Model order reduction for the numerical solution of diffusive inverse problems

Alexander V. Mamonov[1],
Liliana Borcea[2], Vladimir Druskin[3],
Fernando Guevara Vasquez[4] and Mikhail Zaslavsky[3]

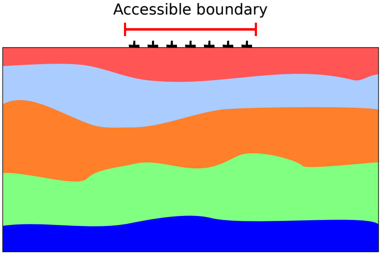
[1]University of Houston, [2]University of Michigan, Ann-Arbor,
[3]Schlumberger-Doll Research Center, [4]University of Utah

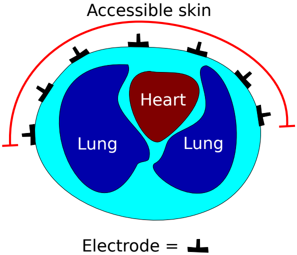
July 14, 2015

Outline

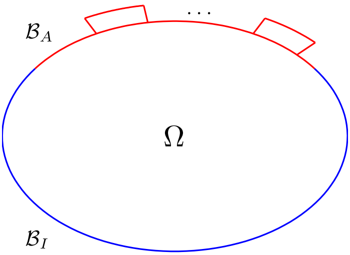
Diffusive inverse problems: motivation

- **General formulation**: determine electrical conductivity inside an object from the electromagnetic excitations and measurements on its boundary
- **Controlled Source Electromagnetic Method (CSEM)**: low frequency EM leads to a parabolic PDE approximation of Maxwell's equations
- **Electrical Impedance Tomography (EIT)**: zero frequency (direct current) leads to an elliptic equation for the potential





Problem formulation: EIT



- Two-dimensional problem $\Omega \subset \mathbb{R}^2$, possibly with partial data
- Equation for electric potential u

$$\nabla \cdot (\sigma \nabla u) = 0, \quad \text{in } \Omega$$

- Dirichlet data $u|_{\mathcal{B}} = \phi$ on $\mathcal{B} = \partial\Omega$
- Dirichlet-to-Neumann (DtN) map $\Lambda_\sigma : H^{1/2}(\mathcal{B}) \to H^{-1/2}(\mathcal{B})$

$$\Lambda_\sigma \phi = \sigma \frac{\partial u}{\partial \nu}\Big|_{\mathcal{B}}$$

Partial data:

- Split the boundary $\mathcal{B} = \mathcal{B}_A \cup \mathcal{B}_I$, accessible $\mathcal{B}_A$, inaccessible $\mathcal{B}_I$
- Similarly to the full DtN map define the partial map

$$\widetilde{\Lambda}_\sigma \widetilde{\phi} = \left(\Lambda_\sigma \widetilde{\phi} \right)\Big|_{\mathcal{B}_A}, \text{ where supp } \widetilde{\phi} \subset \mathcal{B}_A$$

- **Partial data EIT**: find σ given the map $\widetilde{\Lambda}_\sigma$

Problem formulation: CSEM

- Time-dependent diffusion equation for the potential u:

$$u_t = \nabla \cdot (\sigma \nabla u), \quad \text{in } \Omega, \quad t > 0$$

- Also a partial data setting: $\mathcal{B} = \partial\Omega = \mathcal{B}_A \cup \mathcal{B}_I$
- Boundary conditions

$$u|_{\mathcal{B}_I} = 0, \quad \left.\frac{\partial u}{\partial \nu}\right|_{\mathcal{B}_A} = 0$$

- Initial conditions

$$u(x, 0) = \int_{\mathcal{B}_A} \phi(z)\delta(x - z) dS_z, \quad x \in \Omega \cup \mathcal{B}$$

- Measurements $y_\sigma(x, t) = u(x, t)$ for $x \in \mathcal{B}_A$, $t > 0$
- **Partial data CSEM**: find σ given $y_\sigma(x, t)$ for $x \in \mathcal{B}_A$, $t > 0$

