



**Universität Augsburg**

Institut für  
Mathematik

---

---

Dietrich Braess, Ronald H.W. Hoppe, and Joachim Schöberl

**A Posteriori Estimators for Obstacle Problems by the Hypercircle Method**

---

Preprint Nr. 02/2008 — 23. Januar 2008

Institut für Mathematik, Universitätsstraße, D-86135 Augsburg

<http://www.math.uni-augsburg.de/>

---

**Impressum:**

*Herausgeber:*

Institut für Mathematik

Universität Augsburg

86135 Augsburg

<http://www.math.uni-augsburg.de/forschung/preprint/>

*ViSdP:*

Ronald H.W. Hoppe

Institut für Mathematik

Universität Augsburg

86135 Augsburg

*Preprint:* Sämtliche Rechte verbleiben den Autoren © 2008

# A posteriori estimators for obstacle problems by the hypercircle method

Dietrich Braess <sup>\*</sup>    Ronald H.W. Hoppe <sup>†</sup>    Joachim Schöberl <sup>‡</sup>

January 17, 2008

## Abstract

A posteriori error estimates for the obstacle problem are established in the framework of the hypercircle method. To this end, we provide a general theorem of Prager–Synge type. There is now no generic constant in the main term of the estimate. Moreover, the role of edge terms is elucidated, and the analysis also applies to other types of a posteriori error estimators for obstacle problems.

## 1 Introduction

Elliptic obstacle problems often lead to the minimization of a quadratic functional  $J$  on a subspace  $V \subset H^1(\Omega)$

$$J(v) := \frac{1}{2}a(v, v) - (f, v)_0 \tag{1.1}$$

subject to the constraint

$$v(x) \geq \psi(x) \quad \text{for } x \in \Omega \text{ a.e.} \tag{1.2}$$

Here,  $f \in L_2(\Omega)$  and  $\psi \in C(\bar{\Omega})$ . When the problem is solved by the finite element method, the constraint (1.2) is often replaced by pointwise inequalities

$$v_h(x_i) \geq \psi(x_i) \tag{1.3}$$

for all nodal points  $x_i$  of the grid. This discretization is natural, but it implies some extra terms when a posteriori error estimates are computed; see, e.g., [2, 4, 5, 9, 14]. The reason is that Lagrange multipliers for the constraints (1.3) are point functionals. An extension of the functional to  $H^{-1}(\Omega)$  without a violation of the complementarity condition cannot be guaranteed.

The complication is less severe when an a posteriori error estimate is determined by the hypercircle method [11] that was made popular, e.g., by [10]. The procedure known from linear theory can be adapted for the obstacle problem. We note that no extra term occurs

---

<sup>\*</sup>Institute of Mathematics, Ruhr-University of Bochum, D-44801 Bochum, Germany

<sup>†</sup>Department of Mathematics, University of Houston, Houston, TX 77204-3008, U.S.A., and Institute of Mathematics, University of Augsburg, D-86159 Augsburg, Germany

<sup>‡</sup>Dept. of Math. and Center for Comput. Engrg. Science, RWTH Aachen, D-52074 Aachen, Germany

whenever the active point set has some regularity, i.e., if it is the closure of its interior. Otherwise, a generalization of the Prager–Synge theorem for the obstacle problem also yields an additional error term as we find in well-known estimators. Here, the resulting extra term enters into an exact expression for the error. It is therefore clear that it does not spoil the efficiency of the error estimate.

We note that Weiss and Wohlmuth [15] observed a similar phenomenon when they considered inequality constraints on the boundary of the domain  $\Omega$ . On the other hand, Repin [12] considered the hypercircle method without the regularity assumption of the active set, and his result is closer to the classical estimates than to our result with the hypercircle method.

A patch-oriented construction following Braess and Schöberl [7] turns out to be appropriate here, since the Lagrange multipliers for the discretized obstacle problem are associated with the finite element equations on patches. Although there is a strong relation to the determination of estimators by local Neumann problems (see, e.g., [1]), the latter is focused on element-oriented constructions, and the considerations on patches only occur in auxiliary steps.

Often the edge terms are dominating in a posteriori estimates for linear elliptic problems. Those terms, however, may overestimate the error when obstacle problems are considered. We can eliminate this effect in certain cases by relaxing the regularity requirement for the hypercircle method.

The paper is organized as follows. Section 2 will provide a general theorem of Prager–Synge type. The prerequisites from finite element theory are presented in Section 3. The construction of the a posteriori error estimate in Section 4 will be organized such that no extra term arises if possible. The efficiency will be treated in Sections 5 and 6. In two appendices, a discussion of the role of edge terms in a posteriori estimates elucidates the situation not only for the hypercircle method.

## 2 A theorem of Prager–Synge type

For convenience, we restrict ourselves to the Poisson equation with homogeneous Dirichlet boundary conditions, i.e.,

$$a(u, v) := \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx$$

and  $V = H_0^1(\Omega)$ . The energy  $J$  is to be minimized on the convex set

$$\mathcal{K} := \{v \in V \mid v \geq \psi \text{ a.e. in } \Omega\}.$$

The solution  $u$  is known to be characterized by

$$a(u, v - u) \geq (f, v - u)_0, \quad v \in \mathcal{K}. \tag{2.1}$$

The Lagrange multiplier  $\lambda$  is defined by

$$\langle \lambda, v \rangle := a(u, v) - (f, v)_0, \quad v \in V.$$

It follows from (2.1) that for all  $w \in V_+$

$$\langle \lambda, u - \psi \rangle = 0 \text{ and } \langle \lambda, w \rangle \geq 0, \tag{2.2}$$

if we set  $V_+ := \{v \in V \mid v(x) \geq 0 \text{ a.e.}\}$ . Moreover, we have

$$\begin{aligned} J(v) - J(u) & \\ &= \frac{1}{2} \|\nabla v - \nabla u\|_0^2 + \langle \lambda, v - u \rangle \quad \forall v \in V, \end{aligned} \tag{2.3}$$

and each term on the right-hand side of (2.3) is nonnegative, if  $v \in \mathcal{K}$ . Often it is more natural to give the error in terms of the difference  $J(v) - J(u)$  and not by the energy norm. For instance, a convergence analysis of an adaptive P1 conforming finite element approximation of (1.1), (1.2) in the sense of a guaranteed reduction of the objective functional  $J$  has been provided in [5]; cf. also [13] for an approach using quadratic programming techniques.

The dual problem is the maximization of the Trefftz functional

$$J^*(\tau) := -\frac{1}{2} \|\tau\|_0^2 - (\operatorname{div} \tau + f, \psi)_0$$

on the dual convex cone

$$\mathcal{F} := \{\tau \in H(\operatorname{div}) \mid \operatorname{div} \tau + f \leq 0 \text{ a.e.}\}.$$

It is known that there is no duality gap, i.e.,  $J(u) = J^*(\nabla u)$ , and

$$\begin{aligned} J^*(\nabla u) - J^*(\tau) & \\ &= \frac{1}{2} \|\nabla u - \tau\|_0^2 - (\operatorname{div} \tau + f, u - \psi)_0 \\ &\quad \text{for all } \tau \in \mathcal{F}. \end{aligned} \tag{2.4}$$

**Theorem of Prager and Synge type for obstacle problems.**

Let  $v \in \mathcal{K} \subset H_0^1(\Omega)$  and  $\tau \in \mathcal{F} \subset H(\operatorname{div})$ . Then

$$\begin{aligned} 2[J(v) - J(u)] + 2[J^*(\nabla u) - J^*(\tau)] & \\ &= [\|\nabla u - \nabla v\|_0^2 + 2\langle \lambda, v - u \rangle] \\ &\quad + [\|\nabla u - \tau\|_0^2 + 2(\operatorname{div} \tau + f, \psi - u)_0] \\ &= \|\nabla v - \tau\|_0^2 + 2(\operatorname{div} \tau + f, \psi - v)_0. \end{aligned} \tag{2.5}$$

Furthermore, if  $v$  and  $\tau$  satisfy the complementarity condition

$$(\operatorname{div} \tau + f, v - \psi)_0 = 0, \tag{2.6}$$

then

$$\begin{aligned} \|\nabla u - \nabla v\|_0^2 + \|\nabla u - \tau\|_0^2 & \\ \leq 2[J(v) - J(u)] + \|\nabla u - \tau\|_0^2 & \\ \leq \|\nabla v - \tau\|_0^2. & \end{aligned} \tag{2.7}$$

*Remark.* We note that all inner products in (2.5) are nonnegative.

