ERROR ESTIMATES BASED ON ADJOINT PROBLEM.
ERROR CONTROL USING POST-PROCESSING.

Sergey Korotov

Basque Center for Applied Mathematics / IKERBASQUE

`http://www.bcamath.org` & `http://www.ikerbasque.net`

PRELIMINARIES

- *In this section we recall some material from previous lectures.*

Model Problem

Find a function u such that

$$-\mathrm{div}\,(\mathcal{A}\,\nabla u) = f \quad \text{in } \Omega \tag{1}$$

$$u = u_0 \quad \text{on } \Gamma_D \tag{2}$$

$$\mathcal{A}\,\nabla u \cdot n = g \quad \text{on } \Gamma_N \tag{3}$$

- We assume that $a_{ij} \in L^\infty(\Omega)$, $f \in L^2(\Omega)$, $g \in L^2(\Gamma_N)$ and $u_0 \in H^1(\Omega)$. Moreover let for a.a. $x \in \Omega$ one has $a_{ij}(x) = a_{ji}(x)$ and there exist constants $C_1, C_2 > 0$ such that

$$C_1\|\xi\|^2 \le \mathcal{A}(x)\xi \cdot \xi \le C_2\|\xi\|^2 \quad \forall \xi \in \mathbf{R}^d$$

Weak Formulation

We introduce the space of test functions

$$V = \{ w \in H^1(\Omega) \mid w = 0 \text{ on } \Gamma_D \}$$

and define the bilinear and linear forms

$$a(v, w) = \int_\Omega \mathcal{A} \nabla w \cdot \nabla v \, dx, \quad F(v) = \int_\Omega f v \, dx + \int_{\Gamma_N} g v \, ds.$$

Definition: Let $\Omega \in \mathcal{L}$ and $\Gamma_D \neq \emptyset$. A function $u \in H^1(\Omega)$ is called a weak (or generalized) solution of the problem (1)–(3) if

$$u - u_0 \in V \qquad \text{and} \qquad a(w, u) = F(w) \quad \forall w \in V$$

Theorem: Let $\Omega \in \mathcal{L}$ and $\Gamma_D \neq \emptyset$. Then there exists exactly one weak solution $u \in H^1(\Omega)$ of the problem (1)–(3)

Equivalent Norms for Measuring Errors

- The solution $u = u_0 + u_\star$, where $u_\star \in V$. Let $\bar{u} = u_0 + \bar{u}_\star$ ($\bar{u}_\star \in V$) be an approximation of u. We want to find (or estimate) the following value $\|u - \bar{u}\|_{1,\Omega}$
- First, for $w \in H^1(\Omega)$ we have by definition

$$\|w\|_{1,\Omega}^2 := \|w\|_{0,\Omega}^2 + \|\nabla w\|_{0,\Omega}^2$$

- Also, the following Friederichs' inequality holds for $w \in V$

$$\|w\|_{0,\Omega} \leq C_{\Omega,\Gamma_D} \|\nabla w\|_{0,\Omega} \qquad (FI)$$

- The function $u_\star$ minimizes the energy functional over V

$$J(w) = \frac{1}{2} a(w, w) - f(w)$$

Now, since $u - \bar{u} \in V$ we observe that

$$\|u-\bar{u}\|^2_{1,\Omega} \leq \left(1+C^2_{\Omega,\Gamma_D}\right)\|\nabla(u-\bar{u})\|^2_{0,\Omega} \leq \frac{1+C^2_{\Omega,\Gamma_D}}{C_1}a(u-\bar{u},u-\bar{u})$$

and, also, it holds

$$\frac{1}{C_2}a(u-\bar{u},u-\bar{u}) \leq \|\nabla(u-\bar{u})\|^2_{0,\Omega} \leq \|u-\bar{u}\|^2_{1,\Omega}$$

which altogether means that the estimation of the error in the energy norm $\sqrt{a(\cdot,\cdot)}$ is equivalent to the error estimation in the norm $\|\cdot\|_{1,\Omega}$ (see also Lecture 1) with known "equivalence constants"

• Moreover, since $a(u-\bar{u},u-\bar{u}) = 2(J(\bar{u}_\star) - J(u_\star))$, and J often presents the total energy of the system, the measurement of the error in terms of the energy norm is quite natural

• From above we observe that estimation of $\|u-\bar{u}\|_{1,\Omega}$ is also equivalent to estimation of the value $\|\nabla(u-\bar{u})\|^2_{0,\Omega}$

ERROR ESTIMATION BASED ON ADJOINT PROBLEMS

- *In this section we shall give a short overview of known error estimation techniques employing the so-called adjoint problems*

Goal-Oriented Measures of Errors

• Estimation and control of the error in a suitable global norm, e.g., the estimation of $\|\nabla(u - \bar{u})\|_{0,\Omega}$ gives information on the overall quality of the approximation $\bar{u}$ and a general impression about the global behaviour of the solution u in the whole solution domain Ω

• However, it is often required to perform the error control also in another ways depending on more specific goals of the concrete compuational process or on the nature of the problem for which computer simulations are made

• For example, we may wish to control the error $u - \bar{u}$ only locally, in some small subdomains $\omega \subset \Omega$, e.g., near certain points or lines of special interest (called "zones of interest")

• Another problem arising in many real-life problems is that we do not need to know the solution over the whole domain Ω. We can only be interested e.g. in the value of some functional on this unknown solution (often called "quantity of interest"). A relevant example is presented by the J-integral conception in the fracture mechanics

• There exists an unified manner for dealing with two above mentioned problems in the error control. We introduce a suitable linear functional ℓ acting on the space of admissible functions V and associated with "zone of interest", or with "quantity of interest", and obtain an estimate for the value of $\ell(u - \bar{u})$

• The simplest example of a linear bounded functional is

$$\ell(w) = \int_\Omega \varphi w \, dx \,, \qquad w \in H^1(\Omega)$$

where $\varphi \in L^2(\Omega)$ and supp $\varphi = \omega \subset \Omega$

- Thus, the estimation of the value

$$\ell(u-\bar{u}) = \int\limits_{\Omega} \varphi(u-\bar{u})\,dx$$

where supp $\varphi = \omega \subset \Omega$, provides with certain information about the behaviour of the error $u - \bar{u}$ locally – in the subdomain ω

- Estimates for $\ell(u-\bar{u})$ can also be used for estimation of unknown *"quantity of interest"* $\ell(u)$, since

$$\ell(u) = \ell(\bar{u}) + \ell(u-\bar{u})$$

where $\ell(\bar{u})$ is computable and $\ell(u-\bar{u})$ is estimated

- Further, choosing different weight-functions φ in the above presented functional we can create a suitable set of local-type estimates which comlement the information obtained from the global error estimation

• The above presented direction in the error control is called goal-oriented error estimation, since the functional ℓ is usually taken to fit certain computational goal

• The immediate way to estimate the error in terms of the functional ℓ is to use the following inequality

$$|\ell(u - \bar{u})| \leq \|\ell\| \|u - \bar{u}\|_V$$

which can be really done if we have estimated the global error $\|u - \bar{u}\|_V$ and the norm $\|\ell\|$. But in the most of cases such an approach leads to considerable overestimation of the error

• A posteriori estimates for the error in terms of linear functionals are usually derived by employing the so-called adjoint problem, whose right-hand side is formed by the goal-functional ℓ

• Various methods using adjoint problems have been developed in well-known works by R. Rannacher, C. Johnson, R. Becker, J. T. Oden, M. Ainsworth and the other scientists

