Technische Universität Chemnitz
Sonderforschungsbereich 393
Numerische Simulation auf massiv parallelen Rechnern

Gerd Kunert

Robust local problem error estimation for a singularly perturbed problem on anisotropic finite element meshes

Preprint SFB393/01-01

Abstract

Singularly perturbed problems often yield solutions with strong directional features, e.g. with boundary layers. Such anisotropic solutions lend themselves to adapted, anisotropic discretizations. The quality of the corresponding numerical solution is a key issue in any computational simulation.

To this end we present a new robust error estimator for a singularly perturbed reaction–diffusion problem. In contrast to conventional estimators, our proposal is suitable for *anisotropic* finite element meshes. The estimator is based on the solution of a local problem, and yields error bounds uniformly in the small perturbation parameter. The error estimation is efficient, i.e. a lower error bound holds. The error estimator is also reliable, i.e. an upper error bound holds, provided that the anisotropic mesh discretizes the problem sufficiently well.

A numerical example supports the analysis of our anisotropic error estimator.

Keywords: error estimator, anisotropic solution, stretched elements, reaction diffusion equation, singularly perturbed problem

AMS: 65N15, 65N30, 35B25

Preprint-Reihe des Chemnitzer SFB 393

SFB393/01-01 **January 2001**

Contents

Gerd Kunert
TU Chemnitz
Fakultät für Mathematik
09107 Chemnitz
Germany

http://www.tu-chemnitz.de/∼gku
http://www.tu-chemnitz.de/sfb393/

1 Introduction

Adaptive algorithms form nowadays an indispensable tool for most finite element simulations. They basically consist of the ingredients *Solve – Estimate error – Refine mesh* which are repeated until the desired accuracy is achieved, see also [2, 22]. The present work is part of a series of related endeavors in a particular field of finite element analysis. While standard finite element meshes employ *isotropic* (or shape regular) elements, we investigate so–called *anisotropic* (or stretched) elements. They are characterized by a large stretching ratio (also called aspect ratio). Equivalently, the ratio of the diameters of the circumscribed and inscribed spheres can be arbitrarily large. Such anisotropic meshes are particularly useful when the differential equation gives rise to a solution with strong directional features, such as boundary layers or interior layers. Application of anisotropic meshes as well as theoretical investigations can be found for example in [4, 5, 18, 19, 24, 25] and [11, 13, 14, 17, 20].

It is a natural desire to incorporate anisotropic meshes into adaptive algorithms. Clearly, additional ingredients are required then, namely *anisotropic information extraction* (e.g. find the (quasi) optimal stretching direction and stretching ratio of the anisotropic elements), and *anisotropic mesh refinement*. Less obvious but equally important is the *error estimation* part. Unfortunately most of the conventional *a posteriori* error estimators for isotropic meshes fail when applied on anisotropic meshes. Therefore the derivation and analysis of estimators which are suitable for anisotropic elements is of vital importance for any adaptive anisotropic algorithm.

Fortunately this challenging venture has seen some success recently [20, 13, 14, 15, 17, 11]. In case studies of the Poisson model problem it has been shown that anisotropic error estimation is possible, and the methodology and analytical tools have been developed, proposed and refined. Now anisotropic error estimation has to prove its potential for more realistic settings. Singularly perturbed problems offer ideal test fields since they often induce boundary layers where anisotropic elements can be employed favourably.

From now on let us consider a singularly perturbed reaction–diffusion model problem, see (1) below, which usually gives rise to boundary layers whenever a non–vanishing right hand side meets homogeneous Dirichlet boundary data.

Although (1) forms a comparatively simple model problem, the knowledge of *robust* error estimators has been unsatisfactory for a long time. The first estimators with error bounds that are uniform in the small perturbation parameter ε were due to Angermann [3], Verfürth [23] and Ainsworth/Babuška [1]; all of them considered isotropic meshes. Angermann measures the error in a somewhat strange norm (which seems to be mainly of theoretical interest) whereas Verfürth and Ainsworth/Babuška concentrate on the energy norm (which is the most natural norm). For *anisotropic* meshes Kunert [16] recently succeeded in deriving a robust residual error estimator, also for the energy norm. A corollary of that result has been included in [17].

In our present work we propose a new error estimator for the singularly perturbed reaction–diffusion problem (1) which is suitable for anisotropic meshes, and that is based on the solution of a local problem. The roots of this local problem error estimator are twofold. Firstly it relies on the anisotropic residual error estimator [16] whose results are partly the foundation for the present analysis. Secondly we utilize the methodology of local problem error estimation. For the Poisson problem (on isotropic meshes) this is fairly well understood, see e.g. the exposition in [22]; hence the general *framework* of the proofs can be derived relatively easily. The precise definition and analysis of our estimator, however, are much more difficult and technical. This concerns for example the choice of the local problem, the careful calibration of all ingredients, or certain equivalence lemmas. Although we could exploit some experience from anisotropic local problem error estimation for the Poisson problem [15], the 'extension' to the singularly perturbed problem requires several new ingredients and is by no means straight–forward. Note that in [1] also a local problem error estimator is derived (for isotropic meshes). However the local problem there is infinite dimensional whereas our proposal here involves an (at most) five dimensional local space.

When comparing with the anisotropic residual error estimator, our newly proposed local problem error estimator is certainly more expensive since a local problem has to be computed and solved. Nonetheless the disadvantage of any residual based estimator is that the proof of the error bound is based on several intermediate steps, such as interpolation estimates and the Cauchy Schwarz inequality. In contrast to this the local problem error estimator requires less auxiliary steps, and thus contains less constants (which are unknown in general). This can also be observed numerically: The qualitative behaviour of both error estimators is comparable but the local problem error estimation is much closer to the true error.

Finally note that all known anisotropic error estimators require that the anisotropy of the mesh and the anisotropy of the solution correspond sufficiently well. As in previous work, this correspondence is measured by a so–called matching function which is explained in our exposition.

The remainder of the paper is organized as follows. After presenting the model problem in §2, we repeat in §3 some notation, basic tools and lemmas that have been applied successfully in previous anisotropic investigations. The transformation technique and several bubble functions are of particular importance. Furthermore the residual error estimator of [16] is recalled for self–containment. Next, §4 is devoted to the local problem error estimator and its analysis. Reliable upper and lower error bounds are proven and a stable basis of the local problem is presented. Additionally a further, face oriented local problem error estimator is given. Computational aspects are discussed in §5, and the numerical experiments of §6 confirm the analysis. The summary in §7 and a technical proof in appendix A conclude this work.

2 The model problem and its discretization

Our focus is on a singularly perturbed reaction–diffusion model problem with Dirichlet–Neumann boundary conditions

$$\left.\begin{aligned} -\varepsilon\Delta u + u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \Gamma_{\mathrm{D}} \\ \varepsilon \cdot \partial u/\partial n &= g \quad \text{on } \Gamma_{\mathrm{N}} \end{aligned}\right\} \tag{1}$$

in a bounded, polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with boundary $\partial\Omega = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}}$.

Assume $f \in L_2(\Omega)$, $g \in L_2(\Gamma_{\mathrm{N}})$ and $\mathrm{meas}_{d-1}(\Gamma_{\mathrm{D}}) > 0$. The Sobolev space of functions that vanish on Γ_{D} is denoted by $H^1_o(\Omega)$ as usual. The corresponding variational formulation for (1) becomes:

$$\left.\begin{array}{llll} \text{Find } u \in H^1_o(\Omega): & a(u,v) \;=\; \langle \mathfrak{f}, v\rangle & \forall\, v \in H^1_o(\Omega) & \\ \text{with} \quad a(u,v) \;:= \displaystyle\int_\Omega \varepsilon \cdot (\nabla u)^\top \nabla v \;+\; u\, v & & \langle \mathfrak{f}, v\rangle := \displaystyle\int_\Omega f v + \int_{\Gamma_{\mathrm{N}}} g v & . \end{array}\right\} \tag{2}$$

We utilize a family $\mathcal{F} = \{\mathcal{T}_h\}$ of triangulations $\mathcal{T}_h$ of Ω. Let $V_{o,h} \subset H^1_o(\Omega)$ be the space of continuous, piecewise linear functions over $\mathcal{T}_h$ that vanish on Γ_{D}. Then the finite element solution $u_h \in V_{o,h}$ is uniquely defined by

$$a(u_h, v_h) \;=\; \langle \mathfrak{f}, v_h\rangle \qquad \forall\, v_h \in V_{o,h} \quad . \tag{3}$$

Due to the Lax–Milgram Lemma both problems (2) and (3) admit unique solutions.

The main purpose of our analysis is to bound the error $u - u_h$ uniformly in the small perturbation parameter ε. Here we concentrate on the most natural norm related to (2), namely the *energy norm*

$$|||v|||^2 := a(v,v) = \varepsilon\|\nabla v\|^2 \;+\; \|v\|^2$$

which has been used also by other authors [1, 23]. This energy norm is well–suited to produce appropriately refined meshes. This can be easily verified on some 1D model problem, e.g. for $-\varepsilon u'' + u = 0$ in $\Omega = (0,1)$ with $u(0) = 1, u(1) = 0$. Even the optimal order of convergence can be achieved.

3 Notation, basic tools and lemmas

In order to analyse error estimators on anisotropic meshes we will now introduce certain notation as well as important tools, all of which have proven to be advantageous in previous work [13, 15, 16]. All expositions are given for the more technical three dimensional case. The application to the simpler 2D case is readily possible.

From now on, let $\mathbb{P}^k(\omega)$ be the space of polynomials of order k at most over some domain $\omega \subset \mathbb{R}^3$ or $\omega \subset \mathbb{R}^2$. Instead of $x \le c \cdot y$ or $c_1 x \le y \le c_2 x$ (with positive constants independent of x, y and $\varepsilon, \mathcal{T}_h$) we use the abbreviation $x \lesssim y$ and $x \sim y$, respectively. By $\|\cdot\|_\omega$ we denote the L_2 norm of a function over some domain ω. For $\omega = \Omega$ the subscript is omitted. Let $|\omega| := \text{meas}\,(\omega)$ be the measure of a domain ω. Finally for some vector $\mathbf{p}$ let $|\mathbf{p}| := \sqrt{\mathbf{p}^\top \mathbf{p}}$ be its Euclidean norm (i.e. length).

3.1 Tetrahedron – Subdomains – Mesh requirements

Tetrahedron: Let a triangulation $\mathcal{T}_h$ be given which satisfies the usual admissibility conditions (see Ciarlet [10], Chapter 2). The four vertices of an arbitrary tetrahedron $T \in \mathcal{T}_h$ are denoted by $P_0, \ldots, P_3$ such that P_0P_1 is the longest edge of T, $\text{meas}_2(\triangle P_0P_1P_2) \geq \text{meas}_2(\triangle P_0P_1P_3)$, and $\text{meas}_1(P_1P_2) \geq \text{meas}_1(P_0P_2)$.

Additionally define three pairwise orthogonal vectors $\mathbf{p}_i$ with lengths $h_{i,T} := |\mathbf{p}_i|$, see figure 1. Observe $h_{1,T} > h_{2,T} \geq h_{3,T}$ and set $h_{min,T} := h_{3,T}$. The circumscribed hexahedron may facilitate the visualization.

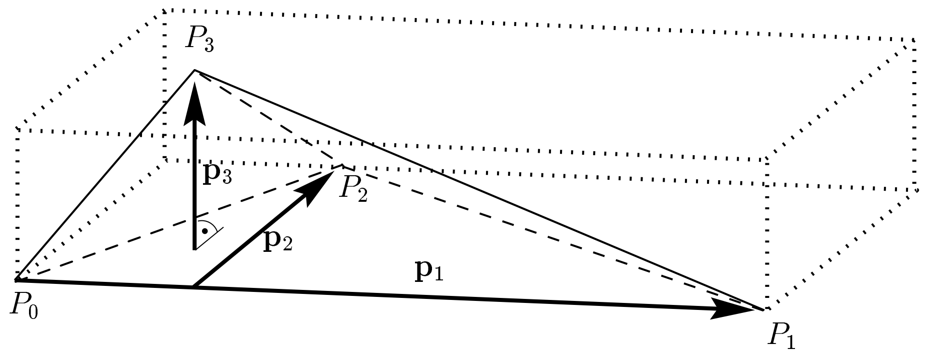

Figure 1: Notation of tetrahedron T

Tetrahedra are denoted by T, T', T'' or T_i. Faces of a tetrahedron are denoted by E or E_i. Let $h_{E,T} := 3|T|/|E|$ be the length of the *height* over a face E.

Auxiliary subdomains: Let $T \in \mathcal{T}_h$ be an arbitrary tetrahedron. Let ω_T be that domain that is formed by T and all tetrahedra that have a common face with T. Note that ω_T consists of less than five tetrahedra if T has a boundary face.

Let E be an inner face (triangle) of $\mathcal{T}_h$, i.e. there are two tetrahedra T_1 and T_2 having the common face E. Set the domain $\omega_E := T_1 \cup T_2$. If E is a boundary face set $\omega_E := T$ with $T \supset E$.

Mesh requirements: In addition to the usual conformity conditions of the mesh (see Ciarlet [10], Chapter 2) we demand the following two assumptions.

1. The number of tetrahedra containing a node x_j is bounded uniformly.
2. The dimensions of adjacent tetrahedra must not change rapidly, i.e.

$$h_{i,T'} \sim h_{i,T} \qquad \forall\, T, T' \text{ with } T \cap T' \neq \emptyset\,,\; i = 1 \ldots d \quad .$$

Remark 1 In certain situations we do not want to use *element based* quantities (such as $h_{min,T}$) but utilize *face related* terms instead. For example consider an interior face $E = T_1 \cap T_2$, and define the terms

$$h_E := (h_{E,T_1} + h_{E,T_2})/2 \qquad , \qquad h_{min,E} := (h_{min,T_1} + h_{min,T_2})/2 \quad .$$

Their advantage is that they are no longer related to T_1 or T_2 but to E. They clearly satisfy $h_E \sim h_{E,T_i}$ and $h_{min,E} \sim h_{min,T_i}$. For a boundary face $E \subset \partial T \cap \Gamma$ define similarly $h_E := h_{E,T}$ and $h_{min,E} := h_{min,T}$. ■

Transformations: The usual transformation technique between a tetrahedron T and a standard tetrahedron plays a vital role in many proofs (cf. [10]). However, a refined analysis has shown that *two different transformations* facilitate matters considerably, see e.g. [13, 14]. Hence define the matrices $H_T := \mathrm{diag}(h_{1,T}, h_{2,T}, h_{3,T})$ and $A_T, C_T \in \mathbb{R}^{3\times 3}$ by

$$A_T := (\ \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \overrightarrow{P_0P_3}\ ) \qquad \text{and} \qquad C_T := (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) \qquad , \tag{4}$$

and introduce affine linear mappings

$$F_A(\mu) := A_T \cdot \mu + \vec{P_0} \qquad \text{and} \qquad F_C(\mu) := C_T \cdot \mu + \vec{P_0} \quad , \ \mu \in \mathbb{R}^3 .$$

These mappings implicitly define the *standard tetrahedron* $\bar{T} := F_A^{-1}(T)$ and the *reference tetrahedron* $\hat{T} := F_C^{-1}(T)$. Then $\bar{T}$ the has vertices $\bar{P}_0 = (0,0,0)^\top$ and $\bar{P}_i = \mathbf{e}_i^\top, i = 1 \ldots 3$, whereas $\hat{T}$ has vertices at $\hat{P}_0 = (0,0,0)^\top$, $\hat{P}_1 = (1,0,0)^\top$, $\hat{P}_2 = (\hat{x}_2, 1, 0)^\top$ and $\hat{P}_3 = (\hat{x}_3, \hat{y}_3, 1)^\top$. The conditions on the P_i yield immediately $0 < \hat{x}_2 \le 1/2$, $0 < \hat{x}_3 < 1$ and $-1 < \hat{y}_3 < 1$. Figures 1 and 2 illustrate this definition.

