Sensitivity Analysis of the Information Gain in Infinite-Dimensional Bayesian Linear Inverse Problems

Abhijit Chowdhary[1] Shanyin Tong[2] Georg Stadler[3] Alen Alexanderian[1]

[1]Department of Mathematics
North Carolina State University

[3]Department of Applied Physics and Applied Mathematics
Columbia University

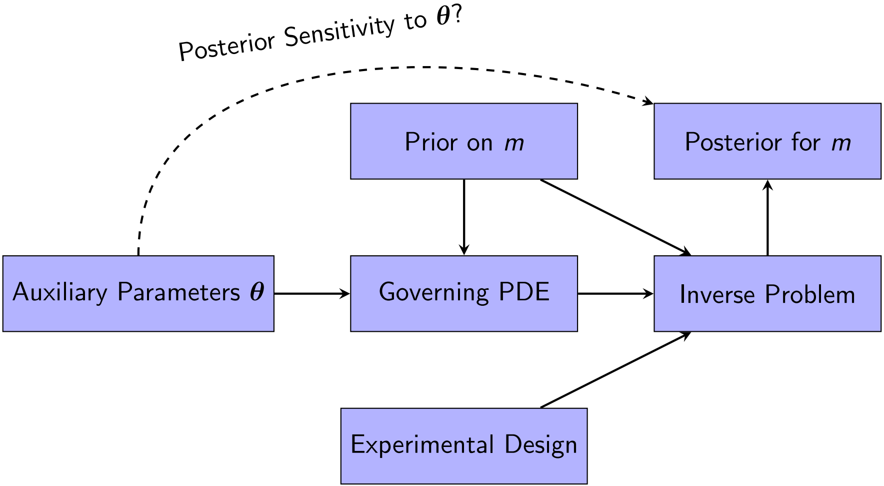
[3]Department of Mathematics
Courant Institute of Mathematical Sciences

SIAM Computational Science and Engineering, 2023

Work done through NSF-DMS-2111044

Table of Contents

Bayesian Inverse Problems Framework



*

$m \in \mathcal{M}$ is the *inversion parameter* and $\boldsymbol{\theta} \in \Theta$ are the *auxiliary parameters*.

Governing PDE Model

For some appropriate function space $\mathcal{V}$, we consider the abstract weak PDE:

$$\mathcal{A}(p, u, \boldsymbol{\theta}) + \mathcal{C}(p, m, \boldsymbol{\theta}) + \mathcal{D}(p, \boldsymbol{\theta}) = 0, \qquad \forall p \in \mathcal{V}$$

- $u \in \mathcal{V}$ is the *state variable*.
- $p \in \mathcal{V}$ is the *adjoint variable*.
- $m \in \mathcal{M}$ is the (potentially infinite dimensional) *inversion parameter*.
- $\boldsymbol{\theta} \in \Theta$ are the (potentially infinite dimensional) *auxiliary parameters*.

Goal

Assuming some measurement data $\mathbf{y}$ obeys the above PDE, infer a probability distribution for m and quantify its sensitivity with respect to $\boldsymbol{\theta}$.

- Let $\boldsymbol{\mathcal{F}} : \mathcal{M} \to \mathbb{R}^{N_{\text{obs}}}$ be the *parameter-to-observable* map.

Bayesian Inversion in Infinite Dimensions

Theorem (Bayes Theorem in Infinite Dimensions)

$$\frac{\mathrm{d}\mu^{\mathbf{y}}_{\mathrm{post}}}{\mathrm{d}\mu_{\mathrm{pr}}} \propto \pi_{\mathrm{like}}(\mathbf{y}|m)$$

- Gaussian Prior $m \sim \mathcal{N}(m_{\mathrm{pr}}, \Gamma_{\mathrm{prior}})$.
- Additive Gaussian noise

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

- Assuming $\mathcal{F}$ linear:

$$\mu^{\mathbf{y}}_{\mathrm{post}} = \mathcal{N}(m_{\mathrm{post}}, \mathcal{C}_{\mathrm{post}})$$

where:

$$m_{\mathrm{post}} = \mathcal{C}_{\mathrm{post}} \left(\mathcal{F}^* \boldsymbol{\Gamma}^{-1}_{\mathrm{noise}} \mathbf{y} + \Gamma^{-1}_{\mathrm{prior}} m_{\mathrm{pr}} \right)$$

