INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 8, Number 1, Pages 1–27

ROBUST A-POSTERIORI ESTIMATORS FOR MULTILEVEL DISCRETIZATIONS OF REACTION–DIFFUSION SYSTEMS

VIVIANE KLEIN AND MAŁGORZATA PESZYŃSKA

Abstract. We define a multilevel finite element discretization for a coupled stationary reaction–diffusion system in which each component can be defined on a separate grid. We prove convergence of the scheme and propose residual a-posteriori estimators for the error in the natural energy norm for the system. The estimators are robust in the coefficients of the system. We prove upper and lower bounds and illustrate the theory with numerical experiments.

Key words. multilevel finite elements, a-priori error estimates, a-posteriori error estimators, reaction-diffusion system

1. Introduction

In this paper we develop a-priori and a-posteriori analysis for finite-element discretizations of stationary reaction-diffusion systems. We are particularly interested in developing results which are uniform, in the sense made precise below, for families of such systems characterized by coefficients of different orders of magnitude. Additionally, since the individual components of such systems may have different variability, we recognize that they should be approximated on multilevel grids. The choice of such grids is guided by the a-posteriori error estimators.

A-posteriori analysis for finite element approximations of scalar self-adjoint elliptic equations is well developed [5, 35, 12]. The various error estimators that have been proposed differ in how closely they estimate the error and in the complexity of implementation and computations. In addition, their properties may depend significantly on the coefficients of the underlying problem.

Consider first the scalar stationary reaction–diffusion equation

$$-\nabla \cdot (a\nabla u) + \kappa u \quad = \quad f, \tag{1}$$

with a solution u. Consider also the corresponding standard Galerkin finite element formulation for (1) with a solution u_h, and an a-posteriori estimator η_s for the error $\mathcal{E}_s = \| u - u_h \|$ in the energy norm $\| \cdot \|$ associated with (1).

In general, the efficiency index $\theta_s := \frac{\eta_s}{\mathcal{E}_s}$ may significantly depend on the parameters in $\mathcal{P}_s = (a, \kappa)$. Standard theory, cf. [10, 5], considers $\mathcal{P}_s = \mathbf{1}^2 := (1, 1)$ and does not extend easily to the families of (1) where the parameters in $\mathcal{P}_s$ vary significantly. The concept of *robustness* [37, 36, 9, 39, 39, 40, 27, 26] allows to study such families of problems (1): the estimator η_s for (1) is *robust* if θ_s is uniform in $\mathcal{P}_s$, i.e., it remains constant or at least stable for a wide range of values in $\mathcal{P}_s$. Robust estimators are applicable, e.g., to singularly perturbed problems.

Now consider the problem of interest in this paper: the *system* of stationary reaction-diffusion equations posed in some domain Ω parametrized by $\mathcal{P} =$

Received by the editors March 2, 2010 and, in revised form, April 21, 2010.
2000 *Mathematics Subject Classification.* 65N30, 65N15.
This research was partially supported by the grants NSF 0511190 and DOE 98089. Peszynska was also partially supported by the Fulbright Scholar Research program 2009-2010.

$(\lambda_1, \lambda_2, a, b, c)$

$$\lambda_1 u - \nabla \cdot (a\nabla u) + c(u - v) = f, \quad x \in \Omega, \tag{2}$$

$$\lambda_2 v - \nabla \cdot (b\nabla v) + c(v - u) = g, \quad x \in \Omega, \tag{3}$$

and complemented by appropriate boundary conditions.

The applications and analysis of general reaction-diffusion systems are considered, among other works, in [6, 33]. The special case of zero'th order coupling term associated with c, and of coefficients in $\mathcal{P}$ that may vary from case to case by orders of magnitude, has several applications. See for example the reaction-diffusion systems as in [23], double diffusion systems [7, 22, 31], ([30], II.5), and singular perturbations or regularizations of degenerate systems such as first-order reactions or adsorption at equilibrium and non-equilibrium [6, 25, 32]. See also pseudo-parabolic systems [24]. Our interest in this paper is in the numerical schemes; applications will be presented elsewhere.

An important observation is true for the families of solutions corresponding to the families of $\mathcal{P}$. In some applications the components u and v of the solution to (2)-(3) may have significantly different variability. In such cases it is natural to approximate the smooth component on a coarse grid and the less-smooth component on a fine grid. Such a multilevel discretization requires appropriate grid transfer operators so that the coupling term can be defined and the convergence ensured.

In addition, note that (2)-(3) can be seen as a prototype of a discretized-in-time parabolic system. While a-posteriori error estimation for parabolic problems can proceed along several paths [21, 34], some involve the consideration of robust estimates for (1) [3], and of the separation of spatial and temporal discretization errors without solving dual problems and/or backward heat equation [38, 8].

The above remarks motivate our work on robust estimators for the system (2)-(3). Our results i) extend the scalar estimators from [37] to the case of a coupled system, and ii) extend the work [4, 2] in which $\mathcal{P}$ was fixed. In addition, to our knowledge, ours is the only result concerning iii) multilevel schemes for (2)–(3).

A separate direction from a-posteriori error estimation is the use of special grids such as Shishkin and equidistributed meshes for resolving boundary layers in singularly perturbed problems [28, 19, 20]. For scalar problems (1) in 1D, it can be shown that with such grids, the dependence of the error of numerical solution on the parameters in $\mathcal{P}_s$ can be eliminated, e.g., by applying the MMPDE [19, 20]. We are unsure however how such grids can be constructed for systems when more than one of the parameters vary; it appears that the methods would not be a straightforward extension of [19].

The paper is organized as follows. We introduce notation and preliminaries in Section 2. In Section 3 we prove a-priori estimates for the multilevel discretization of (2)–(3). The main results of this paper are given in Section 4 where we define appropriate a-posteriori error estimators and prove upper and lower bounds; the estimators that we develop are robust in $\mathcal{P}$. Our theoretical results are illustrated by numerical experiments presented in Section 5.

We close with a few remarks on notation. Throughout the paper C means a generic positive constant; its value is different in each context in which it is used. The symbol $\partial_n w$ denotes the normal component of ∇w with respect to some boundary or edge. In all integrals we omit the symbol of integration variable; this helps to keep the expressions compact. Next, our theoretical results are given for $d = 2, 3$ spatial dimensions. The case $d = 1$ is also covered by the theory but the standard nomenclature and assumptions [14] do not apply; see [26] for robust estimates in $d = 1$.

2. Preliminaries

We consider the equations (2)-(3) defined over an open bounded polygonal domain $\Omega \subset \mathbb{R}^d, d = 2, 3$, with a Lipschitz boundary $\partial\Omega$, on which the homogeneous Dirichlet boundary conditions are posed

$$u|_{\partial\Omega} = v|_{\partial\Omega} \quad = \quad 0, \quad x \in \partial\Omega. \tag{4}$$

For any subset $\omega \subseteq \Omega$ with Lipschitz boundary, we use the standard notation for Lebesgue $L^2(\omega), L^\infty(\omega)$, and Sobolev spaces $H^k(\omega)$, $k \in \mathbb{N}$. These are equipped with the usual seminorms $|\cdot|_{k,\omega}$, norms $\|\cdot\|_{k,\omega} := \|\cdot\|_{H^k(\omega)}$ and the usual scalar product(s) $(f, \psi)_\omega := (f, \psi)_{L^2(\omega)} = \int_\omega f(x)\psi(x)$ [1, 14]. If $\omega = \Omega$, the subscript ω will be omitted. We also set $V := H_0^1(\Omega)$ as the closure of $C_0^\infty(\Omega)$ under the standard norm in $H^1(\Omega)$, $\|w\|_V := \|w\|_1 := (\|w\|_0^2 + \|\nabla w\|_0^2)^{1/2}$.

We assume further that the coefficients of (2)-(3) in $\mathcal{P}$ are positive constants

$$a, b, c, \lambda_1, \lambda_2 > 0. \tag{5}$$

A more general case of variable and degenerate coefficients will be considered elsewhere.

On the product space $V \times V$ we define several norms. First, we denote the Euclidean product norm $\|(u, v)\|_{V\times V} := (\|u\|_V^2 + \|v\|_V^2)^{1/2}$. Next, we consider the scaled norm $\|(u, v)\|_s := (\int_\Omega(\lambda_1 u^2 + a(\nabla u)^2 + \lambda_2 v^2 + b(\nabla v)^2))^{1/2}$, and an additional energy norm to be defined below. It is not hard to see that $\|\cdot\|_{V\times V}$ and $\|\cdot\|_s$ are equivalent.

We define now the functional $A : (V \times V) \times (V \times V) \to \mathbb{R}$

$$A((u, v), (\phi, \psi)) := \int_\Omega (\lambda_1 u\phi + \lambda_2 v\psi) + \int_\Omega (a\nabla u \cdot \nabla\phi + b\nabla v \cdot \nabla\psi) + \int_\Omega (c(u - v)(\phi - \psi)). \tag{6}$$

Clearly, A is symmetric, bilinear, and continuous with respect to the product norm $\|\cdot\|_{V\times V}$

$$|A((u, v), (\phi, \psi))| \le C_A\|(u, v)\|_{V\times V}\|(\phi, \psi)\|_{V\times V},$$

with some constant $C_A > 0$ independent of u, v, ϕ, ψ; this follows easily by an application of Cauchy-Schwarz inequality. We also see by $c(u - v)u + c(v - u)v = c(u - v)^2 \ge 0$ that $A(\cdot, \cdot)$ is coercive in the product norm i.e. there is a constant $\alpha_A > 0$ such that

$$\alpha_A\|(u, v)\|_{V\times V}^2 \le A((u, v), (u, v)).$$

We note that both C_A, α_A depend on $\mathcal{P}$.

Thus $A(\cdot, \cdot)$ can be used en-lieu of the standard inner product on $V \times V$. We associate with $A(\cdot, \cdot)$ the energy norm $\|\cdot\|_e$ on $V \times V$ with respect to which it is naturally coercive and continuous with unit constants

$$\|(u, v)\|_e^2 = A((u, v), (u, v)) = \|(u, v)\|_s^2 + \int_\Omega c(u - v)^2. \tag{7}$$

Remark 2.1. *All the norms $\|\cdot\|_{V\times V}, \|\cdot\|_s, \|\cdot\|_e$ are equivalent and the equivalence constants depend on $\mathcal{P}$.*

Now assume $(f, g) \in H^{-1}(\Omega) \times H^{-1}(\Omega)$ and define a functional $L : V \times V \to \mathbb{R}$

$$L((\phi, \psi)) \quad = \quad \langle f, \phi\rangle + \langle g, \psi\rangle \tag{8}$$

where $\langle \cdot, \cdot\rangle$ denotes the standard duality pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$. It is standard that L is linear and continuous [29].

With the above definitions, we are ready to state the weak form of the problem (2)–(3). We seek $(u, v) \in V \times V$ so that

$$A((u,v),(\phi,\psi)) = L((\phi,\psi)),\ \forall(\phi,\psi) \in V \times V. \tag{9}$$

It follows from Lax-Milgram Theorem [29, 16] that (9) admits a unique solution. In addition, if $f, g \in L^2(\Omega) \times L^2(\Omega)$, then one can easily see [17] that both components u, v are in $H^2(\Omega)$ by regularity of elliptic equations, since each component solves an elliptic problem with a source term in $L^2(\Omega)$.

2.1. Finite element discretization. The notation and nomenclature below is standard [14, 10]; we follow closely [35, 37, 36].

We denote by $\mathcal{T}_h, h > 0$, a family of partitions of Ω into a finite number of elements. We require that the elements in any partition $\mathcal{T}_h$ satisfy the standard admissibility and shape-regularity properties [35]. We denote by $\mathcal{E}_h$ the set of all edges in the partition $\mathcal{T}_h$ that are not contained in $\partial\Omega$. For any element $T \in \mathcal{T}_h$ we let $\tilde{\omega}_T$ be the set of all elements that share a vertex or an edge with T and h_T be the diameter of T. We denote by $h = \max_{T\in\mathcal{T}_h} h_T$. For any edge $E \in \mathcal{E}_h$ we define ω_E to be the set of all elements that contain the edge E and we let h_E denote the diameter of the edge E.

Denote by $\mathcal{P}_k(T)$ the space of polynomials of degree k in $\mathbb{R}^d$ and define the space of approximations

$$V_h = \left\{v_h \in C(\overline{\Omega}) : \forall T \in \mathcal{T}_h, v_h|_T \in \mathcal{P}_k(T),\ v_h|_{\partial\Omega} = 0\right\}.$$

Consider now $h \leq H$ and some two partitions $\mathcal{T}_h$, $\mathcal{T}_H$ with the associated spaces V_h, V_H. Denote $\mathcal{H} = \{h, H\}$. Note that $h = H$ does not necessarily mean $\mathcal{T}_h = \mathcal{T}_H$. We will seek approximations $(u_h, v_H) \in V_h \times V_H$ to $(u, v) \in V \times V$.

