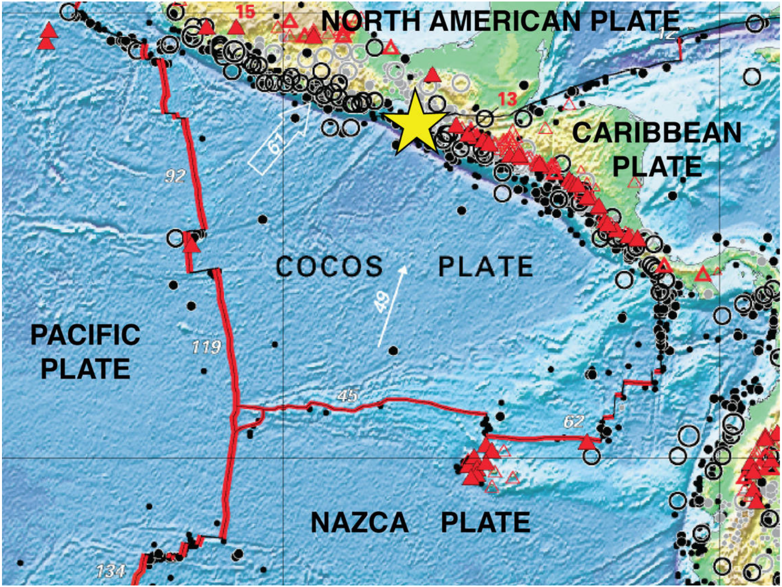
Infinite-dimensional Bayesian inversion for fault slip from surface measurements

Abhijit Chowdhary and Alen Alexanderian

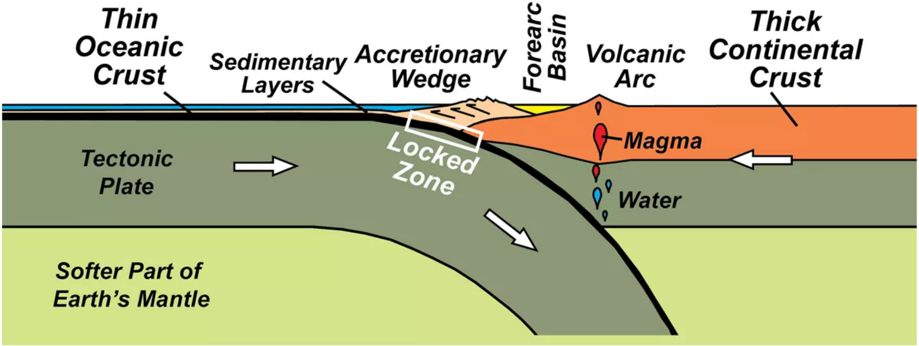
Department of Mathematics
North Carolina State University

April 23, 2022

Work done through NSF-DMS-2111044



Tom Simkin et al. *This dynamic planet: World map of volcanoes, earthquakes, impact craters and plate tectonics*. Tech. rep. 2006. DOI: 10.3133/i2800



Robert J Lillie. *Oregon's Island In The Sky: Geology Road Guide to Marys Peak*. English. OCLC: 979996650. 2017. ISBN: 9781540611963

Goals

Understand the subduction zone from collected observations.

Goals

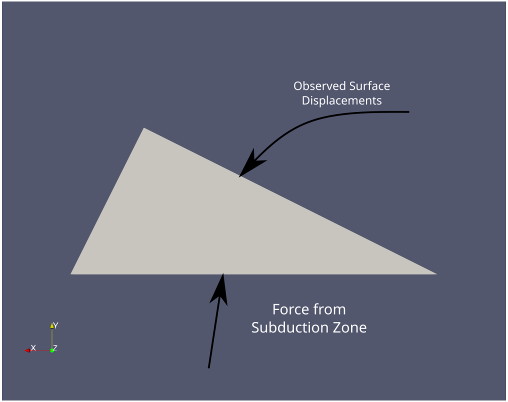
Understand the subduction zone from collected observations.

Do so while quantifying measurement uncertainties.

Table of Contents

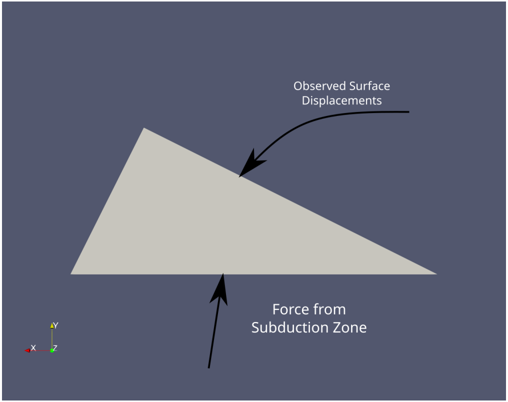
Table of Contents

Model Assumptions



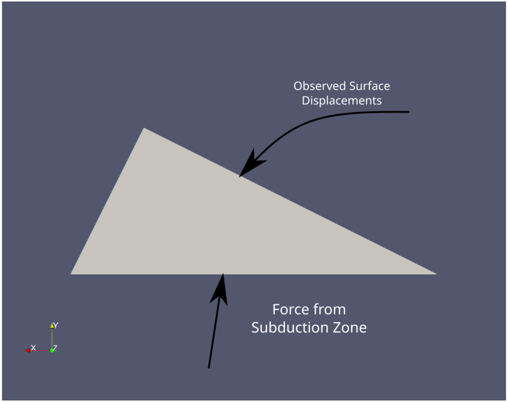
- **Governing PDE** (forward model): Linear elasticity

Model Assumptions



- **Governing PDE** (forward model): Linear elasticity
- **Uncertain parameter**: Displacement on fault plane

Model Assumptions



- **Governing PDE** (forward model): Linear elasticity
- **Uncertain parameter**: Displacement on fault plane
- **Inverse Problem**: Given measurements of surface deformation $\mathbf{u}^{\mathrm{obs}}$ reconstruct fault plane displacement.

Forward Model

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

where:

- $\boldsymbol{\sigma}(\boldsymbol{u}) = \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})$ with
  - $\mathbb{C}[\boldsymbol{\varepsilon}] = 2\mu\boldsymbol{\varepsilon} + \lambda\,\mathrm{tr}(\boldsymbol{\varepsilon})\mathbf{I}$ the fourth-order linear elasticity tensor:
  - $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\left[\boldsymbol{\nabla}\boldsymbol{u} + (\boldsymbol{\nabla}\boldsymbol{u})^{\mathrm{T}}\right]$ the strain tensor.
- μ and λ are known as the Láme constants.

