All-at-once versus reduced formulations of inverse problems and their regularization

Barbara Kaltenbacher

Alpen-Adria-Universität Klagenfurt

joint work with

Christian Clason, Alana Kirchner, Boris Vexler

27 September, 2017

ENUMATH, Voss

Outline

- examples of inverse problems
- regularization: Tikhonov, Newton type and Landweber in
  - reduced formulation
  - all-at-once formulation
- numerical results
- minimization based formulations

examples

Parameter Identification in Differential Equations: Some Examples

- Identify spatially varying coefficients/source a, b, c in linear elliptic boundary value problem on $\Omega \subseteq \mathbb{R}^d$, $d \in \{1,2,3\}$

$$-\nabla(a\nabla u) + cu = b \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = j \text{ on } \partial\Omega\,,$$

from boundary or (restricted) interior observations of u.

Parameter Identification in Differential Equations: Some Examples

- Identify spatially varying coefficients/source a, b, c in linear elliptic boundary value problem on $\Omega \subseteq \mathbb{R}^d$, $d \in \{1, 2, 3\}$

$$-\nabla(a\nabla u) + cu = b \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = j \text{ on } \partial\Omega\,,$$

from boundary or (restricted) interior observations of u.

- Identify source term q in nonlinear elliptic bvp

$$-\Delta u + \xi(u) = q \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = j \text{ on } \partial\Omega\,,$$

from boundary or (restricted) interior observations of u.

Parameter Identification in Differential Equations: Some Examples

- Identify spatially varying coefficients/source a, b, c in linear elliptic boundary value problem on $\Omega \subseteq \mathbb{R}^d$, $d \in \{1, 2, 3\}$

$$-\nabla(a\nabla u) + cu = b \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = j \text{ on } \partial\Omega\,,$$

  from boundary or (restricted) interior observations of u.
- Identify source term q in nonlinear elliptic bvp

$$-\Delta u + \xi(u) = q \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = j \text{ on } \partial\Omega\,,$$

  from boundary or (restricted) interior observations of u.
- Identify parameter ϑ in initial value problem for ODE / PDE

$$\dot{u}(t) = f(t, u(t), \vartheta) \; t \in (0, T)\,, \quad u(0) = u_0$$

  from discrete of continuous observations of u.
  $y_i = g_i(u(t_i))$, $i \in \{1, \ldots, m\}$ or $y(t) = g(t, y(t))$, $t \in (0, T)$

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q, u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

where $q \in X$, $u \in V$, $y \in Y$, $X, V, Y \ldots$ Hilbert (Banach) spaces
$A : X \times V \to W^* \ldots$ differential operator
$C : V \to Y \ldots$ observation operator

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q, u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

where $q \in X$, $u \in V$, $y \in Y$, $X, V, Y \ldots$ Hilbert (Banach) spaces
$A : X \times V \to W^* \ldots$ differential operator
$C : V \to Y \ldots$ observation operator

(a) reduced approach: operator equation for q

$$F(q) = y,$$

$F = C \circ S$ with $S : X \to V$, $q \mapsto u$ parameter-to-state map

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q,u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

where $q \in X$, $u \in V$, $y \in Y$, $X, V, Y \ldots$ Hilbert (Banach) spaces
$A : X \times V \to W^* \ldots$ differential operator
$C : V \to Y \ldots$ observation operator

(a) reduced approach: operator equation for q

$$F(q) = y,$$

$F = C \circ S$ with $S : X \to V$, $q \mapsto u$ parameter-to-state map

(b) all-at once approach: observations and model as system for (q, u)

$$\begin{array}{rcl} A(q,u) & = & 0 \text{ in } W^* \\ C(u) & = & y \text{ in } Y \end{array} \quad \Leftrightarrow \quad \mathbf{F}(q,u) = \mathbf{y}$$

The Parameter-to-State Map S in some Examples

- Identify spatially varying coefficients/source a, b, c in

$$-\nabla(a\nabla u) + cu = b \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega\,,$$

$S : (a, b, c) \mapsto u$ solving the linear elliptic bvp

The Parameter-to-State Map S in some Examples

- Identify spatially varying coefficients/source a, b, c in

$$-\nabla(a\nabla u) + cu = b \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega\,,$$

$S : (a, b, c) \mapsto u$ solving the linear elliptic bvp

- Identify source term q in

$$-\Delta u + \xi(u) = q \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega\,,$$

$S : q \mapsto u$ solving the nonlinear elliptic bvp

The Parameter-to-State Map S in some Examples

- Identify spatially varying coefficients/source a, b, c in

$$-\nabla(a\nabla u) + cu = b \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega\,,$$

$S : (a, b, c) \mapsto u$ solving the linear elliptic bvp

- Identify source term q in

$$-\Delta u + \xi(u) = q \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega\,,$$

$S : q \mapsto u$ solving the nonlinear elliptic bvp

- Identify parameter ϑ in

$$\dot{u}(t) = f(t, u(t), \vartheta) \; t \in (0, T)\,, \quad u(0) = u_0$$

$S : \vartheta \mapsto u$ solving the ivp for an ODE / PDE

The Parameter-to-State Map S in some Examples

- Identify spatially varying coefficients/source a, b, c in

$$-\nabla(a\nabla u) + cu = b \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega\,,$$

$S : (a, b, c) \mapsto u$ solving the linear elliptic bvp

- Identify source term q in

$$-\Delta u + \xi(u) = q \text{ in } \Omega\,, \qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega\,,$$

$S : q \mapsto u$ solving the nonlinear elliptic bvp

- Identify parameter ϑ in

$$\dot{u}(t) = f(t, u(t), \vartheta) \; t \in (0, T)\,, \quad u(0) = u_0$$

$S : \vartheta \mapsto u$ solving the ivp for an ODE / PDE

- generally for model $\quad A(q, u) = 0$:

$$S : q \mapsto u \text{ solving } A(q, S(q)) = 0$$

Motivation for All-at-once Formulation

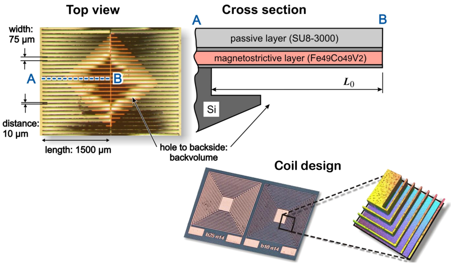
- well-definedness of parameter-to-state map often requires restrictions on ...
  ... parameters (e.g., $a \geq \underline{a} > 0$, $c \geq 0$ in $-\nabla(a\nabla u) + cu = b$)
  ... models (e.g., monotonicity of ξ in $-\Delta u + \xi(u) = q$)