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

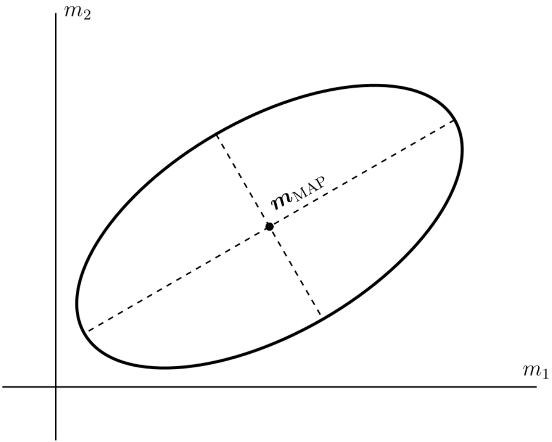


Figure: The confidence region corresponding to bivariate Gaussian posterior.

Table of Contents

Information Gain

Kullback-Leibler Divergence of the Posterior from the Prior

$$D_{\mathrm{kl}}(\mu_{\mathrm{post}}^{\mathbf{y}} || \mu_{\mathrm{pr}}) := \int_{\mathcal{M}} \log \left[\frac{\mathrm{d}\mu_{\mathrm{post}}^{\mathbf{y}}}{\mathrm{d}\mu_{\mathrm{pr}}} \right] \mathrm{d}\mu_{\mathrm{post}}^{\mathbf{y}}$$

- Informally: information gained when updating beliefs to $\mu_{\mathrm{post}}^{\mathbf{y}}$ from μ_{pr}.
- Averaging this over experimental data is the D-optimal design criteria
- In the linear inverse problem setting, possess an analytic expression

*

S. Kullback and R. A. Leibler. "On Information and Sufficiency". In: *The Annals of Mathematical Statistics* 22.1 (Mar. 1951), pp. 79–86. DOI: 10.1214/aoms/1177729694

Adjoint-Based Framework

More generally, m_{post} can be inferred by solving the PDE-constrained optimization problem:

$$\begin{aligned}\mathcal{J}(m, \boldsymbol{\theta}) &:= \frac{1}{2}\|\boldsymbol{\mathcal{F}}(\boldsymbol{\theta})(m) - \mathbf{y}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2}\|m - m_{\text{pr}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{prior}}} \\ &= \frac{1}{2}\|\boldsymbol{\mathcal{Q}} u - \mathbf{y}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2}\|m - m_{\text{pr}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{prior}}}\end{aligned}$$

- u is given by the forward model.
- $\mathbf{y} \in \mathbb{R}^{N_{\text{obs}}}$ where N_{obs} is the number of data points.
- $\boldsymbol{\mathcal{Q}} : \mathcal{V} \to \mathbb{R}^{N_{\text{obs}}}$ is an observation operator.

Gradient and Hessian action implicitly defined by PDEs

Lagrangian Enforcing PDE-Constraints

$$\mathcal{L}(u,p,m) := \frac{1}{2}\|\mathcal{Q}u-\mathbf{y}\|^2_{\Gamma^{-1}_{\text{noise}}} + \frac{1}{2}\|m-m_0\|^2_{\mathcal{C}_0^{-1}} + \mathcal{A}(p,u,\boldsymbol{\theta}) + \mathcal{C}(p,m,\boldsymbol{\theta}) + \mathcal{D}(p,\boldsymbol{\theta})$$

Take variations with respect to (u,p,m) and set to zero to find:

Gradient System

$$\begin{cases} \mathcal{G}(m)(\tilde{m}) = \mathcal{C}(p,\tilde{m},\boldsymbol{\theta}) + \langle m-m_0,\tilde{m}\rangle_{\mathcal{C}_0^{-1}} \\ \langle \mathcal{Q}u-\mathbf{y},\mathcal{Q}\tilde{u}\rangle_{\Gamma^{-1}_{\text{noise}}} + \mathcal{A}(\tilde{u},p,\boldsymbol{\theta}) = 0, & \forall \tilde{u}\in\mathcal{V} \\ \mathcal{A}(\tilde{p},u,\boldsymbol{\theta}) + \mathcal{C}(\tilde{p},m,\boldsymbol{\theta}) + \mathcal{D}(\tilde{p},\boldsymbol{\theta}) = 0, & \forall \tilde{p}\in\mathcal{V} \end{cases}$$