Kimberly Alison McCormack. *Earthquakes, groundwater and surface deformation : Exploring the poroelastic response to megathrust earthquakes*. Aug. 2018. URL: hdl.handle.net/2152/68892

Forward Model

$$
\begin{aligned}
-\boldsymbol{\nabla}\left[\mu(\boldsymbol{\nabla}\boldsymbol{u} + (\boldsymbol{\nabla}\boldsymbol{u})^{\mathrm{T}}) + \lambda\boldsymbol{\nabla}\cdot\boldsymbol{u}\mathbf{I}\right] &= \mathbf{0} \quad \text{in } \Omega, && \text{(1a)}\\
\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n} &= \mathbf{0} \quad \text{on } \Gamma_t && \text{(1b)}\\
\boldsymbol{u} + \beta\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n} &= \boldsymbol{h} \quad \text{on } \Gamma_s && \text{(1c)}\\
\boldsymbol{u}\cdot\boldsymbol{n} &= 0 \quad \text{on } \Gamma_b && \text{(1d)}\\
\delta\mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) + \mathbf{T}\boldsymbol{u} &= \boldsymbol{m} \quad \text{on } \Gamma_b && \text{(1e)}
\end{aligned}
$$

- $\mathbf{T}$ is the tangential operator $\mathbf{T}\boldsymbol{u} = (\mathbf{I} - \boldsymbol{n}\otimes\boldsymbol{n})\boldsymbol{u} = \boldsymbol{u} - (\boldsymbol{n}^{\mathrm{T}}\boldsymbol{u})\boldsymbol{n}$.

Forward Model

$$
\begin{aligned}
-\boldsymbol{\nabla}\left[\mu(\boldsymbol{\nabla}\boldsymbol{u}+(\boldsymbol{\nabla}\boldsymbol{u})^{\mathrm{T}})+\lambda\boldsymbol{\nabla}\cdot\boldsymbol{u}\mathbf{I}\right]&=\mathbf{0} \quad \text{in } \Omega, && \text{(1a)}\\
\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}&=\mathbf{0} \quad \text{on } \Gamma_t && \text{(1b)}\\
\boldsymbol{u}+\beta\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}&=\boldsymbol{h} \quad \text{on } \Gamma_s && \text{(1c)}\\
\boldsymbol{u}\cdot\boldsymbol{n}&=0 \quad \text{on } \Gamma_b && \text{(1d)}\\
\delta\mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n})+\mathbf{T}\boldsymbol{u}&=\boldsymbol{m} \quad \text{on } \Gamma_b && \text{(1e)}
\end{aligned}
$$

- $\mathbf{T}$ is the tangential operator $\mathbf{T}\boldsymbol{u}=(\mathbf{I}-\boldsymbol{n}\otimes\boldsymbol{n})\boldsymbol{u}=\boldsymbol{u}-(\boldsymbol{n}^{\mathrm{T}}\boldsymbol{u})\boldsymbol{n}$.
- The right hand side of (1e) is the displacement on the fault plane that is being inverted for.

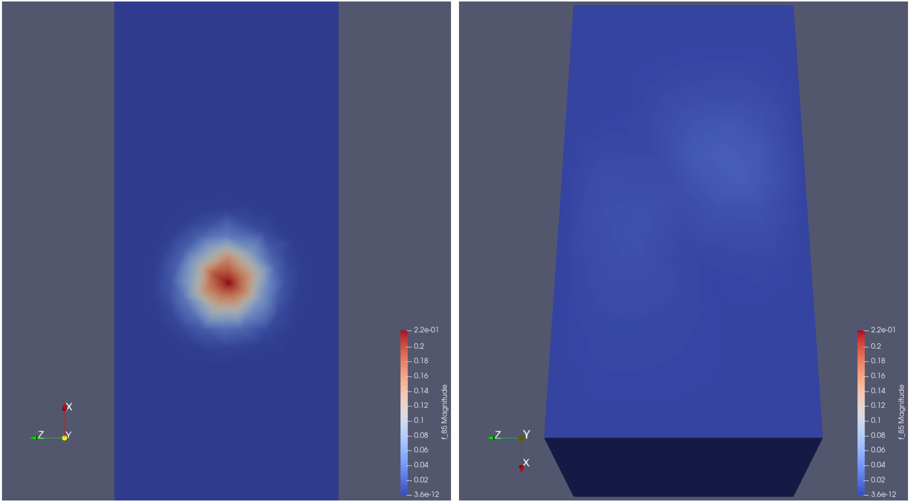
Forward Model

$$
\begin{aligned}
-\boldsymbol{\nabla}\left[\mu(\boldsymbol{\nabla}\boldsymbol{u}+(\boldsymbol{\nabla}\boldsymbol{u})^{\mathrm{T}})+\lambda\boldsymbol{\nabla}\cdot\boldsymbol{u}\mathbf{I}\right]&=\mathbf{0}\quad\text{in }\Omega, &\text{(1a)}\\
\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}&=\mathbf{0}\quad\text{on }\Gamma_t &\text{(1b)}\\
\boldsymbol{u}+\beta\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}&=\boldsymbol{h}\quad\text{on }\Gamma_s &\text{(1c)}\\
\boldsymbol{u}\cdot\boldsymbol{n}&=0\quad\text{on }\Gamma_b &\text{(1d)}\\
\delta\mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n})+\mathbf{T}\boldsymbol{u}&=\boldsymbol{m}\quad\text{on }\Gamma_b &\text{(1e)}
\end{aligned}
$$

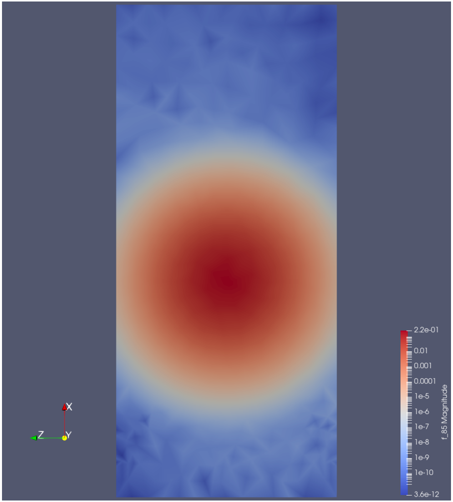
- $\mathbf{T}$ is the tangential operator $\mathbf{T}\boldsymbol{u}=(\mathbf{I}-\boldsymbol{n}\otimes\boldsymbol{n})\boldsymbol{u}=\boldsymbol{u}-(\boldsymbol{n}^{\mathrm{T}}\boldsymbol{u})\boldsymbol{n}$.
- The right hand side of (1e) is the displacement on the fault plane that is being inverted for.
- (1e) can be understood as a regularized Dirichlet condition.

