

Convergence rates for adaptive finite elements

FERNANDO D. GASPOZ[†] AND PEDRO MORIN[‡]

Instituto de Matemática Aplicada del Litoral, Consejo Nacional de Investigaciones Científicas y Técnicas and Universidad Nacional del Litoral, Güemes 3450, S3000GLN Santa Fe, Argentina

[Received on 6 November 2007; revised on 18 April 2008]

In this article, we prove that it is possible to construct, using newest vertex bisection, meshes that equidistribute the error in the H^1 -norm whenever the function to be approximated can be decomposed as a sum of a regular part plus a singular part with singularities around a finite number of points. This decomposition is usual in regularity results of partial differential equations. As a consequence, the meshes turn out to be quasi-optimal, and convergence rates for adaptive finite-element methods using Lagrange finite elements of any polynomial degree are obtained.

Keywords: convergence rates; adaptive finite elements; optimality; regularity.

1. Introduction

The study of adaptive procedures for the numerical solution of partial differential equations (PDEs) started in the late 70s and are now standard tools in science and engineering. The ultimate purpose of adaptivity is to reduce the computational cost through the automatic construction of a sequence of meshes that would eventually equidistribute the approximation errors, leading to (quasi-)optimal meshes. Adaptive methods for stationary problems usually consist of the loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \quad (1.1)$$

Experience strongly suggests that, starting from a coarse mesh, such an iteration converges within any prescribed error tolerance in a finite number of steps, and it does so in an optimal manner, provided that the *a posteriori* error estimators are reliable and efficient. What is observed in fact is that for a large class of problems and data, the solutions $u_{\mathcal{T}}$ and meshes \mathcal{T} obtained with adaptive methods of the form (1.1) satisfy

$$\|u - u_{\mathcal{T}}\|_{H^1} \leq C(\#\mathcal{T})^{-p/d}, \quad (1.2)$$

where u denotes the exact solution, p the polynomial degree of the finite-element space over the mesh \mathcal{T} and d the dimension of the underlying space. This is the same error bound that is obtained with uniformly refined meshes for smooth (regular) solutions $u \in H^{p+1}$ by an application of classical interpolation estimates (see Ciarlet, 1978). The decay rate dictated by (1.2)—which is also observed in practice for the so-called singular solutions belonging to $H^s(\Omega)$ for $s < 2$ —is usually called *optimal error decay*. The precise goal of this paper is to show a broader family of functions for which this so-called optimal decay can be obtained when using adaptive methods. We will prove that this decay holds for functions

[†]Email: fgaspoz@santafe-conicet.gov.ar

[‡]Email: pmorin@santafe-conicet.gov.ar

that can be decomposed as a sum of a regular part plus singular terms, as described in classical regularity results for PDE like those in [Grisvard \(1985, 1992\)](#), [Petzoldt \(2001\)](#) and [Dauge \(1988\)](#).

The first steps towards understanding the optimality of AFEM consisted of studying their convergence. An analysis of (1.1) for linear, elliptic and symmetric problems in one dimension is presented in [Babuška & Vogelius \(1984\)](#). The first multidimensional result is given in [Dörfler \(1996\)](#), where it is proved that, after a pre-adaptation of the data, (1.1) reduces the error below any prescribed tolerance. Proper convergence without conditions on the initial grid is proved in [Morin *et al.* \(2002\)](#), requiring the so-called *interior node property* and an additional marking step driven by *data oscillation*. The latter work has been generalized in various directions. Lately, convergence of adaptive methods with marking strategies other than Dörfler's, for a large class of linear problems with different *a posteriori* error estimators and without requiring the marking due to oscillation or the interior node property, was proved in [Morin *et al.* \(2007\)](#). The result only leads to asymptotic convergence without an error reduction in every step, which seems to be essential to prove optimality though (see [Stevenson, 2007](#); [Cascón *et al.*, 2008](#)).

Regarding complexity, an important result for an algorithm which is very similar to (1.1) is proved in [Stevenson \(2007\)](#). The proof relies on techniques first developed in [Binev *et al.* \(2004\)](#) and new ideas. This result was later improved in several aspects in [Cascón *et al.* \(2008\)](#): the artificial assumptions of interior node and marking due to data oscillation were removed and the result applies to more general elliptic equations.

When considering adaptive methods, the notion of complexity differs from the previous one, which was based on a uniform element size h . It is now defined in terms of the number of elements (or degrees of freedom) necessary to achieve a certain tolerance.

In order to be more specific at this point, we need to introduce some notation. Let us assume that we have a function $u \in H^1(\Omega)$, where Ω is a polygonal domain in \mathbb{R}^2 (polyhedral in \mathbb{R}^3) and $H^1(\Omega)$ denotes the Sobolev space of square integrable functions with square integrable weak derivatives of first order.

We consider an initial triangulation \mathcal{T}_0 of the domain Ω into simplices, and we let the *admissible triangulations* be those obtained from \mathcal{T}_0 using the bisection procedure of [Stevenson \(2008\)](#) (which coincides with the so-called newest vertex bisection when $d = 2$), without hanging nodes. For each admissible triangulation \mathcal{T} , we consider the Lagrange finite-element space

$$\mathbb{V}_{\mathcal{T}} = \{v \in H^1(\Omega) : v|_T \in \mathcal{P}^p \quad \forall T \in \mathcal{T}\},$$

where for $p \in \mathbb{N}$, \mathcal{P}^p denotes the space of polynomials of degree at most p . The best approximation error with complexity N , for $N \in \mathbb{N}$, is defined as follows:

$$\sigma_N^p(u) = \min_{\mathcal{T} \in \mathbb{T}_N} \inf_{v \in \mathbb{V}_{\mathcal{T}}} \|u - v\|_{H^1(\Omega)},$$

where $\mathbb{T}_N := \{\mathcal{T} \text{ admissible} : (\#\mathcal{T} - \#\mathcal{T}_0) \leq N\}$; that is, the minimum over \mathcal{T} is taken over all admissible triangulations obtained with at most N bisections. We now define, for $s > 0$, the approximation classes

$$\mathbb{A}_s^p = \{v \in H^1(\Omega) : \exists C \text{ such that } \sigma_N^p(v) \leq CN^{-s} \quad \forall N \in \mathbb{N}\}$$

or, equivalently,

$$\mathbb{A}_s^p = \left\{ v \in H^1(\Omega) : |v|_{\mathbb{A}_s^p} < \infty \right\} \quad \text{with } |v|_{\mathbb{A}_s^p} := \sup_{N \in \mathbb{N}} \sigma_N^p(v) N^s.$$

The first complexity results for adaptive finite-element methods (AFEM) are presented in [Binev et al. \(2004\)](#) for an algorithm that needs coarsening, which seems unnecessary, at least for symmetric elliptic problems. This and the aforementioned papers on optimality of AFEM (see [Stevenson, 2007](#); [Cascón et al., 2008](#)) study adaptive algorithms for approximating the solution u to an elliptic PDE. Essentially, the following fundamental result is proved: the adaptive algorithms generate a sequence $\{(\mathcal{T}_k, u_k)\}_{k \in \mathbb{N}}$ of triangulations and finite-element approximations $u_k \in \mathbb{V}_{\mathcal{T}_k}^1$ that satisfy the following:

$$\text{if } u \in \mathbb{A}_s^1, \text{ then } \|u - u_k\|_{H^1(\Omega)} \leq \tilde{C}(\#\mathcal{T}_k)^{-s} \quad \forall k \in \mathbb{N}.$$

That is, the sequence of triangulations and approximate solutions have a complexity with the same decay rate as the optimal ones. The interesting aspect of those results is the fact that such a (quasi-)optimal approximation is obtained through a standard adaptive loop for the elliptic problem, without *a priori* knowledge of the exact solution and with a number of operations proportional to the cardinality of the meshes. Note that a simple-minded approach to computing $\sigma_N^p(u)$ with precise knowledge of u could lead to exponential work in terms of N .

The question—already raised in [Cascón et al. \(2008\)](#)—that is still unanswered is what rate s is to be expected in different situations. From the results just described, it is clear that AFEM do a quasi-optimal job among all possible adaptive meshes. What we present in this article is quantitative information about the convergence rate of AFEM. In order to do so, we relate the membership of a function to an approximation class \mathbb{A}_s^p with its regularity, proving rigorously, through the construction of specific meshes, that a certain class of functions is contained in \mathbb{A}_s^p .

In [Binev et al. \(2002\)](#), an almost characterization of these classes is obtained for the case $p = 1$ in terms of Besov regularity for Lipschitz polygonal domains; the proof is based on an adaptive tree approximation algorithm. To illustrate the applicability of this result, we just mention—without giving too much detail—that the Besov space $B_\tau^2(L_\tau(\Omega))$ is contained in $\mathbb{A}_{1/2}^1$ for all $\tau > 1$; see [Binev et al. \(2002, Theorem 5.1\)](#).

The regularity of solutions to Poisson's problem on Lipschitz domains in terms of Besov regularity is studied in [Dahlke & DeVore \(1997\)](#). It is proved that for Poisson's equation $-\Delta u = f$ in a Lipschitz polygonal domain $\Omega \subset \mathbb{R}^2$, with homogeneous Dirichlet boundary values, $u \in B_\tau^2(L_\tau(\Omega))$, for $\frac{1}{2} < \tau < 2$ if $f \in H^1(\Omega)$.

