3 Nonlinear Problems

We introduce a general paradigm for residual-based adaptivity in the approximation of nonlinear variational problems. This provides the abstract framework for the later application to various types of nonlinear problems in solid and fluid mechanics including also eigenvalue and optimal control problems.

****

3.1 Preliminaries

Primal and dual Galerkin approximations of linear problems:

$$a(u,\varphi) = F(\varphi) \quad \forall \varphi \in V, \qquad a(u_h,\varphi_h) = F(\varphi_h) \quad \forall \varphi_h \in V_h$$

$$a(\psi,z) = J(\psi) \quad \forall \psi \in V, \qquad a(\psi_h,z_h) = J(\psi_h) \quad \forall \psi_h \in V_h$$

Primal solution u, dual solution z:

$$J(u) = a(u,z) = F(z)$$

Primal error $e := u - u_h$, dual error $e^* := z - z_h$
(by Galerkin orthogonality):

$$J(e) = a(e,z) = a(e,e^*) = a(u,e^*) = F(e^*)$$

$$J(e) = a(e, z-\psi_h) = \underbrace{F(z-\psi_h) - a(u_h, z-\psi_h)}_{=:\,\rho(u_h)(z-\psi_h)} \quad \psi_h \in V_h$$

$$F(e^*) = a(u-\varphi_h, e^*) = \underbrace{J(u-\varphi_h) - a(u-\varphi_h, z_h)}_{=:\,\rho^*(z_h)(u-\varphi_h)} \quad \varphi_h \in V_h$$

$$\Rightarrow \qquad J(e) = \frac{1}{2}\rho(u_h)(z-\psi_h) + \frac{1}{2}\rho^*(z_h)(u-\varphi_h), \quad \varphi_h, \psi_h \in V_h$$

3.2 Galerkin approximation of nonlinear variational equations

Variational equation in function space V :

$$A(u)(\cdot) = 0, \qquad \text{target quantity } J(u)$$

Galerkin approximation in finite dimensional subspaces $V_h \subset V$:

$$A_h(u_h)(\cdot) = 0, \qquad \text{error measure } J(u) - J(u_h)$$

Example: Nonlinear convection-diffusion equation (Burgers equation)

$$-\nu\Delta u + u\cdot\nabla u = f, \quad V = H_0^1(\Omega)^d$$

$$A(u)(\varphi) := \nu(\nabla u \nabla\varphi) + (u \cdot \nabla u, \varphi) - (f, \varphi)$$

Formal Euler-Lagrange approach:

'Dual' variable z ('Lagrangian multiplier')

Lagrangian functional $\quad Ł(u, z) := J(u) - A(u)(z)$

(P) Stationary point $\{u, z\} \in V \times V$:

$$Ł'(u, z) = 0 \quad \Leftrightarrow \quad \left\{ \begin{array}{l} J'(u)(\varphi) - A'(u)(\varphi, z) \\ -A(u)(\psi) \end{array} \right\} = 0 \quad \forall\{\varphi, \psi\}$$

(P_h) Galerkin approximation $\{u_h, z_h\} \in V_h \times V_h$:

$$Ł'(u_h, z_h) = 0 \quad \Leftrightarrow \quad \left\{ \begin{array}{l} J'(u_h)(\varphi_h) - A'(u_h)(\varphi_h, z_h) \\ -A(u_h)(\psi_h) \end{array} \right\} = 0 \quad \forall\{\varphi_h, \psi_h\}$$

Goal: Estimation of error $J(u) - J(u_h)$ in terms of 'primal' and 'dual' residuals:

$$\rho(u_h)(\cdot) := -A(u_h)(\cdot)$$
$$\rho^*(u_h, z_h)(\cdot) := J'(u_h)(\cdot) - A'(u_h)(\cdot, z_h)$$

Proposition. *We have the error identity*

$$J(u) - J(u_h) \; = \; \frac{1}{2}\underbrace{\rho(u_h)(z-\psi_h)}_{primal} \; + \frac{1}{2}\underbrace{\rho^*(u_h, z_h)(u-\varphi_h)}_{dual} \; + R_h^{(3)},$$

with arbitrary approximations $\varphi_h, \psi_h \in V_h$ *and a remainder* R_h *which is cubic in the primal and dual errors* $e := u-u_h$ *and* $e^* := z-z_h$ *,*

$$R_h^{(3)} = \frac{1}{2}\int_0^1 \big\{ J'''(u_h+se)(e,e,e) - A'''(u_h+se)(e,e,e,z_h+se^*) \\ - 3A''(u_h+se)(e,e,e^*)\big\}\, s(s-1)\big\}\, ds$$