Diffusive inversion stability and optimization

- Both elliptic (EIT) and parabolic (CSEM) inverse problems with boundary data are **ill-posed** due to the **instability**
- At most **logarithmic** stability can be achieved under certain regularity assumptions

$$\|\sigma_1 - \sigma_2\|_\infty \leq C \left|\log \|d_{\sigma_1} - d_{\sigma_2}\|_{\mathcal{B}_A}\right|^{-a},$$

  where the data $d_\sigma = \widetilde{\Lambda}_\sigma$ for EIT and $d_\sigma = y_\sigma$ for CSEM
- Exponential ill-conditioning of any discretization
- Resolution is severely limited by the noise, regularization is required
- Conventional solution method: **non-linear output least squares** (OLS) minimization

$$\underset{\sigma}{\text{minimize}} \, \|d^\star - d_\sigma\|_2^2 + \mu \mathcal{P}(\sigma), \qquad (1)$$

  where $d^\star$ is the measured data, $\mathcal{P}$ is a penalty functional and μ is a penalty parameter
- Due to ill-conditioning (1) is hard to solve, the misfit functional is non-convex, large μ may be needed, convergence is slow

Reduced order models for inversion

- In practice a finite number n of data measurements is taken $\mathcal{M}_n(d_\sigma)$
- Our approach is based on constructing a **reduced order model** (ROM) of size related to n that fits the measured data **exactly**

$$M_n(\gamma) = \mathcal{M}_n(d_\sigma),$$

  here $M_n(\gamma)$ is the discrete response of the ROM parametrized by γ
- The parameters γ are chosen in such way that the mapping

$$\mathcal{Q} : \sigma \to d_\sigma \to \mathcal{M}_n(d_\sigma) \to M_n(\gamma) \to \gamma$$

  is an **approximate identity**
- The optimization problem (1) is replaced by

$$\underset{\sigma}{\text{minimize}} \, \|\gamma^\star - \mathcal{Q}(\sigma)\|_2^2 + \mu \mathcal{P}(\sigma), \tag{2}$$

  where $\gamma^\star$ is computed from data interpolation $M_n(\gamma^\star) = \mathcal{M}_n(d^\star)$
- Since $\mathcal{Q}$ approximates identity, the misfit functional in (2) is close to **quadratic** and thus **convex**, easy to minimize

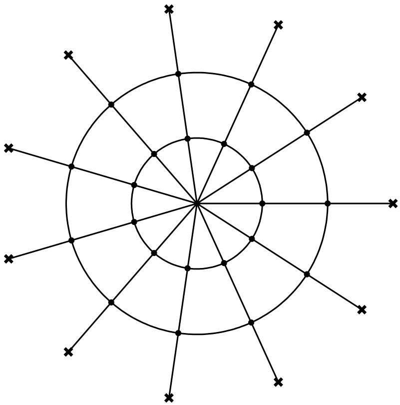
Features of inversion with ROMs

- In practice often a **single Gauss-Newton** iteration is enough to obtain quality reconstructions of σ
- Unlike conventional OLS approach regularization is not required for convergence, but can be added to incorporate prior information about σ
- Optimization (2) is a well-posed problem
- Where did the ill-posedness go?
- It is in the computation of the data fit

$$M_n(\gamma^\star) = \mathcal{M}_n(d^\star)$$

  where we assume that $M_n(\gamma^\star)$ can be inverted for $\gamma^\star$, i.e. we know how to solve the **discrete inverse problem**
- Discrete inversion typically takes a form of **rational interpolation**
- Instability of data fitting is controlled by limiting n
- Also, images can be obtained from ROM parameters $\gamma^\star$ **directly** without optimization using the **optimal grids**

Resistor networks for EIT



Circular planar graph with $n = |\mathcal{B}| = 11$ boundary nodes shown as $\times$

- Appropriate ROMs for EIT in 2D are **resistor networks** with **circular planar graphs**
- Network is a graph $(\mathcal{V}, \mathcal{E})$ with positive weights γ on the edges $\mathcal{E}$
- Vertices $\mathcal{V}$ are split into interior $\mathcal{I}$ and boundary $\mathcal{B}$
- Graph can be embedded into the unit disk $\mathbb{D}$ so that $\mathcal{B}$ are on $\partial\mathbb{D}$
- Discrete derivative D on a graph defines a **Kirchhoff matrix**

$$K = D^T \operatorname{diag}(\gamma) D$$

- **Discrete DtN** map is a Schur complement

$$M_n(\gamma) = K_{\mathcal{B}\mathcal{B}} - K_{\mathcal{B}\mathcal{I}} K_{\mathcal{I}\mathcal{I}}^{-1} K_{\mathcal{I}\mathcal{B}}$$

