Preprint of paper published in Appl. Math. Comput. 219 (2013), no. 13, 7194–7205

COMPLEMENTARY ERROR BOUNDS FOR ELLIPTIC SYSTEMS AND APPLICATIONS

TOMÁŠ VEJCHODSKÝ

Abstract. This contribution derives guaranteed upper bounds of the energy norm of the approximation error for linear elliptic partial differential systems. We generalize the complementarity error estimates known for scalar elliptic problems to general diffusion-convection-reaction linear elliptic systems. For systems we prove analogous properties of these error bounds as for the scalar case. A brief description how the presented general theory applies to linear elasticity is included as well as an application to chemical systems with reactions of at most first order. Numerical experiments showing the sharpness of the obtained upper bounds and their behavior in the adaptive procedure are presented, too.

1. Introduction

Complementarity approach in the calculus of variation is connected with the method of hypercircle which has deep roots going back to 1950, see [26] and also [4]. This approach is based on a formulation of a complementary problem for cogradients of the primal solution. The complementarity can be practically utilized for computation of guaranteed upper bounds of the energy norm of the approximation error.

The guaranteed upper bounds of the error are especially important for reliability of numerical computations. They enable together with an adaptive procedure to solve the problem within the prescribed tolerance. Since the upper bound is guaranteed, the error of the computed approximation is guaranteed to be below this tolerance.

The complementary a posteriori error estimates posses interesting properties. Besides the fact they provide guaranteed upper bounds, they are independent from the way the approximate primal solution is obtained and hence they can be used for arbitrary conforming solution method. Further, provided they are evaluated exactly, the complementary error estimates bound the total error of the approximation – including possible round-off errors, iteration errors in the linear algebraic solver, quadrature errors, etc. Furthermore, certain variants of these error estimates are fully computable in the sense that they contain no problematic constants (like constants form the Friedrichs' and trace inequalities).

On the other hand, the evaluation of the complementarity estimates might be complicated. Moreover, it requires suitable approximation of the complementary solution. This approximation might be also complicated and/or

1991 *Mathematics Subject Classification.* 65N15, 65N30, 74B05, 80A30.

Key words and phrases. a posteriori error estimates, guaranteed error bounds, complementarity, dual finite element methods, elliptic systems, elasticity, diffusion-convection-reaction problems .

expensive to compute. However, in certain cases fast and explicit formulas for the complementary solution exist [2].

The complementary approach to a posteriori error estimates is well established especially for scalar linear elliptic problems. Starting from 1970's we can find several results about the complementary approach (or dual finite element methods), for example in [8, 9, 10, 12, 27]. The complementary approach was worked out by S. Repin and his group into the concept of so-called error majorants, see e.g. [11, 18, 20, 21, 23]. There are also other papers [5, 30], where the complementary idea can be traced. Anyway, the complementarity is not limited to elliptic problems only. There are applications to linear elasticity [16], thermoelasticity [15], Stokes problem [19], Maxwell type problem [3], nonlinear problems [22], etc. In this contribution we generalize the complementary technique to general systems of linear elliptic partial differential equations.

The rest of the paper is organized as follows. Section 2 introduces systems of linear elliptic equations and the needed notation. In Section 3 we derive three variants of the complementary error bounds for linear elliptic systems. In Section 4 we infer the corresponding complementary problems and prove basic properties of the error bounds. In Section 5 we mention how to generalize the presented theory to the system of linear elasticity which is not – strictly speaking – elliptic. Section 6 presents an application to chemical systems. Section 7 provides two numerical experiments and the final Section 8 summarizes the findings, draws conclusions, and mentions further possible generalizations.

2. System of linear elliptic partial differential equations

Let us consider a domain $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary and a system of N linear elliptic partial differential equations in the following general form

$$-\sum_{j=1}^{N} \operatorname{div}\left(\mathcal{A}^{ij}\nabla u^j\right) + \sum_{j=1}^{N} \boldsymbol{b}^{ij}\cdot\nabla u^j + \sum_{j=1}^{N} c^{ij}u^j = f^i \quad \text{in } \Omega, \quad i = 1,2,\dots,N. \tag{1}$$

Functions $u^i \in \mathbb{R}$ represent the solution, $f^i \in \mathbb{R}$ is the right-hand side, and $\mathcal{A}^{ij} \in \mathbb{R}^{d\times d}$, $\boldsymbol{b}^{ij} \in \mathbb{R}^d$, and $c^{ij} \in \mathbb{R}$, $i,j = 1,2,\dots,N$ stand for the diffusion, convection, and reaction coefficients, respectively. Although not explicitly indicated, all these quantities are in general functions of a variable $x \in \Omega$. Symbol ∇ denotes the gradient of a scalar function and div stands for the usual divergence. Further, let us notice that throughout the paper all vectors are understood as columns. The transposition is denoted by $\boldsymbol{v}^T$.

To introduce the boundary conditions, we consider the boundary $\partial\Omega$ to be split into two disjoint parts Γ_{D} and Γ_{N} and we prescribe

$$u^i = g_{\mathrm{D}}^i \quad \text{on } \Gamma_{\mathrm{D}}, \quad i = 1,2,\dots,N, \tag{2}$$

$$\sum_{j=1}^{N} \alpha^{ij}u^j + \sum_{j=1}^{N} \nu^T\mathcal{A}^{ij}\nabla u^i = g_{\mathrm{N}}^i \quad \text{on } \Gamma_{\mathrm{N}}, \quad i = 1,2,\dots,N. \tag{3}$$

Here and below, ν stands for the unit outward normal to the boundary $\partial\Omega$, $g_{\mathrm{D}}^i \in \mathbb{R}$ is a function of $x \in \Gamma_{\mathrm{D}}$ and $g_{\mathrm{N}}^i \in \mathbb{R}$ and $\alpha^{ij} \in \mathbb{R}$ are functions of $x \in \Gamma_{\mathrm{N}}$.

System (1) and boundary conditions (2)–(3) can be reformulated in a vector form:

$$-\,\mathbf{div}(\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}) + \mathbb{B}\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{C}\boldsymbol{u} = \boldsymbol{f} \quad \text{in } \Omega, \tag{4}$$

$$\boldsymbol{u} = \boldsymbol{g}_{\mathbf{D}} \quad \text{on } \Gamma_{\mathrm{D}}, \tag{5}$$

$$\boldsymbol{\alpha}\boldsymbol{u} + (\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u})\nu = \boldsymbol{g}_{\mathbf{N}} \quad \text{on } \Gamma_{\mathrm{N}}, \tag{6}$$

where the column vector $\boldsymbol{u} = \left(u^1, \dots, u^N\right)^T$ has N components, similarly as $\boldsymbol{f}$, $\boldsymbol{g}_{\mathbf{D}}$, and $\boldsymbol{g}_{\mathbf{N}}$. The gradient $\boldsymbol{\nabla}\boldsymbol{u} \in \mathbb{R}^{N\times d}$ is defined in a standard way as $\boldsymbol{\nabla}\boldsymbol{u} = \left(\nabla u^1, \dots, \nabla u^N\right)^T$ as well as the divergence $\mathbf{div}\,\boldsymbol{y} = (\operatorname{div} y^1, \dots, \operatorname{div} y^N)^T \in \mathbb{R}^N$, where y^i stands for the i-th row of the $N \times d$ matrix $\boldsymbol{y}$. The fourth-order tensor $\mathbb{A} \in \mathbb{R}^{N\times d\times N\times d}$ has entries $\mathbb{A}_{ikj\ell} = \mathcal{A}_{k\ell}^{ij}$, the third-order tensor $\mathbb{B} \in \mathbb{R}^{N\times N\times d}$ has entries $\mathbb{B}_{ijk} = \boldsymbol{b}_k^{ij}$, where $i,j = 1,2,\dots,N$ and $k,\ell = 1,2,\dots,d$. The $N \times N$ matrices $\boldsymbol{C} = \left(c^{ij}\right)_{i,j=1}^N$, and $\boldsymbol{\alpha} = \left(\alpha^{ij}\right)_{i,j=1}^N$ are defined in a natural way.

We will concentrate on the weak formulation of problem (4)–(6). Introducing the space

$$V = \left\{\boldsymbol{v} \in \left[H^1(\Omega)\right]^N : \boldsymbol{v} = 0 \text{ on } \Gamma_{\mathrm{D}} \text{ in the sense of traces }\right\}$$

and the Dirichlet lift $\widetilde{\boldsymbol{g}}_{\mathbf{D}} \in \left[H^1(\Omega)\right]^N$ of the Dirichlet data $\boldsymbol{g}_{\mathbf{D}}$, we define the weak solution $\boldsymbol{u} \in \left[H^1(\Omega)\right]^N$ as a function satisfying $\boldsymbol{u} - \widetilde{\boldsymbol{g}}_{\mathbf{D}} \in V$ and

$$\mathcal{B}(\boldsymbol{u},\boldsymbol{v}) = \mathcal{F}(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in V. \tag{7}$$

The bilinear form $\mathcal{B}$ and the linear functional $\mathcal{F}$ are given by

$$\mathcal{B}(\boldsymbol{u},\boldsymbol{v}) = (\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}, \boldsymbol{\nabla}\boldsymbol{v}) + (\mathbb{B}\boldsymbol{\nabla}\boldsymbol{u}, \boldsymbol{v}) + (\boldsymbol{C}\boldsymbol{u}, \boldsymbol{v}) + \langle \boldsymbol{\alpha}\boldsymbol{u}, \boldsymbol{v}\rangle\,, \tag{8}$$

$$\mathcal{F}(\boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) + \langle \boldsymbol{g}_{\mathbf{N}}, \boldsymbol{v}\rangle\,. \tag{9}$$

By symbols $(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle$ we mean the tensor forms of the $L^2(\Omega)$ and $L^2(\Gamma_{\mathrm{N}})$ inner products, respectively, e.g.,

$$(\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}, \boldsymbol{\nabla}\boldsymbol{v}) = \int_\Omega \sum_{i,j=1}^N \sum_{k,\ell=1}^d \mathcal{A}_{k\ell}^{ij} \frac{\partial u^i}{\partial x_k} \frac{\partial v^j}{\partial x_\ell}\,\mathrm{d}x.$$

In addition, we use the colon to denote the entrywise Euclidean scalar product of tensors, e.g., if $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{N\times d}$ then $\boldsymbol{u} : \boldsymbol{v} = \sum_{i=1}^N \sum_{k=1}^d u_{ik} v_{ik}$. We also introduce matrices

$$\boldsymbol{D} = \left(\operatorname{div} \boldsymbol{b}^{ij}\right)_{i,j=1}^N \quad \text{a.e. in } \Omega \quad \text{and} \quad \boldsymbol{E} = \left(\boldsymbol{b}^{ij}\cdot\nu\right)_{i,j=1}^N \quad \text{a.e. on } \Gamma_{\mathrm{N}}.$$

The well-posedness of problem (7) requires the following natural assumptions:

(A1) All entries of tensors $\mathbb{A}$, $\mathbb{B}$, and $\boldsymbol{C}$ are in $L^\infty(\Omega)$, $\boldsymbol{\alpha} \in [L^\infty(\Gamma_{\mathrm{N}})]^{N\times N}$, $\boldsymbol{f} \in [L^2(\Omega)]^N$, and $\boldsymbol{g}_{\mathbf{N}} \in [L^2(\Gamma_{\mathrm{N}})]^N$.

