Adaptive solutions of inverse problems

Mihai Alexe and Adrian Sandu
{mihai,sandu}@vt.edu

DDDAS@ICCS 2011, June 3, 2011

Inverse problems

Inverse problems (IPs):

- Use measurements to infer the values of the defining parameters for a given model.
- Examples: data assimilation (geosciences), parameter estimation (fluid dynamics), image reconstructions (medical imaging).
- Usually set as PDE-constrained optimization problems (one can arrive at this formulation by starting from a probabilistic framework).

The need for adaptivity

- State-of-the-art forward model solvers are adaptive in space and time to maximize efficiency.
- Adaptive solvers can refine the mesh and the time step only where needed, to capture and track phenomena of interest, and to perform as few computations as possible.
- Previous research efforts have preferred the static approach due to the difficulties introduced by adaptive methods.
- However, there is a growing trend towards the use of space time adaptivity in the inverse problem community.

Summary

Problem:

- Quantify the numerical errors in the (optimal) solution of the inverse problem
- Use adaptivity in time-space to control this error

Solution: develop a new, fully discrete framework for the solution of inverse problems.

- Analyze properties of discrete adjoints for adaptive algorithms:
  - h/p-space and time mesh refinements,
  - solution limiters,
  - grid transfer operators, etc
- Discretizations of choice: DG in space + RK in time
- *A priori* and *a posteriori* error analyses of the discrete KKT equations.
- Multiple resolution optimization with error-based adaptive mesh refinement.

Duality framework for space-time inverse problems

Consider the following inverse problem:

$$\min_{\mathbf{u}^0,\mathbf{g},\mathbf{f}} \mathcal{J} \int_0^T \int_\Omega J_\Omega \left[C_\Omega\, \mathbf{u} \right] \mathrm{d}\mathbf{x}\, \mathrm{d}t + \int_0^T \int_\Gamma J_\Gamma \left[C_\Gamma\, \mathbf{u} \right] \mathrm{d}s\, \mathrm{d}t + \int_\Omega K_\Omega \left[E_\Omega\, \mathbf{u} \right]_{t=T} \mathrm{d}\mathbf{x}\,,$$

subject to

$$\begin{aligned} \mathbf{u}_t &= N\left[\mathbf{u}\right] + \mathbf{f}\,,\ \mathbf{x} \in \Omega\,,\ t \in [0, T] \\ B[\mathbf{u}] &= \mathbf{g}\,,\ \mathbf{x} \in \Gamma\,,\ t \in [0, T] \\ \mathbf{u}(t=0,\mathbf{x}) &= \mathbf{u}^0\,,\ \mathbf{x} \in \Omega\,. \end{aligned}$$

Proposition (Alexe and Sandu, 2010)

The adjoint equation is well posed, if the differential operators that define the model and cost functional satisfy a set of three compatibility conditions on $\bar{\Omega} \times [0, T]$.

Discrete adjoints of adaptive time stepping algorithms

Proposition (Alexe and Sandu, 2009b)

Discrete adjoints of adaptive time stepping algorithms are not a priori consistent. Post processing is required to restore the accuracy of the discrete adjoint trajectory.

How can we fix this?

- zero out the spurious adjoint gradients
- implement a correction during post-processing

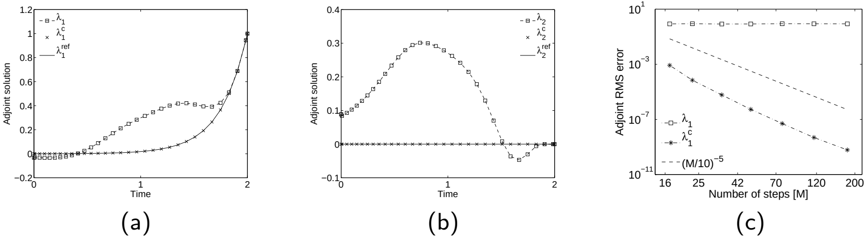


Figure: DA solutions: inconsistent adjoint (a), consistent adjoint (b), and RMS errors (c) for the Prothero - Robinson IVP. The reference solution was was calculated with ode45 (MATLAB).

