Adaptive discretization strategies for the identification of distributed parameters in partial differential equations

Barbara Kaltenbacher, University of Klagenfurt

joint work with

Anke Griesbaum, University of Heidelberg
Alana Kirchner, Technical University of Munich
Slobodan Veljovic, University of Graz
Boris Vexler, Technical University of Munich

Overview

- motivation: parameter identification in PDEs
- ideas on adaptivity for inverse problems
- principles of goal oriented error estimators
- Tikhonov & discrepancy principle
- regularized Newton iterations & discrepancy principle
- outlook

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q:
- at jumps or large gradients or
- at locations with large error contribution

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q:
- at jumps or large gradients or
- at locations with large error contribution

$\longrightarrow$ location of large gradients / large errors a priori unknown

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

$\longrightarrow$ location of large gradients / large errors a priori unknown

$\longrightarrow$ general strategy for mesh generation separately for q and u
(example $-\nabla q(u)\nabla u) = f$)

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

$\longrightarrow$ location of large gradients / large errors a priori unknown

$\longrightarrow$ general strategy for mesh generation separately for q and u
(example $-\nabla q(u)\nabla u) = f$)

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

$\longrightarrow$ location of large gradients / large errors a priori unknown

$\longrightarrow$ general strategy for mesh generation separately for q and u
(example $-\nabla q(u)\nabla u) = f$)

instability $\Rightarrow$ regularization necessary !

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

$\longrightarrow$ location of large gradients / large errors a priori unknown

$\longrightarrow$ general strategy for mesh generation separately for q and u
(example $-\nabla q(u)\nabla u) = f$)

instability $\Rightarrow$ regularization necessary !

Regularization of Inverse Problems

- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $\left(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g)\right)$

Regularization of Inverse Problems

- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $\left(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g)\right)$
- only noisy data $g^\delta \approx g$ available (measurements): $\|g^\delta - g\| \leq \delta$

Regularization of Inverse Problems

- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $\left(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g)\right)$
- only noisy data $g^\delta \approx g$ available (measurements): $\|g^\delta - g\| \leq \delta$
- making $\|F(q) - g^\delta\|$ small $\not\Rightarrow$ good result for q!

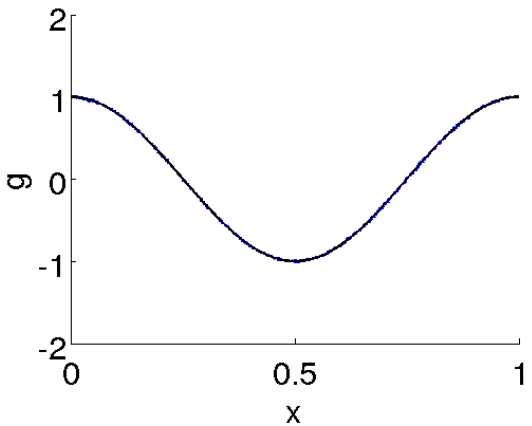
Regularization of Inverse Problems

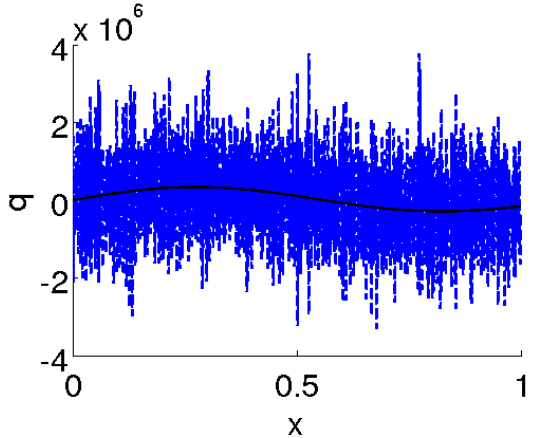
- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $\left(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g)\right)$
- only noisy data $g^\delta \approx g$ available (measurements): $\|g^\delta - g\| \leq \delta$
- making $\|F(q) - g^\delta\|$ small $\not\Rightarrow$ good result for q!

An Example

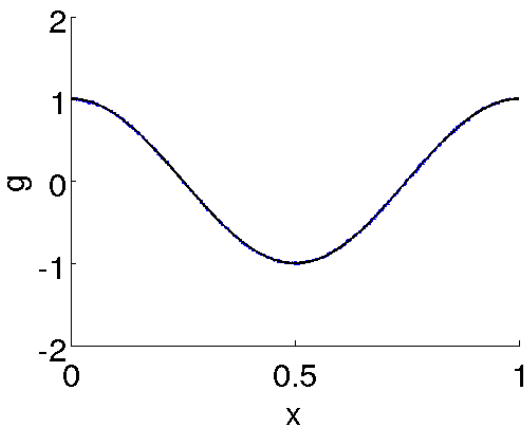
(Identification of a source term in a 1-d differential equation, $\delta = 1\%$)

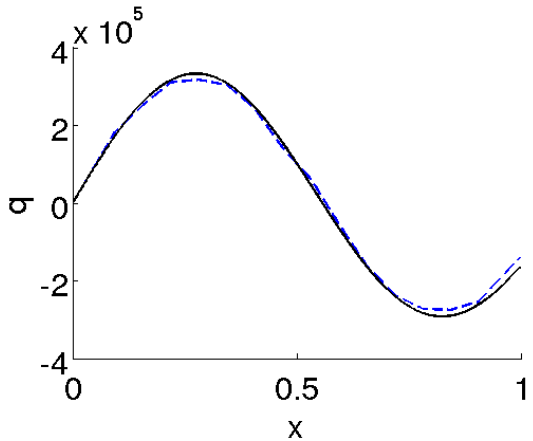
exact and noisy data g, g^δ exact q vs $\tilde{q}$ with $\|F(\tilde{q}) - g^\delta\| = 1.e-14$





exact and noisy data g, g^δ exact q vs $\tilde{q}$ with $\|F(\tilde{q}) - g^\delta\| = 2\delta$





Regularization of Inverse Problems

- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $\left(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g)\right)$
- only noisy data $g^\delta \approx g$ available (measurements): $\|g^\delta - g\| \leq \delta$
- making $\|F(q) - g^\delta\|$ small $\not\Rightarrow$ good result for q!
- regularization means approaching solution along stable path

Regularization of Inverse Problems

- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $\left(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g)\right)$
- only noisy data $g^\delta \approx g$ available (measurements): $\|g^\delta - g\| \leq \delta$
- making $\|F(q) - g^\delta\|$ small $\not\Rightarrow$ good result for q!
- regularization means approaching solution along stable path
- *regularization method*:
  family $(R_\alpha)_{\alpha>0}$ with parameter choice $\alpha = \alpha(g^\delta, \delta)$
  such that worst case convergence as $\delta \to 0$:

$$\sup_{\|g^\delta - g\| \leq \delta} \|R_{\alpha(g^\delta,\delta)}(g^\delta) - F^{-1}(g)\| \to 0 \text{ as } \delta \to 0$$

Regularization of Inverse Problems

- unstable operator equation: $F(q) = g$ with $F : q \mapsto u$ or $C(u)$
- solution $q = F^{-1}(g)$ does not depend continuously on g
  i.e., $\left(\forall (g_n), g_n \to g \;\not\Rightarrow\; q_n := F^{-1}(g_n) \to F^{-1}(g)\right)$
- only noisy data $g^\delta \approx g$ available (measurements): $\|g^\delta - g\| \leq \delta$
- making $\|F(q) - g^\delta\|$ small $\not\Rightarrow$ good result for q!
- regularization means approaching solution along stable path
- *regularization method*:
  family $(R_\alpha)_{\alpha>0}$ with parameter choice $\alpha = \alpha(g^\delta, \delta)$
  such that worst case convergence as $\delta \to 0$:

$$\sup_{\|g^\delta - g\| \leq \delta} \|R_{\alpha(g^\delta,\delta)}(g^\delta) - F^{-1}(g)\| \to 0 \text{ as } \delta \to 0$$

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

$\longrightarrow$ location of large gradients / large errors a priori unknown

$\longrightarrow$ general strategy for mesh generation separately for q and u
(example $-\nabla q(u)\nabla u) = f$)

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

$\longrightarrow$ location of large gradients / large errors a priori unknown

$\longrightarrow$ general strategy for mesh generation separately for q and u
(example $-\nabla q(u)\nabla u) = f$)

computational effort $\Rightarrow$ efficient numerical strategies necessary !

