Numerische Mathematik (2020) 146:159–179
https://doi.org/10.1007/s00211-020-01137-9

Numerische Mathematik

Lower a posteriori error estimates on anisotropic meshes

Natalia Kopteva[1]

Received: 13 July 2019 / Revised: 29 February 2020 / Published online: 30 July 2020

Abstract
Lower a posteriori error bounds obtained using the standard bubble function approach are reviewed in the context of anisotropic meshes. A numerical example is given that clearly demonstrates that the short-edge jump residual terms in such bounds are not sharp. Hence, for linear finite element approximations of the Laplace equation in polygonal domains, a new approach is employed to obtain essentially sharper lower a posteriori error bounds and thus to show that the upper error estimator in the recent paper (Kopteva in Numer Math 137:607–642, 2017) is efficient on partially structured anisotropic meshes.

Keywords Anisotropic triangulation · Lower a posteriori error estimate · Estimator efficiency

Mathematics Subject Classification 65N15 · 65N30

1 Introduction

The purpose of this paper is to address the efficiency of a posteriori error estimators on anisotropic meshes, which essentially reduces to obtaining sharp lower a posteriori error bounds. For shape-regular meshes such lower error bounds can be found in [1,11]. For anisotropic meshes, the situation is more delicate, as we shall now elaborate.

For unstructured anisotropic meshes, both upper and lower a posteriori error estimates were obtained in [6–8] for the Laplace equation and for a singularly perturbed reaction-diffusion equation; see also [11, §4.5]. We also refer the reader to [9], where the reliability and efficiency of a residual-type estimator from [10] based on the Zienkiewicz-Zhu recovery procedure was established on anisotropic meshes under

The author was partially supported by Science Foundation Ireland Grant SFI/12/IA/1683.

✉ Natalia Kopteva
natalia.kopteva@ul.ie
https://staff.ul.ie/natalia/

[1] Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland

an $\eta\%$ superconvergence type condition (explained, e.g., in [1, §4.8]). It should be noted that although the lower error bounds in [6–8] involve the same estimators as the corresponding upper bounds, however the error constants in the upper bounds include the so-called matching functions. The latter depend on the unknown error and take moderate values only when the mesh is either isotropic, or, being anisotropic, is aligned correctly to the solution, while, in general, they may be as large as mesh aspect ratios.

The presence of such matching functions in the estimator is clearly undesirable. It is entirely avoided in the more recent papers [2–4], where upper a posteriori error estimates on anisotropic meshes were obtained for singularly perturbed semilinear reaction-diffusion equations in the energy norm and in the maximum norm.

Interestingly, the efficiency of the estimators in [2–4] cannot be established using the standard bubble function approach, employed in [6–8]. To be more precise, this approach (which will be reviewed in Sect. 2) leads to lower error bounds with significantly smaller weights at the short-edge jump residual terms than those in the upper bounds.

The main findings of the present paper are as follows.

- Lower a posteriori error bounds obtained using the standard bubble function approach, such as in [6–8], will be reviewed in the context of anisotropic meshes. Numerical examples will be given in Sect. 2 that clearly demonstrate that the short-edge jump residual terms in such bounds are not sharp.
- Hence, we shall present a new approach that yields essentially sharper lower a posteriori error bounds and thus shows that the upper error estimator in [3] is efficient on partially structured anisotropic meshes.

Note that mild restrictions on the structure of the mesh are not uncommon in the literature when, for example, recovery type a posterior error estimators are considered [12], and, as discussed in Sect. 5.1, such restrictions are not unreasonable when an anisotropic mesh is generated starting from a regular mesh.

Compared to [2–4], to simplify the presentation, we shall restrict the consideration to the simpler Laplace equation and consider the problem

$$-\triangle u = f(x,y) \quad \text{for} \quad (x,y) \in \Omega, \quad u = 0 \quad \text{on} \quad \partial\Omega, \tag{1.1}$$

posed in a, possibly non-Lipschitz, polygonal domain $\Omega \subset \mathbb{R}^2$. We also assume that $f \in L_\infty(\Omega)$ (for a less smooth f, see Remarks 2.2 and 4.3).

Linear finite element approximations of (1.1) will be considered. Let $S_h \subset H_0^1(\Omega) \cap C(\bar{\Omega})$ be a piecewise-linear finite element space relative to a triangulation $\mathcal{T}$, and let the computed solution $u_h \in S_h$ satisfy

$$\langle \nabla u_h, \nabla v_h \rangle = \langle f, v_h \rangle \quad \forall\, v_h \in S_h, \tag{1.2}$$

where $\langle \cdot, \cdot \rangle$ denotes the $L_2(\Omega)$ inner product.

To give an idea of the results in [3], under the assumptions on the mesh described in Sect. 3, one upper error estimate reduces to [3, Theorems 6.1 and 7.4]

$$\|\nabla(u_h - u)\|_{2;\Omega} \le C \left\{ \sum_{S\in\mathcal{S}\setminus\partial\Omega} |\omega_S| J_S^2 + \sum_{T\in\mathcal{T}} \left\| H_T f^I \right\|_{2;T}^2 + \left\| f - f^I \right\|_{2;\Omega}^2 \right\}^{1/2}, \quad (1.3)$$

where C is independent of the diameters and the aspect ratios of elements in $\mathcal{T}$. Here $\mathcal{S}$ is the set of edges in $\mathcal{T}$, J_S is the standard jump in the normal derivative of u_h across any interior edge $S \in \mathcal{S}\setminus\partial\Omega$, and ω_S is the patch of two elements sharing S. We also use $H_T := \operatorname{diam}(T)$, which may be significantly larger than $h_T := 2H_T^{-1}|T|$, and the standard piecewise-linear Lagrange interpolant $f^I \in S_h$ of f.

Furthermore, under some additional assumptions on the orientation of mesh elements surrounding sequences of anisotropic nodes connected by short edges, a sharper upper estimator was obtained in [3, Theorem 6.2]:

$$\begin{aligned}\|\nabla(u_h - u)\|_{2;\Omega} \le C \Bigg\{ &\sum_{S\in\mathcal{S}\setminus\partial\Omega} |\omega_S| J_S^2 + \sum_{T\in\mathcal{T}} \left\| h_T f^I \right\|_{2;T}^2 + \left\| f - f^I \right\|_{2;\Omega}^2 \\ &+ \sum_{T\in\mathcal{T}} \left\| H_T \operatorname{osc}(f^I; T) \right\|_{2;T}^2 \Bigg\}^{1/2}. \end{aligned} \quad (1.4)$$

To relate (1.3) and (1.4) to interpolation error bounds, as well as to possible adaptive-mesh construction strategies, note that $|J_S|$ may be interpreted as approximating the diameter of ω_S under the metric induced by the squared Hessian matrix of the exact solution (while f^I approximates $\triangle u$).

Our task in this paper will be to establish the efficiency of the upper estimator in (1.4) up to data oscillation. As was already mentioned, the standard bubble function approach yields unsatisfactory lower bounds, with the weight $\frac{|S|}{\operatorname{diam}(\omega_S)}|\omega_S|$ at J_S^2 (rather than a simpler and more natural $|\omega_S|$ in (1.4)). Remark 2.4 sheds some light on our approach to remedying this.

The paper is organized as follows. In Sect. 2, we review lower a posteriori error bounds obtained using the standard bubble function approach. In particular, numerical examples are given that demonstrate that the short-edge jump residual terms in such bounds are not sharp. The remainder of the paper is devoted to obtaining sharper lower error bounds In Sect. 3, we describe basic triangulation assumptions. Then in Sect. 4, we present a version of the analysis for partially structured meshes, while the case of more general anisotropic meshes is addressed in Sect. 5.

Notation We write $a \simeq b$ when $a \lesssim b$ and $a \gtrsim b$, and $a \lesssim b$ when $a \le Cb$ with a generic constant C depending on Ω and f, but not on the diameters and the aspect ratios of elements in $\mathcal{T}$. Also, for $\mathcal{D} \subset \bar{\Omega}$ and $1 \le p \le \infty$, let $\|\cdot\|_{p;\mathcal{D}} = \|\cdot\|_{L_p(\mathcal{D})}$ and $\|\cdot\|_{\mathcal{D}} = \|\cdot\|_{2;\mathcal{D}}$, and also $\operatorname{osc}(v;\mathcal{D}) = \sup_{\mathcal{D}} v - \inf_{\mathcal{D}} v$ for $v \in L_\infty(\mathcal{D})$. Whenever quantities such as $\operatorname{osc}(\cdot;T)$ or H_T appear in volume integrals or related norms, or J_S appears in line integrals or related norms, they are understood as piecewise-constant functions.

2 Standard lower error bounds are not sharp on anisotropic meshes

This section is devoted to lower error bounds, such as in [6–8], obtained using the standard bubble function approach. Numerical examples will be given in Sect. 2.1 that clearly demonstrate that the short-edge jump residual terms in such bounds are not sharp. These examples also suggest that the jump residual terms in our upper estimators (1.3) and (1.4) have correct weights (the efficiency of the latter will be theoretically justified in Sects.4, 5). Furthermore, in Sect. 2.2, we shall review the bubble function approach when applied to anisotropic meshes and discuss its deficiencies with a view of changing the paradigm for deriving upper bounds for jump residuals associated with short edges (in particular, see Remarks 2.3 and 2.4).

2.1 Numerical examples

Our first test problem is (1.1) with the exact solution $u = \sin(\pi a x)$ (for $a = 1, 3$) and the corresponding f in $\Omega = (0,1)^2$. We employ the triangulation obtained by drawing diagonals from the tensor product of the uniform grids $\{\frac{i}{N}\}_{i=0}^{N}$ and $\{\frac{j}{M}\}_{j=0}^{M}$ in the x- and y-directions respectively (with all diagonals having the same orientation). A standard quadrature with f replaced in (1.2) by its Lagrange interpolant $f^I \in S_h$ will be used in numerical experiments.

