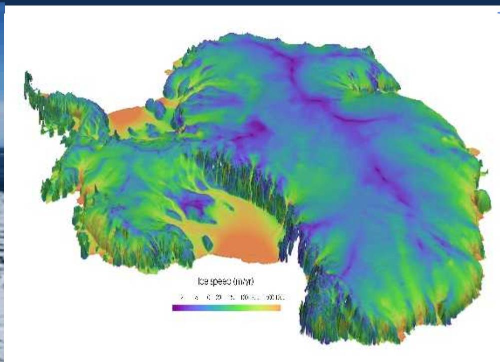
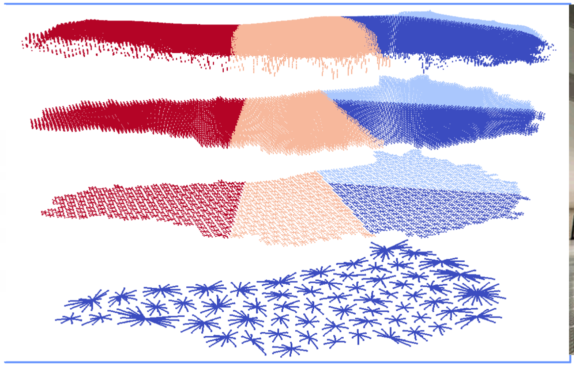
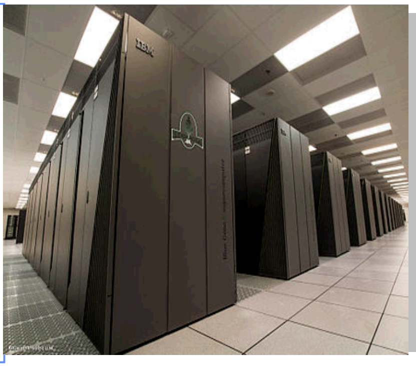
SAND2017-6948C







Large-scale Deterministic Inversion and Bayesian Calibration in Land-Ice Modeling

I. Tezaur[1], J. Jakeman[1], M. Eldred[1], M. Perego[1], S. Price[2], A. Salinger[1]

[1] Sandia National Laboratories, Albuquerque, NM and Livermore, CA, USA.
[2] Los Alamos National Laboratory, Los Alamos, NM, USA.

USNCCM 2017 Montreal, Quebec July 17-20, 2017

Sandia National Laboratories is a multi-mission 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 contract DE-NA0003525.

Outline

1. Background.
   - PISCEES project for land-ice modeling.
   - Land-ice model.
2. UQ problem definition.
3. Inversion/calibration.
   - Deterministic inversion.
   - Bayesian inference.
4. Summary & future work.

Outline

1. **Background.**
   - **PISCEES project for land-ice modeling.**
   - Land-ice model.
2. UQ problem definition.
3. Inversion/calibration.
   - Deterministic inversion.
   - Bayesian inference.
4. Summary & future work.

PISCEES Project for Land-Ice Modeling

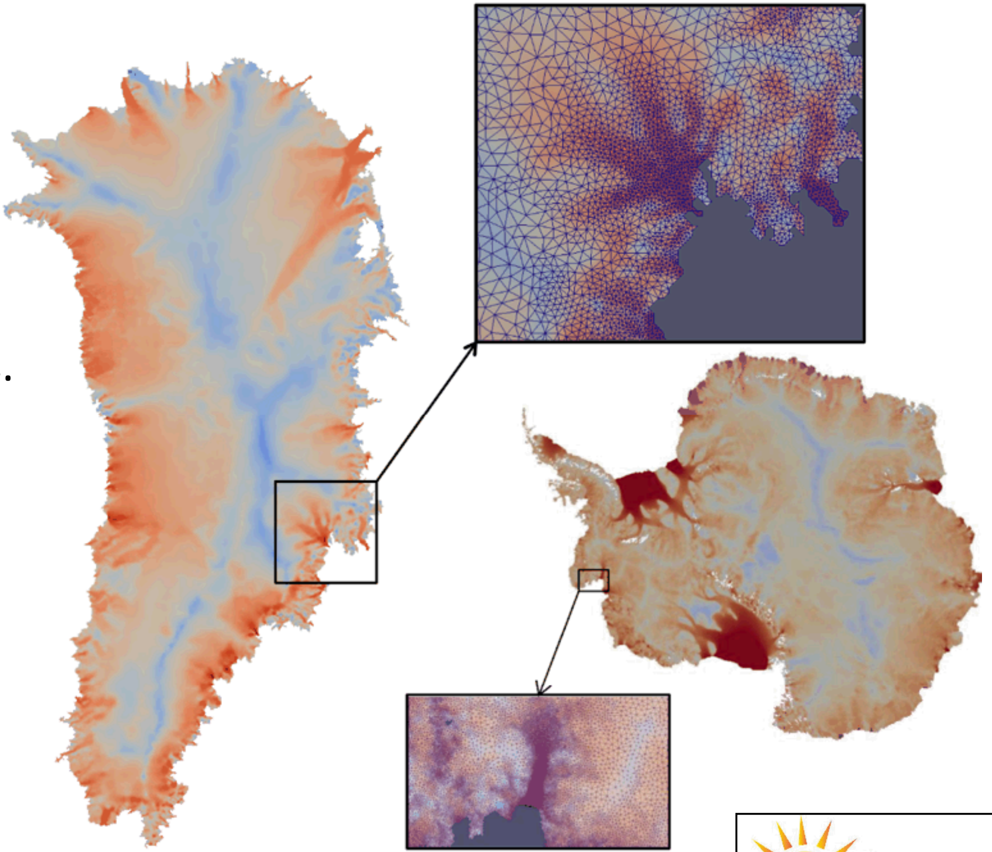
“PISCEES” = Predicting Ice Sheet Climate Evolution at Extreme Scales
5 year SciDAC3 project began in June 2012; proposal for 5 year continuation project submitted to SciDAC4 call.

<u>Sandia’s Role in the PISCEES Project</u>: to **develop** and **support** a robust and scalable land ice solver based on the “First-Order” (FO) Stokes equations → *Albany/FELIX**

<u>Requirements for *Albany/FELIX*</u>:

- ***Unstructured grid*** finite elements.
- ***Scalable, fast*** and ***robust.***
- ***Verified*** and ***validated.***
- ***Portable*** to new architecture machines.
- ***Advanced analysis*** capabilities: deterministic inversion, calibration, uncertainty quantification.

As part of **ACME *DOE Earth System Model***, solver will provide actionable predictions of 21st century sea-level change (including uncertainty bounds).



PISCEES Project for Land-Ice Modeling

"PISCEES" = Predicting Ice Sheet Climate Evolution at Extreme Scales
5 year SciDAC3 project began in June 2012; proposal for 5 year continuation project submitted to SciDAC4 call.

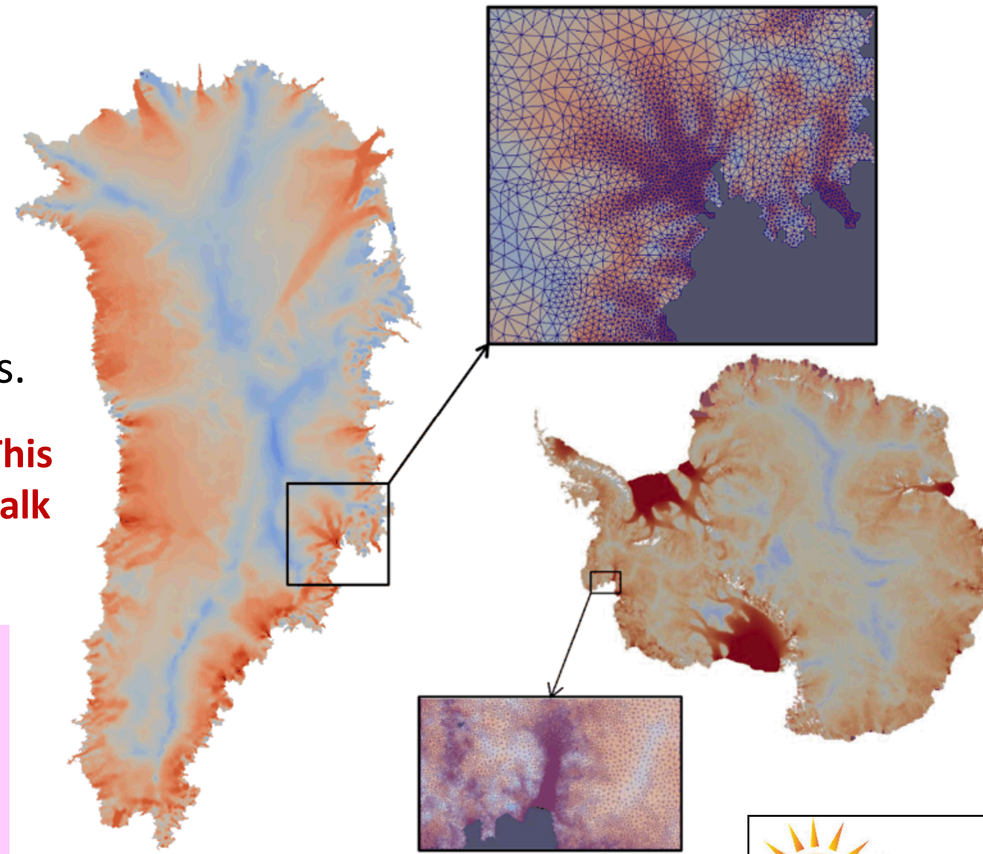
Sandia's Role in the PISCEES Project: to **develop** and **support** a robust and scalable land ice solver based on the "First-Order" (FO) Stokes equations → *Albany/FELIX**

Requirements for *Albany/FELIX*:

- ***Unstructured grid*** finite elements.
- ***Scalable, fast*** and ***robust.***
- ***Verified*** and ***validated.***
- ***Portable*** to new architecture machines.
- ***Advanced analysis*** capabilities: deterministic inversion, calibration, uncertainty quantification.

This talk

As part of **ACME *DOE Earth System Model***, solver will provide actionable predictions of 21st century sea-level change (including **uncertainty bounds**).



*Finite Elements for Land Ice eXperiments

Outline

1. **Background.**
   - PISCEES project for land-ice modeling.
   - **Land-ice model.**

2. UQ problem definition.

3. Inversion/calibration.
   - Deterministic inversion.
   - Bayesian inference.

4. Summary & future work.

The First-Order Stokes Model

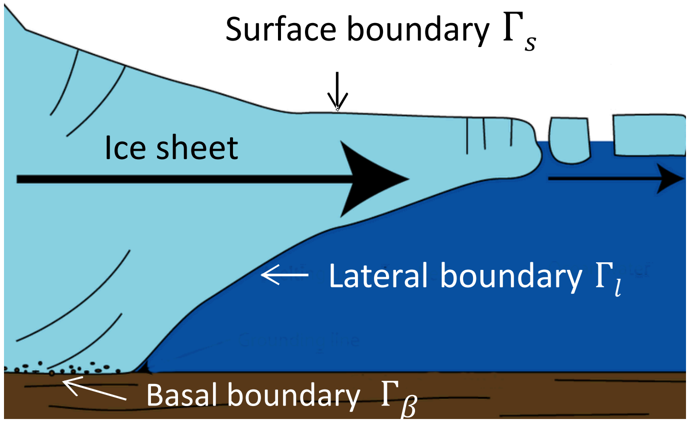
- Ice behaves like a very ***viscous shear-thinning fluid*** (similar to lava flow).
- ***Quasi-static*** model with ***momentum balance*** given by ***"First-Order" Stokes PDEs***: "nice" elliptic approximation* to Stokes' flow equations.

$$\begin{cases} -\nabla \cdot (2\mu \dot{\boldsymbol{\epsilon}}_1) = -\rho g \dfrac{\partial s}{\partial x} \\ -\nabla \cdot (2\mu \dot{\boldsymbol{\epsilon}}_2) = -\rho g \dfrac{\partial s}{\partial y} \end{cases}, \quad \text{in } \Omega$$

$$\dot{\boldsymbol{\epsilon}}_1{}^T = (2\dot{\epsilon}_{11} + \dot{\epsilon}_{22}, \dot{\epsilon}_{12}, \dot{\epsilon}_{13})$$
$$\dot{\boldsymbol{\epsilon}}_2{}^T = (2\dot{\epsilon}_{12}, \dot{\epsilon}_{11} + 2\dot{\epsilon}_{22}, \dot{\epsilon}_{23})$$
$$\dot{\epsilon}_{\text{ij}} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$$

- Viscosity μ is nonlinear function given by "***Glen's law***":

$$\mu = \frac{1}{2} A(T)^{-\frac{1}{n}} \left(\frac{1}{2}\sum_{ij} \dot{\boldsymbol{\epsilon}}_{ij}{}^2\right)^{\left(\frac{1}{2n}-\frac{1}{2}\right)} \qquad (n=3)$$