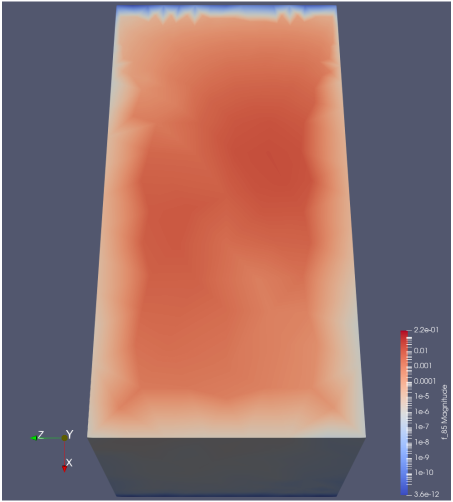
Mesh

Forward Solution



Forward Solution





Weak Formulation

Define:

$$\boldsymbol{V} := \{\boldsymbol{u} \in H^1(\Omega)^3 : \boldsymbol{u} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma_b\}$$

Then the weak formulation of the forward model is given by:

$$\int_{\Gamma_s} \beta^{-1}(\boldsymbol{u} - \boldsymbol{h}) \cdot \boldsymbol{v} \,\mathrm{d}s + \int_{\Gamma_b} \delta^{-1}(\mathbf{T}\boldsymbol{u} - \boldsymbol{m}) \cdot \boldsymbol{v} \,\mathrm{d}s + \int_{\Omega} \boldsymbol{\varepsilon}(\boldsymbol{u}) : \mathbb{C}[\boldsymbol{\varepsilon}(\boldsymbol{v})] \,\mathrm{d}\boldsymbol{x} = 0, \quad \forall v \in \boldsymbol{V} \quad (2)$$

Table of Contents

The Deterministic Inverse Problem

To reconstruct the fault displacement we construct the PDE-constrained optimization problem:

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u}(\boldsymbol{m}) - \mathbf{u}^{\mathrm{obs}}\|^2 + \frac{1}{2}\|\mathcal{A}\boldsymbol{m}\|^2$$

where $\boldsymbol{u}$ is given by the solution of the linear elasticity equation.

- $\boldsymbol{u}(\boldsymbol{m})$ is given by the forward model.
- $\mathcal{B} : (L^2(\Omega))^3 \to \mathbb{R}^N$ is an observation operator.
- $\mathbf{u}^{\mathrm{obs}} \in \mathbb{R}^N$ where N is the number of data points.
- $\mathcal{A}$ is some regularization operator.

[1] In literature, called the *parameter-to-observable operator*.

The Deterministic Inverse Problem

To reconstruct the fault displacement we construct the PDE-constrained optimization problem:

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u}(\boldsymbol{m}) - \mathbf{u}^{\mathrm{obs}}\|^2 + \frac{1}{2}\|\mathcal{A}\boldsymbol{m}\|^2$$

where $\boldsymbol{u}$ is given by the solution of the linear elasticity equation.

- $\boldsymbol{u}(\boldsymbol{m})$ is given by the forward model.
- $\mathcal{B} : (L^2(\Omega))^3 \to \mathbb{R}^N$ is an observation operator.
- $\mathbf{u}^{\mathrm{obs}} \in \mathbb{R}^N$ where N is the number of data points.
- $\mathcal{A}$ is some regularization operator.

If we let $\mathcal{S}$ be the forward model operator, i.e. $\boldsymbol{u} = \mathcal{S}\boldsymbol{m}$, and let[1] $\mathcal{F} = \mathcal{B}\mathcal{S}$, then:

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}\|\mathcal{F}\boldsymbol{m} - \mathbf{u}^{\mathrm{obs}}\|^2 + \frac{1}{2}\|\mathcal{A}\boldsymbol{m}\|^2$$

[1] In literature, called the *parameter-to-observable operator*.

Bayesian Inversion in Finite Dimensions

Theorem (Bayes Theorem in Finite Dimensions)

$$\pi_{\text{post}}(\mathbf{m}|\mathbf{u}^{\text{obs}}) \propto \pi_{\text{like}}(\mathbf{u}^{\text{obs}}|\mathbf{m})\pi_{\text{prior}}(\mathbf{m})$$

- Gaussian Prior $\mathbf{m} \sim \mathcal{N}(\mathbf{m}_{\text{pr}}, \mathbf{\Gamma}_{\text{pr}})$.
- Additive Gaussian noise

$$\mathbf{u}^{\text{obs}} = \mathbf{F}\mathbf{m} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{noise}})$$

- Gaussian posterior with:

$$\mathbf{m}|\mathbf{u}^{\text{obs}} \sim \mathcal{N}(\mathbf{m}_{\text{post}}, \mathbf{\Gamma}_{\text{post}})$$

where:

$$\mathbf{m}_{\text{post}} = \mathbf{\Gamma}_{\text{post}} \left(\mathbf{F}^{\text{T}} \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{u}^{\text{obs}} + \mathbf{\Gamma}_{\text{pr}}^{-1} \mathbf{m}_{\text{pr}} \right)$$

$$\mathbf{\Gamma}_{\text{post}} = \left(\mathbf{F}^{\text{T}} \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{F} + \mathbf{\Gamma}_{\text{pr}}^{-1} \right)^{-1}$$

Bayesian Inversion in Infinite Dimensions

Theorem (Bayes Theorem in Infinite Dimensions)

$$\frac{\mathrm{d}\mu_{\mathrm{post}}^{\mathbf{u}^{\mathrm{obs}}}}{\mathrm{d}\mu_{\mathrm{pr}}} \propto \pi_{\mathrm{like}}(\mathbf{u}^{\mathrm{obs}}|\boldsymbol{m})$$

- Gaussian Prior $\boldsymbol{m} \sim \mathcal{N}(\boldsymbol{m}_{\mathrm{pr}}, \mathcal{C}_0)$.
- Additive Gaussian noise

$$\mathbf{u}^{\mathrm{obs}} = \mathcal{F}\boldsymbol{m} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\mathrm{noise}})$$

- Gaussian posterior:

$$\mu_{\mathrm{post}}^{\mathbf{u}^{\mathrm{obs}}} = \mathcal{N}(\boldsymbol{m}_{\mathrm{post}}, \mathcal{C}_{\mathrm{post}})$$

where:

$$\begin{aligned}
\boldsymbol{m}_{\mathrm{post}} &= \mathcal{C}_{\mathrm{post}} \left(\mathcal{F}^* \boldsymbol{\Gamma}_{\mathrm{noise}}^{-1} \mathbf{u}^{\mathrm{obs}} + \mathcal{C}_{\mathrm{pr}}^{-1} \boldsymbol{m}_{\mathrm{pr}} \right) \\
\mathcal{C}_{\mathrm{post}} &= \left(\mathcal{F}^* \boldsymbol{\Gamma}_{\mathrm{noise}}^{-1} \mathcal{F} + \mathcal{C}_{\mathrm{pr}}^{-1} \right)^{-1}
\end{aligned}$$

Returning to Fault Inversion

Slight Wrinkle: Our parameter-to-observable map is affine.

Returning to Fault Inversion

Slight Wrinkle: Our parameter-to-observable map is affine.

To find $\boldsymbol{m}_{\text{post}}$, need to minimize:

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u} - \mathbf{u}^{\text{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2}\|\mathcal{C}_{\text{pr}}^{-1/2}(\boldsymbol{m} - \boldsymbol{m}_{\text{pr}})\|^2$$

Returning to Fault Inversion

Slight Wrinkle: Our parameter-to-observable map is affine.