Remark 2.2. *If $\mathcal{T}_h \neq \mathcal{T}_H$, we will consider for simplicity only $k = 1$. Our a-posteriori calculations will be carried out however for any k.*

In the analysis below it will be evident that we need to relate the two partitions $\mathcal{T}_h$, $\mathcal{T}_H$ to one another. We say that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$, if every element of $\mathcal{T}_h$ intersects the interior of exactly one element in $\mathcal{T}_H$. Furthermore, let $r \in \mathbb{N}$ be fixed. We call the partition $\mathcal{T}_h$ an r-uniform refinement of $\mathcal{T}_H$ if for every element $K \in \mathcal{T}_H$, the number of $T \in \mathcal{T}_h : T \subseteq K$ equals r. A general case of unrelated partitions $\mathcal{T}_h$, $\mathcal{T}_H$ could be treated but will not be discussed.

In the discrete problem we need the intergrid operators $\lambda : V_H \to V_h$ and $\lambda' : V_h \to V_H$. Various choices for the pair (λ, λ') can be made e.g., via intergrid operators used in multigrid theory or multilevel schems [18, 11, 41]. In this paper we choose λ to be an interpolation operator; it is easily defined between piecewise linear functions from a coarse grid $\mathcal{T}_H$ to its refinement $\mathcal{T}_h$. We choose for λ' the operator

$$(\lambda'\phi_h, \psi_H) := (\phi_h, \lambda\psi_H),\ \forall\psi_H \in V_H, \tag{10}$$

i.e., λ' is adjoint to λ with respect to the $L^2(\Omega)$ product on V_H. This choice eliminates additional error terms that otherwise would arise in error analysis developed below.

Note that if $\mathcal{T}_h = \mathcal{T}_H$, then $V_H = V_h$, and λ and λ' both trivially reduce to identity. Another important observation follows.

Remark 2.3. *Assume $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$. Then $V_H \subseteq V_h$ i.e. $\lambda\psi_H = \psi_H$ for any $\psi_H \in V_H$. Furthermore $\lambda'\lambda\psi_H = \psi_H$ for any ψ_H; in other words, the* *composition $\lambda'\lambda|_{V_H}$ is the identity operator. Moreover, the bilinear form*

$$\tilde{A}((u_h, v_H), (\phi_h, \psi_H)) := \int_\Omega (\lambda_1 u_h \phi_h + \lambda_2 v_H \psi_H) + \int_\Omega (a\nabla u_h \cdot \nabla \phi_h + b\nabla v_H \cdot \nabla \psi_H) + \int_\Omega (c(u_h - \lambda v_H))\phi_h + c(\lambda'\lambda v_H - \lambda' u_h)\psi_H),$$

is a restriction of $A(\cdot,\cdot)$ to $V_h \times V_H$. It follows that it is continuous on $V_h \times V_H$ and positive definite, i.e.,

$$\tilde{A}((u_h, v_H), (u_h, v_H)) \geq \int_\Omega (\lambda_1 (u_h)^2 + \lambda_2 (v_H)^2) + \int_\Omega (a(\nabla u_h)^2 + b(\nabla v_H)^2).$$

Now we define the discrete problem for (9). We seek the approximations $(u_h, v_H) \in V_h \times V_H$ to the solution $(u, v) \in V \times V$ to (9) satisfying

$$\tilde{A}((u_h, v_H), (\phi_h, \psi_H)) = L((\phi_h, \psi_H)), \ \forall (\phi_h, \psi_H) \in V_h \times V_H. \tag{11}$$

Equivalently, the solution $(u_h, v_H) \in V_h \times V_H$ to (11) satisfies

$$\begin{aligned} \int_\Omega \lambda_1 u_h \phi_h + \int_\Omega a\nabla u_h \cdot \nabla \phi_h + \int_\Omega c(u_h - \lambda v_H)\phi_h &= \int_\Omega f\phi_h, \ \ \forall \phi_h \in V_h \\ \int_\Omega \lambda_2 v_H \psi_H + \int_\Omega b\nabla v_H \cdot \nabla \psi_H + \int_\Omega c(\lambda'\lambda v_H - \lambda' u_h)\psi_H &= \int_\Omega g\psi_H, \ \ \forall \psi_H \in V_H. \end{aligned}$$

The problem (11) is square and finite dimensional. By Remark 2.3 it is easy to see that its solution exists and is unique.

3. A priori error analysis

For the error analysis of (11) we first develop the counterpart of Galerkin orthogonality. Thanks to our definition of λ, λ', it follows smoothly without additional consistency errors.

Let $\phi = \phi_h$ and $\psi = \psi_H$ in (9) and subtract it from (11) to get

$$\begin{aligned} (12) \quad 0 &= A((u, \phi_h), (v, \psi_H)) - \tilde{A}((u_h, \phi_h), (v_H, \psi_H)) \\ &= A((u, \phi_h), (v, \psi_H)) - A((u_h, \phi_h), (v_H, \psi_H)) - \int_\Omega c(v_H - \lambda v_H)\phi_h - \int_\Omega c(u_h - \lambda' u_h)\psi_H \\ &= A((u - u_h, \phi_h), (v - v_H, \psi_H)) - \int_\Omega c(v_H - \lambda v_H)\phi_h + \int_\Omega c(u_h - \lambda' u_h)\psi_H. \end{aligned}$$

Now, if $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$, then by (10) and Remark 2.3 the last two terms vanish and we obtain

$$A((u - u_h, \phi_h), (v - v_H, \psi_H)) = 0, \ \ \forall \phi_h \in V_h, \psi_H \in V_H. \tag{13}$$

This is a basic step in proving convergence of the scheme in the energy norm $\|\cdot\|_e$ and of the subsequent a-posteriori estimates.

Theorem 3.1. *Assume that the solution $(u, v) \in V \times V$ of the problem (9) satisfies $(u, v) \in \left(H^2(\Omega) \cap H^1_0(\Omega)\right)^2$. Assume also that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$, $k = 1$, and let $(u_h, v_H) \in V_h \times V_H$ be the two-level solution of the discrete problem (11). Then there exist constants κ_1, κ_2 independent of $\mathcal{H}$ and of u, v, such that*

$$\|(u - u_h, v - v_H)\|_e \leq \kappa_1 h \|u\|_2 + \kappa_2 H \|v\|_2. \tag{14}$$

Proof. Consider the following calculation, similar to the derivation of Céa's lemma in the scalar case [14, 10]. For an arbitrary $z_h \in V_h, w_H \in V_H$ it follows by (13) that

$$\begin{aligned}(15)\quad & \|(u-u_h, v-v_H)\|_e^2 = A((u-u_h, u-u_h),(v-v_H, v-v_H)) \\ &= A((u-u_h, u-u_h),(v-v_H, v-v_H)) + A((u-u_h, u_h-z_h),(v-v_H, v_H-w_H)) \\ &= A((u-u_h, u-z_h),(v-v_H, v-w_H)).\end{aligned}$$

We bound this last term from above and from below using, respectively, continuity and coercivity of $A(\cdot,\cdot)$ in the energy norm. Dividing both sides of the resulting inequality by $\|(u-u_h, v-v_H)\|_e$ yields the standard estimate. Now, since z_h, w_H are arbitrary, we take the inf to get

$$\|(u-u_h, v-v_H)\|_e \le C \inf_{(z_h,w_H)\in V_h\times V_H} \|(u-z_h, v-w_H)\|_e, \tag{16}$$

where C is the ratio of continuity and ellipticity constants. To get the desired convergence estimates, we select the test functions to be the piecewise linear interpolations $(z_h, w_H) = (I_h u, I_H v)$ of the respective components of the analytical solution and set $z = u - I_h u$ and $w = v - I_H v$. Now we have

$$\begin{aligned}\|(u-u_h, v-v_H)\|_e &\le C\|(z,w)\|_e \\ &= C\left[\int_\Omega(\lambda_1 z^2 + \lambda_2 w^2) + \int_\Omega(a(\nabla z)^2 + b(\nabla w)^2) + \int_\Omega c(z-w)^2\right] \\ &\le C\left[\int_\Omega((\lambda_1+2c)z^2 + (\lambda_2+2c)w^2) + \int_\Omega(a(\nabla z)^2 + b(\nabla w)^2)\right] \\ &\le C\left[\max\{a, \lambda_1+2c\}\|z\|_1^2 + \max\{b, \lambda_2+2c\}\|w\|_1^2\right]\end{aligned}$$

The interpolation theory [14, 10] lets us bound the interpolation error $\xi - I_h\xi$ for a smooth enough ξ. For $k = 1$ we have as follows

$$\|\xi - I_h\xi\|_m \le \tilde{c}h^{2-m}|\xi|_{t,\Omega} \quad \text{for} \quad \xi \in H^t(\Omega), \quad 0 \le m \le 2.$$

Applying this bound to z and w we get

$$\begin{aligned}\|(u-u_h, v-v_H)\|_e^2 &\le \max\{a, \lambda_1+2c\}\|u-I_h u\|_1^2 + \max\{b, \lambda_2+2c\}\|v-I_H v\|_1^2 \\ &\le C\max\{a, \lambda_1+2c\}\tilde{c}^2 h^2|u|_2^2 + C\max\{b, \lambda_2+2c)\}\tilde{c}^2 H^2|v|_2^2.\end{aligned}$$

Taking square root of both sides completes the proof. □

This a-priori result shows the structure of the error. First, if $\mathcal{T}_h = \mathcal{T}_H$, then the error converges with the rate $O(h)$, and an easy extension can be formulated for $k > 1$. If $\mathcal{T}_h \ne \mathcal{T}_H$, then the error in (14) is dominated asymptotically by the $O(H)$ terms, at least for $\mathcal{P} = \mathbf{1}^5$. For general $\mathcal{P}$ the individual contributions to the error depend on $\mathcal{P}$. The magnitude of each of the contributions depends on $\mathcal{P}$ and on the variability of u, v. Thus H and h could be adapted to take advantage of this potential disparity.

For example, if $c = O(1)$ is moderate, and $a >> 1$ very large but $b << 1$ very small, one can find $\mathcal{T}_H$ for the component v so that the total error does not increase substantially. Note that with $H > h$ the total number of unknowns decreases. We would proceed similarly if $b = O(1)$ but $|v|_2$ is very small. Conversely, if the error on some coarse grid used for both components is too large for our needs, then one could refine only the grid for the strongly varying component, for example u, for which $|u|_2$ is large. See Section 5 for relevant examples.

To guide the adaptive choice of h, H, i.e., of $\mathcal{T}_h, \mathcal{T}_H$, we need the a-posteriori error analysis provided in the next Section.

4. A posteriori estimate

In this section we define residual-type error estimators for the system (9) and prove the global upper and some lower bounds. The lower bounds, due to the presence of coupling terms, work only if $\mathcal{H}$ are small enough. We develop estimators for both components of the system and an additional estimator for the error in one component only.

We follow the standard technique for residual estimators [35] which, for scalar diffusion problems such as (1), involves the following steps. First, the energy norm of the error $e = u - u_h$ is rewritten using Galerkin orthogonality with an auxiliary function $z_h \in V_h$, and localized by integrating by parts over each element via $\int_\Omega a\nabla e\nabla(e - z_h) = \sum_T \int_T a\nabla e\nabla(e - z_h)$. Thereby the error terms per element and per element boundary are identified; this follows from integration by parts elementwise e.g. $\int_T a\nabla(u - u_h)\nabla(e - z_h) = \int_T -a\Delta(u - u_h)(e - z_h) + \int_{\partial T} a\partial_n(u - u_h)(e - z_h)$. Note that for $k = 1$ the term $\nabla u_h|_T$ vanishes. For the sake of generality and $k \neq 1$ we keep that term.

The first term is then rewritten using (1), and it reveals the residual $f - \kappa u_h$. The second term, when summed over edges of neighboring elements, gives rise to $\int_E [a\partial_n(u - u_h)](e - z_h)$, where $[w]_E$ is the jump of w across the edge $E \subset \partial T$, and further, by continuity of u, e, z_h across the edges, to $-\int_{\partial e}[a\partial_n(u_h)](e - z_h)$. The jump terms dominate in typical circumstances [13] for conforming discretizations. Next, the local interpolation properties of finite element functions are exploited to estimate $e - z_h$ for a particular choice of z_h. This leads to the upper bounds while the lower bounds are derived using bubble functions. The constants in these estimates depend on $\mathcal{P}$, and on the geometry of $\Omega, \mathcal{T}_h$.

