A new approach to a posteriori error estimation for convection-diffusion problems. I. Getting started

Sergey Korotov

Institute of Mathematics, Helsinki University of Technology
P. O. Box 1100, FIN–02015 TKK, Finland
sergey.korotov@hut.fi

Dmitri Kuzmin

Institute of Applied Mathematics, University of Dortmund
Vogelpothsweg 87, D-44227, Dortmund, Germany
kuzmin@math.uni-dortmund.de

Abstract

A new *a posteriori* error estimation technique is applied to the stationary convection-reaction-diffusion equation. In order to estimate the approximation error in the usual energy norm, the underlying bilinear form is decomposed into a computable integral and two other terms which can be estimated from above using elementary tools of functional analysis. Two auxiliary parameter-functions are introduced to construct such a splitting and tune the resulting bound. If these functions are chosen in an optimal way, the exact energy norm of the error is recovered, which proves that the estimate is sharp. The presented methodology is completely independent of the numerical technique used to compute the approximate solution. In particular, it is applicable to approximations which fail to satisfy the Galerkin orthogonality, e.g. due to an inconsistent stabilization, flux limiting, low-order quadrature rules, round-off and iteration errors etc. Moreover, the only constant that appears in the proposed error estimate is global and stems from the Friedrichs-Poincaré inequality.

Key words: convection-reaction-diffusion equations, a posteriori error estimation, adaptivity

MSC: 65N15, 65N50, 76M30

1 Introduction

Many mathematical models are based on (systems of) convection-reaction-diffusion equations which need to be discretized and solved numerically. The error incurred in the course of discretization and iterative solution is responsible for the discrepancy between the computational results and the exact solution of the governing equations. A posteriori error estimation makes it possible to assess this error and reduce it by resorting to adaptive mesh refinement. Currently, reliable error control is feasible, e.g., for finite element approximations of the Poisson equation and other elliptic problems [1, 3, 4, 5, 19, 23]. However, the derivation of reliable error estimates for convection-diffusion equations and systems of hyperbolic conservation laws still represents a challenging open problem, although significant advances were achieved during the past two decades, see, e.g., [8, 9, 11, 12, 22, 24, 25].

An inherent limitation of many a posteriori error estimation techniques is the presence of dubious constants which are difficult to estimate (cf. [6]), especially in the case of complex domains and unstructured meshes. The uncertainty involved in the computation of these constants may seriously reduce the practical utility of the resulting estimates. Moreover, some popular methods rely on the existence of an equivalent minimization problem or assume the Galerkin orthogonality. For the residual to be orthogonal to the space of test functions, the discretization must be performed by a consistent (Petrov-)Galerkin method and the resulting algebraic equations must be solved exactly. These requirements can rarely be satisfied in practice because of numerical quadrature, round-off errors, slack tolerances for iterative solvers and even programming bugs. The use of upwinding [2] or flux/slope limiters [15, 16] in finite element codes may also violate the Galerkin orthogonality.

A very promising approach to error estimation was introduced by Repin *et al.* [19, 20, 21] in the context of diffusion-type problems. Remarkably, it is applicable to any conforming approximation regardless of the numerical method used to compute it. The original derivation is based on rather sophisticated tools of functional analysis (duality theory, Helmholtz decomposition) but a simplified version was recently proposed for reaction-diffusion problems [10, 14]. In the present paper, it is extended to the case of stationary convection-diffusion equations. The resulting upper bound for the error measured in the energy norm is shown be sharp and reduce to the true error if the involved parameter-functions are chosen in an optimal way. Moreover, there is just one global constant which depends solely on the geometry of the computational domain and does not change in the course of mesh adaptation. The derivation of the new estimate and the proof of optimality are followed by a discussion of practical implementation details.

2 Problem statement

Consider the stationary convection-reaction-diffusion problem

$$\begin{cases} -\varepsilon\Delta u + \mathbf{b}\cdot\nabla u + cu = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $\Omega \subset \mathbb{R}^d,\ d \geq 1$, is a bounded domain with a Lipschitz continuous boundary $\partial\Omega$. The constant diffusion coefficient ε, the velocity field $\mathbf{b}$, and the reaction rate c are supposed to satisfy the following conditions

$$\varepsilon > 0, \quad \mathbf{b} \in \mathbf{W}^1_\infty(\Omega), \quad c \in L_\infty(\Omega). \tag{2}$$

