Propagation of Uncertainty from Data through Inference to Prediction for Large-Scale Problems with Application to Ice Sheet Flow[1]

Noémi Petra

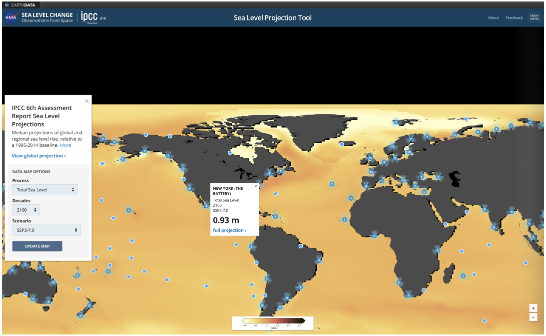
Department of Applied Mathematics
University of California, Merced

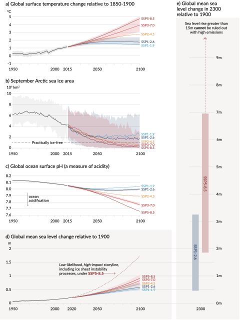
NYSDS 2021: New York Scientific Data Summit 2021:
Data-Driven Discovery in Science and Industry
October 28, 2021

[1]Research supported partially by NSF-DMS-CAREER-1654311 and NSF-OAC-1550547.

Need predictive models with quantified uncertainties to accurately anticipate future sea level rise.

Sea level projections from the IPCC 6th Assessment Report, 2021





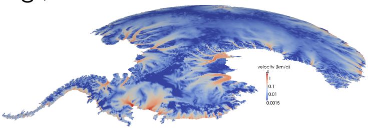
- Several port cities will be at risk from coastal flooding in the future.
- Ice flowing from ice sheets to ocean is primary contributor to sea level rise.

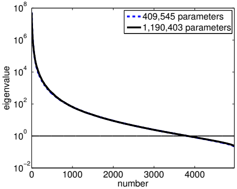
  Details in: Masson-Delmotte, V. et al. "IPCC, 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change", Cambridge University Press. In Press.

Challenges and the need to exploit problem structure and account for model error

Severe mathematical and computational challenges place significant barriers on improving predictability of ice sheet flow models, e.g.,

- complex and very high-aspect ratio (thin) geometry,
- highly nonlinear and anisotropic rheology,
- extremely ill-conditioned and large-scale linear and nonlinear algebraic systems that arise upon discretization,
- uncertain basal sliding parameter, basal topography, geothermal heat flux, and rheology,
- modeling error, etc.

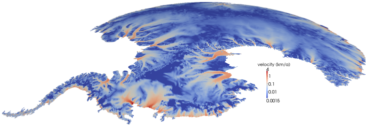


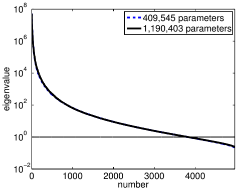


Details in: T. Isaac, N. Petra, G. Stadler, and O. Ghattas. "Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet", Journal of Computational Physics, 296, 348-368 (2015). *Selected as the 2019 SIAM Activity Group on Computational Science and Engineering Best Paper.*

How can we addresss (some of) these challenges?

- Learn models from data, i.e., infer unknown/uncertain parameters from available data, e.g., satellite measurements of surface ice flow velocity (**statistical inverse problems governed by PDEs**)
- Apply/adapt/design fast, mesh-independent, structure exploiting, inner-product- and additional uncertainty-aware methods (**scalable, robust and efficient algorithms**)
- Collect data in an optimal way in order to minimize the uncertainty in the inferred parameters or in some predictive quantity of interest (**optimal experimental design**).





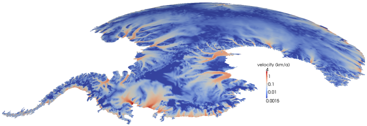
Details in: O. Ghattas and K. Willcox. "Learning physics-based models from data: perspectives from inverse problems and model reduction", Acta Numerica, Cambridge University Press (2021).

