DE GRUYTER Comput. Methods Appl. Math. 2015; 15 (3):259–277

Research Article

Carsten Carstensen and Jun Hu

An Optimal Adaptive Finite Element Method for an Obstacle Problem

Abstract: This paper provides a refined a posteriori error control for the obstacle problem with an affine obstacle which allows for a proof of optimal complexity of an adaptive algorithm. This is the first adaptive mesh-refining finite element method known to be of optimal complexity for some variational inequality. The result holds for first-order conforming finite element methods in any spacial dimension based on shape-regular triangulation into simplices for an affine obstacle. The key contribution is the *discrete reliability* of the a posteriori error estimator from [6] in an edge-oriented modification which circumvents the difficulties caused by the non-existence of a positive second-order approximation [18].

DOI: 10.1515/cmam-2015-0017
Received May 1, 2015; revised May 25, 2015; accepted May 27, 2015

Dedicated to Ralf Kornhuber on the occasion of his 60th birthday

1 Introduction

The obstacle throughout this paper is an affine function χ on the bounded, polygonal or polyhedral domain Ω in $\mathbb{R}^n$ for $n = 2, 3$ with $\chi \leq 0$ on the boundary $\partial\Omega$. This allows to define the non-empty, closed, convex set

$$K := \{v \in V \mid \chi \leq v \text{ a.e. in } \Omega\}$$

in the Hilbert space $V := H_0^1(\Omega)$ (in standard notation for Sobolev spaces) endowed with the energy scalar product

$$a(v, w) := \int_\Omega \nabla v \cdot \nabla w \, dx \quad \text{for all } v, w \in V$$

and its induced norm $|||\cdot||| := a(\cdot,\cdot)^{1/2}$. Given any $f \in L^2(\Omega)$ and the L^2 scale product

$$(f, \cdot)_{L^2(\Omega)} =: F \in V^* \equiv H^{-1}(\Omega)$$

in the dual of V, the weak form of the obstacle problem allows for a unique solution $u \in K$ of [14]

$$F(v - u) \leq a(u, v - u) \quad \text{for all } v \in K. \tag{1.1}$$

The lowest-order conforming finite element approximation replaces K by the set $K(\mathcal{T}_\ell) := K \cap P_1(\mathcal{T}_\ell)$ of some piecewise affine functions with respect to some shape-regular, simplicial triangulation $\mathcal{T}_\ell$ of Ω and its P_1 finite element space $V_\ell = V(\mathcal{T}_\ell) := P_1(\mathcal{T}_\ell) \cap V$. The unique discrete solution $u_\ell \in K(\mathcal{T}_\ell)$ on the level $\ell \in \mathbb{N}_0$ solves [11]

$$F(v_\ell - u_\ell) \leq a(u_\ell, v_\ell - u_\ell) \quad \text{for all } v_\ell \in K_\ell. \tag{1.2}$$

The a priori and a posteriori error controls of the error $u - u_\ell$ in the energy norm have a long history [1, 3, 5, 6, 8, 12, 13, 16, 17, 23]. The adaptive mesh-refining algorithm successively refines the triangulations within the

Carsten Carstensen: Institut für Mathematik, Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany; and Department of Computational Science and Engineering, Yonsei University, 120-749 Seoul, Korea, e-mail: cc@math.hu-berlin.de
Jun Hu: LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, P. R. China, e-mail: hujun@math.pku.edu.cn

steps Solve, Estimate, Mark, Refine. We refer to the seminal contributions [7, 9, 15, 21] for the corresponding variational equality and the first convergence result [6] for variational inequalities. The optimality analysis of this paper has to overcome severe difficulties related to the fact that the Scott–Zhang quasi-interpolation operator is *not* positive and even worse, any positive approximation operator is *not* of second order [18]. The remedy is a refined a posteriori error analysis which involves contributions even from the Lagrange multipliers from the discrete compatibility conditions which are usually just estimated by their upper bound zero. Based on this refined analysis, any quasi-interpolation operator can be employed as long as it allows for the local first-order approximation property and the local projection property (see below for details on those notions).

The resulting *discrete reliability* states for two discrete solutions u_ℓ and $u_{\ell+m}$ based on two admissible triangulations $\mathcal{T}_\ell$ and its refinement $\mathcal{T}_{\ell+m}$ that for their respective sets of sides $\mathcal{E}_\ell$ and $\mathcal{E}_{\ell+m}$, there exists some set $\mathcal{M}_{\ell,\ell+m} \subset \mathcal{E}_\ell$ of sides with cardinality $|\mathcal{M}_{\ell,\ell+m}|$ controlled by the number $|\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}|$ of refined simplices such that

$$E(u_\ell) - E(u_{\ell+m}) + |||u_{\ell+m} - u_\ell|||^2 \le C_{\mathrm{dRel}} \eta_\ell^2(\mathcal{M}_{\ell,\ell+m}). \tag{1.3}$$

The energy difference in (1.3) is defined by

$$E(v) := 1/2\, a(v, v) - F(v) \quad \text{for all } v \in V. \tag{1.4}$$

It is well known that the solution $u \in K$ of (1.1) minimizes E in K (see [14]) and u_ℓ minimizes E in K_ℓ; since $K(\mathcal{T}_\ell) \subset K(\mathcal{T}_{\ell+m})$, the difference $E(u_\ell) - E(u_{\ell+m})$ is non-negative. Notice that (1.3) implies the reliability of the refined estimator from [3, 6] in the limit (1.3) as $m \to \infty$ (with assumed red-refinements).

A side-oriented adaptive finite element method [19] is based on Dörfler marking [9] with the local contribution like

$$\eta_\ell^2(E) := |\omega_E^{(\ell)}|^{1/n} \|[\nabla u_\ell]_E \cdot \nu_E\|_{L^2(E)}^2 + \mathrm{osc}^2(f, \omega_E^{(\ell)}) \quad \text{for } E \in \mathcal{E}_\ell(\Omega) \tag{1.5}$$

for the jump $[\nabla u_\ell]_E$ of the piecewise constant gradients ∇u_ℓ of the discrete solution u_ℓ across the interior side E with normal unit vector ν_E and patch $\omega_E^{(\ell)}$ of volume $|\omega_E^{(\ell)}|$, and the oscillation $\mathrm{osc}(f, \omega_E^{(\ell)})$ defined in Section 3.1 below.

The side $E \in \mathcal{E}_\ell(\partial\Omega)$ is called a *full contact boundary side*, written $E \in \mathrm{FCBS}(\mathcal{E}_\ell, \partial\Omega, \chi)$, if $\chi|_E \equiv 0$. (Details on the notation of $\mathcal{E}_\ell$ and $\mathcal{E}_\ell(\partial\Omega)$ follow in Section 3 below.) Some vaguely related concept of *full contact node* is introduced in [10] for the discrete solution. The oscillation for some boundary side reads

$$\mathrm{Osc}(f, \omega_E^{(\ell)}) := \begin{cases} \mathrm{osc}(f, \omega_E^{(\ell)}) & \text{for any } E \in \mathrm{FCBS}(\mathcal{E}_\ell, \partial\Omega, \chi), \\ |\omega_E^{(\ell)}|^{1/n} \|f\|_{L^2(\omega_E^{(\ell)})} & \text{for any } E \in \mathcal{E}_\ell(\partial\Omega) \setminus \mathrm{FCBS}(\mathcal{E}_\ell, \partial\Omega, \chi), \end{cases}$$

and its square equals the error estimator contribution

$$\eta_\ell^2(E) := \mathrm{Osc}^2(f, \omega_E^{(\ell)}) \quad \text{for all } E \in \mathcal{E}_\ell(\partial\Omega).$$

All those terms form the side-oriented error estimator

$$\eta_\ell^2 := \sum_{E \in \mathcal{E}_\ell} \eta_\ell^2(E) \tag{1.6}$$

which is a refined version of that in [3, 6] and so improves those of [8, 23, 25]. Given $0 < s < \infty$, the non-linear approximation class $\mathcal{A}_s$ for the energy plus data approximation consists of pairs $(u, f) \in K \times L^2(\Omega)$ with

$$|(u, f)|_{\mathcal{A}_s}^2 := \sup_{N \in \mathbb{N}} N^{2s} \inf_{\mathcal{T} \in \mathbb{T}(\mathcal{T}_0, N)} \min_{v_\mathcal{T} \in K(\mathcal{T})} (E(v_\mathcal{T}) - E(u) + \mathrm{Osc}_\mathcal{T}^2) < \infty. \tag{1.7}$$

Here and throughout this paper, $\mathbb{T}(\mathcal{T}_0, N)$ denotes the set of admissible triangulations $\mathcal{T}$ of an initial regular triangulation $\mathcal{T}_0$ with $|\mathcal{T}| \le |\mathcal{T}_0| + N$ by newest-vertex bisections [4, 21] associated to the data oscillation $\mathrm{Osc}_\mathcal{T}^2$ as the counterpart of Osc_ℓ^2 (see (3.1) below for the definition of the oscillation).

The adaptive algorithm of Section 2.1 is quasi-optimal in the sense that (1.7) implies

$$E(u_\ell) - E(u) + \mathrm{Osc}_\ell^2 \lesssim |(u, f)|_{\mathcal{A}_s}^2 (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-2s} \quad \text{for all } \ell = 0, 1, 2, \dots. \tag{1.8}$$

The remaining parts of this paper are organized as follows. Section 2 presents the adaptive algorithm based on the residual-type a posteriori error estimator and a bulk criterion for the side-oriented contributions for refinement. Section 3 proves the discrete reliability (1.3). Section 4 presents the optimality analysis.

Note that optimality of the adaptive algorithm in practice is illustrated in the numerical examples in [19] where a slightly different a posteriori error control was used. That paper proves some contraction property but excludes the optimality analysis.

Throughout this paper, standard notation on Lebesgue and Sobolev spaces is adopted and an inequality $A \leq C\,B$ with some mesh-size independent generic constant $0 \leq C < \infty$ is abbreviated as $A \lesssim B$ while $A \approx B$ reads $A \lesssim B \lesssim A$. All hidden generic factors depend on $\mathcal{T}_0$ and hence do neither depend on data nor on mesh-sizes nor levels nor number of simplices nor sides etc.

2 Algorithm and Main Results

This section is devoted to the adaptive algorithm and its notation as well as the statements of the main results.

2.1 Adaptive Algorithm

The adaptive finite element method consists of successive loops of a cycle involving the steps 'Solve', 'Estimate', 'Mark', and 'Refine' as in [19] as some realization of [6].

Input. Bulk parameter $0 < \theta < 1$ plus a regular triangulation $\mathcal{T}_0$ of the bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ into simplices for $n = 2, 3$ plus refinement edges to allow for admissible refinements by the newest-vertex bisection as in [4, 22].

Loop. For all levels $\ell = 0, 1, 2, \dots$ (until termination) do

Solve. Given triangulation $\mathcal{T}_\ell$ with set of sides $\mathcal{E}_\ell$ with the subset of internal sides $\mathcal{E}_\ell(\Omega)$ and $K_\ell := K(\mathcal{T}_\ell)$ for the piecewise affine $P_1(\mathcal{T}_\ell)$, compute solution u_ℓ of discrete problem (1.2).

Estimate. Compute side contributions (1.5) and error estimator (1.6).

