Reduced, all-at-once, and variational formulations of inverse problems and their solution

Barbara Kaltenbacher

Alpen-Adria-Universität Klagenfurt

CMAM, Vienna, September 1, 2022

joint work with

Kha Van Huynh, AAU

FWF Der Wissenschaftsfonds.

Modeling - Analysis - Optimization
DOC.FUNDS DOCTORAL SCHOOL

UNIVERSITÄT KLAGENFURT

Outline

- examples
- minimization based formulation
  and regularization of inverse problems
- iterative solution methods
  - gradient type
  - Newton type

Outline

- examples
- minimization based formulation
  and regularization of inverse problems
- iterative solution methods
  - gradient type
  - Newton type

Der Wissenschaftsfonds.
FWF project P30054
Solving Inverse Problems without Forward Operators

and

Modeling - Analysis - Optimization
DOC.FUNDS DOCTORAL SCHOOL

FWF doc.funds DOC 78

examples

Parameter Identification in Differential Equations: Some Examples

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

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

  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 $\Omega \subseteq \mathbb{R}^d$, $d \in \{1, 2, 3\}$

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

  from boundary or (restricted) interior observations of u.
- e.g. EIT: identify conductivity σ in

$$-\nabla(\sigma\nabla u_i) = 0 \text{ in } \Omega$$

  from boundary observations
  current $g_i = -\sigma\frac{\partial u_i}{\partial n}$ and voltage $y_i = \phi_i$ on $\partial\Omega \quad i \in \{1, \ldots I\}$

Parameter Identification in Differential Equations: Some Examples

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

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

  from boundary or (restricted) interior observations of u.
- e.g. EIT: identify conductivity σ in

$$-\nabla(\sigma\nabla u_i) = 0 \text{ in } \Omega$$

  from boundary observations
  current $g_i = -\sigma\frac{\partial u_i}{\partial n}$ and voltage $y_i = \phi_i$ on $\partial\Omega \quad i \in \{1, \dots I\}$
- 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.
  $y_i = h_i(u(t_i))$, $i \in \{1, \dots, m\}$ or $y(t) = h(t, y(t))$, $t \in (0, T)$

Abstract Formulation as Operator Equation

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 as Operator Equation

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

- 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 as Operator Equation

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

- 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

- all-at once approach: observations - model 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 as Operator Equation

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

- 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

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

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

- minimization based approach ...

Parameter Identification: Reduced formulation

- Application of classical regularization methods requires repeated evaluation of the parameter-to-state map S

Parameter Identification: Reduced formulation

- Application of classical regularization methods requires repeated evaluation of the parameter-to-state map S
- often numerically too expensive

Parameter Identification: Reduced formulation

- Application of classical regularization methods requires repeated evaluation of the parameter-to-state map S
- often numerically too expensive
- sometimes requires too restrictive assumptions (e.g. for singular PDEs, e.g., static MEMS model $-\Delta u + \frac{q(x)}{(1+u)^2} = 0$)

Parameter Identification: Reduced formulation

- Application of classical regularization methods requires repeated evaluation of the parameter-to-state map S
- often numerically too expensive
- sometimes requires too restrictive assumptions (e.g. for singular PDEs, e.g., static MEMS model $-\Delta u + \frac{q(x)}{(1+u)^2} = 0$)
- destroys time causality for parameter identification in time-dependent PDEs

Parameter Identification: Reduced formulation

- Application of classical regularization methods requires repeated evaluation of the parameter-to-state map S
- often numerically too expensive
- sometimes requires too restrictive assumptions (e.g. for singular PDEs, e.g., static MEMS model $-\Delta u + \frac{q(x)}{(1+u)^2} = 0$)
- destroys time causality for parameter identification in time-dependent PDEs

$\rightsquigarrow$ derive formulations of inverse problems and of their regularization that do not require parameter-to-state map

minimization based formulation of inverse problems

Abstract Formulation as Operator Equation

Identify parameter q in (PDE or ODE) model

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

from observations y 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

- 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

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

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

- minimization based approach ...

Abstract Formulation as Minimization Problem: Basic Idea

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

Abstract Formulation as Minimization Problem: Basic Idea

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

is equivalent to

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

Abstract Formulation as Minimization Problem: Basic Idea

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

is equivalent to

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

or equivalent to

$$\min_{q,u} \|C(u) - y\|^2 + \|A(q,u)\|^2$$

Abstract Formulation as Minimization Problem: Basic Idea

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

is equivalent to

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

or equivalent to

$$\min_{q,u} \|C(u) - y\|^2 + \|A(q,u)\|^2$$

or equivalent to

$$\min_{q,u} \|C(u) - y\|^2 \text{ s.t. } A(q,u) = 0$$

Abstract Formulation as Minimization Problem: Basic Idea

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

is equivalent to

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

or equivalent to

$$\min_{q,u} \|C(u) - y\|^2 + \|A(q,u)\|^2$$

or equivalent to

$$\min_{q,u} \|C(u) - y\|^2 \text{ s.t. } A(q,u) = 0$$

or equivalent to

$$\min_{q,u} \|A(q,u)\|^2 \text{ s.t. } C(u) = y$$

Abstract Formulation as Minimization Problem: Basic Idea

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

is equivalent to

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

or equivalent to

$$\min_{q,u} \|C(u) - y\|^2 + \|A(q,u)\|^2$$

or equivalent to

$$\min_{q,u} \|C(u) - y\|^2 \text{ s.t. } A(q,u) = 0$$

or equivalent to

$$\min_{q,u} \|A(q,u)\|^2 \text{ s.t. } C(u) = y$$

. . . and beyond, e.g., variational formulation of EIT [Kohn&Vogelius'87]

The variational approach to EIT

see, e.g., [Kohn&Vogelius'87, Kohn&McKenny'90, Knowles'98]

The variational approach to EIT

see, e.g., [Kohn&Vogelius'87, Kohn&McKenny'90, Knowles'98]

Identify spatially distributed conductivity $\sigma = \sigma(x)$ in $\Omega \subseteq \mathbb{R}^2$

$$\nabla \cdot \mathbf{J}_i = 0\,, \quad \nabla^\perp \cdot \mathbf{E}_i = 0\,, \quad \mathbf{J}_i = \sigma \mathbf{E}_i \quad \text{in } \Omega\,, \quad i = 1, \ldots, I,$$

from observations y of boundary currents j_i and voltages v_i.

with $\nabla^\perp \psi = (-\frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_1})^T$ so that $\nabla^\perp \cdot = \text{curl}$

The variational approach to EIT

see, e.g., [Kohn&Vogelius'87, Kohn&McKenny'90, Knowles'98]

Identify spatially distributed conductivity $\sigma = \sigma(x)$ in $\Omega \subseteq \mathbb{R}^2$

$$\nabla \cdot \mathbf{J}_i = 0\,, \quad \nabla^\perp \cdot \mathbf{E}_i = 0\,, \quad \mathbf{J}_i = \sigma \mathbf{E}_i \quad \text{in } \Omega\,, \quad i = 1, \ldots, I,$$

from observations y of boundary currents j_i and voltages v_i.

with $\nabla^\perp \psi = (-\frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_1})^T$ so that $\nabla^\perp \cdot = \text{curl}$
Using potentials ϕ_i and ψ_i
for current densities $\mathbf{J}_i$ and electric fields $\mathbf{E}_i$

$$\mathbf{J}_i = -\nabla^\perp \psi_i\,, \quad \mathbf{E}_i = -\nabla \phi_i\,, \quad i = 1, \ldots, I\,,$$

we can rewrite the problem as

$$\sqrt{\sigma} \nabla \phi_i = \frac{1}{\sqrt{\sigma}} \nabla^\perp \psi_i \ \text{ in } \Omega\,;\ \psi_i = \gamma_i\,,\ \phi_i = v_i \ \text{ on } \partial\Omega\,,\ i = 1, \ldots, I\,,$$

where $\gamma_i(x(s)) = -\int_0^s j_i(x(r))\, dr$ for $\partial\Omega = \{x(s)\ :\ s \in (0, \text{length}(\partial\Omega))\}$.

The variational approach to EIT

$$\sqrt{\sigma}\nabla\phi_i = \frac{1}{\sqrt{\sigma}}\nabla^{\perp}\psi_i \text{ in } \Omega\,;\ \psi_i = \gamma_i\,,\ \phi_i = \upsilon_i \text{ on } \partial\Omega\,,\ i = 1,\ldots,I\,,$$

The variational approach to EIT

$$\sqrt{\sigma}\nabla\phi_i = \frac{1}{\sqrt{\sigma}}\nabla^{\perp}\psi_i \text{ in } \Omega\,;\ \psi_i = \gamma_i\,,\ \phi_i = \upsilon_i \text{ on } \partial\Omega\,,\ i = 1,\ldots,I\,,$$

equivalent to $\qquad (\Phi = (\phi_1,\ldots,\phi_I),\ \Psi = (\psi_1,\ldots,\psi_I)\ )$

$$\min_{\sigma,\Phi,\Psi} \sum_{i=1}^{I} \tfrac{1}{2}\int_{\Omega} |\sqrt{\sigma}\nabla\phi_i - \frac{1}{\sqrt{\sigma}}\nabla^{\perp}\psi_i|^2\, dx$$

$$\text{s.t. } \psi_i = \gamma_i\,,\ \phi_i = \upsilon_i \text{ on } \partial\Omega\,, \quad i = 1,\ldots,I$$

The variational approach to EIT

$$\sqrt{\sigma}\nabla\phi_i = \frac{1}{\sqrt{\sigma}}\nabla^\perp\psi_i \text{ in } \Omega\,;\ \psi_i = \gamma_i\,,\ \phi_i = \upsilon_i \text{ on } \partial\Omega\,,\ i = 1,\ldots,I\,,$$

equivalent to $\qquad (\Phi = (\phi_1,\ldots,\phi_I),\ \Psi = (\psi_1,\ldots,\psi_I)\ )$

$$\min_{\sigma,\Phi,\Psi} \sum_{i=1}^{I} \tfrac{1}{2}\int_\Omega |\sqrt{\sigma}\nabla\phi_i - \frac{1}{\sqrt{\sigma}}\nabla^\perp\psi_i|^2\,dx$$

$$\text{s.t. } \psi_i = \gamma_i\,,\ \phi_i = \upsilon_i \text{ on } \partial\Omega\,, \quad i = 1,\ldots,I$$

equivalent to (since $\int_\Omega \nabla\phi_i\cdot\nabla^\perp\psi_i\,dx = \int_{\partial\Omega}\upsilon_i j_i\,dx$)

$$\min_{\sigma,\Phi,\Psi} \sum_{i=1}^{I} \tfrac{1}{2}\int_\Omega \left(\sigma|\nabla\phi_i|^2 + \frac{1}{\sigma}|\nabla^\perp\psi_i|^2\right)dx$$

$$\text{s.t. } \psi_i = \gamma_i\,,\ \phi_i = \upsilon_i \text{ on } \partial\Omega\,, \quad i = 1,\ldots,I$$

Regularized variational EIT

inverse problem (EIT):

$$\min_{\sigma,\Phi,\Psi} \sum_{i=1}^{I} \tfrac{1}{2} \int_\Omega \left(\sigma |\nabla \phi_i|^2 + \frac{1}{\sigma} |\nabla^\perp \psi_i|^2 \right) dx$$

$$\text{s.t. } \psi_i = \gamma_i \,, \ \phi_i = \upsilon_i \ \text{ on } \partial\Omega \,, \quad i = 1, \ldots, I$$

Regularized variational EIT

inverse problem (EIT):

$$\min_{\sigma,\Phi,\Psi} \sum_{i=1}^{I} \tfrac{1}{2} \int_\Omega \Big(\sigma |\nabla \phi_i|^2 + \frac{1}{\sigma} |\nabla^\perp \psi_i|^2 \Big)\, dx$$

$$\text{s.t. } \psi_i = \gamma_i\,,\ \phi_i = \upsilon_i \ \text{ on } \partial\Omega\,, \quad i = 1, \dots, I$$

regularization (RegEIT):

$$\min_{\sigma,\Phi,\Psi} \sum_{i=1}^{I} \left\{ \tfrac{1}{2} \int_\Omega \sum_{i=1}^{I} \tfrac{1}{2} \int_\Omega \Big(\sigma |\nabla \phi_i|^2 + \frac{1}{\sigma} |\nabla^\perp \psi_i|^2 \Big)\, dx + \tfrac{\alpha}{2} \|(\phi_i, \psi)\|^2_{H^{1+\epsilon}(\Omega)^2} \right\}$$

$$\text{s.t.} \quad \underline{\sigma} \le \sigma \le \overline{\sigma} \text{ on } \Omega\,,$$

$$\left.\begin{array}{l} \upsilon_i^\delta - \tau\delta \le \phi_i \le \upsilon_i^\delta + \tau\delta\,, \\ \gamma_i^\delta - \tau\delta \le \psi_i \le \gamma_i^\delta + \tau\delta\,, \end{array}\right\} \quad \text{on } \partial\Omega\,, \quad i = 1, \dots, I\,.$$

with the noise level $\delta \ge \|y - y^\delta\|$ and a safetly factor $\tau > 1$

Convergence as $\delta \to 0$ [BK, SIOPT 2018]

Remarks on EIT example

- cost function: J^δ differentiable;
- constraints: pointwise bounds can be efficiently implemented [Hungerländer, BK and Rendl 2020] and are practically relevant in view of known a prior bounds on σ;
- first order least squares formulation of the PDE model;
- Euler-Lagrange equation for unregularized problem yields second order PDE model $\nabla \cdot (\sigma \nabla \phi_i) = 0$;
- can be extended to complete electrode model CEM [Somersalo, Cheney, and Isaacson, 1992], see [Huynh and BK, IPI 2021];

Some further examples

- crack detection: $\Sigma \subseteq \overline{\Omega}$

$$(\Sigma, \phi_1, \dots \phi_I, \psi_1, \dots \psi_I) \in$$
$$\text{argmin}\{\sum_{i=1}^{I} \int_{\Omega \setminus \Sigma} \left(\frac{\sigma}{2} |\nabla \phi_i|^2 + \frac{1}{2\sigma} |\nabla^\perp \psi_i|^2 \right) dx \; : \; \phi_i|_{\partial\Omega} = v_i \,, \; \psi_i|_{\partial\Omega} = \gamma_i \} \,.$$

Some further examples

- crack detection: $\Sigma \subseteq \overline{\Omega}$

$$(\Sigma, \phi_1, \dots \phi_I, \psi_1, \dots \psi_I) \in$$
$$\operatorname{argmin}\{\sum_{i=1}^{I} \int_{\Omega\setminus\Sigma} \Big(\frac{\sigma}{2}|\nabla\phi_i|^2 + \frac{1}{2\sigma}|\nabla^\perp\psi_i|^2\Big)\, dx \ : \ \phi_i|_{\partial\Omega} = \upsilon_i\,,\ \psi_i|_{\partial\Omega} = \gamma_i\}\,.$$

- magnetostatics: $\mu = \mu(x)$, $B_i = \nabla \times A_i\,,\quad H_i = \nabla\phi_i + A_i^{J^{\text{imp}}}$,

$$(\mu, A_1, \dots, A_I, \phi_1, \dots \phi_I) \in$$
$$\operatorname{argmin}\Big\{\sum_{i=1}^{I} \int_{\Omega} \Big(\frac{\mu}{2}|\nabla\phi_i + A_i^{J^{\text{imp}}}|^2 + \frac{1}{2\mu}|\nabla \times A_i|^2 - J_i^{\text{imp}} \cdot A_i\Big)\, dx \ : \ \phi_i|_{\partial\Omega} = 0\,,\ n \times A_i|_{\partial\Omega} = 0\,,\ \oint_{\partial\Omega_c} A_i \cdot ds = y_i\Big\}\,.$$

Some further examples

- crack detection: $\Sigma \subseteq \overline{\Omega}$

$$(\Sigma, \phi_1, \dots \phi_I, \psi_1, \dots \psi_I) \in$$
$$\operatorname{argmin}\{\sum_{i=1}^{I} \int_{\Omega\setminus\Sigma} \Big(\frac{\sigma}{2}|\nabla\phi_i|^2 + \frac{1}{2\sigma}|\nabla^\perp\psi_i|^2\Big)\, dx \; : \; \phi_i|_{\partial\Omega} = \upsilon_i\,,\; \psi_i|_{\partial\Omega} = \gamma_i\}\,.$$

- magnetostatics: $\mu = \mu(x)$, $B_i = \nabla \times A_i\,,\quad H_i = \nabla\phi_i + A_i^{J^{\text{imp}}}$,

$$(\mu, A_1, \dots, A_I, \phi_1, \dots \phi_I) \in$$
$$\operatorname{argmin}\Big\{\sum_{i=1}^{I} \int_{\Omega} \Big(\frac{\mu}{2}|\nabla\phi_i + A_i^{J^{\text{imp}}}|^2 + \frac{1}{2\mu}|\nabla \times A_i|^2 - J_i^{\text{imp}} \cdot A_i\Big)\, dx \; : \; \phi_i|_{\partial\Omega} = 0\,,\; n \times A_i|_{\partial\Omega} = 0\,,\; \oint_{\partial\Omega_c} A_i \cdot ds = y_i\Big\}\,.$$