Gradient and Hessian action implicitly defined by PDEs

(Meta)-Lagrangian constraining Gradient System

$$\begin{aligned}\mathcal{L}^H(u, p, m, \hat{u}, \hat{p}, \hat{m}) &= \mathcal{C}(p, \hat{m}, \boldsymbol{\theta}) + \langle m - m_0, \hat{m} \rangle_{\mathcal{C}_0^{-1}} \\ &+ \langle \mathcal{Q}u - \mathbf{y}, \mathcal{Q}\hat{u} \rangle_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}} + \mathcal{A}(\hat{u}, p, \boldsymbol{\theta}) \\ &+ \mathcal{A}(\hat{p}, u, \boldsymbol{\theta}) + \mathcal{C}(\hat{p}, m, \boldsymbol{\theta}) + \mathcal{D}(\hat{p}, \boldsymbol{\theta})\end{aligned}$$

Take variations with respect to (u, p, m) and set to zero to find:

Implicit Hessian Action

$$\begin{cases} \mathcal{H}(\hat{m}, \tilde{m}) = \mathcal{C}(\hat{p}, \tilde{m}, \boldsymbol{\theta}) + \langle \tilde{m}, \hat{m} \rangle_{\mathcal{C}_0^{-1}} \\ \langle \mathcal{Q}\tilde{u}, \mathcal{Q}\hat{u} \rangle_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}} + \mathcal{A}(\hat{p}, \tilde{u}, \boldsymbol{\theta}) = 0, \qquad \forall \tilde{u} \in \mathcal{V} \\ \mathcal{A}(\tilde{p}, \hat{u}, \boldsymbol{\theta}) + \mathcal{C}(\tilde{p}, \hat{m}, \boldsymbol{\theta}) = 0, \qquad \forall \tilde{p} \in \mathcal{V} \end{cases}$$

Gradient and Hessian action implicitly defined by PDEs

(Meta)-Lagrangian constraining Gradient System

$$
\begin{aligned}
\mathcal{L}^H(u,p,m,\hat{u},\hat{p},\hat{m}) = \mathcal{C}(p,\hat{m},\boldsymbol{\theta}) &+ \cancel{\langle m - m_0, \hat{m}\rangle_{\mathcal{C}_0^{-1}}} \\
&+ \langle \mathcal{Q}u - \mathbf{y}, \mathcal{Q}\hat{u}\rangle_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}} + \mathcal{A}(\hat{u},p,\boldsymbol{\theta}) \\
&+ \mathcal{A}(\hat{p},u,\boldsymbol{\theta}) + \mathcal{C}(\hat{p},m,\boldsymbol{\theta}) + \mathcal{D}(\hat{p},\boldsymbol{\theta})
\end{aligned}
$$

Take variations with respect to (u, p, m) and set to zero to find:

Implicit Data-Misfit Hessian

$$
\begin{cases}
\mathcal{H}_{\mathrm{m}}(\hat{m},\tilde{m}) = \mathcal{C}(\hat{p},\tilde{m},\boldsymbol{\theta}) + \cancel{\langle \tilde{m}, \hat{m}\rangle_{\mathcal{C}_0^{-1}}} \\
\langle \mathcal{Q}\tilde{u}, \mathcal{Q}\hat{u}\rangle_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}} + \mathcal{A}(\hat{p},\tilde{u},\boldsymbol{\theta}) = 0, \qquad \forall \tilde{u} \in \mathcal{V} \\
\mathcal{A}(\tilde{p},\hat{u},\boldsymbol{\theta}) + \mathcal{C}(\tilde{p},\hat{m},\boldsymbol{\theta}) = 0, \qquad \forall \tilde{p} \in \mathcal{V}
\end{cases}
$$

Information Gain for Linear Bayesian Inverse Problems

Definition (Explicit KL-Divergence)