Mark. Compute some set $\mathcal{M}_\ell$ of sides in $\mathcal{E}_\ell$ of (almost) minimal cardinality $|\mathcal{M}_\ell|$ such that

$$\theta \eta_\ell^2 \leq \eta_\ell^2(\mathcal{M}_\ell) := \sum_{E \in \mathcal{M}_\ell} \eta_\ell^2(E). \tag{2.1}$$

Refine. Bisect all marked sides at least once and add further refinements in some newest vertex bisection to generate the admissible refinement $\mathcal{T}_{\ell+1}$ (see [4, 22]). end do

Output. Sequences of finite element approximations (u_ℓ), the nested conforming sets (K_ℓ) and error estimators (η_ℓ) based on triangulations $(\mathcal{T}_\ell)$.

2.2 Main Results

The main results of the paper are concerned with the following discrete reliability of the estimator and the quasi-optimal convergence of the aforementioned adaptive algorithm.

Theorem 2.1. *The respective solutions u_ℓ and $u_{\ell+m}$ to the discrete problem* (1.2) *with respect to the triangulation $\mathcal{T}_\ell$ and its refinement $\mathcal{T}_{\ell+m}$ satisfy* (1.3) *for some subset $\mathcal{M}_{\ell,\ell+m}$ of $\mathcal{E}_\ell$ with $|\mathcal{M}_{\ell,\ell+m}| \lesssim |\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}|$.*

The full details of the proof of Theorem 2.1 will be given in Section 3. Define the energy difference by

$$\delta_\ell := E(u_\ell) - E(u).$$

The following Theorem 2.2 implies the quasi-optimality (1.8). The proof will be given in Section 4.

Theorem 2.2. *Suppose $(u, f) \in \mathcal{A}_s$ for some $s > 0$ and $\theta < c_{\mathrm{Eff}}/(C_{\mathrm{dRel}} + 1)$. Then the output $(\mathcal{T}_\ell, V_\ell, u_\ell)_{\ell\in\mathbb{N}}$ of the adaptive algorithm of Section 2.1 satisfies*

$$\delta_\ell + \gamma\eta_\ell^2 \lesssim |(u, f)|^2_{\mathcal{A}_s}(|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-2s} \quad \text{for all } \ell = 1, 2, \dots .$$

3 Proof of Discrete Reliability

This section presents the proof of Theorem 2.1 for two discrete solutions $u_{\ell+m}$ and u_ℓ of (1.2) with respect to some admissible refinement $\mathcal{T}_{\ell+m}$ of the shape-regular triangulation $\mathcal{T}_\ell$. The proof is divided into several steps: Sections 3.1–3.7 give a general setting while Sections 3.8–3.11 concern three different cases for some neighborhood of totally or unrefined patches.

3.1 Notation on Triangulations

Given the regular triangulation $\mathcal{T}_\ell$ of the domain into closed triangles/tetrahedrons, let $\mathcal{E}_\ell$ denote the set of all sides of $\mathcal{T}_\ell$, $\mathcal{E}_\ell(\Omega)$ the set of all interior sides, $\mathcal{E}_\ell(\partial\Omega)$ the set of all boundary sides, and $\mathcal{E}_\ell(T)$ the set of sides of a simplex $T \in \mathcal{T}_\ell$. Let $\mathcal{N}_\ell$ denote the set of all nodes in $\mathcal{T}_\ell$, $\mathcal{N}_\ell(\Omega)$ the set of all internal nodes, $\mathcal{N}_\ell(T)$ the set of nodes of $T \in \mathcal{T}_\ell$, and $\mathcal{N}_\ell(E)$ the set of nodes of $E \in \mathcal{E}_\ell$. For any $z \in \mathcal{N}_\ell$, let $\mathcal{E}_\ell(z)$ denote the set of sides in $\mathcal{E}_\ell$ and $\mathcal{T}_\ell(z)$ the set of simplices in $\mathcal{T}_\ell$ that share the vertex z. Let the patch $\omega_z^{(\ell)} := \bigcup_{T\in\mathcal{T}_\ell(z)} \mathrm{int}(T)$ and let $|||\cdot|||_{\omega_z^{(\ell)}}$ denote the restriction of the energy norm over the patch $\omega_z^{(\ell)}$. For interior side E, $[\cdot]_E := \cdot|_{T_+} - \cdot|_{T_-}$ denotes the jump across the side $E = T_+ \cap T_-$ shared by the two elements T_+ and T_-, and $\omega_E^{(\ell)} := \mathrm{int}(T_+ \cup T_-)$. If $E \in \mathcal{E}_\ell(\partial\Omega)$, then $\omega_E^{(\ell)} := \mathrm{int}(T_+)$ where T_+ is the unique element that has E as one side. For any node $z \in \mathcal{N}_\ell$, the nodal basis function $\varphi_z^{(\ell)}$ of the conforming P_1 finite element space V_ℓ satisfies $\varphi_z^{(\ell)}(z) = 1$ and $\varphi_z^{(\ell)}(y) = 0$ for any node $y \in \mathcal{N}_\ell$ other than z. Let h_ℓ denote the piecewise constant mesh-size in $\mathcal{T}_\ell$. Given $\omega \subset \Omega$, the oscillation of $f \in L^2(\Omega)$ over it is defined by

$$\mathrm{osc}(f, \omega) := |\omega|^{1/n}\|f - f_\omega\|_{L^2(\omega)} \quad \text{with the average} \quad f_\omega := \int_\omega f dx/|\omega|.$$

Given any internal side $E \in \mathcal{E}_\ell(\Omega)$, the associated oscillation is defined by

$$\mathrm{Osc}^2(f, \omega_E^{(\ell)}) := \mathrm{osc}^2(f, \omega_E^{(\ell)}).$$

For any set of sides $\mathcal{S}_\ell \subset \mathcal{E}_\ell$, define the oscillation

$$\mathrm{Osc}^2(\mathcal{S}_\ell) := \sum_{E\in\mathcal{S}_\ell} \mathrm{Osc}^2(f, \omega_E^{(\ell)}) \quad \text{and set} \quad \mathrm{Osc}_\ell^2 := \mathrm{Osc}^2(\mathcal{E}_\ell). \tag{3.1}$$

3.2 Overall Notation for the Proof

Throughout this section, let

$$\mathrm{LHS} := E(u_\ell) - E(u_{\ell+m}) + 1/2|||u_{\ell+m} - u_\ell|||^2$$

abbreviate a modified left-hand side of (1.3) and let $e := u_{\ell+m} - u_\ell \in V$ denote the difference with some residual

$$\varrho_z := F(\varphi_z^{(\ell)}) - a(u_\ell, \varphi_z^{(\ell)}) \le 0 \le u_\ell(z) - \chi(z) \quad \text{for any } z \in \mathcal{N}_\ell(\Omega).$$

These definitions imply the following discrete consistency conditions: either $\varrho_z = 0$ or $u_\ell(z) = \chi(z)$ for all $z \in \mathcal{N}_\ell(\Omega)$.

The set $\mathcal{N}_\ell(\Omega)$ of interior nodes is usually split into the contact nodes

$$\mathcal{C}_\ell := \{z \in \mathcal{N}_\ell(\Omega) : u_\ell(z) = \chi(z)\}$$

and its complement. Notice that $\varrho_z = 0$ for any interior node z outside $\mathcal{C}_\ell$.

3.3 Reformulation of LHS

Some elementary algebra reveals that the definition of the energy implies

$$\mathrm{LHS} = F(e) - a(u_\ell, e).$$

With the abbreviations $e_z := 0$ for any boundary node $z \in \mathcal{N}_\ell(\partial\Omega)$ and $e_z := e_\ell(z) := J_\ell(e)(z)$ for any interior node $z \in \mathcal{N}_\ell(\Omega)$, the quasi-interpolation $e_\ell \in V_\ell$ of $e \equiv u_{\ell+m} - u_\ell$ reads

$$e_\ell = \sum_{z \in \mathcal{N}_\ell} e_z \varphi_z^{(\ell)}.$$

It is emphasized that the particular choice of this quasi-interpolation J_ℓ is rather general. This seems to be in contradiction to the analysis in [8] or [3] where some particular design of the quasi-interpolation is seen as the reason for the success of their analysis.

The refined error analysis of this paper merely requires locality in the sense that, for any free node $z \in \mathcal{N}_\ell(\Omega)$, the value $J_\ell(w)(z) := J_{\ell,z}(w)$ exclusively depends on the values $w|_{\omega_z^{(\ell)}}$ of the test function w on the patch $\omega_z^{(\ell)}$ in a linear way plus *first-order approximation* properties in the sense that

$$\|w - J_{\ell,z}(w)\|_{L^2(\omega_z^{(\ell)})} \lesssim h_z \|\nabla w\|_{L^2(\omega_z^{(\ell)})} \quad \text{for all } w \in H^1(\omega_z^{(\ell)})$$

with the diameter $h_z := \operatorname{diam}(\omega_z^{(\ell)})$ of the patch and *exactness for discrete functions* in the sense that

$$J_{\ell,z}(w_\ell) = w_\ell(z) \quad \text{for all } w_\ell \in P_1(\mathcal{T}_\ell(z)) \cap C(\omega_z^{(\ell)}).$$

In particular, the Scott–Zhang quasi-interpolation operator [20] enjoys all those properties, but there is no need for any positivity like in [8] or extra orthogonality like in [3]. Throughout this paper, J_ℓ will be selected as the Scott–Zhang quasi-interpolation operator.

For the Scott–Zhang quasi-interpolation, the selection of some proper face allows for the immediate fulfilment of those boundary condition as it is nowadays standard in the proofs of discrete reliability [7, 21].

Since the nodal basis functions $\varphi_z^{(\ell)}$ form a partition of unity, it follows

$$\mathrm{LHS} = \sum_{z \in \mathcal{N}_\ell} F((e - e_z)\varphi_z^{(\ell)}) + \sum_{z \in \mathcal{N}_\ell(\Omega)} e_z \varrho_z - a(u_\ell, e - e_\ell).$$

3.4 Side Contributions

The first-order approximation property of the quasi-interpolation plus some standard piecewise integration by parts imply

$$-a(u_\ell, e - e_\ell) \lesssim \eta_\ell(\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+m}) |||e|||.$$

Since this analysis is verbatim as in the case of variational equalities as in [1, 2, 24], we omit further details.

