Enabling and interpreting hyper-differential sensitivity analysis for Bayesian inverse problems

Joseph Hart†
with Bart van Bloemen Waanders†

† Sandia National Laboratories[1]
Center for Computing Research
Department for Scientific Machine Learning

SIAM Conference on Uncertainty Quantification

[1] Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under

boratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LL
Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE

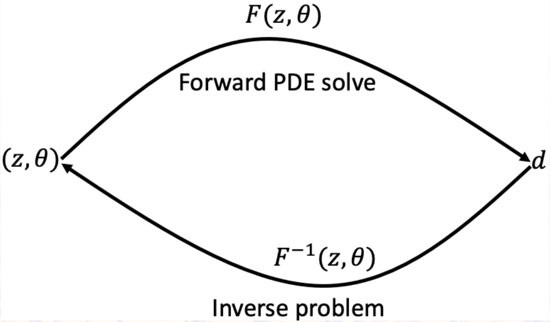
Inverse problems

Find parameters (z, θ) such that

$$F(z, \theta) = \mathcal{Q}(u(z, \theta)) \approx \boldsymbol{d}$$

where

- $\boldsymbol{d}$ are sparse and noisy observations of a state variable $u(z, \theta)$.
- $u(z, \theta)$ is the solution of a PDE and $\mathcal{Q}$ is the observation operator.



The joint Bayesian formulation

- Assume a prior distribution for $(\boldsymbol{z}, \boldsymbol{\theta}) \sim N((\boldsymbol{z}_{\text{prior}}, \boldsymbol{\theta}_{\text{prior}}), \Gamma_{\text{prior}})$.
- Assume that $\boldsymbol{d} = F(\boldsymbol{z}^\star, \boldsymbol{\theta}^\star) + \epsilon$ where $\epsilon \sim N(0, \Gamma_{\text{noise}})$.
- The joint posterior probability density function (PDF) is

$$\pi_{\text{post}}(\boldsymbol{z}, \boldsymbol{\theta}) \propto \pi_{\text{like}}(\boldsymbol{d} | \boldsymbol{z}, \boldsymbol{\theta}) \pi_{\text{prior}}(\boldsymbol{z}, \boldsymbol{\theta})$$

where

- π_{like} is the likelihood function,
- π_{prior} is the prior PDF.

Analyzing properties of $\pi_{\text{post}}(\boldsymbol{z}, \boldsymbol{\theta})$ provide a wealth of information, but may be computationally intractable.

The conditional Bayesian formulation

- The posterior probability density function (PDF) for $\boldsymbol{z}$ given $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{prior}}$ is

$$\pi_{\text{post}}(\boldsymbol{z}|\boldsymbol{\theta}_{\text{prior}}) \propto \pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})\pi_{\text{prior}}(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}}).$$

- The maximum a posteriori probability (MAP) point(s) for $\pi_{\text{post}}(\boldsymbol{z}|\boldsymbol{\theta}_{\text{prior}})$ are local minima of

$$\min_{\boldsymbol{z} \in \mathbb{R}^m} J(\boldsymbol{z}; \boldsymbol{\theta}_{\text{prior}}) := M(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}}) + R(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$$

where $M(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$ and $R(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$ are the negative log likelihood and prior.

The conditional Bayesian formulation

- The posterior probability density function (PDF) for $\boldsymbol{z}$ given $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{prior}}$ is

$$\pi_{\text{post}}(\boldsymbol{z}|\boldsymbol{\theta}_{\text{prior}}) \propto \pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})\pi_{\text{prior}}(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}}).$$

- The maximum a posteriori probability (MAP) point(s) for $\pi_{\text{post}}(\boldsymbol{z}|\boldsymbol{\theta}_{\text{prior}})$ are local minima of

$$\min_{\boldsymbol{z} \in \mathbb{R}^m} J(\boldsymbol{z}; \boldsymbol{\theta}_{\text{prior}}) := M(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}}) + R(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$$

where $M(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$ and $R(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$ are the negative log likelihood and prior.

- How does fixing $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{prior}}$ influence the analysis for $\boldsymbol{z}$?