Dual consistency of grid transfer operators

Consider two spatial meshes Ω^h_n, at t^n (or iteration n), and Ω^h_{n+1} at t^{n+1}.
Define the mesh transfer operators:

$$\mathcal{I}_{n+1\to n} : \Omega^h_n \to \Omega^h_{n+1} \,,$$
$$\mathcal{I}_{n\to n+1} : \Omega^h_{n+1} \to \Omega^h_n$$

Proposition (Alexe and Sandu, 2010, 2011a)

Intergrid transfer operators for h/p-adaptive DG are dual consistent, since they satisfy the transpose relationship

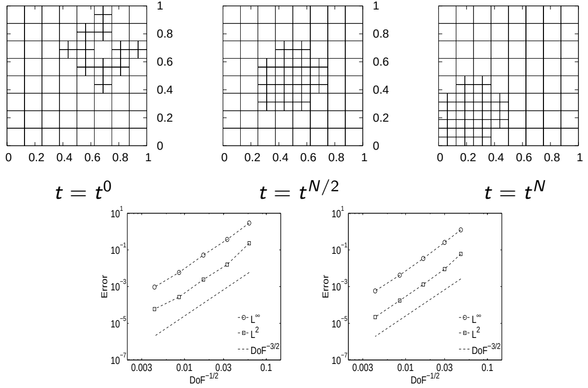
$$\mathcal{I}_{n+1\to n} = \left(\mathcal{I}_{n\to n+1}\right)^T \;.$$

- This implies that adjoint grid transfer operators (generated by AD) are consistent, and yield correct results in practice.
- No need to decouple the solution transfer code from the numerical core when differentiating.
- Proof holds for general mesh refinement (not only for hierarchical structures).

Dual consistency of space-time RK-DG discretizations

Proposition (Alexe and Sandu, 2010)

Space-time RK-DG discretizations that are dual consistent in space are also dual consistent in time. Moreover, they inherit the temporal order of the primal discretization.



Continuous adjoint   Discrete adjoint

Figure: Two-dimensional advection. Time-averaged L^2 and L^∞ errors with $p = 2$.

Mathematical formulation of the inverse problems

Discretize–first

$$
\begin{aligned}
\text{Find } \mathbf{q}_*^h \quad &= \quad \arg \min_{\mathbf{q}^h \in \mathcal{Q}_h, \mathbf{u}^h \in \mathcal{U}_h} \mathcal{J}_h(\mathbf{u}^h, \mathbf{q}^h) \ , \\
&\quad \text{subject to } \mathcal{A}^h(\mathbf{u}^h, \, \mathbf{q}^h)\,(\mathbf{w}^h) = 0 \ , \ \forall \, \mathbf{w}^h \in \mathcal{U}_h \ .
\end{aligned}
$$

Differentiate–first

$$
\begin{aligned}
\text{Find } \mathbf{q}_* \quad &= \quad \arg \min_{\mathbf{q} \in \mathcal{Q}, \mathbf{u} \in \mathcal{U}} \mathcal{J}(\mathbf{u}, \mathbf{q}) \ , \\
&\quad \text{subject to } \mathcal{A}(\mathbf{u}, \, \mathbf{q})\,(\mathbf{w}) = 0 \ , \ \forall \, \mathbf{w} \in \mathcal{U} \ .
\end{aligned}
$$

Goal: control the error in the optimal solution $\mathbf{q}^h$.

Optimality systems

We work under suitable *a priori* structural assumptions (Lions, 1971) on $\mathcal{J}$, $\mathcal{J}_h$, $\mathcal{A}$, and $\mathcal{A}^h$ that guarantee the existence of multipliers $\boldsymbol{\lambda}_* \in \mathcal{U}$, and $\boldsymbol{\lambda}_*^h \in \mathcal{U}_h$, such that, given optimal solution pairs $\{\mathbf{u}_*, \mathbf{q}_*\}$, $\{\mathbf{u}_*^h, \mathbf{q}_*^h\}$, we have:

KKT system in function spaces

$$
\begin{aligned}
\mathcal{A}\,[\mathbf{u}_*, \mathbf{q}_*]\,(\mathbf{w}) &= 0\,,\ \forall\,\mathbf{w} \in \mathcal{U}\,,\\
\mathcal{A}_{\mathbf{u}}\,[\mathbf{u}_*, \mathbf{q}_*]\,(\boldsymbol{\lambda}_*)\,(\psi_{\mathbf{u}}) &= \mathcal{J}_{\mathbf{u}}\,[\mathbf{u}_*, \mathbf{q}_*]\,(\psi_{\mathbf{u}})\,,\ \forall\,\psi_{\mathbf{u}} \in \mathcal{U}\,,\\
\mathcal{A}_{\mathbf{q}}\,[\mathbf{u}_*, \mathbf{q}_*]\,(\boldsymbol{\lambda}_*)\,(\psi_{\mathbf{q}}) &= \mathcal{J}_{\mathbf{q}}\,[\mathbf{u}_*, \mathbf{q}_*]\,(\psi_{\mathbf{q}})\,,\ \forall\,\psi_{\mathbf{q}} \in \mathcal{Q}\,.
\end{aligned}
$$

Discrete KKT system

$$
\begin{aligned}
\mathcal{A}^h[\mathbf{u}_*^h, \mathbf{q}_*^h](\psi_{\mathbf{u}}^h) &= 0\,,\ \forall\,\psi_{\mathbf{u}}^h \in \mathcal{U}_h\,,\\
\mathcal{A}_{\boldsymbol{\lambda}^h}^h[\mathbf{u}_*^h, \mathbf{q}_*^h](\psi_{\mathbf{u}}^h)(\boldsymbol{\lambda}_*^h) &= \mathcal{J}_{\mathbf{u}^h}^h[\mathbf{u}_*^h, \mathbf{q}_*^h](\psi_{\mathbf{u}}^h)\,,\ \forall\,\psi_{\mathbf{u}}^h \in \mathcal{U}_h\,,\\
\mathcal{A}_{\mathbf{q}^h}^h[\mathbf{u}_*^h, \mathbf{q}_*^h](\psi_{\mathbf{q}}^h) &= \mathcal{J}_{\mathbf{q}^h}^h[\mathbf{u}_*^h, \mathbf{q}_*^h](\psi_{\mathbf{q}}^h)\,,\ \forall\,\psi_{\mathbf{q}}^h \in \mathcal{Q}_h\,.
\end{aligned}
$$

Primal, dual, and optimality consistency

Definition (Primal consistency)

The primal discretization is said to be consistent if the exact solutions **u** and **q** to the weak form primal equation satisfy:

$$\mathcal{A}^h[\mathbf{u},\mathbf{q}](\psi_{\mathbf{u}}) = 0 \ , \ \forall \ \psi_{\mathbf{u}} \in \mathcal{U} \ .$$

Definition (Dual consistency)

The primal discretization is said to be dual consistent, if any triplet $\boldsymbol{\xi} = \{\mathbf{u},\boldsymbol{\lambda},\mathbf{q}\} \in \mathcal{X}$ that verifies the weak-form primal and dual equations, also satisfies:

$$\mathcal{A}^h_{\mathbf{u}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{u}})(\boldsymbol{\lambda}) = \mathcal{J}^h_{\mathbf{u}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{u}}) \ , \ \forall \ \psi_{\mathbf{u}} \in \mathcal{U} \ .$$

Definition (Optimality equation consistency)

The optimality equation discretization is said to be consistent, if any triplet $\boldsymbol{\xi} = \{\mathbf{u},\boldsymbol{\lambda},\mathbf{q}\} \in \mathcal{X}$ that verifies the primal and dual equations, also satisfies:

$$\mathcal{A}^h_{\mathbf{q}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{q}})(\boldsymbol{\lambda}) = \mathcal{J}^h_{\mathbf{q}^h}[\mathbf{u},\mathbf{q}](\psi_{\mathbf{q}}) \ , \ \forall \ \psi_{\mathbf{q}} \in \mathcal{Q} \ .$$

Model problem

Cost functional:

$$\mathcal{J}(\mathbf{u},\mathbf{q}) = \frac{1}{2}\|\mathcal{H}\mathbf{u}-\mathbf{o}\|^2_{\mathcal{L}^2(\Omega)} + \frac{1}{2}\|\nabla(\mathbf{q}-\mathbf{q}_B)\|^2_{\mathcal{L}^2(\Omega)} + \frac{\boldsymbol{\beta}}{2}\|\mathbf{q}-\mathbf{q}_B\|^2_{\mathcal{L}^2(\Omega)} .$$