Motivation: Parameter Identification in PDEs

- instability: amplification of numerical errors
- computational effort:
  - large scale: several PDE solves per regularized inversion
  - several reg. inversions to determine regularization parameter

Example $-\Delta u = q$:
refine grid for u and q: • at jumps or large gradients or
• at locations with large error contribution

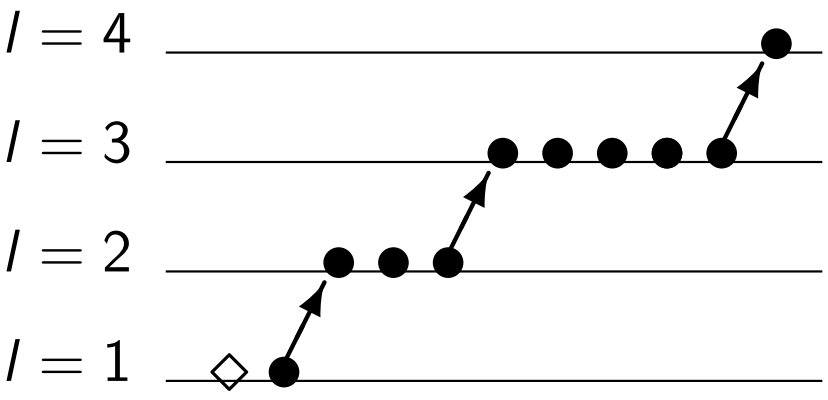
$\longrightarrow$ location of large gradients / large errors a priori unknown

$\longrightarrow$ general strategy for mesh generation separately for q and u
(example $-\nabla q(u)\nabla u) = f$)

computational effort $\Rightarrow$ efficient numerical strategies necessary !

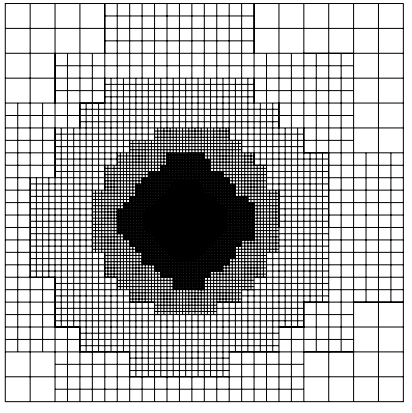
Efficient Methods for PDEs

multilevel iteration:



start with coarse discretization
refine successively

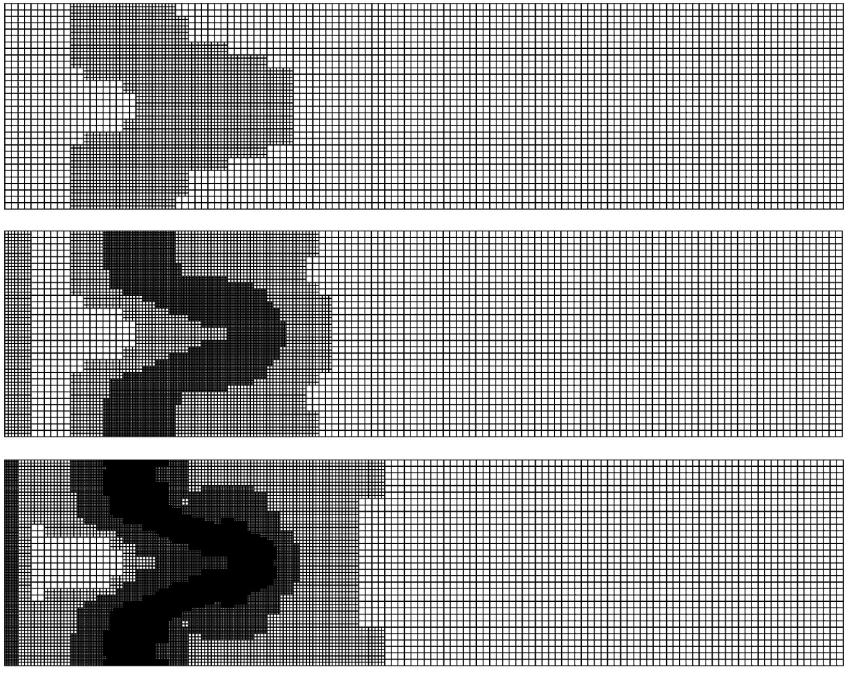
adaptive discretization:



coarse discretization where possible
fine grid only where necessary

Efficient Methods for PDEs

combined multilevel adaptive strategy:



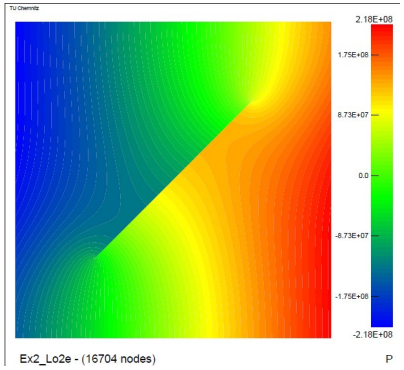
courtesy to [R.Becker&M.Braack&B.Vexler, App.Num.Math., 2005]

start on coarse grid

sucessive adaptive refinement

An Example: Crack Identification in Piezoelectric Ceramics

$$
\begin{aligned}
-DIV\,\vec{\sigma} &= 0 && \text{in } \Omega \\
-\nabla D &= 0 && \text{in } \Omega \\
N^T\vec{\sigma} &= 0 && \text{on } \partial\Omega \cup \Sigma \\
\phi &= \phi^e && \text{on } \Gamma_e = \partial\Omega \\
\vec{n}\cdot D &= 0 && \text{on } \Sigma
\end{aligned}
$$



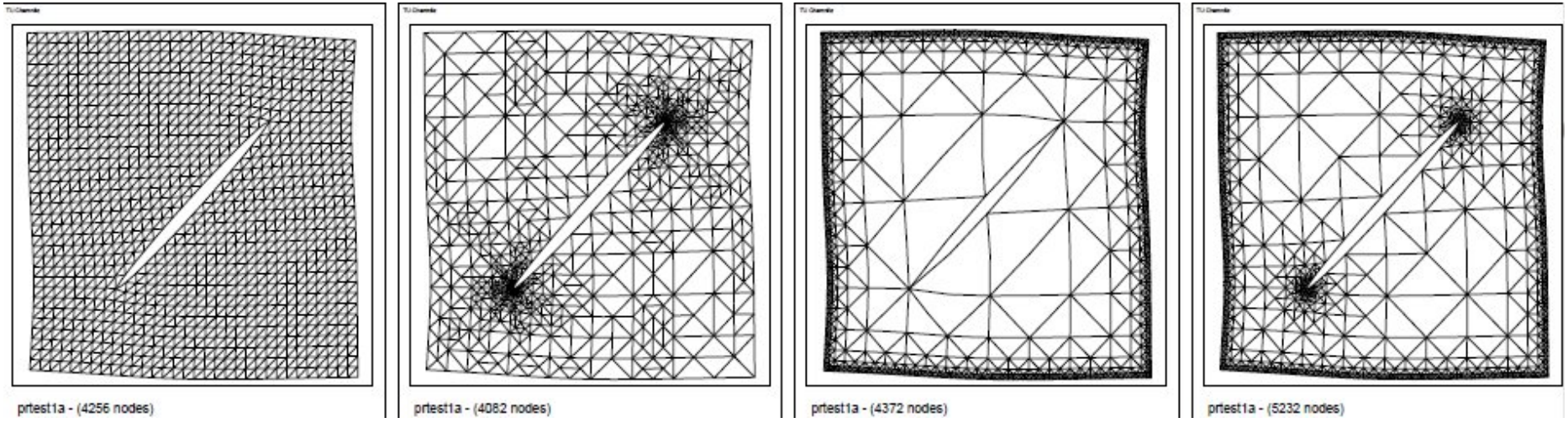
electric potential distribution

$$
\begin{aligned}
\vec{\sigma} &= \mathbf{c}^E DIV^T\vec{d} + \mathbf{e}^T\nabla\phi \ldots \text{stress} \\
\vec{D} &= \mathbf{e}DIV^T\vec{d} - \varepsilon^S\nabla\phi \ldots \text{dielectr. displacement} \\
DIV^T &= \tfrac{1}{2}(\nabla + \nabla^T), \quad \mathbf{c}^E, \mathbf{e}, \varepsilon^S \ldots \text{ material tensors} \\
\vec{d} &= (d_x, d_y, d_z) \ldots \text{mech. displacement} \\
\phi &\quad \ldots \text{electr. potential}
\end{aligned}
$$

Given additional boundary measurements $q^e = \vec{n}\cdot D$ and/or $\vec{d}$ on $\partial\Omega$, determine Σ. [Steinhorst&BK 2011]

