

# A FEM approximation of a two-phase obstacle problem and its a posteriori error estimate



Farid Bozorgnia<sup>a</sup>, Jan Valdman<sup>b,c,\*</sup>

<sup>a</sup> Department of Mathematics, Instituto Superior Técnico, Av. Rovisco Pais, P-1049-001 Lisboa, Portugal

<sup>b</sup> Institute of Mathematics and Biomathematics, Faculty of Science, University of South Bohemia, Branišovská 31, České Budějovice, CZ-37005, Czech Republic

<sup>c</sup> Institute of Information Theory and Automation of the CAS, Pod vodárenskou věží 4, CZ-18208 Praha 8, Czech Republic

## ARTICLE INFO

### Article history:

Received 4 June 2016

Received in revised form 5 October 2016

Accepted 27 November 2016

Available online 20 December 2016

### Keywords:

A free boundary problem

A posteriori error analysis

Finite element method

Dual problem

Discretization

Computer implementation

## ABSTRACT

This paper is concerned with the two-phase obstacle problem, a type of a variational free boundary problem. We recall the basic estimates of Repin and Valdman (2015) and verify them numerically on two examples in two space dimensions. A solution algorithm is proposed for the construction of the finite element approximation to the two-phase obstacle problem. The algorithm is not based on the primal (convex and nondifferentiable) energy minimization problem but on a dual maximization problem formulated for Lagrange multipliers. The dual problem is equivalent to a quadratic programming problem with box constraints. The quality of approximations is measured by a functional a posteriori error estimate which provides a guaranteed upper bound of the difference of approximated and exact energies of the primal minimization problem. The majorant functional in the upper bound contains auxiliary variables and it is optimized with respect to them to provide a sharp upper bound. A space density of the nonlinear related part of the majorant functional serves as an indicator of the free boundary.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

A free boundary problem is a partial differential equation where the equation changes qualitatively across a level set of the equation solution  $u$  so the part of the domain where the equation changes is a priori unknown. A general form of elliptic free boundary problems can be written as

$$\Delta u = f(x, u, \nabla u) \quad \text{in } \Omega, \quad (1)$$

where the right hand side term is piecewise continuous, having jumps at some values of the arguments  $u$  and  $\nabla u$ . Here  $\Omega$  is a bounded open subset of  $\mathbb{R}^n$  with smooth boundary and Dirichlet boundary conditions are considered. In this paper we are concerned about the particular elliptic free boundary problem

$$\begin{cases} \Delta u = \alpha_+ \chi_{\{u>0\}} - \alpha_- \chi_{\{u<0\}} & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \quad (2)$$

Here,  $\chi_A$  denotes the characteristic function of the set  $A$ ,  $\alpha_{\pm} : \Omega \rightarrow \mathbb{R}$  are positive and Lipschitz continuous functions and  $g \in W^{1,2}(\Omega) \cap L^\infty(\Omega)$  and  $g$  changes sign on  $\partial\Omega$ .

\* Corresponding author.

E-mail address: [jan.valdman@gmail.com](mailto:jan.valdman@gmail.com) (J. Valdman).

The boundary

$$(\partial\{x \in \Omega : u(x) > 0\} \cup \partial\{x \in \Omega : u(x) < 0\}) \cap \Omega$$

is called the free boundary. Properties of the solution of the two-phase obstacle problem, regularity of solution and free boundary have been studied in [1,2]. It is known that the differential equation from (2) represents the Euler–Lagrange equation corresponding to the minimizer of the functional

$$J(v) = \int_{\Omega} \left( \frac{1}{2} |\nabla v|^2 + \alpha_+ \max(v, 0) + \alpha_- \max(-v, 0) \right) dx \quad (3)$$

over the affine space

$$K = \{v \in W^{1,2}(\Omega) : v - g \in W_0^{1,2}(\Omega)\}. \quad (4)$$

The functional  $J : K \rightarrow \mathbb{R}$  is convex, coercive on  $K$  and weakly lower semi-continuous, hence the minimum of  $J$  is attained at some  $u \in K$ . The following minimization problem is therefore uniquely solvable.

**Problem 1 (Primal Problem).** Find  $u \in K$  such that

$$J(u) = \inf_{v \in K} J(v). \quad (5)$$

Note that if we let  $\alpha^- = 0$ , and assume that  $g$  is nonnegative on the boundary, then we obtain the well-known one-phase obstacle problem, see e.g. [3,4].

**Remark 1.** The problem (2) describes a complete reaction of two substances coming into contact with an elastic membrane, and sometimes is referred to as the “two-phase membrane” problem. If densities of substances are given as  $\rho_1$  and  $\rho_2$  and the elastic membrane has the density  $\rho_m$  satisfying  $\rho_1 > \rho_m > \rho_2$ , then  $\alpha_+$  is proportional to the difference  $\rho_1 - \rho_m$  and  $\alpha_-$  is proportional to the difference  $\rho_m - \rho_2$ , see [5,2]. In minimization problem (3), if instead of zero obstacle, we consider  $q \in C^2(\Omega)$ , then minimizer satisfies the following problem which generalizes (2) for a non-flat obstacle case.

$$\begin{cases} \Delta u = \alpha_+ \chi_{\{u>q\}} - \alpha_- \chi_{\{u<q\}} + \Delta q \chi_{\{u=q\}} & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \quad (6)$$

There are numerous papers on approximations and error analysis for the one-phase obstacle problem in terms on variational inequalities [6,7]. In [8] a sharp  $L^\infty$  error estimate for semilinear elliptic problems with free boundaries is given. For obstacle problem and combustion problems, the author uses regularization of penalty term combined with piecewise linear finite elements on a triangulation and then shows that the method is accurate in  $L^\infty$ . Using non-degeneracy property of one-phase obstacle problem, a sharp interface error estimate is derived. In [9] error estimates for the finite element approximation of the solution and free boundary of the obstacle problem are presented. Also an optimal error analysis for the thin obstacle problem is derived.

Recently, the numerical approximation of the two-phase membrane problem has attracted much interests. Most approximations are based on the finite difference methods. In [10] different methods to approximate the solution are presented. The first method is based on properties of the given free boundary problem and exploit the disjointness of positive and negative parts of the solution. Regularization method and error estimates are given. The a priori error gives a computable estimate for gradient of the error for regularized solutions in the  $L^2$ . In [11], the authors rewrite the two phase obstacle problem in an equivalent min–max formula then for this new form they introduce the notion of viscosity solution. Discretization of the min–max formula yields a certain linear approximation system. The existence and uniqueness of the solution of the discrete nonlinear system are shown. Also in [12] the author presents a finite difference approximation for a parabolic version of the two-phase membrane problem.

A finite element scheme for solving obstacle problems in divergence form is given in [13]. The authors reformulate the obstacle in terms of an  $L^1$  penalty on the variational problem. The reformulation is an exact regularizer in the sense that for large penalty parameter, it can recover the exact solution. They applied the scheme to approximate classical elliptic obstacle problems, the two-phase membrane problem and the Hele–Shaw model.

We propose a different finite element scheme for solving the two-phase obstacle problem based on the dual maximization problem for Lagrange multipliers. The main focus of the paper is the verification of a posteriori error estimates developed in [14]. For a FEM approximation  $v$ , we explicitly compute the upper bound of the difference  $J(v) - J(u)$  of the approximate energy  $J(v)$  and of the exact unknown minimal energy  $J(u)$ . Since this upper bound is guaranteed we automatically have a lower bound of the exact energy  $J(u)$ . The studied a posteriori error estimates also provide the approximate indication of the exact free boundary. This is demonstrated on two numerical tests in two space dimensions. A MATLAB code is freely available for own testing.

The structure of paper is as follows. In Section 2, we present an overview of basic concepts and mathematical background and recall energy and majorant estimates of [14]. Section 3 deals with discretization using finite elements: construction of the FEM approximation (Algorithm 1) and the optimization of the functional majorant (Algorithm 2). Section 4 reports on numerical examples and Section 5 concludes the work.

## 2. Mathematical background and estimates

Elements of convex analysis are used throughout this paper, in particular the duality method by conjugate functions [15]. For reader's convenience, let us summarize the basic notation used in what follows:

$n = 1, 2, 3$ dimension of the problem,	$\mathcal{T}_h$ uniform regular triangular mesh with mesh size $h$ , $\mathcal{N}_I, \mathcal{N}_D$
$\alpha_{\pm} \geq 0$ problem coefficients,	internal and Dirichlet nodes, $I, D$ their indices, $ \mathcal{E} ,  \mathcal{N} ,  \mathcal{T} $
$v^{\pm}$ positive and negative parts of the function,	number of edges, of nodes and of triangles,
$J : V \times Q \rightarrow \mathbb{R}$ primal functional to be minimized,	$K_h, Q_h^*, \Lambda_h^*$ finite element approximation spaces on $\mathcal{T}_h$
$J^* : Q^* \rightarrow \mathbb{R}$ conjugate functional to be maximized,	$I^* : \Lambda_h^* \rightarrow \mathbb{R}$ discrete dual energy to be maximized,
$J_{\mu}(\cdot)$ perturbed functional with multiplier $\mu$ ,	$\mathbb{K}$ stiffness matrix in $K_h$ ,
$u, v \in K$ exact and arbitrary minimizers of $J$ ,	$\mathbb{K}_{I,I}, \mathbb{K}_{I,D}, \mathbb{K}_{D,D}$ its subblocks with respect to $I$ and $D$ ,
$p^*, q^* \in Q^*$ exact and arbitrary maximizers of $J^*$ ,	$\mathbb{M}$ generalized mass matrix ( $L^2$ - product of $K_h$ and $\Lambda_h$ ),
$\lambda, \mu \in \Lambda$ exact and arbitrary multipliers,	$\mathbb{M}_I, \mathbb{M}_D$ its subblocks with respect to $I$ and $D$ ,
$D_J(\cdot, \cdot) : K \times Q^* \rightarrow \mathbb{R}$ compound functional,	$\lambda, \mu, \mathbf{v}, \mathbf{u}_{\lambda}$ discrete vectors,
$M_+ : K \times \mathbb{R}_+ \times Y^* \times \Lambda \rightarrow \mathbb{R}$ majorant functional,	$\mathbf{v}_I, \mathbf{v}_D, \mathbf{u}_{\lambda_I}, \mathbf{u}_{\lambda_D}$ subvectors with respect to $I$ and $D$ ,
$\eta^* \in Y^*$ flux variable in $M_+$ approximating $p^*$ ,	$u_{ref}$ reference solution,
$C_{\Omega}$ constant from the generalized Friedrichs inequality,	$M_{+1}, M_{+2}, M_{+3}$ majorant functional subparts,

List 1. Summary of the basic notation used through out this paper.