- acoustic sound source localization

Some further examples

- crack detection: $\Sigma \subseteq \overline{\Omega}$

$$(\Sigma, \phi_1, \dots \phi_I, \psi_1, \dots \psi_I) \in$$
$$\operatorname{argmin}\{\sum_{i=1}^{I} \int_{\Omega\setminus\Sigma} \Big(\frac{\sigma}{2}|\nabla\phi_i|^2 + \frac{1}{2\sigma}|\nabla^\perp\psi_i|^2\Big)\, dx \; : \; \phi_i|_{\partial\Omega} = \upsilon_i\,,\; \psi_i|_{\partial\Omega} = \gamma_i\}\,.$$

- magnetostatics: $\mu = \mu(x)$, $B_i = \nabla \times A_i\,,\quad H_i = \nabla\phi_i + A_i^{J^{\text{imp}}}$,

$$(\mu, A_1, \dots, A_I, \phi_1, \dots \phi_I) \in$$
$$\operatorname{argmin}\Big\{\sum_{i=1}^{I} \int_{\Omega} \Big(\frac{\mu}{2}|\nabla\phi_i + A_i^{J^{\text{imp}}}|^2 + \frac{1}{2\mu}|\nabla \times A_i|^2 - J_i^{\text{imp}} \cdot A_i\Big)\, dx \; : \; \phi_i|_{\partial\Omega} = 0\,,\; n \times A_i|_{\partial\Omega} = 0\,,\; \oint_{\partial\Omega_c} A_i \cdot ds = y_i\Big\}\,.$$

- acoustic sound source localization
- spatially varying Lamé parameters in elastostatics

Some further examples

- crack detection: $\Sigma \subseteq \overline{\Omega}$

$$(\Sigma, \phi_1, \dots \phi_I, \psi_1, \dots \psi_I) \in$$
$$\operatorname{argmin}\{\sum_{i=1}^{I} \int_{\Omega\setminus\Sigma} \Big(\frac{\sigma}{2}|\nabla\phi_i|^2 + \frac{1}{2\sigma}|\nabla^\perp\psi_i|^2\Big)\, dx \;: \; \phi_i|_{\partial\Omega} = \upsilon_i\,,\; \psi_i|_{\partial\Omega} = \gamma_i\}\,.$$

- magnetostatics: $\mu = \mu(x)$, $B_i = \nabla \times A_i\,,\quad H_i = \nabla\phi_i + A_i^{J^{\text{imp}}}$,

$$(\mu, A_1, \dots, A_I, \phi_1, \dots \phi_I) \in$$
$$\operatorname{argmin}\Big\{\sum_{i=1}^{I} \int_{\Omega} \Big(\frac{\mu}{2}|\nabla\phi_i + A_i^{J^{\text{imp}}}|^2 + \frac{1}{2\mu}|\nabla \times A_i|^2 - J_i^{\text{imp}} \cdot A_i\Big)\, dx \;: \; \phi_i|_{\partial\Omega} = 0\,,\; n \times A_i|_{\partial\Omega} = 0\,,\; \oint_{\partial\Omega_c} A_i \cdot ds = y_i\Big\}\,.$$

- acoustic sound source localization
- spatially varying Lamé parameters in elastostatics
- ...

Inverse problem formulated via constrained minimization

$$\min J(x) \text{ s.t. } x \in M \tag{1}$$

with solution $x^\dagger$

Inverse problem formulated via constrained minimization

$$\min J(x) \text{ s.t. } x \in M \tag{1}$$

with solution $x^\dagger$

first order optimality condition for a minimizer of (1)

$$\langle \nabla J(x^\dagger), x - x^\dagger \rangle \geq 0 \quad \text{for all } x \in M \tag{2}$$

Inverse problem formulated via constrained minimization

$$\min J(x) \text{ s.t. } x \in M \tag{1}$$

with solution $x^\dagger$

first order optimality condition for a minimizer of (1)

$$\langle \nabla J(x^\dagger), x - x^\dagger \rangle \geq 0 \quad \text{for all } x \in M \tag{2}$$

normalization assumption

$$J \geq 0 \text{ on } M\,, \quad x^\dagger \in M\,, \quad J(x^\dagger) = \min_{x \in M} J(x) = 0 \tag{3}$$

Inverse problem formulated via constrained minimization

$$\min J(x) \text{ s.t. } x \in M \tag{1}$$

with solution $x^\dagger$
first order optimality condition for a minimizer of (1)

$$\langle \nabla J(x^\dagger), x - x^\dagger \rangle \geq 0 \quad \text{for all } x \in M \tag{2}$$

normalization assumption

$$J \geq 0 \text{ on } M\,, \quad x^\dagger \in M\,, \quad J(x^\dagger) = \min_{x \in M} J(x) = 0 \tag{3}$$

noisy data $y^\delta \sim y \rightsquigarrow$

$$\min J^\delta(x) \text{ s.t. } x \in \tilde{M}^\delta$$

$\rightsquigarrow$ replace (3) by

$$J^\delta \geq 0 \text{ on } \tilde{M}^\delta\,, \quad x^\dagger \in \tilde{M}^\delta\,, \quad J^\delta(x^\dagger) \leq \eta(\delta) \quad \text{for all } \delta \in (0, \bar{\delta})\,,$$

where $\eta(\delta) > 0$ and $\eta(\delta) \to 0$ as $\delta \to 0$

Inverse problem formulated via constrained minimization

$$\min J(x) \text{ s.t. } x \in M \tag{1}$$

with solution $x^\dagger$

first order optimality condition for a minimizer of (1)

$$\langle \nabla J(x^\dagger), x - x^\dagger \rangle \geq 0 \quad \text{for all } x \in M \tag{2}$$

normalization assumption

$$J \geq 0 \text{ on } M\,, \quad x^\dagger \in M\,, \quad J(x^\dagger) = \min_{x \in M} J(x) = 0 \tag{3}$$

noisy data $y^\delta \sim y \rightsquigarrow$

$$\min J^\delta(x) \text{ s.t. } x \in \tilde{M}^\delta$$

$\rightsquigarrow$ replace (3) by

$$J^\delta \geq 0 \text{ on } \tilde{M}^\delta\,, \quad x^\dagger \in \tilde{M}^\delta\,, \quad J^\delta(x^\dagger) \leq \eta(\delta) \quad \text{for all } \delta \in (0, \bar{\delta})\,,$$

where $\eta(\delta) > 0$ and $\eta(\delta) \to 0$ as $\delta \to 0$

study convergence as δ tends to zero.

iterative solution methods

define steps by

- gradient descent
- regularized Newton type method

define steps by

- gradient descent
- regularized Newton type method

take into account constraints by

- projection onto M
- minimization over M

define steps by

- gradient descent
- regularized Newton type method

take into account constraints by

- projection onto M
- minimization over M

consider the following combinations:

- projected gradient method (in Hilbert spaces)
- regularized sequential quadratic programming SQP (in general Banach spaces)

define steps by

- gradient descent
- regularized Newton type method

take into account constraints by

- projection onto M
- minimization over M

consider the following combinations:

- projected gradient method (in Hilbert spaces)
- regularized sequential quadratic programming SQP (in general Banach spaces)

. . . taking into account ill-posedness of the underlying problem.

define steps by

- gradient descent
- regularized Newton type method

take into account constraints by

- projection onto M
- minimization over M

consider the following combinations:

- projected gradient method (in Hilbert spaces)
- regularized sequential quadratic programming SQP (in general Banach spaces)

. . . taking into account ill-posedness of the underlying problem.
regularize - then - iterate versus
regularize *by* iterating (and early stopping)

Projected gradient method

aim to solve

$$\min J^{\delta}(x) \text{ s.t. } x \in \tilde{M}^{\delta}$$

Projected gradient method

aim to solve

$$\min J^\delta(x) \text{ s.t. } x \in \tilde{M}^\delta$$

projected gradient descent:

$$\tilde{x}_{k+1} = x_k - \mu_k \nabla J^\delta(x_k)\,, \quad x_{k+1} = P_{\tilde{M}^\delta}(\tilde{x}_{k+1})$$

early stopping according to discrepancy principle (with $\tau > 1$ fixed)

$$k_* = k_*(\delta) = \min\{k \,:\, \|\nabla J^\delta(x_k)\|^2 \leq \tau\eta(\delta)\}$$

Projected gradient method

aim to solve

$$\min J^\delta(x) \text{ s.t. } x \in \tilde{M}^\delta$$

projected gradient descent:

$$\tilde{x}_{k+1} = x_k - \mu_k \nabla J^\delta(x_k)\,, \quad x_{k+1} = P_{\tilde{M}^\delta}(\tilde{x}_{k+1})$$

early stopping according to discrepancy principle (with $\tau > 1$ fixed)

$$k_* = k_*(\delta) = \min\{k \,:\, \|\nabla J^\delta(x_k)\|^2 \le \tau\eta(\delta)\}$$

Assume convexity (i.e., monotonicity of gradient)

$$\langle \nabla J^\delta(x) - \nabla J^\delta(x^\dagger), x - x^\dagger \rangle \ge \gamma \|\nabla J^\delta(x)\|^2 \quad \text{for all } x \in \tilde{M}^\delta$$

and (approximate) stationarity

$$\langle \nabla J^\delta(x^\dagger), x - x^\dagger \rangle \ge -\eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta\,,$$

Projected gradient method

aim to solve

$$\min J^\delta(x) \text{ s.t. } x \in \tilde{M}^\delta$$

projected gradient descent:

$$\tilde{x}_{k+1} = x_k - \mu_k \nabla J^\delta(x_k)\,, \quad x_{k+1} = P_{\tilde{M}^\delta}(\tilde{x}_{k+1})$$

early stopping according to discrepancy principle (with $\tau > 1$ fixed)

$$k_* = k_*(\delta) = \min\{k \,:\, \|\nabla J^\delta(x_k)\|^2 \le \tau\eta(\delta)\}$$

Assume convexity (i.e., monotonicity of gradient)

$$\langle \nabla J^\delta(x) - \nabla J^\delta(x^\dagger), x - x^\dagger \rangle \ge \gamma \|\nabla J^\delta(x)\|^2 \quad \text{for all } x \in \tilde{M}^\delta$$

and (approximate) stationarity

$$\langle \nabla J^\delta(x^\dagger), x - x^\dagger \rangle \ge -\eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta\,,$$

or combined:

$$\langle \nabla J^\delta(x), x - x^\dagger \rangle \ge \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

Projected gradient method

aim to solve

$$\min J^\delta(x) \text{ s.t. } x \in \tilde{M}^\delta$$

projected gradient descent:

$$\tilde{x}_{k+1} = x_k - \mu_k \nabla J^\delta(x_k)\,, \quad x_{k+1} = P_{\tilde{M}^\delta}(\tilde{x}_{k+1})$$

early stopping according to discrepancy principle (with $\tau > 1$ fixed)

$$k_* = k_*(\delta) = \min\{k \,:\, \|\nabla J^\delta(x_k)\|^2 \le \tau\eta(\delta)\}$$

Assume convexity (i.e., monotonicity of gradient)

$$\langle \nabla J^\delta(x) - \nabla J^\delta(x^\dagger), x - x^\dagger \rangle \ge \gamma \|\nabla J^\delta(x)\|^2 \quad \text{for all } x \in \tilde{M}^\delta$$

and (approximate) stationarity

$$\langle \nabla J^\delta(x^\dagger), x - x^\dagger \rangle \ge -\eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta\,,$$

or combined:

$$\langle \nabla J^\delta(x), x - x^\dagger \rangle \ge \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

as well as a continuity and closedness condition on ∇J and M.

Convergence

Theorem (BK and Huynh, COAP 2021)

- *For* $\delta = 0$ *the sequence* $(x_k)_{k\in\mathbb{N}}$ *converges weakly to a solution* $x^* \in M$ *of the first order optimality condition* (2) *as* $k \to \infty$.
- *If* $\delta > 0$ *then the family* $(x_{k_*(\delta)})_{\delta\in(0,\bar{\delta}]}$ *converges weakly subsequentially to a stationary point* $x^\dagger$ *according to* (2) *as* $\delta \to 0$. *If this stationary point is unique, then the whole sequence converges weakly to* $x^\dagger$.

The same assertions hold with stationarity (2) *replaced by*

(a) *minimality, i.e.,* $x^\dagger \in \operatorname{argmin}\{J(x) : x \in M\}$

or by

(b) $\|\nabla J(x^\dagger)\| = 0$

Convergence

Theorem (BK and Huynh, COAP 2021)

- *For $\delta = 0$ the sequence $(x_k)_{k\in\mathbb{N}}$ converges weakly to a solution $x^* \in M$ of the first order optimality condition (2) as $k \to \infty$.*
- *If $\delta > 0$ then the family $(x_{k_*(\delta)})_{\delta\in(0,\bar{\delta}]}$ converges weakly subsequentially to a stationary point $x^\dagger$ according to (2) as $\delta \to 0$. If this stationary point is unique, then the whole sequence converges weakly to $x^\dagger$.*

The same assertions hold with stationarity (2) replaced by

(a) *minimality, i.e., $x^\dagger \in \operatorname{argmin}\{J(x) : x \in M\}$*

or by

(b) $\|\nabla J(x^\dagger)\| = 0$

see also [Kindermann, IPI 2017] for $J(x) = \|F(x) - y\|^2$, $M = X$, generalized convexity

An SQP type constrained Newton method

aim to solve

$$\min J^{\delta}(x) \text{ s.t. } x \in \tilde{M}^{\delta}$$

An SQP type constrained Newton method

aim to solve

$$\min J^\delta(x) \text{ s.t. } x \in \tilde{M}^\delta$$

.

$$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

with

$$G^\delta(x_k) \approx J^{\delta\prime}(x_k), \qquad H^\delta(x_k) \approx J^{\delta\prime\prime}(x_k)$$

$\mathcal{R} \ldots$ regularization functional

Here X is a general Banach space.

An SQP type constrained Newton method – comparison to Gauss-Newton type methods

aim to solve

$$\min J^\delta(x) = \mathcal{S}(F(x), y) \text{ s.t. } x \in \tilde{M}^\delta$$

An SQP type constrained Newton method – comparison to Gauss-Newton type methods

aim to solve

$$\min J^\delta(x) = \mathcal{S}(F(x), y) \text{ s.t. } x \in \tilde{M}^\delta$$

SQP:

$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

IRGNM:

$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} \mathcal{S}(F(x_k) + F'(x_k)(x - x_k), y^\delta) + \alpha_k \mathcal{R}(x)$

An SQP type constrained Newton method – comparison to Gauss-Newton type methods

aim to solve

$$\min J^\delta(x) = \mathcal{S}(F(x), y) \text{ s.t. } x \in \tilde{M}^\delta$$

SQP:

$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

IRGNM:

$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} \mathcal{S}(F(x_k) + F'(x_k)(x - x_k), y^\delta) + \alpha_k \mathcal{R}(x)$

Differently from the iteratively regularized Gauss-Newton method IRGNM [Bakushinskii 1992, BK&Neubauer&Scherzer 1994 ff.] and the Levenberg Marquardt method [Hanke 1995]

- we do not necessarily neglect F'' term in $J^{\delta''}(x_k)$;
- we always solve quadratic programs if $\mathcal{R}$ is quadratic.

An SQP type constrained Newton method

$$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

with

$$G^\delta(x_k) \approx J^{\delta\prime}(x_k), \qquad H^\delta(x_k) \approx J^{\delta\prime\prime}(x_k)$$

$\mathcal{R} \ldots$ regularization functional

An SQP type constrained Newton method

$$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$
$$\text{where } Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \tfrac{1}{2} H^\delta(x_k)(x - x_k)^2$$

with

$$G^\delta(x_k) \approx J^{\delta\prime}(x_k), \qquad H^\delta(x_k) \approx J^{\delta\prime\prime}(x_k)$$
$$\mathcal{R} \ldots \text{ regularization functional}$$

a priori choice of regularization parameters:

$$\alpha_k = \alpha_0 \theta^k$$

(alternatively, a posteriori choice of α_k as in [Hanke 1997])

An SQP type constrained Newton method

$$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$
$$\text{where } Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \tfrac{1}{2} H^\delta(x_k)(x - x_k)^2$$

with

$$G^\delta(x_k) \approx J^{\delta\prime}(x_k), \qquad H^\delta(x_k) \approx J^{\delta\prime\prime}(x_k)$$
$$\mathcal{R} \ldots \text{ regularization functional}$$

a priori choice of regularization parameters:

$$\alpha_k = \alpha_0 \theta^k$$

(alternatively, a posteriori choice of α_k as in [Hanke 1997])

early stopping according to discrepancy principle

$$k_* = k_*(\delta) = \min\{k \,:\, J^\delta(x_k) \le \tau \eta(\delta)\}$$

for some constants $\alpha_0 > 0$, $\theta \in (0,1)$, $\tau > 1$.