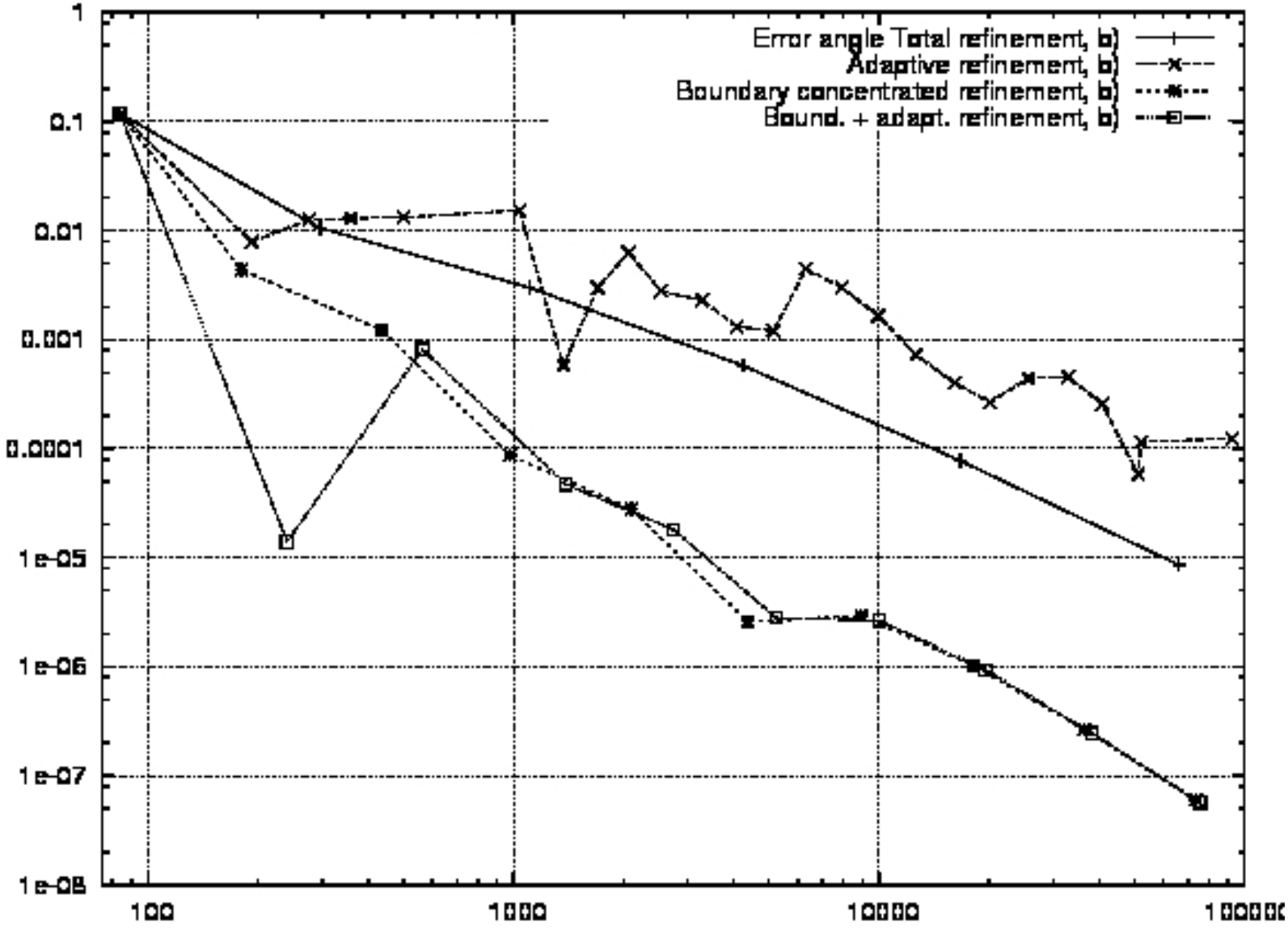
Different Refinement Strategies

$\sim$ 4000 nodes with ...

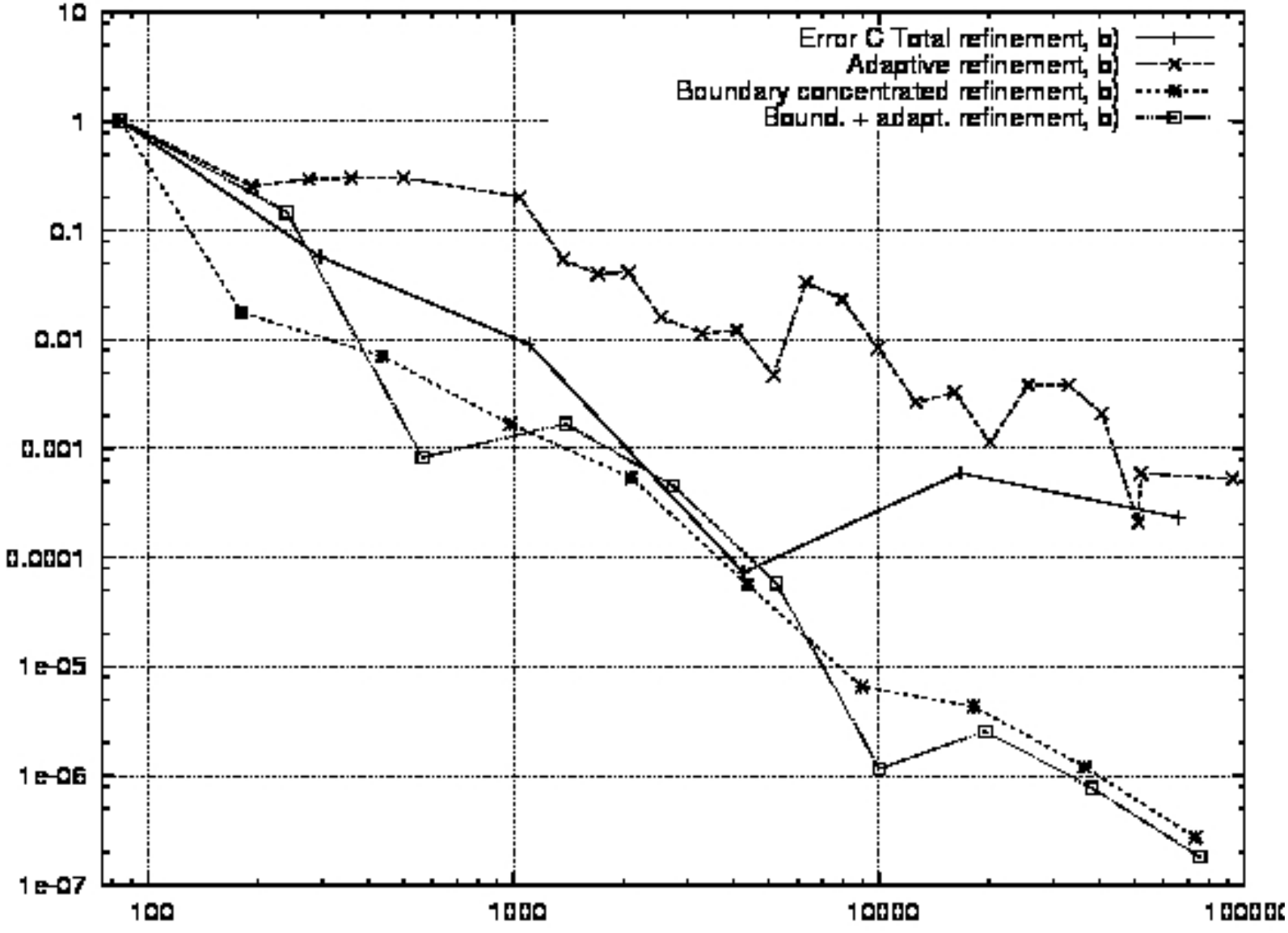


uniform | adaptive (residual based) | boundary concentrated | adaptive & bndy conc.

Error in crack plane angle with different refinement strategies:



Error in crack plane offset with different refinement strategies:



Errors with Different Refinement Strategies

$\sim$ 20.000 nodes:

	uniform	adaptive (residual based)	boundary concentrated	adaptive & bndy conc.
error angle	0.8 e-4	0.3 e-3	1.e-6	1.e-6
error offset	0.7 e-3	0.3 e-2	0.5e-5	0.3e-5

$\sim$ 80.000 nodes:

	uniform	adaptive (residual based)	boundary concentrated	adaptive & bndy conc.
error angle	0.9 e-5	0.2 e-3	0.7 e-7	0.7 e-7
error offset	0.3 e-3	0.5 e-3	0.3 e-6	0.2 e-6

Conventional refinement strategies may lead to bad results!

Errors with Different Refinement Strategies

$\sim$ 20.000 nodes:

	uniform	adaptive (residual based)	boundary concentrated	adaptive & bndy conc.
error angle	0.8 e-4	0.3 e-3	1.e-6	1.e-6
error offset	0.7 e-3	0.3 e-2	0.5e-5	0.3e-5

$\sim$ 80.000 nodes:

	uniform	adaptive (residual based)	boundary concentrated	adaptive & bndy conc.
error angle	0.9 e-5	0.2 e-3	0.7 e-7	0.7 e-7
error offset	0.3 e-3	0.5 e-3	0.3 e-6	0.2 e-6

Conventional refinement strategies may lead to bad results!

Some Ideas on Adaptivity for Inverse Problems

- Haber&Heldmann&Ascher'07: Tikhonov with BV type reg.:
  Refine for u to compute residual term sufficiently precisely;
  Refine for q to compute regularization term sufficiently precisely
- Neubauer'03, '06, '07: moving mesh reg., adaptive grid reg.:
  Tikhonov with BV type regularization:
  Refine where q has jumps or large gradients
- Borcea&Druskin'02: optimal finite difference grids (a priori):
  Refine close to measurements
- Chavent&Bissell'98, Ben Ameur&Chavent&Jaffré'02, BK&Ben Ameur'02:
  refinement and coarsening indicators
- Becker&Vexler'04, Griesbaum&BK&Vexler'07, Bangerth'08, BK&Kirchner&Vexler'11, BK&Kirchner&Veljovic&Vexler'11:
  goal oriented error estimators
- ...

