TU Clausthal

Mathematik-Bericht 2010/1

Residual type a posteriori error estimates for upwinding finite volume approximations of elliptic boundary value problems

L. Angermann

Januar 2010

Institut für Mathematik, TU Clausthal, Erzstraße 1, D-38678 Clausthal-Zellerfeld, Germany

Residual type *a posteriori* error estimates for upwinding finite volume approximations of elliptic boundary value problems

Lutz Angermann*

Abstract: This article describes the extension of recent methods for *a posteriori* error control such as dual-weighted residual methods to node-centered finite volume discretizations of second order elliptic boundary value problems with upwinding.

Keywords: Finite volume methods, a posteriori error control, DWR method

2010 Mathematics Subject Classification: 65 N 08, 65 N 15, 65 N 30, 65 N 50

1 Introduction

In this paper, we give a short overview on recent *a posteriori* error estimates for node-centered finite volume discretizations of second-order elliptic PDEs in $d \in \{2,3\}$ independent variables.

Since finite volume methods do not possess, in general, the so-called *Galerkin-orthogonality* property, special attention is paid to the treatment of the resulting defect term. It is shown that the extension of both the classical residual *a posteriori* error estimates as well as the more recent dual-weighted *a posteriori* error estimates to finite volume discretizations is possible in a reasonable way.

We consider mainly Voronoi and Donald finite volume partitions on simplicial primary partitions of the domain, however the ideas can be extendend to more general primary partitions, in particular quadrilateral or hexahedral partitions (cf., e.g., [Ang06, Sect. 4.2]).

We consider the following boundary value problem:

$$\left\{ \begin{array}{rcll} -\nabla\cdot(A\nabla u) + b\cdot\nabla u + cu &=& f & \text{in } \Omega \\ u &=& 0 & \text{on } \Gamma \end{array} \right. , \tag{1}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded polygonal or hexahedral domain with a Lipschitzian boundary Γ. The coefficients in (1) are assumed to satisfy the following conditions:

(A1.1) $A \in W^1_\infty(\Omega),\ b = (b_1, \dots, b_d)^\top \in [W^1_\infty(\Omega)]^d,\ c \in W^1_\infty(\Omega),$
$f \in W^1_q(\Omega)$ with some $q > d$,

*Technische Universität Clausthal, Institut für Mathematik, Erzstr. 1, D–38678 Clausthal-Zellerfeld, Federal Republic of Germany, e-mail: lutz.angermann@tu-clausthal.de

(A1.2) $A \geq a_0 > 0$ on Ω, where a_0 does not depend on $x \in \Omega$,

(A1.3) $c - \frac{1}{2}\nabla\cdot b \geq a_1 > 0$ on Ω, where a_1 does not depend on $x \in \Omega$.

Using the notations $(w, v \in H^1(\Omega))$

$$\begin{aligned}
(w,v) &:= \int_\Omega wv\,dx,\\
(\nabla w, \nabla v) &:= \int_\Omega \nabla w\cdot\nabla v\,dx,\\
b(w,v) &:= \frac{1}{2}\left[(b\cdot\nabla w, v) - (w, b\cdot\nabla v)\right], &(2)\\
d(w,v) &:= (cw,v) - \frac{1}{2}\left((\nabla\cdot b)w, v\right),\\
a(w,v) &:= (A\nabla w, \nabla v) + b(w,v) + d(w,v), &(3)
\end{aligned}$$

the variational formulation of the problem (1) in the space $V := H_0^1(\Omega)$ reads as follows:

Find $u \in V$ such that

$$\forall v \in V: \quad a(u,v) = (f,v). \tag{4}$$

Under the above assumptions, the bilinear form a is continuous and coercive on $V \times V$, thus a unique solution $u \in V$ of problem (4) exists.

2 The finite volume scheme

2.1 The case of Voronoi diagrams

Let us consider a family of Voronoi diagrams such that their straight-line duals are Delaunay triangulations of Ω consisting of self-centered simplices. Here a simplex T is called *self-centered* if its circumcentre lies in the interior of T or on the boundary ∂T.

Denote by $\overline{\Lambda}$ the index set of all vertices x_i of a particular triangulation $\mathcal{T}$ and by Λ the index set of all inner vertices.

In more detail, let

$$\begin{aligned}
\Omega_i := \Omega_i^V &:= \{x \in \Omega :\ \|x - x_i\| < \|x - x_i\|\ \forall j \in \overline{\Lambda}\setminus\{i\}\},\ i \in \overline{\Lambda},\\
&\ \text{where } \|\cdot\| \text{ denotes the Euclidean norm in } \mathbb{R}^d,\\
m_i &:= \mathrm{meas}_d\,(\Omega_i)\,,\\
&\ \text{where } \mathrm{meas}_d\,(\cdot) \text{ denotes the } d\text{-dimensional volume},\\
\Gamma_{ij} &:= \partial\Omega_i \cap \partial\Omega_j,\ \Gamma_{ij}^T := \Gamma_{ij} \cap T,\ i \in \Lambda, j \in \overline{\Lambda}\setminus\{i\},\ T \in \mathcal{T},\\
m_{ij} &:= \mathrm{meas}_{d-1}\,(\Gamma_{ij})\,,\ m_{ij}^T := \mathrm{meas}_{d-1}\left(\Gamma_{ij}^T\right),\\
d_{ij} &:= \|x_i - x_j\|,\\
\Lambda_i &:= \{j \in \overline{\Lambda}\setminus\{i\} :\ m_{ij} \neq 0\},\\
\Lambda_T &:= \{i \in \overline{\Lambda} :\ x_i \in \partial T\},\\
h &:= \max_{T\in\mathcal{T}} h_T, \quad \text{where } h_T := \mathrm{diam}\, T.
\end{aligned}$$

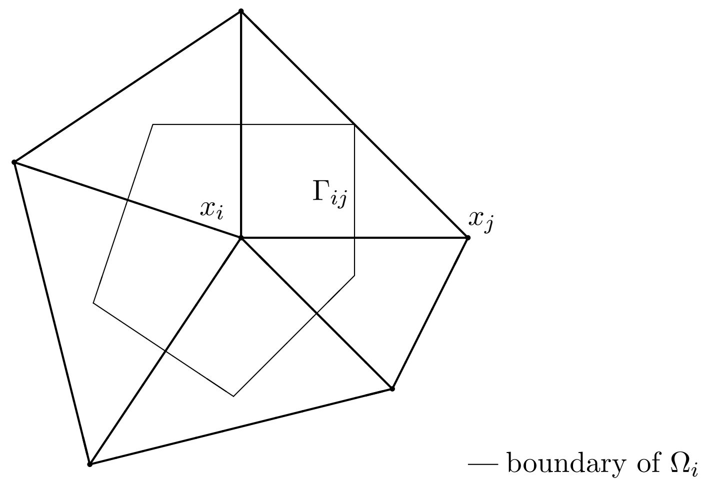

Figure 1: Configuration for the Voronoi-type discretization ($d = 2$)

The finite volume solution will be interpolated in the discrete space

$$V_{\mathcal{T}} := \{v \in V : (\forall T \in \mathcal{T} : v|_T \in \mathcal{P}_1(T))\},$$

where $\mathcal{P}_1(T)$ is the set of all first degree polynomials on T. We introduce a so called *lumping operator*

$$L_{\mathcal{T}} : C(\overline{\Omega}) \to L_\infty(\Omega) \quad \text{acting as} \quad L_{\mathcal{T}} v := \sum_{i \in \overline{\Lambda}} v(x_i)\chi_{\Omega_i},$$

where χ_G denotes the indicator function of a set $G \subset \mathbb{R}^d$.
Due to stability reasons, especially for the case of dominating convection, the class of finite volume methods under consideration is characterized by an additional stabilization technique called *upwinding.* For that purpose we use a weighting function $r : \mathbb{R} \to [0, 1]$, for instance

$$r(z) := 1 - \frac{1}{z}\Big(1 - \frac{z}{e^z - 1}\Big), \tag{5}$$

with the particular values $r_{ij} := r\Big(\frac{\gamma_{ij} d_{ij}}{\mu_{ij}}\Big)$, where

$$A|_{\Gamma_{ij}} \approx \mu_{ij} = \text{const} > 0\,, \quad \nu_{ij} \cdot b|_{\Gamma_{ij}} \approx \gamma_{ij} = \text{const}\,.$$

Precise assumptions w.r.t. the approximations γ_{ij}, μ_{ij} will be given later in Section 3. Furthermore, given a weighting function r, let $K : \mathbb{R} \to [0, \infty)$ be defined by $K(z) := 1 - [1 - r(z)]z$. In case of the weighting function (5), K is the Bernoulli function: $K(z) = z/(e^z - 1)$.