The estimators derived this way are explicit i.e. one can compute them directly from the numerical solution without the need to solve any additional problems. They are also *reliable*, i.e., they bound the true error from above. Unfortunately, they are not very *efficient* i.e. the gap between the estimator and true error can be substantial even for problems with $\mathcal{P} = 1$ and simple domains [12]. Other families of estimators [35, 12] are much more efficient and even asymptotically exact but can be cumbersome in implementation and computationally expensive. We will not study these but mention the work of [4, 2] on systems.

Aside of efficiency, the additional difficulty with residual estimators is the dependence of the efficiency constants on the parameters of the problem. This is directly related to dependence of the ellipticity, continuity, and equivalence constants on the parameters. This issue was brought up in [9, 37], and a remedy involving a particular scaling was proposed; we follow these ideas below.

The proof of the upper bound is tedious but not very complicated as it extends the standard techniques to a system, and involves handling the coupling terms. The lower bound is more delicate to obtain. We develop a global lower bound which is valid for fine enough $\mathcal{H}$, and a local lower bound. We also prove a bound for the error in one component only.

4.1. Residual calculations. Let $Q_h : V \mapsto V_h, Q_H : V \mapsto V_H$ be some quasi-interpolators to be defined later. We first rewrite $\mathcal{E} := \|(e_u, e_v)\|_e$ using (13)

$$\mathcal{E}^2 = A((e_u, e_v), (e_u, e_v)) = A((e_u, e_v), (e_u - Q_h e_u, e_v - Q_H e_v)).$$

Now we follow the standard procedure described above. We rewrite the last term $A((e_u, e_v), (\phi, \psi))$, with $\phi = e_u - Q_h e_u$, $\psi = e_v - Q_H e_v$, replacing $\int_\Omega$ by $\sum_T \int_T$, taking advantage of (2)–(3), and of the continuity of u, v, ψ, ϕ across each edge and integrating by parts on each element T. As before, our calculations work for a general $k \geq 1$. We obtain

$$
\begin{aligned}
(17)\quad & A((e_u, e_v), (\phi, \psi)) \\
&= \sum_T \left\{ \int_T (\lambda_1 e_u \phi + \lambda_2 e_v \psi + c(e_u - e_v)(\phi - \psi)) \right. \\
&\quad \left. - \int_T (a\Delta e_u \phi + b\Delta e_v \psi) + \int_{\partial T} (a\partial_n e_u \phi + b\partial_n e_v \psi) \right\} \\
&= \sum_T \int_T (\lambda_1 u\phi + \lambda_2 v\psi + c(u - v)(\phi - \psi) - a\Delta u\phi - b\Delta v\psi \\
&\quad - \sum_T \int_T (\lambda_1 u_h \phi + \lambda_2 v_H \psi + c(u_h - v_H)(\phi - \psi) - a\Delta u_h \phi - b\Delta v_H \psi) \\
&\quad + \sum_E \int_E ([a\partial_n(u - u_h)]\phi + [b\partial_n(v - v_H)]\psi) \\
&= \sum_T \int_T (f\phi + g\psi) + \sum_T \int_T (-\lambda_1 u_h \phi - \lambda_2 v_H \psi - c(u_h - v_H)(\phi - \psi) \\
&\quad + \sum_E \int_E ([a\partial_n u_h]\phi + [b\partial_n v_H]\psi).
\end{aligned}
$$

Combining the terms we get

$$
\begin{aligned}
(18)\quad \mathcal{E}^2 = & \sum_{T\in\mathcal{T}_h} (R^*_{T,u}, e_u - Q_h e_u)_T + \sum_{E\in\mathcal{E}_h} (R_{E,u}, e_u - Q_h e_u)_E \\
& + \sum_{K\in\mathcal{T}_H} (R^*_{K,v}, e_v - Q_H e_v)_K + \sum_{F\in\mathcal{E}_H} (R_{F,v}, e_v - Q_H e_v)_F,
\end{aligned}
$$

where we have used the element and edge residual terms defined as follows

$$
\begin{aligned}
R^*_{T,u} &:= f - \lambda_1 u_h + a\Delta u_h - c(u_h - v_H) = f - f_h + \overbrace{f_h - \lambda_1 u_h + a\Delta u_h - c(u_h - v_H)}^{R_{T,u}}, \\
R^*_{K,v} &:= g - \lambda_2 v_H + b\Delta v_H - c(v_H - u_h) = g - g_H + \overbrace{g_H - \lambda_2 v_H + b\Delta v_H - c(v_H - u_h)}^{R_{K,v}} \\
R_{E,u} &:= [a\partial_n u_h]_E, \\
R_{F,v} &:= [b\partial_n v_H]_F,
\end{aligned}
$$

and where f_h, g_H are the L_2-projections of f, g onto V_h, V_H respectively.

Now we estimate the terms in (18) by Cauchy-Schwarz inequality to obtain

$$
\begin{aligned}
(19)\quad \mathcal{E}^2 \leq & \sum_{T\in\mathcal{T}_h} \|R^*_{T,u}\|_{0,T} \|e_u - Q_h e_u\|_{0,T} + \sum_{E\subset\mathcal{E}_h} \|R_{E,u}\|_{0,E} \|e_u - Q_h e_u\|_{0,E} \\
& + \sum_{K\in\mathcal{T}_H} \|R^*_{K,v}\|_{0,K} \|e_v - Q_H e_v\|_{0,K} + \sum_{F\subset\mathcal{E}_H} \|R_{F,v}\|_{0,F} \|e_v - Q_H e_v\|_{0,F}.
\end{aligned}
$$

Consider $e_u - Q_h e_u$. The idea is to bound the terms $e_u - Q_h e_u$ from above by the terms involving the energy norm of e_u, without requiring more smoothness than that $e_u \in V$; then the estimate for $\|(e_u, e_v)\|_e$ will follow.

Such estimates are available for various quasi-interpolators [15, 35]. We use the definition and properties of Q_h as modified by Verfürth [35] and quote two basic relevant interpolation estimates which work in any $T \in \mathcal{T}_h$ and any $E \in \mathcal{E}_h$.

The first result ([37], Lemma 3.1) states that for any $w \in H^k(\tilde{\omega}_T), 0 \le k \le 1$

$$\|\nabla^l(w - Q_h w)\|_{0,T} \le C h_T^{k-l} \|\nabla^k w\|_{0,\tilde{\omega}_T} \; 0 \le l \le k \le 1, \tag{20}$$

where the constant C is independent of h, w.

Next, we quote ([36], Lemma 3.1) to estimate the edge terms. Let $E \in \mathcal{E}_h$ and let T be an element in $\mathcal{T}_h$ which has E as an edge. The following trace inequality holds for all $w \in H^1(T)$

$$\|w\|_{0,E} \le c_3 \left(h_T^{-1/2} \|w\|_{0,T} + \|w\|_{0,T}^{1/2} \|\nabla w\|_{0,T}^{1/2} \right), \tag{21}$$

where c_3 is a constant independent of w, h_T.

4.2. Interpolation and scaling techniques. To derive the estimates in the energy norm we find that they involve various equivalence constants dependent on $\mathcal{P}$ between $\|\cdot\|_{V\times V}, \|\cdot\|_s, \|\cdot\|_e$. To prevent the estimates from blowing up when the parameters of the problem change, we define certain scaling factors following [37, 36, 26].

Define for all $T \in \mathcal{T}_h$ and all $K \in \mathcal{T}_H$

$$\theta_{u,T} := \min\{h_T a^{-1/2}, \lambda_1^{-1/2}\}, \tag{22}$$

$$\theta_{v,K} := \min\{H_K b^{-1/2}, \lambda_2^{-1/2}\}, \tag{23}$$

$$\gamma_{u,E} := a^{-1/4} \theta_{u,E}^{1/2}, \tag{24}$$

$$\gamma_{v,F} := b^{-1/4} \theta_{v,F}^{1/2}. \tag{25}$$

Let $T \in \mathcal{T}_h$ and $K \in \mathcal{T}_H$. Clearly $e_u \in H^1(\tilde{\omega}_T)$ and $e_v \in H^1(\tilde{\omega}_K)$. By (20)

$$\|e_u - Q_h e_u\|_{0,T} \le C \lambda_1^{-1/2} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2}, \tag{26}$$

$$\|e_u - Q_h e_u\|_{0,T} \le C h_T a^{-1/2} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2}. \tag{27}$$

We combine these and (22) to get

$$\|e_u - Q_h e_u\|_{0,T} \le C \min\left\{ \lambda_1^{-1/2}, h_T a^{-1/2} \right\} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2}. \tag{28}$$

Similar calculations can be done for e_v, and it follows that we have

$$\|e_u - Q_h e_u\|_{0,T} \le c_1 \theta_{u,T} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2} \tag{29}$$

$$\|e_v - Q_H e_v\|_{0,K} \le c_2 \theta_{v,K} \left\{ \int_{\tilde{\omega}_K} b(\nabla e_v)^2 + \lambda_2 e_v^2 \right\}^{1/2}, \tag{30}$$

where c_1, c_2 are independent of $\mathcal{P}, \mathcal{H}$.

On the edges the calculations are a bit longer. Apply (21) to $w = e_u - Q_h e_u$

$$\|e_u - Q_h e_u\|_{0,E} \le c_3 \left(h_T^{-1/2} \|e_u - Q_h e_u\|_{0,T} + \|e_u - Q_h e_u\|_{0,T}^{1/2} \|\nabla(e_u - Q_h e_u)\|_{0,T}^{1/2} \right).$$

Next, apply (20) to get

$$\|\nabla(e_u - Q_h e_u)\|_{0,T}^{1/2} \le a^{-1/4} \|a^{1/2} \nabla e_u\|_{0,\tilde{\omega}_T}^{1/2} \le a^{-1/4} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/4}.$$

Using (29) and noticing $h_T^{-1/2}\theta_{u,T}^{1/2} + a^{-1/4} \leq 2a^{-1/4}$ we get

$$\|e_u - Q_h e_u\|_{0,E} \leq c_4 \left(h_T^{-1/2}\theta_{u,T} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2} \right.$$
$$\left. + \theta_{u,T}^{1/2} a^{-1/4} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2} \right),$$

and conclude

$$\|e_u - Q_h e_u\|_{0,E} \leq c_4 \gamma_{u,E} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2}. \tag{31}$$

Similar estimates follow for the edges $F \in \mathcal{E}_H$

$$\|e_v - Q_H e_v\|_{0,F} \leq c_5 \gamma_{v,F} \left\{ \int_{\tilde{\omega}_K} b(\nabla e_v)^2 + \lambda_2 e_v^2 \right\}^{1/2}, \tag{32}$$

where c_4, c_5 are independent of $\mathcal{P}, \mathcal{H}$.

We apply the above estimates to (19) and obtain the Lemma.

Lemma 4.1. *The following estimates hold*

$$\|e_u - Q_h e_u\|_{0,T} \leq c_1 \theta_{u,T} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2}, \tag{33}$$

$$\|e_v - Q_H e_v\|_{0,K} \leq c_2 \theta_{v,K} \left\{ \int_{\tilde{\omega}_K} b(\nabla e_v)^2 + \lambda_2 e_v^2 \right\}^{1/2}, \tag{34}$$

$$\|e_u - Q_h e_u\|_{0,E} \leq c_4 \gamma_{u,E} \left\{ \int_{\tilde{\omega}_T} a(\nabla e_u)^2 + \lambda_1 e_u^2 \right\}^{1/2}, \tag{35}$$

$$\|e_v - Q_H e_v\|_{0,F} \leq c_5 \gamma_{v,F} \left\{ \int_{\tilde{\omega}_K} b(\nabla e_v)^2 + \lambda_2 e_v^2 \right\}^{1/2}. \tag{36}$$

4.3. Upper bound. Now we define the local component error estimators

$$\eta_{u,T} := \theta_{u,T}^2 \|R_{T,u}\|_{0,T}^2 + \frac{1}{2} \sum_{E \subset \partial T} \gamma_{u,T}^2 \|R_{E,u}\|_{0,E}^2, \tag{37}$$

$$\eta_{v,K} := \theta_{v,K}^2 \|R_{K,v}\|_{0,K}^2 + \frac{1}{2} \sum_{F \subset \partial K} \gamma_{v,K}^2 \|R_{F,v}\|_{0,F}^2, \tag{38}$$

and the global error estimator for the error in both variables (u, v)

$$\eta := \left\{ \sum_{T \in \mathcal{T}_h} \eta_{u,T} + \sum_{K \in \mathcal{T}_H} \eta_{v,K} \right\}^{1/2}. \tag{39}$$

We recognize the two parts of each local component error estimator (37), (38) as the terms which arise on the right hand side of (19). They are multiplied by the factors which have been estimated in (33)–(36). Taking all these into account, along with an additional application of the discrete Cauchy-Schwarz inequality, yields finally the main result on the upper bound.

