439. doi:10.1007/BF02592192.

[6] C. Hager, B. I. Wohlmuth, Semismooth Newton methods for variational problems with inequality constraints, GAMM-Mitt. 33 (1) (2010) 8–24. doi:10.1002/gamm.201010002.

[7] I. Ben Gharbia, J. C. Gilbert, Nonconvergence of the plain Newton-min algorithm for linear complementarity problems with a P-matrix, Math. Prog. 134 (2) (2012) 349–364. doi:10.1007/s10107-010-0439-6.

[8] J.-P. Dussault, M. Frappier, J. C. Gilbert, A lower bound on the iterative complexity of the Harker and Pang globalization technique of the Newton-min algorithm for solving the linear complementarity problem, EURO J. Comput. Optim. 7 (4) (2019) 359–380. doi:10.1007/s13675-019-00116-6.

[9] J. Dabaghi, G. Delay, A unified framework for high-order numerical discretizations of variational inequalities, Comput. Math. Appl. 92 (2021) 62–75. doi:10.1016/j.camwa.2021.03.011.

[10] C. Kanzow, An active set-type Newton method for constrained nonlinear systems, in: Complementarity: applications, algorithms and extensions (Madison, WI, 1999), Vol. 50 of Appl. Optim., Kluwer Acad. Publ., Dordrecht, 2001, pp. 179–200. doi:10.1007/978-1-4757-3279-5_9.

[11] M. Hintermüller, K. Ito, K. Kunisch, The primal-dual active set strategy as a semismooth Newton method, SIAM J. Optim. 13 (3) (2002) 865–888 (2003). doi:10.1137/S1052623401383558.

[12] N. Xiu, J. Zhang, Some recent advances in projection-type methods for variational inequalities, in: Proceedings of the International Conference on Recent Advances in Computational Mathematics (ICRACM 2001) (Matsuyama), Vol. 152, 2003, pp. 559–585. doi:10.1016/S0377-0427(02)00730-6.

[13] G. Stadler, Semismooth Newton and augmented Lagrangian methods for a simplified friction problem, SIAM J. Optim. 15 (1) (2004) 39–62. doi:10.1137/S1052623403420833.

[14] K. Ito, K. Kunisch, Semi-smooth Newton methods for the Signorini problem, Appl. Math. 53 (5) (2008) 455–468. doi:10.1007/s10492-008-0036-7.

[15] M. Hintermüller, K. Kunisch, Path-following methods for a class of constrained minimization problems in function space, SIAM J. Optim. 17 (1) (2006) 159–187. doi:10.1137/040611598.

[16] G. Stadler, Path-following and augmented Lagrangian methods for contact problems in linear elasticity, J. Comput. Appl. Math. 203 (2) (2007) 533–547. doi:10.1016/j.cam.2006.04.017.

[17] M. H. Wright, The interior-point revolution in optimization: history, recent developments, and lasting consequences, Bull. Amer. Math. Soc. 42 (1) (2005) 39–56. doi:10.1090/S0273-0979-04-01040-7.

[18] J. Gondzio, Interior point methods 25 years later, European J. Oper. Res. 218 (3) (2012) 587–601. doi:10.1016/j.ejor.2011.09.017.

[19] D. T. S. Vu, I. Ben Gharbia, M. Haddou, Q. H. Tran, A new approach for solving nonlinear algebraic systems with complementarity conditions. Application to compositional multiphase equilibrium problems, Mathematics and Computers in Simulation 190 (2021) 1243–1274. doi:10.1016/j.matcom.2021.07.015.

[20] M. C. Ferris, O. L. Mangasarian, J.-S. Pang (Eds.), Complementarity: applications, algorithms and extensions, Vol. 50 of Applied Optimization, Kluwer Academic Publishers, Dordrecht, 2001. doi:10.1007/978-1-4757-3279-5.

[21] F. Facchinei, J.-S. Pang, Finite-dimensional variational inequalities and complementarity problems. Vol. I, Springer Series in Operations Research, Springer-Verlag, New York, 2003. doi:10.1007/b97544.

[22] F. Facchinei, J.-S. Pang, Finite-dimensional variational inequalities and complementarity problems. Vol. II, Springer Series in Operations Research, Springer-Verlag, New York, 2003. doi:10.1007/b97543.

[23] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, C. A. Sagastizábal, Numerical optimization, 2nd Edition, Universitext, Springer-Verlag, Berlin, 2006. doi:10.1007/978-3-540-35447-5.

