Computer tutorial N°4

Identification of error components, adaptivity of linear and nonlinear solvers

Strongly monotone and Lipschitz-continuous nonlinear problem, conforming finite element method, equilibrated flux reconstruction for inexact solvers, error components, stopping criteria

We consider the same geometric setup as in Tutorial N°1: let $\Omega \subset \mathbb{R}^2$ be a polygon with Lipschitz boundary $\partial\Omega = \Gamma_{\rm D} \cup \Gamma_{\rm N}$, $f \in L^2(\Omega)$ a given source term, and $g_{\rm D}$ a given prescribed data on the Dirichlet part of the boundary $\Gamma_{\rm D}$. Additionally, we introduce a scalar-valued function $a : (0,\infty) \to (0,\infty)$, satisfying the assumptions below. The problem of interest in then to find $u : \Omega \to \mathbb{R}$ such that

$$-\nabla\cdot(a(|\nabla u|)\nabla u) = f \qquad \text{in } \Omega, \tag{1a}$$
$$u = g_{\rm D} \qquad \text{on } \Gamma_{\rm D}, \tag{1b}$$
$$-a(|\nabla u|)\nabla u\cdot\boldsymbol{n}_\Omega = 0 \qquad \text{on } \Gamma_{\rm N}. \tag{1c}$$

Note that when $a(s) = 1$, problem (1) becomes the Poisson equation of Tutorial N°1. The weak solution of problem (1) is a function $u \in H^1(\Omega)$ such that $u|_{\Gamma_{\rm D}} = g_{\rm D}$ and

$$(a(|\nabla u|)\nabla u, \nabla v) = (f, v) \qquad \forall v \in H^1_0(\Omega). \tag{2}$$

For a conforming simplicial mesh $\mathcal{T}_\ell$ of Ω and the space V_ℓ^p, with fixed polynomial degree $p \geq 1$,

$$V_\ell^p = \{v_\ell \in H^1(\Omega), v_\ell|_K \in \mathcal{P}_p(K) \quad \forall K \in \mathcal{T}_\ell\} = \mathcal{P}_p(\mathcal{T}_\ell) \cap H^1(\Omega), \tag{3}$$

this leads to the discrete problem: find $u_\ell \in V_\ell^p$ such that $u_\ell|_{\Gamma_{\rm D}} = g_{\rm D}$ and such that

$$(a(|\nabla u_\ell|)\nabla u_\ell, \nabla v_\ell) = (f, v_\ell) \qquad \forall v_\ell \in V_\ell^p \text{ such that } v_\ell|_{\Gamma_{\rm D}} = 0. \tag{4}$$

We consider nonlinear functions a satisfying, for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$,

$$|a(|\boldsymbol{x}|)\boldsymbol{x} - a(|\boldsymbol{y}|)\boldsymbol{y}| \leq L|\boldsymbol{x} - \boldsymbol{y}| \qquad \text{(Lipschitz continuity)}, \tag{5a}$$
$$(a(|\boldsymbol{x}|)\boldsymbol{x} - a(|\boldsymbol{y}|)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) \geq \alpha|\boldsymbol{x} - \boldsymbol{y}|^2 \qquad \text{(strong monotonicity)}, \tag{5b}$$

where α and L are fixed positive real constants. Under these assumptions, problems (2) and (4) are well posed.

We henceforth consider Ω being the unit square, $\Gamma_{\rm D} = \partial\Omega$ and $\Gamma_{\rm N} = \emptyset$, and the exact solution taking the form

$$u(x, y) = x(x - 1)y(y - 1),$$

so that $g_{\rm D} = 0$. We pick the particular form

$$a(s) := 1 + \frac{L - 1}{\sqrt{1 + s^2}}, \tag{6}$$

corresponding to the so-called mean curvature flow. Here $L \geq 1$ is a fixed real parameter. One finds that assumptions (5) are satisfied with this constant L and $\alpha = 1$. The right-hand side f is then obtained by inserting u into (1a). For the illustrations below, we consider $10 \times 10 \times 2$ isosceles triangles in the mesh $\mathcal{T}_\ell$.