**Proof:** Since the boundary terms vanish when partial integration is applied, we have

$$\begin{aligned}
& (\nabla u - \tau, \nabla v - \nabla u)_0 \\
&= (\nabla u, \nabla v - \nabla u)_0 + (\operatorname{div} \tau, v - u)_0 \\
&= (\nabla u, \nabla v - \nabla u)_0 \\
&\quad - (f, v - u)_0 + (\operatorname{div} \tau + f, v - u)_0 \\
&= \langle \lambda, v - u \rangle + (\operatorname{div} \tau + f, v - u)_0.
\end{aligned}$$

Now, the binomial formula is applied to the sum  $[\nabla v - \nabla u] + [\nabla u - \tau]$  to obtain

$$\begin{aligned}
\|\nabla v - \tau\|_0^2 &= \|\nabla v - \nabla u\|_0^2 + \|\nabla u - \tau\|_0^2 \\
&\quad + 2 \langle \lambda, v - u \rangle + 2(\operatorname{div} \tau + f, v - u)_0 \\
&= 2[J(v) - J(u)] + 2[J^*(\nabla u) - J^*(\tau)] \\
&\quad + (\operatorname{div} \tau + f, v - \psi)_0.
\end{aligned}$$

This proves (2.5). The inequality (2.7) follows from (2.3) and (2.4), and the proof is complete.  $\blacksquare$

Obviously, the last term in (2.5) corresponds to the extra term in classical estimators. It will be avoided whenever possible.

**Remark 2.1** *We emphasize that the assumption  $\tau \in H(\operatorname{div})$  may be dropped in Prager and Synge's theorem, if we set  $\langle \operatorname{div} \tau, w \rangle = -(\tau, \nabla w)$  for  $w \in H^1$ . In particular, if  $\tau$  belongs to the broken  $H(\operatorname{div})$  space introduced in Section 4, we have*

$$\begin{aligned}
\langle \operatorname{div} \tau, w \rangle & \tag{2.8} \\
&= \sum_T (\operatorname{div} \tau, w)_{0,T} - \sum_e \int_e [\tau \cdot n] w ds.
\end{aligned}$$

*The benefit of this observation will be elucidated in Appendix B with a one-dimensional example. The additional freedom will be used on edges in the contact zone (also in higher dimensions), in particular, if the obstacle is specified by a non-affine function.*

### 3 The Lagrange multiplier for the finite element solution

The discretization of the obstacle problem means that the linear space is replaced by a finite element space  $V_h$ , which will be here the space of linear elements on a triangulation  $\mathcal{T}_h$  of  $\Omega \subset \mathbb{R}^2$ . As usual,  $\Omega$  is assumed to be a polygonal domain, and the obstacle is given by a piecewise linear function  $\psi \in V_h$ .

The corresponding Lagrange multiplier  $\lambda_h$  is defined by

$$\langle \lambda_h, w \rangle = a(u_h, w) - (f, w)_0 \tag{3.1}$$

for  $w \in V_h$ . Since the right-hand side is defined for all  $w \in V$ , we obtain an extension of  $\lambda_h$  to  $V'$  by (3.1). Partial integration yields the representation

$$\begin{aligned} \langle \lambda_h, w \rangle & \\ &= - \sum_T (f, w)_{0,T} + \sum_e ([\frac{\partial u_h}{\partial n}], w)_{0,e}. \end{aligned} \quad (3.2)$$

It shows that this extension of the Lagrange multiplier contains also the information on the residues outside the coincidence set.

From a computational point of view, it is given by the nonnegative residues of the finite element equations in the contact zone. Let  $\phi_i \in V_h$  be the nodal basis function associated with the nodal point  $x_i$ . Then

$$\lambda_{h,i} := \langle \lambda_h, \phi_i \rangle \geq 0 \quad \text{for all } i, \quad (3.3)$$

is the residue in the finite element inequalities and

$$\lambda_{h,i} = 0, \quad \text{if } u_h(x_i) - \psi(x_i) > 0. \quad (3.4)$$

Therefore, the discrete complementarity condition  $\langle \lambda_h, u_h - \psi \rangle = 0$  holds and

$$\langle \lambda_h, w_h \rangle = \sum_i \lambda_{h,i} w_h(x_i), \quad w_h \in V_h. \quad (3.5)$$

The support of  $\phi_i$  is the patch

$$\omega_i := \bigcup \{ \bar{T} \in \mathcal{T}_h \mid x_i \in \partial T \}.$$

The coincidence set (active set)  $\mathcal{A}_h := \{x \in \Omega \mid u_h(x) = \psi(x)\}$  is called *regular*, if it is the closure of its interior. This means that each nodal point  $x_i \in \mathcal{A}_h$  lies on the boundary of a triangle  $T$  which is contained in the coincidence set  $\mathcal{A}_h$ .

## 4 Equilibration

The main task in the determination of the a posteriori error estimate is the construction of a function  $\sigma$  that satisfies  $\text{div } \sigma \leq -f$  and moreover the complementarity condition (2.6) whenever possible. (The original requirement  $\sigma \in H(\text{div})$  will be relaxed.) The procedure is called *equilibration*. Following [7] we construct an appropriate  $\tau$  in the *broken Raviart-Thomas space*

$$\begin{aligned} RT_{-1} := \{ & \tau \in L_2(\Omega) \mid \tau(x) = a_T + b_T x \\ & \text{with } a_T \in \mathbb{R}^2, b_T \in \mathbb{R} \text{ in each } T \}. \end{aligned}$$

More precisely, the resulting  $\sigma$  will satisfy

$$\begin{aligned} \text{div } \sigma + \bar{f} & \leq 0, \\ [\sigma \cdot n] & \geq 0, \end{aligned} \quad (4.1)$$

where  $\bar{f}$  is the  $L_2$  projection of  $f$  in the space of piecewise constant functions. This means that we separate the *data oscillation* from the main term of the estimate; cf. [9]. As usually, we consider the associated error term

$$ch\|f - \bar{f}\|_0$$

as a term of higher order.

In contrast to the treatment of linear elliptic problems, inequalities are admitted in (4.1). In addition, the complementarity conditions

$$\begin{aligned} (-\operatorname{div} \sigma - \bar{f}, u_h - \psi)_{0,T} &= 0, \\ ([\sigma \cdot n], u_h - \psi)_{0,e} &= 0 \end{aligned} \quad (4.2)$$

will be satisfied at least outside a neighborhood of the coincidence set, since the terms on the left-hand side of (4.2) enter into the error bound. We will silently adopt this point and repeat it only when necessary.

We recall that the finite element functions in the Raviart–Thomas space

$$RT := RT_{-1} \cap H(\operatorname{div}),$$

are specified by their normal components on the edges of the grid. Similarly the functions in the broken Raviart–Thomas space  $RT_{-1}$  are given, if the normal components are known on both sides of the edges.

Obviously,  $\nabla u_h$  is a broken Raviart–Thomas function with zero divergence in each triangle. The required  $\sigma$  will be obtained by a correction that eliminates the jumps of the normal components on the edges. Specifically, the correction

$$\sigma^\Delta := \tau - \nabla u_h \quad (4.3)$$

shall satisfy the following properties:

$$\begin{aligned} [\sigma^\Delta \cdot n] &\geq -\left[\frac{\partial u_h}{\partial n}\right] \text{ on each edge } e, \\ \operatorname{div} \sigma^\Delta &\leq -\bar{f} \text{ on each triangle } T. \end{aligned} \quad (4.4)$$

The desired  $\sigma^\Delta$ , in turn, will be computed as a sum of local corrections with support in the patches  $\omega_i$ ,

$$\sigma^\Delta = \sum_i \sigma_{\omega_i} \text{ with } \operatorname{supp} \sigma_{\omega_i} = \omega_i, \quad (4.5)$$

and

$$\begin{aligned} [\sigma_{\omega_i} \cdot n] &\geq -\frac{1}{2} \left[\frac{\partial u_h}{\partial n}\right], & e \in \omega_i, \\ \operatorname{div} \sigma_{\omega_i} &\leq -f_{T,i}, & T \subset \omega_i, \\ \sigma_{\omega_i} \cdot n &= 0 & \text{on } \partial\omega_i \setminus \partial\Omega, \end{aligned} \quad (4.6)$$

where

$$f_{T,i} := (1/|T|) \int_T f \phi_i dx.$$

Since each interior edge of the triangulation  $\mathcal{T}_h$  belongs to two patches and  $\sum_i \int_T f \phi_i dx = \int_T f 1 dx = |T| \bar{f}$ , the properties (4.6) imply (4.4).