The discrete problem is formulated as follows:

Find $u_{\mathcal{T}} \in V_{\mathcal{T}}$ such that

$$\forall v_{\mathcal{T}} \in V_{\mathcal{T}} : \quad a_{\mathcal{T}}(u_{\mathcal{T}}, v_{\mathcal{T}}) = (f, v_{\mathcal{T}})_{\mathcal{T}}, \tag{6}$$

where

$$\begin{aligned} a_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) &:= \sum_{i\in\Lambda} v_{\mathcal{T}i} \left\{ \sum_{j\in\Lambda_i} \frac{\mu_{ij}}{d_{ij}} K\left(\frac{\gamma_{ij} d_{ij}}{\mu_{ij}}\right)(w_{\mathcal{T}i} - w_{\mathcal{T}j}) m_{ij} + c_i w_{\mathcal{T}i} m_i \right\}, \\ (f, v_{\mathcal{T}})_{\mathcal{T}} &:= \sum_{i\in\Lambda} f_i v_{\mathcal{T}i} m_i \quad \text{and} \quad c_i := c(x_i),\ f_i := f(x_i). \end{aligned}$$

Moreover, we introduce the following norms and seminorms, resp., on $V_{\mathcal{T}}$:

$$\begin{aligned} \|v_{\mathcal{T}}\|_{\mathcal{T}} &:= \sqrt{(v_{\mathcal{T}}, v_{\mathcal{T}})_{\mathcal{T}}} = \|L_{\mathcal{T}} v_{\mathcal{T}}\|_{0,2,\Omega}, \\ |v_{\mathcal{T}}|_V &:= \left\{ \sum_{i\in\Lambda} v_{\mathcal{T}i} \sum_{j\in\Lambda_i} (v_{\mathcal{T}i} - v_{\mathcal{T}j}) \frac{m_{ij}}{d_{ij}} \right\}^{1/2}, \qquad (7) \\ \|v_{\mathcal{T}}\|_V &:= \left\{ |v_{\mathcal{T}}|_V^2 + \|v_{\mathcal{T}}\|_{\mathcal{T}}^2 \right\}^{1/2}. \qquad (8) \end{aligned}$$

For the sake of consistency in the notations, we also use the following abbreviations of wellknown seminorms/norms in the Sobolev space $H^1(\Omega)$:

$$|v_{\mathcal{T}}|_D := |v_{\mathcal{T}}|_{1,2,\Omega}, \quad \|v_{\mathcal{T}}\|_D := \|v_{\mathcal{T}}\|_{1,2,\Omega}. \tag{9}$$

The scheme (6) with the weighting function (5) is often called *exponentially upwinded,* and it can be regarded as a generalization of the Il'in-Allen-Southwell scheme, cf. [Il'69]. It can be defined for other control functions $r : \mathbb{R} \to [0, 1]$, too. However, we have to assume that all of these control functions satisfy the following properties:

(P1) $r(z)$ is monotone for all $z \in \mathbb{R}$,

(P2) $\lim_{z\to-\infty} r(z) = 0, \quad \lim_{z\to\infty} r(z) = 1,$

(P3) $1 + zr(z) \geq 0$ for all $z \in \mathbb{R}$,

(P4) $[1 - r(z) - r(-z)]z = 0$ for all $z \in \mathbb{R}$,

(P5) $\left[r(z) - \frac{1}{2}\right] z \geq 0$ for all $z \in \mathbb{R}$,

(P6) $zr(z)$ is Lipschitz-continuous for all $z \in \mathbb{R}$.

We get from (P4) the relation

$$1 + zr(z) = K(-z). \tag{10}$$

Replacing in (10) the argument z by $-z$, (P3) immediately implies

(P7) $\quad K(z) \geq 0 \quad$ for all $z \in \mathbb{R}$.

EXAMPLE 1 *The function*

$$r(z) = \frac{1}{2}[sign\, z + 1],$$

due to [BT81], has been investigated in [Ris86], [Ris90]. This scheme is called fully upwinded.

EXAMPLE 2 *The choice of the function*

$$r(z) = \frac{1}{2}\left[\frac{z}{2+|z|} + 1\right]$$

goes back to Samarskij [Sam65].

In the sequel, if there is no special reference, we assume that the scheme under consideration is defined for a general function r that possesses the properties (P1) to (P6).

Finally we mention two equivalent representations of the bilinear form $a_{\mathcal{T}}$. First we remember that the leading coefficient $\dfrac{\mu_{ij}}{d_{ij}} K\left(\dfrac{\gamma_{ij} d_{ij}}{\mu_{ij}}\right)$ in $a_{\mathcal{T}}$ can be written, by the definition of K, in the following manner:

$$\frac{\mu_{ij}}{d_{ij}} K\left(\frac{\gamma_{ij} d_{ij}}{\mu_{ij}}\right) = \frac{\mu_{ij}}{d_{ij}}\left\{1 - \frac{\gamma_{ij} d_{ij}}{\mu_{ij}}\left[1 - r\left(\frac{\gamma_{ij} d_{ij}}{\mu_{ij}}\right)\right]\right\} = \frac{\mu_{ij}}{d_{ij}} - (1 - r_{ij})\,\gamma_{ij}\,.$$

Hence we get the representation

$$\begin{aligned} & a_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) \\ = \; & \sum_{i\in\Lambda} v_{\mathcal{T}i}\Big\{ \sum_{j\in\Lambda_i} \{\mu_{ij}(w_{\mathcal{T}i} - w_{\mathcal{T}j})\frac{m_{ij}}{d_{ij}} - (1 - r_{ij})\,(w_{\mathcal{T}i} - w_{\mathcal{T}j})\gamma_{ij} m_{ij}\} \\ & \qquad + c_i w_{\mathcal{T}i} m_i \Big\}. \end{aligned} \tag{11}$$

Furthermore, introducing the notations

$$\begin{aligned} a^0_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) &:= \sum_{i\in\Lambda} v_{\mathcal{T}i} \sum_{j\in\Lambda_i} \mu_{ij}(w_{\mathcal{T}i} - w_{\mathcal{T}j})\frac{m_{ij}}{d_{ij}}, \\ b_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) &:= \sum_{i\in\Lambda} v_{\mathcal{T}i} \sum_{j\in\Lambda_i} \left[(1 - r_{ij}) w_{\mathcal{T}j} - \left(\frac{1}{2} - r_{ij}\right) w_{\mathcal{T}i}\right] \gamma_{ij} m_{ij}, \qquad (12) \\ d_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) &:= \sum_{i\in\Lambda} \left\{ c_i m_i - \frac{1}{2}\sum_{j\in\Lambda_i} \gamma_{ij} m_{ij} \right\} w_{\mathcal{T}i} v_{\mathcal{T}i}, \qquad (13) \end{aligned}$$

we get a splitting of $a_{\mathcal{T}}$ which is comparable with (3):

$$a_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) = a^0_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) + b_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) + d_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}). \tag{14}$$