- Relevant boundary conditions:
  - ***Stress-free BC:*** $2\mu \dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} = 0$, on Γ_s
  - ***Floating ice BC:*** $2\mu \dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} = \begin{cases} \rho g z \boldsymbol{n}, & \text{if } z > 0 \\ 0, & \text{if } z \leq 0 \end{cases}$, on Γ_l
  - ***Basal sliding BC:*** $2\mu \dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} + \beta(x,y) u_i = 0$, on Γ_β

***Assumption:** aspect ratio δ is small and normals to upper/lower surfaces are almost vertical.

The First-Order Stokes Model

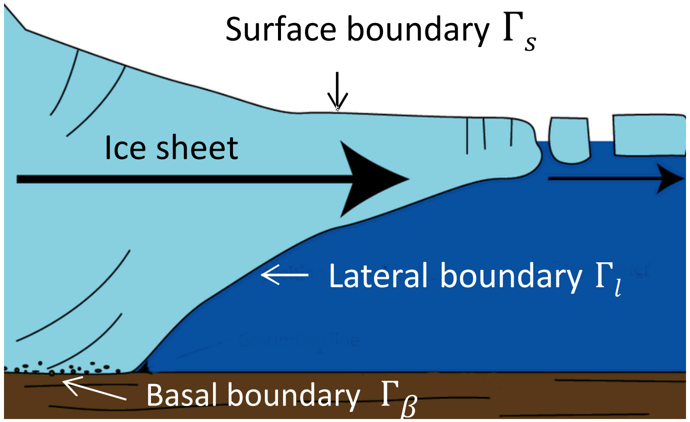
- Ice behaves like a very ***viscous shear-thinning fluid*** (similar to lava flow).
- ***Quasi-static*** model with ***momentum balance*** given by ***"First-Order" Stokes PDEs***: "nice" elliptic approximation* to Stokes' flow equations.

$$\begin{cases} -\nabla \cdot (2\mu \dot{\boldsymbol{\epsilon}}_1) = -\rho g \dfrac{\partial s}{\partial x} \\ -\nabla \cdot (2\mu \dot{\boldsymbol{\epsilon}}_2) = -\rho g \dfrac{\partial s}{\partial y} \end{cases}, \quad \text{in } \Omega$$

$$\dot{\boldsymbol{\epsilon}}_1^{\,T} = (2\dot{\epsilon}_{11} + \dot{\epsilon}_{22}, \dot{\epsilon}_{12}, \dot{\epsilon}_{13})$$
$$\dot{\boldsymbol{\epsilon}}_2^{\,T} = (2\dot{\epsilon}_{12}, \dot{\epsilon}_{11} + 2\dot{\epsilon}_{22}, \dot{\epsilon}_{23})$$
$$\dot{\epsilon}_{\mathrm{ij}} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$$

- Viscosity μ is nonlinear function given by "***Glen's law***":

$$\mu = \frac{1}{2} A(T)^{-\frac{1}{n}} \left(\frac{1}{2}\sum_{ij} \dot{\boldsymbol{\epsilon}}_{ij}^{\,2}\right)^{\left(\frac{1}{2n}-\frac{1}{2}\right)} \qquad (n = 3)$$



- Relevant boundary conditions:
  - ***Stress-free BC:*** $2\mu\dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} = 0$, on Γ_s
  - ***Floating ice BC:*** $2\mu\dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} = \begin{cases} \rho g z \boldsymbol{n}, & \text{if } z > 0 \\ 0, & \text{if } z \le 0 \end{cases}$, on Γ_l
  - ***Basal sliding BC:*** $2\mu\dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} + \beta(x,y) u_i = 0$, on Γ_β

Assumption: aspect ratio δ is small and normals to upper/lower surfaces are almost vertical.

The First-Order Stokes Model

- Ice behaves like a very ***viscous shear-thinning fluid*** (similar to lava flow).
- ***Quasi-static*** model with ***momentum balance*** given by ***"First-Order" Stokes PDEs***: "nice" elliptic approximation* to Stokes' flow equations.

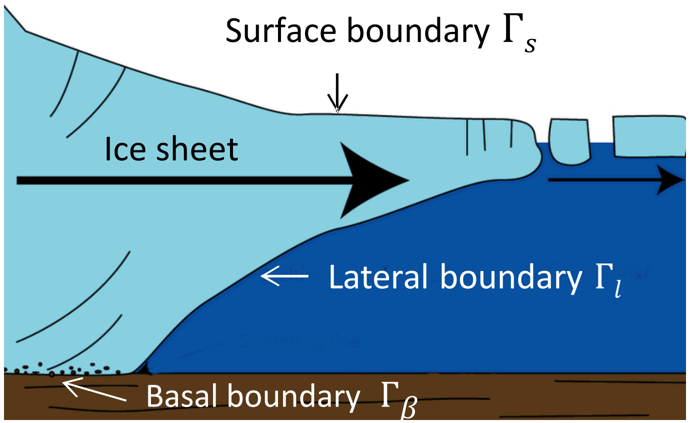
$$\begin{cases} -\nabla \cdot (2\mu \dot{\boldsymbol{\epsilon}}_1) = -\rho g \dfrac{\partial s}{\partial x} \\ -\nabla \cdot (2\mu \dot{\boldsymbol{\epsilon}}_2) = -\rho g \dfrac{\partial s}{\partial y} \end{cases}, \quad \text{in } \Omega$$

$$\dot{\boldsymbol{\epsilon}}_1{}^T = (2\dot{\epsilon}_{11} + \dot{\epsilon}_{22}, \dot{\epsilon}_{12}, \dot{\epsilon}_{13})$$
$$\dot{\boldsymbol{\epsilon}}_2{}^T = (2\dot{\epsilon}_{12}, \dot{\epsilon}_{11} + 2\dot{\epsilon}_{22}, \dot{\epsilon}_{23})$$
$$\dot{\epsilon}_{\text{ij}} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$$

- Viscosity μ is nonlinear function given by "***Glen's law***":

$$\mu = \frac{1}{2} A(T)^{-\frac{1}{n}} \left(\frac{1}{2} \sum_{ij} \dot{\boldsymbol{\epsilon}}_{ij}{}^2 \right)^{\left(\frac{1}{2n} - \frac{1}{2}\right)} \qquad (n = 3)$$

- Relevant boundary conditions:
  - ***Stress-free BC:*** $2\mu \dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} = 0$, on Γ_s
  - ***Floating ice BC:*** $2\mu \dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} = \begin{cases} \rho g z \boldsymbol{n}, & \text{if } z > 0 \\ 0, & \text{if } z \leq 0 \end{cases}$, on Γ_l
  - ***Basal sliding BC:*** $2\mu \dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} + \beta(x, y) u_i = 0$, on Γ_β



$\beta(x, y) =$ basal sliding coefficient

*Assumption: aspect ratio δ is small and normals to upper/lower surfaces are almost vertical.

Thickness & Temperature Equations

Sandia National Laboratories

- Model for ***evolution of the boundaries*** (thickness evolution equation):

$$\frac{\partial H}{\partial t} = -\nabla \cdot (\bar{\boldsymbol{u}} H) + \dot{b}$$

where $\bar{\boldsymbol{u}}$ = vertically averaged velocity, $\dot{b}$ = surface mass balance (conservation of mass).

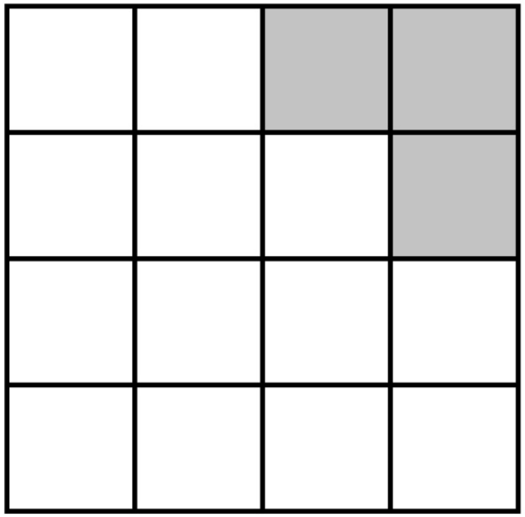
- ***Temperature equation*** (advection-diffusion):

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) - \rho c \boldsymbol{u} \cdot \nabla T + 2 \dot{\boldsymbol{\epsilon}} \boldsymbol{\sigma}$$

(energy balance).

- ***Flow factor*** A in Glen's law depends on temperature T: $A = A(T)$.

- Ice sheet ***grows/retreats*** depending on thickness H.



Ice-covered ("active") cells shaded in white ($H > H_{min}$)

Outline

1. Background.
    - PISCEES project for land-ice modeling.
    - Land-ice model.

2. **UQ problem definition.**

3. Inversion/calibration.
    - Deterministic inversion.
    - Bayesian inference.

4. Summary & future work.

Uncertainty Quantification Problem Definition

QoI in Ice Sheet Modeling: total ice mass loss/gain during 21st century → ***sea level change prediction.***

Uncertainty Quantification Problem Definition

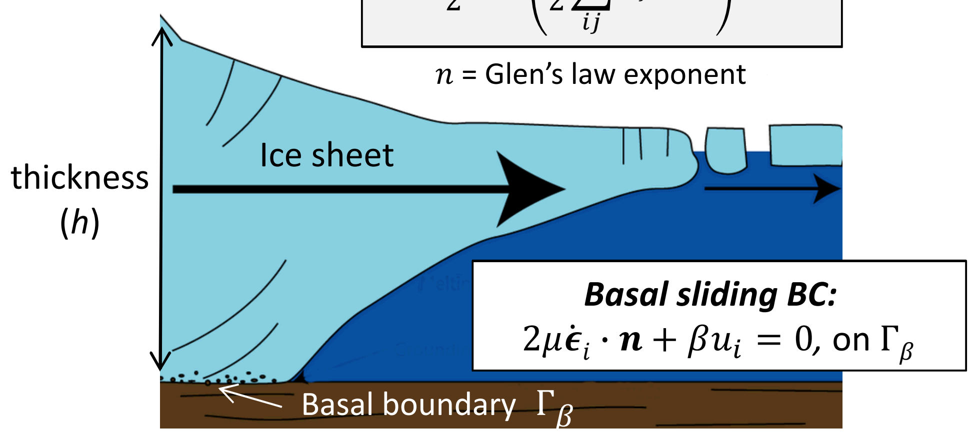
QoI in Ice Sheet Modeling: total ice mass loss/gain during 21st century → ***sea level change prediction.***

Sources of uncertainty affecting this QoI include:

- Climate forcings (e.g., surface mass balance).
- Basal friction (β).
- Ice sheet thickness (h).
- Geothermal heat flux.
- Model parameters (e.g., Glen's flow law exponent).

$$\mu = \frac{1}{2} A^{-\frac{1}{n}} \left(\frac{1}{2} \sum_{ij} \dot{\boldsymbol{\epsilon}}_{ij}{}^{2} + \gamma \right)^{\left(\frac{1}{2n} - \frac{1}{2}\right)}$$

n = Glen's law exponent



Uncertainty Quantification Problem Definition

QoI in Ice Sheet Modeling: total ice mass loss/gain during 21st century → ***sea level change prediction.***

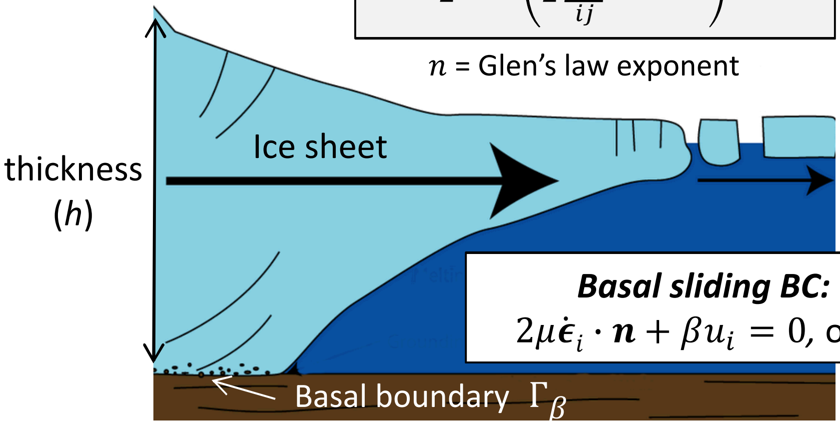
As a first step, we focus on effect of uncertainty in β only.