The forward problem

Nonlinear Stokes ice sheet model (for viscous, shear-thinning, incompressible fluid)

Invoking the balance of mass and linear momentum:

$$
\begin{aligned}
-\boldsymbol{\nabla} \cdot [2\gamma(\boldsymbol{u}, n)\,\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} - \boldsymbol{I}p] = \rho \boldsymbol{g} \quad & \text{in } \Omega \\
\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \quad & \text{in } \Omega \\
\boldsymbol{\sigma}_{\boldsymbol{u}} \boldsymbol{n} = \boldsymbol{0} \quad & \text{on } \Gamma_t \\
\boldsymbol{u} \cdot \boldsymbol{n} = 0,\, \boldsymbol{T}\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n} + \exp(\beta)\boldsymbol{T}\boldsymbol{u} = \boldsymbol{0} \quad & \text{on } \Gamma_b
\end{aligned}
$$



- $\boldsymbol{u}$ ice flow velocity, p pressure
- $\boldsymbol{\sigma}_{\boldsymbol{u}} = -\boldsymbol{I}p + 2\gamma(\boldsymbol{u}, n)\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}$ stress tensor
- $\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} = \frac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\boldsymbol{u}^T)$ strain rate tensor
- $\gamma(\boldsymbol{u}, n) = \frac{1}{2} A^{-\frac{1}{n}}\, \dot{\varepsilon}_{\mathrm{II}}^{\frac{1-n}{2n}}$ effective viscosity
- $\dot{\varepsilon}_{\mathrm{II}} = \frac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)$ second invariant of the strain rate tensor

- ρ density, $\boldsymbol{g}$ gravity
- $\boldsymbol{n}$ unit normal vector
- β log basal sliding coefficient
- $\boldsymbol{T} = \boldsymbol{I} - \boldsymbol{n} \otimes \boldsymbol{n}$ tangential operator
- Γ_t and Γ_b top and base boundaries

The inverse problem

Use available observations/data $\boldsymbol{d}$ to infer the values of the unknown parameter field m that characterize a physical process modeled by PDEs, i.e.,

$$\boldsymbol{d} = \mathcal{F}(m) + \boldsymbol{\eta}.$$

- The map $\mathcal{F} : \mathcal{M} \to \mathbb{R}^q$ is the so-called *parameter-to-observable* map.
- Evaluations of $\mathcal{F}$ involve the solution of the Stokes PDE given m, followed by the application of an observation operator $\mathcal{B} : \mathcal{V} \to \mathbb{R}^q$ to extract the observations from the state.
- $\boldsymbol{\eta}$ accounts for noisy measurements and model errors and is modeled as $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{noise}})$, i.e., a centered Gaussian at $\mathbf{0}$ with covariance $\mathbf{\Gamma}_{\text{noise}}$.

Bayesian formulation of the inverse problem

Describes probability of all models that are consistent with the observations/data and any prior knowledge about the parameters:

$$d\mu_{\text{post}} \propto \exp\left\{ -\frac{1}{2}\|\mathcal{F}(m) - \boldsymbol{d}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} - \frac{1}{2}\|m - m_{\text{pr}}\|^2_{\mathcal{C}^{-1}_{\text{prior}}} \right\}.$$

- The first term in the exponential is the negative log-likelihood.
- The second term represent the negative log-prior (e.g., Gaussian prior, i.e., $m \sim \mathcal{N}(m_{\text{pr}}, \mathcal{C}_{\text{prior}})$).

Goal:

- characterize the posterior statistically (MAP point, mean, covariance, etc.)
- for functions $\boldsymbol{m}$ (large vectors after discretization)
- for expensive $\mathcal{F}(\cdot)$
- exploit connection to PDE-constrained optimization

Inverse problems governed by PDEs

- The *maximum a posteriori* (MAP) point m_{MAP} is defined as the parameter field that maximizes the posterior distribution:

$$
\begin{aligned}
m_{\text{MAP}} :&= \operatorname*{argmin}_{m \in \mathcal{M}} (-\log d\mu_{\text{post}}(m)) \\
&= \operatorname*{argmin}_{m \in \mathcal{M}} \frac{1}{2} \|\mathcal{F}(m) - \boldsymbol{d}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2} \|m - m_{\text{pr}}\|^2_{\mathcal{C}^{-1}_{\text{prior}}} \\
&= \operatorname*{argmin}_{m \in \mathcal{M}} \frac{1}{2} \|\mathcal{B}(u) - \boldsymbol{d}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2} \|m - m_{\text{pr}}\|^2_{\mathcal{C}^{-1}_{\text{prior}}},
\end{aligned}
$$

where u solves the forward Stokes (PDE) problem.