Proof. Set $x := \{u, z\}$, $x_h := \{u_h, z_h\}$, $e := x-x_h$, and $L(x) := \text{Ł}(u,z)$.

$$\begin{aligned} J(u) - J(u_h) &= L(x) - L(x_h) - \underbrace{A(u)(z)}_{=0} + \underbrace{A_h(u_h)(z_h)}_{=0} \\ &= \int_0^1 L'(x_h+se)(e)\, ds - \frac{1}{2}\big\{L'(x_h)(e) + \underbrace{L'(x)(e)}_{=0}\big\} \\ &\quad + \frac{1}{2}L'(x_h)(e) \end{aligned}$$

Error representation of the trapezoidal rule

$$\int_0^1 f(t)\, dt = \frac{1}{2}\big\{f(0) + f(1)\big\} + \frac{1}{2}\int_0^1 f''(s)s(s-1)\, ds$$

yields

$$J(u) - J(u_h) = \frac{1}{2}L'(x_h)(e) + \frac{1}{2}\int_0^1 L'''(x_h+se)(e,e,e)\, s(s-1)\, ds$$

By Galerkin orthogonality

$$\begin{aligned} L'(x_h)(e) &= L'(x_h)(x-y_h) + \underbrace{L'(x_h)(y_h-x_h)}_{=0}, \quad y_h \in V_h \times V_h \\ &= J'(u_h)(u-\varphi_h) - A'(u_h)(u-\varphi_h, z_h) - A(u_h)(z-\psi_h) \end{aligned}$$

Q.E.D.

Remarks:

1. The derivation of the error representation does not require the uniqueness of solutions (important for application to eigenvalue problems). The a priori assumption $x_h \to x$ $(h \to 0)$ makes the result meaningful for cases with non-unique solutions.

2. The evaluation of the error identity requires guesses for the primal and dual solution u and z.

3. The cubic remainder term $R_h^{(3)}$ is usually neglected.

4. The solution of the dual problem takes only a 'linear' work unit.

Proposition. *There holds the simplified error representation:*

$$J(u) - J(u_h) = \rho(u_h)(z - \varphi_h) + R_h^{(2)},$$

for arbitrary $\varphi_h \in V_h$ *, with the quadratic remainder*

$$R_h^{(2)} := \int_0^1 \big\{ A''(u_h + se)(e, e, z) - J''(u_h + se)(e, e) \big\}\, ds.$$

Proof. By integration by parts

$$\begin{aligned} R_h^{(2)} &= -\int_0^1 \big\{ A'(u_h + se)(e, z) - J'(u_h + se)(e) \big\}\, ds \\ &\quad + \underbrace{A'(u)(e, z) - J'(u)(e)}_{=0} \\ &= -A(u)(z) + A(u_h)(z) + J(u) - J(u_h) \\ &= -\rho(u_h)(z) + J(u) - J(u_h) = -\rho(u_h)(z - \varphi_h) + J(u) - J(u_h) \end{aligned}$$

Q.E.D.

Remark: Application of the abstract theory to the Galerkin approximation of the Navier-Stokes equations with the quadratic nonlinearity $(u \cdot \nabla u, \cdot)_{L^2}$ yields remainder terms of the form

$$R_h^{(2)} = 2(e^u \cdot \nabla e^u, z), \qquad R_h^{(3)} = -\frac{1}{2}(e^u \cdot \nabla e^u, e^z)$$

3.3 Nested Solution Approach

For solving the nonlinear problems by a Galerkin finite element method, we employ the following iterative scheme. Starting from a coarse initial mesh $\mathbb{T}_0$, a hierarchy of refined meshes

$$\mathbb{T}_0 \subset \mathbb{T}_1 \subset \cdots \subset \mathbb{T}_l \subset \cdots \subset \mathbb{T}_L$$

and corresponding finite element spaces V_l, $l = 1, ..., L$, is generated by a nested solution process.