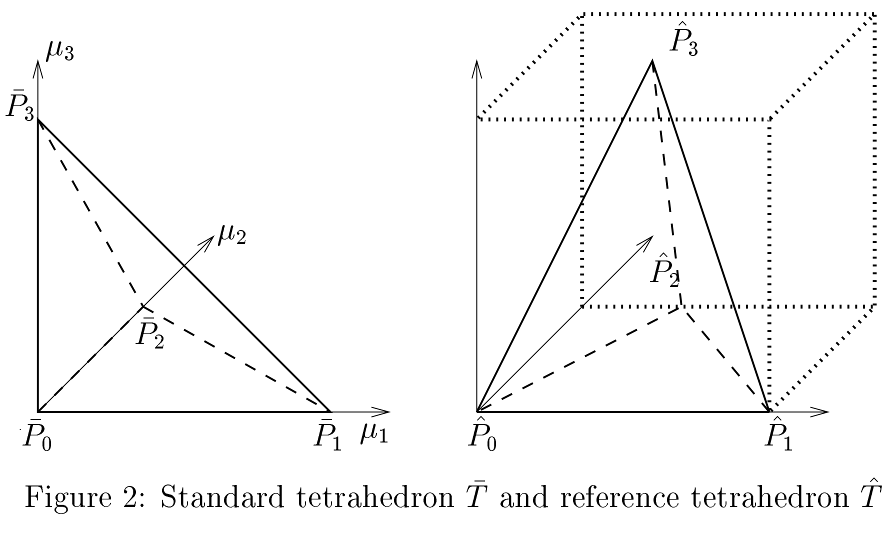

Figure 2: Standard tetrahedron $\bar{T}$ and reference tetrahedron $\hat{T}$

Variables and operators that are related to the standard tetrahedron $\bar{T}$ and the reference tetrahedron $\hat{T}$ are referred to with a *bar* and a *hat*, respectively (e.g. $\bar{\nabla}$, $\hat{v}$). The determinants of both mappings are $|\det(A_T)| = |\det(C_T)| = 6|T|$, and the transformed derivatives satisfy $\bar{\nabla}\bar{v} = A_T^\top \nabla v$ and $\hat{\nabla}\hat{v} = C_T^\top \nabla v$.

Although C_T is naturally associated with our analysis, it transforms $\hat{T}$ into T. Inequality constants would thus depend on $\hat{T}$. To overcome this drawback, the transformation via A_T is used in conjunction with C_T (cf. the compactness arguments in the proof of Lemma 5).

Finally, $H_T^{-1} C_T^\top$ is orthogonal since $C_T^\top \cdot C_T = H_T^2$. Hence

$$\|H_T^{-1} C_T^\top \nabla v\|_T \;=\; \|\nabla v\|_T \qquad . \tag{5}$$

Squeezed tetrahedron $T_{E,\delta}$: The concept of the squeezed tetrahedron has been introduced in [16] and originates from [23] (in a simpler, modified form there). Here we repeat the definition and only state the required results.

Because of the singular perturbation character of the differential equation we can favourably employ a sub–tetrahedron $T_{E,\delta} \subset T$ which depends on a face E of T and a real number $\delta \in (0, 1]$. In an attempt to use a vivid name we will refer to $T_{E,\delta}$ as a *squeezed tetrahedron*. For its precise definition, let T be an arbitrary but fixed tetrahedron, and enumerate temporarily its vertices such that $E = Q_1Q_2Q_3$ and $T = OQ_1Q_2Q_3$, cf. Figure 3. Introduce barycentric coordinates such that λ_0 is related to O, and λ_1, λ_2, λ_3 correspond to Q_1, Q_2, Q_3, respectively.

Let P be that point with barycentric coordinates

$$\lambda_0(P) = \delta \qquad \text{and} \qquad \lambda_1(P) = \lambda_2(P) = \lambda_3(P) = \frac{1-\delta}{3} \qquad .$$

Then let $T_{E,\delta}$ be the tetrahedron with vertices P and Q_1, Q_2, Q_3, i.e. $T_{E,\delta}$ has the same face E as T but the fourth vertex is moved towards E with the rate δ.

An alternative description is as follows. With S_E being the midpoint (i.e. center of gravity) of face E, point P lies on the line S_EO such that $|\vec{S_EP}| = \delta \cdot |\vec{S_EO}|$. Note that for $\delta = 1$ one gets $T_{E,\delta} \equiv T$ whereas in the limiting case $\delta \to 0$ the tetrahedron $T_{E,\delta}$ collapses to the face E.

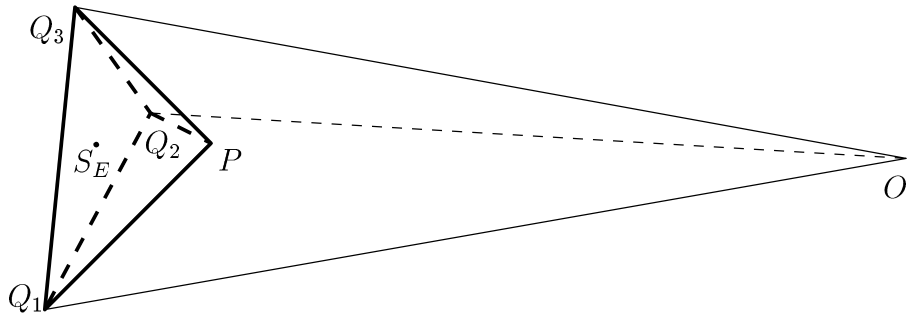

Figure 3: Tetrahedra $T = OQ_1Q_2Q_3$ and $T_{E,\delta} = PQ_1Q_2Q_3$

In order to utilize $T_{E,\delta}$ efficiently, we also require an affine linear transformation $F_{T,E,\delta}$ that maps the standard tetrahedron $\bar{T}$ onto $T_{E,\delta}$. This affine linear mapping is unique (up to permutations of the enumeration of the vertices of $\bar{T}$ and $T_{E,\delta}$). In [16] the following relations has been proven.

Lemma 1 *The radius $\varrho(T_{E,\delta})$ of the largest inscribed sphere of $T_{E,\delta}$ is equivalent to*

$$\varrho(T_{E,\delta}) \ \sim \ \min\{\delta \cdot h_{E,T} \ , \ h_{min,T}\} \ \sim \ h_{min,T_{E,\delta}} \qquad . \tag{6}$$

The norm of the transformation matrix $F^{-1}_{T,E,\delta}$ is bounded by

$$\|F^{-1}_{T,E,\delta}\|_{\mathbb{R}^{3\times 3}} \ \lesssim \ \min\{\delta \cdot h_{E,T} \ , \ h_{min,T}\}^{-1} \qquad .$$

3.2 Bubble functions

Another useful and important tool are so-called bubble functions which are applied, for example, for defining the local problem and its ansatz space but also for the analysis. The bubble functions were already partially introduced in [22] and [16].

Denote by $\lambda_{T,1}, \cdots, \lambda_{T,4}$ the barycentric coordinates of an arbitrary tetrahedron T. The *element bubble function* b_T is defined by

$$b_T := 4^4 \cdot \lambda_{T,1} \cdot \lambda_{T,2} \cdot \lambda_{T,3} \cdot \lambda_{T,4} \in \mathbb{P}^4(T) \qquad \text{on } T \qquad . \tag{7}$$

For simplicity assume that b_T is extended by zero outside its original domain of definition.

Further we require face bubble functions. To this end let $E = T_1 \cap T_2$ be an inner face (triangle) of $\mathcal{T}_h$. Enumerate the vertices of T_1 and T_2 such that the vertices of E are numbered first, and introduce the functions

$$b_{E,T_i} := 3^3 \cdot \lambda_{T_i,1} \cdot \lambda_{T_i,2} \cdot \lambda_{T_i,3} \qquad \text{on } T_i, \ i = 1, 2 \qquad .$$

The *standard face bubble function* $b_E \in C^0(\omega_E)$ is now defined in a piecewise fashion (with support $\omega_E = T_1 \cup T_2$) by

$$b_E \ := \ \begin{cases} b_{E,T_1} & \text{on } T_1 \\ b_{E,T_2} & \text{on } T_2 \\ 0 & \text{otherwise} \end{cases} \qquad ,$$

see also the middle of Figure 4. Note that $0 \leq b_T(\mathbf{x}), b_E(\mathbf{x}) \leq 1$ and $\|b_T\|_\infty = \|b_E\|_\infty = 1$.

For clarity of notation we also introduce a trivial extension operator $F_{ext} : \mathbb{P}^0(E) \to \mathbb{P}^0(\omega_E)$ that maps a constant function over some face E to the same constant function acting on ω_E. If E is a boundary face then b_E and F_{ext} are obviously defined only on the single tetrahedron $T \supset E$.

The following anisotropic equivalences/inverse inequalities can be derived easily, cf. [13].

Lemma 2 (Inverse inequalities I) *Assume that $\varphi_T \in \mathbb{P}^0(T)$ and $\varphi_E \in \mathbb{P}^0(E)$. Then*

$$\|b_T\|_T \ \sim \ |T|^{1/2} \tag{8}$$

$$\|b_T^{1/2} \cdot \varphi_T\|_T \ \sim \ \|\varphi_T\|_T \tag{9}$$

$$\|\nabla(b_T \cdot \varphi_T)\|_T \ \lesssim \ h^{-1}_{min,T} \cdot \|\varphi_T\|_T \tag{10}$$

$$\|b_E^{1/2} \cdot \varphi_E\|_E \ \sim \ \|\varphi_E\|_E \tag{11}$$

The bubble functions above suffice to analyse the error estimator for the *Poisson* equation, cf. [22, 15]. However, for the *singularly perturbed problem* considered here we have to introduce modified face bubble functions, cf. also [13, 23].

Start with some face E and let T_1, T_2 be its two neighbouring tetrahedra, i.e. $\omega_E = T_1 \cup T_2$. For an arbitrary real number $\delta \in (0,1]$ consider both squeezed tetrahedra $T_{1,E,\delta} \subset T_1$ and $T_{2,E,\delta} \subset T_2$, cf. Figures 3 and 4. Now we are ready to present the so–called *squeezed face bubble function* $b_{E,\delta}$ which acts only on $T_{1,E,\delta} \cup T_{2,E,\delta} \subset \omega_E$. Its piecewise definition is

$$b_{E,\delta} := \begin{cases} b_{\bar{E}} \circ F^{-1}_{T_1,E,\delta} & \text{on } T_{1,E,\delta} \\ b_{\bar{E}} \circ F^{-1}_{T_2,E,\delta} & \text{on } T_{2,E,\delta} \\ 0 & \text{on } \omega_E \setminus (T_{1,E,\delta} \cup T_{2,E,\delta}) \end{cases} , \tag{12}$$

with $b_{\bar{E}}$ being the standard face bubble function for the face $\bar{E} = F^{-1}_{T_i,E,\delta}(E)$ of the tetrahedron $\bar{T} = F^{-1}_{T_i,E,\delta}(T_{i,E,\delta})$. Note that the squeezed face bubble function on T_i can equivalently be viewed as the standard face bubble function on the squeezed tetrahedron $T_{i,E,\delta}$, i.e.

$$b_{E,\delta}|_{T_i} \;\equiv\; b_{E,T_{i,E,\delta}} \qquad .$$

Figure 4 may facilitate the understanding of the standard/squeezed face bubble function for the two–dimensional case. For boundary faces one restricts $b_{E,\delta}$ to the unique tetrahedron with $\partial T \supset E$.

Standard scaling arguments for the transformation $F_{T_i,E,\delta} : \bar{T} \to T_{i,E,\delta}$, together with the essential Lemma 1 yield now the inverse inequalities for the squeezed face bubble function.

Lemma 3 (Inverse equivalences II) *Let E be an arbitrary face of T, assume $\varphi_E \in \mathbb{P}^0(E)$, and let $\delta \in (0,1]$ be arbitrary. Then one has*

$$\|b_{E,\delta} \cdot F_{ext}(\varphi_E)\|_T \;\sim\; \delta^{1/2} \cdot h_{E,T}^{1/2} \cdot \|\varphi_E\|_E \tag{13}$$

$$\|\nabla(b_{E,\delta} \cdot F_{ext}(\varphi_E))\|_T \;\sim\; \delta^{1/2} \cdot h_{E,T}^{1/2} \cdot \min\{\delta \cdot h_{E,T} \;,\; h_{min,T}\}^{-1} \cdot \|\varphi_E\|_E \qquad . \tag{14}$$

Proof: Standard scaling arguments for the transformation $F_{T,E,\delta} : \bar{T} \to T_{E,\delta}$ readily imply (13).

For (14) we start with the equivalence

$$\|\nabla b_E\|_T \sim h^{-1}_{min,T} \cdot |T|^{1/2}$$

which has been proven (in a slightly different form) in [13, Lemma 3.5] and [15, Lemma 5]. Above we have realized that the squeezed face bubble function can also be viewed as the standard face bubble function on the squeezed tetrahedron. Thus one can utilize the previous equivalence with the function $b_{E,\delta}$ on the tetrahedron $T_{E,\delta}$. Together with Lemma 1 and $|T_{E,\delta}| = \delta \cdot |T|$ this results in

$$\begin{aligned} \|\nabla b_{E,\delta}\|_T \;&\sim\; h^{-1}_{min,T_{E,\delta}} \cdot |T_{E,\delta}|^{1/2} \\ &\overset{(6)}{\sim}\; \min\{\delta \cdot h_{E,T} \;,\; h_{min,T}\}^{-1} \cdot \delta^{1/2} \cdot (h_{E,T}|E|)^{1/2} \end{aligned}$$

which completes the proof. ∎

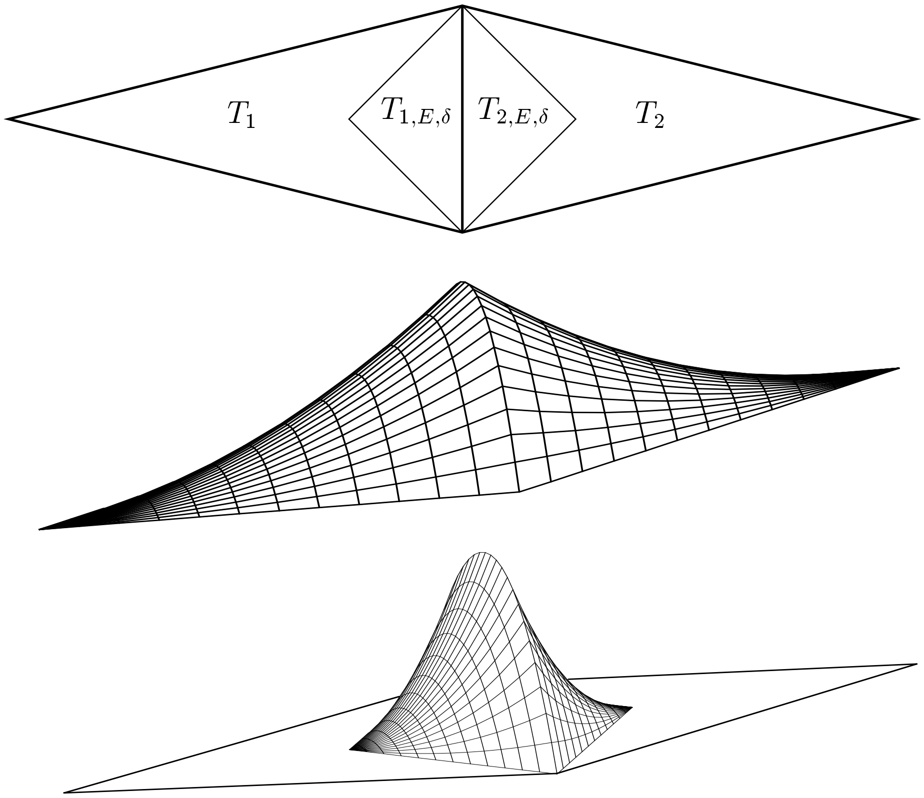

Figure 4: Top: ω_E and squeezed triangles $T_{i,E,\delta}$ (2D case)
Middle: standard face bubble function b_E
Bottom: squeezed face bubble function $b_{E,\delta}$

3.3 Matching function and residual error estimator

Parts of the analysis of the local problem error estimator rely on results for the anisotropic residual error estimator of [16] which are thus repeated now for self–containment. Note that both estimators have been developed in close collaboration to enable certain equivalence properties (cf. Theorem 6 below). Related aspects are discussed here as well.