Data measurements and fitting

- Data measured with disjoint electrode functions ψ_j, supp$\psi_j \subset \mathcal{B}_A$
- Measurement matrix $\mathcal{M}_n(\widetilde{\Lambda}_\sigma) \in \mathbb{R}^{n \times n}$ given by

$$\left[\mathcal{M}_n(\widetilde{\Lambda}_\sigma)\right]_{k,j} = \int_{\mathcal{B}_A} \psi_k \widetilde{\Lambda}_\sigma \psi_j dS, \quad i \neq j$$

  with the diagonal determined by current conservation
- **Morrow, Ingerman, 1998**: $\mathcal{M}_n(\widetilde{\Lambda}_\sigma)$ has the properties of a DtN map of a resistor network
- Thus $M_n(\gamma^\star) = \mathcal{M}_n(\widetilde{\Lambda}^\star)$ for some network
- **Curtis, Ingerman, Morrow, 1998**: $\gamma^\star$ is uniquely recoverable from $M_n(\gamma^\star)$ iff the network's graph is well-connected and critical
- **Well-connected**: certain subsets of $\mathcal{B}$ can be connected with disjoint paths through the network
- **Critical**: removal of any edge breaks some connection
- Constructive direct method for network recovery: **layer peeling**

Sensitivity analysis, optimal grids and reconstructions

- Why is the mapping $\gamma = \mathcal{Q}(\sigma)$ an approximate identity?
- Can be studied by considering the **sensitivity functions**

$$\left[\frac{\delta \mathcal{Q}}{\delta \sigma}\right]_k = \left[\left(\frac{\partial M_n(\gamma)}{\partial \gamma}\right)^{-1} \mathcal{M}_n\left(\frac{\delta \widetilde{\Lambda}_\sigma}{\delta \sigma}\right)\right]_k,$$

  where $M_n(\gamma) = \mathcal{M}_n(\widetilde{\Lambda}_\sigma)$
- Sensitivity functions of resistor networks are **localized**
- Roughly, γ_k is an average of σ near the **optimal grid** node

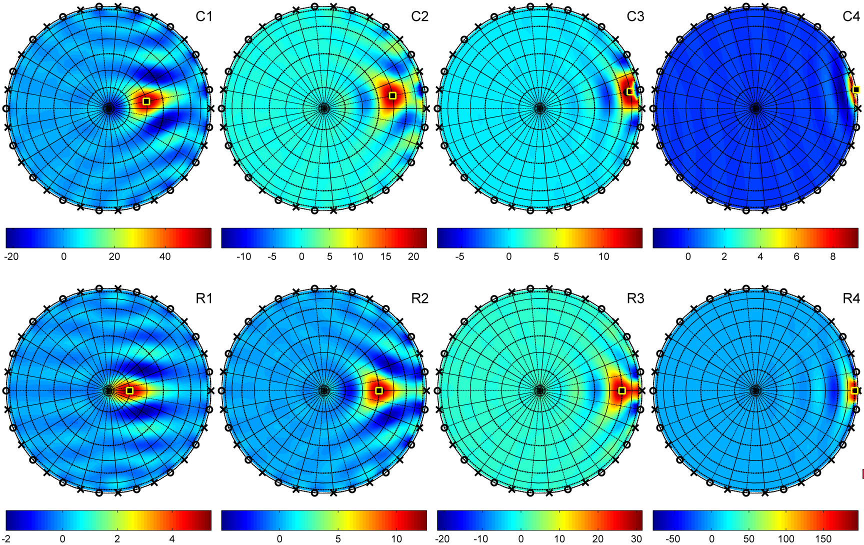
$$x_k = \text{argmax} \left[\frac{\delta \mathcal{Q}}{\delta \sigma}\right]_k$$