1. *Initialization:* For $j = 0$, compute the solution $u_0 \in V_0$ on the coarsest mesh $\mathbb{T}_0$.

2. *Defect correction iteration:* For $l \geq 1$, start with $u_l^{(0)} = u_{l-1} \in V_l$.

3. *Iteration step:* For computed iterate $u_l^{(j)}$ evaluate the defect

   $$(d_l^{(j)}, \varphi) = -A(u_l^{(j)})(\varphi), \quad \varphi \in V_l$$

   and solve the correction equation

   $$\tilde{A}'(u_l^{(j)})(v_l^{(j)}, \varphi) = (d_l^{(j)}, \varphi) \quad \forall \varphi \in V_l$$

   by Krylov-space or multigrid iterations using the hierarchy of already constructed meshes $\{\mathbb{T}_l, ..., \mathbb{T}_0\}$. Update

   $$u_l^{(j+1)} = u_l^{(j)} + v_l^{(j)}$$

   set $j = j + 1$ and go back to (2). This process is repeated until a limit $u_l \in V_l$ is reached within a certain prescribed accuracy.

4. *Error estimation:* Solve the (linearized) discrete dual problem

   $$z_l \in V_l : \quad A'(u_l)(\varphi, z_l) = J(\varphi) \quad \forall \varphi \in V_l$$

   and evaluate the a posteriori error representation and estimate

   $$J(e_l) \approx \tilde{E}(u_l), \quad |\tilde{E}(u_l)| \leq \tilde{\eta}(u_l)$$

   If $|\tilde{E}(u_l)| \leq TOL$, or $N_l \geq N_{max}$, then stop. Otherwise cell-wise mesh adaptation yields the new mesh $\mathbb{T}_{l+1}$. Then, set $l = l + 1$ and go back to (1).

3.4 Application to eigenvalue problems

Eigenvalue problems of particular interest:

- The *symmetric* eigenvalue problem of the Laplace operator:

$$-\Delta u = \lambda u$$

- The *nonsymmetric* eigenvalue problem of a convection-diffusion operator:

$$-\Delta u + b \cdot \nabla u = \lambda u$$

- The *stability* eigenvalue problem governed by the linearized Navier-Stokes operator:

$$-\nu\Delta v + \hat{v} \cdot \nabla v + v \cdot \nabla \hat{v} + \nabla p = \lambda v, \quad \nabla \cdot v = 0$$

  where $\hat{v}$ is some 'base solution' the stability of which is to be investigated.

Eigenvalue problem in (complex) function space V
($m(\cdot,\cdot)$ semi-scalar product):

$$a(u,\varphi) = \lambda\, m(u,\varphi) \quad \forall \varphi \in V, \quad \lambda \in \mathbb{C}, \quad m(u,u) = 1$$

Galerkin approximation in finite dimensional subspaces $V_h \subset V$:

$$a(u_h,\varphi_h) = \lambda_h\, m(u_h\varphi_h) \quad \forall \varphi_h \in V_h, \quad \lambda_h \in \mathbb{C}, \quad m(u_h,u_h) = 1$$

Goal: *Control of error in eigenvalue* $\lambda - \lambda_h$ *in terms of the residual*

$$\rho(u_h,\lambda_h)(\cdot) := \lambda_h\, m(u_h,\cdot) - a(u_h,\cdot)$$

A posteriori error analysis

Embedding into the general framework of variational equations

$$U := \{u, \lambda\} \in \mathcal{V} := V \times \mathbb{C}, \qquad U_h := \{u_h, \lambda_h\} \in \mathcal{V}_h := V_h \times \mathbb{C}$$

Semilinear form for $\Phi = \{\varphi, \mu\} \in \mathcal{V}$:

$$A(U)(\Phi) := \lambda\, m(u,\varphi) - a(u,\varphi) + \overline{\mu}\, \underbrace{\{m(u,u) - 1\}}_{\text{normalization}}$$

Compact variational formulation of eigenvalue problems

$$\begin{aligned} A(U)(\Phi) &= 0 \quad \forall \Phi \in \mathcal{V} \\ A(U_h)(\Phi_h) &= 0 \quad \forall \Phi_h \in \mathcal{V}_h \end{aligned}$$