Theorem 4.2. *Let the assumptions of Theorem 3.1 hold and in particular, let (u, v) be the unique solution of* (9) *and (u_h, v_H) be the unique solution of* (11).

Then the following upper bound holds

$$\mathcal{E} \leq C^*\eta + \left\{ \sum_{T\in\mathcal{T}_h} \theta_{u,T}^2 \|f - f_h\|_{0,T}^2 + \sum_{K\in\mathcal{T}_H} \theta_{v,K}^2 \|g - g_H\|_{0,T}^2 \right\}^{1/2},$$

where C^* *does not depend on* $\mathcal{H}$, $\mathcal{P}$, *or* u, v.

4.4. Lower bound. In this section we want to establish the global lower bound i.e. $C_*\eta \leq \mathcal{E}$, and some appropriate local counterpart, with some constant C_* independent of $\mathcal{H}, \mathcal{P}, u, v$. Due to the coupling terms in our system this is not possible without additional assumptions.

To establish the result, we proceed using the standard approach of bubble functions [35]. Let $T \in \mathcal{T}_h$ be fixed and denote by $\mathcal{N}_T$ the set of its vertices. For $x \in \mathcal{N}_T$ denote by λ_x the nodal basis function from V_h associated with the point x. Define the element bubble $\Psi_T = \Gamma_T \prod_{x\in\mathcal{N}_T} \lambda_x$ where the constant Γ_T is chosen so that Ψ_T equals 1 at the barycenter of T. Now let $E \in \mathcal{E}_h$ and denote by $\mathcal{N}_E$ the set of all vertices of the edge E and define the edge bubble function $\Psi_E = \Gamma_E \prod_{x\in\mathcal{N}_e} \lambda_x$ where the constant Γ_E is chosen so that Ψ_E equals 1 at the barycenter of E.

The element and edge bubbles have the following properties shown in ([37], Lemma 3.3), with generic constants depending only on the shape of the elements; these constants are different from those in Section 4.3. Let $T \in \mathcal{T}_h, E \in \mathcal{E}_h$ and let $w \in \mathcal{P}_1(T)$, $\sigma \in \mathcal{P}_1(E)$ be arbitrary. We have

$$\|\Psi_T\|_\infty \leq 1, \tag{40}$$

$$c_1\|w\|_{0,T}^2 \leq (w, \Psi_T w)_T, \tag{41}$$

$$\|\nabla\Psi_T w\|_{0,T} \leq c_2 h_T^{-1}\|w\|_{0,T}, \tag{42}$$

$$c_3\|\sigma\|_{0,E}^2 \leq (\sigma, \Psi_E\sigma)_E \tag{43}$$

$$\|\nabla\Psi_E\sigma\|_{0,T} \leq c_4 h_E^{-1/2}\|\sigma\|_{0,E}, \tag{44}$$

$$\|\Psi_E\sigma\|_{0,\omega_E} \leq c_5 h_E^{1/2}\|\sigma\|_{0,E}. \tag{45}$$

Now we fix an element T, define $\rho_T := \Psi_T R_{T,u}$, and estimate $R_{T,u}$ from above in the goal to isolate the coupling terms and to get the bounds in terms of the energy norm of the error.

$$\begin{aligned} \|R_{T,u}\|_{0,T}^2 &= \|(f_h + \nabla(a\nabla u_h) - (\lambda_1 + c)u_h + cv_H)\|_{0,T}^2 \\ &\overset{(42)}{\leq} c_1^{-2}\int_T (f_h + a\Delta u_h - (\lambda_1 + c)u_h + cv_H)\rho_T \\ &= c_1^{-2}\left[\int_T (f_h + a\Delta u_h - (\lambda_1 + c)u_h + cv_H)\rho_T + \int_T f\rho_T - \int_T f\rho_T\right]. \end{aligned} \tag{46}$$

Next we integrate by parts over T, use the strong form of (2), i.e., $f = \lambda_1 u - a\Delta u + c(u - v)$, and the fact that $\rho_T|_{\partial T} \equiv 0$ to see from (46) that

$$\begin{aligned} \|R_{T,u}\|_{0,T}^2 = c_1^{-2}\Bigg[\int_T &(a\nabla(u - u_h)\cdot\nabla\rho_T \\ &+ [(\lambda_1 + c)(u - u_h) - c(v - v_H)]\,\rho_T + \int_T (f_h - f)\rho_T\Bigg] \end{aligned} \tag{47}$$

Now we estimate both integrals using Cauchy-Schwarz inequality. For the second integral in (47) we have, using (40) and Cauchy-Schwarz again

$$\int_T (f_h - f)\rho_T \leq \|f_h - f\|_{0,T}\|\rho_T\|_{0,T} \leq \|f_h - f\|_{0,T}\|R_{T,u}\|_{0,T}. \tag{48}$$

The bounds for the first integral in (47) involve $\int_T (c(u-u_h) - c(v-v_H))\rho_T = \int_T \frac{c}{2}(e_u - e_v)2\rho_T$ leading to, by a multiple application of Cauchy-Schwarz to the upper bound for that term, to the expression

$$\left\{\int_T (a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2} \left\{\int_T a(\nabla \rho_T)^2 + (\lambda_1 + 2c)\rho_T^2\right\}^{1/2}.$$

To estimate the second term in this expression from above by a multiple of $\|R_{T,u}\|_{0,T}$, we first observe that by (22)

$$h_T^{-2} a + \lambda_1 \le 2\max\{a h_T^{-2}, \lambda_1\} = \theta_{u,T}^{-2}. \tag{49}$$

Next, we estimate

$$\begin{aligned}\left\{\int_T a(\nabla \rho_T)^2 + (\lambda_1 + 2c)\rho_T^2\right\}^{1/2} &\le a^{1/2}\|\nabla \rho_T\|_{0,T} + (\lambda_1+2c)^{1/2}\|\rho_T\|_{0,T} \\ &\overset{(42),(40)}{\le} a^{1/2} c_2 h_T^{-1}\|R_{T,u}\|_{0,T} + (\lambda_1 + 2c)^{1/2}\|R_{T,u}\|_{0,T} \\ &\le \max\{c_2, 1\}(2\max\{a^{1/2}h_T^{-1}, \lambda_1^{1/2}\} + (2c)^{1/2})\|R_{T,u}\|_{0,T} \\ &\overset{(49)}{\le} \underbrace{2\max\{c_2,1\}}_{\bar{c}_2}(\theta_{u,T}^{-1} + c^{1/2})\|R_{T,u}\|_{0,T}.\end{aligned}$$

Now we combine the estimates following (47) to get, upon dividing by $\|R_{T,u}\|_{0,T}$

$$\|R_{T,u}\|_{0,T} \le c_1^{-2}\left[\bar{c}_2(\theta_{u,T}^{-1} + c^{1/2})\left\{\int_T a(\nabla e_u)^2 + \lambda_1(e_u)^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2} + \|f_h - f\|_{0,T}\right]. \tag{50}$$

Multiplying both sides by $\bar{\theta}_{u,T} := (\theta_{u,T}^{-1} + c^{1/2})^{-1}$ we finally obtain

$$\bar{\theta}_{u,T}\|R_{T,u}\|_{0,T} \le c_1^{-2}\left[\bar{c}_2\left\{\int_T a(\nabla e_u)^2 + \lambda_1(e_u)^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2} + \bar{\theta}_{u,T}\|f_h - f\|_{0,T}\right].$$

Next we estimate the edge residuals. Consider an arbitrary edge $E \in \mathcal{E}_h$ and denote by T_1, T_2 the two elements that it separates. Let $\rho_E := \beta\Psi_E[a\partial_n u_h]_E = \beta\Psi_E R_{E,u}$ with some scaling factor $0 < \beta \le 1$ to be determined later as in [37]. We will estimate $\|R_{E,u}\|_{0,E}$ from above using steps similar to those above: adding and subtracting terms and integrating by parts over $T_1 \cup T_2$ and taking advantage of (2) and of the bubbles ω_E vanishing conveniently at all edges of $T_1 \cup T_2$ other than

E, and estimating by Cauchy-Schwarz inequality

$$
\begin{aligned}
(51)\quad \|R_{E,u}\|_{0,E}^2 &= \|[a\partial_n u_h]_E\|_{0,E}^2 \le c_3^{-2}\int_E [a\partial_n u_h]_E \rho_E \\
&= c_3^{-2}\left[\int_{T_1\cup T_2} a\Delta u_h\rho_E + \int_{T_1\cup T_2} a\nabla u_h\nabla\rho_E + \int_{T_1\cup T_2} f\rho_E - \int_{T_1\cup T_2} f\rho_E\right] \\
&\overset{(2)}{=} c_3^{-2}\left[\int_{T_1\cup T_2} (f + a\Delta u_h)\rho_E - \int_{T_1\cup T_2} a\nabla e_u\nabla\rho_E + (\lambda_1 u + c(u-v))\rho_E\right] \\
&= c_3^{-2}\left[\int_{T_1\cup T_2} (f_h + a\Delta u_h - \lambda_1 u_h + c(u_h - v_H))\,\rho_E + \int_{T_1\cup T_2} (f - f_h)\rho_E \right. \\
&\qquad \left. - \int_{T_1\cup T_2} a\nabla e_u\nabla\rho_E + (\lambda_1 e_u + c(e_u - e_v)\rho_E\right] \\
&\le c_3^{-2}\sum_{i=1}^{2}\Big[\ \|R_{T,u}\|_{0,T_i}\|\rho_E\|_{0,T_i} + \|(f-f_h)\|_{0,T_i}\|\rho_E\|_{0,T_i} \\
&+\left\{\int_{T_i} a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2}\left\{\int_{T_i} a(\nabla\rho_E)^2 + (\lambda_1 + 2c)\rho_E\right\}^{1/2}\Big].
\end{aligned}
$$

In the last inequality we need to bound $\|\rho_E\|_{0,Ti}$ and $\left\{\int_{T_i} a(\nabla\rho_E)^2 + (\lambda_1 + 2c)\rho_E\right\}^{1/2}$ in terms of the edge residuals. We have

$$\|\rho_E\|_{0,Ti} = \|\beta\Psi_E R_{E,u}\|_{0,T_i} \le c_5\beta h_E^{1/2}\|R_{E,u}\|_{0,E}.$$

Also, by (44) and (45)

$$\left\{\int_{T_i} a(\nabla\rho_E)^2 + (\lambda_1 + 2c)\rho_E\right\}^{1/2} \le 2\max\{c_4, c_5\}\beta(h_E^{-1/2}\theta_{u,E}^{-1} + h_E^{1/2}c^{1/2})\|R_{E,u}\|_{0,E}.$$

To remove the dependence of the constants on the right hand side on h_E, we define $\beta := \min\{1, h_E^{-1/2}a^{1/4}\lambda_1^{-1/4}\}$. Now we see $a^{-1/2}h_E\beta^2 = \theta_{u,E}^2$ and further

$$
\begin{aligned}
\beta h_E^{1/2}\theta_{u,E}^{-1} &= \gamma_{u,E}^{-1}, \\
\beta h_E^{1/2} &= \gamma_{u,E}a^{1/2}.
\end{aligned}
$$

We obtain therefore

$$
\begin{aligned}
\|\rho_E\|_{0,Ti} &\le c_5\gamma_{u,E}a^{1/2}\|R_{E,u}\|_{0,E}, \\
\left\{\int_{T_i} a(\nabla\rho_E)^2 + (\lambda_1 + 2c)\rho_E\right\}^{1/2} &\le 2\max\{c_4, c_5\}(\gamma_{u,E}^{-1} + \gamma_{u,E}a^{1/2}c^{1/2})\|R_{E,u}\|_{0,E}.
\end{aligned}
$$

Using the above estimates we get from (51), upon dividing by $\|R_{E,u}\|_{0,E}$

$$
\begin{aligned}
(52)\quad \|R_{E,u}\|_{0,E} &\le c_3^{-2}\sum_{i=1}^{2}\Big[c_5\gamma_{u,E}a^{1/2}\left(\|R_{T,u}\|_{0,T_i} + \|f - f_h\|_{0,T_i}\right) \\
&+2\max\{c_4, c_5\}\left(\gamma_{u,E}^{-1} + \gamma_{u,E}a^{1/2}c^{1/2}\right)\left\{\int_{T_i} a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2}\Big].
\end{aligned}
$$

Substituting (50) in the bound above and with (22)-(23) we arrive at

$$\|R_{E,u}\|_{0,E} \le C \sum_{i=1}^{2} \Big[\gamma_{u,E} a^{1/2} \|f - f_h\|_{0,T_i} + \left(\gamma_{u,E}^{-1} + \gamma_{u,E} a^{1/2} c^{1/2}\right) \left\{ \int_{T_i} a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2 \right\}^{1/2} \Big].$$

Equivalently, with $\bar{\gamma}_{u,E} := \left(\gamma_{u,E}^{-1} + \gamma_{u,E} a^{1/2} c^{1/2}\right)^{-1}$ we have

$$\bar{\gamma}_{u,E} \|R_{E,u}\|_{0,E} \le C \sum_{i=1}^{2} \Big[\gamma_{u,E} a^{1/2} \bar{\gamma}_{u,E} \|f - f_h\|_{0,T_i} + \left\{ \int_{T_i} a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2 \right\}^{1/2} \Big]. \tag{53}$$

One can now prove similar lower bounds for the second component of the system in terms of $\|R_{F,v}\|_{0,F}$ and $\|R_{K,v}\|_{0,K_i}$, and b, λ_2 instead of a, λ_1, respectively. Upon adding the u and v components and by combining $\int_{T_i} a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2$ with $\int_{T_i} b(\nabla e_v)^2 + \lambda_2 e_v^2 + \frac{c}{2}(e_u - e_v)^2$, we recover on the right hand-sides of (50) and (53) the error $\mathcal{E}$. On the left hand side we combine the element and edge residuals corresponding to u and v. This seems superficially like a straightforward procedure leading to the bounds of the type proven in [37].