For this problem, we compare two lower error estimates: obtained using the standard bubble function approach [8] (see also Lemma 2.1 in Sect. 2.2) and the one obtained in Sect. 4 (see Theorem 4.1). They can be described by

$$\mathcal{E} := \Big\{ \sum_{S\in\mathcal{S}\setminus\partial\Omega} \varrho_S\, |\omega_S| J_S^2 + \|h_T f^I\|_\Omega^2 \Big\}^{1/2} \lesssim \|\nabla(u_h - u)\|_\Omega + \|h_T(f - f^I)\|_\Omega, \tag{2.1a}$$

where for the weight ϱ_S for $S \in \mathcal{S}\setminus\partial\Omega$ we consider two choices:

$$\varrho_S = \begin{cases} \frac{|S|}{\operatorname{diam}(\omega_S)}, & \text{[8] using bubble functions (see also Sect. 2.2),} \\ 1, & \text{see Theorem 4.1 in Sect. 4.} \end{cases} \tag{2.1b}$$

(To be more precise, when $\varrho_S = 1$ is used, the term $\|h_T(f - f^I)\|_\Omega$ in the right-hand side of (2.1a) should be replaced by a larger $\|H_T \operatorname{osc}(f;T)\|_\Omega$; see Sect. 4 for details.) Importantly, the choice $\varrho_S = 1$, which will be theoretically justified in Sects. 4 and 5, is consistent with the jump residual terms in our upper error estimates (1.3) and (1.4).

To address whether the lower error estimator $\mathcal{E}$ in (2.1a) is sharp, the errors $\|\nabla(u_h - u)\|_\Omega$ (as well as $\|h_T(f - f^I)\|_\Omega$) are compared with $\mathcal{E}$ in Table 1. (In these computations ∇u and f are replaced, respectively, by their piecewise-linear and piecewise-quadratic interpolants.)

Clearly, the standard lower estimator with $\varrho_S = \frac{|S|}{\operatorname{diam}(\omega_S)}$ is not sharp. Not only its effectivity indices strongly depend on the ratio M/N, but, perhaps more alarmingly, $\mathcal{E}$ converges to zero as M/N increases, i.e. when the mesh is anisotropically refined in the

Table 1 Lower error estimators (2.1) for test problem with $u = \sin(\pi a x)$ in $\Omega = (0, 1)^2$

	$a = 1$			$a = 3$		
	$N = 20$	$N = 40$	$N = 80$	$N = 20$	$N = 40$	$N = 80$
Errors $\|\nabla(u_h - u)\|_\Omega$ (odd rows) and $\|h_T(f - f^I)\|_\Omega$ (even rows)						
$M = 2N$	1.01e−1	5.04e−2	2.52e−2	9.00e−1	4.52e−1	2.27e−1
	3.51e−4	4.39e−5	5.49e−6	2.83e−2	3.55e−3	4.45e−4
$M = 8N$	1.01e−1	5.04e−2	2.52e−2	9.00e−1	4.52e−1	2.27e−1
	9.74e−5	1.22e−5	1.52e−6	7.86e−3	9.86e−4	1.23e−4
$M = 32N$	1.01e−1	5.04e−2	2.52e−2	9.00e−1	4.52e−1	2.27e−1
	2.45e−5	3.07e−6	3.84e−7	1.98e−3	2.48e−4	3.11e−5
$M = 128N$	1.01e−1	5.04e−2	2.52e−2	9.00e−1	4.52e−1	2.27e−1
	6.14e−6	7.67e−7	9.59e−8	4.95e−4	6.21e−5	7.77e−6
$\mathcal{E}$ using $\varrho_S = \frac{\|S\|}{\operatorname{diam}(\omega_S)}$ (odd rows) and Effectivity Indices (even rows)						
$M = 2N$	2.80e−1	1.40e−1	7.02e−2	2.46e+0	1.25e+0	6.31e−1
	2.78	2.79	2.79	2.73	2.77	2.78
$M = 8N$	1.30e−1	6.51e−2	3.26e−2	1.14e+0	5.82e−1	2.93e−1
	1.29	1.29	1.29	1.26	1.29	1.29
$M = 32N$	6.24e−2	3.13e−2	1.57e−2	5.46e−1	2.80e−1	1.41e−1
	0.62	0.62	0.62	0.61	0.62	0.62
$M = 128N$	3.09e−2	1.55e−2	7.74e−3	2.71e−1	1.38e−1	6.95e−2
	0.31	0.31	0.31	0.30	0.31	0.31
$\mathcal{E}$ using $\varrho_S = 1$ (odd rows) and Effectivity Indices (even rows)						
$M = 2N$	3.81e−1	1.91e−1	9.55e−2	3.34e+0	1.71e+0	8.58e−1
	3.79	3.79	3.79	3.71	3.77	3.79
$M = 8N$	3.51e−1	1.76e−1	8.79e−2	3.06e+0	1.57e+0	7.90e−1
	3.48	3.49	3.49	3.40	3.47	3.49
$M = 32N$	3.48e−1	1.74e−1	8.73e−2	3.04e+0	1.56e+0	7.84e−1
	3.46	3.46	3.47	3.38	3.44	3.46
$M = 128N$	3.48e−1	1.74e−1	8.72e−2	3.04e+0	1.56e+0	7.84e−1
	3.46	3.46	3.46	3.38	3.44	3.46

wrong direction (while the error remains almost independent of M/N). By contrast, the estimator of Sect. 4, with $\varrho_S = 1$, performs quite well, with the effectivity indices stabilizing.

When comparing the two estimators, note that their weights are similar when $|S| \simeq \operatorname{diam} \omega_S$; however, they become dramatically different when $|S| \ll \operatorname{diam}(\omega_S)$, i.e. for short edges. Hence, our numerical experiments clearly suggest that it is the short-edge jump residual terms in the standard lower error estimator that are not sharp.

Next, consider a two-scale exact solution $u = \sin(x/\mu)e^{-y/\varepsilon}$ with $\mu = 2\varepsilon^2$, which exhibits a boundary layer in variable y and smaller-scale oscillations in variable x. To simplify the setting, we consider a version of problem (1.1) with this exact solution only in the boundary-layer domain $\Omega = (0, 1) \times (0, \varepsilon)$, with the corresponding f and

Dirichlet boundary conditions. The two lower estimators from (2.1) are compared in Table 2 on the mesh constructed similarly to the first test problem, with $M = 2N$, only now the 1d grid in the y-direction is $\{\varepsilon \frac{j}{M}\}_{j=0}^{M}$. Thus, the mesh is correctly adapted in the y-direction, but ignores the oscillations in the x-direction, i.e. it is anisotropic, but incorrectly aligned. As ε takes smaller values, the errors increase, which is not adequately detected by the estimator with $\varrho_S = \frac{|S|}{\operatorname{diam}(\omega_S)}$, the effectivity indices of which deteriorate (although more moderately than in Table 1; see Remark 2.1 for further discussion). The estimator with $\varrho_S = 1$ again performs quite well, with all effectivity indices close to 3.45.

Finally, in Table 3, the two estimators are tested for a one-scale exact solution $u = \sin((2y - x)/\varepsilon)$ on the anisotropic mesh which is incorrectly aligned in the x-direction only. To simplify the setting, we again consider a version of (1.1) in $\Omega = (0, 1) \times (0, \varepsilon)$ and use $N = M$. As discussed in Remark 2.1 below, both estimators exhibit stable effectivity indices. However, the bulk contribution of the short-edge residuals, computed as $\mathring{\mathcal{E}} := \big\{\sum_{S\in\mathring{\mathcal{S}}} \varrho_S\, |\omega_S| J_S^2\big\}^{1/2}$ with $\mathring{S} := \big\{|S| < \frac{1}{2}\operatorname{diam}(\omega_S)\big\}$, becomes negligible for $\varrho_S = \frac{|S|}{\operatorname{diam}(\omega_S)}$ (unlike the case $\varrho_S = 1$). This is undesirable, as may lead to the erroneous interpretation that the mesh is aligned correctly and possibly requires further refinement only in the y-direction. (As here we compare $\mathring{\mathcal{E}}$ with the overall estimator $\mathcal{E}$, it is important to note for the component $\|h_T f^I\|_\Omega$ of $\mathcal{E}$ that $\|h_T f^I\|_\Omega/\mathring{\mathcal{E}}$ was ≈ 1.43 for the first estimator and $\varepsilon = 2^{-4}$, and did not exceed 1 in all other computations for this problem.)

Remark 2.1 From the point of view of interpolation, if the anisotropic elements are aligned in the x-direction, roughly speaking, one may expect that $|J_S|$ gives an approximation to $O(h_y|\partial_y^2 u| + h_x|\partial_{xy}^2 u|)$ for long edges and $O(h_x|\partial_x^2 u| + h_y|\partial_{xy}^2 u|)$ for short edges, where h_x and h_y are the mesh sizes respectively in the x- and y-directions. In all our computations $h_y \ll h_x$. For the first test problem, $\partial_{xy}^2 u = \partial_y^2 u = 0$, which explains why the contributions of the short-edge residuals with correct weights are crucial for the overall efficiency of the estimator. In our second test, $h_y|\partial_{xy}^2 u|$ is dominated by $h_x|\partial_x^2 u|$, but not as significantly, which is reflected in a more moderate

Table 2 Lower error estimators (2.1) for test problem with $u = \sin(x/\mu)e^{-y/\varepsilon}$ for $\mu = 2\varepsilon^2$ in $\Omega = (0, 1) \times (0, \varepsilon)$ using $M = 2N$

	$\varepsilon = 2^{-2}$	$\varepsilon = 2^{-3}$	$\varepsilon = 2^{-4}$	$\varepsilon = 2^{-2}$	$\varepsilon = 2^{-3}$	$\varepsilon = 2^{-4}$
	Errors $\|\nabla(u_h - u)\|_\Omega$			$\|h_T(f - f^I)\|_\Omega$		
$N = 320$	1.66e−2	1.60e−1	1.74e+0	2.51e−7	2.79e−5	2.67e−3
$N = 640$	8.30e−3	8.01e−2	8.73e−1	3.13e−8	3.49e−6	3.34e−4
	$\mathcal{E}$ using $\varrho_S = \frac{\|S\|}{\operatorname{diam}(\omega_S)}$			Effectivity Indices		
$N = 320$	3.68e−2	2.30e−1	1.48e+0	2.22	1.44	0.85
$N = 640$	1.84e−2	1.15e−1	7.47e−1	2.22	1.44	0.86
	$\mathcal{E}$ using $\varrho_S = 1$			Effectivity Indices		
$N = 320$	5.76e−2	5.55e−1	5.92e+0	3.47	3.46	3.40
$N = 640$	2.88e−2	2.78e−1	3.01e+0	3.47	3.47	3.45

Table 3 Lower error estimators (2.1) for test problem with $u = \sin((2y - x)/\varepsilon)$ in $\Omega = (0, 1) \times (0, \varepsilon)$ using $N = M$