[24] K. Ito, K. Kunisch, Lagrange multiplier approach to variational problems and applications, Vol. 15 of Advances in Design and Control, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. doi:10.1137/1.9780898718614.

[25] M. Ulbrich, Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces, Vol. 11 of MOS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2011. doi:10.1137/1.9781611970692.

[26] I. Ben Gharbia, J. Ferzly, M. Vohralík, S. Yousef, Semismooth and smoothing Newton methods for nonlinear systems with complementarity constraints: adaptivity and inexact resolution, J. Comput. Appl. Math. 420 (2023) 114765. doi:10.1016/j.cam.2022.114765.

[27] F. Facchinei, C. Kanzow, A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, Math. Program. 76 (3, Ser. B) (1997) 493–512. doi:10.1007/BF02614395.

[28] J. M. Martínez, L. Q. Qi, Inexact Newton methods for solving nonsmooth equations, J. Comput. Appl. Math. 60 (1-2) (1995) 127–145. doi:10.1016/0377-0427(94)00088-I.

[29] S.-P. Rui, C.-X. Xu, A smoothing inexact Newton method for nonlinear complementarity problems, J. Comput. Appl. Math. 233 (9) (2010) 2332–2338. doi:10.1016/j.cam.2009.10.018.

[30] Z. Ge, Q. Ni, X. Zhang, A smoothing inexact Newton method for variational inequalities with nonlinear constraints, J. Inequal. Appl. (2017) Paper No. 160, 12. doi:10.1186/s13660-017-1433-9.

[31] M. Picasso, A stopping criterion for the conjugate gradient algorithm in the framework of anisotropic adaptive finite elements, Comm. Numer. Methods Engrg. 25 (4) (2009) 339–355. doi:10.1002/cnm.1120.

[32] R. Becker, D. Capatina, R. Luce, Stopping criteria based on locally reconstructed fluxes, in: Numerical mathematics and advanced applications—ENUMATH 2013, Vol. 103 of Lect. Notes Comput. Sci. Eng., Springer, Cham, 2015, pp. 243–251.

[33] A. Ern, M. Vohralík, Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs, SIAM J. Sci. Comput. 35 (4) (2013) A1761–A1791. doi:10.1137/120896918.

[34] D. A. Di Pietro, E. Flauraud, M. Vohralík, S. Yousef, A posteriori error estimates, stopping criteria, and adaptivity for multiphase compositional Darcy flows in porous media, J. Comput. Phys. 276 (2014) 163–187. doi:10.1016/j.jcp.2014.06.061.

[35] V. Rey, C. Rey, P. Gosselet, A strict error bound with separated contributions of the discretization and of the iterative solver in non-overlapping domain decomposition methods, Comput. Methods Appl. Mech. Engrg. 270 (2014) 293–303. doi:10.1016/j.cma.2013.12.001.

[36] F. Chouly, M. Fabre, P. Hild, J. Pousin, Y. Renard, Residual-based *a posteriori* error estimation for contact problems approximated by Nitsche's method, IMA J. Numer. Anal. 38 (2) (2018) 921–954. doi:10.1093/imanum/drx024.

[37] P. Heid, T. P. Wihler, Adaptive iterative linearization Galerkin methods for nonlinear problems, Math. Comp. 89 (326) (2020) 2707–2734. doi:10.1090/mcom/3545.

[38] S. Giani, L. Grubišić, L. Heltai, O. Mulita, Smoothed-adaptive perturbed inverse iteration for elliptic eigenvalue problems, Comput. Methods Appl. Math. 21 (2) (2021) 385–405. doi:10.1515/cmam-2020-0027.

[39] M. Ainsworth, J. T. Oden, A posteriori error estimation in finite element analysis, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000. doi:10.1002/9781118032824.

[40] S. Repin, A posteriori estimates for partial differential equations, Vol. 4 of Radon Series on Computational and Applied Mathematics, Walter de Gruyter GmbH & Co. KG, Berlin, 2008. doi:10.1515/9783110203042.

[41] R. Verfürth, A posteriori error estimation techniques for finite element methods, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013. doi:10.1093/acprof:oso/9780199679423.001.0001.

[42] S. I. Repin, Functional a posteriori estimates for elliptic variational inequalities, J. Math. Sci. 152 (2008) 702–712. doi:10.1007/s10958-008-9093-4.