(A2) Tensor $\mathbb{A}$ is symmetric and uniformly positive definite, i.e., $\mathcal{A}^{ij}_{k\ell} = \mathcal{A}^{ji}_{\ell k}$ for all $i,j = 1,2,\dots,N$ and $k,\ell = 1,2,\dots,d$ and there exists a constant $\widetilde{\lambda} > 0$ such that

$$(\mathbb{A}(x)\boldsymbol{\xi}) : \boldsymbol{\xi} \geq \widetilde{\lambda}\,\boldsymbol{\xi} : \boldsymbol{\xi} \quad \forall \boldsymbol{\xi} \in \mathbb{R}^{N\times d} \quad \text{and for a.e. } x \in \Omega. \tag{10}$$

(A3) Coefficients $\boldsymbol{b}^{ij}$ satisfy

$$\boldsymbol{b}^{ij} = \boldsymbol{b}^{ji} \quad \forall i,j = 1,2,\dots,N.$$

(A4) Matrices $\boldsymbol{C} - \frac{1}{2}\boldsymbol{D}$ and $\boldsymbol{\alpha} + \frac{1}{2}\boldsymbol{E}$ are symmetric and positive *semidefinite* almost everywhere in Ω and on Γ_{N}, respectively.

Notice that condition (A1) guarantees integrability of the used integrals, condition (A2) provides ellipticity of problem (7), and conditions (A3) and (A4) enable to prove V-ellipticity of the bilinear form $\mathcal{B}$.

The unique solvability of system (7) follows from the Lax-Milgram lemma due to the boundedness and V-ellipticity of the bilinear form $\mathcal{B}$. The boundedness is immediate from the boundedness of the equation coefficients, see (A1), and from the trace inequality

$$\|v\|_{0,\Gamma_{\mathrm{N}}} \leq C^T_{\Omega,\Gamma_{\mathrm{N}}} \|v\|_{1,\Omega} \quad \forall v \in H^1(\Omega). \tag{11}$$

On the other hand, the V-ellipticity of $\mathcal{B}$ (see Proposition 2.1 below) requires one of the following variants of the Friedrichs' inequality:

$$\|v\|^2_{1,\Omega} \leq C^F_{\Omega,\Gamma}\left(\|\nabla v\|^2_{0,\Omega} + \|v\|^2_{0,\Gamma}\right) \quad \forall v \in H^1(\Omega), \tag{12}$$

$$\|v\|^2_{1,\Omega} \leq C^F_{\Omega,B}\left(\|\nabla v\|^2_{0,\Omega} + \|v\|^2_{0,B}\right) \quad \forall v \in H^1(\Omega), \tag{13}$$

where $\Gamma \neq \emptyset$ is a relatively open subset of $\partial\Omega$ and $B \subset \Omega$ is a ball. For proofs of inequalities (11)–(13) we refer for example to [6] and [17].

Proposition 2.1. *Let assumptions* (A1)–(A4) *be fulfilled and let at least one of the following conditions be satisfied:*

(a) Γ_{D} *is a relatively open subset of* $\partial\Omega$*,*
(b) *there exists a constant* $\tau > 0$ *and a ball* $B \subset \Omega$ *such that* $\boldsymbol{\xi}^T(\boldsymbol{C} - \frac{1}{2}\boldsymbol{D})\boldsymbol{\xi} \geq \tau\boldsymbol{\xi}^T\boldsymbol{\xi}$ *a.e. in* B*, for all* $\boldsymbol{\xi} \in \mathbb{R}^N$*,*
(c) *there exists a constant* $\sigma > 0$ *and a relatively open subset* Γ^0_{N} *of* Γ_{N} *such that* $\boldsymbol{\xi}^T(\boldsymbol{\alpha} + \frac{1}{2}\boldsymbol{E})\boldsymbol{\xi} \geq \sigma\boldsymbol{\xi}^T\boldsymbol{\xi}$ *a.e. on* Γ^0_{N}*, for all* $\boldsymbol{\xi} \in \mathbb{R}^N$*.*

Then the bilinear form $\mathcal{B}$ *is* V*-elliptic.*

Proof. This is a standard result for the scalar case, see e.g. [13]. Its generalization to elliptic systems is straightforward. Assumption (A3) and Green's theorem enable to express

$$\mathcal{B}(\boldsymbol{v},\boldsymbol{v}) = (\mathbb{A}\boldsymbol{\nabla}\boldsymbol{v},\boldsymbol{\nabla}\boldsymbol{v}) + \left((\boldsymbol{C} - \tfrac{1}{2}\boldsymbol{D})\boldsymbol{v},\boldsymbol{v}\right) + \left\langle(\boldsymbol{\alpha} + \tfrac{1}{2}\boldsymbol{E})\boldsymbol{v},\boldsymbol{v}\right\rangle \quad \forall \boldsymbol{v} \in V.$$

The V-ellipticity

$$\mathcal{B}(\boldsymbol{v},\boldsymbol{v}) \geq C\,\|\boldsymbol{v}\|^2_{1,\Omega} \quad \forall \boldsymbol{v} \in V$$

then follows from the uniform positive definiteness (A2), from the positive semidefiniteness (A4), and from the Friedrichs' inequalities (12)–(13). □

3. Guaranteed upper bound on the error

In this section we derive the computable guaranteed upper bound on the energy norm of the error of an approximate solution $\boldsymbol{u}_h \in V$. The approach is independent from the particular numerical method and the approximation $\boldsymbol{u}_h \in V$ might be arbitrary.

The derivation of the upper bound is based on the divergence theorem. We will use it in the following form

$$(\mathbf{div}\,\boldsymbol{y}, \boldsymbol{v}) + (\boldsymbol{y}, \boldsymbol{\nabla v}) - \langle \boldsymbol{y}\nu, \boldsymbol{v} \rangle = 0 \quad \forall \boldsymbol{v} \in V \ \forall \boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N, \tag{14}$$

where the space $[\mathbf{H}(\mathrm{div}, \Omega)]^N$ consists of $N \times d$ matrices whose rows lie in $\mathbf{H}(\mathrm{div}, \Omega)$. Hence, for the weak solution $\boldsymbol{u} \in V$ of (7), for any field $\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N$ and for any $\boldsymbol{u}_h \in V$ and $\boldsymbol{v} \in V$ we obtain the identity

$$\begin{aligned} \mathcal{B}(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}) &= (\boldsymbol{f}, \boldsymbol{v}) + \langle \boldsymbol{g}_{\mathbf{N}}, \boldsymbol{v} \rangle - (\mathbb{A}\boldsymbol{\nabla u}_h, \boldsymbol{\nabla v}) - (\mathbb{B}\boldsymbol{\nabla u}_h, \boldsymbol{v}) - (\boldsymbol{C}\boldsymbol{u}_h, \boldsymbol{v}) \\ &\quad - \langle \boldsymbol{\alpha}\boldsymbol{u}_h, \boldsymbol{v} \rangle + (\mathbf{div}\,\boldsymbol{y}, \boldsymbol{v}) + (\boldsymbol{y}, \boldsymbol{\nabla v}) - \langle \boldsymbol{y}\nu, \boldsymbol{v} \rangle \\ &= (\boldsymbol{r}^*, \boldsymbol{\nabla v}) + (\boldsymbol{r}_\Omega, \boldsymbol{v}) + \langle \boldsymbol{r}_{\mathrm{N}}, \boldsymbol{v} \rangle, \end{aligned} \tag{15}$$

where we introduce the quantities

$$\boldsymbol{r}^* = \boldsymbol{y} - \mathbb{A}\boldsymbol{\nabla u}_h, \tag{16}$$

$$\boldsymbol{r}_\Omega = \boldsymbol{f} - \mathbb{B}\boldsymbol{\nabla u}_h - \boldsymbol{C}\boldsymbol{u}_h + \mathbf{div}\,\boldsymbol{y}, \tag{17}$$

$$\boldsymbol{r}_{\mathrm{N}} = \boldsymbol{g}_{\mathbf{N}} - \boldsymbol{\alpha}\boldsymbol{u}_h - \boldsymbol{y}\nu \tag{18}$$

to simplify the exposition. Relation (15) can be used in two ways to derive the upper bound on the error. These two possibilities are presented below as Lemmas 3.2 and 3.3.

For their formulations we introduce the notation $|\!|\!|\boldsymbol{v}|\!|\!|^2 = \mathcal{B}(\boldsymbol{v}, \boldsymbol{v})$ for the energy norm and $\|\boldsymbol{v}\|^2_{\boldsymbol{M}} = (\boldsymbol{M}\boldsymbol{v}, \boldsymbol{v})$ and $\langle\!|\boldsymbol{v}|\!\rangle^2_{\boldsymbol{K}} = \langle \boldsymbol{K}\boldsymbol{v}, \boldsymbol{v} \rangle$ for norms induced by a symmetric and uniformly positive definite tensors $\boldsymbol{M}$ and $\boldsymbol{K}$, respectively. We will use the same notation even if $\boldsymbol{M}$ or $\boldsymbol{K}$ are positive semidefinite only. In this case $\|\boldsymbol{v}\|_{\boldsymbol{M}}$ and $\langle\!|\boldsymbol{v}|\!\rangle_{\boldsymbol{K}}$ are seminorms only. Furthermore, we introduce the sets

$$Q(\boldsymbol{f}, \boldsymbol{u}_h) = \Big\{ \boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N : \\ \boldsymbol{f} - \mathbb{B}\boldsymbol{\nabla u}_h - \boldsymbol{C}\boldsymbol{u}_h + \mathbf{div}\,\boldsymbol{y} \in \big(\mathrm{Ker}(\boldsymbol{C} - \tfrac{1}{2}\boldsymbol{D})\big)^{\perp} \text{ a.e. in } \Omega \Big\},$$

$$G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h) = \Big\{ \boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N : \\ \boldsymbol{g}_{\mathbf{N}} - \boldsymbol{\alpha}\boldsymbol{u}_h - \boldsymbol{y}\nu \in \big(\mathrm{Ker}(\boldsymbol{\alpha} + \tfrac{1}{2}\boldsymbol{E})\big)^{\perp} \text{ a.e. on } \Gamma_{\mathrm{N}} \Big\},$$

where Ker stands for the kernel, e.g., $\mathrm{Ker}\,\boldsymbol{M} = \big\{\boldsymbol{q} \in \mathbb{R}^N : \boldsymbol{M}\boldsymbol{q} = 0\big\}$ is the kernel of a matrix $\boldsymbol{M} \in \mathbb{R}^{N\times N}$. Further, $S^\perp = \big\{\boldsymbol{q} \in \mathbb{R}^N : \boldsymbol{q} \cdot \boldsymbol{w} = 0 \quad \forall \boldsymbol{w} \in S\big\}$ denotes the orthogonal complement of $S \subset \mathbb{R}^N$. As an example, notice that if the matrix $\boldsymbol{C} - \frac{1}{2}\boldsymbol{D}$ is nonsingular then $\big(\mathrm{Ker}(\boldsymbol{C} - \frac{1}{2}\boldsymbol{D})\big)^\perp = \mathbb{R}^N$ and $Q(\boldsymbol{f}, \boldsymbol{u}_h) = [\mathbf{H}(\mathrm{div}, \Omega)]^N$. On the other hand, if $\boldsymbol{C} - \frac{1}{2}\boldsymbol{D} = 0$ then $\big(\mathrm{Ker}(\boldsymbol{C} - \frac{1}{2}\boldsymbol{D})\big)^\perp = \{0\}$ and $Q(\boldsymbol{f}, \boldsymbol{u}_h)$ is a set of those vector fields whose divergence is equal to $-\boldsymbol{f} + \mathbb{B}\boldsymbol{\nabla u}_h + \boldsymbol{C}\boldsymbol{u}_h$ a.e. in Ω. Finally, by $\boldsymbol{M}^\dagger$ we denote the Moore-Penrose pseudoinverse of $\boldsymbol{M}$.