Sources of uncertainty affecting this QoI include:

- Climate forcings (e.g., surface mass balance).
- **Basal friction (β).**
- Ice sheet thickness (h).
- Geothermal heat flux.
- Model parameters (e.g., Glen's flow law exponent).

$$\mu = \frac{1}{2} A^{-\frac{1}{n}} \left(\frac{1}{2} \sum_{ij} \dot{\epsilon}_{ij}^{\,2} + \gamma \right)^{\left(\frac{1}{2n} - \frac{1}{2}\right)}$$

n = Glen's law exponent



Basal sliding BC:

$$2\mu\dot{\epsilon}_i \cdot \boldsymbol{n} + \beta u_i = 0, \text{ on } \Gamma_\beta$$

Uncertainty Quantification Problem Definition

QoI in Ice Sheet Modeling: total ice mass loss/gain during 21st century → ***sea level change prediction.***

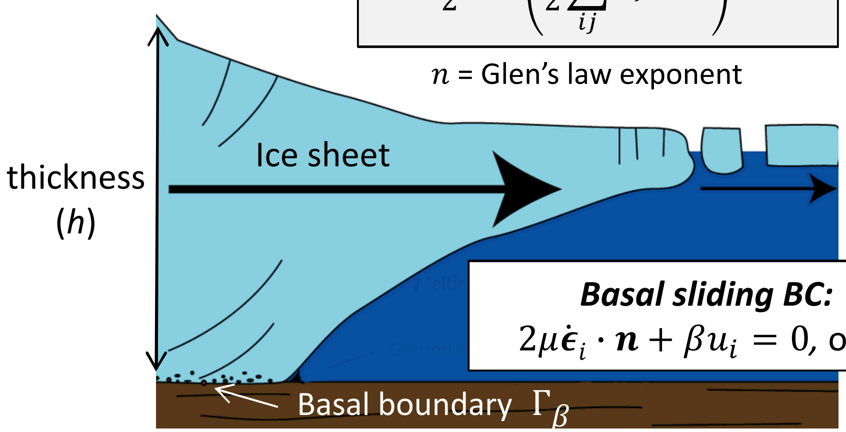
As a first step, we focus on effect of uncertainty in β only.

Sources of uncertainty affecting this QoI include:

- Climate forcings (e.g., surface mass balance).
- **Basal friction (β).**
- Ice sheet thickness (h).
- Geothermal heat flux.
- Model parameters (e.g., Glen's flow law exponent).

$$\mu = \frac{1}{2} A^{-\frac{1}{n}} \left(\frac{1}{2} \sum_{ij} \dot{\epsilon}_{ij}^{\,2} + \gamma \right)^{\left(\frac{1}{2n} - \frac{1}{2}\right)}$$

n = Glen's law exponent



Basal sliding BC:
$2\mu \dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} + \beta u_i = 0$, on Γ_β

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

Uncertainty Quantification Problem Definition

QoI in Ice Sheet Modeling: total ice mass loss/gain during 21st century → ***sea level change prediction.***

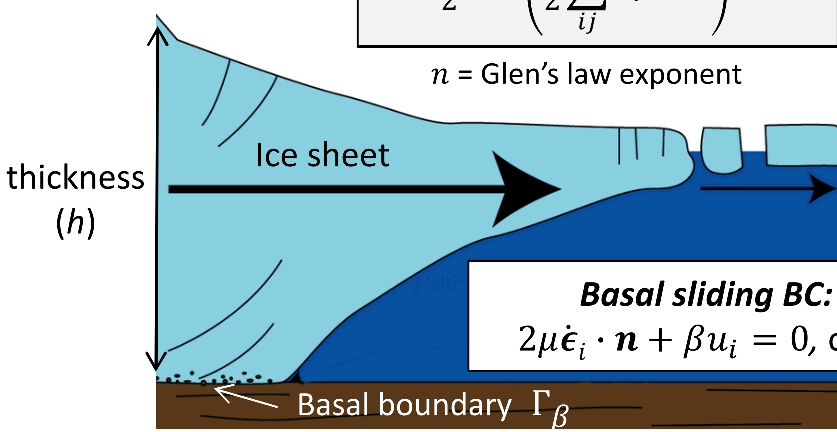
As a first step, we focus on effect of uncertainty in β only.

<u>Sources of uncertainty affecting this QoI include:</u>

- Climate forcings (e.g., surface mass balance).
- **Basal friction (β).**
- Ice sheet thickness (h).
- Geothermal heat flux.
- Model parameters (e.g., Glen's flow law exponent).

$$\mu = \frac{1}{2} A^{-\frac{1}{n}} \left(\frac{1}{2} \sum_{ij} \dot{\epsilon}_{ij}^{\,2} + \gamma \right)^{\left(\frac{1}{2n} - \frac{1}{2}\right)}$$

n = Glen's law exponent



Basal sliding BC:
$2\mu\dot{\epsilon}_i \cdot \boldsymbol{n} + \beta u_i = 0$, on Γ_β

UQ Workflow

Deterministic inversion — ***Stage 1:*** Estimate ice sheet initial condition (MAP point).

Bayesian calibration — ***Stage 2:*** Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Forward propagation — ***Stage 3:*** Propagate uncertain initial condition through ice-sheet evolution model

Uncertainty Quantification Problem Definition

QoI in Ice Sheet Modeling: total ice mass loss/gain during 21st century → ***sea level change prediction.***

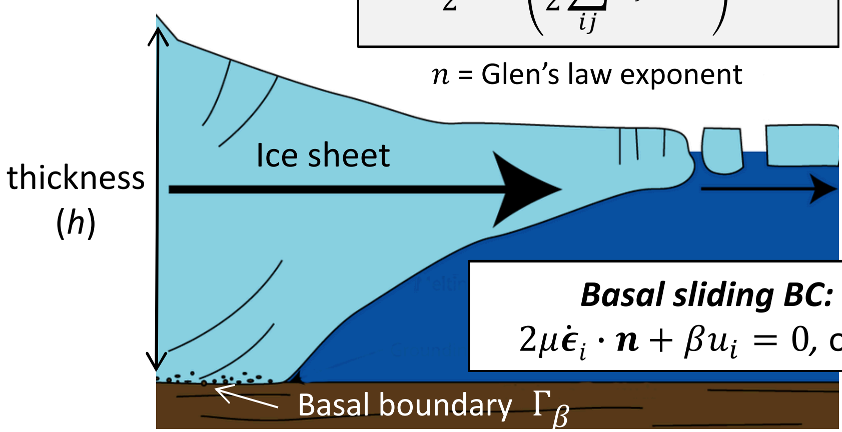
As a first step, we focus on effect of uncertainty in β only.

Sources of uncertainty affecting this QoI include:

- Climate forcings (e.g., surface mass balance).
- **Basal friction ($\boldsymbol{\beta}$).**
- Ice sheet thickness (h).
- Geothermal heat flux.
- Model parameters (e.g., Glen's flow law exponent).

$$\mu = \frac{1}{2} A^{-\frac{1}{n}} \left(\frac{1}{2} \sum_{ij} \dot{\boldsymbol{\epsilon}}_{ij}{}^{2} + \gamma \right)^{\left(\frac{1}{2n} - \frac{1}{2}\right)}$$

n = Glen's law exponent



Basal sliding BC:

$2\mu\dot{\boldsymbol{\epsilon}}_i \cdot \boldsymbol{n} + \beta u_i = 0$, on Γ_β

UQ Workflow

Deterministic inversion — ***Stage 1:*** Estimate ice sheet initial condition (MAP point).

Bayesian calibration — ***Stage 2:*** Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Forward propagation — ***Stage 3:*** Propagate uncertain initial condition through ice-sheet evolution model

Outline

1. Background.
   - PISCEES project for land-ice modeling.
   - Land-ice model.
2. UQ problem definition.
3. **Inversion/calibration.**
   - **Deterministic inversion.**
   - Bayesian inference.
4. Summary & future work.

Estimation of Ice Sheet Initial Condition

Sandia National Laboratories

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

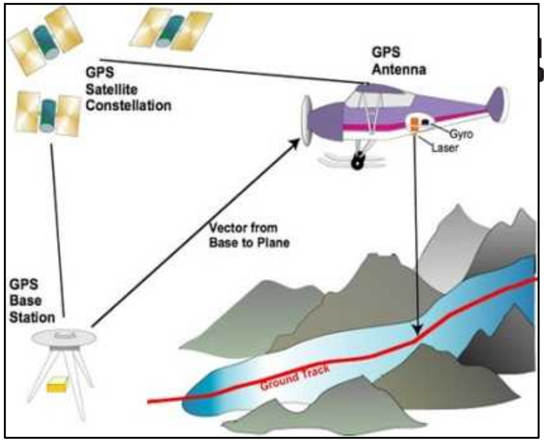
Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

<u>Goal:</u> find ice sheet initial state that:

- matches observations (e.g. surface velocity, temperature, etc.).
- matches present-day geometry (elevation, thickness).
- is in "equilibrium" with climate forcings (SMB).

Available Data & Assumptions



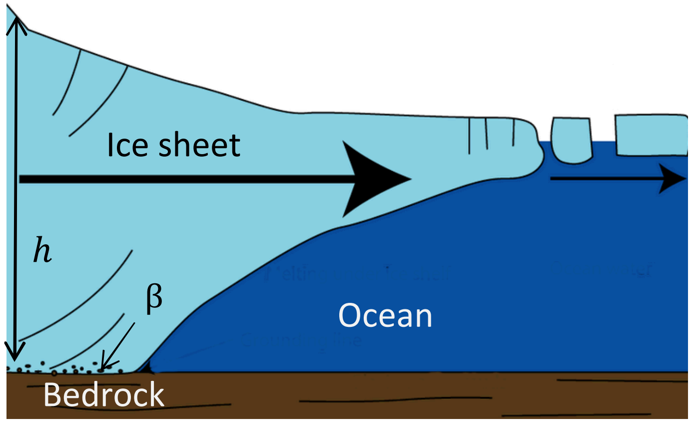
Available data/measurements:

- ice extent and surface topography.
- surface velocity.
- surface mass balance (SMB).
- ice thickness h (sparse measurements).

Sources of data: satellite infrarometry, radar, altimetry, etc.

Fields to be estimated:

- ice thickness h (allowed to vary but weighted by observational uncertainties).
- basal friction β (spatially variable proxy for all basal processes).



Modeling Assumptions:

- ice flow described by nonlinear first-order Stokes equations.
- ice close to mechanical equilibrium.

Deterministic Inversion

Sandia National Laboratories

First-Order Stokes PDE-Constrained optimization problem for initial condition*:

minimize $_{\beta,h}$ $m(\beta,h)$
s.t. FO Stokes PDEs

$\boldsymbol{U}$: computed depth averaged velocity
h: ice thickness
β: basal sliding friction coefficient
τ_s: surface mass balance (SMB)
$\mathcal{R}(\beta,h)$: regularization term

$$m(\beta,h) = \int_\Gamma \frac{1}{\sigma_u^2}|\boldsymbol{u}-\boldsymbol{u}^{obs}|^2 ds$$ surface velocity mismatch

$$+\int_\Gamma \frac{1}{\sigma_\tau^2}|div(\boldsymbol{U}h)-\tau_s|^2 ds$$ SMB mismatch

$$+\int_\Gamma \frac{1}{\sigma_h^2}|h-h^{obs}|^2 ds$$ thickness mismatch

$$+\mathcal{R}(\beta,h)$$ regularization terms

Deterministic Inversion

Sandia National Laboratories

First-Order Stokes PDE-Constrained optimization problem for initial condition*:

minimize $_{\beta,h}$ $m(\beta,h)$
s.t. FO Stokes PDEs

$\boldsymbol{U}$: computed depth averaged velocity
h: ice thickness
β: basal sliding friction coefficient
τ_s: surface mass balance (SMB)
$\mathcal{R}(\beta,h)$: regularization term

$$m(\beta,h) = \int_\Gamma \frac{1}{\sigma_u^2}|\boldsymbol{u}-\boldsymbol{u}^{obs}|^2 ds$$ surface velocity mismatch

$$+\int_\Gamma \frac{1}{\sigma_\tau^2}|div(\boldsymbol{U}h)-\tau_s|^2 ds$$ SMB mismatch

$$+\int_\Gamma \frac{1}{\sigma_h^2}|h-h^{obs}|^2 ds$$ thickness mismatch

$$+\mathcal{R}(\beta,h)$$ regularization terms

Solving FO Stokes PDE-constrained optimization problem for initial condition significantly **reduces non-physical model transients**!

* Perego, Stadler, Price, *JGR*, 2014.