3.5 Contact on Entire Patch of Interior Node

This subsection concerns the control of the contributions $F((e - e_z)\varphi_z^{(\ell)}) + e_z \varrho_z$ in case $z \in \mathcal{C}_\ell$ with $u_\ell \equiv \chi$ on $\omega_z^{(\ell)}$. Since $u_\ell \equiv \chi \in P_1(\omega_z^{(\ell)})$, we have

$$a(u_\ell, \varphi_z^{(\ell)}) = a(\chi, \varphi_z^{(\ell)}) = 0.$$

This and the discrete compatibility conditions imply

$$F((e-e_z)\varphi_z^{(\ell)}) + e_z\varrho_z = F(e\varphi_z^{(\ell)}) \quad \text{and} \quad \varrho_z = F(\varphi_z^{(\ell)}) \le 0.$$

Since $0 \le u_{\ell+m} - \chi = u_{\ell+m} - u_\ell = e$ on $\omega_z^{(\ell)}$, it follows

$$\overline{e} := \frac{\int_{\omega_z^{(\ell)}} \varphi_z^{(\ell)} e\,dx}{\int_{\omega_z^{(\ell)}} \varphi_z^{(\ell)}\,dx} \ge 0.$$

The average $\overline{f} \in P_0(\omega_z^{(\ell)})$ of f satisfies that

$$F(e\varphi_z^{(\ell)}) \le F((e-\overline{e})\varphi_z^{(\ell)}) = \int\limits_{\omega_z^{(\ell)}} (f-\overline{f})(e-\overline{e})\varphi_z^{(\ell)}\,dx.$$

The Poincaré inequality leads to

$$F((e-e_z)\varphi_z^{(\ell)}) + e_z\varrho_z = F(e\varphi_z^{(\ell)}) \lesssim \operatorname{osc}(f, \omega_z^{(\ell)}) |||e|||_{\omega_z^{(\ell)}}.$$

Remark 3.1. Some refined a posteriori error estimate may exclude the oscillations of the set of the *discrete full contact edges*

$$\mathrm{DFCE}_\ell := \{E \in \mathcal{E}_\ell,\ u_\ell|_{\omega_z^{(\ell)}} \equiv \chi \text{ and } f|_{\omega_z^{(\ell)}} \le 0,\ z \in \mathcal{N}_\ell(E)\}$$

and so defines the reliable a posteriori error estimator

$$\eta_\ell^2 := \sum_{E\in\mathcal{E}_\ell} |\omega_E^{(\ell)}|^{1/n} \|[\nabla u_\ell]_E \cdot \nu_E\|_{L^2(E)}^2 + \sum_{E\in\mathcal{E}_\ell\setminus\mathrm{DFCE}_\ell} \operatorname{Osc}^2(f, \omega_E^{(\ell)}). \tag{3.2}$$

However, since possibly $\mathrm{DFCE}_\ell \not\supseteq \mathrm{DFCE}_{\ell+1}$, we experienced difficulty in the proof of the contraction property of a related adaptive mesh-refining algorithm based on refinement indication via (3.2).

In a second scenario assume that ω_z^* is some part of $\omega_z^{(\ell)}$ with $|\omega_z^{(\ell)}| \lesssim |\omega_z^*|$ and $e_z \ge 0$. Then it holds that

$$(f, \varphi_z^{(\ell)}(e-e_z))_{L^2(\omega_z^*)} + e_z\varrho_z \lesssim \operatorname{osc}(f, \omega_z^{(\ell)}) |||e|||_{\omega_z^{(\ell)}}. \tag{3.3}$$

Recall that (3.3) holds for the first scenario where $\omega_z^* := \omega_z^{(\ell)}$ without any sign condition of e_z. The proof of (3.3) in the second scenario follows from the aforementioned arguments and adapts with $\omega_z' := \omega_z^{(\ell)} \setminus \omega_z^*$ and the averages

$$e_z^* := \frac{\int_{\omega_z^*} \varphi_z^{(\ell)} e\,dx}{\int_{\omega_z^*} \varphi_z^{(\ell)}\,dx} \quad \text{and} \quad f_z^* := \frac{\int_{\omega_z^*} f\,dx}{|\omega_z^*|}.$$

The remaining part of this subsection shall present some details of the proof of (3.3). Direct calculations prove

$$(f, \varphi_z^{(\ell)}(e-e_z))_{L^2(\omega_z^*)} + e_z\varrho_z = (f - f_z^*, \varphi_z^{(\ell)}(e-e_z^*))_{L^2(\omega_z^*)} + e_z^*(f, \varphi_z^{(\ell)})_{L^2(\omega_z^*)} + e_z(f, \varphi_z^{(\ell)})_{L^2(\omega_z')}.$$

The last two terms on the right-hand side of the above equation allow for the two following identities:

$$\begin{aligned} M &:= e_z^*(f, \varphi_z^{(\ell)})_{L^2(\omega_z^*)} + e_z(f, \varphi_z^{(\ell)})_{L^2(\omega_z')} \\ &= (e_z^* - e_z)(f, \varphi_z^{(\ell)})_{L^2(\omega_z^*)} + e_z\varrho_z \\ &= (e_z - e_z^*)(f, \varphi_z^{(\ell)})_{L^2(\omega_z')} + e_z^*\varrho_z. \end{aligned}$$

Since $\varrho_z \le 0 \le e_z, e_z^*$, we have

$$M \le \min\{(e_z^* - e_z)(f, \varphi_z^{(\ell)})_{L^2(\omega_z^*)}, (e_z - e_z^*)(f, \varphi_z^{(\ell)})_{L^2(\omega_z')}\}.$$

Since $\min\{(e_z^* - e_z)\bar{f}, (e_z - e_z^*)\bar{f}\} \le 0$, we have

$$M \le |e_z - e_z^*| \|f - \bar{f}\|_{L^2(\omega_z^{(\ell)})} \|\varphi_z^{(\ell)}\|_{L^2(\omega_z^{(\ell)})}.$$

Some elementary analysis and the very definition of e_z^* show that

$$
\begin{aligned}
|\omega_z^*||e_z - e_z^*|^2 &\approx 1/2 \int_{\omega_z^*} \varphi_z^{(\ell)} (e_z - e_z^*)^2\,dx \\
&\le \int_{\omega_z^*} \varphi_z^{(\ell)} (e_z - e)^2\,dx + \int_{\omega_z^*} \varphi_z^{(\ell)} (e - e_z^*)^2\,dx \\
&\le \int_{\omega_z^{(\ell)}} \varphi_z^{(\ell)} (e_z - e)^2\,dx + \int_{\omega_z^*} \varphi_z^{(\ell)} (e - \bar{e})^2\,dx \\
&\le \|J_{\ell,z}(e) - e\|^2_{L^2(\omega_z^{(\ell)})} + \|e - \bar{e}\|^2_{L^2(\omega_z^{(\ell)})}.
\end{aligned}
$$

The first order approximation property of the local quasi-interpolation $J_{\ell,z}$ plus the Poincaré inequality prove the previous upper bounds $\lesssim h_z^2 |||e|||^2_{\omega_z^{(\ell)}}$. This and the aforementioned bound of M conclude the proof of (3.3).

3.6 Boundary Contribution I: $\mathcal{N}_\ell(\partial\Omega) \cap \{\chi < 0\}$

The Friedrichs inequality for patches on the boundary $\partial\Omega$ allows for the proof of

$$
F((e - e_z)\varphi_z^{(\ell)}) = F(e\varphi_z^{(\ell)}) \lesssim \|h_\ell f\|_{L^2(\omega_z^{(\ell)})} |||e|||_{\omega_z^{(\ell)}}
$$

for each $z \in \mathcal{N}_\ell(\partial\Omega) \cap \{\chi < 0\}$ with patch $\omega_z^{(\ell)}$ in the triangulation $\mathcal{T}_\ell$. Given $T \in \mathcal{T}_\ell(z)$, some elementary analysis proves

$$
\|h_\ell f\|_{L^2(\omega_z^{(\ell)})} \approx \|h_\ell f\|_{L^2(T)} + \operatorname{osc}(f, \omega_z^{(\ell)}). \tag{3.4}
$$

In fact, the triangle inequality gives for the integral mean $\overline{f}$ of f over the patch $\omega_z^{(\ell)}$

$$
\begin{aligned}
\|h_\ell f\|_{L^2(\omega_z^{(\ell)})} &\lesssim \|h_\ell (f - \overline{f})\|_{L^2(\omega_z^{(\ell)})} + \|h_\ell \overline{f}\|_{L^2(T)} \\
&\lesssim \|h_\ell (f - \overline{f})\|_{L^2(\omega_z^{(\ell)})} + \|h_\ell (f - \overline{f})\|_{L^2(T)} + \|h_\ell f\|_{L^2(T)} \\
&\lesssim \|h_\ell (f - \overline{f})\|_{L^2(\omega_z^{(\ell)})} + \|h_\ell f\|_{L^2(T)}.
\end{aligned}
$$

The estimate of (3.4) can be employed to reduce $\|h_\ell f\|_{L^2(\omega_z^{(\ell)})}$ to some simplex at the boundary plus some patch oscillations. The combination of the aforementioned estimates shows

$$
\sum_{z \in \mathcal{N}_\ell(\partial\Omega) \cap \{\chi < 0\}} F((e - e_z)\varphi_z^{(\ell)}) \lesssim \operatorname{Osc}_\ell |||e|||.
$$

3.7 Volume Contributions and Their Oscillations for Interior Nodes

Since the other case is already discussed in Section 3.5, it remains the case that $u_\ell \not\equiv \chi$ on $\omega_z^{(\ell)}$ for $z \in \mathcal{N}_\ell(\Omega)$. This case is discussed below where the volume contribution $\|h_\ell f\|_{\omega_z^{(\ell)}}$ arises which is indeed controlled by oscillations in the sense that

$$
\|h_\ell f\|^2_{L^2(\omega_z^{(\ell)})} \lesssim \sum_{y \in \mathcal{N}_\ell(\omega_z^{(\ell)})} (\operatorname{Osc}^2(f, \omega_y^{(\ell)}) + \eta_\ell^2(\mathcal{E}_\ell(y))). \tag{3.5}
$$

Notice that $y \in \mathcal{N}_\ell(\omega_z^{(\ell)})$ means that y is some neighboring node of z or equal to z. Here and throughout, $\operatorname{Osc}^2(f, \omega_y^{(\ell)})$ equals the oscillation $\operatorname{osc}^2(f, \omega_y^{(\ell)})$ over the patch $\omega_y^{(\ell)}$ for an interior node y while it equals $\|h_\ell f\|^2_{L^2(\omega_y^{(\ell)})}$ for a boundary node $y \in \mathcal{N}_\ell(\partial\Omega) \cap \{\chi < 0\}$.

The remaining part of this subsection is devoted to the proof of (3.5) in two cases.

In *the first case* that $z \in \mathcal{N}_\ell(\Omega) \setminus \mathcal{C}_\ell$, recall that $\bar{f}$ denotes the integral mean of the right-hand side f over $\omega_z^{(\ell)}$ of diameter h_z. Then

$$\|h_\ell f\|^2_{L^2(\omega_z^{(\ell)})}/h_z^2 \approx \|f\|^2_{L^2(\omega_z^{(\ell)})} = \|f-\bar{f}\|^2_{L^2(\omega_z^{(\ell)})} + \bar{f}^2|\omega_z^{(\ell)}|.$$

Since

$$\bar{f}|\omega_z^{(\ell)}| \approx \int_{\omega_z^{(\ell)}} \bar{f}\varphi_z^{(\ell)}\,dx = F(\varphi_z^{(\ell)}) - \int_{\omega_z^{(\ell)}} (f-\bar{f})\varphi_z^{(\ell)}\,dx,$$

this implies

$$\|h_\ell f\|_{L^2(\omega_z^{(\ell)})} \lesssim \operatorname{osc}(f,\omega_z^{(\ell)}) + |\bar{f}|h_z|\omega_z^{(\ell)}|^{1/2} \lesssim \operatorname{osc}(f,\omega_z^{(\ell)}) + h_z^{1-n/2}\,|F(\varphi_z^{(\ell)})|.$$

Since $z \in \mathcal{N}_\ell(\Omega) \setminus \mathcal{C}_\ell$, we have $\varrho_z = 0$ and some piecewise integration by parts shows

$$F(\varphi_z^{(\ell)}) = a(u_\ell, \varphi_z^{(\ell)}) = \sum_{E\in\mathcal{E}_\ell(z)} \int_E \varphi_z^{(\ell)}[\partial u_\ell/\partial\nu_E]_E\,ds.$$

The sidewise Cauchy inequality (with $|E| \approx h_z^{n-1}$) and the Cauchy inequality in $\mathbb{R}^m$ for $m = |\mathcal{E}_\ell(z)| \approx 1$ lead to

$$h_z^{2-n}\,|F(\varphi_z^{(\ell)})|^2 \le \sum_{E\in\mathcal{E}_\ell(z)} \eta_\ell^2(E) =: \eta_\ell^2(\mathcal{E}_\ell(z)).$$

The combination of the aforementioned estimates leads to (3.5).

