Adaptive discretization of parameter identification problems in PDEs for variational and iterative regularization

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

OCIP 2013

Overview

- motivation: parameter identification in PDE
- ideas on adaptivity for inverse problems
- principles of goal oriented error estimators
- variational regularization
- iterative regularization
- conclusions and outlook

Motivation: Parameter Identification in PDE

Some Model problems:

- *a*-example: identify the diffusivity $a = a(x)$ in

  $$-\nabla (a(x) \nabla u) = f \quad \text{in } \Omega$$

  from measurements of the state u
  $\rightsquigarrow$ nonlinear inverse problem

- *c*-example: identify the potential $c = c(x)$ in

  $$-\Delta u + c(x) u = f \quad \text{in } \Omega$$

  from measurements of the state u.
  $\rightsquigarrow$ nonlinear inverse problem

Motivation: Parameter Identification in PDE

Some Model problems:

- inverse source problem: identify the source $f = f(x)$ in

  $$-\Delta u = f(x) \quad \text{in } \Omega$$

  from measurements of the state u.
  $\rightsquigarrow$ linear inverse problem
- nonlinearity identification: identify the heat conductivity $q = q(u)$ in

  $$-\nabla \left(q(u)\nabla u\right) = f \quad \text{in } \Omega$$

  from measurements of the state u.
  $\rightsquigarrow$ nonlinear inverse problem

Motivation: Parameter Identification in PDE

Some Model problems:

$$-\nabla(a\nabla u) = f \qquad -\Delta u = f$$

$$-\Delta u + cu = f \qquad -\nabla(q(u)\nabla u) = f$$

- The majority of parameter identification problems in PDEs leads to nonlinear inverse problems

Motivation: Parameter Identification in PDE

Some Model problems:

$$-\nabla\left(a\nabla u\right) = f \qquad\qquad -\Delta u = f$$

$$-\Delta u + cu = f \qquad\qquad -\nabla\left(q(u)\nabla u\right) = f$$

- The majority of parameter identification problems in PDEs leads to nonlinear inverse problems
- From now on all parameters will be called "coefficients" and denoted by q.

Motivation: Parameter Identification in PDE

Some Model problems:

$$-\nabla(a\nabla u) = f \qquad -\Delta u = f$$

$$-\Delta u + cu = f \qquad -\nabla(q(u)\nabla u) = f$$

- The majority of parameter identification problems in PDEs leads to nonlinear inverse problems
- From now on all parameters will be called "coefficients" and denoted by q.
- The forward operator F maps the coefficient q to the measurement u (or Cu)

Motivation: Parameter Identification in PDE

Some Model problems:

$$-\nabla(a\nabla u) = f \qquad -\Delta u = f$$

$$-\Delta u + cu = f \qquad -\nabla(q(u)\nabla u) = f$$

- The majority of parameter identification problems in PDEs leads to nonlinear inverse problems
- From now on all parameters will be called "coefficients" and denoted by q.
- The forward operator F maps the coefficient q to the measurement u (or Cu)

Motivation: Coefficient Identification in PDE

Abstract formulation:

$$A(q,u)(v) = (f,v) \quad \forall v \in V \qquad \dots \text{ PDE in weak form}$$
$$Cu = g \qquad \dots \text{ measurements}$$

or equivalently

$$F(q) = g$$

$F \dots$ forward operator: $F(q) = (C \circ S)(q) = Cu$
where $u = S(q)$ solves PDE; $S \dots$ coefficient-to-state-map

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

Motivation: Coefficient Identification in PDE

Abstract formulation:

$$\begin{aligned} A(q,u)(v) &= (f,v) \quad \forall v \in V && \ldots \text{ PDE in weak form} \\ Cu &= g && \ldots \text{ measurements} \end{aligned}$$

or equivalently

$$F(q) = g$$

$F \ldots$ forward operator: $F(q) = (C \circ S)(q) = Cu$
where $u = S(q)$ solves PDE; $S \ldots$ coefficient-to-state-map

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

inverse problem: identify q from measurements g^δ of g

Motivation: Coefficient Identification in PDE

Abstract formulation:

$$\begin{aligned} A(q,u)(v) = &\ (f,v) \quad \forall v \in V && \ldots \text{ PDE in weak form} \\ Cu = &\ g && \ldots \text{ measurements} \end{aligned}$$

or equivalently

$$F(q) = g$$

$F \ldots$ forward operator: $F(q) = (C \circ S)(q) = Cu$
where $u = S(q)$ solves PDE; $S \ldots$ coefficient-to-state-map

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

inverse problem: identify q from measurements g^δ of g

Motivation: Relation to Optimization

operator equation

$$\begin{aligned} A(q,u)(v) &= (f,v) \quad \forall v \in V \\ Cu &= g \end{aligned}$$

or equivalently

$$F(q) = g$$

minimization problem

$$\min_{q,u} \ \|Cu - g^\delta\|^2$$
$$\text{s.t.} \ \ A(q,u)(v) = (f,v) \quad \forall v \in V$$

or equivalently

$$\min_q \ \|F(q) - g^\delta\|^2$$

. . . plus regularization:

Motivation: Relation to Optimization

operator equation

$$\begin{aligned} A(q,u)(v) &= (f,v) \quad \forall v \in V \\ Cu &= g \end{aligned}$$

or equivalently

$$F(q) = g$$

minimization problem

$$\min_{q,u} \ \|Cu - g^\delta\|^2$$
$$\text{s.t.} \quad A(q,u)(v) = (f,v) \quad \forall v \in V$$

or equivalently

$$\min_q \ \|F(q) - g^\delta\|^2$$

...plus regularization:

Motivation: Relation to Optimization

e.g. **Tikhonov regularization:**[1]

Minimize $J_\alpha(q,u) = \|Cu - g^\delta\|^2 + \alpha\|q\|^2$ over $q \in Q$, $u \in V$
under the constraint $A(q,u)(v) = (f,v) \quad \forall v \in V$

or equivalently

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

$\rightsquigarrow$ PDE constrained optimization

additional issue: choice of α!