Combining these two results, we obtain that $u \in \mathbb{A}_{1/2}^1$ if $f \in H^1$, but a stronger result holds. Using the Sobolev regularity results from [Grisvard \(1985\)](#), we have that result by assuming only $f \in L^2(\Omega)$, $u \in W_p^2(\Omega)$, that is, all derivatives of order up to two are in $L^p(\Omega)$ for all $1 \leq p < 4/3$. This in turn implies that for all $1 < \tau < 4/3$, u belongs to the Besov space $B_\tau^2(L_\tau(\Omega))$, and applying the result from [Binev et al. \(2002\)](#), this implies $u \in \mathbb{A}_{1/2}^1$ under the sole assumption of $f \in L^2(\Omega)$.

The spirit of the results that we present in this article is a combination of the results in [Binev et al. \(2002\)](#) and those in [Dahlke & DeVore \(1997\)](#). However, our approach will not hinge upon regularity in Besov terms but rather upon a decomposition of the functions as a sum of a regular part plus singular terms, as stems from the classical regularity results for PDE like those stated in [Grisvard \(1985, 1992\)](#), [Kellogg \(1975, 1992\)](#), [Petzoldt \(2001\)](#) and [Dauge \(1988\)](#). We obtain results for polygonal domains which are not necessarily Lipschitz (including slit domains) and we generalize to any polynomial degree $p \geq 1$; the proof is elementary and does not make use of sophisticated theory of L^q spaces for $q < 1$, as seems necessary in the approach of [Binev et al. \(2002\)](#). Moreover, our result is directly applicable in some cases where the Besov regularity of the solutions to the PDE is not available; instead, a decomposition into a regular plus a singular part is known to hold.

In Grisvard (1985, 1992), one can find some conditions on the element sizes relative to the distance to the points where the singularities are located in order to obtain an error of order $N^{-1/2}$ when using linear elements in two dimensions. The difference between our result and those is that we present an algorithm for *constructing those meshes using bisection*, and thus show that those meshes are attainable by an adaptive algorithm. Moreover, in view of the results in Stevenson (2007) and Cascón *et al.* (2008), a consequence of our result is that the standard adaptive algorithms proposed there generate a sequence of meshes and discrete solutions $\{\mathcal{T}_k, u_k\}_k$ satisfying $\|u - u_k\|_{H^1(\Omega)} \leq C(\#\mathcal{T}_k)^{-p/d}$. A quantitative answer regarding convergence rates of AFEMs is thus obtained for Lagrange finite elements of any polynomial degree $p \geq 1$.

The rest of the article is organized as follows: In Section 2, we state the main result and present some applications to solutions of elliptic PDE in Section 3. In Section 4, we propose an algorithm for constructing the desired mesh and prove some of its properties. We conclude the proof of the main result by bounding the error in Section 5.

2. Main result

From now on, for any admissible triangulation \mathcal{T} of the domain Ω , we let $\mathbb{V}_{\mathcal{T}}$ denote the finite-element space of continuous piecewise polynomials of degree at most p , where p is a fixed positive integer. The following is the main result of this article which states that a large family of functions, such as those obtained when solving elliptic and other PDE, belongs to $\mathbb{A}_{p/2}^p$.

THEOREM 2.1 Let $\Omega \subset \mathbb{R}^d$ be a polygonal ($d = 2$) or polyhedral ($d = 3$) domain, not necessarily Lipschitz, let \mathcal{T}_0 be an initial triangulation of Ω and suppose that

$$u = \sum_{i=0}^N u_i, \quad (2.1)$$

where:

- $u_0 \in H^1(\Omega)$, with $u_0|_T \in H^{p+1}(T)$, for all $T \in \mathcal{T}_0$,
- for $i = 1, 2, \dots, N$, u_i can be expressed in polar coordinates around x_i as

$$u_i = c_i (\ln(r_i))^{k_i} r_i^{\gamma_i} g_i(\vec{\theta}_i) \chi_i,$$

where:

1. c_i are constants and k_i are non-negative integers;
2. $\{x_i\}_{i=1}^N =: \mathcal{N}$ is a set of points in $\overline{\Omega}$ that are also vertices of \mathcal{T}_0 ;
3. r_i denotes the distance to x_i and
 - $\vec{\theta}_i = \theta_i \in [0, 2\pi)$ is the angle coordinate of x with respect to x_i and a half line starting at x_i when $d = 2$;
 - $\vec{\theta}_i = (\theta_i, \phi_i) \in [0, 2\pi) \times [0, \pi]$, where ϕ_i is the angle coordinate of x with respect to x_i and a half line R starting at x_i , and letting P denote the plane orthogonal to R that contains x_i , θ_i is the angle coordinate of the projection of x on the plane and a half line S starting at x_i contained into P when $d = 3$;
4. γ_i are positive constants;

5. the functions g_i satisfy the following assumptions depending on the dimension d :
 - $g_i \in W_\infty^1(0, 2\pi)$ satisfies the periodicity condition $g_i(0) = g_i(2\pi)$ and is piecewise W_∞^{p+1} in the following sense: there exists a partition \mathfrak{P}_i of $[0, 2\pi]$ into segments such that $g_i|_S \in W_\infty^{p+1}(S)$ for all $S \in \mathfrak{P}_i$ when $d = 2$;
 - $g_i \in W_\infty^1((0, 2\pi) \times (0, \pi))$ satisfies the periodicity conditions $g_i(0, \phi_i) = g_i(2\pi, \phi_i)$, $0 < \phi_i < \pi$, and $g_i(0, 0) = g_i(\theta_i, 0)$, $g_i(0, 2\pi) = g_i(\theta_i, 2\pi)$, $0 < \theta_i < 2\pi$, and is piecewise W_∞^{p+1} in the following sense: there exists a partition \mathfrak{P}_i of $(0, 2\pi) \times (0, \pi)$ into triangles such that $g_i|_S \in W_\infty^{p+1}(S)$ for all $S \in \mathfrak{P}_i$ when $d = 3$;
6. χ_i are $C^\infty(\overline{\Omega})$ cut-off functions;
7. the jumps of ∇u_i (if any) are aligned with the edges of the initial mesh \mathcal{T}_0 .

Then, for any given tolerance $\varepsilon > 0$, there exists a conforming triangulation \mathcal{T} , obtained by newest vertex bisection, starting from \mathcal{T}_0 such that

$$\inf_{u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}} \|u - u_{\mathcal{T}}\|_{1, \Omega} \leq \varepsilon \quad \text{and} \quad \#\mathcal{T} - \#\mathcal{T}_0 \leq \mathbf{C}_{u, \mathcal{T}_0} \frac{1}{\varepsilon^{d/p}}, \quad (2.2)$$

where $\mathbf{C}_{u, \mathcal{T}_0}$ depends on all the parameters that enter the definition of the singular part $\sum_{i=1}^N u_i$, on \mathcal{T}_0 and on u through the broken seminorm

$$|u_0|_{H_{\mathcal{T}_0}^{p+1}(\Omega)} := \left(\sum_{T \in \mathcal{T}_0} \|D^{p+1}u_0\|_{L^2(T)}^2 \right)^{1/2}$$

but not on ε . Therefore, $u \in \mathbb{A}_{p/d}^p$.

It is worth observing that if u satisfies the assumptions of the theorem, then we can only ensure that $u \in H^{1+\varepsilon}(\Omega)$ for all $0 < \varepsilon < \min_{1 \leq i \leq N} \gamma_i$. Uniform global refinements would only lead to $u \in \mathbb{A}_{\varepsilon/d}^p$, but ε could be very small, and this rate is very pessimistic with respect to the one that can be obtained with adaptivity.

REMARK 2.2 In order to shed some light on the assumptions of the theorem, we note that they imply the following:

- If we let $\gamma = \frac{\min_i \gamma_i}{2}$, we are able to control the singular terms through the following bound:

$$Cr_i^\gamma > \ln(r_i)^{k_i} r_i^{\gamma_i} \quad (2.3)$$

and similar ones. They imply that for each of the singular terms u_i , $i = 1, 2, \dots, N$, there exists a constant C such that

$$|u_i| \leq Cr_i^\gamma, \quad |\nabla u_i| \leq Cr_i^{\gamma-1} \quad \text{and} \quad |D^{p+1}u_i| \leq Cr_i^{\gamma-p-1}, \quad (2.4)$$

the last inequality holding only in the interior of the elements of \mathcal{T}_0 and thus also in the interior of any element of any refinement of \mathcal{T}_0 . The constant C depends on c_i , k_i , γ_i , the W_∞^{p+1} -norm of χ_i , the W_∞^1 -norm of g_i and the piecewise W_∞^{p+1} -norm of g_i , that is, on the $W_\infty^{p+1}(S)$ -norm of g_i for all $S \in \mathfrak{P}_i$.

The factor $\frac{1}{2}$ in the definition of γ is imposed to control the logarithmic terms. If all $k_i = 0$, $i = 1, \dots, N$, then γ could be chosen equal to $\min_i \gamma_i$ and the same bounds would hold.