**Lemma 4.1** *Let  $x_i \in \overline{\Omega} \setminus \partial\Omega$  be a node of the triangulation, and let  $\phi_i \in V_h$  be the nodal basis function with  $\phi_i(x_i) = 1$  and  $\phi_i(x) = 0$  for  $x \in \Omega \setminus \omega_i$ . Then*

$$\begin{aligned} & \frac{1}{2} \sum_{e \subset \omega_i} \int_e \left[ \frac{\partial u_h}{\partial n} \right] ds \\ &= \sum_{T \subset \omega_i} \int_T f \phi_i dx + \lambda_{h,i}. \end{aligned} \quad (4.7)$$

**Proof:** Since  $u_h$  is the finite element solution in  $V_h$ , we obtain from (3.1) with  $w = \phi_i$ :

$$\int_{\omega_i} \nabla u_h \nabla \phi_i dx = \int_{\omega_i} f \phi_i dx + \lambda_{h,i}. \quad (4.8)$$

We recall that  $\partial u_h / \partial n$  is constant and  $\phi_i$  is linear on each edge. Partial integration of the left-hand side of (4.8) yields

$$\begin{aligned} \int_{\omega_i} \nabla u_h \nabla \phi_i dx &= \sum_{T \subset \omega_i} \int_{\partial T} \frac{\partial u_h}{\partial n} \phi_i dx \\ &= \sum_{e \subset \omega_i} \int_e \left[ \frac{\partial u_h}{\partial n} \right] \phi_i ds \\ &= \frac{1}{2} \sum_{e \subset \omega_i} \int_e \left[ \frac{\partial u_h}{\partial n} \right] ds. \end{aligned} \quad (4.9)$$

Now, the assertion of the lemma follows from (4.8) and (4.9). ■

Next, we consider a patch around a node  $x_i$ . The desired function  $\sigma_{\omega_i}$  will be specified by the integral fluxes  $\int_e \sigma_{\omega_i}|_T \cdot n ds$  on the two sides of each edge  $e \subset T \subset \omega_i$ . The boundary condition (4.6)<sub>3</sub> will always silently be assumed to hold.

First, we describe a cheap construction and distinguish four cases. In all cases Algorithm 4.2 will be applied. It follows the procedure known from linear theory; see [3, p.184]. Here the input contains some extra parameters  $\Lambda_{T,i}$  and  $\Lambda_{e,i}$  to cope with excess sources and sinks. In particular, the parameters  $\Lambda_{T,i}$  vanish in Case 1 below as in the linear case. The parameters  $\Lambda_{e,i}$  are set to zero in all four cases and will be activated only later. The notation for the algorithm is specified in Fig. 1.

Later, we will present a version with an optimization process in order to improve the efficiency of the estimator.

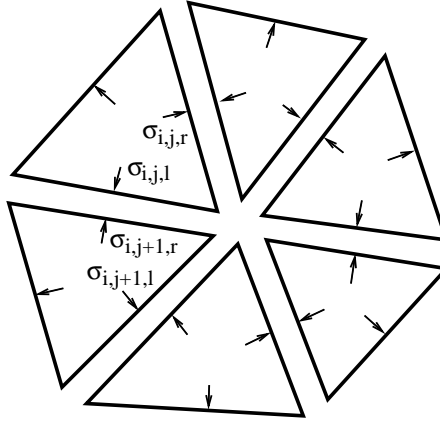


Figure 1: Fluxes in a patch around a vertex  $x_i$ .  $\sigma_{i,j,r}$  and  $\sigma_{i,j,l}$  are the normal components of the fluxes that leave the triangle  $T_j$  on the right and left side, respectively. The triangles are enumerated counter clockwise, and  $e_j = \partial T_j \cap \partial T_{j+1}$  (with indices modulo the number of triangles)

**Algorithm 4.2**

Set  $\sigma_{i,1,r} = 0$ ;

for  $j = 1, 2, \dots$ , until an entire circuit around  $x_i$  is completed (or an edge on  $\partial\Omega$  is met)

{

$$\text{fix } \sigma_{i,j,l} \text{ such that } \int_{e_j} \sigma_{\omega_i} \cdot n ds =$$

$$\int_{T_j} f \phi_i dx - \int_{e_{j-1}} \sigma_{\omega_i} \cdot n ds + \Lambda_{T_j,i};$$

$$\text{fix } \sigma_{i,j+1,r} \text{ such that } [\sigma_{\omega_i} \cdot n] =$$

$$-\frac{1}{2} [\nabla u_h \cdot n] + \Lambda_{e_j,i} \text{ on } e_j;$$

}

The fluxes define a preliminary  $\tilde{\sigma}$  with support  $\omega_i$ . Add a constant  $\alpha$  to all  $\sigma_{i,j,l}$  and  $\sigma_{i,j,r}$  for which  $\|\sigma\|_0$  is minimal. ■

The two rules within the braces care that (4.6)<sub>1</sub> and (4.6)<sub>2</sub> hold. Since an additive constant  $\alpha$  does not change (4.6), it is fixed in the last step for minimizing the  $L_2$ -norm.

**Remark 4.3** *There is an easy interpretation. By Gauss' law the normal components of the fluxes on the three edges of a triangle determine the magnitude of the source or sink in a triangle. Similarly, there is a source or sink between the two sides of an edge that is given by the jump of the flux on the edge and its length. Lemma 4.1 and (4.7) assert that the sum of all of them in a patch vanishes.*

*Case 1.*  $x_i \in \Omega \setminus \partial\Omega$  and  $u_h(x_i) > \psi(x_i)$ .

Here (4.7) holds with  $\lambda_{h,i} = 0$ , and we apply Algorithm 4.2 with

$$\Lambda_{T,i} = 0 \quad \text{for all } T.$$

*Case 2.*  $x_i \in \Omega \setminus \partial\Omega$ ,  $u_h(x_i) = \psi(x_i)$ , and  $u_h = \psi$  holds at least in one triangle  $T \subset \omega_i$ .

Let  $m = m_i$  be the number of triangles in the patch on which  $u_h = \psi$  holds. By assumption,  $m \geq 1$ . We set

$$\Lambda_{T,i} := \begin{cases} \frac{1}{m} \lambda_{h,i}, & \text{if } u_h(x) = \psi(x), \ x \in T, \\ 0, & \text{otherwise.} \end{cases} \quad (4.10)$$

The algorithm now yields a correction that satisfies (4.6)<sub>1</sub>, (4.6)<sub>2</sub> and the complementarity relation

$$\sum_{T \subset \omega_i} (\operatorname{div} \sigma_{\omega_i}, u_h - \psi)_0 = 0.$$

*Case 3.*  $x_i \in \Omega \setminus \partial\Omega$ ,  $u_h(x_i) = \psi(x_i)$ , and  $u_h(x) \neq \psi(x)$  holds for at least one point in each triangle  $T \subset \omega_i$ .

Let  $m$  be the number of triangles in  $\omega_i$ . Set  $\Lambda_{T,i} = \frac{1}{m} \lambda_{h,i}$ . The algorithm yields a correction that satisfies (4.6)<sub>1</sub>, (4.6)<sub>2</sub>, but there will now be a nonzero contribution of the complementarity term to the error estimate.

*Case 4.*  $x_i \in \partial\Omega$ .

The edges and triangles in  $\omega_i$  are enumerated such that the algorithm starts at an edge on  $\partial\Omega$  and stops at the other edge on the boundary. Since the circuit is incomplete, no condition has to be satisfied, and we can perform the algorithm with  $\Lambda_{T,i} = 0$  for all  $T \subset \omega_i$ .

By construction, (4.6) is guaranteed in all four cases.