Motivation for All-at-once Formulation

- well-definedness of parameter-to-state map often requires restrictions on ...
  ... parameters (e.g., $a \geq \underline{a} > 0$, $c \geq 0$ in $-\nabla(a\nabla u) + cu = b$)
  ... models (e.g., monotonicity of ξ in $-\Delta u + \xi(u) = q$)
- singular PDEs: parameter-to space map may exist only on a very restricted set

MicroElectroMechanical Systems (MEMS)

acceleration sensors, microphones, pumps, loudspeakers, . . .

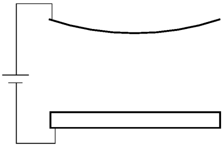


transient MEMS equation

$$u_{tt} + cu_t + du + \rho\Delta^2 u - \eta\Delta u + \frac{b(t)a(x)}{(1+u)^2} = 0$$

u. . . membrane/beam displacement
$b(t)$. . . voltage excitation
$a(x)$. . . dielectric properties

MicroElectroMechanical Systems (MEMS)



$\rightsquigarrow$ control of voltage $b(t)$ and/or design of dielectric properties $a(x)$ to achieve prescribed displacement $y(x,t)$;

$$u_{tt} + cu_t + du + \rho\Delta^2 u - \eta\Delta u + \frac{b(t)a(x)}{(1+u)^2} = 0$$

achieve large displacements $|u|$
avoid pull-in instability at $u = -1$!

parameter-to space map exists only on a very restricted set (too restrictive for certain tracking tasks)

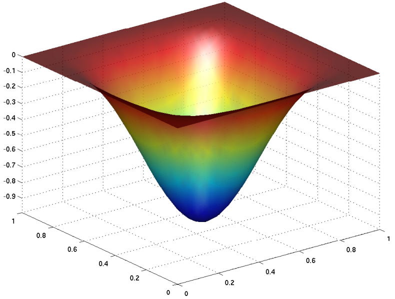
Numerical tests

$$J(a,u) = \frac{1}{2}\,\|u - y_d\|^2_{L^2} + \frac{\alpha}{2}\,\|a\|^2_{L^2}$$

static case

$$-\Delta u + \frac{a(x)}{(1+u)^2} = 0$$

$\Omega = (0,1)^2$, $\alpha = 10^{-6}$, 64×64 grid.

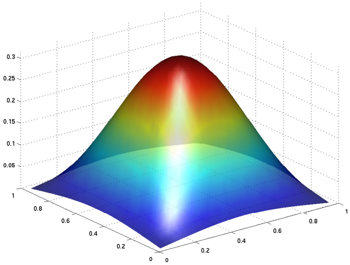


Target y_d (desired maximal deflection: -0.99)

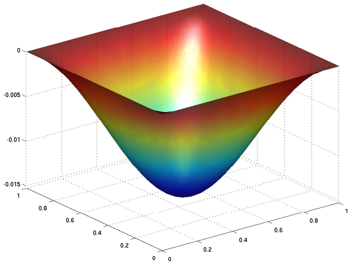
Numerical tests: Using control-to-state map

impose control constraints: $\|a\|_{L^2} \leq \frac{4}{27} = 0.14815\cdots$
to guarantee well-definedness of control-to-state map

optimal control a



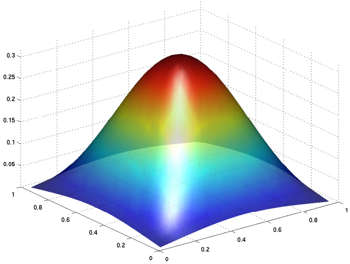
optimal state u



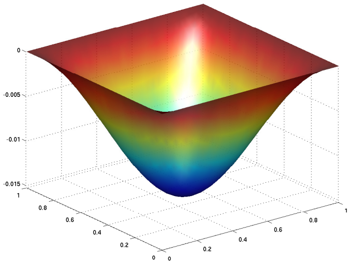
Numerical tests: Using control-to-state map

impose control constraints: $\|a\|_{L^2} \leq \frac{4}{27} = 0.14815\cdots$
to guarantee well-definedness of control-to-state map

optimal control a



optimal state u

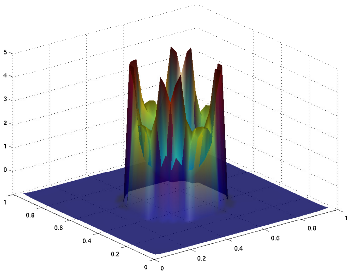


maximal deflection: -0.015!

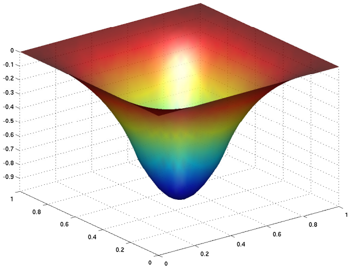
Numerical tests: Not using control-to-state map

impose pointwise state constraints: $u(x) \geq -0.99$
to avoid singularity

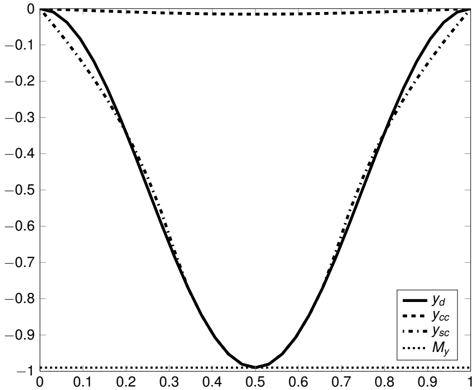
optimal control a



optimal state u



Comparison: with vs without control-to-state map



Cross sections of states for approach with (dashed) and without (dash-dotted) control-to-state map, as well as target y_d (solid) and bound -0.99 (dotted)