REMARK 1 *In the special case $\nabla \cdot b \equiv 0$ on Ω, it makes more sense to use the following representations of $b_{\mathcal{T}}$ and $d_{\mathcal{T}}$:*

$$\begin{aligned} b_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) &= \sum_{i\in\Lambda} v_{\mathcal{T}i} \sum_{j\in\Lambda_i} \left[(1-r_{ij})w_{\mathcal{T}j} + r_{ij}w_{\mathcal{T}i}\right] \gamma_{ij} m_{ij}, \\ d_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) &= \sum_{i\in\Lambda} c_i w_{\mathcal{T}i} v_{\mathcal{T}i} m_i. \end{aligned}$$

2.2 The case of Donald diagrams

Let us now consider a family of admissible triangulations $\mathcal{F} = \{\mathcal{T}\}$. Then, for any $T \in \mathcal{T}$ with local vertices $z_j \equiv x_{i_j},\ i_j \in \Lambda_T,\ j \in [1, d+1]_{\mathbb{N}}$, we define

$$\Omega^D_{i_j,T} := \{x \in T : (\forall k \in [1, d+1]_{\mathbb{N}} \setminus \{j\} : \lambda_k(x) < \lambda_j(x))\},$$

where $\lambda_j(x)$ is the j-th barycentric coordinate of x w.r.t. T. Define for $i \in \overline{\Lambda}$ the sets

$$\Omega^D_i := \operatorname{int}\left(\bigcup_{T:\,\partial T \ni x_i} \overline{\Omega^D_{i,T}}\right).$$

In this way, we get a family of Donald diagrams.
Although it is possible to introduce a discretization like (14), we use the following version:

$$a_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) = (A\nabla w_{\mathcal{T}}, \nabla v_{\mathcal{T}}) + b_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) + d_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}), \qquad (15)$$

where the forms $b_{\mathcal{T}}$, $d_{\mathcal{T}}$ are defined analogously to (12),(13). In particular, $\gamma_{ij} \in \mathbb{R}$ is an approximation to $(\nu \cdot b)|_{\Gamma ij}$.

3 Stability and a priori error estimates

In this section we give a short review of some wellknown properties of the schemes (6) and (15). We start with the formulation of conditions with respect to the approximations μ_{ij} and γ_{ij}.

(A2.1) μ_{ij} is an approximation of the term $m_{ij}^{-1}\int_{\Gamma_{ij}} A\,ds$ satisfying the following conditions:

- (i) $0 \le \mu_{ij} \le \|A\|_{1,\infty,\Omega}$,
- (ii) $\mu_{ij} = \mu_{ji}$,
- (iii) $\left|\mu_{ij} - m_{ij}^{-1}\int_{\Gamma_{ij}} A\,ds\right| \le Ch_T|A|_{1,\infty,\Omega}$, where T is one of the simplices having the vertices x_i, x_j, and $C > 0$ is a constant independent of a, h_T, i, j.

(A2.2) γ_{ij} is an approximation of the term $m_{ij}^{-1}\int_{\Gamma_{ij}} \nu{\cdot}b\,ds$ satisfying the following conditions:

- (i) $|\gamma_{ij}| \le \|b\|_{1,\infty,\Omega}$,

(ii) $\gamma_{ij} = -\gamma_{ji}$,

(iii) $\left|\gamma_{ij} - m_{ij}^{-1} \int_{\Gamma_{ij}} (\nu_{ij} \cdot b) ds\right| \le C h_T |b|_{1,\infty,\Omega}$, where T is one of the simplices having the vertices x_i, x_j, and $C > 0$ is a constant independent of b, h_T, i, j.

The subsequent results are extensions of the theory developed in [Ang91], [Ang95b].

THEOREM 1 (Discrete coercivity) *Let a family $\mathcal{F} = \{\mathcal{T}\}$ of triangulations be given, where in the special case of Voronoi diagrams (i.e. $\Xi = V$) all elements T are self-centered and in the special case of Donald diagrams (i.e. $\Xi = D$) the family is shape-regular. Moreover, let the assumptions (A1.1) – (A1.3) and (A2.1), (A2.2) be satisfied. Then, for $h_0 > 0$ sufficiently small there exist two constants $\overline{a}_0 > 0$ and $\overline{a}_1 > 0$ independent of h such that for all $h \in (0, h_0]$ and $v_\mathcal{T} \in V_\mathcal{T}$ the relation*

$$a_\mathcal{T}(v_\mathcal{T}, v_\mathcal{T}) \ge \overline{a}_0 |v_\mathcal{T}|_\Xi^2 + \overline{a}_1 \|v_\mathcal{T}\|_\mathcal{T}^2$$

holds.

The a priori error estimate is based on this stability property and on the following consistency result.

LEMMA 1 (Discrete consistency) *Let a shape-regular family $\mathcal{F}$ of triangulations $\{\mathcal{T}\}$ be given, where in the special case of Voronoi diagrams (i.e. $\Xi = V$) all elements T are self-centered, and let the assumptions (A1.1) and (A2.1), (A2.2) be satisfied. Then, if $h_0 > 0$ is sufficiently small, for any element $w \in W_2^2(\Omega) \cap V$ and any element $v_\mathcal{T} \in V_\mathcal{T}$ the estimate*

$$|a_\mathcal{T}(I_\mathcal{T} w, v_\mathcal{T}) - (-\nabla \cdot (A \nabla w) + b \cdot \nabla w + cw, L_\mathcal{T} v_\mathcal{T})| \le Ch \|w\|_{2,2,\Omega} \left[|v_\mathcal{T}|_\Xi + \|v_\mathcal{T}\|_\mathcal{T}\right]$$

holds for all $h \in (0, h_0]$, where $C > 0$ is a constant which does not depend on h.

The proof of the following theorem is a modification of the standard proof of Strang's first lemma.

THEOREM 2 (A priori error estimate) *Let a shape-regular family $\mathcal{F}$ of triangulations $\{\mathcal{T}\}$ be given, where in the special case of Voronoi diagrams (i.e. $\Xi = V$) all elements T are self-centered, let the assumptions (A1.1) – (A1.3) and (A2.1), (A2.2) be satisfied and suppose that the solution $u \in V$ of problem (1) additionally belongs to $W_2^2(\Omega)$.*
Then, for sufficiently small $h_0 > 0$ the estimate

$$\|u - u_\mathcal{T}\|_\Xi \le Ch \left[\|u\|_{2,2,\Omega} + |f|_{1,q,\Omega}\right]$$

holds for all $h \in (0, h_0]$, where the constant $C > 0$ is independent of h.

4 *A posteriori* error estimates

In the papers [Ang91], [Ang92], an extension to finite volume methods of Babuška&Rheinboldt's approach to *a posteriori* error estimation for finite element methods has been proposed.
In a subsequent paper ([Ang95a]), for a singularly perturbed model problem a modification was introduced with the aim to get two-sided bounds of the error such that the constants occuring in these bounds are independent of the perturbation parameter. In [AKT98] and [Thi99], residual type error estimates for finite volume discretizations of more complicated problems in two and three space dimensions have been presented.
Here we give a more up-to-date treatment based on the so-called *dual-weighted residual error estimators* (see, e.g., [Joh94], [BR96], [BR03], [Ran05]). The aforementioned *a posteriori* error estimates have two disadvantages: On the one hand, certain global constants, which are not known in general, enter into the bounds. The coercivity constant of the bilinear form a is a typical example of such a global constant. On the other hand, certain scaling factors like h_T and $\sqrt{h_E}$ occur simply by using a particular approximation operator.
In the following, we will outline a method that attempts to circumvent these drawbacks. It is especially appropriate for the estimation of errors of functionals depending linearly on the solution.