PDE constraint (primal problem):

$$\begin{aligned} -\nabla\cdot(\mathbf{q}(\mathbf{x})\,\nabla\mathbf{u}) &= \mathbf{f}(\mathbf{x})\,,\ \mathbf{x}\in\Omega \\ \mathbf{u} &= \mathbf{g}(\mathbf{x})\,,\ \mathbf{x}\in\Gamma=\partial\Omega\,, \end{aligned}$$

Adjoint problem:

$$\begin{aligned} -\nabla\cdot(\mathbf{q}\,\nabla\boldsymbol{\lambda}) &= \mathcal{H}^*(\mathcal{H}\mathbf{u}-\mathbf{o})\,,\ \mathbf{x}\in\Omega \\ \boldsymbol{\lambda} &= 0\,,\ \mathbf{x}\in\Gamma\,. \end{aligned}$$

Optimality condition:

$$\begin{aligned} -\Delta\,\mathbf{q}+\boldsymbol{\beta}\,(\mathbf{q}-\mathbf{q}_B) &= -\Delta\mathbf{q}_B+\nabla\mathbf{u}\cdot\nabla\boldsymbol{\lambda}\,,\ \mathbf{x}\in\Omega\,, \\ \nabla\mathbf{q}\cdot\vec{\mathbf{n}} &= \nabla\mathbf{q}_B\cdot\vec{\mathbf{n}}\,,\ \mathbf{x}\in\Gamma\,. \end{aligned}$$

Primal SIPG DG discretization

Primal problem: Find $\mathbf{u}^h \in \mathcal{U}_h^p$ s.t., $\forall\, \mathbf{w}^h \in \mathcal{U}_h^p$, we have:

$$\begin{aligned}
\mathcal{N}^h(\mathbf{u}^h, \mathbf{w}^h) &= \int_\Omega \mathbf{f}^h\, \mathbf{w}^h\, \mathrm{d}\mathbf{x} + \mathcal{B}^h(\mathbf{g}^h, \mathbf{w}^h) \\
\mathcal{N}^h(\mathbf{u}^h, \mathbf{w}^h) &:= \int_\Omega \mathbf{q}^h\, \nabla \mathbf{u}^h \cdot \nabla \mathbf{w}^h\, \mathrm{d}\mathbf{x} + \int_{\Gamma_\mathcal{I} \cup \Gamma} \phi\, [\![\mathbf{u}^h]\!] \cdot [\![\mathbf{w}^h]\!]\, \mathrm{d}\mathbf{s} \\
&\quad - \int_{\Gamma_\mathcal{I} \cup \Gamma} \left([\![\mathbf{u}^h]\!] \cdot \{\mathbf{q}^h\, \nabla \mathbf{w}^h\} + \{\mathbf{q}^h\, \nabla \mathbf{u}^h\} \cdot [\![\mathbf{w}^h]\!] \right) \mathrm{d}\mathbf{s}\,, \\
\mathcal{B}^h(\mathbf{g}^h, \mathbf{w}^h) &:= -\int_\Gamma \mathbf{q}^h\, \mathbf{g}^h\, \nabla \mathbf{w}^h \cdot \vec{\mathbf{n}}\, \mathrm{d}\mathbf{s} + \int_\Gamma \phi\, \mathbf{g}^h\, \mathbf{w}^h\, \mathrm{d}\mathbf{s}\,.
\end{aligned}$$

$$\|\mathbf{v}\|_{\mathrm{DG}} := \left(\sum_{\kappa \in \mathcal{T}_h} \int_\kappa \mathbf{q}\, \nabla \mathbf{v} \cdot \nabla \mathbf{v}\, \mathrm{d}\mathbf{x} + \sum_{e \in \Gamma_\mathcal{I} \cup \Gamma} \hat{\phi}\, h^{-1} \int_e [\![\mathbf{v}]\!] \cdot [\![\mathbf{v}]\!]\, \mathrm{d}\mathbf{s} \right)^{1/2}, \ \forall\, \mathbf{v} \in \mathcal{U}\,.$$

Theorem

For sufficiently large penalty parameter $\hat{\phi} > 0$, there exists C independent of h such that

$$\left\| \mathbf{u} - \mathbf{u}^h \right\|_{\mathrm{DG}} \leq C\, h^{\min(p+1,s)-1} \|\mathbf{u}\|_{\mathcal{H}^s(\mathcal{T}_h)}\,.$$