The case 1 corresponds to the construction for linear elliptic equations [3, p. 184], and it is optimal in the framework of local procedures. In the cases 2 and 3 the efficiency of the result can be improved. Instead of fixing the variables  $\Lambda_{T,i}$  a priori, they are determined by a small quadratic program. In order to have a unified description and to avoid the distinction of the cases, the optimization is generally included:

$$\begin{aligned} \eta_{PS,i} &:= & (4.11) \\ \min\{ & \|\tau\|_0^2 + 2(\operatorname{div} \tau + f\phi_i, \psi - u_h)_0 \\ & + 2[(\tau + \nabla u_h) \cdot n], u_h - \psi)_{0, \cup e} \} \end{aligned}$$

with  $\tau$  determined by Algorithm 4.2 executed with parameters subject to the constraints

$$\begin{aligned} \sum_{T \subset \omega_i} \Lambda_{T,i} + \sum_{e \subset \omega_i} |e| \Lambda_{e,i} &= \lambda_{h,i}, \\ \Lambda_{T,i} &\geq 0 \quad \text{for all } T \subset \omega_i, \\ \Lambda_{e,i} &\geq 0 \quad \text{for all } e \subset \omega_i. \end{aligned}$$

■

The optimization problem is solvable, since a feasible solution exists. This follows from the procedure with a priori fixed parameters. If the node in the interior of the patch does not

belong to the coincidence set, then  $\lambda_{h,i} = 0$ , all slack variables vanish, and the optimization is trivial. We have not used the short notation with (2.8) in order to see the jumps more clearly.

After summing the corrections on all the patches we obtain the final estimate.

**Theorem 4.4** *Let each  $\sigma_{\omega_i}$  be determined as described above and  $\sigma^\Delta$  by (4.5). Then we have the a posteriori error estimate*

$$\begin{aligned} \|\nabla u - \nabla u_h\|_0^2 &\leq J(u_h) - J(u) \\ &\leq \|\sigma^\Delta\|_0^2 + ch^2 \|f - \bar{f}\|_0^2 \\ &\quad + 2(\operatorname{div} \sigma^\Delta + \bar{f}, \psi - u_h)_0 \\ &\quad + 2([\sigma^\Delta \cdot n], u_h - \psi)_{0, \cup e}. \end{aligned}$$

Here, the term in the third line gets nonzero contributions only via Step 3, and in each triangle  $T$ ,

$$-\operatorname{div} \sigma^\Delta - \bar{f} = \frac{1}{|T|} \sum_i \Lambda_{T,i}.$$

## 5 Relation to residual estimators

By Theorem 4.4, the hypercircle method results in a reliable estimator. For studying its efficiency we will compare the estimator with the classical ones for the obstacle problem. In particular, we focus on residual estimators. The relation to the true error of the finite element solution will be investigated in the next section.

The optimization problem (4.11) on a patch  $\omega_i$  will be modified to achieve a simpler, but equivalent one. For simplicity we drop the index  $i$  whenever there is no danger of confusion. (There is, e.g., the exception  $f_{T,i}$ .) First, (4.11) is rewritten,

$$\eta_{PS} = \min \left\{ \|\tau\|_0^2 + 2 \langle \operatorname{div} \tau + f\phi_i, \psi - u_h \rangle \right\}, \quad (5.1)$$

subject to

$$\begin{aligned} \operatorname{div} \tau &\leq -f_{T,i}, \quad \tau \in RT_{-1}(\omega), \\ [\tau \cdot n] &\geq -(1/2)[\nabla u_h \cdot n], \\ \tau \cdot n &= 0 \quad \text{on } \partial\omega. \end{aligned} \quad (5.2)$$

For  $\tau \in RT_{-1}(\omega)/\ker(\operatorname{div} RT_{-1})$ , a scaling argument shows that

$$c_1 \|\tau\|_0^2 \leq \|\operatorname{div} \tau\|_{-1,h}^2 \leq c_1^{-1} \|\tau\|_0^2,$$

where

$$\begin{aligned} &\|\operatorname{div} \tau\|_{-1,h}^2 \\ &:= \sum_{T \subset \omega} h_T^2 \|\operatorname{div} \tau\|_{0,T}^2 + \sum_{e \subset \omega} h_e \|\tau \cdot n\|_{0,e}^2. \end{aligned}$$

The jumps of  $\tau$  on the edges are now also considered as (distributional) parts of  $\operatorname{div} \tau$  in the spirit of Remarks 2.1 and 4.3. Therefore, we define  $s = \operatorname{div} \tau$  by setting

$$\begin{aligned} s_T &:= (\operatorname{div} \tau)_T, \\ s_e &:= -[\tau \cdot n]_e. \end{aligned}$$

In particular,  $s$  is given by  $2m$  real numbers if  $\omega$  consist of  $m$  triangles and  $\tau \in RT_{-1}(\omega)$ . Hence,

$$\begin{aligned} c_1 \eta_{PS} \leq \eta_s &:= \\ \min \left\{ \|s\|_{-1,h}^2 + 2(s + f\phi_i, \psi - u_h)_{0,\omega} \right. \\ &\quad \left. + 2\left(s + \frac{1}{2}[\nabla u_h \cdot n], u_h - \psi\right)_{0,\cup e} \right\}, \end{aligned} \tag{5.3}$$

subject to

$$\begin{aligned} s_T &\leq -f_{T,i}, \\ s_e &\leq (1/2)[\nabla u_h \cdot n], \\ \langle s, 1 \rangle &= 0. \end{aligned} \tag{5.4}$$

Here, the total divergence on the patch is defined by

$$\langle s, 1 \rangle := \sum_{T \subset \omega} |T| s_T + \sum_{e \subset \omega} |e| s_e.$$

Equation (5.4)<sub>3</sub> was hidden in (5.1) by the condition  $\tau \cdot n = 0$  on  $\partial\omega$ .

The elimination of the condition (5.4)<sub>3</sub> will make the construction simpler.

**Lemma 5.1** *Assume that  $u_h = \psi$  holds in at least one triangle of  $\omega$ , and let  $\omega$  consist of  $m$  triangles. Let*

$$\begin{aligned} \eta_{s,2} &:= \min \left\{ \|s\|_{-1,h}^2 + 2(s + f\phi_i, \psi - u_h)_{0,\omega} \right. \\ &\quad \left. + 2\left(s + \frac{1}{2}[\nabla u_h \cdot n], u_h - \psi\right)_{0,\cup e} \right\}, \end{aligned} \tag{5.5}$$

subject to

$$\begin{aligned} s_T &\leq -f_{T,i}, \\ s_e &\leq (1/2)[\nabla u_h \cdot n] \end{aligned}$$

be the error estimator without the constraint (5.4)<sub>3</sub>. Then

$$\eta_s \geq \eta_{s,2} \geq (1 + \sqrt{2m})^{-2} c_2 \eta_s \tag{5.6}$$

with the constant  $c_2$  depending only on the shape parameter.

**Proof:** The two terms in (5.5)<sub>1</sub> are nonnegative. There is a constant  $c_2$  that depends only on the shape parameter such that

$$c_2 h_T^2 \leq |T| \leq c_2^{-1} h_T^2.$$

Therefore, we consider the minimization of the equivalent expression

$$\begin{aligned} & \sum_{T \subset \omega} (|T|s_T)^2 + \sum_{e \subset \omega} (|e|s_e)^2 \\ & + 2(s + f\phi_i, \psi - u_h)_{0,\omega} \\ & + 2\left(s + \frac{1}{2}[\nabla u_h \cdot n], u_h - \psi\right)_{0,\cup e}. \end{aligned}$$

Let  $s$  be the minimizer of the problem (5.5). We construct a feasible candidate  $\tilde{s}$  that satisfies the averaging constraint (5.4)<sub>3</sub>, and the functional will increase only by the given  $m$ -dependent factor.

Case (a): Assume that  $\langle s, 1 \rangle > 0$ .

Let  $T$  be a triangle with  $u_h = \psi$  on  $T$ . We set  $\tilde{s} := s$  and redefine it on the special triangle

$$\tilde{s}_T := s_T - |T|^{-1} \langle s, 1 \rangle$$

without changing the other values. A straight forward calculation shows that  $\|\tilde{s} - s\|^2 \leq 2m\|s\|^2$ . Hence,  $\|\tilde{s}\| \leq (1 + \sqrt{2m})\|s\|$  holds for the modified  $(-1, h)$ -norm. Obviously,  $\tilde{s}$  is feasible due to the negative correction, and the second term on the right-hand side of (5.5) does not change. The assertion holds in this case.