5 Dual-weighted residual error estimators

Let $J : V \to \mathbb{R}$ denote a linear, continuous functional. We are interested in an estimate of $|J(u) - J(u_{\mathcal{T}})|$, where $u \in V$ is the weak solution of the elliptic boundary value problem (4) and $u_{\mathcal{T}} \in V_{\mathcal{T}}$ is the finite volume approximation from (6).
To do this, the following auxiliary *dual problem* is considered:

Find $z \in V$ such that

$$\forall v \in V : \quad a(v, z) = J(v). \tag{16}$$

The solution $z \in V$ of the dual problem is called *influence function* for the particular choice of J ([AO00]).
Taking $v := u - u_{\mathcal{T}}$ in (16), we get immediately

$$J(u) - J(u_{\mathcal{T}}) = J(u - u_{\mathcal{T}}) = a(u - u_{\mathcal{T}}, z)\,.$$

If $z_{\mathcal{T}} \in V_{\mathcal{T}}$ is an arbitrary element, then

$$J(u) - J(u_{\mathcal{T}}) = a(u - u_{\mathcal{T}}, z - z_{\mathcal{T}}) + a(u - u_{\mathcal{T}}, z_{\mathcal{T}})\,. \tag{17}$$

The first term of the right-hand side is of the same structure as in many well-known finite element methods. It can be regarded as the *conforming residual* of the (primal) finite volume solution weighted be the formal error[1] of the dual solution z. Namely, equation (17) can be rewritten as

$$J(u) - J(u_{\mathcal{T}}) = \langle \varrho(u_{\mathcal{T}}), z - z_{\mathcal{T}} \rangle + a(u - u_{\mathcal{T}}, z_{\mathcal{T}})\,, \tag{18}$$

[1] Note that up to now $z_{\mathcal{T}} \in V_{\mathcal{T}}$ is arbitrary.

where

$$\langle \varrho(u_{\mathcal{T}}), v\rangle := a(u-u_{\mathcal{T}}, v) = (f, v) - a(u_{\mathcal{T}}, v) \quad \forall v \in V$$

by (4). Often this residual is estimated as follows. For arbitrary $v \in V$, it is not difficult to obtain the representation

$$\begin{aligned} a(u-u_{\mathcal{T}}, v) &= (f, v) - a(u_{\mathcal{T}}, v) \\ &= \sum_{T\in\mathcal{T}} \int_T r_T(u_{\mathcal{T}}) v \, dx - \sum_{E\in\mathcal{E}_{\mathcal{T}}} \int_E [\nu_E \cdot (A\nabla u_{\mathcal{T}})]_E \, v \, ds\,, \end{aligned}$$

where

$$r_T(u_{\mathcal{T}}) := (f - (-\nabla\cdot(A\nabla u_{\mathcal{T}}) + b\cdot\nabla u_{\mathcal{T}} + c u_{\mathcal{T}}))\big|_T\,.$$

Here $\mathcal{E}_{\mathcal{T}}$ denotes the set of all interior faces of all elements $T \in \mathcal{T}$, ν_E is a fixed unit normal assigned to any of those faces and $[\cdot]_E$ denotes the jump across the face E in the normal direction ν_E.
Setting $v = z - z_{\mathcal{T}}$ and applying the Cauchy-Schwarz inequality, we immediately obtain an estimate of the first term of the right-hand side of (17):

$$\begin{aligned} |a(u-u_{\mathcal{T}}, z-z_{\mathcal{T}})| \;\le\; & \sum_{T\in\mathcal{T}} \|r_T(u_{\mathcal{T}})\|_{0,2,T} \|z-z_{\mathcal{T}}\|_{0,2,T} \\ & + \sum_{E\in\mathcal{E}_{\mathcal{T}}} \big\|[\nu_E\cdot(A\nabla u_{\mathcal{T}})]_E\big\|_{0,2,E} \|z-z_{\mathcal{T}}\|_{0,2,E}\,. \end{aligned} \tag{19}$$

Concerning the second term of the right-hand side of (17), so we use the following argument:

$$\begin{aligned} a(u-u_{\mathcal{T}}, z_{\mathcal{T}}) &= a(u, z_{\mathcal{T}}) - a(u_{\mathcal{T}}, z_{\mathcal{T}}) \\ &= (f, z_{\mathcal{T}}) - a_{\mathcal{T}}(u_{\mathcal{T}}, z_{\mathcal{T}}) + a_{\mathcal{T}}(u_{\mathcal{T}}, z_{\mathcal{T}}) - a(u_{\mathcal{T}}, z_{\mathcal{T}}) \\ &= (f, z_{\mathcal{T}}) - (f, z_{\mathcal{T}})_{\mathcal{T}} + a_{\mathcal{T}}(u_{\mathcal{T}}, z_{\mathcal{T}}) - a(u_{\mathcal{T}}, z_{\mathcal{T}}). \end{aligned} \tag{20}$$

It is rather obvious that, given $z_{\mathcal{T}} \in V_{\mathcal{T}}$, the last two differences can be locally calculated with sufficient accuracy.
Namely, we have

$$\begin{aligned} (f, z_{\mathcal{T}}) - (f, z_{\mathcal{T}})_{\mathcal{T}} &= \sum_{T\in\mathcal{T}} \{(f, z_{\mathcal{T}})_T - (f, z_{\mathcal{T}})_{l,T}\} \\ &:= \sum_{T\in\mathcal{T}} \Big\{(f, z_{\mathcal{T}})_T - \sum_{i\in\Lambda_T} f_i z_{\mathcal{T}i} m_i^T\Big\}, \end{aligned} \tag{21}$$

where $m_i^T := \operatorname{meas}_d(\Omega_i \cap T)$. Analogously, with

$$\begin{aligned} & a_{\mathcal{T},T}(u_{\mathcal{T}}, z_{\mathcal{T}}) \\ := & \sum_{i\in\Lambda} z_{\mathcal{T}i} \left\{ \sum_{j\in\Lambda_T\setminus\{i\}} \left\{ \mu_{ij} \frac{u_{\mathcal{T}i} - u_{\mathcal{T}j}}{d_{ij}} - \gamma_{ij}\,(1-r_{ij})\,(u_{\mathcal{T}i} - u_{\mathcal{T}j}) \right\} m_{ij}^T + c_i u_{\mathcal{T}i} m_i^T \right\}, \end{aligned}$$

we have

$$a_{\mathcal{T}}(u_{\mathcal{T}}, z_{\mathcal{T}}) - a(u_{\mathcal{T}}, z_{\mathcal{T}}) = \sum_{T\in\mathcal{T}} \{a_{\mathcal{T},T}(u_{\mathcal{T}}, z_{\mathcal{T}}) - a_T(u_{\mathcal{T}}, z_{\mathcal{T}})\}\,. \tag{22}$$

Then from (20) we conclude the estimate

$$|a(u-u_{\mathcal{T}}, z_{\mathcal{T}})| \leq \sum_{T\in\mathcal{T}} \Big|(f, z_{\mathcal{T}})_T - \sum_{i\in\Lambda_T} f_i z_{\mathcal{T}i} m_i^T\Big| + \sum_{T\in\mathcal{T}} \Big|a_{\mathcal{T},T}(u_{\mathcal{T}}, z_{\mathcal{T}}) - a_T(u_{\mathcal{T}}, z_{\mathcal{T}})\Big|.$$

Putting this relation together with (19), we arrive at

$$\begin{aligned} |a(u-u_{\mathcal{T}}, z-z_{\mathcal{T}})| \;\leq\; & \sum_{T\in\mathcal{T}} \|r_T(u_{\mathcal{T}})\|_{0,2,T} \|z-z_{\mathcal{T}}\|_{0,2,T} \\ & + \sum_{E\in\mathcal{E}_{\mathcal{T}}} \big\|[\nu_E\cdot(A\nabla u_{\mathcal{T}})]_E\big\|_{0,2,E} \|z-z_{\mathcal{T}}\|_{0,2,E} \\ & + \sum_{T\in\mathcal{T}} \Big|(f, z_{\mathcal{T}})_T - \sum_{i\in\Lambda_T} f_i z_{\mathcal{T}i} m_i^T\Big| \\ & + \sum_{T\in\mathcal{T}} \Big|a_{\mathcal{T},T}(u_{\mathcal{T}}, z_{\mathcal{T}}) - a_T(u_{\mathcal{T}}, z_{\mathcal{T}})\Big|. \end{aligned}$$

