

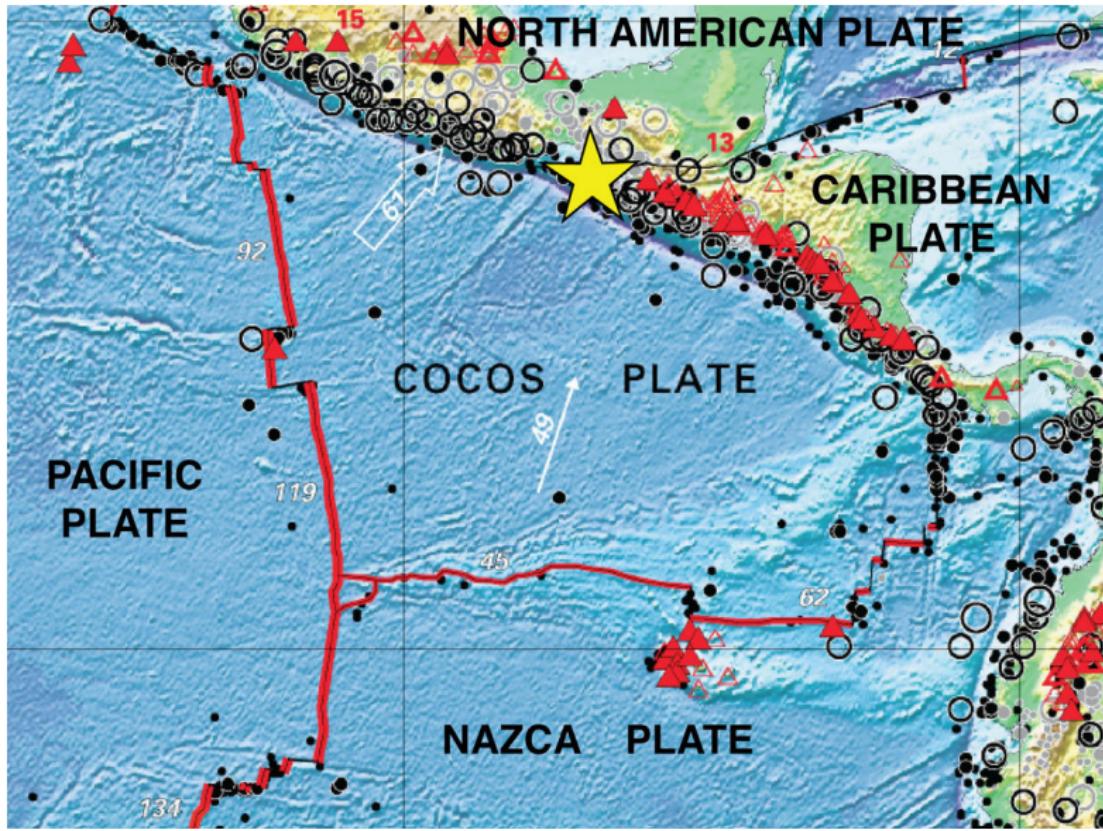
# Infinite-dimensional Bayesian inversion for fault slip from surface measurements

Abhijit Chowdhary and Alen Alexanderian

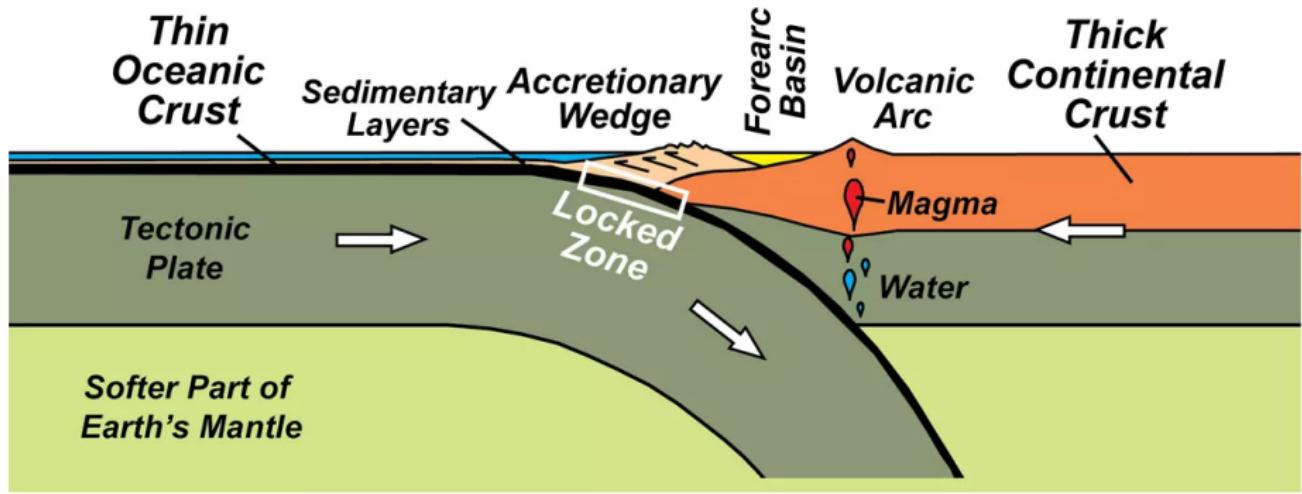
Department of Mathematics  
North Carolina State University

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](https://doi.org/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.