Lemma 3.1. *Let H be a Hilbert space with an inner product $(\cdot,\cdot)$. Let $M : H \mapsto H$ be linear, continuous, symmetric, and positive semidefinite operator. If $p, w \in H$ and $p \in (\operatorname{Ker} M)^\perp$ then*

$$(p, w) \leq \|p\|_{M^\dagger} \|w\|_M ,$$

where $\|p\|_{M^\dagger} = (p, M^\dagger p)$ and $\|w\|_M^2 = (Mw, w)$ are seminorms in general.

Proof. Since M is symmetric and positive semidefinite, there exists an operator $K : H \mapsto H$ such that $M = K^T K$. Further, we set $q = M^\dagger p$. Since $p \in (\operatorname{Ker} M)^\perp$ and the range of M coincides with $(\operatorname{Ker} M)^\perp$, we have $Mq = p$. Now, we can directly compute

$$(p, w) = (K^T K q, w) = (Kq, Kw) \leq (Kq, Kq)^{1/2} (Kw, Kw)^{1/2}$$
$$= (Mq, q)^{1/2} \|w\|_M = \|p\|_{M^\dagger} \|w\|_M .$$

□

Lemma 3.2. *Let assumptions* (A1)–(A4) *be fulfilled. If $\boldsymbol{u} \in V$ stands for the weak solution of problem* (7) *then*

$$|||\boldsymbol{u} - \boldsymbol{u}_h||| \leq \eta(\boldsymbol{u}_h, \boldsymbol{y}) \quad \forall \boldsymbol{u}_h \in V, \ \forall \boldsymbol{y} \in Q(\boldsymbol{f}, \boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h) \tag{19}$$

with

$$\eta^2(\boldsymbol{u}_h, \boldsymbol{y}) = \|\boldsymbol{r}^*\|^2_{\mathbb{A}^{-1}} + \|\boldsymbol{r}_\Omega\|^2_{(\boldsymbol{C}-\frac{1}{2}\boldsymbol{D})^\dagger} + \langle|\boldsymbol{r}_\mathrm{N}|\rangle^2_{(\boldsymbol{\alpha}+\frac{1}{2}\boldsymbol{E})^\dagger}. \tag{20}$$

Proof. The statement follows from the identity (15). If we apply Lemma 3.1 to the term $(\boldsymbol{r}^*, \boldsymbol{\nabla v})$ with $M = \mathbb{A}$ (since $\mathbb{A}$ is invertible we have $\mathbb{A}^\dagger = \mathbb{A}^{-1}$), to the term $(\boldsymbol{r}_\Omega, \boldsymbol{v})$ with $M = \boldsymbol{C} - \frac{1}{2}\boldsymbol{D}$, and to the term $\langle \boldsymbol{r}_\mathrm{N}, \boldsymbol{v}\rangle$ with $M = \boldsymbol{\alpha} + \frac{1}{2}\boldsymbol{E}$, we obtain

$$\mathcal{B}(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}) \leq \|\boldsymbol{r}^*\|_{\mathbb{A}^{-1}} \|\boldsymbol{\nabla v}\|_{\mathbb{A}} + \|\boldsymbol{r}_\Omega\|_{(\boldsymbol{C}-\frac{1}{2}\boldsymbol{D})^\dagger} \|\boldsymbol{v}\|_{\boldsymbol{C}-\frac{1}{2}\boldsymbol{D}} + \langle|\boldsymbol{r}_\mathrm{N}|\rangle_{(\boldsymbol{\alpha}+\frac{1}{2}\boldsymbol{E})^\dagger} \langle|\boldsymbol{v}|\rangle_{\boldsymbol{\alpha}+\frac{1}{2}\boldsymbol{E}}$$
$$\leq \left(\|\boldsymbol{r}^*\|^2_{\mathbb{A}^{-1}} + \|\boldsymbol{r}_\Omega\|^2_{(\boldsymbol{C}-\frac{1}{2}\boldsymbol{D})^\dagger} + \langle|\boldsymbol{r}_\mathrm{N}|\rangle^2_{(\boldsymbol{\alpha}+\frac{1}{2}\boldsymbol{E})^\dagger} \right)^{1/2}$$
$$\times \left(\|\boldsymbol{\nabla v}\|^2_{\mathbb{A}} + \|\boldsymbol{v}\|^2_{\boldsymbol{C}-\frac{1}{2}\boldsymbol{D}} + \langle|\boldsymbol{v}|\rangle^2_{\boldsymbol{\alpha}+\frac{1}{2}\boldsymbol{E}} \right)^{1/2} = \eta(\boldsymbol{u}_h, \boldsymbol{y}) |||\boldsymbol{v}|||.$$

Substitution $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}_h$ yields immediately the statement of the lemma. □

Lemma 3.3. *Let assumptions* (A1)–(A4) *be fulfilled and let at least one of conditions* (a)–(c) *from Proposition 2.1 be satisfied. If $\boldsymbol{u} \in V$ stands for the weak solution of problem* (7) *then*

$$|||\boldsymbol{u} - \boldsymbol{u}_h||| \leq \widehat{\eta}(\boldsymbol{u}_h, \boldsymbol{y}) \quad \forall \boldsymbol{u}_h \in V, \ \forall \boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N, \tag{21}$$

with

$$\widehat{\eta}(\boldsymbol{u}_h, \boldsymbol{y}) = \|\boldsymbol{r}^*\|_{\mathbb{A}^{-1}} + C_0 \|\boldsymbol{r}_\Omega\|_{0,\Omega} + C_1 \|\boldsymbol{r}_\mathrm{N}\|_{0,\Gamma_\mathrm{N}} ,$$

where the constant C_0 is given in terms of constants from the uniform positive definiteness (10), *trace theorem* (11), *and the Friedrichs' inequalities* (12)–(13) *and its value depends on the validity of conditions* (a)–(c) *from Proposition 2.1. If* (a) *is satisfied then $C_0^2 = C^F_{\Omega,\Gamma_\mathrm{D}}/\widetilde{\lambda}$, if* (b) *is satisfied then $C_0^2 = C^F_{\Omega,B} \max\{\widetilde{\lambda}^{-1}, \tau^{-1}\}$, and if* (c) *is satisfied then $C_0^2 = C^F_{\Omega,\Gamma^0_\mathrm{N}} \max\{\widetilde{\lambda}^{-1}, \sigma^{-1}\}$. If more then one of conditions* (a)–(c) *are satisfied simultaneously then C_0 attains the smallest of the possible values. Finally, $C_1^2 = C^T_{\Omega,\Gamma_\mathrm{N}}/\widetilde{\lambda}$, see* (11).

Proof. Let us consider arbitrary $\boldsymbol{v} \in V$. If condition (a) is satisfied then

$$\|\boldsymbol{v}\|_{0,\Omega}^2 \le C^F_{\Omega,\Gamma_\mathrm{D}} \|\boldsymbol{\nabla v}\|_{0,\Omega}^2 \le C^F_{\Omega,\Gamma_\mathrm{D}}/\widetilde{\lambda} |||\boldsymbol{v}|||^2.$$

If condition (b) is satisfied then

$$\|\boldsymbol{v}\|_{0,\Omega}^2 \le C^F_{\Omega,B}\left(\|\boldsymbol{\nabla v}\|_{0,\Omega}^2 + \|\boldsymbol{v}\|_{0,B}^2\right) \le C^F_{\Omega,B}\left(\frac{1}{\widetilde{\lambda}}\|\mathbb{A}\boldsymbol{\nabla v}\|_{0,\Omega}^2 + \frac{1}{\tau}\|\boldsymbol{v}\|^2_{\boldsymbol{C}-\frac{1}{2}\boldsymbol{D}}\right)$$
$$\le C^F_{\Omega,\Gamma_\mathrm{D}} \max\{\widetilde{\lambda}^{-1}, \tau^{-1}\} |||\boldsymbol{v}|||^2.$$

Finally, if condition (c) is satisfied then

$$\|\boldsymbol{v}\|_{0,\Omega}^2 \le C^F_{\Omega,\Gamma^0_\mathrm{N}}\left(\|\boldsymbol{\nabla v}\|_{0,\Omega}^2 + \|\boldsymbol{v}\|_{0,\Gamma^0_\mathrm{N}}^2\right) \le C^F_{\Omega,\Gamma^0_\mathrm{N}}\left(\frac{1}{\widetilde{\lambda}}\|\mathbb{A}\boldsymbol{\nabla v}\|_{0,\Omega}^2 + \frac{1}{\sigma}\langle|\boldsymbol{v}|\rangle^2_{\boldsymbol{\alpha}+\frac{1}{2}\boldsymbol{E}}\right)$$
$$\le C^F_{\Omega,\Gamma^0_\mathrm{N}} \max\{\widetilde{\lambda}^{-1}, \sigma^{-1}\} |||\boldsymbol{v}|||^2.$$

Thus, in any case, we have $\|\boldsymbol{v}\|_{0,\Omega} \le C_0 |||\boldsymbol{v}|||$. Similarly, the trace theorem implies $\|\boldsymbol{v}\|_{0,\Gamma_\mathrm{N}} \le C_1 |||\boldsymbol{v}|||$. These estimates used in (15) yield

$$\mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{v}) \le \left(\|\boldsymbol{r}^*\|_{\mathbb{A}^{-1}} + C_0\|\boldsymbol{r}_\Omega\|_{0,\Omega} + C_1\|\boldsymbol{r}_\mathrm{N}\|_{0,\Gamma_\mathrm{N}}\right) |||\boldsymbol{v}|||.$$

Similarly as before, substitution $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}_h$ gives the desired result. □

The estimates (19) and (21) have their advantages and disadvantages. The value of $\eta(\boldsymbol{u}_h, \boldsymbol{y})$ can be easily computed only if the sets $Q(\boldsymbol{f}, \boldsymbol{u}_h)$ and $G(\boldsymbol{g}_\mathbf{N}, \boldsymbol{u}_h)$ can be handled well. This is the case if $\boldsymbol{C} - \frac{1}{2}\boldsymbol{D}$ and $\boldsymbol{\alpha} - \frac{1}{2}\boldsymbol{E}$ are nonsingular, for example. On the other hand, estimate (21) is valid in general for any $\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N$, but evaluation of $\widehat{\eta}(\boldsymbol{u}_h, \boldsymbol{y})$ requires the knowledge of constants C_0 and C_1 or of their upper bounds.

The error bounds (19) and (21) can be simplified if $\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N$ is chosen in a special form. First of all, it is easy to constrain the $\boldsymbol{y}$ such that $\boldsymbol{r}_\mathrm{N} = 0$ a.e. on Γ_N. It is just a natural boundary condition of the Dirichlet type for vector fields from $[\mathbf{H}(\mathrm{div}, \Omega)]^N$. To handle this constraint we introduce an affine space

$$G_0(\boldsymbol{g}_\mathbf{N}, \boldsymbol{u}_h) = \left\{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N : \boldsymbol{y}\nu = \boldsymbol{g}_\mathbf{N} - \boldsymbol{\alpha}\boldsymbol{u}_h \text{ a.e. on } \Gamma_\mathrm{N}\right\} \subset G(\boldsymbol{g}_\mathbf{N}, \boldsymbol{u}_h).$$

Notice that substitution $\boldsymbol{g}_\mathbf{N} = 0$ and $\boldsymbol{u}_h = 0$ yields a linear space

$$G_0(0,0) = \{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N : \boldsymbol{y}\nu = 0 \text{ a.e. on } \Gamma_\mathrm{N}\}.$$

Clearly, $G_0(\boldsymbol{g}_\mathbf{N}, \boldsymbol{u}_h) = \boldsymbol{y}_G + G_0(0,0)$, where $\boldsymbol{y}_G$ is an arbitrary but fixed element of $G_0(\boldsymbol{g}_\mathbf{N}, \boldsymbol{u}_h)$.