1. Bayesian inverse problems
2. Hyper-differential sensitivity analysis (HDSA)
3. The Bayesian interpretation of HDSA
4. Enabling HDSA for ill-posed problems
5. Numerical Results

Post optimality sensitivity analysis

$$\min_{\boldsymbol{z} \in \mathbb{R}^m} J(\boldsymbol{z}; \boldsymbol{\theta}_{\text{prior}}) := M(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}}) + R(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$$

- Let $\boldsymbol{z}^\star$ denote a local minimum when $\boldsymbol{\theta} = \boldsymbol{\theta}_{prior}$ is fixed,

$$\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star, \boldsymbol{\theta}_{\text{prior}}) = 0 \qquad \text{and} \qquad \nabla_{\boldsymbol{z},\boldsymbol{z}} J(\boldsymbol{z}^\star, \boldsymbol{\theta}_{\text{prior}}) \succ 0$$

Post optimality sensitivity analysis

$$\min_{\boldsymbol{z}\in\mathbb{R}^m} J(\boldsymbol{z};\boldsymbol{\theta}_{\text{prior}}) := M(\boldsymbol{z},\boldsymbol{\theta}_{\text{prior}}) + R(\boldsymbol{z},\boldsymbol{\theta}_{\text{prior}})$$

- Let $\boldsymbol{z}^\star$ denote a local minimum when $\boldsymbol{\theta} = \boldsymbol{\theta}_{prior}$ is fixed,

$$\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star,\boldsymbol{\theta}_{\text{prior}}) = 0 \qquad \text{and} \qquad \nabla_{\boldsymbol{z},\boldsymbol{z}} J(\boldsymbol{z}^\star,\boldsymbol{\theta}_{\text{prior}}) \succ 0$$

- The implicit function theorem gives

$$\mathcal{G} : \mathcal{N}(\boldsymbol{\theta}_{\text{prior}}) \to \mathcal{N}(\boldsymbol{z}^\star),$$

defined on neighborhoods of $\boldsymbol{\theta}_{\text{prior}}$ and $\boldsymbol{z}^\star$, such that

$$\nabla_{\boldsymbol{z}} J(\mathcal{G}(\boldsymbol{\theta}),\boldsymbol{\theta}) = 0 \qquad \forall \boldsymbol{\theta} \in \mathcal{N}(\boldsymbol{\theta}_{\text{prior}})$$

Post optimality sensitivity analysis

$$\mathcal{G} : \mathcal{N}(\boldsymbol{\theta}_{\text{prior}}) \to \mathcal{N}(\boldsymbol{z}^\star)$$

- $\mathcal{G}$ associates parameters $\boldsymbol{\theta}$ with the corresponding MAP points for $\boldsymbol{z}$ given $\boldsymbol{\theta}$
- Further, $\mathcal{G}$ is differentiable at $\boldsymbol{\theta}_{\text{prior}}$ and its Jacobian is given by

$$\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}}) = -\mathcal{H}^{-1}\mathcal{B} \in \mathbb{R}^{m \times n}$$

- $\mathcal{H} = \nabla_{\boldsymbol{z},\boldsymbol{z}} J$ and $\mathcal{B} = \nabla_{\boldsymbol{z},\boldsymbol{\theta}} J$ are second derivatives of the objective $J(\boldsymbol{z}, \boldsymbol{\theta})$

Post optimality sensitivity analysis

$$\mathcal{G} : \mathcal{N}(\boldsymbol{\theta}_{\text{prior}}) \to \mathcal{N}(\boldsymbol{z}^\star)$$

- $\mathcal{G}$ associates parameters $\boldsymbol{\theta}$ with the corresponding MAP points for $\boldsymbol{z}$ given $\boldsymbol{\theta}$
- Further, $\mathcal{G}$ is differentiable at $\boldsymbol{\theta}_{\text{prior}}$ and its Jacobian is given by

$$\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}}) = -\mathcal{H}^{-1}\mathcal{B} \in \mathbb{R}^{m \times n}$$

- $\mathcal{H} = \nabla_{\boldsymbol{z},\boldsymbol{z}} J$ and $\mathcal{B} = \nabla_{\boldsymbol{z},\boldsymbol{\theta}} J$ are second derivatives of the objective $J(\boldsymbol{z}, \boldsymbol{\theta})$

- $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ is a Newton step updating the MAP point given a perturbation of $\boldsymbol{\theta}$.
- HDSA efficiently interrogates $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ using adjoint-based derivative calculations and matrix-free linear algebra.