[43] F. Ben Belgacem, C. Bernardi, A. Blouza, M. Vohralík, On the unilateral contact between membranes. Part 2: *a posteriori* analysis and numerical experiments, IMA J. Numer. Anal. 32 (3) (2012) 1147–1172. doi:10.1093/imanum/drr003.

[44] M. Bürg, A. Schröder, A posteriori error control of *hp*-finite elements for variational inequalities of the first and second kind, Comput. Math. Appl. 70 (12) (2015) 2783–2802. doi:10.1016/j.camwa.2015.08.031.

[45] P. Destuynder, B. Métivet, Explicit error bounds in a conforming finite element method, Math. Comp. 68 (1999) 1379–1396. doi:10.1090/S0025-5718-99-01093-5.

[46] D. Braess, J. Schöberl, Equilibrated residual error estimator for edge elements, Math. Comp. 77 (262) (2008) 651–672. doi:10.1090/S0025-5718-07-02080-7.

[47] M. Ainsworth, Robust a posteriori error estimation for nonconforming finite element approximation, SIAM J. Numer. Anal. 42 (6) (2005) 2320–2341. doi:10.1137/S0036142903425112.

[48] M. Vohralík, Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods, Numer. Math. 111 (1) (2008) 121–158. doi:10.1007/s00211-008-0168-4.

[49] A. Ern, M. Vohralík, Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations, SIAM J. Numer. Anal. 53 (2) (2015) 1058–1081. doi:10.1137/130950100.

[50] F. Ben Belgacem, C. Bernardi, A. Blouza, M. Vohralík, A finite element discretization of the contact between two membranes, M2AN Math. Model. Numer. Anal. 43 (1) (2009) 33–52. doi:10.1051/m2an/2008041.

[51] F. Ben Belgacem, C. Bernardi, A. Blouza, M. Vohralík, On the unilateral contact between membranes. Part 1: Finite element discretization and mixed reformulation, Math. Model. Nat. Phenom. 4 (1) (2009) 21–43. doi:10.1051/mmnp/20094102.

[52] S. Zhang, Y. Yan, R. Ran, Path-following and semismooth Newton methods for the variational inequality arising from two membranes problem, J. Inequal. Appl. 1. doi:10.1186/s13660-019-1955-4.

[53] R. Eymard, T. Gallouët, R. Herbin, Finite volume methods, Handb. Numer. Anal., VII, North-Holland, Amsterdam, 2000. doi:10.1016/S1570-8659(00)07005-8.

[54] M. Vohralík, On the discrete Poincaré-Friedrichs inequalities for nonconforming approximations of the Sobolev space H^1, Numer. Funct. Anal. Optim. 26 (7-8) (2005) 925–952. doi:10.1080/01630560500444533.

[55] M. Bebendorf, A note on the Poincaré inequality for convex domains, Z. Anal. Anwendungen 22 (4) (2003) 751–756. doi:10.4171/ZAA/1170.

[56] H. Brezis, Functional analysis, Sobolev spaces and partial differential equations, Universitext, Springer, New York, 2011. doi:10.1007/978-0-387-70914-7.

[57] R. Eymard, T. Gallouët, R. Herbin, Finite volume approximation of elliptic problems and convergence of an approximate gradient, Appl. Numer. Math. 37 (1-2) (2001) 31–53. doi:10.1016/S0168-9274(00)00024-6.

[58] F. Brezzi, M. Fortin, Mixed and hybrid finite element methods, Vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991. doi:10.1007/978-1-4612-3172-1.

[59] J. Papež, U. Rüde, M. Vohralík, B. Wohlmuth, Sharp algebraic and total a posteriori error bounds for h and p finite elements via a multilevel approach. Recovering mass balance in any situation, Comput. Methods Appl. Mech. Engrg. 371 (2020) 113243, 39. doi:10.1016/j.cma.2020.113243.

[60] Y. Saad, M. H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput. 7 (3) (1986) 856–869. doi:10.1137/0907058.

[61] D. A. Di Pietro, M. Vohralík, S. Yousef, Adaptive regularization, linearization, and discretization and a posteriori error control for the two-phase Stefan problem, Math. Comp. 84 (291) (2015) 153–186. doi:10.1090/S0025-5718-2014-02854-8.

[62] I. B. Gharbia, E. Flauraud, Study of compositional multiphase flow formulation using complementarity conditions, Oil Gas Sci. Technol. Rev. IFP Energies nouvelles. doi:10.2516/ogst/2019012.