Deterministic Inversion Algorithm & Software

First-Order Stokes PDE-Constrained optimization problem for initial condition*:

minimize $_{\beta,h}$ $m(\beta, h)$
s.t. FO Stokes PDEs

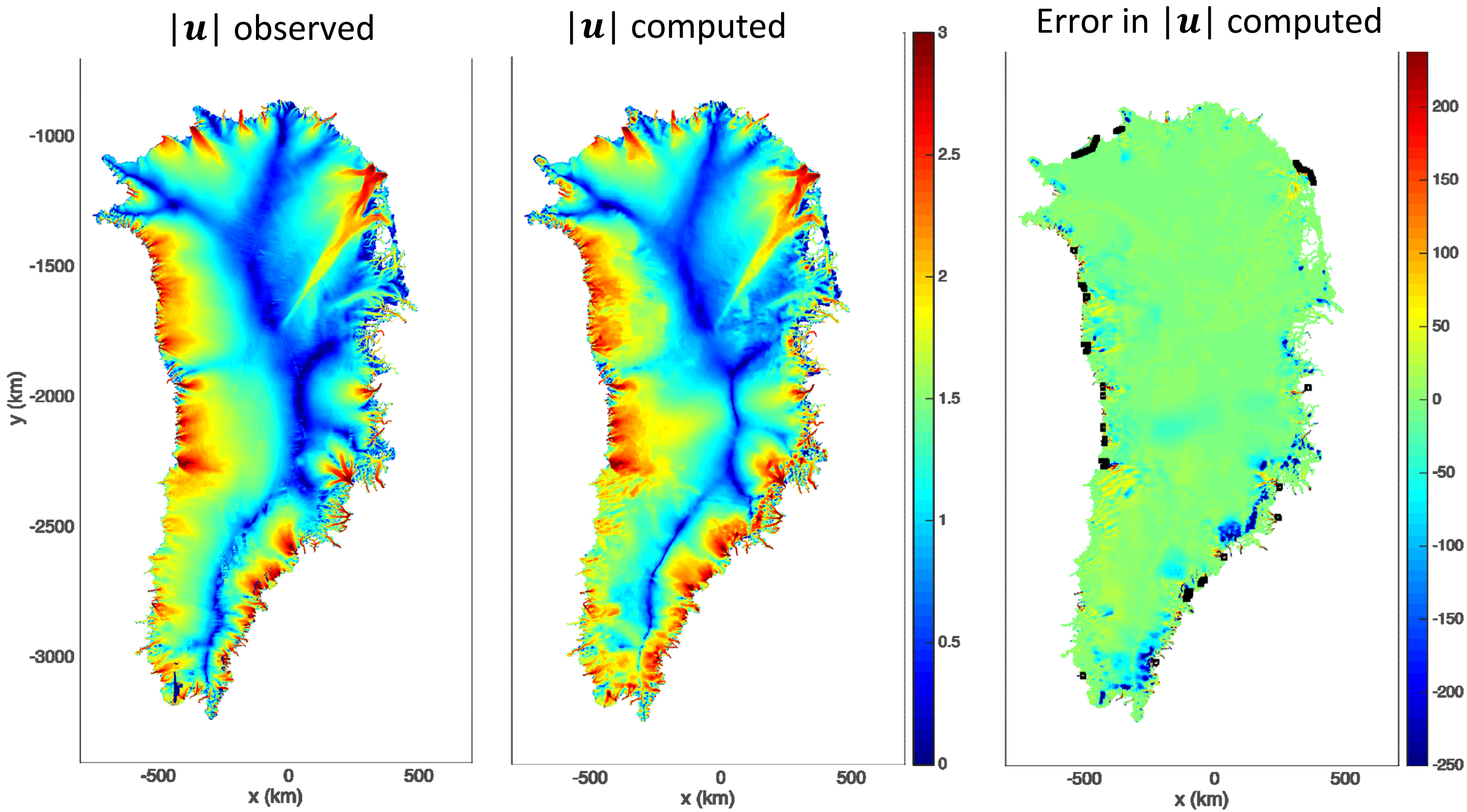
Solved via embedded ***adjoint-based PDE-constrained optimization*** algorithm in Albany/FELIX.

Algorithm	Software
Finite Element Method discretization	Albany
Quasi-Newton optimization (L-BFGS)	ROL
Nonlinear solver (Newton)	NOX
Krylov linear solvers	AztecOO+Ifpack/ML

- **Some details:**
  - ***Regularization:*** Tikhonov.
  - Total derivatives of objective functional $m(\beta, h)$ computed using ***adjoints*** and ***automatic differentiation*** (Sacado package of Trilinos).
  - ***Gradient-based optimization***: limited memory BFGS initialized with Hessian of regularization terms (ROL) with backtrack linesearch.

* Perego, Stadler, Price, *JGR*, 2014.

Deterministic Inversion: 1km Greenland Initial Condition



Outline

1. Background.
   - PISCEES project for land-ice modeling.
   - Land-ice model.
2. UQ problem definition.
3. **Inversion/calibration.**
   - Deterministic inversion.
   - **Bayesian inference.**
4. Summary & future work.

Bayesian Inference

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

Goal: solve inverse problem for ice sheet initial state but in ***Bayesian framework***

Bayesian Inference

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

Goal: solve inverse problem for ice sheet initial state but in ***Bayesian framework***

- ***Naïve parameterization***: represent each degree of freedom on mesh be an uncertain variable

$$\beta(\boldsymbol{x}) = (z_1, z_2, \ldots, z_{n_{\mathrm{dof}}})$$

Bayesian Inference

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

Goal: solve inverse problem for ice sheet initial state but in ***Bayesian framework***

- ***Naïve parameterization***: represent each degree of freedom on mesh be an uncertain variable

$$\beta(\boldsymbol{x}) = (z_1, z_2, \dots, z_{n_{\mathrm{dof}}})$$

Intractable due to **curse of dimensionality**: $n_{\mathrm{dof}} = O(100K)$!

Bayesian Inference

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

Goal: solve inverse problem for ice sheet initial state but in ***Bayesian framework***

- ***Naïve parameterization***: represent each degree of freedom on mesh be an uncertain variable

$$\beta(\boldsymbol{x}) = (z_1, z_2, \dots, z_{n_{\mathrm{dof}}})$$

Intractable due to **curse of dimensionality**: $n_{\mathrm{dof}} = O(100K)$!

- ***To circumvent this difficulty:*** assume $\beta(\boldsymbol{x})$ can be represented in ***reduced basis*** (e.g., KLE modes, Hessian eigenvectors*) centered around mean $\bar{\beta}(\boldsymbol{x})$:

$$\log(\beta(\boldsymbol{x})) = \log(\bar{\beta}) + \sum_{i=1}^{d} \sqrt{\lambda_i}\, \phi_i(\boldsymbol{x}) z_i$$

* Isaac, Petra, Stadler, Ghattas, *JCP,* 2015.

Bayesian Inference

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

<u>Goal:</u> solve inverse problem for ice sheet initial state but in ***Bayesian framework***

- ***Naïve parameterization***: represent each degree of freedom on mesh be an uncertain variable

$$\beta(\boldsymbol{x}) = (z_1, z_2, \dots, z_{n_{\text{dof}}})$$

Intractable due to **curse of dimensionality**: $n_{\text{dof}} = O(100K)$!

- ***To circumvent this difficulty:*** assume $\beta(\boldsymbol{x})$ can be represented in ***reduced basis*** (e.g., KLE modes, Hessian eigenvectors*) centered around mean $\bar{\beta}(\boldsymbol{x})$:

$$\log(\beta(\boldsymbol{x})) = \log(\bar{\beta}) + \sum_{i=1}^{d} \sqrt{\lambda_i}\, \phi_i(\boldsymbol{x}) z_i$$

- Mean field $\bar{\beta}(\boldsymbol{x})$ = initial condition.

* Isaac, Petra, Stadler, Ghattas, *JCP,* 2015.

Bayesian Inference

Sandia National Laboratories

UQ Workflow

Stage 1:
Estimate ice sheet initial condition (MAP point).

Stage 2:
Update prior uncertainty in ice sheet initial condition using observational data and steady state model

Stage 3:
Propagate uncertain initial condition through ice-sheet evolution model

Deterministic inversion is consistent with Bayesian analog: it is used to find the MAP point of posterior.

<u>Goal:</u> solve inverse problem for ice sheet initial state but in ***Bayesian framework***

- ***Naïve parameterization***: represent each degree of freedom on mesh be an uncertain variable

$$\beta(\boldsymbol{x}) = (z_1, z_2, \dots, z_{n_{\text{dof}}})$$

Intractable due to **curse of dimensionality**: $n_{\text{dof}} = O(100K)$!

- ***To circumvent this difficulty:*** assume $\beta(\boldsymbol{x})$ can be represented in ***reduced basis*** (e.g., KLE modes, Hessian eigenvectors*) centered around mean $\bar{\beta}(\boldsymbol{x})$:

$$\log(\beta(\boldsymbol{x})) = \log(\bar{\beta}) + \sum_{i=1}^{d} \sqrt{\lambda_i}\, \phi_i(\boldsymbol{x}) z_i$$

- Mean field $\bar{\beta}(\boldsymbol{x})$ = initial condition.

* Isaac, Petra, Stadler, Ghattas, *JCP*, 2015.

Bayesian Inference Assumptions

- Additive ***Gaussian noise*** model: $\boldsymbol{y}^{\text{obs}} = \boldsymbol{f}(\boldsymbol{z}) + \epsilon, \quad \epsilon \sim N(\boldsymbol{0}, \boldsymbol{\Gamma}_{\text{obs}})$

⇒ Mismatch functional to be minimized:

$$m(\boldsymbol{z}) = \frac{1}{2}\left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)^T \boldsymbol{\Gamma}_{\text{obs}}^{-1} \left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)$$

- ***Gaussian prior*** with exponential covariance.

Notation*:

$\boldsymbol{y}^{\text{obs}}$= observations

$\boldsymbol{z}$ = random params

$\boldsymbol{f}(\boldsymbol{z})$ = deterministic map from params to observables.

Bayesian Inference Assumptions

- Additive ***Gaussian noise*** model: $\boldsymbol{y}^{\text{obs}} = \boldsymbol{f}(\boldsymbol{z}) + \epsilon, \quad \epsilon \sim N(\boldsymbol{0}, \boldsymbol{\Gamma}_{\text{obs}})$

⇒ Mismatch functional to be minimized:

$$m(\boldsymbol{z}) = \frac{1}{2}\left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)^T \boldsymbol{\Gamma}_{\text{obs}}^{-1}\left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)$$

- ***Gaussian prior*** with exponential covariance.

Bayes' rule

+ linearization of $\boldsymbol{f}(\boldsymbol{z})$ around $\boldsymbol{z}_{MAP}$

Notation*:

$\boldsymbol{y}^{\text{obs}}$= observations

$\boldsymbol{z}$ = random params

$\boldsymbol{f}(\boldsymbol{z})$ = deterministic map from params to observables.

- ***Likelihood*** is: $\hat{\pi}_{\text{lhood}}(\boldsymbol{z}) = e^{-m_{\text{lin}}(\boldsymbol{z})}$
- ***Normal Laplace posterior*** given by: $\pi_{\text{pos}}(\boldsymbol{z}) = C_{\text{evid}}^{-1}\hat{\pi}_{\text{lhood}}(\boldsymbol{z})\pi_{\text{pr}}(\boldsymbol{z})$

  where $C_{evid} = \int \hat{\pi}_{\text{lhood}}(\boldsymbol{z})\pi_{\text{pr}}(\boldsymbol{z})d\boldsymbol{z}$.

Bayesian Inference Assumptions

- Additive ***Gaussian noise*** model: $\boldsymbol{y}^{\text{obs}} = \boldsymbol{f}(\boldsymbol{z}) + \epsilon, \quad \epsilon \sim N(\boldsymbol{0}, \boldsymbol{\Gamma}_{\text{obs}})$

⇒ Mismatch functional to be minimized:

$$m(\boldsymbol{z}) = \frac{1}{2}\left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)^T \boldsymbol{\Gamma}_{\text{obs}}^{-1} \left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)$$

- ***Gaussian prior*** with exponential covariance.

Bayes' rule

+ linearization of $\boldsymbol{f}(\boldsymbol{z})$ around $\boldsymbol{z}_{MAP}$

Covariance of Gaussian posterior related to ***inverse of misfit Hessian*** at MAP point**.

Notation*:

$\boldsymbol{y}^{\text{obs}}$= observations

$\boldsymbol{z}$ = random params

$\boldsymbol{f}(\boldsymbol{z})$ = deterministic map from params to observables.

- ***Likelihood*** is: $\hat{\pi}_{\text{lhood}}(\boldsymbol{z}) = e^{-m_{\text{lin}}(\boldsymbol{z})}$
- ***Normal Laplace posterior*** given by: $\pi_{\text{pos}}(\boldsymbol{z}) = C_{\text{evid}}^{-1}\hat{\pi}_{\text{lhood}}(\boldsymbol{z})\pi_{\text{pr}}(\boldsymbol{z})$

  where $C_{evid} = \int \hat{\pi}_{\text{lhood}}(\boldsymbol{z})\pi_{\text{pr}}(\boldsymbol{z})d\boldsymbol{z}$.

* Constantine, Kent, Bui-Thanh, *SISC*, 2016. **Bui-Thanh, Ghattas, Martin, Stadler, *SISC*, 2013.

Bayesian Inference Assumptions

- Additive ***Gaussian noise*** model: $\boldsymbol{y}^{\text{obs}} = \boldsymbol{f}(\boldsymbol{z}) + \epsilon, \quad \epsilon \sim N(\boldsymbol{0}, \boldsymbol{\Gamma}_{\text{obs}})$

⇒ Mismatch functional to be minimized:

$$m(\boldsymbol{z}) = \frac{1}{2}\left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)^T \boldsymbol{\Gamma}_{\text{obs}}^{-1} \left(\boldsymbol{y}^{\text{obs}} - \boldsymbol{f}(\boldsymbol{z})\right)$$

Evaluation of misfit Hessian is **expensive!** ⇒ further approximation required.