This is the starting point for the practical computation.
In contrast to traditional approaches, here the norms of $z - z_{\mathcal{T}}$ will not be theoretically analyzed but numerically approximated. This can be done by an approximation of the influence function z. There are several (more or less heuristic) ways to do this. A practically successful approach consists in the so-called *higher-order recovery*, where z is replaced by an element that is recovered from the finite element solution $z_{\mathcal{T}} \in V_{\mathcal{T}}$ of the auxiliary problem. The recovered element approximates z with higher order than $z_{\mathcal{T}}$ does (see, e.g., [BR03, Sect. 4.1] or [Ran05, Sect. 3.2]).
A different view on the left-hand side of (18) is obtained if the term $(f, z - z_{\mathcal{T}})$ is rewritten as follows:

$$\begin{aligned} (f, z-z_{\mathcal{T}}) &= a(u, z-z_{\mathcal{T}}) \\ &= a(u-u_{\mathcal{T}}, z-z_{\mathcal{T}}) + a(u_{\mathcal{T}}, z-z_{\mathcal{T}}) \\ &= a(u-u_{\mathcal{T}}, z) - a(u-u_{\mathcal{T}}, z_{\mathcal{T}}) + a(u_{\mathcal{T}}, z-z_{\mathcal{T}}) \\ &= J(u) - J(u_{\mathcal{T}}) - a(u-u_{\mathcal{T}}, z_{\mathcal{T}}) + a(u_{\mathcal{T}}, z-z_{\mathcal{T}}), \end{aligned}$$

where we have used (4) and (16). The first three terms on the right-hand side can be interpreted as the conforming residual of the approximation $z_{\mathcal{T}}$ of the dual solution z weighted by the error of the finite volume solution. That is, with

$$\langle \varrho^*(z_{\mathcal{T}}), v\rangle := J(v) - a(v, z_{\mathcal{T}}) \quad \forall v \in V \tag{23}$$

we have

$$(f, z-z_{\mathcal{T}}) = \langle \varrho^*(z_{\mathcal{T}}), u-u_{\mathcal{T}}\rangle + a(u_{\mathcal{T}}, z-z_{\mathcal{T}}).$$

From (23) we get that

$$J(u) - J(u_{\mathcal{T}}) = \langle \varrho^*(z_{\mathcal{T}}), u-u_{\mathcal{T}}\rangle + a(u-u_{\mathcal{T}}, z_{\mathcal{T}}).$$

This relation together with (18) leads to

$$J(u) - J(u_{\mathcal{T}}) = \beta\langle \varrho(u_{\mathcal{T}}), z-z_{\mathcal{T}}\rangle + (1-\beta)\langle \varrho^*(z_{\mathcal{T}}), u-u_{\mathcal{T}}\rangle + a(u-u_{\mathcal{T}}, z_{\mathcal{T}}) \tag{24}$$

for any parameter $\beta \in [0,1]$. In this way we get a representation of the error under consideration as a convex combination of the primal and dual residuals plus the orthogonality defect.
The values of the unknown primal solution u in the expression $\langle \varrho^*(z_\mathcal{T}), u - u_\mathcal{T}\rangle$ can be approximated in the same way as the values of the influence function z, for instance by means of recovery techniques as indicated above.
At the end of this section we want to mention how the method could be used to estimate certain norms of the error. In the case where the norms are induced by particular inner products, there is a simple, formal way. For example, for the L_2-norm we have

$$\|u - u_\mathcal{T}\|_{0,2,\Omega} = \frac{(u - u_\mathcal{T}, u - u_\mathcal{T})}{\|u - u_\mathcal{T}\|_{0,2,\Omega}} .$$

Keeping both u and $u_\mathcal{T}$ fixed, we get with the definition

$$J(v) := \frac{(v, u - u_\mathcal{T})}{\|u - u_\mathcal{T}\|_{0,2,\Omega}} \tag{25}$$

a linear, continuous functional $J : H^1(\Omega) \to \mathbb{R}$ such that $J(u) - J(u_\mathcal{T}) = \|u - u_\mathcal{T}\|_{0,2,\Omega}$.
The practical difficulty of this approach consists in the fact that in order to be able to find the solution z of the auxiliary problem we have to know the values of J, but they depend on the unknown element $u - u_\mathcal{T}$. However, if there is a higher recovery of u at hand, then it could be used to approximate u in the definition (25) of J.

6 Residual *a posteriori* estimates of the error in the energy norm

The "traditional" residual error estimates start from the relation

$$\alpha \|u - u_\mathcal{T}\|^2_{1,2,\Omega} \le a(u - u_\mathcal{T}, u - u_\mathcal{T}) ,$$

where $\alpha > 0$ is the coercivity constant of a. Without loss of generality we may suppose $u - u_\mathcal{T} \in V \setminus \{0\}$, hence

$$\|u - u_\mathcal{T}\|_{1,2,\Omega} \le \frac{1}{\alpha} \frac{a(u - u_\mathcal{T}, u - u_\mathcal{T})}{\|u - u_\mathcal{T}\|_{1,2,\Omega}} \le \frac{1}{\alpha} \sup_{v \in V} \frac{a(u - u_\mathcal{T}, v)}{\|v\|_{1,2,\Omega}} . \tag{26}$$

Now, the term $a(u - u_\mathcal{T}, v)$ is treated as the term $a(u - u_\mathcal{T}, z)$ in the previous section, i.e. we have, for an arbitrary element $v_\mathcal{T} \in V_\mathcal{T}$,

$$a(u - u_\mathcal{T}, v) = a(u - u_\mathcal{T}, v - v_\mathcal{T}) + a(u - u_\mathcal{T}, v_\mathcal{T}) \tag{27}$$

and so we get

$$\begin{aligned} |a(u - u_\mathcal{T}, v - v_\mathcal{T})| \;\le\; & \sum_{T \in \mathcal{T}} \|r_T(u_\mathcal{T})\|_{0,2,T} \|v - v_\mathcal{T}\|_{0,2,T} \\ & + \sum_{E \in \mathcal{E}_\mathcal{T}} \big\|[\nu_E \cdot (A\nabla u_\mathcal{T})]_E\big\|_{0,2,E} \|v - v_\mathcal{T}\|_{0,2,E} , \end{aligned} \tag{28}$$

$$|a(u-u_{\mathcal{T}},v_{\mathcal{T}})| \leq \sum_{T\in\mathcal{T}} \Big|(f,v_{\mathcal{T}})_T - \sum_{i\in\Lambda_T} f_i v_{\mathcal{T}i} m_i^T\Big| + \sum_{T\in\mathcal{T}} \Big|a_{\mathcal{T},T}(u_{\mathcal{T}},v_{\mathcal{T}}) - a_T(u_{\mathcal{T}},v_{\mathcal{T}})\Big|. \tag{29}$$

To get the first bound (28) as small as possible, the function $v_{\mathcal{T}} \in V_{\mathcal{T}}$ is chosen such that the element $v \in V$ is approximated adequately in both spaces $L^2(T)$ and $L^2(E)$. Typically, certain quasi-interpolation procedures such as Clément's ([Clé75]) or Scott-Zhang's ([SZ90]) quasi-interpolation operators are applied. This leads to an estimate of the form

$$\begin{aligned} & a(u-u_{\mathcal{T}}, v-v_{\mathcal{T}}) \\ \leq\ & C\Big\{\sum_{T\in\mathcal{T}} h_T^2\|r_T(u_{\mathcal{T}})\|_{0,2,T}^2 + \sum_{E\in\mathcal{E}_{\mathcal{T}}} h_E\big\|[\nu_E\cdot(A\nabla u_{\mathcal{T}})]_E\big\|_{0,2,E}^2\Big\}^{1/2} |v-v_{\mathcal{T}}|_{1,2,\Omega}\,. \end{aligned}$$

It remains to give a decomposition of the second bound (29) with this particular choice of $v_{\mathcal{T}}$ such that we obtain the structure

$$\text{error estimator} \times \|v_{\mathcal{T}}\|_{1,2,\Omega}.$$

Then the boundedness of the quasi-interpolation operator implies

$$\text{error estimator} \times C\|v\|_{1,2,\Omega},$$

and this can be used in (26) to complete the estimation.
This decomposition will be given in the following section.