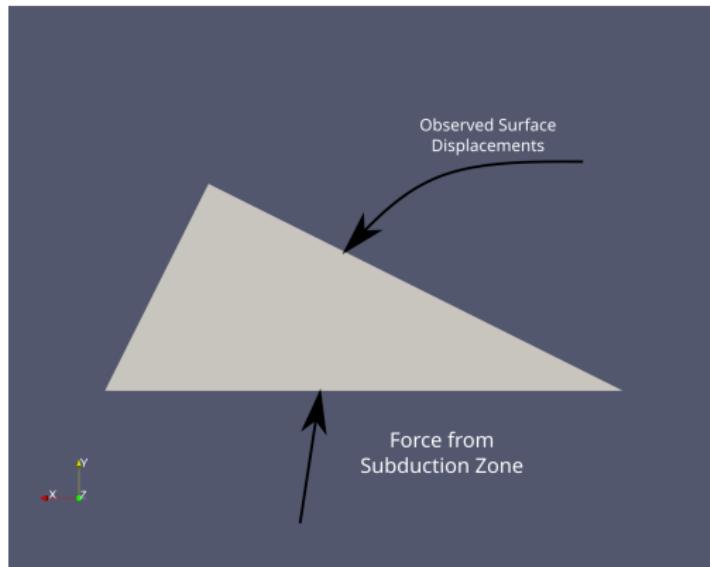
# Table of Contents

- 1 Seismic Inversion Model Problem
- 2 Infinite-Dimensional Inverse Problem Setting
- 3 Numerical Discretization and Considerations
- 4 Results
- 5 Future Work
- 6 References

# Table of Contents

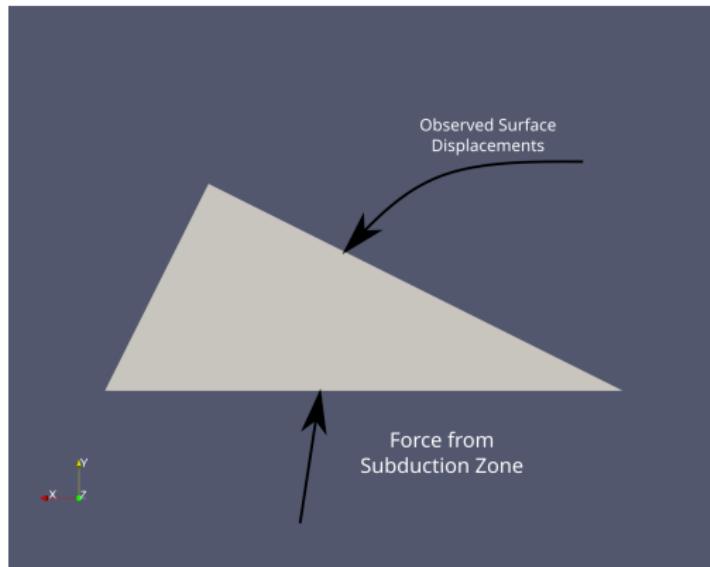
- 1 Seismic Inversion Model Problem
- 2 Infinite-Dimensional Inverse Problem Setting
- 3 Numerical Discretization and Considerations
- 4 Results
- 5 Future Work
- 6 References

# Model Assumptions



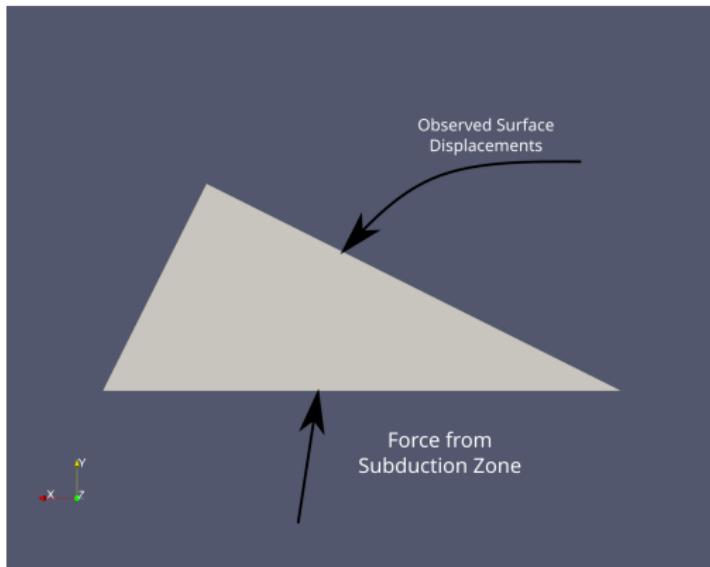
- **Governing PDE** (forward model): Linear elasticity

# 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  $\mathbf{u}^{\text{obs}}$  reconstruct fault plane displacement.

# Forward Model

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

where:

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

# Forward Model

$$-\nabla \left[ \mu(\nabla \mathbf{u} + (\nabla \mathbf{u})^T) + \lambda \nabla \cdot \mathbf{u} \mathbf{I} \right] = \mathbf{0} \quad \text{in } \Omega, \quad (1a)$$

$$\sigma(\mathbf{u})\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_t \quad (1b)$$

$$\mathbf{u} + \beta \sigma(\mathbf{u})\mathbf{n} = \mathbf{h} \quad \text{on } \Gamma_s \quad (1c)$$

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_b \quad (1d)$$

$$\delta \mathbf{T}(\sigma(\mathbf{u})\mathbf{n}) + \mathbf{T}\mathbf{u} = \mathbf{m} \quad \text{on } \Gamma_b \quad (1e)$$

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

# Forward Model

$$-\nabla \left[ \mu(\nabla \mathbf{u} + (\nabla \mathbf{u})^T) + \lambda \nabla \cdot \mathbf{u} \mathbf{I} \right] = \mathbf{0} \quad \text{in } \Omega, \quad (1a)$$

$$\sigma(\mathbf{u})\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_t \quad (1b)$$

$$\mathbf{u} + \beta \sigma(\mathbf{u})\mathbf{n} = \mathbf{h} \quad \text{on } \Gamma_s \quad (1c)$$

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_b \quad (1d)$$

$$\delta \mathbf{T}(\sigma(\mathbf{u})\mathbf{n}) + \mathbf{T}\mathbf{u} = \mathbf{m} \quad \text{on } \Gamma_b \quad (1e)$$

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

# Forward Model

$$-\nabla \left[ \mu(\nabla \mathbf{u} + (\nabla \mathbf{u})^T) + \lambda \nabla \cdot \mathbf{u} \mathbf{I} \right] = \mathbf{0} \quad \text{in } \Omega, \quad (1a)$$

$$\sigma(\mathbf{u})\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_t \quad (1b)$$