- ***Gaussian prior*** with exponential covariance.

Bayes' rule

+ linearization of $\boldsymbol{f}(\boldsymbol{z})$ around $\boldsymbol{z}_{MAP}$

Covariance of Gaussian posterior related to ***inverse of misfit Hessian*** at MAP point**.

Notation*:

$\boldsymbol{y}^{\text{obs}}$= observations

$\boldsymbol{z}$ = random params

$\boldsymbol{f}(\boldsymbol{z})$ = deterministic map from params to observables.

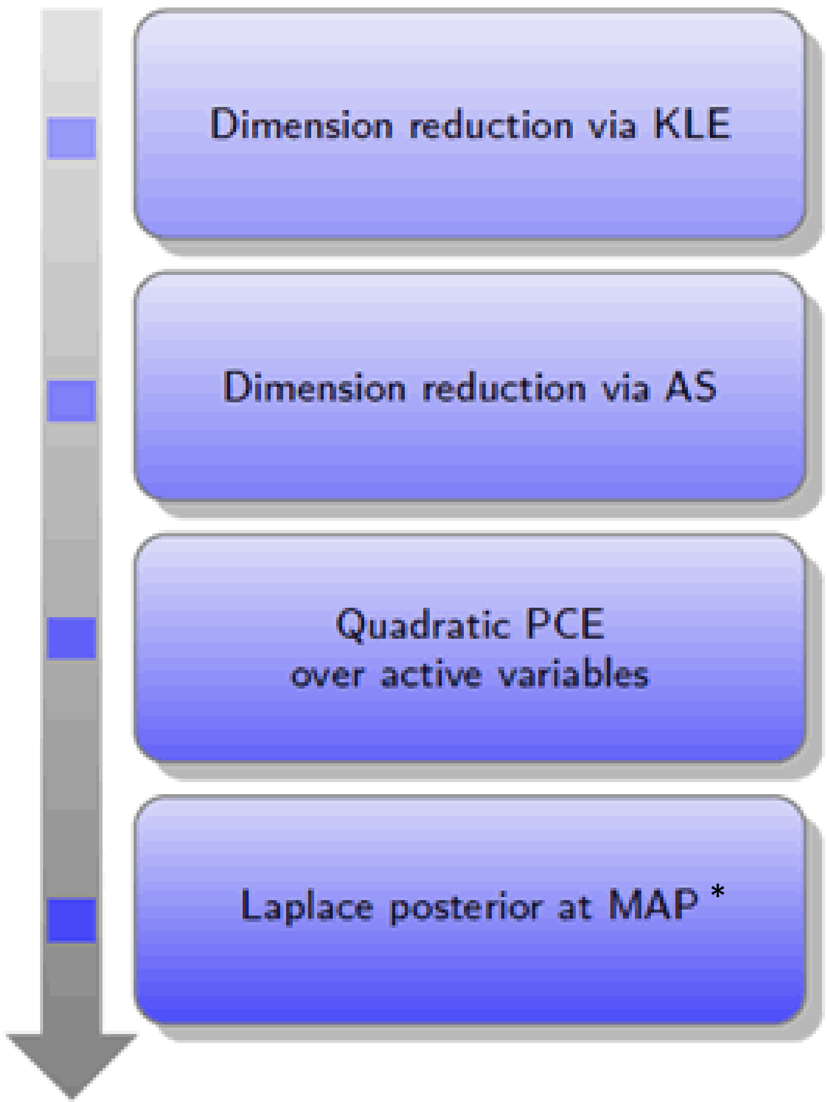
- ***Likelihood*** is: $\hat{\pi}_{\text{lhood}}(\boldsymbol{z}) = e^{-m_{\text{lin}}(\boldsymbol{z})}$
- ***Normal Laplace posterior*** given by: $\pi_{\text{pos}}(\boldsymbol{z}) = C_{\text{evid}}^{-1}\hat{\pi}_{\text{lhood}}(\boldsymbol{z})\pi_{\text{pr}}(\boldsymbol{z})$

  where $C_{evid} = \int \hat{\pi}_{\text{lhood}}(\boldsymbol{z})\pi_{\text{pr}}(\boldsymbol{z})d\boldsymbol{z}$.

* Constantine, Kent, Bui-Thanh, *SISC*, 2016. **Bui-Thanh, Ghattas, Martin, Stadler, *SISC*, 2013.

Bayesian Inference Workflow

KLE = Karhunen-Loeve Expansion
AS = Active Subspace
PCE = Polynomial Chaos Expansion
MAP = Maximum a Posteriori

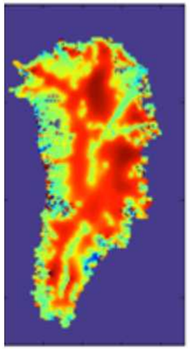


Two-part ***dimension reduction*** procedure to obtain modes $\phi_i(\boldsymbol{x})$

Procedure for computing ***covariance of normal posterior***, $\boldsymbol{\Gamma}_{\text{post}}$

* Bui-Thanh, Ghattas, Martin, Stadler, *SISC*, 2013.

Karhunen-Loeve Expansion (KLE)

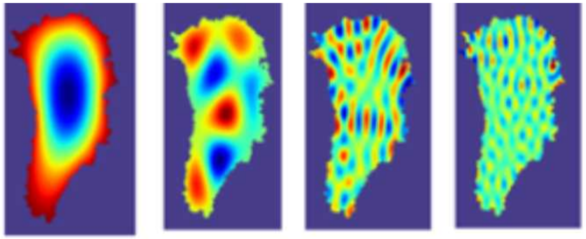


Best fit $\bar{\beta}$

$O(100K)$ dimensional inversion problem can be reduced to smaller dimensional problem using ***Karhunen-Loeve Expansion (KLE)***

$$\log(\beta(\boldsymbol{x})) = \log(\bar{\beta}) + \sum_{i=1}^{d} \sqrt{\lambda_i}\, \phi_i(\boldsymbol{x}) z_i$$

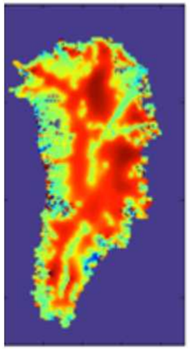
KLE modes ϕ_i



- KLE modes $\phi_i(\boldsymbol{x})$ are eigenvectors of assumed ***exponential covariance kernel***:

$$C(r_1, r_2) = \exp\left(-\frac{(r_1 - r_2)^2}{L^2}\right)$$

Karhunen-Loeve Expansion (KLE)

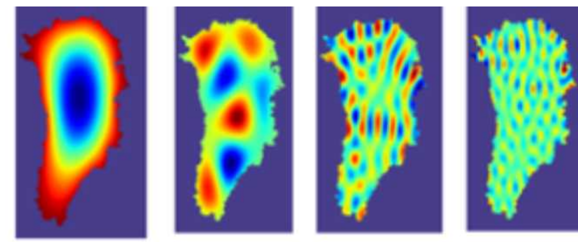


Best fit $\bar{\beta}$

$O(100K)$ dimensional inversion problem can be reduced to smaller dimensional problem using ***Karhunen-Loeve Expansion (KLE)***

$$\log(\beta(\boldsymbol{x})) = \log(\bar{\beta}) + \sum_{i=1}^{d} \sqrt{\lambda_i}\, \phi_i(\boldsymbol{x}) z_i$$

KLE modes ϕ_i

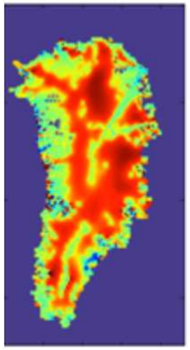


- KLE modes $\phi_i(\boldsymbol{x})$ are eigenvectors of assumed ***exponential covariance kernel***:

$$C(r_1, r_2) = \exp\left(-\frac{(r_1 - r_2)^2}{L^2}\right)$$

- ***Parameters to be selected***: correlation length L, basis size d.

Karhunen-Loeve Expansion (KLE)

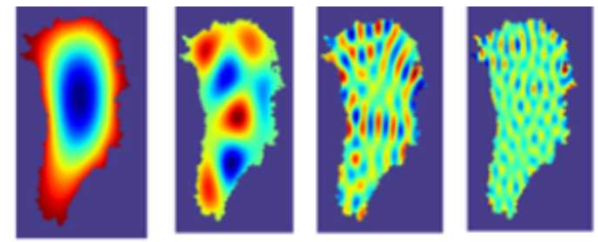


Best fit $\bar{\beta}$

$O(100K)$ dimensional inversion problem can be reduced to smaller dimensional problem using ***Karhunen-Loeve Expansion (KLE)***

$$\log(\beta(\boldsymbol{x})) = \log(\bar{\beta}) + \sum_{i=1}^{d} \sqrt{\lambda_i}\, \phi_i(\boldsymbol{x}) z_i$$

KLE modes ϕ_i

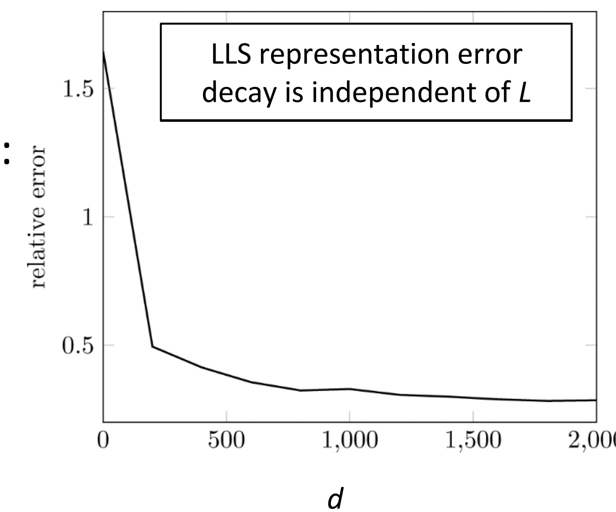


- KLE modes $\phi_i(\boldsymbol{x})$ are eigenvectors of assumed ***exponential covariance kernel***:

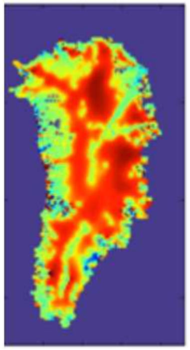
$$C(r_1, r_2) = \exp\left(-\frac{(r_1 - r_2)^2}{L^2}\right)$$

- ***Parameters to be selected***: correlation length L, basis size d.
- L and d can be estimated “rigorously” by solving ***LLS problem***:

$$min_{L,d} \left\| \exp\left(\bar{\beta}^{opt}(\min m(\beta)) - \bar{\beta}^{opt}(\min m(\beta, h)) - \sum_{k=1}^{d} \sqrt{\lambda_k} \phi_k z_k \right) \right\|$$



Karhunen-Loeve Expansion (KLE)

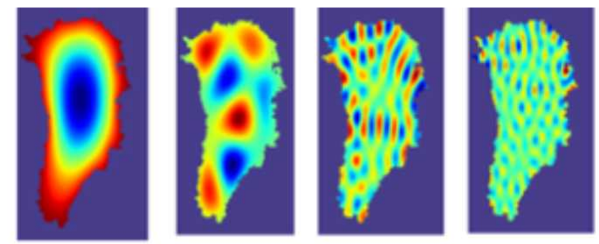


Best fit $\bar{\beta}$

$O(100K)$ dimensional inversion problem can be reduced to smaller dimensional problem using ***Karhunen-Loeve Expansion (KLE)***

$$\log(\beta(\boldsymbol{x})) = \log(\bar{\beta}) + \sum_{i=1}^{d} \sqrt{\lambda_i}\, \phi_i(\boldsymbol{x}) z_i$$

KLE modes ϕ_i



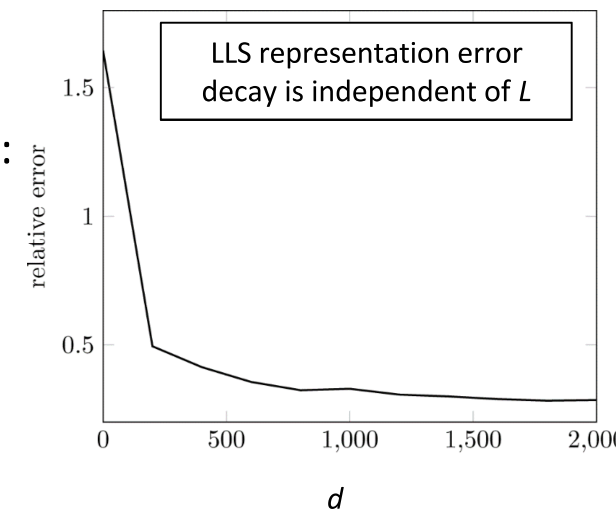
- KLE modes $\phi_i(\boldsymbol{x})$ are eigenvectors of assumed ***exponential covariance kernel***:

$$C(r_1, r_2) = \exp\left(-\frac{(r_1 - r_2)^2}{L^2}\right)$$

- ***Parameters to be selected***: correlation length L, basis size d.
- L and d can be estimated "rigorously" by solving ***LLS problem***:

$$min_{L,d} \left\| \exp\left(\bar{\beta}^{opt}(\min m(\beta)) - \bar{\beta}^{opt}(\min m(\beta, h)) - \sum_{k=1}^{d} \sqrt{\lambda_k} \phi_k z_k \right) \right\|$$

$\Rightarrow$ d should be $O(1000)$



TI1

Active Subspaces (AS)

- KLE eigenvalue analysis suggests $d = O(1000)$ – **still large for MCMC**!