An SQP type constrained Newton method

$$x_{k+1} \in \text{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

An SQP type constrained Newton method

$$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

assume

$$G^\delta(x_k) : X \to \mathbb{R} \text{ linear }, \quad H^\delta(x_k) : X^2 \to \mathbb{R} \text{ bilinear },$$

$$\mathcal{R} : X \to [0, \infty] \text{ proper with domain } \operatorname{dom}(\mathcal{R}) \supseteq \bigcup_{\delta \in (0, \bar{\delta})} \tilde{M}^\delta \cup M$$

An SQP type constrained Newton method

$$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

assume

$$G^\delta(x_k) : X \to \mathbb{R} \text{ linear}, \quad H^\delta(x_k) : X^2 \to \mathbb{R} \text{ bilinear},$$

$$\mathcal{R} : X \to [0, \infty] \text{ proper with domain } \operatorname{dom}(\mathcal{R}) \supseteq \bigcup_{\delta \in (0, \bar{\delta})} \tilde{M}^\delta \cup M$$

restriction on nonlinearity

$$\underline{a} J^\delta(x_+) - \underline{b} J^\delta(x) \le G^\delta(x)(x_+ - x) + \tfrac{1}{2} H^\delta(x)(x_+ - x)^2 \le \overline{a} J^\delta(x_+) - \overline{b} J^\delta(x)$$

$(*)$ for all $x, x_+ \in \tilde{M}^\delta, \quad \delta \in (0, \bar{\delta})$,

Taylor $\rightsquigarrow$ $\underline{a}, \overline{a}, \underline{b}, \overline{b}, \sim 1$

An SQP type constrained Newton method

$$x_{k+1} \in \operatorname{argmin}_{x \in \tilde{M}^\delta} Q_k^\delta(x) + \alpha_k \mathcal{R}(x)$$

where $Q_k^\delta(x) = J^\delta(x_k) + G^\delta(x_k)(x - x_k) + \frac{1}{2} H^\delta(x_k)(x - x_k)^2$

assume

$$G^\delta(x_k) : X \to \mathbb{R} \text{ linear}, \quad H^\delta(x_k) : X^2 \to \mathbb{R} \text{ bilinear},$$

$$\mathcal{R} : X \to [0, \infty] \text{ proper with domain } \operatorname{dom}(\mathcal{R}) \supseteq \bigcup_{\delta \in (0, \bar{\delta})} \tilde{M}^\delta \cup M$$

restriction on nonlinearity

$$\underline{a} J^\delta(x_+) - \underline{b} J^\delta(x) \le G^\delta(x)(x_+ - x) + \tfrac{1}{2} H^\delta(x)(x_+ - x)^2 \le \overline{a} J^\delta(x_+) - \overline{b} J^\delta(x)$$

$(*)$ for all $x, x_+ \in \tilde{M}^\delta, \quad \delta \in (0, \bar{\delta})$,

Taylor $\rightsquigarrow$ $\underline{a}, \overline{a}, \underline{b}, \overline{b}, \sim 1$

and some "usual" continuity/closedness/compactness assumptions on J^δ, G^δ, H^δ, $\tilde{M}^\delta$ in some topology $\mathcal{T}$.

Convergence

Theorem (BK and Huynh, COAP 2021)

- *For any* $\delta \in (0, \bar{\delta})$*, and any* $x_0 \in \bigcap_{\delta \in (0,\bar{\delta})} \tilde{M}^\delta \cap M$*,*
  - *the iterates* x_k *are well-defined for all* $k \leq k_*(\delta)$
  - $k_*(\delta)$ *is finite;*
  - *for all* $k \in \{1, \ldots, k_*(\delta)\}$ *we have*

$$J^\delta(x_k) \leq \frac{\underline{b}-\bar{b}}{\underline{a}} J^\delta(x_{k-1}) + \frac{1}{\underline{a}} \alpha_k \mathcal{R}(x^\dagger) + \frac{\bar{a}}{\underline{a}} \eta;$$

- *As* $\delta \to 0$*, the final iterates* $x_{k_*(\delta)}$ *tend to a solution of the inverse problem* $\mathcal{T}$*-subsequentially, i.e., every sequence* $x_{k_*(\delta_j)}$ *with* $\delta_j \to 0$ *as* $j \to \infty$ *has a* $\mathcal{T}$ *convergent subsequence and the limit of every* $\mathcal{T}$ *convergent subsequence solves* (1).

Application in diffusion/impedance identification

Identify $\sigma = \sigma(x)$ in the elliptic PDE

$$\nabla \cdot (\sigma \nabla \phi) = 0 \text{ in } \Omega$$

from observations of ϕ.

Application in diffusion/impedance identification

Identify $\sigma = \sigma(x)$ in the elliptic PDE

$$\nabla \cdot (\sigma \nabla \phi) = 0 \text{ in } \Omega$$

from observations of ϕ.

- electrical impedance tomography EIT (Calderon's problem)
  $v = \phi|_{\partial\Omega}$ (the voltage at the boundary)
- impedance acoustic tomography IAT
  $\mathcal{H} = \sigma|\nabla\phi|^2$ (the power density)
- simplified version of inverse groundwater filtration GWF (Darcy's problem)
  $p = \phi$ (the hydraulic head)

Minimization based formulations

$\mathbf{E} = \nabla\phi, \quad \mathbf{J} = \nabla^{\perp}\psi$

- Cost function part incorporating the model: Kohn-Vogelius functional

$$J_{mod}^{KV}(\sigma, \mathbf{E}, \mathbf{J}) = \frac{1}{2}\int_{\Omega} \left|\sqrt{\sigma}\mathbf{E} - \frac{1}{\sqrt{\sigma}}\mathbf{J}\right|^2 d\Omega,$$

Minimization based formulations

$\mathbf{E} = \nabla\phi, \quad \mathbf{J} = \nabla^{\perp}\psi$

- Cost function part incorporating the model: Kohn-Vogelius functional

$$J_{mod}^{KV}(\sigma, \mathbf{E}, \mathbf{J}) = \frac{1}{2}\int_{\Omega} \left|\sqrt{\sigma}\mathbf{E} - \frac{1}{\sqrt{\sigma}}\mathbf{J}\right|^2 d\Omega,$$

- Cost function part incorporating the observations:

$$J_{obs}^{EIT}(\phi; v) = \frac{1}{2}\int_{\partial\Omega} (\phi - v)^2 \, d\Omega$$

$$J_{obs}^{IAT}(\mathbf{E}, \mathbf{J}; \mathcal{H}) = \frac{1}{2}\int_{\Omega} (\mathbf{J}\cdot\mathbf{E} - \mathcal{H})^2 \, d\Omega \text{ or } \frac{1}{2}\int_{\Omega} (\sigma|\mathbf{E}|^2 - \mathcal{H})^2 \, d\Omega$$

$$J_{obs}^{GWF}(\phi; p) = \frac{1}{2}\|\phi - p\|^2_{H^s(\Omega)}$$

Minimization based formulations

several excitations $(j_1, \ldots, j_I)$

$\rightsquigarrow$ several states $\Phi = (\phi_1, \ldots, \phi_I)$, $\Psi = (\psi_1, \ldots, \psi_I)$:

- Cost function part incorporating the model: Kohn-Vogelius functional

$$J_{mod}^{KV}(\sigma, \mathbf{E}, \mathbf{J}) = \frac{1}{2} \sum_{i=1}^{I} \int_{\Omega} \left| \sqrt{\sigma} \mathbf{E}_i - \frac{1}{\sqrt{\sigma}} \mathbf{J}_i \right|^2 d\Omega,$$

- Cost function part incorporating the observations:

$$J_{obs}^{EIT}(\Phi; \vec{v}) = \frac{1}{2} \sum_{i=1}^{I} \int_{\partial\Omega} (\phi_i - v_i)^2 \, d\Omega$$

$$J_{obs}^{IAT}(\mathbf{E}, \mathbf{J}; \vec{\mathcal{H}}) = \frac{1}{2} \sum_{i=1}^{I} \int_{\Omega} (\mathbf{J}_i \cdot \mathbf{E}_i - \mathcal{H}_i)^2 \, d\Omega \text{ or } \frac{1}{2} \sum_{i=1}^{I} \int_{\Omega} (\sigma |\mathbf{E}_i|^2 - \mathcal{H}_i)^2 \, d\Omega$$

$$J_{obs}^{GWF}(\Phi; \vec{p}) = \frac{1}{2} \|\phi_i - p_i\|_{H^s(\Omega)}^2$$

Minimization based formulations

Three versions of treating σ, Φ, Ψ

- all-at-once: minimize with respect to σ, Φ, Ψ simultaneously

Minimization based formulations

Three versions of treating σ, Φ, Ψ

- all-at-once: minimize with respect to σ, Φ, Ψ simultaneously
- eliminate σ, minimize with respect to Φ, Ψ, setting

$$\sigma(\Phi, \Psi) = \min\left\{\overline{\sigma}, \max\left\{\underline{\sigma}, \sqrt{\frac{\sum_{i=1}^{I} |\nabla^{\perp}\psi_i|^2}{\sum_{i=1}^{I} |\nabla\phi_i|^2}}\right\}\right\}.$$

Minimization based formulations

Three versions of treating σ, Φ, Ψ

- all-at-once: minimize with respect to σ, Φ, Ψ simultaneously
- eliminate σ, minimize with respect to Φ, Ψ, setting
$$\sigma(\Phi, \Psi) = \min\left\{\overline{\sigma}, \max\left\{\underline{\sigma}, \sqrt{\frac{\sum_{i=1}^{I} |\nabla^{\perp}\psi_i|^2}{\sum_{i=1}^{I} |\nabla\phi_i|^2}}\right\}\right\}.$$
- eliminate Φ, Ψ, minimize with respect to σ, setting $\phi_i(\sigma)$ $\psi_i(\sigma)$, $\mathbf{E}_i(\sigma)$ according to:

$$\phi(\sigma) \text{ solves } \left\{\begin{array}{rcll} \nabla\cdot(\sigma\nabla\phi) & = & 0 & \text{in } \Omega \\ \phi & = & v & \text{on } \partial\Omega \end{array}\right.$$

$$\phi_N(\sigma) \text{ solves } \left\{\begin{array}{rcll} \nabla\cdot(\sigma\nabla\phi) & = & 0 & \text{in } \Omega \\ \nabla\phi\cdot\nu & = & j & \text{on } \partial\Omega \end{array}\right. \qquad \int_\Omega \phi \, d\Omega = 0$$

$$\psi(\sigma) \text{ solves } \left\{\begin{array}{rcll} \nabla^{\perp}\cdot(\frac{1}{\sigma}\nabla^{\perp}\psi) & = & 0 & \text{in } \Omega \\ \psi & = & \alpha & \text{on } \partial\Omega \end{array}\right.$$

$$\mathbf{E}(\sigma) = \frac{\nabla^{\perp}\psi(\sigma)}{\sigma} \text{ pointwise in } \Omega$$

Minimization based formulations

EIT:

$$(i)\ \min_{\sigma,\phi,\psi}\{J_{mod}(\sigma,\nabla\phi,\nabla^{\perp}\psi)+\beta J_{obs}^{EIT}(\phi;v)\ :\ \sigma\in L^2_{[\underline{\sigma},\overline{\sigma}]}(\Omega)\,,\ \phi\in H^1_{\diamond}(\Omega),\psi\in H^1_0(\Omega)+\psi_0\}$$

$$(ii)\ \min_{\sigma,\phi,\psi}\{J_{mod}(\sigma,\nabla\phi,\nabla^{\perp}\psi)\ :\ \sigma\in L^2_{[\underline{\sigma},\overline{\sigma}]}(\Omega)\,,\ \phi\in H^1_0(\Omega)+\phi_0\,,\ \psi\in H^1_0(\Omega)+\psi_0\}$$

$$(iii)\ \min_{\phi,\psi}\{J_{mod}(\sigma(\nabla\phi,\nabla^{\perp}\psi),\nabla\phi,\nabla^{\perp}\psi)+\beta J_{obs}^{EIT}(\phi;v)\ :\ \phi\in H^1_{\diamond}(\Omega),\psi\in H^1_0(\Omega)+\psi_0\}$$

$$(iv)\ \min_{\phi,\psi}\{J_{mod}(\sigma(\nabla\phi,\nabla^{\perp}\psi),\nabla\phi,\nabla^{\perp}\psi)\ :\ \phi\in H^1_0(\Omega)+\phi_0,\psi\in H^1_0(\Omega)+\psi_0\}$$

$$(v)\ \min_{\sigma}\{J_{mod}(\sigma,\nabla\phi(\sigma),\nabla^{\perp}\psi(\sigma))\ :\ \sigma\in L^2_{[\underline{\sigma},\overline{\sigma}]}(\Omega)\}$$

$$(vi)\ \min_{\sigma}\{J_{obs}^{EIT}(\phi_N(\sigma);v)\ :\ \sigma\in L^2_{[\underline{\sigma},\overline{\sigma}]}(\Omega)\}$$

IAT:

$$(i)\ \min_{\sigma,\phi,\psi}\{J_{mod}(\sigma,\nabla\phi,\nabla^{\perp}\psi)+\beta\left\{\begin{array}{l}J_{obs_1}^{IAT}(\nabla\phi,\nabla^{\perp}\psi;\mathcal{H})\\ J_{obs_2}^{IAT}(\sigma,\nabla\phi;\mathcal{H})\end{array}\right.\ :\ \sigma\in L^2_{[\underline{\sigma},\overline{\sigma}]}(\Omega)\,,\ \phi\in H^1_{\diamond}(\Omega),\psi\in H^1_0(\Omega)+\psi_0\}$$

$$(ii)\ \min_{\phi,\psi}\{J_{mod}(\sigma(\nabla\phi,\nabla^{\perp}\psi),\nabla\phi,\nabla^{\perp}\psi)+\beta\left\{\begin{array}{l}J_{obs_1}^{IAT}(\nabla\phi,\nabla^{\perp}\psi;\mathcal{H})\\ J_{obs_2}^{IAT}(\sigma,\nabla\phi;\mathcal{H})\end{array}\right.\ :\ \phi\in H^1_{\diamond}(\Omega),\psi\in H^1_0(\Omega)+\psi_0\}$$

$$(iii)\ \min_{\sigma}\{\left\{\begin{array}{l}J_{obs_1}^{IAT}(\mathbf{E}(\sigma),\nabla^{\perp}\psi(\sigma);\mathcal{H})\\ J_{obs_2}^{IAT}(\sigma,\mathbf{E}(\sigma);\mathcal{H})\end{array}\right.\ :\ \sigma\in L^2_{[\underline{\sigma},\overline{\sigma}]}(\Omega)\}$$

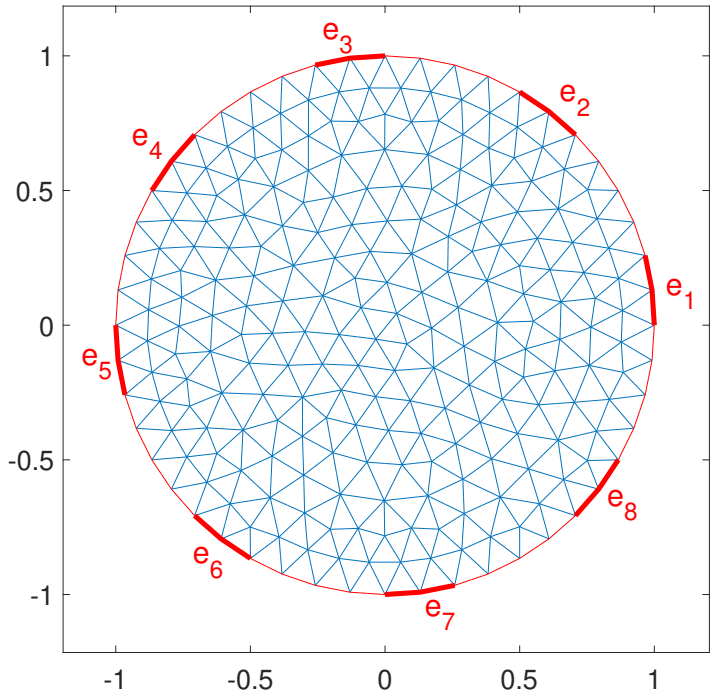
GWF:

$$(i)\ \min_{\sigma,\phi,\psi}\{J_{mod}(\sigma,\nabla\phi,\nabla^{\perp}\psi)+\beta\left\{\begin{array}{l}J_{obs_1}^{GWF}(\phi;p)\\ J_{obs_2}^{GWF}(\nabla\phi;g)\end{array}\right.\ :\ \sigma\in L^2_{[\underline{\sigma},\overline{\sigma}]}(\Omega)\,,\ \phi\in H^1_{\diamond}(\Omega),\psi\in H^1_0(\Omega)+\psi_0\}$$

$$(ii)\ \min_{\phi,\psi}\{J_{mod}(\sigma(\nabla\phi,\nabla^{\perp}\psi),\nabla\phi,\nabla^{\perp}\psi)+\beta\left\{\begin{array}{l}J_{obs_1}^{GWF}(\phi;p)\\ J_{obs_2}^{GWF}(\nabla\phi;g)\end{array}\right.\ :\ \phi\in H^1_{\diamond}(\Omega),\psi\in H^1_0(\Omega)+\psi_0\}$$

Computational setup

8 electrodes $\rightsquigarrow$ 28 possible excitation combinations; we consider:



- $I = 1$, with $j_{1,1} = 1$, $j_{5,1} = -1$ and $j_{k,1} = 0$ otherwise;
- $I = 2$, with $j_{1,1} = j_{3,2} = 1$, $j_{5,1} = j_{7,2} = -1$, and $j_{k,i} = 0$ otherwise;
- $I = 4$, with $j_{1,1} = j_{3,2} = j_{2,3} = j_{4,4} = 1$, $j_{5,1} = j_{7,2} = j_{6,3} = j_{8,4} = -1$ and $j_{k,i} = 0$ otherwise.
- $I = 28$, with all $\binom{8}{2}$ combinations of setting $j_{k,i} = 1$, $j_{\ell,i} = -1$ for $k \neq \ell \in \{1, \ldots, 8\}$