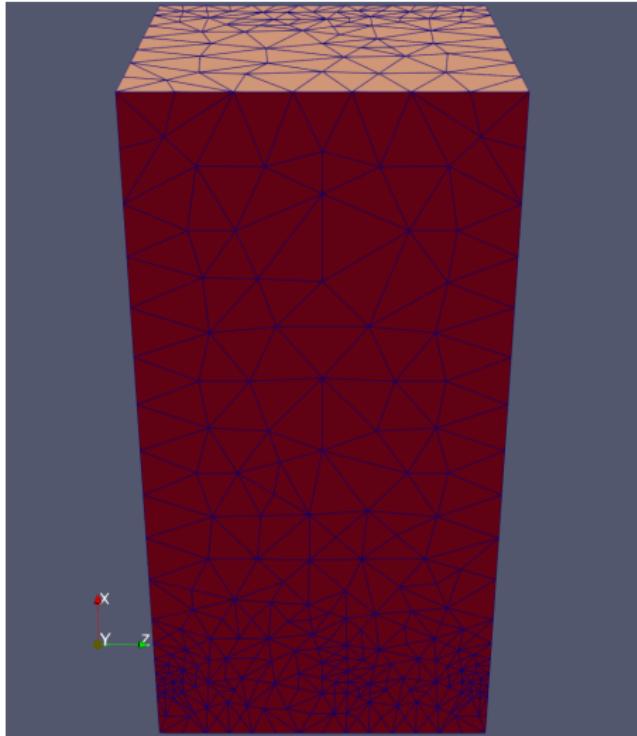
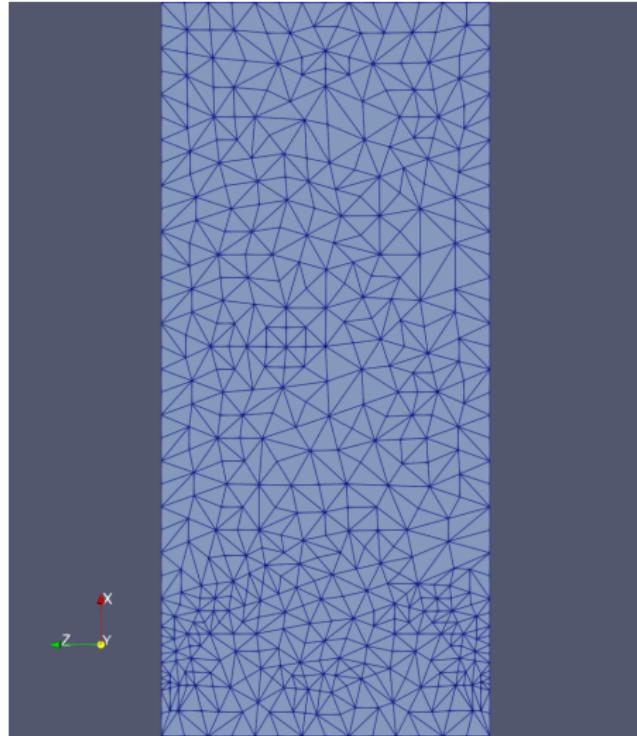
$$\mathbf{u} + \beta \sigma(\mathbf{u})\mathbf{n} = \mathbf{h} \quad \text{on } \Gamma_s \quad (1c)$$

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_b \quad (1d)$$

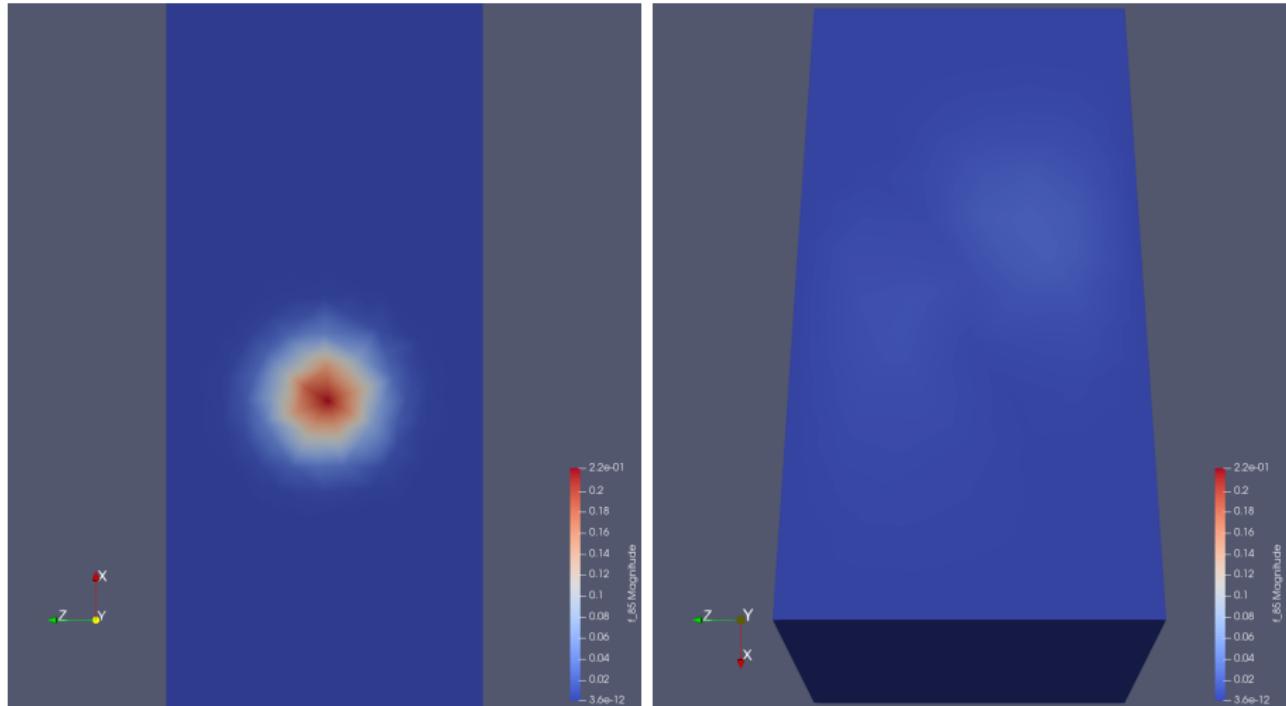
$$\delta \mathbf{T}(\sigma(\mathbf{u})\mathbf{n}) + \mathbf{T}\mathbf{u} = \mathbf{m} \quad \text{on } \Gamma_b \quad (1e)$$