7 Analysis of the orthogonality defect

Here we investigate the structure of the defect terms $a(u-u_{\mathcal{T}}, z_{\mathcal{T}})$ and $a(u-u_{\mathcal{T}}, v_{\mathcal{T}})$ in (24) and (27), respectively.
Using (21), (22), we get from (20) the following decomposition of the orthogonality defect:

$$\begin{aligned} & a(u-u_{\mathcal{T}}, z_{\mathcal{T}}) \\ =\ & \sum_{i\in\Lambda} \int_{\Omega_i} [f z_{\mathcal{T}} - f_i z_{\mathcal{T}i}]dx \\ & + \sum_{i\in\Lambda} z_{\mathcal{T}i} \sum_{j\in\Lambda_i} \mu_{ij}(u_{\mathcal{T}i} - u_{\mathcal{T}j})\frac{m_{ij}}{d_{ij}} - (A\nabla u_{\mathcal{T}}, \nabla z_{\mathcal{T}}) \\ & + \sum_{i\in\Lambda}\sum_{j\in\Lambda_i} (1-r_{ij})\gamma_{ij}(u_{\mathcal{T}j} - u_{\mathcal{T}i}) z_{\mathcal{T}i} m_{ij} - (b\cdot\nabla u_{\mathcal{T}}, z_{\mathcal{T}}) \\ & + \sum_{i\in\Lambda} \int_{\Omega_i} [c_i u_{\mathcal{T}i} z_{\mathcal{T}i} - c u_{\mathcal{T}} z_{\mathcal{T}}]dx \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in\Lambda}\left\{\int_{\Omega_i} f(z_{\mathcal{T}} - z_{\mathcal{T}i})dx + \int_{\Omega_i}(f-f_i)z_{\mathcal{T}i}dx\right\}\\
&\quad + \sum_{i\in\Lambda} z_{\mathcal{T}i}\sum_{j\in\Lambda_i}\mu_{ij}(u_{\mathcal{T}i}-u_{\mathcal{T}j})\frac{m_{ij}}{d_{ij}} - (A\nabla u_{\mathcal{T}}, \nabla z_{\mathcal{T}})\\
&\quad + \sum_{i\in\Lambda}\Big\{\sum_{j\in\Lambda_i}(1-r_{ij})\gamma_{ij}(u_{\mathcal{T}j}-u_{\mathcal{T}i})z_{\mathcal{T}i}m_{ij} \; - \; \int_{\Omega_i}(b\cdot\nabla u_{\mathcal{T}})z_{\mathcal{T}i}dx\Big\}\\
&\quad - \sum_{i\in\Lambda}\int_{\Omega_i}(b\cdot\nabla u_{\mathcal{T}})(z_{\mathcal{T}} - z_{\mathcal{T}i})dx\\
&\quad + \sum_{i\in\Lambda}\left\{\int_{\Omega_i}[c_i u_{\mathcal{T}i} - cu_{\mathcal{T}}]z_{\mathcal{T}i}dx \; - \; \int_{\Omega_i} cu_{\mathcal{T}}(z_{\mathcal{T}} - z_{\mathcal{T}i})dx\right\}\\
&= \delta_0 + \delta_1 + \delta_2 + \delta_3
\end{aligned}
$$

with

$$
\begin{aligned}
\delta_0 &:= \sum_{i\in\Lambda} z_{\mathcal{T}i}\sum_{j\in\Lambda_i}\mu_{ij}(u_{\mathcal{T}i}-u_{\mathcal{T}j})\frac{m_{ij}}{d_{ij}} - (A\nabla u_{\mathcal{T}}, \nabla z_{\mathcal{T}}),\\
\delta_1 &:= \sum_{i\in\Lambda}\int_{\Omega_i}[f - b\cdot\nabla u_{\mathcal{T}} - cu_{\mathcal{T}}](z_{\mathcal{T}} - z_{\mathcal{T}i})dx,\\
\delta_2 &:= \sum_{i\in\Lambda} z_{\mathcal{T}i}\Big\{\int_{\Omega_i}[f - f_i + (\nabla\cdot b - c)u_{\mathcal{T}} + c_i u_{\mathcal{T}i}]dx \; - \; \sum_{j\in\Lambda_i} u_{\mathcal{T}i}\gamma_{ij}m_{ij}\Big\},\\
\delta_3 &:= \sum_{i\in\Lambda}\sum_{j\in\Lambda_i}\int_{\Gamma_{ij}}[(r_{ij}u_{\mathcal{T}i} + (1-r_{ij})u_{\mathcal{T}j})\gamma_{ij} - (\nu_{ij}\cdot b)u_{\mathcal{T}}]z_{\mathcal{T}i}ds.
\end{aligned}
$$

Here we have used that $b\cdot\nabla u_{\mathcal{T}} = \nabla\cdot(bu_{\mathcal{T}}) - (\nabla\cdot b)u_{\mathcal{T}}$.
In the case of Donald diagrams, $\delta_0 = 0$.
In order to treat δ_0 in the case of Voronoi diagrams, we introduce a piecewise constant (w.r.t. $\mathcal{T}$) approximation $A_{\mathcal{T}}$ to A by $A_{\mathcal{T}}|_T := \dfrac{1}{\mathrm{meas}_d(T)}\displaystyle\int_T A\,dx$, $T\in\mathcal{T}$. Then we can write

$$
\begin{aligned}
\delta_0 &= \sum_{i\in\Lambda} z_{\mathcal{T}i}\sum_{j\in\Lambda_i}\left(\mu_{ij} - \frac{1}{m_{ij}}\int_{\Gamma_{ij}} A_{\mathcal{T}}ds\right)(u_{\mathcal{T}i}-u_{\mathcal{T}j})\frac{m_{ij}}{d_{ij}}\\
&\quad + \sum_{i\in\Lambda} z_{\mathcal{T}i}\sum_{j\in\Lambda_i}\left(\int_{\Gamma_{ij}} A_{\mathcal{T}}ds\right)\frac{u_{\mathcal{T}i}-u_{\mathcal{T}j}}{d_{ij}} - (A\nabla u_{\mathcal{T}}, \nabla z_{\mathcal{T}}).
\end{aligned}
$$

It is wellknown that, for arbitrary $u_{\mathcal{T}}, z_{\mathcal{T}} \in V_{\mathcal{T}}$,

$$
\sum_{i\in\Lambda} z_{\mathcal{T}i}\sum_{j\in\Lambda_i}\left(\int_{\Gamma_{ij}} A_{\mathcal{T}}ds\right)\frac{u_{\mathcal{T}i}-u_{\mathcal{T}j}}{d_{ij}} = (A_{\mathcal{T}}\nabla u_{\mathcal{T}}, \nabla z_{\mathcal{T}}).
$$

Hence

$$
\delta_0 = \sum_{i\in\Lambda} z_{\mathcal{T}i}\sum_{j\in\Lambda_i}\left(\mu_{ij} - \frac{1}{m_{ij}}\int_{\Gamma_{ij}} A_{\mathcal{T}}ds\right)(u_{\mathcal{T}i}-u_{\mathcal{T}j})\frac{m_{ij}}{d_{ij}} + ((A_{\mathcal{T}}-A)\nabla u_{\mathcal{T}}, \nabla z_{\mathcal{T}}).
$$