Inverse problems governed by PDEs

- The *maximum a posteriori* (MAP) point m_{MAP} is defined as the parameter field that maximizes the posterior distribution:

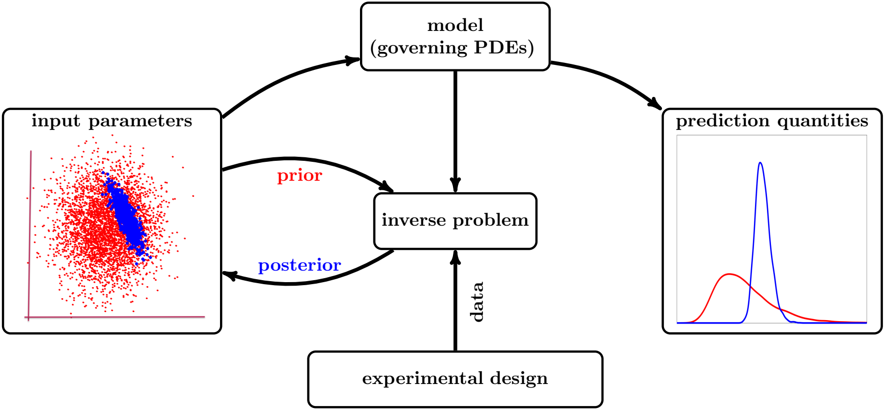
$$\begin{aligned} m_{\text{MAP}} :&= \underset{m \in \mathcal{M}}{\operatorname{argmin}} (-\log d\mu_{\text{post}}(m)) \\ &= \underset{m \in \mathcal{M}}{\operatorname{argmin}} \frac{1}{2} \|\mathcal{F}(m) - \boldsymbol{d}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2} \|m - m_{\text{pr}}\|^2_{\mathcal{C}^{-1}_{\text{prior}}} \\ &= \underset{m \in \mathcal{M}}{\operatorname{argmin}} \frac{1}{2} \|\mathcal{B}(u) - \boldsymbol{d}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2} \|m - m_{\text{pr}}\|^2_{\mathcal{C}^{-1}_{\text{prior}}}, \end{aligned}$$

  where u solves the forward Stokes (PDE) problem.
- When $\mathcal{F}$ is linear, due to the particular choice of prior and noise model, the posterior measure is Gaussian, $\mathcal{N}(m_{\text{MAP}}, \mathcal{C}_{\text{post}})$

$$m_{\text{MAP}} = \mathcal{C}_{\text{post}}(\mathcal{F}^* \boldsymbol{\Gamma}^{-1}_{\text{noise}} \boldsymbol{d} + \mathcal{C}^{-1}_{\text{prior}} m_{\text{pr}}), \qquad \mathcal{C}_{\text{post}} = \mathcal{H}^{-1} = (\mathcal{F}^* \boldsymbol{\Gamma}^{-1}_{\text{noise}} \mathcal{F} + \mathcal{C}^{-1}_{\text{prior}})^{-1},$$

  where $\mathcal{F}^* : \mathbb{R}^q \to \mathcal{M}$ is the adjoint of $\mathcal{F}$, and $\mathcal{H}$ is the Hessian (second derivative) of the negative-log posterior.
- **Note:** In the general case of nonlinear parameter-to-observable map $\mathcal{F}$ the posterior distribution is not Gaussian.