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

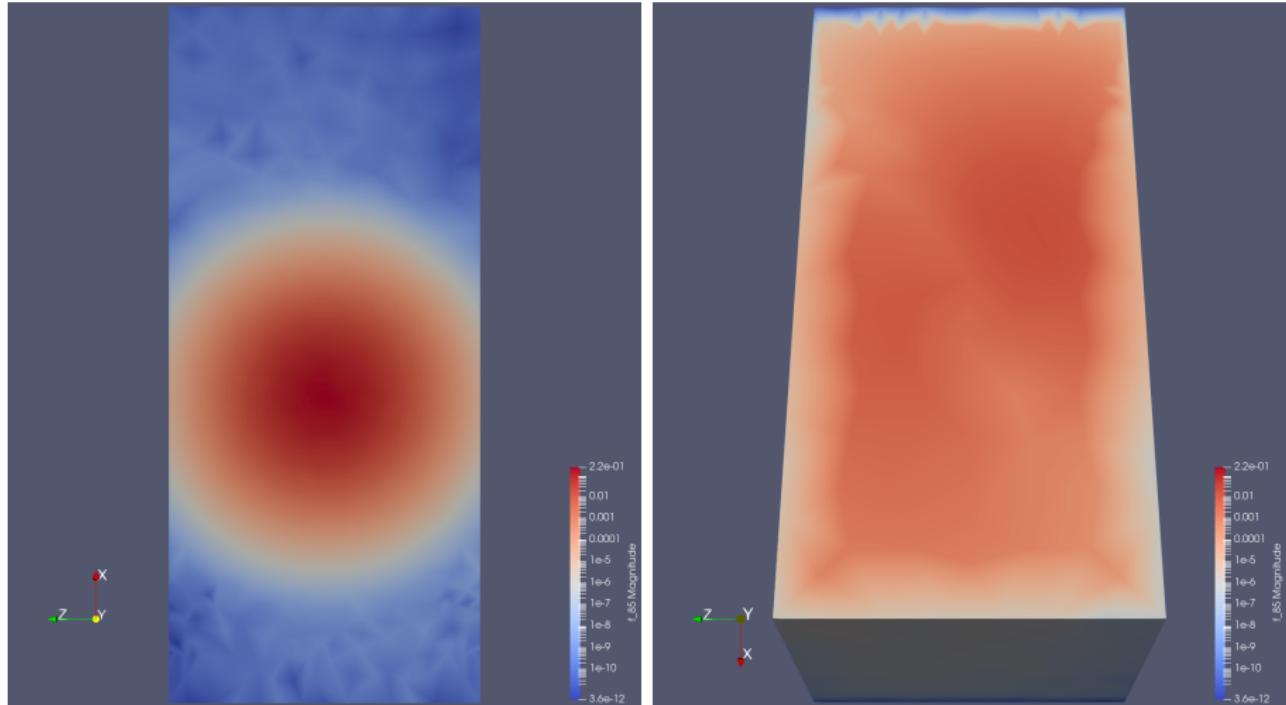
# Mesh



# Forward Solution



# Forward Solution



## Weak Formulation

Define:

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

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

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

# Table of Contents

- 1 Seismic Inversion Model Problem
- 2 Infinite-Dimensional Inverse Problem Setting
- 3 Numerical Discretization and Considerations
- 4 Results
- 5 Future Work
- 6 References

# The Deterministic Inverse Problem

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

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

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

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

---

<sup>1</sup>In literature, called the *parameter-to-observable operator*.

# The Deterministic Inverse Problem

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

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

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

- $\mathbf{u}(\mathbf{m})$  is given by the forward model.
- $\mathcal{B} : (L^2(\Omega))^3 \rightarrow \mathbb{R}^N$  is an observation operator.
- $\mathbf{u}^{\text{obs}} \in \mathbb{R}^N$  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.  $\mathbf{u} = \mathcal{S}\mathbf{m}$ , and let<sup>1</sup>  $\mathcal{F} = \mathcal{B}\mathcal{S}$ , then:

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

---

<sup>1</sup>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  $\mathbf{m} \sim \mathcal{N}(\mathbf{m}_{\text{pr}}, \boldsymbol{\Gamma}_{\text{pr}})$ .
- Additive Gaussian noise

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

- Gaussian posterior with:

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

where:

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

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

# Bayesian Inversion in Infinite Dimensions

## Theorem (Bayes Theorem in Infinite Dimensions)

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

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

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

- Gaussian posterior:

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

where:

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

$$\mathcal{C}_{\text{post}} = \left( \mathcal{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathcal{F} + \mathcal{C}_{\text{pr}}^{-1} \right)^{-1}$$

## 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  $\mathbf{m}_{\text{post}}$ , need to minimize:

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

# Returning to Fault Inversion

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

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

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

Define:

$$\begin{aligned} \mathcal{L}(\mathbf{m}, \mathbf{u}, \mathbf{v}) = & \frac{1}{2} \|\mathcal{B}\mathbf{u} - \mathbf{u}^{\text{obs}}\|_{\Gamma_{\text{noise}}^{-1}}^2 + \frac{1}{2} \|\mathcal{C}_{\text{pr}}^{-1/2}(\mathbf{m} - \mathbf{m}_{\text{pr}})\|^2 + \int_{\Gamma_s} \beta^{-1}(\mathbf{u} - \mathbf{h}) \cdot \mathbf{v} \, ds \\ & + \int_{\Gamma_b} \delta^{-1}(\mathbf{T}\mathbf{u} - \mathbf{m}) \cdot \mathbf{v} \, ds + \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \mathbb{C}[\boldsymbol{\varepsilon}(\mathbf{v})] \, dx \end{aligned}$$