Exercice 1. (The Picard iteration)

In this exercise, we will investigate solving the model problem (1) using the *Picard (fixed-point) iteration*: given $u_\ell^0 \in V_\ell^p$, we seek to solve iteratively, for $k \geq 1$, for $u_\ell^k \in V_\ell^p$ such that $u_\ell|_{\Gamma_D} = g_D$ and such that

$$(a(|\nabla u_\ell^{k-1}|)\nabla u_\ell^k, \nabla v_\ell) = (f, v_\ell) \qquad \forall v_\ell \in V_\ell^p \text{ such that } v_\ell|_{\Gamma_D} = 0. \tag{7}$$

In this exercise we will focus on the code inside the block `if (DoEx1)`. In order to check that the Picard method is converging, we will compute the following error between consecutive iterates

$$\|\nabla(u_\ell^k - u_\ell^{k-1})\|. \tag{8}$$

The criteria to stop the solver is that the quantity (8) is less than $1e-7$. This is given in the code by `real SolverTol = 1e-7`. You may experiment with different values.
We also compute the H^1-seminorm error as in Exercise 4 of Tutorial N°1:

$$\|\nabla(u - u_\ell^k)\|. \tag{9}$$

To run this exercise, set the flag `bool DoEx1 = 1`. After the script is finished, you should have the gnuplot script file `Exercise1.plt` in the same directory. Run `>gnuplot Exercise1.plt` whereupon you can open the plot as `Exercise1.png`. What do you observe?

Answer 1. (The Picard iteration)

The result is given in Figure 1. We observe that the H^1-seminorm error more or less stabilizes after three iterations. The difference of iterates appears to decrease quite rapidly at an exponential rate. In the following exercises, we will explore error estimates for this problem as well as adaptive stopping criteria.

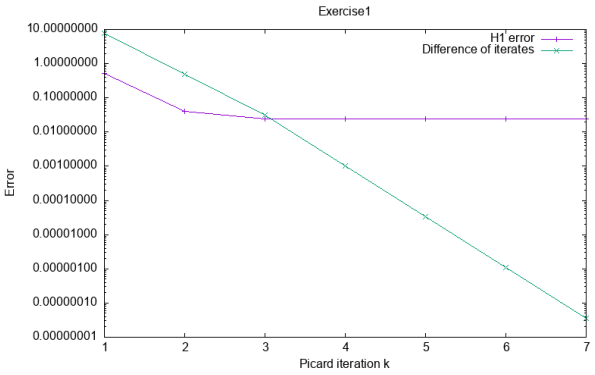

Figure 1: Difference of iterates given by (8) and the error in the H^1-seminorm (9) for the Picard iteration (7), $L = 10$

Exercice 2. (A posteriori estimates for the linearized problem)

We now consider the problem of error estimation for the nonlinear problem. In particular, we consider the equilibrated flux of Exercise 3 in Tutorial N°1, with a slight difference highlighted in green,

$$\boldsymbol{\sigma}_\ell^{k,\boldsymbol{a}} := \arg \min_{\substack{\boldsymbol{v}_\ell \in \boldsymbol{\mathcal{RT}}_{p'}(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}_0(\mathrm{div},\omega_{\boldsymbol{a}}) \\ \nabla\cdot\boldsymbol{v}_\ell = \Pi_{p'}(f\psi^{\boldsymbol{a}} - a(|\nabla u_\ell^{k-1}|)\nabla u_\ell^k\cdot\nabla\psi^{\boldsymbol{a}})}} \|\psi^{\boldsymbol{a}} a(|\nabla u_\ell^{k-1}|)\nabla u_\ell^k + \boldsymbol{v}_\ell\|_{\omega_{\boldsymbol{a}}}, \tag{10a}$$