Motivation for All-at-once Formulation

- well-definedness of parameter-to-state map often requires restrictions on ...
  ... parameters (e.g., $a \geq \underline{a} > 0$ in $-\nabla(a\nabla u) = f$)
  ... models (e.g., monotonicity of ξ in $-\Delta u + \xi(u) = q$)
- singular PDEs: parameter-to space map may exist only on a very restricted set, e.g. MEMS equation

$$u_{tt} + cu_t + du + \rho\Delta^2 u - \eta\Delta u + \frac{b(t)a(x)}{(1+u)^2} = 0$$

Motivation for All-at-once Formulation

- well-definedness of parameter-to-state map often requires restrictions on ...
  ... parameters (e.g., $a \geq \underline{a} > 0$ in $-\nabla(a\nabla u) = f$)
  ... models (e.g., monotonicity of ξ in $-\Delta u + \xi(u) = q$)
- singular PDEs: parameter-to space map may exist only on a very restricted set, e.g. MEMS equation

$$u_{tt} + cu_t + du + \rho\Delta^2 u - \eta\Delta u + \frac{b(t)a(x)}{(1+u)^2} = 0$$

- it can make a difference in implementation and in the analysis (convergence conditions)

reduced formulation

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q,u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

(a) reduced approach: operator equation for q

$$F(q) = y,$$

$F = C \circ S$ with $S : X \to V$, $q \mapsto u$ parameter-to-state map
existence of parameter-to-state map S requires condition of the type

$$A_u(q,u)^{-1} \text{ exists and } \|A_u(q,u)^{-1}\| \leq C_A$$

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q,u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

(a) reduced approach: operator equation for q

$$F(q) = y,$$

$F = C \circ S$ with $S : X \to V$, $q \mapsto u$ parameter-to-state map
existence of parameter-to-state map S requires condition of the type

$$A_u(q,u)^{-1} \text{ exists and } \|A_u(q,u)^{-1}\| \leq C_A$$

ill-posedness (F not continuously invertible)
only noisy measurements $y^\delta \approx y$ given
$\Rightarrow$ regularization needed

Tikhonov Regularization

regularization functional $\mathcal{R} : X \to \overline{\mathbb{R}}$ (proper, convex)
regularization parameter $\alpha > 0$

$$\min_q \|F(q) - y^\delta\|^2 + \alpha \mathcal{R}(q)$$

Tikhonov Regularization

regularization functional $\mathcal{R} : X \to \overline{\mathbb{R}}$ (proper, convex)
regularization parameter $\alpha > 0$

$$\min_q \|F(q) - y^\delta\|^2 + \alpha\mathcal{R}(q)$$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equivalent to

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \alpha\mathcal{R}(q) \quad \text{s.t. } A(q, u) = 0$$

Tikhonov Regularization

regularization functional $\mathcal{R} : X \to \overline{\mathbb{R}}$ (proper, convex)
regularization parameter $\alpha > 0$

$$\min_q \|F(q) - y^\delta\|^2 + \alpha\mathcal{R}(q)$$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equivalent to

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \alpha\mathcal{R}(q) \quad \text{s.t. } A(q,u) = 0$$

[Seidman&Vogel '89, Engl&Kunisch&Neubauer '89,...] in Hilbert space
[Burger& Osher'04, Resmerita & Scherzer'06, Scherzer et al. '08, Hofmann&Pöschl&BK&Scherzer '07, Pöschl '09, Flemming '11, Werner '12,...] in Banach space

Regularized Gauss-Newton Method

q^k fixed, one Gauss-Newton step:

$$\min_q \|F(q^k) + F'(q^k)(q - q^k) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q)$$

$\rightsquigarrow q^{k+1}$

Regularized Gauss-Newton Method

q^k fixed, one Gauss-Newton step:

$$\min_q \|F(q^k) + F'(q^k)(q - q^k) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q)$$

$\rightsquigarrow q^{k+1}$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equivalent to

$$\min_{q,u,\tilde{u}} \|C(\tilde{u}) + C'(\tilde{u})(u - \tilde{u}) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q)$$

$$\text{s.t. } A(q^k, \tilde{u}) + A'_u(q^k, \tilde{u})(u - \tilde{u}) + A'_q(q^k, \tilde{u})(q - q^k) = 0$$

$$\text{and } A(q^k, \tilde{u}) = 0$$

Regularized Gauss-Newton Method

q^k fixed, one Gauss-Newton step:

$$\min_q \|F(q^k) + F'(q^k)(q - q^k) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q)$$

$\rightsquigarrow q^{k+1}$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equivalent to

$$\min_{q,u,\tilde{u}} \|C(\tilde{u}) + C'(\tilde{u})(u - \tilde{u}) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q)$$

$$\text{s.t. } A(q^k, \tilde{u}) + A'_u(q^k, \tilde{u})(u - \tilde{u}) + A'_q(q^k, \tilde{u})(q - q^k) = 0$$

$$\text{and } A(q^k, \tilde{u}) = 0$$

[Bakushinskii '92, Hohage '97, BK&Neubauer&Scherzer '97,...] in Hilbert space

e.g., [Bakushinskii&Kokurin'04, BK&Schöpfer&Schuster '08, Jin '12, Hohage&Werner '13,...] in Banach space

Gradient Methods

gradient steps for

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

$\rightsquigarrow$ Landweber iteration (steepest descent, mimimal error)

$$q^{k+1} = q^k - \mu^k F'(q^k)^*(F(q^k) - y^\delta)$$

Gradient Methods

gradient steps for

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

$\rightsquigarrow$ Landweber iteration (steepest descent, mimimal error)

$$q^{k+1} = q^k - \mu^k F'(q^k)^*(F(q^k) - y^\delta)$$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equiv. to

$$q^{k+1} = q^k - \mu^k F'(q^k)^* \left(F(q^k) - y^\delta\right)$$

Gradient Methods

gradient steps for

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

$\rightsquigarrow$ Landweber iteration (steepest descent, mimimal error)

$$q^{k+1} = q^k - \mu^k F'(q^k)^*(F(q^k) - y^\delta)$$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equiv. to

$$\begin{aligned} q^{k+1} &= q^k - \mu^k F'(q^k)^* \left(F(q^k) - y^\delta\right) \\ &= q^k - \mu^k (C'(S(q^k))S'(q^k))^* \left(C(S(q^k)) - y^\delta\right) \end{aligned}$$