**Procedure:** Obtain  $\mathbf{m}_{\text{post}}$  by enforcing  $\mathcal{L}_m = \mathcal{L}_u = \mathcal{L}_v = 0$ .

# Returning to Fault Inversion

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

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

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

Define:

$$\begin{aligned} \mathcal{L}(\mathbf{m}, \mathbf{u}, \mathbf{v}) = & \frac{1}{2} \|\mathcal{B}\mathbf{u} - \mathbf{u}^{\text{obs}}\|_{\Gamma_{\text{noise}}^{-1}}^2 + \frac{1}{2} \|\mathcal{C}_{\text{pr}}^{-1/2}(\mathbf{m} - \mathbf{m}_{\text{pr}})\|^2 + \int_{\Gamma_s} \beta^{-1}(\mathbf{u} - \mathbf{h}) \cdot \mathbf{v} \, ds \\ & + \int_{\Gamma_b} \delta^{-1}(\mathbf{T}\mathbf{u} - \mathbf{m}) \cdot \mathbf{v} \, ds + \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \mathbb{C}[\boldsymbol{\varepsilon}(\mathbf{v})] \, dx \end{aligned}$$

**Procedure:** Obtain  $\mathbf{m}_{\text{post}}$  by enforcing  $\mathcal{L}_m = \mathcal{L}_u = \mathcal{L}_v = 0$ .

Choice of Prior:

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

# Returning to Fault Inversion

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

We want to minimize

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

Define:

$$\begin{aligned} \mathcal{L}(\mathbf{m}, \mathbf{u}, \mathbf{v}) = & \frac{1}{2} \|\mathcal{B}\mathbf{u} - \mathbf{u}^{\text{obs}}\|_{\Gamma_{\text{noise}}^{-1}}^2 + \frac{1}{2} \|\mathcal{C}_{\text{pr}}^{-1/2}(\mathbf{m} - \mathbf{m}_{\text{pr}})\|^2 + \int_{\Gamma_s} \beta^{-1}(\mathbf{u} - \mathbf{h}) \cdot \mathbf{v} \, ds \\ & + \int_{\Gamma_b} \delta^{-1}(\mathbf{T}\mathbf{u} - \mathbf{m}) \cdot \mathbf{v} \, ds + \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \mathbb{C}[\boldsymbol{\varepsilon}(\mathbf{v})] \, dx \end{aligned}$$

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

Choice of Prior:

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

# Table of Contents

- 1 Seismic Inversion Model Problem
- 2 Infinite-Dimensional Inverse Problem Setting
- 3 Numerical Discretization and Considerations**
- 4 Results
- 5 Future Work
- 6 References

# FEM Discretization

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

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

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

- Finite dimensional Hilbert space

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

- From the Lagrangian we find

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

where  $\mathbf{A}\mathbf{u} = \mathbf{C}\mathbf{m} + \mathbf{D}\mathbf{h}$  is the discretized weak form.

# Reduced Linear System

## Parameter-to-Observable Map

From the discretized weak form we can define:

$$\mathbf{A}\mathbf{u} = \mathbf{Cm} + \mathbf{Dh} \implies \mathcal{F}(\mathbf{m}) = \mathbf{B}\mathbf{A}^{-1} [\mathbf{Cm} + \mathbf{Dh}] = \mathcal{G}\mathbf{m} + \mathbf{g}$$

# Reduced Linear System

## Parameter-to-Observable Map

From the discretized weak form we can define:

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

With this:

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

can be reduced to

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

# Tractability of Computation

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

How do we make this computation tractable?

# Tractability of Computation

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

How do we make this computation tractable?

- Exploit the dimensionality reduction in the parameter space

# Tractability of Computation

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

How do we make this computation tractable?

- Exploit the dimensionality reduction in the parameter space
- Use symmetry aware solvers

# Tractability of Computation

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

How do we make this computation tractable?

- Exploit the dimensionality reduction in the parameter space
- Use symmetry aware solvers
- Leverage sparsity of

# Tractability of Computation

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

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

# Tractability of Computation

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

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

# Tractability of Computation

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

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

---

Martin Alnæs et al. "The FEniCS Project Version 1.5". en. In: *Archive of Numerical Software* Vol 3 (2015). DOI: 10.11588/ANS.2015.100.20553

Umberto Villa, Noemi Petra, and Omar Ghattas. "HIPPYlib: An Extensible Software Framework for Large-Scale Inverse Problems Governed by PDEs: Part I: Deterministic Inversion and Linearized Bayesian Inference". In: *ACM Trans. Math. Softw.* 47.2 (Apr. 2021). ISSN: 0098-3500. DOI: 10.1145/3428447

# The discretized parameter space

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

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

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

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

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

- FEniCS data structures not very interoperable 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  $T : (\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}}) \rightarrow (\mathcal{Y}, \langle \cdot, \cdot \rangle_{\mathcal{Y}})$ , 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,  $(\cdot)^T \neq (\cdot)^*$ !

# Symmetry: Dealing with the inner product

For  $T : (\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}}) \rightarrow (\mathcal{Y}, \langle \cdot, \cdot \rangle_{\mathcal{Y}})$ , 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,  $(\cdot)^T \neq (\cdot)^*$ ! Consider the parameter-to-observable map:

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

Since

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

So

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

# Symmetry: Dealing with the inner product

For  $T : (\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}}) \rightarrow (\mathcal{Y}, \langle \cdot, \cdot \rangle_{\mathcal{Y}})$ , 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,  $(\cdot)^T \neq (\cdot)^*$ ! Consider the parameter-to-observable map:

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

Since

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

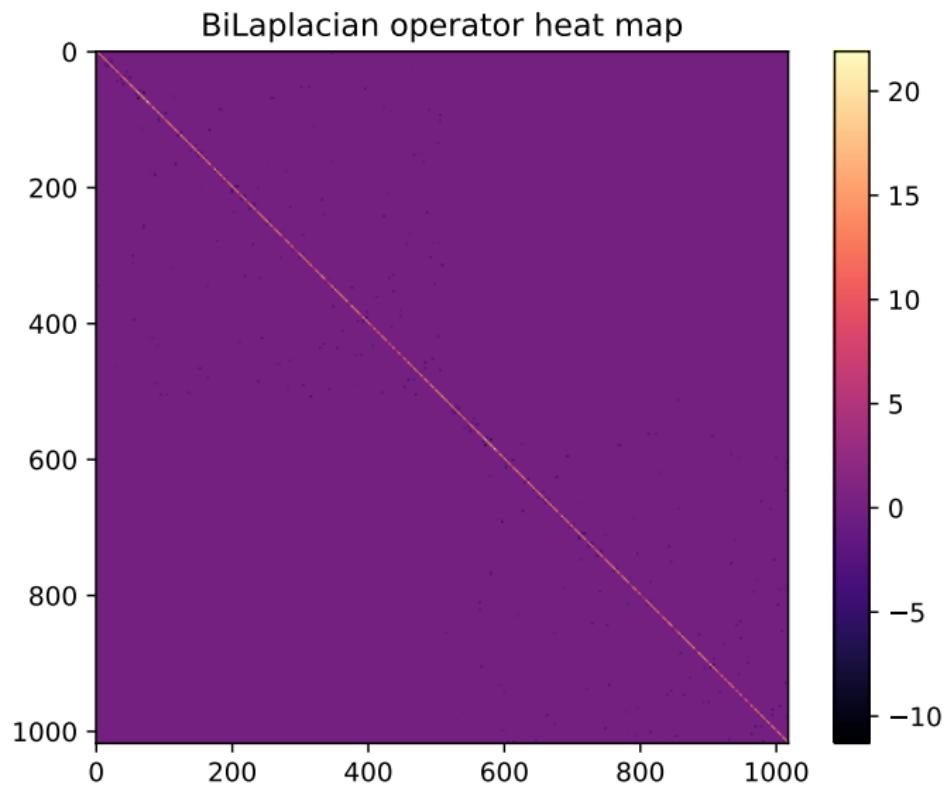
So

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

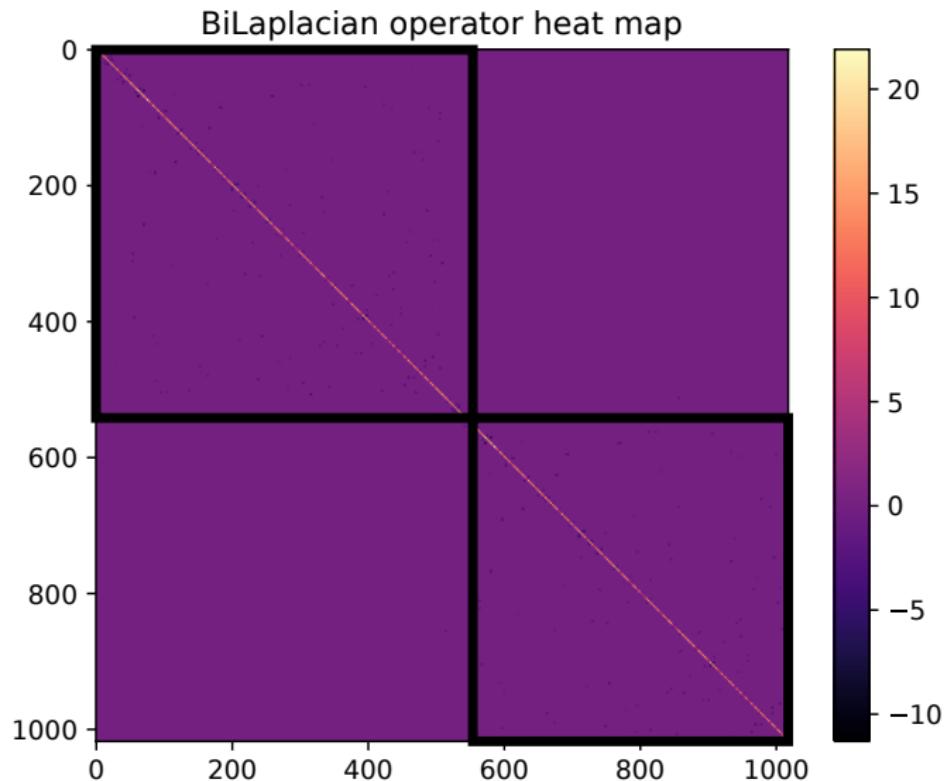
Note:

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

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

# Sparsity: Observations

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

Recall:

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

Note:

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

Construct:

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

Construct:

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

Via the Sherman-Morrison-Woodbury formula:

$$\begin{aligned}\mathbf{D}_r &= \text{Diag} \left( \frac{\lambda_1}{1 + \lambda_1}, \dots, \frac{\lambda_r}{1 + \lambda_r} \right) \\ (\tilde{\mathbf{H}}_{\text{misfit}} + \mathbf{I})^{-1} &= \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* + \mathcal{O} \left( \sum_{k=r+1}^m \frac{\lambda_k}{1 + \lambda_k} \right) \\ &\approx \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^*\end{aligned}$$

Construct:

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

Via the Sherman-Morrison-Woodbury formula:

$$\begin{aligned}\mathbf{D}_r &= \text{Diag} \left( \frac{\lambda_1}{1 + \lambda_1}, \dots, \frac{\lambda_r}{1 + \lambda_r} \right) \\ (\tilde{\mathbf{H}}_{\text{misfit}} + \mathbf{I})^{-1} &= \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* + \mathcal{O} \left( \sum_{k=r+1}^m \frac{\lambda_k}{1 + \lambda_k} \right) \\ &\approx \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^*\end{aligned}$$

Hence:

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

Construct:

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

Via the Sherman-Morrison-Woodbury formula:

$$\begin{aligned}\mathbf{D}_r &= \text{Diag} \left( \frac{\lambda_1}{1 + \lambda_1}, \dots, \frac{\lambda_r}{1 + \lambda_r} \right) \\ (\tilde{\mathbf{H}}_{\text{misfit}} + \mathbf{I})^{-1} &= \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^* + \mathcal{O} \left( \sum_{k=r+1}^m \frac{\lambda_k}{1 + \lambda_k} \right) \\ &\approx \mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^*\end{aligned}$$