- If \mathcal{T} is any refinement of \mathcal{T}_0 and $T \in \mathcal{T}$ with $T \cap \mathcal{N} = \emptyset$, then $u_i|_T \in H^{p+1}(T)$, $i = 0, 1, \dots, N$.
- Since $p \geq 1$ and $d \leq 3$, the Sobolev embedding theorem and the fact that $\gamma_i > 0$, $i = 1, 2, \dots, N$, imply that each component u_i , $i = 0, \dots, N$, is continuous in $\overline{\Omega}$, and consequently also u is continuous.

This consequences of the assumptions are the main ingredients that will be used in the proof of our results below.

NOTATION 2.3 From now on, the letter C will denote a constant, not always equal, depending on the given function u of the assumption of Theorem 2.1, through the $H^1(\Omega)$ -norm of u_0 , the broken seminorm

$$|u_0|_{H_{\mathcal{T}_0}^{p+1}(\Omega)} := \left(\sum_{T \in \mathcal{T}_0} \|D^{p+1}u_0\|_{L^2(T)}^2 \right)^{1/2}$$

and the parameters and functions defining the singular terms u_i , $i = 1, 2, \dots, N$, of u as in the second item of the previous remark. We will reserve the notation $a \lesssim b$ to denote $a \leq cb$ with a constant c depending only on shape regularity or the geometry of the domain. And $a \simeq b$ will indicate that $a \lesssim b$ and $b \lesssim a$.

3. Applications to elliptic PDE on polygons

In this section, we state two applications to elliptic PDE in two dimensions in order to illustrate the applicability of our result.

3.1 Poisson equation

Let Ω be a polygonal domain in \mathbb{R}^2 , not necessarily Lipschitz. Let u be the (weak) solution to

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{3.1}$$

As a consequence of Theorem 3.1 in Kellogg (1992) (see also Dauge, 1988, or Nochetto *et al.*, 2008, Theorem 3.1), it holds that if $f \in H^{p-1+\varepsilon}(\Omega)$ for some $\varepsilon > 0$, then u can be written as in the assumptions of Theorem 2.1, where $\mathcal{N} = \{x_i\}_{i=1}^N$ is the set of vertices of Ω , and $k_i = 0$, $i = 1, 2, \dots, N$.

In the case of $p = 1$, ε can be taken to be zero, i.e. $f \in L^2(\Omega)$, the set \mathcal{N} contains only the vertices of Ω with inner angle ω_i greater than π ($c_i = 0$ for the other vertices) and $g_i(t) = \sin(\pi t/\omega_i)$ for all $i = 1, 2, \dots, N$.

In the case of $p > 1$, the set \mathcal{N} contains all the vertices of Ω . In order to avoid the pathological cases where at least one inner angle α of Ω satisfies $\alpha p/\pi \in \mathbb{N}$, we assume that $f \in H^{p-1+\varepsilon}(\Omega)$ for some $\varepsilon > 0$ instead of $H^{p-1}(\Omega)$, but this is not such a big restriction in practice. Moreover, this hypothesis can be weakened to ask that $f \in L^2(\Omega)$ and $f|_T \in H^{p-1+\varepsilon}(T)$ for all $T \in \mathcal{T}_0$.

We conclude then that if $f \in H^{p-1+\varepsilon}(\Omega)$ (piecewise over \mathcal{T}_0), then the solution u to Poisson's equation (3.1) belongs to $\mathbb{A}_{p/2}^p$.

3.2 Interface problems for the Laplacian

Let Ω be a polygonal domain, not necessarily Lipschitz, that can be decomposed into disjoint subdomains Ω_i , $i = 1, \dots, n_d$, with polygonal boundaries $\overline{\Omega} = \bigcup_{i=1}^{n_d} \overline{\Omega}_i$. We define the interface

$$\Gamma = \overline{\left(\bigcup_{i=1}^{n_d} (\partial\Omega_i \setminus \partial\Omega) \right)}.$$

Denote by $a(x) = \sum_{i=1}^{n_d} a_i \chi_{\Omega_i}(x)$ the global weight function, which is constant and positive on each subdomain Ω_i .

We want to solve the following problem, written in variational form:

$$\text{find } u \in \mathbb{V} : \int_{\Omega} a \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in \mathbb{V}, \quad (3.2)$$

where $f \in L^2(\Omega)$, $\mathbb{V} = H_D^1(\Omega) = \{v \in H^1(\Omega) : v|_{\Gamma_D} = 0\}$ and $\Gamma_D \subset \partial\Omega$ is the Dirichlet boundary. This problem is usually called the *interface problem for the Laplacian* and corresponds to the following strong form:

$$\begin{aligned} -\nabla \cdot (a(x) \nabla u) &= f \quad \text{in } \Omega_i, \quad i = 1, 2, \dots, n_d, \\ u &= 0 \quad \text{on } \Gamma_D, \end{aligned}$$

$$\frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma_N = \partial\Omega \setminus \Gamma_D,$$

$$a_i \frac{\partial u|_{\Omega_i}}{\partial n_i} = -a_j \frac{\partial u|_{\Omega_j}}{\partial n_j} \quad \text{on } \partial\Omega_i \cap \partial\Omega_j,$$

where n denotes the outer unit normal to Ω and n_i that of Ω_i .

Following the original ideas from Kellogg (1992), Petzoldt (2001, Chapter 2, and references therein) has proved that the solution u to (3.2) satisfies the assumptions of Theorem 2.1 for $p = 1$ if the mesh \mathcal{T}_0 matches the boundaries of the subdomains Ω_i and the points on $\partial\Omega$ where the boundary condition changes are vertices of \mathcal{T}_0 . The points x_ℓ correspond to the vertices of the interface Γ , $\partial\Omega$ and those points on $\partial\Omega$ where the boundary condition changes.

We conclude that if $f \in L^2(\Omega)$, then by Theorem 2.1 the solution u belongs to $\mathbb{A}_{1/2}^1$ and the optimal error decay is recovered.

It is worth mentioning that for certain singular points x_ℓ , the value of γ_ℓ can be as close to zero as desired, depending on the values of $a(x)$ around x_ℓ , providing very singular examples for the classical theory. In order to illustrate on this, we write down the formulas derived in Kellogg (1975) to construct an exact solution of an elliptic problem with piecewise constant coefficients and vanishing right-hand side f ; for the particular case $\Omega = (-1, 1)^2$, $a = a_1$ in the first and third quadrants and $a = a_2$ in the second and fourth quadrants. An exact solution u to (3.2) for $f \equiv 0$ (and nonhomogeneous Dirichlet boundary values) is given in polar coordinates by $u(r, \theta) = r^\gamma \mu(\theta)$, where

$$\mu(\theta) = \begin{cases} \cos((\pi/2 - \sigma)\gamma) \cdot \cos((\theta - \pi/2 + \rho)\gamma) & \text{if } 0 \leq \theta \leq \pi/2, \\ \cos(\rho\gamma) \cdot \cos((\theta - \pi + \sigma)\gamma) & \text{if } \pi/2 \leq \theta \leq \pi, \\ \cos(\sigma\gamma) \cdot \cos((\theta - \pi - \rho)\gamma) & \text{if } \pi \leq \theta < 3\pi/2, \\ \cos((\pi/2 - \rho)\gamma) \cdot \cos((\theta - 3\pi/2 - \sigma)\gamma) & \text{if } 3\pi/2 \leq \theta \leq 2\pi, \end{cases}$$

and the numbers γ , ρ and σ satisfy the nonlinear relations

$$\begin{cases} R := a_1/a_2 = -\tan((\pi/2 - \sigma)\gamma) \cdot \cot(\rho\gamma), \\ 1/R = -\tan(\rho\gamma) \cdot \cot(\sigma\gamma), \\ R = -\tan(\sigma\gamma) \cdot \cot((\pi/2 - \rho)\gamma), \\ 0 < \gamma < 2, \\ \max\{0, \pi\gamma - \pi\} < 2\gamma\rho < \min\{\pi\gamma, \pi\}, \\ \max\{0, \pi - \pi\gamma\} < -2\gamma\sigma < \min\{\pi, 2\pi - \pi\gamma\}. \end{cases} \quad (3.3)$$

Choosing $\gamma = 0.1$ and solving (3.3) for R , ρ and σ using Newton's method, we obtain $R = a_1/a_2 \cong 161.4476$, $\rho = \pi/4$ and $\sigma \cong -14.92256$. A smaller γ would lead to a larger ratio R , but in principle γ may be as close to 0 as desired.

This function u belongs to the Sobolev space $H^{1+\gamma}(\Omega)$, and is thus *barely in* $H^1(\Omega)$, but—according to our results—is still in $\mathbb{A}_{p/2}^p$ for all $p \geq 1$. That is, an adaptive finite-element approximation to a solution like this, using Lagrange finite elements of degree p , will lead to a sequence of meshes and discrete solutions $\{\mathcal{T}_k, u_k\}_k$ satisfying $\|u - u_k\|_{H^1(\Omega)} \leq C(\#\mathcal{T}_k)^{-p/2}$. On the other hand, the Besov regularity of the solutions to (3.2) is not well established, and thus the results of [Binev et al. \(2002\)](#) are not yet applicable to the interface problem for the Laplacian. Until the Besov regularity of solutions to PDE is further developed, our result—which is far from being a near characterization of the class of functions that can be approximated with optimal decay $N^{-p/d}$ —still provides a useful tool to investigate the convergence rate of AFEM for PDE.