Gradient Methods

gradient steps for

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

$\rightsquigarrow$ Landweber iteration (steepest descent, mimimal error)

$$q^{k+1} = q^k - \mu^k F'(q^k)^*(F(q^k) - y^\delta)$$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equiv. to

$$\begin{aligned} q^{k+1} &= q^k - \mu^k F'(q^k)^* \left(F(q^k) - y^\delta\right) \\ &= q^k - \mu^k (C'(S(q^k))S'(q^k))^* \left(C(S(q^k)) - y^\delta\right) \\ &= q^k + \mu^k A'_q(q^k, \tilde{u})^* p \end{aligned}$$

where

$$\begin{cases} A(q^k, \tilde{u}) = 0 \\ A'_u(q^k, \tilde{u})^* p = -C'(\tilde{u})^*(C(\tilde{u}) - y^\delta) \end{cases}$$

Gradient Methods

gradient steps for

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

$\rightsquigarrow$ Landweber iteration (steepest descent, mimimal error)

$$q^{k+1} = q^k - \mu^k F'(q^k)^*(F(q^k) - y^\delta)$$

with $F = C \circ S$, S parameter-to-state map, $A(q, S(q)) = 0$, equiv. to

$$\begin{aligned} q^{k+1} &= q^k - \mu^k F'(q^k)^* \Big(F(q^k) - y^\delta\Big) \\ &= q^k - \mu^k (C'(S(q^k))S'(q^k))^* \Big(C(S(q^k)) - y^\delta\Big) \\ &= q^k + \mu^k A'_q(q^k, \tilde{u})^* p \end{aligned}$$

where

$$\begin{cases} A(q^k, \tilde{u}) = 0 \\ A'_u(q^k, \tilde{u})^* p = -C'(\tilde{u})^*(C(\tilde{u}) - y^\delta) \end{cases}$$

[Hanke&Neubauer&Scherzer '95,...] in Hilbert space
[BK&Schöpfer&Schuster '08,...] in Banach space

all-at-once formulation

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q,u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

(b) all-at once approach: observations and model as system for (q,u)

$$\begin{array}{rcl} A(q,u) & = & 0 \text{ in } W^* \\ C(u) & = & y \text{ in } Y \end{array} \quad \Leftrightarrow \quad \mathbf{F}(q,u) = \mathbf{y}$$

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q,u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

(b) all-at once approach: observations and model as system for (q,u)

$$\begin{array}{rcl} A(q,u) & = & 0 \text{ in } W^* \\ C(u) & = & y \text{ in } Y \end{array} \quad \Leftrightarrow \quad \mathbf{F}(q,u) = \mathbf{y}$$

for other all-at-once type approaches see, e.g.,
[Kupfer & Sachs '92, Shenoy & Heinkenschloss & Cliff '98, Haber & Ascher '01, Burger & Mühlhuber '02,...]

Abstract Formulation

Identify parameter q in (PDE or ODE) model

$$A(q,u) = 0$$

from observations of the state u

$$C(u) = y\,,$$

(b) all-at once approach: observations and model as system for (q,u)

$$\begin{array}{rcl} A(q,u) & = & 0 \text{ in } W^* \\ C(u) & = & y \text{ in } Y \end{array} \quad \Leftrightarrow \quad \mathbf{F}(q,u) = \mathbf{y}$$

for other all-at-once type approaches see, e.g.,
[Kupfer & Sachs '92, Shenoy & Heinkenschloss & Cliff '98, Haber & Ascher '01, Burger & Mühlhuber '02,...]

ill-posedness (**F** not continuously invertible)
only noisy measurements $y^\delta \approx y$ given
$\Rightarrow$ regularization needed

Tikhonov Regularization

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \|A(q,u)\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u)$$

Tikhonov Regularization

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \|A(q,u)\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u)$$

first order optimality condition:

$$A'_q(q,u)^* A(q,u) + \alpha\partial\mathcal{R}(q) = 0$$
$$C'(u)^*(C(u) - y^\delta) + A'_u(q,u)^* A(q,u) + \alpha\partial\tilde{\mathcal{R}}(u) = 0$$

Tikhonov Regularization

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \|A(q,u)\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u)$$

first order optimality condition:

$$\begin{aligned} A'_q(q,u)^* A(q,u) + \alpha\partial\mathcal{R}(q) &= 0 \\ C'(u)^*(C(u) - y^\delta) + A'_u(q,u)^* A(q,u) + \alpha\partial\tilde{\mathcal{R}}(u) &= 0 \end{aligned}$$

i.e., with $p = A(q,u)$:

$$\begin{cases} A(q,u) = p \quad \text{(state equation)} \\ A'_q(q,u)^* p + \alpha\partial\mathcal{R}(q) = 0 \quad \text{(gradient equation)} \\ A'_u(q,u)^* p = -C'(u)^*(C(u) - y^\delta) - \alpha\partial\tilde{\mathcal{R}}(u) \quad \text{(adjoint equation)} \end{cases}$$

Tikhonov Regularization

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \rho\|A(q,u)\| + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u)$$

Tikhonov Regularization

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \rho\|A(q,u)\| + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u)$$

i.e., (exact penalization) with ρ sufficiently large

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u) \text{ s.t. } A(q,u) = 0$$

i.e., reduced Tikhonov.

Tikhonov Regularization

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \rho\|A(q,u)\| + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u)$$

i.e., (exact penalization) with ρ sufficiently large

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u) \text{ s.t. } A(q,u) = 0$$

i.e., reduced Tikhonov.