Some Ideas on Adaptivity for Inverse Problems

- Haber&Heldmann&Ascher'07: Tikhonov with BV type reg.:
  Refine for u to compute residual term sufficiently precisely;
  Refine for q to compute regularization term sufficiently precisely
- Neubauer'03, '06, '07: moving mesh reg., adaptive grid reg.:
  Tikhonov with BV type regularization:
  Refine where q has jumps or large gradients
- Borcea&Druskin'02: optimal finite difference grids (a priori):
  Refine close to measurements
- Chavent&Bissell'98, Ben Ameur&Chavent&Jaffré'02, BK&Ben Ameur'02:
  refinement and coarsening indicators
- Becker&Vexler'04, Griesbaum&BK&Vexler'07, Bangerth'08, BK&Kirchner&Vexler'11, BK&Kirchner&Veljovic&Vexler'11:
  goal oriented error estimators
- ...

Parameter identification as nonlinear operator equation

$$F(q) = g$$

$g^\delta \approx g \ldots$ given data; noise level $\delta \geq \|g^\delta - g\|$

$F \ldots$ forward operator: $F(q) = (C \circ S)(q) = C(u)$

where $C \ldots$ observation operator

where $S \ldots$ parameter-to-solution operator, i.e., $u = S(q)$ solves

$$A(q, u)(v) = (f, v) \quad \forall v \in V \quad \ldots \text{ PDE in weak form}$$

Hilbert spaces Q, V, G: $\quad q \in Q \xrightarrow{S} u \in V \xrightarrow{C} g \in G$

Tikhonov Regularization and the Discrepancy Principle

$$\text{Minimize} \quad j_\alpha(q) = \|F(q) - g^\delta\|^2 + \alpha\|q\|^2 \text{ over } q \in Q\,,$$

$$\Leftrightarrow$$

$$\begin{aligned} &\text{Minimize} \quad J_\alpha(q,u) = \|C(u) - g^\delta\|^2 + \alpha\|q\|^2 \text{ over } q \in Q\,,\ u \in V \\ &\text{under the constraint} \quad A(q,u)(v) = (f,v) \quad \forall v \in V \end{aligned}$$

Tikhonov Regularization and the Discrepancy Principle

$$\text{Minimize} \quad j_\alpha(q) = \|F(q) - g^\delta\|^2 + \alpha\|q\|^2 \text{ over } q \in Q\,,$$

Choice of α: discrepancy principle (fixed constant $\tau \geq 1$)

$$\|F(q_{\alpha_*}^\delta) - g^\delta\| = \tau\delta$$

Convergence analysis: [Engl& Hanke& Neubauer 1996] and the references therein

Goal Oriented Error Estimators in PDE Constrained Optimization

[Becker&Kapp&Rannacher'00], [Becker&Rannacher'01], [Becker&Vexler '04, '05]

$$\begin{aligned} &\text{Minimize} \quad J(q,u) \quad \text{over } q \in Q\,,\ u \in V \\ &\text{under the constraints} \quad A(q,u)(v) = f(v) \quad \forall v \in V\,, \end{aligned}$$

Lagrange functional:

$$\mathcal{L}(q,u,z) = J(q,u) + f(z) - A(q,u)(z)\,.$$

First order optimality conditions:

$$\mathcal{L}'(q,u,z)[(p,v,y)] = 0 \quad \forall (p,v,y) \in Q \times V \times V \qquad (1)$$

Discretization $Q_h \subseteq Q$, $V_h \subseteq V \rightsquigarrow$ discretized version of (1).

Estimate the error due to discretization in some quantity of interest I:

$$I(q,u) - I(q_h,u_h) \leq \eta$$

Goal Oriented Error Estimators (II)

Auxiliary functional:

$$\mathcal{M}(q,u,z,p,v,y) = I(q,u) + \mathcal{L}'(q,u,z)[(p,v,y)]$$

Consider additional equations:

$$\mathcal{M}'(x_h)(dx_h) = 0 \quad \forall dx_h \in X_h = (Q_h \times V_h \times V_h)^2$$

Theorem (Becker&Vexler, J. Comp. Phys., 2005):

$$I(q,u) - I(q_h,u_h) = \underbrace{\frac{1}{2}\mathcal{M}'(x_h)(x-\tilde{x}_h)}_{=:\eta} + O(\|x-x_h\|^3) \quad \forall \tilde{x}_h \in X_h\,.$$

Error estimator η is a sum of local contributions due to either q, u, z, ...

$$\eta = \sum_{i=1}^{N_q} \eta_i^q + \sum_{i=1}^{N_u} \eta_i^u + \sum_{i=1}^{N_z} \eta_i^z + \sum_{i=1}^{N_p} \eta_i^p + \sum_{i=1}^{N_v} \eta_i^v + \sum_{i=1}^{N_y} \eta_i^y$$

$\rightsquigarrow$ local refinement separately for $q \in Q_h$, $u \in V_h$, $z \in V_h$, ...

Choice of Quantity of Interest ?

aim: recover infinite dimensional convergence results
for Tikhonov + discrepancy principle
in the adaptively discretized setting

challenge: carrying over infinite dimensional results is

. . . straightforward if we can guarantee smallness of operator norm $\|F_h - F\|$

$\rightsquigarrow$ huge number of quantities of interest!

Choice of Quantity of Interest ?

aim: recover infinite dimensional convergence results
for Tikhonov + discrepancy principle
in the adaptively discretized setting

challenge: carrying over infinite dimensional results is

. . . straightforward if we can guarantee smallness of operator norm
$\|F_h - F\|$
$\rightsquigarrow$ huge number of quantities of interest!

. . . not too hard if we can guarantee smallness of
$\|F_h(q^\dagger) - F(q^\dagger)\|$
$\rightsquigarrow$ large number of quantities of interest!

Choice of Quantity of Interest ?

aim: recover infinite dimensional convergence results
for Tikhonov + discrepancy principle
in the adaptively discretized setting

challenge: carrying over infinite dimensional results is

. . . straightforward if we can guarantee smallness of operator norm
$\|F_h - F\|$
$\rightsquigarrow$ huge number of quantities of interest!

. . . not too hard if we can guarantee smallness of
$\|F_h(q^\dagger) - F(q^\dagger)\|$
$\rightsquigarrow$ large number of quantities of interest!

. . . but we only want to guarantee precision of ≤ 5 quantities of interest

Choice of Quantity of Interest ?

aim: recover infinite dimensional convergence results
for Tikhonov + discrepancy principle
in the adaptively discretized setting

challenge: carrying over infinite dimensional results is

. . . straightforward if we can guarantee smallness of operator norm $\|F_h - F\|$
$\rightsquigarrow$ huge number of quantities of interest!

. . . not too hard if we can guarantee smallness of $\|F_h(q^\dagger) - F(q^\dagger)\|$
$\rightsquigarrow$ large number of quantities of interest!

. . . but we only want to guarantee precision of ≤ 5 quantities of interest

Convergence Analysis $\rightsquigarrow$ Choice of Quantities of Interest

Theorem [Griesbaum&BK& Vexler'07], [BK&Kirchner&Vexler'11]:

$\alpha_* = \alpha_*(\delta, g^\delta)$ and $Q_h \times V_h \times V_h$ such that for

$$I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$\underline{\tau}^2\delta^2 \leq I(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \leq \overline{\tau}\delta^2$$

(i) If additionally

$$|I(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) - I(q^\delta_{\alpha_*}, u^\delta_{\alpha_*})| \leq cI(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*})$$

for some sufficiently small constant $c > 0$ then $q^\delta_{\alpha_*} \to q^\dagger$ as $\delta \to 0$.

Optimal rates under source conditions (logarithic/Hölder).

Convergence Analysis $\rightsquigarrow$ Choice of QoI

Theorem [Griesbaum&BK&Vexler'07], [BK&Kirchner&Vexler'11]:

$\alpha_* = \alpha_*(\delta, g^\delta)$ and $Q_h \times V_h \times V_h$ such that for

$$I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$\underline{\tau}^2\delta^2 \leq I(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \leq \overline{\tau}\delta^2$$

(ii) If additionally for

$$I_2(q,u) := J_\alpha(q,u)$$

$$|I_2(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) - I_2(q^\delta_{\alpha_*}, u^\delta_{\alpha_*})| \leq \sigma\delta^2$$

for some constant $C > 0$ with $\underline{\tau}^2 \geq 1 + \sigma$, then $q^\delta_{h,\alpha_*} \to q^\dagger$ as $\delta \to 0$.
Optimal rates under source conditions (logarithic/Hölder).

see also [Neubauer&Scherzer 1990]

J as quantity of interest $\rightsquigarrow$ [Becker&Kapp&Rannacher'00], [Becker&Rannacher'01]

Convergence Analysis $\rightsquigarrow$ Choice of QoI

Theorem [Griesbaum&BK&Vexler'07], [BK&Kirchner&Vexler'11]:

$\alpha_* = \alpha_*(\delta, g^\delta)$ and $Q_h \times V_h \times V_h$ such that for

$$I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$\underline{\tau}^2\delta^2 \leq I(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) \leq \overline{\tau}\delta^2$$