Let  $V$  and  $Q$  be two normed spaces,  $V^*$  and  $Q^*$  their dual spaces and let  $\langle \cdot, \cdot \rangle$  denote the duality pairing. Assume that there exists a continuous linear operator  $l$  from  $V$  to  $Q$ ,  $l \in \mathcal{L}(V, Q)$ . The adjoint operator  $l^* \in \mathcal{L}(Q^*, V^*)$  of the operator  $l$  is defined through the relation

$$\langle l^* q^*, v \rangle = \langle q^*, lv \rangle \quad \forall v \in V, q^* \in Q^*.$$

Let  $J : V \times Q \rightarrow \bar{\mathbb{R}}$  be a convex functional mapping in the space of extended reals  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ . Consider the minimization problem

$$\inf_{v \in V} J(v, lv) \tag{7}$$

and its dual conjugate problem

$$\sup_{q^* \in Q^*} [-J^*(l^* q^*, -q^*)], \tag{8}$$

where the convex conjugate function of  $J$  is given by

$$J^*(v^*, q^*) = \sup_{v \in V, q \in Q} [\langle v, v^* \rangle + \langle q, q^* \rangle - J(v, q)], \quad v^* \in V, q^* \in Q^*.$$

The relation between (7) and (8) is stated in the following theorem.

**Theorem 1** (Theorem 2.38 of [16]). Assume that  $V$  is a reflexive Banach space and  $Q$  is a normed vector space, and let  $l \in \mathcal{L}(V, Q)$ . Let  $J : V \times Q \rightarrow \bar{\mathbb{R}}$  be proper lower semi continuous, strictly convex such that

1. There exists  $v_0 \in V$ , such that  $J(v_0, lv_0) < \infty$  and  $q \mapsto J(v_0, q)$  is continuous at  $lv_0$ .
2.  $J(v, lv) \rightarrow +\infty$ , as  $\|v\| \rightarrow \infty$ ,  $v \in V$ .

Then problem (7) has a solution  $u \in V$  also problem (8) has a solution  $p^* \in Q^*$ , and

$$J(u, lu) = -J^*(l^* p^*, -p^*). \tag{9}$$

In the case that the function  $J$  is of a separated form, i.e.,

$$J(v, q) = F(v) + G(q) \quad v \in V, q \in Q,$$

then the conjugate of  $J$  is

$$J^*(v^*, q^*) = F^*(v^*) + G^*(q^*),$$

where  $F^*$  and  $G^*$  are the conjugate functions of  $F$  and  $G$ , respectively. To calculate the conjugate function when the functional is defined by an integral, we use the following theorems which can be found in [16].

**Theorem 2** (Theorem 2.35 of [16]). Assume  $h : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$  is a Carathéodory function with  $h \in L^1(\Omega)$  and suppose

$$G(q) = \int_{\Omega} h(x, q(x)) \, dx.$$

Then the conjugate function of  $G$  is

$$G^*(q^*) = \int_{\Omega} h^*(x, q^*(x)) \, dx \quad \forall q^* \in Q^*,$$

where

$$h^*(x, y) = \sup_{\xi \in \mathbb{R}^n} [y \cdot \xi - h(x, \xi)].$$

The compound functional  $D_J(v, q; v^*, q^*) : (V \times Q) \times (V^* \times Q^*) \rightarrow \bar{\mathbb{R}}$  is defined by

$$D_J(v, q; v^*, q^*) := J(v, q) + J^*(v^*, q^*) - \langle v, v^* \rangle - \langle q, q^* \rangle. \tag{10}$$

It holds  $D_J(v, q; v^*, q^*) \geq 0$  for all  $(v, q) \in V \times Q, (v^*, q^*) \in V^* \times Q^*$  and  $D_J(v, q; v^*, q^*) = 0$  only if the function  $(v, q)$  belongs to set of subdifferential  $\partial J^*(v^*, q^*)$  and  $(v^*, q^*)$  belongs to set of subdifferential  $\partial J(v, q)$ , see Proposition 1.2 of [17].

2.1. Energy identity

For simplicity of notation, we introduce the positive and negative parts of a function  $v$

$$v^+ := \max(v, 0), \quad v^- := \max(-v, 0),$$

so it holds  $v = v^+ - v^-$  and  $|v| = v^+ + v^-$ . The Euclidean norm in  $\mathbb{R}^n$  is denoted by  $|\cdot|$ . We write (3) in the form  $J(v) = F(v) + G(lv)$ , where

$$F(v) := \int_{\Omega} (\alpha_+ v^+ + \alpha_- v^-) \, dx, \quad G(lv) := \frac{1}{2} \int_{\Omega} \nabla v \cdot \nabla v \, dx \tag{11}$$

and  $l : K \rightarrow Q = L^2(\Omega, \mathbb{R}^d)$  is the gradient operator  $lv = \nabla v$ .

**Remark 2.** Theorem 1 assumes  $J : V \times Q \rightarrow \bar{\mathbb{R}}$ , where  $V$  is a normed space. It can be shown that all results are also valid for  $J : K \times Q \rightarrow \bar{\mathbb{R}}$  from above.

The corresponding dual conjugate functionals are

$$F^*(v^*) = \int_{\Omega} h^*(v^*) \, dx, \quad G^*(q^*) = \frac{1}{2} \int_{\Omega} q^* \cdot q^* \, dx, \tag{12}$$

where  $h^*(z^*)(x) = 0$  for  $z^*(x) \in [-\alpha_-(x), \alpha_+(x)]$  otherwise  $h^*(z^*)(x) = +\infty$  for  $x \in \Omega$ . Since  $lv = \nabla v$ , the dual operator is represented by the divergence operator  $-l^*q^* = \text{div}q^*$ . Combining (11) and (12) we derive compound functionals

$$D_F(v, v^*) = \int_{\Omega} (\alpha_+ v^+ + \alpha_- v^- - v^* v) \, dx, \quad D_G(lv, q^*) = \frac{1}{2} \int_{\Omega} (lv - q^*) \cdot (lv - q^*) \, dx, \tag{13}$$

where the form for  $D_F(\cdot)$  is valid if the condition

$$v^*(x) \in [-\alpha_-(x), \alpha_+(x)] \tag{14}$$

is satisfied almost everywhere in  $\Omega$ , otherwise  $D_F(v, v^*) = +\infty$ . For the gradient type problem, it holds

$$p^* = \nabla u, \tag{15}$$

i.e.,  $p^*$  represents the exact flux (gradient of the exact solution). Compound functionals appear in the energy identity (Proposition (7.2.13) of [18])

$$D_F(v, -l^*p^*) + D_G(lv, p^*) = J(v) - J(u) \quad \text{for all } v \in K. \tag{16}$$

The terms in the left part of (16) are

$$D_F(v, -l^*p^*) = \int_{\Omega} (\alpha_+ v^+ + \alpha_- v^- - (\text{div}p^*)v) \, dx, \tag{17}$$

$$D_G(lv, p^*) = \frac{1}{2} \int_{\Omega} \nabla(u - v) \cdot \nabla(u - v) \, dx = \frac{1}{2} \|\nabla(u - v)\|_{L^2(\Omega)}^2 \tag{18}$$

and  $D_F(v, v^*)$  is always finite since the condition  $\text{div}p^*(x) \in [-\alpha_-(x), \alpha_+(x)]$  is satisfied almost everywhere in  $\Omega$ .

**Remark 3** (*Gap In The Energy Estimate*). If we drop the nonnegative term  $D_F(v, v^*)$  we get

$$\frac{1}{2} \|\nabla(u - v)\|_{L^2(\Omega)}^2 \leq J(v) - J(u) \quad \text{for all } v \in K \tag{19}$$

which is well known in connection to class of nonlinear problems related to variational inequalities. For the two-phase obstacle problem, it was derived in [10]. The gap in the sharpness of the estimate (19) is exactly measured by the term  $D_F(v, -l^*p^*)$ . By respecting  $D_F(v, -l^*p^*)$  we can get the equality formulated by the main estimate (16). The contribution of  $D_F(v, -l^*p^*)$  is expected not to be very high for good quality approximation  $v \in K$  to the exact solution  $u$ . An example, when the gap becomes significantly large (for a bad approximation  $v$ ) is given in Section 4 of [14]. The form of  $D_F(v, -l^*p^*)$  represents a certain measure of the error associated with free boundary and it is further discussed in Section 2 of [14].