However, due to the presence of the coupling terms, the scaling in the residuals such as in (50), (53) involves the factors $\bar{\theta}_{u,T}$ and $\bar{\gamma}_{u,E}$ instead of $\theta_{u,T}$ and $\gamma_{u,E}$, respectively. Since these scaling constants are dependent on additional parameters as well as on the grid discretization, we cannot obtain the "usual" lower bounds without additional mild assumptions. The main idea to get the lower bound which is robust in $\mathcal{P}$ and $\mathcal{H}$ is then to find a lower bound for $\bar{\theta}_{u,T}$ and $\bar{\gamma}_{u,E}$ in terms of $\theta_{u,T}$ and $\gamma_{u,E}$. These can be established in various ways, for example, assuming h, H are small enough.

Theorem 4.3. *Let the assumptions of Theorem 3.1 hold. Assume also that*

$$h \le \sqrt{a} \min\{\lambda_1^{-1/2}, c^{-1/2}\}, \tag{54}$$

$$H \le \sqrt{b} \min\{\lambda_2^{-1/2}, c^{-1/2}\}. \tag{55}$$

Then there is a constant C_* such that

$$C_* \eta \le \|(u - u_h, v - v_H)\|_e + \left\{ \sum_{T \in \mathcal{T}_h} \theta_{u,T}^2 \|f - f_h\|_{0,T}^2 + \sum_{K \in \mathcal{T}_H} \theta_{v,K}^2 \|g - g_H\|_{0,T}^2 \right\}^{1/2}.$$

Proof. We note that by (54) we have from (22)

$$\theta_{u,T} \sqrt{c} = \frac{h_T}{\sqrt{a}} \sqrt{c} \le 1. \tag{56}$$

Next, from (24) we see that by (56) we have

$$\gamma_{u,E} = a^{-1/4} \sqrt{\theta_{u,E}} = \frac{\sqrt{h_E}}{\sqrt{a}}. \tag{57}$$

Thus

$$\bar{\theta}_{u,T} := (\theta_{u,T}^{-1} + c^{1/2})^{-1} = \frac{\theta_{u,T}}{1 + \theta_{u,T}\sqrt{c}} \ge \frac{\theta_{u,T}}{2}, \tag{58}$$

if only $\theta_{u,T}\sqrt{c} \leq 1$ which follows from (56). Similarly, we obtain that

$$\bar{\gamma}_{u,E} = (\gamma_{u,E}^{-1} + \gamma_{u,E} a^{1/2} c^{1/2})^{-1} = \frac{\gamma_{u,E}}{1 + \gamma_{u,E}^2 \sqrt{ac}} \geq \frac{\gamma_{u,E}}{2}, \tag{59}$$

as long as $\gamma_{u,E}^2 \sqrt{ac} \leq 1$ which in turn is guaranteed by (57).

Analogous estimates hold for the v component as a consequence of (55)

$$\theta_{v,K} \geq \bar{\theta}_{v,K} \geq \frac{\theta_{v,K}}{2}, \tag{60}$$

$$\gamma_{v,F} \geq \bar{\gamma}_{v,F} \geq \frac{\gamma_{v,F}}{2}. \tag{61}$$

The rest of the proof is straightforward. We collect (50), (53), apply (58) and (59) to see that

$$\theta_{u,T}\|R_{T,u}\|_{0,T} \leq C\left[\left\{\int_T (a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2} + \theta_{u,T}\|f_h - f\|_{0,T}\right], \tag{62}$$

$$\gamma_{u,E}\|R_{E,u}\|_{0,E} \leq C\sum_{i=1}^{2}\left[\theta_{u,E}\|f - f_h\|_{0,T_i} + 2\left\{\int_{T_i} a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2}\right]. \tag{63}$$

We repeat the same steps to estimate $\|R_{K,v}\|_{0,T}, \|R_{E,v}\|_{0,E}$ and obtain

$$\theta_{v,K}\|R_{K,v}\|_{0,K} \leq C\left[\left\{\int_K (b(\nabla e_v)^2 + \lambda_2 e_v^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2} + \theta_{v,K}\|g - g_H\|_{0,K}\right], \tag{64}$$

$$\gamma_{v,F}\|R_{F,v}\|_{0,F} \leq C\sum_{i=1}^{2}\left[\theta_{v,F}\|g - g_H\|_{0,K_i} + 2\left\{\int_{K_i} b(\nabla e_v)^2 + \lambda_2 e_v^2 + \frac{c}{2}(e_u - e_v)^2\right\}^{1/2}\right]. \tag{65}$$

Adding these equations and summing over all elements T and all elements K completes the proof of the global lower bound. The constant $C_* := \max\{\tilde{c}_2, \frac{c_5\tilde{c}_2 + 2\bar{c}}{c_3^2}\}$, where $\tilde{c}_2 := \frac{\max\{c_2, 1\}}{c_1^2}$ and $\bar{c} := \max\{c_4, c_5\}$ is independent of $\mathcal{P}$ and $\mathcal{H}$. □

The bound in Theorem 4.3 is a global lower bound. We would like also to prove some local lower bounds in analogy to those obtained for scalar equations. Define $\omega_T := \cup\{\omega_E : E \subset \partial T\}$. The local lower bound follows by adding the terms similar to those in (62)-(63) over all edges E of the element T which in turn requires adding the contributions from ω_T. The lower bound involves on the right hand side the energy norm restricted to the neighborhood ω_T of T.

On multilevel grids, in order to obtain a local lower bound between the error and estimator over an element T, we must be able to combine on the left hand side the contributions from all edges of T, and over all edges of K. On the right hand-side a recombination in terms of energy norm is only possible if the summation is done over all elements $T : T \subset K$ and over all corresponding edges E, F. However, no result for a lower bound local to T is available.

Corollary 4.4. *Let $\mathcal{T}_H \ni K = \bigcup_{i=1}^n T_i$ where $T_i \in \mathcal{T}_h$ and assume* (54), (55) *hold. Then the following local lower bound holds*

$$\left\{\eta_{v,K} + \frac{1}{n}\sum_{i=1}^n \eta_{u,T_i}\right\}^{1/2} \leq \|(u-u_h, v-v_H)\|_{\omega_K} + \left\{\frac{1}{n}\sum_{i=1}^n \sum_{T_i'\in\omega_{T_i}} \theta^2_{u,T_i'}\|f-f_h\|^2_{0,T_i'} + \sum_{K'\in\omega_K}\theta^2_{v,K'}\|g-g_H\|^2_{0,K'}\right\}^{1/2}. \tag{66}$$

Proof. To shorten the exposition let $A_p := a(\nabla e_u)^2 + \lambda_1 e_u^2 + \frac{c}{2}(e_u - e_v)^2$ and $B_p := b(\nabla e_v)^2 + \lambda_2 e_v^2 + \frac{c}{2}(e_u-e_v)^2$.

The equation (63) can be written as:

$$\gamma_{u,E}\|R_{E,u}\|_{0,E} \quad\leq\quad C\left[\left\{\int_{\omega_E} A_p\right\}^{1/2} + \sum_{T'\in\omega_E}\theta_{u,E}\|f-f_h\|_{0,T'}\right]. \tag{67}$$

Adding the square of (62) and (67), and using (37) we have

$$\eta_{u,T} \leq C\left[\int_T A_p + \theta^2_{u,T}\|f_h - f\|^2_{0,T} + \frac{1}{2}\sum_{E\subset\partial T}\left(\int_{\omega_E} A_p + \sum_{T'\in\omega_E}\theta^2_{u,E}\|f-f_h\|^2_{0,T'}\right)\right].$$

Note that $\sum_{E\subset\partial T}\int_{\omega_E} w \leq s\int_{\omega_T} w$ for any positive-valued w, where s is the number of sides of T. Also $\sum_{E\subset\partial T}\sum_{T'\in\omega_E} w = \sum_{T'\in\omega_T} w$ for any w. Thus

$$\eta_{u,T} \leq C\left[\int_{\omega_T} A_p + \sum_{T'\in\omega_T}\theta^2_{u,T'}\|f_h - f\|^2_{0,T'}\right]. \tag{68}$$

Similarly,

$$\eta_{v,K} \leq C\left[\int_{\omega_K} B_p + \sum_{K'\in\omega_K}\theta^2_{v,K'}\|g_H - g\|^2_{0,K'}\right]. \tag{69}$$

Let $K = \bigcup_{i=1}^n T_i$. Note that $\omega_K \supset \bigcup_{i=1}^n \omega_{T_i}$. Then

$$\begin{aligned}\sum_{i=1}^n \eta_{u,T_i} &\leq C\sum_{i=1}^n\left[\int_{\omega_{T_i}} A_p + \sum_{T_i'\in\omega_{T_i}}\theta^2_{u,T_i'}\|f_h-f\|^2_{0,T_i'}\right]\\ &\leq C\left[n\int_{\bigcup_{i=1}^n \omega_{T_i}} A_p + \sum_{i=1}^n\sum_{T_i'\in\omega_{T_i}}\theta^2_{u,T_i'}\|f_h-f\|^2_{0,T_i'}\right]\\ &\leq C\left[n\int_{\omega_K} A_p + \sum_{i=1}^n\sum_{T_i'\in\omega_{T_i}}\theta^2_{u,T_i'}\|f_h-f\|^2_{0,T_i'}\right].\end{aligned}$$

Thus

$$\eta_{v,K} + \frac{1}{n}\sum_{i=1}^n \eta_{u,T_i} \leq C\left[\int_{\omega_K} A_p + B_p + \frac{1}{n}\sum_{i=1}^n\sum_{T_i'\in\omega_{T_i}}\theta^2_{u,T_i'}\|f_h-f\|^2_{0,T_i'} + \sum_{K'\in\omega_K}\theta^2_{v,K'}\|g_H-g\|^2_{0,K'}\right],$$

which completes the proof. □

4.5. Upper bound for the error in only one of the unknowns. In some instances it may be known a priori that one of the unknowns is smoother than the other and that its numerical approximation has a smaller error associated with it. In such a case, we can use a multilevel grid. For the purposes of grid adaptation, it is also useful to estimate the error only in the variable that contributes the bulk fraction of the error. For example, if v is smoother than u, then it is natural to use a multilevel grid with $H >> h$, and define an estimate for the error in u only.

Throughout this section we will assume that there is a constant $0 < \alpha \ll 1$ such that $c\alpha < 1$ and

$$\|e_v\|_{0,T} < \alpha \|e_u\|_{0,T}, \qquad \forall T \in \mathcal{T}_h. \tag{70}$$

We will consider the form on $V \times V$

$$a(u, \phi) = \int_\Omega (\lambda_1 + c)u\phi + a\nabla u \cdot \nabla \phi,$$

and the norm $\|u\|_*^2 := a(u, u)$ on V. With these we prove the following result, which resembles the scalar estimates in [37] for $a = 1$.