Case (b). Assume that  $\langle s, 1 \rangle < 0$ . Set

$$\hat{s}_T := f_{T,i}, \quad \hat{s}_e := \hat{s}_e := s_e.$$

Since there exists a feasible solution of the minimization problem, it follows that  $\langle \hat{s}, 1 \rangle \geq 0$ . For  $T \in \omega$ , the convex combination

$$\tilde{s}_T := s_T + \frac{-\langle s, 1 \rangle}{\langle \hat{s}, 1 \rangle - \langle s, 1 \rangle} |T|^{-1} (\hat{s}_T - s_T)$$

yields a feasible solution. Since we have distributed the mean value on one or more triangles, the  $\|\cdot\|_{1,h}$  norm is not more increased than in case 1. The second term in (5.5) was diminished or unchanged by the choice above, and the proof of the nontrivial part is complete.

The inequality  $\eta_s \geq \eta_{s,2}$  is obvious. ■

Since the quadratic terms in (5.5) are diagonal, the variables are now separated in the problem, and the minimizer is easily determined. Adding now the label of the patch, we have

$$\begin{aligned} s_{T,i} = R_{T,i} & := \min\{f_{T,i}, \overline{h^{-2}u_h - \psi}\}, \\ s_{e,i} = R_{e,i} & := \min\left\{\frac{1}{2}\left[\frac{\partial u_h}{\partial n}\right], \overline{h^{-1}u_h - \psi}\right\}. \end{aligned} \tag{5.7}$$

Here and in the sequel, an overlined quantity refers to the mean value on the subset under consideration. We associate to the choice (5.7) residual type error estimators. There are element terms (area-based terms)

$$\begin{aligned} \eta_{T,i} & \\ & := h_T^2 (-f_{T,i}, \min\{-f_{T,i}, \overline{h^{-2}u_h - \psi}\})_{0,T} \\ & = \begin{cases} h_T^2 \|f_{T,i}\|_{0,T}^2, & \text{if } -f_{T,i} \leq \overline{h_T^{-2}u_h - \psi}, \\ (-f_{T,i}, \overline{u_h - \psi})_{0,T}, & \text{otherwise} \end{cases} \end{aligned} \tag{5.8}$$

and, with the abbreviation  $\lambda_e := [\frac{\partial u_h}{\partial n}]$ , the edge terms

$$\begin{aligned}\eta_{e,i} &:= h_e(\lambda_e, \min\{\lambda_e, h^{-1} \overline{u_h - \psi^e}\})_{0,e} \\ &= \begin{cases} h_e \|\lambda_e\|_{0,e}^2, & \text{if } h_e \lambda_e \leq \overline{u_h - \psi^e} \\ (\lambda_e, u_h - \psi)_{0,e}, & \text{otherwise.} \end{cases} \end{aligned} \quad (5.9)$$

The quantities above reflect the fact that a continuous transition between the coincidence set and the points in its neighborhood is reasonable.

**Theorem 5.2** *The Prager-Synge error estimator is equivalent to the residual error estimator, i.e.,*

$$\eta_{PS,i} \approx \sum_{T \subset \omega_i} \eta_{T,i} + \sum_{e \subset \omega_i} \eta_{e,i}.$$

**Proof:** We present the proof for the element terms, the edge terms can be treated in the same way. Moreover, we recall the equivalence of  $\eta_{PS,i}$  with the variants in Lemma 5.1.

*Case 1.*  $-f_{T,i} \leq h_T^{-2} \overline{u_h - \psi}$ . Then we have  $s_T = -f_{T,i}$  and the contribution of the element  $T$  to  $\eta_s$  in (5.3) is

$$\begin{aligned}\eta_{s,T} &= h_T^2 \|f_{T,i}\|_{0,T}^2 \\ &\quad + 2(-f_{T,i} + f_{T,i}, \psi - u_h)_{0,T} \\ &= h_T^2 \|f_{T,i}\|_{0,T}^2.\end{aligned}$$

*Case 2.*  $-f_{T,i} > h_T^{-2} \overline{u_h - \psi}$ . Then we have  $s_T = h_T^{-2} \overline{u_h - \psi}$ . Note that the two contributions are nonnegative. Hence, we may multiply by a factor of two for eliminating inconvenient terms

$$\begin{aligned}\eta_{s,T} &= h_T^{-2} \|\overline{u_h - \psi}\|_{0,T}^2 + \\ &\quad + 2(-h^{-2} \overline{u_h - \psi} - f_{T,i}, u_h - \psi)_{0,T} \\ &\leq 2h_T^{-2} \|\overline{u_h - \psi}\|_{0,T}^2 + \\ &\quad + 2(-h^{-2} \overline{u_h - \psi} - f_{T,i}, \overline{u_h - \psi})_{0,T} \\ &= 2(-f_{T,i}, \overline{u_h - \psi})_{0,T}.\end{aligned}$$

Similarly, by taking half of the second term we obtain  $\eta_{s,T} \geq (-f_{T,i}, \overline{u_h - \psi})_{0,T}$ . ■

## 6 Efficiency

We proceed with the analysis of the efficiency and focus our attention on the hypercircle method. However, the results will be of interest for residual-type estimators as well. First, we see that solving local Dirichlet problems is efficient.

**Lemma 6.1** *Assume that  $v_i \in H_0^1(\omega_i)$  and  $v_i \geq \psi - u_h$ . Then*

$$\begin{aligned}\sum_i \{J(u_h) - J(u_h + v_i)\} \\ \leq 3 \{J(u_h) - J(u)\}.\end{aligned}$$

**Proof:** Let each  $v_i$  satisfy the assumption of the lemma, and let  $m$  be the maximal number of overlapping patches. Obviously,  $m \leq 3$  holds in 2-space. The element  $w := \frac{1}{m} \sum v_i$  is in the convex set, and thus

$$J(u_h) - J(u_h + w) \leq J(u_h) - J(u) \quad (6.1)$$

From Young's inequality it follows that

$$\|w\|_a^2 = \left\| \frac{1}{m} \sum v_i \right\|_a^2 \leq \frac{1}{m} \sum \|v_i\|_a^2.$$

The differences of the energies evaluates to

$$\begin{aligned} J(u_h) - J(u_h + v) & \quad (6.2) \\ &= \frac{1}{2}a(u_h, u_h) - (f, u_h)_0 - \frac{1}{2}a(u_h, u_h) \\ & \quad - a(u_h, v) - \frac{1}{2}a(v, v) + (f, u_h + v)_0 \\ &= -a(u_h, v) + (f, v)_0 - \frac{1}{2}a(v, v) \\ &= -\langle \lambda_h, v \rangle - \frac{1}{2}a(v, v). \end{aligned}$$

By applying (6.2) to  $v = w$  and  $v = v_i$  and recalling (6.2) we obtain

$$\begin{aligned} & \sum_i \{J(u_h) - J(u_h + v_i)\} \\ &= \sum_i \left\{ -\langle \lambda_h, v_i \rangle - \frac{1}{2}\|v_i\|^2 \right\} \\ &\leq -\langle \lambda_h, mw \rangle - m \frac{1}{2}\|w\|^2 \\ &= m \{J(u_h) - J(u_h + w)\} \\ &\leq m \{J(u_h) - J(u)\}. \end{aligned}$$

■

Following Lemma 6.1 we will construct a correction  $v \geq \psi - u_h$  such that the improvement (6.2) dominates the residual error estimator. This shows its efficiency. Some elementary properties of the element bubble functions  $b_T$  and the edge bubble functions  $b_e$  are required. They are defined in terms of the barycentric coordinates

$$b_T := \lambda_1 \lambda_2 \lambda_3, \quad b_e := \lambda_1 \lambda_2.$$

**Lemma 6.2** (1) *Let  $g$  be a linear function that is non-negative on a triangle  $T$  and  $\bar{g}$  be its mean-value on  $T$ . Then*

$$12\bar{g}b_T(x) \leq g(x) \quad \text{for all } x \in T. \quad (6.3)$$

(2) *There is a constant  $c \approx 1$  ( $c \leq 12$ ) such that*

$$c \leq \frac{\int b_T}{\|\nabla b\|^2 h^2} \quad (6.4)$$



and

$$\int b_T = \frac{1}{60}|T|.$$

**Proof:** (1) Let  $\alpha_i$  denote the non-negative value of  $g$  at the vertex  $i$ . We write  $g(x) = \sum_{i=1}^3 \alpha_i \lambda_i$  and note that  $\bar{g} = (1/3) \sum_{i=1}^3 \alpha_i$ . If the indices  $i, j, k$  are in cyclic order, we obtain

$$\begin{aligned} g(x) &= \sum_{i=1}^3 \alpha_i \lambda_i \geq \sum_{i=1}^3 \alpha_i \lambda_i (4\lambda_j \lambda_k) \\ &= 4 \sum_{i=1}^3 \alpha_i b_T = 12\bar{g} b_T. \end{aligned}$$