The weak form of the above problem reads: Find $u \in H^1_0(\Omega)$ such that

$$a(u, w) = F(w), \qquad \forall w \in H^1_0(\Omega), \tag{3}$$

where the bilinear form $a(\cdot,\cdot)$ and the linear functional $F(\cdot)$ are given by

$$a(u, w) = \int_\Omega \varepsilon \nabla u \cdot \nabla w \, dx + \int_\Omega \mathbf{b} \cdot \nabla u \, w \; dx + \int_\Omega cuw \; dx \tag{4}$$

$$F(w) = \int_\Omega fw \; dx, \qquad u, w \in H^1_0(\Omega). \tag{5}$$

It is well known that the weak solution $u \in H^1_0(\Omega)$ of variational problem (3) exists and is unique provided that the following extra condition holds

$$c - \frac{1}{2}\nabla \cdot \mathbf{b} \geq 0. \tag{6}$$

Under this constraint, the bilinear form $a(\cdot,\cdot)$ is coercive

$$\begin{aligned} a(w, w) &= \int_\Omega \varepsilon \nabla w \cdot \nabla w \, dx + \int_\Omega (\mathbf{b} \cdot \nabla w) \, w \, dx + \int_\Omega cw^2 \, dx \\ &= \varepsilon \int_\Omega |\nabla w|^2 \, dx + \int_\Omega \left(c - \frac{1}{2}\nabla \cdot \mathbf{b} \right) w^2 \, dx \geq C\|w\|^2_{1,\Omega}, \end{aligned} \tag{7}$$

where C is a positive constant and $\|\cdot\|_{1,\Omega}$ denotes the standard norm in $H^1(\Omega)$. The coercivity of the bilinear form $a(\cdot,\cdot)$ implies the unique solvability of problem (3) due to the Lax-Milgram lemma (see, e.g., [7]).

3 Error estimation

The goal of *a posteriori* error estimation is to quantify the discrepancy between the exact and the numerical solution of the problem at hand. To this end, computable upper/lower bounds are to be derived for the approximation error measured in a suitably defined norm. This information makes it possible to assess the quality of the numerical results and represents a crucial ingredient of adaptive mesh refinement techniques.

Let $\bar{u}$ be some function from $H_0^1(\Omega)$ which is supposed to be an approximate solution of problem (3). Unlike other error estimation techniques for convection-diffusion equations, the methodology to be presented below is independent of the numerical method employed to compute $\bar{u}$. In particular, iteration errors and/or violation of the Galerkin orthogonality are admissible, so that $\bar{u}$ can be an arbitrary function from the admissible class.

In this paper, we will estimate the error $e := u - \bar{u}$ in the energy norm

$$|||e|||_\Omega^2 := \varepsilon \int\limits_\Omega |\nabla e|^2 \, dx + \int\limits_\Omega \left(c - \frac{1}{2} \nabla \cdot \mathbf{b} \right) e^2 \, dx = a(e, e). \tag{8}$$

In order to derive a computable upper bound for the error e, let us consider the following representation of the above expression

$$\begin{aligned} a(e,e) &= a(u - \bar{u}, u - \bar{u}) \\ &= \varepsilon \int\limits_\Omega \nabla(u - \bar{u}) \cdot \nabla(u - \bar{u}) \, dx + \int\limits_\Omega \mathbf{b} \cdot \nabla(u - \bar{u})(u - \bar{u}) \, dx + \int\limits_\Omega c(u - \bar{u}) \, (u - \bar{u}) \, dx \\ &= \int\limits_\Omega f(u - \bar{u}) \, dx - \varepsilon \int\limits_\Omega \nabla\bar{u} \cdot \nabla(u - \bar{u}) \, dx - \int\limits_\Omega \mathbf{b} \cdot \nabla\bar{u} \, (u - \bar{u}) \, dx - \int\limits_\Omega c \, \bar{u} \, (u - \bar{u}) \, dx, \end{aligned} \tag{9}$$

which follows from the integral identity (3) with test function $w = u - \bar{u}$.

Furthermore, it is worthwhile to regroup some terms in relation (9) and introduce an auxiliary vector function $\mathbf{y}^\star \in H(\mathrm{div}, \Omega)$ so that

$$\begin{aligned} a(u - \bar{u}, u - \bar{u}) &= \int\limits_\Omega [f - \mathbf{b} \cdot \nabla\bar{u} - c\,\bar{u}](u - \bar{u}) \, dx - \int\limits_\Omega \mathbf{y}^\star \cdot \nabla(u - \bar{u}) \, dx \\ &+ \int\limits_\Omega [\mathbf{y}^\star - \varepsilon\nabla\bar{u}] \cdot \nabla(u - \bar{u}) \, dx. \end{aligned} \tag{10}$$