	$N = 160$	$N = 320$	$N = 640$	$N = 160$	$N = 320$	$N = 640$
	Errors $\|\nabla(u_h - u)\|_\Omega$			$\|h_T(f - f^I)\|_\Omega$		
$\varepsilon = 2^{-4}$	2.29e−1	1.14e−1	5.72e−2	7.17e−5	8.97e−6	1.12e−6
$\varepsilon = 2^{-5}$	6.67e−1	3.34e−1	1.67e−1	4.29e−4	5.36e−5	6.71e−6
$\varepsilon = 2^{-6}$	1.90e+0	9.59e−1	4.80e−1	2.49e−3	3.12e−4	3.90e−5
	$\mathcal{E}$ using $\varrho_S = \frac{\|S\|}{\text{diam}(\omega_S)}$ (odd rows)			Corresponding $\mathring{\mathcal{E}}$ (odd rows)		
	Effectivity Indices (even rows)			$\mathring{\mathcal{E}}/\mathcal{E}$ (even rows)		
$\varepsilon = 2^{-4}$	7.59e−1	3.80e−1	1.90e−1	6.16e−2	3.09e−2	1.54e−2
	3.32	3.32	3.32	0.08	0.08	0.08
$\varepsilon = 2^{-5}$	2.19e+0	1.10e+0	5.50e−1	1.31e−1	6.61e−2	3.31e−2
	3.28	3.29	3.29	0.06	0.06	0.06
$\varepsilon = 2^{-6}$	6.20e+0	3.13e+0	1.57e+0	2.67e−1	1.36e−1	6.82e−2
	3.26	3.27	3.27	0.04	0.04	0.04
	$\mathcal{E}$ using $\varrho_S = 1$ (odd rows)			Corresponding $\mathring{\mathcal{E}}$ (odd rows)		
	Effectivity Indices (even rows)			$\mathring{\mathcal{E}}/\mathcal{E}$ (even rows)		
$\varepsilon = 2^{-4}$	7.96e−1	3.98e−1	1.99e−1	2.46e−1	1.24e−1	6.18e−2
	3.48	3.48	3.48	0.31	0.31	0.31
$\varepsilon = 2^{-5}$	2.31e+0	1.16e+0	5.80e−1	7.43e−1	3.74e−1	1.87e−1
	3.46	3.47	3.47	0.32	0.32	0.32
$\varepsilon = 2^{-6}$	6.55e+0	3.31e+0	1.66e+0	2.14e+0	1.09e+0	5.46e−1
	3.44	3.46	3.46	0.33	0.33	0.33

deterioration of the estimator efficiency whenever $\varrho_S \ll 1$ for short edges. For the final test, $h_x|\partial^2_{xy}u| \simeq h_x|\partial^2_x u|$, so $|J_S|$ takes similar in magnitude values for short and long edges; hence, even when the bulk contribution of short-edge residuals is almost nullified by $\varrho_S \ll 1$, the overall estimator efficiency remains adequate.

2.2 Lower error bounds using the standard bubble approach

Here, for completeness, and with a view of motivating the new approach of Sects. 4 and 5, we prove a version of the lower error bound from [8, Theorem 5.1]; see also [11, Theorem 4.37]. Similar bounds can also be found in [6, Theorem 2] for the 3d case, and in [7, Theorem 4.3] for a singularly perturbed equation; see also [11, §4.5]. Note also that Lemma 2.1 below gives a version of the lower error bounds from [7, Theorem 4.3], while in the earlier literature the weight $\varrho_S = \frac{|S|}{\text{diam}(\omega_S)}$ in the bounds of type (2.2b) was replaced by the smaller ϱ_S^2.

Lemma 2.1 *Let $\mathcal{T}$ satisfy the maximum angle condition, and let $|T| \simeq |\omega_S|\ \forall\, T \subset \omega_S$, $S \in \mathcal{S}\backslash\partial\Omega$. Then for a solution u of (1.1) and any $u_h \in S_h$, one has*

$$h_T \|f^I\|_T \lesssim \|\nabla(u_h - u)\|_T + h_T \|f - f^I\|_T \qquad \forall T \in \mathcal{T}, \tag{2.2a}$$

$$\tfrac{|S|}{\operatorname{diam}(\omega_S)} |\omega_S| J_S^2 \lesssim \|\nabla(u_h - u)\|_{\omega_S}^2 + \|h_T(f - f^I)\|_{\omega_S}^2 \qquad \forall S \in \mathcal{S}\backslash\partial\Omega. \tag{2.2b}$$

Proof (i) On any $T \in \mathcal{T}$, consider $w := f^I\, \phi_1\phi_2\phi_3$, where $\{\phi_i\}_{i=1}^3$ are the standard hat functions associated with the three vertices of T. Now, a standard calculation yields $\|f^I\|_T^2 \simeq \langle f^I, w\rangle$. Note also that, in view of (1.1) and also $\Delta u_h = 0$ on T, one has $\langle f^I, w\rangle = \langle \nabla(u - u_h), \nabla w\rangle - \langle f - f^I, w\rangle$. Next, invoking $\|\nabla w\|_T \lesssim h_T^{-1}\|w\|_T$, one arrives at

$$\|f^I\|_T^2 \lesssim \left(h_T^{-1}\|\nabla(u_h - u)\|_T + \|f - f^I\|_T\right)\|w\|_T.$$

The first desired result (2.2a) follows in view of $\|w\|_T \lesssim \|f^I\|_T$.

(ii) For each of the two triangles $T \subset \omega_S$, introduce a triangle $\widetilde{T} \subseteq T$ with an edge S such that $|\widetilde{T}| \simeq h_T|S|$. Next, set $w := J_S\, \widetilde{\phi}_1\widetilde{\phi}_2$, where $\widetilde{\phi}_1$ and $\widetilde{\phi}_2$ are the hat functions associated with the end points of S on the obtained triangulation $\{\widetilde{T}\}_{T\subset\omega_S}$ (with $w := 0$ on each $T\backslash\widetilde{T}$ for $T \subset \omega_S$). A standard calculation using $\Delta u_h = 0$ in $T \subset \omega_S$ and (1.1), yields

$$|S|J_S^2 \simeq \int_S w\,[\partial_{\boldsymbol{\nu}} u_h]_S = \langle \nabla u_h, \nabla w\rangle = \langle \nabla(u_h - u), \nabla w\rangle + \langle f, w\rangle.$$

Next, invoking $\|\nabla w\|_T \lesssim h_T^{-1}\|w\|_T$ for any $T \subset \omega_S$, we arrive at

$$|S|J_S^2 \lesssim \sum_{T\in\omega_S} \underbrace{\left(h_T^{-1}\|\nabla(u_h - u)\|_T + \|f\|_T\right)}_{\lesssim h_T^{-1}\,\mathcal{Y}^T_{(2.2a)}} \underbrace{\|w\|_T}_{\simeq (h_T|S|)^{1/2}|J_S|}, \tag{2.3}$$

where $\mathcal{Y}^T_{(2.2a)}$ denotes the right-hand side of (2.2a), and the latter bound was also employed for the estimation of $\|f\|_T$. The second desired bound (2.2b) follows in view of $h_T = 2|T|/H_T \simeq |\omega_S|/\operatorname{diam}(\omega_S)$. □

Remark 2.2 The piecewise-linear Lagrange interpolant f^I of f used in (2.2) may be replaced by any, possibly discontinuous, quasi-interpolant of f (such as the piecewise-constant approximation of f by its element average values).

Remark 2.3 (Deficiency of the bubble function approach) An inspection of the above proof shows that it is sharp in the sense that it cannot be tweaked to remove the weight $\frac{|S|}{\operatorname{diam}(\omega_S)}$ in (2.2b). More precisely, for such an improvement, one would need $h_T \simeq |\omega_S|/|S|$ in (2.3), which is not the case for short edges.

Remark 2.4 (Preview of the new approach). The bubble function in the proof of (2.2b) may be viewed as a simplest local cut-off function. However, in the case of anisotropic

mesh elements, its gradient is not consistent with the diameter of the local patch. To remedy this, when dealing with short edges in Sects. 4 and 5 below, we shall switch to a cut-off function, the support of which comprises a larger local patch of anisotropic elements (rather than a two-triangle patch) and has an interior diameter $\simeq \operatorname{diam}(\omega_S)$. (Such local patches are highlighted in grey in Fig. 1 (left) and Fig. 2.) Unsurprisingly, this approach brings new challenges. For example, we have to deal with multiple edges inside this larger patch; in particular, we need to find a way to (almost) eliminate the jump residuals associated with the long edges. But this change of the paradigm will lead to essentially sharper lower error bounds of type (2.1a) with $\varrho_S = 1$.

3 Basic triangulation assumptions

In the remainder of the paper, we shall use $z = (x_z, y_z)$, S and T to denote particular mesh nodes, edges and elements, respectively, while $\mathcal{N}$, $\mathcal{S}$ and $\mathcal{T}$ will denote their respective sets. For each $z \in \mathcal{N}$, let ω_z be the patch of elements surrounding z, $\mathcal{S}_z$ the set of edges originating at z, and

$$H_z := \operatorname{diam}(\omega_z), \quad h_z := H_z^{-1}|\omega_z|, \quad \gamma_z := \mathcal{S}_z \setminus \partial\Omega.$$

Throughout the paper we make the following triangulation assumptions.

- *Maximum angle condition* Let the maximum interior angle in any triangle $T \in \mathcal{T}$ be uniformly bounded by some positive $\alpha_0 < \pi$.
- *Local element orientation condition* For any $z \in \mathcal{N}$, there is a rectangle $\omega_z^* \supset \omega_z$ such that $|\omega_z^*| \simeq |\omega_z|$.
- Also, let the number of triangles containing any node be uniformly bounded.

Note that the above conditions are automatically satisfied by shape-regular triangulations.

Additionally, we restrict our analysis to the following two node types defined using a fixed small constant c_0 (to distinguish between anisotropic and isotropic elements), with the notation $a \ll b$ for $a < c_0 b$.

(1) *Anisotropic nodes*, the set of which is denoted by $\mathcal{N}_{\text{ani}}$, are such that

$$h_z \ll H_z, \quad \text{and} \quad |T| \simeq |\omega_z| \quad \forall\, T \subset \omega_z. \tag{3.1}$$

Note that the above implies that $\mathcal{S}_z$ contains at most two edges of length $\lesssim h_z$ (see also Fig. 2).

(2) *Regular nodes*, the set of which is denoted by $\mathcal{N}_{\text{reg}}$, are those surrounded by shape-regular mesh elements.

The above imposes a gradual transition between anisotropic and isotropic elements, i.e. the set $\mathcal{N}_{\text{ani}} \cap \mathcal{N}_{\text{reg}}$ is not necessarily empty. (To simplify the presentation, here we exclude more general node types, such as in [2–4], with both anisotropic and isotropic mesh elements allowed to appear within the same patch ω_z.)