Post optimality sensitivity analysis

$$\mathcal{G} : \mathcal{N}(\boldsymbol{\theta}_{\text{prior}}) \to \mathcal{N}(\boldsymbol{z}^\star)$$

- $\mathcal{G}$ associates parameters $\boldsymbol{\theta}$ with the corresponding MAP points for $\boldsymbol{z}$ given $\boldsymbol{\theta}$
- Further, $\mathcal{G}$ is differentiable at $\boldsymbol{\theta}_{\text{prior}}$ and its Jacobian is given by

$$\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}}) = -\mathcal{H}^{-1}\mathcal{B} \in \mathbb{R}^{m \times n}$$

- $\mathcal{H} = \nabla_{\boldsymbol{z},\boldsymbol{z}} J$ and $\mathcal{B} = \nabla_{\boldsymbol{z},\boldsymbol{\theta}} J$ are second derivatives of the objective $J(\boldsymbol{z}, \boldsymbol{\theta})$

- $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ is a Newton step updating the MAP point given a perturbation of $\boldsymbol{\theta}$.
- HDSA efficiently interrogates $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ using adjoint-based derivative calculations and matrix-free linear algebra.

1. What is the Bayesian interpretation of $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$?

Post optimality sensitivity analysis

$$\mathcal{G} : \mathcal{N}(\boldsymbol{\theta}_{\text{prior}}) \to \mathcal{N}(\boldsymbol{z}^\star)$$

- $\mathcal{G}$ associates parameters $\boldsymbol{\theta}$ with the corresponding MAP points for $\boldsymbol{z}$ given $\boldsymbol{\theta}$
- Further, $\mathcal{G}$ is differentiable at $\boldsymbol{\theta}_{\text{prior}}$ and its Jacobian is given by

$$\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}}) = -\mathcal{H}^{-1}\mathcal{B} \in \mathbb{R}^{m \times n}$$

- $\mathcal{H} = \nabla_{\boldsymbol{z},\boldsymbol{z}} J$ and $\mathcal{B} = \nabla_{\boldsymbol{z},\boldsymbol{\theta}} J$ are second derivatives of the objective $J(\boldsymbol{z}, \boldsymbol{\theta})$

- $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ is a Newton step updating the MAP point given a perturbation of $\boldsymbol{\theta}$.
- HDSA efficiently interrogates $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ using adjoint-based derivative calculations and matrix-free linear algebra.

1. What is the Bayesian interpretation of $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$?

2. Can I still compute/interpret $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ for ill-posed problems where the optimizer may struggle to solve the MAP point estimation problem to optimality (satisfaction of the first order optimality condition)?

The linear case

Theorem

If $F(\boldsymbol{z},\boldsymbol{\theta}) = A\boldsymbol{z} + B\boldsymbol{\theta}$ *then the posterior is Gaussian with covariance*

$$\Sigma_{post} = \begin{pmatrix} \Sigma_{\boldsymbol{z},\boldsymbol{z}} & \Sigma_{\boldsymbol{z},\boldsymbol{\theta}} \\ \Sigma_{\boldsymbol{\theta},\boldsymbol{z}} & \Sigma_{\boldsymbol{\theta},\boldsymbol{\theta}} \end{pmatrix}$$

and the post-optimality sensitivity is given by

$$\mathcal{G}'(\boldsymbol{\theta}_{prior}) = \Sigma_{\boldsymbol{z},\boldsymbol{\theta}}\Sigma_{\boldsymbol{\theta},\boldsymbol{\theta}}^{-1}.$$

- The post-optimality sensitivity is a correlation between $\boldsymbol{z}$ and $\boldsymbol{\theta}$.