Slide 41

TI1 Tezaur, Irina, 6/1/2017

TI1

Active Subspaces (AS)

- KLE eigenvalue analysis suggests $d = O(1000)$ – **still large for MCMC**!

Idea: combine KLE with Active Subspace (AS) information for further (and better) ***data-informed*** dimension reduction.

Slide 42

TI1 Tezaur, Irina, 6/1/2017

TI1

Active Subspaces (AS)

- KLE eigenvalue analysis suggests $d = O(1000)$ – **still large for MCMC**!

<u>Idea:</u> combine KLE with Active Subspace (AS) information for further (and better) ***data-informed*** dimension reduction.

- ***Active Subspace (AS)*** = directions along which objective function has strongest variability.

Slide 43

TI1 Tezaur, Irina, 6/1/2017

TI1

Active Subspaces (AS)

Sandia National Laboratories

- KLE eigenvalue analysis suggests $d = O(1000)$ – **still large for MCMC**!

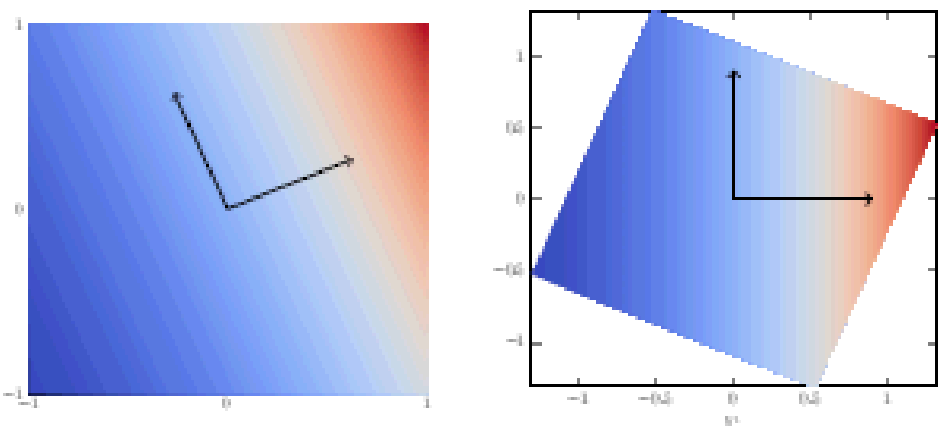
Idea: combine KLE with Active Subspace (AS) information for further (and better) ***data-informed*** dimension reduction.

- ***Active Subspace (AS)*** = directions along which objective function has strongest variability.
- ***Active subspace approach:*** mismatch approximated by related function of fewer variables g:

$$m(\boldsymbol{z}) = \frac{1}{2}(\boldsymbol{d} - \boldsymbol{f}(\boldsymbol{z}))^T \boldsymbol{\Gamma}_{\text{obs}}^{-1}(\boldsymbol{d} - \boldsymbol{f}(\boldsymbol{z})) \approx g(\boldsymbol{W}_1^T \boldsymbol{z})$$

$\boldsymbol{W}_1^T \boldsymbol{z}$ = "active variables"
$\boldsymbol{W}_1^T$ = rotation of coords

Example*: $m(\boldsymbol{z}) = \exp(0.7 z_1 + 0.3 z_2)$



Dimension reduction via AS:

(i) Rotate coords s.t. directions of strongest variation are aligned with the rotated coords.

(ii) Construct response surface using only most important rotated coords.

→ Bivariate function $m(\boldsymbol{z})$ is effectively ***univariate*** in rotated coordinate system

* Constantine, Dow, Wang, *SISC*, 2014.

Slide 44

TI1 Tezaur, Irina, 6/1/2017

TI1

Active Subspaces (AS)

- KLE eigenvalue analysis suggests $d = O(1000)$ – **still large for MCMC**!

Idea: combine KLE with Active Subspace (AS) information for further (and better) ***data-informed*** dimension reduction.

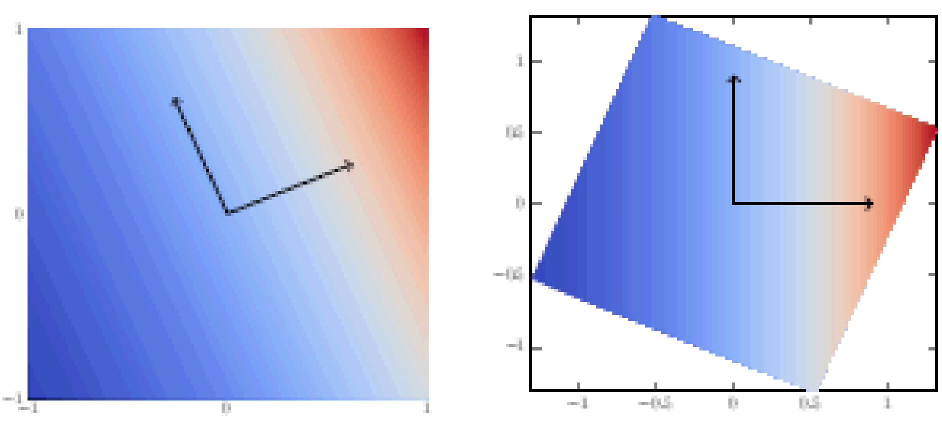
- ***Active Subspace (AS)*** = directions along which objective function has strongest variability.
- ***Active subspace approach:*** mismatch approximated by related function of fewer variables g:

$$m(\boldsymbol{z}) = \frac{1}{2}(\boldsymbol{d} - \boldsymbol{f}(\boldsymbol{z}))^T \boldsymbol{\Gamma}_{\text{obs}}^{-1}(\boldsymbol{d} - \boldsymbol{f}(\boldsymbol{z})) \approx g(\boldsymbol{W}_1^T \boldsymbol{z})$$

$\boldsymbol{W}_1^T \boldsymbol{z}$ = "active variables"
$\boldsymbol{W}_1^T$ = rotation of coords

- AS identified using ***gradients of mismatch function*** ∇m: $\int_{\mathbb{R}^d} \nabla m(z) \nabla m(z)^T d\rho(z) = \boldsymbol{W \Lambda W}^T$

Example*: $m(\boldsymbol{z}) = \exp(0.7 z_1 + 0.3 z_2)$



Dimension reduction via AS:

(i) Rotate coords s.t. directions of strongest variation are aligned with the rotated coords.

(ii) Construct response surface using only most important rotated coords.

→ Bivariate function $m(\boldsymbol{z})$ is effectively ***univariate*** in rotated coordinate system

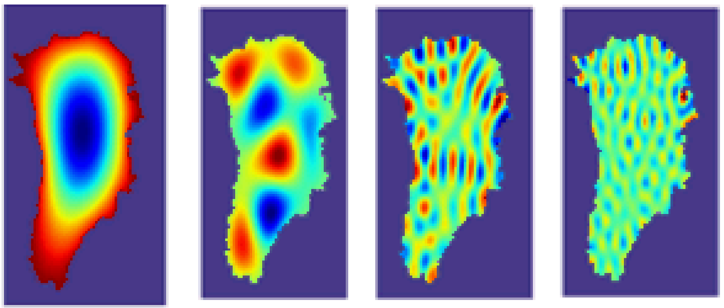
* Constantine, Dow, Wang, *SISC*, 2014.

Slide 45

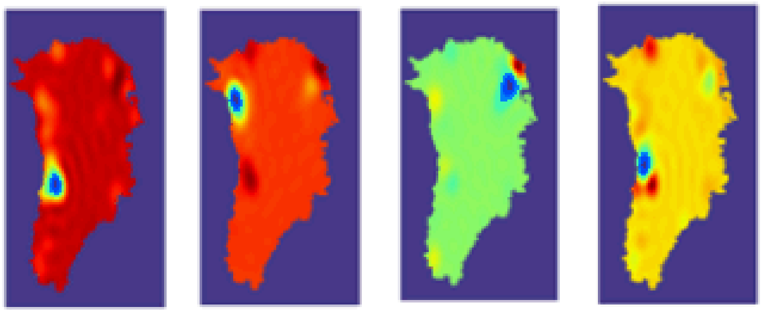
TI1 Tezaur, Irina, 6/1/2017

Greenland Bayesian Inference via KLE + AS

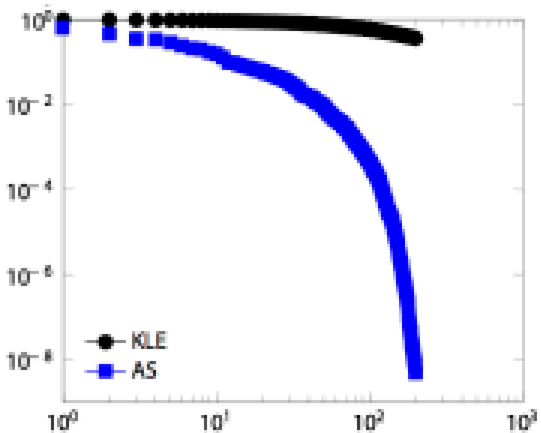
KLE modes



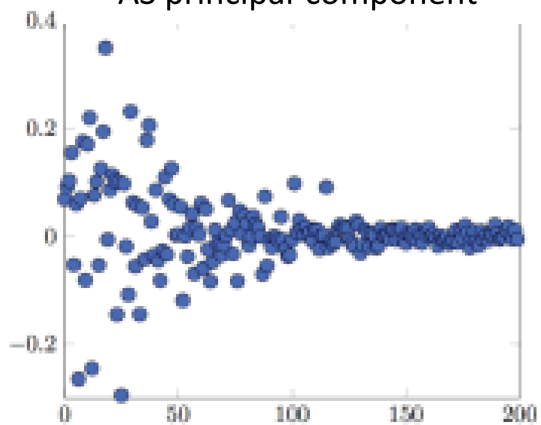
Data-informed (AS) directions (d=73*)



KLE and AS amplitudes



AS principal component



index

Gradients of mismatch function obtained via ***adjoint solve*** in Albany/FELIX.

- ***Above, left***: fewer modes are needed to build the basal friction parameter map when using KLE + AS methods than when using straight KLE.
- ***Above, right***: relative clustering of large values towards smaller indices implies KLE coefficients corresponding to larger singular values contribute most to variability in $m(\boldsymbol{z})$.

* Value of d was obtained via cross-validation.

Active Subspaces for Inference

Revert to prior in uninformed directions

$$\pi_{\text{pos}}(\mathbf{z}) \approx C_{\text{evid}}^{-1} \underbrace{\exp(-\hat{m}_s(\mathbf{W}_1^T\mathbf{z}))\pi_{\text{pr}}(\mathbf{W}_1^T\mathbf{z})} \overbrace{\pi_{\text{pr}}(\mathbf{W}_2^T\mathbf{z})}$$

Approximate posterior only in directions informed by data

Various levels of approximation can be employed:

- Reduce dimension but no surrogate of misfit
  - Perform MCMC in active subspace to improve mixing
- Surrogate of misfit with rotation but no dimension reduction
  - Leverage increased sparsity induced by rotation
- Surrogate of misfit and dimension reduction
  - **Combine MCMC in active subspaces with surrogates that adaptively target regions of high probability**

Quadratic PCE over Active Variables

Idea: approximate misfit $m(\boldsymbol{z})$ using quadratic PCE for efficient computation of misfit Hessian.