Data→Inference→Prediction



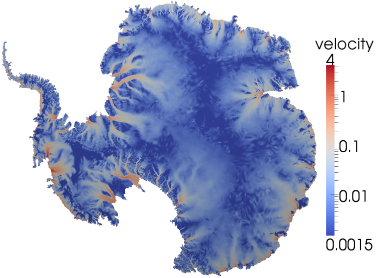
[1]Schematic courtesy of Alen Alexanderian.

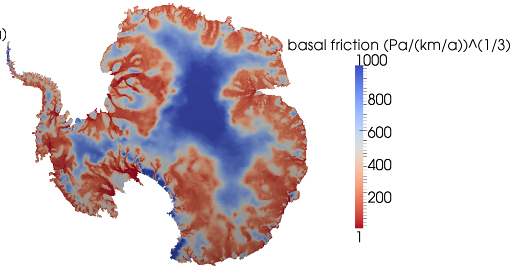
Data→Inference→Prediction

- While one can formulate a **data-to-prediction framework** to **quantify uncertainties from data to inferred model parameters to predictions** with an underlying model of non-Newtonian ice sheet flow, attempting to execute this framework for the Antarctic ice sheet (or other large-scale complex models) is intractable for high-dimensional parameter fields using current algorithms.
- Yet, quantifying the uncertainties in predictions of ice sheet models is essential if these models are to play a significant role in projections of future sea level.
- **Goal:** design an integrated framework and efficient, scalable algorithms (under Gaussian approximations of the posterior and prediction) for carrying out this data-to-prediction process.
- **Scalable:** the cost-measured in number of PDE solves-is independent of not only the number of processor cores, but importantly the state variable dimension, the parameter dimension, and the data dimension.

Details in: T. Isaac, N. Petra, G. Stadler, and 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", Journal of Computational Physics, 296, 348-368 (2015).

Antarctic ice sheet inversion for basal sliding field InSAR data, dim: 400k, fwd solver: ymir (T. Isaac *et al.*)

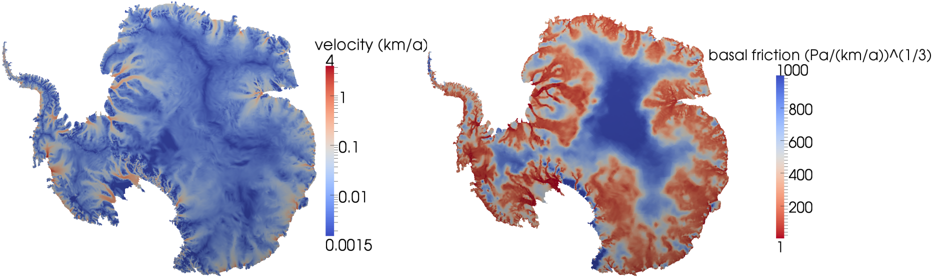




Left: InSAR-based Antarctica ice surface velocity observations
Right: Inferred basal sliding field (MAP point)

Details in: T. Isaac, N. Petra, G. Stadler, and 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", Journal of Computational Physics, 296, 348-368 (2015).

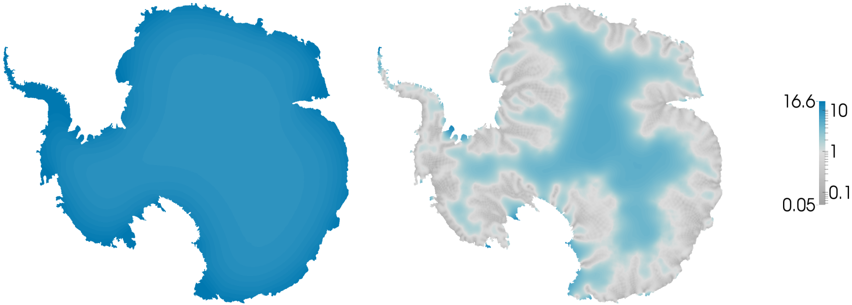
Antarctic ice sheet inversion for basal sliding field
InSAR data, dim: 400k, fwd solver: ymir (T. Isaac *et al.*)



Left: Recovered ice surface velocity observations
Right: Inferred basal sliding field (MAP point)

Details in: T. Isaac, N. Petra, G. Stadler, and 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", Journal of Computational Physics, 296, 348-368 (2015).