To find $\boldsymbol{m}_{\text{post}}$, need to minimize:

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u} - \mathbf{u}^{\text{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2}\|\mathcal{C}_{\text{pr}}^{-1/2}(\boldsymbol{m} - \boldsymbol{m}_{\text{pr}})\|^2$$

Define:

$$\mathcal{L}(\boldsymbol{m}, \boldsymbol{u}, \boldsymbol{v}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u} - \mathbf{u}^{\text{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2}\|\mathcal{C}_{\text{pr}}^{-1/2}(\boldsymbol{m} - \boldsymbol{m}_{\text{pr}})\|^2 + \int_{\Gamma_s} \beta^{-1}(\boldsymbol{u} - \boldsymbol{h}) \cdot \boldsymbol{v}\,\mathrm{d}s + \int_{\Gamma_b} \delta^{-1}(\mathbf{T}\boldsymbol{u} - \boldsymbol{m}) \cdot \boldsymbol{v}\,\mathrm{d}s + \int_{\Omega} \boldsymbol{\varepsilon}(\boldsymbol{u}) : \mathbb{C}[\boldsymbol{\varepsilon}(\boldsymbol{v})]\,\mathrm{d}\boldsymbol{x}$$

Procedure: Obtain $\boldsymbol{m}_{\text{post}}$ by enforcing $\mathcal{L}_{\boldsymbol{m}} = \mathcal{L}_{\boldsymbol{u}} = \mathcal{L}_{\boldsymbol{v}} = 0$.

Returning to Fault Inversion

Slight Wrinkle: Our parameter-to-observable map is affine.

To find $\boldsymbol{m}_{\mathrm{post}}$, need to minimize:

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u} - \mathbf{u}^{\mathrm{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\mathrm{noise}}} + \frac{1}{2}\|\mathcal{C}_{\mathrm{pr}}^{-1/2}(\boldsymbol{m} - \boldsymbol{m}_{\mathrm{pr}})\|^2$$

Define:

$$\mathcal{L}(\boldsymbol{m}, \boldsymbol{u}, \boldsymbol{v}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u} - \mathbf{u}^{\mathrm{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\mathrm{noise}}} + \frac{1}{2}\|\mathcal{C}_{\mathrm{pr}}^{-1/2}(\boldsymbol{m} - \boldsymbol{m}_{\mathrm{pr}})\|^2 + \int_{\Gamma_s} \beta^{-1}(\boldsymbol{u} - \boldsymbol{h}) \cdot \boldsymbol{v}\, \mathrm{d}s \\ + \int_{\Gamma_b} \delta^{-1}(\mathbf{T}\boldsymbol{u} - \boldsymbol{m}) \cdot \boldsymbol{v}\, \mathrm{d}s + \int_{\Omega} \boldsymbol{\varepsilon}(\boldsymbol{u}) : \mathbb{C}[\boldsymbol{\varepsilon}(\boldsymbol{v})]\, \mathrm{d}\boldsymbol{x}$$

Procedure: Obtain $\boldsymbol{m}_{\mathrm{post}}$ by enforcing $\mathcal{L}_{\boldsymbol{m}} = \mathcal{L}_{\boldsymbol{u}} = \mathcal{L}_{\boldsymbol{v}} = 0$.

Choice of Prior:

In our case, we choose $\mathcal{C}_{\mathrm{pr}} = \mathcal{A}^{-2} = (-\gamma\Delta + \delta\mathbf{I})^{-2}$.

Returning to Fault Inversion

Slight Wrinkle: Our parameter-to-observable map is affine.

We want to minimize

$$\mathcal{J}(\boldsymbol{m}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u} - \mathbf{u}^{\mathrm{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\mathrm{noise}}} + \frac{1}{2}\|\mathcal{C}_{\mathrm{pr}}^{-1/2}(\boldsymbol{m} - \boldsymbol{m}_{\mathrm{pr}})\|^2$$

Define:

$$\mathcal{L}(\boldsymbol{m}, \boldsymbol{u}, \boldsymbol{v}) = \frac{1}{2}\|\mathcal{B}\boldsymbol{u} - \mathbf{u}^{\mathrm{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\mathrm{noise}}} + \frac{1}{2}\|\mathcal{C}_{\mathrm{pr}}^{-1/2}(\boldsymbol{m} - \boldsymbol{m}_{\mathrm{pr}})\|^2 + \int_{\Gamma_s} \beta^{-1}(\boldsymbol{u} - \boldsymbol{h}) \cdot \boldsymbol{v}\, \mathrm{d}s$$
$$+ \int_{\Gamma_b} \delta^{-1}(\mathbf{T}\boldsymbol{u} - \boldsymbol{m}) \cdot \boldsymbol{v}\, \mathrm{d}s + \int_{\Omega} \boldsymbol{\varepsilon}(\boldsymbol{u}) : \mathbb{C}[\boldsymbol{\varepsilon}(\boldsymbol{v})]\, \mathrm{d}\boldsymbol{x}$$

Procedure: Obtain $\boldsymbol{m}_{\mathrm{post}}$ by enforcing $\mathcal{L}_{\boldsymbol{m}} = \mathcal{L}_{\boldsymbol{u}} = \mathcal{L}_{\boldsymbol{v}} = 0$.

Choice of Prior:

In our case, we choose $\mathcal{C}_{\mathrm{pr}} = \mathcal{A}^{-2} = (-\gamma\Delta + \delta\mathbf{I})^{-2}$.

A. M. Stuart. "Inverse problems: A Bayesian perspective". In: *Acta Numerica* 19 (May 2010), pp. 451–559. DOI: 10.1017/s0962492910000061

Table of Contents

FEM Discretization

- Consider finite element space $\mathcal{V}_h = \mathrm{Span}(\phi_1, \ldots, \phi_n)$.
- ϕ_k: Lagrange basis functions.
- For $\mathbf{f} \in \mathcal{V}_h$:

$$\mathbf{f} = \sum_{k=1}^{n} f_k \phi_k$$

In particular, construct $\mathbf{u}, \mathbf{v}, \mathbf{m}, \mathbf{m}_{\mathrm{pr}}, \mathbf{h} \in \mathcal{V}_h$ from $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{m}, \boldsymbol{m}_{\mathrm{pr}}$ and $\boldsymbol{h}$.

- Finite dimensional Hilbert space

$$(\mathcal{V}_h, \langle \cdot, \cdot \rangle_{L^2}) \cong (\mathbb{R}^n, \langle \cdot, \cdot \rangle_M), \qquad \langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^{\mathrm{T}} \mathbf{M} \mathbf{v}$$

- From the Lagrangian we find

$$\begin{bmatrix} \mathbf{C}_{\mathrm{pr}}^{-1} & & -\mathbf{C}^* \\ & \mathbf{B}^* \mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathbf{B} & \mathbf{A}^* \\ -\mathbf{C} & \mathbf{A} & \end{bmatrix} \begin{bmatrix} \mathbf{m}_{\mathrm{post}} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{\mathrm{pr}}^{-1} \mathbf{m}_{\mathrm{pr}} \\ \mathbf{B}^* \mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathbf{u}^{\mathrm{obs}} \\ \mathbf{D} \mathbf{h} \end{bmatrix}$$

where $\mathbf{Au} = \mathbf{Cm} + \mathbf{Dh}$ is the discretized weak form.