### 2.2. Majorant estimate

The exact energy  $J(u)$  in the energy identity (16) and the energy inequality (19) is not computable without the knowledge of the exact solution  $u$ . However, we can get its computable lower bound using a perturbed functional

$$J_\mu(v) := G(lv) + F_\mu(v), \quad F_\mu(v) := \int_\Omega \mu v \, dx, \tag{20}$$

where a multiplier  $\mu \in \Lambda$  belongs to the space

$$\Lambda := \{ \mu \in L^\infty(\Omega) : \mu(x) \in [-\alpha_-, \alpha_+] \text{ a.e. in } \Omega \}.$$

The perturbed functional  $J_\mu(v)$  replaces the non-differentiable functional  $J(v)$  at the cost of a new variable  $\mu \in \Lambda$  in  $F_\mu(\cdot)$ . It holds

$$J(u) = \inf_{v \in K} \sup_{\mu \in \Lambda} J_\mu(v) = \sup_{\mu \in \Lambda} \inf_{v \in K} J_\mu(v) \geq \inf_{v \in K} J_\mu(v) =: J_\mu(u_\mu) \quad \text{for all } \mu \in \Lambda, \tag{21}$$

where  $u_\mu \in K$  is unique. In view of (21), the minimal perturbed energy  $J_\mu(u_\mu)$  serves as the lower bound of  $J(u)$ . We find a computable lower bound of  $J_\mu(u_\mu)$  by means of the dual counterpart of the perturbed problem. The dual problem is generated by the Lagrangian

$$L_\mu(v, q^*) := \langle q^*, lv \rangle - G^*(q^*) + F_\mu(v).$$

We note  $v = g + w$ , where  $w \in H_0^1(\Omega)$  and estimate

$$\begin{aligned} J_\mu(u_\mu) &= \inf_{v \in K} J_\mu(v) = \inf_{v \in K} \sup_{q^* \in Q^*} L_\mu(v, q^*) = \sup_{q^* \in Q^*} \inf_{v \in K} L_\mu(v, q^*) \\ &= \sup_{q^* \in Q^*} \inf_{w \in H_0^1(\Omega)} \left\{ -G^*(q^*) + \int_\Omega (q^* \cdot \nabla g - \mu g) \, dx + \int_\Omega (q^* \cdot \nabla w - \mu w) \, dx \right\} \\ &\geq \sup_{q^* \in Q_\mu} \left\{ -G^*(q^*) + \int_\Omega (q^* \cdot \nabla g - \mu g) \, dx \right\} \\ &= \sup_{q^* \in Q_\mu} J_\mu^*(q^*) \geq J_\mu^*(q^*) \quad \text{for all } q^* \in Q_\mu^*, \end{aligned} \tag{22}$$

where

$$J_\mu^*(q^*) := -G^*(q^*) + \int_\Omega (q^* \cdot \nabla g - \mu g) \, dx \tag{23}$$

and

$$Q_\mu^* = \left\{ q^* \in Q^* : \int_\Omega (q^* \cdot \nabla w - \mu w) \, dx = 0 \text{ for all } w \in H_0^1(\Omega) \right\}. \tag{24}$$

Due to (21) and (22), we obtain the estimate

$$J(v) - J(u) \leq J(v) - J_\mu^*(q^*) = \frac{1}{2} \|\nabla v - q^*\|_{L^2(\Omega)}^2 + \int_\Omega (\alpha_+ v^+ + \alpha_- v^- - \mu v) \, dx \tag{25}$$

valid for all  $v \in K, \mu \in \Lambda, q^* \in Q_\mu^*$ . The right-hand side of (25) is fully computable, but it requires the constraint  $q^* \in Q_\mu^*$ . To bypass this constraint we introduce a new variable

$$\eta^* \in Y^* := H(\Omega, \text{div})$$

and project it to  $Q_\mu^*$ . The space  $H(\Omega, \text{div})$  is a subspace of  $L^2(\Omega, \mathbb{R}^n)$  that contains vector-valued functions with square-summable divergence. There holds a projection-type inequality (see, e.g., Chapter 3 of [17])

$$\inf_{q^* \in Q_\mu^*} \|\eta^* - q^*\|_{L^2(\Omega)} \leq C_\Omega \|\text{div} \eta^* + \mu\|_{L^2(\Omega)} \quad \text{for all } \eta^* \in Y^*,$$

where the constant  $C_\Omega > 0$  originates from the generalized Friedrichs inequality

$$\|w\|_{L^2(\Omega)} \leq C_\Omega \|\nabla w\|_{L^2(\Omega)} \quad \text{for all } w \in H_0^1(\Omega).$$

Then, the  $q^*$ -dependent term in (25) satisfies

$$\begin{aligned} \inf_{q^* \in Q_\mu^*} \|\nabla v - q^*\|_{L^2(\Omega)}^2 &\leq (\|\nabla v - \eta^*\|_{L^2(\Omega)} + \inf_{q^* \in Q_\mu^*} \|\eta^* - q^*\|_{L^2(\Omega)})^2 \\ &\leq (\|\nabla v - \eta^*\|_{L^2(\Omega)} + C_\Omega \|\text{div} \eta^* + \mu\|_{L^2(\Omega)})^2 \\ &\leq (1 + \beta) \|\nabla v - \eta^*\|_{L^2(\Omega)}^2 + \left(1 + \frac{1}{\beta}\right) C_\Omega^2 \|\text{div} \eta^* + \mu\|_{L^2(\Omega)}^2, \end{aligned} \tag{26}$$

where we used Young’s inequality with a parameter  $\beta > 0$  in the last inequality. Hence the combination of (25) and (26) yields the majorant estimate

$$J(v) - J(u) \leq M_+(v; \beta, \eta^*, \mu) \quad \text{for all } v \in K, \mu \in \Lambda, \eta^* \in Y^*, \beta > 0, \tag{27}$$

where the nonnegative functional

$$M_+(v; \beta, \eta^*, \mu) := \mathcal{M}_{+1}(v; \beta, \eta^*) + \mathcal{M}_{+2}(\beta, \eta^*, \mu) + \mathcal{M}_{+3}(v; \mu) \tag{28}$$

represents a functional error majorant with three additive subparts

$$\begin{aligned} \mathcal{M}_{+1}(v; \beta, \eta^*) &:= \frac{1}{2} (1 + \beta) \|\nabla v - \eta^*\|_{L^2(\Omega)}^2, \\ \mathcal{M}_{+2}(\beta, \eta^*, \mu) &:= \frac{1}{2} \left(1 + \frac{1}{\beta}\right) C_\Omega^2 \|\text{div} \eta^* - \mu\|_{L^2(\Omega)}^2, \\ \mathcal{M}_{+3}(v; \mu) &:= \int_\Omega (\alpha_+ v^+ + \alpha_- v^- - \mu v) \, dx. \end{aligned} \tag{29}$$

**Remark 4.** The final form of the functional error majorant (28) is slightly different to formula (3.13) in [14]. In (28), only one multiplier variable  $\mu$  is introduced replacing two multipliers  $\mu_-, \mu_+$  of [14]. Another simplification in this paper is that the variable diffusions coefficient matrix  $A$  is not considered here and we treat the quadratic part of the energy  $\frac{1}{2} \int_\Omega \nabla v \cdot \nabla v \, dx$  instead of  $\frac{1}{2} \int_\Omega A \nabla v \cdot \nabla v \, dx$ .

### 3. Discretization

We assume a domain  $\Omega \subset \mathbb{R}^2$  with a polygonal boundary discretized by a uniform regular triangular mesh  $\mathcal{T}_h$  in the sense of Ciarlet [19], where  $h$  denotes the mesh size. Let  $\mathcal{E}$  denote the set of all edges and  $\mathcal{N}$  the set of all nodes in  $\mathcal{T}_h$ . By

$$|\mathcal{E}|, |\mathcal{N}|, |\mathcal{T}|$$

we mean the number of edges, of nodes and of triangles of  $\mathcal{T}_h$ . The following lowest order finite elements (FE) approximations are considered:

- The exact solution  $u \in K$  of the two phase obstacle problem is approximated by

$$u_h \in K_h := K \cap P_1(\mathcal{T}_h),$$

where  $P_1(\mathcal{T}_h)$  denotes the space of elementwise nodal and continuous functions defined on  $\mathcal{T}_h$ .

- The exact multiplier  $\lambda \in \Lambda$  is approximated by

$$\lambda_h \in \Lambda_h := \Lambda \cap P_0(\mathcal{T}_h),$$

where  $P_0(\mathcal{T}_h)$  denotes the space of element wise constant functions defined on  $\mathcal{T}_h$ .

- The flux variable  $\eta^* \in Q$  in the functional majorant is approximated by

$$\eta_h^* \in Y_h^* := RT_0(\mathcal{T}_h),$$

where  $RT_0(\mathcal{T}_h)$  is the space of the lowest order Raviart–Thomas functions.

Note that dimensions of these approximation spaces are

$$\dim(K_h) = |\mathcal{N}|, \quad \dim(\Lambda_h) = |\mathcal{T}|, \quad \dim(Y_h^*) = |\mathcal{E}|.$$

We are interested in two computation tasks: First obtaining a discrete solution  $u_h$  or its approximation  $v_h$  and then measuring its quality by the optimized functional error majorant.