Using the Green formula, we can integrate the second term in the right-hand side of the above formula by parts, which yields

$$\begin{aligned} a(u - \bar{u}, u - \bar{u}) &= \int\limits_\Omega [f - \mathbf{b} \cdot \nabla\bar{u} - c\,\bar{u} + \nabla \cdot \mathbf{y}^\star](u - \bar{u}) \, dx \\ &+ \int\limits_\Omega [\mathbf{y}^\star - \varepsilon\nabla\bar{u}] \cdot \nabla(u - \bar{u}) \, dx. \end{aligned} \tag{11}$$

Finally, let us introduce another auxiliary function $v \in H_0^1(\Omega)$ and consider the following decomposition of the energy norm

$$a(u - \bar{u}, u - \bar{u}) = I_1 + I_2 + I_3, \tag{12}$$

where the terms $I_1,\ I_2$ and I_3 are defined as follows

$$I_1 = \int\limits_\Omega \left[f - \mathbf{b}\cdot\nabla\bar{u} - c\bar{u} + \nabla\cdot(\mathbf{y}^\star - \mathbf{b}v) + cv\right](u-\bar{u})\ dx, \tag{13}$$

$$I_2 = \int\limits_\Omega \left[\mathbf{y}^\star - \varepsilon\nabla(\bar{u}-v)\right]\cdot\nabla(u-\bar{u})\ dx, \tag{14}$$

$$I_3 = \int\limits_\Omega \left[(\nabla\cdot(\mathbf{b}v) - c\,v)(u-\bar{u}) - \varepsilon\nabla v\cdot\nabla(u-\bar{u})\right]\ dx. \tag{15}$$

Integration by parts using Green's formula yields

$$\begin{aligned} I_3 &= \int\limits_\Omega \left[v(\mathbf{b}\cdot\nabla\bar{u} + c\,\bar{u}) + \varepsilon\nabla v\cdot\nabla\bar{u}\right]\ dx \\ &- \int\limits_\Omega \left[v(\mathbf{b}\cdot\nabla u + c\,u) + \varepsilon\nabla v\cdot\nabla u\right]\ dx \\ &= \int\limits_\Omega \left[v(\mathbf{b}\cdot\nabla\bar{u} + c\,\bar{u} - f) + \varepsilon\nabla v\cdot\nabla\bar{u}\right]\ dx \qquad (16) \\ &= a(\bar{u},v) - F(v) = R(v,\bar{u}), \end{aligned}$$

where $R(v,\bar{u})$ is the residual of problem (3) for $w = v$ and $\bar{u}$ in place of u.

Hence, the term I_3 is computable and it remains to derive an upper bound for the integrals I_1 and I_2. The Cauchy-Schwarz inequality yields

$$I_1 \le \|f - \mathbf{b}\cdot\nabla\bar{u} - c\bar{u} + \nabla\cdot(\mathbf{y}^\star - \mathbf{b}v) + cv\|_{0,\Omega}\|u-\bar{u}\|_{0,\Omega} \tag{17}$$

and the Friedrichs-Poincaré inequality which reads

$$\|w\|_{0,\Omega} \le C_\Omega\|\nabla w\|_{0,\Omega}, \quad \forall w\in H_0^1(\Omega), \tag{18}$$

where C_Ω is a positive constant and $\|\cdot\|_{0,\Omega}$ is the standard L_2-norm. Thus,

$$I_1 \le C_\Omega\|f - \mathbf{b}\cdot\nabla\bar{u} - c\bar{u} + \nabla\cdot(\mathbf{y}^\star - \mathbf{b}v) + cv\|_{0,\Omega}\|\nabla(u-\bar{u})\|_{0,\Omega}. \tag{19}$$

Similarly, the Cauchy-Schwarz inequality yields the estimate

$$I_2 \le \|\mathbf{y}^\star - \varepsilon\nabla(\bar{u}-v)\|_{0,\Omega}\|\nabla(u-\bar{u})\|_{0,\Omega}. \tag{20}$$

Combining inequalities (19) and (20) we obtain an estimate of the form

$$I_1 + I_2 \le \Lambda(v,\mathbf{y}^\star,\bar{u})\|\nabla(u-\bar{u})\|_{0,\Omega}, \tag{21}$$

where the functional $\Lambda(v,\mathbf{y}^\star,\bar{u})$ is given by the relation

$$\begin{aligned} \Lambda(v,\mathbf{y}^\star,\bar{u}) &= C_\Omega\|f - \mathbf{b}\cdot\nabla\bar{u} - c\bar{u} + \nabla\cdot(\mathbf{y}^\star - \mathbf{b}v) + cv\|_{0,\Omega} \\ &+ \|\mathbf{y}^\star - \varepsilon\nabla(\bar{u}-v)\|_{0,\Omega}. \qquad (22) \end{aligned}$$