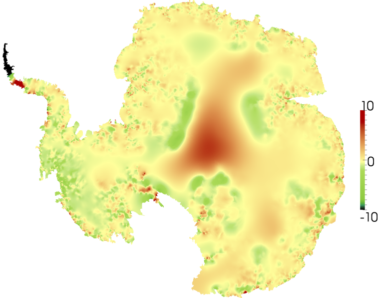
Gaussian approximation of the posterior



The standard deviations of the pointwise marginals of the prior distribution (left) and of (the Gaussian approximation of) the posterior distribution (right).

Details in: T. Isaac, N. Petra, G. Stadler, and 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", Journal of Computational Physics, 296, 348-368 (2015).

Data→Inference→Prediction for ice sheet flow

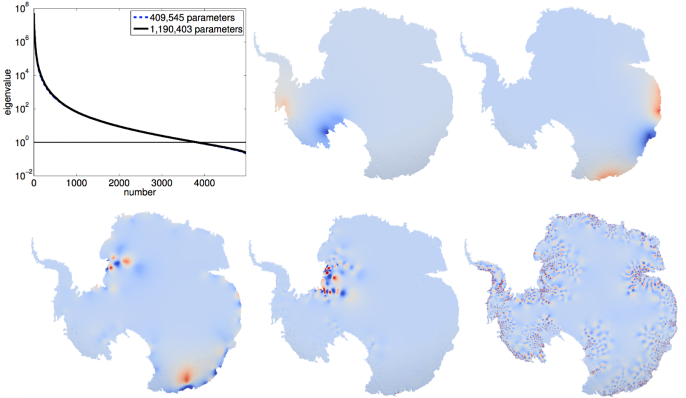


The gradient (left) and the "influential direction" in parameter space (right) for the ice mass flux from Totten Glacier to ocean. The mean and standard deviation of the prediction probability distribution for the ice mass flux is 71.24 ± 0.30 Gt/a.

Details in: T. Isaac, N. Petra, G. Stadler, and 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", Journal of Computational Physics, 296, 348-368 (2015).

Hessian for Antarctic ice flow inverse problem

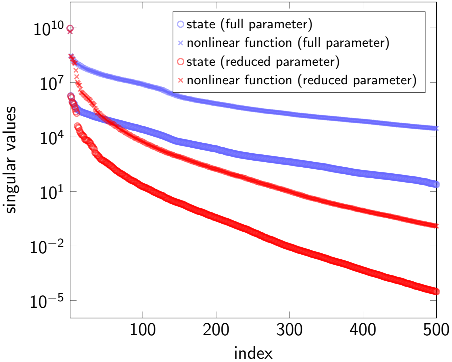
Eigenvalue decay of prior-preconditioned data misfit Hessian



- The data are informative about only a low-dimensional subspace within the high-dimensional parameter space (here only 5000 out of 1.19 million!).
- The data-to-prediction process is sensitive only to the true information contained within the data, as opposed to the data or parameter dimensions.

Joint parameter and state dimension reduction

via proper orthogonal decomposition and discrete empirical interpolation

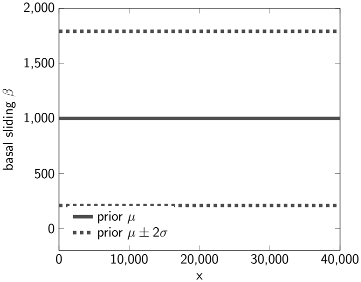


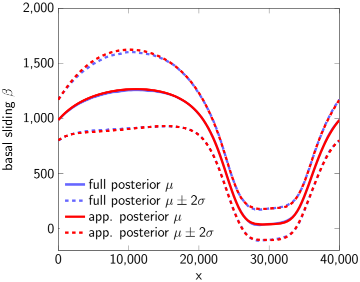
The singular values of the snapshots.

