Finite element approximation of penalized elastoplastic torsion problem with nonconstant source term

Franz Chouly, Tom Gustafsson, Patrick Hild

Finite element approximation of penalized elastoplastic torsion problem with nonconstant source term

Franz Chouly[a,*], Tom Gustafsson[b], Patrick Hild[c]

[a] *Center of Mathematics, University of the Republic, 11400 Montevideo, Uruguay*
[b] *Department of Mathematics and Systems Analysis, Aalto University, Otakaari 1 F, Espoo, Finland*
[c] *Institut de Mathématiques de Toulouse - UMR CNRS 5219, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 9, France*

Abstract

This study is concerned with the finite element approximation of the elastoplastic torsion problem. We focus on the case of a nonconstant source term, which cannot be easily recast into an obstacle problem as can be done in the case of a constant source term. We present a simple formulation that penalizes the constraint directly on the gradient norm of the solution. We study its well-posedness, derive error estimates and present numerical results to illustrate the theory.

Keywords: variational inequalities; elastoplastic torsion problem; finite elements; penalty; error estimates.
2020 MSC: 65N15, 65N30.

---

1. Introduction

Problems involving variational inequalities represent various nonlinear phenomena which occur in mechanics and physics such as contact [14, 24], friction [14, 16, 24], plasticity [16, 23] and non-Newtonian flow [16, 27]. These examples may involve a constraint on the value of the unknown, or on its gradient, representing the two main categories of variational inequalities. We focus on an archetype example from the latter category, the elastoplastic torsion problem, as presented in, *e.g.*, [19, 20]. The aforementioned reference [19] presents a direct piecewise affine Lagrange finite element approximation of the variational inequality together with a convergence result (Theorem 3.3), and two error estimates in the H^1-norm, in dimension one (Theorem 3.4) and in dimension two (Theorem 3.5). The error estimate in dimension one is optimal ($\mathcal{O}(h)$), whereas the estimate in dimension two is suboptimal, as it is of order $\mathcal{O}(h^{\frac{1}{2}-\frac{1}{p}})$ for a source term in L^p, $p > 2$. Among the earliest results are weak and strong convergence results [26], and error estimates of order $\mathcal{O}(h)$ for the L^2-norm of the gradient of the solution and, under suitable restrictive assumptions, for mixed finite element approximations using $\mathbb{P}_1/\mathbb{P}_0$ finite elements [17] or Raviart-Thomas finite elements [3].

A common approach for solving the elastoplastic torsion problem relies on a reformulation of the problem as an "obstacle" problem which bounds the unknown value by a distance field to the boundary, to avoid directly constraining the gradient. This approach works only for a constant source term and will not generalize to more complex variational inequalities such as those encountered, *e.g.*, in contact mechanics with friction or three-dimensional elastoplasticity. Consequently in this paper, we focus on a torsion problem that has a nonconstant source term as in

*Corresponding author
Email addresses: `fchouly@cmat.edu.uy` (Franz Chouly), `tom.gustafsson@aalto.fi` (Tom Gustafsson), `patrick.hild@math.univ-toulouse.fr` (Patrick Hild)

[19, 20]. In a previous paper [13], we proposed direct finite element approximation of the variational inequality in the case of a constant source term, using piecewise affine, continuous, Lagrange finite elements, and in which the constraint involving the distance field is imposed nodewise. In the case of a convex domain, error estimates were established in any dimension $n = 1, 2, 3$, with an optimal error bound of $\mathcal{O}(h)$, for a regular enough solution. In the case of a nonconvex domain, an error bound of $\mathcal{O}(h^{\frac{3}{4}})$ was proven for a solution of regularity H^{α}, $\alpha \geq 7/4$.

In another previous paper [12], we proposed a new method that combines both the reformulation using the distance field, as in [13], and a Nitsche term that weakly incorporates the inequality constraint, following [9] and the related works on Nitsche's method for variational inequalities, see, *e.g.*, [11] and the references therein. For such Nitsche's method, we managed to derive optimal error estimates, using linear or quadratic finite elements, even in the case of a nonconvex domain, which improves the results of [13]. Moreover, the method is convenient to implement into modern finite element libraries. As an example, we provided numerical experiments which confirm the expected theoretical convergence rates. Besides, another possibilities combined with adaptive meshing are discussed in [2, 21].

These existing techniques cannot be applied when the source term is general and possibly nonconstant. Hence in this paper, we describe a simple penalty technique combined with low order Lagrange finite elements for approximating the elastoplastic torsion problem in its general form. Our penalization differs from the one suggested by R. Glowinski in [20], for technical reasons in order to obtain error estimates.

In section 2 we recall the exact problem whose solution is denoted u. The penalized problem is introduced in section 3 where we prove that it admits a unique solution u_ε. Section 4 is concerned with the strong formulation of the penalized problem and section 5 deals with a first nonoptimal (and nontrivial) error estimate for $u - u_\varepsilon$. The finite element problem is set in section 6 ; its unique solution is denoted u_{ε_h} and we prove optimal convergence rates for $u_\varepsilon - u_{\varepsilon_h}$. Note that surprisingly we are not able to obtain an error estimate for $u - u_{\varepsilon_h}$ depending only on u, ε and h. Nevertheless all the results and error estimates presented in the present paper are new to our knowledge. Finally, we show numerical experiments to illustrate the optimal rate of section 6, these numerical results confirm the theoretical rate and also suggest other expected rates (like L^2 error bounds) which we are not able to prove theoretically.

As usual, we denote by $H^s(\cdot)$, $s \in \mathbb{R}$, the Sobolev spaces. The usual norm of $H^s(D)$, $D \subset \mathbb{R}^n$, $n \geq 1$, is denoted by $\| \cdot \|_{s,D}$, and the corresponding semi-norm is denoted by $| \cdot |_{s,D}$. The space $H^1_0(D)$ is the subspace of functions in $H^1(D)$ with vanishing trace on ∂D. The letter C stands for a generic constant, independent of the mesh size h and the penalty parameter ε, with possibly different value at different occurences.

2. The elastoplastic torsion problem

Let $\Omega \subset \mathbb{R}^n$, $n \geq 1$, be an open bounded polytope, connected and with Lipschitz boundary. Let us denote $V := H^1_0(\Omega)$ and introduce the notation K, that represents the nonempty closed convex set of admissible stress potentials:

$$K := \{ v \in V \ : \ |\nabla v| \leq 1 \text{ a.e. in } \Omega \},$$

where $| \cdot |$ denotes the euclidean norm in $\mathbb{R}^n$.

We consider the variational inequality which, for $n = 2$ and a constant source term f, models the torsion of an infinitely long elastoplastic cylinder of cross section Ω and plasticity yield $r > 0$. To simplify the presentation, we assume that $r = 1$. The problem is to find the stress potential u such that

$$u \in K \ : \qquad a(u, v - u) \geq L(v - u) \qquad \forall\, v \in K, \tag{1}$$

where $a : V \times V \to \mathbb{R}$ is the bilinear form given by:

$$a(u, v) := \int_\Omega \nabla u \cdot \nabla v \quad \forall u, v \in V,$$

and

$$L(v) := \int_\Omega fv \quad \forall v \in V,$$

with $f \in L^2(\Omega)$. From Stampacchia's theorem we deduce that problem (1) admits a unique solution (see also, *e.g.*, [16, 19, 20, 25]), that is also the unique minimizer in K of the strongly convex continuous quadratic functional

$$\mathcal{J} : V \ni v \mapsto \frac{1}{2}a(v,v) - L(v) \in \mathbb{R}.$$

Remark 2.1. *We recall some regularity results for (1): if $\Omega \subset \mathbb{R}^n$ is open, bounded and convex, with Lipschitz boundary, and for $f \in L^p(\Omega)$ with $n < p < +\infty$, then $u \in W^{2,p}(\Omega) \cap \mathscr{C}^{1,\alpha}(\overline{\Omega})$, where $\alpha = 1 - n/p$ [8]. When the domain is nonconvex the $W^{2,p}(\Omega)$ regularity can be obtained but the boundary needs to be more regular ($\mathscr{C}^{1,1}$ more precisely, see [18]) so reentrant corners of polytopes are not allowed. When reentrant corners of polytopes are considered, the loss of $W^{2,p}$-regularity is only located near these corners [10].*

3. A penalized elastoplastic problem

First we present the continuous formulation, obtained through minimization of a functional with a penalty term. The weak form is then obtained as the first-order optimality condition associated with this functional. We can prove that the penalized problem is well-posed by establishing the Lipschitz continuity and the strong convexity of the functional.

3.1. Penalized formulation as a minimization problem and well-posedness

Let us first introduce the notation $[\cdot]_+$ for the positive part ($[x]_+ = \frac{1}{2}(x + |x|)$, for $x \in \mathbb{R}$) and

$$P(v) := |\nabla v| - 1, \tag{2}$$

for $v \in V$. As usual in constrained optimization problems, we can penalize the constraint on the gradient of the solution and find an approximate solution to problem (1) by minimizing on V the following functional:

$$\mathcal{J}_P(v) := \mathcal{J}(v) + \mathcal{J}_\varepsilon(v)$$

where $\mathcal{J}$ has been introduced above and where

$$\mathcal{J}_\varepsilon(v) := \frac{1}{2}\int_\Omega \frac{1}{\varepsilon}[P(v)]_+^2,$$

defined for $v \in V$, is the penalty functional to impose weakly the condition $|\nabla v| \leq 1$ on the gradient of v, with the penalty parameter $\varepsilon > 0$.