4. Construction

From now on, we assume that u is as in the assumptions of [Theorem 2.1](#) and we will present an algorithm to construct, via newest vertex bisection, a mesh fulfilling the properties stated in the theorem.

Before we introduce the algorithm, we will present a heuristic idea with the ideal properties that the optimal mesh should have. This will motivate the precise definition of the algorithm, which is rather technical, but achieves with controlled complexity the goal of *equidistribution*.

4.1 Heuristic idea

Everything in this subsection will be heuristics, and is presented here—following the arguments in [Grisvard \(1985\)](#), [Liao & Nochetto \(2003\)](#) and [Babuška et al. \(1996\)](#)—in order to motivate the choice of properties that the optimal mesh should fulfil. The precise construction of the optimal mesh and the rigorous proof of [Theorem 2.1](#) will be given in the following sections after the algorithm for constructing the mesh has been presented.

In order to introduce the basic idea, consider the simplest case of a function u written in polar coordinates as $u = r^\gamma \sin(\gamma\theta)$ on a two-dimensional domain with a re-entrant corner of inner angle π/γ at the origin. Suppose that we approximate u with continuous piecewise linear finite elements ($p = 1$) on a mesh \mathcal{T} . The H^1 -seminorm $|u - I_{\mathcal{T}}u|_{1,T}$ of the error between u and its Lagrange interpolant $I_{\mathcal{T}}u$ on each element is bounded by $h\|D^2u\|_{L^2(T)}$ if $0 \notin T$ and by $\|Du\|_{L^2(T)}$ if $0 \in T$. These quantities

(squared) also satisfy the following:

$$\begin{aligned} h_T^2 \|D^2 u\|_{L^2(T)}^2 &\cong h_T^2 r_T^{2(\gamma-2)} |T| \cong h_T^4 r_T^{2(\gamma-2)} \quad \text{if } 0 \notin T, \\ \|Du\|_{L^2(T)}^2 &\cong \int_0^{h_T} r^{2(\gamma-1)} r \, dr \cong h_T^{2\gamma} \quad \text{if } 0 \in T, \end{aligned}$$

where r_T denotes the distance of T to the origin and $h_T := |T|^{1/2} \cong \text{diam}(T)$. In order to achieve the equidistribution of the local error bounds, we then require for the mesh \mathcal{T} that, given a small parameter $h > 0$, the elements satisfy

$$h_T^4 r_T^{2(\gamma-2)} \cong h^{2\gamma} \quad \text{if } 0 \notin T \quad \text{and} \quad h_T \cong h \quad \text{if } 0 \in T.$$

Suppose now that this goal is achievable. More precisely, we can classify the elements into rings at dyadic distance from the origin, by defining

$$D_k = \{T \in \mathcal{T} : 2^{-k-1} \leq r_T < 2^{-k}\}$$

for $k \in \mathbb{N}$, $k < K := \lfloor \log_2(1/h) \rfloor$ and $D_K = \{T \in \mathcal{T} : r_T < 2^{-K}\}$.

Then, on the one hand, the elements $T \in D_k$ have size $|T| = h_T^2 \cong h^\gamma r_T^{-(\gamma-2)} \cong h^\gamma 2^{k(\gamma-2)}$, and thus $\#D_k \cong \frac{2^{-2k}}{h^\gamma 2^{k(\gamma-2)}} = h^{-\gamma} 2^{-k\gamma}$, which implies that

$$\#\mathcal{T} \cong \sum_{k \leq K} \#D_k \cong h^{-\gamma} \sum_k 2^{-k\gamma} \cong h^{-\gamma}.$$

On the other hand, the error satisfies

$$\|u - u_h\|_{1,\Omega}^2 \cong \#\mathcal{T} h^{2\gamma} \cong h^{-\gamma} h^{2\gamma} = h^\gamma \cong (\#\mathcal{T})^{-1}.$$

This finally implies that $\|u - u_h\|_{1,\Omega} \lesssim (\#\mathcal{T})^{-1/2}$, and thus $u \in \mathbb{A}_{1/2}^1$.

In the case $d = 3$ if u has a singularity like r^γ as in the previous example, the bound $\|u - u_h\|_{1,\Omega} \lesssim (\#\mathcal{T})^{-1/3}$ would be obtained if

$$h_T^5 r_T^{2(\gamma-2)} \cong h^{2\gamma+1} \quad \text{if } 0 \notin T \quad \text{and} \quad h_T \cong h \quad \text{if } 0 \in T.$$

These conditions coincide with the grading assumptions presented in [Apel et al. \(1996, Section 3.1\)](#) replacing h by h^μ , with $\mu = \frac{\gamma}{2}$ if $d = 2$ and $\mu = \frac{2\gamma+1}{5}$ if $d = 3$.

4.2 Algorithm

In this section, we will introduce the algorithm that will achieve, using newest vertex bisection, a mesh with the precise grading stated in [Section 4.1](#), generalized to polynomials of degree p .

From now on, we will use the notation

$$r_X = \min_{x_i \in \mathcal{N}} \text{dist}(x_i, X)$$

defined for X compact; typically X is a triangle T or a point x , where \mathcal{N} denotes the finite set where the singularities are located (as in the assumptions of [Theorem 2.1](#)).

We choose and fix $\gamma = \frac{\min_i \gamma_i}{2}$. This choice allows us to bound the singular terms as in (2.4).

Let \mathcal{T}_0 be the given initial mesh and let $\delta > 0$ be a small parameter so that $\#\mathcal{T}_0 \leq \delta^{-d}$. Later, δ will be chosen such that $\delta^p \approx \varepsilon$, where ε is the error to be achieved between u and $u_{\mathcal{T}}$, a discrete approximation to u in $\mathbb{V}_{\mathcal{T}}$, and \mathcal{T} is the mesh generated by the algorithm (see proof of Theorem 2.1 in Section 5.3). Now, let $K \in \mathbb{N}$ be such that

$$2^{-\frac{(K+1)(2\gamma+d-2)}{2p+d}} \leq \delta < 2^{-\frac{K(2\gamma+d-2)}{2p+d}}. \quad (4.1)$$

Denoting, for any element T , the element size by $h_T = |T|^{1/d}$, the constructive algorithm reads as follows.

```

 $\mathcal{T}_{0,0}^c \leftarrow \mathcal{T}_0$ 
 $j = 0$ 
% initial (global) refinement to control the error of  $u_0$ 
% FIRST LOOP
do
   $\mathcal{M}_{0,j} = \{T \in \mathcal{T}_{0,j}^c : h_T > \delta\}$ 
   $\mathcal{T}_{0,j+1} \leftarrow \text{refine}(\mathcal{T}_{0,j}^c, \mathcal{M}_{0,j})$ 
   $\mathcal{T}_{0,j+1}^c \leftarrow \text{complete}(\mathcal{T}_{0,j+1})$ 
   $j \leftarrow j + 1$ 
until  $\mathcal{M}_{0,j-1} = \emptyset$ 
 $J = j$ 
 $\mathcal{T}_1^c \leftarrow \mathcal{T}_{0,j}^c$ 
 $\ell = 1$ 
% Selective refinement according to distance to singularities
% SECOND LOOP
while ( $\ell < d(K + 1)$ )
   $\Omega_\ell = \cup \left\{ T \mid T \in \mathcal{T}_\ell^c \wedge r_T \leq 2^{-\frac{\ell}{d}} \right\}$ 
   $\mathcal{M}_\ell = \left\{ T \subset \Omega_\ell : h_T > \delta 2^{\frac{2\ell(\gamma-p-1)}{d(2p+d)}} \right\}$ 
   $\mathcal{T}_{\ell+1} \leftarrow \text{refine}(\mathcal{T}_\ell^c, \mathcal{M}_\ell)$ 
   $\mathcal{T}_{\ell+1}^c \leftarrow \text{complete}(\mathcal{T}_{\ell+1})$ 
   $\ell \leftarrow \ell + 1$ 
end

```

The algorithm makes use of two routines that need further explanation. The first one,

$$\mathcal{T}_{\text{new}} \leftarrow \text{refine}(\mathcal{T}_{\text{old}}, \mathcal{M}),$$

receives a mesh \mathcal{T}_{old} , usually admissible, and a set \mathcal{M} of *marked* elements from \mathcal{T}_{old} . It returns a new mesh \mathcal{T}_{new} that is obtained after bisecting once the marked elements according to the newest vertex bisection rule. The new mesh is not necessarily admissible (it may have hanging nodes), but it clearly holds that

$$\#\mathcal{T}_{\text{new}} = \#\mathcal{T}_{\text{old}} + \#\mathcal{M},$$

i.e. $\#\mathcal{T}_{\text{new}} - \#\mathcal{T}_{\text{old}} = \#\mathcal{M}$.