Next, recall that ω_S is the patch of two elements sharing S, and introduce the set of short edges

$$\mathring{\mathcal{S}} := \{S \in \mathcal{S}\backslash\partial\Omega : |S| \ll \operatorname{diam}(\omega_S)\}.$$

Now, motivated by our upper error estimate (1.4), for any open domain $\mathcal{D} \subset \Omega$, define

$$\mathcal{Y}_{\mathcal{D}} := \|\nabla(u_h - u)\|_{\mathcal{D}} + \|H_T \operatorname{osc}(f\,;T)\|_{\mathcal{D}}\,, \tag{3.2a}$$

$$\mathcal{E}_{\mathcal{D}} := \Big\{ \sum_{S\in\mathcal{S}\cap\mathcal{D}} |\omega_S| J_S^2 + \|h_T\, f^I\|_{\mathcal{D}} \Big\}^{1/2}, \tag{3.2b}$$

$$\mathring{\mathcal{E}}_{\mathcal{D}} := \Big\{ \sum_{S\in\mathring{\mathcal{S}}\cap\mathcal{D}} |\omega_S| J_S^2 \Big\}^{1/2}. \tag{3.2c}$$

Remark 3.1 By Lemma 2.1, one has $\big\{\sum_{S\subset\mathcal{D}\backslash\mathring{\mathcal{S}}} |\omega_S| J_S^2 + \|h_T\, f^I\|_{\mathcal{D}}\big\}^{1/2} \lesssim \mathcal{Y}_{\mathcal{D}}$. Indeed, this follows from (2.2) combined with $\mathcal{Y}^T_{(2.2a)} \le \mathcal{Y}_T$, where $\mathcal{Y}^T_{(2.2a)}$ denotes the right-hand side of (2.2a). Hence, for $\mathcal{E}_{\mathcal{D}} \lesssim \mathcal{Y}_{\mathcal{D}}$, it suffices to prove that $\mathring{\mathcal{E}}_{\mathcal{D}} \lesssim \mathcal{Y}_{\mathcal{D}}$.

4 Estimator efficiency on a partially structured anisotropic mesh

4.1 Lower error bound on a partially structured anisotropic mesh

To illustrate our approach in a simpler setting, we first present a version of the analysis for a simpler, partially structured, anisotropic mesh in a square domain $\Omega = (0, 1)^2$. So, throughout this section, we make the following triangulation assumptions.

A1. Let $\{x_i\}_{i=0}^n$ be an arbitrary mesh on the interval $(0, 1)$ in the x direction. Then, let each $T \in \mathcal{T}$, for some i,

(i) have the shortest edge on the line segment $\mathcal{P}_i := \{x = x_i,\ y \in [0, 1]\}$;

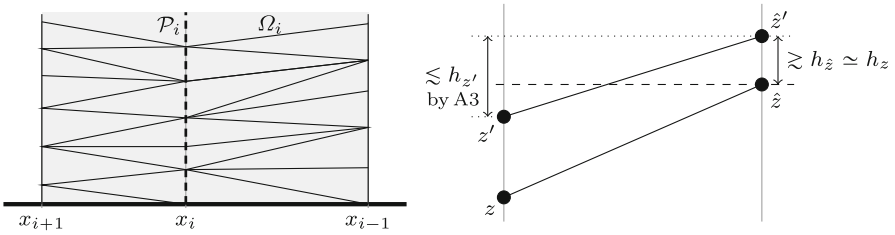

Fig. 1 Partially structured anisotropic mesh (left); illustration for Remark 4.2 (right): for any fixed edge $z\hat{z}$ and any edge $z'\hat{z}'$ intercepting the dashed horizontal line via $\hat{z}$, the figure shows that $h_z \lesssim h_{z'}$, so there is a uniformly bounded number of edges of type $z'\hat{z}'$, so $\omega_z^* \subset \omega_z^{(k)}$ with $k \lesssim 1$

(ii) have a vertex on $\mathcal{P}_{i+1}$ or $\mathcal{P}_{i-1}$ (see Fig. 1, left).

A2. Let $\mathcal{N} = \mathcal{N}_{\text{ani}}$, i.e. each mesh node z satisfies (3.1).
A3. *Global element orientation condition* For any $z \in \mathcal{N}$, there is a rectangle $\omega_z^* \supset \omega_z$ with sides parallel to the coordinate axes such that $|\omega_z^*| \simeq |\omega_z|$.

These conditions essentially imply that all mesh elements are anisotropic and aligned in the x-direction.

Theorem 4.1 *Let u and u_h solve, respectively, (1.1) and (1.2) under conditions A1–A3. Then in $\Omega_i := (x_{i-1}, x_{i+1}) \times (0, 1)$, using the notation (3.2), one has*

$$\mathring{\mathcal{E}}_{\Omega_i} \lesssim \mathcal{Y}_{\Omega_i} \quad \forall\, i = 1, \ldots, n-1. \tag{4.1}$$

The remainder of this section will be devoted to the proof of this result.

Corollary 4.2 *Under the conditions of Theorem 4.1, with Ω_0 and Ω_n defined using $x_{-1} := x_0$ and $x_{n+1} := x_n$, one has*

$$\mathcal{E}_{\Omega_i} \lesssim \mathcal{Y}_{\Omega_i} \quad \forall\, i = 0, \ldots, n, \quad \mathcal{E}_{\Omega} \lesssim \mathcal{Y}_{\Omega}.$$

Proof Combining (4.1) with $\mathring{\mathcal{E}}_{\Omega_0} = \mathring{\mathcal{E}}_{\Omega_n} = 0$ (as there are no short edges in $\Omega_0 \cup \Omega_n$) and Remark 3.1, we conclude that $\mathcal{E}_{\Omega_i} \lesssim \mathcal{Y}_{\Omega_i}\ \forall\, i$. The final bound $\mathcal{E}_{\Omega} \lesssim \mathcal{Y}_{\Omega}$ follows. □

Remark 4.1 (Estimator efficiency) It follows from [3, Theorems 5.1 and 7.4] that if $f(0, y) = f(1, y)$, then $\|\nabla(u_h - u)\|_{\Omega} \lesssim \mathcal{E}_{\Omega} + \|H_T \operatorname{osc}(f; T)\|_{\Omega} + \|f - f^I\|_{\Omega}$. Comparing this upper error bound with $\mathcal{E}_{\Omega} \lesssim \mathcal{Y}_{\Omega}$ from Corollary 4.2, we conclude that the error estimator $\mathcal{E}_{\Omega}$ is efficient up to data oscillation.

4.2 Preliminary results for partially structured meshes

The following result will be useful in the proof of Theorem 4.1.

Lemma 4.3 *(i) If $z \in \mathcal{P}_i \setminus \partial\Omega$ for some $1 \le i \le n-1$, with $\gamma_z \cap \mathcal{P}_i$ formed by the two edges S^- and S^+, then*

$$\left|J_{S^+} - J_{S^-}\right| \lesssim h_z H_z^{-1} \sum_{S \in \gamma_z \setminus \mathcal{P}_i} |J_S|. \tag{4.2}$$

(ii) If $z \in \mathcal{P}_i \cap \partial\Omega$ for some $1 \le i \le n-1$, with $\gamma_z \cap \mathcal{P}_i$ formed by a single edge S^+, then (4.2) holds true with J_{S^-} replaced by 0.

Proof (i) As $z \notin \partial\Omega$, so $\sum_{S \in \gamma_z} [\![\nabla u_h]\!]_S = 0$, where $[\![\nabla u_h]\!]_S$ denotes the jump in ∇u_h across any edge S in γ_z evaluated in the anticlockwise direction about z. Multiplying this relation by the unit vector $\mathbf{i}_x$ in the x-direction, and noting that $[\![\nabla u_h]\!]_{S^\pm} \cdot \mathbf{i}_x = \pm J_{S^\pm}$, one gets the desired assertion. Here we also use the observation that for $S \in$

$\gamma_z \backslash \mathcal{P}_i$, one has $|[\![\nabla u_h]\!]_S \cdot \mathbf{i}_x| \simeq |J_S\, \boldsymbol{\nu}_S \cdot \mathbf{i}_x|$, where $\boldsymbol{\nu}_S$ is a unit normal vector to S, for which A3 implies $|\boldsymbol{\nu}_S \cdot \mathbf{i}_x| \lesssim h_z H_z^{-1}$.

(ii) Now $z \in \partial\Omega$, so extend u_h to $\mathbb{R}^2 \backslash \Omega$ by 0 and imitate the above proof with the modification that now $\sum_{S \in \mathcal{S}_z} [\![\nabla u_h]\!]_S = 0$. When dealing with the two edges on $\partial\Omega$, note that for $S \in \mathcal{S}_z \cap \partial\Omega$, one gets $\boldsymbol{\nu}_S \cdot \mathbf{i}_x = 0$. □

Corollary 4.4 *Under the conditions of Lemma 4.3, one has*

$$|\omega_z| \left| \tfrac{H_z}{h_z} (J_{S+} - J_{S-}) \right|^2 \lesssim \mathcal{Y}_{\omega_z}^2, \tag{4.3}$$

where $\mathcal{Y}_{\omega_z}$ is from (3.2a), and if $z \in \mathcal{P}_i \cap \partial\Omega$, then J_{S-} in (4.3) is replaced by 0.

Proof In view of (4.2), the left-hand side in (4.3) is $\lesssim \sum_{S \in \gamma_z \backslash \mathcal{P}_i} |\omega_S| J_S^2$, where we also used $|\omega_S| \simeq |\omega_z|\ \forall\, S \in \gamma_z$. Next, note that the set of edges $\{S \in \gamma_z \backslash \mathcal{P}_i\}$ can be described as $\{S \subset \omega_z \backslash \mathring{\mathcal{S}}\}$, so, by Remark 3.1, the desired assertion follows. □

Remark 4.2 The minimal rectangle ω_z^* from condition A3 is defined by $\omega_z^* = (x_{i-1}, x_{i+1}) \times (y_z^-, y_z^+)$, where (y_z^-, y_z^+) is the range of y within ω_z. For this rectangle, the above conditions (in particular A3) imply that $y_z^+ - y_z^- \simeq h_z$. Furthermore, there is $k \lesssim 1$ such that $\omega_z^* \subset \omega_z^{(k)}\ \forall\, z \in \mathcal{N}$, where $\omega_z^{(0)} := \omega_z$, and $\omega_z^{(j+1)}$ denotes the patch of elements in/touching $\omega_z^{(j)}$. This conclusion is illustrated on Fig. 1 (right). (Note that $k = 1$ if our partially structured triangulation is non-obtuse.)