Lagrange function

$$\mathcal{L}(q,u,p) = \|C(u) - y^\delta\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u) + \langle A(q,u), p\rangle$$

first order optimality condition:

$$\begin{cases} A(q,u) = 0 \quad \text{(state equation)} \\ A'_q(q,u)^* p + \alpha\partial\mathcal{R}(q) = 0 \quad \text{(gradient equation)} \\ A'_u(q,u)^* p = -C'(u)^*(C(u) - y^\delta) - \alpha\partial\tilde{\mathcal{R}}(u) \quad \text{(adjoint equation)} \end{cases}$$

Tikhonov Regularization

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \rho\|A(q,u)\| + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u)$$

i.e., (exact penalization) with ρ sufficiently large

$$\min_{q,u} \|C(u) - y^\delta\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u) \text{ s.t. } A(q,u) = 0$$

i.e., reduced Tikhonov.

Lagrange function

$$\mathcal{L}(q,u,p) = \|C(u) - y^\delta\|^2 + \alpha\mathcal{R}(q) + \alpha\tilde{\mathcal{R}}(u) + \langle A(q,u), p\rangle$$

first order optimality condition:

$$\begin{cases} A(q,u) = 0 \quad \text{(state equation)} \\ A_q'(q,u)^* p + \alpha\partial\mathcal{R}(q) = 0 \quad \text{(gradient equation)} \\ A_u'(q,u)^* p = -C'(u)^*(C(u) - y^\delta) - \alpha\partial\tilde{\mathcal{R}}(u) \quad \text{(adjoint equation)} \end{cases}$$

i.e., reduced and all-at-once Tikhonov regularization are basically the same.

Regularized Gauss-Newton Method

(q^k, u^k) fixed, one Gauss-Newton step:

$$\min_{q,u} \|C(u^k) + C'(u^k)(u - u^k) - y^\delta\|^2 + \alpha_k \mathcal{R}(q) + \alpha_k \tilde{\mathcal{R}}(u)$$
$$+ \|A(q^k, u^k) + A'_u(q^k, u^k)(u - u^k) + A'_q(q^k, u^k)(q - q^k)\|^2$$

$\rightsquigarrow (q^{k+1}, u^{k+1})$

Regularized Gauss-Newton Method

(q^k, u^k) fixed, one Gauss-Newton step:

$$\min_{q,u} \|C(u^k) + C'(u^k)(u - u^k) - y^\delta\|^2 + \alpha_k \mathcal{R}(q) + \alpha_k \tilde{\mathcal{R}}(u)$$
$$+ \|A(q^k, u^k) + A'_u(q^k, u^k)(u - u^k) + A'_q(q^k, u^k)(q - q^k)\|^2$$

$\rightsquigarrow (q^{k+1}, u^{k+1})$

first order optimality condition:
with $p = A(q^k, u^k) + A'_u(q^k, u^k)(u - u^k) + A'_q(q^k, u^k)(q - q^k)$:

$$\begin{cases} A'_u(q^k, u^k)(u - u^k) = -A(q^k, u^k) - A'_q(q^k, u^k)(q - q^k) + p & \text{(linear state equation)} \\ A'_q(q^k, u^k)^* p + \alpha \partial \mathcal{R}(q) = 0 & \text{(gradient equation)} \\ A'_u(q^k, u^k)^* p = -C'(u^k)^*(C(u^k) + C'(u_k)(u - u_k) - y^\delta) + \alpha \partial \tilde{\mathcal{R}}(u) & \text{(adjoint equation)} \end{cases}$$

Regularized Gauss-Newton Method

(q^k, u^k) fixed, one Gauss-Newton step:

$$\min_{q,u} \|C(u^k) + C'(u^k)(u - u^k) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q) + \alpha_k \tilde{\mathcal{R}}(u)$$
$$+ \rho \|A(q^k, u^k) + A'_u(q^k, u^k)(u - u^k) + A'_q(q^k, u^k)(q - q^k)\|$$

i.e. (exact penalization) with ρ sufficiently large

$$\min_{q,u} \|C(u^k) + C'(u^k)(u - u^k) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q) + \alpha_k \tilde{\mathcal{R}}(u)$$
$$\text{s.t. } A(q^k, u^k) + A'_u(q^k, u^k)(u - u^k) + A'_q(q^k, u^k)(q - q^k) = 0$$

Regularized Gauss-Newton Method

(q^k, u^k) fixed, one Gauss-Newton step:

$$\min_{q,u} \|C(u^k) + C'(u^k)(u - u^k) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q) + \alpha_k \tilde{\mathcal{R}}(u)$$
$$+ \rho \|A(q^k, u^k) + A'_u(q^k, u^k)(u - u^k) + A'_q(q^k, u^k)(q - q^k)\|$$

i.e. (exact penalization) with ρ sufficiently large

$$\min_{q,u} \|C(u^k) + C'(u^k)(u - u^k) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q) + \alpha_k \tilde{\mathcal{R}}(u)$$
$$\text{s.t. } A(q^k, u^k) + A'_u(q^k, u^k)(u - u^k) + A'_q(q^k, u^k)(q - q^k) = 0$$

first order optimality condition:

$$\begin{cases} A'_u(q^k, u^k)(u - u^k) = -A(q^k, u^k) - A'_q(q^k, u^k)(q - q^k) & \text{(linear state eq.)} \\ A'_q(q^k, u^k)^* p + \alpha \partial \mathcal{R}(q) = 0 & \text{(gradient equation)} \\ A'_u(q^k, u^k)^* p = -C'(u^k)^*(C(u^k) + C'(u_k)(u - u_k) - y^\delta) - \alpha \partial \tilde{\mathcal{R}}(u) & \text{(adj.eq.)} \end{cases}$$

Regularized Gauss-Newton Method

The latter is **not** reduced regularized Gauss-Newton!

Regularized Gauss-Newton Method

The latter is **not** reduced regularized Gauss-Newton!
So what would then reduced regularized Gauss-Newton mean?