(2) The estimate (6.4) follows by standard scaling arguments. The last equation is obtained by simple computation of the integral.  $\blacksquare$

Next we refer to the lower bounds of the error that result from the local Dirichlet problems on elements or edges and their neighborhood

$$\begin{aligned} \mathcal{E}_{D,T} &:= \sup_{\substack{v \in H_0^1(T) \\ v \geq \psi - u_h}} J(u_h) - J(u_h + v) \\ \mathcal{E}_{D,e} &:= \sup_{\substack{v \in H_0^1(\omega_e) \\ v \geq \psi - u_h}} J(u_h) - J(u_h + v) \end{aligned}$$

In particular, Lemma 6.1 yields

$$\sum_T \mathcal{E}_{D,T} + \sum_e \mathcal{E}_{D,e} \leq c \left\{ J(u_h) - J(u) \right\}.$$

**Theorem 6.3** *There exists a constant  $c$  such that the area portion of the estimator  $\eta_{PS}$  satisfies*

$$\eta_T \leq c \mathcal{E}_{D,T}.$$

**Proof:** Given  $T \subset \omega_i$ , let

$$v := cb_T \max\{h_T^2 f_{T,i}, \overline{\psi - u_h}\} \in H_0^1(T),$$

where  $c$  is the constant in Lemma 6.2. By definition,  $v \geq cb_T \overline{\psi - u_h} \geq 12b_T \overline{\psi - u_h} \geq \psi - u_h$ . Hence,  $u_h + v \geq \psi$ .

Since the support of  $v$  is contained in the element  $T$ , it follows from (3.1) that

$$\begin{aligned} \mathcal{E}_{D,T} &\geq J(u_h) - J(u_h + v) \\ &= -\frac{1}{2} \|\nabla v\|_0^2 - \langle \lambda_h, v \rangle \\ &= -\frac{1}{2} \|\nabla v\|_0^2 - (-f_T, v). \end{aligned}$$

We distinguish two cases.

*Case 1.*  $h_T^2 f_T < \overline{\psi - u_h}$ .

Then  $v = cb_T \overline{\psi - u_h}$  and  $f_T$  is negative.

$$\begin{aligned} \mathcal{E}_{D,T} &\geq -\frac{1}{2}c^2 \|\nabla b_T\|^2 |\overline{\psi - u_h}|^2 \\ &\quad - \int b_T c f_T \overline{\psi - u_h} \geq -\frac{1}{2}c \int b_T |\overline{\psi - u_h}|^2 \\ &\quad - c \int b_T f_T \overline{\psi - u_h} \geq \frac{c}{120} |T| |f_T| \overline{u_h - \psi}. \end{aligned}$$

*Case 2.*  $h_T^2 f_T \geq \overline{\psi - u_h}$ .

Then  $v = cb_T h_T^2 f_T$  and

$$\begin{aligned} \mathcal{E}_{D,T} &\geq -\frac{1}{2}c^2 \|\nabla b_T\|^2 |h_T^2 \lambda_T|^2 \\ &\quad + \int b_T c h_T^2 f_T^2 \geq \frac{1}{2}c \int b_T |h_T^2 f_T|^2 \\ &= \frac{c}{120} h_T^2 \|f_T\|_{0,T}^2. \end{aligned}$$

In both cases, the local improvement  $\mathcal{E}_{D,T}$  dominates a multiple of the residual error estimator  $\eta_T$ .  $\blacksquare$

**Theorem 6.4** *There exists a constant  $c$  such that the edge portion of the estimator  $\eta_{PS}$  satisfies*

$$\eta_e \leq c \chi_e \mathcal{E}_{D,e}.$$

with the efficiency measure  $\chi_e$  defined as

$$\chi_e = 1 \quad \text{unless} \quad \left[ \frac{\partial u_h}{\partial n} \right] < 0 \quad \text{and} \quad h_T^2 f_T < -\overline{u_h - \psi} \quad (6.5)$$

and

$$\chi_e = \max \left\{ 1, \max_{T \subset \omega_e} \min \left\{ \frac{h f_T}{\left[ \frac{\partial u_h}{\partial n} \right]}, \frac{h^2 |f_T|}{\overline{u_h - \psi}} \right\} \right\}. \quad (6.6)$$

The proof is postponed to Appendix C.

The theorems show that the element terms of the estimator due to Prager and Synge are efficient, but that we have a weaker result for the ingredients of the edges. We may summarize the results as

$$\begin{aligned} J(u_h) - J(u) &\geq c_1 \|\sigma^\Delta\|_0^2 \\ &\quad - c_2 \sum_T h_T^2 \|f - \bar{f}\|_0^2 \\ &\quad - c_3 \sum'_{e,T} h_e \left[ \frac{\partial u_h}{\partial n} \right]_e \int_T f_T, \end{aligned}$$

where the prime at the last sum indicates that it runs over those pairs with  $e \subset \bar{T}$ ,  $[\frac{\partial u_h}{\partial n}] < 0$ , and  $h_T^2 f_T < -\overline{u_h - \psi}$ .

An example in Appendix A shows the loss of efficiency with exactly the factor (6.6), since the extra terms are much larger than the true error. The discussion of the example also elucidates that there is an inherent handicap with obstacle problems. Fortunately, this is no drawback in actual computations, if local refinements take care of extra terms. These extra terms can be data oscillations on patches [2] or associated with the error in the Lagrange multipliers [14]. In order to achieve an error reduction, the refinement has to be organized in such a way that not only the estimator but also the extra terms are reduced within the adaptive cycle; see, e.g., [5] and [6].

On the other hand, we demonstrate in Appendix B that the estimator due to Prager–Synge deals very well with other phenomena of non-affine obstacles if one-sided jumps are admitted with the equilibrated fluxes.

## A A Counterexample

The handicap of a posteriori error estimates for obstacle problems and their efficiency is elucidated by a one-dimensional example. The obstacle will be even affine linear.

Let  $b \gg d > 0$ , and consider the variational problem in  $H^1(-1, +1)$ :

$$\frac{1}{2} \int_{-1}^{+1} v'(x)^2 dx + b \int_{-1}^0 v(x) dx \longrightarrow \min! \quad (\text{A.1})$$

with the constraint  $\psi = 0$  and the boundary conditions  $v(-1) = 0$ ,  $v(1) = d$ . A boundary point  $-z$  of the contact zone is given by  $z + \frac{1}{2}z^2 = d/b$ , i.e.,  $z \approx d/b$ . The solution of the variational problem is

$$u(x) = \begin{cases} 0, & -1 \leq x \leq -z, \\ \frac{1}{2}b(z+x)^2, & -z \leq x \leq 0, \\ \frac{1}{2}bz^2 + bzx, & 0 \leq x \leq 1; \end{cases}$$

see Figure 2. The corresponding finite element solution with one node at 0 is

$$u_h(x) = \begin{cases} 0, & -1 \leq x \leq 0, \\ x, & 0 \leq x \leq 1. \end{cases}$$

A straight-forward calculation yields

$$\begin{aligned} & |u - u_h|_1^2 \\ &= \frac{1}{2} \int_{-z}^0 b^2(x+z)^2 dx + (d-bz)^2 \int_0^1 dx \\ &= \frac{1}{3}b^2z^3 + \left(\frac{1}{2}bz^2\right)^2 \leq d^3/b. \end{aligned}$$

Since the jump of  $u'_h$  equals  $d$ , the error bound is  $\approx d^2$ . Hence, the quotient of the error estimate and the true error is  $\approx b/d \approx f_T/\lambda_e$ , and the formula (6.6) for the efficiency measure is sharp (modulo a constant). Note that the same edge term as in the estimate due

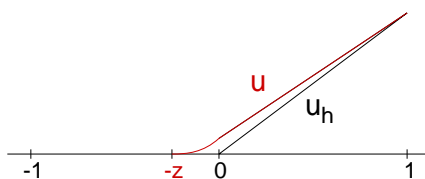


Figure 2: Exact and finite element solution of the problem in Appendix A.