In *the second case* $z \in \mathcal{C}_\ell$, the present hypothesis $u_\ell \not\equiv \chi$ on $\omega_z^{(\ell)}$ guarantees the existence of some node $y \in \mathcal{N}_\ell$, which is a neighbor of z in the sense that the convex hull $\operatorname{conv}\{z, y\}$ is some edge in the triangulation $\mathcal{T}_\ell$ with $\chi(y) < u_\ell(y)$.

In case that $y \in \mathcal{N}_\ell(\partial\Omega)$ belongs to the boundary, the very definition of $\{\chi = 0\}$ plus the homogeneous boundary condition of u_ℓ imply $y \in \{\chi < 0\}$. Then, the arguments of Section 3.6 show

$$\|h_\ell f\|^2_{L^2(\omega_z^{(\ell)})} \lesssim \|h_\ell f\|^2_{L^2(\omega_y^{(\ell)})} + \operatorname{osc}^2(f,\omega_z^{(\ell)}).$$

This implies (3.5). In the remaining case, the interior node $y \in \mathcal{N}_\ell(\Omega) \setminus \mathcal{C}_\ell$ is analyzed as z in the first case. This leads to

$$\|h_\ell f\|^2_{L^2(\omega_y^{(\ell)})} \lesssim \operatorname{osc}^2(f,\omega_y^{(\ell)}) + \eta_\ell^2(\mathcal{E}_\ell(y)).$$

The argument of Section 3.6 on some element domain $T \in \mathcal{T}_\ell(z) \cap \mathcal{T}_\ell(y)$ can be employed to verify

$$\|h_\ell f\|^2_{L^2(\omega_z^{(\ell)})} \lesssim \|h_\ell f\|^2_{L^2(T)} + \operatorname{osc}^2(f,\omega_z^{(\ell)}) + \operatorname{osc}^2(f,\omega_y^{(\ell)}).$$

The combination of the previous two estimates implies (3.5) in the final case as well and concludes the proof.

3.8 Boundary Contribution II: $\mathcal{N}_\ell(\partial\Omega) \cap \{\chi = 0\}$

In the *first case* assume the existence of some interior node $y \in \mathcal{N}_\ell(\Omega)$, which is a neighbor of $z \in \mathcal{N}_\ell(\partial\Omega) \cap \{\chi = 0\}$ in the sense that the convex hull $\operatorname{conv}\{z, y\}$ is some interior edge in the triangulation $\mathcal{T}_\ell$ with $\chi(y) < u_\ell(y)$. A similar argument to that of Section 3.7 proves

$$\begin{aligned} F((e-e_z)\varphi_z^{(\ell)}) &= F(e\varphi_z^{(\ell)}) \lesssim \|h_\ell f\|_{L^2(\omega_z^{(\ell)})} |||e|||_{\omega_z^{(\ell)}} \\ &\lesssim \Big(\operatorname{osc}(f,\omega_y^{(\ell)}) + \operatorname{osc}(f,\omega_z^{(\ell)}) + \sum_{E\in\mathcal{E}_\ell(y)} |\omega_E^{(\ell)}|^{1/2n}\|[\nabla u_\ell]_E\cdot\nu_E\|_{L^2(E)}\Big)|||e|||_{\omega_z^{(\ell)}}. \end{aligned}$$

A *second case* will assume that $u_\ell \equiv \chi$ on $\omega_z^{(\ell)}$ and $\bar{f} := \int_{\omega_z^{(\ell)}} f\,dx/|\omega_z^{(\ell)}| \le 0$. Hence $e \ge 0$ on $\omega_z^{(\ell)}$, which leads to

$$F((e-e_z)\varphi_z^{(\ell)}) = \int_{\omega_z^{(\ell)}} ef\varphi_z^{(\ell)}\,dx \le \int_{\omega_z^{(\ell)}} e(f-\bar{f})\varphi_z^{(\ell)}\,dx \lesssim \operatorname{osc}(f,\omega_z^{(\ell)})|||e|||_{\omega_z^{(\ell)}}.$$

A *third case* will suppose again $u_\ell \equiv \chi$ on $\omega_z^{(\ell)}$ but $\bar{f} > 0$. Without loss of generality, we may assume that there exists an interior neighbor node y of z such that $u_\ell|_{\omega_y^{(\ell)}} \equiv \chi|_{\omega_y^{(\ell)}}$; otherwise, the *first case* applies to some further circle of interior neighboring nodes of z. This, in particular, leads to

$$F(\varphi_y^{(\ell)}) := \int_{\omega_y^{(\ell)}} f\varphi_y^{(\ell)}\,dx \leq 0.$$

The standard bubble-function technique [24] yields

$$\|\bar{f}\|^2_{L^2(\omega_z^{(\ell)})} \approx |\bar{f}|^2 \int_{\omega_y^{(\ell)}} \varphi_y^{(\ell)}\,dx \leq \bar{f} \int_{\omega_y^{(\ell)}} (\bar{f} - f)\varphi_y^{(\ell)}\,dx \lesssim \|\bar{f}\|_{L^2(\omega_z^{(\ell)})} \|f - \bar{f}\|_{L^2(\omega_y^{(\ell)})}.$$

Recall that $\bar{f} \equiv \int_{\omega_z^{(\ell)}} f dx / |\omega_z^{(\ell)}|$ and set $\widetilde{f} := \int_{\omega_y^{(\ell)}} f dx / |\omega_y^{(\ell)}|$ so that

$$\begin{aligned}\|\bar{f}\|_{L^2(\omega_z^{(\ell)})} \leq \|f - \bar{f}\|_{L^2(\omega_y^{(\ell)})} &\lesssim \|f - \widetilde{f}\|_{L^2(\omega_y^{(\ell)})} + |\widetilde{f} - \bar{f}|\,|\omega_y^{(\ell)} \cap \omega_z^{(\ell)}|^{1/2} \\ &\leq \|f - \bar{f}\|_{L^2(\omega_z^{(\ell)})} + \|f - \widetilde{f}\|_{L^2(\omega_y^{(\ell)})}.\end{aligned}$$

This implies

$$\begin{aligned}|\omega_z^{(\ell)}|^{2/n} \|f\|^2_{L^2(\omega_z^{(\ell)})} &= |\omega_z^{(\ell)}|^{2/n} \|f - \bar{f}\|^2_{L^2(\omega_z^{(\ell)})} + |\omega_z^{(\ell)}|^{2/n} \|\bar{f}\|^2_{L^2(\omega_z^{(\ell)})} \\ &\lesssim \operatorname{osc}^2(f, \omega_z^{(\ell)}) + \operatorname{osc}^2(f, \omega_y^{(\ell)}).\end{aligned}$$

Together with the preceding results, this leads to

$$F((e - e_z)\varphi_z^{(\ell)}) \lesssim (\operatorname{osc}(f, \omega_z^{(\ell)}) + \operatorname{osc}(f, \omega_y^{(\ell)})) |||e|||_{\omega_z^{(\ell)}}.$$

The *remaining case* is $\chi \not\equiv 0$ on $\partial\omega_z^{(\ell)} \cap \partial\Omega$ and there does not exist an interior neighboring node $y \in \mathcal{N}_\ell(\Omega)$ with $u_\ell(y) > \chi(y)$. This case can be analyzed as in Section 3.6, which will yield some volume terms associated to some boundary side $E \in \mathcal{E}_\ell(\partial\Omega)$ with $\chi|_E \not\equiv 0$.

3.9 Three Sets of Interior Nodes and Their Reduced Patches

The refined analysis of the discrete reliability is concerned with the decomposition

$$\mathcal{N}_\ell(\Omega) = \mathcal{U}_\ell \cup \mathcal{I}_\ell \cup \mathcal{R}_\ell$$

into the three pairwise disjoint (possibly empty) sets of interior nodes

$$\begin{aligned}\mathcal{U}_\ell &:= \{z \in \mathcal{N}_\ell(\Omega) : \mathcal{T}_\ell(z) = \mathcal{T}_{\ell+m}(z)\} && \text{(neighborhood unrefined)},\\ \mathcal{I}_\ell &:= \mathcal{N}_\ell(\Omega) \setminus (\mathcal{U}_\ell \cup \mathcal{R}_\ell) && \text{(intermediate refinement)},\\ \mathcal{R}_\ell &:= \{z \in \mathcal{N}_\ell(\Omega) : \mathcal{T}_\ell(z) \subset \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}\} && \text{(neighborhood totally refined)}.\end{aligned}$$

The boundary conditions on $\partial\Omega$ and on $\mathcal{T}_{\ell+m} \cap \mathcal{T}_\ell$ for the quasi-interpolation e_ℓ are arranged so that it holds $e_z := e_\ell(z) = e(z)$ for any node $z \in \mathcal{U}_\ell \cup \mathcal{I}_\ell$. In case $z \in \mathcal{C}_\ell$, we have $e_z = u_{\ell+m}(z) - \chi(z) \geq 0$ while $\varrho_z = 0$ otherwise. Altogether,

$$e_z \varrho_z \leq 0 \quad \text{for all } z \in \mathcal{U}_\ell \cup \mathcal{I}_\ell. \tag{3.6}$$

The quasi-interpolation error vanishes on any simplex that belongs to either refinement,

$$e = e_\ell \quad \text{on any } T \in \mathcal{T}_{\ell+m} \cap \mathcal{T}_\ell. \tag{3.7}$$

We therefore may neglect all volume contributions from the union

$$\Omega' := \bigcup_{z \in \mathcal{U}_\ell} \omega_z^{(\ell)}$$

of all patches of $\mathcal{U}_\ell$ (which includes $\mathcal{T}_{\ell+m} \cap \mathcal{T}_\ell$) and define the reduced patches

$$\omega_z^* := \omega_z^{(\ell)} \setminus \Omega' \quad \text{for all } z \in \mathcal{N}_\ell.$$

This reduced patch vanishes, $\omega_z^* = \emptyset$, for any node $z \in \mathcal{U}_\ell$ and is unchanged, $\omega_z^* = \omega_z^{(\ell)}$, for any $z \in \mathcal{R}_\ell$. In the intermediate case $z \in \mathcal{I}_\ell$, some simplices in $\omega_z^{(\ell)}$ are refined and some are unrefined. Hence ω_z^* contains at least one simplex so that the shape regularity implies

$$|\omega_z^*| \approx |\omega_z^{(\ell)}| \quad \text{for all } z \in \mathcal{I}_\ell \cup \mathcal{R}_\ell.$$

(Notice that $\omega_z^* = \omega_z^{(\ell)}$ is not excluded in case $z \in \mathcal{I}_\ell$. The definition of ω_z^* applies to boundary nodes as well.)