### 3.1. Dual based algorithm for Lagrange multipliers

As we mentioned in 2.2 due to the non-differentiability term in  $J(\cdot)$ , we do not obtain the approximative solution  $u_h \in K_h$  from the relation  $J(u_h) = \inf_{v \in K_h} J(v)$  directly. We estimate

$$J(u_h) = \inf_{v_h \in K_h} \sup_{\mu \in \Lambda} J_{\mu}(v) \geq \inf_{v_h \in K_h} \sup_{\mu_h \in \Lambda_h} J_{\mu_h}(v_h) \geq \sup_{\mu_h \in \Lambda_h} \inf_{v_h \in K_h} J_{\mu_h}(v_h) =: J_{\lambda_h}(u_{\lambda_h}) \tag{30}$$

and look for an approximation pair  $(\lambda_h, u_{\lambda_h}) \in \Lambda_h \times K_h$  instead. Note, in general  $u_{\lambda_h} \neq u_h$  and it holds  $J(u_{\lambda_h}) \geq J(u_h)$  only. The saddle point problem on the right-hand side of (30) can be further reformulated as a dual problem for a Lagrange multiplier

$$I^*(\lambda_h) = \sup_{\mu_h \in \Lambda_h} I^*(\mu_h), \quad \text{where } I^*(\mu_h) := \inf_{v_h \in K_h} J_{\mu_h}(v_h). \tag{31}$$

Approximations  $v_h \in K_h$  and  $\mu_h \in \Lambda_h$  from (20) are equivalently represented by discrete column vectors  $\mathbf{v} \in \mathbb{R}^{|\mathcal{N}|}$ ,  $\boldsymbol{\mu} \in \mathbb{R}^{|\mathcal{E}|}$  and we can rewrite  $J_{\mu_h}(\cdot)$  from (20) as

$$J_{\mu}(\mathbf{v}) = \frac{1}{2} \mathbf{v}^T \mathbb{K} \mathbf{v} + \mathbf{v}^T \mathbb{M} \boldsymbol{\mu}, \tag{32}$$

where  $\mathbb{K} \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{N}|}$  and  $\mathbb{M} \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{E}|}$ . The square matrix  $\mathbb{K}$  represents a stiffness matrix from a discretization of the Laplace operator in  $K_h$ . The rectangular matrix  $\mathbb{M}$  represents the  $L^2$ -scalar product of functions from spaces  $P_1(\mathcal{T}_h)$  and  $P_0(\mathcal{T}_h)$ . It holds

$$\int_{\Omega} \nabla v_h \cdot \nabla v_h \, dx = \mathbf{v}^T \mathbb{K} \mathbf{v}, \quad \int_{\Omega} v_h \cdot \mu \, dx = \mathbf{v}^T \mathbb{M} \boldsymbol{\mu}$$

for all  $v_h \in P_1(\mathcal{T}_h)$ ,  $\mu \in P_0(\mathcal{T}_h)$  and corresponding column vectors  $\mathbf{v} \in \mathbb{R}^{|\mathcal{N}|}$ ,  $\boldsymbol{\mu} \in \mathbb{R}^{|\mathcal{E}|}$ . If we order nodes  $\mathcal{N}$  in a way that internal nodes  $\mathcal{N}_I$  precede Dirichlet nodes  $\mathcal{N}_D$  (no Neumann nodes are assumed for simplicity), we have the decomposition

$$\mathbf{v} = (\mathbf{v}_I, \mathbf{v}_D) \in \mathbb{R}^{|\mathcal{N}_I|} \times \mathbb{R}^{|\mathcal{N}_D|}$$

in Dirichlet and internal components and  $|\mathcal{N}| = |\mathcal{N}_I| + |\mathcal{N}_D|$ . Then (32) can be rewritten as

$$J_{\mu}(\mathbf{v}) = \frac{1}{2} \begin{pmatrix} \mathbf{v}_I \\ \mathbf{v}_D \end{pmatrix}^T \begin{pmatrix} \mathbb{K}_{I,I} & \mathbb{K}_{I,D}^T \\ \mathbb{K}_{I,D} & \mathbb{K}_{D,D} \end{pmatrix} \begin{pmatrix} \mathbf{v}_I \\ \mathbf{v}_D \end{pmatrix} + \begin{pmatrix} \mathbf{v}_I \\ \mathbf{v}_D \end{pmatrix}^T \begin{pmatrix} \mathbb{M}_I \\ \mathbb{M}_D \end{pmatrix} \boldsymbol{\mu}. \tag{33}$$

Note that the rectangular matrix  $\mathbb{K}_{ID}$  is the restriction of  $\mathbb{K}$  to its subblock with rows  $\mathcal{N}_I$  and columns  $\mathcal{N}_D$ . Therefore  $\mathbb{K}_{I,I}$  and  $\mathbb{K}_{D,D}$  are not diagonal matrices. The rectangular matrices  $\mathbb{M}_I$  and  $\mathbb{M}_D$  are then restrictions of  $\mathbb{M}$  to subblocks with rows  $\mathcal{N}_I$  and  $\mathcal{N}_D$  with all columns left. The value of  $\mathbf{v}_D$  is known and given by Dirichlet boundary conditions. The direct computation reveals

$$I^*(\boldsymbol{\mu}) = \frac{1}{2} \mathbf{v}_D^T \mathbb{K}_{D,D} \mathbf{v}_D + \mathbf{v}_D^T \mathbb{M}_D \boldsymbol{\mu} - \frac{1}{2} (\mathbb{K}_{I,D} \mathbf{v}_D + \mathbb{M}_I \boldsymbol{\mu}) \mathbb{K}_{I,I}^{-1} (\mathbb{K}_{I,D} \mathbf{v}_D + \mathbb{M}_I \boldsymbol{\mu}). \tag{34}$$

In addition to it, for a given  $\boldsymbol{\mu}$ , the component  $\mathbf{v}_I$  minimizing the functional (33) satisfies

$$\mathbf{v}_I = -\mathbb{K}_{I,I}^{-1} (\mathbb{K}_{I,D} \mathbf{v}_D + \mathbb{M}_I \boldsymbol{\mu}) \quad \text{or equivalently} \quad -\mathbb{K}_{I,I} \mathbf{v}_I = (\mathbb{K}_{I,D} \mathbf{v}_D + \mathbb{M}_I \boldsymbol{\mu}). \tag{35}$$

This formula is applied for the reconstruction of  $\mathbf{v}_I$  from  $\boldsymbol{\mu}$ . Since  $\mathbb{K}_{I,I}$  is a sparse matrix, its inverse  $\mathbb{K}_{I,I}^{-1}$  is a dense matrix. Then, the right part of (35) is exploited in practical evolutions including (34). The matrix  $\mathbb{K}_{I,I}^{-1}$  is positive definite as well as  $\mathbb{K}_{I,I}$  and the functional  $I^*(\cdot)$  is concave and contains quadratic and linear terms only. Thus the functional  $-I^*(\cdot)$  is convex and its minimum  $\lambda_h$  from (31) is represented by a column vector  $\boldsymbol{\lambda} \in \mathbb{R}^{|\mathcal{E}|}$  and solves a quadratic programming (QP) problem with box constraints

$$-I^*(\boldsymbol{\lambda}) = \min -I^*(\boldsymbol{\mu}), \quad \text{where } \{\boldsymbol{\mu}\}_j \in [-\{\boldsymbol{\alpha}_-\}_j, \{\boldsymbol{\alpha}_+\}_j] \quad \text{for all } j \in \{1, \dots, |\mathcal{E}|\}. \tag{36}$$

Here, vectors  $\boldsymbol{\alpha}_-, \boldsymbol{\alpha}_+ \in \mathbb{R}^{|\mathcal{E}|}$  represent  $P_0(\mathcal{T}_h)$  discretizations of functions  $\alpha_-, \alpha_+$ . The corresponding solution  $u_{\lambda_h} \in K_h$  is then represented by a column vector  $\mathbf{u}_{\lambda} = (\mathbf{u}_{\lambda_I}, \mathbf{u}_{\lambda_D}) \in \mathbb{R}^{|\mathcal{N}|}$  and  $\mathbf{u}_{\lambda_I}$  solves (35) for  $\mathbf{v} = \boldsymbol{\lambda}$ . The solutions steps above are summarized in Algorithm 1.