Regularized Gauss-Newton Method (reduced)

q^k fixed, one reduced Gauss-Newton step:

$$\min_{q,u,\tilde{u}} \|C(\tilde{u}) + C'(\tilde{u})(u-\tilde{u}) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q)$$

$$\text{s.t. } A(q^k,\tilde{u}) + A'_u(q^k,\tilde{u})(u-\tilde{u}) + A'_q(q^k,\tilde{u})(q-q^k) = 0$$

$$\text{and } A(q^k,\tilde{u}) = 0$$

Regularized Gauss-Newton Method (reduced)

q^k fixed, one reduced Gauss-Newton step:

$$\min_{q,u,\tilde{u}} \|C(\tilde{u}) + C'(\tilde{u})(u-\tilde{u}) - y^\delta\|^2 + \alpha_k \mathcal{R}_k(q)$$

$$\text{s.t. } A(q^k,\tilde{u}) + A'_u(q^k,\tilde{u})(u-\tilde{u}) + A'_q(q^k,\tilde{u})(q-q^k) = 0$$

$$\text{and } A(q^k,\tilde{u}) = 0$$

first order optimality condition:

$$\begin{cases} A(q^k,\tilde{u}) = 0 & \text{(nonlinear decoupled state equation)} \\ A'_u(q^k,\tilde{u})(u-\tilde{u}) = -A(q^k,\tilde{u}) - A'_q(q^k,\tilde{u})(q-q^k) & \text{(linear state eq.)} \\ A'_q(q^k,\tilde{u})^* p + \alpha\partial\mathcal{R}(q) = 0 & \text{(gradient equation)} \\ A'_u(q^k,\tilde{u})^* p = -C'(\tilde{u})^*(C(\tilde{u}) + C'(\tilde{u})(u-\tilde{u}) - y^\delta) & \text{(adjoint equation)} \end{cases}$$

Comparison of optimality conditions for reduced and all-at-once Newton

reduced:

$$\begin{cases} A(q^k,\tilde{u})=0 & \text{(nonlinear decoupled state equation)} \\ A'_u(q^k,\tilde{u})(u-\tilde{u})=-A(q^k,\tilde{u})-A'_q(q^k,\tilde{u})(q-q^k) & \text{(linear state eq.)} \\ A'_q(q^k,\tilde{u})^*p+\alpha\partial\mathcal{R}(q)=0 & \text{(gradient equation)} \\ A'_u(q^k,\tilde{u})^*p=-C'(\tilde{u})^*(C(\tilde{u})+C'(\tilde{u})(u-\tilde{u})-y^\delta) & \text{(adjoint equation)} \end{cases}$$

all-at-once:

$$\begin{cases} A'_u(q^k,u^k)(u-u^k)=-A(q^k,u^k)-A'_q(q^k,u^k)(q-q^k) & \text{(linear state eq.)} \\ A'_q(q^k,u^k)^*p+\alpha\partial\mathcal{R}(q)=0 & \text{(gradient equation)} \\ A'_u(q^k,u^k)^*p=-C'(u^k)^*(C(u^k)+C'(u_k)(u-u_k)-y^\delta)-\alpha\partial\tilde{\mathcal{R}}(u) & \text{(adj.eq.)} \end{cases}$$

Gradient Methods (reduced)

q^k fixed, one Landweber step

$$
\begin{aligned}
q^{k+1} &= q^k - \mu^k F'(q^k)^* \left(F(q^k) - y^\delta\right) \\
&= q^k - \mu^k (C'(S(q^k))S'(q^k))^* \left(C(S(q^k)) - y^\delta\right) \\
&= q^k + \mu^k A_q'(q^k, \tilde{u})^* p
\end{aligned}
$$

where

$$
\begin{cases}
A(q^k, \tilde{u}) = 0 & \text{(nonlinear decoupled state equation)} \\
A_u'(q^k, \tilde{u})^* p = -C'(\tilde{u})^*(C(\tilde{u}) - y^\delta) & \text{(adjoint equation)}
\end{cases}
$$

Gradient Methods (all-at-once)

(q^k, u^k) fixed, one Landweber step for $\mathbf{F}\begin{pmatrix} q \\ u \end{pmatrix} = \begin{pmatrix} A(q,u) \\ C(u) \end{pmatrix}$:

$$\begin{pmatrix} q^{k+1} \\ u^{k+1} \end{pmatrix} = \begin{pmatrix} q^k \\ u^k \end{pmatrix} - \mu_k \mathbf{F}' \begin{pmatrix} q^k \\ u^k \end{pmatrix}^* \left(\mathbf{F} \begin{pmatrix} q^k \\ u^k \end{pmatrix} - \mathbf{y}^\delta \right)$$

$$= \begin{pmatrix} q^k \\ u^k \end{pmatrix} - \mu_k \begin{pmatrix} A'_q(q^k, u^k) & A'_u(q^k, u^k) \\ 0 & C'(u^k) \end{pmatrix}^* \begin{pmatrix} A(q^k, u^k) \\ C(u^k) - y^\delta \end{pmatrix}$$

Gradient Methods (all-at-once)

(q^k, u^k) fixed, one Landweber step for $\mathbf{F}\begin{pmatrix} q \\ u \end{pmatrix} = \begin{pmatrix} A(q,u) \\ C(u) \end{pmatrix}$:

$$\begin{pmatrix} q^{k+1} \\ u^{k+1} \end{pmatrix} = \begin{pmatrix} q^k \\ u^k \end{pmatrix} - \mu_k \mathbf{F}' \begin{pmatrix} q^k \\ u^k \end{pmatrix}^* \left(\mathbf{F} \begin{pmatrix} q^k \\ u^k \end{pmatrix} - \mathbf{y}^\delta \right)$$
$$= \begin{pmatrix} q^k \\ u^k \end{pmatrix} - \mu_k \begin{pmatrix} A'_q(q^k, u^k) & A'_u(q^k, u^k) \\ 0 & C'(u^k) \end{pmatrix}^* \begin{pmatrix} A(q^k, u^k) \\ C(u^k) - y^\delta \end{pmatrix}$$

i.e.

$$\begin{cases} q^{k+1} = A'_q(q^k, u^k)^* A(q^k, u^k) \\ u^{k+1} = C'(u^k)^*(C(u^k) - y^\delta) + A'_u(q^k, u^k)^* A(q^k, u^k) \end{cases}$$

completely explicit, no model to solve!