Discrete dual problem

- Discrete cost functional may have an asymptotically vanishing term introduced for consistency of the dual (Hartmann, 2007):

$$
\begin{aligned}
\mathcal{J}_h(\mathbf{u}^h, \mathbf{q}^h) \quad = \quad & \frac{1}{2}\int_\Omega (\mathcal{H}^h \mathbf{u}^h - \mathbf{o}^h)^T (\mathcal{H}^h \mathbf{u}^h - \mathbf{o}^h)\, \mathrm{d}\mathbf{x} \\
& + \frac{1}{2}\int_\Omega (\nabla \mathbf{q}^h - \nabla \mathbf{q}_B^h)^T (\nabla \mathbf{q}^h - \nabla \mathbf{q}_B^h)\, \mathrm{d}\mathbf{x} \\
& + \frac{\beta}{2}\int_\Omega (\mathbf{q}^h - \mathbf{q}_B^h)^T (\mathbf{q}^h - \mathbf{q}_B^h)\, \mathrm{d}\mathbf{x} + \mathcal{R}_h(\mathbf{u}^h, \mathbf{q}^h)\ .
\end{aligned}
$$

- Dual problem:

$$
\begin{gathered}
\text{Find } \boldsymbol{\lambda} \in \mathcal{U}_h^p \text{ such that:} \\
\mathcal{N}^{h\,*}(\mathbf{w}^h, \boldsymbol{\lambda}^h) = \frac{\partial \mathcal{J}_h}{\partial \mathbf{u}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{w}^h)\ ,\ \forall\, \mathbf{w}^h \in \mathcal{U}_h\ .
\end{gathered}
$$

  - For the SIPG primal discretization, the discrete dual is consistent without the need of $\mathcal{R}_h$ (Hartmann, 2007).

Discrete optimality equation

Proposition

With a suitable consistent modification to $\mathcal{J}_h$, the discrete optimality condition can be made consistent with its continuous counterpart.

Proof:

- Let $\widehat{\mathbf{q}}^h := \mathbf{q}^h - \mathbf{q}_B^h$.
- Differentiate the primal discretization and cost functional along $\mathbf{z}^h \in \mathcal{Q}_h^r$, to get:

$$\int_\Omega \nabla\widehat{\mathbf{q}}^h \cdot \nabla \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \beta \int_\Omega \widehat{\mathbf{q}}^h \, \mathbf{z}^h \, \mathrm{d}\mathbf{x} + \frac{\partial \mathcal{R}_h}{\partial \widehat{\mathbf{q}}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h)$$
$$= \int_\Omega \left(\nabla \mathbf{u}^h \cdot \nabla \boldsymbol{\lambda}^h\right) \mathbf{z}^h \, \mathrm{d}\mathbf{x} - \int_{\Gamma_{\mathcal{I}}} \left([\![\mathbf{u}^h]\!] \cdot \{\mathbf{z}^h \nabla \boldsymbol{\lambda}^h\} + \{\mathbf{z}^h \nabla \mathbf{u}^h\} \cdot [\![\boldsymbol{\lambda}^h]\!] \right) \mathrm{d}\mathbf{s}$$
$$- \int_\Gamma \left(\mathbf{u}^h - \mathbf{g}^h\right) \mathbf{z}^h \nabla \boldsymbol{\lambda}^h \cdot \vec{\mathbf{n}} \, \mathrm{d}\mathbf{s} - \int_\Gamma \boldsymbol{\lambda}^h \nabla \mathbf{u}^h \cdot \vec{\mathbf{n}} \, \mathbf{z}^h \, \mathrm{d}\mathbf{s} \, , \ \forall \, \mathbf{z}^h \in \mathcal{Q}_h^r \, .$$

Strictly speaking, this discretization is consistent, but it has no stabilization (penalty) terms!