3.10 Auxiliary Result for Interior Patch in Contact

The estimate of $e_z \varrho_z$ for some interior node $z \in \mathcal{R}_\ell \cap \mathcal{C}_\ell$ in the subsequent subsection requires some elementary result which is stated for some interior patch $\omega_z^{(\ell)}$ and any non-negative continuous piecewise affine function $v_\ell \in P_1(\mathcal{T}_\ell(z)) \cap C(\omega_z^{(\ell)})$ with $v_\ell(z) = 0$ and $0 \le v_\ell$ on $\omega_z^{(\ell)}$ as

$$\min_{g \in P_1(\omega_z^{(\ell)})} \|v_\ell - g\|_{L^2(\omega_z^{(\ell)})} \approx \|v_\ell\|_{L^2(\omega_z^{(\ell)})}. \tag{3.8}$$

This equivalence is certainly known to the experts but key to the analysis of the entire proof. Since the analysis in the related [3, Lemma 7] utilizes an equivalence of norms in $\mathbb{R}^J$ but leaves the interaction with the non-negativity of the coefficients $a_1, \dots, a_J$ rather unclear, this section investigates the equivalence constants in some detail in an explicit hard analysis proof.

The left-hand side of (3.8) is clearly smaller than or equal to its right-hand side. The point is the converse estimate with the focus on the generic multiplicative positive constant $C \approx 1$ in

$$\|v_\ell\|_{L^2(\omega_z^{(\ell)})} \le C \min_{g \in P_1(\omega_z^{(\ell)})} \|v_\ell - g\|_{L^2(\omega_z^{(\ell)})}. \tag{3.9}$$

This constant C does not depend on v_ℓ but may depend on the shape of the patch in the following sense. Let $0 < \underline{r} \le \overline{r}$ denote the radii of the inclusion circle and the circumcircle in the sense that $\underline{r}$ (resp. $\overline{r}$) is maximal (resp. minimal) with

$$B(z, \underline{r}) \subset \omega_z^{(\ell)} \subset \overline{B(z, \overline{r})}.$$

The ratio $0 < \kappa := \underline{r}/\overline{r} \le 1$ exclusively depends on the interior angles of the triangulation and satisfies $\kappa \approx 1$.

The remaining part of this subsection is devoted to some elementary proof of (3.9) and the dependence of C on the shape regularity and on κ. The patch geometry is depicted in Figure 1 with vertices $P_1, \dots, P_J$ around the interior node z and edges $E_j := \operatorname{conv}\{z, P_j\}$ for $j = 1, \dots, J$. Recall that $v_\ell(z) = 0$ and $a_j := v_\ell(P_j) \ge 0$ for all $j = 1, \dots, J$. Some elementary calculations with mass matrices show that, for any $g \in P_1(\omega_z^{(\ell)})$, it suffices to prove

$$\frac{\kappa^2}{4J(J+1)} \sum_{j=1}^J a_j^2 \le M := \min_{g \in P_1(\omega_z^{(\ell)})} \Big(g(z)^2 + \sum_{j=1}^J (a_j - g(P_j))^2 \Big). \tag{3.10}$$

(The remaining contributions to C in (3.9) depend on the dimension n and the number J of simplices in $\omega_z^{(\ell)}$.) The optimal affine function g depends on $n+1$ real parameters $\alpha, \beta_1, \dots, \beta_n$, e.g.,

$$g(x) := \alpha + (\beta_1, \dots, \beta_n) \cdot (x - z) \quad \text{for all } x \in \omega_z^{(\ell)} \subset \mathbb{R}^n.$$

The partial derivatives of the left-hand side

$$\alpha^2 + \sum_{j=1}^J (a_j - g(P_j))^2 = \alpha^2 + \sum_{j=1}^J (a_j - \alpha - (\beta_1, \dots, \beta_n) \cdot (P_j - z))^2$$

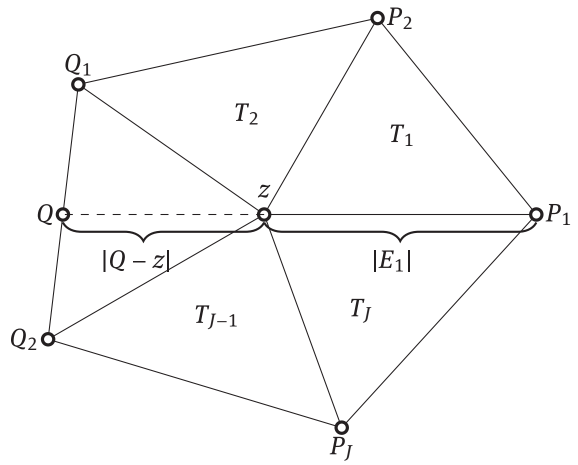

Figure 1. Geometry of a patch $\omega_z^{(\ell)} \subseteq \mathbb{R}^2$ for $k = 1$ in Section 3.10 to illustrate the convex coefficient μ in $z = \mu P_1 + (1-\mu)Q$ for $n = 2$.

with respect to $\alpha, \beta_1, \dots, \beta_n$ vanish for their optimal values. With (the Euclidian scalar product $\cdot$ and) the abbreviation

$$b_j := a_j - (\beta_1, \dots, \beta_n) \cdot (P_j - z) \quad \text{for } j = 1, \dots, J,$$

the first optimality condition reads $\alpha = (b_1 + \cdots + b_J)/(J+1)$ and

$$M = \alpha^2 + \sum_{j=1}^{J} (b_j - \alpha)^2 = (J+1)\alpha^2 - 2\alpha \sum_{j=1}^{J} b_j + \sum_{k=1}^{J} b_k^2 = \sum_{k=1}^{J} b_k^2 - \Big(\sum_{k=1}^{J} b_k \Big)^2 / (J+1).$$

The Cauchy inequality $(\sum_{k=1}^{J} b_k)^2 \le J \sum_{k=1}^{J} b_k^2$ proves that the previous expression is greater than or equal to $\sum_{k=1}^{J} b_k^2/(J+1)$. Thus

$$b_1^2 + \cdots + b_J^2 \le (J+1)M. \tag{3.11}$$

Recall that $a_1, \dots, a_J$ are non-negative and fix an index $k = 1, \dots, J$ with $a_k = \max\{a_1, \dots, a_J\}$. The typical geometry on the patch $\omega_z^{(\ell)}$ is depicted in Figure 1 for $k = 1$ in $n = 2$ dimensions. The node P_k lies inside the ball $B(z, \overline{r})$, whence the length $|E_k|$ of the edge $E_k = \operatorname{conv}\{z, P_k\}$ is $|E_k| \le \overline{r}$. The straight line S through P_k and z hits twice the boundary of the interior ball

$$B(z, \underline{r}) \subset \overline{\omega_z^{(\ell)}}$$

and twice the boundary $\partial\omega_z^{(\ell)}$ of the interior patch $\omega_z^{(\ell)}$ outside $B(z, \underline{r})$. The intersection point Q of S and $\partial\omega_z^{(\ell)}$ opposite E_k satisfies

$$Q \in \operatorname{conv}\{Q_1, \dots, Q_n\} \quad \text{and} \quad \underline{r} \le |z - Q| \le \overline{r}$$

for some vertices $Q_1, \dots, Q_n \in \{P_1, \dots, P_J\} \setminus \{P_k\}$ different from P_k. In other words, the $n+1$ pairwise different vertices $P_k, Q_1, \dots, Q_n$ of $\partial\omega_z^{(\ell)}$ satisfy

$$z = \mu P_k + (1-\mu) Q \quad \text{and} \quad \mu = |z - Q|/|Q - P_k| \ge \underline{r}/(\overline{r} + \underline{r}) = \kappa/(1+\kappa).$$

Some remaining non-negative coefficients $\lambda_1, \dots, \lambda_n$ satisfy

$$\mu P_k + \lambda_1 Q_1 + \cdots + \lambda_n Q_n = z \quad \text{and} \quad \mu + \lambda_1 + \cdots + \lambda_n = 1.$$

This implies that

$$0 = \mu(P_k - z) + \lambda_1(Q_1 - z) + \cdots + \lambda_n(Q_n - z).$$

The scalar product of this with $(\beta_1, \dots, \beta_n)$ leads to terms of the form

$$a_j - b_j = (\beta_1, \dots, \beta_n) \cdot (P_j - z) = (\beta_1, \dots, \beta_n) \cdot (Q_m - z)$$

with appropriate indices j and m for $Q_m =: P_j$. With the abbreviations $\tilde{a}_m := a_j$ and $\tilde{b}_m := b_j$ for $m = 1, \dots, n$ and $j \in \{1, \dots, J\} \setminus \{k\}$ with $Q_m = P_j$, this reads

$$\mu a_k + \lambda_1 \tilde{a}_1 + \cdots + \lambda_n \tilde{a}_n = \mu b_k + \lambda_1 \tilde{b}_1 + \cdots + \lambda_n \tilde{b}_n.$$

This, the non-negativity of $a_1, \dots, a_J$, plus the above bound $\kappa/2 \le \mu$ show

$$\begin{aligned}\kappa^2\, a_k^2/4 \le (\mu a_k)^2 &\le (\mu a_k + \lambda_1 \tilde{a}_1 + \cdots + \lambda_n \tilde{a}_n)^2 \\ &= (\mu b_k + \lambda_1 \tilde{b}_1 + \cdots + \lambda_n \tilde{b}_n)^2 \\ &\le \mu b_k^2 + \lambda_1 \tilde{b}_1^2 + \cdots + \lambda_n \tilde{b}_n^2.\end{aligned}$$

Since $(b_k, \tilde{b}_1, \dots, \tilde{b}_n)$ is some permutation of some sub-list of $(b_1, b_2, \dots, b_J)$ of length $n+1$ and with (3.11) in the end, we deduce

$$\kappa^2\, a_k^2/4 \le \mu b_k^2 + \lambda_1 \tilde{b}_1^2 + \cdots + \lambda_n \tilde{b}_n^2 \le b_1^2 + \cdots + b_J^2 \le (J+1)M.$$

Since a_k is maximal, we have

$$a_1^2 + \cdots + a_J^2 \le J a_k^2 \le 4J(J+1)/\kappa^2\, M.$$

This concludes the proof of (3.10).