Ley us start with an important difference between error estimation on *isotropic* and *anisotropic* meshes. For isotropic meshes the error bounds hold unconditionally, whereas this is no longer the case for anisotropic meshes. For such meshes apparently all estimators require the anisotropy of the mesh to be aligned with the anisotropy of the solution. Heuristically this means that anisotropic elements (e.g. tetrahedra) are stretched in that direction where the solution shows little variation.

In order to investigate this matter mathematically, let us recall the proposals from known (analytically based) anisotropic error estimators. Siebert [20] restricts the set of treatable anisotropic functions. Kunert [13, 14, 16] introduces a so–called matching function $m_1(v, \mathcal{T}_h)$ that measures the alignment of an anisotropic function v and an anisotropic mesh $\mathcal{T}_h$. Lastly, in Dobrowolski/Gräf/Pflaum [11] a saturation assumption is necessary that implies a similar correspondence.

For a rigorous analysis it is advantageous to *measure* the alignment of mesh and function. To this end the *matching function* has been proposed by Kunert [13, 14]:

Definition 1 (Matching function) *Let $v \in H^1(\Omega)$, and $\mathcal{T}_h \in \mathcal{F}$ be a triangulation of Ω. Define the* matching function $m_1 : H^1(\Omega) \times \mathcal{F} \mapsto \mathbb{R}$ *by*

$$m_1(v, \mathcal{T}_h) \;:=\; \Big(\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^\top \nabla v\|_T^2 \Big)^{1/2} \Big/ \; \|\nabla v\| \qquad . \tag{15}$$

Note that the entries of the vector $C_T^\top \nabla v \equiv (\mathbf{p}_1^\top \nabla v, \mathbf{p}_2^\top \nabla v, \mathbf{p}_3^\top \nabla v)^\top$ can also be viewed as scaled directional derivatives along the orthogonal directions $\mathbf{p}_i$ (recall $|\mathbf{p}_i| = h_{i,T}$).

To deepen the understanding of the matching function let us briefly discuss its behaviour and influence. More details and a comprehensive discussion can be found in [13, 14].

By defining temporarily $h_{max,T} := h_{1,T}$, one obtains

$$1 \;\le\; m_1(v, \mathcal{T}_h) \lesssim \max_{T \in \mathcal{T}_h} \frac{h_{max,T}}{h_{min,T}} \qquad .$$

Although this crude upper bound is useless for practical purposes it implies $m_1 \sim 1$ on isotropic meshes. Then m_1 merges with other constants and becomes invisible; in this sense (15) is an extension of the theory for isotropic meshes. If an anisotropic mesh $\mathcal{T}_h$ is *well aligned* with an anisotropic function v then one also obtains $m_1(v, \mathcal{T}_h) \sim 1$. If, however, the anisotropic meshes are *not aligned* with the function then the matching function can be arbitrarily large, $m_1(v, \mathcal{T}_h) \gg 1$.

The influence of the matching function m_1 can be seen in the error bound (19) of Lemma 4 and in the discussion afterwards.

Next the residual error estimator will be presented. The methodology to obtain a lower error bound requires a modification of f and g [22, 13]. Hence we replace f by a function f_h which is piecewise constant over the elements. Analogously g is replaced by g_h which is piecewise constant over the Neumann faces. Then the definitions are as follows.

Element and face residual: Define the *element residual* over an element T by

$$r_T := f_h \;-\; (-\varepsilon \Delta u_h + u_h) \qquad \text{on } T.$$

For $x \in E$ define the *face residual* r_E by

$$r_E(x) := \begin{cases} \varepsilon \cdot \lim\limits_{t \to +0} \left[\dfrac{\partial u_h}{\partial n_E}(x + t n_E) - \dfrac{\partial u_h}{\partial n_E}(x - t n_E) \right] & \text{if } E \subset \Omega \setminus \Gamma \\ g_h - \varepsilon \cdot \partial u_h / \partial n & \text{if } E \subset \Gamma_{\mathrm{N}} \\ 0 & \text{if } E \subset \Gamma_{\mathrm{D}} \quad . \end{cases}$$

Here $n_E \perp E$ is any of the two unitary normal vectors whereas $n \perp E \subset \Gamma_{\mathrm{N}}$ denotes the outer unitary normal vector.

Residual scaling factor: The residuals are often accompanied by the factor

$$\alpha_T := \min\{1, \varepsilon^{-1/2} \cdot h_{min,T}\} \qquad . \tag{16}$$

For some interior face $E = T_1 \cap T_2$ we define the corresponding *face related* term by

$$\alpha_E := (\alpha_{T_1} + \alpha_{T_2})/2 \qquad . \tag{17}$$

For boundary faces E set similarly $\alpha_E := \alpha_T$ for $E \subset \partial T$. Note that the mesh requirements imply $\alpha_E \sim \alpha_{T_1} \sim \alpha_{T_2}$, cf. also remark 1.

Local residual error estimator: For a tetrahedron T, define it by

$$\eta_{\varepsilon,\mathrm{R},T} := \Big(\alpha_T^2 \cdot \|r_T\|_T^2 + \varepsilon^{-1/2} \cdot \alpha_T \cdot \sum_{E \subset \partial T \setminus \Gamma_{\mathrm{D}}} \|r_E\|_E^2\Big)^{1/2} . \tag{18}$$

Local approximation term: To shorten the notation, introduce

$$\zeta_{\varepsilon,T} := \Big(\alpha_T^2 \cdot \|f - f_h\|_{\omega_T}^2 \; + \; \varepsilon^{-1/2} \cdot \alpha_T \cdot \sum_{E \subset \partial T \cap \Gamma_{\mathrm{N}}} \|g - g_h\|_E^2\Big)^{1/2} .$$

Finally, define the *global* terms

$$\eta_{\varepsilon,\mathrm{R}}^2 := \sum_{T \in \mathcal{T}_h} \eta_{\varepsilon,\mathrm{R},T}^2 \qquad \text{and} \qquad \zeta_\varepsilon^2 := \sum_{T \in \mathcal{T}_h} \zeta_{\varepsilon,T}^2 \qquad .$$

The following residual error estimation essentially has been proven by Kunert [16]. Here we have included the treatment of Neumann boundary conditions.

Lemma 4 *The error is bounded locally from below for all $T \in \mathcal{T}_h$ by*

$$\eta_{\varepsilon,\mathrm{R},T} \lesssim |||u - u_h|||_{\omega_T} \; + \; \zeta_{\varepsilon,T} \qquad .$$

The error is bounded globally from above by

$$|||u - u_h||| \lesssim m_1(u - u_h, \mathcal{T}_h) \cdot \big[\eta_{\varepsilon,\mathrm{R}}^2 + \zeta_\varepsilon^2\big]^{1/2} \qquad . \tag{19}$$

Both error bounds are uniform in ε.

We remark that only the upper error bound contains the matching function m_1. Hence only (19) is influenced by the degree of the alignment of mesh and function, i.e. the global error estimator $\eta_{\varepsilon,\mathrm{R}}$ is emphasized by the factor $m_1(u - u_h, \mathcal{T}_h)$. When $m_1 \sim 1$ the lower and upper error bound possess the same quality. Obviously, the smaller m_1 the better the upper error bound. In the case of $m_1 \gg 1$ however both error bounds differ by a large factor, thus rendering the error estimation useless.

Remark 2 Note that the upper error bound of (19) can not be computed exactly as it contains $m_1(u - u_h, \mathcal{T}_h)$ and thus the (unknown) error $u - u_h$. As a remedy m_1 can be approximated, e.g. by means of a recovered gradient $\nabla^R u_h \approx \nabla u$:

$$\begin{aligned} m_1(u - u_h, \mathcal{T}_h) &\equiv \Big(\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^\top \nabla(u - u_h)\|_T^2 \Big)^{1/2} \Big/ \|\nabla(u - u_h)\| \\ &\approx \Big(\sum_{T \in \mathcal{T}_h} h_{min,T}^{-2} \cdot \|C_T^\top (\nabla^R u_h - \nabla u_h)\|_T^2 \Big)^{1/2} \Big/ \|\nabla^R u_h - \nabla u_h\| \\ &=: m_1^R(u_h, \mathcal{T}_h) \qquad , \end{aligned} \tag{20}$$

cf. [14] for a more comprehensive discussion. All numerical experiments so far indicate that m_1^R is a robust approximation to m_1, see also the numerical experiments below. ■

4 Local problem error estimators

4.1 Definition of the error estimator $\eta_{\varepsilon,\mathrm{D}}$

The main ideas behind local problem error estimation have been known for a long time [2, 6, 8, 21, 22]. Basically the problem is solved locally but with higher accuracy, and the difference between the new solution and the original finite element solution serves as error estimator.

In each one of the aforementioned sources an *isotropic* mesh is assumed. In contrast to this the author has shown in [15] that reliable local problem error estimation is possible on *anisotropic* meshes as well. There a Poisson model problem has been investigated, the methodology of the analysis has been presented, and some important new tools and results have been developed.

In our work here we demonstrate that anisotropic local problem error estimation is not restricted to the Poisson problem but that it can be extended to the singularly perturbed reaction–diffusion problem (1). We propose a new error estimator for the latter problem. Note that the only other local problem error estimator for a singularly perturbed reaction–diffusion problem is due to [1] where an isotropic mesh is assumed, and where the local problem is infinite dimensional.

While the general structure of the proofs here is similar to the ones for the local problem error estimator for the Poisson problem [15], the actual ingredients differ. This mainly concerns the squeezed tetrahedron and its properties as well as the squeezed face bubble functions which play a vital role in almost all analysis. The definitions of the error estimators $\eta_{\varepsilon,\mathrm{D},T}$ and $\eta_{\varepsilon,\mathrm{R},T}$ require a very careful balancing of all scaling factors (e.g. α_T from (16)) and of the 'squeezing' parameter δ_E from (21). Consequently the proof of the vital Lemma 5 is even more technical than in [15], see also Appendix A. The derivation of a stable basis for the local problem is different from [15]. Furthermore special care has to be taken to obtain a feasible implementation of the estimator. Hence computational aspects and difficulties are addressed.

The remainder of this section is devoted to the definition of the local problem and the error estimator. Then Lemma 5 gives two central inequalities for the local space. Next, Theorem 6 states the equivalence of the local problem error estimator $\eta_{\varepsilon,\mathrm{D},T}$ and the residual error estimator $\eta_{\varepsilon,\mathrm{R},T}$. The main results, namely lower and upper bounds of the error, are given in Theorem 7. In Theorem 8 it is shown that a certain basis of the local space V_T is stable (i.e. the local Dirichlet problem is well–conditioned). Finally other choices of local problem error estimators are feasible as well. This is demonstrated exemplarily in Section 4.4 for a face based estimator.

When deriving the error estimator, the corresponding local problem should be cheap to solve but simultaneously be rich enough to extract information on the error $e := u - u_h$. Here the subdomain of the local problem is chosen to be ω_T. Let

$$H^1_o(\omega_T) := \{v \in H^1(\Omega)\ : \mathrm{supp}\ v \subseteq \omega_T, \quad v = 0 \ \text{ on } \partial\omega_T \setminus \Gamma_\mathrm{N}\} \qquad .$$

For an arbitrary function $v \in H^1_o(\omega_T)$ the error then satisfies

$$a(u - u_h, v) = \int\limits_{\omega_T} f \cdot v + \int\limits_{\partial\omega_T \cap \Gamma_\mathrm{N}} g \cdot v - \int\limits_{\omega_T} \varepsilon(\nabla u_h)^\top \nabla v - \int\limits_{\omega_T} u_h v \qquad .$$

The local problem is obtained by approximating the space $H^1_o(\omega_T)$ by some local, finite dimensional space $V_T \subset H^1_o(\omega_T)$ which is spanned by an element bubble function and some squeezed face bubble functions. Their 'squeezing' parameters δ_E (cf. 12) are now specified to be

$$\delta_E := \min\left\{1, \frac{h_{min,E}}{h_E}, \frac{\sqrt{\varepsilon}}{h_E}\right\} \qquad . \tag{21}$$

Note that $\delta_E \sim \sqrt{\varepsilon} h_E^{-1} \alpha_E$. The local space V_T now becomes

$$V_T := \mathrm{span}\{b_T, b_{E,\delta_E} : E \subset \partial T \setminus \Gamma_\mathrm{D}\} \qquad . \tag{22}$$

Analogously to the residual error estimator one replaces f and g by f_h and g_h, respectively.

Definition 2 (Local Dirichlet problem error estimator)
Find a solution $e_T \in V_T$ of the local variational problem:

$$\begin{aligned} a(e_T, v_T) \ &\equiv \int\limits_{\omega_T} \varepsilon(\nabla e_T)^\top \nabla v_T + e_T\, v_T \\ &\overset{!}{=} \int\limits_{\omega_T} f_h \cdot v_T + \int\limits_{\partial\omega_T \cap \Gamma_\mathrm{N}} g_h \cdot v_T - \int\limits_{\omega_T} \varepsilon(\nabla u_h)^\top \nabla v_T - \int\limits_{\omega_T} u_h v_T \end{aligned} \tag{23}$$

for all $v_T \in V_T$. The local *and* global error estimators *then become*

$$\eta_{\varepsilon,\mathrm{D},T} := |||e_T|||_{\omega_T} \qquad \text{and} \qquad \eta^2_{\varepsilon,\mathrm{D}} := \sum_{T \in \mathcal{T}_h} \eta^2_{\varepsilon,\mathrm{D},T} \qquad . \tag{24}$$

Note that the particular choice of the local ansatz space V_T (namely $v_T = 0$ on $\partial\omega_T \setminus \partial T$) reduces certain boundary integrals and norms. Equivalent formulations of the local problem are derived by partial integration.

Alternative 1: Find $e_T \in V_T$ such that

$$a(e_T, v_T) \;=\; \sum_{T' \in \omega_T} \int_{T'} r_{T'} \cdot v_T \;+\; \sum_{E \subset \partial T \setminus \Gamma_{\rm D}} \int_E r_E \cdot v_T \qquad \forall\, v_T \in V_T \qquad . \tag{25}$$

Alternative 2: Find $e_T \in V_T$ such that

$$a(e_T, v_T) \;=\; a(u - u_h, v_T) - \int\limits_{\omega_T} (f - f_h)\, v_T \;-\; \int\limits_{\partial T \cap \Gamma_{\rm N}} (g - g_h)\, v_T \qquad \forall\, v_T \in V_T \quad . \tag{26}$$

4.2 Equivalence and bounds of the local problem estimator

The methodology of the error estimator partly utilizes ideas that have already been introduced for the anisotropic local problem estimator for the Poisson problem [15], and for the anisotropic residual estimator for a singularly perturbed reaction–diffusion equation [16]. All the details however are original and new. The first lemma plays a central role in the analysis of the estimator.