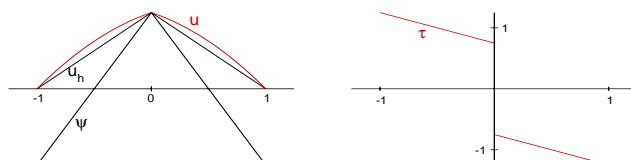


Figure 3: Exact and finite element solution of the problem in Appendix B and the equilibrated flux.

to Prager–Synge is encountered in the typical classical estimators [2, 4, 5, 9, 14]. It is no drawback in actual computations; cf. Section 6.

On purpose, we have chosen an example with an affine obstacle. We get a similar example with zero data oscillation if non-affine obstacles are chosen. If we extend the load in (A.1) to the complete domain and set  $\psi(x) = x$  on  $[0, 1]$ , then the finite element solution is the same. A symmetry argument shows that the exact solution changes so little, that the efficiency problem is the same. Obviously the kink of the obstacle implies the deterioration here.

## B Effects of edge terms with inequalities

Another one-dimensional example shows that the hypercircle method can cope with non-affine obstacles better than some well-known estimators. We gain appropriate flexibility by admitting equilibrated fluxes  $\tau \notin H(\text{div})$  as stated in Remark 2.1. This is positive in contrast to the example in the preceding appendix, but the situation is different, since the jump there has the opposite sign. – The discussion of the example may be of interest independently of the hypercircle method. Therefore, some arguments of Section 2 are repeated.

Let  $0 < b \ll 1$ , and consider the variational problem in  $H_0^1(-1, +1)$ :

$$\frac{1}{2} \int_{-1}^{+1} v'(x)^2 dx - 2b \int_{-1}^{+1} v(x) dx \longrightarrow \min!$$

with the constraint

$$\psi(x) = 1 - 2|x|$$

and homogeneous Dirichlet boundary conditions.

The solution is obviously  $u(x) = 1 - |x| + b|x|(1 - |x|)$ ; see Figure 3. The finite element solution with linear elements and one node at the midpoint of the interval is

$$u_h(x) = 1 - |x|.$$

Let  $\tau$  be a piecewise polynomial with a possible jump at  $x = 0$ . We have

$$\|u'_h - u'\|_0^2 + \|u' - \tau\|_0^2 \leq \|u'_h - \tau\|_0^2 \quad (\text{B.1})$$

if  $(u' - \tau, u'_h - u')_0 \geq 0$ . We start as in the proof of (2.7), but proceed in the spirit of Remark 2.1. In this example, we have  $f = 2b$  and

$$\begin{aligned} & (u' - \tau, u'_h - u')_0 \quad (\text{B.2}) \\ &= (u', u'_h - u')_0 - (f, u_h - u)_0 \\ & - \int_{-1}^1 \tau(u'_h - u') dx + (f, u_h - u)_0 \\ &= \langle \lambda, u_h - u \rangle + \int_{-1}^1 (\tau + 2b)(u_h - u) dx \\ & - [\tau(0+) - \tau(0-)](u - u_h)(0). \end{aligned}$$

From the characterization of the exact solution we know that the first term in (B.2) is nonnegative. The second term vanishes, if we have pointwise  $\tau' = -b$ . Since  $x = 0$  belongs to the active point set, we have  $(u - u_h)(0) \geq 0$ , and the last term is nonnegative whenever the jump of  $\tau$  is nonpositive. Therefore, the appropriate equilibration leads to

$$\tau(x) = \begin{cases} \rho - bx, & x < 0, \\ -\rho - bx, & x \geq 0 \end{cases}$$

with an arbitrary  $\rho \geq 0$ . The resulting estimator  $\|u'_h - \tau\|_0$  attains its minimum for  $\rho = 1 - b/2$ . Here  $\|u'_h - \tau\|_0 = b/\sqrt{6}$  and the estimator equals  $\|u'_h - u'\|_0$ . This proves the efficiency.

Error estimators which contain jump terms of  $\nabla u_h$  or of  $\nabla(u_h - \psi)$  cannot be efficient for small values of the parameter  $b$ .

## C Proof of Theorem 6.4

**Lemma C.1** (1) Let  $g$  be a linear function that is non-negative on a triangle  $T$  and  $\bar{g}^e$  be its mean-value on the edge  $e \subset \bar{T}$ . Then

$$2\bar{g}^e b_e(x) \leq g(x) \quad \text{for all } x \in T. \quad (\text{C.1})$$

(2) There is a constant  $c \approx 1$  ( $c \leq 2$ ) such that

$$c \leq \frac{\int b_e}{\|\nabla b_e\|^2 h}. \quad (\text{C.2})$$

(3) Let  $0 < z < 1$  and

$$b_e^{(z)} = \max\{0, (\lambda_1 - \lambda_3/z)(\lambda_2 - \lambda_3/z)\} \quad (\text{C.3})$$

be a bubble function whose support is reduced to a strip of width  $2zh$ . Then

$$\begin{aligned} \|b_e^{(z)}\|_{0,T}^2 &\approx h^2 z, \\ \|\nabla b_e^{(z)}\|_{0,T}^2 &\approx z^{-1}, \\ \|b_e^{(z)}\|_{0,e}^2 &\approx h. \end{aligned}$$

**Proof:** (1) Let  $\alpha_i$  denote the non-negative value of  $g$  at the vertex  $i$ . We write  $g(x) = \sum_{i=1}^3 \alpha_i \lambda_i$  and note that  $\bar{g}^e = (1/2) \sum_{i=1}^2 \alpha_i$ . We obtain

$$\begin{aligned} g(x) &= \sum_{i=1}^3 \alpha_i \lambda_i \geq \sum_{i=1}^2 \alpha_i \lambda_i \\ &\geq \sum_{i=1}^2 \alpha_i \lambda_1 \lambda_2 = 2\bar{g}^e b_e. \end{aligned}$$

(2) The estimate (C.2) follows by standard scaling arguments. Moreover, we have  $\int_e b_e = (1/6)|e|$ .

(3) The estimates follow by standard scaling arguments.  $\blacksquare$

*Proof of the theorem.*

We recall  $\lambda_e = \left[ \frac{\partial u_h}{\partial n} \right]$ . Given  $e \subset \omega_i$ , let

$$v := cb_e \max\{-h_e \lambda_e, \overline{\psi - u_h}^e\} \in H_0^1(\omega_e),$$

where  $c$  is the constant in Lemma C.1. By definition,  $v \geq cb_e \overline{\psi - u_h}^e \geq 2b_e \overline{\psi - u_h}^e \geq \psi - u_h$ . Hence,  $u_h + v \geq \psi$ .

From (3.1) we obtain for  $v \in H_0^1(\omega_e)$ :

$$\begin{aligned} \mathcal{E}_{D,e} &\geq J(u_h) - J(u_h + v) \\ &= -\frac{1}{2} \|\nabla v\|_0^2 - \langle \lambda_h, v \rangle \\ &= -\frac{1}{2} \|\nabla v\|_0^2 - (-f, v)_{\omega_e} - (\lambda_e, v)_e. \end{aligned} \tag{C.4}$$

We distinguish two cases.

*Case 1.*  $h_e \lambda_e \geq \overline{u_h - \psi}^e$ . Then  $\lambda_e$  is positive and  $v = -cb_e \overline{u_h - \psi}^e$  is negative. We have

$$\eta_e = (\lambda_e, \overline{u_h - \psi}^e)_{0,e} \geq |\overline{u_h - \psi}^e|^2.$$

Now (C.4) yields

$$\begin{aligned} \mathcal{E}_{D,e} &\geq -\frac{1}{2} c^2 \|\nabla b_e\|^2 |\overline{\psi - u_h}^e|^2 \\ &\quad + \int_{\omega_e} f v - \int_e b_e c \lambda_e \overline{\psi - u_h}^e \\ &\geq -\frac{1}{2} c \int_e b_e h^{-1} |\overline{\psi - u_h}^e|^2 + \int_{\omega_e} f v \\ &\quad + c \int_e b_e \lambda_e \overline{\psi - u_h}^e \\ &\geq \frac{c}{12} |e| |\lambda_e| \overline{u_h - \psi}^e + \int_{\omega_e} f v \\ &\geq c \eta_e + \sum_{T \subset \omega_e} \int_T f_T v. \end{aligned} \tag{C.5}$$

Case 2.  $h_e \lambda_e < \overline{u_h - \psi^e}$ . In this case we have

$$\eta_e = h^2 |\lambda_e|^2$$

and consider a test function with the modified bubble function

$$v = -\alpha b_e^{(z)} h_e \lambda_e.$$

The parameters  $\alpha > 0$  and  $z < 1$  will be fixed later. Since  $b_e^{(z)} \leq b_e$ , we get  $u_h + v \geq \psi$  as above.