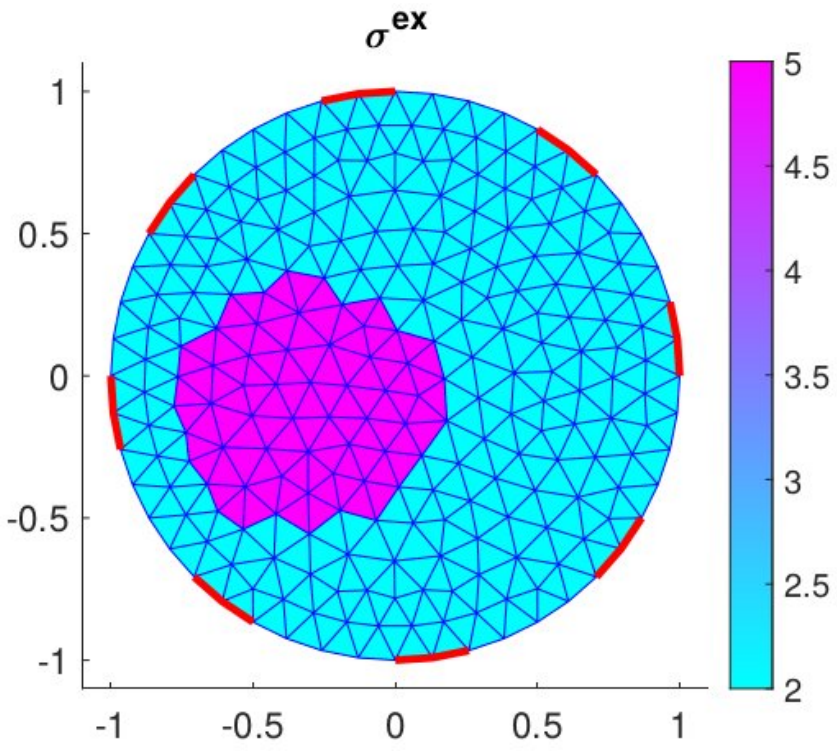
starting values: $\sigma_0 = \frac{1}{2}(\underline{\sigma} + \overline{\sigma})$; $\Phi_0 = \Phi(\sigma_0)$, $\Psi_0 = \Psi(\sigma_0)$

numerical results with

projected gradient method

by Kha Van Huynh

Numerical Tests

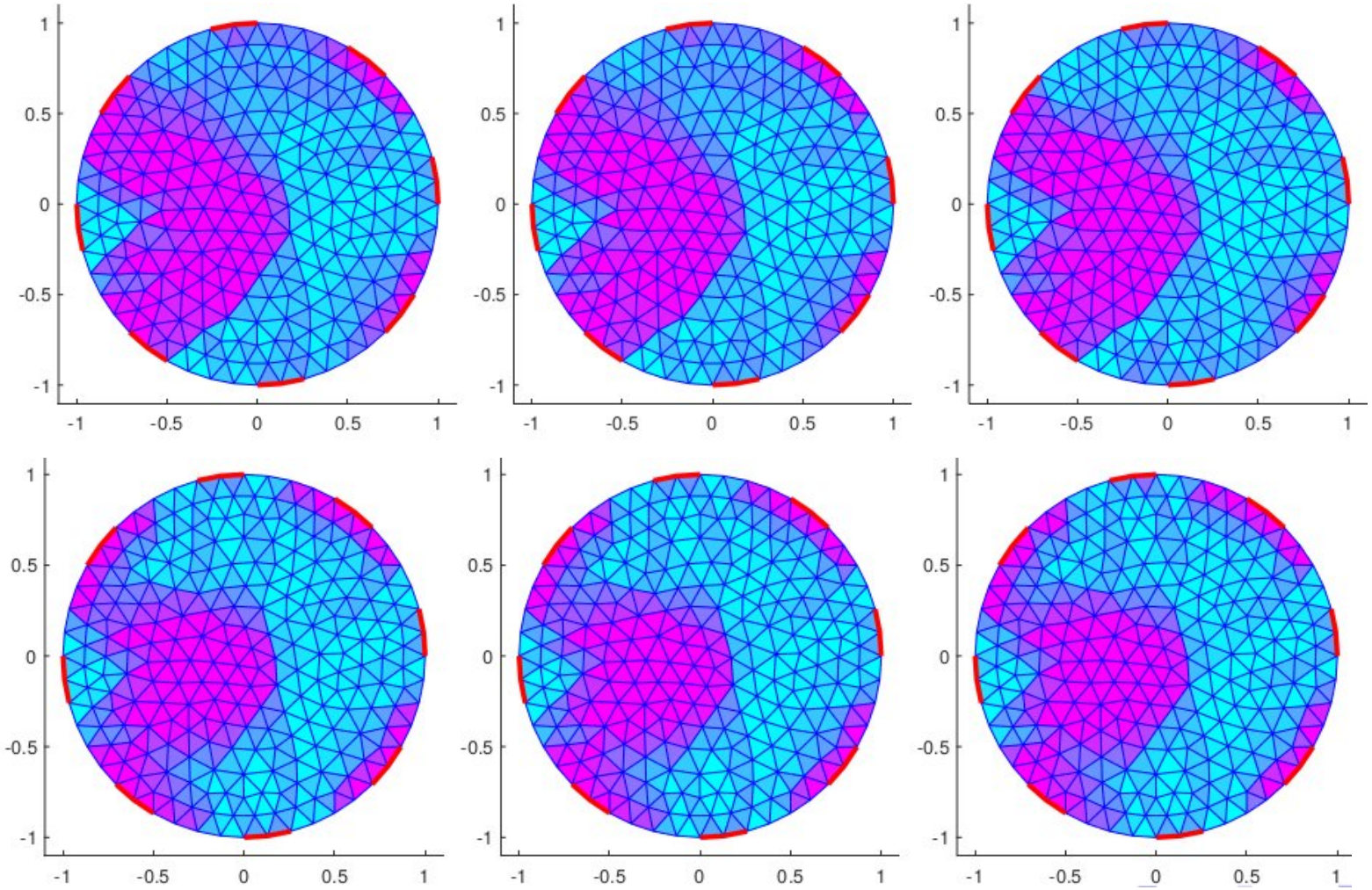


test case 0: exact σ

GWF: all-at-once reconstructions with $l \in \{1, 2\}$

$l = 1$:

$l = 2$:

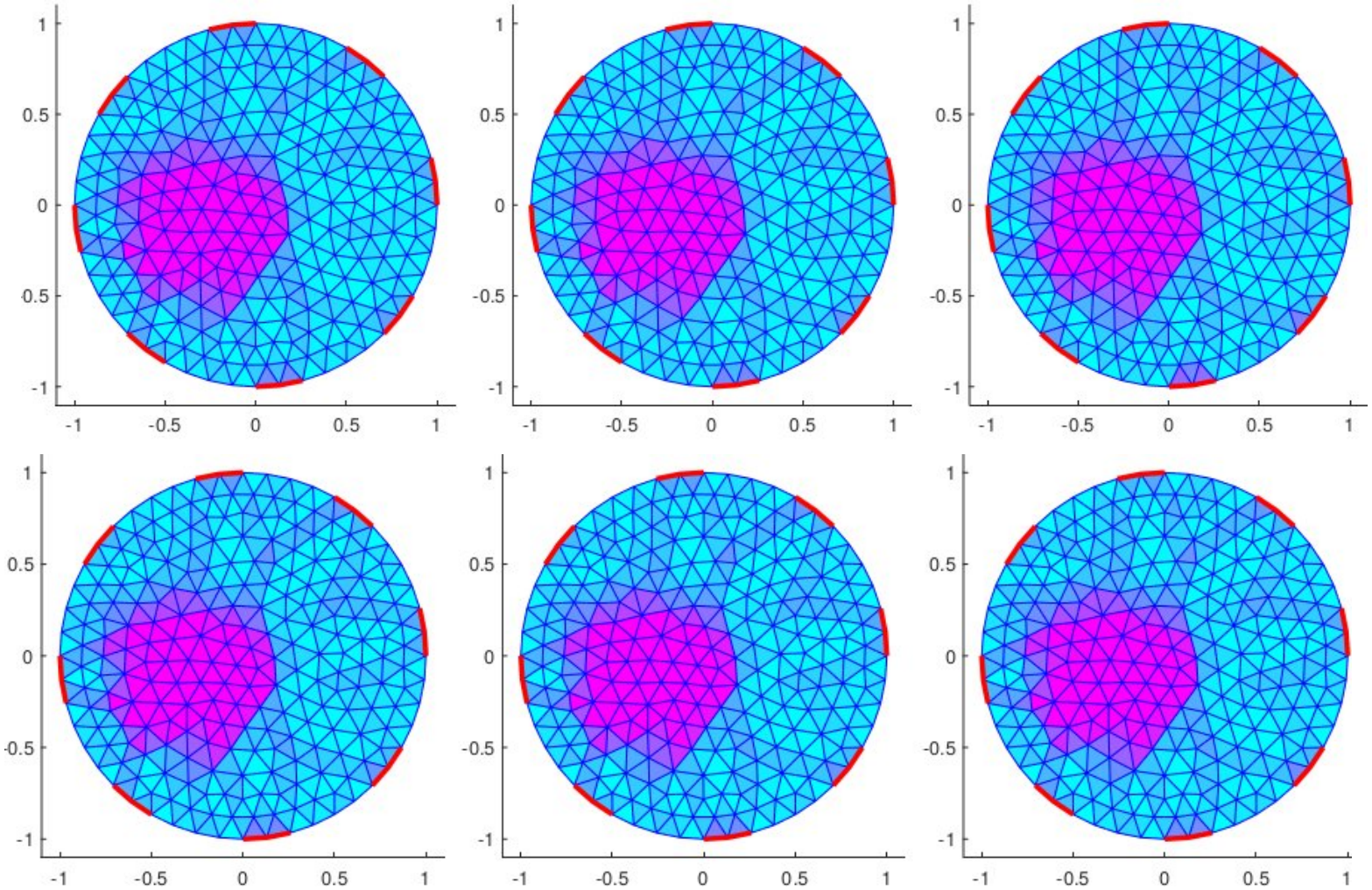


$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

GWF: all-at-once reconstructions with $l \in \{4, 28\}$

$l = 4$:

$l = 28$:

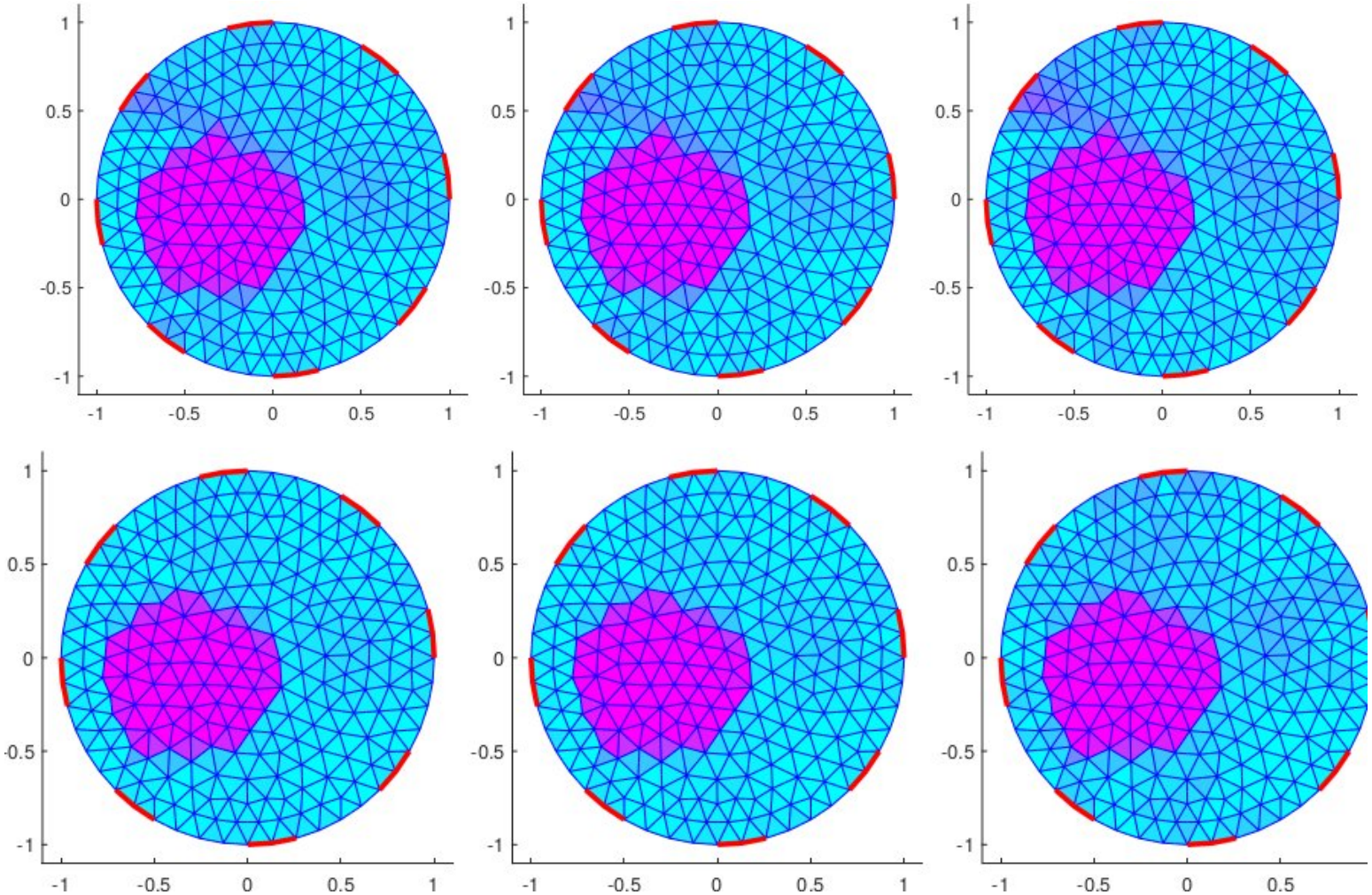


$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: all-at-once reconstructions with $l \in \{1, 2\}$

$l = 1$:

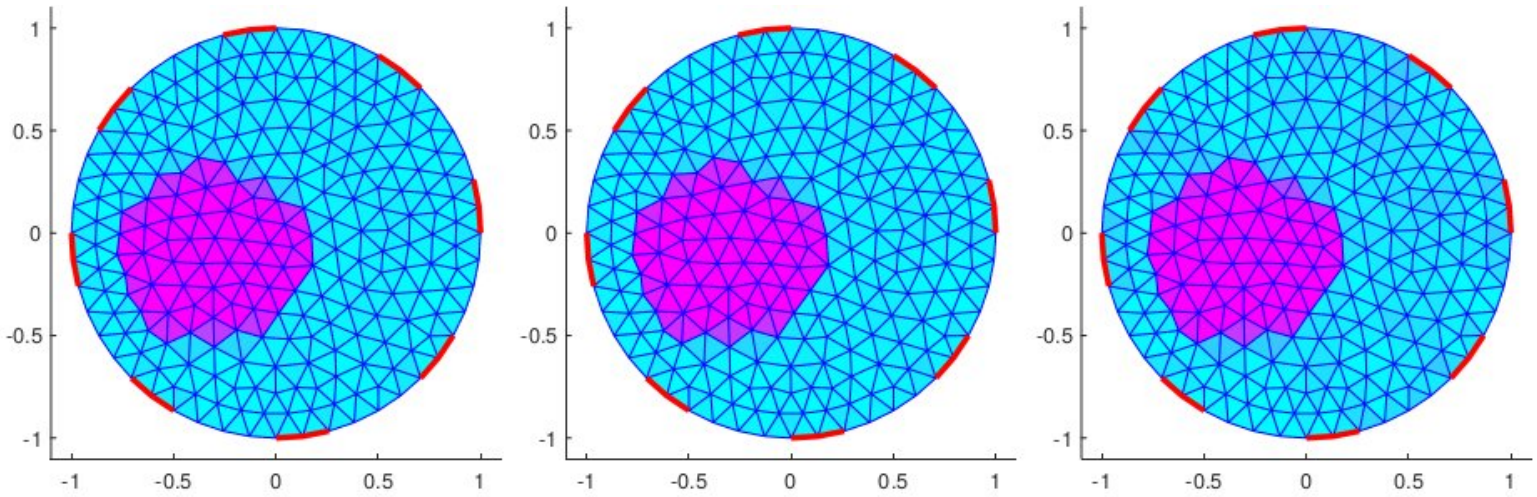
$l = 2$:



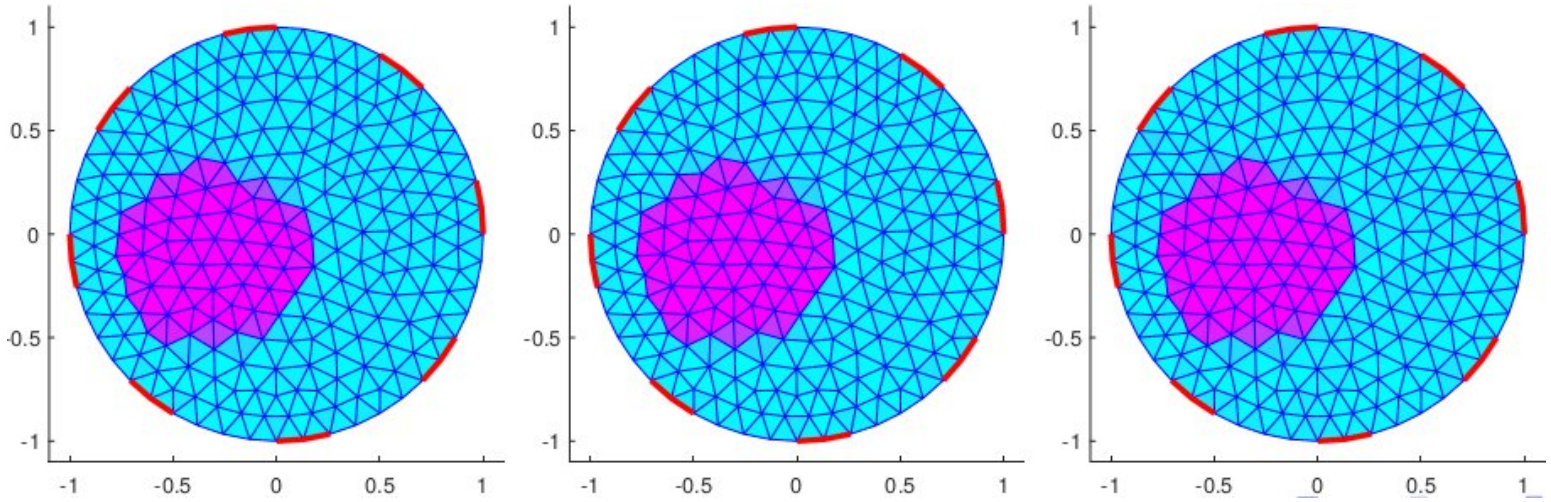
$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: all-at-once reconstructions with $I \in \{4, 28\}$

$I = 4$:



$I = 28$:

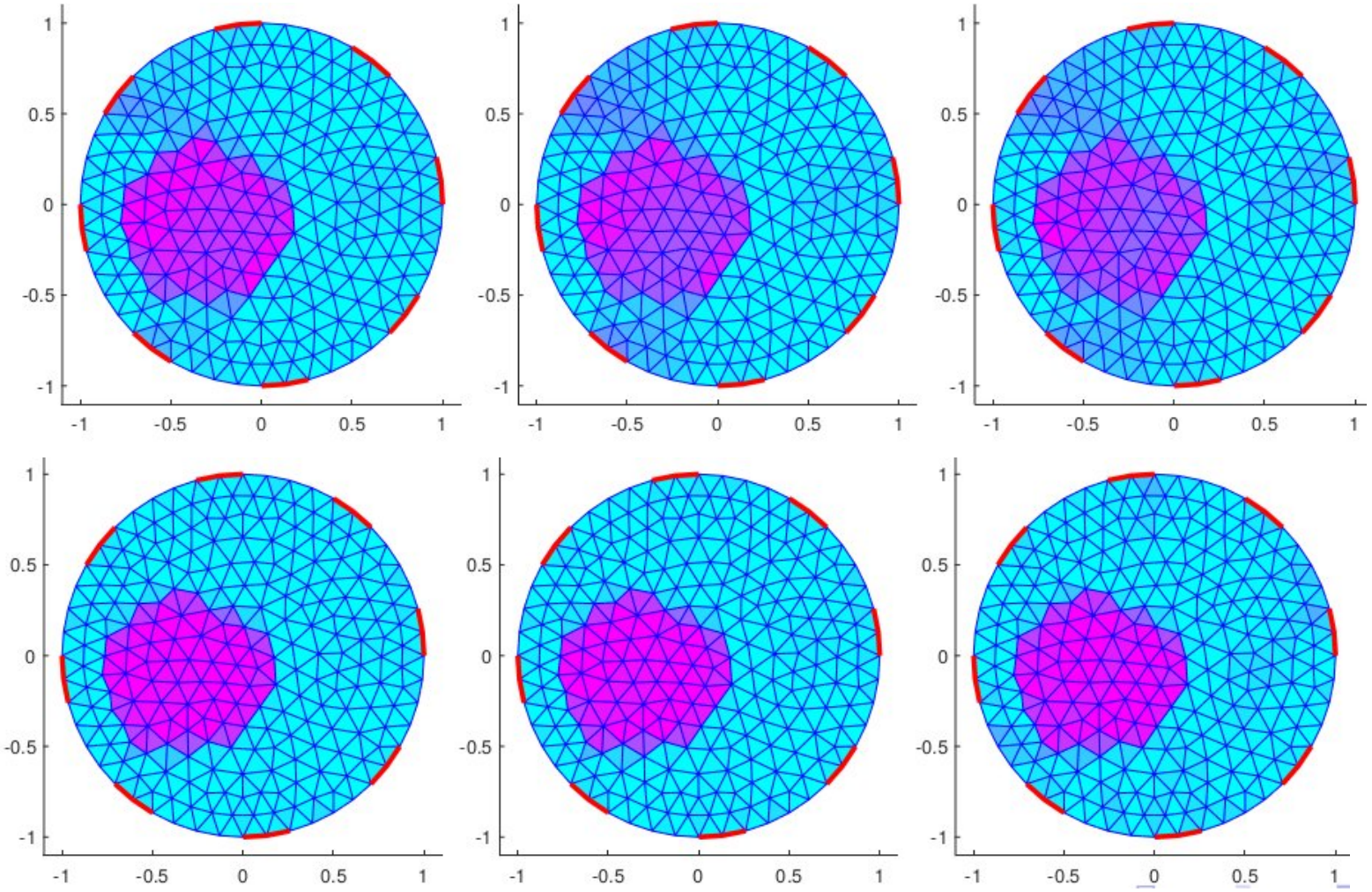


$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: eliminating σ reconstructions with $l \in \{1, 2\}$

$l = 1$:

$l = 2$:

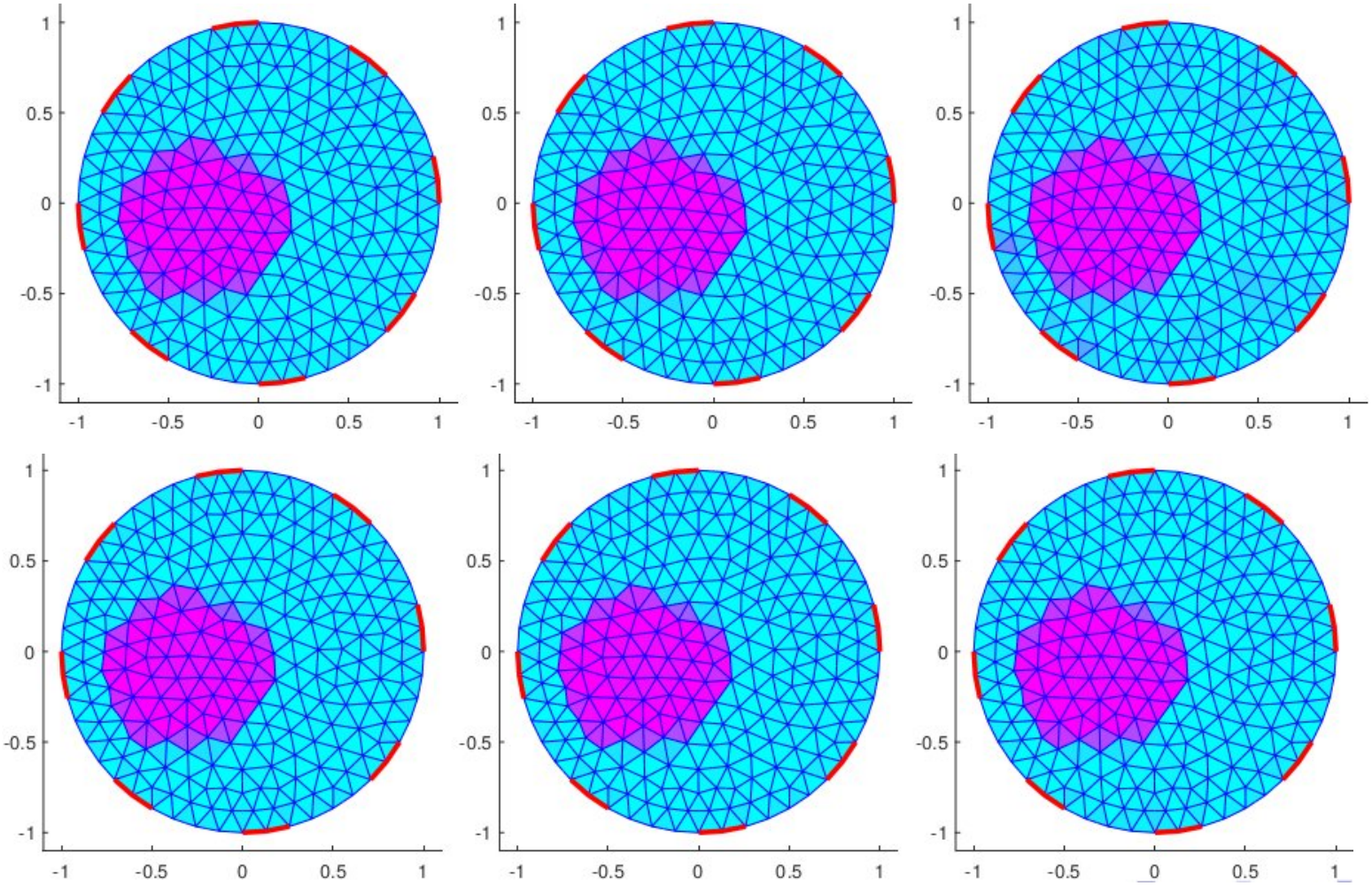


$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: eliminating σ reconstructions with $l \in \{4, 28\}$

$l = 4$:

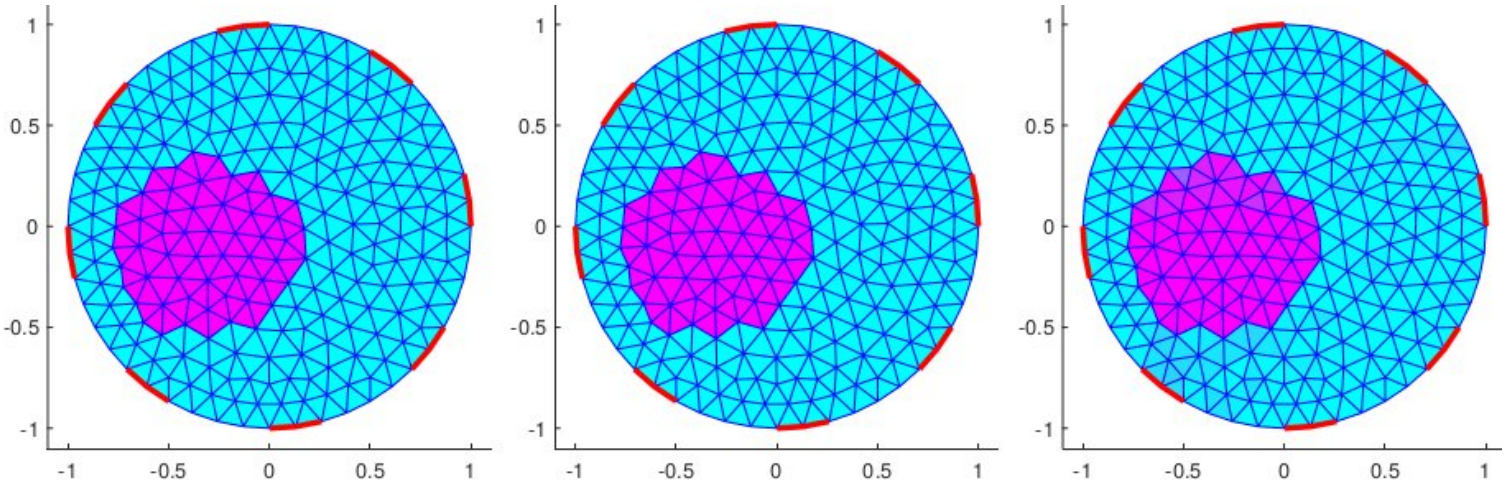
$l = 28$:



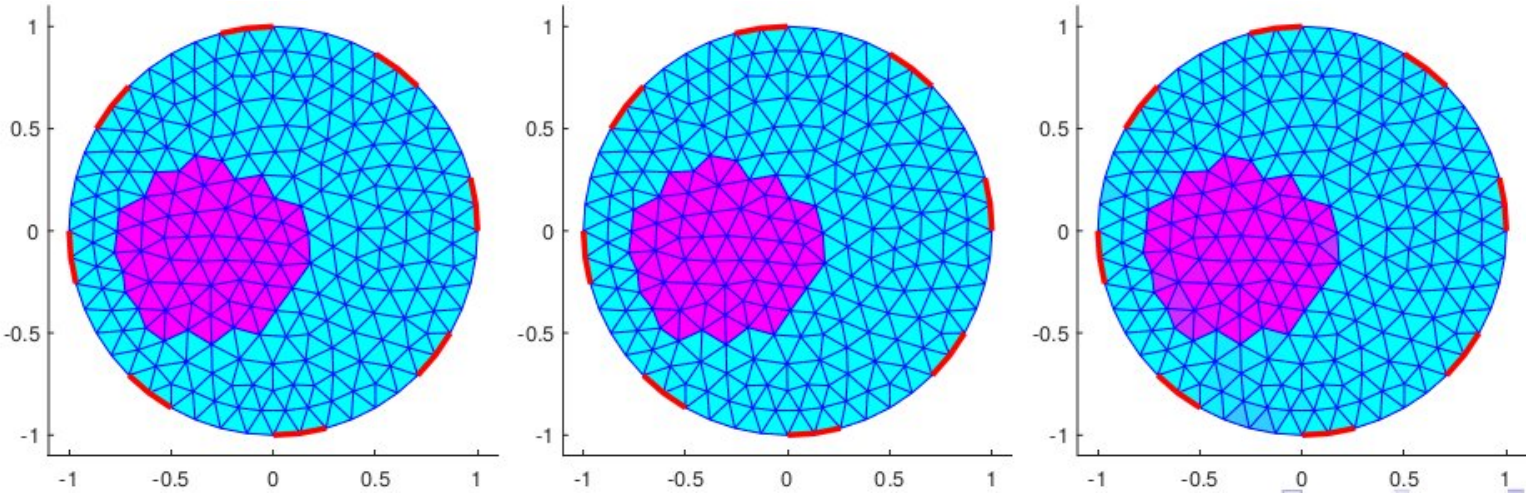
$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: eliminating Φ, Ψ reconstructions with $l \in \{1, 2\}$

$l = 1$:



$l = 2$:

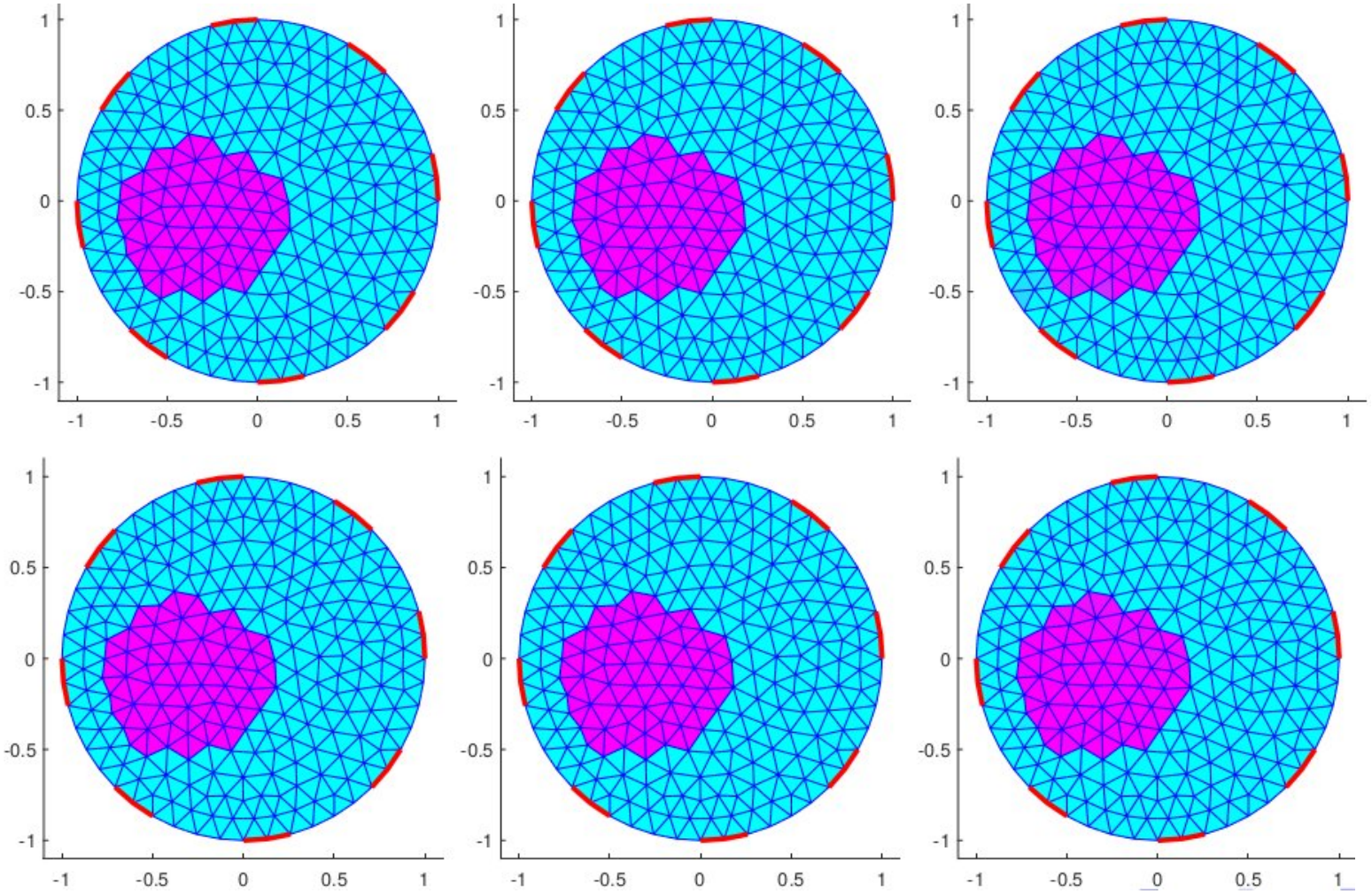


$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: eliminating Φ, Ψ reconstructions with $l \in \{4, 28\}$