**Algorithm 1** Quadratic programming for Lagrange multipliers.

Let discretization matrices  $\mathbb{K} \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{N}|}$  and  $\mathbb{M} \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{T}|}$  be given with their subblocks  $\mathbb{K}_{I,I} \in \mathbb{R}^{|\mathcal{N}_I| \times |\mathcal{N}_I|}$ ,  $\mathbb{K}_{D,D} \in \mathbb{R}^{|\mathcal{N}_D| \times |\mathcal{N}_D|}$ ,  $\mathbb{K}_{I,D} \in \mathbb{R}^{|\mathcal{N}_I| \times |\mathcal{N}_D|}$  and  $\mathbb{M}_I \in \mathbb{R}^{|\mathcal{N}_I| \times |\mathcal{T}|}$ ,  $\mathbb{M}_D \in \mathbb{R}^{|\mathcal{N}_D| \times |\mathcal{T}|}$ . Let  $\mathbf{v}_D \in \mathbb{R}^{|\mathcal{N}_D|}$  be vector of prescribed Dirichlet values in nodes  $\mathcal{N}_D$ . Then:

(i) find the vector of Lagrange multipliers from the quadratic minimization problem

$$\boldsymbol{\lambda} = \operatorname{argmin}_{\boldsymbol{\mu} \in \mathbb{R}^{|\mathcal{T}|}} \left( -\frac{1}{2} \mathbf{v}_D^T \mathbb{K}_{D,D} \mathbf{v}_D - \mathbf{v}_D^T \mathbb{M}_D \boldsymbol{\mu} + \frac{1}{2} (\mathbb{K}_{I,D} \mathbf{v}_D + \mathbb{M}_I \boldsymbol{\mu}) \mathbb{K}_{I,I}^{-1} (\mathbb{K}_{I,D} \mathbf{v}_D + \mathbb{M}_I \boldsymbol{\mu}) \right)$$

under box constraints  $\{\boldsymbol{\mu}\}_j \in [-\{\boldsymbol{\alpha}_-\}_j, \{\boldsymbol{\alpha}_+\}_j]$  for all  $j \in \{1, \dots, |\mathcal{T}|\}$ .

(ii) reconstruct the solution vector  $\mathbf{v}_I \in \mathbb{R}^{|\mathcal{N}_I|}$  from  $-\mathbb{K}_{I,I} \mathbf{v}_I = (\mathbb{K}_{I,D} \mathbf{v}_D + \mathbb{M}_I \boldsymbol{\mu})$ .

(iii) output  $\lambda_h$  and  $u_{\lambda_h}$  represented by vectors  $\boldsymbol{\lambda}$  and  $\mathbf{v} = (\mathbf{v}_I, \mathbf{v}_D)$ .

**Algorithm 2** Majorant minimization algorithm.

Let  $k = 0$  and let initial  $\beta_0 > 0$  and  $\mu_0 \in \Lambda_h$  be given. Then:

- (i) find an iteration  $\eta_{k+1}^* \in Y_h^*$  such that  $\eta_{k+1}^* = \operatorname{argmin}_{\eta \in Y_h^*} \mathcal{M}_+(u_{\lambda_h}; \beta_k, \mu_k, \eta^*)$ ,
- (ii) find  $\mu_{k+1} \in \Lambda_h$  such that  $\mu_{k+1} = \operatorname{argmin}_{\mu \in \Lambda_h} \mathcal{M}_+(u_{\lambda_h}; \beta_k, \mu, \eta_{k+1}^*)$ ,
- (iii) find  $\beta_{k+1} > 0$  such that  $\beta_{k+1} = \operatorname{argmin}_{\beta \in \mathbb{R}_+} \mathcal{M}_+(u_{\lambda_h}; \beta, \mu_{k+1}, \eta_{k+1}^*)$ ,
- (iv) set  $k := k + 1$  and repeat (i)–(iii) until convergence. Then, output  $\eta_h^* := \eta_{k+1}^*$  and  $\mu_h := \mu_{k+1}$ .

3.2. Minimization of the functional error majorant

For a given approximation  $u_{\lambda_h} \in K_h$ , the majorant value  $\mathcal{M}_+(u_{\lambda_h}; \beta, \eta^*, \mu)$  majorizes the value  $J(u_{\lambda_h}) - J(u)$ . The majorant  $\mathcal{M}_+(u_{\lambda_h}; \beta, \eta^*, \mu)$  can be minimized with respect to its free arguments  $\beta > 0$ ,  $\eta^* \in Y^*$ ,  $\mu \in \Lambda$  in order to obtain the sharp upper bound. The fields  $\eta^* \in Y^*$ ,  $\mu \in \Lambda$  can be sought on a mesh  $\mathcal{T}_{\tilde{h}}$  with a different mesh size  $\tilde{h}$ . Choosing very small mesh size  $\tilde{h} \ll h$  leads to sharper bounds but higher computational costs. Here we consider the same mesh size  $\tilde{h} = h$  for simplicity,

$$\eta^* \in Y_h^*, \quad \mu \in \Lambda_h.$$

We use the successive minimization algorithm described in Algorithm 2. The step (i) corresponds to the solution of a linear system of equations

$$\left[ (1 + \beta_k) \mathbb{M}^{RT0} + C_\Omega^2 \left( 1 + \frac{1}{\beta_k} \right) \mathbb{K}^{RT0} \right] \boldsymbol{\eta}_{k+1}^* = (1 + \beta_k) \mathbf{c} + C_\Omega^2 \left( 1 + \frac{1}{\beta_k} \right) \mathbf{d}_k \tag{37}$$

for a column vector  $\boldsymbol{\eta}_{k+1}^* \in \mathbb{R}^{|\mathcal{E}|}$ . Here,  $\mathbb{K}^{RT0}, \mathbb{M}^{RT0} \in \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}$  are stiffness and mass matrices corresponding to  $RT_0(\mathcal{T}_h)$  elements. Vectors  $\mathbf{c}, \mathbf{d}_k \in \mathbb{R}^{|\mathcal{E}|}$  are column vectors with components constructed as

$$c_i = \int_\Omega \nabla u_{\lambda_h} \cdot \boldsymbol{\eta}_i \, dx, \quad d_{k,i} = \int_\Omega \mu_k \operatorname{div} \boldsymbol{\eta}_i \, dx \tag{38}$$

for  $i \in \{1, \dots, |\mathcal{E}|\}$ . No boundary conditions are imposed on  $\mathbf{y}$  in (37) since the discrete solution  $v$  satisfies Dirichlet boundary conditions only. The minimal argument  $\mu_{k+1} \in \Lambda_h$  in step (ii) is locally computed on every triangle  $T \in \mathcal{T}_h$  from the formula

$$\mu_{k+1}|_T = \mathcal{P}_{[-\alpha_-, \alpha_+]} \left( \operatorname{div} \boldsymbol{\eta}_{k+1}^*|_T + \frac{\bar{u}_{\lambda_h}|_T}{C_\Omega^2 \left( 1 + \frac{1}{\beta_k} \right)} \right), \tag{39}$$

where  $\mathcal{P}_{[-\alpha_-, \alpha_+]}$  is the projection on the convex set  $[-\alpha_-, \alpha_+]$  and  $\bar{u}_{\lambda_h}|_T$  means an averaged value of  $u_{\lambda_h}$  over a triangular element  $T$ . The minimization in step (iii) leads to the explicit relation

$$\beta_{k+1} = \frac{\|\operatorname{div} \boldsymbol{\eta}_{k+1}^* - \mu_{k+1}\|_\Omega}{\|\nabla v - \boldsymbol{\eta}_{k+1}^*\|_\Omega}. \tag{40}$$

**Remark 5.** Algorithm 2 is a subsequent minimization algorithm and majorant values of iterates (in fact their subsequence) convergence to the minimal majorant value in case of classical obstacle problems, see [20] for details. Here, a convergence has not been studied yet. A practical convergence can be speeded up significantly, if the approximation  $\lambda_h$  is taken as an initial approximation  $\mu_0$ .

**Table 1**  
Computations of Example I on various uniform triangular meshes.

Level	$ \mathcal{N} $	$J(u_{\lambda_h}) (I^*(\lambda_h))$	$J(u_{\lambda_h}) - J(u)$	$\mathcal{M}_+(u_{\lambda_h}, \cdot)$	$\mathcal{M}_{+1}(\cdot)$	$\mathcal{M}_{+2}(\cdot)$	$\mathcal{M}_{+3}(\cdot)$	$\frac{\mathcal{M}_+(u_{\lambda_h}, \cdot)}{J(u_{\lambda_h}) - J(u)}$
1	15	6.0383 (5.9975)	7.05e-01	1.79e+00	1.43e+00	2.88e-02	3.29e-01	2.54
2	45	5.5030 (5.5000)	1.70e-01	3.91e-01	3.75e-01	1.08e-03	1.48e-02	2.30
3	153	5.3779 (5.3750)	4.46e-02	9.38e-02	8.88e-02	3.72e-05	5.00e-03	2.11
4	561	5.3449 (5.3437)	1.15e-02	2.32e-02	2.16e-02	1.24e-06	1.62e-03	2.01
5	2145	5.3409 (5.3355)	7.55e-03	1.07e-02	5.16e-03	6.40e-08	5.54e-03	1.42
6	8385	5.3354 (5.3340)	2.04e-03	2.77e-03	1.32e-03	2.09e-08	1.45e-03	1.36

**4. Numerical examples**

In this section we elaborate two numerical examples, i.e. Example I, and Example II, with known and unknown exact solutions.