Convergence Analysis

- Existence of minimizers, stability, convergence, rates under (variational, approximate) source conditions follow as corollaries of existing results for Tikhonov, IRGNM, Landweber, when regularizing with respect to q and u
- Case of regularization $\alpha\mathcal{R}(q)$ of q only:
  Recover bounds on u via solvability condition $\|A_u(q,u)^{-1}\| \leq C_A$

Convergence Analysis

- Existence of minimizers, stability, convergence, rates under (variational, approximate) source conditions follow as corollaries of existing results for Tikhonov, IRGNM, Landweber, when regularizing with respect to q and u
- Case of regularization $\alpha\mathcal{R}(q)$ of q only:
  Recover bounds on u via solvability condition $\|A_u(q,u)^{-1}\| \leq C_A$
- Case of additional regularization $\beta\tilde{\mathcal{R}}(u)$ of u:
  solvability condition $\|A_u(q,u)^{-1}\| \leq C_A$ not needed!

Convergence Analysis

- Existence of minimizers, stability, convergence, rates under (variational, approximate) source conditions follow as corollaries of existing results for Tikhonov, IRGNM, Landweber, when regularizing with respect to q and u
- Case of regularization $\alpha\mathcal{R}(q)$ of q only:
  Recover bounds on u via solvability condition $\|A_u(q,u)^{-1}\| \leq C_A$
- Case of additional regularization $\beta\tilde{\mathcal{R}}(u)$ of u:
  solvability condition $\|A_u(q,u)^{-1}\| \leq C_A$ not needed!
- Getting rid of solvability condition allows to skip constraints on parameters (e.g. $a \geq \underline{a} > 0$ in a-problem $-\nabla(a\nabla u) = b$)!

numerical results

Numerical Tests

nonlinear inverse source problem:

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

Identify q from distributed measurements of u in Ω

Comparison of reduced and all-at-once Landweber

ζ	it_{aao}	it_{red}	cpu_{aao}	cpu_{red}	$\frac{\|b^\delta_{k_*(\delta),\text{aao}} - b^\dagger\|_X}{\|b^\dagger\|_X}$	$\frac{\|b^\delta_{k_*(\delta),\text{red}} - b^\dagger\|_X}{\|b^\dagger\|_X}$
0.5	5178	2697	2.97	18.07	0.0724	0.1047
5	$> 2 \cdot 10^6$	48510	1293.60	482.19	0.7837	0.1633
10	$> 2 \cdot 10^6$	$> 10^5$	1257.50	639.87	0.9621	0.1632
-0.5	10895	2016	8.85	14.55	0.1406	0.2295
-1	18954	-	11.42	-	0.2313	-

(1% Gaussian noise)

Comparison of reduced and all-at-once IRGNM

ζ	it_{aao}	it_{red}	cpu_{aao}	cpu_{red}	$\frac{\|b^\delta_{k_*(\delta),\text{aao}} - b^\dagger\|_X}{\|b^\dagger\|_X}$	$\frac{\|b^\delta_{k_*(\delta),\text{red}} - b^\dagger\|_X}{\|b^\dagger\|_X}$
0	34	32	0.14	0.10	0.0149	0.0151
10	43	43	0.20	0.55	0.0996	0.1505
100	55	56	0.28	0.82	0.0721	0.0770
1000	68	68	0.42	1.07	0.0543	0.0588
-0.5	33	32	0.13	0.35	0.1174	0.2165
-1.	35	-	0.23	-	0.2023	-
-10	44	-	0.23	-	0.0768	-
-100	77	-	0.59	-	0.2246	-
-1000	70	-	0.49	-	0.0321	-

(1% Gaussian noise)

Numerical Tests in 2-d with Adaptive Discretization

nonlinear inverse source problem:

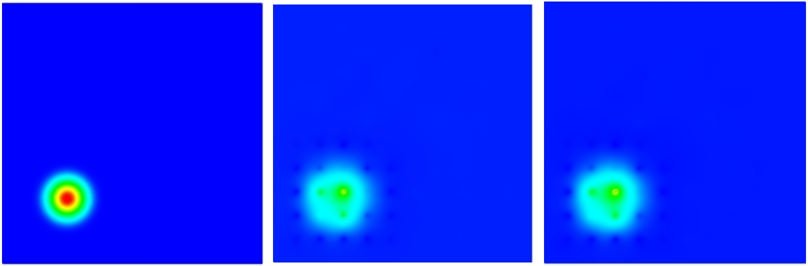
$$-\Delta u + \zeta u^3 = q \text{ in } \Omega = (0,1)^2 \quad \& \text{ homogeneous Dirichlet BC}$$

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

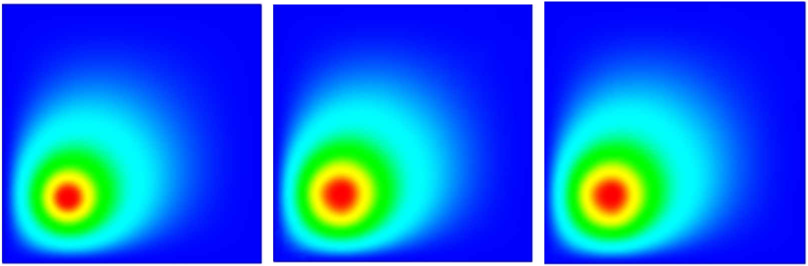
$$q^\dagger = \frac{c}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\left(\frac{sx-\mu}{\sigma}\right)^2 + \left(\frac{sy-\mu}{\sigma}\right)^2\right)\right)$$

with $c = 10$, $\mu = 0.5$, $\sigma = 0.1$, and $s = 2$.