(ii) If additionally for

$$I_2(q,u) := J_\alpha(q,u)$$

$$|I_2(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}) - I_2(q^\delta_{\alpha_*}, u^\delta_{\alpha_*})| \leq \sigma\delta^2$$

for some constant $C > 0$ with $\underline{\tau}^2 \geq 1 + \sigma$, then $q^\delta_{h,\alpha_*} \to q^\dagger$ as $\delta \to 0$.
Optimal rates under source conditions (logarithic/Hölder).

see also [Neubauer&Scherzer 1990]

J as quantity of interest $\rightsquigarrow$ [Becker&Kapp&Rannacher'00], [Becker&Rannacher'01]

Efficient Computation of α $\rightsquigarrow$ Choice of QoI

Choice of α: discrepancy principle (fixed constant $\tau \geq 1$)

$$\|F(q_{\alpha_*}^{\delta}) - g^{\delta}\| = \tau\delta$$

$\rightsquigarrow$ 1-d nonlinear equation

$\rightsquigarrow$ solve by Newton's method

Efficient Computation of $\alpha \rightsquigarrow$ Choice of QoI

Theorem [Griesbaum&BK&Vexler'07], [BK& Kirchner&Vexler'11]:

Define

$$i(\tfrac{1}{\alpha}) := I(q,u) := \|C(u) - g^\delta\|_G^2 = \|F(q) - g^\delta\|_G^2$$

$$I_2(q,u) := i'(\tfrac{1}{\alpha})$$

β_* sol. to $i(\beta_*) = \tau^2\delta^2$ (discr.princ.) $\quad \beta^{k+1} = \beta^k - \dfrac{i_h^k - \tau^2\delta^2}{i'^k_h}$ (Newton)

for $k \le k_* - 1$ with $\quad k_* = \min\{k \in \mathbb{N} \mid i_h^k - \tau^2\delta^2 \le 0\}$ with

$$|i(\beta^k) - i_h^k| \le \varepsilon^k\,, \qquad |i'(\beta^k) - i'^k_h| \le \varepsilon'^k\,,$$

$\varepsilon^k, \varepsilon'^k$ sufficiently small.
Then β^k satisfies quadratic convergence estimate and

$$(\tau^2 - \tilde{\tau}^2)\delta^2 \le i(\beta^{k_*}) \le (\tau^2 + \tilde{\tau}^2)\delta^2$$

Remarks

- computation of error estimators for $i(\beta)$:
  just one more SQP type step;
- evaluation of $i'(\beta)$:
  can be extracted from quantities computed for error estimators for $i(\beta)$
- error estimators for $i'(\beta)$:
  stationary point of another auxiliary functional

Newton type Regularization and the Discrepancy Principle

Newton's method for $F(q) = g$:

$$q_{k+1} = q_k - F'(q_k)^{-1}(F(q_k) - g)$$

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^{\delta} = q_k^{\delta} - (F'(q_k^{\delta})^* F'(q_k^{\delta}) + \alpha_k I)^{-1}(F'(q_k^{\delta})^*(F(q_k^{\delta}) - g^{\delta}) + \alpha_k(q_k^{\delta} - q_0))$$

Newton type Regularization and the Discrepancy Principle

Newton's method for $F(q) = g$:

$$q_{k+1} = q_k - F'(q_k)^{-1}(F(q_k) - g)$$

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^\delta = q_k^\delta - (F'(q_k^\delta)^* F'(q_k^\delta) + \alpha_k I)^{-1}(F'(q_k^\delta)^*(F(q_k^\delta) - g^\delta) + \alpha_k(q_k^\delta - q_0))$$

or equivalently

$$q_{k+1}^\delta = \arg\min_q \|F'(q_k^\delta)(q - q_k^\delta) + F(q_k^\delta) - g^\delta\|^2 + \alpha_k\|q - q_0\|^2$$

Newton type Regularization and the Discrepancy Principle

Newton's method for $F(q) = g$:

$$q_{k+1} = q_k - F'(q_k)^{-1}(F(q_k) - g)$$

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^\delta = q_k^\delta - (F'(q_k^\delta)^* F'(q_k^\delta) + \alpha_k I)^{-1}(F'(q_k^\delta)^*(F(q_k^\delta) - g^\delta) + \alpha_k(q_k^\delta - q_0))$$

or equivalently

$$q_{k+1}^\delta = \arg\min_q \|F'(q_k^\delta)(q - q_k^\delta) + F(q_k^\delta) - g^\delta\|^2 + \alpha_k\|q - q_0\|^2$$

Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^{\delta} = \arg\min_{q} \|F'(q_k^{\delta})(q - q_k^{\delta}) + F(q_k^{\delta}) - g^{\delta}\|^2 + \alpha_k \|q - q_0\|^2$$

or equivalently

$$\text{Minimize} \quad J_k(q,u,w) = \|C(w+u) - g^{\delta}\|_G^2 + \alpha_k \|q - q_0\|^2 \text{ over } \begin{matrix} q \in Q \\ u \in V \\ w \in V \end{matrix}$$

under the constraints

$$A'_u(q_k^{\delta}, u)[w](v) + A'_q(q_k^{\delta}, u)[q - q_k^{\delta}](v) = 0 \quad \forall v \in V\,,$$

$$A(q_k^{\delta}, u)(v) = f(v) \quad \forall v \in V\,,$$

Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^{\delta} = \arg\min_{q} \|F'(q_k^{\delta})(q - q_k^{\delta}) + F(q_k^{\delta}) - g^{\delta}\|^2 + \alpha_k \|q - q_0\|^2$$

or equivalently

$$\text{Minimize} \quad J_k(q,u,w) = \|C(w+u) - g^{\delta}\|_G^2 + \alpha_k \|q - q_0\|^2 \text{ over } \begin{matrix} q \in Q \\ u \in V \\ w \in V \end{matrix}$$

under the constraints

$$A'_u(q_k^{\delta}, u)[w](v) + A'_q(q_k^{\delta}, u)[q - q_k^{\delta}](v) = 0 \quad \forall v \in V\,,$$

$$A(q_k^{\delta}, u)(v) = f(v) \quad \forall v \in V\,,$$

Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^{\delta} = q_k^{\delta} - (F'(q_k^{\delta})^* F'(q_k^{\delta}) + \alpha_k I)^{-1}(F'(q_k^{\delta})^*(F(q_k^{\delta}) - g^{\delta}) + \alpha_k(q_k^{\delta} - q_0))$$

a posteriori selection of α_k (inexact Newton)

$$\underline{\tilde{\theta}}\|F(q_k) - g^{\delta}\| \leq \|F'(q_k)(q_{k+1} - q_k) + F(q_k) - g^{\delta}\| \leq \bar{\tilde{\theta}}\|F(q_k) - g^{\delta}\|$$

Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^{\delta} = q_k^{\delta} - (F'(q_k^{\delta})^* F'(q_k^{\delta}) + \alpha_k I)^{-1}(F'(q_k^{\delta})^*(F(q_k^{\delta}) - g^{\delta}) + \alpha_k(q_k^{\delta} - q_0))$$

a posteriori selection of α_k (inexact Newton)

$$\underline{\tilde{\theta}}\|F(q_k) - g^{\delta}\| \leq \|F'(q_k)(q_{k+1} - q_k) + F(q_k) - g^{\delta}\| \leq \tilde{\bar{\theta}}\|F(q_k) - g^{\delta}\|$$

a posteriori selection of k_* (discrepancy principle)

$$k^* = \min\{k \in \mathbb{N} \ : \ \|F(q_k) - g^{\delta}\| \leq \tau\delta\}$$

Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$q_{k+1}^{\delta} = q_k^{\delta} - (F'(q_k^{\delta})^* F'(q_k^{\delta}) + \alpha_k I)^{-1}(F'(q_k^{\delta})^*(F(q_k^{\delta}) - g^{\delta}) + \alpha_k(q_k^{\delta} - q_0))$$

a posteriori selection of α_k (inexact Newton)

$$\underline{\tilde{\theta}}\|F(q_k) - g^{\delta}\| \leq \|F'(q_k)(q_{k+1} - q_k) + F(q_k) - g^{\delta}\| \leq \tilde{\bar{\theta}}\|F(q_k) - g^{\delta}\|$$

a posteriori selection of k_* (discrepancy principle)

$$k^* = \min\{k \in \mathbb{N} \ : \ \|F(q_k) - g^{\delta}\| \leq \tau\delta\}$$

Convergence Analysis $\rightsquigarrow$ Choice of QoI

$$|I_{i,h}^{k+1} - I_i^{k+1}| \leq \eta_i^{k+1}, \quad i \in \{1,2,3,4\} \qquad (*)$$

by goal oriented adaptivity, where

$$\begin{aligned}
I_{1,h}^{k+1} &= \|F_h'(q_k^H)(q_{k+1}^h - q_k^H) + F_h(q_k^H) - g^\delta\|^2 + \alpha_k \|q_{k+1}^h - q_0\|^2 \\
I_{2,h}^{k+1} &= \|F_h'(q_k^H)(q_{k+1}^h - q_k^H) + F_h(q_k^H) - g^\delta\|^2 \\
I_{3,h}^{k+1} &= \|F_h(q_k^H) - g^\delta\|^2 \\
I_{4,h}^{k+1} &= \|F_h(q_{k+1}^h) - g^\delta\|^2,
\end{aligned}$$

Theorem [BK&Kirchner&Veljovic&Vexler'11]:
Let $(*)$ hold with η_i^{k+1} sufficiently small. Then $q_{h,k_*}^\delta \to q^\dagger$ as $\delta \to 0$.
Optimal rates under source conditions (logarithic/Hölder).