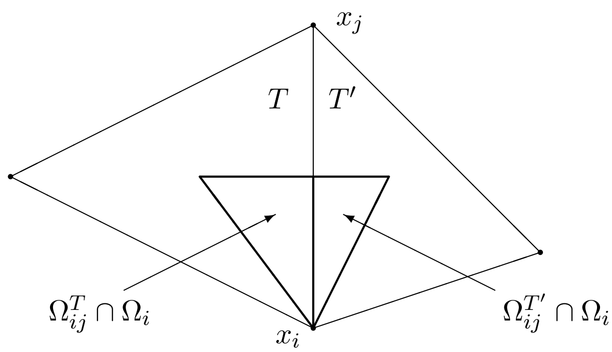

Figure 2: The auxiliary simplices in the case $d = 2$

Since both $\nabla u_{\mathcal{T}}, \nabla z_{\mathcal{T}}$ are constant on every element $T \in \mathcal{T}$, the second term vanishes. By a symmetry argument, we arrive at

$$\delta_0 = \frac{1}{2} \sum_{i \in \Lambda} \sum_{j \in \Lambda_i} \left(\mu_{ij} - \frac{1}{m_{ij}} \int_{\Gamma_{ij}} A_{\mathcal{T}} ds \right) (u_{\mathcal{T}i} - u_{\mathcal{T}j})(z_{\mathcal{T}i} - z_{\mathcal{T}j}) \frac{m_{ij}}{d_{ij}}.$$

Now the Cauchy-Schwarz inequality implies

$$\begin{aligned} |\delta_0| \quad \leq \quad & \frac{1}{2} \left\{ \sum_{i \in \Lambda} \sum_{j \in \Lambda_i} \left(\mu_{ij} - \frac{1}{m_{ij}} \int_{\Gamma_{ij}} A_{\mathcal{T}} ds \right)^2 (u_{\mathcal{T}i} - u_{\mathcal{T}j})^2 \frac{m_{ij}}{d_{ij}} \right\}^{1/2} \\ & \times \left\{ \sum_{i \in \Lambda} \sum_{j \in \Lambda_i} (z_{\mathcal{T}i} - z_{\mathcal{T}j})^2 \frac{m_{ij}}{d_{ij}} \right\}^{1/2} . \end{aligned}$$

The last factor can be bounded by $C_1 |z_{\mathcal{T}}|_{1,2,\Omega}$, therefore we get

$$|\delta_0| \leq C_1 \eta_0 |z_{\mathcal{T}}|_{1,2,\Omega}, \tag{30}$$

where

$$\eta_0^2 := \sum_{i \in \Lambda} \eta_{0i}^2 \quad \text{with} \quad \eta_{0i}^2 := \frac{1}{4} \sum_{j \in \Lambda_i} \left(\mu_{ij} - \frac{1}{m_{ij}} \int_{\Gamma_{ij}} A_{\mathcal{T}} ds \right)^2 (u_{\mathcal{T}i} - u_{\mathcal{T}j})^2 \frac{m_{ij}}{d_{ij}} \, .$$

Setting $g := f - b \cdot \nabla u_{\mathcal{T}} - c u_{\mathcal{T}}$ and $\delta_{1i} := \int_{\Omega_i} g(z_{\mathcal{T}} - z_{\mathcal{T}i}) dx$, we can write (cf. Figure 2 for the case $d = 2$):

$$\delta_{1i} = \sum_{j \in \Lambda_i} \sum_{T \in \mathcal{T} \, : \, m_{ij}^T > 0} \int_{\Omega_{ij}^T \cap \Omega_i} g(z_{\mathcal{T}} - z_{\mathcal{T}i}) dx.$$

On each simplex T, it holds

$$z_{\mathcal{T}} = z_{\mathcal{T}i} + \nabla z_{\mathcal{T}} \cdot (x - x_i),$$

where $\nabla z_{\mathcal{T}}$ is constant on Ω_{ij}^T.

It follows

$$
\begin{aligned}
\delta_{1i} &= \sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \int_{\Omega_{ij}^T\cap\Omega_i} g\nabla z_{\mathcal{T}}\cdot(x-x_i)dx \\
&\le \sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \int_{\Omega_{ij}^T\cap\Omega_i} |g|\|\nabla z_{\mathcal{T}}\|\|x-x_i\|dx \\
&\le \left\{\sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \int_{\Omega_{ij}^T\cap\Omega_i} |g|^2\|x-x_i\|^2dx\right\}^{1/2} \\
&\quad \times\left\{\sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \int_{\Omega_{ij}^T\cap\Omega_i} \|\nabla z_{\mathcal{T}}\|^2dx\right\}^{1/2} \\
&\le \left\{\sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \|x_T^V-x_i\|^2\int_{\Omega_{ij}^T\cap\Omega_i} |g|^2dx\right\}^{1/2} |z_{\mathcal{T}}|_{1,2,\Omega_i} \\
&\le \eta_{1i}|z_{\mathcal{T}}|_{1,2,\Omega_i},
\end{aligned}
$$

where

$$\eta_{1i}^2 := \sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \|x_T^V-x_i\|^2\int_{\Omega_{ij}^T\cap\Omega_i} g^2dx.$$

Thus we arrive at

$$\delta_1 \le \eta_1|z_{\mathcal{T}}|_{1,2,\Omega}. \tag{31}$$

For the third term δ_2, with

$$\theta_i := \int_{\Omega_i}[f-f_i+(\nabla\cdot b-c)u_{\mathcal{T}}+c_iu_{\mathcal{T}i}]dx\ -\sum_{j\in\Lambda_i}u_{\mathcal{T}i}\gamma_{ij}m_{ij},$$

we have

$$\delta_2 = \sum_{i\in\Lambda} z_{\mathcal{T}i}\theta_i.$$

Because of

$$z_{\mathcal{T}i}\theta_i \le \eta_{2i}|z_{\mathcal{T}i}|\sqrt{m_i},$$

where $\eta_{2i} := |\theta_i|/\sqrt{m_i}$, it follows

$$\delta_2 \le \eta_2\|z_{\mathcal{T}}\|_{\mathcal{T}}.$$

In view of the equivalence of the L_2-norm and the lumped L_2-norm on $V_{\mathcal{T}}$, we obtain

$$\delta_2 \le C_2\eta_2\|z_{\mathcal{T}}\|_{0,2,\Omega}. \tag{32}$$

For the remaining term δ_3 we have (by the symmetry argument)

$$\delta_3 = \sum_{i\in\Lambda}\delta_{3i},$$

where

$$\delta_{3i} := \frac{1}{2} \sum_{j\in\Lambda_i} \int_{\Gamma_{ij}} \zeta_{\mathcal{T}ij}(z_{\mathcal{T}i} - z_{\mathcal{T}j})ds$$

with

$$\zeta_{\mathcal{T}ij} := [r_{ij}u_{\mathcal{T}i} + (1-r_{ij})u_{\mathcal{T}j}]\gamma_{ij} \;-\; (\nu_{ij}\cdot b)u_{\mathcal{T}}.$$

In view of $z_{\mathcal{T}i} - z_{\mathcal{T}j} = d_{ij}(\nu_{ij}\cdot\nabla z_{\mathcal{T}})$ on Ω_{ij}^T we get

$$\delta_{3i} = \frac{1}{2} \sum_{j\in\Lambda_i} d_{ij} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \int_{\Gamma_{ij}^T} \zeta_{\mathcal{T}ij}(\nu_{ij}\cdot\nabla z_{\mathcal{T}})ds.$$

It follows (remember that $\nu_{ij}\cdot\nabla z_{\mathcal{T}}$ is constant on Γ_{ij}^T and $\nabla z_{\mathcal{T}}$ is constant on $\Omega_{ij}^T \cap \Omega_i$)

$$\begin{aligned}
\delta_{3i} &\le \frac{1}{2} \sum_{j\in\Lambda_i} d_{ij} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \left| \int_{\Gamma_{ij}^T} \zeta_{\mathcal{T}ij} ds \right| \|\nabla z_{\mathcal{T}}\| \\
&= \frac{1}{2} \sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \frac{d_{ij}}{\sqrt{\operatorname{meas}_d\left(\Omega_{ij}^T\cap\Omega_i\right)}} \left| \int_{\Gamma_{ij}^T} \zeta_{\mathcal{T}ij} ds \right| \|\nabla z_{\mathcal{T}}\| \sqrt{\operatorname{meas}_d\left(\Omega_{ij}^T\cap\Omega_i\right)}.
\end{aligned}$$