*4.1. Example I with known exact solution*

This example is introduced in [10]; it is also tested for one dimensional case in [14]. Here, we consider it in two dimensional and assume a rectangular domain

$$\Omega = X \times Y := (-1, 1) \times (0, 1), \tag{41}$$

and constant coefficients

$$\alpha_- = \alpha_+ = 8. \tag{42}$$

The two phase obstacle problem (2) is supplied with the Dirichlet boundary conditions

$$u(-1, y) = -1, \quad u(1, y) = 1 \quad \forall y \in Y \tag{43}$$

and homogeneous Neumann boundary conditions

$$\frac{\partial u}{\partial x}(x, 0) = \frac{\partial u}{\partial x}(x, 1) = 0 \quad \forall x \in X. \tag{44}$$

The exact solution  $u \in K$  is given by the relation independent of  $y \in Y$ ,

$$u(x, y) = \begin{cases} -4x^2 - 4x - 1, & x \in X_- := [-1, -0.5], \\ 0, & x \in X_0 := [-0.5, 0.5], \\ 4x^2 - 4x + 1, & x \in X_+ := [0.5, 1] \end{cases} \tag{45}$$

and its (exact) energy is  $J(u) = 5\frac{1}{3}$ . The (exact) free boundary is characterized by two lines

$$(\pm 0.5, y), \quad \text{where } y \in Y.$$

The (exact) Lagrange multiplier  $\lambda \in \Lambda$  is then given by

$$\lambda(x, y) = \begin{cases} -\alpha_-, & x \in X_-, \\ 0, & x \in X_0, \\ \alpha_+, & x \in X_+ \end{cases} \tag{46}$$

and it is a discontinuous function with a jump on the free boundary. We compute approximation pairs  $(\lambda_h, u_{\lambda_h}) \in \Lambda_h \times K_h$  for a sequence of nested uniformly refined meshes. Levels 1 and 2 meshes are depicted in Fig. 1. Since some discretization nodes are lying exactly on the free boundary, there might be a chance to reconstruct the free boundary exactly from approximative solutions. A finer (level 5) approximation pair  $(\lambda_h, u_{\lambda_h}) \in \Lambda_h \times K_h$  computed from the dual-based solver is shown in Fig. 2. The approximative Lagrange multiplier field  $\lambda_h$  however only approximates the exact free boundary. To the given approximation pair  $(\lambda_h, u_{\lambda_h})$ , a functional majorant is optimized using a fixed number of 10 iterations of Algorithm 1 (we set  $\mu_0 = \lambda_h$ ). To get more insight on the majorant behaviour, we display space densities of all three additive majorant subparts

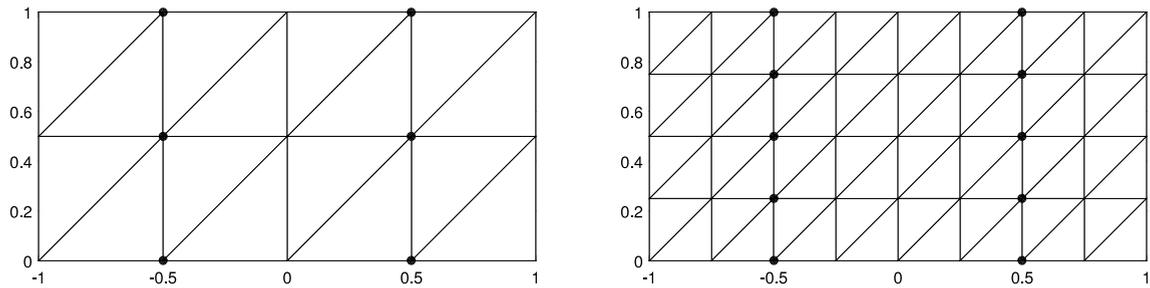
$$\mathcal{M}_{+1}(u_{\lambda_h}; \beta, \eta^*), \quad \mathcal{M}_{+2}(\beta, \eta^*, \mu), \quad \mathcal{M}_{+3}(u_{\lambda_h}; \mu)$$

separately in Fig. 3. The amplitudes of  $\mathcal{M}_{+2}$  are significantly lower than amplitudes of  $\mathcal{M}_{+1}$  and  $\mathcal{M}_{+3}$ , but the value of  $\mathcal{M}_{+3}$  is still relatively high. The high value of  $\mathcal{M}_{+3}$  indicates that the exact free boundary is not sufficiently resolved yet and the density of  $\mathcal{M}_{+3}$  seem to be a reasonable indicator of the exact free boundary.

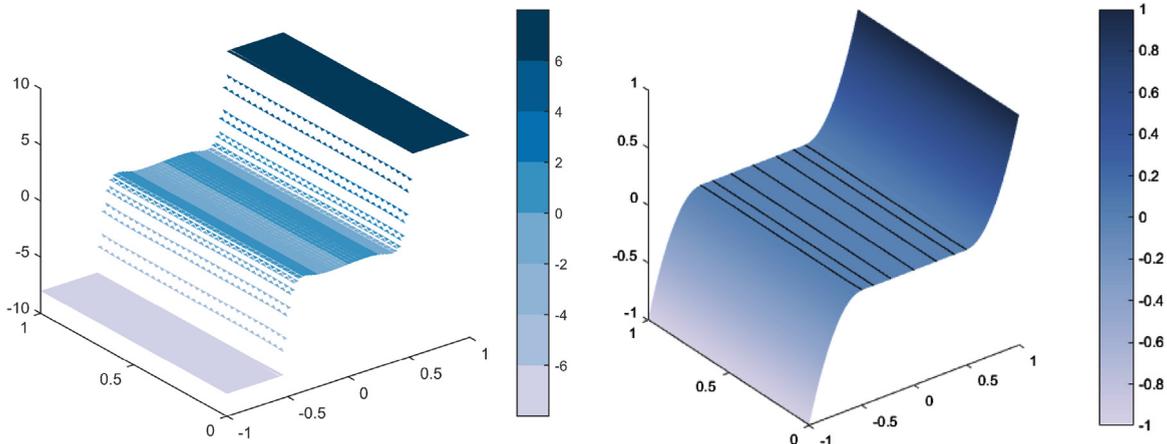
Computations on all nested uniformly refined triangular meshes are summarized in Table 1. The dual energy  $I^*(\lambda_h)$  and primal energy  $J(u_{\lambda_h})$  converge to the exact energy  $J(u)$  as  $h \rightarrow 0$ . Since we work with nested meshes, we additionally have

$$J(u_{\lambda_h}) \searrow J(u) \quad \text{or equivalently} \quad J(u_{\lambda_h}) - J(u) \searrow 0.$$

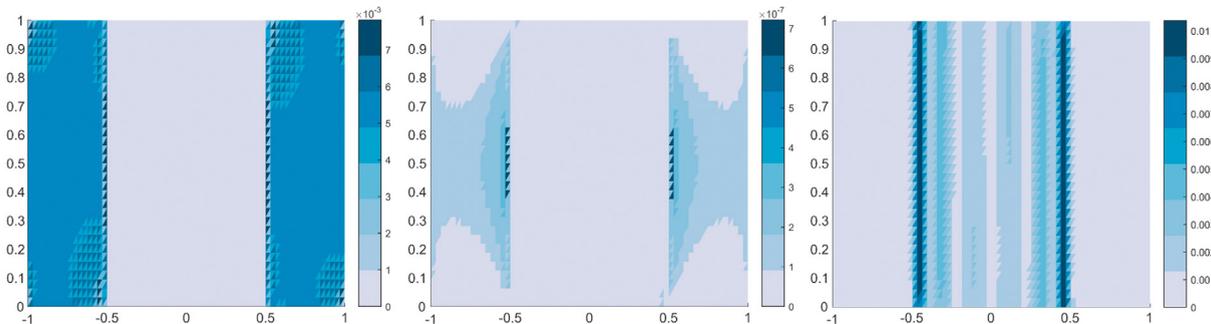
The difference of energies  $J(u_{\lambda_h}) - J(u) \searrow 0$  is bounded from above by the majorant value  $\mathcal{M}_+(u_{\lambda_h}, \dots)$  as stated in the majorant estimate (27).



**Fig. 1.** Example I—level 1 (left) and level 2 (right) nested triangular meshes. Note that there are nodes lying on the exact free boundary given by lines  $x = \pm 0.5$ .



**Fig. 2.** Example I—approximations: multiplier  $\lambda_h \in A_h$  (left) and the corresponding solution  $u_{\lambda_h} \in K_h$  (right) computed on level 5 triangular mesh (referred to as level 5 in Table 1). The multiplier approximation  $\lambda_h$  indicates an approximative free boundary, the exact free boundary is given by lines  $x = \pm 0.5$ . Full contour lines of  $u_{\lambda_h}$  at values  $\pm 0.0001$  are additionally displayed (right).



**Fig. 3.** Example I—distribution of the majorant parts  $\mathcal{M}_{+1}$  (left),  $\mathcal{M}_{+2}$  (middle),  $\mathcal{M}_{+3}$  (right) computed on level 5 triangular mesh (referred to as level 5 in Table 1).

**Remark 6 (Extension To Mixed Dirichlet–Neumann Boundary Conditions).** This example assumes both Dirichlet and Neumann boundary conditions, but only Dirichlet boundary conditions are considered in  $K$ . The dual based solver for a double-phase problem can still be applied, with Neumann nodes  $\mathcal{N}_N$  being added to internal nodes  $\mathcal{N}_I$ . The majorant estimate (27) is valid with the same majorant form (28), but the flux  $\eta^* \in Q_h$  must satisfy an extra condition  $\eta^* \cdot n = 0$  on a Neumann boundary, where  $n$  is a normal vector to the boundary. This condition means that components of  $\eta_{k+1}^*$  from (37) corresponding to Neumann edges  $\mathcal{E}_N$  must be equal to zero.

4.2. Example II

The second example is also taken from [10] and considers a square domain

$$\Omega = X \times Y := (-1, 1) \times (-1, 1), \tag{47}$$

**Table 2**

Computations of Example II on various uniform triangular meshes. Note that  $J(u_{ref}) = J(u_{\lambda_h})$  for  $u_{\lambda_h}$  computed on level 6 mesh.