Remark 3.1. *Note that the formulation is different from the saddle point formulation presented in [20], where the authors use a term*

$$P_{GLT}(v) := |\nabla v|^2 - 1,$$

for $v \in V$ and a Lagrange multiplier to enforce the nonpositivity of $P_{GLT}(v)$. In the present paper we choose the definition of $P(v)$ given in (2) *in order to obtain error estimates which is an open question to our knowledge.*

The existence and uniqueness of a minimizer to $\mathcal{J}_P$ on V is stated below.

Theorem 3.1. *For every value of $\varepsilon > 0$, there exists one unique minimizer of $\mathcal{J}_P$ on V.*

Proof. We first prove that the functional $\mathcal{J}_\varepsilon$ is continuous on V. We have for any v, w in V:

$$\begin{aligned}
|[P(v)]_+^2 - [P(w)]_+^2| &= |[P(v)]_+ - [P(w)]_+|\,|[P(v)]_+ + [P(w)]_+| \\
&\leq |P(v) - P(w)|\,|[|\nabla v| - 1]_+ + [|\nabla w| - 1]_+| \\
&\leq ||\nabla v| - |\nabla w||\,||\nabla v| + |\nabla w|| \\
&\leq |\nabla(v-w)|\,||\nabla(v-w)| + 2|\nabla w|| \\
&= |\nabla(v-w)|^2 + 2|\nabla w|\,|\nabla(v-w)|.
\end{aligned}$$

In the previous calculation we pass from the first to the second line using property $|[x]_+ - [y]_+| \leq |x-y|$, we use the inequality $[|x|-1]_+ \leq |x|$ to obtain the expression in the third line and later two triangular norm inequalities to obtain the expression in the fourth line. Note that in the fifth line, all the vertical bars represent the canonical norm in $\mathbb{R}^n$. So we get by integration on Ω of $[P(v)]_+^2 - [P(w)]_+^2$ (or its opposite value $[P(w)]_+^2 - [P(v)]_+^2$) and Cauchy-Schwarz inequality, the same upper bound, i.e.:

$$\begin{aligned}
2\varepsilon|\mathcal{J}_\varepsilon(v) - \mathcal{J}_\varepsilon(w)| &\leq \|\nabla(v-w)\|_{0,\Omega}^2 + 2\int_\Omega |\nabla w|\,|\nabla(v-w)| \\
&\leq \|\nabla(v-w)\|_{0,\Omega}^2 + 2\|\nabla w\|_{0,\Omega}\|\nabla(v-w)\|_{0,\Omega} \\
&\leq \|v-w\|_{1,\Omega}^2 + 2\|w\|_{1,\Omega}\|v-w\|_{1,\Omega}.
\end{aligned}$$

The continuity of $\mathcal{J}_\varepsilon$ on V follows. Since $\mathcal{J}$ is continuous we deduce that $\mathcal{J}_P = \mathcal{J}_\varepsilon + \mathcal{J}$ is continuous on V. In addition $\mathcal{J}(v) \to +\infty$ as $\|v\|_{1,\Omega} \to +\infty$ and $\mathcal{J}_\varepsilon$ is nonnegative so $\mathcal{J}_P(v) \to +\infty$ as $\|v\|_{1,\Omega} \to +\infty$. Finally $\mathcal{J}_P$ is convex as the sum of two convex functions (the convexity of $\mathcal{J}_\varepsilon$ is easy to check). This ensures the existence of a minimizer for $\mathcal{J}_P$ (see [1, 7, 24]).

To obtain uniqueness we mention that $\mathcal{J}$ is smooth and its second order derivative is $V-$elliptic:

$$\mathcal{J}''(u; w, w) \geq c\|w\|_{1,\Omega}^2,$$

for $u, w \in V$. This implies strong convexity, i.e.,

$$\mathcal{J}\left(\frac{1}{2}u + \frac{1}{2}v\right) \leq \frac{1}{2}\mathcal{J}(u) + \frac{1}{2}\mathcal{J}(v) - \frac{c}{8}\|u-v\|_{1,\Omega}^2$$

for $u, v \in V$. So $\mathcal{J}_P$ is strongly convex as the sum of a convex function and a strongly convex function. This proves the existence of a unique minimizer of $\mathcal{J}_P$. □

3.2. Equivalent weak formulation and well-posedness

Next we derive the discrete weak form, as the first order optimality condition associated with $\mathcal{J}_P$. For this purpose, we need to compute the derivative of $\mathcal{J}_\varepsilon$. First we provide below the expression of P':

$$P'(v; w) = \frac{\nabla v \cdot \nabla w}{|\nabla v|}.$$

As a result, the discrete weak form reads: find $u_\varepsilon \in V$ satisfying

$$\mathcal{J}_P'(u_\varepsilon; v) = \int_\Omega \nabla u_\varepsilon \cdot \nabla v - \int_\Omega fv + \int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ \left(\frac{\nabla u_\varepsilon \cdot \nabla v}{|\nabla u_\varepsilon|}\right) = 0 \quad \forall v \in V. \tag{3}$$

To alleviate the notation, we set

$$s(x) := \frac{x}{|x|}$$

(and $s(0) = 0$ which is not a restrictive choice since $[P(u)]_+ = 0$ when ∇u=0) and we introduce the functional B such that

$$B(w; v) = \mathcal{J}_\varepsilon'(w; v) = \int_\Omega \frac{1}{\varepsilon}[P(w)]_+ s(\nabla w) \cdot \nabla v.$$

As a result, the weak form of the penalized problem reads: find $u_\varepsilon \in V$ such that

$$\int_\Omega \nabla u_\varepsilon \cdot \nabla v + \frac{1}{\varepsilon} B(u_\varepsilon; v) = \int_\Omega f v \quad \forall v \in V. \tag{4}$$

Theorem 3.2. *For every value of $\varepsilon > 0$, there exists one unique solution u_ε of* (4).

Proof. Of course the minimizing problem in Theorem 3.1 and (4) are equivalent. The main aim of this proof is to prove that problem (4) is well-posed by using different techniques (e.g., the next lemma).

Lemma 3.1. *For all v and w in V, we have*

$$\frac{1}{\varepsilon}\left(B(v; v-w) - B(w; v-w)\right) \geq \|\varepsilon^{-1/2}([P(v)]_+ - [P(w)]_+)\|_{0,\Omega}^2 \tag{5}$$

Proof of the lemma. We take arbitrary v and w. Then

$$\begin{aligned} B(v; v-w) - B(w; v-w) &= \int_\Omega [P(v)]_+ s(\nabla v) \cdot \nabla(v-w) - \int_\Omega [P(w)]_+ s(\nabla w) \cdot \nabla(v-w) \\ &= \int_\Omega [P(v)]_+ |\nabla v| + \int_\Omega [P(w)]_+ |\nabla w| \\ &\quad - \int_\Omega [P(v)]_+ \left(s(\nabla v) \cdot \nabla w\right) - \int_\Omega [P(w)]_+ \left(s(\nabla w) \cdot \nabla v\right). \end{aligned}$$

We use Cauchy Schwarz inequality, add and subtract terms, and get:

$$\begin{aligned} B(v; v-w) - B(w; v-w) &\geq \int_\Omega [P(v)]_+ |\nabla v| + \int_\Omega [P(w)]_+ |\nabla w| \\ &\quad - \int_\Omega [P(v)]_+ |\nabla w| - \int_\Omega [P(w)]_+ |\nabla v| \\ &= \int_\Omega [P(v)]_+ (|\nabla v| - 1) + \int_\Omega [P(w)]_+ (|\nabla w| - 1) \\ &\quad - \int_\Omega [P(v)]_+ (|\nabla w| - 1) - \int_\Omega [P(w)]_+ (|\nabla v| - 1). \end{aligned}$$

Then we reformulate and use the monotonicity of the positive part:

$$\begin{aligned} B(v; v-w) - B(w; v-w) &\geq \int_\Omega [P(v)]_+ P(v) + \int_\Omega [P(w)]_+ P(w) \\ &\quad - \int_\Omega [P(v)]_+ P(w) - \int_\Omega [P(w)]_+ P(v) \\ &= \int_\Omega ([P(v)]_+ - [P(w)]_+)(P(v) - P(w)) \\ &= \int_\Omega ([P(v)]_+ - [P(w)]_+)^2. \end{aligned}$$

Dividing by ε, we obtain (5).
End of the proof of the lemma.

Using the Riesz representation theorem, we define a (nonlinear) operator $A : V \to V$ with the following formula:

$$(Av, w)_{1,\Omega} := a(v, w) + \frac{1}{\varepsilon} B(v; w), \quad \forall v, w \in V,$$

where $(.,.)_{1,\Omega}$ denotes the inner product in $H^1(\Omega)$. Note that Problem (4) is well-posed if and only if A is a one-to-one operator.