This constraint on $\boldsymbol{y}$ simplifies estimates (19) and (21) as follows

$$|||\boldsymbol{u} - \boldsymbol{u}_h|||^2 \le \eta^2(\boldsymbol{u}_h, \boldsymbol{y}) = \|\boldsymbol{r}^*\|^2_{\mathbb{A}^{-1}} + \|\boldsymbol{r}_\Omega\|^2_{(\boldsymbol{C}-\frac{1}{2}\boldsymbol{D})^\dagger} \tag{22}$$

for any $\boldsymbol{u}_h \in V$ and $\boldsymbol{y} \in Q(\boldsymbol{f}, \boldsymbol{u}_h) \cap G_0(\boldsymbol{g}_\mathbf{N}, \boldsymbol{u}_h)$ and

$$|||\boldsymbol{u} - \boldsymbol{u}_h||| \le \widehat{\eta}(\boldsymbol{u}_h, \boldsymbol{y}) = \|\boldsymbol{r}^*\|_{\mathbb{A}^{-1}} + C_0\|\boldsymbol{r}_\Omega\|_{0,\Omega} \tag{23}$$

for any $\boldsymbol{u}_h \in V$ and $\boldsymbol{y} \in G_0(\boldsymbol{g}_\mathbf{N}, \boldsymbol{u}_h)$.

Further, it is possible to constrain $\boldsymbol{y}$ even more such that $\boldsymbol{r}_\Omega = 0$ a.e. in Ω holds. This approach is advantagenous in particular if $\boldsymbol{C}$ and $\mathbb{B}$ vanish (or if they are small) and if $\boldsymbol{f}$ is a simple function (e.g. a constant). Then it is easy to construct such $\boldsymbol{y}$ that $\boldsymbol{r}_\Omega$ vanishes in Ω and the resulting upper bound provides sharp results. Formally, we introduce an affine space

$$Q_0(\boldsymbol{f},\boldsymbol{u}_h) = \left\{ \boldsymbol{y} \in [\mathbf{H}(\mathrm{div},\Omega)]^N : \mathbf{div}\,\boldsymbol{y} = -\boldsymbol{f} + \mathbb{B}\boldsymbol{\nabla}\boldsymbol{u}_h + \boldsymbol{C}\boldsymbol{u}_h \text{ in } \Omega \right\} \subset Q(\boldsymbol{f},\boldsymbol{u}_h)$$

and observe that both $\eta(\boldsymbol{u}_h,\boldsymbol{y})$ and $\widehat{\eta}(\boldsymbol{u}_h,\boldsymbol{y})$ collapse to

$$\widetilde{\eta}(\boldsymbol{u}_h,\boldsymbol{y}) = \|\boldsymbol{y} - \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}_h\|_{\mathbb{A}^{-1}} \quad \text{for all } \boldsymbol{u}_h \in V \text{ and } \boldsymbol{y} \in Q_0(\boldsymbol{f},\boldsymbol{u}_h) \cap G_0(\boldsymbol{g}_{\mathbf{N}},\boldsymbol{u}_h) \tag{24}$$

which is still an upper bound on the energy norm of the error.

The error estimate $\widetilde{\eta}$ has certain advantages. There are no constants C_0 and C_1 as in $\widehat{\eta}$. It is applicable in general and no Moore-Penrose pseudoinverse of $\boldsymbol{C} - \frac{1}{2}\boldsymbol{D}$ or $\boldsymbol{\alpha} + \frac{1}{2}\boldsymbol{E}$ is needed. However, there are also disadvantages. It might be complicated to construct suitable $\boldsymbol{y} \in Q_0(\boldsymbol{f},\boldsymbol{u}_h)$ in general. Moreover, if the coefficients $\boldsymbol{C}$ or $\mathbb{B}$ dominates $\mathbb{A}$ than the upper bound $\widetilde{\eta}$ is inaccurate. For more details how to handle the spaces $Q(\boldsymbol{f},\boldsymbol{u}_h)$ and $Q_0(\boldsymbol{f},\boldsymbol{u}_h)$ see [12, 28].

4. The complementary problem

For practical utilization of estimates (19), (21), and (24), it is necessary to specify a suitable value of $\boldsymbol{y}$. This value must be easily computable and should lead to a sharp estimate of the error. A natural approach is to consider fixed $\boldsymbol{u}_h \in V$ and approximately minimize the quantity $\eta(\boldsymbol{u}_h,\boldsymbol{y})$ with respect to $\boldsymbol{y} \in Q(\boldsymbol{f},\boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}},\boldsymbol{u}_h)$, the quantity $\widehat{\eta}(\boldsymbol{u}_h,\boldsymbol{y})$ with respect to $\boldsymbol{y} \in [\mathbf{H}(\mathrm{div},\Omega)]^N$ and the quantity $\widetilde{\eta}(\boldsymbol{u}_h,\boldsymbol{y})$ with respect to $\boldsymbol{y} \in Q_0(\boldsymbol{f},\boldsymbol{u}_h) \cap G_0(\boldsymbol{g}_{\mathbf{N}},\boldsymbol{u}_h)$.

Let us start with the minimization of η^2. The minimization problem reads: find $\boldsymbol{y}^* \in Q(\boldsymbol{f},\boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}},\boldsymbol{u}_h)$ such that

$$\eta(\boldsymbol{u}_h,\boldsymbol{y}^*) \le \eta(\boldsymbol{u}_h,\boldsymbol{y}) \quad \forall \boldsymbol{y} \in Q(\boldsymbol{f},\boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}},\boldsymbol{u}_h). \tag{25}$$

Since $\eta^2(\boldsymbol{u}_h,\boldsymbol{y})$ is quadratic in $\boldsymbol{y}$, it is easy to see that this minimization problem is equivalent to the variational problem: find $\boldsymbol{y}^* \in Q(\boldsymbol{f},\boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}},\boldsymbol{u}_h)$ such that

$$\mathcal{B}^*(\boldsymbol{y}^*,\boldsymbol{w}) = \mathcal{F}^*_{\boldsymbol{u}_h}(\boldsymbol{w}) \quad \forall \boldsymbol{w} \in Q(0,0) \cap G(0,0), \tag{26}$$

where the bilinear form $\mathcal{B}^*$ and the linear functional $\mathcal{F}^*_{\boldsymbol{u}_h}$ are given by

$$\mathcal{B}^*(\boldsymbol{y}^*,\boldsymbol{w}) = \left((\boldsymbol{C} - \tfrac{1}{2}\boldsymbol{D})^\dagger\, \mathbf{div}\,\boldsymbol{y}^*, \mathbf{div}\,\boldsymbol{w} \right) + \left(\mathbb{A}^{-1}\boldsymbol{y}^*, \boldsymbol{w} \right) + \left\langle (\boldsymbol{\alpha} + \tfrac{1}{2}\boldsymbol{E})^\dagger \boldsymbol{y}\nu, \boldsymbol{w}\nu \right\rangle,$$

$$\begin{aligned} \mathcal{F}^*_{\boldsymbol{u}_h}(\boldsymbol{w}) &= \left((\boldsymbol{C} - \tfrac{1}{2}\boldsymbol{D})^\dagger (-\boldsymbol{f} + \boldsymbol{C}\boldsymbol{u}_h + \mathbb{B}\boldsymbol{\nabla}\boldsymbol{u}_h), \mathbf{div}\,\boldsymbol{w} \right) + (\boldsymbol{\nabla}\boldsymbol{u}_h, \boldsymbol{w}) \\ &\quad + \left\langle (\boldsymbol{\alpha} + \tfrac{1}{2}\boldsymbol{E})^\dagger (\boldsymbol{g}_{\mathbf{N}} - \boldsymbol{\alpha}\boldsymbol{u}_h), \boldsymbol{w}\nu \right\rangle. \end{aligned}$$

The upper bound $\widehat{\eta}(\boldsymbol{u}_h,\boldsymbol{y})$ can also be minimized with respect to $\boldsymbol{y} \in [\mathbf{H}(\mathrm{div},\Omega)]^N$, but it is not a simple quadratic minimization. However, following [20], we can estimate $\widehat{\eta}^2(\boldsymbol{u}_h,\boldsymbol{y})$ for $\boldsymbol{y} \in [\mathbf{H}(\mathrm{div},\Omega)]^N$ as follows

$$\begin{aligned} \widehat{\eta}^2(\boldsymbol{u}_h,\boldsymbol{y}) &= \left(\|\boldsymbol{r}^*\|_{\mathbb{A}^{-1}} + C_0 \|\boldsymbol{r}_\Omega\|_{0,\Omega} + C_1 \|\boldsymbol{r}_{\mathrm{N}}\|_{0,\Gamma_{\mathrm{N}}} \right)^2 \le \widehat{\eta}^2_{\beta,\gamma}(\boldsymbol{u}_h,\boldsymbol{y}), \\ \widehat{\eta}^2_{\beta,\gamma}(\boldsymbol{u}_h,\boldsymbol{y}) &= (1+\beta^{-1}) \|\boldsymbol{r}^*\|^2_{\mathbb{A}^{-1}} + (1+\beta)(1+\gamma) C_0^2 \|\boldsymbol{r}_\Omega\|^2_{0,\Omega} \\ &\quad + (1+\beta)(1+\gamma^{-1}) C_1^2 \|\boldsymbol{r}_{\mathrm{N}}\|^2_{0,\Gamma_{\mathrm{N}}} \qquad \forall \beta > 0,\ \gamma > 0. \end{aligned}$$

For a fixed $\beta > 0$ and $\gamma > 0$, the quantity $\widehat{\eta}^2_{\beta,\gamma}(\boldsymbol{u}_h, \boldsymbol{y})$ is already a quadratic functional in $\boldsymbol{y}$. Formally, it is in the same form as $\eta(\boldsymbol{u}_h, \boldsymbol{y})$. As before, the minimizer $\widehat{\boldsymbol{y}}^*$ of $\widehat{\eta}^2_{\beta,\gamma}(\boldsymbol{u}_h, \boldsymbol{y})$ solves the following variational problem: find $\widehat{\boldsymbol{y}}^* \in [\mathbf{H}(\mathrm{div}, \Omega)]^N$ such that

$$\widehat{\mathcal{B}}^*(\widehat{\boldsymbol{y}}^*, \boldsymbol{w}) = \widehat{\mathcal{F}}^*_{\boldsymbol{u}_h}(\boldsymbol{w}) \quad \forall \boldsymbol{w} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N, \tag{27}$$

where the bilinear form $\widehat{\mathcal{B}}^*$ and the linear functional $\widehat{\mathcal{F}}^*_{\boldsymbol{u}_h}$ are given by

$$\begin{aligned}
\widehat{\mathcal{B}}^*(\boldsymbol{y}, \boldsymbol{w}) &= (1+\beta)(1+\gamma)C_0^2\,(\mathbf{div}\,\boldsymbol{y}, \mathbf{div}\,\boldsymbol{w}) + (1+\beta^{-1})\left(\mathbb{A}^{-1}\boldsymbol{y}, \boldsymbol{w}\right) \\
&\quad + (1+\beta)(1+\gamma^{-1})C_1^2\,\langle \boldsymbol{y}\nu, \boldsymbol{w}\nu\rangle\,, \\
\widehat{\mathcal{F}}^*_{\boldsymbol{u}_h}(\boldsymbol{w}) &= (1+\beta)(1+\gamma)C_0^2\,(-\boldsymbol{f} + \boldsymbol{C}\boldsymbol{u}_h + \mathbb{B}\boldsymbol{\nabla}\boldsymbol{u}_h, \mathbf{div}\,\boldsymbol{w}) + (1+\beta^{-1})(\boldsymbol{\nabla}\boldsymbol{u}_h, \boldsymbol{w}) \\
&\quad + (1+\beta)(1+\gamma^{-1})C_1^2\,\langle \boldsymbol{g}_{\mathbf{N}} - \boldsymbol{\alpha}\boldsymbol{u}_h, \boldsymbol{w}\nu\rangle\,.
\end{aligned}$$