3.11 Completely Refined Interior Patch in Contact

This subsection is devoted to an upper bound of $e_z \varrho_z$ for some interior node $z \in \mathcal{R}_\ell \cap \mathcal{C}_\ell$. Since the other case is already discussed in Section 3.5, it remains the case that

$$u_\ell \not\equiv \chi \text{ on } \omega_z^{(\ell)} \text{ while } u_\ell(z) = \chi(z). \tag{3.12}$$

Since the patch is completely refined, the quasi-interpolation e_ℓ of $e := u_{\ell+m} - u_\ell$ computes

$$e_z \equiv e_\ell(z) = J_{\ell,z}(e|_{\omega_z^{(\ell)}})$$

via some linear function $J_{\ell,z}$ with the first-order approximation property and the exactness for discrete functions. Since u_ℓ and χ belong to $P_1(\mathcal{T}_\ell(z)) \cap C(\omega_z^{(\ell)})$, the difference $w := u_{\ell+m} - \chi$ satisfies

$$e_z = J_{\ell,z}(u_{\ell+m}|_{\omega_z^{(\ell)}}) - u_\ell(z) = J_{\ell,z}(u_{\ell+m}|_{\omega_z^{(\ell)}}) - \chi(z) = J_{\ell,z}(w|_{\omega_z^{(\ell)}}).$$

Since $w \ge 0$, the constant average value $\overline{w} := \int_{\omega_z^{(\ell)}} w\, dx / |\omega_z^{(\ell)}| \ge 0$ satisfies

$$-e_z = J_{\ell,z}(\overline{w} - w) - \overline{w} \le J_{\ell,z}(\overline{w} - w).$$

The Poincaré inequality on the patch $\omega_z^{(\ell)}$ of diameter $h_z := \operatorname{diam}(\omega_z^{(\ell)})$ reads

$$\|w - \overline{w}\|_{L^2(\omega_z^{(\ell)})} \lesssim h_z\, \|\nabla w\|_{L^2(\omega_z^{(\ell)})}.$$

The combination with the local first-order approximation property implies that the constant $J_{\ell,z}(\overline{w} - w) = \overline{w} - J_{\ell,z}(w)$ satisfies

$$\begin{aligned}\|J_{\ell,z}(\overline{w} - w)\|_{L^2(\omega_z^{(\ell)})} &\le \|w - J_{\ell,z}(w)\|_{L^2(\omega_z^{(\ell)})} + \|w - \overline{w}\|_{L^2(\omega_z^{(\ell)})} \\ &\lesssim h_z\, \|\nabla w\|_{L^2(\omega_z^{(\ell)})}.\end{aligned}$$

In other words,

$$J_{\ell,z}(\overline{w} - w) \lesssim h_z^{1-n/2}\, \|\nabla w\|_{L^2(\omega_z^{(\ell)})}.$$

Since $\varrho_z \le 0$, the combination of the previous estimates shows

$$e_z \varrho_z \lesssim h_z^{1-n/2}\, |\varrho_z|\, \|\nabla w\|_{L^2(\omega_z^{(\ell)})}. \tag{3.13}$$

Since $u_\ell \geq \chi$ and χ is an affine function, the piecewise affine function $v_\ell := u_\ell - \chi \in P_1(\mathcal{T}_\ell(z)) \cap C(\omega_z^{(\ell)})$ is non-negative and vanishes at the interior node z. Hence (3.9) is applicable. This and some inverse estimate show

$$h_z \|\nabla v_\ell\|_{L^2(\omega_z^{(\ell)})} \leq \|v_\ell\|_{L^2(\omega_z^{(\ell)})} \lesssim \min_{g \in P_1(\omega_z^{(\ell)})} \|v_\ell - g\|_{L^2(\omega_z^{(\ell)})},$$

which, together with the Poincaré inequality, leads to

$$\|\nabla v_\ell\|_{L^2(\omega_z^{(\ell)})} \lesssim \min_{q \in \mathbb{R}^n} \|\nabla v_\ell - q\|_{L^2(\omega_z^{(\ell)})}.$$

Some equivalence of norms on the finite-dimensional vector space $P_1(\mathcal{T}_\ell(z)) \cap C(\omega_z^{(\ell)})$ plus a scaling argument establish the estimate

$$\min_{q \in \mathbb{R}^n} \|\nabla v_\ell - q\|^2_{L^2(\omega_z^{(\ell)})} \lesssim \sum_{E \in \mathcal{E}_\ell(z)} |\omega_E^{(\ell)}|^{1/n} \, \|[\partial v_\ell / \partial \nu_E]_E\|^2_{L^2(E)}.$$

Recall $v_\ell = u_\ell - \chi$ and hence

$$\sum_{E \in \mathcal{E}_\ell(z)} |\omega_E^{(\ell)}|^{1/n} \, \|[\partial v_\ell / \partial \nu_E]_E\|^2_{L^2(E)} = \eta_\ell^2(\mathcal{E}_\ell(z)).$$

Altogether,

$$\|\nabla v_\ell\|_{L^2(\omega_z^{(\ell)})} \lesssim \eta_\ell(\mathcal{E}_\ell(z)).$$

This and the triangle inequality,

$$\|\nabla w\|_{L^2(\omega_z^{(\ell)})} \leq \|\nabla e\|_{L^2(\omega_z^{(\ell)})} + \|\nabla(u_\ell - \chi)\|_{L^2(\omega_z^{(\ell)})},$$

lead in (3.13) to

$$h_z^{n/2-1} e_z \varrho_z \lesssim \|\nabla w\|_{L^2(\omega_z^{(\ell)})} |\varrho_z| \leq \big(\|\nabla e\|_{L^2(\omega_z^{(\ell)})} + \eta_\ell(\mathcal{E}_\ell(z))\big)|\varrho_z|.$$

The Cauchy inequality for the volume term $F(\varphi_z^{(\ell)})$ plus some integration by parts argument from Section 3.7 for the side contributions from $a(u_\ell, \varphi_z^{(\ell)})$ lead to

$$|\varrho_z| \leq |F(\varphi_z^{(\ell)})| + |a(u_\ell, \varphi_z^{(\ell)})| \lesssim \big(\|h_\ell f\|_{L^2(\omega_z^{(\ell)})} + \eta_\ell(\mathcal{E}_\ell(z))\big) h_z^{n/2-1}.$$

The volume control of (3.5) (recall (3.12)) shows

$$h_z^{2-n} |\varrho_z|^2 \lesssim \eta_\ell^2(\mathcal{E}_\ell(z)) + \sum_{y \in \mathcal{N}_\ell(\omega_z^{(\ell)})} \operatorname{Osc}^2(f, \omega_y^{(\ell)}).$$

A summary concludes that

$$e_z \varrho_z \lesssim \Big(\eta_\ell(\mathcal{E}_\ell(z)) + \sum_{y \in \mathcal{N}_\ell(\overline{\omega_z^{(\ell)}})} \operatorname{Osc}(f, \omega_y^{(\ell)})\Big)\big(|||e|||_{\omega_z^{(\ell)}} + \eta_\ell(\mathcal{E}_\ell(z))\big). \tag{3.14}$$

3.12 Reduction of LHS

The reduced patches, (3.6), (3.7), and the fact that $e_z \varrho_z \leq 0$ for $z \in \mathcal{U}_z$ result in the estimate

$$\text{LHS} \leq \sum_{z \in \mathcal{N}_\ell} \int_{\omega_z^*} f(e - e_z)\varphi_z^{(\ell)} \, dx + \sum_{z \in (\mathcal{R}_\ell \cup \mathcal{I}_\ell) \cap \mathcal{C}_\ell} e_z \varrho_z - a(u_\ell, e - e_\ell).$$

The arguments of Sections 3.6 and 3.8 apply to the reduced patches as well and yield

$$\sum_{z \in \mathcal{N}_\ell(\partial\Omega)} \int_{\omega_z^*} f(e - e_z)\varphi_z^{(\ell)} \, dx \lesssim \operatorname{Osc}_\ell(\mathcal{E}_\ell \backslash \mathcal{E}_{\ell+m}) \, |||e|||.$$

This, Section 3.4, and the empty reduced patches of $z \in \mathcal{U}_\ell$ show that

$$\text{LHS} \lesssim \eta_\ell(\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+m}) \,|\!|\!| e |\!|\!| + \sum_{z \in \mathcal{R}_\ell \cup \mathcal{I}_\ell} \int_{\omega_z^*} f(e - e_z)\varphi_z^{(\ell)}\,dx + \sum_{z \in (\mathcal{R}_\ell \cup \mathcal{I}_\ell) \cap \mathcal{C}_\ell} e_z \varrho_z.$$

The first case where $z \notin \mathcal{C}_\ell$ has already been analyzed in Section 3.7, which reads

$$\int_{\omega_z^*} f(e - e_z)\varphi_z^{(\ell)}\,dx \lesssim (\operatorname{osc}(f, \omega_z^{(\ell)}) + \eta_\ell(\mathcal{E}_\ell(z)))|\!|\!| e |\!|\!|_{\omega_z^{(\ell)}}.$$

The second case, where $z \in (\mathcal{R}_\ell \cup \mathcal{I}_\ell) \cap \mathcal{C}_\ell$ such that $u_\ell \equiv \chi$ on $\omega_z^{(\ell)}$ (note that $0 \le e_z$ for $z \in \mathcal{I}_\ell \cap \mathcal{C}_\ell$), has been considered in Section 3.5, namely,

$$\int_{\omega_z^*} f\varphi_z^{(\ell)}(e - e_z)\,dx + e_z\varrho_z \lesssim \operatorname{osc}(f, \omega_z^{(\ell)})|\!|\!| e |\!|\!|_{\omega_z^{(\ell)}}.$$

Since $\varrho_z \le 0 \le e_z$ for the case where $z \in \mathcal{I}_\ell \cap \mathcal{C}_\ell$ such that $u_\ell \not\equiv \chi$ on $\omega_z^{(\ell)}$, it follows from the estimate of Section 3.7 that

$$\int_{\omega_z^*} f\varphi_z^{(\ell)}(e - e_z)\,dx + e_z\varrho_z \le \int_{\omega_z^*} f\varphi_z^{(\ell)}(e - e_z)\,dx \lesssim \sum_{y \in \mathcal{N}_\ell(\omega_z^{(\ell)})} (\operatorname{Osc}(f, \omega_y^{(\ell)}) + \eta_\ell(\mathcal{E}_\ell(y)))|\!|\!| e |\!|\!|_{\omega_z^{(\ell)}}.$$

The last case, where $z \in \mathcal{R}_\ell \cap \mathcal{C}_\ell$ such that $u_\ell \not\equiv \chi$ on $\omega_z^{(\ell)}$, can be bounded by the combination of Section 3.7 and the estimate (3.14), which leads to

$$\int_{\omega_z^*} f\varphi_z^{(\ell)}(e - e_z)\,dx + e_z\varrho_z \lesssim \sum_{y \in \mathcal{N}_\ell(\omega_z^{(\ell)})} (\operatorname{Osc}(f, \omega_y^{(\ell)}) + \eta_\ell(\mathcal{E}_\ell(y)))\Big(|\!|\!| e |\!|\!|_{\omega_z^{(\ell)}} + \sum_{y \in \mathcal{N}_\ell(\omega_z^{(\ell)})} \eta_\ell(\mathcal{E}_\ell(y))\Big).$$

In conclusion, the set

$$\mathcal{M}_{\ell,\ell+m} := \{E \in \mathcal{E}_\ell(\Omega) : \exists F \in \mathcal{E}_\ell \,\exists G \in \mathcal{E}_\ell \setminus \mathcal{E}_{\ell+m} \text{ such that } E \cap F \ne \emptyset \ne F \cap G\} \tag{3.15}$$

contains the above sides as well as those from $\eta_\ell(\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+m})$ and satisfies

$$\text{LHS} \lesssim (\eta_\ell(\mathcal{M}_{\ell,\ell+m}) + \operatorname{Osc}_\ell(\mathcal{M}_{\ell,\ell+m}))(|\!|\!| e |\!|\!| + \eta_\ell(\mathcal{M}_{\ell,\ell+m})). \tag{3.16}$$

It is not hard to prove that the number $|\mathcal{M}_{\ell,\ell+m}|$ of sides in $\mathcal{M}_{\ell,\ell+m}$ is controlled by the number $|\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}|$ of refined simplices $\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}$ in the sense that

$$|\mathcal{M}_{\ell,\ell+m}| \lesssim |\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+m}|.$$

3.13 Finish of the Proof

The error term $|\!|\!| e |\!|\!|^2 \le \text{LHS}$ can be absorbed and (3.16) proves

$$\text{LHS} \lesssim \eta_\ell^2(\mathcal{M}_{\ell,\ell+m}) + \operatorname{Osc}_\ell^2(\mathcal{M}_{\ell,\ell+m}).$$

4 Optimal Convergence Rates

4.1 Efficiency of the Error Estimator

The efficiency of the error estimator will be stated in terms of the energy functional E from (1.4), and *not* in terms of the energy norm $|\!|\!| \cdot |\!|\!|$. In this way we circumvent the lack of Galerkin orthogonality and focus on the energy difference

$$0 \le \delta_\ell := E(u_\ell) - E(u).$$

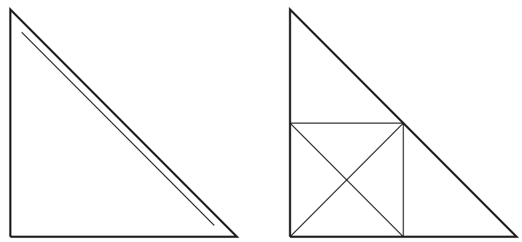

Figure 2. Reference triangle T (left) with reference edge marked, and bisec5(T) (right) in Section 4.1.