[1] There exist many other regularization methods!

Motivation: Relation to Optimization

e.g. **Tikhonov regularization:**[1]

Minimize $J_\alpha(q,u) = \|Cu - g^\delta\|^2 + \alpha\|q\|^2$ over $q \in Q\,,\ u \in V$
under the constraint $A(q,u)(v) = (f,v) \quad \forall v \in V$

or equivalently

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

$\rightsquigarrow$ PDE constrained optimization

additional issue: choice of α!

[1]There exist many other regularization methods!

e.g., **Discrepancy Principle**[2]

assume noise level $\delta \geq \|g - g^\delta\|$ to be known;
fix constant $\tau \geq 1$ independent of δ;
determine $\alpha = \alpha_*$ such that

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

(relaxed version $\underline{\tau}^2\delta^2 \leq \|F(q^\delta_{\alpha_*}) - g^\delta\|^2 \leq \overline{\tau}^2\delta^2$)
where q^δ_α is the Tikhonov minimizer

$\rightsquigarrow$ nonlinear 1-d equation $\phi(\alpha) = 0$ for α;

each evaluation of ϕ requires minimization of Tikhonov functional!

[2] There exist many other regularization parameter choice strategies!

Motivation: Coefficient Identification in PDE

computational issues:

- instability:
  amplification of numerical errors
- computational effort:
  several reg. inversions to determine regularization parameter

Motivation: Coefficient Identification in PDE

adaptive discretization:

- examples $-\nabla(q\nabla u) = f$; $-\Delta u + qu = f$; $-\Delta u = q$: refine grid for u and q
  - at jumps or large gradients
  - close to measurements
  - at locations with large error contribution

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

Motivation: Coefficient Identification in PDE

adaptive discretization:

- examples $-\nabla(q\nabla u) = f$; $-\Delta u + qu = f$; $-\Delta u = q$: refine grid for u and q
  - at jumps or large gradients
  - close to measurements
  - at locations with large error contribution

  $\rightarrow$ location of large gradients / large errors a priori unknown
- example $-\nabla(q(u)\nabla u) = f$:
  - no direct relation between refinement regions for u and q

  $\rightarrow$ general strategy for mesh generation separately for q and u

Motivation: Coefficient Identification in PDE

adaptive discretization:

- examples $-\nabla(q\nabla u) = f$; $-\Delta u + qu = f$; $-\Delta u = q$: refine grid for u and q
  - at jumps or large gradients
  - close to measurements
  - at locations with large error contribution

  $\rightarrow$ location of large gradients / large errors a priori unknown
- example $-\nabla(q(u)\nabla u) = f$:
  - no direct relation between refinement regions for u and q

  $\rightarrow$ general strategy for mesh generation separately for q and u

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.:
  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:
  Refine to reduce data misfit and coarsen to reduce number of dofs
  (refinement and coarsening indicators)
- Becker&Vexler'04, Griesbaum&BK&Vexler'07, Bangerth&Joshi'08,
  Beilina et. al.'05,'06,'09,'10,'11,'12, BK&Kirchner&Vexler'11,
  BK&Kirchner&Veljovic&Vexler'13:
  Refine to obtain sufficient precision in some quantity of interest
  (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.:
  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:
  Refine to reduce data misfit and coarsen to reduce number of dofs
  (refinement and coarsening indicators)
- Becker&Vexler'04, Griesbaum&BK&Vexler'07, Bangerth&Joshi'08, Beilina et. al.'05,'06,'09,'10,'11,'12, BK&Kirchner&Vexler'11, BK&Kirchner&Veljovic&Vexler'13:
  Refine to obtain sufficient precision in some quantity of interest
  (goal oriented error estimators)
- ...

Variational Regularization

Coefficient Identification in PDE as Operator Equation

$$\begin{aligned} A(q,u)(v) &= (f,v) \quad \forall v \in V && \ldots \text{ PDE in weak form} \\ Cu &= g && \ldots \text{ measurements} \end{aligned}$$

or equivalently

$$F(q) = g$$

$F \ldots$ forward operator: $F(q) = (C \circ S)(q) = Cu$
where $u = S(q)$ solves PDE; $S \ldots$ coefficient-to-state-map

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

inverse problem: identify q from measurements g^δ of g

Tikhonov Regularization

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

or equivalently

Minimize $J_\alpha(q,u) = \|Cu - g^\delta\|^2 + \alpha\|q\|^2$ over $q \in Q$, $u \in V$
under the constraint $A(q,u)(v) = (f,v) \quad \forall v \in V$

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 $\tau \geq 1$)

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

$\leadsto$ nonlinear 1-d equation $\phi(\alpha) = 0$ for α;
evaluation of ϕ requires minimization of Tikhonov functional

Convergence analysis as $\delta \to 0$:
[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]

Minimize $J(q, u)$ over $q \in Q$, $u \in V$
under the constraint $A(q, u)(v) = f(v) \quad \forall v \in V$

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:

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

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

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

After computing a stationary point (q_h, u_h, z_h), computation of $x_h = (q_h, u_h, z_h, p_h, v_h, y_h)$ from $(*)$ only requires one more Newton step:

$$0 \stackrel{!}{=} \mathcal{M}'_{(q,u,z)}(x_h) \quad = \quad I'_{(q,u,z)}(q_h,u_h) + \mathcal{L}''_{(q,u,z)}(q_h,u_h,z_h)[(p_h,v_h,y_h)]$$

$$0 = \mathcal{M}'_{(p,v,y)}(x_h) \quad = \quad \mathcal{L}'(q_h,u_h,z_h)[(p_h,v_h,y_h)] \text{ since } (q_h,u_h,z_h) \text{ stat. point}$$

Goal Oriented Error Estimators (III)

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

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 at large error contributions η_i^j
separately for $q \in Q_h$, $u \in V_h$, $z \in V_h$, . . .