4.3 Proof of Theorem 4.1

Proof Throughout the proof we shall use the somewhat simplified notation $\mathcal{Y}_i := \mathcal{Y}_{\Omega_i}$ and $\mathring{\mathcal{E}}_i := \mathring{\mathcal{E}}_{\Omega_i}$, and also will frequently drop the index i and write $\mathcal{P} := \mathcal{P}_i = \{x = x_i,\ y \in [0,1]\}$, and $H := H_i := \frac{1}{2}(x_{i+1} - x_{i-1})$. With this notation, $\mathring{\mathcal{S}} \cap \Omega_i = \mathcal{P}$, so, taking into consideration the structure of the mesh (see Fig. 1, left), (3.2c) and (3.2a) with $\mathcal{D} = \Omega_i$ can be rewritten as

$$\mathring{\mathcal{E}}_i^2 = \sum_{S \subset \mathcal{P}} |\omega_S| J_S^2 = H \int_{\mathcal{P}} J_S^2\,, \quad \mathcal{Y}_i = \|\nabla(u_h - u)\|_{2;\Omega_i} + H \|\mathrm{osc}(f\,;T)\|_{2;\Omega_i}\,. \tag{4.4}$$

Next, note that for any $v \in H_0^1(\Omega)$ and $v_h \in S_h$, a standard calculation using (1.1), (1.2) yields

$$\begin{aligned} \underbrace{\langle \nabla(u_h - u), \nabla v \rangle}_{=:\,\psi_1} &= \langle \nabla u_h, \nabla v \rangle - \langle f, v \rangle \\ &= \underbrace{\langle \nabla u_h, \nabla(v - \tfrac{1}{2} v_h) \rangle}_{=:\,\Psi + \frac{1}{2} H^{-1} \mathring{\mathcal{E}}_i^2} - \underbrace{\langle f, v - \tfrac{1}{2} v_h \rangle}_{=:\,\psi_2}\,. \end{aligned} \tag{4.5}$$

As this immediately implies $\mathring{\mathcal{E}}_i^2 \lesssim H(|\psi_1| + |\psi_2| + |\Psi|)$, it suffices to prove that

$$H\big(|\psi_1| + |\psi_2| + |\Psi|\big) \lesssim \mathcal{Y}_i\,(\mathring{\mathcal{E}}_i + \mathcal{Y}_i). \tag{4.6}$$

The desired assertion (4.1) will indeed follow, in view of $\mathcal{Y}_i\,\mathring{\mathcal{E}}_i \le \theta\mathring{\mathcal{E}}_i^2 + \frac{1}{4}\theta^{-1}\mathcal{Y}_i^2$ with a sufficiently small positive constant θ.

The remainder of the proof is split into three parts. In part (i), we shall describe appropriate non-standard v_h and v, which will be crucial for (4.6) to hold true. Certain sufficient conditions for the latter will be established in part (ii), and then shown to be satisfied in part (iii).

(i) Crucially, in (4.5), we require that $v_h \in S_h$ (with the convention that $v_h = 0$ in $\mathbb{R}\backslash\Omega$) and $v := \hat{v}_h \notin S_h$ both have support in Ω_i and satisfy

$$v_h(z) := \tfrac{1}{2}\sum_{S\in\gamma_z\cap\mathcal{P}} J_S \quad \forall\, z \in \mathcal{P}\backslash\partial\Omega, \qquad \hat{v}_h(x, y) := v_h\big(x_i + 2[x - x_i],\, y\big). \tag{4.7}$$

Note that $\gamma_z\cap\mathcal{P}$, which appears in the definition of nodal values of v_h, includes exactly two short edges, while, to be more precise, $\hat{v}_h$ has support in $\hat{\Omega}_i := (x_{i-1/2}, x_{i+1/2}) \times (0, 1) \subset \Omega_i$.

(ii) We claim that for (4.6), and hence for the desired assertion (4.1), it suffices to prove that the following conditions are satisfied:

$$\Big|\int_S (\hat{v}_h - \tfrac{1}{2}v_h)\Big| \lesssim \|\partial_y v_h\|_{1;\omega_z^*} \quad \forall\, S \in \gamma_z\backslash\mathcal{P},\ z \in \mathcal{P}, \tag{4.8a}$$

$$\|H\nabla v_h\|_{2;\Omega_i} + \|v_h\|_{2;\Omega_i} \lesssim \mathring{\mathcal{E}}_i + \mathcal{Y}_i\,, \tag{4.8b}$$

$$\sum_{S\subset\mathcal{P}} |\omega_S|\Big\{\tfrac{H}{|S|}\mathrm{osc}(v_h\,;\,S)\Big\}^2 \lesssim \mathcal{Y}_i^2\,. \tag{4.8c}$$

Indeed, for ψ_1 from (4.5), by (4.4), one immediately has $|\psi_1| \lesssim \mathcal{Y}_i\|\nabla v\|_{2;\Omega_i}$. Here, by (4.7), $\|\nabla v\|_{2;\Omega_i} = \|\nabla\hat{v}_h\|_{2;\Omega_i} \simeq \|\nabla v_h\|_{2;\Omega_i}$, for which we have (4.8b). Combining these observations, one gets the desired bound on ψ_1 in (4.6).

Next, for ψ_2 from (4.5), set $\hat{f}(x, y) := f(x_i + 2[x - x_i], y)$ (similarly to $\hat{v}_h$ in (4.7)). Then $\frac{1}{2}\langle f, v_h\rangle = \langle\hat{f}, \hat{v}_h\rangle = \langle\hat{f}, v\rangle$, so

$$|\psi_2| = |\langle f - \hat{f}, v\rangle| \le \|f - \hat{f}\|_{2;\hat{\Omega}_i}\,\|v\|_{2;\hat{\Omega}_i} \lesssim \|\mathrm{osc}(f\,;T)\|_{2;\Omega_i}\,\|v_h\|_{2;\Omega_i}\,. \tag{4.9}$$

Here we also used $\|v\|_{2;\Omega_i} = \|\hat{v}_h\|_{2;\Omega_i} \simeq \|v_h\|_{2;\Omega_i}$ (in view of (4.7)), while the bound on $\|f - \hat{f}\|_{2;\hat{\Omega}_i}$ follows from Remark 4.2. Combining the above with $H\|\mathrm{osc}(f\,;T)\|_{2;\Omega_i} \lesssim \mathcal{Y}_i$ (in view of (4.4)) and the bound in (4.8b) on $\|v_h\|_{2;\Omega_i}$ yields the desired bound on ψ_2 in (4.6).

Finally, consider Ψ, the most delicate term in (4.5). To check that the corresponding bound in (4.6) follows from (4.8), note that in each triangle $T \in \mathcal{T}\cap\Omega_i$, one has $\triangle u_h = 0$, so $\int_T \nabla u_h\cdot\nabla(v - \frac{1}{2}v_h) = \int_{\partial T}\nabla u_h\cdot\boldsymbol{\nu}(v - \frac{1}{2}v_h)$. Note also that $v = v_h = 0$

on $\partial\Omega_i$, so $\langle \nabla u_h, \nabla(v-\frac12 v_h)\rangle = \sum_{S\subset\Omega_i}\int_S J_S(v-\frac12 v_h)$. It also follows from (4.7) that $v-\frac12 v_h = \frac12 v_h$ on $\mathcal{P}$. Combining these observations, one gets

$$\langle \nabla u_h, \nabla(v-\tfrac12 v_h)\rangle = \tfrac12\int_{\mathcal{P}} J_S v_h \;+\; \sum_{S\subset\Omega_i\setminus\mathcal{P}}\int_S J_S(v-\tfrac12 v_h). \tag{4.10}$$

Now, subtracting $\frac12 H^{-1}\mathring{\mathcal{E}}_i^2 = \frac12\int_{\mathcal{P}} J_S^2$ (in view of (4.4)) yields

$$\Psi = \tfrac12\int_{\mathcal{P}} J_S(v_h - J_S) + \sum_{S\subset\Omega_i\setminus\mathcal{P}} J_S\int_S (v-\tfrac12 v_h).$$

So, using (4.4) for the first term, and (4.8a) combined with Remark 4.2 for the second, one gets

$$|\Psi| \lesssim H^{-1/2}\mathring{\mathcal{E}}_i\;\|v_h - J_S\|_{2;\mathcal{P}} + \Big\{\sum_{S\subset\Omega_i\setminus\mathcal{P}}|\omega_S|J_S^2\Big\}^{1/2}\|\partial_y v_h\|_{2;\Omega_i}. \tag{4.11}$$

When dealing with the second term, we also used $|\omega_z^*| \simeq |\omega_z| \simeq |\omega_S|$ for any edge S originating at $z\in\mathcal{P}$. For the first term in (4.11), in view of (4.7), $\|v_h - J_S\|_{2;\mathcal{P}} \lesssim \|\mathrm{osc}(v_h\,;S)\|_{2;\mathcal{P}} \lesssim H^{-1/2}\mathcal{Y}_i$, where the latter bound follows from (4.8c) combined with $\frac{H}{|S|}\gtrsim 1$ and $|\omega_S|\simeq H|S|\;\forall\, S\subset\mathcal{P}$. The second term in (4.11) is bounded by $\mathcal{Y}_i\cdot H^{-1}(\mathring{\mathcal{E}}_i+\mathcal{Y}_i)$, where we used Remark 3.1 and (4.8b). Combining these findings yields the desired bound on Ψ in (4.6).

(iii) To complete the proof, it remains to establish the three bounds on v_h in (4.8). To establish (4.8a), for any $S\subset\gamma_z\setminus\mathcal{P}$ starting at $z=(x_i,y_z)$, let $S' := \mathrm{proj}_{y=y_z}S$, the projection of S onto the line $y=y_z$. Then, by (4.7), $\int_{S'}\hat v_h = \frac12\int_{S'} v_h$. On the other hand, by A3, one has $\big|\int_S v_h - \frac{|S|}{|S'|}\int_{S'} v_h\big| \lesssim \|\partial_y v_h\|_{1;\omega_z^*}$ and a similar bound on $\hat v_h$ (see, e.g., [3, Lemma 7.1]). Combining these observations, and also noting that $\|\partial_y\hat v_h\|_{1;\omega_z^*}\simeq\|\partial_y v_h\|_{1;\omega_z^*}$, yields (4.8a).