$m(\boldsymbol{z}) \approx \widehat{m}(\boldsymbol{z})$ = quadratic PCE function

- Approximate misfit over active variables using a quadratic function obtained via compressed sensing (using M= 733 samples and a PCE with 20,301 terms)*:

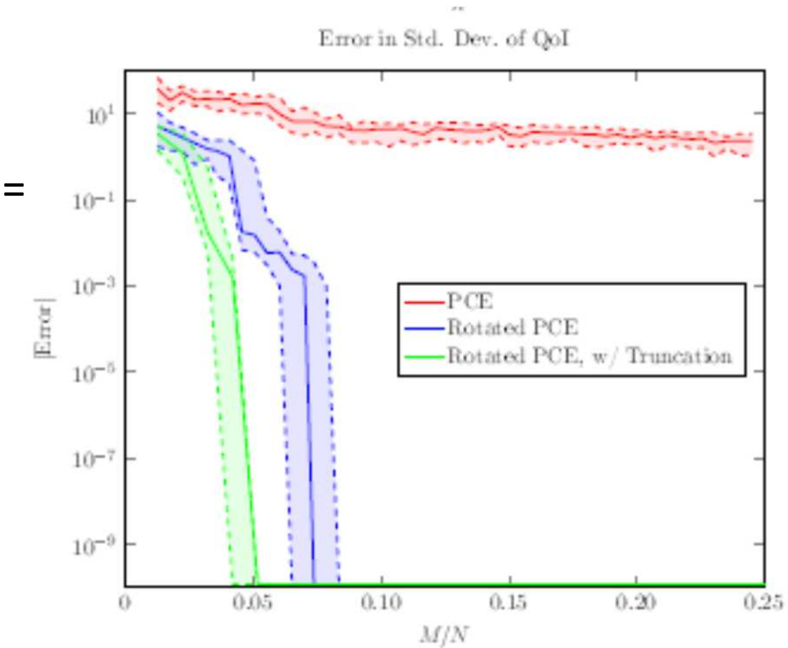
$$\frac{||m(\boldsymbol{z}) - \widehat{m}(\boldsymbol{z})||_{l^2_\rho}}{||m(\boldsymbol{z}) - \sum_{i=1}^{M} m(\boldsymbol{z}^{(i)})\,||_{l^2_\rho}} \approx 0.981$$

- Approximate misfit with quadratic PCE in rotated d = 200 space:

$$\frac{||m(\boldsymbol{z}) - \widehat{m}(\boldsymbol{W}^T\boldsymbol{z})||_{l^2_\rho}}{||m(\boldsymbol{z}) - \sum_{i=1}^{M} m(\boldsymbol{z}^{(i)})\,||_{l^2_\rho}} \approx 0.190$$

- Approximate misfit with ***quadratic PCE*** in ***rotated and truncated*** d = 73 space:

$$\frac{||m(\boldsymbol{z}) - \widehat{m}_{s=73}\,(\boldsymbol{W}_1^T\boldsymbol{z})||_{l^2_\rho}}{||m(\boldsymbol{z}) - \sum_{i=1}^{M} m(\boldsymbol{z}^{(i)})\,||_{l^2_\rho}} \approx 0.136$$



rank$(A) = 5$, $d = 15$, $z_i \sim N(1/2, 1/5)$

* Ratios are improvements relative to using mean of data; want ratio close to 0.

Low Rank Laplace-Based Covariance*

$$\pi_{\text{pos}}\left(\boldsymbol{z} \mid \boldsymbol{y}^{\text{obs}}\right) = N(\boldsymbol{z}_{\text{MAP}}, \boldsymbol{\Gamma}_{\text{post}})$$

- ***Linearize*** parameter-to-observable map around MAP point:

$$\boldsymbol{y}^{\text{obs}} = \boldsymbol{f}(\boldsymbol{z}) + \epsilon \approx \boldsymbol{f}(\boldsymbol{z}_{\text{MAP}}) + \boldsymbol{F}(\boldsymbol{z} - \boldsymbol{z}_{\text{MAP}}) + \epsilon$$

where $\boldsymbol{F}$ = Frechet derivative of $\boldsymbol{f}$.

- ***Covariance*** of Gaussian ***posterior*** given by:

$$\boldsymbol{\Gamma}_{\text{post}} = \left(\boldsymbol{H}_{PCE} + \boldsymbol{\Gamma}_{\text{prior}}^{-1}\right)^{-1}$$

Low Rank Laplace-Based Covariance*

$\boldsymbol{\Gamma}_{\text{post}}$ is dense!
⇒ ***prohibitively expensive*** to store & construct.

$$\pi_{\text{pos}}\left(\boldsymbol{z} \mid \boldsymbol{y}^{\text{obs}}\right) = N(\boldsymbol{z}_{\text{MAP}}, \boldsymbol{\Gamma}_{\text{post}})$$

- ***Linearize*** parameter-to-observable map around MAP point:

$$\boldsymbol{y}^{\text{obs}} = \boldsymbol{f}(\boldsymbol{z}) + \epsilon \approx \boldsymbol{f}(\boldsymbol{z}_{\text{MAP}}) + \boldsymbol{F}(\boldsymbol{z} - \boldsymbol{z}_{\text{MAP}}) + \epsilon$$

  where $\boldsymbol{F}$ = Frechet derivative of $\boldsymbol{f}$.

- ***Covariance*** of Gaussian ***posterior*** given by:

$$\boldsymbol{\Gamma}_{\text{post}} = \left(\boldsymbol{H}_{PCE} + \boldsymbol{\Gamma}_{\text{prior}}^{-1}\right)^{-1}$$

Low Rank Laplace-Based Covariance*

$$\pi_{\text{pos}}(\boldsymbol{z} \mid \boldsymbol{y}^{\text{obs}}) = N(\boldsymbol{z}_{\text{MAP}}, \boldsymbol{\Gamma}_{\text{post}})$$

$\boldsymbol{\Gamma}_{\text{post}}$ is dense!
⇒ ***prohibitively expensive*** to store & construct.

- ***Linearize*** parameter-to-observable map around MAP point:

$$\boldsymbol{y}^{\text{obs}} = \boldsymbol{f}(\boldsymbol{z}) + \epsilon \approx \boldsymbol{f}(\boldsymbol{z}_{\text{MAP}}) + \boldsymbol{F}(\boldsymbol{z} - \boldsymbol{z}_{\text{MAP}}) + \epsilon$$

  where $\boldsymbol{F}$ = Frechet derivative of $\boldsymbol{f}$.

- ***Covariance*** of Gaussian ***posterior*** given by:

$$\boldsymbol{\Gamma}_{\text{post}} = \left(\boldsymbol{H}_{PCE} + \boldsymbol{\Gamma}_{\text{prior}}^{-1}\right)^{-1}$$

- ***Low-rank approximation*** of $\boldsymbol{\Gamma}_{\text{post}}$ obtained using Sherman-Morrison-Woodbury formula:

$$\boldsymbol{\Gamma}_{\text{post}} \approx \boldsymbol{\Gamma}_{\text{prior}} - \widetilde{\boldsymbol{V}}_r \boldsymbol{D}_r \widetilde{\boldsymbol{V}}_r^{\Diamond}$$

Symbols*:

$\boldsymbol{V}_r, \boldsymbol{D}_r$: eigenvecs, eigenvals of $\widetilde{\boldsymbol{H}}_{\text{misfit}}$

$\widetilde{\boldsymbol{H}}_{\text{misfit}}$ = prior-preconditioned Hessian of data misfit = $\boldsymbol{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{H}_{\text{misfit}} \boldsymbol{\Gamma}_{\text{prior}}^{1/2}$

$\boldsymbol{H}_{\text{misfit}}$ = Gauss-Newton portion of Hessian misfit = $\boldsymbol{F}^{\square} \boldsymbol{\Gamma}_{\text{obs}}^{-1} \boldsymbol{F}$

$\widetilde{\boldsymbol{V}}_r$ = $\boldsymbol{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{V}_r$, $\widetilde{\boldsymbol{V}}_r^{\Diamond}$ = adjoint of $\widetilde{\boldsymbol{V}}_r$

$\boldsymbol{\Gamma}_{\text{prior}}^{-1}$ = $\boldsymbol{M}^{-1}\boldsymbol{K}$, $\boldsymbol{K}$= Laplace stiffness.

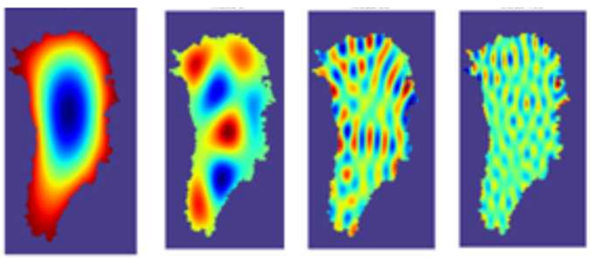
- $\widetilde{\boldsymbol{H}}_{\text{misfit}}$ and its EV decomposition can be computed efficiently using a parallel ***matrix-free Lanczos method***.

- ***Rank of*** $\boldsymbol{\Gamma}_{\text{post}}$ = # of directions that informed directions of posterior.

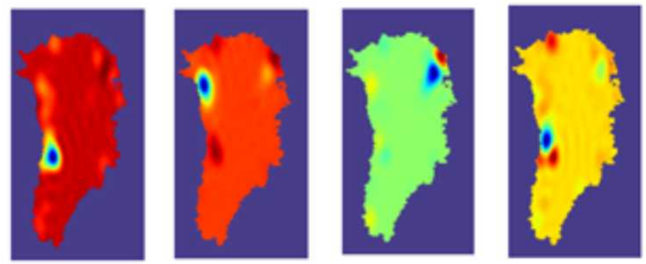
Greenland Bayesian Inference via KLE + AS

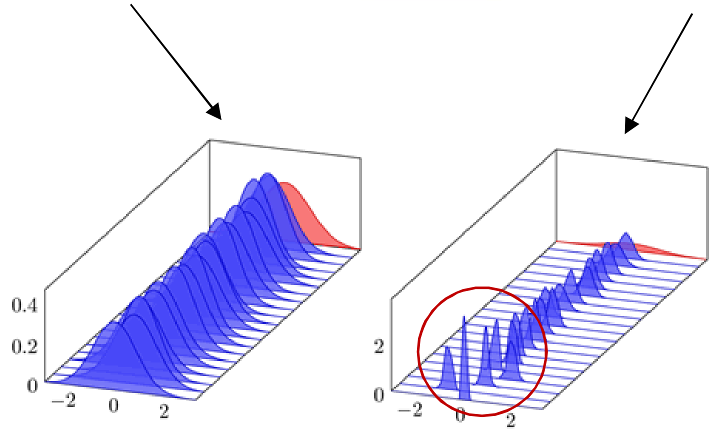
Sandia National Laboratories

KLE modes



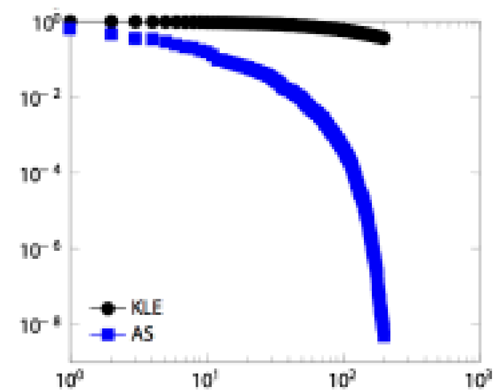
Data-informed (AS) directions (d=73)



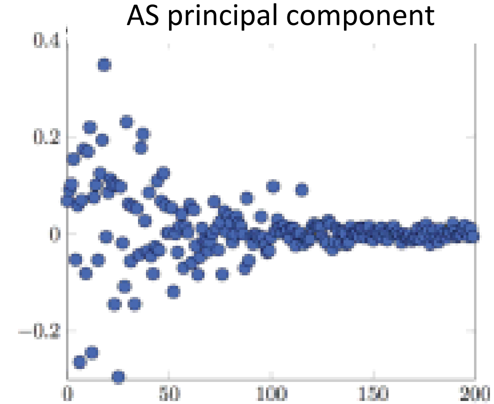


Gradients of mismatch function obtained via ***adjoint solve*** in Albany/FELIX.

KLE and AS amplitudes



AS principal component



- ***Above:*** marginal distributions of Gaussian posterior computed using KLE vs. KLE+AS.
    - Data-informed eigenvectors have smaller variance and are most shifted w.r.t. prior distribution.

Outline

1. Background.
    - PISCEES project for land-ice modeling.
    - Land-ice model.
2. UQ problem definition.
3. Inversion/calibration.
    - Deterministic inversion.
    - Bayesian inference.
4. **Summary & future work.**

Summary & future work

This talk described our ***workflow*** for quantifying uncertainties in expected aggregate ice sheet mass change and its ***demonstration*** on some Greenland ice sheet problems, focusing on ***inversion***.

Summary & future work

This talk described our ***workflow*** for quantifying uncertainties in expected aggregate ice sheet mass change and its ***demonstration*** on some Greenland ice sheet problems, focusing on ***inversion***.

- **<u>Future work:</u>**
  - Execute ***full UQ workflow*** (inversion + forward propagation) on realistic Greenland/Antarctic ice sheet problems.
  - ***Squared Laplace covariance operator approach**** (no KLE) → less expensive than building PCE, allows higher dimensional parameter spaces.
  - Can use ***cheaper physical models*** (e.g., the shallow ice model or SIA) or ***low resolution solves*** to reduce the cost of building the emulator.
  - Incorporate effects of ***other sources of uncertainty***, e.g., surface height, surface mass balance.

* Bui-Thanh, Ghattas, Martin, Stadler, *SISC*, 2013.

Summary & future work

This talk described our ***workflow*** for quantifying uncertainties in expected aggregate ice sheet mass change and its ***demonstration*** on some Greenland ice sheet problems, focusing on ***inversion***.