The Young inequality implies that for any $p > 0$ and $q > 0$

$$pq \le \frac{\sigma}{2} p^2 + \frac{1}{2\sigma} q^2, \qquad \sigma > 0. \tag{23}$$

This enables us to estimate the product in the right-hand side of (21) in terms of the energy norm (8) which resides in the left-hand side of (12)

$$\begin{aligned} I_1 + I_2 &\le \frac{\sigma}{2} \Lambda(v, \mathbf{y}^\star, \bar{u}) \|\nabla(u - \bar{u})\|^2_{0,\Omega} + \frac{1}{2\sigma} \Lambda(v, \mathbf{y}^\star, \bar{u}) & (24) \\ &\le \frac{\sigma}{2\varepsilon} \Lambda(v, \mathbf{y}^\star, \bar{u}) |||u - \bar{u}|||^2_\Omega + \frac{1}{2\sigma} \Lambda(v, \mathbf{y}^\star, \bar{u}). & (25) \end{aligned}$$

Finally, we substitute this inequality into (12) and recall (16)

$$|||u - \bar{u}|||^2_\Omega \le \frac{\sigma}{2\varepsilon} \Lambda(v, \mathbf{y}^\star, \bar{u}) |||u - \bar{u}|||^2_\Omega + \frac{1}{2\sigma} \Lambda(v, \mathbf{y}^\star, \bar{u}) + R(v, \bar{u}). \tag{26}$$

Thus, the energy norm of the error is bounded from above by

$$|||u - \bar{u}|||^2_\Omega \le \frac{R(v, \bar{u}) + \frac{1}{2\sigma} \Lambda(v, \mathbf{y}^\star, \bar{u})}{1 - \frac{\sigma}{2\varepsilon} \Lambda(v, \mathbf{y}^\star, \bar{u})}, \tag{27}$$

where the free parameter $\sigma > 0$ is to be chosen so that

$$\frac{\sigma}{2\varepsilon} \Lambda(v, \mathbf{y}^\star, \bar{u}) < 1. \tag{28}$$

Thus, the desired upper bound for the energy norm of the error reads

$$|||e|||^2_\Omega \le \mathbf{EST}(\sigma, \mathbf{y}^\star, v, \bar{u}), \tag{29}$$

where **EST** denotes the (computable) right-hand side of (27).

Recall that estimate (29) is valid for an arbitrary choice of $\mathbf{y}^\star \in H(\mathrm{div}, \Omega)$, $v \in H^1_0(\Omega)$ and $\sigma > 0$ satisfying (28). Clearly, these parameters should be designed so as to minimize the functional **EST** as far as possible. Let the corresponding optimal values be denoted by $\mathbf{y}^\star_{opt}$, v_{opt} and σ_{opt}, respectively. In the next section we will show that the optimal upper bound

$$\underline{\mathbf{EST}} := \mathbf{EST}(\sigma_{opt}, \mathbf{y}^\star_{opt}, v_{opt}, \bar{u}) \tag{30}$$

reduces to the energy norm (8), which means that estimate (29) is sharp.

Remark. Clearly, the standard energy norm (8) does not provide a proper control of the error if the diffusion coefficient ε is small as compared to $|\mathbf{b}|$, i.e., if the Peclet number is large. Such convection-dominated problems call for the use of (artificial) streamline diffusion in the variational formulation, as explained, e.g., in the book by Johnson ([13], Chapter 9). In order to construct a proper norm for a singularly perturbed problem with $\varepsilon \ll 1$ or $\varepsilon = 0$ (pure convection), it is worthwhile to add a certain amount of streamline diffusion to (3) even if the approximate solution $\bar{u}$ is computed using a different stabilization technique, e.g., finite volume upwinding or algebraic flux correction [15]. The use of streamline diffusion in a stabilized version of estimate (27) will be addressed in a sequel to this paper.

4 Sharpness of the estimate

In this section, we will demonstrate that there exist certain values of the free parameters $\mathbf{y}^\star$, v and σ such that (29) holds as equality, i.e., the optimal upper bound (30) yields the real error e measured in the energy norm.