For (4.8b), first, note that $v_h\in S_h$ has support in Ω_i, so $\|v_h\|_{2;\Omega_i}^2 \simeq H\|v_h\|_{2;\mathcal{P}}^2 \lesssim H\|J_S\|_{2;\mathcal{P}}^2 = \mathring{\mathcal{E}}_i^2$, where we used (4.7) and then (4.4). Furthermore, $\|\nabla v_h\|_{2;\Omega_i}\lesssim \|\partial_y v_h\|_{2;\Omega_i} + H^{-1}\|v_h\|_{2;\Omega_i}$. So it remains to bound $\|\partial_y v_h\|_{2;\Omega_i}^2$, for which we note that $|\partial_y v_h| = |S|^{-1}\mathrm{osc}(v_h\,;S)$ on any T having an edge $S\subset\mathcal{P}$ (while otherwise $\partial_y v_h=0$). Assuming that (4.8c) is true, one then gets $\|H\,\partial_y v_h\|_{2;\Omega_i}^2\lesssim\mathcal{Y}_i^2$. Combining our findings, we conclude that (4.8b) follows from (4.8c).

Finally, to establish (4.8c), recall (4.3) and combine it with the definition of v_h in (4.7) and the observation that $\sum_{z\in\mathcal{P}_i}\mathcal{Y}_{\omega_z}^2\lesssim\mathcal{Y}_i^2$. □

Remark 4.3 (Non-smooth f) An inspection of the above proof shows that (4.1) remains valid if in $\mathcal{Y}_{\Omega_i}$, which appears in the right-hand side, the term $\|H_T\,\mathrm{osc}(f\,;T)\|_{\Omega_i}$ is replaced by $\|H_T(f-\bar f_i)\|_{\Omega_i} + \|h_T(f-f^I)\|_{\Omega_i}$, where $\bar f_i = \bar f_i(y)$ is an arbitrary function of variable y. To be more precise, the bound (4.9) for ψ_2 can be replaced by

$|\psi_2| = |\langle f - \bar{f}_i, v - \frac{1}{2} v_h \rangle| \lesssim \|f - \bar{f}_i\|_{2;\Omega_i} \|v_h\|_{2;\Omega_i}$. Additionally, we use a sharper version of the bound (4.3), with $H_T \operatorname{osc}(f;T)$ in the right-hand side term $\mathcal{Y}_{\omega_z}$ replaced by $h_T(f - f^I)$. (This version of (4.3) holds true as a similar improvement applies to the bound of Remark 3.1.) Note that if $f \in H^1(\Omega)$, then $\bar{f}_i(y)$ may be chosen equal to a 1d local average of f, and if $f \in L_2(\Omega)$, then $\bar{f}_i(y)$ may be piecewise-constant with local 2d average values, while f^I may be a quasi-interpolant described in Remark 2.2.

5 Estimator efficiency on more general meshes

5.1 Main result

Under the triangulation assumptions of Sect. 3, each anisotropic node shares the local patch orientation with its anisotropic neighbours. So it is not unreasonable to expect that an anisotropic mesh may include small clusters of anisotropic elements sharing the same orientation. In fact, (in particular, if a locally anisotropic mesh was generated starting from some regular mesh) one may assume that the entire anisotropic part of the mesh can be split into sufficiently large, and possibly overlapping, clusters of anisotropic elements with interior cluster diameters $\simeq \operatorname{diam}(\omega_z)$.

Hence, in this section, lower error bounds will be given for small patches of elements surrounding what will be called a local anisotropic path (also see Fig. 2 (left)).

DEFINITION A *local anisotropic path* $\mathcal{P}$ is a simple polygonal curve formed by a subset of short edges, together with their endpoints, that does not touch any corners of Ω, has 2 endpoints (the set of the latter is denoted $\partial\mathcal{P}$), and satisfies the following conditions:

- Any node $z \in \mathcal{P}$ is anisotropic in the sense (3.1) and satisfies $H_z \simeq H_{\mathcal{P}}$ for some $H_{\mathcal{P}}$ associated with $\mathcal{P}$, and also $|\gamma_z \cap \mathcal{P}| \simeq h_z$ (so $\gamma_z \cap \mathcal{P}$ is formed by at most two short edges).
- *Path element orientation condition* There exists a path-specific cartesian coordinate system $(\xi, \eta) = (\xi_{\mathcal{P}}, \eta_{\mathcal{P}})$ such that for any node $z \in \mathcal{P}$, there is a rectangle $\omega_z^* \supset \omega_z$ with sides parallel to the coordinate axes and $|\omega_z^*| \simeq |\omega_z|$.

Theorem 5.1 (Short-edge jump residual terms) *Suppose that $\mathcal{P}_0 \subset \mathcal{P}$, where $\mathcal{P}_0$ and $\mathcal{P}$ are local anisotropic paths that share a coordinate system (ξ, η), and also $\partial\mathcal{P} \cap$*

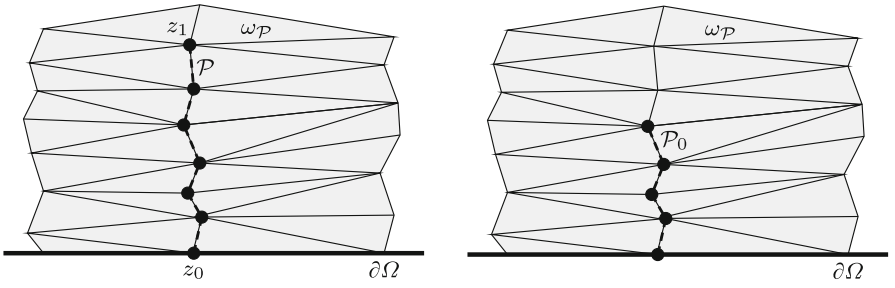

Fig. 2 A local anisotropic path $\mathcal{P}$ with endpoints z_0 and z_1 (left); $\mathcal{P}_0 \subset \mathcal{P}$ from Theorem 5.1 (right)

$\partial\Omega \subset \partial\mathcal{P}_0$ and $\operatorname{dist}(\partial\mathcal{P}\backslash\partial\Omega,\ \partial\mathcal{P}_0\backslash\partial\Omega) \simeq H_{\mathcal{P}} \lesssim |\mathcal{P}|$. *Then for u and u_h satisfying respectively (1.1) and (1.2), with the notation (3.2), one has*

$$\sum_{S\subset\mathcal{P}_0} |\omega_S| J_S^2 \lesssim \mathcal{Y}^2_{\omega_{\mathcal{P}}}, \quad \textit{where} \quad \omega_{\mathcal{P}} := \cup_{z\in\mathcal{P}}\, \omega_z\,. \tag{5.1}$$

The remainder of this section is devoted to the proof of this result.

Corollary 5.2 *Under the conditions of Theorem 5.1, one has*

$$\mathcal{E}_{\omega_{\mathcal{P}_0}} \lesssim \mathcal{Y}_{\omega_{\mathcal{P}}}\,.$$

Proof As (5.1) is equivalent to $\mathring{\mathcal{E}}_{\omega_{\mathcal{P}_0}} \lesssim \mathcal{Y}_{\omega_{\mathcal{P}}}$, combining the latter with Remark 3.1 immediately yields the desired result. □

Remark 5.1 (Estimator efficiency) It follows from [3, §6.1 and Theorem 7.4] that under conditions on the mesh described in Sect. 3 and some additional assumptions on the orientation of anisotropic mesh elements, the error bound (1.4) holds true, i.e. $\|\nabla(u_h-u)\|_\Omega \lesssim \mathcal{E}_\Omega + \|H_T \operatorname{osc}(f;T)\|_\Omega + \|f-f^I\|_\Omega$. Note that for any regular node $z \in \mathcal{N}_{\rm reg}$, (2.2) yields a standard bound $\mathcal{E}_{\omega_z} \lesssim \mathcal{Y}_{\omega_z}$. Now, suppose that all anisotropic nodes in $\mathcal{N}_{\rm ani}\backslash\mathcal{N}_{\rm reg}$ can be split into disjoint sets, each forming a local anisotropic path of type $\mathcal{P}_0$ in Theorem 5.1, and any node in $\mathcal{N}$ belongs to at most a finite number of the respective paths of type $\mathcal{P}$. Then, in view of Corollary 5.2, one gets $\mathcal{E}_\Omega \lesssim \mathcal{Y}_\Omega$, i.e. the error estimator $\mathcal{E}_\Omega$ is efficient up to data oscillation.

Remark 5.2 (Singular perturbation case) Note that the upper a posteriori error bounds in [3] were obtained for more general singularly perturbed semilinear reaction-diffusion equations, solutions of which typically exhibit sharp boundary and interior layers, so anisotropic meshes are frequently employed in their numerical solution. With regard to the lower error bounds for such equations, the standard bubble-function approach was employed in [7], and, as was shown in Sect. 2, the resulting estimates are not sharp even in the regular regime. Sharper lower bounds of type (5.1) will be generalized to this case in a forthcoming paper [5].

5.2 Preliminary results for a local anisotropic path

To prove Theorem 5.1, we shall use a version of Lemma 4.3, in which we shall consider the normalized version of J_S defined by

$$J'_S := J_S|\boldsymbol{\nu}_S\cdot\mathbf{i}_\xi| \quad \forall S\subset\mathcal{P} \quad\Rightarrow\quad J'_S \simeq J_S. \tag{5.2}$$

Here $\mathcal{P}$ is a local anisotropic path associated with the coordinate system (ξ,η), $\mathbf{i}_\xi$ is the unit vector in the ξ-direction, and $\boldsymbol{\nu}_S$ is a unit vector normal to S, while $|\boldsymbol{\nu}_S\cdot\mathbf{i}_\xi| \simeq 1$ follows from S being a short edge and the path element orientation condition. It may be helpful to note that J'_S equals a signed jump of $\partial_\xi u_h$ across S.

Lemma 5.3 *Let $\mathcal{P}$ be a local anisotropic path associated with the coordinate system (ξ, η), and J'_S from (5.2).*
(i) For any node $z \in \mathcal{P}\backslash\partial\mathcal{P}$, with $\gamma_z \cap \mathcal{P}$ formed by two edges S^- and S^+,

$$\left|J'_{S^+} - J'_{S^-}\right| \lesssim h_z H_z^{-1} \sum_{S\in\gamma_z\backslash\mathcal{P}} |J_S|. \tag{5.3}$$

(ii) If $z \in \partial\mathcal{P} \cap \partial\Omega$, with $\gamma_z \cap \mathcal{P}$ formed by a single edge S^+, then (5.3) holds true with J'_{S^-} replaced by 0.