Usage of Adjoint Problem – Simplest Case

Let $\mathbf{A}$ be a positive definite $(d \times d)$ matrix, and let $\mathbf{f} \in \mathbf{R}^d$ be a given vector. Consider a system of linear algebraic equations

$$\mathbf{A}\mathbf{u} = \mathbf{f}$$

Let $\bar{\mathbf{u}}$ be an approximate solution. Assume that we want to estimate the quantity $\mathbf{l} \cdot (\mathbf{u} - \bar{\mathbf{u}})$, where $\mathbf{l}$ is a given vector. Let us define $\mathbf{v}$ as a solution of the *adjoint problem*

$$\mathbf{A}^\star \mathbf{v} = \mathbf{l}$$

where $\mathbf{A}^\star$ is adjoint to $\mathbf{A}$. Then we observe that

$$\mathbf{l} \cdot (\mathbf{u} - \bar{\mathbf{u}}) = \mathbf{A}^\star \mathbf{v} \cdot (\mathbf{u} - \bar{\mathbf{u}}) = \mathbf{v} \cdot (\mathbf{f} - \mathbf{A}\bar{\mathbf{u}})$$

- The above relation means that we can easily find the desired quantity $\mathbf{l} \cdot (\mathbf{u} - \bar{\mathbf{u}})$ if the solution of adjoint problem $\mathbf{v}$ is known

How to Use Adjoint Problem for BVPs

Consider now the case of BVP using our model problem with homogenuous Dirichlet boundary condition: Find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \mathcal{A}\nabla u \cdot \nabla w \, dx = \int_\Omega f w \, dx \qquad \forall w \in H_0^1(\Omega)$$

Assume that we want to control the computational error in terms of the following linear functional (possibly with local suport) $\ell(w) = \int\limits_\Omega \varphi w \, dx,$ where $\varphi \in L^2(\Omega)$ (and $\operatorname{supp} \varphi = \omega \subset \Omega$)

Adjoint problem reads then as follows: Find $v \in H_0^1(\Omega)$ such that

$$\int_\Omega \mathcal{A}\nabla v \cdot \nabla w \, dx = \int_\Omega \varphi w \, dx \qquad \forall w \in H_0^1(\Omega)$$

Assume now that $u_h, v_h \in H_0^1(\Omega)$ are Galerkin approximations of the original and adjoint problems, respectively, built using the finite-dimensional subspace $V_h \subset H_0^1(\Omega)$, i.e.,

$$\int_\Omega \mathcal{A}\nabla u_h \cdot \nabla w_h \, dx = \int_\Omega f w_h \, dx \qquad \forall w_h \in V$$

$$\int_\Omega \mathcal{A}\nabla v_h \cdot \nabla w_h \, dx = \int_\Omega \varphi w_h \, dx \qquad \forall w_h \in V$$

Further, we observe that

$$\int_\Omega \varphi(u-u_h) \, dx = \int_\Omega \mathcal{A}\nabla v \cdot \nabla(u-u_h) \, dx = \int_\Omega \mathcal{A}\nabla(v-v_h)\cdot\nabla(u-u_h) \, dx$$

We can write

$$\int_\Omega \mathcal{A}\nabla(v - v_h)\cdot\nabla(u - u_h)\,dx = \int_\Omega \mathcal{A}\nabla(u - u_h)\cdot\nabla(v - v_h)\,dx =$$

$$= -\sum_{i=1}^{N}\int_{\triangle_i} \mathrm{div}(\mathcal{A}\nabla(u - u_h))(v - v_h)\,dx +$$

$$+\sum_{k=1}^{K}\sum_{l<k,E_{kl}\subset\Omega}^{K}\int_{E_{kl}} j(\nu_{kl}\cdot\mathcal{A}\nabla(u - u_h))(v - v_h)\,ds \le$$

$$\le \sum_{i=1}^{N}\|f + \mathrm{div}\sigma_h\|_{0,\triangle_i}\|v - v_h\|_{0,\triangle_i} +$$

$$+\sum_{k=1}^{K}\sum_{l<k,E_{kl}\subset\Omega}^{K}\|j(\nu_{kl}\cdot\sigma_h)\|_{0,E_{kl}}\|v - v_h\|_{0,E_{kl}}$$

• Just presented estimate is similar to that one obtained by the explicit residual method, where the weights are given now by the norms of $v - v_h$

• Actually, if the solution of the adjoint problem $v \in H^2(\Omega)$ and v_h is computed via linear finite elements then

$$\|v - v_h\|_{0,\triangle_i} \leq C_{1,i}(v)h$$

and

$$\|v - v_h\|_{0,E_{kl}} \leq C_{2,kl}(v)h^{1/2}$$

• It is not easy to compute the constants $C_{1,i}, C_{2,kl}$ in the above

• For more details and another variants of the approach see the works of the above-mentioned authors

ERROR ESTIMATION METHODS

BASED ON POST-PROCESSING

Post-Processing

• Plenty of interesting material on various post-processing procedures can be found in Chapter 8 of the monograph "Finite Element Approximation of Variational Problems and Applications" (1990) by M. Křížek and P. Neittaanmäki and also in these authors' papers on superconvergence and gradient averaging

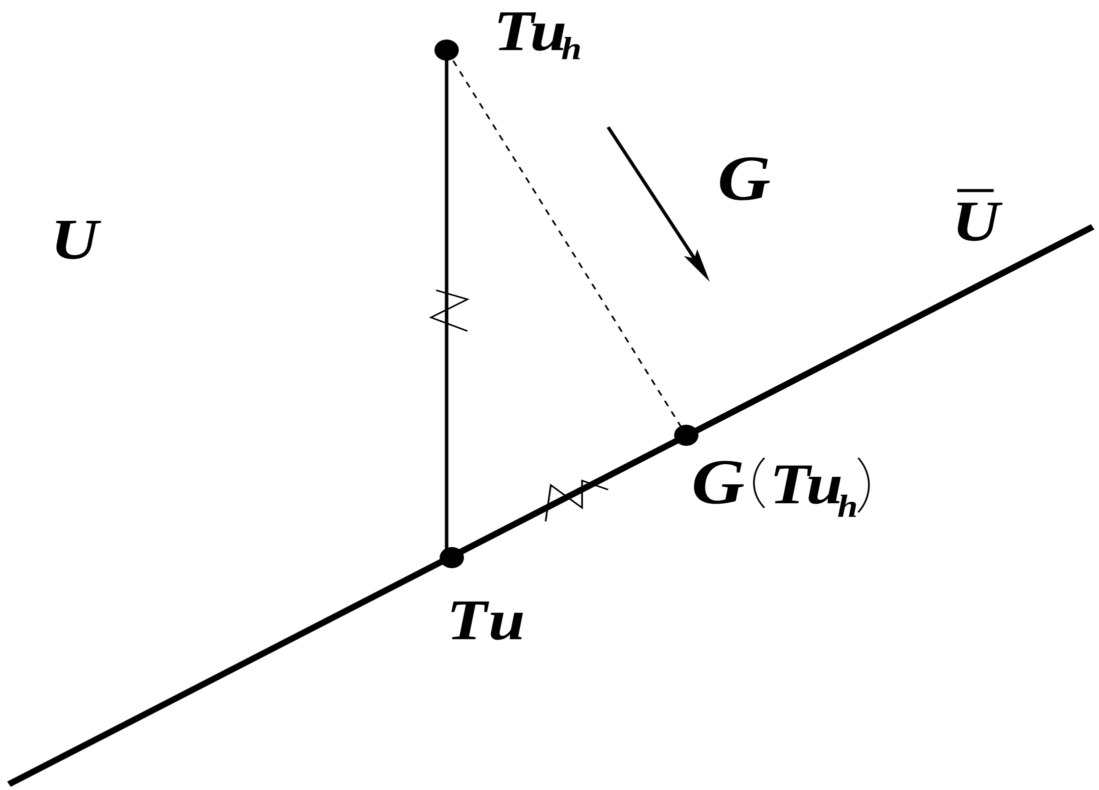
"Definition": *Post-processing* is a special numerical procedure (often "computationally cheap") designed for further modifying already computed approximation in such a way that the post-processed function fits some a priori known property much better than the original approximation