$$\boldsymbol{\sigma}_\ell^k := \sum_{\boldsymbol{a}\in\mathcal{V}_\ell} \boldsymbol{\sigma}_\ell^{k,\boldsymbol{a}} : \tag{10b}$$

only the flux ∇u_ℓ^k from the linear problem of Tutorial N°1 is replaced by the current flux, featuring in addition the multiplication by the nonlinear function a. Note that the function a is evaluated at $|\nabla u_\ell^{k-1}|$ of the previous Picard iterate. The reason is that the term $a(|\nabla u_\ell^{k-1}|)$ is taken from (7), where $a(|\nabla u_\ell^k|)$ could not appear, since it is not known yet.

As in Tutorial N°1, problem (10) is equivalent to the following problem: find $\boldsymbol{\sigma}_\ell^{k,\boldsymbol{a}} \in \boldsymbol{\mathcal{RT}}_{p'}(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}_0(\mathrm{div},\omega_{\boldsymbol{a}})$, together with the additional scalar-valued piecewise polynomial $\gamma_\ell^{\boldsymbol{a}} \in \mathcal{P}_{p'}(\mathcal{T}_{\boldsymbol{a}})$, such that

$$\begin{aligned} &(\boldsymbol{\sigma}_\ell^{k,\boldsymbol{a}}, \boldsymbol{v}_\ell)_{\omega_{\boldsymbol{a}}} \\ -(\gamma_\ell^{\boldsymbol{a}}, \nabla\cdot\boldsymbol{v}_\ell)_{\omega_{\boldsymbol{a}}} &= -(\psi^{\boldsymbol{a}} a(|\nabla u_\ell^{k-1}|)\nabla u_\ell^k, \boldsymbol{v}_\ell)_{\omega_{\boldsymbol{a}}} \qquad \forall \boldsymbol{v}_\ell \in \boldsymbol{\mathcal{RT}}_{p'}(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}_0(\mathrm{div},\omega_{\boldsymbol{a}}), \end{aligned} \tag{11a}$$

$$(\nabla\cdot\boldsymbol{\sigma}_\ell^{\boldsymbol{a}}, q_\ell)_{\omega_{\boldsymbol{a}}} = (f\psi^{\boldsymbol{a}} - a(|\nabla u_\ell^{k-1}|)\nabla u_\ell^k\cdot\nabla\psi^{\boldsymbol{a}}, q_\ell)_{\omega_{\boldsymbol{a}}} \quad \forall q_\ell \in \mathcal{P}_{p'}(\mathcal{T}_{\boldsymbol{a}}). \tag{11b}$$

Recall from the lectures the residual stemming from the weak formulation (2): for the current Picard iterate u_ℓ^k, $k \geq 1$,

$$\langle \mathcal{R}(u_\ell^k), v\rangle := (f, v) - (a(|\nabla u_\ell^k|)\nabla u_\ell^k, \nabla v).$$

Its dual norm is then given by

$$\|\mathcal{R}(u_\ell^k)\|_{-1} := \sup_{v\in H^1_{0,\mathrm{D}}(\Omega), \|\nabla v\|=1} \langle \mathcal{R}(u_\ell^k), v\rangle. \tag{12}$$

Then, the guaranteed upper bound

$$\|\mathcal{R}(u_\ell^k)\|_{-1} \leq \eta_\ell^k := \left\{ \sum_{K\in\mathcal{T}_\ell} \left[\|a(|\nabla u_\ell^k|)\nabla u_\ell^k + \boldsymbol{\sigma}_\ell^k\|_K + \frac{h_K}{\pi}\|f - \Pi_{p'} f\|_K \right]^2 \right\}^{\frac{1}{2}} \tag{13}$$

is valid, and we know from the theory that the converse inequality, the efficiency, is also true, and this for a generic constant independent of both α and L, i.e., structurally, up to data oscillation and quadrature errors,

$$\eta_\ell^k \lesssim \|\mathcal{R}(u_\ell^k)\|_{-1}. \tag{14}$$

Moreover, (14) also holds locally, on patches of elements.