However...
need to solve

$$
\begin{aligned}
2\alpha_k(q_{k+1}-q_0)+A'_q(q_k^H,u_{old})^*v^1 &= 0\\
2C^*(C(w+u_{old})-g^\delta)+A''_{uu}(q_k^H,u_{old})[w,\cdot](v^1)&\\
+A''_{qu}(q_k^H,u_{old})[q-q_k^H,\cdot](v^1)+A'_u(q_k^H,u_{old})^*v^2 &= 0\\
2C^*(C(w+u_{old})-g^\delta)+A'_u(q_k^H,u_{old})^*v^1 &= 0\\
A'_u(q_k^H,u_{old})[w]+A'_q(q_k^H,u_{old})[q-q_k^H] &= 0\\
A(q_k^H,u_{old})-f &= 0\\
A(q,u)-f &= 0
\end{aligned}
$$

- second order derivatives of PDEs appear
- nonlinear PDEs have to be solved

An all-at once Formulation (I)

$$\begin{array}{lrcl} \text{measurements:} & C(u) & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array} \quad \text{i.e., } \mathbf{F}(u,q) = \mathbf{g}$$

$\rightsquigarrow$ Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \begin{pmatrix} q_k^\delta \\ u_k^\delta \end{pmatrix} - \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* \mathbf{F}'(q_k^\delta, u_k^\delta) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1}$$
$$\times \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* (\mathbf{F}(q_k^\delta, u_k^\delta) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q_k^\delta - q_0 \\ 0 \end{pmatrix} \right)$$

An all-at once Formulation (I)

$$\begin{array}{lrcl} \text{measurements:} & C(u) & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array} \quad \text{i.e., } \mathbf{F}(u,q) = \mathbf{g}$$

$\rightsquigarrow$ Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \begin{pmatrix} q_k^\delta \\ u_k^\delta \end{pmatrix} - \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* \mathbf{F}'(q_k^\delta, u_k^\delta) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1}$$
$$\times \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* (\mathbf{F}(q_k^\delta, u_k^\delta) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q_k^\delta - q_0 \\ 0 \end{pmatrix} \right)$$

or equivalently: unconstrained quadratic minimization

$$(q_{k+1}^\delta, u_{k+1}^\delta)$$
$$= \arg\min_{q,u} \quad \|A_q'(q_k,u_k)(q-q_k) + A_u'(q_k,u_k)(u-u_k) + A(q_k,u_k) - f\|_{V^*}^2$$
$$+ \|Cu - g^\delta\|_G^2 + \alpha_k \|q - q_0\|_Q^2 \,.$$

An all-at once Formulation (I)

$$\begin{array}{rrcl} \text{measurements:} & C(u) & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array} \quad \text{i.e., } \mathbf{F}(u,q) = \mathbf{g}$$

$\rightsquigarrow$ Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \begin{pmatrix} q_k^\delta \\ u_k^\delta \end{pmatrix} - \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* \mathbf{F}'(q_k^\delta, u_k^\delta) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1}$$
$$\times \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* (\mathbf{F}(q_k^\delta, u_k^\delta) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q_k^\delta - q_0 \\ 0 \end{pmatrix} \right)$$

or equivalently: unconstrained quadratic minimization

$$(q_{k+1}^\delta, u_{k+1}^\delta)$$
$$= \arg\min_{q,u} \quad \|A_q'(q_k,u_k)(q-q_k) + A_u'(q_k,u_k)(u-u_k) + A(q_k,u_k) - f\|_{V^*}^2$$
$$+ \|Cu - g^\delta\|_G^2 + \alpha_k \|q - q_0\|_Q^2 .$$

see also [Burger&Mühlhuber'02]: SQP type approach

An all-at once Formulation (I)

$$\begin{array}{rrcl} \text{measurements:} & C(u) & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array} \quad \text{i.e., } \mathbf{F}(u,q) = \mathbf{g}$$

$\rightsquigarrow$ Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \begin{pmatrix} q_k^\delta \\ u_k^\delta \end{pmatrix} - \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* \mathbf{F}'(q_k^\delta, u_k^\delta) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1}$$
$$\times \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* (\mathbf{F}(q_k^\delta, u_k^\delta) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q_k^\delta - q_0 \\ 0 \end{pmatrix} \right)$$

or equivalently: unconstrained quadratic minimization

$$(q_{k+1}^\delta, u_{k+1}^\delta)$$
$$= \arg\min_{q,u} \quad \|A_q'(q_k,u_k)(q-q_k) + A_u'(q_k,u_k)(u-u_k) + A(q_k,u_k) - f\|_{V^*}^2$$
$$+ \|Cu - g^\delta\|_G^2 + \alpha_k \|q - q_0\|_Q^2 \,.$$

see also [Burger&Mühlhuber'02]: SQP type approach

An all-at once Formulation (II)

$$
\begin{aligned}
&(q^{\delta}_{k+1}, u^{\delta}_{k+1}) \\
&\quad = \arg\min_{q,u} \quad \|L(q-q_k) + K(u-u_k) + A(q_k,u_k) - f\|^2_{V^*} \\
&\qquad\qquad\qquad + \|Cu - g^{\delta}\|^2_G + \alpha_k \|q - q_0\|^2_Q \,.
\end{aligned}
$$

with $K = A'_u(q_k, u_k)$, $L = A'_q(q_k, u_k)$.

K regular $\Rightarrow$ Hessian $\begin{pmatrix} L^*L + \alpha_k I & L^*K \\ K^*L & C^*C + K^*K \end{pmatrix}$ positive definite.

An all-at once Formulation (II)

$$
\begin{aligned}
&(q_{k+1}^{\delta}, u_{k+1}^{\delta}) \\
&\quad = \arg\min_{q,u} \quad \|L(q-q_k) + K(u-u_k) + A(q_k,u_k) - f\|_{V^*}^2 \\
&\qquad\qquad\qquad + \|Cu - g^{\delta}\|_G^2 + \alpha_k \|q - q_0\|_Q^2 \,.
\end{aligned}
$$

with $K = A'_u(q_k, u_k)$, $L = A'_q(q_k, u_k)$.

K regular $\Rightarrow$ Hessian $\begin{pmatrix} L^*L + \alpha_k I & L^*K \\ K^*L & C^*C + K^*K \end{pmatrix}$ positive definite.

First order optimality system:

$$
\begin{pmatrix} \alpha_k I & 0 & L^* \\ 0 & C^*C & K^* \\ L & K & I \end{pmatrix}
\begin{pmatrix} q_{k+1} \\ u_{k+1} \\ z_{k+1} \end{pmatrix}
=
\begin{pmatrix} \alpha_k q_k \\ -C^*(Cu_k - g^{\delta}) \\ Lq_k + Ku_k - A(q_k,u_k) + f \end{pmatrix}
$$

An all-at once Formulation (II)

$$
\begin{aligned}
&(q^\delta_{k+1}, u^\delta_{k+1}) \\
&\quad = \arg\min_{q,u} \quad \|L(q-q_k) + K(u-u_k) + A(q_k,u_k) - f\|^2_{V^*} \\
&\qquad\qquad\qquad + \|Cu - g^\delta\|^2_G + \alpha_k \|q - q_0\|^2_Q \,.
\end{aligned}
$$

with $K = A'_u(q_k, u_k)$, $L = A'_q(q_k, u_k)$.

$$
K \text{ regular} \quad \Rightarrow \quad \text{Hessian} \begin{pmatrix} L^*L + \alpha_k I & L^*K \\ K^*L & C^*C + K^*K \end{pmatrix} \text{ positive definite.}
$$

First order optimality system:

$$
\begin{pmatrix} \alpha_k I & 0 & L^* \\ 0 & C^*C & K^* \\ L & K & I \end{pmatrix} \begin{pmatrix} q_{k+1} \\ u_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} \alpha_k q_k \\ -C^*(Cu_k - g^\delta) \\ Lq_k + Ku_k - A(q_k,u_k) + f \end{pmatrix}
$$

All-At-Once-Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \begin{pmatrix} q_k^\delta \\ u_k^\delta \end{pmatrix} - \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* \mathbf{F}'(q_k^\delta, u_k^\delta) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1} \times \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* (\mathbf{F}(q_k^\delta, u_k^\delta) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q_k^\delta - q_0 \\ 0 \end{pmatrix} \right)$$

a posteriori selection of α_k (inexact Newton)

$$\underline{\tilde{\theta}} \|\mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\| \leq \|\mathbf{F}'(q_k,u_k)\begin{pmatrix} q_{k+1}-q_k \\ u_{k+1}-u_k \end{pmatrix} + \mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\| \leq \bar{\tilde{\theta}} \|\mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\|$$