$l = 4$:

$l = 28$:



$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: comparison

relative errors:

	all-at-once			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.130	0.076	0.061	0.060
0.01	0.131	0.077	0.062	0.060
0.1	0.159	0.118	0.086	0.064

	elim. σ			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.115	0.071	0.064	0.065
0.01	0.163	0.074	0.062	0.065
0.1	0.174	0.087	0.084	0.065

	elim. Φ, Ψ			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.	0.	0.	0.
0.01	0.004	0.003	0.003	0.001
0.1	0.044	0.034	0.024	0.013

IAT: comparison

relative errors:

	all-at-once			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.130	0.076	0.061	0.060
0.01	0.131	0.077	0.062	0.060
0.1	0.159	0.118	0.086	0.064

	elim. σ			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.115	0.071	0.064	0.065
0.01	0.163	0.074	0.062	0.065
0.1	0.174	0.087	0.084	0.065

	elim. Φ, Ψ			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.	0.	0.	0.
0.01	0.004	0.003	0.003	0.001
0.1	0.044	0.034	0.024	0.013

runtimes (in hours):

all-at-once			
$l = 1$	$l = 2$	$l = 4$	$l = 28$
262	74	20.4	22.8
259	74	20.8	22.9
274	84.5	21.8	34.4

elim. σ			
$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.16	0.33	0.23	0.42
0.08	0.29	0.34	0.39
0.07	0.36	0.18	0.39

elim. Φ, Ψ			
$l = 1$	$l = 2$	$l = 4$	$l = 28$
41.1	8.38	2.9	13.6
23.1	17.0	7.5	26.7
10.6	11.5	5.5	33.6

IAT: comparison

relative errors:

	all-at-once			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.130	0.076	0.061	0.060
0.01	0.131	0.077	0.062	0.060
0.1	0.159	0.118	0.086	0.064

	elim. σ			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.115	0.071	0.064	0.065
0.01	0.163	0.074	0.062	0.065
0.1	0.174	0.087	0.084	0.065

	elim. Φ, Ψ			
δ	$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.	0.	0.	0.	0.
0.01	0.004	0.003	0.003	0.001
0.1	0.044	0.034	0.024	0.013

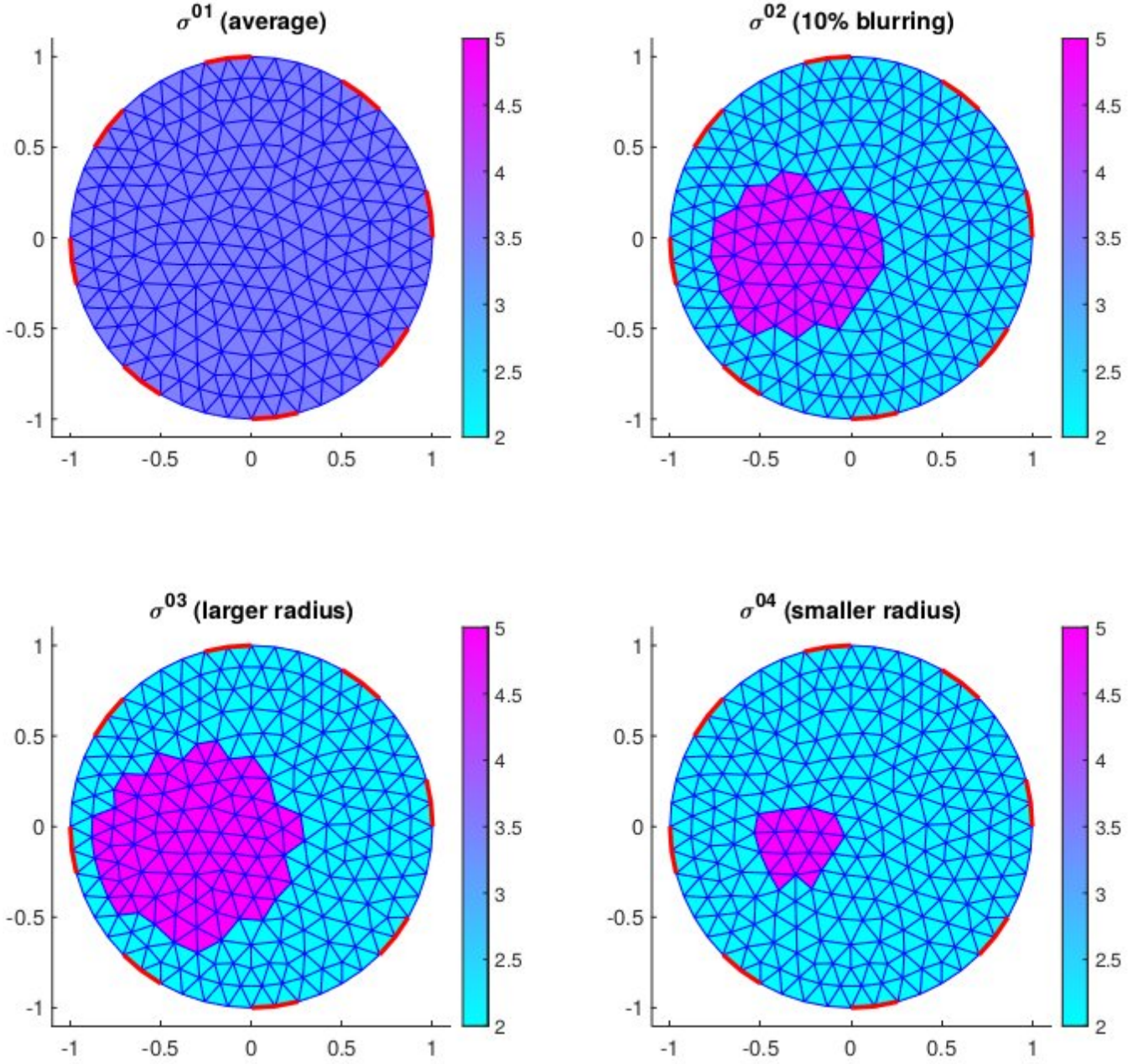
runtimes (in hours):

all-at-once			
$l = 1$	$l = 2$	$l = 4$	$l = 28$
262	74	20.4	22.8
259	74	20.8	22.9
274	84.5	21.8	34.4

elim. σ			
$l = 1$	$l = 2$	$l = 4$	$l = 28$
0.16	0.33	0.23	0.42
0.08	0.29	0.34	0.39
0.07	0.36	0.18	0.39

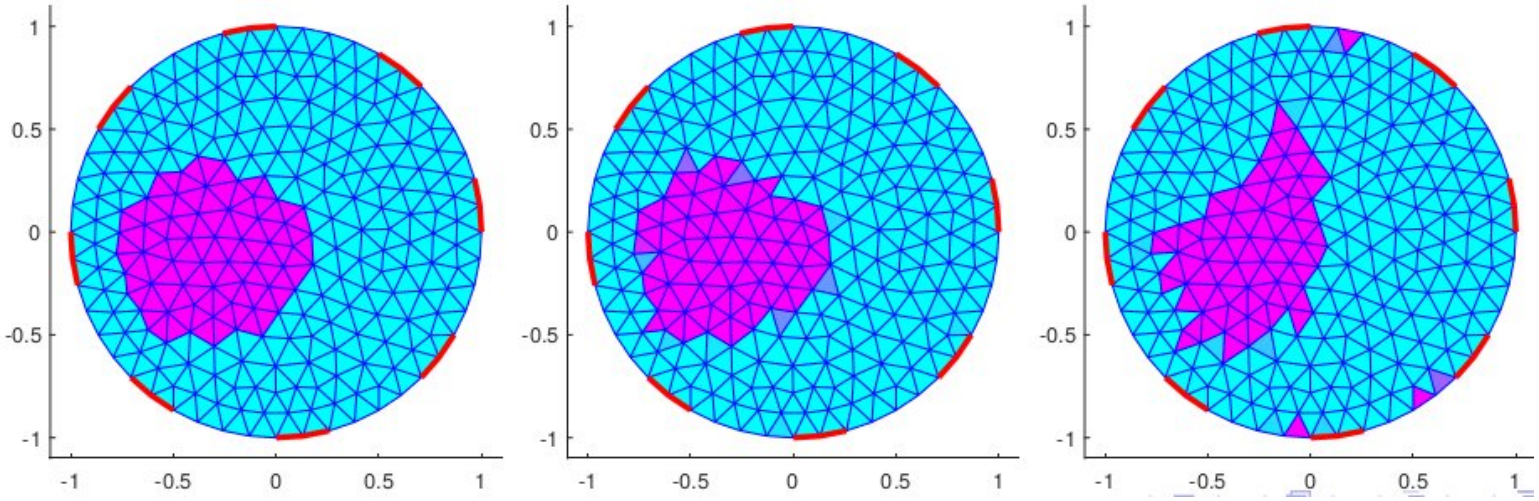
elim. Φ, Ψ			
$l = 1$	$l = 2$	$l = 4$	$l = 28$
41.1	8.38	2.9	13.6
23.1	17.0	7.5	26.7
10.6	11.5	5.5	33.6

EIT: four different starting values



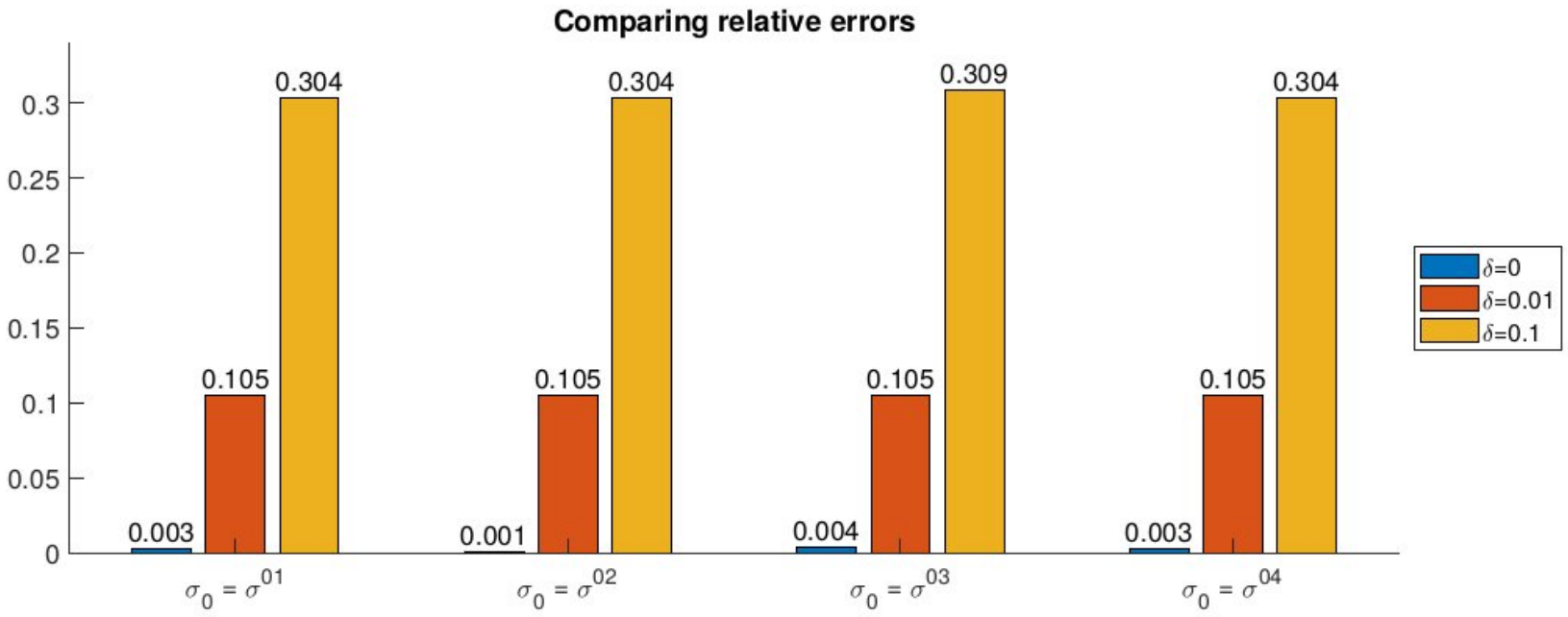
EIT: eliminating Φ, Ψ reconstructions with $l = 28$

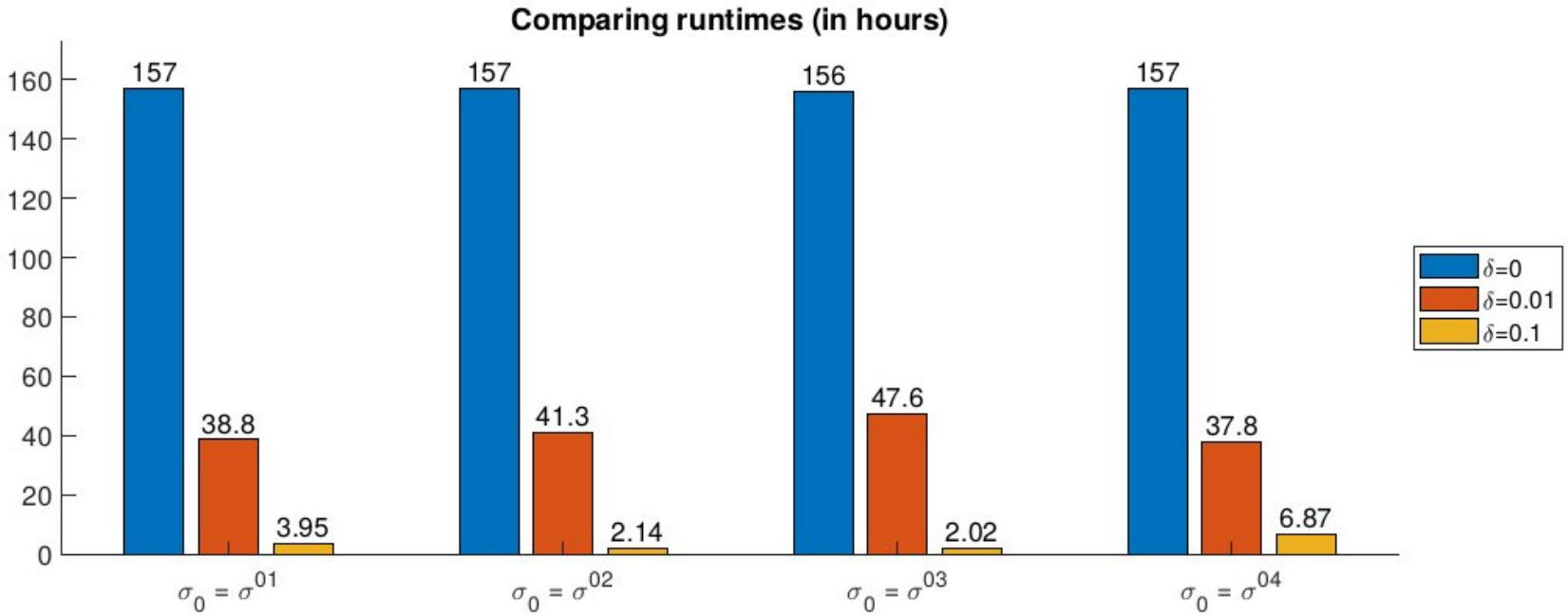
$l = 28$:



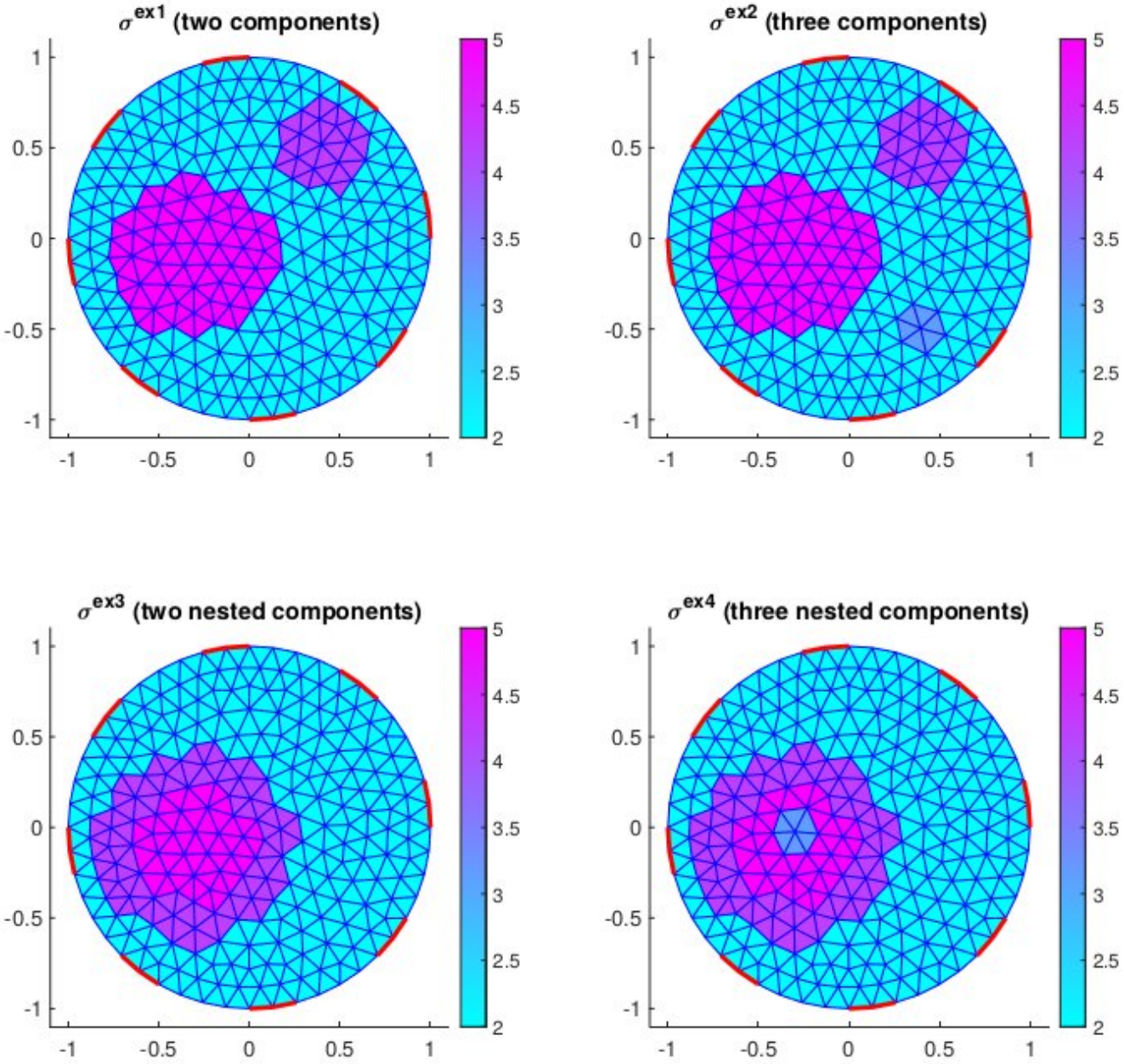
$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

EIT: comparison





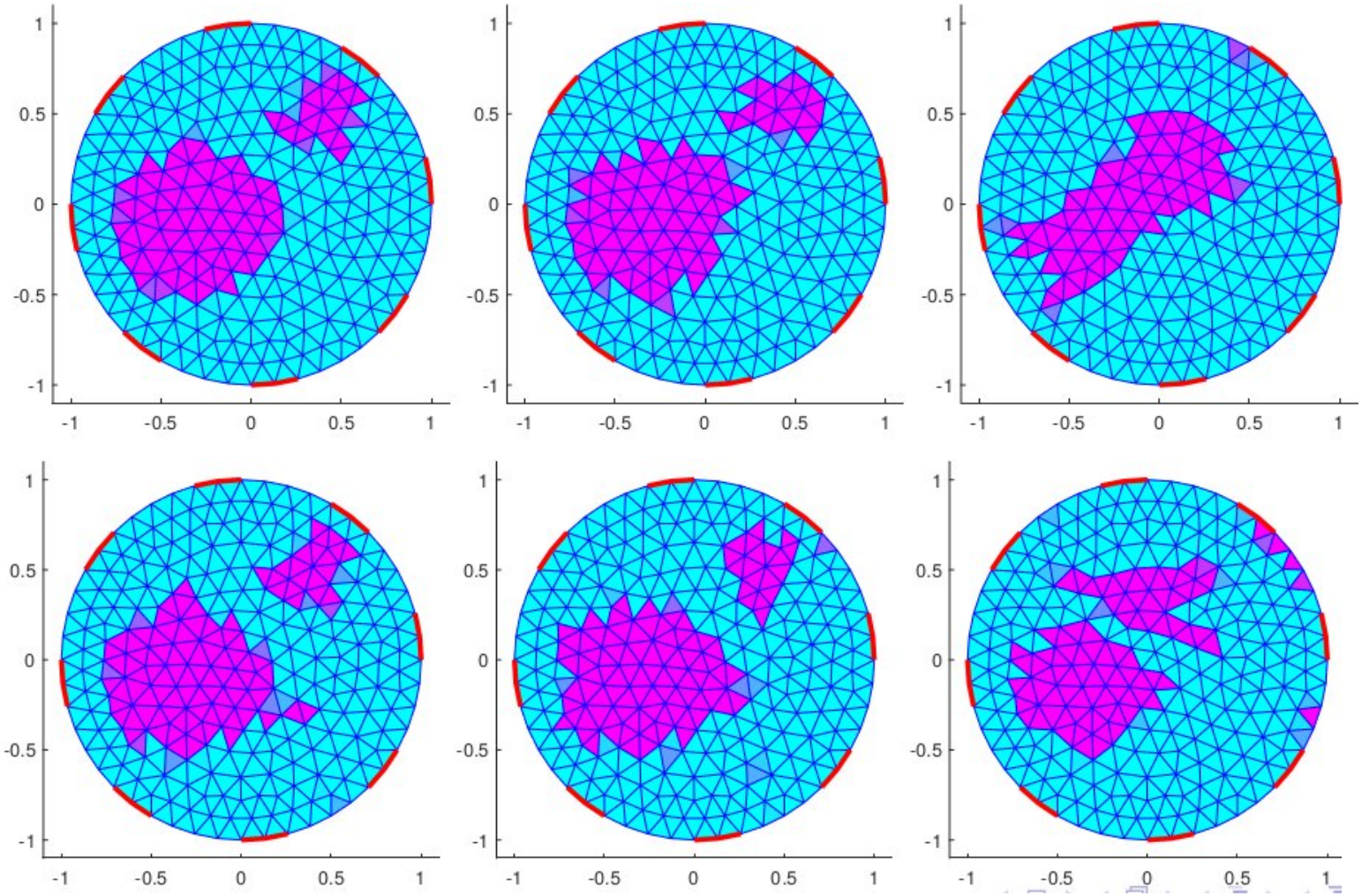
further test cases



EIT: reconstructions of σ^{ex1}, σ^{ex2}

σ^{ex1}:

σ^{ex2}:

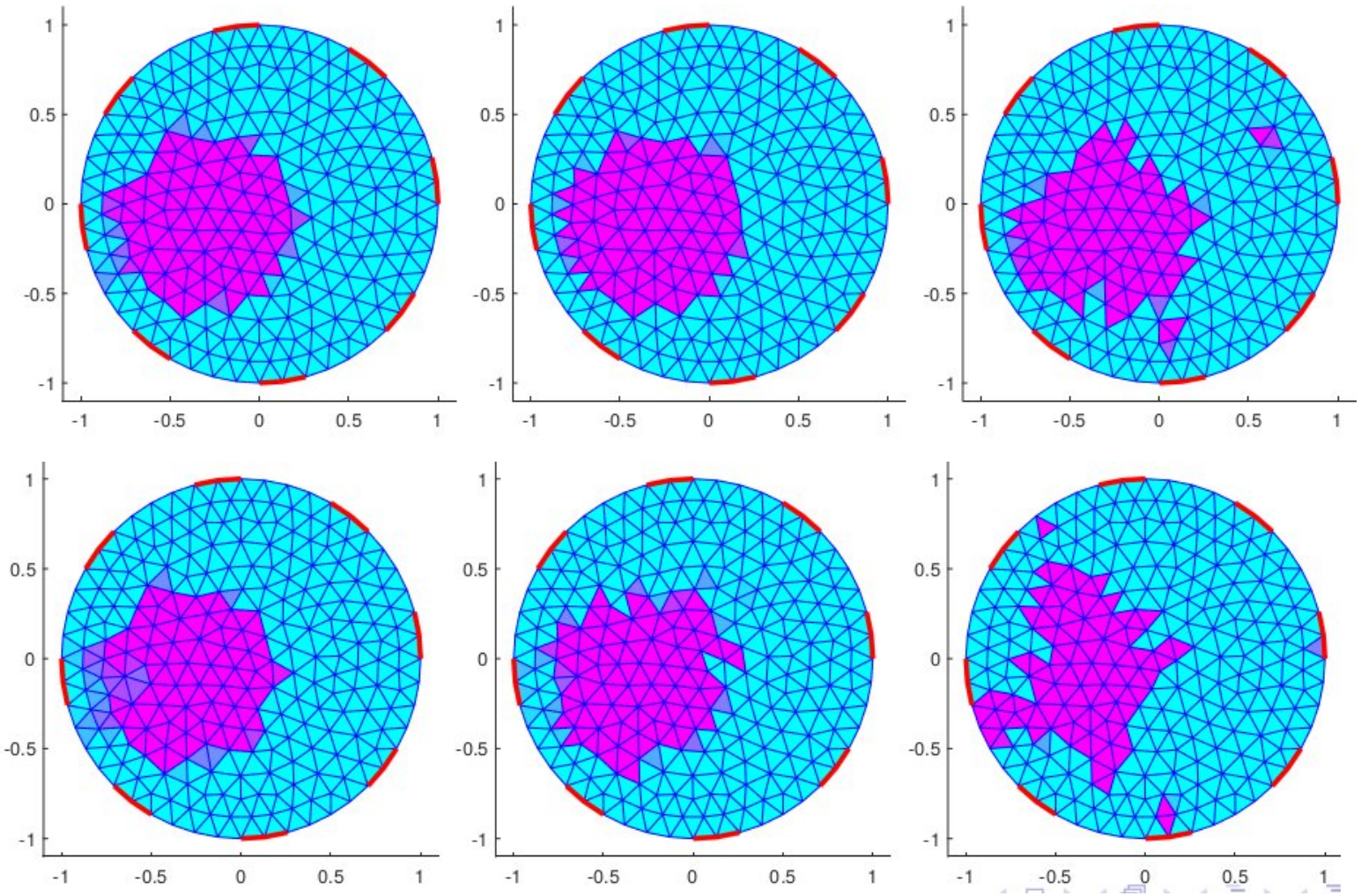


$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

EIT: reconstructions of σ^{ex3}, σ^{ex4}

σ^{ex3}:

σ^{ex4}:

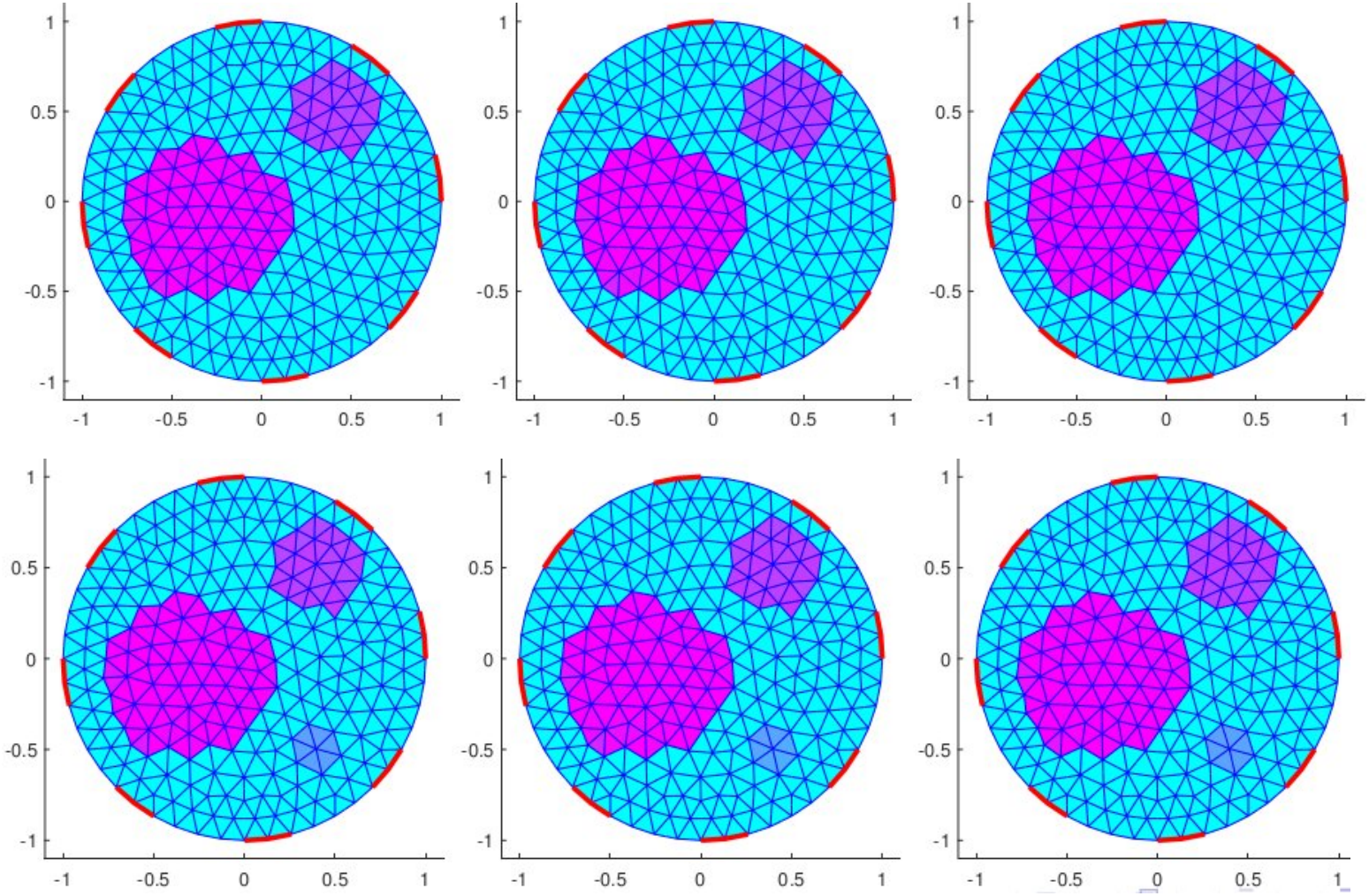


$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: reconstructions of σ^{ex1}, σ^{ex2}

σ^{ex1}:

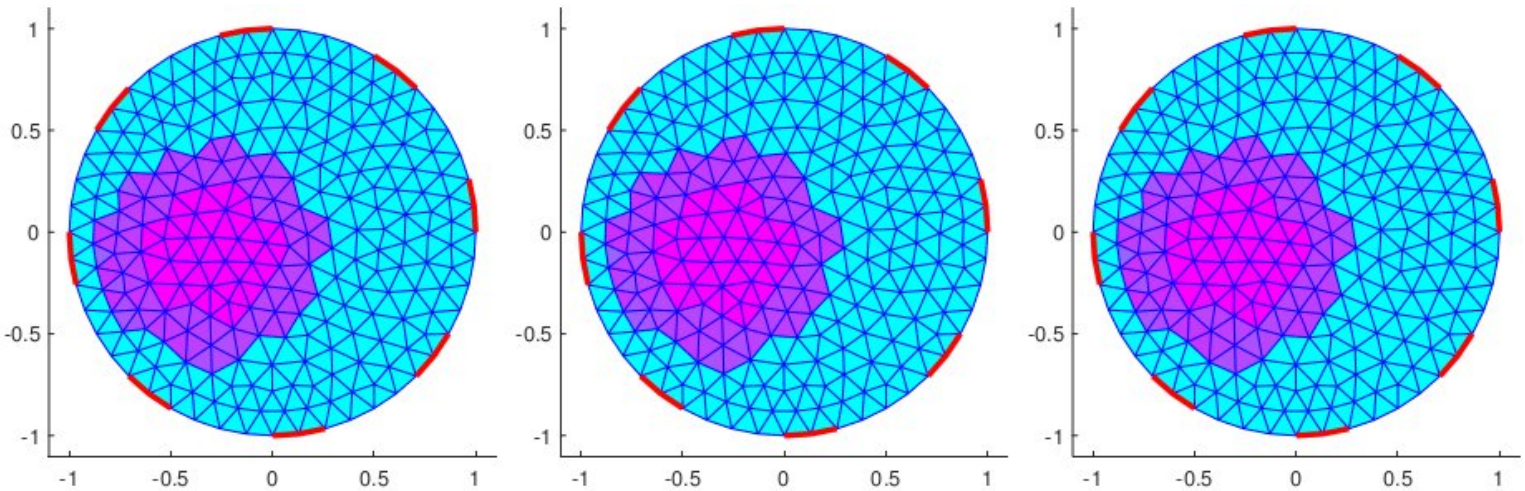
σ^{ex2}:



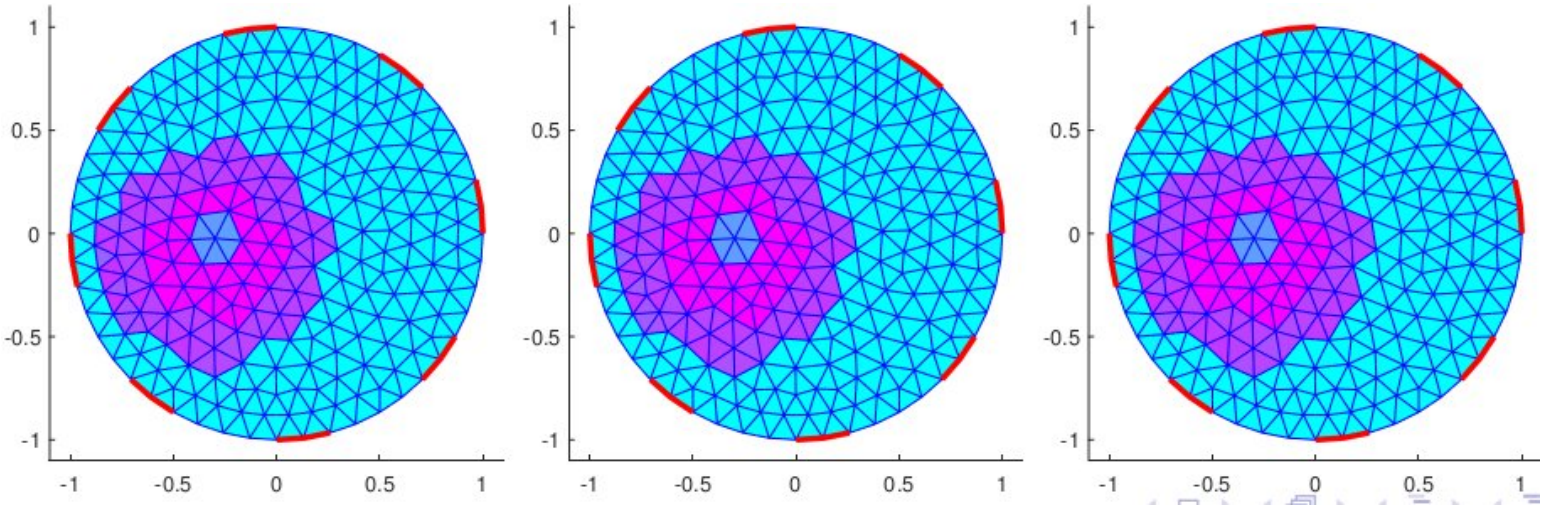
$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

IAT: reconstructions of σ^{ex3}, σ^{ex4}

σ^{ex3}:



σ^{ex4}:



$\delta = 0$ $\delta = 0.01$ $\delta = 0.1$

Conclusions

- all-at-once and minimization based formulations provide more freedom in formulating and regularizing inverse problems

Conclusions

- all-at-once and minimization based formulations provide more freedom in formulating and regularizing inverse problems
- all-at-once gradient type methods never solve PDE models
- all-at-once Newton type methods only solve linear PDE models

Conclusions

- all-at-once and minimization based formulations provide more freedom in formulating and regularizing inverse problems
- all-at-once gradient type methods never solve PDE models
- all-at-once Newton type methods only solve linear PDE models
- further applications: sound source localization, distributed or nonlinear permeabilities in magentostatics, Lamé parameters in elastostatics, cracks...

Thank you for your attention!

- BK. *Regularization based on all-at-once formulations for inverse problems.* SIAM J. Numer. Anal. 54, 2594–2618 (2016)
- BK. *Minimization based formulations of inverse problems and their regularization*. SIAM J. Opt. 28, 620–645, (2018).
- Kha Van Huynh and BK. *Some application examples of minimization based formulations of inverse problems and their regularization.* Inverse Problems and Imaging 15, 415–443 (2021)
- BK and Kha Van Huynh. *Iterative regularization for constrained minimization formulations of nonlinear inverse problems.* Computational Optimization and Applications 81, 569–611 (2021)

Idea of proof I

Nonexpansivity of $P_{\tilde{M}^\delta} \Rightarrow$

$$\begin{aligned}\|x_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 &= \|P_{\tilde{M}^\delta}(\tilde{x}_{k+1}) - P_{\tilde{M}^\delta}(x^\dagger)\|^2 - \|x_k - x^\dagger\|^2 \\ &\leq \|\tilde{x}_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2\end{aligned}$$

Idea of proof I

Nonexpansivity of $P_{\tilde{M}^\delta} \Rightarrow$

$$\begin{aligned}
\|x_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 &= \|P_{\tilde{M}^\delta}(\tilde{x}_{k+1}) - P_{\tilde{M}^\delta}(x^\dagger)\|^2 - \|x_k - x^\dagger\|^2 \\
&\leq \|\tilde{x}_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 \\
&= \|\tilde{x}_{k+1} - x_k\|^2 + 2\langle \tilde{x}_{k+1} - x_k, x_k - x^\dagger \rangle \\
&= \mu_k^2 \|\nabla J^\delta(x_k)\|^2 - 2\mu_k \langle \nabla J^\delta(x_k), x_k - x^\dagger \rangle \,.
\end{aligned}$$

Idea of proof I

Nonexpansivity of $P_{\tilde{M}^\delta} \Rightarrow$

$$\begin{aligned}
\|x_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 &= \|P_{\tilde{M}^\delta}(\tilde{x}_{k+1}) - P_{\tilde{M}^\delta}(x^\dagger)\|^2 - \|x_k - x^\dagger\|^2 \\
&\leq \|\tilde{x}_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 \\
&= \|\tilde{x}_{k+1} - x_k\|^2 + 2\langle \tilde{x}_{k+1} - x_k, x_k - x^\dagger\rangle \\
&= \mu_k^2 \|\nabla J^\delta(x_k)\|^2 - 2\mu_k \langle \nabla J^\delta(x_k), x_k - x^\dagger\rangle \,.
\end{aligned}$$

Under combined convexity and approximate stationarity condition

$$\langle \nabla J^\delta(x), x - x^\dagger\rangle \geq \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

which for $k < k_* = \min\{k \,:\, \|\nabla J^\delta(x_k)\|^2 \leq \tau\eta(\delta)\}$ implies

$$\begin{aligned}
&(\gamma\tau - 1)\eta(\delta) \leq \langle \nabla J^\delta(x_k), x_k - x^\dagger\rangle, \\
&(1 + \tfrac{1}{\gamma\tau - 1})\langle \nabla J^\delta(x_k), x_k - x^\dagger\rangle \geq \gamma \|\nabla J^\delta(x_k)\|^2 \,,
\end{aligned}$$

Idea of proof I

Nonexpansivity of $P_{\tilde{M}^\delta} \Rightarrow$

$$\begin{aligned}\|x_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 &= \|P_{\tilde{M}^\delta}(\tilde{x}_{k+1}) - P_{\tilde{M}^\delta}(x^\dagger)\|^2 - \|x_k - x^\dagger\|^2 \\ &\le \|\tilde{x}_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 \\ &= \|\tilde{x}_{k+1} - x_k\|^2 + 2\langle \tilde{x}_{k+1} - x_k, x_k - x^\dagger\rangle \\ &= \mu_k^2 \|\nabla J^\delta(x_k)\|^2 - 2\mu_k \langle \nabla J^\delta(x_k), x_k - x^\dagger\rangle\,.\end{aligned}$$

Under combined convexity and approximate stationarity condition

$$\langle \nabla J^\delta(x), x - x^\dagger\rangle \ge \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

which for $k < k_* = \min\{k \,:\, \|\nabla J^\delta(x_k)\|^2 \le \tau\eta(\delta)\}$ implies

$$\begin{aligned}&(\gamma\tau - 1)\eta(\delta) \le \langle \nabla J^\delta(x_k), x_k - x^\dagger\rangle, \\ &(1 + \tfrac{1}{\gamma\tau - 1})\langle \nabla J^\delta(x_k), x_k - x^\dagger\rangle \ge \gamma \|\nabla J^\delta(x_k)\|^2\,,\end{aligned}$$

$\tau > \frac{1}{\gamma}$, $0 < \underline{\mu} \le \mu_k \le \overline{\mu} < \frac{2\gamma(\gamma\tau - 1)}{\gamma\tau}$, we get monotonicity of the error

$$\begin{aligned}\|x_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 &\le -\mu_k \tilde{C} \langle \nabla J^\delta(x_k), x_k - x^\dagger\rangle \\ &\le -\mu_k C \|\nabla J^\delta(x_k)\|^2 \le 0\end{aligned}$$

Idea of proof II

monotonicity estimate

$$\begin{aligned}\|x_{k+1} - x^\dagger\|^2 - \|x_k - x^\dagger\|^2 &\leq -\mu_k \tilde{C} \langle \nabla J^\delta(x_k), x_k - x^\dagger \rangle \\ &\leq -\mu_k C \|\nabla J^\delta(x_k)\|^2 \leq 0\end{aligned}$$

implies summability of residuals

$$\sum_{k=0}^{k_*} \langle \nabla J^\delta(x_k), x_k - x^\dagger \rangle \leq \frac{1}{\underline{\mu} C} \|x_0 - x^\dagger\|^2 .$$

$$\sum_{k=0}^{k_*} \|\nabla J^\delta(x_k)\|^2 \leq \frac{1}{\underline{\mu} \tilde{C}} \|x_0 - x^\dagger\|^2 .$$

Convexity Condition

special case $J^\delta(x) = \frac{1}{2}\|F(x) - y^\delta\|^2$: Condition

$$\langle \nabla J^\delta(x), x - x^\dagger \rangle \geq \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

becomes

$$\langle F(x) - y^\delta, F'(x)(x - x^\dagger) \rangle \geq \gamma \|F'(x)^*(F(x) - y^\delta)\|^2 - \eta(\delta) \quad (4)$$

Convexity Condition

special case $J^\delta(x) = \frac{1}{2}\|F(x) - y^\delta\|^2$: Condition

$$\langle \nabla J^\delta(x), x - x^\dagger \rangle \geq \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

becomes

$$\langle F(x) - y^\delta, F'(x)(x - x^\dagger) \rangle \geq \gamma \|F'(x)^*(F(x) - y^\delta)\|^2 - \eta(\delta) \quad (4)$$

which follows, e.g., from the weak tangential cone and boundedness conditions

$\|F'(x)\| \leq 1$ and

$\langle F(x) - F(x^\dagger) - F'(x)(x - x^\dagger), F(x) - y^\delta \rangle \leq (1 - \gamma - \kappa)\|F(x) - y^\delta\|^2$

with $\|F(x^\dagger) - y^\delta\|^2 \leq 4\kappa\eta(\delta)$.

Convexity Condition

special case $J^\delta(x) = \frac{1}{2}\|F(x) - y^\delta\|^2$: Condition

$$\langle \nabla J^\delta(x), x - x^\dagger \rangle \geq \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

becomes

$$\langle F(x) - y^\delta, F'(x)(x - x^\dagger) \rangle \geq \gamma \|F'(x)^*(F(x) - y^\delta)\|^2 - \eta(\delta) \quad (4)$$

which follows, e.g., from the weak tangential cone and boundedness conditions

$\|F'(x)\| \leq 1$ and

$\langle F(x) - F(x^\dagger) - F'(x)(x - x^\dagger), F(x) - y^\delta \rangle \leq (1 - \gamma - \kappa)\|F(x) - y^\delta\|^2$

with $\|F(x^\dagger) - y^\delta\|^2 \leq 4\kappa\eta(\delta)$.

cf. normalization and tangential cone conditions for Landweber iteration, see, e.g., [Hanke&Neubauer&Scherzer 1995].

Convexity Condition

special case $J^\delta(x) = \frac{1}{2}\|F(x) - y^\delta\|^2$: Condition

$$\langle \nabla J^\delta(x), x - x^\dagger \rangle \geq \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

becomes

$$\langle F(x) - y^\delta, F'(x)(x - x^\dagger) \rangle \geq \gamma \|F'(x)^*(F(x) - y^\delta)\|^2 - \eta(\delta) \quad (4)$$

which follows, e.g., from the weak tangential cone and boundedness conditions

$\|F'(x)\| \leq 1$ and

$\langle F(x) - F(x^\dagger) - F'(x)(x - x^\dagger), F(x) - y^\delta \rangle \leq (1 - \gamma - \kappa)\|F(x) - y^\delta\|^2$

with $\|F(x^\dagger) - y^\delta\|^2 \leq 4\kappa\eta(\delta)$.

cf. normalization and tangential cone conditions for Landweber iteration, see, e.g., [Hanke&Neubauer&Scherzer 1995].

For more general conditions on gradiend descent regularization methods see [Kindermann 2017].

Convexity Condition

special case $J^\delta(x) = \frac{1}{2}\|F(x) - y^\delta\|^2$: Condition

$$\langle \nabla J^\delta(x), x - x^\dagger \rangle \geq \gamma \|\nabla J^\delta(x)\|^2 - \eta(\delta) \quad \text{for all } x \in \tilde{M}^\delta$$

becomes

$$\langle F(x) - y^\delta, F'(x)(x - x^\dagger) \rangle \geq \gamma \|F'(x)^*(F(x) - y^\delta)\|^2 - \eta(\delta) \quad (4)$$

which follows, e.g., from the weak tangential cone and boundedness conditions

$\|F'(x)\| \leq 1$ and

$\langle F(x) - F(x^\dagger) - F'(x)(x - x^\dagger), F(x) - y^\delta \rangle \leq (1 - \gamma - \kappa)\|F(x) - y^\delta\|^2$

with $\|F(x^\dagger) - y^\delta\|^2 \leq 4\kappa\eta(\delta)$.

cf. normalization and tangential cone conditions for Landweber iteration, see, e.g., [Hanke&Neubauer&Scherzer 1995].

For more general conditions on gradiend descent regularization methods see [Kindermann 2017].

This applies to both reduced $F(q) = CS(q)$

and all-at-once $F(q, u) = (A(q, u), Cu)$ type formulation.

Nonlinearity Restriction

$$\underline{a}J^\delta(x_+) - \underline{b}J^\delta(x) \le G^\delta(x)(x_+ - x) + \tfrac{1}{2}H^\delta(x)(x_+ - x)^2 \le \overline{a}J^\delta(x_+) - \overline{b}J^\delta(x)$$

$(*)$ for all $x, x_+ \in \tilde{M}^\delta\,, \quad \delta \in (0, \bar{\delta})\,,$

with $\underline{a}, \underline{b}, \overline{a}, \overline{b} \ge 0$;
motivated by (with equality in case of quadratic J^δ)

$$G^\delta(x)(x_+ - x) + \tfrac{1}{2}H^\delta(x)(x_+ - x)^2 \approx J^\delta(x_+) - J^\delta(x)\,,$$

Nonlinearity Restriction

$$\underline{a}J^\delta(x_+) - \underline{b}J^\delta(x) \le G^\delta(x)(x_+ - x) + \tfrac{1}{2}H^\delta(x)(x_+ - x)^2 \le \overline{a}J^\delta(x_+) - \overline{b}J^\delta(x)$$

$(*)$ for all $x, x_+ \in \tilde{M}^\delta\,, \quad \delta \in (0, \bar{\delta})\,,$

with $\underline{a}, \underline{b}, \overline{a}, \overline{b} \ge 0$;
motivated by (with equality in case of quadratic J^δ)

$$G^\delta(x)(x_+ - x) + \tfrac{1}{2}H^\delta(x)(x_+ - x)^2 \approx J^\delta(x_+) - J^\delta(x)\,,$$

sufficient for $(*)$ (with $\underline{a} = 1 - \tilde{c}$, $\underline{b} = 1 + \tilde{c}$, $\overline{a} = 1 + \tilde{c}$, $\overline{b} = 1 - \tilde{c}$) is

$$|J^\delta(x_+) - J^\delta(x) - G^\delta(x)(x_+ - x) - \tfrac{1}{2}H^\delta(x)(x_+ - x)^2| \le \tilde{c}(J^\delta(x_+) + J^\delta(x))$$

$(**)$ for all $x, x_+ \in \tilde{M}^\delta\,, \quad \delta \in (0, \bar{\delta})\,,$

Nonlinearity Restriction

$$\underline{a}J^\delta(x_+) - \underline{b}J^\delta(x) \le G^\delta(x)(x_+ - x) + \tfrac{1}{2}H^\delta(x)(x_+ - x)^2 \le \overline{a}J^\delta(x_+) - \overline{b}J^\delta(x)$$

$(*)$ for all $x, x_+ \in \tilde{M}^\delta\,, \quad \delta \in (0, \bar{\delta})\,,$

with $\underline{a}, \underline{b}, \overline{a}, \overline{b} \ge 0$;
motivated by (with equality in case of quadratic J^δ)

$$G^\delta(x)(x_+ - x) + \tfrac{1}{2}H^\delta(x)(x_+ - x)^2 \approx J^\delta(x_+) - J^\delta(x)\,,$$

sufficient for $(*)$ (with $\underline{a} = 1 - \tilde{c}$, $\underline{b} = 1 + \tilde{c}$, $\overline{a} = 1 + \tilde{c}$, $\overline{b} = 1 - \tilde{c}$) is

$$|J^\delta(x_+) - J^\delta(x) - G^\delta(x)(x_+ - x) - \tfrac{1}{2}H^\delta(x)(x_+ - x)^2| \le \tilde{c}(J^\delta(x_+) + J^\delta(x))$$

$(**)$ for all $x, x_+ \in \tilde{M}^\delta\,, \quad \delta \in (0, \bar{\delta})\,,$

sufficient for $(**)$ in special case $J^\delta(x) = \frac{1}{2}\|F(x) - y^\delta\|^2$, X Hilbert space, is

$$|\langle F(x_+) - F(x) - F'(x)(x_+ - x), F(x) - y^\delta\rangle| \le c_{tc}\|F(x_+) - F(x)\|\,\|F(x) - y^\delta\|$$

for all $x, x_+ \in \tilde{M}^\delta$, $\delta \in (0, \bar{\delta})$,