The linear case

Theorem

If $F(\boldsymbol{z},\boldsymbol{\theta}) = A\boldsymbol{z} + B\boldsymbol{\theta}$ then the posterior is Gaussian with covariance

$$\Sigma_{post} = \begin{pmatrix} \Sigma_{\boldsymbol{z},\boldsymbol{z}} & \Sigma_{\boldsymbol{z},\boldsymbol{\theta}} \\ \Sigma_{\boldsymbol{\theta},\boldsymbol{z}} & \Sigma_{\boldsymbol{\theta},\boldsymbol{\theta}} \end{pmatrix}$$

and the post-optimality sensitivity is given by

$$\mathcal{G}'(\boldsymbol{\theta}_{prior}) = \Sigma_{\boldsymbol{z},\boldsymbol{\theta}} \Sigma_{\boldsymbol{\theta},\boldsymbol{\theta}}^{-1}.$$

- The post-optimality sensitivity is a correlation between $\boldsymbol{z}$ and $\boldsymbol{\theta}$.
- Connection between optimization/analysis and Bayesian statistics.

The linear case

Theorem

If $F(\boldsymbol{z},\boldsymbol{\theta}) = A\boldsymbol{z} + B\boldsymbol{\theta}$ *then the posterior is Gaussian with covariance*

$$\Sigma_{post} = \begin{pmatrix} \Sigma_{\boldsymbol{z},\boldsymbol{z}} & \Sigma_{\boldsymbol{z},\boldsymbol{\theta}} \\ \Sigma_{\boldsymbol{\theta},\boldsymbol{z}} & \Sigma_{\boldsymbol{\theta},\boldsymbol{\theta}} \end{pmatrix}$$

and the post-optimality sensitivity is given by

$$\mathcal{G}'(\boldsymbol{\theta}_{prior}) = \Sigma_{\boldsymbol{z},\boldsymbol{\theta}}\Sigma_{\boldsymbol{\theta},\boldsymbol{\theta}}^{-1}.$$

- The post-optimality sensitivity is a correlation between $\boldsymbol{z}$ and $\boldsymbol{\theta}$.
- Connection between optimization/analysis and Bayesian statistics.
- Local correlation for nonlinear inverse problems (Laplace approximation).

Handling ill-posedness

- For ill-posed inverse problems, $\mathcal{H}$ may be ill-conditioned yielding high sensitivity as a result of lacking information/data.
- Analyzing $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ results in sensitivities dominated by what the data does not tell you.

[2] T. Cui, J. Martin, Y. M. Marzouk, A. Solonen, and A. Spantini, Likelihood-informed dimension reduction for nonlinear inverse problems, Inverse Problems, 30 (2014), pp. 1–28.

Handling ill-posedness

- For ill-posed inverse problems, $\mathcal{H}$ may be ill-conditioned yielding high sensitivity as a result of lacking information/data.
- Analyzing $\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}})$ results in sensitivities dominated by what the data does not tell you.

Proposed Approach: Compute sensitivities

$$\mathcal{P}\mathcal{G}'(\boldsymbol{\theta}_{\text{prior}}) = -\mathcal{P}\mathcal{H}^{-1}\mathcal{B}$$

where $\mathcal{P}$ projects onto the likelihood informed subspace [2] defined by the leading eigenvectors of

$$\mathcal{H}_M \boldsymbol{v}_j = \lambda_j \mathcal{H}_R \boldsymbol{v}_j.$$

The eigenvalues

$$\lambda_j = \frac{\boldsymbol{v}_j^T \mathcal{H}_M \boldsymbol{v}_j}{\boldsymbol{v}_j^T \mathcal{H}_R \boldsymbol{v}_j}$$

measure the ratio of the likelihood and prior in the direction of $\boldsymbol{v}_j$.

[2] T. Cui, J. Martin, Y. M. Marzouk, A. Solonen, and A. Spantini, Likelihood-informed dimension reduction for nonlinear inverse problems, Inverse Problems, 30 (2014), pp. 1–28.