For linear Bayesian inverse problems, the KL-Divergence has explicit form

$$\Phi_{\mathrm{KL}} = \frac{1}{2}\left[\log\det\left(\tilde{\mathcal{H}}_{\mathrm{m}} + \mathcal{I}\right) - \mathrm{tr}\left(\tilde{\mathcal{H}}_{\mathrm{m}}(\mathcal{I} + \tilde{\mathcal{H}}_{\mathrm{m}})^{-1}\right) + \left\| m_{\mathsf{post}} - m_{\mathsf{pr}} \right\|^2_{\Gamma^{-1}_{\mathsf{prior}}}\right]$$

where

$$\tilde{\mathcal{H}}_{\mathrm{m}} = \Gamma^{1/2}_{\mathsf{prior}} \mathcal{H}_{\mathrm{m}} \Gamma^{1/2}_{\mathsf{prior}}$$

is the prior-preconditioned data-misfit Hessian

For local sensitivity, we want to compute $\frac{\partial}{\partial \theta_j}\Phi_{\mathrm{KL}}$ for every j.

Information Gain for Linear Bayesian Inverse Problems

Definition (Explicit KL-Divergence)

For linear Bayesian inverse problems, the KL-Divergence has explicit form

$$\Phi_{\mathrm{KL}} = \frac{1}{2}\left[\log\det\left(\tilde{\mathcal{H}}_{\mathrm{m}} + \mathcal{I}\right) - \mathrm{tr}\left(\tilde{\mathcal{H}}_{\mathrm{m}}(\mathcal{I} + \tilde{\mathcal{H}}_{\mathrm{m}})^{-1}\right) + \left\| m_{\text{post}} - m_{\text{pr}} \right\|^2_{\Gamma^{-1}_{\text{prior}}}\right]$$

where

$$\tilde{\mathcal{H}}_{\mathrm{m}} = \Gamma^{1/2}_{\text{prior}} \mathcal{H}_{\mathrm{m}} \Gamma^{1/2}_{\text{prior}}$$

is the prior-preconditioned data-misfit Hessian

For local sensitivity, we want to compute $\frac{\partial}{\partial \theta_j}\Phi_{\mathrm{KL}}$ for every j.

Q: How do we differentiate this expression?

Table of Contents

Low Rank Approximations

Definition ($\tilde{\mathcal{H}}_{\mathrm{m}}$ Low-Rank Approximation)

Let $\{(\lambda_i, \psi_i)\}_{i=1}^{r}$ be the dominant eigenpairs of $\tilde{\mathcal{H}}_{\mathrm{m}}$, then:

$$\tilde{\mathcal{H}}_{\mathrm{m}}\phi \approx \sum_{i=1}^{r} \lambda_i \langle \psi_i, \phi \rangle \psi_i$$

is a low rank approximation of $\tilde{\mathcal{H}}_{\mathrm{m}}$.

Definition (Kullback-Leibler Divergence Low-Rank Approximation)

$$\hat{\Phi}_{\mathrm{KL}} := \frac{1}{2}\sum_{j=1}^{r}\left[\log\left(\lambda_j + 1\right) - \frac{\lambda_j}{1+\lambda_j}\right] + \frac{1}{2}\left\| m_{\mathrm{post}} - m_{\mathrm{pr}} \right\|^2_{\Gamma^{-1}_{\mathrm{prior}}}$$

Hyper Differential Sensitivity Analysis

Q: How do we differentiate this expression?

$$\hat{\Phi}_{\mathrm{KL}}(\boldsymbol{\theta}) = \underbrace{\frac{1}{2}\sum_{j=1}^{r}\left[\log\left(\lambda_j(\boldsymbol{\theta})+1\right) - \frac{\lambda_j(\boldsymbol{\theta})}{1+\lambda_j(\boldsymbol{\theta})}\right]}_{(1)} + \underbrace{\frac{1}{2}\left\| m_{\text{post}}(\boldsymbol{\theta}) - m_{\text{pr}}\right\|^2_{\Gamma^{-1}_{\text{prior}}}}_{(2)}$$

- (1): Eigenvalue Sensitivities
- (2): Post-Optimal Sensitivity Analysis

Adjoint-Based Eigenvalue Sensitivities

Q: How do we differentiate this expression?

$$\hat{\Phi}^{(1)}_{\mathrm{KL}}(\boldsymbol{\theta}) := \sum_{j=1}^{k} \frac{1}{2}\left[\log\left(\lambda_j(\boldsymbol{\theta}) + 1\right) - \frac{\lambda_j(\boldsymbol{\theta})}{1 + \lambda_j(\boldsymbol{\theta})}\right]$$