The second routine that is used,

$$\mathcal{T}^c \leftarrow \text{complete}(\mathcal{T}),$$

receives a mesh \mathcal{T} that is not necessarily admissible, and returns a new mesh \mathcal{T}^c which is made admissible by refining the smallest number of necessary elements, again by the newest vertex bisection rule. The study of complexity of this routine is not as easy as that of the previous one, and it is not true that there exists a constant \mathcal{C} such that

$$\#\mathcal{T}_{\ell+1}^c \leq \#\mathcal{T}_\ell^c + \mathcal{C}(\#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c).$$

The complexity result that holds—regarding the *spreading of refinement* implied by the completion algorithm—is the following one, which is a little bit weaker, but is fundamental and sufficient for the purposes of studying optimality of AFEM. It was first proved in [Binev et al. \(2004\)](#) for triangles, later extended to simplicial meshes of any dimension in [Stevenson \(2008\)](#) and requires the following.

DEFINITION 4.1 We say that an initial mesh \mathcal{T}_0 is *properly flagged* if besides being admissible it satisfies the following condition, written in the language of [Stevenson \(2008\)](#):

Any two neighboring tagged simplices $T = (x_0, \dots, x_n)_{\text{type}}$, $T' = (x'_0, \dots, x'_n)_{\text{type}}$ from \mathcal{T}_0 match in the sense that if $\overline{x_0x_n}$ or $\overline{x'_0x'_n}$ is on $T \cap T'$, then T and T' are reflected neighbours. Otherwise, the pair of neighbouring children of T and T' are reflected neighbours.

In two dimensions, this is equivalent to requiring that whenever an interior edge is a refinement edge, it is the common refinement edge for all adjacent elements. We refer the reader to [Stevenson \(2008\)](#) for details in the three-dimensional case, as well as for a proof that an initial mesh satisfying this condition can always be obtained.

THEOREM 4.2 Let $\mathcal{T}_0 = \mathcal{T}_0^c$ be an initial admissible mesh of a polygonal (polyhedral) domain Ω in \mathbb{R}^2 (\mathbb{R}^3), which is properly flagged. If the sequence $\{\mathcal{T}_\ell^c\}_{\ell \geq 1}$ is obtained by subsequent calls to

$$\begin{aligned} \mathcal{T}_{\ell+1} &\leftarrow \text{refine}(\mathcal{T}_\ell^c), \\ \mathcal{T}_{\ell+1}^c &\leftarrow \text{complete}(\mathcal{T}_{\ell+1}), \end{aligned}$$

then for $k \geq 1$ we have that

$$\#\mathcal{T}_k^c - \#\mathcal{T}_0 \leq \mathcal{C} \left(\sum_{\ell=1}^k (\#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c) \right),$$

where \mathcal{C} is a constant depending only on \mathcal{T}_0 .

As a consequence of this theorem, we have that if now $\mathcal{T}_{0,j}$, $\mathcal{T}_{0,j}^c$, \mathcal{T}_ℓ and \mathcal{T}_ℓ^c are the meshes obtained by our algorithm, it holds that

$$\#\mathcal{T}_{d(K+1)}^c - \#\mathcal{T}_0 \leq \mathcal{C} \left(\sum_{\ell=1}^{d(K+1)-1} (\#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c) + \sum_{j=0}^{J-1} (\#\mathcal{T}_{0,j+1} - \#\mathcal{T}_{0,j}^c) \right). \quad (4.2)$$

REMARK 4.3 Before proceeding to the proof of our result, some remarks are in order.

- The idea of the algorithm is to achieve an equidistribution of the error following the heuristics stated in [Section 4.1](#). Since the refinement is stronger closer to the singularity points, our approach considers a sequence of regions Ω_ℓ around them with geometrically decreasing radii given by $2^{-\ell/d}$. The denominator d in the exponent is related to the fact that we perform only one bisection to marked elements in ‘refine’, and d are necessary to achieve a halving of h_T .

- The algorithm does not take into account the different sizes of the powers γ_i ; it just looks at a worst-case scenario taking a unified value $\gamma = \frac{\min_i \gamma_i}{2}$. As we will see later, the property $\gamma > 0$ is the only one used in the proof. In the same manner, the distance to the singularity points x_i is unified by taking the minimum distance symbolized by r_T . It may look as though the simplification introduced by this *unification* will lead to suboptimal meshes, and it is true that the constant $\mathbf{C}_{u, \mathcal{T}_0}$ in (2.2) may be bigger than necessary with this approach. But this is an *a priori* approach where we want to show the *membership* of certain functions to the spaces $\mathbb{A}_{p/d}^p$, not caring about the size of their norm.
- If an efficient construction of the mesh is desired, the algorithm could be improved by marking separately according to the different strengths of the singularities. This would lead to a better constant $\mathbf{C}_{u, \mathcal{T}_0}$, but the overall theoretical result will be the same. We decided to present this unified approach for ease of presentation.
- One important property of the newest vertex bisection rule is that it leads to a sequence of meshes with a uniformly bounded shape-regularity constant, which depends only on that from the initial mesh \mathcal{T}_0 and the new vertex flagging of the initial mesh. We thus have that all the meshes \mathcal{T}_ℓ^c obtained by the application of our algorithm are shape regular with a uniform constant.

4.3 Properties of the algorithm

In this section, we will bound, through a series of lemmas, the complexity of the resulting mesh $\mathcal{T}_{d(K+1)}^c$, and in Section 5, we will relate this complexity to the error of the best approximation to u through finite-element functions over $\mathcal{T}_{d(K+1)}^c$.

The following lemma is related to the termination of the first loop of the algorithm in a finite number of steps and a control on the number of elements added. The termination of the second loop is straightforward, since it is just a *for* loop in disguise.

LEMMA 4.4 The first loop of the algorithm terminates after J iterations with $J \leq \log_2 \left(\frac{\max_{T \in \mathcal{T}_0} |T|}{\delta^d} \right) + 1$ and there exists a constant $\mathbb{C}_1 = 2|\Omega|$ such that

$$\sum_{j=0}^{J-1} (\#\mathcal{T}_{0,j+1} - \#\mathcal{T}_{0,j}^c) \leq \mathbb{C}_1 \delta^{-d}. \quad (4.3)$$

This implies that for all $T \in \mathcal{T}_1^c$, $|T| < \delta^d$.

Proof. Observe first that if one bisects an element $T \in \mathcal{T}_0$, J times with $J \geq \log_2 \left(\frac{\max_{T \in \mathcal{T}_0} |T|}{\delta^d} \right) + 1$, then the measure of the resulting sub-elements will be strictly less than δ^d and the marking step will no mark them longer. This proves the first part of the statement.

In order to prove the bound (4.3), we define, for $i \geq 0$,

$$\mathcal{F}_i = \left\{ T \mid T \in \bigcup_k \mathcal{T}_{0,k}^c \wedge 2^i \delta^d \leq |T| < 2^{i+1} \delta^d \right\}.$$

It is easy to see that even though \mathcal{F}_i contains elements belonging to different meshes, they do not overlap, and then

$$|\Omega| \geq \sum_{T \in \mathcal{F}_i} |T| \geq \sum_{T \in \mathcal{F}_i} \delta^d 2^i = \delta^d 2^i (\#\mathcal{F}_i),$$

which implies that $\#\mathcal{F}_i \leq |\Omega| \delta^{-d} 2^{-i}$.

Now, applying these estimates and using that

$$\bigcup_{i=0}^{\infty} \mathcal{F}_i = \left\{ T \mid T \in \bigcup_{k=0}^J \mathcal{T}_{0,k} \wedge |T| \geq \delta^d \right\} = \bigcup_{j=0}^{J-1} \mathcal{M}_{0,j},$$

we obtain that

$$\sum_{j=0}^{J-1} (\#\mathcal{T}_{0,j+1} - \#\mathcal{T}_{0,j}^c) = \sum_{j=0}^{J-1} \#\mathcal{M}_{0,j} = \sum_{i=0}^{\infty} \#\mathcal{F}_i \leq 2|\Omega| \delta^{-d},$$

and the claim is proved. \square

REMARK 4.5 This proof is a little complicated due to the way the algorithm was proposed in order to take into account any previous grading of the mesh. Observe that in the first loop, we do not refine all the elements, but only those which are bigger than the threshold δ , instead of doing just uniform global refinements. If we did this, the proof would be simpler, but the number of elements in \mathcal{T}_1^c would be unnecessarily bigger.

The following lemma is just an observation of the fact that if a point z is a vertex of a shape-regular triangulation, then the distance of the elements to z is an upper bound to the diameter of the element, unless of course the distance is zero. This means that the diameter of the elements can grow at most linearly with the distance to a point.

LEMMA 4.6 Let \mathcal{T} be a regular mesh such that z is a node; then $\forall T \in \mathcal{T}$ with $\text{dist}(z, T) \neq 0$, we have that $|T| \lesssim \text{dist}(z, T)^d$ or $h_T \lesssim \text{dist}(z, T)$.

This result may be familiar to some practitioners, but it is not completely obvious. A stronger result was proved in [Nochetto *et al.* \(1991, Lemma 5.1\)](#), but we decided to include its proof here for the sake of completeness.