Let $v \in H_0^1(\Omega)$ be the weak solution of the adjoint problem

$$a^*(v, w) = R(w, \bar{u}), \qquad \forall w \in H_0^1(\Omega), \tag{31}$$

where $a^*(\cdot, \cdot)$ is a bilinear form such that $a^*(v, w) = a(w, v)$, i.e.,

$$a^*(v, w) = \int\limits_\Omega \varepsilon \nabla v \cdot \nabla w \, dx - \int\limits_\Omega [\nabla \cdot (\mathbf{b} v) - c\, v] w \, dx. \tag{32}$$

The linear functional $R(w, \bar{u}) = a(\bar{u}, w) - F(w)$ represents the residual of the primal problem (3) evaluated using $\bar{u}$ instead of u. That is,

$$R(w, \bar{u}) = \int\limits_\Omega \varepsilon \nabla \bar{u} \cdot \nabla w \, dx + \int\limits_\Omega [\mathbf{b} \cdot \nabla \bar{u} + c\, \bar{u} - f] w \, dx. \tag{33}$$

The unique solvability of problem (31) follows from the Lax-Milgram lemma.

Furthermore, let us define the free parameter $\mathbf{y}^\star$ as follows

$$\mathbf{y}^\star = \varepsilon \nabla (\bar{u} - v). \tag{34}$$

In order to prove that the so-defined $\mathbf{y}^\star$ does belong to the space $H(\mathrm{div}, \Omega)$, let us represent (31) in the following form

$$\int\limits_\Omega \varepsilon \nabla (\bar{u} - v) \cdot \nabla w \, dx + \int\limits_\Omega g(\bar{u}, v)\, w \, dx = 0, \quad \forall w \in H_0^1(\Omega), \tag{35}$$

where

$$g(\bar{u}, v) = \mathbf{b} \cdot \nabla \bar{u} + c\bar{u} + \nabla \cdot (\mathbf{b} v) - cv - f \quad \in L_2(\Omega). \tag{36}$$

Plugging (34) into (35), we obtain the integral identity

$$\int\limits_\Omega \mathbf{y}^\star \cdot \nabla w \, dx + \int\limits_\Omega g(\bar{u}, v)\, w \, dx = 0, \quad \forall w \in H_0^1(\Omega), \tag{37}$$

which shows that $\mathbf{y}^\star \in H(\mathrm{div}, \Omega)$ and its divergence is implicitly defined as

$$\nabla \cdot \mathbf{y}^\star = g(\bar{u}, v). \tag{38}$$

Hence, the integral I_1 vanishes for the above choice of v and $\mathbf{y}^\star$

$$\begin{aligned} I_1 &= \int\limits_\Omega [f - \mathbf{b}\cdot\nabla\bar{u} - c\bar{u} + \nabla\cdot(\mathbf{y}^\star - \mathbf{b}v) + cv]\,(u-\bar{u})\;dx \\ &= \int\limits_\Omega [\nabla\cdot\mathbf{y}^\star - g(\bar{u},v)]\,(u-\bar{u})\;dx = 0. \end{aligned} \tag{39}$$

Moreover, definition (34) renders the integral I_2 equal to zero

$$I_2 = \int\limits_\Omega [\mathbf{y}^\star - \varepsilon\nabla(\bar{u}-v)]\cdot\nabla(u-\bar{u})\;dx = 0. \tag{40}$$

It follows from (12) that the energy norm of the error is given by

$$a(u-\bar{u},u-\bar{u}) = R(v,\bar{u}). \tag{41}$$

In view of definitions (34) and (38), the contributions of I_1 and I_2 to the upper bound $\underline{\mathbf{EST}}$ vanish as well, i.e., $\Lambda(v,\mathbf{y}^\star,\bar{u}) = 0$. Therefore, the parameter $\sigma > 0$ can be chosen arbitrarily and we have

$$\underline{\mathbf{EST}} = R(v,\bar{u}). \tag{42}$$

Since the right-hand sides of (41) and (42) are equal, we obtain

$$|||e|||_\Omega^2 = a(u-\bar{u},u-\bar{u}) = \underline{\mathbf{EST}}, \tag{43}$$

which proves that the upper bound $\underline{\mathbf{EST}}$ is optimal and cannot be improved.

5 Practical implementation

Some remarks are in order regarding the practical use of estimate (27) for mesh adaptation purposes. In this section, we provide some guidelines for the choice of free parameters and decompose the upper bound **EST** into a sum of element contributions, as required to assess the local mesh resolution. In addition, we present a simple estimate for the Friedrichs constant C_Ω.

5.1 Parameter settings

In practice, the optimal values of v and $\mathbf{y}^\star$ are not available, since the adjoint problem is as difficult to solve as the primal one. However, usable approximations can be obtained by solving the adjoint problem numerically. In the finite element framework, the discrete counterpart of (31) reads

$$a^*(v_h,w_h) = R(w_h,\bar{u}), \qquad \forall w_h \in V_h^*, \tag{44}$$

where V_h^* is a finite-dimensional subspace of $H_0^1(\Omega)$. Thus, it is natural to consider $\bar{v} := v_h \in V_h^*$ but any other approximation of v_{opt} is also admissible.