Let $v, w \in V$, it follows from the V-ellipticity of $a(\cdot,\cdot)$ and the lemma above that there exists $\alpha > 0$ such that:

$$(Av - Aw, v - w)_{1,\Omega} \geq \alpha \|v - w\|^2_{1,\Omega}. \tag{6}$$

Let us also show that the operator A is hemicontinuous, which means that for all $v, w \in V$, the real function

$$[0,1] \ni t \mapsto \varphi(t) := (A(v - tw), w)_{1,\Omega} \in \mathbb{R}$$

is continuous. For $r, t \in [0,1]$, we have:

$$\begin{aligned}
|\varphi(t) - \varphi(r)| &= |(A(v - tw) - A(v - rw), w)_{1,\Omega}| \\
&\leq |a(v - tw, w) - a(v - rw, w)| \\
&\quad + \frac{1}{\varepsilon}\left|\int_\Omega ([P(v - tw)]_+ s(\nabla(v - tw)) - [P(v - rw)]_+ s(\nabla(v - rw))) \cdot \nabla w\right|
\end{aligned}$$

The bounding of the integral term, in particular the term $[P(v - tw)]_+ s(\nabla(v - tw)) - [P(v - rw)]_+ s(\nabla(v - rw))$ is a long work that is done hereafter (see section 6) when bounding the term $[P(u_\varepsilon)]_+ s(\nabla(u_\varepsilon)) - [P(u_{\varepsilon_h})]_+ s(\nabla(u_{\varepsilon_h}))$ in the proof of the error estimates for the penalized problem. To be brief, we obtain

$$|\varphi(t) - \varphi(r)| \quad \leq \quad |t - r| a(w, w) + 3|t - r| \|\nabla w\|^2_{0,\Omega}.$$

It follows that φ is Lipschitz, so continuous. The operator A is then hemicontinuous. Since (6) also holds, we can apply the Corollary 15 (p.126) of [4] to conclude that A is a one-to-one operator from V to V. That concludes the proof of the theorem. $\square$

4. Penalized strong formulation

In this section we derive the strong formulation corresponding to (4) that will be useful in the forthcoming section. We start from (3) and apply the Green formula

$$\int_\Omega \Phi \cdot \nabla v = -\int_\Omega (\text{div } \Phi) v,$$

so

$$\int_\Omega \nabla u_\varepsilon \cdot \nabla v + \int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon) \cdot \nabla v = -\int_\Omega \Delta u_\varepsilon v - \int_\Omega \frac{1}{\varepsilon} \text{div}\left([P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\right) v, \quad \forall v \in V.$$

Therefore the strong form of (4) is:

$$-\Delta u_\varepsilon - \frac{1}{\varepsilon} \text{div}\left([P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\right) = f.$$

Heuristically, when $|\nabla u_\varepsilon| < 1$, then $[P(u_\varepsilon)]_+ = 0$ and where the constraint is not activated we recover $-\Delta u_\varepsilon = f$. Conversely if $|\nabla u_\varepsilon| \geq 1$, then we get

$$[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon) = (|\nabla u_\varepsilon| - 1)\frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} = \nabla u_\varepsilon - \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|}.$$

Observe that, the strong form can be reformulated as

$$-\text{div}\left(\nabla u_\varepsilon + \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\right) = f. \tag{7}$$

Now consider the Lagrange multiplier formulation described in [5, 15], that reads:

$$-\mathrm{div}((2\mu + 1)\nabla u) = f, \tag{8}$$

with the conditions $\mu \geq 0$ and $\mu P(u) = 0$. If we compare the mixed formulation (8) with (7) the Lagrange multiplier μ is approximated by

$$\mu_\varepsilon = \frac{[P(u_\varepsilon)]_+}{2\varepsilon|\nabla u_\varepsilon|}.$$

Note that the expression does not seem to have an obvious limit when the penalty parameter ε vanishes. Finally note that a term similar to $\mu_\varepsilon \nabla(u_\varepsilon)$ appears in the forthcoming convergence analysis of u_ε towards u (see Theorem 5.1). Its limit is also not obvious.

5. Convergence of the continuous penalized formulation

This section is concerned with error bounds for $u - u_\varepsilon$ as ε vanishes. As far as we know such results are not available for the elastoplastic torsion problem. Suppose that $f \in L^p(\Omega)$, $p \geq 2$ and that u satisfies the (technical) assumption (S) in [15, Theorem 3.4]. We know from, e.g., [6] that the unique solution to the elastoplastic torsion problem satisfies: $u \in K \cap W^{2,p}(\Omega)$. Furthermore, the reference [15] ensures that there exists a multiplier $\mu \in L^\infty(\Omega)$ such that

$$\mu \geq 0 \ (i), \quad \mu P_{GLT}(u) = 0 \ (ii), \quad -\Delta u - f = 2\,\mathrm{div}(\mu \nabla u) \ (iii), \tag{9}$$

where the first two relationships hold a.e., and the last one holds in the distributional sense (and, in fact, a.e. due to the assumption on f). Note that P_{GLT} (see Remark 3.1) can be replaced by P above and that the regularity assumption on u implies $\Delta u \in L^2(\Omega)$.

Theorem 5.1. *Let u and u_ε be the solutions of (9) and (4) respectively. Suppose that $\mu \in L^\infty(\Omega)$. Then there is a constant C independent of ε such that:*

$$\|u - u_\varepsilon\|_{1,\Omega} + \sqrt{\varepsilon}\left\| \frac{1}{\varepsilon} s(\nabla u_\varepsilon)[P(u_\varepsilon)]_+ - 2\mu\nabla u \right\|_{0,\Omega} \leq C\sqrt{\varepsilon}\|\mu\|_{\infty,\Omega}\|u\|_{1,\Omega}.$$

Proof. We recall that $V = H_0^1(\Omega)$ and that the weak form of the penalized problem reads

$$u_\varepsilon \in V \ : \ \int_\Omega \nabla u_\varepsilon \cdot \nabla v + \frac{1}{\varepsilon} B(u_\varepsilon; v) = \int_\Omega f v \quad \forall v \in V. \tag{10}$$

We start with the Poincaré inequality, integration by parts and (10), and get

$$\begin{aligned}
& \alpha\|u - u_\varepsilon\|_{1,\Omega}^2 \\
\leq \ & \int_\Omega \nabla(u - u_\varepsilon) \cdot \nabla(u - u_\varepsilon) \\
= \ & \int_\Omega \nabla u \cdot \nabla(u - u_\varepsilon) - \int_\Omega \nabla u_\varepsilon \cdot \nabla(u - u_\varepsilon) \\
= \ & -\int_\Omega \Delta u(u - u_\varepsilon) - \int_\Omega f(u - u_\varepsilon) + \int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon) \cdot \nabla(u - u_\varepsilon) \\
= \ & \int_\Omega (-\Delta u - f)(u - u_\varepsilon) + \int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon) \cdot \nabla u - \int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon) \cdot \nabla u_\varepsilon \\
=: \ & T.
\end{aligned}$$

Using the last equation (iii) in (9) with the definition of the weak divergence, we get after integration by parts and the use of the second equation (ii) in (9):

$$
\begin{aligned}
T &= -\int_\Omega 2\mu\nabla u\cdot\nabla(u-u_\varepsilon)+\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u_\varepsilon.\\
&= -\int_\Omega 2\mu|\nabla u|^2+\int_\Omega 2\mu\nabla u\cdot\nabla u_\varepsilon+\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u_\varepsilon\\
&= -\int_\Omega 2\mu+\int_\Omega 2\mu\nabla u\cdot\nabla u_\varepsilon+\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u_\varepsilon. \quad (11)
\end{aligned}
$$

We transform (with the convention $s(0)=0$ and the notation $[\cdot]_-$ for the negative part, $[x]_-=\frac{1}{2}(-x+|x|)$, for $x\in\mathbb{R}$):

$$
\begin{aligned}
\nabla u_\varepsilon &= s(\nabla u_\varepsilon)(|\nabla u_\varepsilon|-1)+s(\nabla u_\varepsilon)\\
&= s(\nabla u_\varepsilon)P(u_\varepsilon)+s(\nabla u_\varepsilon)\\
&= s(\nabla u_\varepsilon)[P(u_\varepsilon)]_+-s(\nabla u_\varepsilon)[P(u_\varepsilon)]_-+s(\nabla u_\varepsilon).
\end{aligned}
$$

And we use this rewriting to transform the second term in the last expression (11) of T:

$$
\begin{aligned}
T &= -\int_\Omega 2\mu+\int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon)[P(u_\varepsilon)]_+-\int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon)[P(u_\varepsilon)]_-+\int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon).\\
&\quad+\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u_\varepsilon. \quad (12)
\end{aligned}
$$

First we consider the expression T_1 involving terms number 1, number 3 and number 4 in (12) and we gather terms number 3 and number 4 as follows:

$$
\begin{aligned}
T_1 &= -\int_\Omega 2\mu-\int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon)[P(u_\varepsilon)]_-+\int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon)\\
&= -\int_\Omega 2\mu+\int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon)(1-[P(u_\varepsilon)]_-).
\end{aligned}
$$

Because $0\le 1-[P(u_\varepsilon)]_-\le 1$, there holds (using (i) in (9))

$$
|2\mu\nabla u\cdot s(\nabla u_\varepsilon)(1-P(u_\varepsilon)_-)|\le 2\mu|\nabla u|\le 2\mu
$$

and so $T_1\le 0$.
Then we rewrite term number 6 in (12):

$$
\begin{aligned}
-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u_\varepsilon &= -\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+|\nabla u_\varepsilon|\\
&= -\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+(P(u_\varepsilon)+1)\\
&= -\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+^2-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+, \quad (13)
\end{aligned}
$$

where we used $a[a]_+=[a]_+^2$. As a result, there remains to bound in (12) the terms number 2 and number 5 complemented with both terms above in (13). This yields:

$$
T \le \int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon)[P(u_\varepsilon)]_++\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+-\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+^2.
$$

Since

$$
|[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u|\le [P(u_\varepsilon)]_+
$$

we deduce

$$\int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)\cdot\nabla u - \int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+ \leq 0$$

and finally

$$T \quad \leq \quad \int_\Omega 2\mu\nabla u\cdot s(\nabla u_\varepsilon)[P(u_\varepsilon)]_+ - \int_\Omega \frac{1}{\varepsilon}[P(u_\varepsilon)]_+^2. \tag{14}$$