Example: Consider our model elliptic problem. If the finite element approximations u_h are computed by the linear finite elements then their gradients ∇u_h and also $\mathcal{A}\nabla u_h$ are piecewise constant vector-functions, i.e., belong to the space $(L^\infty(\Omega))^d$. However, we know that $\mathcal{A}\nabla u \in H(\mathrm{div}\,,\Omega)$ if $f \in L^2(\Omega)$

- But $(L^\infty(\Omega))^d$ is, in fact, larger than the space $H(\mathrm{div}\,,\Omega)$ **!**
- Sometimes the solution u, and therefore $\mathcal{A}\nabla u$, can be even more smooth ...

Post-processing and error estimation for FEMs: Let for the FE approximation $u_h \in V_h$ the function Tu_h, where T is some linear operator, lies in the space U. Assume that we know a priori that $Tu \in \overline{U} \subset U$, and that we have some computationally inexpensive mapping G such that $G(Tv_h) \in \overline{U}$ for any $v_h \in V_h$. Then we may expect that the (post-processed) function $G(Tu_h)$ would be much closer to Tu than Tu_h

Graphical Illustration



Post-processing scheme

Error Indicator

• Most of post-processing computational procedures used for FE approximations are based on the above argumentation

• If the error caused by violation of a priori regularity properties is dominant and the operator G is properly designed we expect that

$$\|G(Tu_h) - Tu\| \leq \varkappa \|Tu_h - Tu\| \qquad (\varkappa)$$

where $\varkappa << 1$ and $\| \cdot \|$ stands for a norm in the space U

• If this is the case, then directly computable value

$$\|G(Tu_h) - Tu_h\|$$

can be used as a *reliable error indicator* for the error $\|Tu_h - Tu\|$

Really,

$$\|Tu_h - Tu\| \leq \frac{1}{1-\varkappa}\Big(\|Tu_h - Tu\| - \|G(Tu_h) - Tu\|\Big) \leq \frac{1}{1-\varkappa}\|G(Tu_h) - Tu_h\|$$

and

$$\|Tu_h - Tu\| \geq \frac{1}{1+\varkappa}\Big(\|G(Tu_h) - Tu\| + \|Tu_h - Tu\|\Big) \geq \frac{1}{1+\varkappa}\|G(Tu_h) - Tu_h\|$$

That is, if $\varkappa << 1$, then

$$\|Tu_h - Tu\| \approx \|G(Tu_h) - Tu_h\|$$

and $\|G(Tu_h) - Tu_h\|$ can serve as the error indicator

- Small values of $\varkappa$ mean better quality of the indicator

Examples of Post-Processing Procedures

Model Problem: Find a function u such that

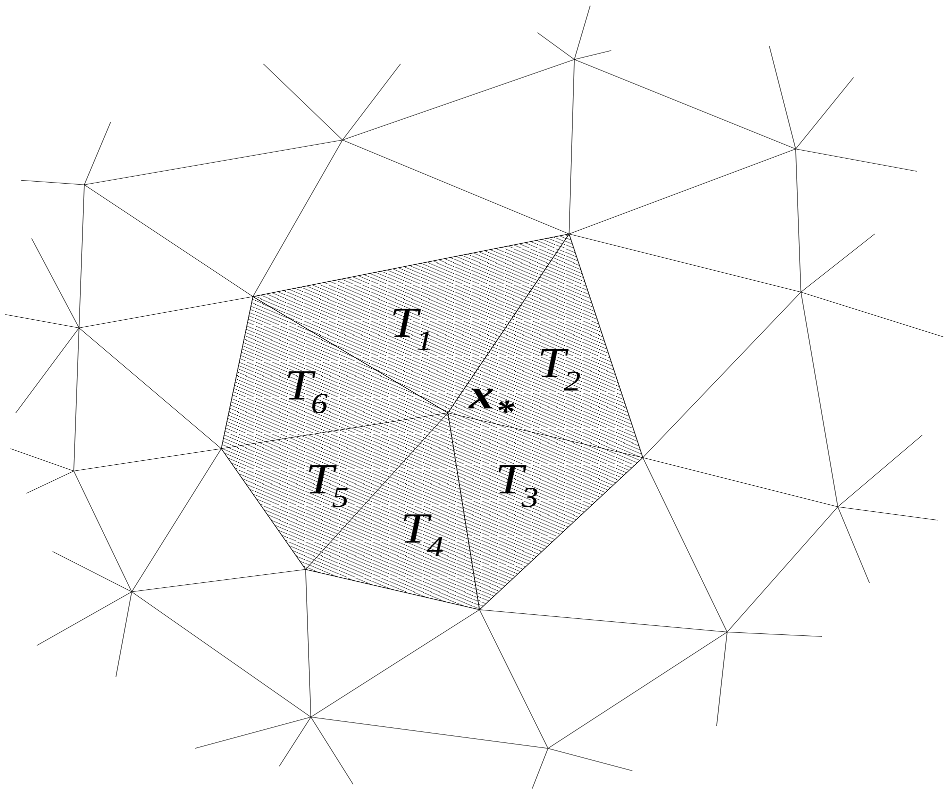
$$-\mathrm{div}\,(\mathcal{A}\,\nabla u) = f \quad \text{in } \Omega \quad \& \quad u = 0 \quad \text{on } \partial\Omega$$

We take $T = \mathcal{A}\nabla$ or ∇, and $\overline{U} = H(\mathrm{div}\,, \Omega)$. If we compute piecewise polynomial continuous approximations $v_h \in V_h$ then

$$\mathcal{A}\nabla v_h \in U = (L^\infty(\Omega))^d$$

- Obviously, $\overline{U} \subset U$. How to define post-processing operator G **?**
- Various post-processing procedures for FE approximations are often constructed using the so-called *patches* of computational mesh

Mesh and Patch



Computational mesh and patch around the node x_*

Integral Gradient Averaging on Patch

General scheme: Let Ω_s be the patch associated with the node x_s of the mesh $\mathcal{T}_h$ on which the approximation u_h is computed. Further, let $(\mathbf{P}^k(\Omega_s))^d$ be the space of vector-valued polynomial functions of degree less or equal to k defined on Ω_s. We define a polynomial $p_s \in (\mathbf{P}^k(\Omega_s))^d$ as the solution of the problem

$$\inf_{p \in (\mathbf{P}^k(\Omega_s))^d} \int_{\Omega_s} |p - Tu_h|^2 \, dx$$

• The minimizer p_s of the above problem can be used to define value of $G(Tu_h)$ at the node x_s

$$G(Tu_h)(x_s) = p_s(x_s)$$

• Further, those nodal values can be used to define $G(Tu_h)$ over the whole solution domain Ω by some prolongation procedure