Error control functional

$$J(\Phi) := \mu\, m(\varphi, \varphi), \qquad \Phi = \{\varphi, \mu\}$$

$$m(u,u) = 1 \quad \Rightarrow \quad J(U) = \lambda$$

Dual solutions $Z = \{z, \pi\} \in \mathcal{V}$, $Z_h = \{z_h, \pi_h\} \in \mathcal{V}_h$:

$$\begin{aligned} A'(U)(\Phi, Z) &= J'(U)(\Phi) \qquad \forall \Phi \in \mathcal{V} \\ A'(U_h)(\Phi_h, Z_h) &= J'(U_h)(\Phi_h) \quad \forall \Phi_h \in \mathcal{V}_h \end{aligned}$$

Detailed form of adjoint equations: $U = \{u, \lambda\},\ \Phi = \{\varphi, \mu\},\ Z = \{z, \pi\}$

$$\begin{aligned} A'(U)(\Phi, Z) \;&=\; \lambda\, m(\varphi, z) - a(\varphi, z) + \mu\, m(u, z) + 2\bar{\pi}\mathrm{Re}\, m(\varphi, u) \\ &\quad + \mu\{m(u, u) - 1\} \end{aligned}$$

$$J'(U)(\Phi) \;=\; \mu\, m(u, u) + 2\lambda \mathrm{Re}\, m(\varphi, u)$$

Observing that $m(u, u) = 1$, the dual problem takes the form

$$\bar{\lambda}\, m(z, \varphi) - a(\varphi, z) + \mu\, \{m(u, z) - 1\} + 2\{\bar{\pi} - \lambda\}\mathrm{Re}\, m(\varphi, u) \;=\; 0$$

for all $\Phi = \{\varphi, \mu\}$. This system is equivalent to the 'dual' eigenvalue problem

$$a(\varphi, z) = \bar{\pi}\, m(\varphi, z) \quad \forall \varphi \in V, \quad \bar{\pi} = \lambda, \quad m(u, z) = 1$$

or identifying $z = u^*$ and $\pi = \lambda^*$,

$$a(\varphi, u^*) = \bar{\lambda}^*\, m(\varphi, u^*) \quad \forall \varphi \in V, \quad m(u, u^*) = 1$$

The discrete adjoint problem is equivalent to

$$a(\varphi_h, u_h^*) = \bar{\lambda}_h^*\, m(\varphi_h, u_h^*) \quad \forall \varphi_h \in V_h, \quad m(u_h, u_h^*) = 1$$

Associated dual residual:

$$\rho^*(u_h^*, \lambda_h^*)(\cdot) := \bar{\lambda}_h^*\, m(\cdot, u_h^*) - a(\cdot, u_h^*)$$

Proposition. *We have the error representation*

$$\lambda - \lambda_h \;=\; \frac{1}{2}\rho(u_h, \lambda_h)(u^* - \psi_h) + \frac{1}{2}\rho^*(u_h^*, \lambda_h^*)(u - \varphi_h) + \; R_h$$

for arbitrary $\psi_h, \varphi_h \in V_h$, *with the cubic remainder term*

$$R_h \;=\; \frac{1}{2}(\lambda - \lambda_h)\, m(u - u_h, u^* - u_h^*)$$

Proof: Setting $E := \{u-u_h, \lambda-\lambda_h\}$ and $E^* := \{u^*-u_h^*, \lambda^*-\lambda_h^*\}$, the general remainder term from the abstract theory has the form

$$R_h^{(3)} = \frac{1}{2}\int_0^1 \big\{J'''(U_h+sE)(E,E,E) - A'''(U_h+sE)(E,E,E,Z_h+sE^*) \\ - 3A''(U_h+sE)(E,E,E^*)\big\}\, s(s-1)\, ds$$

In the present case, by a simple calculation, we have

$$\begin{aligned} J'''(U_h+sE)(E,E,E) &= 6(\lambda-\lambda_h)m(u-u_h, u-u_h), \\ A'''(U_h+sE)(E,E,E,Z_h+sE^*) &= 0 \\ -3A''(U_h+sE)(E,E,E^*) &= -6(\lambda-\lambda_h)\, m(u-u_h, u^*-u_h^*) \\ &\quad - 6(\overline{\lambda^*-\lambda_h^*})\, m(u-u_h, u-u_h) \end{aligned}$$

Consequently, noting that $\lambda-\lambda_h = \overline{\lambda^*-\lambda_h^*}$, it follows that

$$\begin{aligned} R_h^{(3)} &= -3\int_0^1 (\lambda-\lambda_h)\, m(u-u_h, u^*-u_h^*) s(s-1)\, ds \\ &= \tfrac{1}{2}(\lambda-\lambda_h)\, m(u-u_h, u^*-u_h^*) \end{aligned}$$

which completes the proof.