By Cauchy's inequality, we have

$$\begin{aligned}
\delta_{3i} &\le \frac{1}{2} \left\{ \sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \frac{d_{ij}^2}{\operatorname{meas}_d\left(\Omega_{ij}^T\cap\Omega_i\right)} \left(\int_{\Gamma_{ij}^T} \zeta_{\mathcal{T}ij} ds \right)^2 \right\}^{1/2} |z_{\mathcal{T}}|_{1,2,\Omega_i} \\
&\le \eta_{3i} |z_{\mathcal{T}}|_{1,2,\Omega_i},
\end{aligned}$$

where

$$\begin{aligned}
\eta_{3i}^2 &:= \frac{1}{4} \sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \frac{d_{ij}^2}{\operatorname{meas}_d\left(\Omega_{ij}^T\cap\Omega_i\right)} \left(\int_{\Gamma_{ij}^T} \zeta_{\mathcal{T}ij} ds \right)^2 \\
&= \frac{d}{4} \sum_{j\in\Lambda_i} \sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \frac{d_{ij}}{m_{ij}^T} \left(\int_{\Gamma_{ij}^T} \zeta_{\mathcal{T}ij} ds \right)^2 .
\end{aligned}$$

Thus it holds

$$\delta_3 \le \eta_3 |z_{\mathcal{T}}|_{1,2,\Omega}. \tag{33}$$

Summarizing the estimates (30) – (33), we obtain

$$a(u-u_{\mathcal{T}}, z_{\mathcal{T}}) \le (C_1\eta_0 + \eta_1 + \eta_3)|z_{\mathcal{T}}|_{1,2,\Omega} + C_2\eta_2\|z_{\mathcal{T}}\|_{0,2,\Omega}.$$

The indicators have the following structure:

$$\eta_l = \left\{ \sum_{i\in\Lambda} \eta_{li}^2 \right\}^{1/2}, \qquad l \in \{0,1,2,3\},$$

where

$$\eta_{0i} = \frac{1}{2}\left\{\sum_{j\in\Lambda_i}\left(\mu_{ij} - \frac{1}{m_{ij}}\int_{\Gamma_{ij}} A_{\mathcal{T}}ds\right)^2 (u_{\mathcal{T}i} - u_{\mathcal{T}j})^2 \frac{m_{ij}}{d_{ij}}\right\}^{1/2}$$

in case of Voronoi diagrams and $\eta_{0i} = 0$ in case of Donald diagrams,

$$\eta_{1i} = \left\{\sum_{j\in\Lambda_i}\sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0} \|x_T^V - x_i\|^2 \int_{\Omega_{ij}^T\cap\Omega_i} [f - b\cdot\nabla u_{\mathcal{T}} - cu_{\mathcal{T}}]^2 dx\right\}^{1/2},$$

$$\eta_{2i} = \frac{1}{\sqrt{m_i}}\left|\int_{\Omega_i}[f - f_i + (\nabla\cdot b - c)u_{\mathcal{T}} + c_i u_{\mathcal{T}i}]dx - \sum_{j\in\Lambda_i} u_{\mathcal{T}i}\gamma_{ij}m_{ij}\right|,$$

$$\eta_{3i} = \left\{\frac{d}{2}\sum_{j\in\Lambda_i}\sum_{T\in\mathcal{T}\,:\,m_{ij}^T>0}\frac{d_{ij}}{m_{ij}^T}\left(\int_{\Gamma_{ij}^T}[(r_{ij}u_{\mathcal{T}i} + (1-r_{ij})u_{\mathcal{T}j})\gamma_{ij} - (\nu_{ij}\cdot b)u_{\mathcal{T}}]ds\right)^2\right\}^{1/2}.$$

REMARK 2 *(i) We mention that all the indicators η_l can be rewritten in such a way that the resulting local indicators are related to the elements $T\in\mathcal{T}$.*
(ii) It can be shown that the indicators η_l are order-consistent *with the* a priori *error estimate (Theorem 2) in the following sense :*
If $f\in W_q^1(\Omega)$ with some $q>d$ and $u\in W_2^2(\Omega)$, then there is a constant $C_c>0$ such that

$$\sum_{l=0}^{3}\eta_l \le C_c h\left[\|u\|_{2,2} + \|f\|_{1,r}\right],$$

see [Ang92, Thm. 4] for a special case.

References

[AKT98] L. Angermann, P. Knabner, and K. Thiele. An error estimator for a finite volume discretization of density driven flow in porous media. *Appl. Numer. Math.*, 26(1-2):179–191, 1998.

[Ang91] L. Angermann. A modified error estimator of Babuška-Rheinboldt's type for singularly perturbed elliptic problems. In H.-G. Roos, A. Felgenhauer, and L. Angermann, editors, *Numerical methods in singularly perturbed problems*, pages 1–12. TU Dresden, 1991.

[Ang92] L. Angermann. An a-posteriori estimation for the solution of elliptic boundary value problems by means of upwind FEM. *IMA J. Numer. Anal.*, 12:201–215, 1992.

[Ang95a] L. Angermann. Balanced a-posteriori error estimates for finite volume type discretizations of convection-dominated elliptic problems. *Computing*, 55(4):305–323, 1995.

[Ang95b] L. Angermann. Error estimates for the finite-element solution of an elliptic singularly perturbed problem. *IMA J. Numer. Anal.*, 15:161–196, 1995.

[Ang06] L. Angermann. Transport-stabilized semidiscretizations of the incompressible Navier-Stokes equations. *Comput. Methods Appl. Math.*, 6(3):239–263, 2006.

[AO00] M. Ainsworth and J.T. Oden. *A posteriori error estimation in finite element analysis*. Wiley, New York, 2000.

[BR96] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East-West J. Numer. Math.*, 4(4):237–264, 1996.

[BR03] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[BT81] K. Baba and M. Tabata. On a conservative upwind finite element scheme for convective diffusion equations. *RAIRO Anal. Numér.*, 15(1):3–25, 1981.

[Clé75] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.*, 9(R-2):77–84, 1975.

[Il'69] A.M. Il'in. Differencing scheme for a differential equation with a small parameter affecting the highest derivative. *Mat. Zametki*, 6:237–248, 1969. English transl. in Math. Notes Acad. Sci. USSR 6:596-602, 1969.

[Joh94] C. Johnson. A new paradigm for adaptive finite element methods. In *The mathematics of finite elements and applications (Uxbridge, 1993)*, pages 105–120. Wiley, Chichester, 1994.

[Ran05] R. Rannacher. Adaptive finite element methods in flow computations. In Z.-C. Shi, Z. Chen, T. Tang, and D. Yu, editors, *Recent Advances in Adaptive Computation. Contemporary Mathematics, vol. 383*, pages 183–176. AMS, Providence, Rhode Island, 2005.

[Ris86] U. Risch. Ein hybrides upwind-FEM-Verfahren und dessen Anwendung auf schwach gekoppelte elliptische Differentialgleichungssysteme mit dominanter Konvektion. Dissertation, TH Magdeburg, 1986.

[Ris90] U. Risch. An upwind finite element method for singularly perturbed elliptic problems and local estimates in the L_∞-norm. M^2AN, 24(2):235–264, 1990.

[Sam65] A.A. Samarskij. Monotonic difference schemes for elliptic and parabolic equations in the case of a non-selfadjoint elliptic operator. *U.S.S.R. Comput. Maths. Math. Physics*, 5(3):212–217, 1965.

[SZ90] L.R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[Thi99] K. Thiele. Adaptive finite volume discretization of density driven flows in porous media. Dissertation, Naturwissenschaftliche Fakultät I, Universität Erlangen-Nürnberg, 1999.