We define

$$Q_\varepsilon = s(\nabla u_\varepsilon)[P(u_\varepsilon)]_+.$$

Now note that $|Q_\varepsilon|^2 = [P(u_\varepsilon)]_+^2$ and we continue bounding starting from (14) (we use $\mu\nabla u \in L^2(\Omega)$):

$$\begin{aligned} T &\leq \int_\Omega 2\mu\nabla u\cdot Q_\varepsilon - \int_\Omega \frac{1}{\varepsilon}Q_\varepsilon^2 \\ &= -\int_\Omega \left(\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u\right)\cdot\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon \\ &= -\int_\Omega \left(\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u\right)\cdot\left(\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u + 2\sqrt{\varepsilon}\mu\nabla u\right). \end{aligned}$$

Now we simplify:

$$\begin{aligned} &-\int_\Omega \left(\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u\right)\cdot\left(\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u + 2\sqrt{\varepsilon}\mu\nabla u\right) \\ &= -\int_\Omega \left|\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u\right|^2 - \int_\Omega \left(\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u\right)\cdot(2\sqrt{\varepsilon}\mu\nabla u). \end{aligned} \tag{15}$$

The first term in (15) that becomes:

$$-\int_\Omega \left|\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u\right|^2 \quad = \quad -\varepsilon\int_\Omega \left|\frac{1}{\varepsilon}Q_\varepsilon - 2\mu\nabla u\right|^2 =: -\varepsilon\|e\|_{0,\Omega}^2.$$

The second term in (15) can be transformed as follows and bounded using Cauchy–Schwarz and Young inequalities:

$$\begin{aligned} -\int_\Omega \left(\frac{1}{\sqrt{\varepsilon}}Q_\varepsilon - 2\sqrt{\varepsilon}\mu\nabla u\right)\cdot(2\sqrt{\varepsilon}\mu\nabla u) &= -\varepsilon\int_\Omega \left(\frac{1}{\varepsilon}Q_\varepsilon - 2\mu\nabla u\right)\cdot(2\mu\nabla u) \\ &\leq \varepsilon\|e\|_{0,\Omega}\|2\mu\nabla u\|_{0,\Omega} \\ &\leq \frac{\varepsilon}{2}\|e\|_{0,\Omega}^2 + 2\varepsilon\|\mu\nabla u\|_{0,\Omega}^2. \end{aligned}$$

We combine the above results and get:

$$\alpha\|u-u_\varepsilon\|_{1,\Omega}^2 + \frac{\varepsilon}{2}\|e\|_{0,\Omega}^2 \leq 2\varepsilon\|\mu\nabla u\|_{0,\Omega}^2.$$

and so

$$\|u-u_\varepsilon\|_{1,\Omega} + \sqrt{\varepsilon}\|e\|_{0,\Omega} \leq C\sqrt{\varepsilon}\|\mu\|_{\infty,\Omega}\|u\|_{1,\Omega}.$$

This ends the proof. □

Remark 5.1. *The two following observations can be made about the above result:*

1. *This error analysis seems to be the first for a penalized torsion problem. The convergence rate of order $\sqrt{\varepsilon}$ for $u-u_\varepsilon$ may be improved using more regularity on u and μ. Yet we do not see how a better regularity on u and μ could be used to improve the rate. Moreover such extra regularity result for μ does not seem to be available.*
2. *We are not able to obtain any convergence result for the $(Q_\varepsilon)/\varepsilon - 2\mu\nabla u$ term. Moreover it is not clear to us if it is really an error term or not.*

6. Finite element problem and convergence towards the penalized problem

This section introduces a simple finite element approximation for the penalized elastoplastic torsion problem followed by an error estimate.

6.1. Finite element problem

Let $V_h \subset V$ be a family of finite dimensional vector spaces indexed by h coming from a family $\mathcal{T}_h$ of triangulations (triangles in $2D$ and tetrahedra in $3D$) of the domain Ω ($h = \max_{T\in\mathcal{T}_h} h_T$ where h_T is the diameter of T). The family of triangulations is supposed regular, i.e., there exists $\sigma > 0$ such that $\forall T \in \mathcal{T}_h, h_T/\rho_T \leq \sigma$ where ρ_T denotes the radius of the inscribed ball in T. We choose standard continuous and piecewise affine functions, i.e.:

$$V_h = \left\{ v_h \in C(\overline{\Omega}) : v_h|_T \in \mathbb{P}_1(T), \forall T \in \mathcal{T}_h, v_h = 0 \text{ on } \partial\Omega \right\}.$$

The discrete version of the penalty method (4) reads:

$$u_{\varepsilon_h} \in V_h : a(u_{\varepsilon_h}, v_h) + \frac{1}{\varepsilon} B(u_{\varepsilon_h}; v_h) = L(v_h), \qquad \forall v_h \in V_h. \tag{16}$$

Using exactly the same argument as for the continuous problem in Theorem 3.2, we see that this problem admits one unique solution, which is the unique minimizer of $\mathcal{J}_P$ on the finite element space V^h.

6.2. Finite element error estimate for the penalized problem

When the solution u_ε to the continuous penalized problem is regular enough, we can derive the following a priori error bound:

Theorem 6.1. *Let u_ε and u_{ε_h} be the solutions to problems of (4) and (16), respectively. Suppose that $u_\varepsilon \in H^2(\Omega)$. Then there is a constant C independent of h and ε such that:*

$$\|u_\varepsilon - u_{\varepsilon_h}\|_{1,\Omega} + \|\varepsilon^{-1/2}[P(u_\varepsilon)]_+ - [P(u_{\varepsilon_h})]_+\|_{0,\Omega} \leq C(1+\varepsilon^{-1}) h \|u_\varepsilon\|_{2,\Omega}. \tag{17}$$

Proof. We start using the ellipticity of $a(\cdot,\cdot)$: let u_ε and u_{ε_h} be the solutions of Problem (4) and Problem (16), respectively, then

$$\begin{aligned}
\alpha\|u_\varepsilon - u_{\varepsilon_h}\|^2_{1,\Omega} &\leq a(u_\varepsilon - u_{\varepsilon_h}, u_\varepsilon - u_{\varepsilon_h}) \\
&= a(u_\varepsilon - u_{\varepsilon_h}, u_\varepsilon - v_h + v_h - u_{\varepsilon_h}) \\
&= a(u_\varepsilon - u_{\varepsilon_h}, u_\varepsilon - v_h) + a(u_\varepsilon - u_{\varepsilon_h}, v_h - u_{\varepsilon_h}), \qquad \forall v_h \in V_h.
\end{aligned}$$

As usual the first right term is handled with Cauchy-Schwarz and Young inequalities (we recall that C denotes a generic constant).

$$C\|u_\varepsilon - u_{\varepsilon_h}\|^2_{1,\Omega} \leq \|u_\varepsilon - v_h\|^2_{1,\Omega} + a(u_\varepsilon - u_{\varepsilon_h}, v_h - u_{\varepsilon_h}), \qquad \forall v_h \in V_h. \tag{18}$$

Now, let us have a look at the second term. We use (4) and (16), the conformity property $V_h \subset V$ and do the appropriate splittings: we add and subtract u_ε. For any $v_h \in V_h$:

$$\begin{aligned}
& a(u_\varepsilon - u_{\varepsilon_h}, v_h - u_{\varepsilon_h}) \\
=\ & a(u_\varepsilon, v_h - u_{\varepsilon_h}) - a(u_{\varepsilon_h}, v_h - u_{\varepsilon_h}) \\
=\ & \left(B(u_{\varepsilon_h}; v_h - u_{\varepsilon_h}) - B(u_\varepsilon; v_h - u_{\varepsilon_h})\right)/\varepsilon \\
=\ & \left(B(u_{\varepsilon_h}; v_h - u_\varepsilon) + B(u_{\varepsilon_h}; u_\varepsilon - u_{\varepsilon_h}) - B(u_\varepsilon; v_h - u_\varepsilon) - B(u_\varepsilon; u_\varepsilon - u_{\varepsilon_h})\right)/\varepsilon.
\end{aligned} \tag{19}$$

We first treat the following two terms (number 2 and number 4 above), using the definition of B and then the property (5), we get:

$$\left(B(u_{\varepsilon_h}; u_\varepsilon - u_{\varepsilon_h}) - B(u_\varepsilon; u_\varepsilon - u_{\varepsilon_h})\right)/\varepsilon \leq -\|\varepsilon^{-1/2}[P(u_\varepsilon)]_+ - [P(u_{\varepsilon_h})]_+\|^2_{0,\Omega}. \tag{20}$$

The technical term that remains is:

$$B(u_{\varepsilon_h}; v_h - u_\varepsilon) - B(u_\varepsilon; v_h - u_\varepsilon).$$

We evaluate it on each element T, and therefore we need to estimate

$$\int_T [P(u_{\varepsilon_h})]_+ s(\nabla u_{\varepsilon_h}) \cdot \nabla(v_h - u_\varepsilon) - [P(u_\varepsilon)]_+ s(\nabla u_\varepsilon) \cdot \nabla(v_h - u_\varepsilon). \tag{21}$$

We add and subtract $[P(u_\varepsilon)]_+ s(\nabla u_{\varepsilon_h})$ to the above expression and get

$$\int_T [([P(u_{\varepsilon_h})]_+ - [P(u_\varepsilon)]_+) s(\nabla u_{\varepsilon_h}) + [P(u_\varepsilon)]_+ (s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon))] \cdot \nabla(v_h - u_\varepsilon). \tag{22}$$

We call Y_T the integrand:

$$Y_T := [([P(u_{\varepsilon_h})]_+ - [P(u_\varepsilon)]_+) s(\nabla u_{\varepsilon_h}) + [P(u_\varepsilon)]_+ (s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon))] \cdot \nabla(v_h - u_\varepsilon). \tag{23}$$

First remark that some terms in Y_T are easy to bound, *e.g.*:

$$|[P(u_{\varepsilon_h})]_+ - [P(u_\varepsilon)]_+| \leq ||\nabla u_{\varepsilon_h}| - |\nabla u_\varepsilon|| \leq |\nabla(u_\varepsilon - u_{\varepsilon_h})| \tag{24}$$

and

$$|s(\nabla u_{\varepsilon_h})| \leq 1. \tag{25}$$

Now we have to estimate another term in Y:

$$|s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)|.$$

* If both $\nabla u_{\varepsilon_h} = 0$ and $\nabla u_\varepsilon = 0$ then this term is 0.