Failure to satisfy the first order optimality condition?

It may not be practical to solve the MAP point estimation problem

$$\min_{\boldsymbol{z} \in \mathbb{R}^m} J(\boldsymbol{z}; \boldsymbol{\theta}_{\text{prior}}) := M(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}}) + R(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$$

to optimality if ill-conditioning yields slow convergence.

- Early iterations refine features which are well informed by the data.
- Ill-conditioning may yield slow convergence in the uniformed subspace.

Failure to satisfy the first order optimality condition?

It may not be practical to solve the MAP point estimation problem

$$\min_{\boldsymbol{z} \in \mathbb{R}^m} J(\boldsymbol{z}; \boldsymbol{\theta}_{\text{prior}}) := M(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}}) + R(\boldsymbol{z}, \boldsymbol{\theta}_{\text{prior}})$$

to optimality if ill-conditioning yields slow convergence.

- Early iterations refine features which are well informed by the data.
- Ill-conditioning may yield slow convergence in the uniformed subspace.

Question: HDSA assumes satisfaction of the optimality criteria. What can I do when converging the optimization is impractical/unnecessary?

Failure to satisfy the first order optimality condition?

Idea: Compute sensitivities of a nearby problem which is solved to optimality.

Failure to satisfy the first order optimality condition?

Idea: Compute sensitivities of a nearby problem which is solved to optimality.

- Assume that $\boldsymbol{z}^\star$ is an approximation of the MAP point but

$$\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star, \boldsymbol{\theta}_{\text{prior}}) \neq 0.$$

- Find a minimum norm perturbation $\tilde{R}$ so that $\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star; \boldsymbol{\theta}_{\text{prior}}) + \nabla_{\boldsymbol{z}} \tilde{R}(\boldsymbol{z}^\star) = 0$,

$$\min_{\tilde{R} \in Q} ||\tilde{R}||_{L^1(\mu)}$$
$$\text{s.t. } \nabla_{\boldsymbol{z}} \tilde{R}(\boldsymbol{z}^\star) = -\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star; \boldsymbol{\theta}_{\text{prior}})$$

Failure to satisfy the first order optimality condition?

Idea: Compute sensitivities of a nearby problem which is solved to optimality.

- Assume that $\boldsymbol{z}^\star$ is an approximation of the MAP point but

$$\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star, \boldsymbol{\theta}_{\text{prior}}) \neq 0.$$

- Find a minimum norm perturbation $\tilde{R}$ so that $\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star; \boldsymbol{\theta}_{\text{prior}}) + \nabla_{\boldsymbol{z}} \tilde{R}(\boldsymbol{z}^\star) = 0$,

$$\min_{\tilde{R} \in Q} ||\tilde{R}||_{L^1(\mu)}$$

$$\text{s.t. } \nabla_{\boldsymbol{z}} \tilde{R}(\boldsymbol{z}^\star) = -\nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star; \boldsymbol{\theta}_{\text{prior}})$$

- Judicious linear algebra gives a closed form solution

$$\tilde{R}(\boldsymbol{z}) = \frac{\alpha}{2}||\boldsymbol{g}||_2 - (\boldsymbol{z} - \boldsymbol{z}^\star)^T \boldsymbol{g} + \frac{1}{2}(\boldsymbol{z} - \boldsymbol{z}^\star)^T \frac{1}{\alpha||\boldsymbol{g}||_2} \boldsymbol{g}\boldsymbol{g}^T (\boldsymbol{z} - \boldsymbol{z}^\star),$$

where $\boldsymbol{g} = \nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star; \boldsymbol{\theta}_{\text{prior}})$.

The perturbed MAP point problem

$$\min_{\boldsymbol{z} \in \mathbb{R}^m} J(\boldsymbol{z}; \boldsymbol{\theta}_{\text{prior}}) + \tilde{R}(\boldsymbol{z})$$

- $\boldsymbol{z}^\star$ satisfies the first order optimality condition.
- Post-optimality sensitivities are well defined for this perturbed problem.