****

Remarks:

- Simultaneous solution of primal and adjoint eigenvalues problems is necessary within an optimal multigrid solver of nonsymmetric eigenvalue problems.
- Case of multiple eigenvalues can easily be treated.
- No assumption about the multiplicity of the eigenvalue necessary.
- Error estimates for functionals $j(u)$ of eigenfunctions.
- Case of additional approximation in operator $\mathcal{A} = \mathcal{A}(\hat{u})$ (stability eigenvalue problem).

Evaluation of eigenvalue residuals

On a polygonal/hedral domain $\Omega \subset \mathbb{R}^d$ consider the eigenvalue problem

$$\mathcal{A}v := -\Delta v + b\cdot\nabla v \;=\; \lambda v \quad \text{in } \Omega, \quad \mathrm{v}_{|\partial\Omega} = 0$$

with a smooth (or even constant) transport coefficient b, and $\mathcal{M} := \mathrm{id}$. This eigenvalue problem is approximated by the Galerkin method using piecewise linear or d-linear finite elements on meshes $\mathbb{T}_h = \{K\}$.

Within this setting, we can proceed analogously as before, obtaining

$$\begin{aligned}
\rho(u_h,\lambda_h)(\psi) &= \lambda_h\, m(u_h,\psi) - a(u_h,\psi)\\
&= \sum_{K\in\mathbb{T}_h} \big\{(\lambda_h\mathcal{M}u_h - \mathcal{A}u_h,\psi)_K - (\partial_n^{\mathcal{A}}u_h,\psi)_{\partial K}\big\}\\
&= \sum_{K\in\mathbb{T}_h} \big\{(\lambda_h\mathcal{M}u_h - \mathcal{A}u_h,\psi)_K - \tfrac{1}{2}([\partial_n^{\mathcal{A}}u_h],\psi)_{\partial K}\big\}
\end{aligned}$$

and

$$\begin{aligned}
\rho^*(u_h^*,\lambda_h^*)(\varphi) &= \lambda_h^*\, m(\varphi,z_h) - a(\psi,z_h)\\
&= \sum_{K\in\mathbb{T}_h} \big\{(\varphi,\lambda_h^*\mathcal{M}z_h - \mathcal{A}^*z_h)_K - (\varphi,\partial_n^{\mathcal{A}^*}z_h)_{\partial K}\big\}\\
&= \sum_{K\in\mathbb{T}_h} \big\{(\varphi,\lambda_h^*\mathcal{M}z_h - \mathcal{A}^*z_h)_K - \tfrac{1}{2}(\varphi,[\partial_n^{\mathcal{A}^*}z_h])_{\partial K}\big\}
\end{aligned}$$

Hence, using again the notation of 'equation' and 'jump residuals', the primal residual admits the estimate

$$|\rho(u_h,\lambda_h)(u^* - i_h u^*)| \le \sum_{K\in\mathbb{T}_h} \rho_K\omega_K^*$$

with the cell-residuals and weights defined by

$$\begin{aligned}
\rho_K &:= \big(\|R(u_h,\lambda_h)\|_K^2 + h_K^{-1/2}\|r(u_h)\|_{\partial K}^2\big)^{1/2}\\
\omega_K^* &:= \big(\|u^* - I_h u^*\|_K^2 + h_K^{1/2}\|u^* - I_h u^*\|_{\partial K}^2\big)^{1/2}
\end{aligned}$$