Theorem 4.5. *Let u, u_h, v, v_H be as in Theorem 3.1. and suppose that* (70) *holds. Then*

$$\mathcal{E}^* := \|e_u\|_* \le \bar{C}\eta^*, \tag{71}$$

where $\bar{C} := (1 - c\alpha)^{-1} \max\{c_1, c_4\}$ *and*

$$\eta^* := \left\{\textstyle\sum_{T\in\mathcal{T}_h} (\theta^*_{u,T})^2 \|R^*_{T,u}\|^2_{0,T} + \frac{1}{2}\sum_{E\subset\partial T} (\gamma^*_{u,E})^2 \|R_{E,u}\|^2_{0,E}\right\}^{1/2} \tag{72}$$

with

$$\begin{aligned} \theta^*_{u,T} &:= \min\{h_S a^{-1/2}, (\lambda_1 + c)^{-1/2}\} \quad \forall S \in \mathcal{T}_h \cup \mathcal{E}_h, \\ \gamma^*_{u,E} &:= a^{-1/4}(\theta^*_{u,E})^{1/2}. \end{aligned}$$

Proof. Subtracting the first component of (11) from the respective one of (9) with $\phi = \phi_h$ we get

$$a(e_u, \phi_h) = \int_\Omega c(v - \lambda v_H)\phi_h \overset{Remark\ (2.3)}{=} \int_\Omega c e_v \phi_h. \tag{73}$$

By letting $\phi_h = I_h e_u$ in (73) we get

$$\|e_u\|_*^2 = a(e_u, e_u) \overset{(73)}{=} a(e_u, e_u) - a(e_u, I_h e_u) + (ce_v, I_h e_u) = a(e_u, e_u - I_h e_u) + (ce_v, I_h e_u).$$

Now we integrate by parts, use (2), and add and subtract $c(v_H, e_u)$ in the second identity to get

$$\|e_u\|_*^2 = \sum_{T\in\mathcal{T}_h} (f - (\lambda_1 + c)u_h + \nabla(a\nabla u_h) + cv, e_u - I_h e_u)_T + (ce_v, I_h e_u)$$

$$+ \sum_{E\subset\mathcal{E}_h} ([a\partial_n u_h], e_u - I_h e_u)_E = \sum_{T\in\mathcal{T}_h} (R^*_{T,u}, e_u - I_h e_u)_T + (ce_v, e_u) + \sum_{E\subset\mathcal{E}_h} (R_{E,u}, e_u - I_h e_u)_E.$$

Next we estimate the terms in this identity with Cauchy-Schwarz inequality

$$\|e_u\|_*^2 \le \sum_{T\in\mathcal{T}_h} \|R^*_{T,u}\|_{0,T} \|e_u - I_h e_u\|_{0,T} + c\|e_v\|_{0,T}\|e_u\|_{0,T} + \sum_{E\subset\mathcal{E}_h} \|R_{E,u}\|_{0,E}\|e_u - I_h e_u\|_{0,E},$$

and by (70) we obtain

$$(1 - c\alpha)\|e_u\|_*^2 \le \sum_{T\in\mathcal{T}_h} \|R^*_{T,u}\|_{0,T}\|e_u - I_h e_u\|_{0,T} + \sum_{E\subset\mathcal{E}_h} \|R_{E,u}\|_{0,E}\|e_u - I_h e_u\|_{0,E}.$$

To conclude, we apply (29) and (31) replacing λ_1 with $\lambda_1 + c$ to get the following estimate

$$\begin{aligned}(1-c\alpha)\|e_u\|_*^2 &\leq \sum_{T\in\mathcal{T}_h} c_1\theta^*_{u,T}\|R^*_{T,u}\|_{0,T}\|e_u\|_{*,\tilde{\omega}_T} + \sum_{E\subset\mathcal{E}_h} c_4\gamma^*_{u,E}\|R_{E,u}\|_{0,e}\|e_u\|_{*,\tilde{\omega}_T} \\ &\leq \max\{c_1,c_4\}\left[\sum_{T\in\mathcal{T}_h}\theta^*_{u,T}\|R^*_{T,u}\|_{0,T}\|e_u\|_* + \sum_{E\subset\mathcal{E}_h}\gamma^*_{u,E}\|R_{E,u}\|_{0,E}\|e_u\|_*\right].\end{aligned}$$

Dividing both sides by $(1-c\alpha)$ concludes the proof. □

5. Examples

In this section we provide examples illustrating the theory developed above. We demonstrate the robustness of a-posteriori estimators with respect to $\mathcal{P} = (\lambda_1, \lambda_2, a, b, c)$ and illustrate how the multilevel scheme and error estimation work together.

On multilevel grids, we use grid parameters h and H with the number of elements in $\mathcal{T}_h$ and $\mathcal{T}_H$ denoted, respectively, by n and N. In all examples except Example 6 we consider r-refinements $\mathcal{T}_h$ of $\mathcal{T}_H$ i.e. $r = H/h$ where $1 \leq r \in \mathbb{N}$. When $r > 1$ we refer to $\mathcal{T}_h$ as the fine mesh and to $\mathcal{T}_H$ as the coarse mesh.

In each case, we obtain (u_h, v_H) by solving the linear system associated with (11), and compute the error $\mathcal{E}$ using the known analytical solution (u, v). If (u, v) are not known, then we estimate it from the finest grid possible or by Richardson's extrapolation.

We recall that the efficiency index $\Theta := \frac{\eta}{\mathcal{E}}$. For various implicit estimators, asymptotically, $\Theta \downarrow 1$. However, for residual estimators $\Theta >> 1$ [12, 27]. For perspective, we show typical values of Θ for scalar and non-scalar model problems below. In this paper our concern is in showing that Θ remains constant or at least stable for a large range of values in $\mathcal{P}$.

We start with a scalar example demonstrating the typical values of the efficiency index Θ_s for the scalar problem.

Example 1. *Let $\Omega = (0,1)$, $f(x) = x$, $a = 1$. We solve* (1) *with homogeneous boundary conditions imposed. Let u be the solution and u_h be the corresponding finite element solution. We compute $\mathcal{E}_s := \|u - u_h\|_*$ and η^*, using notation from Section 4.5, with $\kappa = \lambda_1$ and $c \equiv 0$. Let $\Theta_s = \frac{\eta^*}{\mathcal{E}_s}$; its values for a range of values of κ are shown in Table 1.*

From Table 1 it is clear that Θ_s is stable for all κ but not constant for large κ. For a large κ we have a singularly perturbed problem and a developing boundary layer which requires a small enough h for convergence of numerical method and for efficiency of the estimator.

Next, we consider the numerical solution (11) to the coupled system (9). We demonstrate that the algorithm converges on multilevel grids and that Θ remains essentially constant. The latter is thanks to the appropriate scaling in the definition of the estimator.

Example 2. *Let $\Omega = (0,1)$ and $u(x) = x^2\sin(\pi x)$, $v(x) = x - x^3$, be the exact solution of* (9) *with $\mathcal{P} = \mathbf{1}^5$. We compute the corresponding f and g, and solve for the numerical solutions u_h, v_H. We consider here various uniform multilevel grids with $r = 1, 2, 5, 100$. Table 2 shows the value of the error and of the error estimate as well as of the efficiency index Θ.*

TABLE 1. Efficiency index Θ_s for the scalar equation (1) from Example 1 and $\mathcal{P}_s = \{1, \kappa\}$

n	$\kappa = 10^{-5}$	$\kappa = 10^{-4}$	$\kappa = 10^{-3}$	$\kappa = 10^{-2}$	$\kappa = 10^{-1}$	$\kappa = 1$
8	4.6810	4.6810	4.6810	4.6808	4.6790	4.6613
16	4.7871	4.7871	4.7871	4.7870	4.7858	4.7747
32	4.8423	4.8423	4.8423	4.8422	4.8416	4.8354
64	4.8705	4.8705	4.8705	4.8704	4.8701	4.8668
128	4.8847	4.8847	4.8847	4.8847	4.8845	4.8828
256	4.8918	4.8918	4.8918	4.8918	4.8917	4.8909
512	4.8954	4.8954	4.8954	4.8954	4.8953	4.8949
1024	4.8972	4.8972	4.8972	4.8972	4.8972	4.8970

n	$\kappa = 10$	$\kappa = 10^2$	$\kappa = 10^3$	$\kappa = 10^4$	$\kappa = 10^5$
8	4.5231	3.5583	1.5819	1.1502	1.0573
16	4.6831	4.4072	2.4450	1.3288	1.0952
32	4.7835	4.5898	4.3229	1.7912	1.1940
64	4.8393	4.7267	4.4699	2.9479	1.4383
128	4.8686	4.8082	4.6431	4.3557	2.0793
256	4.8837	4.8524	4.7597	4.5333	3.5948
512	4.8913	4.8754	4.8264	4.6893	4.4102
1024	4.8951	4.8871	4.8619	4.7870	4.5926

We see that for any grid level r, the error and the estimator converge linearly with H: the error decreases by 1/2 when N is halved. This example also shows robustness of the estimator with respect to h and r: Θ remains essentially constant in all Tables. The value $\Theta \approx 7$ is typical for the coupled system and should be compared to $\Theta_s \approx 4$ in Example 1.

Next we discuss the error for a fixed H and varying r, in order to understand the merits of multilevel discretizations. For example we compare the error and the estimator for $N = 160$ i.e. fourth row in the list for each r in Table 2. We see that the error decreases quite a bit initially between $r = 1$ and $r = 2$ but that it remains dominated by the $O(H)$ component for large r.

These results illustrate in what instances it makes sense to refine the grid in one component only. In general, the refinement in u-component increases the total number of unknowns from $N + N = 2N$ to $rN + N = (r + 1)N$. If useful, this should be accompanied by a proportional decrease in the error by a factor of $(1 + r)/2$. In Example 2 for large r this is not true since for small h the error remains bounded by the $O(H)$ contribution. However, for $r = 2$ we have the desired proportional decrease in the error. Here the number of unknowns between $r = 1$ and $r = 2$ increases by a factor of 1.5 while the error decreases by the factor of $0.0136/0.00870 \approx 1.563$.

The computational cost of a multilevel algorithm obviously is case-dependent since the error components depend on u, v, h, H. Example 2 can be seen as the "worst case scenario" since the components u and v have comparable variability and $\mathcal{P} = \mathbf{1}^5$. However, the usefullness of multilevel grids is evident in other cases to follow, and in particular in the next example which is a variation on Example 2.

Example 3. *Here we modify Example 2 and choose a fast oscillating u component $u(x) = x^2 \sin(10\pi x)$. We keep $v(x) = x - x^3$ with $\mathcal{P} = (1, 1, 1, 1, 1)$. Table 3 shows the results.*

TABLE 2. Results for Example 2

N	$\mathcal{E}$	η	Θ	N	$\mathcal{E}$	η	Θ
$r=1$				$r=2$			
20	0.109	0.814	7.46	20	0.0696	0.526	7.56
40	0.0545	0.415	7.61	40	0.0348	0.266	7.62
80	0.0272	0.209	7.68	80	0.0174	0.134	7.70
160	0.0136	0.105	7.71	160	0.00870	0.0672	7.72
320	0.0068	0.0526	7.73	320	0.00435	0.0336	7.73
640	0.00341	0.0264	7.74	640	0.00218	0.0168	7.74
1280	0.00170	0.0132	7.74	1280	0.00109	0.00842	7.74
2560	0.00085	0.0066	7.74	2560	0.000544	0.00421	7.74
5120	0.00042	0.0033	7.74	5120	0.000272	0.00211	7.74

N	$\mathcal{E}$	η	Θ	N	$\mathcal{E}$	η	Θ
$r=5$				$r=100$			
20	0.0536	0.404	7.54	20	0.0500	0.376	7.52
40	0.0268	0.205	7.64	40	0.0250	0.191	7.63
80	0.0134	0.103	7.69	80	0.0125	0.0961	7.69
160	0.00670	0.052	7.72	160	0.00625	0.0482	7.72
320	0.00335	0.026	7.73	320	0.00312	0.0242	7.73
640	0.00168	0.013	7.74	640	0.00156	0.0121	7.74
1280	0.000838	0.0065	7.74				
2560	0.000419	0.0032	7.74				
5120	0.000209	0.0016	7.74				

TABLE 3. Results for Example 3

n	$\mathcal{E}$	η	Θ	n	$\mathcal{E}$	η	Θ
$r=1$				$r=5$			
20	4.4206	30.049	6.7975	20	4.4274	30.094	6.7971
40	2.2704	16.996	7.4860	40	2.2737	17.019	7.4850
80	1.1433	8.7785	7.6780	80	1.1450	8.7904	7.6774
160	0.57269	4.4259	7.7281	160	0.57351	4.4320	7.7278
320	0.28648	2.2176	7.7411	320	0.28688	2.2208	7.7409
640	0.14325	1.1094	7.7446	640	0.14346	1.1110	7.7445
1280	0.071629	0.55480	7.7455	1280	0.071731	0.55559	7.7455
2560	0.035815	0.27741	7.7458	2560	0.035866	0.27781	7.7458
5120	0.017907	0.13871	7.7459	5120	0.017933	0.13891	7.7459

Consider the error in Example 3 for a fixed h and varying r. For instance, focus on $n = 160$ and the fourth row in each r in Table 3. The error remains almost constant between $r = 1$ and $r = 5$ despite the fact that we are coarsening the H grid by a factor of 5. This is because the error is dominated by the $O(h)$ terms for small r. The number of elements between $r = 1$ and $r = 5$ decreases from 320 down to 192, i.e. , by the factor of 1.66 while the error increases only by less than one percent since $0.5735/0.5727 \approx 1.0014$.