Choice of Quantity of Interest $I(q,u)$?

First guess:
Since we wish to reconstruct the coefficient $q = q(x)$,
all $I_x(q,u) := q(x), \;\; x \in \Omega$ are quantities of interest.

These are by far too many!

Choice of Quantity of Interest $I(q,u)$?

First guess:
Since we wish to reconstruct the coefficient $q = q(x)$,
all $I_x(q,u) := q(x), \;\; x \in \Omega$ are quantities of interest.

These are by far too many!

Choice of Quantity of Interest $I(q, u)$?

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

challenge: carrying over infinite dimensional results is ...

Choice of Quantity of Interest $I(q,u)$?

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 $I(q,u)$?

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 $I(q,u)$?

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
≤ 3 quantities of interest

Choice of Quantity of Interest $I(q, u)$?

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
≤ 3 quantities of interest

$\rightsquigarrow$ have a look into the proofs

Choice of Quantity of Interest $I(q,u)$?

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
≤ 3 quantities of interest

$\rightsquigarrow$ have a look into the proofs

Idea of proof for Tikhonov & Discrepancy Principle [3]

- minimality in Q of Tikhonov minimizer $q^\delta_{\alpha_*}$ and $q^\dagger \in Q$

$$\Rightarrow \quad J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) \leq \|F(q^\dagger) - g^\delta\|^2 + \alpha_*\|q^\dagger\|^2 \leq \delta^2 + \alpha_*\|q^\dagger\|^2$$

[3] $q^\dagger \ldots$ exact solution of inverse problem $F(q^\dagger) = Cu^\dagger = g$

q^δ_α Tikhonov minimizer, $\quad u^\delta_\alpha = S(q^\delta_\alpha) \ldots$ corresponding state

$\alpha_* \ldots$ regularization parameter according to discrepancy principle

Idea of proof for Tikhonov & Discrepancy Principle [3]

- minimality in Q of Tikhonov minimizer $q^\delta_{\alpha_*}$ and $q^\dagger \in Q$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) \leq \|F(q^\dagger) - g^\delta\|^2 + \alpha_*\|q^\dagger\|^2 \leq \delta^2 + \alpha_*\|q^\dagger\|^2$

- discrepancy principle $\underline{\underline{\tau}}^2\delta^2 \leq \|F(q^\delta_{\alpha_*}) - g^\delta\|^2$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) = \|F(q^\delta_{\alpha_*}) - g^\delta\|^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2 \geq \underline{\underline{\tau}}^2\delta^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2$

---

[3] $q^\dagger \ldots$ exact solution of inverse problem $F(q^\dagger) = Cu^\dagger = g$

q^δ_α Tikhonov minimizer, $u^\delta_\alpha = S(q^\delta_\alpha) \ldots$ corresponding state

$\alpha_* \ldots$ regularization parameter according to discrepancy principle

Idea of proof for Tikhonov & Discrepancy Principle [3]

- minimality in Q of Tikhonov minimizer $q^{\delta}_{\alpha_*}$ and $q^{\dagger} \in Q$

$\Rightarrow$ $J_{\alpha_*}(q^{\delta}_{\alpha_*}, u^{\delta}_{\alpha_*}) \leq \|F(q^{\dagger}) - g^{\delta}\|^2 + \alpha_* \|q^{\dagger}\|^2 \leq \delta^2 + \alpha_* \|q^{\dagger}\|^2$

- discrepancy principle $\underline{\underline{\tau}}^2 \delta^2 \leq \|F(q^{\delta}_{\alpha_*}) - g^{\delta}\|^2$

$\Rightarrow$ $J_{\alpha_*}(q^{\delta}_{\alpha_*}, u^{\delta}_{\alpha_*}) = \|F(q^{\delta}_{\alpha_*}) - g^{\delta}\|^2 + \alpha_* \|q^{\delta}_{\alpha_*}\|^2 \geq \underline{\underline{\tau}}^2 \delta^2 + \alpha_* \|q^{\delta}_{\alpha_*}\|^2$

- sum up, use $\underline{\underline{\tau}}^2 \delta^2 > \delta^2$, divide by α_* $\Rightarrow$ $\|q^{\delta}_{\alpha_*}\|^2 \leq \|q^{\dagger}\|^2$

$\Rightarrow$ $\exists$ weakly convergent subsequence $q^{\delta}_{\alpha_*} \rightharpoonup \bar{q}$ as $\delta \to 0$.

---

[3] $q^{\dagger}$... exact solution of inverse problem $F(q^{\dagger}) = Cu^{\dagger} = g$

q^{δ}_{α} Tikhonov minimizer, $u^{\delta}_{\alpha} = S(q^{\delta}_{\alpha})$... corresponding state

α_*... regularization parameter according to discrepancy principle

Idea of proof for Tikhonov & Discrepancy Principle [3]

- minimality in Q of Tikhonov minimizer $q^\delta_{\alpha_*}$ and $q^\dagger \in Q$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) \leq \|F(q^\dagger) - g^\delta\|^2 + \alpha_*\|q^\dagger\|^2 \leq \delta^2 + \alpha_*\|q^\dagger\|^2$

- discrepancy principle $\underline{\underline{\tau}}^2\delta^2 \leq \|F(q^\delta_{\alpha_*}) - g^\delta\|^2$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) = \|F(q^\delta_{\alpha_*}) - g^\delta\|^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2 \geq \underline{\underline{\tau}}^2\delta^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2$

- sum up, use $\underline{\underline{\tau}}^2\delta^2 > \delta^2$, divide by α_* $\Rightarrow$ $\|q^\delta_{\alpha_*}\|^2 \leq \|q^\dagger\|^2$

$\Rightarrow$ $\exists$ weakly convergent subsequence $q^\delta_{\alpha_*} \rightharpoonup \bar{q}$ as $\delta \to 0$.