Level	$ \mathcal{N} $	$J(u_{\lambda_h}) (I^*(\lambda_h))$	$J(u_{\lambda_h}) - J(u_{ref})$	$\mathcal{M}_+(u_{\lambda_h}, \cdot)$	$\mathcal{M}_{+1}(\cdot)$	$\mathcal{M}_{+2}(\cdot)$	$\mathcal{M}_{+3}(\cdot)$	$\frac{\mathcal{M}_+(u_{\lambda_h}, \cdot)}{J(u_{\lambda_h}) - J(u_{ref})}$
1	13	13.6667 (13.6667)	6.64e-01	1.84e+00	1.53e+00	1.17e-01	1.98e-01	2.77
2	41	13.1924 (13.1924)	1.90e-01	4.99e-01	4.74e-01	4.44e-03	2.01e-02	2.63
3	145	13.0492 (13.0489)	4.67e-02	1.21e-01	1.19e-01	1.69e-04	1.08e-03	2.58
4	545	13.0135 (13.0134)	1.10e-02	2.96e-02	2.93e-02	6.13e-06	2.19e-04	2.69
5	2113	13.0044 (13.0042)	1.85e-03	7.26e-03	7.01e-03	2.35e-07	2.56e-04	3.92
6	8321	13.0020 (13.0019)		Not evaluated				

**Table 3**

Example II: Lower and upper bounds of the exact energy  $J(u)$  computed for various triangular meshes.

Level	1	2	3	4	5	6
Lower bound of energy $J(u)$	11.8338	12.7075	12.9319	12.9841	12.9971	Not evaluated
Upper bound of energy $J(u)$	13.6667	13.1924	13.0491	13.0137	13.0045	13.0020

constant coefficients

$$\alpha_+ = \alpha_- = 4. \tag{48}$$

The Dirichlet boundary conditions as assumed in the form

$$u(x, y) = \begin{cases} x + 1 & x \in [-1, 1] \text{ and } y = 1, \\ x - 1 & x \in [-1, 1] \text{ and } y = -1, \\ y + 1 & y \in [-1, 1] \text{ and } x = 1, \\ y - 1 & y \in [-1, 1] \text{ and } x = -1. \end{cases} \tag{49}$$

The exact solution  $u \in K$  is not known for this example. Consequently, no apriori information about the shape of the free boundary or the value of the exact energy  $J(u)$  is provided. We compute approximation pairs  $(\lambda_h, u_{\lambda_h}) \in \Lambda_h \times K_h$  again for a sequence of nested uniformly refined meshes. Levels 1 and 2 meshes are depicted in Fig. 5. An approximative solutions pair  $(\lambda_h, u_{\lambda_h}) \in \Lambda_h \times K_h$  obtained by the dual-based solver is depicted in Fig. 6. The approximative Lagrange multiplier field  $\lambda_h$  presumably indicates the exact free boundary. Space distributions of majorant subparts are visualized in Fig. 7. We assume that the density of  $\mathcal{M}_{+3}$  serves as an indicator of the exact free boundary. Table 2 summarizes computations on all nested uniformly refined triangular meshes. The exact energy  $J(u)$  is not known but it is replaced by the energy  $J(u_{ref})$  of a reference solution  $u_{ref}$  in Table 2. The reference solution  $u_{ref}$  is computed as  $u_{\lambda_h}$  on the mesh one level higher (level 6 uniformly refined triangular mesh here).

**Remark 7** (Lower Bound of Difference of Energies Based on a Reference Solution). If a reference solution  $u_{ref}$  is available, its energy  $J(u_{ref})$  satisfies  $J(u_{\lambda_h}) \geq J(u_{ref}) \geq J(u)$  and

$$J(u_{\lambda_h}) - J(u_{ref}) \leq J(u_{\lambda_h}) - J(u) \leq \mathcal{M}_+(u_{\lambda_h}, \dots). \tag{50}$$

The inequality (50) provides actually guaranteed lower and upper bounds of the difference of energies  $J(u_{\lambda_h}) - J(u)$ . Fig. 4 displays convergence of both bounds of  $J(u_{\lambda_h}) - J(u)$  for considered 5 levels approximations  $u_{\lambda_h}$ . By reformulating (50) we get guaranteed bounds of the exact energy

$$J(u_{\lambda_h}) - \mathcal{M}_+(u_{\lambda_h}, \dots) \leq J(u) \leq J(u_{ref}) \tag{51}$$

valid for every approximation  $u_{\lambda_h} \in K_h$ . For this example, lower bounds of (51) create an increasing sequence with respect to levels of refinement as reported in Table 3. For convenience, corresponding upper bounds from Table 2 are also added. The sharpest bounds of Table 3 provide the inequality

$$12.9971 \leq J(u) \leq 13.0020$$

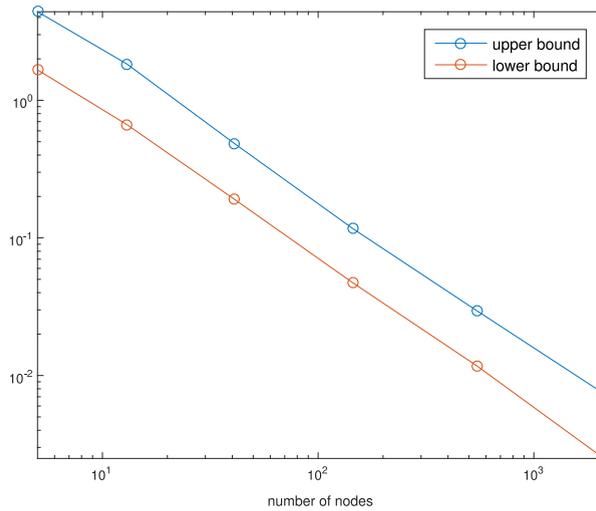
and it suggests  $J(u) = 13$  although there is no analytical proof of it.

#### 4.3. Implementation details and performance

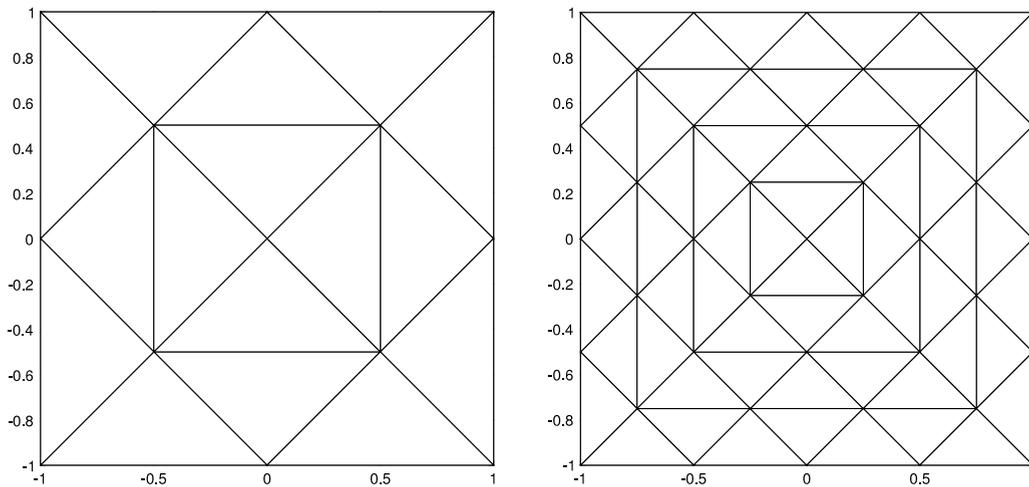
Both numerical examples are implemented in MATLAB and the code available for download at <http://www.mathworks.com/matlabcentral/fileexchange/57232>.

The code is based on vectorization techniques of [21,22]. The main file 'start.m' is located in the directory 'solver\_two\_phase\_obstacle'. The following parameters can be adjusted:

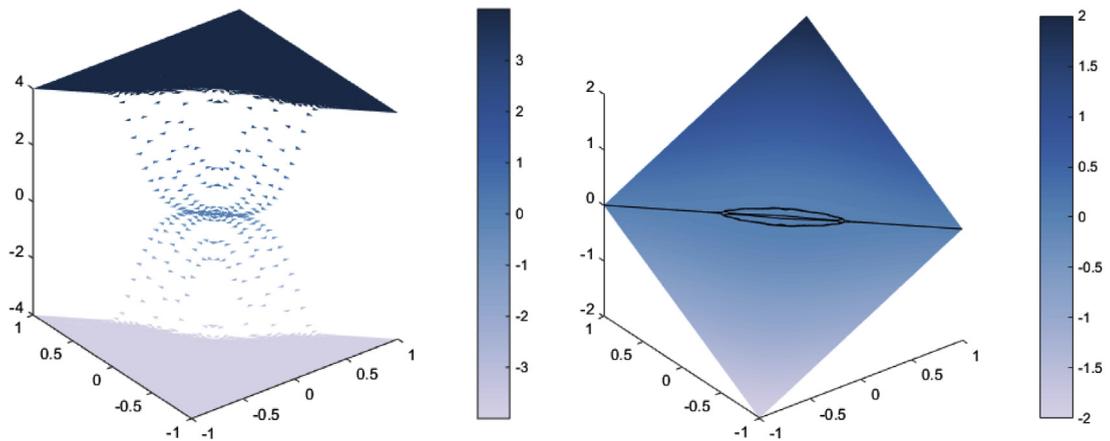
'example'—the considered example ('1' or '2')



**Fig. 4.** Example II—convergence of the difference of energies  $J(u_{\lambda_h}) - J(u)$  is controlled by its computable upper bound  $\mathcal{M}_+(u_{\lambda_h}, \dots)$  and its computable lower bound  $J(u_{\lambda_h}) - J(u_{ref})$ .