Next we illustrate other properties of the estimator. First we verify how η behaves for a system close to being degenerate.

Example 4. *Let* $\Omega = (0, 1)$ *and* $u(x) = v(x) = x^2 \sin(\pi x)$ *be the exact solution of* (2)-(3). *Let* $\mathcal{P} = (1, 1, 1, 10^{-5}, 10)$. *We use* $r = 1, 5, 40$*, see Table 4.*

Table 4. Results for Example 4 with degenerate $\mathcal{P}$

N	$\mathcal{E}$	η	Θ	N	$\mathcal{E}$	η	Θ
$r=1$				$r=5$			
20	0.049999	0.37583	7.52	20	0.010069	0.077007	7.64
40	0.025000	0.19076	7.63	40	0.0050085	0.038615	7.71
80	0.012500	0.096101	7.68	80	0.0025013	0.019336	7.73
160	0.0062500	0.048231	7.71	160	0.0012503	0.0096758	7.74
320	0.0031250	0.024161	7.73	320	0.00062509	0.0048400	7.74
640	0.0015625	0.012092	7.73	640	0.00031254	0.0024205	7.74
1280	0.00078125	0.0060487	7.74	1280	0.00015627	0.0012103	7.74
2560	0.00039063	0.0030251	7.74	2560	7.81×10^{-5}	0.00060520	7.74
5120	0.00019531	0.0015127	7.74	5120	3.91×10^{-5}	0.00030261	7.74

N	$\mathcal{E}$	η	Θ
$r=40$			
20	0.0017252	0.0098209	5.69
40	0.00069009	0.0048530	7.03
80	0.00032240	0.0024250	7.52
160	0.00015841	0.0012146	7.67
320	7.89×10^{-5}	0.00060995	7.73
640	3.94×10^{-5}	0.00030498	7.74
1280	1.97×10^{-5}	0.00015249	7.74

In Example 4 we demonstrate the robustness of the estimator with respect to H, r, i.e., we show that the estimator converges and that Θ is essentially constant with respect to H and r. This Example also provides yet another motivation for the use of multilevel grids. The error between $r = 1$ and $r = 5$ appears to decrease by the factor $0.0625/0.0015 \approx 4.1$ while the number of unknowns increases by $(1+5)/2 = 3$ (cf. row 4). With $r > 5$ the advantages of multilevel grids deteriorate slowly as r increases because the error in u gets resolved better and it slowly stops dominating the total error. For $r = 40$ we have a decrease of the error by $0.0625/0.00015 \approx 40$ in row 4 but only $0.0499/0.0017 \approx 29$ while the number of unknowns increases by about 20.5. This suggests that $r = 40$ is close to the final value of r beyond which no decrease of the error can happen.

Now we demonstrate the robustness of the estimator with respect to $\mathcal{P}$.

Example 5. *Here we have $\Omega = (0, 1)$ and f, g as in Example 2. We now vary the coefficients in $\mathcal{P}$. Since an analytical solution is not easily obtained for such a problem, we approximate $(u, v) \approx (u_*, v_*)$ where the latter is obtained on a grid with $n_* = 5210$ elements. We fix the grid and set $N = 160$ elements. and let the various coefficients in $\mathcal{P}$ vary, one at a time, by several orders of magnitude. In Table 5 we show the variation of the efficiency index Θ with respect to $\mathcal{P}$ and r.*

We observe in Table 5 that the ratio Θ is essentially constant i.e. the estimator η is quite robust with respect to r and c, λ_1. It is also stable with respect to a, b. However, for $a, b \downarrow 0$ the efficiency index Θ varies, even though it changes only by a factor less than 4 when a, b change by a factor of 10^{10}. This variability arises in a context similar to the singularly perturbed problem in Example 1; for a small enough H, h it disappears.

Next we consider application of the multilevel estimator to grid adaptivity.

TABLE 5. Efficiency index Θ in Example 5. Each row corresponds to a different value of a parameter from $\mathcal{P}$ as indicated while other parameters are kept fixed with value 1. Each column corresponds to the different r

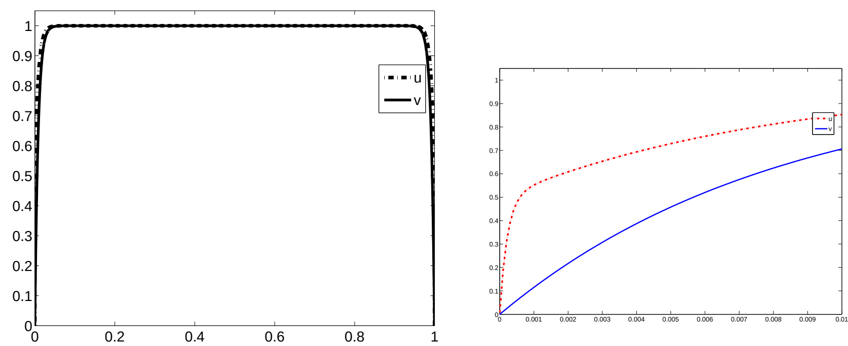
a / r	1	2	4	16
10^{10}	7.76	7.75	7.76	7.72
10^{5}	7.76	7.75	7.76	7.72
1	7.76	7.82	7.92	7.73
10^{-5}	3.06	5.59	6.89	8.08
10^{-10}	1.42	1.63	1.48	2.26

b / r	1	2	4	16
10^{10}	7.76	7.89	8.67	8.94
10^{5}	7.76	7.89	8.67	8.94
1	7.76	7.8199	7.92	7.73
10^{-5}	3.19	3.11	3.10	3.06
10^{-10}	1.53	1.53	1.42	1.43

c / r	1	2	4	16
10^{10}	7.76	6.94	6.24	5.67
10^{5}	7.76	7.61	7.60	7.51
1	7.76	7.82	7.92	7.73
10^{-5}	7.76	7.82	7.92	7.74
10^{-10}	7.76	7.82	7.92	7.74

λ_1 / r	1	2	4	16
10^{10}	7.76	7.75	7.7635	7.72
10^{5}	7.69	7.74	7.7629	7.72
1	7.76	7.82	7.92	7.73
10^{-5}	7.76	7.82	7.92	7.74
10^{-10}	7.76	7.82	7.92	7.74

FIGURE 1. Solutions u, v in Example 6. Left: plot over $(0, 1)$. Right: zoomed in boundary layer for u, v with an additional boundary layer for u.



Remark 5.1. *Various adaptive strategies based on a-posteriori error estimates can be defined.*
Component-based strategy: In each step we mark for refinement those elements $T' \in \mathcal{T}_h$ for which the local estimator

$$\eta_{T',u} > 0.5 \max_T \eta_{T,u}.$$

Analogously we mark the elements $K \in \mathcal{T}_H$. The actual choice of new grid elements honors the requirement that $\mathcal{T}_h$ be a refinement of $\mathcal{T}_H$.

Alternative strategy: we refine those $T' \in \mathcal{T}_h$ for which

$$\eta_{T',u} > 0.5 \min(\max_T \eta_{T,u}, \max_K \eta_{K,v}),$$

with a natural K-analogue.

These two strategies frequently mark the same elements for refinement. However, in some cases the alternative method leads to a faster decrease in $\mathcal{E}$ than the component-based strategy.

Example 6. *Let $\mathcal{P} = (1, 1, a, b, 1)$ and $\Omega = (0, 1)$, $f = g \equiv 1$. This example is from ([23], Example 1) where it is shown that both u, v have both boundary layers of*

TABLE 6. Refinement at each step (recall symmetry of the domain)

	1^{st} Step	2^{nd} Step	3^{rd} Step
$\max \eta_{u,T}$	0.2318	0.0518	0.0178
$\max \eta_{v,K}$	0.2315	0.1087	0.0444
η	0.4818	0.2120	0.140
# of elements	5 + 5 =10	43+43 = 86	79 + 61 = 140

width $O(b^{1/2}\ln b)$*, and that* u *has an extra layer of width* $O(a^{1/2}\ln a)$*. Let* $a = 10^{-7}$*,* $b = 10^{-4}$*. The solution is shown in Figure 1.*

Starting from a uniform grid $\mathcal{T}_h^0 = \mathcal{T}_H^0$ *with* $h = 0.2$ *and* $n = N = 5$*, we use our a-posteriori estimator* η *to guide the appropriate grid refinement in the boundary layer as in Remark 5.1. We show details of the first few steps of this strategy, refering only to the boundary layer on the left hand side; the other side follows by symmetry. Table 6 summarizes the quantitative information and Figure 2 illustrates its effects.*

(1) *After we compute the solution* (u_h, v_H) *and the local error estimator for* $\mathcal{T}_h^0 = \mathcal{T}_H^0$ *we find that we need to refine the grid in the intervals* $[0, 0.2]$ *for both* u *and* v *components, according to both strategies in Remark 5.1. We denote this grid by* $\mathcal{T}_h^1 = \mathcal{T}_H^1$*.*
(2) *Compute the solution and the local error estimator for* $\mathcal{T}_h^1 = \mathcal{T}_H^1$*, here* $n = N = 43$*. We find that we need to refine the elements in* $[0, 0.1]$ *for* u *components and in* $[0, 0.1]$*,* $[0.1, 0.2]$ *for* v *component. The marking is the same for both strategies. We denote the resulting grid by* $\mathcal{T}_h^2, \mathcal{T}_H^2$*. Note that* $\mathcal{T}_h^2 \neq \mathcal{T}_H^2$*.*
(3) *Compute the solution and the local error estimator for* $\mathcal{T}_h^2, \mathcal{T}_H^2$*. Here we have* $n = 79, N = 61$*. Using the component-based strategy, we find that we need to refine in the interval* $[0, 0.001]$ *for* u *and in* $[0.02, 0.03]$ *for* v *component.*

    The alternative strategy marks $[0, 0.001]$ *for* u *and* $[0.01, 0.02], [0.02, 0.03]$ *for* v*.*
(4) *Continue ...*

Analyzing this last Example we see that the a-posteriori error estimators suggest after Step 2 that a multilevel rather than identical grid should be used for the two components. In order to refine separately the u- and v- grids, we need an ability to estimate the error in each component separately and at best locally.

Our next example shows the application of the error estimate in one variable only.

Example 7. *Let* $\Omega = (0, 1)$ *and* $u(x) = x^2 \sin(10\pi x)$*,* $v(x) = x - x^2$ *be the exact solution of* (2)–(3) *with* $\mathcal{P} = (1, 1, 10^{-3}, 1, 100)$*. We note that* $c = 100$ *indicates a rather strong coupling in the system. We use* $N = 320$ *and let* r *vary. Table 7 shows the application of the global estimator* η *and of the estimator* η^* *for* u*-variable only.*

Studying the first row of Table 7 and comparing $\mathcal{E}$ with $\mathcal{E}^*$ and η with η^* we notice that that they are close, i.e., error is dominated by the error in u. It is important to notice that since both estimators are robust in r, we can use the estimators instead of the error information as a tool to determine the dominating component. To decrease that component of the error, we refine the mesh on which u is computed. This helps to decrease the error significantly while making the total number of unknowns grow by a factor smaller than $2r$ between each grid steps.

FIGURE 2. Illustration of adaptive steps from Example 6: plot of solution (u_h, v_H). Top: solution in steps 1 (left) and step 2 (right). Bottom: step 3 with original strategy (left) and with alternative strategy (right). Zoom is indicated by the range of x.