Lemma 5 *The following relations below hold for all $v_T \in V_T$.*

$$\|v_T\|_{\omega_T} \;\lesssim\; h_{min,T} \cdot \|\nabla v_T\|_{\omega_T} \tag{27}$$

$$\|v_T\|_E \;\lesssim\; h_E^{-1/2}\, \delta_E^{-1/2} \cdot \min\{h_{min,T}, \delta_E\, h_E\} \cdot \|\nabla v_T\|_{\omega_T} \qquad \forall\, E \subset \partial T \quad , \tag{28}$$

with δ_E from (21). The inequalities are uniform in δ_E.

If T has at least two Neumann boundary faces then the constants can depend on the angles of the corners and edges of Ω (but do not depend on the triangulation $\mathcal{T}_h$ nor on T).

Proof: The technical proof is postponed to the appendix. ■

Theorem 6 (Equivalence with residual error estimator) *The local problem error estimator $\eta_{\varepsilon,\mathrm{D},T}$ is equivalent to the residual error estimator $\eta_{\varepsilon,\mathrm{R},T}$ in the following sense:*

$$\eta^2_{\varepsilon,\mathrm{D},T} \;\lesssim\; \sum_{T' \subset \omega_T} \eta^2_{\varepsilon,\mathrm{R},T'} \tag{29}$$

$$\eta^2_{\varepsilon,\mathrm{R},T} \;\lesssim\; \sum_{T' \subset \omega_T} \eta^2_{\varepsilon,\mathrm{D},T'} \qquad . \tag{30}$$

Both inequalities are uniform in ε.

If T has at least two Neumann boundary faces then the constant in (29) can depend on the angles of the corners and edges of Ω (but do not depend on the triangulation $\mathcal{T}_h$ nor on T).

Proof: Recall the definition (24) of $\eta_{\varepsilon,\mathrm{D},T}$, observe that $e_T = 0$ on $\partial\omega_T \setminus \partial T$, and take into account the modifications for boundary faces. By integration by parts one obtains

$$\begin{aligned}\eta^2_{\varepsilon,\mathrm{D},T} &= |||e_T|||^2_{\omega_T} = a(e_T, e_T) \overset{(25)}{=} \sum_{T'\in\omega_T} \int_{T'} r_{T'} \cdot e_T + \sum_{E\subset\partial T\setminus\Gamma_\mathrm{D}} \int_E r_E \cdot e_T \\ &\leq \left(\sum_{T'\subset\omega_T} \|r_{T'}\|^2_{T'}\right)^{1/2} \cdot \|e_T\|_{\omega_T} + \sum_{E\subset\partial T\setminus\Gamma_\mathrm{D}} \|r_E\|_E \cdot \|e_T\|_E .\end{aligned}$$

Now $\|e_T\|_{\omega_T}$ and $\|e_T\|_E$, $E \subset \partial T$, are to be bounded. Recall the definition of α_T and δ_E and apply Lemma 5 to obtain

$$\begin{aligned} & \|e_T\|_{\omega_T} &\leq\ & |||e_T|||_{\omega_T} \\ & \|e_T\|_{\omega_T} &\overset{(27)}{\lesssim}\ & h_{min,T} \cdot \|\nabla e_T\|_{\omega_T} \leq h_{min,T} \cdot \varepsilon^{-1/2}\, |||e_T|||_{\omega_T} \\ \Rightarrow\ & \|e_T\|_{\omega_T} &\lesssim\ & \min\{1, \varepsilon^{-1/2} \cdot h_{min,T}\} \cdot |||e_T|||_{\omega_T} \equiv \alpha_T \cdot |||e_T|||_{\omega_T} \end{aligned} \tag{31}$$

$$\begin{aligned} \text{and}\quad \|e_T\|_E &\overset{(28)}{\lesssim} h_E^{-1/2}\, \delta_E^{-1/2} \cdot \min\{h_{min,T}, \delta_E\, h_E\} \cdot \|\nabla e_T\|_{\omega_T} \\ &\sim \varepsilon^{1/4}\, \alpha_E^{1/2}\, \|\nabla e_T\|_{\omega_T} \\ &\lesssim \varepsilon^{-1/4}\, \alpha_T^{1/2}\, |||e_T|||_{\omega_T} \quad . \end{aligned} \tag{32}$$

Inserting these inequalities and utilizing $\alpha_T \sim \alpha_{T'}$ for neighboring tetrahedra results in

$$\eta^2_{\varepsilon,\mathrm{D},T} \lesssim \left(\sum_{T'\subset\omega_T} \alpha^2_{T'} \cdot \|r_{T'}\|^2_{T'} + \varepsilon^{-1/2}\, \alpha_T \cdot \sum_{E\subset\partial T\setminus\Gamma_\mathrm{D}} \|r_E\|_E\right)^{1/2} \cdot |||e_T|||_{\omega_T}$$

which, together with $|||e_T|||_{\omega_T} = \eta_{\varepsilon,\mathrm{D},T}$, proves (29).

For the proof of (30) we require bounds of $\eta_{\varepsilon,\mathrm{R},T}$, and thus of $\|r_T\|_T$ and $\|r_E\|_E$. The structure of the proof is similar to our analysis for the Poisson equation [15].

We first bound the term $\|r_{T'}\|_{T'}$, with $T' \subset \omega_T$ being an arbitrary tetrahedron. Recall definition (7) of the bubble function $b_{T'}$ and set $v_{T'} := b_{T'} \cdot r_{T'}$. Then $b_{T'}$ and $v_{T'}$ belong to the finite element space $V_{T'}$. Hence the local problem related to T' has to be invoked. The local problem (25) and equivalence (9) imply

$$\begin{aligned}\|r_{T'}\|^2_{T'} &\overset{(9)}{\sim} \|b^{1/2}_{T'} \cdot r_{T'}\|^2_{T'} = \int_{T'} r_{T'} \cdot v_{T'} \qquad \text{since } v_{T'} \in H^1_o(T') \\ &\overset{(25)}{=} a(e_{T'}, v_{T'}) \leq |||e_{T'}|||_{T'} \cdot |||v_{T'}|||_{T'} \qquad ,\end{aligned}$$

where $e_{T'} \in V_{T'}$ denotes the solution of the local problem over $\omega_{T'}$. Inequality (10) results in

$$\begin{aligned}|||v_{T'}|||^2_{T'} &= \varepsilon\, \|\nabla(b_{T'} \cdot r_{T'})\|^2_{T'} + \|b_{T'} \cdot r_{T'}\|^2_{T'} \\ &\overset{(10)}{\sim} \varepsilon\, h^{-2}_{min,T'} \cdot \|r_{T'}\|^2_{T'} + \|r_{T'}\|^2_{T'} \sim \alpha^{-2}_{T'}\, \|r_{T'}\|^2_{T'} \qquad .\end{aligned}$$

Combining both inequalities yields

$$\|r_{T'}\|_{T'} \lesssim \alpha_T^{-1} \cdot |||e_{T'}|||_{T'} \le \alpha_T^{-1} \cdot \eta_{\varepsilon,\mathrm{D},T'} \qquad \forall\, T' \in \omega_T \tag{33}$$

since $\alpha_{T'}$ does not change rapidly across adjacent tetrahedra T'.

The norm of $r_E \in \mathbb{P}^0(E)$ for an interior face $E \subset \partial T \setminus \Gamma$ is bounded similarly. Recall the definition (12) of the squeezed face bubble function $b_{E,\delta}$, and set $v_E := b_{E,\delta} \cdot F_{ext}(r_E) \in V_T \cap H_o^1(\omega_E)$. Integration by parts and $v_E = 0$ on $\partial T \cap \Gamma$ imply

$$\begin{aligned}
\|r_E\|_E^2 \;&\overset{(11)}{\sim}\; \|b_E^{1/2} \cdot r_E\|_E^2 \;=\; \int_E r_E \cdot v_E \\
&\overset{(25)}{=}\; a(e_T, v_E) \;-\; \sum_{T' \subset \omega_E} \int_{T'} r_{T'}\, v_E \\
&\le\; |||e_T|||_{\omega_E} \cdot |||v_E|||_{\omega_E} \;+\; \sum_{T' \subset \omega_E} \|r_{T'}\|_{T'} \cdot \|v_E\|_{T'} \qquad .
\end{aligned}$$

Now the norms of v_E are bounded by means of inverse inequalities, and by using the specific value of δ_E from (21). This leads to

$$\begin{aligned}
\|v_E\|_{T'} \;&=\; \|b_{E,\delta} \cdot F_{ext}(r_E)\|_{T'} \;\overset{(13)}{\lesssim}\; \delta_E^{1/2} \cdot h_{E,T'}^{1/2} \cdot \|r_E\|_E \;\overset{(21)}{\sim}\; \varepsilon^{1/4}\, \alpha_T^{1/2} \cdot \|r_E\|_E \\
\|\nabla v_E\|_{T'} \;&=\; \|\nabla(b_{E,\delta} \cdot F_{ext}(r_E))\|_{T'} \\
&\overset{(14)}{\lesssim}\; \delta_E^{1/2} \cdot h_{E,T'}^{1/2} \cdot \min\{\delta_E \cdot h_{E,T'}, h_{min,T'}\}^{-1} \cdot \|r_E\|_E \\
&\overset{(21)}{\sim}\; \varepsilon^{-1/4}\, \alpha_T^{-1/2} \cdot \|r_E\|_E \\
\Rightarrow \qquad |||v_E|||_{\omega_E} \;&=\; (\varepsilon \|\nabla(v_E)\|_{\omega_E}^2 \;+\; \|v_E\|_{\omega_E}^2)^{1/2} \;\sim\; \varepsilon^{1/4}\, \alpha_T^{-1/2} \cdot \|r_E\|_E \qquad .
\end{aligned}$$

Next one utilizes the previous bound (33) of $\|r_{T'}\|_{T'}$ for both tetrahedra $T' \subset \omega_E$. Combining all estimates yields

$$\|r_E\|_E \lesssim \varepsilon^{1/4}\, \alpha_T^{-1/2} \cdot \sum_{T' \subset \omega_E} \eta_{\varepsilon,\mathrm{D},T'} \qquad \forall\, E \subset \partial T \setminus \Gamma \qquad . \tag{34}$$

The norm of $r_E \in \mathbb{P}^0(E)$ for a Neumann boundary face $E \subset \partial T \cap \Gamma_\mathrm{N}$ is bounded similarly (cf. [15]) and gives analogously

$$\|r_E\|_E \lesssim \varepsilon^{1/4}\, \alpha_T^{-1/2} \cdot \eta_{\varepsilon,\mathrm{D},T} \qquad \forall\, E \subset \partial T \cap \Gamma_\mathrm{N} \qquad .$$

Collecting all the results for $\|r_T\|_T$ and $\|r_E\|_E$ and inserting them into the definition of $\eta_{\varepsilon,\mathrm{R},T}$ gives (30). ■

With the help of Theorem 6 we easily derive the main result, namely upper and lower error bounds by means of the local problem error estimator.

Theorem 7 (Local problem error estimation)
The error is bounded locally from below by

$$\eta_{\varepsilon,\mathrm{D},T} \;\leq\; |||u - u_h|||_{\omega_T} \;+\; c \cdot \zeta_{\varepsilon,T} \qquad \forall\, T \in \mathcal{T}_h \qquad . \tag{35}$$

The error is bounded globally from above by

$$|||u - u_h||| \;\lesssim\; m_1(u - u_h, \mathcal{T}_h) \;\cdot\; \left[\eta^2_{\varepsilon,\mathrm{D}} \;+\; \zeta^2_{\varepsilon}\right]^{1/2} \qquad . \tag{36}$$

Both inequalities are uniform in ε.

If T has at least two Neumann boundary faces then the constant in (35) can depend on the angles of the corners and edges of Ω (but do not depend on the triangulation $\mathcal{T}_h$ nor on T).

Note that in analogy to Lemma 4 only the upper error bound is influenced by the matching function $m_1(u - u_h, \mathcal{T}_h)$.

Proof: For (35), apply formulation (26) of the local problem, recall how ω_T and V_T are modified if T has a boundary face, and observe in particular that $e_T = 0$ on $\partial\omega_T \setminus \partial T$. Then one obtains

$$\begin{aligned}
\eta^2_{\varepsilon,\mathrm{D},T} \;&=\; |||e_T|||^2_{\omega_T} \;=\; a(e_T, e_T) \\
&\overset{(26)}{=}\; a(u - u_h, e_T) - \int_{\omega_T} (f - f_h) \cdot e_T \;-\; \int_{\Gamma_\mathrm{N} \cap \partial T} (g - g_h) \cdot e_T \\
&\leq\; |||u - u_h|||_{\omega_T} \cdot |||e_T|||_{\omega_T} \;+\; \|f - f_h\|_{\omega_T} \cdot \|e_T\|_{\omega_T} + \;\|g - g_h\|_{\Gamma_\mathrm{N} \cap \partial T} \cdot \|e_T\|_{\Gamma_\mathrm{N} \cap \partial T} .
\end{aligned}$$

With the previous bounds (31) and (32) one readily obtains the desired estimate (35). Finally inequality (36) follows immediately from the error bound (19) of the residual error estimator, and from relation (30) between $\eta_{\varepsilon,\mathrm{R},T}$ and $\eta_{\varepsilon,\mathrm{D},T}$. ■

4.3 A stable basis for the local problem

Here we will present a stable basis for the local problem under consideration. An equivalent description of this aim is that the variational problem is well-conditioned, i.e. the condition number of the corresponding finite element matrix is bounded independently of the perturbation parameter ε and of aspect ratio of the elements.

Recall that the local ansatz space is $V_T := \mathrm{span}\{b_T, b_{E,\delta_E} : E \subset \partial T \setminus \Gamma_\mathrm{D}\}$. As a basis of V_T we choose

$$\Phi \;:=\; (b_T \;,\; \delta_E^{-1/2} \cdot b_{E,\delta} \;:\; E \subset \partial T \setminus \Gamma_\mathrm{D}) \qquad . \tag{37}$$

For simplicity of notation enumerate the faces of T such that interior and Neumann faces come first, and denote them by E_i, $i = 1, \ldots, m$, $m \leq 4$. Denote the parameter of the squeezed face bubble functions temporarily by $\delta_i := \delta_{E_i}$. Hence any function $v_T \in V_T$ can be expressed as

$$\begin{aligned} v_T &= \beta_0 b_T + \sum_{i=1}^{m} \beta_i \cdot \delta_i^{-1/2} \cdot b_{E_i,\delta_i} = \Phi \cdot \mathbf{v} \\ \text{with} \qquad \mathbf{v} &:= (\beta_0, \beta_1 \ldots \beta_m)^\top \qquad . \end{aligned}$$

The stiffness matrix $K_T \in \mathbb{R}^{(1+m)\times(1+m)}$ of the local problem is given by means of the finite element isomorphism $a(v_T, w_T) = (K_T\mathbf{v}, \mathbf{w})$ for all $w_T = \Phi \cdot \mathbf{w} \in V_T$.