Reduced Linear System

Parameter-to-Observable Map

From the discretized weak form we can define:

$$\mathbf{Au} = \mathbf{Cm} + \mathbf{Dh} \implies \mathcal{F}(\mathbf{m}) = \mathbf{BA}^{-1}\left[\mathbf{Cm} + \mathbf{Dh}\right] = \mathcal{G}\mathbf{m} + \mathbf{g}$$

Reduced Linear System

Parameter-to-Observable Map

From the discretized weak form we can define:

$$\mathbf{A}\mathbf{u} = \mathbf{C}\mathbf{m} + \mathbf{D}\mathbf{h} \implies \mathcal{F}(\mathbf{m}) = \mathbf{B}\mathbf{A}^{-1}\left[\mathbf{C}\mathbf{m} + \mathbf{D}\mathbf{h}\right] = \mathcal{G}\mathbf{m} + \mathbf{g}$$

With this:

$$\begin{bmatrix} \mathbf{C}_{\mathrm{pr}}^{-1} & & -\mathbf{C}^* \\ & \mathbf{B}^*\mathbf{\Gamma}_{\mathrm{noise}}^{-1}\mathbf{B} & \mathbf{A}^* \\ -\mathbf{C} & \mathbf{A} & \end{bmatrix} \begin{bmatrix} \mathbf{m}_{\mathrm{post}} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{\mathrm{pr}}^{-1}\mathbf{m}_{\mathrm{pr}} \\ \mathbf{B}^*\mathbf{\Gamma}_{\mathrm{noise}}^{-1}\mathbf{u}^{\mathrm{obs}} \\ \mathbf{D}\mathbf{h} \end{bmatrix}$$

can be reduced to

$$\left(\mathbf{C}_{\mathrm{pr}}^{-1} + \mathcal{G}^*\mathbf{\Gamma}_{\mathrm{noise}}^{-1}\mathcal{G}\right)\mathbf{m}_{\mathrm{post}} = \mathcal{G}^*\mathbf{\Gamma}_{\mathrm{noise}}^{-1}\left(\mathbf{g} + \mathbf{u}^{\mathrm{obs}}\right) + \mathbf{C}_{\mathrm{pr}}^{-1}\mathbf{m}_{\mathrm{pr}}$$

Tractability of Computation

$$\underbrace{\left[\mathbf{C}_{\mathrm{pr}}^{-1} + \mathcal{G}^* \boldsymbol{\Gamma}_{\mathrm{noise}}^{-1} \mathcal{G}\right]}_{\mathbf{C}_{\mathrm{post}}^{-1}} \mathbf{m}_{\mathrm{post}} = \left[\mathcal{G}^* \boldsymbol{\Gamma}_{\mathrm{noise}}^{-1} \left(\mathbf{g} + \mathbf{u}^{\mathrm{obs}}\right) + \mathbf{C}_{\mathrm{pr}}^{-1} \mathbf{m}_{\mathrm{pr}}\right]$$

How do we make this computation tractable?

Tractability of Computation

$$\underbrace{\left[\mathbf{C}_{\mathrm{pr}}^{-1} + \mathcal{G}^* \boldsymbol{\Gamma}_{\mathrm{noise}}^{-1} \mathcal{G}\right]}_{\mathbf{C}_{\mathrm{post}}^{-1}} \mathbf{m}_{\mathrm{post}} = \left[\mathcal{G}^* \boldsymbol{\Gamma}_{\mathrm{noise}}^{-1} \left(\mathbf{g} + \mathbf{u}^{\mathrm{obs}}\right) + \mathbf{C}_{\mathrm{pr}}^{-1} \mathbf{m}_{\mathrm{pr}}\right]$$

How do we make this computation tractable?

- Exploit the dimensionality reduction in the parameter space
- Use symmetry aware solvers
- Leverage sparsity of
  - matrices induced by differential operators
  - observations

The discretized parameter space

Recall the boundary condition describing the inversion parameter (1e)

$$\delta \mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) + \mathbf{T}\boldsymbol{u} = \boldsymbol{m}, \quad \text{on } \Gamma_b \subset \mathbb{R}^2$$

By the definition of $\mathbf{T}$:

$$\boldsymbol{m} = \begin{pmatrix} m_x \\ 0 \\ m_z \end{pmatrix}$$

Construct a second finite element space $\mathcal{V}_h'$ for Γ_b and instead choose $\mathbf{m}, \mathbf{m}_{\mathrm{pr}} \in (\mathcal{V}_h', \langle \cdot, \cdot \rangle_{L^2}) \cong (\mathbb{R}^m, \langle \cdot, \cdot \rangle_{\mathbf{M}_p})$.

The discretized parameter space

Recall the boundary condition describing the inversion parameter (1e)

$$\delta \mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) + \mathbf{T}\boldsymbol{u} = \boldsymbol{m}, \quad \text{on } \Gamma_b \subset \mathbb{R}^2$$

By the definition of $\mathbf{T}$:

$$\boldsymbol{m} = \begin{pmatrix} m_x \\ 0 \\ m_z \end{pmatrix}$$

Construct a second finite element space $\mathcal{V}_h'$ for Γ_b and instead choose $\mathbf{m}, \mathbf{m}_{\mathrm{pr}} \in (\mathcal{V}_h', \langle \cdot, \cdot \rangle_{L^2}) \cong (\mathbb{R}^m, \langle \cdot, \cdot \rangle_{\mathbf{M}_p})$.

- FEniCS data structures not very interoporable over mixed dimensions.

The discretized parameter space

Recall the boundary condition describing the inversion parameter (1e)

$$\delta \mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) + \mathbf{T}\boldsymbol{u} = \boldsymbol{m}, \quad \text{on } \Gamma_b \subset \mathbb{R}^2$$

By the definition of $\mathbf{T}$:

$$\boldsymbol{m} = \begin{pmatrix} m_x \\ 0 \\ m_z \end{pmatrix}$$

Construct a second finite element space $\mathcal{V}_h'$ for Γ_b and instead choose $\mathbf{m}, \mathbf{m}_{\mathrm{pr}} \in (\mathcal{V}_h', \langle \cdot, \cdot \rangle_{L^2}) \cong (\mathbb{R}^m, \langle \cdot, \cdot \rangle_{\mathbf{M}_p})$.

- FEniCS data structures not very interoporable over mixed dimensions.
- Bijective maps between degrees of freedom and vertices are supported only for 1st order Lagrange elements.