**Fig. 5.** Example II—level 1 (left) and level 2 (right) nested triangular meshes.



**Fig. 6.** Example II—approximations: multiplier  $\lambda_h \in A_h$  (left) and the corresponding solution  $u_{\lambda_h} \in K_h$  (right) computed on level 5 triangular mesh (referred to as level 5 in Table 2). The multiplier approximation  $\lambda_h$  (left) indicates an approximative free boundary, the exact free boundary is unknown. Full contour lines of  $u_{\lambda_h}$  at values  $\pm 0.0001$  are additionally displayed (right).

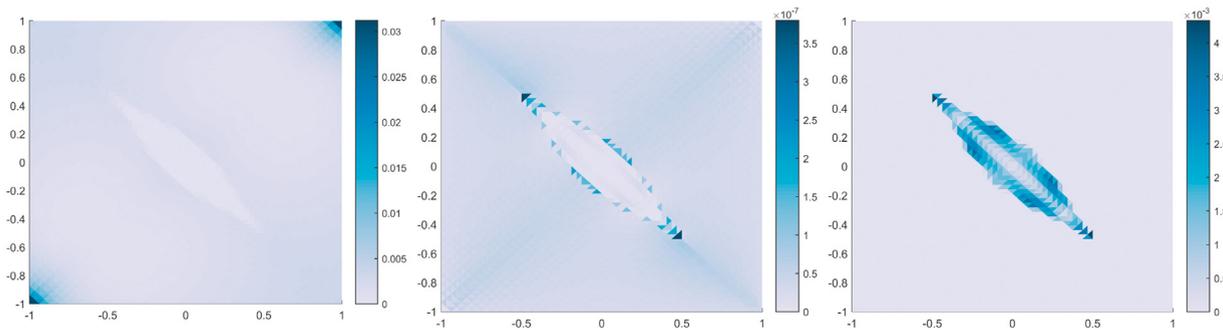


Fig. 7. Example II—distribution of the majorant parts  $\mathcal{M}_{+1}$  (left),  $\mathcal{M}_{+2}$  (middle),  $\mathcal{M}_{+3}$  (right) computed on level 5 triangular mesh (referred to as level 5 in Table 2).

Table 4

Example II: energy evaluations inside QP solver of Algorithm 1 and indices of efficiency for various number of iterations of Algorithm 2.

Level	1	2	3	4	5	6
Energy evaluations in QP	88	268	1202	3323	4133	10 424
$I_{\text{eff}}$ for 10 majorant iterations	2.7713	2.6260	2.5799	2.6926	3.9192	Not evaluated
$I_{\text{eff}}$ for 100 majorant iterations	2.7598	2.5537	2.5126	2.6585	3.9121	Not evaluated
$I_{\text{eff}}$ for 1000 majorant iterations	2.7598	2.5489	2.5005	2.6242	3.9058	Not evaluated
$I_{\text{eff}}$ for 10 000 majorant iterations	2.7598	2.5489	2.4976	2.6196	3.9025	Not evaluated

‘levels\_energy\_error’—the number of the finest uniform triangular level (default is ‘5’)

‘iterations\_majorant’—the number of iterations of Algorithm 2 (default is ‘10’)

The dual based solver of Section 3.1 is implemented in ‘optimize\_energy\_dual\_mu\_constant\_compact.m’ and the underlying quadratic programming function ‘quadprog’ requires the optimization toolbox of MATLAB to be available. We have exploited the inbuilt trust region reflective algorithm and Hessian multiply function option for lower memory. Evaluation of the primal energy  $J(u_{\lambda_h})$  for a given function  $u_{\lambda_h} \in K$  is done in the function ‘energy’. This function is able to provide an exact quadrature [23] of the energy  $J(v)$  for any function  $v \in K_h$ , including nondifferentiable terms  $\int_{\Omega} v^+ dx, \int_{\Omega} v^- dx$ .

Numerical performance of both algorithms in case of Example II is further reported in Table 4. The number of function evaluations in the quadratic programming of Algorithm 1 (corresponding to solution of the linear system (35)) increases significantly for finer meshes. It indicates that the QP solver is not scalable with respect to the problem size and it performs slowly for finer meshes. A future more efficient implementation deserves some extra testing and preconditioning. We have used default MATLAB tolerances and stopping criteria. It also seems that too high number of iterations of Algorithm 2 has only a marginal impact on the improvement of the index of efficiency

$$I_{\text{eff}} = \frac{\mathcal{M}_+(u_{\lambda_h}, \cdot)}{J(u_{\lambda_h}) - J(u_{\text{ref}})}$$

measuring how much the optimized majorant value over-determines the lower bound of the difference of energies.

### 5. Conclusions and future outlook

A dual based solution algorithm (Algorithm 1) to provide a finite element approximation of the Lagrange multiplier of the perturbed problem was described and tested on two benchmarks in 2D. The finite elements approximation of the primal minimization problem can be easily reconstructed from Lagrange multipliers by solving one linear system of equations. The quality of such approximation is measured in terms of a fully computational functional majorant. A nonlinear part of the optimized functional majorant (by Algorithm 2) seems to work as an indicator of the free boundary. In the future, it would be interesting to consider some preconditioning in order to reduce number of evaluations in the quadratic programming of Algorithm 1. A subsequent minimization of Algorithm 2 might be replaced by some variant of ‘all at one’ coupled algorithm that might converge faster. We also plan to extend the Matlab solver to allow for a mesh adaptivity based on a majorant distribution.

### Acknowledgements

We thank anonymous referees for valuable comments and remarks which improved final exposition of our work. The second author acknowledges the support by GAČR through project GF16-34894L and by MŠMT ČR through project 7AMB16AT015.

## References

- [1] N.N. Uraltseva, Two-phase obstacle problem, *J. Math. Sci.* 106 (3) (2001) 3073–3077.
- [2] G.S. Weiss, The two-phase obstacle problem: Pointwise regularity of the solution and an estimate of the hausdorff dimension of the free boundary, *Interfaces Free Bound.* 3 (2) (2001) 121–128.
- [3] L. Caffarelli, The obstacle problem revisited, *J. Fourier Anal. Appl.* 4 (1998) 383–402.
- [4] D. Kinderlehrer, G. Stampacchia, *An Introduction to Variational Inequalities and their Applications*, Academic Press, 1980.
- [5] H. Shahgholian, N.N. Uraltseva, G.S. Weiss, The two-phase membrane problem – regularity of the free boundaries in higher dimensions, *Int. Math. Res. Not.* (2007).
- [6] R.S. Falk, Error estimates for the approximation of a class of variational inequalities, *Math. Comp.* 28 (128) (1974) 963–971.
- [7] R. Glowinski, J.L. Lions, R. Trémolieres, *Numerical Analysis of Variational Inequalities*, North-Holland, 1981.
- [8] R.H. Nochetto, Sharp  $L^\infty$ -error estimates for semilinear elliptic problems with free boundaries, *Numer. Math.* 54 (3) (1989) 243–255.
- [9] R.H. Nochetto, E. Otárola, A.J. Salgado, Convergence rates for the classical, thin and fractional elliptic obstacle problems, *Phil. Trans. R. Soc. A* 373 (2050) (2015).
- [10] F. Bozorgnia, Numerical solutions of a two-phase membrane problem, *Appl. Numer. Math.* 61 (2011) 92–107.
- [11] A. Arakelyan, R. Barkhudaryan, M. Poghosyan, A finite difference method for two-phase parabolic obstacle-like problem, *Armen. J. Math.* 2 (2015) 164–182.
- [12] A. Arakelyan, A finite difference method for two-phase parabolic obstacle-like problem, *Armen. J. Math.* 7 (2015) 32–49.
- [13] G. Tran, H. Schaeffer, W. Feldman, S. Osher, An  $L^1$  penalty method for general obstacle problems, *SIAM J. Appl. Math.* 75 (4) (2007) 1424–1444.
- [14] S. Repin, J. Valdman, A posteriori error estimates for two-phase obstacle problem, *J. Math. Sci.* 20 (2) (2015) 324–336.
- [15] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, North-Holland, 1976.
- [16] W. Han, *A Posteriori Error Analysis Via Duality Theory*, Springer, 2005.
- [17] S. Repin, *A Posteriori Estimates for Partial Differential Equations*, Walter de Gruyter, 2008.
- [18] P. Neittaanmäki, S. Repin, *Reliable Methods for Computer Simulation (Error Control and a Posteriori Estimates)*, Elsevier, 2004.
- [19] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.
- [20] P. Harasim, J. Valdman, Verification of functional a posteriori error estimates for obstacle problem in 1D, *Kybernetika* 49 (5) (2013) 738–754.
- [21] I. Anjam, J. Valdman, Fast matlab assembly of fem matrices in 2D and 3D: edge elements, *Appl. Math. Comput.* 267 (2015) 252–263.
- [22] T. Rahman, J. Valdman, Fast matlab assembly of fem matrices in 2D and 3D: nodal elements, *Appl. Math. Comput.* 219 (2013) 7151–7158.
- [23] J. Kadlec, J. Valdman, Quadrature of the absolute value of a function. Matlab package. <http://www.mathworks.com/matlabcentral/fileexchange/54183>.