We consider all the remaining cases. To summarize there are three cases and our aim is to obtain a common upper bound of $|s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)|$, see (28).

* If $\nabla u_{\varepsilon_h} \neq 0$ and $\nabla u_\varepsilon = 0$, then

$$|s(\nabla u_{\varepsilon_h}) - s(\nabla u\epsilon)| = |s(\nabla u_{\varepsilon_h})| = 1 \leq 2 = \frac{2|\nabla u_{\varepsilon_h}|}{|\nabla u_{\varepsilon_h}|} = \frac{2|\nabla(u_\varepsilon - u_{\varepsilon_h})|}{|\nabla u_\varepsilon| + |\nabla u_{\varepsilon_h}|}.$$

* If $\nabla u_{\varepsilon_h} = 0$ and $\nabla u_\varepsilon \neq 0$, the same happens since the expression is symmetric in u_ε and u_{ε_h}.
* If $\nabla u_{\varepsilon_h} \neq 0$ and $\nabla u_\varepsilon \neq 0$, we first introduce the quantity Z:

$$\begin{aligned} Z &= (|\nabla u_\varepsilon| + |\nabla u_{\varepsilon_h}|)(s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)) \\ &= (|\nabla u_\varepsilon| + |\nabla u_{\varepsilon_h}|)\left(\frac{\nabla u_{\varepsilon_h}}{|\nabla u_{\varepsilon_h}|} - \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|}\right) \\ &= |\nabla u_\varepsilon| \frac{\nabla u_{\varepsilon_h}}{|\nabla u_{\varepsilon_h}|} - |\nabla u_{\varepsilon_h}| \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} - \nabla u_\varepsilon + \nabla u_{\varepsilon_h}. \end{aligned} \tag{26}$$

We use then the remarkable property

$$\left| |\nabla u_\varepsilon| \frac{\nabla u_{\varepsilon_h}}{|\nabla u_{\varepsilon_h}|} - |\nabla u_{\varepsilon_h}| \frac{\nabla u_\varepsilon}{|\nabla u_\varepsilon|} \right|^2 = |\nabla u_\varepsilon|^2 + |\nabla u_{\varepsilon_h}|^2 - 2\nabla u_\varepsilon \cdot \nabla u_{\varepsilon_h} = |\nabla(u_\varepsilon - u_{\varepsilon_h})|^2. \tag{27}$$

Now we conclude by using a triangular inequality in (26) and estimate (27):

$$|Z| = (|\nabla u_\varepsilon| + |\nabla u_{\varepsilon_h}|)|s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)| \leq 2|\nabla(u_\varepsilon - u_{\varepsilon_h})|.$$

So there always hold

$$|s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)| \le \frac{2|\nabla(u_\varepsilon - u_{\varepsilon_h})|}{|\nabla u_\varepsilon| + |\nabla u_{\varepsilon_h}|}. \tag{28}$$

Now we bound $\left|\int_T Y_T\right|$, *e.g.*, the term in (22).

Using the bounds in (24) and (25), we get :

$$\begin{aligned}
\left|\int_T Y_T\right| &\le \int_T |([P(u_{\varepsilon_h})]_+ - [P(u_\varepsilon)]_+)s(\nabla u_{\varepsilon_h}) + [P(u_\varepsilon)]_+(s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon))|\,|\nabla(v_h - u_\varepsilon)| \\
&\le \int_T |\nabla(u_\varepsilon - u_{\varepsilon_h})||\nabla(v_h - u_\varepsilon)| + |[P(u_\varepsilon)]_+(s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon))|\,|\nabla(v_h - u_\varepsilon)| \\
&\le \|\nabla(u_\varepsilon - u_{\varepsilon_h})\|_{0,T}\|\nabla(u_\varepsilon - v_h)\|_{0,T} + \int_T |[P(u_\varepsilon)]_+(s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon))|\,|\nabla(v_h - u_\varepsilon)|
\end{aligned}$$

We now consider two cases for estimating the last integral term.
* Case 1: suppose $\nabla u_{\varepsilon_h} \not\equiv 0$ on T (we recall that ∇u_{ε_h} is constant on the triangle T for $\mathbb{P}_1$ finite elements).

a. If $\nabla u_\varepsilon \neq 0$ (at point $x \in T$), then we use (28) and $|\nabla u_\varepsilon| + |\nabla u_{\varepsilon_h}| \ge |\nabla u_\varepsilon|$ from which we deduce

$$|s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)| \le \frac{2|\nabla(u_\varepsilon - u_{\varepsilon_h})|}{|\nabla u_\varepsilon| + |\nabla u_{\varepsilon_h}|} \le \frac{2|\nabla(u_\varepsilon - u_{\varepsilon_h})|}{|\nabla u_\varepsilon|}.$$

We can bound

$$\frac{[P(u_\varepsilon)]_+}{|\nabla u_\varepsilon|} = \frac{[|\nabla u_\varepsilon| - 1]_+}{|\nabla u_\varepsilon|} = \left[1 - \frac{1}{|\nabla u_\varepsilon|}\right]_+ \le 1 \tag{29}$$

and we obtain

$$[P(u_\varepsilon)]_+|s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)| \le [P(u_\varepsilon)]_+\frac{2|\nabla(u_\varepsilon - u_{\varepsilon_h})|}{|\nabla u_\varepsilon|} \le 2|\nabla(u_\varepsilon - u_{\varepsilon_h})|.$$

b. If $\nabla u_\varepsilon = 0$ (at point $x \in T$) then $[P(u_\varepsilon)]_+ = 0$ and still we have

$$[P(u_\varepsilon)]_+|s(\nabla u_{\varepsilon_h}) - s(\nabla u_\varepsilon)| \le 2|\nabla(u_\varepsilon - u_{\varepsilon_h})|.$$

Combining (sub)cases a. and b., we get:

$$\begin{aligned}
\left|\int_T Y_T\right| &\le \|\nabla(u_\varepsilon - u_{\varepsilon_h})\|_{0,T}\|\nabla(u_\varepsilon - v_h)\|_{0,T} + 2\int_T |\nabla(u_\varepsilon - u_{\varepsilon_h})||\nabla(v_h - u_\varepsilon)| \\
&\le 3\|\nabla(u_\varepsilon - u_{\varepsilon_h})\|_{0,T}\|\nabla(u_\varepsilon - v_h)\|_{0,T}.
\end{aligned} \tag{30}$$

* Case 2: suppose, otherwise, that $\nabla u_{\varepsilon_h} \equiv 0$ on T.

In this case $s(\nabla u_{\varepsilon_h}) = 0$ on T ; looking back at the definition of Y_T in (23), we get:

$$\begin{aligned}
\left|\int_T Y_T\right| &\le \int_T |[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)|\,|\nabla(v_h - u_\varepsilon)| \\
&= \int_X |[P(u_\varepsilon)]_+ s(\nabla u_\varepsilon)|\,|\nabla(v_h - u_\varepsilon)|
\end{aligned}$$

where X is the set of points $x \in T$ such that $\nabla u_\varepsilon(x) \neq 0$. Finally we reintroduce $s(\nabla u_{\varepsilon_h}) = 0$ on X and use estimates (28) and (29) which are valid on X : so

$$\begin{aligned}\left|\int_T Y_T\right| &\leq \int_X |[P(u_\varepsilon)]_+(s(\nabla u_\varepsilon) - s(\nabla u_{\varepsilon_h}))|\,|\nabla(v_h - u_\varepsilon)| \\ &\leq \int_X [P(u_\varepsilon)]_+ \frac{2|\nabla(u_\varepsilon - u_{\varepsilon_h})|}{|\nabla u_\varepsilon|}|\nabla(v_h - u_\varepsilon)| \\ &\leq \int_X 2|\nabla(u_\varepsilon - u_{\varepsilon_h})||\nabla(v_h - u_\varepsilon)| \\ &\leq 2\|\nabla(u_\varepsilon - u_{\varepsilon_h})\|_{0,X}\|\nabla(u_\varepsilon - v_h)\|_{0,X} \\ &\leq 2\|\nabla(u_\varepsilon - u_{\varepsilon_h})\|_{0,T}\|\nabla(u_\varepsilon - v_h)\|_{0,T}. \end{aligned} \tag{31}$$

Let us sum things up now. Putting together (18), (19), (20) and (21), we first have shown that:

$$\|u_\varepsilon - u_{\varepsilon_h}\|_{1,\Omega}^2 + \|\varepsilon^{-1/2}[P(u_\varepsilon)]_+ - [P(u_{\varepsilon_h})]_+\|_{0,\Omega}^2 \leq C\left(\|u_\varepsilon - v_h\|_{1,\Omega}^2 + \frac{1}{\varepsilon}\sum_T \int_T Y_T\right).$$