The discretized parameter space

Recall the boundary condition describing the inversion parameter (1e)

$$\delta \mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) + \mathbf{T}\boldsymbol{u} = \boldsymbol{m}, \quad \text{on } \Gamma_b \subset \mathbb{R}^2$$

By the definition of $\mathbf{T}$:

$$\boldsymbol{m} = \begin{pmatrix} m_x \\ 0 \\ m_z \end{pmatrix}$$

Construct a second finite element space $\mathcal{V}_h'$ for Γ_b and instead choose $\mathbf{m}, \mathbf{m}_{\mathrm{pr}} \in (\mathcal{V}_h', \langle \cdot, \cdot \rangle_{L^2}) \cong (\mathbb{R}^m, \langle \cdot, \cdot \rangle_{\mathbf{M}_p})$.

- FEniCS data structures not very interoporable over mixed dimensions.
- Bijective maps between degrees of freedom and vertices are supported only for 1st order Lagrange elements.
- Need to build custom projector between mixed dimensional topologies and function spaces (Special thanks to Tucker Hartland from UC Merced).

Symmetry: Dealing with the inner product

For $T : (\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}}) \to (\mathcal{Y}, \langle \cdot, \cdot \rangle_{\mathcal{Y}})$, the adjoint T^* is defined by the identity:

$$\langle x, T^* y \rangle_{\mathcal{X}} = \langle Tx, y \rangle_{\mathcal{Y}}, \quad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}$$

Hence, in our setting, $(\cdot)^{\mathrm{T}} \neq (\cdot)^*$!

Symmetry: Dealing with the inner product

For $T : (\mathcal{X}, \langle\cdot,\cdot\rangle_{\mathcal{X}}) \to (\mathcal{Y}, \langle\cdot,\cdot\rangle_{\mathcal{Y}})$, the adjoint T^* is defined by the identity:

$$\langle x, T^* y\rangle_{\mathcal{X}} = \langle Tx, y\rangle_{\mathcal{Y}}, \quad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}$$

Hence, in our setting, $(\cdot)^{\mathrm{T}} \neq (\cdot)^*$! Consider the parameter-to-observable map:

$$\mathcal{G} : (\mathbb{R}^m, \langle\cdot,\cdot\rangle_{\mathbf{M}_p}) \to (\mathbb{R}^q, \langle\cdot,\cdot\rangle)$$

Since

$$\langle \mathcal{G}\mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^{\mathrm{T}}\mathcal{G}^{\mathrm{T}}\mathbf{y} = \mathbf{x}^{\mathrm{T}}\mathbf{M}_p\mathbf{M}_p^{-1}\mathcal{G}^{\mathrm{T}}\mathbf{y} = \langle \mathbf{x}, \mathbf{M}_p^{-1}\mathcal{G}^{\mathrm{T}}\mathbf{y}\rangle$$

So

$$\mathcal{G}^* = \mathbf{M}_p^{-1}\mathcal{G}^{\mathrm{T}}$$

Symmetry: Dealing with the inner product

For $T : (\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}}) \to (\mathcal{Y}, \langle \cdot, \cdot \rangle_{\mathcal{Y}})$, the adjoint T^* is defined by the identity:

$$\langle x, T^* y \rangle_{\mathcal{X}} = \langle Tx, y \rangle_{\mathcal{Y}}, \quad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}$$

Hence, in our setting, $(\cdot)^{\mathrm{T}} \neq (\cdot)^*$! Consider the parameter-to-observable map:

$$\mathcal{G} : (\mathbb{R}^m, \langle \cdot, \cdot \rangle_{\mathbf{M}_p}) \to (\mathbb{R}^q, \langle \cdot, \cdot \rangle)$$

Since

$$\langle \mathcal{G}\mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathrm{T}} \mathcal{G}^{\mathrm{T}} \mathbf{y} = \mathbf{x}^{\mathrm{T}} \mathbf{M}_p \mathbf{M}_p^{-1} \mathcal{G}^{\mathrm{T}} \mathbf{y} = \langle \mathbf{x}, \mathbf{M}_p^{-1} \mathcal{G}^{\mathrm{T}} \mathbf{y} \rangle$$

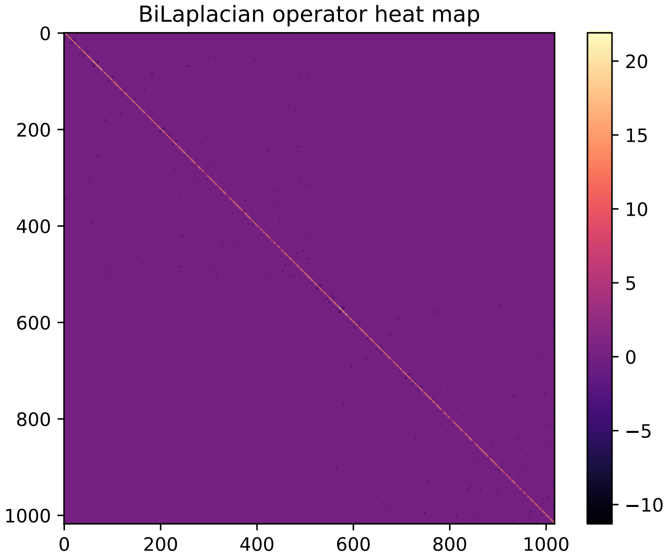
So

$$\mathcal{G}^* = \mathbf{M}_p^{-1} \mathcal{G}^{\mathrm{T}}$$

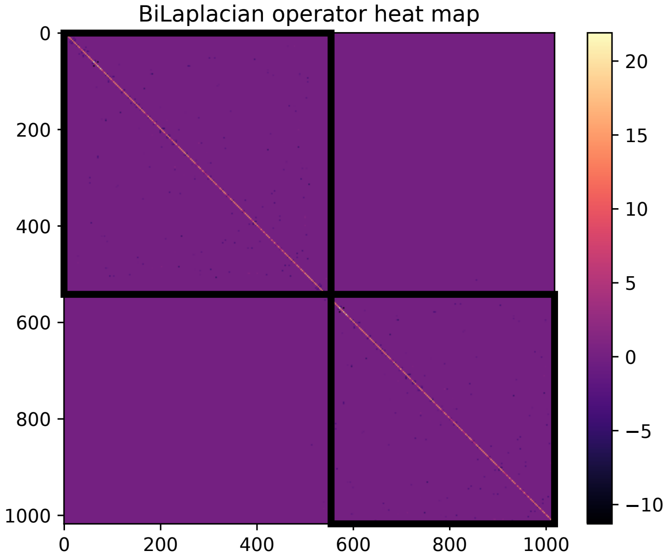
Note:

There are two mass matrices, $\mathbf{M}$ and $\mathbf{M}_p$. Thankfully, only $\mathbf{M}_p$ is necessary.