All-At-Once-Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \begin{pmatrix} q_k^\delta \\ u_k^\delta \end{pmatrix} - \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* \mathbf{F}'(q_k^\delta, u_k^\delta) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1} \times \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* (\mathbf{F}(q_k^\delta, u_k^\delta) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q_k^\delta - q_0 \\ 0 \end{pmatrix} \right)$$

a posteriori selection of α_k (inexact Newton)

$$\underline{\tilde{\theta}} \|\mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\| \leq \|\mathbf{F}'(q_k,u_k)\begin{pmatrix} q_{k+1}-q_k \\ u_{k+1}-u_k \end{pmatrix} + \mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\| \leq \bar{\tilde{\theta}} \|\mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\|$$

a posteriori selection of k_* (discrepancy principle)

$$k^* = \min\{k \in \mathbb{N} \ : \ \|\mathbf{F}(q_k, u_k) - \mathbf{g}^\delta\| \leq \tau\delta\}$$

All-At-Once-Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \begin{pmatrix} q_k^\delta \\ u_k^\delta \end{pmatrix} - \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* \mathbf{F}'(q_k^\delta, u_k^\delta) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1} \times \left(\mathbf{F}'(q_k^\delta, u_k^\delta)^* (\mathbf{F}(q_k^\delta, u_k^\delta) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q_k^\delta - q_0 \\ 0 \end{pmatrix} \right)$$

a posteriori selection of α_k (inexact Newton)

$$\underline{\tilde{\theta}} \|\mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\| \leq \|\mathbf{F}'(q_k,u_k)\begin{pmatrix} q_{k+1}-q_k \\ u_{k+1}-u_k \end{pmatrix} + \mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\| \leq \bar{\tilde{\theta}} \|\mathbf{F}(q_k,u_k) - \mathbf{g}^\delta\|$$

a posteriori selection of k_* (discrepancy principle)

$$k^* = \min\{k \in \mathbb{N} \ : \ \|\mathbf{F}(q_k, u_k) - \mathbf{g}^\delta\| \leq \tau\delta\}$$

Convergence Analysis $\rightsquigarrow$ Choice of QoI

$$|I_{i,h}^{k+1} - I_i^{k+1}| \leq \eta_i^{k+1}, \quad i \in \{1,2,3,4\} \qquad (*)$$

by goal oriented adaptivity, where

$$\begin{aligned}
I_{1,h}^{k+1} &= \|\mathbf{F}_h'(q_k^H,u_k^H)\begin{pmatrix} q_{k+1}^h - q_k^H \\ u_{k+1}^h - u_k^H \end{pmatrix} + \mathbf{F}_h(q_k^H,u_k^H) - \mathbf{g}^\delta\| + \alpha_k\|q_{k+1}^h - q_0\|^2 \\
I_{2,h}^{k+1} &= \|\mathbf{F}_h'(q_k^H,u_k^H)\begin{pmatrix} q_{k+1}^h - q_k^H \\ u_{k+1}^h - u_k^H \end{pmatrix} + \mathbf{F}_h(q_k^H,u_k^H) - \mathbf{g}^\delta\| \\
I_{3,h}^{k+1} &= \|\mathbf{F}_h(q_k^H, u_k^H) - \mathbf{g}^\delta\|^2 \\
I_{4,h}^{k+1} &= \|\mathbf{F}_h(q_{k+1}^h, u_{k+1}^h) - \mathbf{g}^\delta\|^2,
\end{aligned}$$

Theorem [BK&Kirchner&Veljovic&Vexler'11]:
Let $(*)$ hold with η_i^{k+1} sufficiently small. Then $q_{h,k_*}^\delta \to q^\dagger$ as $\delta \to 0$.
Optimal rates under source conditions (logarithic/Hölder).

Convergence Analysis $\rightsquigarrow$ Choice of QoI

$$|I_{i,h}^{k+1} - I_i^{k+1}| \leq \eta_i^{k+1}, \quad i \in \{1,2,3,4\} \qquad (*)$$

by goal oriented adaptivity, where

$$\begin{aligned}
I_{1,h}^{k+1} &= \|\mathbf{F}_h'(q_k^H,u_k^H)\begin{pmatrix} q_{k+1}^h - q_k^H \\ u_{k+1}^h - u_k^H \end{pmatrix} + \mathbf{F}_h(q_k^H,u_k^H) - \mathbf{g}^\delta\| + \alpha_k\|q_{k+1}^h - q_0\|^2 \\
I_{2,h}^{k+1} &= \|\mathbf{F}_h'(q_k^H,u_k^H)\begin{pmatrix} q_{k+1}^h - q_k^H \\ u_{k+1}^h - u_k^H \end{pmatrix} + \mathbf{F}_h(q_k^H,u_k^H) - \mathbf{g}^\delta\| \\
I_{3,h}^{k+1} &= \|\mathbf{F}_h(q_k^H, u_k^H) - \mathbf{g}^\delta\|^2 \\
I_{4,h}^{k+1} &= \|\mathbf{F}_h(q_{k+1}^h, u_{k+1}^h) - \mathbf{g}^\delta\|^2,
\end{aligned}$$

- only first order derivatives of PDEs appear
- only residuals of nonlinear PDEs have to be evaluated

Numerical Tests (Tikhonov & Discrepancy Principle)

nonlinear inverse source problem:

$$-\Delta u + 1000u^3 = q \text{ in } \Omega = (0,1)^2 \quad + \text{ homogeneous Dirichlet BC}$$

Identify q from distributed measurements of u at 10×10 points in Ω

$$(a) \quad q^\dagger(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-\frac{5}{11})^2+(y-\frac{5}{11})^2}{2\sigma^2}\right), \ \sigma = 0.01$$

$$(b) \quad q^\dagger(x,y) = q_1(x,y) + q_2(x,y)$$

$$q_i = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\left(\frac{s_i x - \frac{1}{2}}{\sigma}\right)^2 + \left(\frac{s_i y - \frac{1}{2}}{\sigma}\right)^2\right)\right), \quad \begin{array}{l}\sigma = 0.1 \\ s_1 = 2, \\ s_2 = 0.8\end{array}$$

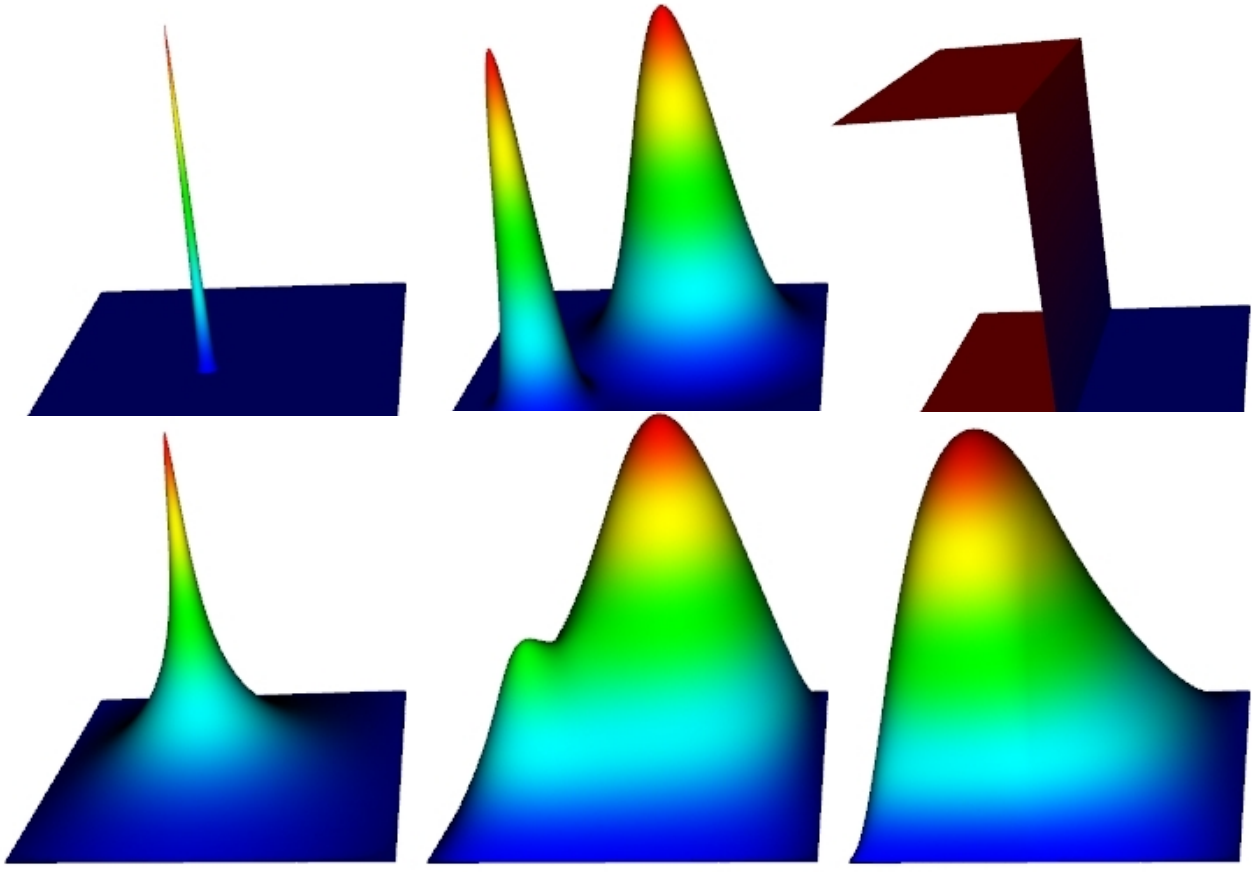
$$(c) \qquad q^\dagger(x,y) = \begin{cases} 1 & x \le \frac{1}{2} \\ 0 & x > \frac{1}{2} \end{cases}$$

Computations with *Gascoigne* and *RoDoBo*.