Proof. Let T be an element of \mathcal{T} and let us define $\omega_T = \cup\{\bar{T} \mid \bar{T} \in \mathcal{T} \wedge T \cap \bar{T} \neq \emptyset\}$. If $z \notin \omega_T$, then by shape regularity, $\text{dist}(z, T) \geq ch_T$. If $z \in \omega_T \setminus T$, then z is a vertex of a neighbouring element T' and thus $\text{dist}(z, T) \approx h_{T'} \approx h_T$. \square

The next result implies that the desired grading of the mesh was achieved by the algorithm.

LEMMA 4.7 Let $\mathcal{T} = \mathcal{T}_{d(K+1)}^c$; then for $0 \leq \ell \leq d(K+1)$, the following property holds:

$$T \in \mathcal{T} \quad \text{and} \quad r_T < 2^{-\frac{\ell}{d}} \implies |T| < \delta^d 2^{\frac{2\ell(\gamma-p-1)}{2p+d}}.$$

Proof. We first claim that for each $0 \leq \ell < d(K+1)$, the following holds for the intermediate triangulations $\mathcal{T}_{\ell+1}$:

$$T \in \mathcal{T}_{\ell+1}^c \quad \text{and} \quad r_T < 2^{-\frac{\ell}{d}} \implies |T| < \delta^d 2^{\frac{2\ell(\gamma-p-1)}{2p+d}}. \quad (4.4)$$

We prove this by induction on ℓ : by Lemma 4.4, it holds for $\ell = 0$. Before proceeding, observe that if $T' \in \mathcal{T}_{\ell}^c$ and $T \in \mathcal{T}_k^c$ with $k > \ell$,

$$T \subset T' \implies r_T \geq r_{T'}. \quad (4.5)$$

Suppose now that (4.4) holds for ℓ and let us prove it for $\ell + 1$. If $T \in \mathcal{T}_{\ell+2}^c$ and $r_T < 2^{-\frac{\ell+1}{d}}$, there exist $T' \in \mathcal{T}_{\ell+1}^c$ such that $T \subset T'$, whence $r_{T'} < 2^{-\frac{\ell}{d}}$, and by the inductive assumption $|T'| < \delta^d 2^{\frac{2\ell(\gamma-p-1)}{2p+d}}$. Now, if already $|T'| < \delta^d 2^{\frac{2(\ell+1)(\gamma-p-1)}{2p+d}}$, then the results hold because $|T| \leq |T'|$. Otherwise, $T' \in \mathcal{M}_{\ell+1}$ and we have that

$$|T| \leq \frac{1}{2}|T'| < \frac{\delta^d 2^{\frac{2\ell(\gamma-p-1)}{2p+d}}}{2} < \delta^d 2^{\frac{2(\ell+1)(\gamma-p-1)}{2p+d}}$$

because $\gamma > 0$ and $d \geq 2$. Thus, (4.4) is proved for $\ell + 1$.

We now proceed to prove the claim of the lemma: Let $T \in \mathcal{T}$ such that $r_T < 2^{-\frac{\ell}{d}}$; then there exist $T' \supset T$, $T' \in \mathcal{T}_{\ell+1}^c$, and then by (4.5), $r_{T'} < 2^{-\frac{\ell}{d}}$ and by (4.4), $|T| \leq |T'| < \delta^d 2^{\frac{2\ell(\gamma-p-1)}{2p+d}}$. \square

The claim of the previous lemma could have been achieved by simple uniform refinement, but this would have destroyed the complexity of the mesh. The next lemma shows that the number of marked elements in each iteration is reasonably bounded in such a way that the overall complexity of the final mesh is under control.

LEMMA 4.8 There exists a constant \mathbb{C}_2 , depending only on shape regularity, such that for $1 \leq \ell < d(K+1)$,

$$\#\mathcal{M}_\ell = \#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c \leq \mathbb{C}_2 \delta^{-d} 2^{-\frac{\ell(2\gamma+d-2)}{2p+d}}. \quad (4.6)$$

Proof. Recall that in the algorithm we define $\Omega_\ell = \cup \{T \mid T \in \mathcal{T}_\ell^c \wedge r_T \leq 2^{-\frac{\ell}{d}}\}$, and since $\mathcal{T}_{\ell+1}$ is obtained from \mathcal{T}_ℓ^c by refinement only, we have that $\Omega_\ell = \cup \{T \mid T \in \mathcal{T}_{\ell+1} : T \subset \Omega_\ell\}$, whence

$$|\Omega_\ell| = \sum_{T \in \mathcal{T}_{\ell+1}, T \subset \Omega_\ell} h_T^d = \sum_{T \in \mathcal{T}_{\ell+1} \setminus \mathcal{T}_\ell^c} h_T^d + \sum_{\substack{T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell^c \\ T \subset \Omega_\ell}} h_T^d \geq \sum_{T \in \mathcal{T}_{\ell+1} \setminus \mathcal{T}_\ell^c} h_T^d.$$

But if $T \in \mathcal{T}_{\ell+1} \setminus \mathcal{T}_\ell^c$, then T is half of an element $T' \in \mathcal{M}_\ell$, and thus by the definition of \mathcal{M}_ℓ in the algorithm,

$$2h_T^d = h_{T'}^d \geq \delta^d 2^{\frac{2\ell(\gamma-p-1)}{2p+d}},$$

which in turn implies that

$$|\Omega_\ell| \geq \frac{\delta^d 2^{\frac{2\ell(\gamma-p-1)}{2p+d}}}{2} (\#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c) = \frac{\delta^d 2^{-\ell} 2^{\frac{\ell(2\gamma+d-2)}{2p+d}}}{2} (\#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c).$$

By Lemma 4.6, we have that $|\Omega_\ell| \leq \mathbb{C}2^{-\ell}$ and then

$$\#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c \leq 2|\Omega_\ell| 2^\ell \delta^{-d} 2^{-\frac{\ell(2\gamma+d-2)}{2p+d}} \leq \mathbb{C}_2 \delta^{-d} 2^{-\frac{\ell(2\gamma+d-2)}{2p+d}},$$

and the lemma is proved. \square

The next lemma makes use of the complexity result (4.2) of the completion procedure for the newest vertex bisection rule to bound the complexity of the final mesh.

LEMMA 4.9 There exists a constant \mathbb{C}_3 , depending only on shape regularity, the polynomial degree p , the dimension d , the function u through γ and \mathcal{T}_0 such that

$$\#\mathcal{T}_{d(K+1)}^c - \#\mathcal{T}_0 \leq \mathbb{C}_3 \delta^{-d}. \quad (4.7)$$

Proof. Using (4.2) and Lemmas 4.4 and 4.8, we have that

$$\begin{aligned} \#\mathcal{T}_{d(K+1)}^c - \#\mathcal{T}_0 &\leq \mathcal{C} \left(\sum_{\ell=1}^{d(K+1)-1} (\#\mathcal{T}_{\ell+1} - \#\mathcal{T}_\ell^c) + \sum_{j=0}^{J-1} (\#\mathcal{T}_{0,j+1} - \#\mathcal{T}_{0,j}^c) \right) \\ &\leq \mathcal{C} \left(\sum_{\ell=1}^{d(K+1)-1} \mathbb{C}_2 \delta^{-d} 2^{-\frac{\ell(2\gamma+d-2)}{2p+d}} + \mathbb{C}_1 \delta^{-d} \right) \\ &\leq \mathcal{C} \delta^{-d} \left(\mathbb{C}_2 \sum_{\ell=1}^{\infty} 2^{-\frac{\ell(2\gamma+d-2)}{2p+d}} + \mathbb{C}_1 \right). \end{aligned}$$

Since $\gamma > 0$, the sum $\sum_{\ell=1}^{\infty} 2^{-\frac{\ell(2\gamma+d-2)}{2p+d}}$ is finite and the claim follows taking $\mathbb{C}_3 = \mathcal{C} \left(\mathbb{C}_2 \sum_{\ell=1}^{\infty} 2^{-\frac{\ell(2\gamma+d-2)}{2p+d}} + \mathbb{C}_1 \right)$. \square

5. Error

In this section, we bound the best error with finite-element functions in terms of the complexity of the mesh.

THEOREM 5.1 There exist two constants \mathbf{A}_1 and \mathbf{A}_2 that may depend on u through the broken seminorm

$$|u_0|_{H_{\mathcal{T}_0}^{p+1}(\Omega)} := \left(\sum_{T \in \mathcal{T}_0} \|D^{p+1}u_0\|_{L^2(T)}^2 \right)^{1/2},$$

c_i, k_i, γ_i , the $\|\chi_i\|_{W_{\infty}^{p+1}(\Omega)}$, $\|g_i\|_{W_{\infty}^1(\Omega)}$, the $W_{\infty}^{p+1}(S)$ -norm of g_i , $S \in \mathfrak{P}_i$, $i = 1, \dots, N$, the polynomial degree p , the dimension d , shape regularity and \mathcal{T}_0 , but otherwise independent of K and δ such that if $\mathcal{T} = \mathcal{T}_{d(K+1)}^c$, then

$$\inf_{u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}} \|u - u_{\mathcal{T}}\|_{1,\Omega} \leq \mathbf{A}_1 \delta^p, \quad (5.1)$$

$$\inf_{u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}} \|u - u_{\mathcal{T}}\|_{1,\Omega} \leq \mathbf{A}_2 (\#\mathcal{T} - \#\mathcal{T}_0)^{-\frac{p}{d}}. \quad (5.2)$$

In order to prove this theorem, we will consider the regular part u_0 of u and the singular part given by $\sum_{i=1}^N u_i$.