Ideally, the concomitant function $\bar{\mathbf{y}}^\star \in H(\mathrm{div}, \Omega)$ should be chosen so as to minimize the functional $\Lambda(\bar{v}, \mathbf{y}^\star, \bar{u})$ which was shown to vanish for the optimal choice of v and $\mathbf{y}^\star$. Using the following inequality

$$(p+q)^2 \leq (1+\beta)p^2 + \left(1 + \frac{1}{\beta}\right) q^2, \quad \forall \beta > 0 \tag{45}$$

the square of $\Lambda(\bar{v}, \mathbf{y}^\star, \bar{u})$ as defined in (22) can be estimated as follows

$$\begin{aligned}[\Lambda(\bar{v}, \mathbf{y}^\star, \bar{u})]^2 \;&\leq\; (1+\beta)C_\Omega^2 \|f - \mathbf{b}\cdot\nabla\bar{u} - c\bar{u} + \nabla\cdot(\mathbf{y}^\star - \mathbf{b}\bar{v}) + c\bar{v}\|_{0,\Omega}^2 \\ &+\; \left(1 + \frac{1}{\beta}\right) \|\mathbf{y}^\star - \varepsilon\nabla(\bar{u} - \bar{v})\|_{0,\Omega}^2 = \eta(\bar{v}, \mathbf{y}^\star, \bar{u}, \beta). \end{aligned} \tag{46}$$

For a fixed $\bar{v} \in H_0^1(\Omega)$ and $\beta > 0$, it is possible to determine $\bar{\mathbf{y}}^\star$ by solving a minimization problem for the quadratic functional $\eta(\bar{v}, \mathbf{y}^\star, \bar{u}, \beta)$, as explained in [10, 14, 19] in the context of reaction-diffusion equations. As soon as $\bar{v}$ and $\bar{\mathbf{y}}^\star$ are available, the remaining free parameter σ can be chosen so as to minimize the upper bound **EST** subject to constraint (28).

On the other hand, the use of optimization tools is computationally expensive and will hardly pay off in real-life applications. Another way to compute $\bar{\mathbf{y}}^\star$ for a given $\bar{v}$ is to use definition (34) and a suitable gradient averaging technique such as the standard L_2-projection

$$\int_\Omega \bar{\mathbf{y}}^\star w \, dx = \int_\Omega \varepsilon\nabla(\bar{u} - \bar{v}) w \, dx, \qquad \forall w \in V_h^\star. \tag{47}$$

This relation gives rise to an algebraic system with a consistent mass matrix which can be 'lumped' in the usual way. Furthermore, slope limiters can be applied to the averaged gradient, so as to enforce the natural upper and lower bounds provided by the values of the underlying discontinuous function at interelement boundaries [18]. This sort of postprocessing is to be recommended for problems with steep gradients.

It is worth mentioning that if $\bar{u} = u_h \in V_h \subset H_0^1(\Omega)$ is a true Galerkin solution of the primal problem, then $R(w_h, \bar{u}) = 0, \; \forall w_h \in V_h$. In particular, the term $I_3 = R(\bar{v}, \bar{u})$ is equal to zero if $\bar{v} \in V_h$. Likewise, if $V_h^* = V_h$ in (44), then the right-hand side vanishes and the solution is trivial

$$\bar{v} = 0, \qquad \bar{\mathbf{y}}^\star = G_h(\varepsilon\nabla\bar{u}), \tag{48}$$

where G_h denotes the gradient averaging operator. This parameter constellation may result in a rather crude estimate. In order to improve it, the adjoint problem should be solved on a finer/adapted mesh, so as to obtain a nontrivial solution $\bar{v}$ even if $\bar{u}$ enjoys the Galerkin orthogonality.

Alternatively, it is possible to take $\bar{\mathbf{y}}^\star = G_h(\varepsilon\nabla\bar{u})$ and define $\bar{v}$ as an approximate solution to the weak form of equation (38)

$$\int_\Omega [\nabla\cdot(\mathbf{b}v) - cv]w\,dx = \int_\Omega [f - \mathbf{b}\cdot\nabla\bar{u} - c\bar{u} + \nabla\cdot\bar{\mathbf{y}}^\star]w\,dx, \quad \forall w\in V_h \tag{49}$$

which can be solved on the same mesh as the primal problem. The resulting solution $\bar{v}$ is supposed to take the convective transport of errors into account. Note that the right-hand side of the hyperbolic problem (49) contains a weighted residual of (1), where the diffusive flux $\varepsilon\nabla u$ is approximated by $\bar{\mathbf{y}}^\star$.