Hence:

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

## Sampling from resulting distributions

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

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

$$\mathbf{s} = \mathbf{m} + \mathbf{Ln}$$

# Sampling from resulting distributions

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

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

$$\mathbf{s} = \mathbf{m} + \mathbf{Ln}$$

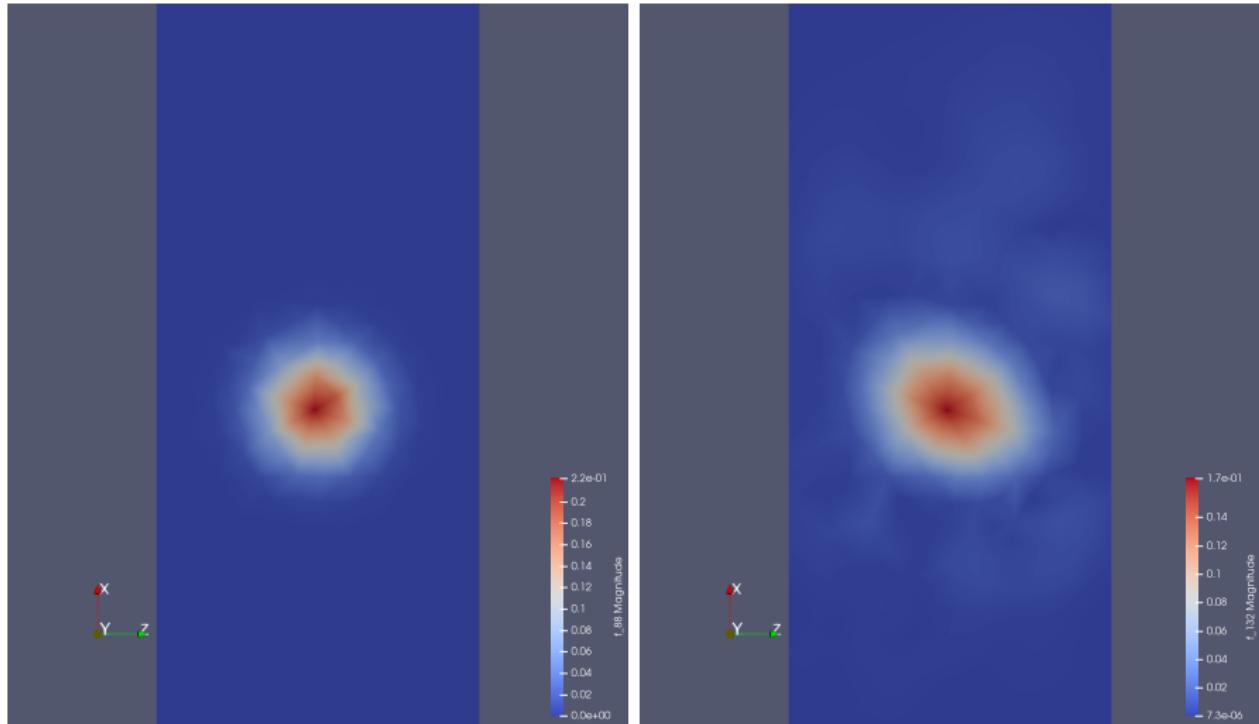
Since  $\mathcal{C}_{\text{pr}} = \mathbf{A}^{-2}$ , prior is easy. For posterior:

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

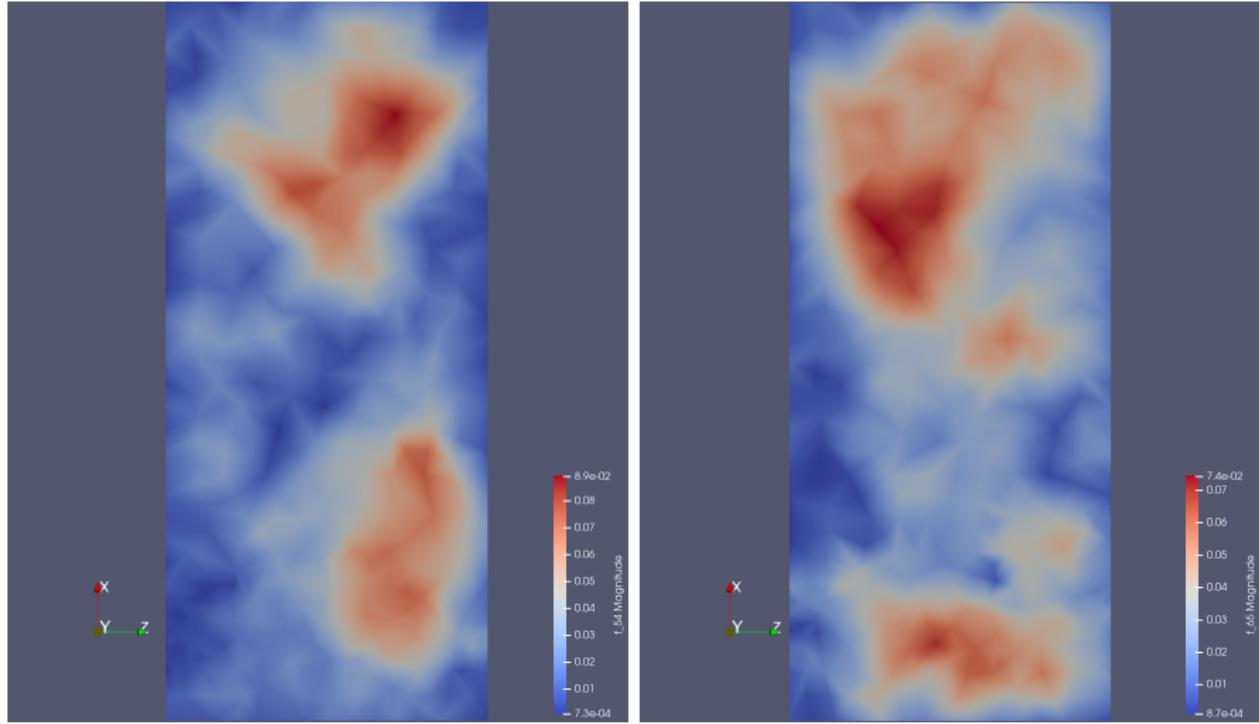
# Table of Contents

- 1 Seismic Inversion Model Problem
- 2 Infinite-Dimensional Inverse Problem Setting
- 3 Numerical Discretization and Considerations
- 4 Results
- 5 Future Work
- 6 References