Recall: Structure around $\mathcal{H}_{\mathrm{m}}$

Assuming dominant eigenpairs are not repeated, $\mathcal{H}_{\mathrm{m}}$ satisfies the eigenproblem

$$\mathcal{H}_{\mathrm{m}}(\psi_i, \phi) = \lambda_i \langle \phi, \psi_i \rangle_{\Gamma^{-1}_{\text{prior}}} \quad \text{and} \quad \langle \psi_i, \psi_i \rangle_{\Gamma^{-1}_{\text{prior}}}$$

for every $\phi \in \mathcal{M}$ and $i = 1, \ldots, r$. Likewise, it has implicit structure

$$\begin{cases} \mathcal{H}_{\mathrm{m}}(\hat{m}, \tilde{m}) = \mathcal{C}(\hat{p}, \tilde{m}, \boldsymbol{\theta}) \\ \langle \boldsymbol{\mathcal{Q}}\tilde{u}, \boldsymbol{\mathcal{Q}}\hat{u} \rangle_{\Gamma^{-1}_{\text{noise}}} + \mathcal{A}(\hat{p}, \tilde{u}, \boldsymbol{\theta}) = 0, \qquad \forall \tilde{u} \in \mathcal{V} \\ \mathcal{A}(\tilde{p}, \hat{u}, \boldsymbol{\theta}) + \mathcal{C}(\tilde{p}, \hat{m}, \boldsymbol{\theta}) = 0, \qquad \forall \tilde{p} \in \mathcal{V} \end{cases}$$

Adjoint-Based Eigenvalue Sensitivities

Lagrangian Constraining $\mathcal{H}_{\mathrm{m}}$ and Eigenproblem

$$
\begin{aligned}
&\mathcal{L}^{\lambda}\left(\{\psi_i\}_{i=1}^{r}, \{\hat{u}_i\}_{i=1}^{r}, \{\hat{p}_i\}_{i=1}^{r}, \{\gamma_i^*\}_{i=1}^{r}, \{\hat{u}_i^*\}_{i=1}^{r}, \{\hat{p}_i^*\}_{i=1}^{r}; \boldsymbol{\theta}\right) \\
&\quad = \frac{1}{2}\sum_{i=1}^{r}\left[\log\left(1+\mathcal{C}(\hat{p}_i, \psi_i; \boldsymbol{\theta})\right) - \frac{\mathcal{C}(\hat{p}_i, \psi_i; \boldsymbol{\theta})}{1+\mathcal{C}(\hat{p}_i, \psi_i; \boldsymbol{\theta})}\right] \\
&\qquad + \sum_{i=1}^{r}\left[\mathcal{A}(\hat{p}_i^*, \hat{u}_i; \boldsymbol{\theta}) + \mathcal{C}(\hat{p}_i^*, \psi_i; \boldsymbol{\theta}) + \mathcal{A}(\hat{p}_i, \hat{u}_i^*; \boldsymbol{\theta}) + \langle \boldsymbol{\mathcal{Q}}\hat{u}_i^*, \boldsymbol{\mathcal{Q}}\hat{u}_i\rangle_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}}\right] \\
&\qquad + \sum_{i=1}^{r}\gamma_i^*\left[\langle \psi_i, \psi_i\rangle_{\boldsymbol{\Gamma}_{\text{prior}}^{-1}} - 1\right]
\end{aligned}
$$

Adjoint-Based Eigenvalue Sensitivities

Resulting Expressions

With Lagrange multipliers

$$\hat{u}_i^* = \frac{\lambda_i}{2(1+\lambda_i)^2}\hat{u}(\psi_i) \quad \text{and} \quad \hat{p}_i^* = \frac{\lambda_i}{2(1+\lambda_i)^2}\hat{p}(\psi_i)$$