In light of the above, there is a lot of flexibility in the choice of the free parameters $\bar{v}$ and $\bar{\mathbf{y}}^\star$. In particular, they can be defined by

- solving (44) followed by a minimization problem for $\eta(\bar{v},\mathbf{y}^\star,\bar{u},\beta)$;
- computing approximations to v_{opt} and $\mathbf{y}^\star_{opt}$ from (44)/(47) or (48);
- using the averaged gradient $\bar{\mathbf{y}}^\star = G_h(\varepsilon\nabla\bar{u})$ and solving (49) for $\bar{v}$.

A detailed numerical study of these (and other) parameter settings is beyond the scope of this paper and will be presented in a forthcoming publication.

5.2 Local error control

In order to assess the local mesh quality, it is necessary to identify individual element contributions to the global bound given by (27). The term $R(\bar{v},\bar{u})$ can be readily decomposed into a sum of element residuals but the treatment of the functional $\Lambda(\bar{v},\bar{\mathbf{y}}^\star,\bar{u})$, which makes a positive contribution to the nominator and a negative contribution to the denominator of **EST**, calls for further explanation. For our purposes, it is instructive to consider

$$\sigma = \frac{\alpha}{\Lambda(\bar{v},\bar{\mathbf{y}}^\star,\bar{u})}, \qquad \text{where} \quad 0<\alpha<2\varepsilon. \tag{50}$$

By virtue of (46), the resulting value of **EST** satisfies the inequality

$$\mathbf{EST} \le \frac{1}{1-\frac{\alpha}{2\varepsilon}}\left[R(\bar{v},\bar{u}) + \frac{1}{2\alpha}\eta(\bar{v},\bar{\mathbf{y}}^\star,\bar{u},\beta)\right] = \overline{\mathbf{EST}}. \tag{51}$$

The adjustable weighting factor β provides an additional degree of freedom which can be used to vary the shares of the two squared L_2-norms that contribute to the quadratic functional $\eta(\bar{v},\bar{\mathbf{y}}^\star,\bar{u},\beta)$.

Let $\mathcal{T}_h$ be a triangulation of the computational domain Ω, so that $\overline{\mathbf{EST}}$ admits the following decomposition into a sum of element contributions

$$\overline{\mathbf{EST}} = \frac{1}{1-\frac{\alpha}{2\varepsilon}}\left[\sum_{K\in\mathcal{T}_h} R(\bar{v},\bar{u})|_K + \frac{1}{2\alpha}\sum_{K\in\mathcal{T}_h}\eta(\bar{v},\bar{\mathbf{y}}^\star,\bar{u},\beta)|_K\right]. \tag{52}$$

An adaptive mesh for the primal and/or adjoint problem can be constructed using the principle of error equidistribution. An element $K \in \mathcal{T}_h$ for which (the absolute value of) $R(\bar{v},\bar{u})|_K$ and/or $\eta(\bar{v},\bar{\mathbf{y}}^\star,\bar{u},\beta)|_K$ is much greater/smaller than the average value over all elements of the current mesh $\mathcal{T}_h$ is a natural candidate for refinement/coarsening, respectively.

Remark. Note that the functional $\eta(\bar{v},\bar{\mathbf{y}}^\star,\bar{u},\beta)$ provides a perfect error indicator for the adjoint problem, since it equals zero for the exact solution of (31) combined with $\mathbf{y}^\star_{opt}$ as defined in (34). At the same time, the term $R(\bar{v},\bar{u})$ is associated with the approximation of the primal problem and approaches the squared energy norm of the error as $\bar{v} \to v_{opt}$ and $\bar{\mathbf{y}}^\star \to \mathbf{y}^\star_{opt}$.

5.3 Estimation of C_Ω

Finally, we remark that the constant C_Ω that appears in (27) is global and depends solely on the geometry of the domain. In the case of homogeneous Dirichlet boundary conditions considered in this paper, a usable estimate can be readily obtained by enclosing the domain Ω into a rectangular box as proposed by Mikhlin (see [17], p. 18). The resulting upper bound reads

$$C_\Omega \leq \frac{1}{\pi\sqrt{\frac{1}{a_1^2}+\cdots+\frac{1}{a_d^2}}}, \tag{53}$$

where $a_1,\ldots,a_d$ are the dimensions of the box. Note that C_Ω is independent of the mesh and needs to be evaluated only once for each particular domain.