Correspondingly, for the dual residual:

$$|\rho^*(u_h^*, \lambda_h^*)(u - i_h u)| \le \sum_{K\in\mathbb{T}_h} \rho_K^* \omega_K$$

with

$$\rho_K^* := \big(\|R^*(u_h^*, \lambda_h^*)\|_K^2 + h_K^{-1/2}\|r^*(u_h^*)\|_{\partial K}^2\big)^{1/2}$$
$$\omega_K := \big(\|u - I_h u\|_K^2 + \tfrac{1}{2} h_K^{1/2}\|u - I_h u\|_{\partial K}^2\big)^{1/2}$$

Proposition. *Within the above setting, assuming that*

$$|m(u - u_h, u^* - u_h^*)| \le 1$$

we have the a posteriori error estimate

$$|\lambda - \lambda_h| \;\le\; \eta_\lambda^\omega := \sum_{K\in\mathbb{T}_h} \big\{\rho_K \omega_K^* + \rho_K^* \omega_K\big\}$$

Proof: By the above estimates,

$$|\lambda - \lambda_h| \le \tfrac{1}{2} \sum_{K\in\mathbb{T}_h} \big\{\rho_K \omega_K^* + \rho_K^* \omega_K\big\} + R_h$$

Since, by the assumption,

$$|R_h| = \tfrac{1}{2}|(\lambda - \lambda_h)\, m(u - u_h, u^* - u_h^*)| \le \tfrac{1}{2}\,|\lambda - \lambda_h|$$

the asserted estimate follows.

Remark: From the above 'weighted' a posteriori error estimate, we can derive the following 'weight-free' estimate

$$|\lambda - \lambda_h| \;\le\; \eta_\lambda^{(1)} := c_\lambda \sum_{K\in\mathbb{T}_h} \big\{\rho_K^2 + \rho_K^{*2}\big\}$$

with a constant $c_\lambda = \mathcal{O}(|\lambda|)$.

Numerical example (V. Heuveline 2001)

Convection-diffusion problem:

$$-\Delta v + b\cdot\nabla v \;=\; \lambda v \quad \text{in } \Omega, \quad \text{v} = 0 \;\; \text{auf } \partial\Omega$$

$b = (0, b_y)^T\,,\;\; \Omega = (-1,1)\times(-1,3)$

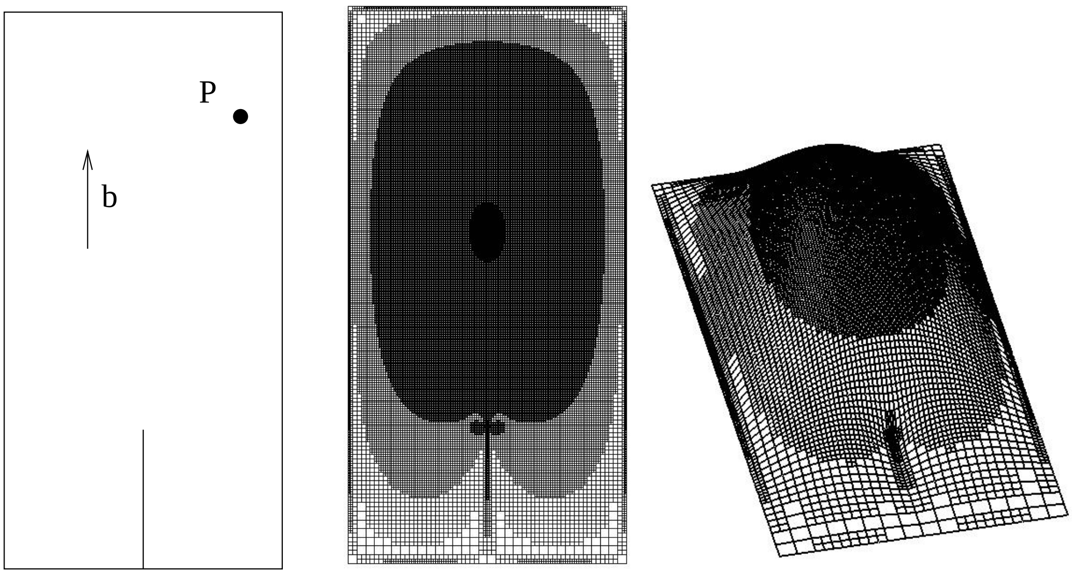

Figure. Configuration for $b \equiv 0$ (left), adapted mesh with $12,000$ cells (middle), normlized eigenfunctions (right).