- Thus, γ_k may be used to define an interpolated (e.g. piecewise linear) **reconstruction** on the optimal grid

$$\sigma(x_k) \approx \frac{\gamma_k}{\gamma_k^{(1)}},$$

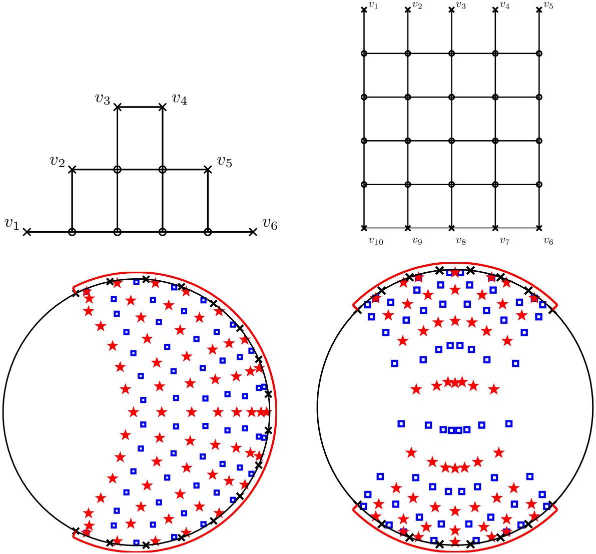
  where $\gamma^{(1)} = \mathcal{Q}(1)$, i.e. resistors computed for $\sigma^{(1)} \equiv 1$

Sensitivity functions



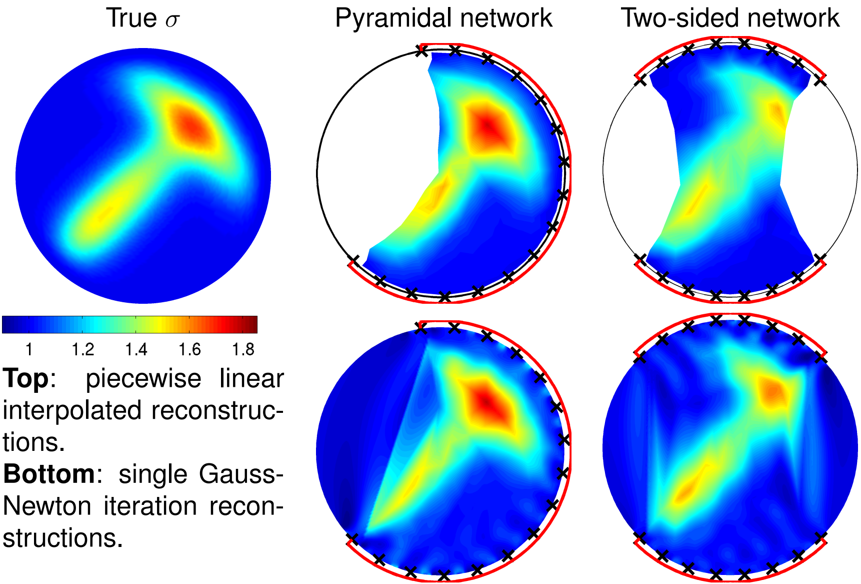
Network topologies and optimal grids

Pyramidal network | Two-sided network



- Circular planar networks do not have to look circular
- Other topologies are better suited for partial data problem
- **Pyramidal**: if $\mathcal{B}_A$ is simply connected
- **Two-sided**: if $\mathcal{B}_A$ is doubly connected
- Both are well-connected and critical
- **Top**: network topology; **Bottom**: optimal grid.