- discrepancy principle $\|F(q^\delta_{\alpha_*}) - g^\delta\|^2 \leq \overline{\overline{\tau}}\delta^2$ $\delta \to 0$

$\Rightarrow$ $F(\bar{q}) = g$

---

[3] $q^\dagger$... exact solution of inverse problem $F(q^\dagger) = Cu^\dagger = g$

q^δ_α Tikhonov minimizer, $u^\delta_\alpha = S(q^\delta_\alpha)$... corresponding state

α_*... regularization parameter according to discrepancy principle

Idea of proof for Tikhonov & Discrepancy Principle [3]

- minimality in Q of Tikhonov minimizer $q^\delta_{\alpha_*}$ and $q^\dagger \in Q$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) \leq \|F(q^\dagger) - g^\delta\|^2 + \alpha_*\|q^\dagger\|^2 \leq \delta^2 + \alpha_*\|q^\dagger\|^2$

- discrepancy principle $\underline{\underline{\tau}}^2\delta^2 \leq \|F(q^\delta_{\alpha_*}) - g^\delta\|^2$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) = \|F(q^\delta_{\alpha_*}) - g^\delta\|^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2 \geq \underline{\underline{\tau}}^2\delta^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2$

- sum up, use $\underline{\underline{\tau}}^2\delta^2 > \delta^2$, divide by α_* $\Rightarrow$ $\|q^\delta_{\alpha_*}\|^2 \leq \|q^\dagger\|^2$

$\Rightarrow$ $\exists$ weakly convergent subsequence $q^\delta_{\alpha_*} \rightharpoonup \bar{q}$ as $\delta \to 0$.

- discrepancy principle $\|F(q^\delta_{\alpha_*}) - g^\delta\|^2 \leq \overline{\overline{\tau}}\delta^2$ $\delta \to 0$

$\Rightarrow$ $F(\bar{q}) = g$

[3] $q^\dagger \ldots$ exact solution of inverse problem $F(q^\dagger) = Cu^\dagger = g$

q^δ_α Tikhonov minimizer, $u^\delta_\alpha = S(q^\delta_\alpha)\ldots$ corresponding state

$\alpha_* \ldots$ regularization parameter according to discrepancy principle

Idea of proof for Tikhonov & Discrepancy Principle

- minimality in Q of Tikhonov minimizer $q^\delta_{\alpha_*}$ and $q^\dagger \in Q$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) \leq \|F(q^\dagger) - g^\delta\|^2 + \alpha_*\|q^\dagger\|^2 \leq \delta^2 + \alpha_*\|q^\dagger\|^2$

- discrepancy principle $\underline{\underline{\tau}}^2\delta^2 \leq \|F(q^\delta_{\alpha_*}) - g^\delta\|^2$

$\Rightarrow$ $J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) = \|F(q^\delta_{\alpha_*}) - g^\delta\|^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2 \geq \underline{\underline{\tau}}^2\delta^2 + \alpha_*\|q^\delta_{\alpha_*}\|^2$

- sum up, use $\underline{\underline{\tau}}^2\delta^2 > \delta^2$, divide by α_* $\Rightarrow$ $\|q^\delta_{\alpha_*}\|^2 \leq \|q^\dagger\|^2$

$\Rightarrow$ $\exists$ weakly convergent subsequence $q^\delta_{\alpha_*} \rightharpoonup \bar{q}$ as $\delta \to 0$.

- discrepancy principle $\|F(q^\delta_{\alpha_*}) - g^\delta\|^2 \leq \overline{\overline{\tau}}\delta^2$ $\delta \to 0$

$\Rightarrow$ $F(\bar{q}) = g$

Idea of proof for Tikhonov & Discrepancy Principle

- minimality in Q of Tikhonov minimizer $q^\delta_{\alpha_*}$ and $q^\dagger \in Q$

with $|J_{\alpha_*}(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}) - |J_{\alpha_*}(q^\delta_{\alpha_* h}, u^\delta_{\alpha_* h})| \leq \eta_1 \leq (\underline{\underline{\tau}}^2 - 1)\delta^2$

$\Rightarrow\ J_{\alpha_*}(q^\delta_{\alpha_* h}, u^\delta_{\alpha_* h}) \leq \|F(q^\dagger) - g^\delta\|^2 + \alpha_*\|q^\dagger\|^2 + \eta_1 \leq \delta^2 + \alpha_*\|q^\dagger\|^2 + \eta_1$

- discrepancy principle $\underline{\underline{\tau}}^2\delta^2 \leq \|F_h(q^\delta_{\alpha_* h}) - g^\delta\|^2$

$\Rightarrow J_{\alpha_*}(q^\delta_{\alpha_* h}, u^\delta_{\alpha_* h}) = \|F_h(q^\delta_{\alpha_* h}) - g^\delta\|^2 + \alpha_*\|q^\delta_{\alpha_* h}\|^2 \geq \underline{\underline{\tau}}^2\delta^2 + \alpha_*\|q^\delta_{\alpha_* h}\|^2$

- sum up, use $\underline{\underline{\tau}}^2\delta^2 > \delta^2 + \eta_1$, divide by α_* $\Rightarrow$ $\|q^\delta_{\alpha_* h}\|^2 \leq \|q^\dagger\|^2$

$\Rightarrow\ \exists$ weakly convergent subsequence $q^\delta_{\alpha_* h} \rightharpoonup \bar{q}$ as $\delta \to 0$.