6 Conclusions

Building on the methodology presented in [10, 14, 19, 20, 21], a new approach to a posteriori error estimation for stationary convection-reaction-diffusion equations was developed. The resulting upper bound for the energy norm of the error can be optimized by tuning two auxiliary functions related to the solution of the associated adjoint problem. A further advantage is the simplicity of derivation and the lack of dubious constants. An extension to the case of mixed boundary conditions can be readily performed as explained in [10, 14] in the framework of reaction-diffusion equations. Some promising directions for further research include the derivation of a better estimate for the integrals I_1 and I_2, error estimation in different norms, and the use of streamline diffusion for convection-dominated problems.

Acknowledgements: The first author was supported by the Academy Research Fellowship no. 208628 from the Academy of Finland. The authors are grateful to Prof. Dr. Stefan Turek for valuable remarks.

References

[1] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis.* John Wiley & Sons, 2000.

[2] L. Angermann, A posteriori error estimates for FEM with violated Galerkin orthogonality. *Numer. Methods Partial Differential Equations* **18** (2002) 241–259.

[3] I. Babuška and T. Strouboulis. *The Finite Element Method and Its Reliability.* Oxford University Press Inc., New York, 2001.

[4] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations.* Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2003.

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

[6] C. Carstensen and S. A. Funken, Constants in Clément-interpolation error and residual based a posteriori error estimates in finite element methods. *East-West J. Numer. Math.* **8** (2000) 153–175.

[7] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam, 1978.

[8] B. Cockburn and H. Gau, A posteriori error estimates for general numerical methods for scalar conservation laws. *Comput. Appl. Math.* **14** (1995) 37–47.

[9] K. Eriksson and C. Johnson, Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems. *Math. Comp.* **60** (1993) 167–188.

[10] A. Hannukainen and S. Korotov, Computational technologies for reliable control of global and local errors for linear elliptic type boundary value problems. Technical Report **A494**, Helsinki University of Technology, 2006. Submitted to *J. Numer. Anal. Comput. Math.*

[11] R. Hartmann and P. Houston. Adaptive discontinuous Galerkin finite element methods for nonlinear hyperbolic conservation laws. *SIAM J. Sci. Comput.* **24** (2002) 979–1004.

[12] V. John. A numerical study of a posteriori error estimators for convection-diffusion equations. *Comput. Methods Appl. Mech. Engrg.* **190** (2000) 757–781.

[13] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method.* Studentlitteratur, Lund, 1987.

[14] S. Korotov. A posteriori error estimation for linear elliptic problems with mixed boundary conditions. Technical Report **A495**, Helsinki University of Technology, 2006. To appear in *Appl. Math.*

[15] D. Kuzmin, R. Löhner and S. Turek (eds.) *Flux-Corrected Transport: Principles, Algorithms, and Applications.* Springer, 2005.

[16] R. Löhner, K. Morgan, J. Peraire and M. Vahdati, Finite element flux-corrected transport (FEM-FCT) for the Euler and Navier-Stokes equations. *Int. J. Numer. Meth. Fluids* **7** (1987) 1093–1109.

[17] S. G. Mikhlin. Constants in Some Inequalities of Analysis. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1986.

[18] M. Möller and D. Kuzmin, Adaptive mesh refinement for high-resolution finite element schemes. *Int. J. Numer. Meth. Fluids* **52** (2006) 545-569.

[19] P. Neittaanmäki and S. Repin, *Reliable Methods for Computer Simulation: Error Control and A Posteriori Estimates.* Elsevier, 2004.

[20] S. Repin, S. Sauter and A. Smolianski, A posteriori error estimation for the Dirichlet problem with account of the error in the approximation of boundary conditions. *Computing* **70** (2003) 205–233.

[21] S. Repin, S. Sauter and A. Smolianski, A posteriori error estimation for the Poisson equation with mixed Dirichlet/Neumann boundary conditions. *J. Comput. Appl. Math.* **164/165** (2004) 601–612.

[22] E. Süli and P. Houston, Adaptive finite element approximation of hyperbolic problems. In: T.J. Barth and H. Deconinck (eds.), *Error estimation and adaptive discretization methods in computational fluid dynamics.* LNCSE **25**, Springer, 2003, 269–344.

[23] R. Verfürth. A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques, Wiley-Teubner, 1996.

[24] R. Verfürth. A posteriori error estimators for convection-diffusion equations. *Numer. Math.* **80** (1998) 641–663.

[25] R. Verfürth. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.* **43** (2005) 1766–1782.