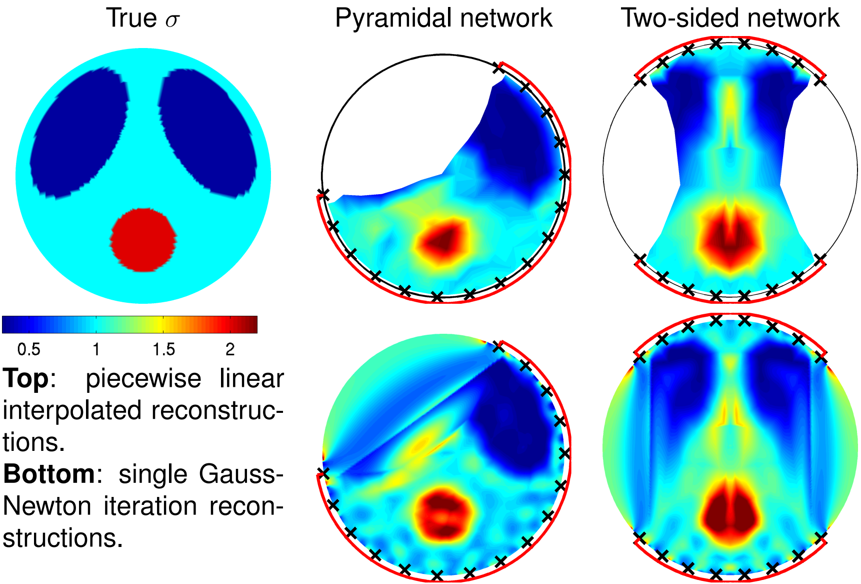
Reconstructions: smooth σ, n=16



Top: piecewise linear interpolated reconstructions.

Bottom: single Gauss-Newton iteration reconstructions.

Reconstructions: piecewise constant σ, n=16



Top: piecewise linear interpolated reconstructions.
Bottom: single Gauss-Newton iteration reconstructions.

Single measurement CSEM

- Recall the CSEM equation

$$u_t = \nabla \cdot (\sigma \nabla u) = A_\sigma u, \quad \text{in } \Omega, \quad t > 0$$

with boundary conditions

$$u|_{\mathcal{B}_I} = 0, \quad \left. \frac{\partial u}{\partial \nu} \right|_{\mathcal{B}_A} = 0$$

and an initial condition

$$u(x,0) = \int_{\mathcal{B}_A} \phi(z)\delta(x-z)dS_z = \int_\Omega b(z)\delta(x-z)dz, \quad x \in \Omega \cup \mathcal{B},$$

with a **transducer** function $b(z)$ satisfying $\operatorname{supp} b \subseteq \mathcal{B}_A$

- Let us consider a **single measurement**

$$y_\sigma(t) = \int_\Omega b(z)u(z,t)dz$$

Projection-based model order reduction

- Define the **transfer function** via Laplace transform

$$g_\sigma(s) = \int_0^{+\infty} y_\sigma(t) e^{-st} dt = b^*(sI - A_\sigma)^{-1} b, \quad s > 0$$

- Transfer function of a reduced model $A_n \in \mathbb{R}^{n \times n}$, $b_n \in \mathbb{R}^n$

$$g_n(s) = b_n^*(sI_n - A_n)^{-1} b_n$$

- Projection-based model reduction

$$A_n = V^* A_\sigma V, \quad b_n = V^* b, \quad V^* V = I_n$$

- The n “columns” of V span the projection **subspace**
- Choice of subspace is determined by **matching conditions**

$$[\mathcal{M}_n(y_\sigma)]_{k,j} = \left. \frac{\partial^k g_\sigma}{\partial s^k} \right|_{s=s_j} = \left. \frac{\partial^k g_n}{\partial s^k} \right|_{s=s_j}, \quad j = 1, \ldots, m, \quad k = 1, \ldots, 2k_j - 1$$

at **interpolation nodes** $s_j \in [0, +\infty)$ with

$$n = \sum_{j=1}^{m} k_j$$

Rational Krylov model order reduction

- Partial fraction expansion

$$g_n(s) = \sum_{j=1}^{n} \frac{c_j}{s+\theta_j}, \quad c_j > 0, \quad \theta_j > 0,$$

with negative **poles** $-\theta_j$ and positive **residues** c_j

- Rational g_n, hence **rational interpolation**
- Typical choices of projection subspaces in model reduction: rational **Krylov** subspaces

$$\mathcal{K}_n(\mathbf{s}) = \text{span}\left\{(s_j I - A_\sigma)^{-k} b \mid j = 1, \ldots, m;\ k = 1, \ldots, k_j\right\}$$