Theorem 8 (Stable basis) *The basis (37) of* V_T *is stable, i.e. the condition number* $\kappa(K_T)$ *of the local problem stiffness matrix* K_T *is bounded uniformly in* ε *and* T*:*

$$\kappa(K_T) \sim 1 \qquad \forall\, T \in \mathcal{T}_h \qquad .$$

Proof: The condition number is given by

$$\kappa(K_T) = \Big[\max_{\mathbf{v}\neq\mathbf{0}} (K_T\mathbf{v}, \mathbf{v})/(\mathbf{v}, \mathbf{v})\Big] \Big/ \Big[\min_{\mathbf{w}\neq\mathbf{0}} (K_T\mathbf{w}, \mathbf{w})/(\mathbf{w}, \mathbf{w})\Big] \qquad .$$

Thus investigate the scalar product $(K_T\mathbf{v}, \mathbf{v})$ which equals

$$(K_T\mathbf{v}, \mathbf{v}) = a(v_T, v_T) = |||v_T|||^2_{\omega_T} = \varepsilon\|\nabla v_T\|^2_{\omega_T} + \|v_T\|^2_{\omega_T} \qquad .$$

We start by bounding $|||v_T|||_{\omega_T}$ from above. The triangle inequality readily implies

$$|||v_T|||_{\omega_T} \le |\beta_0| \cdot |||b_T|||_{\omega_T} + \sum_{i=1}^{m} |\beta_i| \cdot \delta_i^{-1/2}\, |||b_{E_i,\delta_i}|||_{\omega_T} \qquad .$$

Using inverse inequalities (8) and (10) one derives

$$|||b_T|||^2_{\omega_T} = \varepsilon\|\nabla b_T\|^2_T + \|b_T\|^2_T \lesssim \varepsilon h^{-2}_{min,T}\, |T| + |T| \overset{(16)}{\sim} \alpha_T^{-2} \cdot |T| \qquad .$$

The second inverse equivalences (13) and (14) and the particular choice of $\delta_i \equiv \delta_{E_i}$ from (21) yield

$$\begin{aligned} |||b_{E_i,\delta_i}|||^2_{\omega_T} &= \varepsilon\|\nabla b_{E_i,\delta_i}\|^2_{\omega_E} + \|b_{E_i,\delta_i}\|^2_{\omega_E} \\ &\sim \delta_i \cdot |T| \cdot \Big(\varepsilon \min\{\delta_i \cdot h_{E,T}, h_{min,T}\}^{-2} + 1\Big) \\ &\overset{(21)}{\sim} \delta_i \cdot |T| \cdot (1 + \alpha_T^{-2}) \sim \delta_i \cdot |T| \cdot \alpha_T^{-2} \qquad . \end{aligned}$$

Altogether one obtains

$$|||v_T|||_{\omega_T} \lesssim |\beta_0| \cdot \alpha_T^{-1}\, |T|^{1/2} + \sum_{i=1}^{m} |\beta_i| \cdot \alpha_T^{-1}\, |T|^{1/2} \sim \alpha_T^{-1}\, |T|^{1/2} \cdot \|\mathbf{v}\|_{\mathbb{R}^{1+m}} \qquad .$$

To bound $|||v_T|||_{\omega_T}$ from below, apply Lemma 5 giving

$$\begin{array}{rrcl} & \|\nabla v_T\|_{\omega_T} & \overset{(27)}{\gtrsim} & h_{min,T}^{-1} \cdot \|v_T\|_{\omega_T} \\ \text{and} & |||v_T|||_{\omega_T}^2 & \gtrsim & (\varepsilon\, h_{min,T}^{-2} + 1) \cdot \|v_T\|_{\omega_T}^2 \sim \alpha_T^{-2} \cdot \|v_T\|_{\omega_T}^2 \quad . \end{array}$$

In the proof of Lemma 5 in Appendix A it is shown that

$$\|v_T\|_{\omega_T}^2 \overset{(40)}{\sim} |T| \cdot \sum_{i=0}^{m} \beta_i^2 \sim |T| \cdot \|\mathbf{v}\|_{\mathbb{R}^{1+m}}^2$$

which completes the lower bound of $|||v_T|||_{\omega_T}$.

Summarizing all results, one ends up with

$$(K_T\mathbf{v}, \mathbf{v}) = |||v_T|||_{\omega_T}^2 \sim \alpha_T^{-2}\, |T| \cdot \|\mathbf{v}\|_{\mathbb{R}^{1+m}}^2$$

which immediately yields $\lambda_{min}(K_T) \sim \lambda_{max}(K_T) \sim \alpha_T^{-2}\, |T|$ and the desired assertion $\kappa(K_T) \sim 1$. ■

4.4 A further, face based local problem error estimator

With the methodology presented so far one can derive further local problem error estimators. This will be demonstrated here for a *face based* local problem error estimator. Such an estimator can be advantageous when other ingredients of an adaptive algorithm are face based too (e.g. the refinement procedure).

We start again with a corresponding residual error estimator. For an arbitrary but fixed face E define the face based residual error estimator and the approximation term by

$$\eta_{\varepsilon,\mathrm{R},E} := \Big(\alpha_E^2 \cdot \sum_{T\subset\omega_E} \|r_T\|_T^2 + \varepsilon^{-1/2}\, \alpha_E \cdot \|r_E\|_E^2\Big)^{1/2} \tag{38}$$

$$\zeta_{\varepsilon,E} := \alpha_E \cdot \|f - f_h\|_{\omega_E} + \varepsilon^{-1/4}\, \alpha_E^{1/2} \cdot \|g - g_h\|_{E\cap\Gamma_\mathrm{N}} \quad , \tag{39}$$

respectively (the norm $\|\cdot\|_{E\cap\Gamma_\mathrm{N}}$ here is to be evaluated only when $E \subset \Gamma_\mathrm{N}$).

Utilizing the techniques and most of the results of [16] one can comparatively easily prove the following residual error estimation.

Lemma 9 *The error is bounded locally from below for all faces E of $\mathcal{T}_h$ by*

$$\eta_{\varepsilon,\mathrm{R},E} \lesssim |||u - u_h|||_{\omega_E} + \zeta_{\varepsilon,E} \quad .$$

The error is bounded globally from above by

$$|||u - u_h||| \lesssim m_1(u - u_h, \mathcal{T}_h) \cdot \left(\sum_{E\in\mathcal{T}_h} \eta_{\varepsilon,\mathrm{R},E}^2 + \zeta_{\varepsilon,E}^2\right)^{1/2} \quad ,$$

where the sum over $E \in \mathcal{T}_h$ includes interior and boundary faces of the triangulation. Both error bounds are uniform in ε.

Note that the residual error estimator can be modified such that it contains only the face residual but not the element residuals. Then a very similar result is achieved, cf. [17]. Since this modification is not suitable for our subsequent analysis, we omit a detailed description.

The local space associated to a face E is now set to

$$V_E := \operatorname{span}\{b_{E,\delta_E} \text{ if } E \not\subset \Gamma_\mathrm{D}\ ,\ b_T\ \forall\, T \subset \omega_E\} \qquad ,$$

i.e. V_E is three dimensional for interior faces E. The local problem is: Find $e_E \in V_E$ such that

$$a(e_E, v_E) \ \stackrel{!}{=} \int\limits_{\omega_E} f_h \cdot v_E \ + \int\limits_{E\cap\Gamma_\mathrm{N}} g_h \cdot v_E - \int\limits_{\omega_E} \varepsilon(\nabla u_h)^\top \nabla v_E - \int\limits_{\omega_E} u_h v_E$$

for all $v_E \in V_E$. The *local face error estimator* then becomes

$$\eta_{\varepsilon,\mathrm{D},E} := |||e_E|||_{\omega_E} \qquad .$$

Again two alternative, equivalent descriptions of the local problem are possible and advantageous. Find $e_E \in V_E$ such that

$$\begin{aligned} \text{Alternative 1:} \qquad a(e_E, v_E) \ &= \ \sum_{T\in\omega_E} \int_T r_T \cdot v_E \ + \ \int_E r_E \cdot v_E \\ \text{Alternative 2:} \qquad a(e_E, v_E) \ &= \ a(u - u_h, v_E) - \int\limits_{\omega_E} (f - f_h)\, v_E \ - \int\limits_{E\cap\Gamma_\mathrm{N}} (g - g_h)\, v_E \end{aligned}$$

holds for all $v_E \in V_E$.

Using the techniques and even some results of the previous analysis of the element based local problem error estimator the following theorem can be shown. Because of the similarities of the proofs we only state the result.

Theorem 10 (Face based local problem error estimator)
The face based residual error estimator and local problem error estimator are equivalent:

$$\eta_{\varepsilon,\mathrm{D},E} \ \sim \ \eta_{\varepsilon,\mathrm{R},E} \qquad \forall\, E \in \mathcal{T}_h \qquad .$$

The error is bounded locally from below for all faces E of $\mathcal{T}_h$ by

$$\eta_{\varepsilon,\mathrm{D},E} \ \leq \ |||u - u_h|||_{\omega_E} \ + \ c \cdot \zeta_{\varepsilon,E} \qquad .$$

The error is bounded globally from above by

$$|||u - u_h||| \lesssim m_1(u - u_h, \mathcal{T}_h) \cdot \left(\sum_{E\in\mathcal{T}_h} \eta^2_{\varepsilon,\mathrm{D},E} + \zeta^2_{\varepsilon,E} \right)^{1/2} \qquad .$$

All relations are uniform in ε.

5 Computational implementation

5.1 Difficulties and their solution

It is a major demand that the local problem can be constructed and solved as fast as possible since usually the error estimation is as expensive as the assembly of the global finite element stiffness matrix and the solution process for u_h. Therefore one encounters two main problems when applying our error estimator. Both difficulties are related to the computation of the local stiffness matrix K_T which arises from the bilinear form $a(\cdot,\cdot)$, see also Section 4.3.

Problem 1: The support of the squeezed face bubble function $b_{E,\delta}$ is not ω_E but some (δ–dependent) part of it. For example the computation of $a(b_{E,\delta},\cdot)$ implies a comparatively complicated domain of integration. In particular when evaluating $a(b_{E_1,\delta_{E_1}}, b_{E_2,\delta_{E_2}})$ for two different squeezed face bubble functions, the domain of integration becomes

$$\mathrm{supp}(b_{E_1,\delta_{E_1}}) \cap \mathrm{supp}(b_{E_2,\delta_{E_2}})$$

which might be empty, or a single tetrahedron, or the union of two tetrahedra, depending on δ_{E_1} and δ_{E_2} (cf. also Figures 3 and 4). Even to determine and describe the domain of integration is not trivial, save the actual integration.

Remedy: We modify the parameter for the squeezed face bubble function to be

$$\tilde{\delta}_E := \min\left\{\frac{1}{4}, \frac{h_{min,E}}{h_E}, \frac{\sqrt{\varepsilon}}{h_E}\right\} \;\equiv\; \min\left\{\frac{1}{4}, \delta_E\right\} \;\sim\; \delta_E \qquad .$$

Then all results remain valid, only the inequality constants may be slightly worse (but they are still uniform in ε). The main advantage now is that

$$\mathrm{supp}(b_{E_1,\tilde{\delta}_{E_1}}) \cap \mathrm{supp}(b_{E_2,\tilde{\delta}_{E_2}}) \;=\; \emptyset \qquad .$$

Hence the computation of the modified local matrix $\tilde{K}_T$ is less expensive, as the matrix now contains several zero entries. Even more, the sparsity pattern

$$\tilde{K}_T = \tilde{K}_T^\top \;=\; \begin{bmatrix} * & * & * & * & * \\ * & * & 0 & 0 & 0 \\ * & 0 & * & 0 & 0 \\ * & 0 & 0 & * & 0 \\ * & 0 & 0 & 0 & * \end{bmatrix}$$

allows a particularly fast and simple solution of the local problem.

Problem 2: The basis functions of V_T are polynomials of a relatively high degree. Hence numerical integration rules to compute $a(\cdot,\cdot)$ are far too expensive and thus unsuitable.

Remedy: Instead we propose a direct computation of the integrals involved. The procedure is explained exemplarily for $a(b_T, b_T)$. Using the transformation technique via $F_A : \bar{T} \to T$, one obtains

$$\begin{aligned} a(b_T, b_T) &= \varepsilon \int_T (\nabla b_T)^\top \cdot \nabla b_T \; + \; \int_T b_T^2 \\ &= 6|T|\,\varepsilon \int_{\bar{T}} (\bar{\nabla} b_{\bar{T}})^\top \cdot A_T^{-1} A_T^{-\top} \cdot \bar{\nabla} b_{\bar{T}} \; + \; 6|T| \int_{\bar{T}} b_{\bar{T}}^2 \quad . \end{aligned}$$

with $b_{\bar{T}}$ being the element bubble function for the standard tetrahedron $\bar{T}$. A straight–forward computation yields

$$\int_{\bar{T}} b_{\bar{T}}^2 = \frac{4096}{155\,925} \quad .$$

In order to obtain the remaining integral, define the matrices

$$\begin{aligned} & M := (m_{ij})_{i,j=1}^3 = A_T^{-1} A_T^{-\top} \\ \text{and} \quad & N := (n_{ij})_{i,j=1}^3 = \int_{\bar{T}} \bar{\nabla} b_{\bar{T}} \cdot (\bar{\nabla} b_{\bar{T}})^\top \quad , \qquad \text{i.e.} \quad n_{ij} = \int_{\bar{T}} \bar{\partial}_{\bar{x}_1} b_{\bar{T}} \cdot \bar{\partial}_{\bar{x}_i} b_{\bar{T}} \end{aligned}$$

and observe that

$$\int_{\bar{T}} (\bar{\nabla} b_{\bar{T}})^\top \cdot A_T^{-1} A_T^{-\top} \cdot \bar{\nabla} b_{\bar{T}} \; = \; N : M \; = \; \sum_{i,j=1}^{3} n_{ij} \cdot m_{ij} \quad .$$

The matrices A_T and M are determined by the geometry of T whereas N can be computed directly giving

$$N = \frac{2048}{2835} \cdot \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad .$$

Hence one only has to determine A_T and $|T|$ and perform the aforementioned operations to obtain $a(b_T, b_T)$.

The remaining values $a(\cdot,\cdot)$ are computed similarly, thus a brief description of the result should suffice.

$\boldsymbol{a(b_T, b_{E,\delta})}$: Use an affine linear transformation $F_{\tilde{A}} : \bar{T} \to T$ such that $\bar{E} = F_{\tilde{A}}^{-1}(E)$ lies in the $\bar{x}_1\bar{x}_2$ plane. The parameter $\delta \in (0,1]$ of the squeezed face bubble function can be arbitrary here. Then

$$\begin{aligned} \int_{\bar{T}} b_{\bar{T}} \cdot b_{\bar{E},\delta} &= \frac{4}{4725}\delta^2(-2\delta^3 + 15\delta^2 - 42\delta + 47) \\ \int_{\bar{T}} \bar{\nabla} b_{\bar{T}} \cdot (\bar{\nabla} b_{\bar{E},\delta})^\top &= \begin{bmatrix} 2c & c & c \\ c & 2c & c \\ c & c & d \end{bmatrix} \\ \text{with} \quad c &:= -\frac{4}{35}\delta^2(\delta - 4) \quad \text{and} \quad d := \frac{8}{105}(\delta - 1)(\delta^2 - 7\delta + 18) \quad . \end{aligned}$$

$\boldsymbol{a(b_{E,\delta}, b_{E,\delta})}$: Utilize the same transformation $F_{\tilde{A}}$ as before which implies for arbitrary $\delta \in (0, 1]$

$$\begin{aligned} \int_{\bar{T}} b^2_{\bar{E},\delta} &= \frac{9}{560}\delta \\ \int_{\bar{T}} \bar{\nabla} b_{\bar{E},\delta} \cdot (\bar{\nabla} b_{\bar{E},\delta})^\top &= \begin{bmatrix} 2c & c & c \\ c & 2c & c \\ c & c & d \end{bmatrix} \\ \text{with} \qquad c &:= \frac{81}{280}\delta \qquad \text{and} \qquad d := \frac{27}{140}(\delta + 2/\delta) \qquad . \end{aligned}$$

$\boldsymbol{a(b_{E_1,\delta_1}, b_{E_2,\delta_2})}$: Since we propose to use $\delta_i := \tilde{\delta}_{E_i} \leq 1/4$, the supports of both squeezed face bubble functions are distinct, thus $a(b_{E_1,\delta_1}, b_{E_2,\delta_2}) = 0$.