Example: Let $T = \nabla$ and let u_h be piecewise linear continuous approximations. Then ∇u_h is a piecewise constant function on $\mathcal{T}_h$. Let $\Omega_s = \cup T_{ij}^s$, and let us define

$$(\nabla u_h)_{ij}^s := \nabla u_h|_{T_{ij}^s}$$

Setting $k = 0$ we find $p_s \in (\mathbf{P}^0(\Omega_s))^d$ such that

$$\int_{\Omega_s} |p_s - \nabla u_h|^2\,dx = \inf_{p \in (\mathbf{P}^0(\Omega_s))^d} \int_{\Omega_s} |p - \nabla u_h|^2\,dx =$$

$$= \inf_{p \in (\mathbf{P}^0(\Omega_s))^d} \Big(|p|^2 \text{meas}\,\Omega_s - 2p \cdot \sum_{i,j} (\nabla u_h)_{ij}^s \text{meas}\,T_{ij}^s + \sum_{i,j} ((\nabla u_h)_{ij}^s)^2 \text{meas}\,T_{ij}^s \Big)$$

The minimizer in above can be found

$$p_s = \sum_{i,j} \frac{\text{meas}\,T_{ij}^s}{\text{meas}\,\Omega_s} (\nabla u_h)_{ij}^s$$

We define

$$G(\nabla u_h)(x_s) = p_s$$

• Using this procedure for all the nodes of $\mathcal{T}_h$ and prolongating $G(\nabla u_h)$ linearly we thus define the vector-valued piecewise linear continuous function over the whole domain Ω

• For regular meshes we, obviously, obtain

$$p_s = \sum_{i,j} \frac{1}{M_s} (\nabla u_h)_{ij}^s$$

where M_s is the number of elements in the patch Ω_s

• However, for the nodes on the boundary $\partial\Omega$ we should choose special weights, see the paper by I. Hlaváček and M. Křížek "On a superconvergent finite element scheme for elliptic systems". I, II, III, Aplikace Matematiky **32** 131–154, 200–213, 276–289 (1987)

Discrete Gradient Averaging on Patch

General scheme: We define a polynomial $p_s \in (\mathbf{P}^k(\Omega_s))^d$ as the solution of the problem

$$\inf_{p \in (\mathbf{P}^k(\Omega_s))^d} \sum_{i=1}^{m_s} |p(x_i^s) - Tu_h(x_i^s)|^2 \, dx$$

where x_i^s are the so-called *superconvergence points* in Ω_s and use p_s (and suitable prolongation) in order to define $G(Tu_h)$

Example: If $k = 0$ and $T = \nabla$ then

$$p_s = \frac{1}{m_s} \sum_{i=1}^{m_s} \nabla u_h(x_i^s)$$

Post-Processing of Global Type

• Sometimes it can be desirable to make post-processing globally over the whole solution domain at once

General scheme: Let Tu_h be square summable, we are looking for $\bar{p}_h \in V_h(\Omega) \subset \overline{U}$ such that

$$\|\bar{p}_h - Tu_h\|_{0,\Omega} = \inf_{p_h \in V_h(\Omega)} \|p_h - Tu_h\|_{0,\Omega}$$

Then the function $\bar{p}_h$ is percieved as $G(Tu_h)$

• For more detail on this type of averaging see the recent works by C. Carstensen, S. Bartels "Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids" (Parts I and II) Mathematics of Computation **71** (2002)

Superconvergence Phenomena

• Mathematical proofs of the required properties of post-processing procedures described just above are often based on the so-called *superconvergence phenomena*

• Let u_h be a piecewise linear continuous finite element approximation. Then we have the well-known estimates

$$\|u - u_h\|_{0,\Omega} \leq C_1 h^2 \quad \& \quad \|\nabla(u - u_h)\|_{0,\Omega} \leq C_2 h$$

• Under suitable conditions on regularity we have similar estimates for the maximum norm $\|\cdot\|_{C(\overline{\Omega})}$

$$\max_{x\in\overline{\Omega}} |u(x) - u_h(x)| \leq C_1' h^2 \quad \& \quad \max_{x\in\overline{\Omega}} |\nabla(u(x) - u_h(x))| \leq C_2' h$$

• The constants in above estimates depend on u and properties of $\mathcal{T}_h$. The estimates are optimal, i.e., cannot be improved

• However it was discovered that in many situations the above presented rates of convergence can be higher. For example, we can often prove

$$|u(x_i) - u_h(x_i)| \leq Ch^{2+\alpha}, \quad \alpha > 0$$

for some special points $x_i \in \Omega$. We call then those points as supeconvergence points for the sequence $\{u_h\}$ and use them in gradient averagings

• One popular example of integral gradient averaging on patches was proposed and mathematically analysed by M. Křížek and P. Neittaanmäki in the paper "Superconvergence phenomenon in the finite element method arising from averaging gradients", Numerische Mathematik **45** 105–116 (1984) and later generalized in e.g. I. Hlaváček and M. Křížek "On a superconvergent finite element scheme for elliptic systems". I, II, III, Aplikace Matematiky **32** 131–154, 200–213, 276–289 (1987)

• The postprocessing (gradient averaging) G proposed there allows to get superconvergence estimates of the type

$$\|\nabla u - G(\nabla u_h)\|_{0,\tilde{\Omega}} \leq Ch^{1+\alpha}, \quad \alpha > 0$$

while

$$\|\nabla u - \nabla u_h\|_{0,\Omega} \leq Ch$$

where $\tilde{\Omega} \subseteq \Omega$

• Using the above superconvergence property we can see that the desired inequality

$$\|G(\nabla u_h) - \nabla u\|_{0,\Omega} \leq \varkappa \|\nabla u_h - \nabla u\|_{0,\Omega} \qquad (\varkappa)$$

can be derived with some $\varkappa = \varkappa(h) \to 0$ as $h \to 0$

• Thus, $\|G(\nabla u_h) - \nabla u_h\|_{0,\Omega}$ is a good indicator for $\|\nabla u - \nabla u_h\|_{0,\Omega}$

A POSTERIORI ERROR ESTIMATION

IN TERMS OF LINEAR FUNCTIONALS

- *In this section we describe a new approach for the error control in terms of linear functionals, which, in fact, uses both previous ideas - usage of the adjoint problem and of post-processing procedures*
- *References:* Korotov, S., Neittaanmäki, P., Repin, S.: A posteriori error estimation of goal-oriented quantities by the superconvergence patch recovery. Journal of Numerical Mathematics **11** 33–59 (2003), and Korotov, S.: A posteriori error estimation of goal-oriented quantities for elliptic type BVPs. Journal of Computational and Applied Mathematics **191** 216–227 (2006)

Description of New Approach

1. Model Problem & Principal Estimate

Find u such that

$$\begin{aligned} -\mathrm{div}(\mathcal{A}\nabla u) &= f \quad \text{in} \quad \Omega &(1)\\ u &= 0 \quad \text{on} \quad \Gamma_D &(2)\\ \mathcal{A}\nabla u \cdot n &= g \quad \text{on} \quad \Gamma_N &(3) \end{aligned}$$

where $\Omega \subset \mathbb{R}^d$, n – outward normal to $\partial\Omega = \Gamma_D \cup \Gamma_N$, matrix $\mathcal{A} = \{a_{ij}(x)\}_{i,j=1}^d$ is such that

$$\mathcal{A} = \mathcal{A}^T,\ a_{ij}(x) \in L^\infty(\Omega),\ \mathrm{A}(x)\xi \cdot \xi \geq C\,\|\xi\|^2\ \forall \xi \in \mathbb{R}^d\ \forall x \in \overline{\Omega} \quad (4)$$

$$\text{and} \qquad f \in L^2(\Omega), \quad g \in L^2(\Gamma_N) \qquad (5)$$

• Let $H^1_D = \{w \in H^1(\Omega) | w = 0 \quad \text{on} \quad \Gamma_D\}$. Weak formulation or