Throughout this section, we will use the Lagrange interpolator $I_{\mathcal{T}}u_i$ of u_i , which is the finite-element function that coincides with u_i at all the nodes and is well defined for each $i = 0, 1, \dots, N$, since by the assumptions of Theorem 4.2, all the u_i functions are continuous in $\overline{\Omega}$; see Remark 2.2.

5.1 Estimation of the regular part

THEOREM 5.2 There exist two constants \mathbb{C}_4 and \mathbb{C}_5 , depending on the broken seminorm

$$|u_0|_{H_{\mathcal{T}_0}^{p+1}(\Omega)} := \left(\sum_{T \in \mathcal{T}_0} \|D^{p+1}u_0\|_{L^2(T)}^2 \right)^{1/2},$$

the polynomial degree p , shape regularity and \mathcal{T}_0 , but otherwise independent of K and δ such that if $\mathcal{T} = \mathcal{T}_{d(K+1)}^c$, then

$$|u_0 - I_{\mathcal{T}}u_0|_{1,\Omega} \leq \mathbb{C}_4\delta^p,$$

$$|u_0 - I_{\mathcal{T}}u_0|_{1,\Omega} \leq \mathbb{C}_5(\#\mathcal{T} - \#\mathcal{T}_0)^{-\frac{p}{d}}.$$

Proof. Since $u_0|_T \in H^{p+1}(T)$ for all $T \in \mathcal{T}_0$ and \mathcal{T} was obtained only by refinement, $u_0|_T \in H^{p+1}(T)$ for all $T \in \mathcal{T}$ and standard interpolation estimates (see [Ciarlet, 1978](#)) yield

$$\begin{aligned} |u_0 - I_{\mathcal{T}}u_0|_{1,\Omega}^2 &= \sum_{T \in \mathcal{T}} |u_0 - I_{\mathcal{T}}u_0|_{1,T}^2 \lesssim \sum_{T \in \mathcal{T}} h_T^{2p} \|D^{p+1}u_0\|_{L^2(T)}^2 \\ &\leq \delta^{2p} |u_0|_{H_{\mathcal{T}_0}^{p+1}(\Omega)}^2, \end{aligned}$$

where the last inequality is a consequence of Lemma 4.4 and the first loop of the algorithm.

Then, by Lemma 4.9,

$$|u_0 - I_{\mathcal{T}}u_0|_{1,\Omega} \leq \mathbb{C}_4\delta^p \leq \mathbb{C}_5(\#\mathcal{T} - \#\mathcal{T}_0)^{-\frac{p}{d}},$$

and the theorem is proved. \square

5.2 Estimation of the singular part

Throughout this section, we will denote by u one of the singular terms u_i defining u in (2.1); that is, it is defined in polar coordinates around a point x_i in $\overline{\Omega}$ as

$$u = c_i (\ln(r_i))^{k_i} r_i^{\gamma_i} g_i(\vec{\theta}_i) \chi_i \quad (5.3)$$

for some $i = 1, 2, \dots, N$ and $c_i, r_i, k_i, \gamma_i, g_i, \vec{\theta}_i, \chi_i$ as in the assumptions of Theorem 2.1.

The three bounds of (2.4) are the only features of u that will be used in the proof of the following theorem.

THEOREM 5.3 There exist two constants \mathbb{C}_6 and \mathbb{C}_7 that depend on the parameters defining u in (5.3), shape regularity and \mathcal{T}_0 , but are otherwise independent of K and δ such that if $\mathcal{T} = \mathcal{T}_{d(K+1)}^c$, then

$$|u - I_{\mathcal{T}}u|_{1,\Omega} \leq \mathbb{C}_6\delta^d,$$

$$|u - I_{\mathcal{T}}u|_{1,\Omega} \leq \mathbb{C}_7(\#\mathcal{T} - \#\mathcal{T}_0)^{-\frac{p}{d}}.$$

Proof. Let $D_\ell = \cup \left\{ T \mid T \in \mathcal{T} \wedge 2^{-\frac{\ell+1}{d}} < \text{dist}(x_i, T) \leq 2^{-\frac{\ell}{d}} \right\}$ for $0 \leq \ell < d(K+1)$ and $D_{d(K+1)} = \cup \{ T \mid T \in \mathcal{T} \wedge \text{dist}(x_i, T) \leq 2^{-(K+1)} \}$. Then we obtain

$$\begin{aligned} |u - I_{\mathcal{T}}u|_{1,\Omega}^2 &= \sum_{T \in \mathcal{T}} |u - I_{\mathcal{T}}u|_{1,T}^2 \\ &= \sum_{\ell=0}^{d(K+1)-1} \sum_{T \subset D_\ell} |u - I_{\mathcal{T}}u|_{1,T}^2 + |u - I_{\mathcal{T}}u|_{1,D_{d(K+1)}}^2. \end{aligned} \quad (5.4)$$

The second term in (5.4) can be bounded as follows:

$$\begin{aligned} |u - I_{\mathcal{T}}u|_{1,D_{d(K+1)}}^2 &\leq |u|_{1,D_{d(K+1)}}^2 + |I_{\mathcal{T}}u|_{1,D_{d(K+1)}}^2 \\ &= |u|_{1,D_{d(K+1)}}^2 + \sum_{\substack{T \subset D_{d(K+1)}, \\ x_i \in T}} |I_{\mathcal{T}}u|_{1,T}^2 + \sum_{\substack{T \subset D_{d(K+1)}, \\ x_i \notin T}} |I_{\mathcal{T}}u|_{1,T}^2 =: B_1 + B_2 + B_3. \end{aligned}$$

From (2.4) and Lemma 4.6, we obtain

$$\begin{aligned} B_1 &= |u|_{1,D_{d(K+1)}}^2 \leq |u|_{1,B(x_i, c2^{-(K+1)})}^2 \leq 2\pi C \int_0^{c2^{-(K+1)}} r^{2(\gamma-1)} r^{d-1} dr \\ &= 2\pi C \int_0^{c2^{-(K+1)}} r^{2\gamma+d-3} dr \simeq C2^{-(K+1)(2\gamma+d-2)}. \end{aligned}$$

For the term B_2 , we use the fact that on a reference element \hat{T} , $|\widehat{I_{\mathcal{T}}u}|_{1,\hat{T}} \lesssim \|\widehat{I_{\mathcal{T}}u}\|_{L^\infty(\hat{T})} = \|I_{\mathcal{T}}u\|_{L^\infty(T)}$. By (2.4), if $x_i \in T$ and $T \subset D_{d(K+1)}$, $\|I_{\mathcal{T}}u\|_{L^\infty(T)} \leq Ch_T^\gamma$. An appropriate scaling leads to

$$\begin{aligned} B_2 &= \sum_{\substack{T \subset D_{d(K+1)}, \\ x_i \in T}} |I_{\mathcal{T}}u|_{1,T}^2 \approx \sum_{\substack{T \subset D_{d(K+1)}, \\ x_i \in T}} h_T^{d-2} |\widehat{(I_{\mathcal{T}}u)}|_{1,\hat{T}}^2 \lesssim C \sum_{\substack{T \subset D_{d(K+1)}, \\ x_i \in T}} h_T^{2\gamma+d-2} \\ &\leq \#\{T \subset D_{d(K+1)} : x_i \in T\} |D_{d(K+1)}|^{\frac{2\gamma+d-2}{d}}. \end{aligned}$$

Since for these T values, $r_T = 0$, Lemma 4.7 leads to

$$B_2 \lesssim \#\{T \subset D_{d(K+1)} : x_i \in T\} (2^{-d(K+1)})^{\frac{2\gamma+d-2}{d}} \lesssim 2^{-(K+1)(2\gamma+d-2)},$$

where we have used the fact that the number of elements which have x_i as a vertex is bounded by a constant depending only on mesh regularity.