Finally, we introduce the minimization of $\widetilde{\eta}^2(\boldsymbol{u}_h, \boldsymbol{y})$ with respect to $\boldsymbol{y} \in Q_0(\boldsymbol{f}, \boldsymbol{u}_h) \cap G_0(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h)$. The minimization problem: find $\widetilde{\boldsymbol{y}}^* \in Q_0(\boldsymbol{f}, \boldsymbol{u}_h) \cap G_0(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h)$ such that

$$\widetilde{\eta}^2(\boldsymbol{u}_h, \widetilde{\boldsymbol{y}}^*) \le \widetilde{\eta}^2(\boldsymbol{u}_h, \boldsymbol{y}) \quad \forall \boldsymbol{y} \in Q_0(\boldsymbol{f}, \boldsymbol{u}_h) \cap G_0(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h)$$

is equivalent to the variational problem: find $\widetilde{\boldsymbol{y}}^* \in Q_0(\boldsymbol{f}, \boldsymbol{u}_h) \cap G_0(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h)$ such that

$$\left(\mathbb{A}^{-1}\widetilde{\boldsymbol{y}}^*, \boldsymbol{w}\right) = (\boldsymbol{\nabla}\boldsymbol{u}_h, \boldsymbol{w}) \quad \forall \boldsymbol{w} \in Q_0(0,0) \cap G_0(0,0). \tag{28}$$

Problems (26), (27), and (28) are called complementary problems to (7). Consistently, we call $\boldsymbol{y}^*$, $\widehat{\boldsymbol{y}}^*$, and $\widetilde{\boldsymbol{y}}^*$ complementary solutions, $\mathcal{B}^*$, $\widehat{\mathcal{B}}^*$, and $\widetilde{\mathcal{B}}^*$ the complementary bilinear forms, etc. For the further reference we introduce the complementary energy norm $\|\!|\boldsymbol{w}|\!\|_*^2 = \mathcal{B}^*(\boldsymbol{w}, \boldsymbol{w})$. Notice that the unique solvability of the complementary problems (26), (27), and (28) can be verified by the Lax-Milgram lemma.

In a special case, when the convection coefficients matrix $\mathbb{B}$ vanishes, the complementary problem (26) has interesting properties. First of all, if $\mathbb{B}$ vanishes then the complementary bilinear form $\mathcal{B}^*$ and the linear function $\mathcal{F}^*$ simplify to

$$\begin{aligned}
\mathcal{B}^*(\boldsymbol{y}^*, \boldsymbol{w}) &= \Big(\boldsymbol{C}^\dagger\,\mathbf{div}\,\boldsymbol{y}^*, \mathbf{div}\,\boldsymbol{w}\Big) + \left(\mathbb{A}^{-1}\boldsymbol{y}^*, \boldsymbol{w}\right) + \Big\langle \boldsymbol{\alpha}^\dagger\boldsymbol{y}\nu, \boldsymbol{w}\nu\Big\rangle\,, \\
\mathcal{F}^*_{\boldsymbol{u}_h}(\boldsymbol{w}) = \mathcal{F}^*(\boldsymbol{w}) &= \Big(-\boldsymbol{C}^\dagger\boldsymbol{f}, \mathbf{div}\,\boldsymbol{w}\Big) + \langle \boldsymbol{g}_{\mathbf{D}}, \boldsymbol{w}\nu\rangle_{\Gamma_{\mathrm{D}}} + \Big\langle \boldsymbol{\alpha}^\dagger\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{w}\nu\Big\rangle\,.
\end{aligned}$$

Notice that in this case the complementary problem is independent from the approximate solution $\boldsymbol{u}_h \in V$. The following theorem summarizes the properties of the complementary solution $\boldsymbol{y}^*$ of (26).

Theorem 4.1. *Let assumptions* (A1)–(A4) *be fulfilled. Let* $\mathbb{B} = 0$. *Further, let* $\boldsymbol{u} \in V$ *be the exact solution to the primal problem* (7) *and let* $\boldsymbol{y}^* \in Q(\boldsymbol{f}, \boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h)$ *be the exact solution to the complementary problem*

(26). *If* $\boldsymbol{u}_h \in V$ *is arbitrary but fixed then*

$$\boldsymbol{y}^* = \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}, \tag{29}$$

$$\eta(\boldsymbol{u}_h, \boldsymbol{y}^*) = |||\boldsymbol{u} - \boldsymbol{u}_h|||, \tag{30}$$

$$\eta(\boldsymbol{u}, \boldsymbol{y}_h) = |||\boldsymbol{y}^* - \boldsymbol{y}_h|||_* \quad \forall \boldsymbol{y}_h \in Q(\boldsymbol{f}, \boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h), \tag{31}$$

$$|||\boldsymbol{u} - \boldsymbol{u}_h|||^2 + |||\boldsymbol{y}^* - \boldsymbol{y}_h|||_*^2 = \eta^2(\boldsymbol{u}_h, \boldsymbol{y}_h) \quad \forall \boldsymbol{y}_h \in Q(\boldsymbol{f}, \boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h). \tag{32}$$

Proof. From the weak formulatin (7), from the facts that $\boldsymbol{f} - \boldsymbol{C}\boldsymbol{u} \in L^2(\Omega)$ and from the definition of the distributional divergence, we immediately conclude that $\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N$. Hence the traces $(\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u})\nu \in L^2(\Gamma_{\mathrm{N}})$ are well defined and we have $(\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u})\nu = \boldsymbol{g}_{\mathbf{N}} - \boldsymbol{\alpha}\boldsymbol{u}$, see (6). Using $\boldsymbol{y} = \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}$ in (16)–(18) we obtain

$$\boldsymbol{r}^* = \mathbb{A}\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h), \quad \boldsymbol{r}_\Omega = \boldsymbol{C}(\boldsymbol{u} - \boldsymbol{u}_h), \quad \boldsymbol{r}_{\mathrm{N}} = \boldsymbol{\alpha}(\boldsymbol{u} - \boldsymbol{u}_h), \tag{33}$$

and, hence, $\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u} \in Q(\boldsymbol{f}, \boldsymbol{u}_h)$ and $\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u} \in G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h)$ clearly hold. In addition, relations (33) immediately yield

$$\eta(\boldsymbol{u}_h, \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}) = |||\boldsymbol{u} - \boldsymbol{u}_h|||. \tag{34}$$

Due to (19), we see that $\boldsymbol{y} = \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u} \in Q(\boldsymbol{f}, \boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h)$ is a minimizer of $\eta(\boldsymbol{u}_h, \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u})$. Thanks to the equivalence of problems (25) and (26) we conclude that the complementary solution to (26) is $\boldsymbol{y}^* = \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}$.

Equality (30) was already shown in (34). Equality (31) can be shown similarly. Indeed, if we use $\boldsymbol{u}_h = \boldsymbol{u}$ and $\boldsymbol{y} = \boldsymbol{y}_h$ in (16)–(18) we find that

$$\begin{aligned}
\boldsymbol{r}^* &= \boldsymbol{y}_h - \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u} = \boldsymbol{y}_h - \boldsymbol{y}^*, \\
\boldsymbol{r}_\Omega &= \boldsymbol{f} - \boldsymbol{C}\boldsymbol{u} + \mathbf{div}\, \boldsymbol{y}_h = -\mathbf{div}(\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}) + \mathbf{div}\, \boldsymbol{y}_h = \mathbf{div}(\boldsymbol{y}_h - \boldsymbol{y}^*), \\
\boldsymbol{r}_{\mathrm{N}} &= \boldsymbol{g}_{\mathbf{N}} - \boldsymbol{\alpha}\boldsymbol{u} - \boldsymbol{y}_h\nu = (\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u})\nu - \boldsymbol{y}_h\nu = (\boldsymbol{y}^* - \boldsymbol{y}_h)\nu.
\end{aligned}$$

These relations together with the definition of the complementary energy norm immediately prove (31).

Finally, the relation (32) can be verified by a direct inspection. □

In the context of Theorem 4.1, we point out two important special cases. First, if tensors $\boldsymbol{C}$ and $\boldsymbol{\alpha}$ are nonsingular then $Q(\boldsymbol{f}, \boldsymbol{u}_h) \cap G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h) = [\mathbf{H}(\mathrm{div}, \Omega)]^N$, Moore-Penrose pseudoinverse of these tensors turns into usual inverse and the error estimate assumes more-less simple form. Second, if both $\boldsymbol{C}$ and $\boldsymbol{\alpha}$ vanish then $Q(\boldsymbol{f}, \boldsymbol{u}_h) = \{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N : \boldsymbol{f} + \mathbf{div}\, \boldsymbol{y} = 0 \text{ a.e. in } \Omega\}$, $G(\boldsymbol{g}_{\mathbf{N}}, \boldsymbol{u}_h) = \{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^N : \boldsymbol{y}\nu = \boldsymbol{g}_{\mathbf{N}} \text{ a.e. on } \Gamma_{\mathrm{N}}\}$, and the Moore-Penrose pseudoinverse is not needed. This case is well treatable, especially if $\boldsymbol{f}$ and $\boldsymbol{g}_{\mathbf{N}}$ are simple functions, e.g. constants, see e.g. [12, 28, 29].