Primal Problem $(\mathcal{P})$: Find $u \in H^1_D$ such that

$$\int_\Omega \mathcal{A}\nabla u \cdot \nabla w \, dx = \int_\Omega f w \, dx + \int_{\Gamma_N} g w \, ds \quad \forall w \in H^1_D \tag{6}$$

Let $\bar{u} \in H^1_D$ be an approximation of u , we are interested in estimation of the following value

$$\ell(u - \bar{u}) = \int_\Omega \varphi (u - \bar{u}) \, dx \tag{7}$$

where $\varphi \in L^2(\Omega)$ and $\text{supp } \varphi \subset \omega \subseteq \Omega$

• Estimation of (7) for one (or several) a priori given φ allows to control the error $u - \bar{u}$ locally, in ω, and also to estimate the (unknown) "quantity of interest" $\int_\Omega \varphi u$

Adjoint Problem $(\mathcal{P}_a)$: Find $v \in H^1_D$ such that

$$\int_\Omega \mathcal{A}\nabla v \cdot \nabla w \, dx \; = \; \ell(w) \quad \forall w \in H^1_D \tag{8}$$

Suppose that $\bar{v} \in H^1_D$ is an approximation of v

Theorem 1 (Principal Estimate): We observe that

$$\ell(u - \bar{u}) \; = \; E_0(\bar{u}, \bar{v}) \; + \; E_1(\nabla(u - \bar{u}), \nabla(v - \bar{v})) \tag{9}$$

where

$$E_0(\bar{u}, \bar{v}) = \int_\Omega f\bar{v} \, dx + \int_{\Gamma_N} g\bar{v} \, ds - \int_\Omega \mathcal{A}\nabla\bar{v} \cdot \nabla\bar{u} \, dx \tag{10}$$

$$E_1(\nabla(u - \bar{u}), \nabla(v - \bar{v})) = \int_\Omega \mathcal{A}\nabla(v - \bar{v}) \cdot \nabla(u - \bar{u}) \, dx \tag{11}$$

Proof: From (8), we have

$$\ell(u-\bar{u}) = \int\limits_{\Omega} \mathcal{A}\nabla v \cdot \nabla(u-\bar{u})\ dx =$$

$$= \int\limits_{\Omega} \mathcal{A}\nabla(v-\bar{v}) \cdot \nabla(u-\bar{u})\ dx + \int\limits_{\Omega} \mathcal{A}\nabla\bar{v} \cdot \nabla(u-\bar{u})\ dx =$$

$$= E_1(\nabla(u-\bar{u}), \nabla(v-\bar{v})) - \int\limits_{\Omega} \mathcal{A}\nabla\bar{v} \cdot \nabla\bar{u}\ dx + \int\limits_{\Omega} \mathcal{A}\nabla u \cdot \nabla\bar{v}\ dx =$$

$$= E_1(\nabla(u-\bar{u}), \nabla(v-\bar{v})) - \int\limits_{\Omega} \mathcal{A}\nabla\bar{v} \cdot \nabla\bar{u}\ dx + \int\limits_{\Omega} f\bar{v}\ dx + \int\limits_{\Gamma_N} g\bar{v}\ ds =$$

$$= E_0(\bar{u}, \bar{v}) + E_1(\nabla(u-\bar{u}), \nabla(v-\bar{v}))$$

• Term $E_0(\bar{u},\bar{v})$ is explicitly computable

• If $\bar{v} \approx v$ then $E_1 \approx 0$ and $E_0 \approx \ell(u-\bar{u})$, i.e. term E_0 contains a major part of the error in the "quantity of interest"

• In practice, $\bar{u} = u_h$ – Galerkin approximation from $V_h \subset H^1_D$, and $\bar{v} = v_\tau$ – Galerkin approximation from $V_\tau \subset H^1_D$

• If $V_h \equiv V_\tau$ (as always assumed in the above cited works) or if $V_\tau \subset V_h$, then

$$E_0(u_h, v_\tau) = \int_\Omega f v_\tau \, dx + \int_{\Gamma_N} g v_\tau \, ds - \int_\Omega \mathcal{A}\nabla v_\tau \cdot \nabla u_h \, dx \equiv 0$$

and the major term

$$E_0 = \int_\Omega \mathcal{A}\nabla v_\tau \cdot (\nabla u - \nabla u_h) \, dx \le C \|\nabla u - \nabla u_h\|_{0,\Omega} = \mathcal{O}(h)$$

is lost and what remains is the higher order term $E_1 = \mathcal{O}(h\tau)$

2. New Modus Operandi

1) Primal and Adjoint Problems are solved in finite-dimensional subspaces V_h and V_τ constructed in such a manner that *V_τ is not a subset of or equal to V_h* (i.e. $E_0(u_h, v_\tau)$ is not necessarily zero) and Principal Estimate (9) is used

2) Space V_τ is adopted to "quantity of interest", i.e. to function φ

3) Term E_1 is estimated by the explicitly computable term $\tilde{E}_1$ using suitable *gradient averaging procedures*

• In what follows, we present the estimator for $\ell(u - \bar{u})$ in explicit form for the model problem and analyse its asymptotic behaviour

3. Construction of Estimator

Let $V_h, V_\tau \subset H^1_D$ be constructed by the Courant type finite element discretization, and Galerkin approximations u_h, v_τ be defined by

Problem $(\mathcal{P}^h)$: Find $u_h \in V_h$ such that

$$\int_\Omega \mathcal{A}\nabla u_h \cdot \nabla w_h \; dx = \int_\Omega f w_h \; dx + \int_{\Gamma_N} g w_h \; ds \quad \forall w_h \in V_h \tag{12}$$

Problem $(\mathcal{P}^\tau_a)$: Find $v_\tau \in V_\tau$ such that

$$\int_\Omega \mathcal{A}\nabla v_\tau \cdot \nabla w_\tau \; dx = \ell(w_\tau) \; dx \quad \forall w_\tau \in V_\tau \tag{13}$$

From Theorem 1

$$\ell(u - u_h) \; dx = E_0(u_h, v_\tau) + E_1(\nabla(u - u_h), \nabla(v - v_\tau)) \tag{14}$$

• For the considered class of (elliptic) problems and linear finite elements, it is known that

$$\|\nabla(u - u_h)\|_{0,\Omega} = \mathcal{O}(h) \quad \text{and} \quad \|\nabla(v - v_\tau)\|_{0,\Omega} = \mathcal{O}(\tau)$$

and that the above convergence rate estimates are optimal

• Let G_h and G_τ be suitable gradient averaging operators, then usually we have

$$\|\nabla u - G_h(\nabla u_h)\|_{0,\Omega} = \mathcal{O}(h^{1+\alpha}), \; \alpha > 0$$

$$\|\nabla v - G_\tau(\nabla v_\tau)\|_{0,\Omega} = \mathcal{O}(\tau^{1+\beta}), \; \beta > 0$$

i.e., (computable) averaged gradients $G_h(\nabla u_h)$ and $G_\tau(\nabla v_\tau)$ are *asymptotically much closer* to the exact (unknown) gradients ∇u and ∇v than to ∇u_h and ∇v_τ, respectively (*superconvergence*)

The above considerations suggest to replace the term

$$E_1 = \int_\Omega \mathcal{A}\nabla(u - u_h) \cdot \nabla(v - v_\tau)\, dx$$

with unknown gradients ∇u , ∇v by a computable term

$$\widetilde{E}_1(u_h, v_\tau) = \int_\Omega \mathcal{A}(G_h(\nabla u_h) - \nabla u_h) \cdot (G_\tau(\nabla v_\tau) - \nabla v_\tau)\, dx \qquad (15)$$

and define the estimator for quantity $\ell(u - u_h)$ as follows

$$\boxed{\widetilde{E}(u_h, v_\tau) := E_0(u_h, v_\tau) + \widetilde{E}_1(u_h, v_\tau)} \qquad (16)$$