Lemma 4.1 (Efficiency). *There exists some $c_{\mathrm{Eff}} \approx 1$ with*

$$c_{\mathrm{Eff}}\eta_\ell^2 \le \delta_\ell + \mathrm{Osc}_\ell^2 .$$

Proof. We only show the result for the two-dimensional case based on the discrete efficiency from [6] and claim that similar arguments prove it for the three-dimensional case as well. The discrete solution $\hat{u}_\ell$ of (1.2) with respect to the refined mesh $\hat{\mathcal{T}}_\ell := \mathrm{bisec5}(\mathcal{T}_\ell)$ in Figure 2 satisfies [6, Theorem 2]

$$\begin{aligned}\sum_{E\in\mathcal{E}_\ell(\Omega)} |\omega_E^{(\ell)}|^{1/n}\|[\nabla u_\ell]_E\cdot\nu_E\|_{L^2(E)}^2 &\lesssim \sum_{E\in\mathcal{E}_\ell(\Omega)} \big(\|\nabla(u_\ell-\hat{u}_\ell)\|_{L^2(\omega_E^{(\ell)})}^2 + \mathrm{osc}^2(f,\omega_E^{(\ell)})\big)\\ &\lesssim |||u_\ell-\hat{u}_\ell|||^2 + \mathrm{Osc}_\ell^2 .\end{aligned}$$

Since $0 \le \hat{\delta}_\ell := E(\hat{u}_\ell) - E(u)$, it follows

$$1/2|||u_\ell-\hat{u}_\ell|||^2 \le E(u_\ell) - E(\hat{u}_\ell) = \delta_\ell - \hat{\delta}_\ell \le \delta_\ell .$$

The combination of the aforementioned inequalities completes the proof. □

4.2 Contraction Property

This subsection analyzes the convergence of the adaptive algorithm of Section 2.1.

Theorem 4.2. *There exist constants $\gamma > 0$ and $0 < q < 1$ such that*

$$\delta_{\ell+1} + \gamma\eta_{\ell+1}^2 \le q(\delta_\ell + \gamma\eta_\ell^2) \quad \textit{for all } \ell = 0, 1, 2, \dots . \tag{4.1}$$

Proof. The proof is based on two observations. This first observation reads

$$\begin{aligned}&\sum_{E\in\mathcal{E}_{\ell+1}(\Omega)} |\omega_E^{(\ell+1)}|^{1/n}\|[\nabla u_\ell]_E\cdot\nu_E\|_{L^2(E)}^2\\ &\quad\le \sum_{E\in\mathcal{E}_\ell(\Omega)\setminus\mathcal{M}_\ell} |\omega_E^{(\ell)}|^{1/n}\|[\nabla u_\ell]_E\cdot\nu_E\|_{L^2(E)}^2 + (1/2)^{1/n}\sum_{E\in\mathcal{E}_\ell(\Omega)\cap\mathcal{M}_\ell} |\omega_E^{(\ell)}|^{1/n}\|[\nabla u_\ell]_E\cdot\nu_E\|_{L^2(E)}^2 .\end{aligned}$$

The second observation is that

$$\mathrm{Osc}_{\ell+1}^2 \le \mathrm{Osc}_\ell^2 - \rho_1 \sum_{E\in\mathcal{E}_\ell\setminus\mathcal{E}_{\ell+1}} \mathrm{Osc}^2(f,\omega_E^{(\ell)}) \quad \text{for some } 0 < \rho_1 < 1.$$

These two estimates are proved in [19] as for the linear Poisson model problem [7]. Whence, with some constants $0 < \rho_2 < 1$ and $\Lambda > 0$, the triangle inequality plus the bulk criterion lead to

$$\eta_{\ell+1}^2 \le \rho_2\eta_\ell^2 + \Lambda|||u_{\ell+1} - u_\ell|||^2 .$$

Since

$$\delta_{\ell+1} \le \delta_\ell - 1/2|||u_{\ell+1} - u_\ell|||^2$$

and using the combination of the previous two inequalities plus $\gamma = \frac{1}{2\Lambda}$, the contraction property (4.1) follows from the reliability of the estimator η_ℓ with $0 < \rho_2 < q < 1$; see [19] for more details in two dimensions. □

This subsection concludes with an application of the aforementioned discrete reliability which indicates that the bulk criterion is in some sense a necessary condition for the reduction of the energy norm between two levels.

Lemma 4.3. *Let $\mathcal{T}_{\ell+m}$ be some refinement of $\mathcal{T}_\ell$ with*

$$\delta_{\ell+m} + \mathrm{Osc}^2_{\ell+m} \le q(\delta_\ell + \mathrm{Osc}^2_\ell) \tag{4.2}$$

for some $0 < q < 1$. *Then it holds*

$$c_{\mathrm{Eff}}(1-q)\eta^2_\ell \le (1 + C_{\mathrm{dRel}})\eta^2_\ell(\mathcal{M}_{\ell,\ell+m}). \tag{4.3}$$

Proof. The efficiency of Lemma 4.1 proves that

$$c_{\mathrm{Eff}}\eta^2_\ell \le \delta_\ell + \mathrm{Osc}^2_\ell\,.$$

The discrete reliability of Section 3 leads to

$$\delta_\ell - \delta_{\ell+1} \le C_{\mathrm{dRel}}\eta^2_\ell(\mathcal{M}_{\ell,\ell+m}).$$

Note, for any $E \in \mathcal{E}_\ell \cap \mathcal{E}_{\ell+m}$ with $\omega^{(\ell)}_E = \omega^{(\ell+m)}_E$, that

$$\mathrm{Osc}(f, \omega^{(\ell)}_E) = \mathrm{Osc}(f, \omega^{(\ell+m)}_E).$$

Since $\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+m} \subseteq \mathcal{M}_{\ell,\ell+m}$ (from the definition (3.15) of $\mathcal{M}_{\ell,\ell+m}$), we have

$$\mathrm{Osc}^2_\ell - \mathrm{Osc}^2_{\ell+m} \le \sum_{E\in\mathcal{M}_{\ell,\ell+m}} \mathrm{Osc}^2(f, \omega^{(\ell)}_E) \le \eta^2_\ell(\mathcal{M}_{\ell,\ell+m}).$$

This implies (4.3). □

4.3 Optimality

With the discrete admissible set $K(\mathcal{T}) := K \cap P_1(\mathcal{T})$ with respect to the triangulation $\mathcal{T}$ and the solution u of (1.1), recall the definition (1.7) of the seminorm, and define

$$\mathbb{E}(\mathcal{T}_0, N; u, f) := \inf_{\mathcal{T}\in\mathbb{T}(\mathcal{T}_0,N)} \min_{v_\mathcal{T}\in K(\mathcal{T})} (E(v_\mathcal{T}) - E(u) + \mathrm{Osc}^2_\mathcal{T}).$$

Lemma 4.4. *Suppose that* $(u, f) \in K \times L^2(\Omega)$ *satisfies* $|(u, f)|_{\mathcal{A}_s} < \infty$ *for some* $0 < s < \infty$ *and suppose that* $0 < \theta < c_{\mathrm{Eff}}/(C_{\mathrm{dRel}} + 1)$ *from Lemma 4.3. Then,*

$$|\mathcal{M}_\ell|^2 \lesssim |(u, f)|^{2/s}_{\mathcal{A}_s}(\delta_\ell + \mathrm{Osc}^2_\ell)^{-1/s} \quad \textit{for all } \ell = 0, 1, 2, \dots. \tag{4.4}$$

Proof. Given any $\ell \in \mathbb{N}$, set

$$q := \min\left\{1 - \frac{\theta(C_{\mathrm{dRel}} + 1)}{c_{\mathrm{Eff}}}, \frac{\mathbb{E}(\mathcal{T}_0, N_0 + 1; u, f)}{2(\delta_\ell + \mathrm{Osc}^2_\ell)}\right\} < 1.$$

To prove $q^{-1/s} \lesssim 1$, notice that

$$1 \approx \frac{\mathbb{E}(\mathcal{T}_0, N_0 + 1; u, f)}{\delta_0 + \mathrm{Osc}^2_0} \lesssim \frac{\mathbb{E}(\mathcal{T}_0, N_0 + 1; u, f)}{2(\delta_\ell + \mathrm{Osc}^2_\ell)}$$

implies that $1 \lesssim q$ is uniformly bounded away from zero.

Since $\mathbb{E}(\mathcal{T}_0, N; u, f) \to 0$ as $N \to \infty$, the number

$$L := \min\{N \in \{N_0 + 1, N_0 + 2, \dots\} \mid \mathbb{E}(\mathcal{T}_0, N; u, f) \le q(\delta_\ell + \mathrm{Osc}^2_\ell)\}$$

is well-defined. Since $q(\delta_\ell + \mathrm{Osc}_\ell^2) < \mathbb{E}(\mathcal{T}_0, N_0 + 1; u, f)$, it follows $L \geq N_0 + 2$. Hence

$$\mathbb{E}(\mathcal{T}_0, L; u, f) \leq q(\delta_\ell + \mathrm{Osc}_\ell^2) < \mathbb{E}(\mathcal{T}_0, L - 1; u, f). \tag{4.5}$$

The definition of $|(u, f)|_{\mathcal{A}_s}$ in (1.7) and the estimate (4.5) yield

$$\begin{aligned}(L-1)^2 &\leq |(u,f)|_{\mathcal{A}_s}^{2/s} \mathbb{E}(\mathcal{T}_0, L-1; u, f)^{-1/s} \\ &\leq q^{-1/s} |(u,f)|_{\mathcal{A}_s}^{2/s} (\delta_\ell + \mathrm{Osc}_\ell^2)^{-1/s} \\ &\lesssim |(u,f)|_{\mathcal{A}_s}^{2/s} (\delta_\ell + \mathrm{Osc}_\ell^2)^{-1/s}.\end{aligned}$$

The remaining part of the proof shows that $|\mathcal{M}_\ell| \lesssim L$. Let $\mathcal{T}_\varepsilon$ be some optimal refinement by newest-vertex bisection of $\mathcal{T}_0$ with $|\mathcal{T}_\varepsilon| \leq L + |\mathcal{T}_0|$ such that

$$\mathbb{E}(\mathcal{T}_0, L; u, f) = E(u_\varepsilon) - E(u) + \mathrm{Osc}_\varepsilon^2 \leq q(\delta_\ell + \mathrm{Osc}_\ell^2).$$