We can take derivatives of $\hat{\Phi}_{\mathrm{KL}}^{(1)}$ with the expression:

$$\mathcal{L}_{\theta_j}^{\lambda} = \sum_{i=1}^{r}\left[\frac{\mathcal{C}(\hat{p}_i,\psi_i;\boldsymbol{\theta})}{2(1+\mathcal{C}(\hat{p}_i,\psi_i;\boldsymbol{\theta}))^2}\mathcal{C}_{\theta_j}(\hat{p}_i,\psi_i;\boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(\hat{p}_i^*,\hat{u}_i;\boldsymbol{\theta}) + \mathcal{C}_{\theta_j}(\hat{p}_i^*,\psi_i;\boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(\hat{p}_i,\hat{u}_i^*;\boldsymbol{\theta})\right]$$

Post-Optimal Sensitivity Analysis

Q: How do we differentiate this expression?

$$\hat{\Phi}^{(2)}_{\mathrm{KL}}(\boldsymbol{\theta}) := \frac{1}{2}\| m_{\mathsf{post}}(\boldsymbol{\theta}) - m_{\mathsf{pr}} \|^2_{\Gamma^{-1}_{\mathsf{prior}}}$$

Post-Optimal Sensitivity Analysis

Q: How do we differentiate this expression?

$$\hat{\Phi}^{(2)}_{\mathrm{KL}}(\boldsymbol{\theta}) := \frac{1}{2}\|m_{\mathsf{post}}(\boldsymbol{\theta}) - m_{\mathsf{pr}}\|^2_{\Gamma^{-1}_{\mathsf{prior}}}$$

Post-Optimal Sensitivity Analysis

Previous work demonstrated that

$$\frac{\partial}{\partial \theta_j} m_{\mathsf{post}}(\boldsymbol{\theta}) = \mathcal{H}^{-1}(\boldsymbol{\theta})\mathcal{B}_j(\boldsymbol{\theta})$$

where $\mathcal{B}_j : \Theta_j \to \mathcal{E}$ is the operator describing the Frechet derivative of the gradient $\mathcal{G}$ with respect to θ_j evaluated at m_{post}, i.e. $\mathcal{B} := J_{m,\theta_j}(m_{\mathsf{post}}, \boldsymbol{\theta})$. Thus

$$\frac{\mathrm{d}}{\mathrm{d}\theta_j}\hat{\Phi}^{(2)}_{\mathrm{KL}}(\boldsymbol{\theta}) = 2\langle m_{\mathsf{post}} - m_{\mathsf{pr}}, \frac{\mathrm{d}}{\mathrm{d}\theta_j} m_{\mathsf{post}}\rangle_{\Gamma^{-1}_{\mathsf{prior}}} = -2\langle m_{\mathsf{post}} - m_{\mathsf{pr}}, \mathcal{H}^{-1}\mathcal{B}_j\rangle_{\Gamma^{-1}_{\mathsf{prior}}}$$

Isaac Sunseri et al. "Hyper-differential sensitivity analysis for nonlinear Bayesian inverse problems". In: *International Journal for Uncertainty Quantification* Accepted (2023)

Post-Optimal Sensitivity Analysis

Lagrangian Constraining Gradient System

$$\mathcal{L}^{(2)}(u, p, m, \hat{u}, \hat{p}, \hat{m}, \theta) = \langle m - m_0, \hat{m} \rangle_{\mathcal{C}_0^{-1}} + \mathcal{C}(p, \hat{m}, \boldsymbol{\theta}) + \langle \mathcal{Q}u - \mathbf{y}, \mathcal{Q}\hat{u} \rangle_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}} + \mathcal{A}(\hat{u}, p, \boldsymbol{\theta}) + \mathcal{A}(\hat{p}, u, \boldsymbol{\theta}) + \mathcal{C}(\hat{p}, m, \boldsymbol{\theta}) + \mathcal{D}(\hat{p}, \boldsymbol{\theta})$$

Derivative of $\hat{\Phi}^{(2)}$ with respect to θ_j

$$\mathcal{L}^{(2)}_{\theta_j} = \mathcal{B}_j(\hat{m}, \boldsymbol{\theta}) := \mathcal{C}_{\theta_j}(p, \hat{m}; \boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(p, \hat{u}; \boldsymbol{\theta}) + \mathcal{C}_{\theta_j}(p, \hat{m}; \boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(\hat{p}, u; \boldsymbol{\theta}) + \mathcal{D}_{\theta_j}(\hat{p}; \boldsymbol{\theta})$$