Sparsity: Matrices induced by differential operators



Sparsity: Matrices induced by differential operators



Sparsity: Observations

$$
\begin{aligned}
\mathbf{C}_{\mathrm{post}} &= \left(\mathcal{G}^* \mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathcal{G} + \mathbf{C}_{\mathrm{pr}}^{-1}\right) \\
&= \mathbf{C}_{\mathrm{pr}}^{1/2} \left(\mathbf{C}_{\mathrm{pr}}^{1/2} \mathcal{G}^* \mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathcal{G} \mathbf{C}_{\mathrm{pr}}^{1/2} + \mathbf{I}\right) \mathbf{C}_{\mathrm{pr}}^{1/2} \\
&= \mathbf{C}_{\mathrm{pr}}^{1/2} \left(\tilde{\mathbf{H}}_{\mathrm{misfit}} + \mathbf{I}\right) \mathbf{C}_{\mathrm{pr}}^{1/2}
\end{aligned}
$$

Sparsity: Observations

$$\begin{aligned}
\mathbf{C}_{\mathrm{post}} &= \left(\mathcal{G}^*\boldsymbol{\Gamma}_{\mathrm{noise}}^{-1}\mathcal{G} + \mathbf{C}_{\mathrm{pr}}^{-1}\right) \\
&= \mathbf{C}_{\mathrm{pr}}^{1/2}\left(\mathbf{C}_{\mathrm{pr}}^{1/2}\mathcal{G}^*\boldsymbol{\Gamma}_{\mathrm{noise}}^{-1}\mathcal{G}\mathbf{C}_{\mathrm{pr}}^{1/2} + \mathbf{I}\right)\mathbf{C}_{\mathrm{pr}}^{1/2} \\
&= \mathbf{C}_{\mathrm{pr}}^{1/2}\left(\tilde{\mathbf{H}}_{\mathrm{misfit}} + \mathbf{I}\right)\mathbf{C}_{\mathrm{pr}}^{1/2}
\end{aligned}$$

Recall:

$$\mathcal{G} : (\mathbb{R}^m, \langle\cdot,\cdot\rangle_{\mathbf{M}_p}) \to (\mathbb{R}^q, \langle\cdot,\cdot\rangle)$$

Note:

Observations are sparse, so $q \ll m$. Thus $\tilde{\mathbf{H}}_{\mathrm{misfit}}$ is low rank.

Construct:

$$\tilde{\mathbf{H}}_{\mathrm{misfit}} = \mathbf{V}_r \mathbf{\Lambda}_r \mathbf{V}_r^* + \mathcal{O}\left(\sum_{k=r+1}^{m} \lambda_k \right)$$

Construct:

$$\tilde{\mathbf{H}}_{\mathrm{misfit}} = \mathbf{V}_r \boldsymbol{\Lambda}_r \mathbf{V}_r^* + \mathcal{O}\left(\sum_{k=r+1}^{m} \lambda_k\right)$$

Via the Sherman-Morrison-Woodbury formula:

$$\mathbf{D}_r = \mathrm{Diag}\left(\frac{\lambda_1}{1+\lambda_1}, \cdots, \frac{\lambda_r}{1+\lambda_r}\right)$$

$$\begin{aligned}\left(\tilde{\mathbf{H}}_{\mathrm{misfit}} + \mathbf{I}\right)^{-1} &= \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* + \mathcal{O}\left(\sum_{k=r+1}^{m} \frac{\lambda_k}{1+\lambda_k}\right) \\ &\approx \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^*\end{aligned}$$

Construct:

$$\tilde{\mathbf{H}}_{\mathrm{misfit}} = \mathbf{V}_r \mathbf{\Lambda}_r \mathbf{V}_r^* + \mathcal{O}\left(\sum_{k=r+1}^{m} \lambda_k \right)$$

Via the Sherman-Morrison-Woodbury formula:

$$\mathbf{D}_r = \mathrm{Diag}\left(\frac{\lambda_1}{1+\lambda_1}, \cdots, \frac{\lambda_r}{1+\lambda_r} \right)$$

$$\left(\tilde{\mathbf{H}}_{\mathrm{misfit}} + \mathbf{I} \right)^{-1} = \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* + \mathcal{O}\left(\sum_{k=r+1}^{m} \frac{\lambda_k}{1+\lambda_k} \right)$$

$$\approx \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^*$$

Hence:

$$\mathbf{C}_{\mathrm{post}} = \mathbf{C}_{\mathrm{pr}} - \mathbf{C}_{\mathrm{pr}}^{1/2} \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* \mathbf{C}_{\mathrm{pr}}^{1/2}$$

Construct:

$$\tilde{\mathbf{H}}_{\mathrm{misfit}} = \mathbf{V}_r \mathbf{\Lambda}_r \mathbf{V}_r^* + \mathcal{O}\left(\sum_{k=r+1}^{m} \lambda_k \right)$$

Via the Sherman-Morrison-Woodbury formula:

$$\mathbf{D}_r = \mathrm{Diag}\left(\frac{\lambda_1}{1+\lambda_1}, \cdots, \frac{\lambda_r}{1+\lambda_r} \right)$$

$$\left(\tilde{\mathbf{H}}_{\mathrm{misfit}} + \mathbf{I}\right)^{-1} = \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* + \mathcal{O}\left(\sum_{k=r+1}^{m} \frac{\lambda_k}{1+\lambda_k} \right)$$

$$\approx \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^*$$

Hence:

$$\mathbf{C}_{\mathrm{post}} = \mathbf{C}_{\mathrm{pr}} - \mathbf{C}_{\mathrm{pr}}^{1/2} \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* \mathbf{C}_{\mathrm{pr}}^{1/2}$$

Tan Bui-Thanh et al. "A Computational Framework for Infinite-Dimensional Bayesian Inverse Problems Part I: The Linearized Case, with Application to Global Seismic Inversion". In: *SIAM Journal on Scientific Computing* 35.6 (Jan. 2013), A2494–A2523. DOI: 10.1137/12089586x

Sampling from resulting distributions

Note, in order to draw samples from $\mathcal{N}(\mathbf{m}, \mathbf{C})$:

- Find $\mathbf{L}$ where $\mathbf{C} = \mathbf{L}\mathbf{L}^*$
- Sample $n \in \mathcal{N}(\mathbf{0}, \mathbf{I})$.
- Construct sample $\mathbf{s}$ as:

$$\mathbf{s} = \mathbf{m} + \mathbf{L}\mathbf{n}$$

Sampling from resulting distributions

Note, in order to draw samples from $\mathcal{N}(\mathbf{m}, \mathbf{C})$:

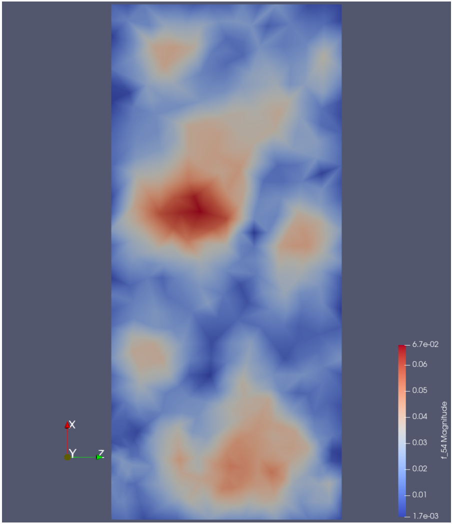
- Find $\mathbf{L}$ where $\mathbf{C} = \mathbf{L}\mathbf{L}^*$
- Sample $n \in \mathcal{N}(\mathbf{0}, \mathbf{I})$.
- Construct sample $\mathbf{s}$ as:
$$\mathbf{s} = \mathbf{m} + \mathbf{L}\mathbf{n}$$

Since $\mathcal{C}_{\mathrm{pr}} = \mathcal{A}^{-2}$, prior is easy. For posterior:

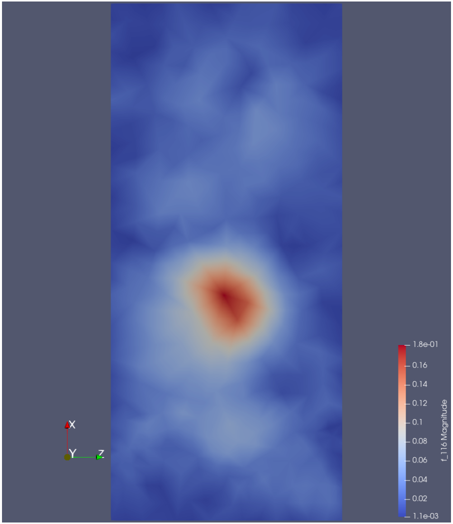
$$\mathbf{L} = \mathbf{C}_{\mathrm{pr}}^{1/2}(\mathbf{V}_r\mathbf{P}_r\mathbf{V}_r^* + \mathbf{I})\mathbf{M_p}^{-1/2}$$
$$\mathbf{P}_r = \mathrm{Diag}\left(\frac{1}{\sqrt{\lambda_1 + 1} - 1}\right)$$

Table of Contents

Results

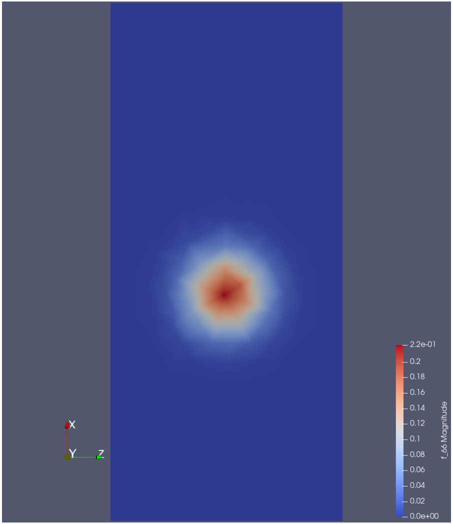


X
Y
Z
6.7e-02
0.06
0.05
0.04
0.03
0.02
0.01
1.7e-03
f_54 Magnitude

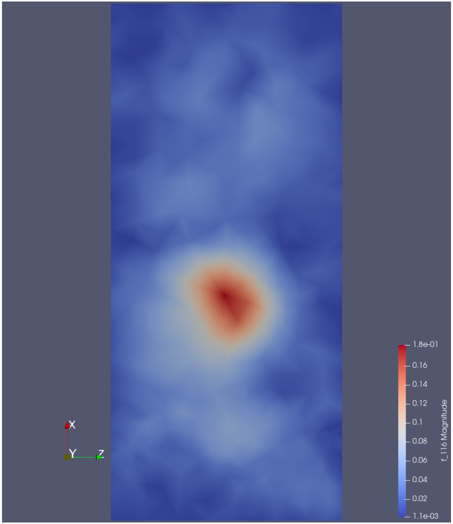


X
Y
Z
1.8e-01
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
1.1e-03
f_116 Magnitude

Results



X
Y
Z
2.2e-01
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0.0e+00
f_66 Magnitude



X
Y
Z
1.8e-01
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
1.1e-03
f_116 Magnitude

Table of Contents

Future Work

This problem is intended as a model problem for future uncertainty quantification endavors. This problem has three interesting qualities:

- It's a reasonable approximation to well-studied physical problem.
- Mixed dimensional formulation in topology and function space.
- It's posterior is provably Gaussian, hence a priori analysis is possible.

The intention is to use this as an additional example for work in

- Different criteria for optimal experimental design
- Hyper-Differential Sensitivity Analysis

Alen Alexanderian et al. "Optimal Design of Large-scale Bayesian Linear Inverse Problems Under Reducible Model Uncertainty: Good to Know What You Don't Know". In: *SIAM/ASA Journal on Uncertainty Quantification* 9.1 (Jan. 2021), pp. 163–184. DOI: 10.1137/20m1347292. URL: https://doi.org/10.1137/20m1347292

Isaac Sunseri et al. "Hyper-differential sensitivity analysis for inverse problems constrained by partial differential equations". In: *Inverse Problems* 36.12 (Dec. 2020), p. 125001. DOI: 10.1088/1361-6420/abaf63

Table of Contents

References I

[1] Alen Alexanderian et al. "Optimal Design of Large-scale Bayesian Linear Inverse Problems Under Reducible Model Uncertainty: Good to Know What You Don't Know". In: *SIAM/ASA Journal on Uncertainty Quantification* 9.1 (Jan. 2021), pp. 163–184. DOI: 10.1137/20m1347292. URL: https://doi.org/10.1137/20m1347292.

[2] Tan Bui-Thanh et al. "A Computational Framework for Infinite-Dimensional Bayesian Inverse Problems Part I: The Linearized Case, with Application to Global Seismic Inversion". In: *SIAM Journal on Scientific Computing* 35.6 (Jan. 2013), A2494–A2523. DOI: 10.1137/12089586x.

[3] Robert J Lillie. *Oregon's Island In The Sky: Geology Road Guide to Marys Peak*. English. OCLC: 979996650. 2017. ISBN: 9781540611963.

[4] Kimberly Alison McCormack. *Earthquakes, groundwater and surface deformation : Exploring the poroelastic response to megathrust earthquakes*. Aug. 2018. URL: hdl.handle.net/2152/68892.

References II

[5] Tom Simkin et al. *This dynamic planet: World map of volcanoes, earthquakes, impact craters and plate tectonics*. Tech. rep. 2006. DOI: 10.3133/i2800.

[6] A. M. Stuart. "Inverse problems: A Bayesian perspective". In: *Acta Numerica* 19 (May 2010), pp. 451–559. DOI: 10.1017/s0962492910000061.

[7] Isaac Sunseri et al. "Hyper-differential sensitivity analysis for inverse problems constrained by partial differential equations". In: *Inverse Problems* 36.12 (Dec. 2020), p. 125001. DOI: 10.1088/1361-6420/abaf63.