4. Superconvergence of Averaged Gradients

Theorem 2 [Hlavaček, Křížek, 1987]: Let Ω be a polygon, $\Gamma_N = \emptyset$, $u, v \in H^3(\Omega)$, and triangulations $\mathcal{T}_h$, $\mathcal{T}_\tau$ be uniform, then

$$\|\nabla u - G_h(\nabla u_h)\|_{0,\Omega} = \mathcal{O}(h^2) \quad \text{and} \quad \|\nabla v - G_\tau(\nabla v_\tau)\|_{0,\Omega} = \mathcal{O}(\tau^2)$$

where the gradient averagings are built on patches

- *However, in practice, the superconvergence for the gradient averagings on patches can be observed under much weaker conditions on the meshes employed and with less smooth exact solutions. The same observations hold also for another types of averaging*

5. Asymptotic Behaviour of Estimator

- $E_0 = \mathcal{O}(h)$, $E_1 = \mathcal{O}(h\tau)$, and $\widetilde{E}_1 = \mathcal{O}(h\tau)$ (in the presence of superconvergence), $E_1 - \widetilde{E}_1 = ?$

Theorem 3 [Korotov, Neittaanmäki, Repin, 2003]: Under assumptions of Theorem 2, for sufficiently small h and τ $(\tau < h)$

$$|E_1(u_h, v_\tau) - \widetilde{E}_1(u_h, v_\tau)| \leq C\,(h\tau^{\frac{3}{2}} + \tau h^{\frac{3}{2}} + h\tau(h+\tau)^{1-\frac{2}{m}} + \tau^2) + \mu(h,\tau) \qquad (17)$$

where m is any positive integer greater than 2, $\mu(h,\tau)$ contains higher order terms, and C does not depend on h, τ

- $\Rightarrow$ change of $E_1(u_h, v_\tau)$ by $\widetilde{E}_1(u_h, v_\tau)$ is asymptotically correct

Numerical Experiments

Ex1 (Various Adjoint Meshes): Find u such that

$$-\Delta u = 10 \quad \text{in} \quad \Omega, \quad u = 0 \quad \text{on} \quad \Gamma_D, \quad \nabla u \cdot n = 0 \quad \text{on} \quad \Gamma_N \quad (18)$$

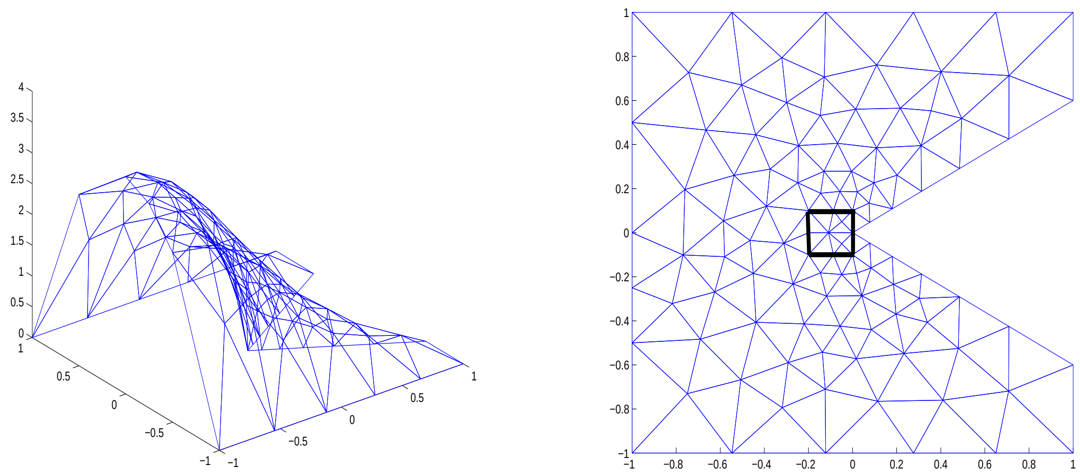


Figure 1: u_h, Ω, ω and $\mathcal{T}_h$ with 107 nodes

We are interested in the estimation of $\int\limits_{\omega} (u - u_h) \, dx$

$$I_{eff} := \frac{|\tilde{E}(u_h, v_\tau)|}{|\int\limits_{\omega} (u - u_h) \, dx|} \qquad \text{- effectivity index} \qquad (19)$$

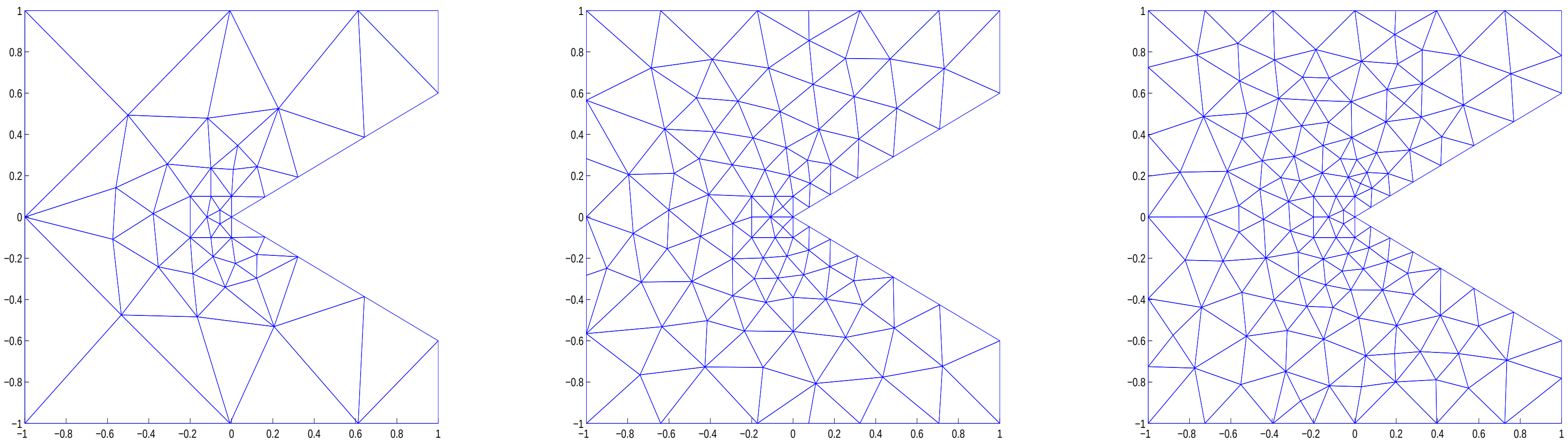


Figure 2: Typical adjoint meshes with 48, 105, 146 nodes

Table 1: Performance of $\tilde{E}$ for Ex1

Pr	Ad	E_0	$\tilde{E}_1$	$\tilde{E}$	$\int_\omega (u - u_h)dx$	I_{eff}
107	48	0.001616	0.003146	0.004762	0.004514	1.05
107	57	0.001585	0.002759	0.004344	0.004514	0.96
107	68	0.001804	0.002620	0.004424	0.004514	0.98
107	107	0.000000	0.004220	0.004220	0.004514	0.93
107	134	0.000566	0.003612	0.004178	0.004514	0.93
107	146	0.001662	0.002902	0.004564	0.004514	1.01
107	171	0.002947	0.002025	0.004972	0.004514	1.10
107	244	0.003311	0.001228	0.004539	0.004514	1.01
107	923	0.003901	0.000646	0.004547	0.004514	1.01