Let $\mathcal{T}_{\ell+\varepsilon}$ be the overlay of the triangulations $\mathcal{T}_\ell$ and $\mathcal{T}_\varepsilon$ (which is the smallest common refinement of $\mathcal{T}_\ell$ and $\mathcal{T}_\varepsilon$ obtained by newest-vertex bisections). Since

$$E(u_{\ell+\varepsilon}) - E(u) \leq E(u_\varepsilon) - E(u) \quad \text{and} \quad \mathrm{Osc}_{\ell+\varepsilon}^2 \leq \mathrm{Osc}_\varepsilon^2,$$

it follows

$$E(u_{\ell+\varepsilon}) - E(u) + \mathrm{Osc}_{\ell+\varepsilon}^2 \leq q(\delta_\ell + \mathrm{Osc}_\ell^2).$$

This is (4.2) and Lemma 4.3 implies

$$\frac{(1-q)c_{\mathrm{Eff}}}{1 + C_{\mathrm{dRel}}} \eta_\ell(u_\ell, \mathcal{T}_\ell) \leq \eta_\ell(u_\ell, \mathcal{M}_{\ell,\ell+\varepsilon}).$$

The definition of q shows $\theta(1 + C_{\mathrm{dRel}})/c_{\mathrm{Eff}} \leq 1 - q$ and so

$$\theta \eta_\ell(u_\ell, \mathcal{T}_\ell) \leq \eta_\ell(u_\ell, \mathcal{M}_{\ell,\ell+\varepsilon}).$$

Hence, the set $\mathcal{M}_{\ell,\ell+\varepsilon}$ thus satisfies the bulk criterion (2.1). Therefore

$$|\mathcal{M}_\ell| \leq |\mathcal{M}_{\ell,\ell+\varepsilon}| \lesssim |\mathcal{T}_{\ell+\varepsilon} \setminus \mathcal{T}_\ell| \leq |\mathcal{T}_{\ell+\varepsilon}| - |\mathcal{T}_\ell|.$$

It is well known [7, Lemma 3.7] that the overlay satisfies

$$|\mathcal{T}_{\ell+\varepsilon}| - |\mathcal{T}_\ell| \leq |\mathcal{T}_\varepsilon| - |\mathcal{T}_0| \leq L - 1. \qquad \square$$

Proof of Theorem 2.2. Given the marked elements $\mathcal{M}_0, \mathcal{M}_1, \ldots$, the articles [4, 21, 22] imply

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim \sum_{k=0}^{\ell-1} |\mathcal{M}_k| \quad \text{for all } \ell \in \mathbb{N}.$$

This and (4.4) show

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim |(u,f)|_{\mathcal{A}_s}^{1/s} \sum_{k=0}^{\ell-1} (\delta_k + \mathrm{Osc}_k^2)^{-1/2s} \leq |(u,f)|_{\mathcal{A}_s}^{1/s} \max(1, 1/\gamma)^{-1/s} \sum_{k=0}^{\ell-1} (\delta_k + \gamma\eta_k^2)^{-1/2s}.$$

The contraction of Theorem 4.2 yields a constant $0 < q < 1$ such that, for all $k = 1, \ldots, \ell$,

$$\delta_\ell + \gamma\eta_\ell^2 \leq q^{\ell-k}(\delta_k + \gamma\eta_k^2).$$

The combination of the previous two estimates leads to

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim |(u,f)|_{\mathcal{A}_s}^{1/s} (\delta_\ell + \gamma\eta_\ell^2)^{-1/2s} \sum_{k=1}^{\ell} q^{k/(2s)}.$$

Since the geometrical sum

$$\sum_{k=1}^{\ell} q^{k/(2s)} \leq \frac{1}{1 - q^{1/(2s)}} \lesssim 1$$

is bounded for all $\ell \in \mathbb{N}$, we deduce

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \lesssim |(u,f)|_{\mathcal{A}_s}^{1/s} (\delta_\ell + \gamma\eta_\ell^2)^{-1/2s}. \qquad \square$$

4.4 Comments on the Boundary Layer Volume Contribution

Some comments are in order for the volume terms involved in the oscillation Osc_ℓ. The aim of the subsection is to prove that Osc_ℓ can be bounded by the usual side-oriented oscillation up to a higher order term. Define

$$\mathcal{F}_\ell(\partial\Omega) := \{T \in \mathcal{T}_\ell,\ \exists\, E \in \mathcal{E}_\ell(\partial\Omega) \cap \mathcal{E}_\ell(T) \text{ such that } 0|_E \equiv u|_E \not\equiv \chi|_E\},$$
$$\mathcal{Z}_\ell(\partial\Omega) := \mathcal{F}_\ell(\partial\Omega) \setminus \{T \in \mathcal{T}_\ell,\ \exists\, S \subset T \text{ such that } \chi|_S < u|_S \text{ and } |T| \lesssim |S|\}.$$

Theorem 4.5. *It holds that*

$$\mathrm{Osc}_\ell^2 \lesssim \sum_{E\in\mathcal{E}_\ell} \mathrm{osc}^2(f, \omega_E^{(\ell)}) + \sum_{T\in\mathcal{Z}_\ell(\partial\Omega)} |T|^{2/n}\|f\|_{L^2(T)}^2 + |||u - u_\ell|||^2. \tag{4.6}$$

Remark 4.6. Since the obstacle χ is affine and since $u = 0 \geq \chi$ on the boundary $\partial\Omega$, there are only several elements in the set $\mathcal{Z}_\ell(\partial\Omega)$. Hence the second term on the right-hand side of (4.6) is of higher order.

Proof. Since there exists a unique $T \in \mathcal{T}_\ell$ with $\omega_E^{(\ell)} = T$ and $0 \equiv u|_E \not\equiv \chi|_E$ for $E \in \mathcal{E}_\ell(\partial\Omega) \setminus \mathrm{FCBS}(\mathcal{E}_\ell, \partial\Omega, \chi)$, we have

$$\{\omega_E^{(\ell)},\ E \in \mathcal{E}_\ell(\partial\Omega) \setminus \mathrm{FCBS}(\mathcal{E}_\ell, \partial\Omega, \chi)\} = \mathcal{F}_\ell(\partial\Omega).$$

For

$$T \in \mathcal{F}_\ell(\partial\Omega) \cap \{T \in \mathcal{T}_\ell,\ \exists\, S \subset T \text{ such that } u|_S > \chi|_S \text{ and } |T| \lesssim |S|\},$$

the bubble technique [24] proves the efficiency of the volume term in the sense that

$$|\omega_E^{(\ell)}|^{2/n}\|f\|_{L^2(\omega_E^{(\ell)})}^2 \lesssim \mathrm{osc}^2(f, \omega_E^{(\ell)}) + |||u - u_\ell|||_{\omega_E^{(\ell)}}^2.$$

Since the other terms of the oscillation Osc_ℓ^2 except those for sides $E \in \mathcal{E}_\ell(\partial\Omega) \setminus \mathrm{FCBS}(\mathcal{E}_\ell, \partial\Omega, \chi)$ are the usual side-based oscillations, this completes the proof. □

Funding: This work was supported by DFG Research Center MATHEON. The work of the first author was partly supported by the WCU program through KOSEF (R31-2008-000-10049-0). The work of the second author was supported by the NSFC projects 11271035, 91430213 and 11421101.

References

[1] M. Ainsworth, J. T. Oden and C. Y. Lee, Local a posteriori error estimators for variational inequalities, *Numer. Methods Partial Differential Equations* **9** (1993), 23–33.

[2] I. Babuška and T. Strouboulis, *The Finite Element Method and its Reliability*, Clarendon Press, Oxford, 2001.

[3] S. Bartels and C. Carstensen, Averaging techniques yield reliable a posteriori finite element error control for obstacle problems, *Numer. Math.* **99** (2004), 225–249.

[4] P. Binev, W. Dahmen and R. DeVore, Adaptive finite element methods with convergence rates, *Numer. Math.* **97** (2004), 219–268.

[5] D. Braess, A posteriori error estimators for obstacle problems – Another look, *Numer. Math.* **101** (2005), 415–421.

[6] D. Braess, C. Carstensen and R. Hoppe, Convergence analysis of a conforming adaptive finite element method for an obstacle problem, *Numer. Math.* **107** (2007), 455–471.

[7] J. M. Cascon, C. Kreuzer, R. H. Nochetto, K. G. Siebert, Quasi-optimal convergence rate for an adaptive finite element method, *SIAM J. Numer. Anal.* **46** (2008), 2524–2550.

[8] Z. Chen and R. Nochetto, Residual type a posteriori error estimates for elliptic obstacle problems, *Numer. Math.* **84** (2000), 527–548.

[9] W. Dörfler, A convergent adaptive algorithm for Poisson's equation, *SIAM J. Numer. Anal.* **33** (1996), 1106–1124.

[10] F. Fierro and A. Veeser, A posteriori error estimators for regularized total variation of characteristic functions, *SIAM J. Numer. Anal.* **41** (2003), 2032–2055.

[11] R. Glowinski, J. L. Lions and R. Trémolières, *Numerical Analysis of Variational Inequalities*, North-Holland, Amsterdam, 1981.

[12] R. H. W. Hoppe and R. Kornhuber, Adaptive multilevel methods for obstacle problems, *SIAM J. Numer. Anal.* **31** (1994), 301–323.
[13] C. Johnson, Adaptive finite element methods for the obstacle problem, *Math. Models Methods Appl. Sci.* **2** (1992), 483–487.
[14] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities*, Academic Press, New York, 1980.
[15] P. Morin, R. H. Nochetto and K. G. Siebert, Data oscillation and convergence of adaptive FEM, *SIAM J. Numer. Anal.* **38** (2000), 466–488.
[16] R. H. Nochetto, K. G. Siebert and A. Veeser, Pointwise a posteriori error control for elliptic obstacle problems, *Numer. Math.* **95** (2003), 163–195.
[17] R. H. Nochetto, K. G. Siebert and A. Veeser, Fully localized a posteriori error estimators and barrier sets for contact problems, *SIAM J. Numer. Anal.* **42** (2005), 2118–2135.
[18] R. H. Nochetto and L. B. Wahlbin, Positivity preserving finite element approximation, *Math. Comp.* **71** (2002), 1405–1419.
[19] M. Page and D. Praetorius, Convergence of adaptive FEM for some elliptic obstacle problem, *Appl. Anal.* **92** (2013), 595–615.
[20] L. R. Scott and S. Y. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comp.* **54** (1990), 483–493.
[21] R. Stevenson, Optimality of a standard adaptive finite element method, *Found. Comput. Math.* **7** (2007), 245–269.
[22] R. Stevenson, The completion of locally refined simplicial partitions created by bisection, *Math. Comp.* **77** (2007), 227–241.
[23] A. Veeser, Efficient and reliable a posteriori error estimators for elliptic obstacle problems, *SIAM J. Numer. Anal.* **39** (2001), 146–167.
[24] R. Verfürth, *A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, New York, 1996.
[25] Q. Zou, A. Veeser, R. Kornhuber and C. Gräser, Hierarchical error estimates for the energy functional in obstacle problems, *Numer. Math.* **117** (2011), 653–677.