Numerical Tests

exact par. q:
(a), (b), (c)

exact state u:
(a), (b), (c)

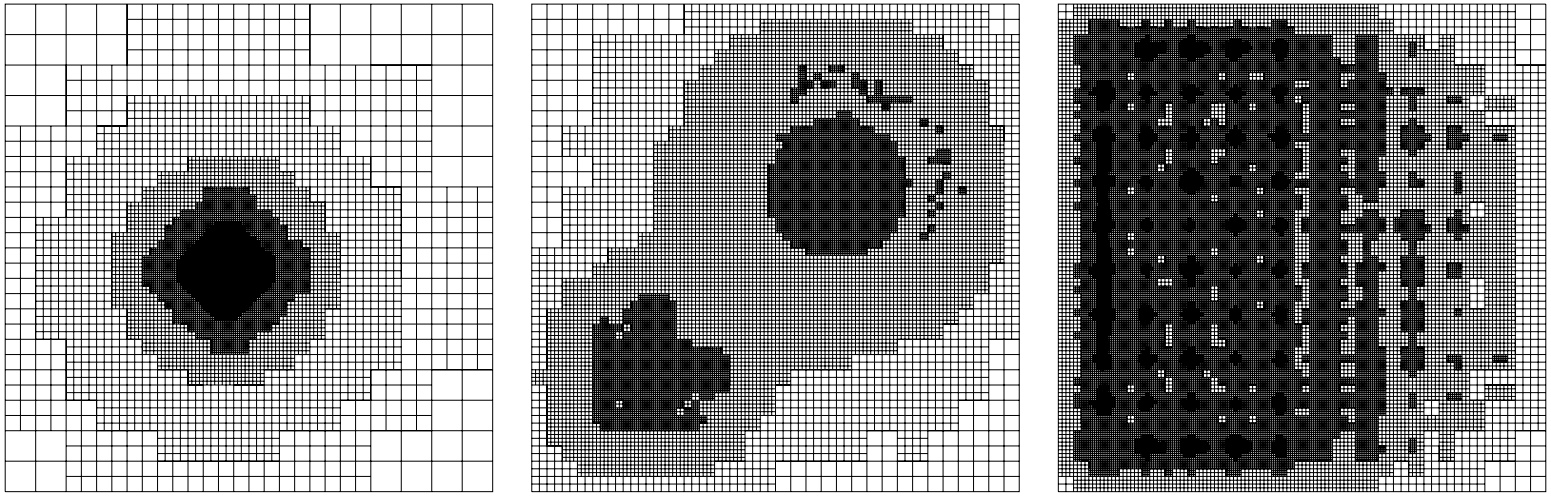


Numerical Results (I)

Computations with 1% random noise:
number of nodes on finest grid:

	(a)	(b)	(c)
uniform	263169	66049	66049
adaptive	14157	18035	56409
reduction of CPU time	92%	53%	10%

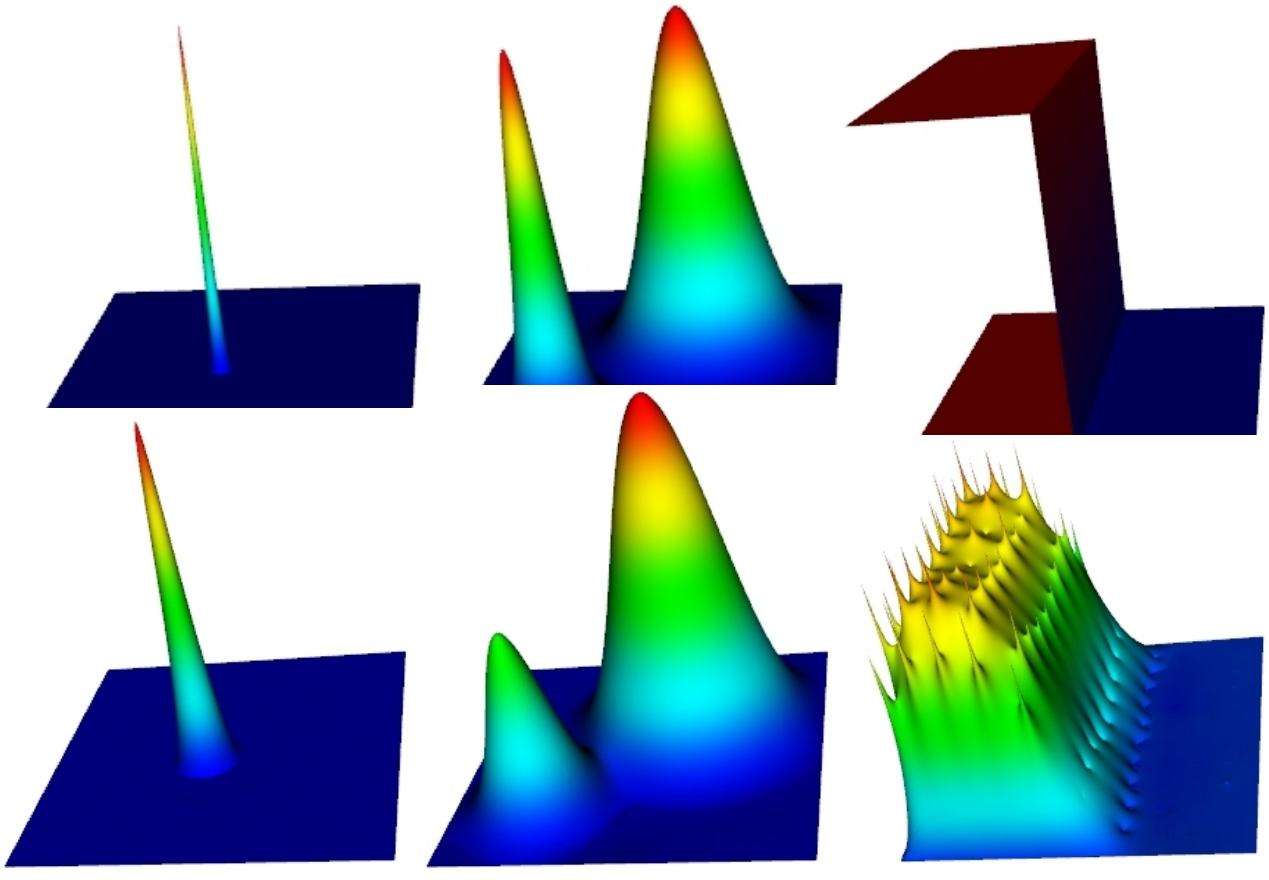
adaptive grids:



Numerical Results (II)

exact par. q:
(a), (b), (c)

computed par. q
(a), (b), (c)



Numerical Results (III)

Convergence as $\delta \to 0$ for example (a),
linear inverse source problem

with $\sigma = 0.05$

δ	$\frac{\|q^\delta_{\alpha_*} - q^\dagger\|}{\|q^\dagger\|}$	$1/\alpha_*$
8%	0.761	156.390
4%	0.592	660.930
2%	0.414	2426.109
1%	0.288	7047.472
0.5%	0.229	17042.825

with $\sigma = 0.01$

δ	$\frac{\|q^\delta_{\alpha_*} - q^\dagger\|}{\|q^\dagger\|}$	$1/\alpha_*$
8%	0.869	2396.281
4%	0.776	9044.374
2%	0.744	24364.894
1%	0.734	55017.364
0.5%	0.731	117560.866

Summary and Outlook

- goal oriented error estimators allow to control the error in some quantity of interest
- Tikhonov& discrepancy principle:
  - sufficiently small error in residual norm and Tikh. func.
    $\Rightarrow$ convergence and rates
  - sufficiently small error in residual norm $i(\frac{1}{\alpha})$ and in $i'(\frac{1}{\alpha})$
    $\Rightarrow$ fast convergence of Newton's method for finding α_* (discr.princ.)
    $\rightsquigarrow$ coarse grids at the beginning of Newton's method
    $\rightarrow$ save computational effort
- regularized Newton iterations& discrepancy principle:
  - sufficiently small error in nonlinear residual norm, linearized residual norm, and linearized Tikhonov functional
    $\Rightarrow$ convergence and rates
  - all-at-once-approach

$\rightarrow$ other regularization methods

$\rightarrow$ other parameter choice strategies: e.g., balancing principle

Thank you for your attention!