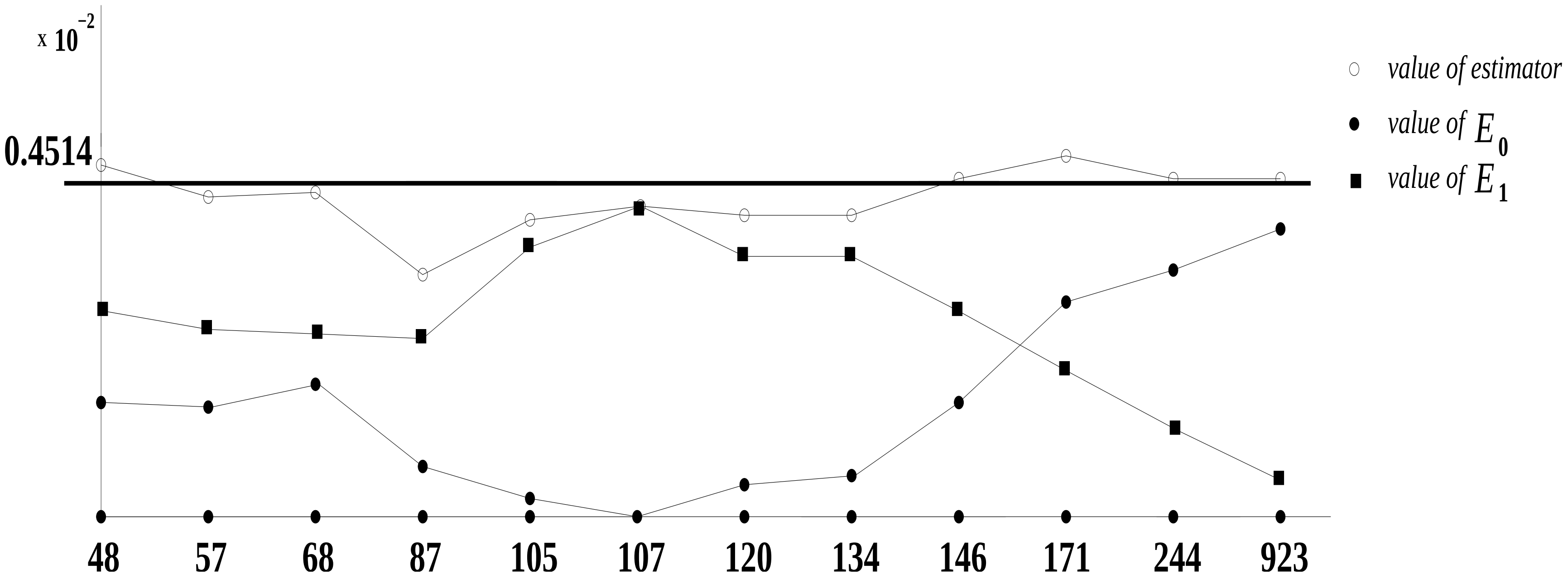


Figure 3: Behaviour of estimator $\tilde{E}$ and its parts, E_0 and $\tilde{E}_1$, for various choices of the adjoint meshes in Ex 1

Ex2 (Sharp Reentrant Corners): Find u such that

$$-\Delta u = 10 \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \partial\Omega \tag{20}$$

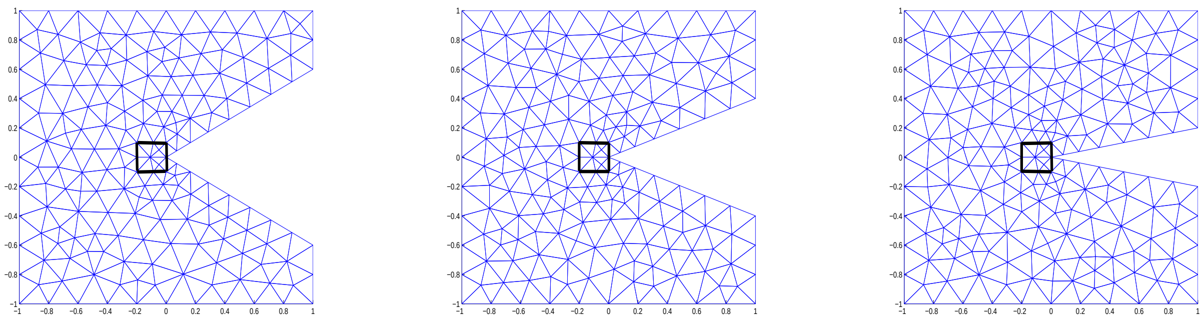


Figure 4: Domains Ω and primal meshes with 183, 191, 205 nodes

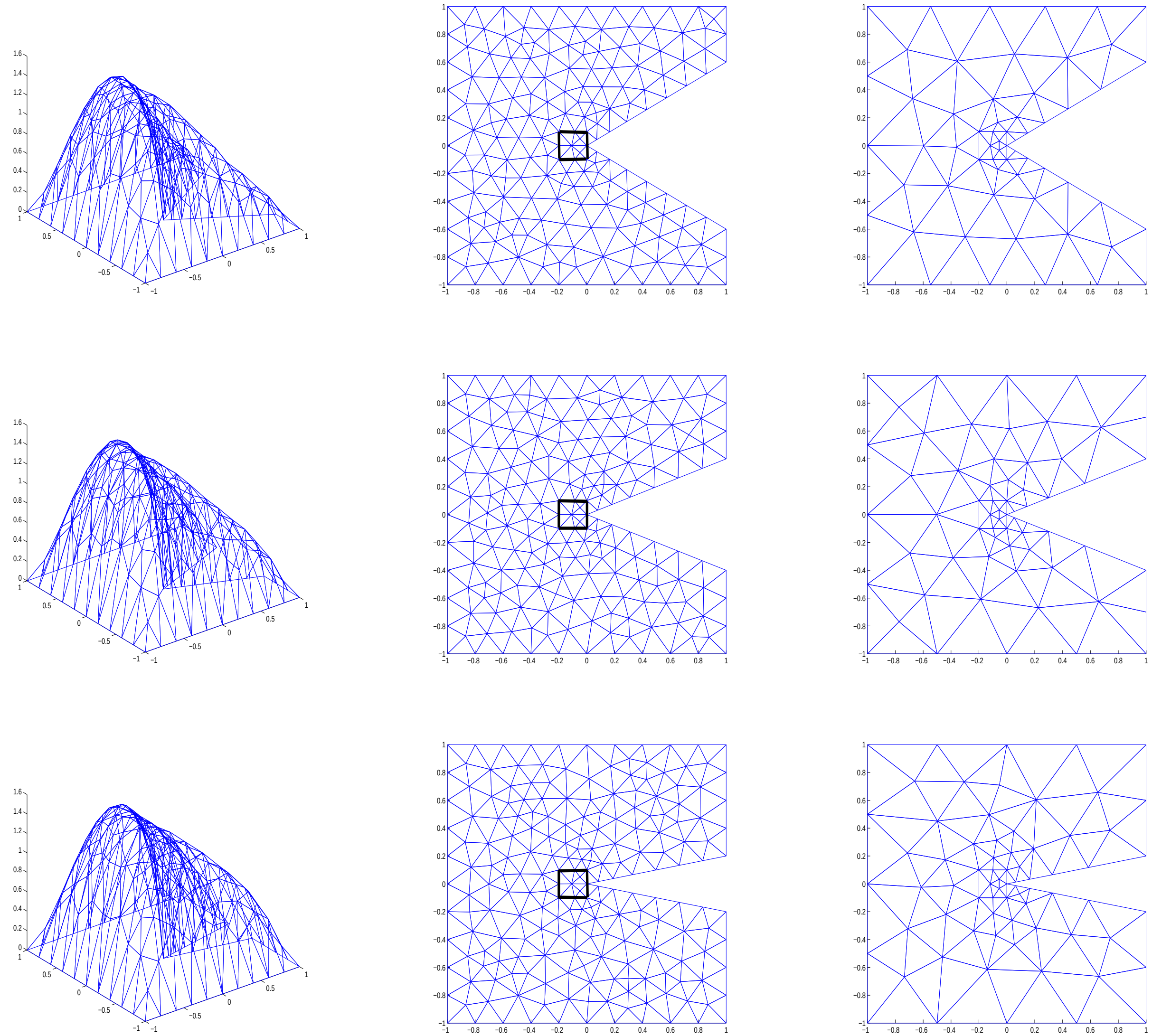
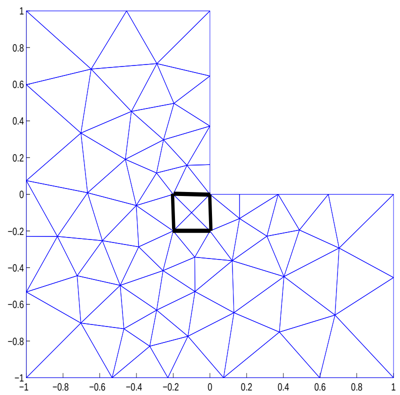