The perturbed MAP point problem

$$\min_{\boldsymbol{z} \in \mathbb{R}^m} J(\boldsymbol{z}; \boldsymbol{\theta}_{\text{prior}}) + \tilde{R}(\boldsymbol{z})$$

- $\boldsymbol{z}^\star$ satisfies the first order optimality condition.
- Post-optimality sensitivities are well defined for this perturbed problem.

Some important questions:

- What is the Bayesian interpretation of $\tilde{R}$?
- How does the perturbation influence the sensitivities?

A perturbed Gaussian prior

- What is the Bayesian interpretation of $\tilde{R}$?
- How does the perturbation influence the sensitivities?

Theorem

The perturbed inverse problem has a Gaussian prior with mean

$$\tilde{\boldsymbol{z}}_{prior} = \boldsymbol{z}_{prior} + \frac{\alpha - (\boldsymbol{z}^\star - \boldsymbol{z}_{prior})^T \boldsymbol{s}}{\alpha - \boldsymbol{v}^T \boldsymbol{s}} \boldsymbol{v}$$

and covariance

$$\tilde{\Gamma}_{prior} = \Gamma_{prior} - \frac{1}{||\boldsymbol{g}||_2} \frac{1}{\alpha - \boldsymbol{v}^T \boldsymbol{s}} \boldsymbol{v}\boldsymbol{v}^T$$

where

$$\boldsymbol{g} = \nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star; \boldsymbol{\theta}_{prior}), \qquad \boldsymbol{s} = -\frac{\boldsymbol{g}}{||\boldsymbol{g}||_2}, \qquad \text{and} \qquad \boldsymbol{v} = \Gamma_{prior} \boldsymbol{g}.$$

The perturbation shifts the mean and reduces uncertainty in the direction $\boldsymbol{v}$.

Difference in sensitivities

- What is the Bayesian interpretation of $\tilde{R}$?
- How does the perturbation influence the sensitivities?

Theorem

The quantities

$$S(\overline{\boldsymbol{\theta}}) = \left|\left|\mathcal{P}\mathcal{H}^{-1}\mathcal{B}\overline{\boldsymbol{\theta}}\right|\right|_{W_z} \qquad \text{and} \qquad \tilde{S}(\overline{\boldsymbol{\theta}}) = \left|\left|\mathcal{P}\tilde{\mathcal{H}}^{-1}\tilde{\mathcal{B}}\overline{\boldsymbol{\theta}}\right|\right|_{W_z}$$

satisfy

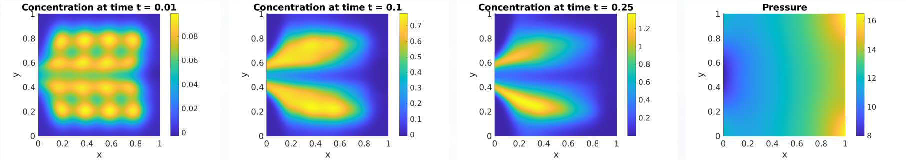
$$\frac{|\tilde{S}(\overline{\boldsymbol{\theta}}) - S(\overline{\boldsymbol{\theta}})|}{\left|\left|\mathcal{H}^{-1}\mathcal{B}\overline{\boldsymbol{\theta}}\right|\right|_{W_z}} \leq \frac{||\mathcal{P}\boldsymbol{n}||_{W_z}}{\boldsymbol{s}^T\boldsymbol{n} + \alpha},$$

where

$$\boldsymbol{g} = \nabla_{\boldsymbol{z}} J(\boldsymbol{z}^\star; \boldsymbol{\theta}_{prior}), \qquad \boldsymbol{s} = -\frac{\boldsymbol{g}}{||\boldsymbol{g}||_2}, \qquad \text{and} \qquad \boldsymbol{n} = -\mathcal{H}^{-1}\boldsymbol{g}.$$