The dual norm of the residual (12) is closely related to the error in the fluxes

$$\|a(\nabla u)\nabla u - a(|\nabla u_\ell^k|)\nabla u_\ell^k\|. \tag{15}$$

Indeed, (15) is an upper bound to (12), as seen from

$$\begin{aligned}\|\mathcal{R}(u_\ell^k)\|_{-1} &= \sup_{v\in H^1_{0,\mathrm{D}}(\Omega),\|\nabla v\|=1} (f,v) - (a(|\nabla u_\ell^k|)\nabla u_\ell^k, \nabla v)\\ &= \sup_{v\in H^1_{0,\mathrm{D}}(\Omega),\|\nabla v\|=1} (a(|\nabla u|)\nabla u - a(|\nabla u_\ell^k|)\nabla u_\ell^k, \nabla v)\\ &\le \sup_{v\in H^1_{0,\mathrm{D}}(\Omega),\|\nabla v\|=1} \big\{\|a(|\nabla u|)\nabla u - a(|\nabla u_\ell^k|)\nabla u_\ell^k\|\,\|\nabla v\|\big\}\\ &\le \|a(|\nabla u|)\nabla u - a(|\nabla u_\ell^k|)\nabla u_\ell^k\|.\end{aligned} \tag{16}$$

Set `bool DoEx2 = 1` and run the script. This will produce `Exercise2.plt` which can be run with `>gnuplot Exercise2.plt`. This will produce `Exercise2.png`. The plot includes the flux error (15), the estimator from (13), and the H^1-seminorm error from (9) as a function of the Picard iterations. What do you observe about the estimator? Can you justify, based on theory, which of the two errors the estimator tracks better? Try redoing the experiment with various values of the Lipschitz constant $L \ge 1$.

Answer 2. (A posteriori estimates for the linearized problem)

The result is presented in Figure 2. We observe that the plotted quantities do not change substantially after three iterations, but additional iterations are needed to satisfy our stopping criteria (the same as in Exercise 1). Next we notice that the estimator η_ℓ^k from (13) provides an excellent upper bound on the flux error (15), whereas the H^1-seminorm error (9) is much smaller than η_ℓ^k. This will be explored in more detail in the following exercises.

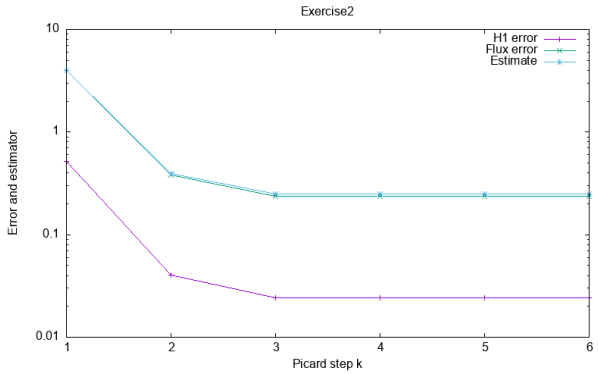

Figure 2: The H^1-seminorm error (9), the flux error (15), and the estimator (13), $L = 10$

Exercice 3. (Closer investigation along the iterations)

In the last exercise, we saw that the estimator (13) does a good job of tracking the flux error (15). We now want to study more carefully the approximate solutions, equilibrated fluxes, errors, and estimators that are produced at each Picard iterate. We also want to stress that the equilibrated flux $\boldsymbol{\sigma}_\ell^k$ physical, ensuring both mass balance over the mesh faces and the equilibrium constraint, and this on any Picard iteration $k \ge 1$. To run this exercise, set the flag `DoEx3 = true`.