Collecting all the previous results, the local stiffness matrix K_T can now be assembled. The right–hand side is computed similarly (actually, the procedure is even simpler since the integrals do not involve derivatives). In the next paragraphs we show that this direct computation of the local problem is indeed much cheaper than the numerical integration rules.

5.2 Computational effort

The comparison will not investigate every detail and every possible optimization, as the difference between both approaches will turn out to be overwhelming. Even more, a precise operation count would be computer dependent. For example present processors may be able to combine one multiplication and one addition to a single operation.

Computational effort for direct computation

The suggested approach utilizes four different transformations $F_{\tilde{A}}$ on the element T and four transformations on each neighbouring tetrahedron (i.e. to compute $b_{E,\delta}$ on this neighbour). Hence $|\det\tilde{A}|$ has to be computed five times, and $\tilde{A}^{-1}$ and $\tilde{A}^{-1} \cdot \tilde{A}^{-\top}$ are to be computed eight times. The operation count is roughly

Operation	Operation count	Total
$\det\tilde{A}$	$5 \times (\ 9^* \ \ 5+) =$	$(\ 45^* \ \ 25+)$
$\tilde{A}^{-1}$	$8 \times (19^* \ \ 9+) =$	$(152^* \ \ 72+)$
$\tilde{A}^{-1} \cdot \tilde{A}^{-\top}$ (symmetric)	$8 \times (18^* \ 12+) =$	$(144^* \ \ 96+)$
	$\Sigma:$	$(341^* \ 193+)$

where $5 \times (9^* \ 5+)$ stands for 9 multiplications and 5 additions which are performed five times.

Next, the values of $\int_{\bar{T}} b_{\bar{T}}^2$, $\int_{\bar{T}} \bar{\nabla} b_{\bar{T}} \cdot (\bar{\nabla} b_{\bar{T}})^\top$ etc are determined. Some of these values (which contain $b_{E,\delta}$) depend on δ. The computational effort is roughly

$\int_{\bar{T}} b_{\bar{T}}^2$		$(\ 0^* \ \ 0+)$
$\int_{\bar{T}} b_{\bar{T}} \cdot b_{\bar{E},\delta}$	$4 \times (8^* \ 3+) =$	$(32^* \ 12+)$
$\int_{\bar{T}} b_{\bar{E},\delta}^2$	$4 \times (1^* \ 0+) =$	$(\ 4^* \ \ 0+)$
$\int_{\bar{T}} \bar{\nabla} b_{\bar{T}} \cdot (\bar{\nabla} b_{\bar{T}})^\top$		$(\ 0^* \ \ 0+)$
$\int_{\bar{T}} \bar{\nabla} b_{\bar{T}} \cdot (\bar{\nabla} b_{\bar{E},\delta})^\top$	$4 \times (6^* \ 4+) =$	$(24^* \ 16+)$
$\int_{\bar{T}} \bar{\nabla} b_{\bar{E},\delta} \cdot (\bar{\nabla} b_{\bar{E},\delta})^\top$	$4 \times (1^* \ 1+) =$	$(16^* \ \ 4+)$
	$\Sigma :$	$(76^* \ 32+)$

Subsequently $N : M$ is to be determined, with N, M being symmetric matrices. The computational effort is approximately $(7^* \ 5+)$ which has to be repeated 9 times (i.e. once for each matrix entry of K_T). The final value of $a(\cdot,\cdot)$ is obtained by adding both sub–integrals and multiplying it by $|\mathrm{det}\tilde{A}|$. This adds $9 \times (1^* \ 1+)$.

Summarizing all results, the total effort required to assemble the local stiffness matrix is approximately

$$(341^* \ 193+) \ + \ (76^* \ 32+) \ + \ (63^* \ 45+) \ + \ (9^* \ 9+) \ = \ \mathbf{(489 * \quad 279+)} \qquad .$$

Computational effort for numerical integration

Here we will exemplarily investigate $a(b_T, b_T) = \int_T b_T^2 + \varepsilon \int_T (\nabla b_T)^2$. Computation by means of numerical integration is based on

$$\begin{aligned} \int_T b_T^2 &= 6|T| \int_{\bar{T}} b_{\bar{T}}^2 \ \approx \ 6|T| \cdot \sum_i \omega_i \cdot b_{\bar{T}}^2(\bar{x}_i) \\ \int_T (\nabla b_T)^2 &= 6|T| \int_{\bar{T}} (\bar{\nabla} b_{\bar{T}})^\top \cdot A_T^{-1} A_T^{-\top} \cdot \bar{\nabla} b_{\bar{T}} \\ &\approx 6|T| \cdot \sum_i \omega_i \cdot \Big((\bar{\nabla} b_{\bar{T}})^\top \cdot A_T^{-1} A_T^{-\top} \cdot \bar{\nabla} b_{\bar{T}} \Big)(\bar{x}_i) \end{aligned}$$

where $(\omega_i, \bar{x}_i)_i$ denotes some numerical integration rule for the standard tetrahedron $\bar{T}$ with weights ω_i and evaluation points $\bar{x}_i$. Exactly as for the direct computation above, one requires the matrices A_T and A_T^{-1} (computational effort is $(19^* \ \ 9+)$) as well as $6|T| = |\mathrm{det} A_T|$ (which leads to $(9^* \ 5+)$).

Consider $\int \bar{T} b_{\bar{T}}^2$ next. Since $b_{\bar{T}} \in \mathbb{P}^4(\bar{T})$ one requires an integration rule which is exact for $\mathbb{P}^8(\bar{T})$. The simplest rule that we know of involves 43 evaluation points [9]. The evaluation of

$$\omega_i \cdot b_{\bar{T}}^2(\bar{x}_i) \ = \ \omega_i \cdot (256 \cdot \lambda_1 \lambda_2 \lambda_3 \lambda_4)^2(\bar{x}_i)$$

at a single point $\bar{x}_i$ requires $(6^* \ 3+)$. Hence the total amount for $\int_T b_T^2$ is about $43 \times (6^* \ 4+) = (258^* \ 172+)$. Similarly $\int_T (\nabla b_T)^2$ is investigated where an integration rule with 24 evaluation points suffices [12]. After some consideration one ends up with an effort of roughly (480* 384+). Thus the total effort to compute $a(b_T, b_T)$ amounts approximately to (750* 550+). We expect the other eight scalar products $a(\cdot,\cdot)$ to be cheaper because of the smaller polynomial degree. Then, however, the computational domain involves $T_{E,\delta}$ which requires more considerations. Altogether the effort to compute the whole matrix of the local problem is likely to be of order

$$(\mathcal{O}(5000) * \quad \mathcal{O}(4000)+) \qquad .$$

Even without a precise calculation of the computational effort for the numerical integration it is absolutely clear that this approach is far too expensive. Equivalently, if the computational effort should be of the same size ($\mathcal{O}(500) * \mathcal{O}(300)+$) as for the direct computation then only one tenth of the required evaluation points can be used. This would render the integrals and, subsequently, the matrix to be very inaccurate and thus useless. Hence direct computation of the local problem is a must.

Furthermore we note that our procedure for the direct computation is very similar to the computation of the local problem for the *Poisson equation*, cf. [15]. The computational effort is roughly the same, i.e. the singularly perturbed character of our differential equation here is no disadvantage.

6 Numerical experiments

Here we investigate the performance of the local problem error estimator $\eta_{\varepsilon,\mathrm{D},T}$ of (24) by means of numerical experiments. We utilize a model problem which has already been applied in [16] to analyse the element based residual error estimator $\eta_{\varepsilon,\mathrm{R},T}$, and which has been employed in [17] to investigate a face based residual error estimator. Thus even the interesting comparison between different types of error estimators is possible.

Let us consider the 3D model problem

$$-\varepsilon\Delta u + u = 0 \quad \text{in } \Omega := (0,1)^3 \qquad , \qquad u = u_0 \quad \text{on } \Gamma_\mathrm{D} := \partial\Omega$$

where the perturbation parameter is set to $\varepsilon = 10^{-4}$. Prescribe the exact solution

$$u = e^{-x/\sqrt{\varepsilon}} + e^{-y/\sqrt{\varepsilon}} + e^{-z/\sqrt{\varepsilon}} \qquad .$$

which displays typical boundary layers along the planes $x = 0$, $y = 0$, and $z = 0$. The Dirichlet boundary data u_0 are chosen accordingly.

We apply the finite element method with a sequence of meshes, each of which is the tensor product of three one–dimensional Bakhvalov–like meshes [7] with 2^k intervals in [0,1], $k = 1\ldots6$. To describe the 1D nodal distribution properly, denote the transition point of the boundary layer by $\tau := \sqrt{\varepsilon}|\ln\sqrt{\varepsilon}|$. Then 2^{k-1} nodes are *exponentially* distributed in the boundary layer interval $[0,\tau]$ whereas the remaining interval $[\tau, 1]$ is divided into 2^{k-1} *equidistant* intervals, cf. Figure 5. More precisely, the (1D) nodal coordinate of the m-th node is

$$x_m := \begin{cases} -\beta\sqrt{\varepsilon}\ln\left[1 - \frac{m}{2^{k-1}}(1 - e^{-\tau/\beta/\sqrt{\varepsilon}})\right] & \text{for } m = 0\ldots 2^{k-1}, \beta = 3/2 \\ \tau + (1-\tau)\cdot\left(\frac{m}{2^{k-1}} - 1\right) & \text{for } m = 2^{k-1}+1\ldots 2^k \end{cases} .$$

Note that the original (1D) Bakhvalov mesh utilizes a slightly different transition point τ. Furthermore we do not know whether these tensor product type meshes are optimal (which, of course, also depends on the optimality criterion).

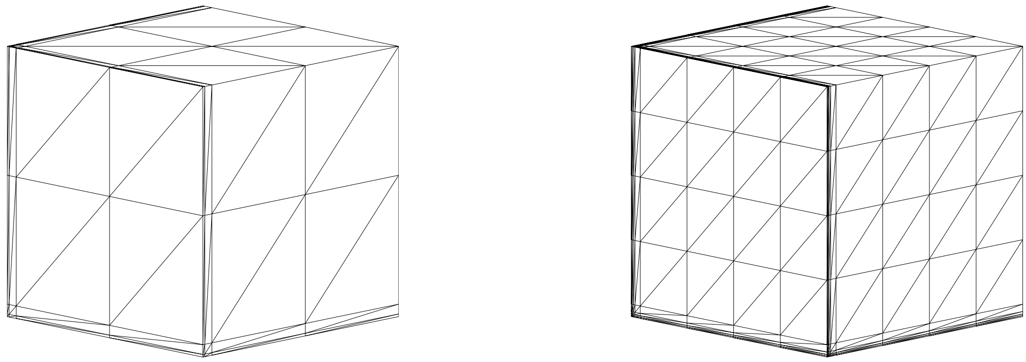

Figure 5: Mesh 2 – Mesh 3

The first table below presents some information about the meshes and their maximum aspect ratio. The last two columns give the exact value of the matching function $m_1(u - u_h, \mathcal{T}_h)$ as well as its approximation $m_1^R(u_h, \mathcal{T}_h)$ from (20).

Mesh k	# Elements	Aspect ratio	$m_1(u-u_h,\mathcal{T}_h)$	$m_1^R(u_h,\mathcal{T}_h)$
1	48	29.4	1.55	1.68
2	384	69.5	1.62	1.52
3	3 072	82.6	1.69	1.69
4	24 576	88.6	1.88	1.86
5	196 608	91.5	2.37	2.03
6	1 572 864	92.9	3.04	2.29

Since the size of m_1 is comparatively small and grows only mildly, the chosen meshes discretize the problem sufficiently well. Additionally the approximation m_1^R is satisfactorily close to the exact value. Hence the matching function and its approximation are useful tools for the theoretical analysis as well as for assessing the mesh quality in numerical computations. This topic has already been discussed for the Poisson equation in [14].

Next our main analytical results are to be confirmed numerically, namely the error bounds of Theorem 7. Therefore we present the ratios of left–hand side and right–hand side of (36) and (35), respectively, in the table below. Note that the approximation terms $\zeta_{\varepsilon,T}$ and ζ_ε vanish here.

Mesh k	$\|\|\|u-u_h\|\|\|$	$\dfrac{\|\|\|u-u_h\|\|\|}{m_1\cdot\eta_{\varepsilon,\mathrm{D}}}$	$\max\limits_{T\in\mathcal{T}_h}\dfrac{\eta_{\varepsilon,\mathrm{D},T}}{\|\|\|u-u_h\|\|\|_{\omega_T}}$
1	$0.154E+0$	1.102	0.411
2	$0.536E-1$	0.810	0.568
3	$0.229E-1$	0.705	0.611
4	$0.110E-1$	0.605	0.607
5	$0.553E-2$	0.474	0.597
6	$0.282E-2$	0.371	0.582

Start with the second column which yields a convergence rate of the error $|||u-u_h|||$ of approximately $N^{-0.324}$, with N being the number of elements. This is almost the optimal rate of $N^{-1/3}$ which indicates that the meshes under consideration discretize the singular problem well. Next, the ratios of the third and fourth column are related to the upper and lower error bound, respectively. These ratios are bounded from above and thus confirm the predictions of Theorem 7. Note that from a practical point of view the moderately decreasing values of the upper error bound (third column) imply that the error is increasingly overestimated.

In the last table we examine the equivalence of the local problem error estimator and the residual error estimator, as described in Theorem 6. Again we compute the ratios related to (29) and (30). Since all values are bounded from above, this impressively underpins our analytical results.

Mesh k	$\max\limits_{T\in\mathcal{T}_h}\dfrac{\eta_{\varepsilon,\mathrm{D},T}}{\Big(\sum\limits_{T'\subset\omega_T}\eta^2_{\varepsilon,\mathrm{R},T'}\Big)^{1/2}}$	$\max\limits_{T\in\mathcal{T}_h}\dfrac{\eta_{\varepsilon,\mathrm{R},T}}{\Big(\sum\limits_{T'\subset\omega_T}\eta^2_{\varepsilon,\mathrm{D},T'}\Big)^{1/2}}$
1	0.266	1.102
2	0.280	4.977
3	0.338	4.889
4	0.293	4.819
5	0.203	4.731
6	0.178	4.743

Since the same numerical example has been considered for the residual error estimator of [16] we can easily compare both estimator. Qualitatively both estimators behave similarly whereas from a quantitative viewpoint one observes roughly $\eta_{\varepsilon,\mathrm{R}}\approx 4\cdot\eta_{\varepsilon,\mathrm{D}}$. Furthermore the residual error estimator $\eta_{\varepsilon,\mathrm{R}}$ overestimates the true error more than the local problem error estimator $\eta_{\varepsilon,\mathrm{D}}$ does. This indeed can be expected since the derivation of $\eta_{\varepsilon,\mathrm{R}}$ requires more intermediate steps (such as interpolation estimates and Cauchy Schwarz inequalities).

7 Summary

We have considered a singularly perturbed reaction–diffusion problem and proposed a new error estimator that can be applied to *anisotropic* finite element meshes. The rigorous analysis confirms that the error estimation is uniform in the small perturbation parameter. Furthermore tight error bounds are obtained provided the anisotropic mesh is chosen according to the anisotropy of the solution. Thus reliable and efficient error estimation is possible on anisotropic meshes.

Then a stable basis of the local problem has been derived and an additional, face oriented local problem error estimator has been proposed. Finally implementational aspects have been discussed and analysed. A numerical experiment complements the theory.

A Proof of Lemma 5

First we state Lemma 5 again.