Putting it all together

Local Information Gain Sensitivity with respect to θ_j

$$\frac{\mathrm{d}}{\mathrm{d}\theta_j}\hat{\Phi}_{\mathrm{KL}} = \sum_{i=1}^{r}\left[\tfrac{\mathcal{C}(\hat{p}_i,\psi_i;\boldsymbol{\theta})}{2(1+\mathcal{C}(\hat{p}_i,\psi_i;\boldsymbol{\theta}))^2}\mathcal{C}_{\theta_j}(\hat{p}_i,\psi_i;\boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(\hat{p}_i^*,\hat{u}_i;\boldsymbol{\theta}) + \mathcal{C}_{\theta_j}(\hat{p}_i^*,\psi_i;\boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(\hat{p}_i,\hat{u}_i^*;\boldsymbol{\theta})\right]$$
$$-2\langle m_{\mathrm{post}} - m_{\mathrm{pr}}, \mathcal{H}^{-1}\mathcal{B}_j\rangle_{\Gamma_{\mathrm{prior}}^{-1}}$$

where

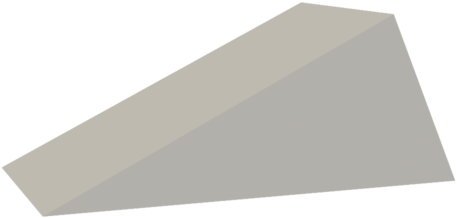
$$\mathcal{B}_j(\hat{m},\boldsymbol{\theta}) := \mathcal{C}_{\theta_j}(p,\hat{m};\boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(p,\hat{u};\boldsymbol{\theta}) + \mathcal{C}_{\theta_j}(p,\hat{m};\boldsymbol{\theta}) + \mathcal{A}_{\theta_j}(\hat{p},u;\boldsymbol{\theta}) + \mathcal{D}_{\theta_j}(\hat{p};\boldsymbol{\theta})$$

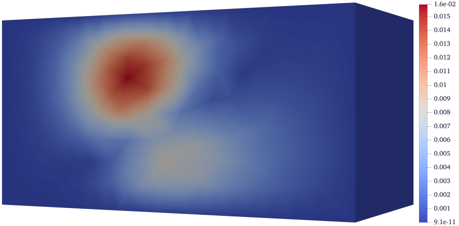
Table of Contents

Fault-Slip Inversion

Governing PDE

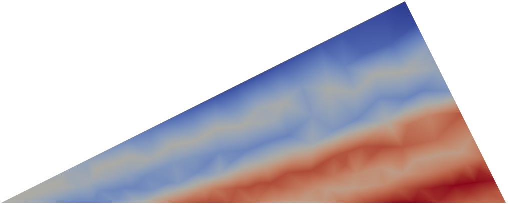
$$
\begin{aligned}
-\nabla \cdot [\mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T) + \lambda \nabla \cdot \boldsymbol{u}\mathbf{I}] &= \mathbf{0} && \text{in } \Omega, \\
\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n} &= 0 && \text{on } \Gamma_t, \\
\boldsymbol{u} + \beta_k \boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n} &= \mathbf{0} && \text{on } \Gamma_k, \\
\boldsymbol{u} + \beta_s \boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n} &= \boldsymbol{h} && \text{on } \Gamma_s, \\
\boldsymbol{u} \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma_b, \\
\delta \mathbf{T}(\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{n}) + \mathbf{T}\boldsymbol{u} &= \boldsymbol{m} && \text{on } \Gamma_b.
\end{aligned}
$$