Proof (i) As $z \in \mathcal{N}_{\rm ani}\backslash\partial\Omega$, so $\sum_{S\in\gamma_z} [\![\nabla u_h]\!]_S = 0$, where $[\![\nabla u_h]\!]_S$ denotes the jump in ∇u_h across any edge S in γ_z evaluated in the anticlockwise direction about z. Multiply this relation by the unit vector $\mathbf{i}_\xi$ in the ξ-direction, and note that the quantities $\boldsymbol{\nu}_S \cdot \mathbf{i}_\xi$ for $S = S^\pm$ have opposite signs (in view of the path element orientation condition combined with the maximum angle condition), so $|([\![\nabla u_h]\!]_{S^-} + [\![\nabla u_h]\!]_{S^+}) \cdot \mathbf{i}_\xi| = |J'_{S^+} - J'_{S^-}|$. Note also that for $S \in \gamma_z\backslash\mathcal{P}$, one has $|S| \simeq H_z$ and $|\boldsymbol{\nu}_S \cdot \mathbf{i}_\xi| \lesssim h_z H_z^{-1}$ (again, in view of the path element orientation condition combined with the maximum angle condition), so $|[\![\nabla u_h]\!]_S \cdot \mathbf{i}_\xi| = |J_S\, \boldsymbol{\nu}_S \cdot \mathbf{i}_\xi| \lesssim h_z H_z^{-1} |J_S|$. Combining theses observations yields the desired assertion (5.3).

(ii) Now $z \in \mathcal{N}_{\rm ani} \cap \partial\Omega$, and z is not a corner of $\partial\Omega$. First, suppose that $\mathcal{S}_z \cap \partial\Omega$ is parallel to the ξ-axis. Then extend u_h to $\mathbb{R}^2\backslash\Omega$ by 0 and imitate the above proof with the modification that now $\sum_{S\in\mathcal{S}_z} [\![\nabla u_h]\!]_S = 0$. When dealing with the two edges on $\partial\Omega$, note that for $S \in \mathcal{S}_z \cap \partial\Omega$, one gets $\boldsymbol{\nu}_S \cdot \mathbf{i}_\xi = 0$.

Finally, suppose $\mathcal{S}_z \cap \partial\Omega$ is not parallel to the ξ-axis; then introduce a $\widetilde{\xi}$-axis parallel to $\mathcal{S}_z \cap \partial\Omega$. Now the above argument yields a version of (5.3) with $J'_{S^+} - J'_{S^-}$ replaced by $\widetilde{J}'_S := J_S |\boldsymbol{\nu}_S \cdot \mathbf{i}_{\widetilde{\xi}}|$. The desired result follows as $\widetilde{J}'_S \simeq J_S \simeq J'_S$. The latter follows from $|\boldsymbol{\nu}_S \cdot \mathbf{i}_{\widetilde{\xi}}| \simeq 1$, in view of the path element orientation condition combined with the maximum angle condition. □

Corollary 5.4 *Under the conditions of Lemma 5.3, one has*

$$|\omega_z| \left|\tfrac{H_z}{h_z}(J'_{S^+} - J'_{S^-})\right|^2 \lesssim \mathcal{Y}^2_{\omega_z}, \tag{5.4}$$

where $\mathcal{Y}_{\omega_z}$ is from (3.2a), and if $z \in \mathcal{P} \cap \partial\Omega$, then J_{S^-} in (5.4) is replaced by 0.

Proof Imitate the proof of Corollary 4.4. □

Remark 5.3 Similarly to the case of a partially structured mesh (see Remark 4.2 and Fig. 1 (right)), there is $k \lesssim 1$ such that each rectangle ω_z^* from the above path element orientation condition satisfies $\omega_z^* \cap \omega_{\mathcal{P}} \subset \omega_z^{(k)}$ for all $z \in \mathcal{P}$.

5.3 Proof of Theorem 5.1

We generalize the proof of Theorem 4.1.

Proof Without loss of generality, let $\partial\mathcal{P} = \{z_0, z_1\}$ such that $z_0 \in \partial\Omega$ and $z_1 \notin \partial\Omega$ (see Fig. 2). Also, to simplify the presentation, let the ξ-axis be parallel to $\partial\Omega$ at z_0 (otherwise, see Remark 5.4).

Set $H := H_{\mathcal{P}} \simeq H_{\mathcal{P}_0}$. A certain weight $\rho_S \in [0, 1]$ will be associated with each $S \subset \mathcal{P}$, and it will be imposed that $\rho_S = 1 \; \forall\, S \subset \mathcal{P}_0$. Hence, it suffices to prove that

$$\widetilde{\mathcal{E}}^2_{\mathcal{P}} := \sum_{S\subset\mathcal{P}} H|S|\,\rho_S J_S J'_S \lesssim \mathcal{Y}^2_{\omega_{\mathcal{P}}}\,, \tag{5.5}$$

where J'_S is from (5.2). Then, indeed, in view of $H|S| \simeq |\omega_S|$ and $J'_S \simeq J_S$, (5.5) immediately implies the desired assertion (5.1).

Next, note that for any $v \in H^1_0(\Omega)$ and $v_h \in S_h$, a standard calculation using (1.1), (1.2) yields

$$\begin{aligned}\underbrace{\langle\nabla(u_h - u), \nabla v\rangle}_{=:\psi_1} &= \langle\nabla u_h, \nabla v\rangle - \langle f, v\rangle \\ &= \underbrace{\langle\nabla u_h, \nabla(v - \tfrac12 v_h)\rangle}_{=:\Psi + \frac12 H^{-1}\widetilde{\mathcal{E}}^2_{\mathcal{P}}} - \underbrace{\langle f, v - \tfrac12 v_h\rangle}_{=:\psi_2}\,.\end{aligned} \tag{5.6}$$

As this immediately implies $\widetilde{\mathcal{E}}^2_{\mathcal{P}} \lesssim H(|\psi_1| + |\psi_2| + |\Psi|)$, to get (5.5) (and hence the desired assertion (5.1)), it suffices to prove that

$$H\big(|\psi_1| + |\psi_2| + |\Psi|\big) \lesssim \mathcal{Y}_{\omega_{\mathcal{P}}}\,(\widetilde{\mathcal{E}}_{\mathcal{P}} + \mathcal{Y}_{\omega_{\mathcal{P}}}). \tag{5.7}$$

The remainder of the proof is split into three parts. In part (i), we shall describe appropriate weights $\{\rho_S\}$ and non-standard functions v_h and v, which will be crucial for (5.7) to hold true. Certain sufficient conditions for the latter will be established in part (ii), and then shown to be satisfied in part (iii).

(i) We start by introducing a smooth monotone cut-off function ρ of the arc-length parameter l of $\mathcal{P}$ such that

$$\rho = 1 \;\text{ on } \mathcal{P}_0, \quad \rho = 0 \;\text{ on } \partial\mathcal{P}\backslash\partial\Omega, \quad |\rho'(l)| \lesssim H^{-1}\sqrt{\rho(l)} \;\; \forall l. \tag{5.8}$$

Here for the final relation, recall that $\operatorname{dist}(\partial\mathcal{P}\backslash\partial\Omega, \partial\mathcal{P}_0\backslash\partial\Omega) \simeq H$ and let ρ be quadratic near its zeros.

Next, introduce

$$\rho_S := \tfrac12\sum_{z\in\partial S}\rho(z), \quad \delta_S := \operatorname{osc}(J'_S; \mathcal{P}_S) \quad \forall\, S \subset \mathcal{P}, \tag{5.9}$$

where $J'_S = J_S|\boldsymbol{\nu}_S \cdot \mathbf{i}_\xi|$ is from (5.2) (and also appears in (5.5)), and $\mathcal{P}_S$ denotes the patch of (at most three) edges in $\mathcal{P}$ touching S (so $S \subset \mathcal{P}_S \subset \mathcal{P}$).

Finally, in (5.6), we let $v_h \in S_h$, with support in $\omega_{\mathcal{P}}$ (and with the convention that $v_h = 0$ in $\mathbb{R}\backslash\Omega$), and $v := \hat{v}_h \in H_0^1(\Omega)$ satisfy

$$v_h(z) := \tfrac{1}{2}\rho(z)\sum_{S\in\gamma_z\cap\mathcal{P}} J'_S \quad \forall\, z \in \mathcal{P}\backslash\partial\Omega, \qquad \hat{v}_h(\xi,\eta) := v_h\big(\bar{\xi}_{\mathcal{P}}(\eta) + 2[\xi - \bar{\xi}_{\mathcal{P}}(\eta)], \eta\big). \tag{5.10}$$

Here the function $\bar{\xi} = \bar{\xi}_{\mathcal{P}}(\eta) \in C(\mathbb{R})$ describes the curve $\mathcal{P}$ for the range of η in $\mathcal{P}$, and is constant outside this range. Without loss of generality, $\hat{v}_h$ has support in $\omega_{\mathcal{P}}$. (Otherwise, in view of Remark 5.3, shorten the support of ρ by k short edges starting from z_1, where $k \lesssim 1$.)

For $\bar{\xi}_{\mathcal{P}}(\eta)$ in (5.10), note that $|\bar{\xi}'_{\mathcal{P}}| \lesssim 1$ (in view of the path element orientation condition combined with the maximum angle condition). This observation implies that $\hat{v}_h$ is well-defined in Ω, and $\|\nabla\hat{v}_h\|_{2;\Omega} \simeq \|\nabla v_h\|_{2;\omega_{\mathcal{P}}}$, as well as $\|\hat{v}_h\|_{2;\Omega} \simeq \|v_h\|_{2;\omega_{\mathcal{P}}}$.

Note also a few useful properties, which follow from (5.9) and (5.10):

$$\Big|\int_S (v_h - \rho_S J'_S)\Big| \le |S|\rho_S\delta_S, \tag{5.11a}$$

$$\sup_S |v_h| \lesssim \sup_{S\subset\mathcal{P}_S} |\rho_S J'_S|, \tag{5.11b}$$

$$\operatorname{osc}(v_h\,;S) \lesssim \tfrac{|S|}{H}\sqrt{\rho_S}\,|J_S| + \delta_S \quad \forall\, S \subset \mathcal{P}. \tag{5.11c}$$

To check (5.11a), note that v_h is linear on $S \subset \mathcal{P}$, so $|S|^{-1}\int_S v_h$ is between $\rho_S \min_{\mathcal{P}_S}\{J'_S\}$ and $\rho_S \max_{\mathcal{P}_S}\{J'_S\}$, so this assertion follows. For (5.11b), we note that $\frac{1}{2}\rho(z) \le \rho_S$ for any $S \in \gamma_z \cap \mathcal{P}$, so $|v_h(z)| \le \sum_{S\in\gamma_z\cap\mathcal{P}} |\rho_S J'_S|$. Finally, $\forall\, z \in \partial S$, where $S \subset \mathcal{P}$, one has $|v_h(z) - \rho(z)J'_S| \le \rho(z)\delta_S \le \delta_S$, so $\operatorname{osc}(v_h\,;S) \le \operatorname{osc}(\rho\,;S)|J'_S| + 2\delta_S$. Here $|J'_S| \le |J_S|$, while the final relationship in (5.8) yields $\operatorname{osc}(\rho\,;S) \lesssim \frac{|S|}{H}\sqrt{\rho_S}$ (where we also used $\sup_S \rho \le 2\rho_S$ as $\rho(l)$ is monotone).