5. Linear elasticity system

In this section we briefly mention how the system of linear elasticity fits into the general setting of elliptic systems. For simplicity, we restrict ourselves to two-dimensional problem, i.e., $d = 2$. The classical formulation of the problem of elasticity for an elastic body $\Omega \subset \mathbb{R}^2$, reads as follows: find the displacement $\boldsymbol{u}$ such that

$$-\operatorname{\mathbf{div}} \boldsymbol{\sigma}(\boldsymbol{u}) = \boldsymbol{f}^{\boldsymbol{u}} \quad \text{in } \Omega, \tag{35}$$

$$\boldsymbol{u} = \boldsymbol{g} \quad \text{on } \Gamma_{\mathrm{D}}^{\boldsymbol{u}}, \tag{36}$$

$$\boldsymbol{\sigma}(\boldsymbol{u})\nu = \boldsymbol{t} \quad \text{on } \Gamma_{\mathrm{N}}^{\boldsymbol{u}}. \tag{37}$$

The meaning of the above symbols is standard. The stress tensor $\boldsymbol{\sigma}(\boldsymbol{u}) \in \mathbb{R}^{2\times 2}$ is defined by

$$\boldsymbol{\sigma}(\boldsymbol{u}) = 2\mu\boldsymbol{\epsilon}(\boldsymbol{u}) + \lambda(\operatorname{div} \boldsymbol{u})\boldsymbol{I}$$

with $\boldsymbol{\epsilon}(\boldsymbol{u}) = \frac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u} + (\boldsymbol{\nabla}\boldsymbol{u})^T)$ denoting the symmetric gradient of $\boldsymbol{u}$ and $\boldsymbol{I}$ being the 2×2 identity matrix. Further, $\lambda \in \mathbb{R}$ and $\mu \in \mathbb{R}$ are the Lamé constants, $\boldsymbol{f}^{\boldsymbol{u}} \in \mathbb{R}^2$ is the density of the volume force, $\boldsymbol{g} \in \mathbb{R}^2$ is the prescribed displacement on the part of the boundary $\Gamma_{\mathrm{D}}^{\boldsymbol{u}}$, and $\boldsymbol{t} \in \mathbb{R}^2$ is the traction on the part of the boundary $\Gamma_{\mathrm{N}}^{\boldsymbol{u}}$.

Elasticity problem (35)–(37) can be seen as a special case of the general elliptic system (4)–(6). Indeed, setting $N = 2$,

$$\mathbb{A}_{1111} = \mathbb{A}_{2222} = 2\mu+\lambda, \quad \mathbb{A}_{1122} = \mathbb{A}_{2211} = \lambda, \quad \mathbb{A}_{1212} = \mathbb{A}_{1221} = \mathbb{A}_{2112} = \mathbb{A}_{2121} = \mu$$

and the other entries of $\mathbb{A}$ as zeros then $\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u} = \mathbb{A}\boldsymbol{\epsilon}(\boldsymbol{u}) = \boldsymbol{\sigma}(\boldsymbol{u})$. Further, putting $\mathbb{B} = 0$, $\boldsymbol{C} = 0$, $\boldsymbol{\alpha} = 0$, $\boldsymbol{f} = \boldsymbol{f}^{\boldsymbol{u}}$, $\boldsymbol{g}_{\mathbf{D}} = \boldsymbol{g}$, $\Gamma_{\mathrm{D}} = \Gamma_{\mathrm{D}}^{\boldsymbol{u}}$, $\boldsymbol{g}_{\mathbf{N}} = \boldsymbol{t}$, and $\Gamma_{\mathrm{N}} = \Gamma_{\mathrm{N}}^{\boldsymbol{u}}$, the general system (4)–(6) transforms to (35)–(37).

The bilinear form (8) and the linear functional (9) are then

$$\mathcal{B}^{\boldsymbol{u}}(\boldsymbol{u}, \boldsymbol{v}) = (\mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}, \boldsymbol{\nabla}\boldsymbol{v}) = (\mathbb{A}\boldsymbol{\epsilon}(\boldsymbol{u}), \boldsymbol{\epsilon}(\boldsymbol{v})) = (\boldsymbol{\sigma}(\boldsymbol{u}), \boldsymbol{\epsilon}(\boldsymbol{v})),$$
$$\mathcal{F}^{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{f}^{\boldsymbol{u}}, \boldsymbol{v}) + \langle \boldsymbol{t}, \boldsymbol{v} \rangle$$

and the energy norm $\|\!|\boldsymbol{u}|\!\|_{\boldsymbol{u}}^2 = \mathcal{B}^{\boldsymbol{u}}(\boldsymbol{u}, \boldsymbol{u}) = (\mathbb{A}\boldsymbol{\epsilon}(\boldsymbol{u}), \boldsymbol{\epsilon}(\boldsymbol{u}))$.

However, strictly speaking the elasticity problem is not elliptic – assumption (A2) is not satisfied, because the tensor $\mathbb{A}$ is singular. Indeed, the kernel of $\mathbb{A}$ consists of antisymmetric matrices:

$$\operatorname{Ker} \mathbb{A} = \mathbb{R}_{anti}^{2\times 2} = \left\{ \boldsymbol{u} = \begin{pmatrix} 0 & \xi \\ -\xi & 0 \end{pmatrix}, \quad \xi \in \mathbb{R} \right\}.$$

Theoretically, there is a simple remedy to this problem. To generalize the estimates (19), (21), and (24), we can handle the positive semidefinite tensor $\mathbb{A}$ in the same way as the positive semidefinite matrices $\boldsymbol{C} - \frac{1}{2}\boldsymbol{D}$ and $\boldsymbol{\alpha} + \frac{1}{2}\boldsymbol{E}$. In particular, we will restrict the possible complementary solutions $\boldsymbol{y}$ to those who are in the range of $\mathbb{A}$. We define

$$R = \left\{ \boldsymbol{y} \in [\mathbf{H}(\operatorname{div}, \Omega)]^2 : \boldsymbol{y} \in (\operatorname{Ker} \mathbb{A})^{\perp} \right\} = \left\{ \boldsymbol{y} \in [\mathbf{H}(\operatorname{div}, \Omega)]^2 : \boldsymbol{y} \in \mathbb{R}_{sym}^{2\times 2} \right\},$$

where $\mathbb{R}_{sym}^{2\times 2}$ stands for the space of 2×2 symmetric matrices.

In general, estimates (19), (21), and (24) remain valid even in the case of positive semidefinite tensor $\mathbb{A}$, but the inverse $\mathbb{A}^{-1}$ has to be replaced by the Moore-Penrose pseudoinverse $\mathbb{A}^{\dagger}$ and the admissible $\boldsymbol{y}$ must lie in R.

In the case of linear elasticity, the Moore-Penrose pseudoinverse $\mathbb{A}^{\dagger}$ can be expressed as $\mathbb{A}^{\dagger} = (4\mu(\mu + \lambda))^{-1}\mathbb{M}$, where

$$\mathbb{M}_{1111} = \mathbb{M}_{2222} = 2\mu + \lambda, \qquad \mathbb{M}_{1122} = \mathbb{M}_{2211} = -\lambda,$$
$$\mathbb{M}_{1212} = \mathbb{M}_{1221} = \mathbb{M}_{2112} = \mathbb{M}_{2121} = \mu + \lambda$$

and the remaining entries of $\mathbb{M}$ vanish.

Since the convection and reaction coefficients are not present in the linear elasticity system, estimate (19) collapses to (24). For the singular tensor $\mathbb{A}$ of the linear elasticity coefficients we obtain

$$|\!|\!|\boldsymbol{u} - \boldsymbol{u}_h|\!|\!|_{\boldsymbol{u}} \leq \widetilde{\eta}_{\boldsymbol{u}}(\boldsymbol{u}_h, \boldsymbol{y}),$$

where

$$\widetilde{\eta}_{\boldsymbol{u}}(\boldsymbol{u}_h, \boldsymbol{y}) = \|\boldsymbol{y} - \mathbb{A}\boldsymbol{\nabla}\boldsymbol{u}_h\|_{\mathbb{A}^\dagger} \quad \text{for all } \boldsymbol{u}_h \in V \text{ and } \boldsymbol{y} \in R \cap Q_0(\boldsymbol{f}^{\boldsymbol{u}}, \boldsymbol{u}_h) \cap G_0(\boldsymbol{t}, \boldsymbol{u}_h),$$

$Q_0(\boldsymbol{f}^{\boldsymbol{u}}, \boldsymbol{u}_h) = \{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^2 : \boldsymbol{f}^{\boldsymbol{u}} + \mathbf{div}\,\boldsymbol{y} = 0 \text{ a.e. in } \Omega\}$, and $G_0(\boldsymbol{t}, \boldsymbol{u}_h) = \{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^2 : \boldsymbol{y}\nu = \boldsymbol{t} \text{ a.e. on } \Gamma_{\mathrm{N}}\}$.

The complementary problem (26) for linear elasticity system (35)–(37) reads as follows: find $\boldsymbol{y} \in R \cap Q_0(\boldsymbol{f}^{\boldsymbol{u}}, \boldsymbol{u}_h) \cap G_0(\boldsymbol{t}, \boldsymbol{u}_h)$ such that

$$\frac{1}{2\mu}(\boldsymbol{y}, \boldsymbol{w}) - \frac{\lambda}{4\mu(\mu + \lambda)}(\mathrm{tr}\,\boldsymbol{y}, \mathrm{tr}\,\boldsymbol{w}) = (\boldsymbol{\epsilon}(\boldsymbol{u}_h), \boldsymbol{w}) \quad \forall \boldsymbol{w} \in R \cap Q_0(0,0) \cap G_0(0,0).$$

Notice that $Q_0(0,0) = \{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^2 : \mathbf{div}\,\boldsymbol{y} = 0 \text{ a.e. in } \Omega\}$, and $G_0(0,0) = \{\boldsymbol{y} \in [\mathbf{H}(\mathrm{div}, \Omega)]^2 : \boldsymbol{y}\nu = 0 \text{ a.e. on } \Gamma_{\mathrm{N}}\}$. We point out that of the space $R \cap Q_0(0,0) \cap G_0(0,0)$ corresponds to symmetric tensors with vanishing divergence in Ω and with vanishing normal components on Γ_{N}. This space might be practically problematic to handle.

Another possibility is to use a variant of (21). This approach is treated in detail in [16, 20].

6. Application to chemical systems

The guaranteed upper bounds derived in Sections 3 for a general elliptic problem can be directly applied to the diffusion–convection–reaction problem in chemistry. The same approach might be equally well applied to the modeling of air pollution in the presence of convection (the wind) and chemical reactions between various pollutants. A typical example is the traffic pollution. The exhalations from the combustion engines undergo various chemical reactions in the air. These reactions have various rates and the result might be high concentrations of pollutants (e.g. of ozone) quite far away from the original source.

The general elliptic system (4)–(6) describes the steady state concentrations $u^1, u^2, \dots, u^N$ of chemical species $S_1, S_2, \dots, S_N$, which undergo the following chemical reactions:

$$\emptyset \xrightarrow{q^i} S_i, \qquad i = 1, 2, \dots, N, \tag{38}$$

$$S_i \xleftrightarrow{k^{ij}} S_j, \qquad i, j = 1, 2, \dots, N, \quad i \neq j, \tag{39}$$

$$S_i \xrightarrow{k^{ii}} \emptyset, \qquad i = 1, 2, \dots, N. \tag{40}$$

Reaction (38) is the production of S_i with the constant rate q^i. Reaction (39) is the conversion of S_i to S_j and vice versa. Both directions have the same rate constant k^{ij}. Reaction (40) is the degradation of S_i with the rate constant k^{ii}. All these rate constants are nonnegative.

The concentrations u^1, u^2, …, u^N can be computed by the following diffusion-reaction-convection system

$$-\delta^i \Delta u^i + \operatorname{div}(u^i \widetilde{\boldsymbol{b}}) + \sum_{j=1}^{N} k^{ij} u^i - \sum_{j=1, j\neq i}^{N} k^{ij} u^j = q^i \quad \text{in } \Omega, \tag{41}$$

where $i = 1, 2, \dots, N$, δ^i is the diffusivity of S_i, and $\widetilde{\boldsymbol{b}}$ describes the velocity field.