Since we obtained the bounds (30) and (31), we can write:

$$\frac{1}{\varepsilon}\left|\int_T Y_T\right| \leq \frac{3}{\varepsilon}\|\nabla(u_\varepsilon - u_{\varepsilon_h})\|_{0,T}\|\nabla(u_\varepsilon - v_h)\|_{0,T} \leq \frac{\beta}{2}\|u_\varepsilon - u_{\varepsilon_h}\|_{1,T}^2 + \frac{9}{2\beta\varepsilon^2}\|\nabla(u_\varepsilon - v_h)\|_{0,T}^2,$$

with $\beta > 0$, that holds in any situation. We make the sum on each triangle, use Young inequality and obtain:

$$\|u_\varepsilon - u_{\varepsilon_h}\|_{1,\Omega}^2 + \|\varepsilon^{-1/2}[P(u_\varepsilon)]_+ - [P(u_{\varepsilon_h})]_+\|_{0,\Omega}^2 \leq C\left(\|u_\varepsilon - v_h\|_{1,\Omega}^2 + \frac{1}{\varepsilon^2}\|\nabla(u_\varepsilon - v_h)\|_{0,\Omega}^2\right).$$

We take $v_h = \mathcal{I}_h u$, the Lagrange interpolation of u on V_h, use the standard interpolation estimate and take the square root to finally get:

$$\|u_\varepsilon - u_{\varepsilon_h}\|_{1,\Omega} + \|\varepsilon^{-1/2}[P(u_\varepsilon)]_+ - [P(u_{\varepsilon_h})]_+\|_{0,\Omega} \leq C(1 + \varepsilon^{-1})h\|u_\varepsilon\|_{2,\Omega},$$

which is (17). □

Remark 6.1. *Two comments on the above results are:*

1. *This error analysis seems to be the first for a discrete penalized torsion problem (for a fixed ε). The convergence rate of order h for $u_\varepsilon - u_{\varepsilon_h}$ is optimal. We do not know if the constant ε^{-1} in the error bound could be improved or not.*
2. *The error analysis for $u - u_{\varepsilon_h}$ is out of the scope of this paper. We do not have the technical tools to obtain such an estimate for this problem which is much more difficult than e.g. the unilateral contact problem where the analysis exists, see [14] and references therein.*

7. Numerical results

The following numerical experiments have been computed using scikit-fem [22]. The first and the second two-dimensional examples use the square domain $\Omega = (0,1)^2$. The first example in Figure 1 uses the loading $f(x,y) = 10x$. The corresponding convergence rates in H^1- and L^2-norms are given in Figure 2. We observe that the observed convergence rates are practically identical for all of the penalty parameter values $\varepsilon = 100, 10, 1, 0.1$. In particular, the observed convergence rate is $O(h)$ in the H^1-norm and $O(h^2)$ in the L^2-norm. The same conclusion holds for the second example which uses the piecewise-constant loading $f(x,y) = 10$ if $(x,y) \in (0.2, 0.8)^2$

and $f(x, y) = 0$ otherwise. The results of the second example are given in Figure 3, with the corresponding convergence rates in Figure 4. In the absence of an analytical solution, all reference solutions use a piecewise-quadratic finite element and a suitably refined mesh.

The third example uses the three-dimensional cubical domain $\Omega = (0, 1)^3$ and a direct generalization of the piecewise-constant loading, i.e. $f(x, y, z) = 10$ if $(x, y, z) \in (0.2, 0.8)^3$ and $f(x, y, z) = 0$ otherwise. For the three-dimensional example, we present results using the penalty parameter value $\varepsilon = 0.1$ only. The results of the third example are given in Figure 5. The observed convergence rate in the H^1-norm is again $\mathcal{O}(h)$ but slightly less than $\mathcal{O}(h^2)$ in the L^2-norm. Despite this, all of the presented experiments are consistent with the estimate (17).

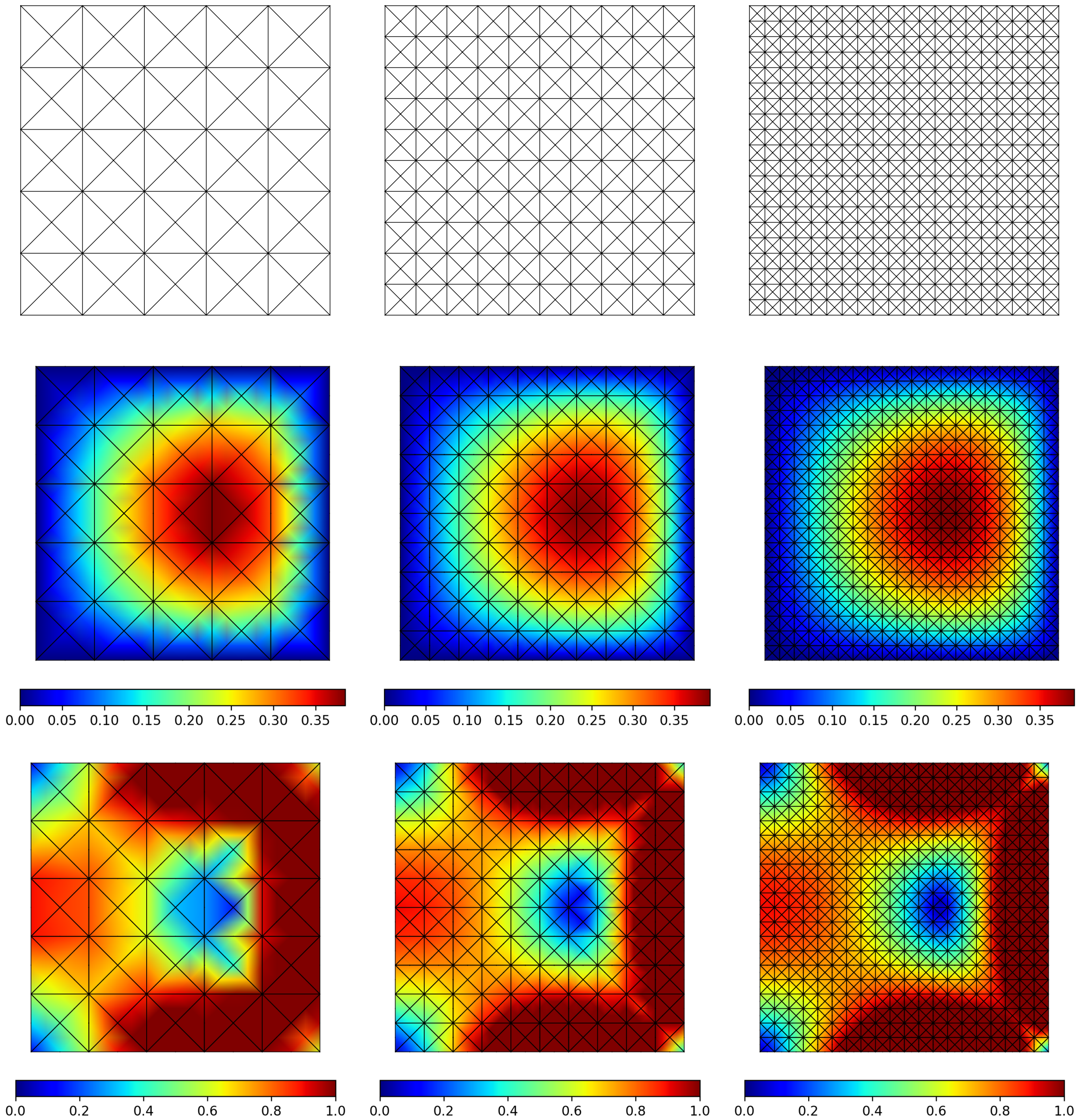

Figure 1: The sequence of meshes (top), the discrete solutions u_{ε_h} (middle) and the gradient lengths $|\nabla(u_{\varepsilon_h})|$ (bottom) when the loading is given by $f(x, y) = 10x$. The solutions in this figure have been computed using $\varepsilon = 100$.