Lemma 5 *The following relations below hold for all $v \in V_T$.*

$$\begin{aligned} \|v\|_{\omega_T} &\lesssim h_{min,T} \cdot \|\nabla v\|_{\omega_T} && (27)\\ \|v\|_E &\lesssim h_E^{-1/2}\, \delta_E^{-1/2} \cdot \min\{h_{min,T}, \delta_E\, h_E\} \cdot \|\nabla v\|_{\omega_T} \qquad \forall\, E \subset \partial T \quad , && (28)\end{aligned}$$

with δ_E from (21).

If T has at least two Neumann boundary faces then the constants can depend on the angles of the corners and edges of Ω (but do not depend on the triangulation $\mathcal{T}_h$ nor on T).

Proof: The proof here utilizes some key ideas that were already applied in [13, Lemma 3.5] and [15]. Our exposition here requires several non–trivial extensions which are due to the singularly perturbed problem, and the use of *squeezed* face bubble functions in particular. In order to facilitate the understanding of the proof, each major step will be given a distinctive name.

Set $T_0 := T$ and enumerate the remaining tetrahedra of $\omega_T \setminus T$ by $T_1 \dots T_k$. If T has boundary faces then $k < 4$. The faces of T are denoted accordingly by $E_i := T_i \cap T$.

Transformation: In order to prove (27) consider the tetrahedron T_i and rewrite $\|v\|_{T_i}$ by means of the transformation A_{T_i}, and $\|\nabla v\|_{T_i}$ via the transformations C_{T_i}, H_{T_i}. Utilizing $|T| \sim |T_i|$, and with certain abbreviations given below this yields

$$\begin{aligned} \|v\|^2_{T_i} &= 6|T_i| \cdot \|\bar{v}\|^2_{\bar{T}_i} \sim |T| \cdot r_i \\ \text{and} \qquad \|v\|^2_{\omega_T} &\sim |T| \cdot \sum_{i=0}^{k} r_i = |T| \cdot r \end{aligned}$$

where we have introduced

$$r_i := \|\bar{v}\|^2_{\bar{T}_i} \geq 0 \qquad \text{and} \qquad r := \sum_{i=0}^{k} r_i \geq 0 \qquad .$$

Similarly the term $\|\nabla v\|_{\omega_T}$ is transformed to give

$$\|\nabla v\|^2_{T_i} \overset{(5)}{=} \|H_{T_i}^{-1} C_{T_i}^\top \nabla v\|^2_{T_i} \;=\; 6|T_i| \cdot \|H_{T_i}^{-1} \cdot \hat{\nabla}\hat{v}\|^2_{\hat{T}_i}$$

$$\text{which implies} \qquad \|\nabla v\|^2_{\omega_T} \;=\; \sum_{i=0}^{k} \|\nabla v\|^2_{T_i} \;=\; \sum_{i=0}^{k} 6|T_i| \cdot \|H_{T_i}^{-1} \cdot \hat{\nabla}\hat{v}\|^2_{\hat{T}_i}$$

$$= \; 6 \cdot \sum_{i=0}^{k} |T_i| \cdot h^{-2}_{min,T_i} \cdot \|\mathrm{diag}(\gamma_{1,i}, \gamma_{2,i}, 1) \cdot \hat{\nabla}\hat{v}\|^2_{\hat{T}_i}$$

$$\sim \; h^{-2}_{min,T} \cdot |T| \cdot \sum_{i=0}^{k} s_i \;=\; h^{-2}_{min,T} \cdot |T| \cdot s$$

$$\text{with} \qquad \gamma_{1,i} \;:=\; h_{min,T_i}/h_{1,T_i} \qquad \text{and} \qquad \gamma_{2,i} \;:=\; h_{min,T_i}/h_{2,T_i}$$

$$s_i \;:=\; \|\mathrm{diag}(\gamma_{1,i}, \gamma_{2,i}, 1) \cdot \hat{\nabla}\hat{v}\|^2_{\hat{T}_i} \;\geq 0 \quad \text{and} \quad s \;:=\; \sum_{i=0}^{k} s_i \;\geq 0\,.$$

A rough outline of the proof is as follows. Realize first that r and s depend on various variables (e.g. on the geometry of T, the parameters $\delta_i := \delta_{E_i}$ etc). Then consider r, s over some *compact* set of variables. Since both terms turn out to be continuous, one can investigate their maxima and minima which eventually provide the assertion.

Extend domain of definition to a compact set: Let us start with the case where T has no Neumann boundary face. Assume further that T has m interior faces and $4 - m$ Dirichlet faces. The local space V_T is spanned by b_T and b_{E_i,δ_i}, $i = 1 \ldots m$. For our purpose we utilize an expansion of $v \in V_T$ where the squeezed face bubble functions are additionally scaled by $\delta_i^{-1/2}$, namely

$$v \;=\; \beta_0 \cdot b_T \;+\; \sum_{i=1}^{m} \beta_i \cdot \delta_i^{-1/2} b_{E_i,\delta_i} \qquad \beta_i \in \mathbb{R} \qquad .$$

Without loss of generality assume $v \not\equiv 0$ and $\sum_{i=0}^{m} \beta_i^2 \;=\; 1$. After transformation via A_{T_i} the representation of $\bar{v}$ becomes

$$\bar{v}|_{\bar{T}_i} = v|_{T_i} \circ A_{T_i} \;=\; \begin{cases} \beta_0 \cdot b_{\bar{T}} \;+\; \sum\limits_{i=1}^{m} \beta_i \cdot \delta_i^{-1/2} b_{\bar{E}_i,\delta_i} & \text{for } i = 0 \text{ (i.e. on } \bar{T}) \\ \beta_i \cdot \delta_i^{-1/2} b_{\bar{E}_i,\delta_i} & \text{for } i = 1 \ldots m \text{ (i.e. on } \bar{T}_i) \end{cases} \quad .$$

Hence $\bar{v}$ depends on $\beta_0 \ldots \beta_m$ and $\delta_1 \ldots \delta_m$, $\delta_i \in (0, 1]$. Note further that δ_i influences $r_0 = \|\bar{v}\|^2_{\bar{T}}$ and $r_i = \|\bar{v}\|^2_{\bar{T}_i}$ but not the other values r_j.

Next $\bar{v}$ is to be considered over a *compact* set. Thus introduce

$$B := \Big\{ (\beta_0, \ldots, \beta_m) \;:\; \sum_{i=0}^{m} \beta_i^2 \;=\; 1 \Big\} \qquad \text{and} \qquad D := \Big\{ (\delta_1, \ldots, \delta_m) \;:\; \delta_i \in [0, 1]\, \forall\, i \Big\} \quad .$$

The case $\delta_i = 0$ requires additional consideration. While the function $\delta_i^{-1/2} b_{E_i,\delta_i}$ has a well–defined meaning for $\delta_i \in (0,1]$, this is no longer true for $\delta_i \to 0$. Then $\operatorname{supp}(b_{E_i,\delta_i}|_{T_i}) = T_{i,E_i,\delta_i}$ degenerates, and $\delta_i^{-1/2} \to \infty$. Therefore the value of $r_j = \|\bar{v}\|^2_{\bar{T}_j}$ for $\delta_i = 0$ is defined as the limit for $\delta_i \to 0$:

$$r_j(\delta_i = 0) := \lim_{\delta_i \to 0} r_j(\delta_i) \qquad .$$

This limit is well–defined since the vanishing support T_{i,E_i,δ_i} of the squeezed face bubble function and its scaling factor $\delta_i^{-1/2}$ are exactly balanced. For the outer tetrahedra this can be easily seen by utilizing two transformations, namely via $A_{T_i} : \bar{T} \to T_i$ and via $F^{-1}_{T_i,E_i,\delta_i} : T_{i,E_i,\delta_i} \to \bar{T}_i$. By using $|\det(A_{T_i})| = 6|T|$ and $|\det(F_{T_i,E_i,\delta_i})| = 6\delta_i|T|$ one obtains

$$\begin{array}{rcl} r_i & = & \beta_i^2\, \|\delta_i^{-1/2} b_{\bar{E}_i,\delta_i}\|^2_{\bar{T}_i} \\ & \overset{A_{T_i}}{=} & \beta_i^2\, \delta_i^{-1} \cdot (6|T|)^{-1} \|b_{E_i,\delta_i}\|^2_{T_i} \;=\; \beta_i^2\, \delta_i^{-1} \cdot (6|T|)^{-1} \|b_{E_i,\delta_i}\|^2_{T_{i,E_i,\delta_i}} \\ & \overset{F^{-1}_{T_i,E_i,\delta_i}}{=} & \beta_i^2\, \delta_i^{-1} \cdot (6|T|)^{-1} \cdot 6\delta_i |T| \cdot \|b_{\bar{E}_i}\|^2_{\bar{T}_i} \;=\; \dfrac{9}{560}\beta_i^2 \;\sim\; \beta_i^2 \\ \text{and} \qquad \lim\limits_{\delta_i \to 0} r_i & = & \dfrac{9}{560}\beta_i^2 \qquad . \end{array}$$

Hence $\lim_{\delta_i \to 0} r_i$ exists. For r_0 proceed similarly.

Consider Maximum and Minimum: As a consequence we can consider r_i and r on $B \times D$, and r_i, r vary continuously over that compact set. Therefore r attains its maximum and minimum. To show that this minimum is positive, assume the contrary which implies $r_i = 0$ for all $i = 0 \ldots m$. On the outer tetrahedra T_i, $i = 1 \ldots m$, proceed exactly as in the last paragraph to obtain

$$0 = r_i \;=\; \frac{9}{560}\beta_i^2$$

which implies $\beta_i = 0$, $i = 1 \ldots m$. On the main tetrahedron T then $\bar{v}$ is reduced to $\bar{v} = \beta_0\, b_{\bar{T}}$ giving

$$0 = r_0 = \|\bar{v}\|^2_{\bar{T}} \;=\; \frac{4096}{155\,925}\beta_0^2 \;\sim\; \beta_0^2$$

and $\beta_0 = 0$ too. This contradicts $\sum_{i=0}^{m} \beta_i^2 = 1$, hence

$$\min_{B \times D} r \;>\; 0 \qquad .$$

Together with $\max\limits_{B \times D} r \sim 1$ we obtain

$$r \sim 1$$

or, equivalently,

$$\|v\|^2_{\omega_T} \;\sim\; |T| \cdot \sum_{i=0}^{m} \beta_i^2 \qquad . \tag{40}$$

Investigation of s: The investigation of s and s_i relies on the same basic ideas as before. The details however are much more technical because derivatives are involved (i.e. ∇v) and the transformation C_{T_i} is applied.

Consider $s_i = \|\mathrm{diag}(\gamma_{1,i}, \gamma_{2,i}, 1) \cdot \hat\nabla \hat v\|^2_{\hat T_i} \geq 0$ which depends on $\hat T_i$, $\gamma_{j,i}$ and $\hat v$. The restrictions on $\hat T_i$ and T_i imply $0 < \gamma_{1,i}, \gamma_{2,i} \leq 1$ and, for the nodal coordinates of $\hat T_i$, $0 < \hat x_{2,T_i} \leq 1/2, \quad 0 < \hat x_{3,T_i} < 1, \quad -1 < \hat y_{3,T_i} < 1$. Similar as before we omit the actual meaning that stands behind s_i, and view it instead as a purely analytical term that depends on $\hat x_{j,T_i}$, $\hat y_{3,T_i}$, $\gamma_{j,i}$, δ_j and β_j. Next consider s_i over the compact set $X_i \times G_i \times D \times B$, with

$$
\begin{aligned}
X_i &:= \Big\{ (\hat x_{2,T_i}, \hat x_{3,T_i}, \hat y_{3,T_i}) \,:\, 0 \leq \hat x_{2,T_i} \leq \frac{1}{2}\,,\, 0 \leq \hat x_{3,T_i} \leq 1\,,\, -1 \leq \hat y_{3,T_i} \leq 1 \Big\},\\
G_i &:= \Big\{ (\gamma_{1,i}, \gamma_{2,i}) \,:\quad 0 \leq \gamma_{1,i}, \gamma_{2,i} \leq 1 \Big\} \qquad .
\end{aligned}
$$

It is obvious that s_i is continuous on X_i, G_i, B and for $\delta_i \in (0,1]$. Note again that δ_i influences only s_0 and s_i. The only cause for discontinuity of s_i is $\delta_i \to 0$ which may lead to $s_i \to \infty$ (because $\delta_i^{-1/2} \to \infty$ and $|\hat\nabla \hat b_{E_i,\delta_i}| \to \infty$, see below). Nevertheless such a discontinuity does not disturb our analysis since we want to bound s_i from below. For a precise investigation we define again

$$
s_j(\delta_i = 0) := \lim_{\delta_i \to 0} s_j(\delta_i)
$$

and consider then the term $\min\{1, s_i\}$ which is continuous for $\delta_i \in [0,1]$.

Since $s = \sum_{i=0}^m s_i$, this term is continuous as well, and it attains its minimum over the compact set

$$
K \;:=\; \mathop{\times}_{i=0}^{m} X_i \quad\times\quad \mathop{\times}_{i=0}^{m} G_i \quad\times\quad D \quad\times\quad B \qquad .
$$

In order to show that this minimum is positive assume the contrary, namely $s = s_i = 0$ for all $i = 0 \ldots m$. Start with any of the outer tetrahedra $T_i, i = 1 \ldots m$. The representation of $\hat v$ there is $\hat v|_{\hat T_i} = \beta_i \delta_i^{-1/2} \cdot \hat b_{E_i,\delta_i}|_{\hat T_i}$. Then

$$
\begin{aligned}
0 = s_i \;&=\; \|\mathrm{diag}(\gamma_{1,i}, \gamma_{2,i}, 1) \cdot \hat\nabla \hat v\|^2_{\hat T_i}\\
&\geq\; \|\mathbf{e}_3^\top \cdot \hat\nabla \hat v\|^2_{\hat T_i} \;=\; \beta_i^2\, \delta_i^{-1}\, \|\mathbf{e}_3^\top \cdot \hat\nabla \hat b_{E_i,\delta_i}\|^2_{\hat T_i} \qquad ,
\end{aligned}
$$

with $\mathbf{e}_3 := (0,0,1)^\top$. The latter norm is analysed similarly as for r_i by using two transformations via $C_{T_i} : \hat T \to T_i$ and via $F^{-1}_{T_i,E_i,\delta_i} : T_{i,E_i,\delta_i} \to \bar T_i$. In contrast to r_i, however, we cannot evaluate s_i exactly but bound it instead. From $|\det(C_{T_i})| = 6|T|$ and $|\det(F_{T_i,E_i,\delta_i})| = 6\delta_i|T|$ one derives

$$
\begin{aligned}
\delta_i^{-1}\, \|\mathbf{e}_3^\top \cdot \hat\nabla \hat b_{E_i,\delta_i}\|^2_{\hat T_i} \;&\overset{C_{T_i}}{=}\; \delta_i^{-1} \cdot (6|T|)^{-1} \|\mathbf{e}_3^\top \cdot C_{T_i}^\top \nabla b_{E_i,\delta_i}\|^2_{T_i}\\
&=\; \delta_i^{-1} \cdot (6|T|)^{-1} \|\mathbf{e}_3^\top \cdot C_{T_i}^\top \nabla b_{E_i,\delta_i}\|^2_{T_{i,E_i,\delta_i}}\\
&\overset{F^{-1}_{T_i,E_i,\delta_i}}{=}\; \delta_i^{-1} \cdot (6|T|)^{-1} \cdot 6\delta_i|T| \cdot \|\mathbf{e}_3^\top \cdot C_{T_i}^\top F^{-\top}_{T_i,E_i,\delta_i} \bar\nabla b_{\bar E_i}\|^2_{\bar T_i}\\
&=\; \|(F^{-1}_{T_i,E_i,\delta_i} C_{T_i} \cdot \mathbf{e}_3)^\top \cdot \bar\nabla b_{\bar E_i}\|^2_{\bar T_i} \qquad .
\end{aligned}
$$