Answer 3. (Closer investigation along the iterations)

We plot here the various objects for $L = 10$ at the third Picard iteration $k = 3$ in Figures 3–6. We first remark that in Figure 3, the exact solution is of course the same as in the linear case, and the approximate solution is also not too different. In Figure 4, we see that now the fluxes obtained from the true and approximate solutions have a different size with respect to the linear case of Tutorial N°1. This is the first place we see a difference with the linear case, since even for the true solution, we have the coefficient $a(|\nabla u|)$. It can be seen that the true solution is in $\boldsymbol{H}(\mathrm{div}, \Omega)$, as it has continuous normal components over the inter-element boundaries. The approximate flux, on the other hand, does not have continuous normal components, same as in the linear case.

Next, we have probably the most interesting plot, the divergence misfit. We remark that the divergence misfit is essentially machine precision. This is to be expected theoretically from the fact that, up to oscillation errors, the equilibrated flux satisfies $\nabla{\cdot}\boldsymbol{\sigma}_\ell^k = \Pi_{p'} f$, and this at each Picard iteration. Thus, crucially, we construct an object that respects the mass balance at each step of the linearization. As a consequence, we can terminate the linearization procedure at any time, with the conservation of mass guaranteed.

Finally, we consider in Figure 6 the elementwise contributions to the two different notions of error, as well as the elementwise estimators. We observe that the estimator actually captures the distribution of the error well in both cases, but interestingly, the estimator and the flux error are about an order of magnitude greater than the H^1-seminorm error. We will explore this phenomenon more in the next exercise.

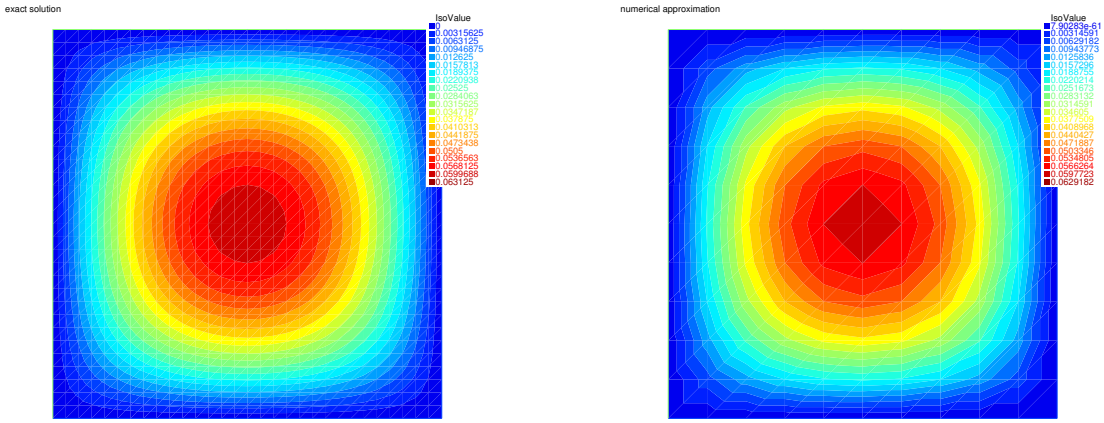

Figure 3: Exact solution u (left) and approximate solution u_ℓ^3 (right, $p = 1$) at the third Picard iteration, $L = 10$

Exercice 4. (Scaling with respect to the Lipschitz constant L)

In this exercise, we will observe the scaling of the estimator and of both notions or error with respect to the Lipschitz constant L. To run this exercise, set the flag `bool DoEx1 = 4`. After the script is finished, you should have the gnuplot script file `Exercise4.plt` in the same directory. Run `>gnuplot Exercise4.plt` whereupon you can open the plot as `Exercise4.png`. What do you observe?