Details in: K. Kim, B. Peherstorfer, T. Cui, Y. Marzouk, K. Willcox, O. Ghattas, and N. Petra. "Joint Parameter and Model Dimension Reduction for Bayesian Inverse Problems with Application to a Nonlinear Stokes Ice Sheet Flow". In preparation.

Joint parameter and state dimension reduction

The prior and posterior variance (ISMIP-HOM benchmark problem)



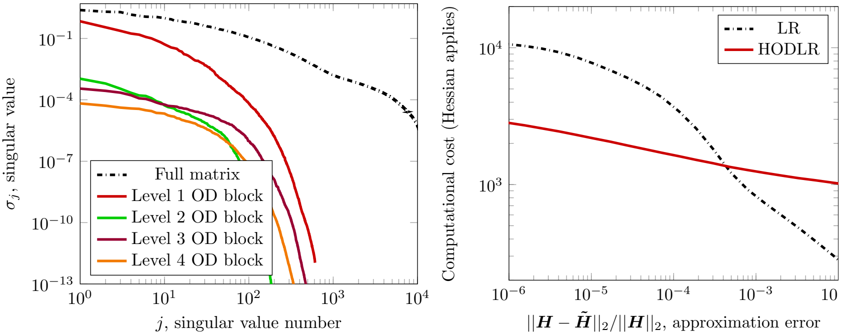


- Full posterior: parameter dimension: 120, state dimension: 4,920
- Jointly approximated posterior: parameter dimension: 10, state dimension: 30 (200 dim. DEIM)
- 60,000 samples from the full posterior, 15,000 samples from the jointly-approximated posterior
- H-pCN (Hessian-informed Crank-Nicolson) MCMC method is used to sample.

  Details in: K. Kim, U. Villa, M. Parno, N. Petra, Y. Marzouk, and O. Ghattas.. "hIPPYlib-MUQ: Scalable Markov chain Monte Carlo sampling methods for large-scale Bayesian inverse problems governed by PDEs". To be submitted. (https://github.com/hippylib/hippylib2muq)

Fast high-rank Hessian approximation for Bayesian ice sheet inverse problems via $\mathcal{H}$-matrices

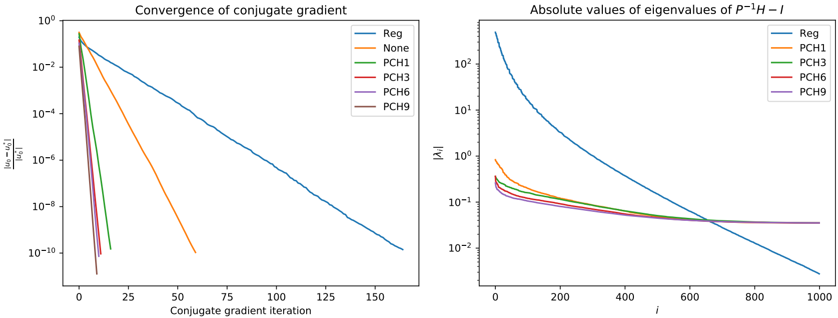
Greenland's Humboldt glacier



Left: Block spectra of the Hessian. Right: Comparison of the hierarchical off-diagonal low-rank (HODLR) and global low-rank (LR) approaximations.

Details in: T. Hartland, G. Stadler, M. Perego and K. Liegeois, and N. Petra. "Hierarchical Off-Diagonal Approximation of Hessians in Inverse Problems", In preparation.

Fast high-rank Hessian approximation for Bayesian ice sheet inverse problems via $\mathcal{H}$-matrices



Left: Error versus Krylov iterations.
Right: Spectrum of the preconditioned Hessian operator. ($\mathcal{H}$1-matrix format)

Details in: N. Alger, T. Hartland, N. Petra and O. Ghattas. "Fast matrix-free approximation of smoothly varying blur operators, with application to Hessians in PDE-constrained inverse problems with highly informative data", In preparation.

Inverse problem governed by random PDE forward problem

In reality, models have multiple sources of uncertainties and randomness (e.g., models have multiple uncertain coefficients, unknown or random source terms, and parameters that cannot be inferred.)