Recalling the definition of C_{T_i} from (4) yields $C_{T_i}\mathbf{e}_3 = \mathbf{p}_{3,T_i}$ which is a vector from a vertice to the opposite face in the tetrahedron T_i, see Figure 1. Hence $F^{-1}_{T_i,E_i,\delta_i} C_{T_i} \cdot \mathbf{e}_3$ is a vector from a vertice to the opposite face in the tetrahedron $F^{-1}_{T_i,E_i,\delta_i} T_i$. Using $T_i \supset T_{i,E_i,\delta_i}$ one obtains $F^{-1}_{T_i,E_i,\delta_i} T_i \supset F^{-1}_{T_i,E_i,\delta_i} T_{i,E_i,\delta_i} \equiv \bar{T}_i$ and thus

$$|F^{-1}_{T_i,E_i,\delta_i} C_{T_i} \cdot \mathbf{e}_3|_{\mathbb{R}^3} \;\geq\; \varrho(\bar{T}_i) \;=\; \frac{1}{3+\sqrt{3}} \;\sim\; 1$$

where $\varrho(\bar{T}_i)$ denotes the diameter of the inscribed ball of $\bar{T}_i$. Then

$$\|(F^{-1}_{T_i,E_i,\delta_i} C_{T_i} \cdot \mathbf{e}_3)^\top \cdot \bar{\nabla} b_{\bar{E}_i}\|^2_{\bar{T}_i} \;\gtrsim\; \min_{|\mathbf{q}|_{\mathbb{R}^3}=1} \|\mathbf{q}^\top \cdot \bar{\nabla} b_{\bar{E}_i}\|^2_{\bar{T}_i} \;=\; 81/280 \;\sim\; 1 \qquad .$$

Summarizing the previous results, we end up with

$$0 = s_i \;\geq\; \beta_i^2\, \delta_i^{-1}\, \|\mathbf{e}_3^\top \cdot \hat{\nabla} \hat{b}_{E_i,\delta_i}\|^2_{\hat{T}_i} \;\gtrsim\; \beta_i^2 \qquad .$$

This holds for $\delta_i \in (0,1]$ and therefore also for the limit $\delta_i = 0$. Hence one concludes

$$\beta_i \;=\; 0 \qquad \forall i = 1 \dots m \qquad .$$

Next consider the main tetrahedron T where v is now reduced to $v = \beta_0 \cdot b_T$. Then

$$0 = s_0 \;\geq\; \beta_0^2\, \|\mathbf{e}_3^\top \cdot \hat{\nabla} b_{\hat{T}}\|^2_{\hat{T}}$$

immediately implies $\beta_0 = 0$ which contradicts the assumption $\sum_{i=1}^m \beta_i^2 = 1$. Therefore the minimum of s is positive giving

$$s \;\gtrsim\; 1 \;\sim\; r$$

which provides the assertion.

***T* has Neumann faces:** In this case ω_T consists of less than five tetrahedra, and $\dim V_T < 5$. Although the representation of v changes as well, the main ideas from above can still be applied to show the assertion. Thus we omit the proof.

It is noteworthy that the case of two or more Neumann boundary faces of T gives rise to a particular phenomenon. If one can guarantee $\delta_i \leq \delta_* < 1\ \forall\, i$ (with some parameter δ_* which is the same for all elements) then the resulting inequality is as before. Otherwise the inequality constant in (27) may depend on the angles of the corners and edges of Ω but do not depend on the triangulation $\mathcal{T}_h$ nor on T, cf. also [13, Lemma 3.5].

Proof of (28): Assume first that E_i is an interior face, and consider the corresponding outer tetrahedron T_i. Apply (14) with $\varphi_E \equiv 1$ to obtain

$$\|\nabla b_{E_i,\delta_i}\|_{T_i} \;\sim\; \delta_i^{1/2} \cdot h^{1/2}_{E_i,T_i} \cdot \min\{\delta_i\, h_{E_i,T_i}\,,\, h_{min,T_i}\}^{-1} \cdot |E_i|^{1/2} \qquad .$$

Together with $\|b_{E_i,\delta_i}\|_{E_i} = \|b_{E_i}\|_{E_i} \sim |E_i|^{1/2}$ this yields immediately

$$\|b_{E_i,\delta_i}\|_{E_i} \;\sim\; \delta_i^{-1/2} \cdot h^{-1/2}_{E_i,T_i} \cdot \min\{\delta_i\, h_{E_i,T_i}\,,\, h_{min,T_i}\} \cdot \|\nabla b_{E_i,\delta_i}\|_{T_i} \qquad .$$

From $v|_{T_i} = \beta_i \cdot b_{E_i,\delta_i}|_{T_i}$ and $h_{E_i,T_i} \sim h_{E_i}$, $h_{min,T_i} \sim h_{min,T}$ one concludes

$$\begin{aligned} \|v\|_{E_i} &\sim \delta_i^{-1/2} \cdot h_{E_i,T_i}^{-1/2} \cdot \min\{\delta_i\, h_{E_i,T_i}\,,\; h_{min,T_i}\} \cdot \|\nabla v\|_{T_i} \\ &\lesssim \delta_i^{-1/2} \cdot h_{E_i}^{-1/2} \cdot \min\{\delta_i\, h_{E_i}\,,\; h_{min,T}\} \cdot \|\nabla v\|_{\omega_T} \end{aligned}$$

which proves the assertion.

If E_i is a Dirichlet face then $v|_{E_i} \equiv 0$, and (28) holds trivially. Finally, if E_i is a Neumann face then the proof becomes more technical since no outer tetrahedron T_i exists. Then one has to utilize similar ideas as for proving (27). The details are omitted here. ■

References

[1] M. Ainsworth and I. Babuška. Reliable and robust a posteriori error estimation for singularly perturbed reaction–diffusion problems. *SIAM J. Num. Anal.*, 36(2):331–353, 1999.

[2] M. Ainsworth and J. Oden. *A posteriori error estimation in finite element analysis.* Wiley, 2000.

[3] L. Angermann. Balanced a-posteriori error estimates for finite volume type discretizations of convection-dominated elliptic problems. *Computing*, 55(4):305–323, 1995.

[4] T. Apel and G. Lube. Anisotropic mesh refinement in stabilized Galerkin methods. *Numer. Math.*, 74(3):261–282, 1996.

[5] T. Apel and S. Nicaise. The finite element method with anisotropic mesh grading for elliptic problems in domains with corners and edges. *Math. Methods Appl. Sci.*, 21:519–549, 1998.

[6] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15(4):736–754, 1978.

[7] N. S. Bakhvalov. Optimization of methods for the solution of boundary value problems in the presence of a boundary layer. *Zh. Vychisl. Mat. Mat. Fiz.*, 9:841–859, 1969. In Russian.

[8] R. E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comput.*, 44(170):283–301, 1985.

[9] M. Beckers. *Numerical integration in high dimensions.* PhD thesis, Katholieke Universiteit Leuven, 1992.

[10] P. G. Ciarlet. *The finite element method for elliptic problems.* North-Holland, Amsterdam, 1978.

[11] M. Dobrowolski, S. Gräf, and C. Pflaum. On a posteriori error estimators in the finite element method on anisotropic meshes. *Electronic Transactions Num. Anal.*, 8:36–45, 1999.

[12] P. Keast. Moderate–degree tetrahedral quadrature formulas. *Comput. Methods Appl. Mech. Engrg.*, 55:339–348, 1986.

[13] G. Kunert. *A posteriori error estimation for anisotropic tetrahedral and triangular finite element meshes.* Logos Verlag, Berlin, 1999. Also PhD thesis, TU Chemnitz, http://archiv.tu-chemnitz.de/pub/1999/0012/index.html.

[14] G. Kunert. An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes. *Numer. Math.*, 86(3):283–303, 2000. DOI 10.1007/s002110000170.

[15] G. Kunert. A local problem error estimator for anisotropic tetrahedral finite element meshes. *accepted by SIAM J. Numer. Anal.*, 2000.

[16] G. Kunert. Robust a posteriori error estimation for a singularly perturbed reaction–diffusion equation on anisotropic tetrahedral meshes. *submitted to Adv. Comp. Math.*, 2000.

[17] G. Kunert and R. Verfürth. Edge residuals dominate a posteriori error estimates for linear finite element methods on anisotropic triangular and tetrahedral meshes. *Numer. Math.*, 86(2):283–303, 2000. DOI 10.1007/s002110000152.

[18] J. Peraire, M. Vahdati, K. Morgan, and O. C. Zienkiewicz. Adaptive remeshing for compressible flow computation. *J. Comp. Phys.*, 72:449–466, 1987.

[19] W. Rick, H. Greza, and W. Koschel. FCT-solution on adapted unstructured meshes for compressible high speed flow computations. In E. H. Hirschel, editor, *Flow simulation with high-performance computers I*, volume 38 of *Notes on Num. Fluid Mechanics*, pages 334–438. Vieweg, 1993.

[20] K. G. Siebert. An a posteriori error estimator for anisotropic refinement. *Numer. Math.*, 73(3):373–398, 1996.

[21] R. Verfürth. A posteriori error estimation and adaptive mesh–refinement techniques. *Journal of Computational and Applied Mathematics*, 50:67–83, 1994.

[22] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh–refinement techniques.* Wiley-Teubner, Chichester; Stuttgart, 1996.

[23] R. Verfürth. Robust a posteriori error estimators for singularly perturbed reaction–diffusion equations. *Numer. Math.*, 78:479–493, 1998.

[24] R. Vilsmeier and D. Hänel. Computational aspects of flow simulation in three dimensional, unstructured, adaptive grids. In E. H. Hirschel, editor, *Flow simulation with high-performance computers II*, volume 52 of *Notes on Num. Fluid Mechanics*, pages 431–446. Vieweg, 1996.

[25] O. C. Zienkiewicz and J. Wu. Automatic directional refinement in adaptive analysis of compressible flows. *Internat. J. Numer. Methods Engrg.*, 37:2189–2210, 1994.

Other titles in the SFB393 series:

00-01 G. Kunert. Anisotropic mesh construction and error estimation in the finite element method. January 2000.

00-02 V. Mehrmann, D. Watkins. Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils. January 2000.

00-03 X. W. Guan, U. Grimm, R. A. Römer, M. Schreiber. Integrable impurities for an open fermion chain. January 2000.

00-04 R. A. Römer, M. Schreiber, T. Vojta. Disorder and two-particle interaction in low-dimensional quantum systems. January 2000.

00-05 P. Benner, R. Byers, V. Mehrmann, H. Xu. A unified deflating subspace approach for classes of polynomial and rational matrix equations. January 2000.

00-06 M. Jung, S. Nicaise, J. Tabka. Some multilevel methods on graded meshes. February 2000.

00-07 H. Harbrecht, F. Paiva, C. Perez, R. Schneider. Multiscale Preconditioning for the Coupling of FEM-BEM. February 2000.

00-08 P. Kunkel, V. Mehrmann. Analysis of over- and underdetermined nonlinear differential-algebraic systems with application to nonlinear control problems. February 2000.

00-09 U.-J. Görke, A. Bucher, R. Kreißig, D. Michael. Ein Beitrag zur Lösung von Anfangs-Randwert-Problemen einschließlich der Materialmodellierung bei finiten elastisch-plastischen Verzerrungen mit Hilfe der FEM. März 2000.

00-10 M. J. Martins, X.-W. Guan. Integrability of the D_n^2 vertex models with open boundary. March 2000.

00-11 T. Apel, S. Nicaise, J. Schöberl. A non-conforming finite element method with anisotropic mesh grading for the Stokes problem in domains with edges. March 2000.

00-12 B. Lins, P. Meade, C. Mehl, L. Rodman. Normal Matrices and Polar Decompositions in Indefinite Inner Products. March 2000.

00-13 C. Bourgeois. Two boundary element methods for the clamped plate. March 2000.

00-14 C. Bourgeois, R. Schneider. Biorthogonal wavelets for the direct integral formulation of the heat equation. March 2000.

00-15 A. Rathsfeld, R. Schneider. On a quadrature algorithm for the piecewise linear collocation applied to boundary integral equations. March 2000.

00-16 S. Meinel. Untersuchungen zu Druckiterationsverfahren für dichteveränderliche Strömungen mit niedriger Machzahl. März 2000.

00-17 M. Konstantinov, V. Mehrmann, P. Petkov. On Fractional Exponents in Perturbed Matrix Spectra of Defective Matrices. April 2000.

00-18 J. Xue. On the blockwise perturbation of nearly uncoupled Markov chains. April 2000.

00-19 N. Arada, J.-P. Raymond, F. Tröltzsch. On an Augmented Lagrangian SQP Method for a Class of Optimal Control Problems in Banach Spaces. April 2000.

00-20 H. Harbrecht, R. Schneider. Wavelet Galerkin Schemes for 2D-BEM. April 2000.

00-21 V. Uski, B. Mehlig, R. A. Römer, M. Schreiber. An exact-diagonalization study of rare events in disordered conductors. April 2000.

00-22 V. Uski, B. Mehlig, R. A. Römer, M. Schreiber. Numerical study of eigenvector statistics for random banded matrices. May 2000.

00-23 R. A. Römer, M. Raikh. Aharonov-Bohm oscillations in the exciton luminescence from a semiconductor nanoring. May 2000.

00-24 R. A. Römer, P. Ziesche. Hellmann-Feynman theorem and fluctuation-correlation analysis of i the Calogero-Sutherland model. May 2000.

00-25 S. Beuchler. A preconditioner for solving the inner problem of the p-version of the FEM. May 2000.

00-26 C. Villagonzalo, R.A. Römer, M. Schreiber, A. MacKinnon. Behavior of the thermopower in amorphous materials at the metal-insulator transition. June 2000.

00-27 C. Mehl, V. Mehrmann, H. Xu. Canonical forms for doubly structured matrices and pencils. June 2000. S. I. Solov'ev. Preconditioned gradient iterative methods for nonlinear eigenvalue problems. June 2000.

00-29 A. Eilmes, R. A. Römer, M. Schreiber. Exponents of the localization lengths in the bipartite Anderson model with off-diagonal disorder. June 2000.

00-30 T. Grund, A. Rösch. Optimal control of a linear elliptic equation with a supremum-norm functional. July 2000.

00-31 M. Bollhöfer. A Robust ILU Based on Monitoring the Growth of the Inverse Factors. July 2000.

00-32 N. Arada, E. Casas, F. Tröltzsch. Error estimates for a semilinear elliptic control problem. July 2000.

00-33 T. Penzl. LYAPACK Users Guide. August 2000.

00-34 B. Heinrich, K. Pietsch. Nitsche type mortaring for some elliptic problem with corner singularities. September 2000.

00-35 P. Benner, R. Byers, H. Faßbender, V. Mehrmann, D. Watkins. Cholesky-like Factorizations of Skew-Symmetric Matrices. September 2000.

00-36 C. Villagonzalo, R. A. Römer, M. Schreiber, A. MacKinnon. Critical Behavior of the Thermoelectric Transport Properties in Amorphous Systems near the Metal-Insulator Transition. September 2000.

00-37 F. Milde, R. A. Römer, M. Schreiber. Metal-insulator transition in anisotropic systems. October 2000.

00-38 T. Stykel. Generalized Lyapunov Equations for Descriptor Systems: Stability and Inertia Theorems. October 2000.

00-39 G. Kunert. Robust a posteriori error estimation for a singularly perturbed reaction-diffusion equation on anisotropic tetrahedral meshes. November 2000.

The complete list of current and former preprints is available via
`http://www.tu-chemnitz.de/sfb393/preprints.html`.