Refinement indicators:

$$\eta_\lambda^{(1)} := \sum_{K\in\mathbb{T}_h} h_K^2\big\{\rho_K^2 + \rho_K^{*2}\big\}$$

$$\eta_\lambda^{red} := \sum_{K\in\mathbb{T}_h} h_K^2\rho_K^2$$

$$\eta_\lambda^{(2)} := \Big(\sum_{K\in\mathbb{T}_h} h_K^4\big\{\rho_K^2 + \rho_K^{*2}\big\}\Big)^{1/2}$$

$$\eta_\lambda^{\omega} := \sum_{K\in\mathbb{T}_h} h_K^2\big\{\rho_K\tilde{\omega}_K^* + \rho_K^*\tilde{\omega}_K\big\}$$

Test case 1: Symmetric problem

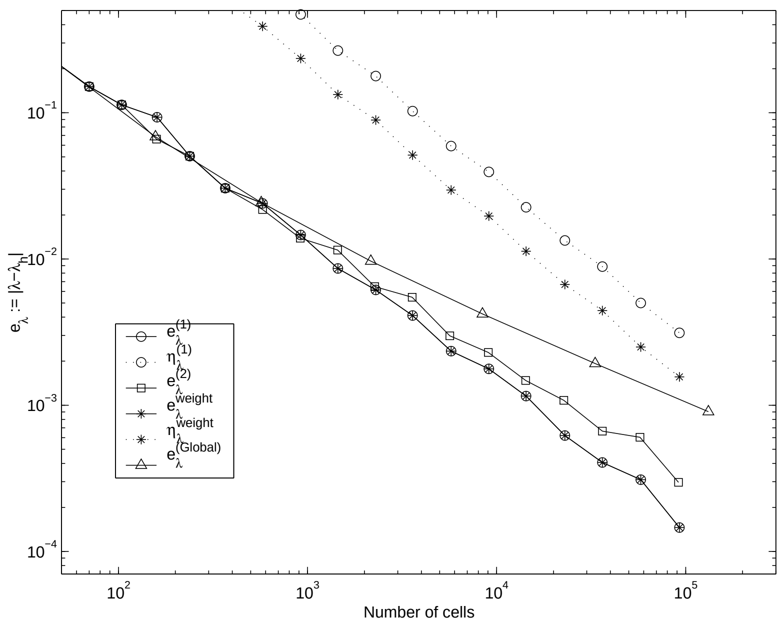

Figure $\eta_\lambda^{(1)}$ ('O'), $\eta_\lambda^{(2)}$ ('□'), η_λ^{weight} ('$*$'), uniform ('△').

Test case 2: Vertical transport $b_y = 3$

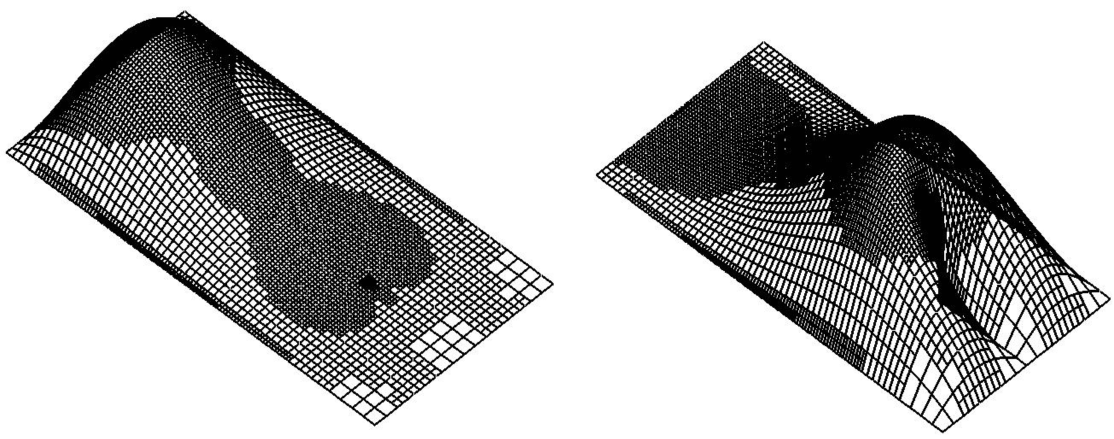

Figure. Primal (left) and dual eigenfunction (right) .