This system is readily in the form (1). We have $\mathcal{A}^{ii} = \delta^i \boldsymbol{I}$ for $i = 1, 2, \dots, N$, $\mathcal{A}^{ij} = 0$ for $i \neq j$, $\boldsymbol{b}^{ii} = \widetilde{\boldsymbol{b}}$ and $\boldsymbol{b}^{ij} = 0$ for $i \neq j$, $i, j = 1, 2, \dots, N$, $c^{ii} = \operatorname{div} \widetilde{\boldsymbol{b}} + \sum_{j=1}^{N} k^{ij}$ for $i = 1, 2, \dots, N$, and $c^{ij} = -k^{ij}$ for $i \neq j$.

If the diffusivity coefficients δ^i do not vanish then the corresponding tensor $\mathbb{A}$ is diagonal and invertible. The matrices $\boldsymbol{D} = (\operatorname{div} \widetilde{\boldsymbol{b}})\boldsymbol{I}$ and $\boldsymbol{E} = (\widetilde{\boldsymbol{b}} \cdot \nu)\boldsymbol{I}$ are just multiples of the identity matrix $\boldsymbol{I}$ in this case. If $k^{ii} + \frac{1}{2} \operatorname{div} \widetilde{\boldsymbol{b}} > 0$ for all $i = 1, 2, \dots, N$ then the matrix $\boldsymbol{C} - \frac{1}{2}\boldsymbol{D}$ is diagonally dominant. Hence, since it is symmetric, it is positive definite. This elliptic system then satisfies assumptions (A1)–(A4) and we can directly apply the presented guaranteed upper bounds.

7. Numerical examples

In this section we present two numerical examples. In the first example the exact solution is known and we test the sharpness of the guaranteed upper bound (19). In the second example the exact solution is unknown and we present an adaptive procedure which together with the guaranteed upper bounds enables to compute the solution with guaranteed accuracy.

Example 1. Let us consider an elliptic system of $N = 3$ equations in $d = 2$ dimensions in the form (4)–(6). The domain $\Omega = (0, 3/2) \times (0, 1)$ is a rectangle. The diffusion terms consist of Laplacians, i.e., $\mathbb{A}_{ikj\ell} = \mathcal{A}^{ij}_{k\ell}$ with $\mathcal{A}^{ij} = \boldsymbol{I}$, $i, j = 1, 2, 3$, $k, \ell = 1, 2$. The velocity field $\widetilde{\boldsymbol{b}}(x_1, x_2) = \rho^2(x_2(1 - x_2), 0)^T$ is divergence free and $\mathbb{B}_{iik} = \widetilde{\boldsymbol{b}}_k$ and $\mathbb{B}_{ijk} = 0$ for $i \neq j$, $i, j = 1, 2, 3$, $k = 1, 2$. The constant ρ is a parameter. The reaction coefficients are given by the following matrix

$$\boldsymbol{C} = \kappa^2 \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix},$$

where the constant κ is the second parameter. We prescribe the Dirichlet boundary conditions on the edge $x_1 = 0$ and the Neumann boundary conditions on the remaining part of $\partial\Omega$. The functions $\boldsymbol{f}$, $\boldsymbol{g}_{\mathbf{D}}$, and $\boldsymbol{g}_{\mathbf{N}}$ are chosen in such a way that the exact solution is $\boldsymbol{u} = u\mathbf{1}$, where $u(x_1, x_2) = \sin(\pi x_1)\cos(\pi x_2)$ and $\mathbf{1} = (1, 1, 1)^T$. Thus, $\boldsymbol{g}_{\mathbf{D}} = 0$, $\boldsymbol{g}_{\mathbf{N}} = 0$ on edges $x_2 = 0$ and $x_2 = 1$, $\boldsymbol{g}_{\mathbf{N}} = -\pi \sin(\pi x_1)$ on the edge $x_1 = 3/2$, and $\boldsymbol{f} = (2\pi^2 + \kappa^2)\boldsymbol{u} + \rho^2(\widetilde{\boldsymbol{b}} \cdot \nabla u)\mathbf{1}$.

We solve this problem by the lowest-order finite element method to obtain a piecewise linear approximation of $\boldsymbol{u}_h$. The used triangular mesh is shown in Figure 1 (right). We use the guaranteed upper bound $\eta(\boldsymbol{u}_h, \boldsymbol{y}_h)$ given by (20). The complementary solution $\boldsymbol{y}_h$ is computed as the finite element approximation of the complementary problem (26). We use piecewise linear

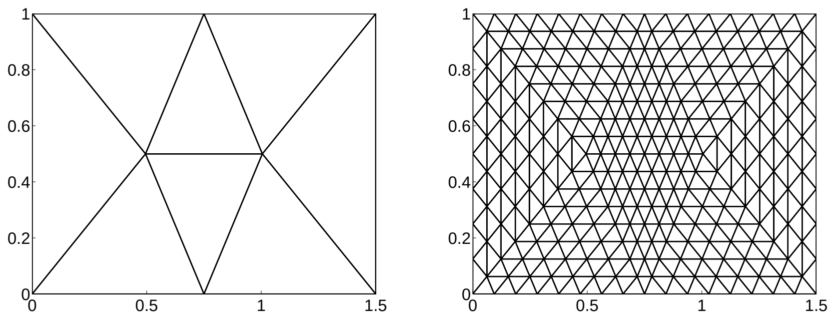

FIGURE 1. The initial finite element mesh (left) and its uniform refinement (right).

ρ	$\kappa=1$	$\kappa=5$	$\kappa=10$	$\kappa=15$	$\kappa=20$	$\kappa=25$
0	8.062	1.459	1.069	1.025	1.014	1.009
5	7.986	1.498	1.095	1.037	1.020	1.013
10	16.150	2.530	1.386	1.136	1.062	1.034
15	29.373	4.375	2.121	1.446	1.198	1.101
20	42.600	6.256	3.042	1.967	1.476	1.249
25	55.490	8.145	3.940	2.580	1.874	1.492

TABLE 1. Index of effectivity I_{eff} for the error bound (19) for various ρ and κ (Example 1).

approximation $\boldsymbol{y}_h$ on the same mesh. Since the exact solution is known, we can test the sharpness of the computed upper bounds $\eta(\boldsymbol{u}_h, \boldsymbol{y}_h)$. More precisely, we evaluate the index of effectivity $I_{\text{eff}} = \eta(\boldsymbol{u}_h, \boldsymbol{y}_h)/|\!|\!|\boldsymbol{u}-\boldsymbol{u}_h|\!|\!|$. The results are presented in Table 1 for various values of κ and ρ.

In [2] and [29] we investigate the complementary error estimates for a scalar diffusion-reaction problem. These results show that the upper bound (20) provides sharp results if the reaction term dominates the diffusion. Results in Table 1 confirm that this is the case even for systems and for problems with convection. The results clearly show that the error estimate (20) is very sharp for elliptic systems if the reaction term dominates both the diffusion and convection. On the other hand, if reaction does not dominate then the estimate gives quite inaccurate results. To obtain sharp results even for small values of the reaction coefficient κ, it is necessary to implement the error bound (21) or (24).

We also point out that all indices of effectivity in Table 1 are greater or equal to one. Hence, as Lemma 3.2 predicts, the computed error estimates are really greater than the energy norm of the error.

Example 2. Let us consider a system of three chemical reactions of the form (38)–(40) with the rate constants $k^{ij} = \kappa^2$ and $q^i = \kappa^2/2$ for $i,j = 1,2,3$. The diffusivity is considered as $\delta^i = 1$ for $i = 1,2,3$. The domain Ω and the velocity field $\widetilde{\boldsymbol{b}}$ are considered the same as in Example 1. Using these data in (41) and translating them to the form (4)–(6), we obtain the same $\mathbb{A}$, $\mathbb{B}$, and $\boldsymbol{C}$ as in Example 1. On the other hand, the source terms $\boldsymbol{f}$, $\boldsymbol{g}_{\mathbf{D}}$, and $\boldsymbol{g}_{\mathbf{N}}$ differ. The production coefficients q^i yield $\boldsymbol{f} = (\kappa^2/2)\mathbf{1}$. On the Dirichlet edge $x_1 = 0$ we prescribe

$$g_{\mathrm{D}}^i(x_1, x_2) = \exp\left(-9(4x_2 - i)^2\right), \quad i = 1, 2, 3.$$

These are Gaussian functions with peaks at $x_2 = i/4$, $i = 1, 2, 3$. On the remaining parts of the boundary $\partial\Omega$ we consider homogeneous Neumann boundary conditions:

$$(\boldsymbol{\nabla u})\nu = 0.$$

The goal of the presented numerical computations is to obtain an approximate solution with relative error at most 5 %. This can be achieved by a standard adaptive procedure, in combination with the guaranteed upper bound (20). A general adaptive procedure follows these steps, see e.g. [7]:

(1) Construct the initial mesh $\mathcal{T}_h$.
(2) Find the finite element solution $\boldsymbol{u}_h$ on $\mathcal{T}_h$.
(3) Find the error indicators η_K for all elements $K \in \mathcal{T}_h$.
(4) Stop, if $\eta^2 = \sum_{K\in\mathcal{T}_h} \eta_K^2$ is under the prescribed tolerance.
(5) Mark those elements for which $\eta_K < \theta \max_{K\in\mathcal{T}_h} \eta_K$.
(6) Refine the marked elements and create a new mesh $\mathcal{T}_h$.
(7) Go to 2.

The guaranteed upper bound (20) can be well used both as the local error indicators η_K and the global error estimator η. Indeed, $\eta^2(\boldsymbol{u}_h, \boldsymbol{y}_h) = \sum_{K\in\mathcal{T}_h} \eta_K^2(\boldsymbol{u}_h, \boldsymbol{y}_h)$, where

$$\eta_K^2(\boldsymbol{u}_h, \boldsymbol{y}_h) = \|\boldsymbol{r}^*\|^2_{\mathbb{A}^{-1},K} + \|\boldsymbol{r}_\Omega\|^2_{(\boldsymbol{C}-\frac{1}{2}\boldsymbol{D})^\dagger,K} + \langle|\boldsymbol{r}_{\mathrm{N}}|\rangle^2_{(\boldsymbol{\alpha}+\frac{1}{2}\boldsymbol{E})^\dagger,K}$$

and the local (semi)norms are defined in a natural way as

$$\|\boldsymbol{v}\|^2_{\boldsymbol{M},K} = \int_K \boldsymbol{v}^T \boldsymbol{M} \boldsymbol{v} \,\mathrm{d}x \quad \text{and} \quad \langle|\boldsymbol{v}|\rangle^2_{\boldsymbol{M},K} = \int_{\Gamma_{\mathrm{N}}\cap\partial K} \boldsymbol{v}^T \boldsymbol{M} \boldsymbol{v} \,\mathrm{d}x.$$

As above, the complementary solution $\boldsymbol{y}_h$ is computed by the finite element method using the piecewise linear approximation on the same mesh as for $\boldsymbol{u}_h$.

The initial mesh for the adaptive procedure is depicted in Figure 1 (left). The relative error tolerance of 5 % was met in the seventh adaptive step. More precisely, after seven adaptive steps we obtained the ratio $\eta(\boldsymbol{u}_h, \boldsymbol{y}_h)/|\!|\!|\boldsymbol{u}_h|\!|\!|$ less then 5 %. Thus, Lemma 3.2 guarantees that the energy norm of the true relative error $|\!|\!|\boldsymbol{u} - \boldsymbol{u}_h|\!|\!|/|\!|\!|\boldsymbol{u}_h|\!|\!|$ is below the prescribed tolerance.