- discrepancy principle $\|F_h(q^\delta_{\alpha_* h}) - g^\delta\|^2 \leq \overline{\overline{\tau}}\delta^2$ $\delta \to 0$

with $|\|F(q^\delta_{\alpha_* h}) - g^\delta\|^2 - \|F_h(q^\delta_{\alpha_* h}) - g^\delta\|^2| \leq \eta_2 \to 0$ as $\delta \to 0$

$\Rightarrow\ F(\bar{q}) = g$

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

$\underline{\underline{\tau}}^2\delta^2 \leq \|F_h(q^\delta_{h,\alpha_*}) - g^\delta\|^2_G = \|Cu^\delta_{h,\alpha_*} - g^\delta\|^2_G \leq \overline{\overline{\tau}}\delta^2$

$I_1(q,u,\alpha) = J_\alpha(q,u) = \|Cu - g^\delta\|^2_G + \alpha\|q\|^2$

satisfies $|I_1(q^\delta_{\alpha_*}, u^\delta_{\alpha_*}, \alpha_*) - I_1(q^\delta_{h,\alpha_*}, u^\delta_{h,\alpha_*}, \alpha_*)| \leq (\underline{\underline{\tau}}^2 - 1)\delta^2$

$I_2(u,\alpha) := \|F(q_{h,\alpha_*}) - g^\delta\|^2_G = \|Cu - g^\delta\|^2_G$

satisfies $|I_2(u^\delta_{\alpha_*}, \alpha_*) - I_2(u^\delta_{h,\alpha_*}, \alpha_*)| \leq cI_2(u^\delta_{h,\alpha_*}, \alpha_*)$

Then $q^\delta_{\alpha_*} \to q^\dagger$ as $\delta \to 0$.

(Optimal rates under source conditions of logarithmic/Hölder type.)

Remarks

- Also works for stationary points q^{δ}_{h,α_*} instead of global minimizers if F is not too nonlinear
- Also works in Banach spaces with general data misfit and (convex) regularization term [4]
  $J_\alpha(q,u) = \mathcal{S}(Cu, g^\delta) + \alpha\mathcal{R}(q)$

[4] see, e.g., the PhD theses of Christiane Pöschl 2008 (Otmar Scherzer), Jens Flemming 2011 (Bernd Hofmann), Frank Werner 2012 (Thorsten Hohage) for the continuous setting.

Efficient Computation of α $\rightsquigarrow$ Choice of QoI

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

$$\|F(q^{\delta}_{\alpha_*}) - g^{\delta}\| \approx \tau\delta$$

$\rightsquigarrow$ 1-d nonlinear equation $\phi(\alpha) = 0$ for α
"less nonlinear" version $\psi(\beta) = \phi(\frac{1}{\beta}) = 0$ for β

$\rightsquigarrow$ solve by Newton's method

$$\beta^{k+1} = \beta^k - \frac{\psi(\beta^k)}{\psi'(\beta^k)}$$

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

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

$I_1(q,u) := \psi(\beta) = \psi(\frac{1}{\alpha}) = \|F(q) - g^\delta\|_G^2 - \tau^2\delta^2 = \|Cu - g^\delta\|_G^2 - \tau^2\delta^2$

$I_2(q,u) := \psi'(\beta)$

$\beta^{k+1} = \beta^k - \frac{\psi_h^k}{{\psi'}_h^k}$ (approximate Newton method)

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

$|\psi(\beta^k) - \psi_h^k| \leq \varepsilon^k\,,\qquad |\psi'(\beta^k) - {\psi'}_h^k| \leq {\varepsilon'}^k\,,\quad \varepsilon^k, {\varepsilon'}^k$ sufficiently small.

Then β^k satisfies quadratic convergence estimate and

$\underline{\underline{\tau}}^2\delta^2 \leq \|F_h(q_h, \frac{1}{\beta_{k_*}}) - g^\delta\|_G^2 \leq \overline{\overline{\tau}}^2\delta^2$

Remarks

- computation of error estimators for $\psi(\beta)$:
  just one more SQP type step after $\mathcal{L}'(q,u,z)[(p,v,y)] = 0$;
- evaluation of $\psi'(\beta)$:
  can be directly extracted from quantities computed for error estimators for $\psi(\beta)$
- error estimators for $\psi'(\beta)$:
  stationary point of another auxiliary functional
  by another SQP step

Numerical Tests

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

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

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

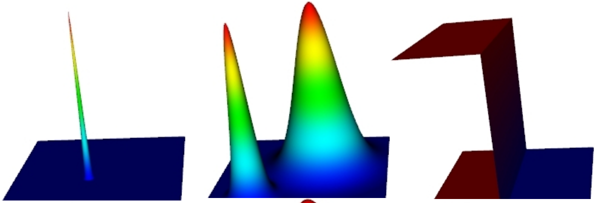
$$q_i = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\left(\frac{s_i x - \frac{1}{2}}{\sigma}\right)^2 + \left(\frac{s_i y - \frac{1}{2}}{\sigma}\right)^2\right)\right), \quad \begin{array}{l} \sigma = 0.1 \\ s_1 = 2, \\ s_2 = 0.8 \end{array}$$

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

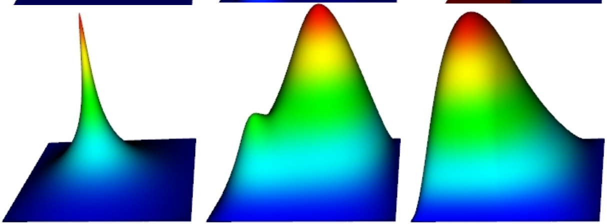
Computations with *Gascoigne* and *RoDoBo*.