Consistent modification to $\mathcal{J}_h$

- To introduce the stabilization terms we modify the target functional as follows:

$$\mathcal{R}_h(\mathbf{u}^h, \widehat{\mathbf{q}}^h) := \frac{1}{2}\int_{\Gamma_{\mathcal{I}}^q} \phi^q \, [\![\widehat{\mathbf{q}}^h]\!] \cdot [\![\widehat{\mathbf{q}}^h]\!] \, \mathrm{d}\mathbf{s} - \int_{\Gamma_{\mathcal{I}}^q} [\![\widehat{\mathbf{q}}^h]\!] \cdot \{\nabla \widehat{\mathbf{q}}^h\} \, \mathrm{d}\mathbf{s} \, .$$

- Then:

$$\frac{\partial \mathcal{R}_h}{\partial \mathbf{q}^h}[\mathbf{u}^h, \mathbf{q}^h](\mathbf{z}^h) \quad = \quad \int_{\Gamma_{\mathcal{I}}^q} \phi^q \, [\![\widehat{\mathbf{q}}^h]\!] \cdot [\![\mathbf{z}^h]\!] \, \mathrm{d}\mathbf{s} - \int_{\Gamma_{\mathcal{I}}^q} \Big([\![\mathbf{z}^h]\!] \cdot \{\nabla \widehat{\mathbf{q}}^h\} + [\![\widehat{\mathbf{q}}^h]\!] \cdot \{\nabla \mathbf{z}^h\} \Big) \, \mathrm{d}\mathbf{s} \, .$$

Proposition (Alexe and Sandu, 2011)

For all $p \geq 1$ and $s > 3/2$, the discrete optimality equation (for the modified cost function) is asymptotically consistent.

A priori error analysis for the optimal solution

Proposition (Alexe and Sandu, 2011b)

The following a priori bound holds for the optimal solution error:

$$\|\mathbf{q}_*^h - \mathbf{q}_*\|_{\mathcal{H}^s(\mathcal{T}_h^q)} \quad \leq \quad C(r,p)\, h^{\min(p+1,s)-3/2} \left(\|\mathbf{u}\|_{\mathcal{H}^s(\mathcal{T}_h)} + \|\boldsymbol{\lambda}\|_{\mathcal{H}^s(\mathcal{T}_h)}\right) \ .$$

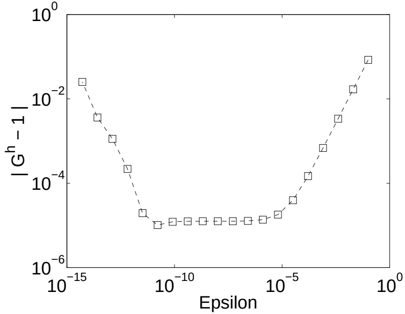
Proof. The equation for the discrete optimal solution error is a perturbed SIPG discretization. Bound these perturbations and asses their impact on the solution via Lax Milgram theorem.

Gradient consistency on adaptive meshes

$$\mathbf{q}_B = \mathbf{q}_* := 5 + \sin\left(\frac{\pi x}{2}\right)\cos(2\pi y)\ ,\ \mathbf{q}_0 := 1\ ,$$

$$\mathbf{u}_*(\mathbf{x}) := 10\ \exp\left(-\frac{(10x-5)^2}{4}\right)\exp\left(-\frac{(10y-5)^2}{4}\right), \Omega \equiv [0,1]^2\ .$$

$$\lim_{\varepsilon\to 0} G^h(\varepsilon) := \lim_{\varepsilon\to 0}\frac{\mathcal{J}^h(\mathbf{q}^h+\varepsilon\,\boldsymbol{\delta}\mathbf{q}^h)-\mathcal{J}^h(\mathbf{q}^h)}{\varepsilon\,\left\langle\nabla_{\mathbf{q}^h}\mathcal{J}^h,\,\boldsymbol{\delta}\mathbf{q}^h\right\rangle_\Omega} = 1\ .$$



A posteriori error estimation based on an error functional $E[\mathbf{q}]$

Let the error functional be defined as:

$$E[\mathbf{q}] : \mathcal{Q} \to \mathbb{R}$$

We seek an error representation of the type:

$$E[\mathbf{q}] - E[\mathbf{q}^h] \approx e_h + \text{h.o.t.}\ .$$

Let $\mathcal{M}$ be the associated Lagrangian for the error functional $E[\mathbf{q}]$:

$$\begin{aligned} \mathcal{M}[\boldsymbol{\xi}, \boldsymbol{\sigma}] &= E[\mathbf{q}] - \mathcal{L}_{\boldsymbol{\xi}}\left[\xi\right]\left(\boldsymbol{\sigma}_{\mathbf{u}}, \boldsymbol{\sigma}_{\boldsymbol{\lambda}}, \boldsymbol{\sigma}_{\mathbf{q}}\right) \\ &= E[\mathbf{q}] - \mathcal{L}_{\mathbf{q}}\left[\xi\right]\left(\boldsymbol{\sigma}_{\mathbf{q}}\right) - \mathcal{L}_{\mathbf{u}}\left[\xi\right]\left(\boldsymbol{\sigma}_{\mathbf{u}}\right) \end{aligned}$$