- Popular special cases for **forward** modeling: moment matching

$$\begin{aligned} \mathcal{K}_n(+\infty) &= \text{span}\left\{b, A_\sigma b, \ldots, A_\sigma^{n-1} b\right\} \\ \mathcal{K}_n(0) &= \text{span}\left\{A_\sigma^{-1} b, A_\sigma^{-2} b, \ldots, A_\sigma^{-n} b\right\} \end{aligned}$$

- $K_n(+\infty)$ is bad for inversion

Connection to resistor networks: S-fraction form

- Write the reduced model response as a Stieltjes **continued fraction** (S-fraction)

$$g_n(s;\gamma) = \cfrac{1}{\widehat{\gamma}_1^{-1} s + \cfrac{1}{\gamma_1^{-1} + \cfrac{1}{\ddots + \cfrac{1}{\widehat{\gamma}_n^{-1} s + \gamma_n}}}}$$

- This is a boundary response $w_1(s)$ of a **second-order finite difference** scheme

$$\widehat{\gamma}_j \left(\gamma_j(w_{j+1} - w_j) - \gamma_{j-1}(w_j - w_{j-1})\right) - s w_j = 0$$

- The coefficients $\gamma = \{\gamma_j, \widehat{\gamma}_j\}_{j=1}^n$ are the analogue of the resistor network coefficients
- They are exactly the same for a rotationally symmetric circular network
- Once we have γ we can define

$$[M_n(\gamma)]_{k,j} = \left. \frac{\partial^k g_n(\,\cdot\,;\gamma)}{\partial s^k} \right|_{s=s_j}$$

CSEM with multiple measurements: backscattering

- To deal with multiple measurements consider many **transducer** functions $b^\alpha(z)$, $\alpha = 1, \ldots, p$ with disjoint supports $\operatorname{supp} b^\alpha \subseteq \mathcal{B}_A$
- For each $\alpha = 1, \ldots, p$ perform a **rational interpolation**

$$M_n(\gamma^\alpha) = \mathcal{M}_n(y_\sigma^\alpha)$$

  and express the interpolant $g_n^\alpha(s; \gamma^\alpha)$ as an S-fraction to obtain the coefficients γ^α
- Form a **joint misfit functional** out of all S-fraction coefficients

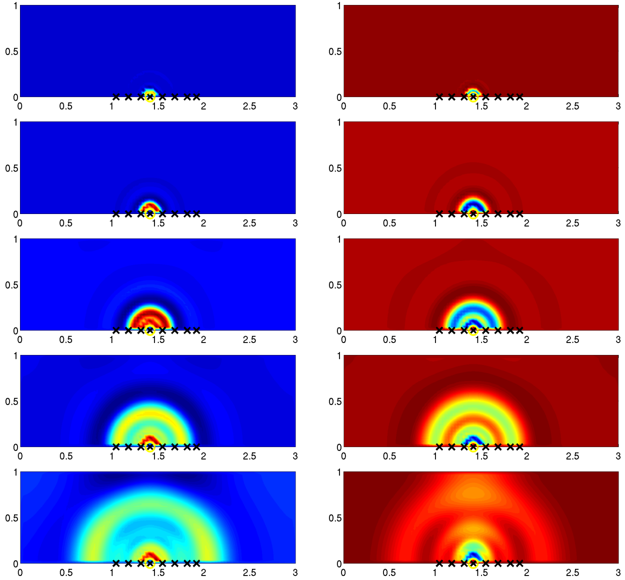
$$\underset{\sigma}{\text{minimize}} \sum_{\alpha=1}^{p} \|\gamma^\alpha - \mathcal{Q}^\alpha(\sigma)\|_2^2 + \mu \mathcal{P}(\sigma),$$

  and solve with (a single step of) Gauss-Newton iteration
- Reminder: the mapping $\mathcal{Q}^\alpha$ is defined as a chain

$$\mathcal{Q}^\alpha : \sigma \to y_\sigma^\alpha \to \mathcal{M}_n(y_\sigma^\alpha) \to M_n(\gamma^\alpha) \to \gamma^\alpha$$