Numerical Tests

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



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



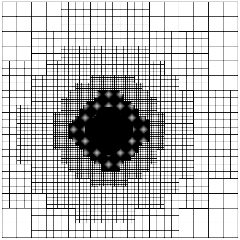
Numerical Results (I)

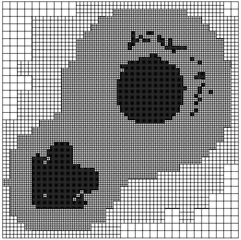
Computations with 1% random noise:

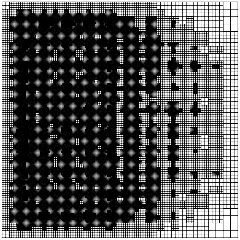
number of nodes on finest grid:

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

adaptive grids:







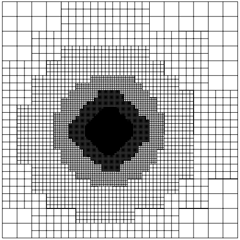
wrong regularization term

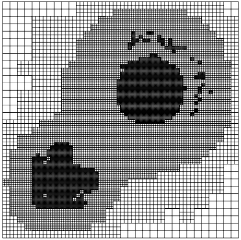
Numerical Results (I)

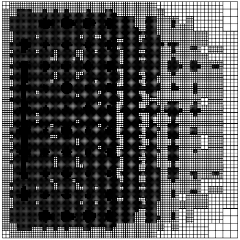
Computations with 1% random noise:
number of nodes on finest grid:

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

adaptive grids:





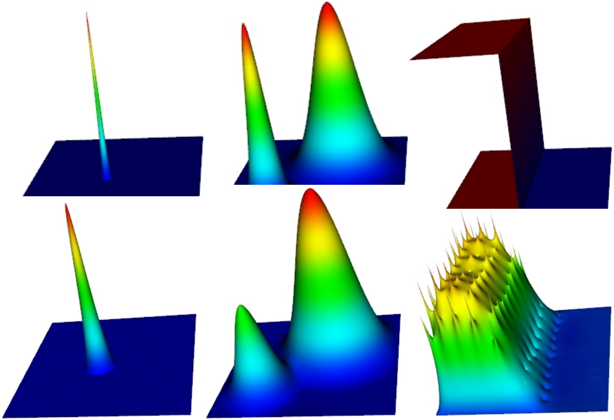


wrong regularization term

Numerical Results (II)

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

computed par.
(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

Iterative Regularization

Coefficient Identification in PDE as Operator Equation

$$\begin{aligned} A(q,u)(v) = & \quad (f,v) \quad \forall v \in V & \quad \ldots \text{ PDE in weak form} \\ Cu = & \quad g & \quad \ldots \text{ measurements} \end{aligned}$$

or equivalently

$$F(q) = g$$

$F \ldots$ forward operator: $F(q) = (C \circ S)(q) = Cu$
where $u = S(q)$ solves PDE; $\quad S \ldots$ coefficient-to-state-map

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

inverse problem: identify q from measurements g^δ of g

Newton type Regularization

Newton step as least squares problem:

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

Iteratively Regularized Gauss-Newton Method IRGNM

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

or equivalently

Minimize

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

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

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

a posteriori selection of α_k (inexact Newton)

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

Newton type Regularization and the Discrepancy Principle

Iteratively Regularized Gauss-Newton Method IRGNM

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

a posteriori selection of α_k (inexact Newton)

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

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

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

a posteriori selection of α_k (inexact Newton)

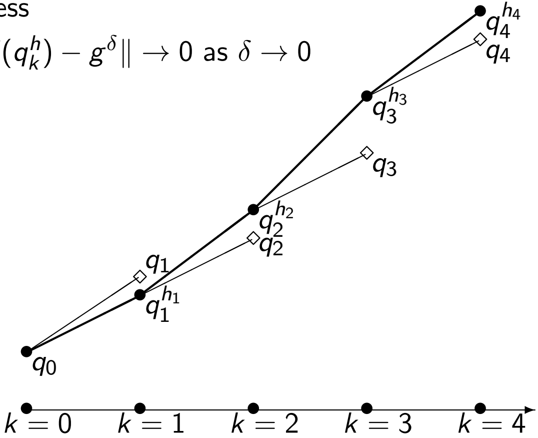
$$\underline{\tilde{\theta}}\|F(q_k) - g^\delta\| \leq \|F'(q_k)(q_{k+1} - q_k) + F(q_k) - g^\delta\| \leq \bar{\tilde{\theta}}\|F(q_k) - g^\delta\|$$

a posteriori selection of k_* (discrepancy principle)

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

Idea of Proof

- minimality of q_k in Q, compare with $q^\dagger$
  $\rightsquigarrow$ boundedness
- show that $\|F(q_k^h) - g^\delta\| \to 0$ as $\delta \to 0$



Convergence Analysis $\rightsquigarrow$ Choice of QoI

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

where for fixed q_k^H and variable q_{k+1}^h, w_k^h, u_k^h, u_{k+1}^h

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

Theorem [BK&Kirchner&Veljovic&Vexler'12]:
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 of logarithmic/Hölder type).

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

where for fixed q_k^H and variable q_{k+1}^h, w_k^h, u_k^h, u_{k+1}^h

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

$$I_{2,h}^{k+1} = \|F_h'(q_k^H)(q_{k+1}^h - q_k^H) + F_h(q_k^H) - g^\delta\|^2$$

⇝ solution of a linear PDE

$$I_{3,h}^{k+1} = \|F_h(q_k^H) - g^\delta\|^2$$

⇝ solution of a nonlinear PDE

⇝ all-at once formulations

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

where for fixed q_k^H and variable q_{k+1}^h, w_k^h, u_k^h, u_{k+1}^h

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

$$I_{2,h}^{k+1} = \|F_h'(q_k^H)(q_{k+1}^h - q_k^H) + F_h(q_k^H) - g^\delta\|^2$$

⇝ solution of a linear PDE

$$I_{3,h}^{k+1} = \|F_h(q_k^H) - g^\delta\|^2$$

⇝ solution of a nonlinear PDE

⇝ all-at once formulations

A Least Squares Formulation (I)

$$\left.\begin{array}{llcl} \text{measurements:} & Cu & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array}\right\} \Leftrightarrow: \ \mathbf{F}(u,q) = \mathbf{g}$$