- The last equality holds true if $\mathbf{u} = U[\mathbf{q}]$.
- U is the solution operator for the primal equation.
- If $(\boldsymbol{\xi}_*, \boldsymbol{\sigma}_*)$ is a stationary point of $\mathcal{M}$, then $\boldsymbol{\xi}_*$ is a stationary point of $\mathcal{L}$, since the first order optimality conditions for $\mathcal{L}$ are imposed as constraints in $\mathcal{M}$.

A posteriori estimation

The equation for $\boldsymbol{\delta}\mathbf{q}$ reads

$$j_{\mathbf{q},\mathbf{q}}[\mathbf{q}_*]\,(\phi, \boldsymbol{\sigma}_\mathbf{q}) = E_\mathbf{q}[\mathbf{q}_*]\,(\phi)\ , \quad \forall \phi \in \mathcal{Q}\,.$$

- In computations, we replace $(j_{\mathbf{q},\mathbf{q}}[\mathbf{q}_*])^{-1}$ by a BFGS approximation.

The equation for $\boldsymbol{\delta\lambda}$ is the tangent linear model evaluated at the optimal solution:

$$0 = \mathcal{A}_\mathbf{u}\,[\boldsymbol{\xi}_*]\,(\boldsymbol{\sigma}_\mathbf{u}) + \mathcal{A}_\mathbf{q}\,[\boldsymbol{\xi}_*]\,(\boldsymbol{\sigma}_\mathbf{q}) \quad \Leftrightarrow \quad \boldsymbol{\sigma}_\mathbf{u} = U'[\mathbf{q}_*]\,\boldsymbol{\sigma}_\mathbf{q}\,.$$

The equation for $\boldsymbol{\delta}\mathbf{u}$ is the second order adjoint model evaluated at the optimal solution:

$$\begin{aligned} 0 \quad &= \quad \mathcal{J}_{\mathbf{u},\mathbf{u}}\,[\boldsymbol{\xi}_*]\,(\boldsymbol{\sigma}_\mathbf{u}) + \mathcal{J}_{\mathbf{q},\mathbf{u}}\,[\boldsymbol{\xi}_*]\,(\boldsymbol{\sigma}_\mathbf{q}) \\ & \quad\ -\mathcal{A}_\mathbf{u}\,[\boldsymbol{\xi}_*]\,(\boldsymbol{\sigma}_\lambda) - \mathcal{A}_{\mathbf{u},\mathbf{u}}\,[\boldsymbol{\xi}_*]\,(\boldsymbol{\sigma}_\mathbf{u}) - \mathcal{A}_{\mathbf{q},\mathbf{u}}\,[\boldsymbol{\xi}_*]\,(\boldsymbol{\sigma}_\mathbf{q})\,. \end{aligned}$$

The resulting error estimate

Proposition (Alexe and Sandu, 2011b)

Combining the equations we are lead to the following error estimate:

$$
\begin{aligned}
E[\xi_*^h] - E[\xi_*] &= \mathcal{L}_{\xi}\left[\xi_*^h\right](\boldsymbol{\sigma}) + h.o.t. \\
&= \mathcal{A}[\mathbf{u}^h, \mathbf{q}^h](\Delta\boldsymbol{\sigma}_{\mathbf{u}}) \\
&\quad + \mathcal{J}_{\mathbf{u}}[\mathbf{u}^h, \mathbf{q}^h](\Delta\boldsymbol{\sigma}_{\boldsymbol{\lambda}}) - \mathcal{A}_{\mathbf{u}}[\mathbf{u}^h, \mathbf{q}^h](\Delta\boldsymbol{\sigma}_{\boldsymbol{\lambda}}, \boldsymbol{\lambda}^h) \\
&\quad + \mathcal{J}_{\mathbf{q}}[\mathbf{u}^h, \mathbf{q}^h](\Delta\boldsymbol{\sigma}_{\mathbf{q}}) - \mathcal{A}_{\mathbf{q}}[\mathbf{u}^h, \mathbf{q}^h](\Delta\boldsymbol{\sigma}_{\mathbf{q}}, \boldsymbol{\lambda}^h) + h.o.t.
\end{aligned}
$$

For the particular case of our elliptic problem, we obtain the estimate in (Becker and Vexler, 2004)!