The three components of the finite element solution $\boldsymbol{u}_h$ and the adapted mesh in the final adaptive step are shown in Figure 2. As expected, the mesh is refined close to the edge $x_1 = 0$, where the solution possesses steep gradients. On the other hand, the solution is almost constant in the opposite half of the domain and we observe very coarse mesh there. This, confirms that the approximate complementary solution $\boldsymbol{y}_h$ used in $\eta_K(\boldsymbol{u}_h, \boldsymbol{y}_h)$ provides quality indicator of the local behaviour of the error.

8. Conclusions

In this paper we generalized the complementary a posteriori error estimates to systems of linear elliptic problems. We introduced three variants of these error estimates: (19), (21), and (24). We proved the upper bound property for these variants, we derived the corresponding complementary

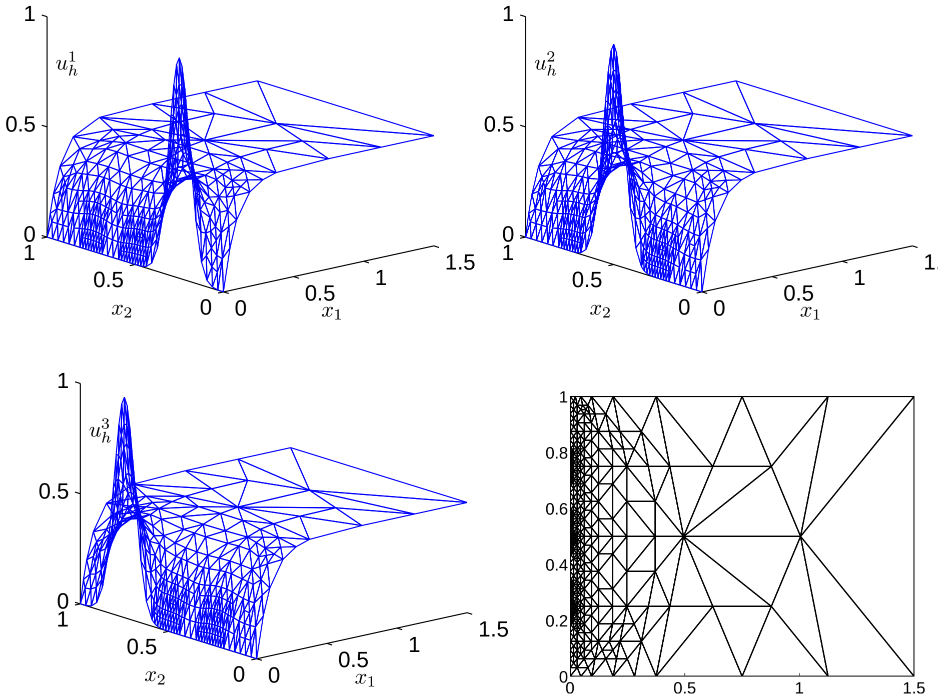

FIGURE 2. The three components of the finite element solution $\boldsymbol{u}_h$ at the last adaptive step and the corresponding adapted mesh (Example 2).

systems, and for the case (19) we proved properties (29)–(32). All these properties are analogous to the scalar case, see [29]. However, a nontrivial feature of systems which is not present in the scalar case is the fact that the equation coefficients might be nonzero and singular matrices. This technical difficulty was solved by the use the Moore-Penrose pseudoinverse.

There are also other properties of the complementary a posteriori error estimates known from the scalar case which are not treated in this paper. For example, if the coefficients $\mathbb{B}$, $\boldsymbol{C}$, and $\boldsymbol{\alpha}$ vanish the upper bound (24) possesses properties analogous to those listed in Theorem 4.1. Another result known from the scalar case is the so-called method of hypercircle. It enables to construct an approximation whose error is known exactly [4, 14, 27, 29]. It is very likely that all these results and properties generalize to systems as well.

Further, we point out that the approximate complementary solution $\boldsymbol{y}_h$ was computed as finite element approximation of the corresponding complementary problem. This approach is not practical due to its high computational cost. If we use the same mesh for both primal and complementary problem, we need several times more degrees of freedom to solve the complementary problem than the primal one. Computationally cheap approximate complementary solution $\boldsymbol{y}_h$ can be found by suitable postprocessing of $\boldsymbol{u}_h$ and its gradient. One possibility is the method of equilibrated residuals [1]. This method was employed in [2] and a fast, robust, and guaranteed upper bound for a scalar diffusion-reaction problem was derived there. Generalization of this result to systems of elliptic equations is possible as well.

Anyway, the presented numerical experiments show the capability of the complementary error estimates to provide sharp upper bounds even for elliptic systems at least if the reaction term dominates. In addition, the experiments confirm that the localized version of the complementary bounds may serve as precise local error indicators for guidance of the adaptive process.

Furthermore, an efficient software for solution of systems of partial differential equations has to approximate each component of solution on its own, individually adapted mesh. An automatically *hp*-adaptive strategy of this kind is developed in [24, 25]. We stress that the complementary error estimates can be well used even in this case. Evenmore, the complementary approach is completely independent from the way how the approximate solution $\boldsymbol{u}_h$ is obtained and the complementary error estimates are valid for arbitrary approximation of the complementary solution $\boldsymbol{y}_h$. This raises a question suitable for further research: how to construct an optimal $\boldsymbol{y}_h$ yielding sharp, robust, and fast complementary error bounds provided $\boldsymbol{u}_h$ has been computed in a particular way.

Acknowledgement

The author has been supported by Grant no. IAA100760702 of the Grant Agency of the Academy of Sciences of the Czech Republic and by the institutional research plan no. AV0Z10190503 of the Academy of Sciences of the Czech Republic. This support is gratefully acknowledged.

References

[1] M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000.

[2] M. Ainsworth and T. Vejchodský, *Fully computable robust a posteriori error bounds for singularly perturbed reaction–diffusion problems*, Numer. Math., (2010). submitted.

[3] I. Anjam, O. Mali, A. Muzalevsky, P. Neittaanmäki, and S. Repin, *A posteriori error estimates for a Maxwell type problem*, Russian J. Numer. Anal. Math. Modelling, 24 (2009), pp. 395–408.

[4] J. P. Aubin and H. G. Burchard, *Some aspects of the method of the hypercircle applied to elliptic variational problems*, in Numerical Solution of Partial Differential Equations, II (SYNSPADE 1970) (Proc. Sympos., Univ. of Maryland, College Park, Md., 1970), Academic Press, New York, 1971, pp. 1–67.

[5] I. Cheddadi, R. Fučík, M. I. Prieto, and M. Vohralík, *Guaranteed and robust a posteriori error estimates for singularly perturbed reactiondiffusion problems*, M2AN Math. Model. Numer. Anal., 43 (2009), pp. 867–888.

[6] R. Dautray and J.-L. Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 2*, Springer-Verlag, Berlin, 1988. Functional and variational methods, With the collaboration of Michel Artola, Marc Authier, Philippe Bénilan, Michel Cessenat, Jean Michel Combes, Hélène Lanchon, Bertrand Mercier, Claude Wild and Claude Zuily, Translated from the French by Ian N. Sneddon.

[7] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.

[8] J. Haslinger and I. Hlaváček, *Convergence of a finite element method based on the dual variational formulation*, Apl. Mat., 21 (1976), pp. 43–65.

[9] I. Hlaváček, *Some equilibrium and mixed models in the finite element method*, in Mathematical models and numerical methods (Papers, Fifth Semester, Stefan Banach Internat. Math. Center, Warsaw, 1975), vol. 3 of Banach Center Publ., PWN, Warsaw, 1978, pp. 147–165.

[10] I. Hlaváček and M. Křížek, *Internal finite element approximations in the dual variational method for second order elliptic problems with curved boundaries*, Apl. Mat., 29 (1984), pp. 52–69.

[11] S. Korotov, *Two-sided a posteriori error estimates for linear elliptic problems with mixed boundary conditions*, Appl. Math., 52 (2007), pp. 235–249.

[12] M. Křížek, *Conforming equilibrium finite element methods for some elliptic plane problems*, RAIRO Anal. Numér., 17 (1983), pp. 35–65.

[13] M. Křížek and P. Neittaanmäki, *Finite element approximation of variational problems and applications*, vol. 50 of Pitman Monographs and Surveys in Pure and Applied Mathematics, Longman Scientific & Technical, Harlow, 1990.

[14] M. Křížek and P. Neittaanmäki, *Mathematical and numerical modelling in electrical engineering*, vol. 1 of Mathematical Modelling: Theory and Applications, Kluwer Academic Publishers, Dordrecht, 1996. Theory and applications, With a foreword by Ivo Babuška.

[15] A. V. Muzalevskiĭ and S. I. Repin, *On error estimates for approximate solutions in problems of the linear theory of thermoelasticity*, Izv. Vyssh. Uchebn. Zaved. Mat., (2005), pp. 64–72.

[16] A. V. Muzalevsky and S. I. Repin, *On two-sided error estimates for approximate solutions of problems in the linear theory of elasticity*, Russian J. Numer. Anal. Math. Modelling, 18 (2003), pp. 65–85.

[17] J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Masson et Cie, Éditeurs, Paris, 1967.

[18] P. Neittaanmäki and S. Repin, *Reliable methods for computer simulation*, vol. 33 of Studies in Mathematics and its Applications, Elsevier Science B.V., Amsterdam, 2004. Error control and a posteriori estimates.

[19] S. Repin, *On a posteriori error estimates for the stationary Navier-Stokes problem*, J. Math. Sci. (N. Y.), 150 (2008), pp. 1885–1889. Problems in mathematical analysis. No. 36.

[20] S. Repin, *A posteriori estimates for partial differential equations*, vol. 4 of Radon Series on Computational and Applied Mathematics, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

[21] S. Repin and S. Sauter, *Functional a posteriori estimates for the reaction-diffusion problem*, C. R. Math. Acad. Sci. Paris, 343 (2006), pp. 349–354.

[22] S. Repin and J. Valdman, *Functional a posteriori error estimates for problems with nonlinear boundary conditions*, J. Numer. Math., 16 (2008), pp. 51–81.

[23] S. I. Repin, *A posteriori error estimation for variational problems with uniformly convex functionals*, Math. Comp., 69 (2000), pp. 481–500.

[24] P. Šolín, L. Dubcová, and J. Kruis, *Adaptive hp-FEM with dynamical meshes for transient heat and moisture transfer problems*, J. Comput. Appl. Math., 233 (2010), pp. 3103–3112.

[25] P. Šolín, K. Segeth, and I. Doležel, *Space-time adaptive hp-FEM: methodology overview*, in Programs and algorithms of numerical mathematics 14, Acad. Sci. Czech Repub. Inst. Math., Prague, 2008, pp. 185–200.

[26] J. L. Synge, *The hypercircle in mathematical physics: a method for the approximate solution of boundary value problems*, Cambridge University Press, New York, 1957.

[27] J. Vacek, *Dual variational principles for an elliptic partial differential equation*, Apl. Mat., 21 (1976), pp. 5–27.

[28] T. Vejchodský, *Guaranteed and locally computable a posteriori error estimate*, IMA J. Numer. Anal., 26 (2006), pp. 525–540.

[29] T. Vejchodský, *Complementarity based a posteriori error estimates and their properties*, Math. Comput. Simulation, (2010). submitted.

[30] M. Vohralík, *A posteriori error estimation in the conforming finite element method based on its local conservativity and using local minimization*, C. R. Math. Acad. Sci. Paris, 346 (2008), pp. 687–690.

Institute of Mathematics, Czech Academy of Sciences, Žitná 25, CZ–115 67 Prague 1, Czech Republic

E-mail address: vejchod@math.cas.cz