$\rightsquigarrow$ Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q^\delta_{k+1} \\ u^\delta_{k+1} \end{pmatrix} = \begin{pmatrix} q^\delta_k \\ u^\delta_k \end{pmatrix} - \left(\mathbf{F}'(q^\delta_k, u^\delta_k)^* \mathbf{F}'(q^\delta_k, u^\delta_k) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\right)^{-1} \times \left(\mathbf{F}'(q^\delta_k, u^\delta_k)^* (\mathbf{F}(q^\delta_k, u^\delta_k) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q^\delta_k - q_0 \\ 0 \end{pmatrix}\right)$$

A Least Squares Formulation (I)

$$\left.\begin{array}{llcl} \text{measurements:} & Cu & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array}\right\} \Leftrightarrow: \ \mathbf{F}(u,q) = \mathbf{g}$$

$\rightsquigarrow$ Iteratively Regularized Gauss-Newton Method IRGNM

$$\begin{pmatrix} q^\delta_{k+1} \\ u^\delta_{k+1} \end{pmatrix} = \begin{pmatrix} q^\delta_{k} \\ u^\delta_{k} \end{pmatrix} - \left(\mathbf{F}'(q^\delta_k,u^\delta_k)^*\mathbf{F}'(q^\delta_k,u^\delta_k) + \alpha_k \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\right)^{-1} \times \left(\mathbf{F}'(q^\delta_k,u^\delta_k)^*(\mathbf{F}(q^\delta_k,u^\delta_k) - \mathbf{g}^\delta) + \alpha_k \begin{pmatrix} q^\delta_k - q_0 \\ 0 \end{pmatrix}\right)$$

or equivalently: unconstrained quadratic minimization

$$\begin{pmatrix} q^\delta_{k+1} \\ u^\delta_{k+1} \end{pmatrix} = \arg\min_{q,u} \|A'_q(q_k,u_k)(q-q_k) + A'_u(q_k,u_k)(u-u_k) + A(q_k,u_k) - f\|^2_{V^*} + \|Cu - g^\delta\|^2_G + \alpha_k\|q-q_0\|^2_Q$$

A Least Squares Formulation (I)

$$\left.\begin{array}{llcl} \text{measurements:} & Cu & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array}\right\} \Leftrightarrow: \ \mathbf{F}(u,q) = \mathbf{g}$$

$\rightsquigarrow$ 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

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

A Least Squares 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 $K = A'_u(q_k,u_k),\ L = A'_q(q_k,u_k)$.

$$
\left.\begin{array}{l} K \text{ regular} \\ \alpha > 0 \end{array}\right\} \Rightarrow \text{ Hessian } \begin{pmatrix} L^*L + \alpha I & L^*K \\ K^*L & C^*C + K^*K \end{pmatrix} \text{ positive definite.}
$$

A Least Squares Formulation (II)

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

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

$$
\left.\begin{array}{l} K \text{ regular} \\ \alpha > 0 \end{array}\right\} \Rightarrow \text{ Hessian } \begin{pmatrix} L^*L + \alpha I & L^*K \\ K^*L & C^*C + K^*K \end{pmatrix} \text{ positive definite.}
$$

Least Squares 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)

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

Least Squares 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)

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

a posteriori selection of k_* (discrepancy principle)

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

Least Squares 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)

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

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

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

where for fixed q_k^H and variable q_{k+1}^h, w_k^h, u_k^h, u_{k+1}^h

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

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

$$I_{3,h}^{k+1} = \|\mathbf{F}_h(q_k^H, u_k^H) - \mathbf{g}^\delta\|^2$$

$$I_{4,h}^{k+1} = \|\mathbf{F}_h(q_{k+1}^h, u_{k+1}^h) - \mathbf{g}^\delta\|^2$$

Theorem [BK&Kirchner&Veljovic&Vexler'12]:
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 of logarithmic/Hölder type).

Convergence Analysis ⇝ Choice of QoI

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

where for fixed q_k^H and variable q_{k+1}^h, w_k^h, u_k^h, u_{k+1}^h

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

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

⇝ evaluate residual of a linear PDE

$$I_{3,h}^{k+1} = \|\mathbf{F}_h(q_k^H, u_k^H) - \mathbf{g}^\delta\|^2$$

⇝ evaluate residual of a nonlinear PDE

Remarks

- Also works in Banach spaces with general data misfit and (convex) regularization term [5]

$$J_k(q,u) = \mathcal{S}_1(Cu, g^\delta) + \mathcal{S}_2(L_k(q-q_k) + K_k(u-u_k) + A(q_k,u_k), f) + \alpha_k \mathcal{R}(q,u)$$

- Use this to avoid equal treatment of both equations in

$$\left.\begin{array}{llcl} \text{measurements:} & Cu & = & g \text{ in } G \\ \text{PDE:} & A(q,u) & = & f \text{ in } V^* \end{array}\right\}$$

by least-squares approach.
("more confidence in PDE than in measurements $g^\delta \approx g$")

[5] see, e.g., the PhD thesis of Frank Werner 2012 (Thorsten Hohage) for the continuous setting.

A Generalized Newton Method (I)

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \arg\min_{q,u} \frac{1}{2}\|Cu - g^\delta\|_G^2 + \frac{\alpha_k}{2}\|q - q_0\|_Q^2$$
$$\text{s.t.} \quad A'_q(q_k^\delta, u_k^\delta)(q - q_k^\delta) + A'_u(q_k^\delta, u_k^\delta)(u - u_k^\delta) + A(q_k^\delta, u_k^\delta) = f$$

or equivalently (by exactness of l^1 penalty):

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \arg\min_{q,u} \frac{1}{2}\|Cu - g^\delta\|_G^2 + \frac{\alpha_k}{2}\|q - q_0\|_Q^2$$
$$+\rho\|A'_q(q_k, u_k)(q - q_k) + A'_u(q_k, u_k)(u - u_k) + A(q_k, u_k) - f\|_{V^*}$$

with ρ sufficiently large (but finite).