- **Future work:**

We are well-positioned to **do these efforts in parallel!**

  - Execute ***full UQ workflow*** (inversion + forward propagation) on realistic Greenland/Antarctic ice sheet problems.
  - ***Squared Laplace covariance operator approach**** (no KLE) → less expensive than building PCE, allows higher dimensional parameter spaces.
  - Can use ***cheaper physical models*** (e.g., the shallow ice model or SIA) or ***low resolution solves*** to reduce the cost of building the emulator.
  - Incorporate effects of ***other sources of uncertainty***, e.g., surface height, surface mass balance.

* Bui-Thanh, Ghattas, Martin, Stadler, *SISC*, 2013.

Funding/acknowledgements

Support for this work was provided through Scientific Discovery through Advanced Computing (**SciDAC**) projects funded by the U.S. Department of Energy, Office of Science (**OSCR**), Advanced Scientific Computing Research and Biological and Environmental Research (**BER**) → **PISCEES SciDAC Application Partnership.**

PISCEES team members: *K. Evans, M. Gunzburger, M. Hoffman, C. Jackson, P. Jones, W. Lipscomb, M. Perego, S. Price, A. Salinger, I. Tezaur, R. Tuminaro, P. Worley.*

Trilinos/DAKOTA collaborators*: M. Eldred, J. Jakeman, E. Phipps, L. Swiler.*

Computing resources: *NERSC, OLCF.*

References

[1] M.A. Heroux *et al.* "An overview of the Trilinos project." *ACM Trans. Math. Softw.* **31**(3) (2005).

[2] A. Salinger, *et al.* "Albany: Using Agile Components to Develop a Flexible, Generic Multiphysics Analysis Code", *Int. J. Multiscale Comput. Engng*. 14(4) (2016) 415-438.

[3] **I. Tezaur**, M. Perego, A. Salinger, R. Tuminaro, S. Price. "*Albany/FELIX*: A Parallel, Scalable and Robust Finite Element Higher-Order Stokes Ice Sheet Solver Built for Advanced Analysis", *Geosci. Model Develop.* 8 (2015) 1-24.

[4] M. Perego, S. Price, G. Stadler. "Optimal initial conditions for coupling ice sheet models to Earth system models". *J. Geophys. Res. Earth Surf.* 119 (2014) 1894-1917.

[5] P. Constantine, E. Dow, Q. Wang. "Active subspace methods in theory and practice: applications to Kriging surfaces". *SIAM J. Sci. Comput.* 36(4) (2014) A1500-A1524.

[6] P. Constantine, C. Kent, T. Bui-Thanh. "Accelerating Markov Chain Monte Carlo with Active Subspaces". *SIAM J. Sci. Comput.* 38 (5) (2016) A2779-A2805.

[7] T. Bui-Thanh, O. Ghattas. J. Martin. G. Stadler. "A computational framework for infinite-dimensional Bayesian inverse problems part I: the linearized case, with application to global seismic inversion". *SIAM J. Sci. Comput.* 35(6) (2013) A2494-2523.

[8] T. Isaac, N. Petra, G. Stadler, O. Ghattas. "Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet". *J. Comput. Phys.* 296 (2015) 348-368.

[9] J. Jakeman, M. Eldred, K. Sargsyan. "Enhancing l1-minimization estimates of polynomial chaos expansions using basis selection", *J. Comp. Phys.* 289 (2015) 18-34.

[10] N. Petra, J. Martin, G. Stadler, O. Ghattas. "A computational framework for infinite dimensional Bayesian inverse problems. Part II: stochastic Newton with application to ice sheet flow inverse problems". *SIAM J. Sci. Comput.* 36 (4) (2014) A1525-1555.

[11] J. Jakeman, **I. Tezaur**, M. Perego, S. Price. "Probabalistic Projections of Sea-Level Change from the Greenland Ice Sheet", in preparation.

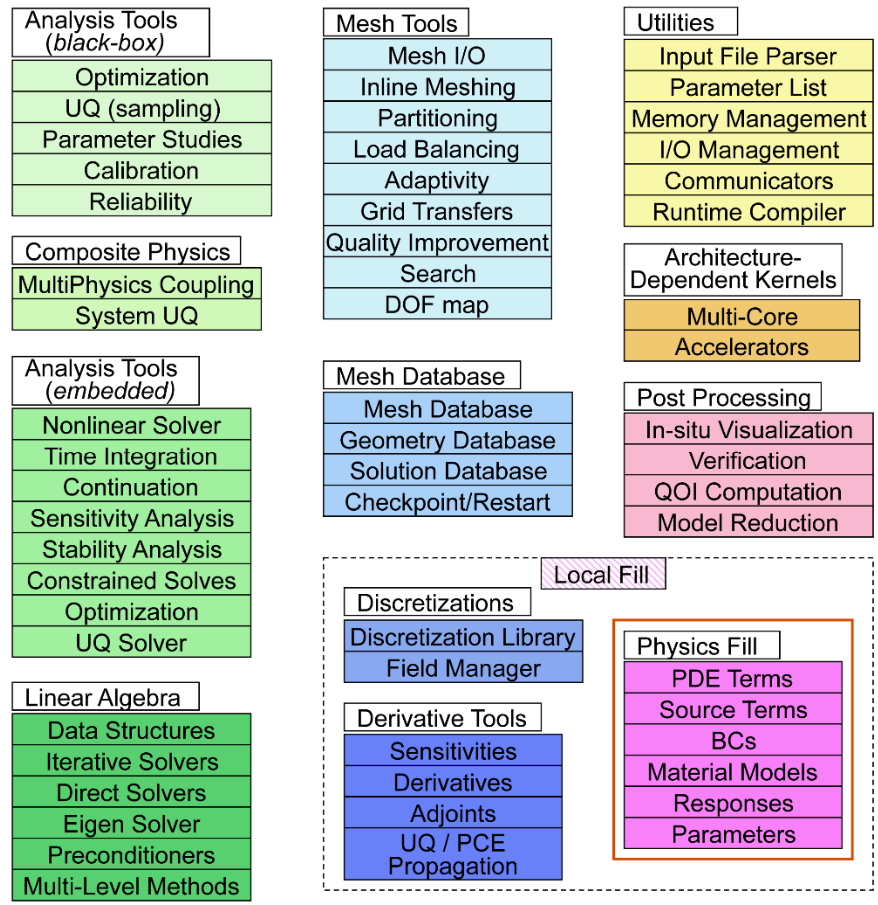
Albany Multiphysics Code

The ***Albany/FELIX land-ice solver*** is implemented within the ***Albany multi-physics code***.

Albany = Sandia open-source* parallel, C++, multi-physics finite element code.

- ***Component-based*** design for rapid development of new physics & capabilities.
- Extensive use of libraries from the open-source ***Trilinos*** project:
  - Automatic differentiation.
  - Discretizations/meshes, mesh adaptivity.
  - Solvers, time-integration schemes.
  - Performance-portable kernels.
- ***Advanced analysis*** capabilities:
  - Parameter estimation.
  - Uncertainty quantification (DAKOTA).
  - Optimization (DAKOTA, ROL).
  - Sensitivity analysis.

40+ packages; 120+ libraries



* https://github.com/gahansen/Albany.

Computing the Active Subspace

Gradients of mismatch $\nabla_\beta m$ can be used to identify subspace that controls variation in likelihood function (active subspace)

- Mismatch ***approximated*** by related function of fewer variables g:

$W_1^T \boldsymbol{z}$ = "active variables"

$$m(\boldsymbol{z}) = \frac{1}{2}(\boldsymbol{d} - \boldsymbol{f}(\boldsymbol{z}))^T \boldsymbol{\Gamma}_{\text{obs}}^{-1}(\boldsymbol{d} - \boldsymbol{f}(\boldsymbol{z})) \approx g(\underbrace{\boldsymbol{W}_1^T \boldsymbol{z}}_{\text{Linear transformation (rotation) of coords}})$$

- Active subspace computed using $\int_{\mathbb{R}^d} \nabla m(z) \nabla m(z)^T d\rho(z) = \boldsymbol{W\Lambda W}^T$
  - Sample gradient using MC: $\left[\nabla m\left(z^{(1)}\right), \ldots, \nabla m\left(z^{(M)}\right)\right]$.
  - Form Gauss-Newton approx. of Hessian averaged over prior:

$$\boldsymbol{C} = \frac{1}{M}\sum_{i=1}^{M} \nabla m\left(z^{(i)}\right) \nabla m\left(z^{(i)}\right)^T$$

  - Compute eigenvalue decomposition: $\boldsymbol{C} = \boldsymbol{W\Lambda W}^T$
    → eigenvectors $\boldsymbol{W}$ define rotation of $\mathbb{R}^M$.

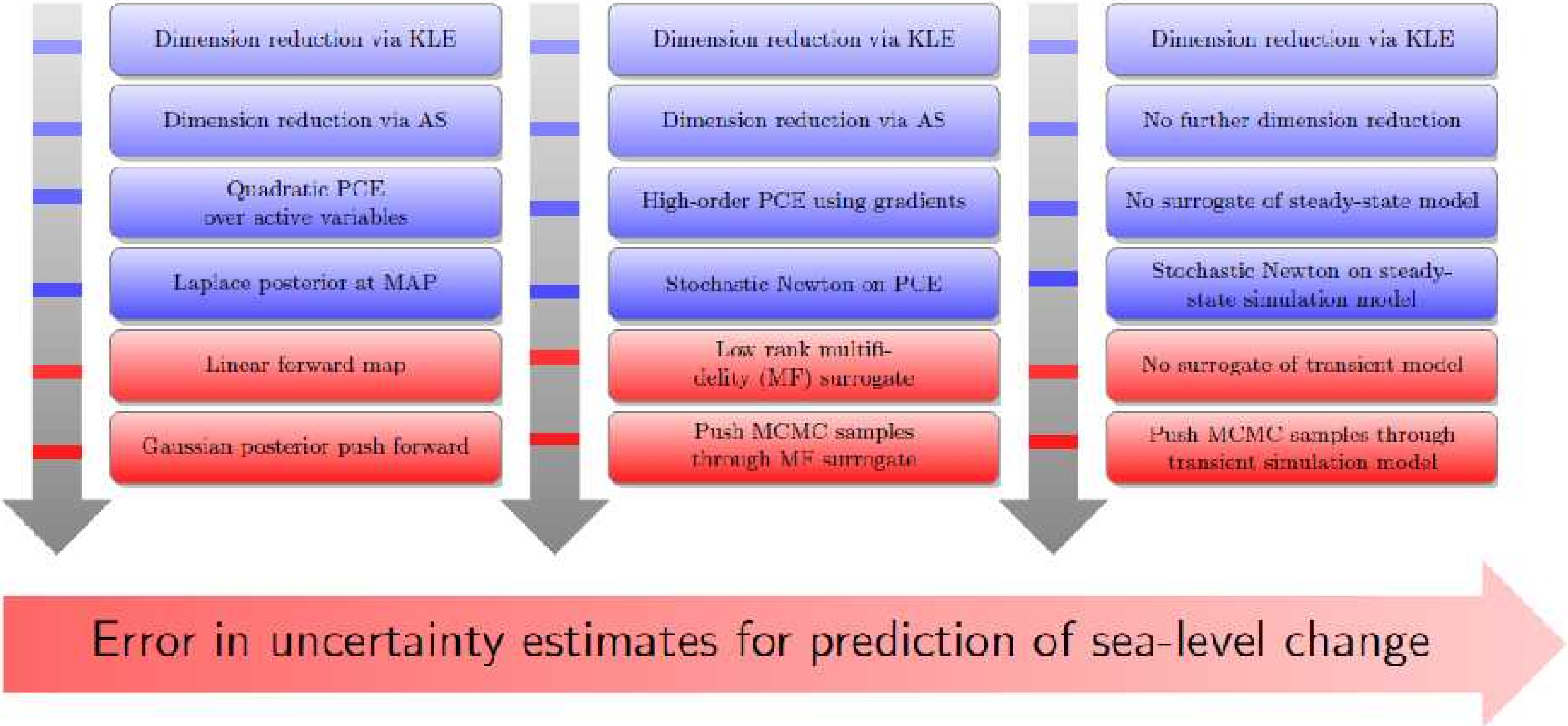
Perturbing $m(\boldsymbol{z})$ along columns of $\boldsymbol{W}_1$ changes $m(\boldsymbol{z})$ more.

- Partition $\boldsymbol{z}$ into ***active*** and ***inactive*** variables:

$$\boldsymbol{z} = \boldsymbol{W}_1^T \boldsymbol{z} + \boldsymbol{W}_2^T \boldsymbol{z}, \quad \boldsymbol{W} = [\boldsymbol{W}_1 \quad \boldsymbol{W}_2]$$

Full UQ Workflow: Varying Levels of Approx.

Sandia National Laboratories



<u>**As with Bayesian inference:**</u>

- ***Future work:*** compare errors as accuracy of approximation is increased to gain insight into viability of lower-dimensional approximations.
- Lessons can be learned by avoiding use of highest fidelity model.