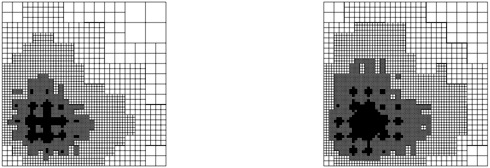
- goal-oriented, dual weighted residual estimators
- computations with *Gascoigne* and *RoDoBo*
- joint work with Alana Kirchner and Boris Vexler (TU Munich)



left: exact source $q^\dagger$,
middle: reconstruction by reduced Tikhonov (RT),
right: reconstruction by all-at-once Gauss-Newton (AGN),
with $\zeta = 100$, 1% noise



left: exact state $u^\dagger$,
middle: reconstruction by reduced Tikhonov (RT),
right: reconstruction by all-at-once Gauss-Newton (AGN),
with $\zeta = 100$, 1% noise



adaptively refined meshes,
left: by reduced Tikhonov (RT),
right: by all-at-once Gauss-Newton (AGN),
with $\zeta = 100$, 1% noise

Table: all-at-once Gauss-Newton (AGN) versus reduced Tikhonov (RT) for different choices of ζ with 1% noise. .
ctr: Computation time reduction using (AGN) in comparison to (RT)

ζ	RT			AGN			ctr
	error	β	# nodes	error	β	# nodes	
1	0.418	2985	2499	0.412	4600	3873	-65%
10	0.417	3194	2473	0.411	4918	3965	-59%
100	0.408	5014	6653	0.417	6773	9813	39%
500	0.418	9421	11851	0.404	13756	821	97%
1000	0.439	11486	44391	0.426	16355	793	99%

Conclusions and Outlook

- Tikhonov:
  reduced $\sim$ all-at-once
- Newton:
  reduced: solve nonlinear and linear models in each step
  all-at-once: only solve linearized models
- Landweber:
  reduced: solve nonlinear and linear models in each step
  all-at-once: never solve models!

Conclusions and Outlook

- Tikhonov:
  reduced $\sim$ all-at-once
- Newton:
  reduced: solve nonlinear and linear models in each step
  all-at-once: only solve linearized models
- Landweber:
  reduced: solve nonlinear and linear models in each step
  all-at-once: never solve models!

$\rightarrow$ time dependent problems

$\rightarrow$ regularization parameter choice

$\rightarrow$ restrictions on nonlinearity of F / $\mathbf{F}$

$\rightarrow$ convergence rates under source conditions

$\rightarrow$ minimization based inverse problems formulations and regularizations

minimization based formulations

Minimization based formulations

$$F(q) = y \quad \text{i.e.,} \quad \begin{cases} A(x,u) = 0 \\ C(u) = y \end{cases}$$

Minimization based formulations

$$F(q) = y \quad \text{i.e.,} \quad \begin{cases} A(x,u) = 0 \\ C(u) = y \end{cases}$$

is equivalent to

$$\min_q \|F(q) - y\|$$

Minimization based formulations

$$F(q) = y \quad \text{i.e.,} \quad \begin{cases} A(x, u) = 0 \\ C(u) = y \end{cases}$$

is equivalent to

$$\min_q \|F(q) - y\|$$

or equivalent to

$$\min_{q,u} \|C(u) - y\| \text{ s.t. } A(q, u) = 0$$

Minimization based formulations

$$F(q) = y \quad \text{i.e.,} \quad \begin{cases} A(x,u) = 0 \\ C(u) = y \end{cases}$$

is equivalent to

$$\min_q \|F(q) - y\|$$

or equivalent to

$$\min_{q,u} \|C(u) - y\| \text{ s.t. } A(q,u) = 0$$

or equivalent to

$$\min_{q,u} \|A(q,u)\| \text{ s.t. } C(u) = y$$

Minimization based formulations

$$F(q) = y \quad \text{i.e.,} \quad \begin{cases} A(x,u) = 0 \\ C(u) = y \end{cases}$$

is equivalent to

$$\min_q \|F(q) - y\|$$

or equivalent to

$$\min_{q,u} \|C(u) - y\| \text{ s.t. } A(q,u) = 0$$

or equivalent to

$$\min_{q,u} \|A(q,u)\| \text{ s.t. } C(u) = y$$

. . . and beyond, e.g., variation formulation of EIT, [Kohn&Vogelius'89]

Minimization based formulations

generally: formulate inverse problem as

$$\min_{q,u} J(q,u;y) \text{ s.t. } (q,u) \in M_{\mathrm{ad}}(y)$$

and regularize it by solving

$$\min_{q,u} J(q,u;y) + \alpha \mathcal{R}(q,u) \text{ s.t. } (q,u) \in M_{\mathrm{ad}}^{\delta}(y^{\delta})$$

where, e.g., $M_{\mathrm{ad}}^{\delta}(y^{\delta}) \subseteq \{(q,u) \,:\, \tilde{\mathcal{R}}(q,u) \leq \varrho\}$

Minimization based formulations

generally: formulate inverse problem as

$$\min_{q,u} J(q,u;y) \text{ s.t. } (q,u) \in M_{\mathrm{ad}}(y)$$

and regularize it by solving

$$\min_{q,u} J(q,u;y) + \alpha \mathcal{R}(q,u) \text{ s.t. } (q,u) \in M_{\mathrm{ad}}^{\delta}(y^{\delta})$$

where, e.g., $M_{\mathrm{ad}}^{\delta}(y^{\delta}) \subseteq \{(q,u) \,:\, \tilde{\mathcal{R}}(q,u) \leq \varrho\}$

[Kindermann '17] (reduced case),
[BK '17] (avoid parameter-to-state map)

C. Clason and B. Kaltenbacher.

On the use of state constraints in optimal control of singular PDEs. *System & Control Letters*, 62:48–54, 2013.

B. Kaltenbacher.

Regularization based on all-at-once formulations for inverse problems. *SIAM Journal on Numerical Analysis*, 54:2594–2618, 2016. arXiv:1603.05332v1 [math.NA].

B. Kaltenbacher.

All-at-once versus reduced iterative methods for time dependent inverse problems. *Inverse Problems*, 33:064002, 2017.

B. Kaltenbacher.

Minimization based formulations of inverse problems and their regularization. submitted.

B. Kaltenbacher, A. Kirchner, and B. Vexler.

Goal oriented adaptivity in the IRGNM for parameter identification in PDEs II: all-at once formulations. *Inverse Problems*, 30, 2014. submitted.

Thank you for your attention!

IRGNM

$$\min_{q,u} \|L(q-q_k)+K(u-u_k)+A(q_k,u_k)-f\|^2+\|Cu-g^\delta\|^2+\alpha_k\|q-q_0\|^2$$

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

IRGNM

$$\min_{q,u} \|L(q-q_k)+K(u-u_k)+A(q_k,u_k)-f\|^2+\|Cu-g^\delta\|^2+\alpha_k\|q-q_0\|^2$$

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

First order optimality system:

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