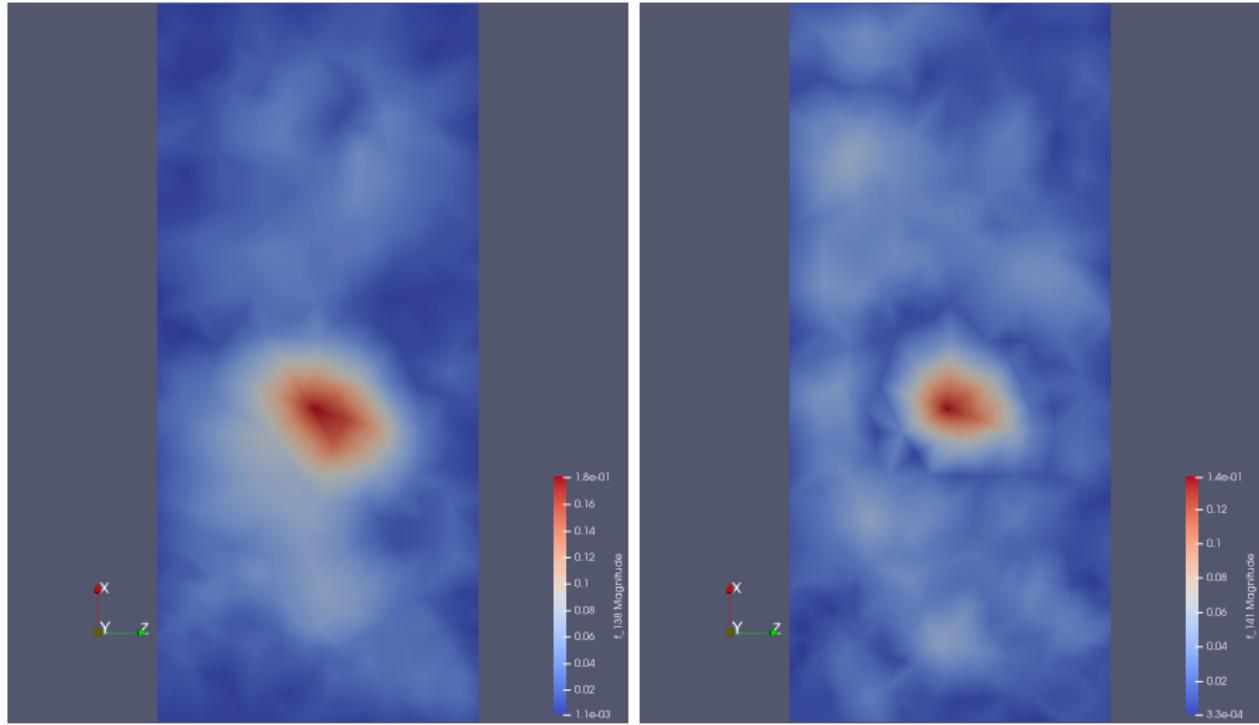
# MAP Estimation vs. True Parameter



# Prior Samples



# Posterior Samples



# Table of Contents

- 1 Seismic Inversion Model Problem
- 2 Infinite-Dimensional Inverse Problem Setting
- 3 Numerical Discretization and Considerations
- 4 Results
- 5 Future Work
- 6 References

# Future Work

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

- It's a reasonable approximation to well-studied physical problem.
- Mixed dimensional formulation in topology and function space.
- Its 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](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](https://doi.org/10.1088/1361-6420/abaf63)

# Table of Contents

- 1 Seismic Inversion Model Problem
- 2 Infinite-Dimensional Inverse Problem Setting
- 3 Numerical Discretization and Considerations
- 4 Results
- 5 Future Work
- 6 References

# 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](https://doi.org/10.1137/20m1347292).
- [2] Martin Alnæs et al. "The FEniCS Project Version 1.5". en. In: *Archive of Numerical Software* Vol 3 (2015). DOI: [10.11588/ANS.2015.100.20553](https://doi.org/10.11588/ANS.2015.100.20553).
- [3] 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](https://doi.org/10.1137/12089586x).
- [4] Robert J Lillie. *Oregon's Island In The Sky: Geology Road Guide to Marys Peak*. English. OCLC: 979996650. 2017. ISBN: 9781540611963.
- [5] Kimberly Alison McCormack. *Earthquakes, groundwater and surface deformation : Exploring the poroelastic response to megathrust earthquakes*. Aug. 2018. URL: [hdl.handle.net/2152/68892](https://hdl.handle.net/2152/68892).

## References II

- [6] Tom Simkin et al. *This dynamic planet: World map of volcanoes, earthquakes, impact craters and plate tectonics*. Tech. rep. 2006. DOI: [10.3133/i2800](https://doi.org/10.3133/i2800).
- [7] A. M. Stuart. "Inverse problems: A Bayesian perspective". In: *Acta Numerica* 19 (May 2010), pp. 451–559. DOI: [10.1017/s0962492910000061](https://doi.org/10.1017/s0962492910000061).
- [8] 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](https://doi.org/10.1088/1361-6420/abaf63).
- [9] Umberto Villa, Noemi Petra, and Omar Ghattas. "HIPPYlib: An Extensible Software Framework for Large-Scale Inverse Problems Governed by PDEs: Part I: Deterministic Inversion and Linearized Bayesian Inference". In: *ACM Trans. Math. Softw.* 47.2 (Apr. 2021). ISSN: 0098-3500. DOI: [10.1145/3428447](https://doi.org/10.1145/3428447).