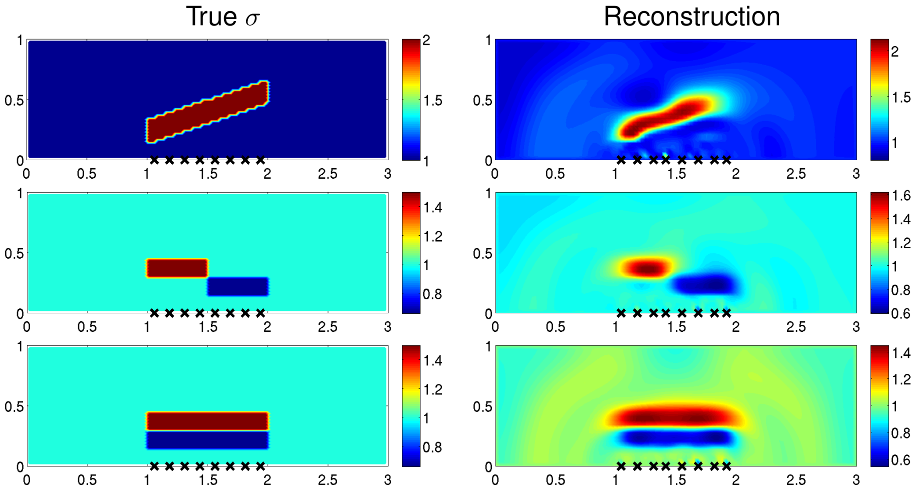
- Similarly to the resistor networks we can consider the **sensitivity functions** $\left[\frac{\partial \mathcal{Q}^\alpha}{\partial \sigma}\right]_j$, $j = 1, \ldots, n$

Sensitivity functions



Sensitivity functions of $\widehat{\gamma}_j^{-1}$ (left) and γ_j^{-1} (right) for $j = 1, \ldots, n$ (top to bottom), $n = 5$ for a single transducer ($\alpha = 4$, yellow $\circ$) out of $p = 8$ (black $\times$). Simple Pade approximant at $s = 60$. Sensitivities resemble propagating spherical waves. Higher s means lower speed of propagation. Should avoid reflections from boundaries.

Reconstructions: piecewise constant σ



Reconstuctions after a single Gauss-Newton iteration with a constant intial guess $\sigma_0 \equiv 1$. Locations of $p = 8$ transducers are black $\times$.

Conclusions and future work

Conclusions:

- A framework of ROM-based inversion for diffusive problems is proposed
- Ill-posed inverse problem is separated into **two stages**: ROM construction and reconstruction from ROM parameters
- The instability is confined to ROM construction, it is controlled by ROM size
- The reconstruction stage is formulated as a stable problem of minimizing the ROM parameter misfit
- The parameters are chosen so that they depend almost linearly on the unknown PDE coefficient
- Thus the ROM parameter misfit minimization is close to quadratic and can be solved with a single step of Gauss-Newton iteration

Future work:

- EIT with resistor networks currently works in 2D or for limited subsets of 3D data, a full 3D approach is yet to be developed
- ROM-based CSEM inversion works in any dimension, but uses only the backscattering data

References

1. *Electrical impedance tomography with resistor networks*, L. Borcea, V. Druskin and F. Guevara Vasquez. Inverse Problems 24(3):035013, 2008.
2. *Circular resistor networks for electrical impedance tomography with partial boundary measurements*, L. Borcea, V. Druskin and A.V. Mamonov, Inverse Problems 26(4):045010, 2010.
3. *Pyramidal resistor networks for electrical impedance tomography with partial boundary measurements*. L. Borcea, V. Druskin, A.V. Mamonov and F. Guevara Vasquez, Inverse Problems 26(10):105009, 2010.
4. *Study of noise effects in electrical impedance tomography with resistor networks*, L. Borcea, F. Guevara Vasquez and A.V. Mamonov, Inverse Problems and Imaging, 7(2):417-443, 2013.
5. *Resistor network approaches to electrical impedance tomography*, L. Borcea, V. Druskin, F. Guevara Vasquez and A.V. Mamonov, Inverse Problems and Applications: Inside Out II, Cambridge University Press, 2012.
6. *A model reduction approach to numerical inversion for a parabolic partial differential equation*, L. Borcea, V. Druskin, A.V. Mamonov and M. Zaslavsky, Inverse Problems 30(12):125011, 2014.