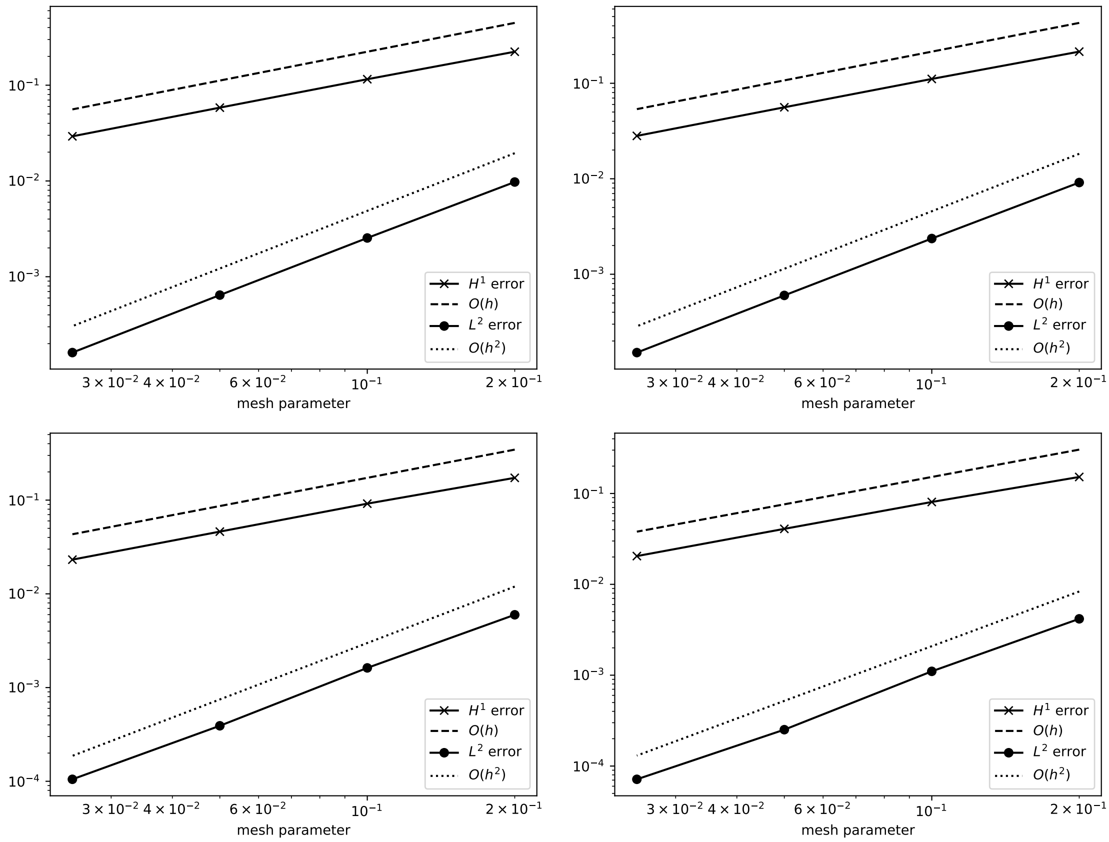

Figure 2: The convergence of the error $u_\varepsilon - u_{\varepsilon_h}$ when $f(x,y) = 10x$, using a sequence of decreasing penalty parameters $\varepsilon = 100, 10, 1, 0.1$. A reference solution is calculated using higher polynomial degree and a suitably refined mesh.

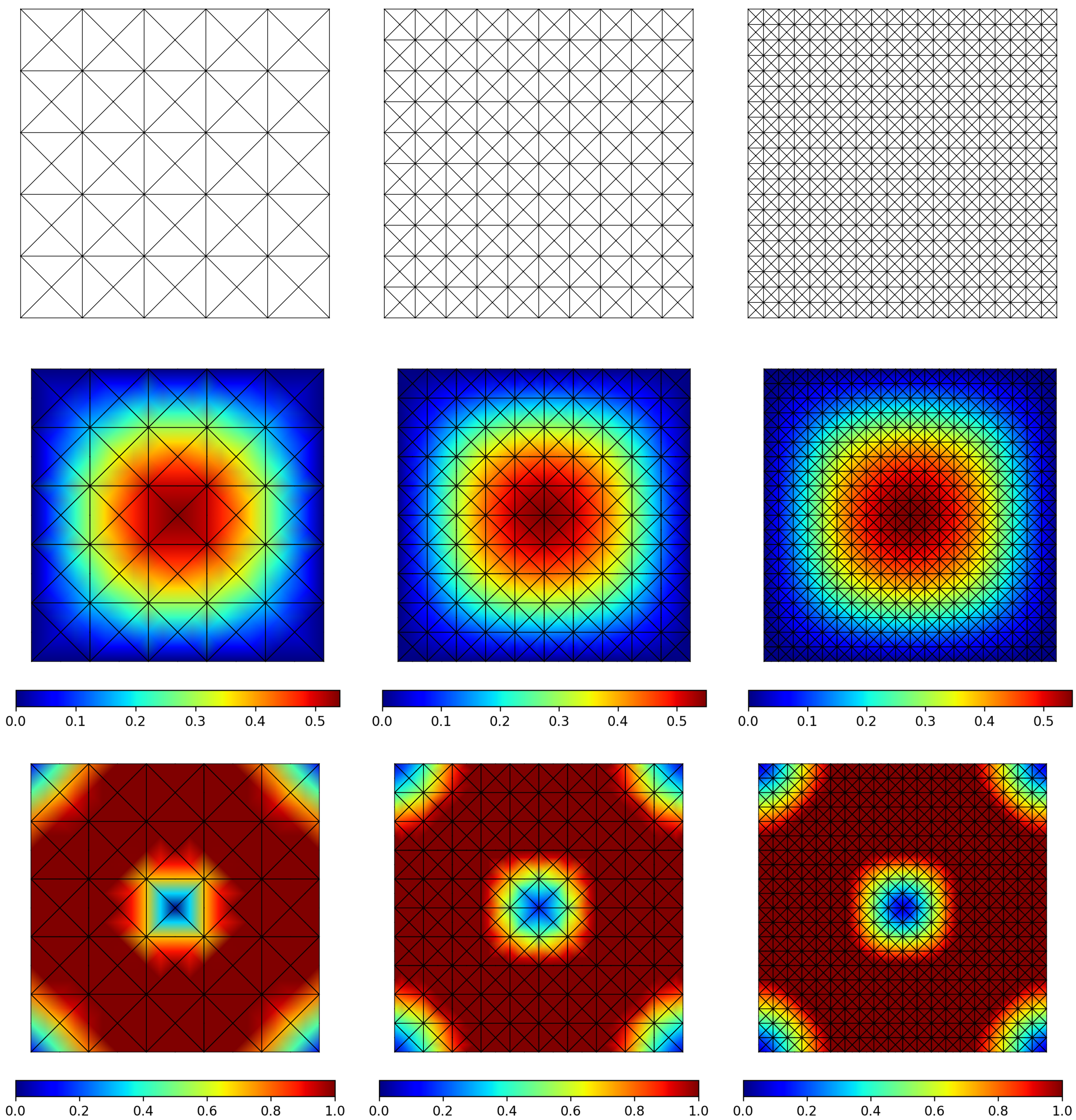

Figure 3: The sequence of meshes (top), the discrete solutions u_{ε_h} (middle) and the gradient lengths $|\nabla(u_{\varepsilon_h})|$ (bottom) when the loading is given by $f(x,y) = 10$ if $(x,y) \in (0.2, 0.8)^2$ and $f(x,y) = 0$ otherwise. The solutions in this figure have been computed using $\varepsilon = 100$.

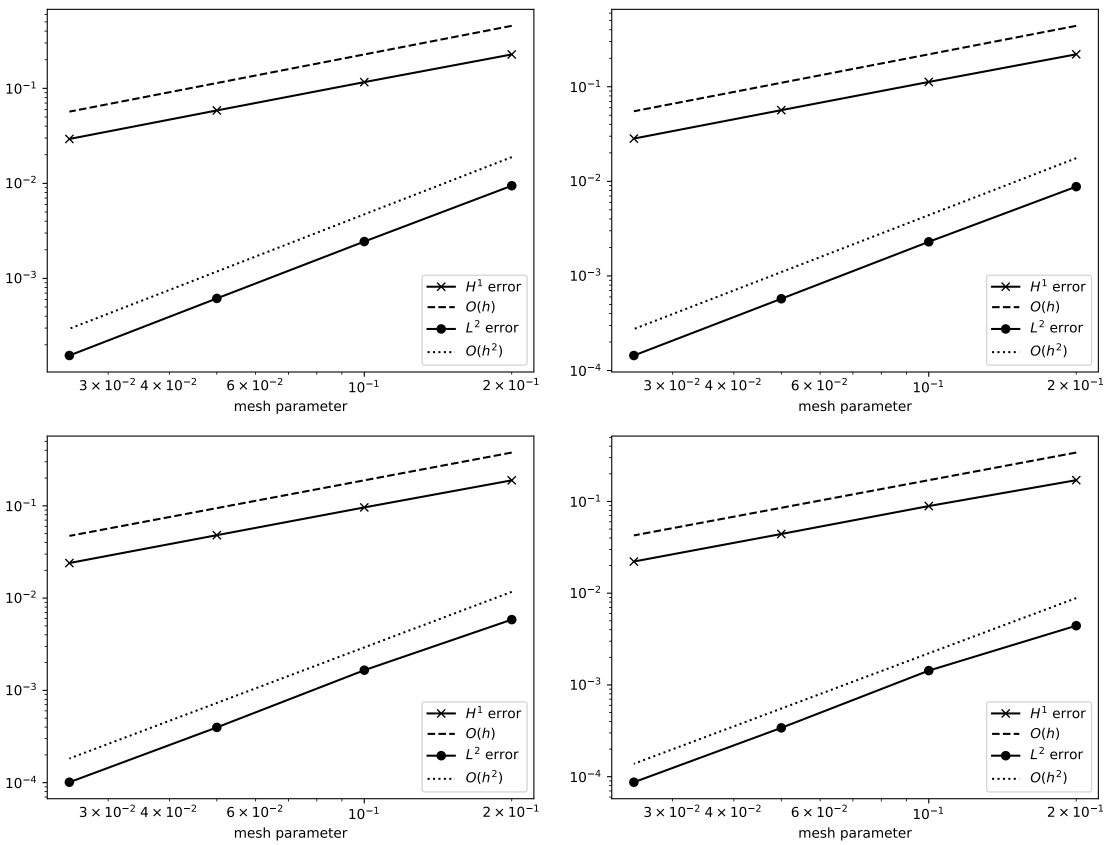

Figure 4: The convergence of the error $u_\varepsilon - u_{\varepsilon_h}$ when $f(x,y) = 10$ if $(x,y) \in (0.2,0.8)^2$ and $f(x,y) = 0$ otherwise, using a sequence of decreasing penalty parameters $\varepsilon = 100, 10, 1, 0.1$. A reference solution is calculated using higher polynomial degree and a suitably refined mesh.