- Inverting for all unknown/uncertain parameters at once is not practical/feasable (due to high(er) parameter dimensions, more severe ill-posedness).)
- Goal: choose a *primary* parameter of interest to infer for and account for *additional uncertainties* using stochastic forward models:

$$r(u(\cdot,\xi), m(\cdot); \xi(\cdot)) = 0 \text{ a.s.}$$

  - $r \in \mathcal{D}$ a.e. a stochastic PDE, $\mathcal{D} \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ domain;
  - $u \in \overline{\mathcal{D}} \times \Omega \to \mathbb{R}$ state/forward variables;
  - $m \in \mathcal{M}$ inversion parameters;
  - $\xi : \Omega \to \mathbb{R}^p$ $(p = 2, 3)$ defined by means of a probability space $(\Omega, F, \mathbb{P})$, where Ω is the sample space (the set of all possible events), F is the $\sigma-$algebra of events, and $\mathbb{P} : F \to [0, 1]$ is a probability measure.

Premarginalization over secondary uncertain parameters

using the Bayesian Approximation Error (BAE) approach

- **Key idea:** carry out estimation of the primary uncertain parameter β while **taking into account the uncertainty in the auxiliary parameter** a:

$$\boldsymbol{d} = \tilde{\mathcal{F}}(\beta, a) + \boldsymbol{\eta} = (\mathcal{F}(\beta) + \boldsymbol{\varepsilon}) + \boldsymbol{\eta} = \mathcal{F}(\beta) + \boldsymbol{\nu}$$

  - $\tilde{\mathcal{F}}(\beta, a)$: the accurate parameter-to-observable mapping
  - $\mathcal{F}(\beta) = \tilde{\mathcal{F}}(\beta, a_*)$: the *approximative parameter-to-observable mapping*, where the auxiliary parameter a is fixed to the mean value a_* (assume a is Gaussian distributed)

$$\begin{aligned}
\boldsymbol{\eta} &= \boldsymbol{d} - \tilde{\mathcal{F}}(\beta, a) & \boldsymbol{\eta} &\sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{noise}}) & &\text{(data noise)}\\
\boldsymbol{\varepsilon} &= \tilde{\mathcal{F}}(\beta, a) - \mathcal{F}(\beta) & \boldsymbol{\varepsilon} &\sim \mathcal{N}(\boldsymbol{\varepsilon}_*, \mathbf{\Gamma}_{\boldsymbol{\varepsilon}}) & &\text{(approximation/model error)}\\
\boldsymbol{\nu} &= \boldsymbol{\varepsilon} + \boldsymbol{\eta} & \boldsymbol{\nu} &\sim \mathcal{N}(\boldsymbol{\nu}_*, \mathbf{\Gamma}_{\boldsymbol{\nu}}) & &\text{(total error)}
\end{aligned}$$

  - $\boldsymbol{\nu}_* = \boldsymbol{\varepsilon}_* + \boldsymbol{\eta}_*$: the mean; $\mathbf{\Gamma}_{\boldsymbol{\nu}} = \mathbf{\Gamma}_{\boldsymbol{\varepsilon}} + \mathbf{\Gamma}_{\text{noise}}$: the *total error* covariance

Details in:

- Jari Kaipio and Erkki Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2005
- O. Babaniyi, R. Nicholson, Umberto Villa and Noemi Petra. "Inferring the basal sliding coefficient for the Stokes ice sheet model under rheological uncertainty", Cryosphere, 2021.

Premarginalization over secondary uncertain parameters

using the Bayesian Approximation Error (BAE) approach

- **Key idea:** carry out estimation of the primary uncertain parameter β while **taking into account the uncertainty in the auxiliary parameter** a:

$$\boldsymbol{d} = \tilde{\mathcal{F}}(\beta, a) + \boldsymbol{\eta} = (\mathcal{F}(\beta) + \boldsymbol{\varepsilon}) + \boldsymbol{\eta} = \mathcal{F}(\beta) + \boldsymbol{\nu}$$