By the inverse inequalities in Lemma C.1 and (C.4) we get

$$\begin{aligned} \mathcal{E}_{D,e} &\geq -\frac{1}{2} \|\nabla v\|_0^2 - \langle \lambda_h, v \rangle \\ &= -\frac{1}{2} \alpha^2 h^2 \lambda_e^2 \|\nabla b_e^{(z)}\|_0^2 \\ &\quad + \alpha h \lambda_e^2 \int_e b_e^{(z)} + \int_{\omega_e} f v \\ &\geq -c_1 \frac{\alpha^2 h^2}{z} \lambda_e^2 \\ &\quad + c_2 \alpha h^2 \lambda_e^2 + \int_{\omega_e} f v. \end{aligned}$$

Now we choose  $\alpha = \frac{c_2 z}{2c_1}$  to absorb the first term by the second one and obtain

$$\mathcal{E}_{D,e} \geq c_3 z \eta_e + \sum_{T \subset \omega_e} \int_T f_T v. \quad (\text{C.6})$$

The interaction of the edge bubbles with the element bubbles is given by the last terms in (C.5) and (C.6). The terms will be absorbed by the observation

$$\mathcal{E}_{D,T} \leq \mathcal{E}_{D,e} \quad \text{if } e \subset \bar{T} \quad (\text{C.7})$$

with the exception specified in Theorem 6.2.

By definition, the test function  $v$  has the opposite sign as  $\lambda_e$ . Therefore, we can drop the term if  $\lambda_e f_T < 0$ . Otherwise we distinguish three cases. In all of them  $\eta_T = h_T^2 \|f_T\|_{0,T}^2$ , and we set  $z = 1$ .

Case a)  $h_e \lambda_e \geq \overline{u_h - \psi^e}$  and  $f_T > 0$ .

A standard scaling argument yields

$\|h^{-1} b_e\|_{0,T} \leq c$ . Moreover,  $\eta_e = h \lambda_e \overline{u_h - \psi^e} \leq (\overline{u_h - \psi^e})^2$ , and

$$\begin{aligned} - \int_T f_T v &\leq \|h f_T\|_{0,T} \|h^{-1} v\|_{0,T} \\ &= \|h f_T\|_{0,T} \|h^{-1} b_e\|_{0,T} c |\overline{\psi - u_h^e}| \\ &\leq c \eta_T^{1/2} \eta_e^{1/2}. \end{aligned}$$

Case b)  $0 \leq h_e \lambda_e < \overline{u_h - \psi^e}$  and  $f_T > 0$ .

A similar scaling argument and  $\eta_e = h^2 \lambda_e^2$  yields

$$\begin{aligned} - \int_T f_T v &\leq \|h f_T\|_{0,T} \|h^{-1} v\|_{0,T} \\ &= \|h f_T\|_{0,T} \|h^{-1} b_e^{(z)}\|_{0,T} \alpha |h \lambda_e| \\ &\leq \alpha \eta_T^{1/2} \eta_e^{1/2}. \end{aligned}$$

Case c)  $\lambda_e < 0$  and  $-\overline{u_h - \psi} \leq h_T^2 f_T < 0$ .

Here  $\eta_e = h^2 \lambda_e^2$  and  $-\int_T f_T v$  can be bounded as in Case b).

In any of the three cases, by Young's inequality it follows that

$$\begin{aligned} \mathcal{E}_{D,e} &\geq c \eta_e - c \sum_{T \in \omega_e} \eta_T^{1/2} \eta_e^{1/2} \\ &\geq c \eta_e - c \sum_{T \in \omega_e} \eta_T. \end{aligned} \tag{C.8}$$

Finally, combining Theorem 6.3 and (C.7) we absorb the last term in (C.5) and (C.6) to obtain

$$\eta_e \leq c \mathcal{E}_{D,e}.$$

As a precaution we recall that the generic constant  $c$  can attain different values at different places.

In the case that was excluded, i.e.,

$$\lambda_e < 0 \quad \text{and} \quad h_T^2 f_T < -\overline{u_h - \psi}$$

we obtain only a weaker bound of  $h_T^2 f_T$  in terms of  $\eta_T$ ,

$$h^2 |f_T|^2 \leq \frac{h^2 |f_T|}{u_h - \psi} \eta_T.$$

The factor on the right-hand side leads to the last term in the efficiency measure (6.6). It guarantees efficiency if we are far away from the obstacle.

We alternatively restart with (C.6) and insert  $v$  and choose  $\alpha = \frac{c_2 z}{2c_1}$ :

$$\begin{aligned} \mathcal{E}_{D,e} &\geq c_3 z \eta_e - \alpha h |\lambda_e| \sum_{T \in \omega_e} |(f_T, b_e^{(z)})_{0,T}| \\ &\geq c_3 z \eta_e - c_4 z^2 \eta_e^{1/2} \sum_{T \in \omega_e} h^2 |f_T|. \end{aligned}$$

We proceed with a Young inequality to get

$$\mathcal{E}_{D,e} \geq c_5 z \eta_e - c_6 z^3 h^4 |f_T|^2$$

and set  $z = \frac{h \lambda_e}{2h^2 |f_T|} \sqrt{c_5/c_6}$  to absorb the second term by the first one to get the bound

$$\mathcal{E}_{D,e} \geq c \frac{\lambda_e}{h f_T} \eta_e = c \frac{\eta_e^{1/2}}{h^2 f_T} \eta_e.$$

This estimate is advantageous if the estimator is large compared to the load. ■



## References

- [1] M. Ainsworth and T.J. Oden, *A Posteriori Error Estimation in Finite Element Analysis*. Wiley, Chichester 2000.
- [2] S. Bartels and C. Carstensen. *Averaging techniques yield reliable a posteriori finite element error control for obstacle problems*. Numer. Math. **99**, 225–249 (2004).
- [3] D. Braess, *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*. 3rd edition. Cambridge University Press 2007.
- [4] D. Braess. *A posteriori error estimators for obstacle problems – another look*. Numer. Math. **101**, 415–421 (2005).
- [5] D. Braess, C. Carstensen and R.H.W. Hoppe, *Convergence analysis of a conforming adaptive finite element method for an obstacle problem*. Numer. Math. **107**, 455–471 (2007).
- [6] D. Braess, C. Carstensen and R.H.W. Hoppe, *Error reduction in adaptive finite element approximation of elliptic obstacle problems*. (in preparation).
- [7] D. Braess and J. Schöberl, *Equilibrated residual error estimator for Maxwell’s equations*. Math. Comp. (to appear).
- [8] R. Luce and B. Wohlmuth, *A local a posteriori error estimator based on equilibrated fluxes*. SIAM J. Numer. Anal. **42**, 1394–1414 (2004).
- [9] P. Morin, R.H. Nochetto and K.G. Siebert. *Data oscillation and convergence of adaptive FEM*. SIAM J. Numer. Anal. **38**, 466–488 (2000).
- [10] P. Neittaanmäki and S. Repin, *Reliable Methods for Computer Simulation. Error Control and A Posteriori Estimates*. Elsevier. Amsterdam 2004.
- [11] W. Prager and J.L. Synge, *Approximations in elasticity based on the concept of function spaces*. Quart. Appl. Math. **5**, 241–269 (1947).
- [12] S. I. Repin. *Estimates of deviations from exact solutions of elliptic variational inequalities*. J. Math. Sciences **115**, 2811–2819 (2003).
- [13] K. Siebert and A. Veese, *A unilaterally constrained quadratic minimization with adaptive finite elements*. SIAM J. Optimization **18**, 260–289 (2007).
- [14] A. Veese. *Efficient and reliable a posteriori error estimators for elliptic obstacle problems*. SIAM J. Numer. Anal. **39**, 146–167 (2001).
- [15] A. Weiss and B. Wohlmuth, *A posteriori error estimator and error control for contact problems*. IANS report 12/2007, University of Stuttgart.

Institute of Mathematics, Ruhr-University of Bochum, D-44801 Bochum, Germany,  
Department of Mathematics, University of Houston, Houston, TX 77204-3008, U.S.A. and  
Institute of Mathematics, University of Augsburg, D-86159 Augsburg, Germany  
Dept. of Math. and Center for Comput. Engrg. Science, RWTH Aachen, D-52074 Aachen,  
Germany