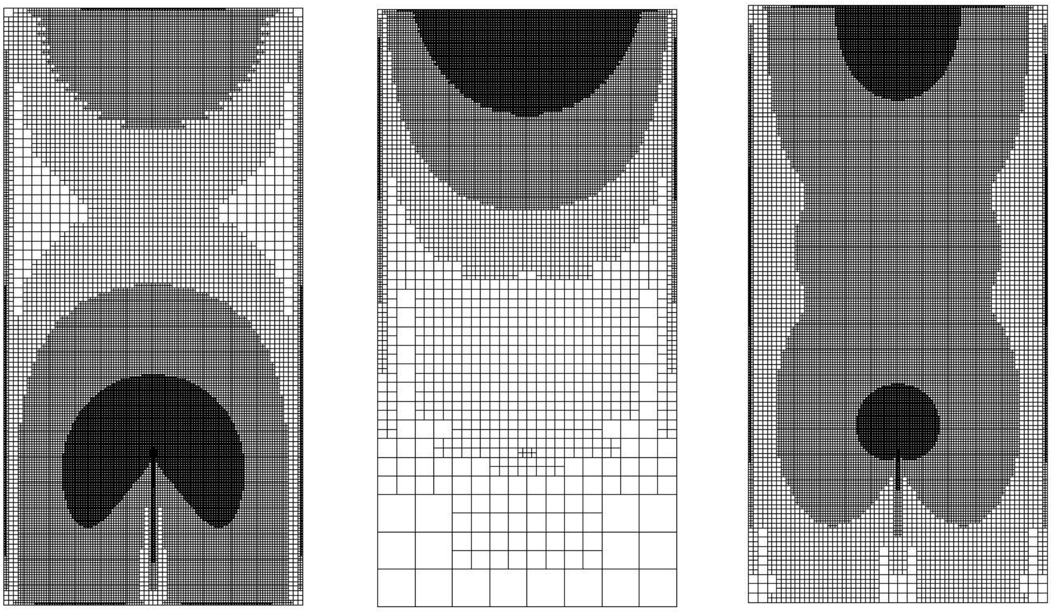

Figure. Adapted mesh with 10,000 cells by $\eta_\lambda^{(1)}$ (left), η_λ^{red} (middle), η_λ^{weight} (right) (fixed-rate strategy with $X = 20\%$.

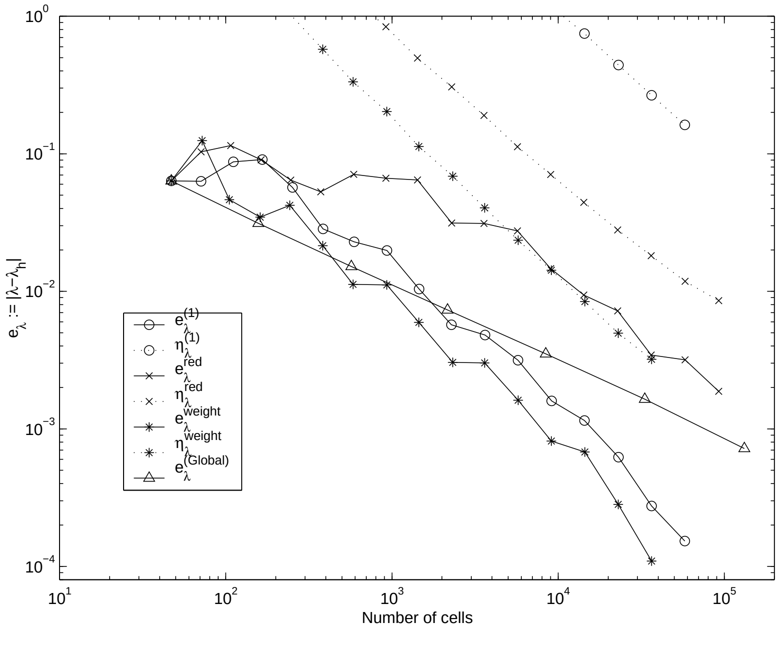

Figure. $\eta_\lambda^{(1)}$ ('O'), η_λ^{red} ('$\times$'), η_λ^{weight} ('$*$'), uniform ('$\triangle$').