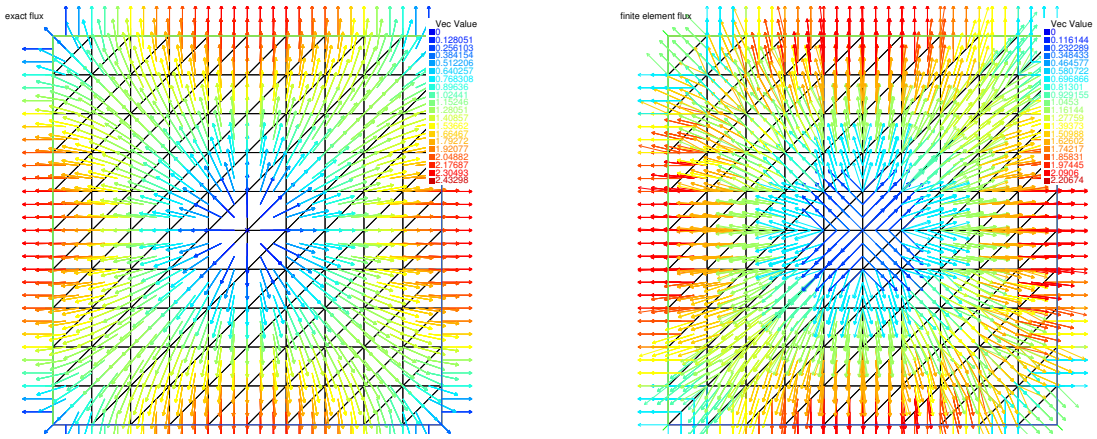

Figure 4: Flux of the exact solution $-a(|\nabla u|)\nabla u$ (left) and flux of the approximation $-a(|\nabla u_\ell^3|)\nabla u_\ell^3$ (right, $p = 1$) at the third Picard iteration, $L = 10$

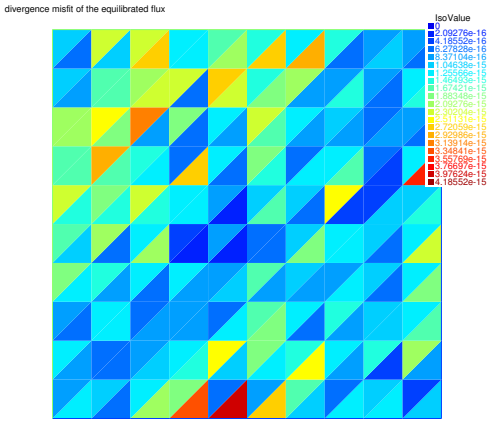

Figure 5: Divergence misfit of the equilibrated flux $\boldsymbol{\sigma}_\ell^3$ at the third Picard iteration, $L = 10$

Answer 4. (Scaling with respect to the Lipschitz constant L)

The result is reported in Figure 7. We recall from the theory that the H^1-seminorm and the dual norm of the residual are related, for any $v \in H^1_{0,\mathrm{D}}(\Omega)$, by

$$\alpha\|\nabla(v-u)\| \le \|\mathcal{R}(v)\|_{-1} \le L\|\nabla(v-u)\|. \tag{17}$$

In our problem, $\alpha = 1$, so the scaling between the dual norm of the residual and the H^1-seminorm is exactly L. We also know that the flux error is an upper bound for the error in the dual norm of the residual, see (16) where merely the Cauchy–Schwarz inequality has been used, so that from (13)–(14), we can expect the ratio of the flux error to the estimator to be stable, i.e., robust. In contrast, the ratio of the flux error to the H^1-seminorm of the error is expected to diverge linearly in the Lipschitz constant L.