(ii) We claim that for (5.7), and hence for the desired assertion (5.1), it suffices to prove that the following conditions (which give a version of (4.8)) are satisfied:

$$\Big|\int_S (\hat{v}_h - \tfrac{1}{2}v_h)\Big| \lesssim \|\nabla v_h\|_{1;\omega_z^*\cap\omega_{\mathcal{P}}} \quad \forall\, S \in \gamma_z\backslash\mathcal{P},\ z \in \mathcal{P}, \tag{5.12a}$$

$$\|H\nabla v_h\|_{2;\omega_{\mathcal{P}}} + \|v_h\|_{2;\omega_{\mathcal{P}}} \lesssim \widetilde{\mathcal{E}}_{\mathcal{P}} + \mathcal{Y}_{\omega_{\mathcal{P}}}, \tag{5.12b}$$

$$\sum_{S\subset\mathcal{P}} |\omega_S|\Big\{\tfrac{H}{|S|}\delta_S\Big\}^2 \lesssim \mathcal{Y}^2_{\omega_{\mathcal{P}}}. \tag{5.12c}$$

Note that ψ_1 and ψ_2 are shown to satisfy (5.7) using (5.12) in a very similar manner to the corresponding bounds in part (ii) of the proof of Theorem 4.1, only for ψ_2 we now employ $\hat{f} := f\big(\bar{\xi}_{\mathcal{P}}(\eta) + 2[\xi - \bar{\xi}_{\mathcal{P}}(\eta)], \eta\big)$ and then Remark 5.3.

To show that Ψ also satisfies (5.7), first, we get a version of (4.10) with Ω_i replaced by $\omega_{\mathcal{P}}$. Next, subtracting $\frac{1}{2}H^{-1}\widetilde{\mathcal{E}}^2_{\mathcal{P}} = \frac{1}{2}\sum_{S\subset\mathcal{P}} |S|\rho_S J_S J'_S$ (in view of the definition of $\widetilde{\mathcal{E}}_{\mathcal{P}}$ in (5.5)) yields

$$\Psi = \sum_{S\subset\mathcal{P}} \tfrac{1}{2} J_S \underbrace{\int_S (v_h - \rho_S J'_S)}_{\le |S|\rho_S\delta_S \text{ by (5.11}a\text{)}} + \sum_{S\subset\omega_\mathcal{P}\setminus\mathcal{P}} J_S \int_S (v - \tfrac{1}{2} v_h).$$

So, using the definition of $\widetilde{\mathcal{E}}_\mathcal{P}$ combined with $J_S \simeq J'_S$ for the first term, and (5.12a) combined with Remark 5.3 for the second, one gets

$$|\Psi| \lesssim H^{-1/2}\widetilde{\mathcal{E}}_\mathcal{P}\ \|\delta_S\|_{2;\mathcal{P}} + \Big\{ \sum_{S\subset\omega_\mathcal{P}\setminus\mathcal{P}} |\omega_S| J_S^2 \Big\}^{1/2} \|\nabla v_h\|_{2;\omega_\mathcal{P}}\,. \tag{5.13}$$

When dealing with the second term, we also used $|\omega_z^*| \simeq |\omega_z| \simeq |\omega_S|$ for any edge S originating at $z \in \mathcal{P}$. For the first term in (5.13), $\|\delta_S\|_{2;\mathcal{P}} \lesssim H^{-1/2}\mathcal{Y}_{\omega_\mathcal{P}}$, which follows from (5.12c) combined with $\frac{H}{|S|} \gtrsim 1$ and $|\omega_S| \simeq H|S|\ \forall\, S \subset \mathcal{P}$. The second term in (5.13) is bounded by $\mathcal{Y}_{\omega_\mathcal{P}} \cdot H^{-1}(\widetilde{\mathcal{E}}_\mathcal{P} + \mathcal{Y}_{\omega_\mathcal{P}})$, where we used Remark 3.1 and (5.12b). Combining these findings yields the desired bound on Ψ in (5.7).

(iii) To complete the proof, it remains to establish the three bounds on v_h in (5.12). The first bound (5.12a) is obtained similarly to (4.8a). Only now for any $S \subset \gamma_z\setminus\mathcal{P}$ starting at $z = (\xi_z, \eta_z)$, we use $S' := \mathrm{proj}_{\eta=\eta_z} S$, the projection of S (or, possibly, its sufficiently long extension) onto the line $\eta = \eta_z$, and also $\|\partial_\eta \hat{v}_h\|_{1;\omega_z^*} \lesssim \|\nabla v_h\|_{1;\omega_z^*} = \|\nabla v_h\|_{1;\omega_z^*\cap\omega_\mathcal{P}}$.

For (5.12b), first, note that $v_h \in S_h$ with support in $\omega_\mathcal{P}$, so $\|v_h\|^2_{2;\omega_\mathcal{P}} \simeq H\|v_h\|^2_{2;\mathcal{P}} \lesssim H\|\rho_S J'_S\|^2_{2;\mathcal{P}} \le \widetilde{\mathcal{E}}^2_\mathcal{P}$, where we used (5.11b) and also the definition of $\widetilde{\mathcal{E}}_\mathcal{P}$ in (5.5). Furthermore, on any $S \subset \mathcal{P}$ one has $|H\,\partial_l u_h| = \frac{H}{|S|}\mathrm{osc}(v_h\,;S)$, so

$$\big\|H\nabla v_h\big\|^2_{2;\omega_\mathcal{P}} \lesssim H\ \Big\|\underbrace{\tfrac{H}{|S|}\mathrm{osc}(v_h\,;S)}_{\lesssim\sqrt{\rho_S}|J_S|+\frac{H}{|S|}\delta_S \text{ by (5.11}c\text{)}}\Big\|^2_{2;\mathcal{P}} + \big\|v_h\big\|^2_{2;\omega_\mathcal{P}}.$$

Here $H\|\sqrt{\rho_S}J_S\|^2_{2;\mathcal{P}} \lesssim \widetilde{\mathcal{E}}_\mathcal{P}$, while $H\|\frac{H}{|S|}\delta_S\|^2_{2;\mathcal{P}} \lesssim \mathcal{Y}^2_{\omega_\mathcal{P}}$ assuming that (5.12c) is true. Combining our findings, we conclude that (5.12b) follows from (5.12c).

Finally, (5.12c) is obtained similarly to (4.8c). To be more precise we recall (5.4) and combine it with the definition of δ_S in (5.9) and the observation that $\sum_{T\subset\omega_\mathcal{P}} \mathcal{Y}_T^2 \lesssim \mathcal{Y}^2_{\omega_\mathcal{P}}$. □

Remark 5.4 If in the proof of Theorem 5.1 $z_0 \in \partial\Omega \cap \partial\mathcal{P}$ is such that the ξ-axis is not parallel to $\partial\Omega$ at z_0, then one needs to tweak the definition of v_h so that its support is in $\omega_\mathcal{P}\setminus\omega^*_{z_0}$ (rather than in $\omega_\mathcal{P}$). This modification is required to ensure that $\hat{v}_h$ has support in $\omega_\mathcal{P}$. For this, ρ remains unchanged (i.e. equal to 1 on $\mathcal{P}$ near $\partial\Omega$), while we now set $v_h(z) := 0$ for any $z \in \mathcal{P} \cap \omega^*_{z_0}$. Note that the evaluations will remain without changes, except that the proof of (5.12c) is now slightly more delicate. To be more precise, as $\mathcal{P} \cap \omega^*_{z_0}$ includes a finite number of edges (in view of Remark 5.3), so $\mathrm{osc}(v_h\,;S)$ for the edge $S \subset \mathcal{P}\setminus\omega^*_{z_0}$ closest to $\partial\Omega$ will involve $\mathrm{osc}(J'_S\,;\mathcal{P}\cap\omega^*_{z_0})$, the estimation of which will require a finite number of applications of (5.4).

6 Conclusion

We have reviewed lower a posteriori error bounds obtained using the standard bubble function approach in the context of anisotropic meshes. Numerical examples have been given in Sect. 2 that clearly demonstrate that the short-edge jump residual terms in such bounds are not sharp. Hence, in Sects. 4 and 5, for linear finite element approximations of the Laplace equation in polygonal domains, a new approach has been presented that yields essentially sharper lower a posteriori error bounds and thus shows that the upper error estimator (1.4) from the recent paper [3] is efficient on partially structured anisotropic meshes.

References

1. Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis. Wiley, New York (2000)
2. Kopteva, N.: Maximum-norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes. SIAM J. Numer. Anal. **53**, 2519–2544 (2015)
3. Kopteva, N.: Energy-norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes. Numer. Math. **137**, 607–642 (2017)
4. Kopteva, N.: Fully computable a posteriori error estimator using anisotropic flux equilibration on anisotropic meshes. arXiv:1704.04404 (2017)
5. Kopteva, N.: Lower a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes (in preparation) (2020)
6. Kunert, G.: An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes. Numer. Math. **86**, 471–490 (2000)
7. Kunert, G.: Robust a posteriori error estimation for a singularly perturbed reaction-diffusion equation on anisotropic tetrahedral meshes. Adv. Comput. Math. **15**, 237–259 (2001)
8. Kunert, G., Verfürth, R.: Edge residuals dominate a posteriori error estimates for linear finite element methods on anisotropic triangular and tetrahedral meshes. Numer. Math. **86**, 283–303 (2000)
9. Micheletti, S., Perotto, S.: Reliability and efficiency of an anisotropic Zienkiewicz-Zhu error estimator. Comput. Methods Appl. Mech. Eng. **195**, 799–835 (2006)
10. Picasso, M.: Numerical study of the effectivity index for an anisotropic error indicator based on Zienkiewicz-Zhu error estimator. Commun. Numer. Methods Eng. **19**, 13–23 (2003)
11. Verfürth, R.: A Posteriori Error Estimation Techniques for Finite Element Methods. Oxford University Press, Oxford (2013)
12. Xu, J., Zhang, Z.: Analysis of recovery type a posteriori error estimators for mildly structured grids. Math. Comput. **73**, 1139–1152 (2004)