Sandia National Laboratories

Subsurface permeability inversion



$$-\nabla\cdot(e^{\kappa}\nabla p)=0 \qquad \text{in } \Omega$$

$$c_t-\nabla\cdot\Big(\epsilon(\boldsymbol{\theta})\nabla c\Big)+\nabla\cdot\big(-e^{\kappa}\nabla pc\big)=g(\boldsymbol{\theta}) \qquad \text{in } [0,T]\times\Omega$$

$$p=p_1(\boldsymbol{\theta}) \qquad \text{on } \Gamma_1$$

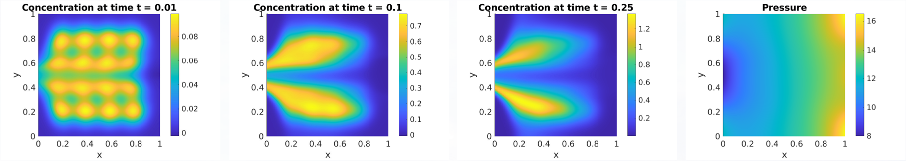
$$p=p_2(\boldsymbol{\theta}) \qquad \text{on } \Gamma_3$$

$$e^{\kappa}\nabla p\cdot n=0 \qquad \text{on } \Gamma_0\cup\Gamma_2$$

$$\nabla c\cdot n=0 \qquad \text{on } [0,T]\times\{\Gamma_0\cup\Gamma_1\cup\Gamma_2\cup\Gamma_3\}$$

$$c(0,x)=0 \qquad \text{in } \Omega$$

Subsurface permeability inversion



$$-\nabla \cdot (e^{\kappa} \nabla p) = 0 \qquad \text{in } \Omega$$

$$c_t - \nabla \cdot \Big(\epsilon(\boldsymbol{\theta}) \nabla c\Big) + \nabla \cdot \big(- e^{\kappa} \nabla p c \big) = g(\boldsymbol{\theta}) \qquad \text{in } [0,T] \times \Omega$$

$$p = p_1(\boldsymbol{\theta}) \qquad \text{on } \Gamma_1$$

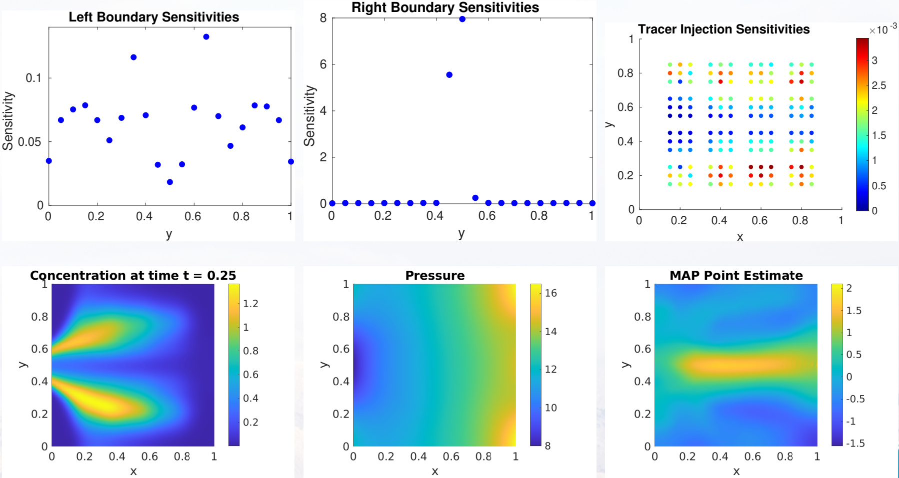
$$p = p_2(\boldsymbol{\theta}) \qquad \text{on } \Gamma_3$$

$$e^{\kappa} \nabla p \cdot n = 0 \qquad \text{on } \Gamma_0 \cup \Gamma_2$$

$$\nabla c \cdot n = 0 \qquad \text{on } [0,T] \times \{\Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_3\}$$

$$c(0,x) = 0 \qquad \text{in } \Omega$$

Sensitivities



Summary

- Established the Bayesian interpretation of post-optimality sensitivity analysis.
- Addressed ill-posedness by projecting on likelihood informed subspaces.
- Theoretically justified HDSA when optimization fails to converge.
- Provided strong error bounds establishing the robust of the analysis.

Joseph Hart, Sandia National Laboratories
joshart@sandia.gov