The term B_3 can be bounded using the fact that if $\text{dist}(x_i, T) > 0$, then by Lemma 4.6, $\text{dist}(x_i, T) \simeq |x - x_i| \forall x \in T$ and thus (2.4) yields

$$|\nabla I_{\mathcal{T}}u(x)| \lesssim C \text{dist}(x_i, T)^{\gamma-1} \lesssim C|x - x_i|^{\gamma-1} \quad \forall x \in T,$$

which implies that $\int_T |\nabla I_{\mathcal{T}} u|^2 \lesssim C \int_T |x - x_i|^{2(\gamma-1)} dx$, and consequently

$$\begin{aligned} B_3 &= \sum_{\substack{T \subset D_{d(K+1)}, \\ x_i \notin T}} \int_T |\nabla I_{\mathcal{T}} u|^2 \lesssim C \int_{D_{d(K+1)}} |x - x_i|^{2(\gamma-1)} dx \\ &\lesssim C \int_0^{c2^{-(K+1)}} r^{2(\gamma-1)} r^{d-1} dr \simeq C 2^{-(K+1)(2\gamma+d-2)}. \end{aligned}$$

Combining the three estimates for B_1 , B_2 and B_3 , we obtain the following bound for the second term of (5.4):

$$|u - I_{\mathcal{T}} u|_{1, D_{d(K+1)}}^2 \lesssim C 2^{-(K+1)(2\gamma+d-2)} \leq C \delta^{2p+d}. \quad (5.5)$$

Using the usual estimates for the Lagrange interpolator and the fact that $u|_T \in H^{p+1}(T) \forall T \subset \Omega \setminus D_{d(K+1)}$ (see Remark 2.2), we can bound the first term of (5.4) by

$$\sum_{\ell=0}^{d(K+1)-1} \sum_{T \subset D_\ell} |u - I_{\mathcal{T}} u|_{1, T}^2 \lesssim \sum_{\ell=0}^{d(K+1)-1} \sum_{T \subset D_\ell} h_T^{2p} \|D^{p+1} u\|_{L^2(T)}^2. \quad (5.6)$$

Finally, by (2.4), if $x \in T$, $|D^{p+1} u(x)| \leq C|x - x_i|^{\gamma-p-1}$ and thus $\|D^{p+1} u\|_{L^2(T)}^2 \leq C \text{dist}(x_i, T)^{2(\gamma-p-1)} h_T^d$; by Lemma 4.7, $h_T < \delta 2^{\frac{2\ell(\gamma-p-1)}{d(2p+d)}}$ if $T \in D_\ell$ and again by Lemma 4.6, we have that

$$\begin{aligned} \sum_{\ell=0}^{d(K+1)-1} \sum_{T \subset D_\ell} |u - I_{\mathcal{T}} u|_{1, T}^2 &\lesssim \sum_{\ell=0}^{d(K+1)-1} \sum_{T \subset D_\ell} h_T^{2p} \|D^{p+1} u\|_{L^2(T)}^2 \\ &\lesssim C \sum_{\ell=0}^{d(K+1)-1} \sum_{T \subset D_\ell} \text{dist}(x_i, T)^{2(\gamma-p-1)} h_T^{2p+d} \\ &\lesssim C \sum_{\ell=0}^{d(K+1)-1} \sum_{T \subset D_\ell} 2^{-\frac{2\ell(\gamma-p-1)}{d}} \delta^{2p+d} 2^{\frac{2\ell(\gamma-p-1)}{d}} \\ &\leq C \delta^{2p+d} \sum_{\ell=0}^{d(K+1)-1} \#D_\ell = C \delta^{2p+d} (\#\mathcal{T}). \end{aligned}$$

Summing up, by (5.4–5.6) and Lemma 4.9,

$$\begin{aligned} |u - I_{\mathcal{T}} u|_{1, \Omega}^2 &\lesssim C \delta^{2p+d} (\#\mathcal{T}) = C \delta^{2p+d} ((\#\mathcal{T} - \#\mathcal{T}_0) + \#\mathcal{T}_0) \\ &\lesssim C \delta^{2p+d} (\delta^{-d} + \#\mathcal{T}_0) \\ &\lesssim C \delta^{2p} \lesssim C (\#\mathcal{T} - \#\mathcal{T}_0)^{-\frac{2p}{d}}, \end{aligned}$$

where we have used the fact that δ was chosen so that $\#\mathcal{T}_0 \leq \delta^{-d}$. \square

5.3 Proof of the main result

Proof of Theorem 5.1. Using the estimates of Theorems 5.2 and 5.3, we obtain

$$\begin{aligned} \inf_{u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}} \|u - u_{\mathcal{T}}\|_{1,\Omega} &\lesssim \inf_{u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}} |u - u_{\mathcal{T}}|_{1,\Omega} \lesssim C|u - I_{\mathcal{T}}u|_{1,\Omega} \\ &= C \left| \sum_{i=0}^n (u_i - I_{\mathcal{T}}u_i) \right| \\ &\leq C \sum_{i=0}^N |(u_i - I_{\mathcal{T}}u_i)|_{1,\Omega} \lesssim CN\delta^p, \end{aligned}$$

and then, using Lemma 4.9, we have that

$$\inf_{u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}} \|u - u_{\mathcal{T}}\|_{1,\Omega} \lesssim CN(\#\mathcal{T} - \#\mathcal{T}_0)^{-\frac{p}{d}}.$$

□

Proof of Theorem 2.1. This is a corollary of Theorem 5.1. It is sufficient to choose $\varepsilon = \mathbf{A}_1\delta^p$. This implies the claim for ε small enough, which immediately implies the result for all $\varepsilon > 0$. □

REMARK 5.4 (Red–Green refinement) Regarding the other well-known algorithm for adaptive mesh refinement in two dimensions, namely, the so-called red–green refinement, the main result presented in this article is still open. However, the algorithm stated here can still be used for the construction of the quasi-optimal mesh with obvious modifications due to the fact that a red subdivision splits the elements into four sub-elements instead of two. The only remaining issue that needs to be solved is to determine whether a complexity result bounding the *spreading* of refined elements similar to Theorem 4.2 holds.

Acknowledgements

The authors want to thank Ricardo H. Nochetto for many interesting discussions and suggestions. They also want to express their gratitude to the anonymous referees, who through their comments and suggestions helped us to substantially improve the manuscript.

Funding

The authors were partially supported by CONICET (PIP 5478, PIP 5811); Universidad Nacional del Litoral (CAI+D 12H421).

REFERENCES

- APEL, T., SÄNDIG, A.-M. & WHITEMAN, J. R. (1996) Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains. *Math. Methods Appl. Sci.*, **19**, 63–85.
- BABUŠKA, I., ANDERSSON, B., GUO, B., MELENK, J. M. & OH, H. S. (1996) Finite element method for solving problems with singular solutions. *J. Comput. Appl. Math.*, **74**, 51–70.
- BABUŠKA, I. & VOGELIUS, M. (1984) Feedback and adaptive finite element solution of one-dimensional boundary value problems. *Numer. Math.*, **44**, 75–102.

- BINEV, P., DAHMEN, W. & DEVORE, R. (2004) Adaptive finite element methods with convergence rates. *Numer. Math.*, **97**, 219–268.
- BINEV, P., DAHMEN, W., DEVORE, R. & PETRUSHEV, P. (2002) Approximation classes for adaptive methods. *Serdica Math. J.*, **28**, 391–416.
- CASCÓN, J. M., KREUZER, CH., NOCHETTO, R. H. & SIEBERT, K. G. (2008) Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, **46**, 2524–2550.
- CIARLET, P. G. (1978) *The Finite Element Method for Elliptic Problems*. New York: North-Holland.
- DAHLKE, S. & DEVORE, R. (1997) Besov regularity for elliptic boundary value problems. *Commun. Partial Differ. Equ.* **22**, 1–16.
- DAUGE, M. (1988) *Elliptic Boundary Value Problems on Corner Domains*. Lecture Notes in Mathematics, vol. 1341. Berlin: Springer.
- DÖRFLER, W. (1996) A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, **33**, 1106–1124.
- GRISVARD, P. (1985) *Elliptic Problems in Nonsmooth Domains*. Boston, MA: Pitman.
- GRISVARD, P. (1992) *Singularities in Boundary Value Problems*. Paris: Masson.
- KELLOGG, R. B. (1975) On the Poisson equation with intersecting interfaces. *Appl. Anal.*, **4**, 101–129.
- KELLOGG, R. B. (1992) Notes on piecewise smooth elliptic boundary value problems. *Technical Note BN-1137*. College Park, MD: Institute for Physical Science and Technology, University of Maryland.
- LIAO, X. & NOCHETTO, R. H. (2003) Local a posteriori error estimates and adaptive control of pollution effects. *Numer. Methods Partial Differ. Equ.*, **19**, 421–442.
- MEKCHAY, K. & NOCHETTO, R. H. (2005) Convergence of adaptive finite element methods for general second order linear elliptic PDE. *SIAM J. Numer. Anal.*, **43**, 1803–1827.
- MORIN, P., NOCHETTO, R. H. & SIEBERT, K. G. (2002) Convergence of adaptive finite element methods. *SIAM Rev.*, **44**, 631–658.
- MORIN, P., SIEBERT, K. G. & VEESER, A. (2007) A basic convergence result for conforming adaptive finite element methods. *Math. Models Methods Appl. Sci.*, **18**, 707–737.
- NOCHETTO, R. H., PAOLINI, M. & VERDI, C. (1991) An adaptive finite element method for the two-phase Stefan problems in two space dimensions. Part I: stability and error estimates. *Math. Comput.*, **57**, 73–108; Supplement, *Math. Comput.* **57**, S1–S11.
- NOCHETTO, R. H., VEESER, A. & VERANI, M. (2008) A safeguarded dual weighted residual method. *IMA J. Numer. Anal.*, doi:10.1093/imanum/drm026.
- PETZOLDT, M. (2001) Regularity and error estimators for elliptic problems with discontinuous coefficients. *Ph.D. Thesis*. Berlin: Freie Universität Berlin.
- STEVENSON, R. (2007) Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, **7**, 245–269.
- STEVENSON, R. (2008) The completion of locally refined simplicial partitions created by bisection. *Math. Comput.*, **77**, 227–241.