Table 2: Performance of $\tilde{E}$ for Ex 2

Pr	Ad	E_0	$\tilde{E}_1$	$\tilde{E}$	$\int_\omega (u - u_h)dx$	I_{eff}
183	61	0.001259	0.001682	0.002941	0.003158	0.93
191	64	0.001460	0.002035	0.003495	0.003859	0.91
205	72	0.001660	0.002754	0.004414	0.004888	0.90

Ex3 (Adaptivity): Find u such that

$$-\Delta u = 10 \quad \text{in} \quad \Omega, \quad u = 0 \quad \text{on} \quad \partial\Omega \tag{21}$$



In fact, the estimator is an integral taken over Ω

$$E(u_h, v_\tau) = \sum_{T \in \mathcal{T}_h} I_T \tag{22}$$

where each contribution I_T is a value of integral over a particular element T of $\mathcal{T}_h$. This suggests a straightforward adaptive strategy

- 4 refinement steps for $\mathcal{T}_h$ have been performed with threshold 70%
- $\mathcal{T}_\tau$ in all the cases has been the same

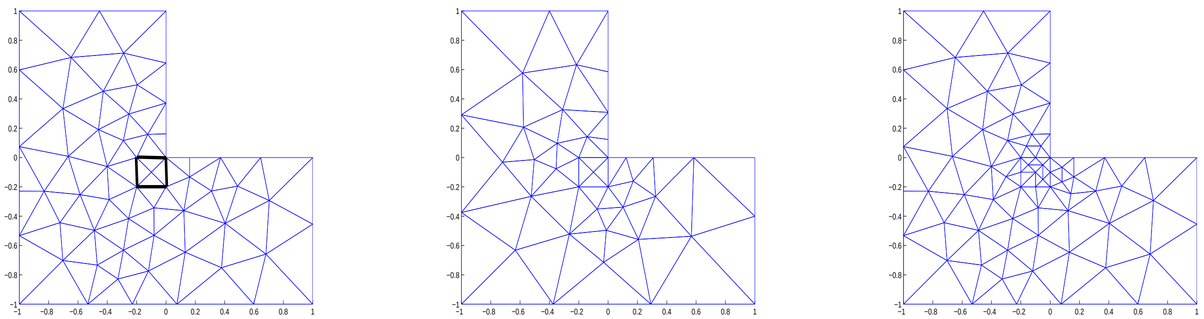


Figure 5: $\mathcal{T}_h^{(0)}$ (59 nodes), $\mathcal{T}_\tau$ (42 nodes), $\mathcal{T}_h^{(1)}$ (72 nodes)

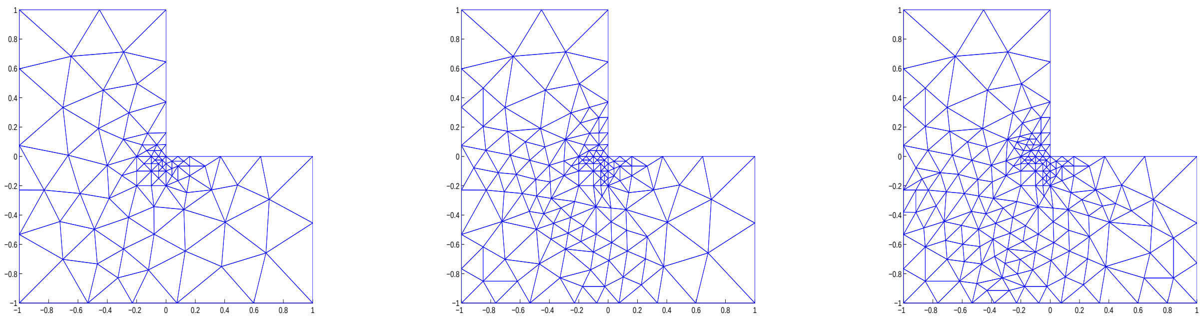


Figure 6: $\mathcal{T}_h^{(2)}$ (90 nodes), $\mathcal{T}_h^{(3)}$ (149 nodes), $\mathcal{T}_h^{(4)}$ (186 nodes)

Table 3: The results for Ex 3

Pr	Ad	E_0	$\tilde{E}_1$	$\tilde{E}$	$\int_\omega (u - u_h)dx$	I_{eff}
59	42	0.000785	0.003433	0.004218	0.005428	0.78
72	42	0.000586	0.001774	0.002360	0.002739	0.86
90	42	0.000582	0.001217	0.001799	0.001804	0.99
149	42	0.000470	0.000939	0.001409	0.001540	0.91
186	42	0.000420	0.000824	0.001244	0.001338	0.93

- As expected, dense mesh refinements appear near reentrant vertex
- Table 3 shows that both, exact error and estimator's values are very close and monotonically decrease as adaptive procedure proceeds

Conclusions

I. Our approach is different from the other techniques proposed, where it is always assumed that the primal and adjoint problems are solved on coinciding meshes. Using our technique one can obtain reliable estimates also for the case when the number of nodes in the mesh used for the adjoint problem is considerably smaller than the number of nodes in the mesh used for the primal problem.

II. The technology proposed can be directly applied to another linear elliptic equations, e.g. to problems in the linear elasticity, provided that the averaged gradients of their solutions demonstrate certain superconvergenve phenomena.

III. The effectivity of the proposed technique, strongly increases when one is interested not in a single solution of the primal problem for a concrete data, but analyzes a series of approximate solutions for a certain set of boundary conditions and various right-hand sides (which is typical in the engineering design when it is necessary to model the behavior of a construction for various working regimes). In this case, the adjoint problem must be solved *only once* for each "quantity of interest", and its solution can be further used in testing the accuracy of approximate solutions of various primal problems.

IV. The technology presented can be easily coded and attached as an independent *programme-checker* to the most of existing educational and industrial codes that use the finite element method as a computational tool.

Another Relevant Publications:

1. Rüter, M., Korotov, S., Steenbock, Ch.: Goal-oriented error estimates based on different FE-solution spaces for the primal and the dual problem with application to linear elastic fracture mechanics. Computational Mechanics, 2007.

2. Korotov, S.: Error control in terms of linear functionals based on gradient averaging techniques. Computing Letters, 2007.