The error estimation algorithm

We obtain the following error estimation algorithm (Alexe and Sandu, 2011b):

Algorithm 8.1 Error estimation using the second order adjoint solution

1: Solve the Hessian equation for $\boldsymbol{\sigma}_{\mathbf{q}}$ using a quasi-Newton approximation of $j_{\mathbf{q},\mathbf{q}}$.
2: Given $\boldsymbol{\sigma}_{\mathbf{q}}$, solve the tangent linear model to obtain $\boldsymbol{\sigma}_{\mathbf{u}}$.
3: Given $\boldsymbol{\sigma}_{\mathbf{q}}$ and $\boldsymbol{\sigma}_{\mathbf{u}}$, solve the second order adjoint model to obtain $\boldsymbol{\sigma}_{\boldsymbol{\lambda}}$.
4: Estimate the element-wise error using the dual weighted residual formula.

Numerical Results: Convergence

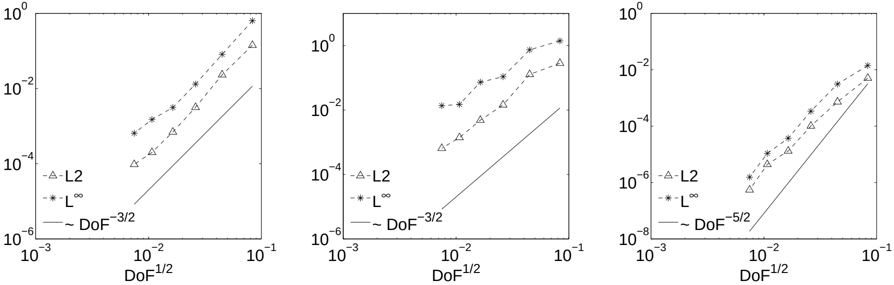


Figure: Convergence of the discrete optimal (left), primal (center), and dual (right) solutions for test B. The errors correspond to the converged solutions on each mesh level, and are plotted versus $h \sim \mathrm{DoF}^{-1/2}$.

Meshes generated by the aposteriori error control

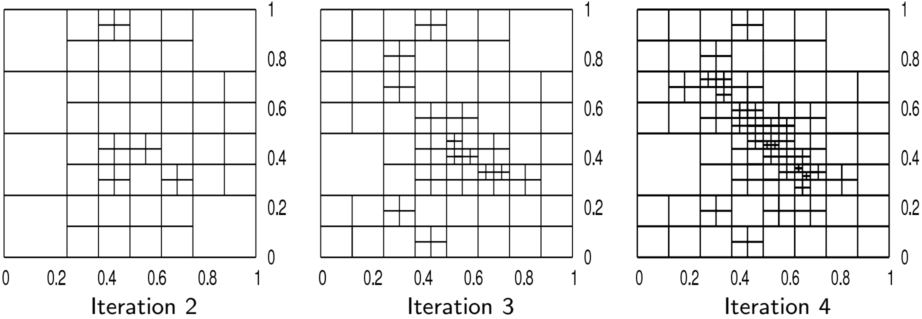


Figure: Optimization meshes generated by aposteriori error estimstion algorithm for numerical test B.

Conclusions

- The proposed discrete duality framework opens the possibility to perform inverse studies using *exclusively the discrete adjoint method*, throughout the entire inversion process.
- The discrete approach is very attractive in practice:
  - the discrete sensitivity equations can be generated with low effort using automatic differentiation.
- We analyze in detail the properties of discrete adjoints for the main ingredients of state–of–the–art adaptive algorithms:
  - Space and time mesh refinements,
  - Solution limiters,
  - Grid transfer operators,
  - Both *a priori* and *a posteriori* error analysis and estimation.
- Enable the construction of consistent and stable discretizations of inverse problems that benefit from all the adaptive features listed above.
- First effort at a completely discrete solution algorithm for adaptive inverse problems.
- The rules and guidelines given herein are applicable to models of high complexity.