Exercice 5. (Adaptive stopping criteria)

In this exercise, we will consider a decomposition of the already established estimator η_ℓ^k from (13) into error components. To this end, we consider two new fluxes,

$$\boldsymbol{l}_\ell^k := a(|\nabla u_\ell^k|)\nabla u_\ell^k - a(|\nabla u_\ell^{k-1}|)\nabla u_\ell^k, \tag{18a}$$

$$\boldsymbol{d}_\ell^k := a(|\nabla u_\ell^{k-1}|)\nabla u_\ell^k + \boldsymbol{\sigma}_\ell^k. \tag{18b}$$

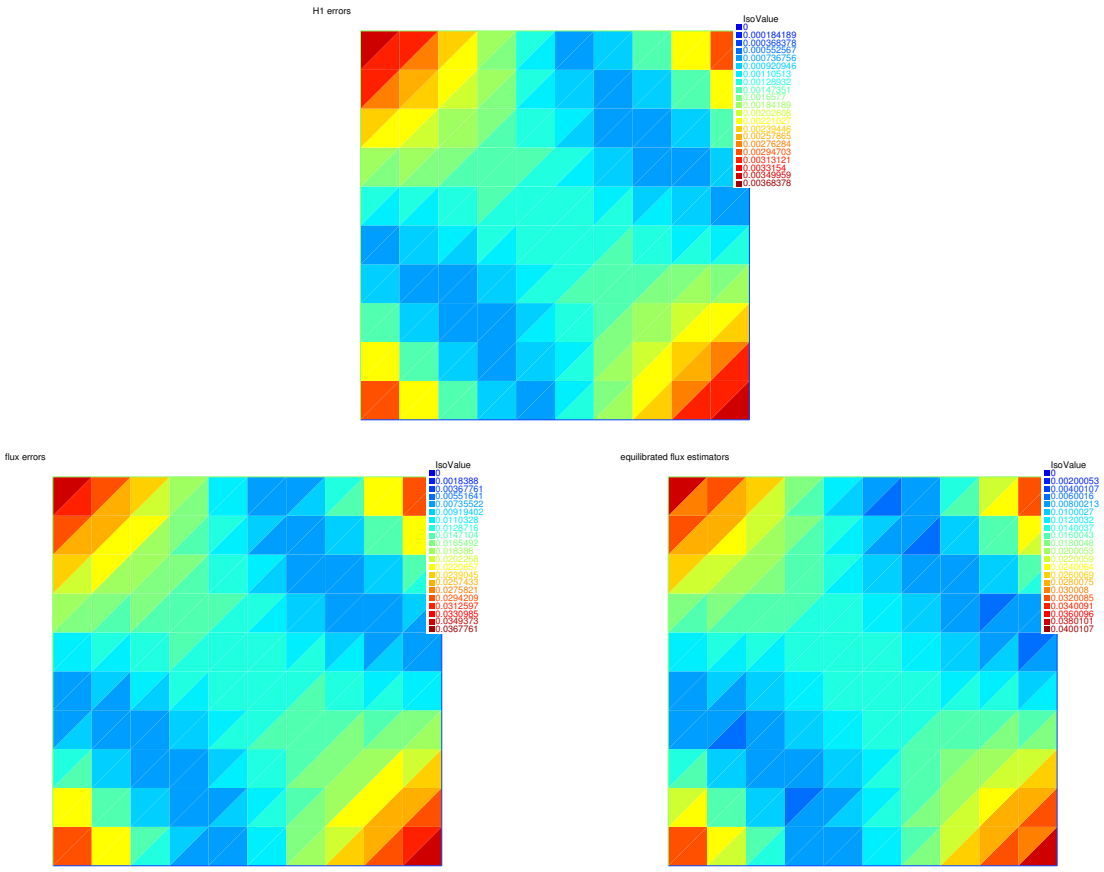

Figure 6: Elementwise H^1-seminorm errors $\|\nabla(u-u_\ell^3)\|_K$ (top), flux errors $\|a(|\nabla u|)\nabla u - a(|\nabla u_\ell^3|)\nabla u_\ell^3\|_K$ (bottom left), and the estimators $\big[\|a(|\nabla u_\ell^3|)\nabla u_\ell^3+\boldsymbol{\sigma}_\ell^3\|_K+\frac{h_K}{\pi}\|f-\Pi_{p'}f\|_K\big]$ (bottom right) at the third Picard iteration, $L=10$