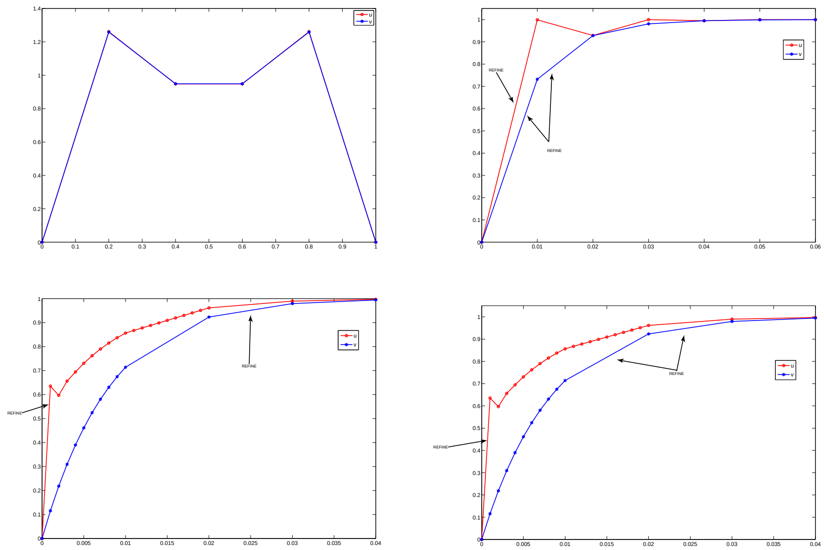


TABLE 7. Robustness and use of estimators η, η^* in Example 7. Shown on the left are the error, estimate, and efficiency index corresponding to the usual estimator (39). On the right we show the corresponding values for the quantities computed in the u variable only (72) and in particular η^* and $\mathcal{E}^* := \|u - u_h\|_*$, and $\Theta^* := \frac{\eta^*}{\mathcal{E}^*}$

rN	η	$\mathcal{E}$	Θ	rN	η^*	$\mathcal{E}^*$	$\eta^*/\mathcal{E}^*$
320	0.071569	0.0093135	7.6844	320	0.070195	0.0091370	7.6825
640	0.037764	0.0048850	7.7305	640	0.035090	0.0045396	7.7296
1280	0.022419	0.0028967	7.7396	1280	0.017544	0.0022662	7.7416
2560	0.016486	0.0021303	7.7387	2560	0.0087722	0.0011327	7.7443
5120	0.014631	0.0018910	7.7371	5120	0.0043861	0.00056644	7.7433
10240	0.014129	0.0018263	7.7365	10240	0.0021931	0.00028347	7.7365

However, in row 6 the errors and the estimators vary already by a factor of ≈ 6 that is, we have decreased the dominating component of the error.

As for computational complexity, we see that the error decreases by a factor of almost 2 between the first row and the second while the number of unknowns increased by a factor of $3/2 = 1.5$. Without multilevel grids, we would have to refine grid in u, v simultaneously, i.e., double the total number of unknowns. We conclude that multilevel grids are quite useful in this example.

In summary, Example 7 is a nice illustration of applicability of the estimator associated with the strongly varying component only.

Next we consider a few examples in $d = 2$ dimensions.

Example 8. *Use as exact solution to* (9) *the functions* $u(x, y) = \sin(2\pi x)(y^2 - y)$, $v(x, y) = (x^2 - x)(y^2 - x)$. *Let* $\Omega = (0, 1)^2$. *The coefficients are set to be* $\mathcal{P} =$

TABLE 8. Convergence of the error and estimator for Example 8, $N = n$ that is $r = 1$ (left) and $N = 4n$ or $r = 4$ (right)

$\frac{1}{h}$	$\mathcal{E}$	η	Θ	$\frac{1}{h}$	$\mathcal{E}$	η	Θ
16	0.13244	1.6367	12.358	4	0.13966	1.6731	11.979
32	0.066409	0.83876	12.630	8	0.067145	0.84301	12.555
64	0.033228	0.42340	12.742	16	0.033321	0.42397	12.724
128	0.016617	0.21255	12.791	32	0.016632	0.21265	12.786

TABLE 9. Efficiency index Θ for Example 9

$a/\ \lambda_2$	10^{-10}	10^{-5}	1	10^5	10^{10}
10^{-10}	3.8324	3.8324	3.7714	1.3563	1.3563
10^{-5}	4.5796	4.5796	4.5117	1.8375	1.8375
1	12.237	12.237	12.231	12.076	12.076
10^5	13.839	13.839	13.839	12.025	12.076
10^{10}	13.840	13.840	13.839	2.2928	12.039

$(1, 1, 1, 10^{-3}, 10)$. *We calculate the corresponding* f, g. *Next, we solve for* (u_h, v_H) *and consider the rate of convergence of the energy error and of the estimator. We use* $N = rn$ *for* $r = 1, 4$. *In Table 8 we can see that* Θ *changes a little with* N *but not much with* r.

Example 9. *Now we vary* $\mathcal{P}$ *in Example 8. Since the analytical solution for general* $\mathcal{P}$ *is not easy to find, we use Richardson extrapolation with* $n_* = 131072$ *elements to approximate the true error. We are interested in the behavior of* Θ *for* $\mathcal{P} = (1, \lambda_2, a, 1, 10)$ *when* a *and* λ_2 *decrease; this example is relevant to a steady-state pseudo-parabolic system* [24]. *The solution* u_h, v_H *is computed with* $n = 8192, N = 2048$ *elements. The results are presented in Table 9*

Our last example shows the application of the global error estimate to adapt the grid uniformly in the goal to satisfy a prescribed tolerance. Specifically, we want to ensure

$$\|(e_u, e_v)\|_e \leq \tau \tag{74}$$

for a given τ. This follows of course if we ensure $\eta \leq \tau$.

Example 10. *We consider* $\mathcal{P} = (0.1, 1, 1, 10^{-3}, 10)$ *and* $\Omega = (0, 1)^2$ *in* (2)-(3). *The analytical solution is given by* $u(x, y) = \sin(2\pi x)(y^2 - y)$, $v(x, y) = (x^2 - x)(y^2 - x)$.

To satisfy (74) *with* $\tau = 0.02$ *we can use either* $\mathcal{T}_h = \mathcal{T}_H$ *and* $h = H = 1/128$ *with a total* $16641 + 16641 = 33282$ *nodes. On the other hand, to satisfy the same tolerance with multilevel mesh it suffices to have* $h = 1/128$, $H = 1/16$, $r = 8$ *and* $16641 + 289 = 16930$ *nodes.*

For $\tau = 0.05$, *we find that* $4225+4225 = 8450$ *nodes are necessary while* $4225+81$ *nodes of multilevel mesh will suffice. Here* $h = 1/64 = 0.015625$, *and* $H = 1/8 = 0.0125$, *so* $r = H/h = 8$.

6. Summary

Above we have defined a convergent multilevel scheme for a stationary system of reaction-diffusion systems coupled by zero'th order terms. We also proposed a-posteriori error estimators and have shown them to be efficient, reliable, and robust.

Our current work involves extensions of these results. In particular, in a forthcoming paper we address the time-dependent case. Analysis of (2)–(3) is a prototype of considerations for more complex systems and more complicated discretizations to be addressed in the future.

References

[1] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[2] M. Ainsworth and J. T. Oden. A posteriori error estimators for second order elliptic systems. II. An optimal order process for calculating self-equilibrating fluxes. *Comput. Math. Appl.*, 26(9):75–87, 1993.

[3] Mark Ainsworth and Ivo Babuska. Reliable and robust a posteriori error estimating for singularly perturbed reaction-diffusion problems. *SIAM J. Numer. Anal.*, 36(2):331–353 (electronic), 1999.

[4] Mark Ainsworth and J. Tinsley Oden. A posteriori error estimators for second order elliptic systems. I. Theoretical foundations and a posteriori error analysis. *Comput. Math. Appl.*, 25(2):101–113, 1993.

[5] Mark Ainsworth and J. Tinsley Oden. A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Engrg.*, 142(1-2):1–88, 1997.

[6] R. Aris. *The mathematical theory of diffusion and reaction in permeable catalysts*. Clarendon Press, Oxford University Press, 1975.

[7] G. I. Barenblatt, I. P. Zheltov, and I. N. Kochina. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks (strata). *J. Appl. Math. Mech.*, 24:1286–1303, 1960.

[8] A. Bergam, C. Bernardi, and Z. Mghazli. A posteriori analysis of the finite element discretization of some parabolic equations. *Math. Comp.*, 74(251):1117–1138 (electronic), 2005.

[9] C. Bernardi and R. Verfürth. Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math.*, 85(4):579–608, 2000.

[10] Dietrich Braess. *Finite elements*. Cambridge University Press, Cambridge, second edition, 2001. Theory, fast solvers, and applications in solid mechanics, Translated from the 1992 German edition by Larry L. Schumaker.

[11] William L. Briggs, Van Emden Henson, and Steve F. McCormick. *A multigrid tutorial*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2000.

[12] Carsten Carstensen. Some remarks on the history and future of averaging techniques in a posteriori finite element error analysis. *ZAMM Z. Angew. Math. Mech.*, 84(1):3–21, 2004.

[13] Carsten Carstensen and Rüdiger Verfürth. Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM J. Numer. Anal.*, 36(5):1571–1587 (electronic), 1999.

[14] Philippe G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR **58** #25001].

[15] Ph. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique*, 9(R-2):77–84, 1975.

[16] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[17] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1983.

[18] W. Hackbusch. *Multigrid methods and applications*, volume 4 of *Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[19] Weizhang Huang. Convergence analysis of finite element solution of one-dimensional singularly perturbed differential equations on equidistributing meshes. *Int. J. Numer. Anal. Model.*, 2(1):57–74, 2005.

[20] Weizhang Huang, Jingtang Ma, and Robert D. Russell. A study of moving mesh PDE methods for numerical simulation of blowup in reaction diffusion equations. *J. Comput. Phys.*, 227(13):6532–6552, 2008.

[21] Claes Johnson. *Numerical solution of partial differential equations by the finite element method*. Cambridge University Press, Cambridge, 1987.

[22] Alexander I. Lee and James M. Hill. On the general linear coupled system for diffusion in media with two diffusivities. *J. Math. Anal. Appl.*, 89(2):530–557, 1982.

[23] Niall Madden and Martin Stynes. A uniformly convergent numerical method for a coupled system of two singularly perturbed linear reaction-diffusion problems. *IMA J. Numer. Anal.*, 23(4):627–644, 2003.

[24] M. Peszynska, R.E. Showalter, and S.-Y.Yi. Homogenization of a pseudoparabolic system. *Applicable Analysis*, 88(9):1265–1282, 2009.

[25] D. M. Ruthven. *Principles of adsorption and adsorption processes*. Wiley, 1984.

[26] Giancarlo Sangalli. An almost-robust a posteriori estimator for the one-dimensional advection-diffusion-reaction problem. In *Applied and industrial mathematics in Italy*, volume 69 of *Ser. Adv. Math. Appl. Sci.*, pages 496–507. World Sci. Publ., Hackensack, NJ, 2005.

[27] Giancarlo Sangalli. Robust a-posteriori estimator for advection-diffusion-reaction problems. *Math. Comp.*, 77(261):41–70 (electronic), 2008.

[28] Grigory I. Shishkin and Lidia P. Shishkina. *Difference methods for singular perturbation problems*, volume 140 of *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*. CRC Press, Boca Raton, FL, 2009.

[29] R. E. Showalter. *Hilbert space methods for partial differential equations*. Electronic Monographs in Differential Equations, San Marcos, TX, 1994. Electronic reprint of the 1977 original.

[30] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[31] R. E. Showalter and D. B. Visarraga. Double-diffusion models from a highly-heterogeneous medium. *J. Math. Anal. Appl.*, 295(1):191–210, 2004.

[32] Ralph E. Showalter, Thomas D. Little, and Ulrich Hornung. Parabolic PDE with hysteresis. *Control Cybernet.*, 25(3):631–643, 1996. Distributed parameter systems: modelling and control (Warsaw, 1995).

[33] Joel Smoller. *Shock waves and reaction-diffusion equations*, volume 258 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1994.

[34] Vidar Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[35] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley & Teubner, 1996.

[36] R. Verfürth. A posteriori error estimators for convection-diffusion equations. *Numer. Math.*, 80(4):641–663, 1998.

[37] R. Verfürth. Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation. *Numer. Math.*, 78(3):479–493, 1998.

[38] R. Verfürth. A posteriori error estimates for finite element discretizations of the heat equation. *Calcolo*, 40(3):195–212, 2003.

[39] R. Verfürth. Robust a posteriori error estimates for nonstationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1783–1802 (electronic), 2005.

[40] R. Verfürth. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1766–1782 (electronic), 2005.

[41] Jinchao Xu. Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J. Numer. Anal.*, 33(5):1759–1777, 1996.

Department of Mathematics, Oregon State University, Corvallis, OR 97331
E-mail: kleinv@math.oregonstate.edu and mpesz@math.oregonstate.edu