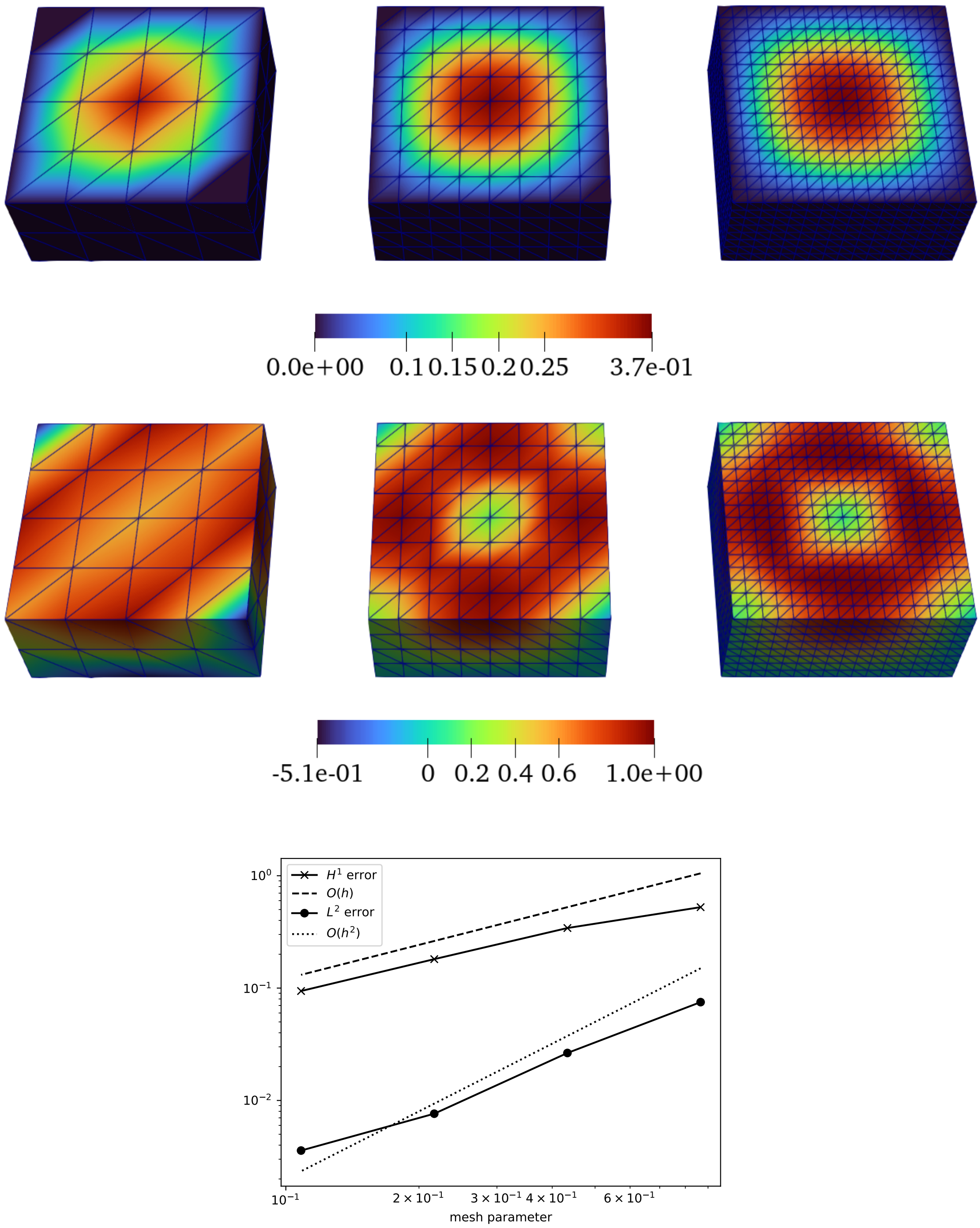

Figure 5: The sequence of discrete solutions (top), the gradient lengths (middle) and the convergence of the error (bottom) when the loading is given by $f(x, y, z) = 10$ if $(x, y, z) \in (0.2, 0.8)^3$ and $f(x, y, z) = 0$ otherwise. The solutions in this figure have been computed using $\varepsilon = 0.1$.

References

[1] G. Allaire. *Numerical analysis and optimization. An introduction to mathematical modelling and numerical simulation. Translation from the French by Alan Craig.* Numer. Math. Sci. Comput. Oxford: Oxford University Press, 2007.

[2] S. Bartels and A. Kaltenbach. Exact error control for variational problems via convex duality and explicit flux reconstruction. *Advances in Applied Mechanics (AAMS)*, 58:281–361, 2024. `https://doi.org/10.1016/bs.aams.2024.04.001`.

[3] A. Bermúdez de Castro López. A mixed method for the elastoplastic torsion problem. *IMA J. Numer. Anal.*, 2(3):325–334, 1982.

[4] H. Brézis. Équations et inéquations non linéaires dans les espaces vectoriels en dualité. *Ann. Inst. Fourier (Grenoble)*, 18(fasc. 1):115–175, 1968.

[5] H. Brézis. Multiplicateur de lagrange en torsion élasto-plastique. *Arch. Rational Mech. Anal.*, 49:32–40, 1972.

[6] H. Brézis. Problèmes unilatéraux. *J. Math. Pures Appl. (9)*, 51:1–168, 1972.

[7] H. Brézis. *Analyse fonctionnelle, théorie et applications.* Collection Mathématiques appliquées pour la maîtrise. Masson, Paris, New York, Barcelone, Milan, Mexico, Sao Paolo, second edition, 1987.

[8] H. Brézis and G. Stampacchia. Sur la régularité de la solution d'inéquations elliptiques. *Bull. Soc. Math. France*, 96:153–180, 1968.

[9] E. Burman, P. Hansbo, M. G. Larson, and R. Stenberg. Galerkin least squares finite element method for the obstacle problem. *Comput. Meth. Appl. Mech. Engrg.*, 313:362–374, September 2017.

[10] L. A. Caffarelli and A. Friedman. The free boundary for elastic-plastic torsion problems. *Trans. Amer. Math. Soc.*, 252:65–97, 1979.

[11] F. Chouly, M. Fabre, P. Hild, R. Mlika, J. Pousin, and Y. Renard. An overview of recent results on Nitsche's method for contact problems. In *Geometrically unfitted finite element methods and applications*, volume 121 of *Lect. Notes Comput. Sci. Eng.*, pages 93–141. Springer, Cham, 2017.

[12] F. Chouly, T. Gustafsson, and P. Hild. A Nitsche method for the elastoplastic torsion problem. *ESAIM Math. Model. Numer. Anal.*, 57(3):1731–1746, 2023.

[13] F. Chouly and P. Hild. On a finite element approximation for the elastoplastic torsion problem. *Appl. Math. Lett.*, 132:Paper No. 108191, 6, 2022.

[14] F. Chouly, P. Hild, and Y. Renard. *Finite element approximation of contact and friction in elasticity*, volume 48 of *Adv. Mech. Math.* Cham: Birkhäuser, 2023.

[15] P. Daniele, S. Giuffrè, A. Maugeri, and F. Raciti. Duality theory and applications to unilateral problems. *J. Optim. Theory Appl.*, 162(3):718–734, 2014.

[16] G. Duvaut and J.-L. Lions. *Inequalities in mechanics and physics*, volume 219 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin-New York, 1976.

[17] R. S. Falk and B. Mercier. Error estimates for elasto-plastic problems. *RAIRO Anal. Numér.*, 11(2):135–144, 219, 1977.

[18] C. Gerhardt. Regularity of solutions of nonlinear variational inequalities with a gradient bound as constraint. *Arch. Rational Mech. Anal.*, 58(4):309–315, 1975.

[19] R. Glowinski. *Lectures on numerical methods for nonlinear variational problems*, volume 65 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics.* Tata Institute of Fundamental Research, Bombay; Springer-Verlag, Berlin-New York, 1980.

[20] R. Glowinski, J.-L. Lions, and R. Trémolières. *Numerical analysis of variational inequalities*, volume 8 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam-New York, 1981.

[21] T. Gustafsson. Adaptive finite elements for obstacle problems. *Advances in Applied Mechanics (AAMS)*, 58:191–229, 2024. `https://doi.org/10.1016/bs.aams.2024.03.004`.

[22] T. Gustafsson and G. D. McBain. scikit-fem: A Python package for finite element assembly. *Journal of Open Source Software*, 5(52):2369, August 2020.

[23] W. Han and B. D. Reddy. *Plasticity. Mathematical theory and numerical analysis.*, volume 9 of *Interdiscip. Appl. Math.* New York, NY: Springer, 2nd ed. edition, 2013.

[24] N. Kikuchi and J. T. Oden. *Contact problems in elasticity: a study of variational inequalities and finite element methods*, volume 8 of *SIAM Studies in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1988.

[25] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications*, volume 88 of *Pure and Applied Mathematics.* Academic Press, Inc., New York-London, 1980.

[26] K. Mouallif. Approximation du problème de la torsion élasto-plastique d'une barre cylindrique par régularisation et discrétisation d'un problème inf-sup sur $H_0^1(\Omega) \times L_+^\infty(\Omega)$. *Travaux Sém. Anal. Convexe*, 12(1):exp. no. 1, 24, 1982.

[27] P. Saramito. *Complex fluids. Modeling and algorithms*, volume 79 of *Math. Appl. (Berl.).* Cham: Springer, 2016.