We remark that $\|\boldsymbol{l}_\ell^k\|$ converges to 0 as the Picard iteration index k goes to infinity, because $|\nabla u_\ell^{k-1}| - |\nabla u_\ell^k| \to 0$. This allows us to split the upper bound as (neglecting the data oscillation error)

$$\eta_\ell^k = \|a(|\nabla u_\ell^k|)\nabla u_\ell^k + \boldsymbol{\sigma}_\ell^k\| \leq \underbrace{\|\boldsymbol{d}_\ell^k\|}_{\text{discretization est. } \eta_{\ell,\mathrm{dis}}^k} + \underbrace{\|\boldsymbol{l}_\ell^k\|}_{\text{linearization est. } \eta_{\ell,\mathrm{lin}}^k}.$$

We then can stop the solver based on the criterion

$$\eta_{\ell,\mathrm{lin}}^{\overline{k}} < \gamma \eta_{\ell,\mathrm{dis}}^{\overline{k}}, \tag{19}$$

where $\gamma > 0$ is a user-defined parameter for the desired residual part of the linearization estimator $\eta_{\ell,\mathrm{lin}}^{\overline{k}}$ that is in the script given by `real gamma = 0.1`. To run this exercise, set `bool DoEx5 = 1`. Then run the script. This will produce `Exercise5.plt` which can be executed by `>gnuplot Exercise5.plt`. Remark on the number of iterates required compared to the fixed criteria based on the difference of consecutive iterates. Try several different values of γ to see if you can change the number of iterates.

Answer 5. (Adaptive stopping criteria)

The answer is given in Figure 8. We first remark that the two quantities $\eta_{\ell,\mathrm{lin}}^k$ and $\eta_{\ell,\mathrm{dis}}^k$ are computable, so this is a valid criterion. Next we observe that the criterion is achieved in only 3 iterations, exactly half of what was required for the naive criteria solely based on the difference of consecutive iterates.

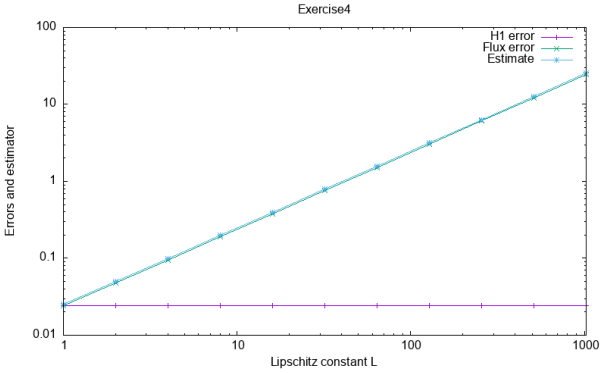

Figure 7: The H^1-seminorm error (9), the flux error (15), and the estimator (13) at Picard "convergence", when the error between the consecutive iterates (8) is below $1e-7$, in function of the Lipschitz constant L

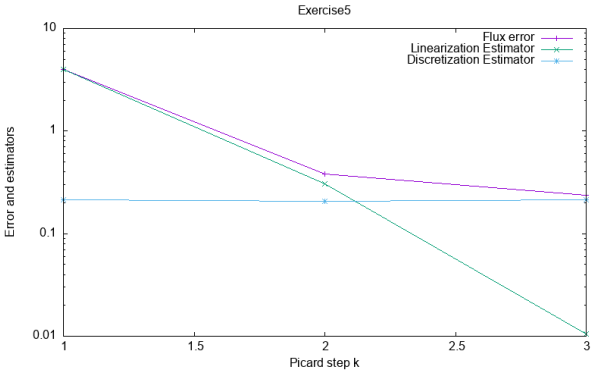

Figure 8: Adaptive stopping criterion (19) with the parameter $\gamma = 0.1$