A Generalized Newton Method (II)

$$
\begin{aligned}
(q_{k+1}^\delta, u_{k+1}^\delta) = \arg\min_{q,u} \frac{1}{2}\|Cu - g^\delta\|_G^2 + \frac{\alpha_k}{2}\|q - q_0\|_Q^2 \\
\text{s.t. } L(q - q_k) + K(u - u_k) + A(q_k, u_k) - f = 0
\end{aligned}
$$

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

First order optimality system:

$$
\begin{pmatrix} \alpha_k I & 0 & L^* \\ 0 & C^*C & K^* \\ L & K & 0 \end{pmatrix}
\begin{pmatrix} q_{k+1} \\ u_{k+1} \\ z_{k+1} \end{pmatrix}
=
\begin{pmatrix} \alpha_k q_0 \\ C^* g^\delta \\ Lq_k + Ku_k - A(q_k, u_k) + f \end{pmatrix}
$$

A Generalized Newton Method (II)

$$(q^\delta_{k+1}, u^\delta_{k+1}) = \arg\min_{q,u} \frac{1}{2}\|Cu - g^\delta\|^2_G + \frac{\alpha_k}{2}\|q - q_0\|^2_Q$$
$$\text{s.t. } L(q - q_k) + K(u - u_k) + A(q_k, u_k) - f = 0$$

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

First order optimality system:

$$\begin{pmatrix} \alpha_k I & 0 & L^* \\ 0 & C^*C & K^* \\ L & K & 0 \end{pmatrix} \begin{pmatrix} q_{k+1} \\ u_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} \alpha_k q_0 \\ C^* g^\delta \\ Lq_k + Ku_k - A(q_k, u_k) + f \end{pmatrix}$$

Generalized Newton Type Regularization and the Discrepancy Principle

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \arg\min_{q,u} \frac{1}{2}\|Cu - g^\delta\|_G^2 + \frac{\alpha_k}{2}\|q - q_0\|_Q^2$$

s.t. $A'_q(q_k^\delta, u_k^\delta)(q - q_k^\delta) + A'_u(q_k^\delta, u_k^\delta)(u - u_k^\delta) + A(q_k^\delta, u_k^\delta) = f$

a posteriori selection of α_k

$$\underline{\tilde{\theta}}(\|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_k) - f\|) \leq \|C(u_{k+1}) - g^\delta\|_G^2$$
$$\leq \tilde{\tilde{\overline{\theta}}}(\|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_k) - f\|)$$

Generalized Newton Type Regularization and the Discrepancy Principle

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \arg\min_{q,u} \tfrac{1}{2}\|Cu - g^\delta\|_G^2 + \tfrac{\alpha_k}{2}\|q - q_0\|_Q^2$$

$$\text{s.t.} \quad A'_q(q_k^\delta, u_k^\delta)(q - q_k^\delta) + A'_u(q_k^\delta, u_k^\delta)(u - u_k^\delta) + A(q_k^\delta, u_k^\delta) = f$$

a posteriori selection of α_k

$$\underline{\tilde{\theta}}(\|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_k) - f\|) \le \|C(u_{k+1}) - g^\delta\|_G^2 \le \bar{\tilde{\theta}}(\|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_k) - f\|)$$

a posteriori selection of k_* (discrepancy principle)

$$k^* = \min\{k \in \mathbb{N} \;:\; \|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_{k+1}) - f\| \le \tau^2\delta^2\}$$

Generalized Newton Type Regularization and the Discrepancy Principle

$$\begin{pmatrix} q_{k+1}^\delta \\ u_{k+1}^\delta \end{pmatrix} = \arg\min_{q,u} \frac{1}{2}\|Cu - g^\delta\|_G^2 + \frac{\alpha_k}{2}\|q - q_0\|_Q^2$$

$$\text{s.t.} \quad A'_q(q_k^\delta, u_k^\delta)(q - q_k^\delta) + A'_u(q_k^\delta, u_k^\delta)(u - u_k^\delta) + A(q_k^\delta, u_k^\delta) = f$$

a posteriori selection of α_k

$$\underline{\tilde{\theta}}(\|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_k) - f\|) \leq \|C(u_{k+1}) - g^\delta\|_G^2 \leq \bar{\tilde{\theta}}(\|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_k) - f\|)$$

a posteriori selection of k_* (discrepancy principle)

$$k^* = \min\{k \in \mathbb{N} \ : \ \|C(u_k) - g^\delta\|^2 + \rho\|A(q_k, u_{k+1}) - f\| \leq \tau^2\delta^2\}$$

Convergence Analysis ⇝ Choice of QoI

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

where for fixed q_k^H and variable q_{k+1}^h, w_k^h, u_k^h, u_{k+1}^h

$$\begin{aligned}
I_{1,h}^{k+1} &= \|C(u_{k+1}^h) - g^\delta\|_G^2 + \alpha_k \|q_{k+1}^h - q_0\|_Q^2 \\
I_{2,h}^{k+1} &= \|C(u_{k+1}^h) - g^\delta\|_G^2 \\
I_{3,h}^{k+1} &= \|C(u_k^h) - g^\delta\|^2 + \rho \|A(q_k^H, u_k^H) - f\| \\
I_{4,h}^{k+1} &= \|C(u_{k+1}^h) - g^\delta\|_G^2 + \rho \|A(q_{k+1}^H, u_{k+1}^h) - f\|_{V^*}
\end{aligned}$$

Theorem [BK&Kirchner&Veljovic&Vexler'12]:
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 of logarithmic/Hölder type).

Idea of proof: equivalence to exact l^1 penalty formulation

Outlook

- $\rightarrow$ other regularization methods: e.g., regularization by discretization
- $\rightarrow$ other parameter choice strategies: e.g., balancing principle
- $\rightarrow$ other noise models: e.g., Poisson noise
- $\rightarrow$ other PDEs: e.g., time adaptivity
- $\rightarrow$ other error estimators: e.g., functional estimators $\rightsquigarrow$ residuals
- $\rightarrow$ ...

Thank you for your attention!