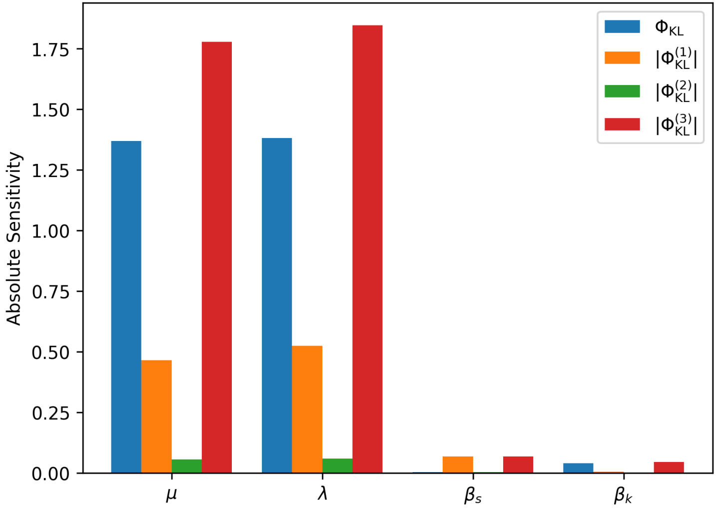


KL-Expansions in the Láme Parameters

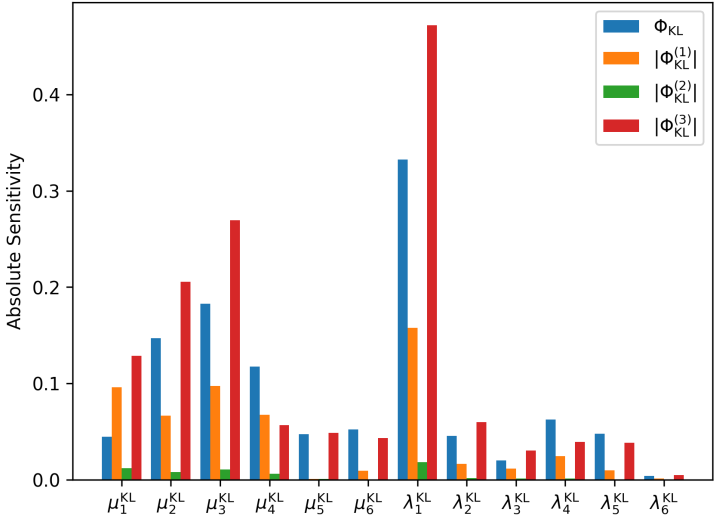
$$\lambda(x,\omega) = \overline{\lambda} + \sum_{i=1}^{\infty} \sqrt{\gamma_i} e_i(x) \xi_i^{\lambda}(\omega) \qquad \text{and} \qquad \mu(x,\omega) = \overline{\mu} + \sum_{i=1}^{\infty} \sqrt{\gamma_i} e_i(x) \xi_i^{\mu}(\omega)$$



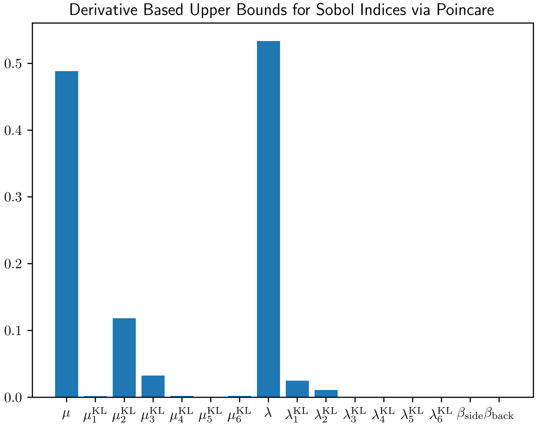
Sensitivity Results



Sensitivity Results



Derivative-Based Sobol Index Upper Bounds



$$S_i^{\text{tot}} \leq \frac{C(F_i)}{V} \int \left(\frac{\partial}{\partial \theta_i} \Phi_{\text{KL}}(\boldsymbol{\theta}) \right)^2 \mathrm{d}F(\boldsymbol{\theta})$$

*

Sergey Kucherenko and Bertrand Iooss. "Derivative-Based Global Sensitivity Measures". In: (2016). Ed. by Roger Ghanem, David Higdon, and Houman Owhadi, pp. 1–24. DOI: 10.1007/978-3-319-11259-6_36-1

Conclusion

In summary,

- Exact expressions scalable to infinite-dimensional parameters
- Feasible to use to perform global sensitivity analysis
- Only requires variational derivatives of the weak form with respect to $\boldsymbol{\theta}$.

*

Abhijit Chowdhary et al. "Sensitivity Analysis of the Information Gain in Infinite-Dimensional Bayesian Linear Inverse Problems". en. In: *International Journal for Uncertainty Quantification* (2024). ISSN: 2152-5080. DOI: 10.1615/Int.J.UncertaintyQuantification.2024051416. URL: http://dl.begellhouse.com/journals/52034eb04b657aea,forthcoming,51416.html