  - $\tilde{\mathcal{F}}(\beta, a)$: the accurate parameter-to-observable mapping
  - $\mathcal{F}(\beta) = \tilde{\mathcal{F}}(\beta, a_*)$: the *approximative parameter-to-observable mapping*, where the auxiliary parameter a is fixed to the mean value a_* (assume a is Gaussian distributed)

$$\begin{aligned}
\boldsymbol{\eta} &= \boldsymbol{d} - \tilde{\mathcal{F}}(\beta, a) & \boldsymbol{\eta} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\text{noise}}) & &\text{(data noise)}\\
\boldsymbol{\varepsilon} &= \tilde{\mathcal{F}}(\beta, a) - \mathcal{F}(\beta) & \boldsymbol{\varepsilon} &\sim \mathcal{N}(\boldsymbol{\varepsilon}_*, \boldsymbol{\Gamma}_{\boldsymbol{\varepsilon}}) & &\text{(approximation/model error)}\\
\boldsymbol{\nu} &= \boldsymbol{\varepsilon} + \boldsymbol{\eta} & \boldsymbol{\nu} &\sim \mathcal{N}(\boldsymbol{\nu}_*, \boldsymbol{\Gamma}_{\boldsymbol{\nu}}) & &\text{(total error)}
\end{aligned}$$

  - $\boldsymbol{\nu}_* = \boldsymbol{\varepsilon}_* + \boldsymbol{\eta}_*$: the mean; $\boldsymbol{\Gamma}_{\boldsymbol{\nu}} = \boldsymbol{\Gamma}_{\boldsymbol{\varepsilon}} + \boldsymbol{\Gamma}_{\text{noise}}$: the *total error* covariance

Details in:

- Jari Kaipio and Erkki Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2005
- O. Babaniyi, R. Nicholson, Umberto Villa and Noemi Petra. "Inferring the basal sliding coefficient for the Stokes ice sheet model under rheological uncertainty", Cryosphere, 2021.

Additional uncertainty-aware inversion

Approximate posterior covariances

- The reference case (**REF**):

$$\pi_{\text{like}}(\boldsymbol{d}|\beta) \propto \exp\left\{-\frac{1}{2}\left\|\tilde{\mathcal{F}}(\beta, a_{\text{true}}) - \boldsymbol{d}\right\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}}\right\}$$
$$\boldsymbol{\Gamma}_{\text{post}} = (\tilde{\mathbf{F}}^T(\beta_{\text{MAP}}, a_{\text{true}})\boldsymbol{\Gamma}^{-1}_{\text{noise}}\tilde{\mathbf{F}}(\beta_{\text{MAP}}, a_{\text{true}}) + \boldsymbol{\Gamma}^{-1}_{\text{prior}})^{-1}$$

- The conventional error model (**CEM**):

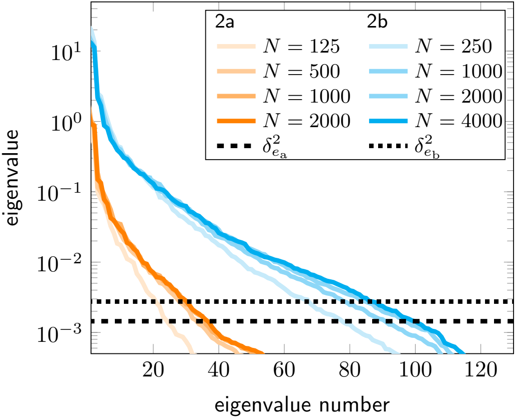
$$\pi_{\text{like}}(\boldsymbol{d}|\beta) \propto \exp\left\{-\frac{1}{2}\left\|\mathcal{F}(\beta) - \boldsymbol{d}\right\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}}\right\}$$
$$\boldsymbol{\Gamma}_{\text{post}} = (\mathbf{F}^T(\beta_{\text{MAP}})\boldsymbol{\Gamma}^{-1}_{\text{noise}}\mathbf{F}(\beta_{\text{MAP}}) + \boldsymbol{\Gamma}^{-1}_{\text{prior}})^{-1}$$

- The BAE approximation error approach (**BAE**):

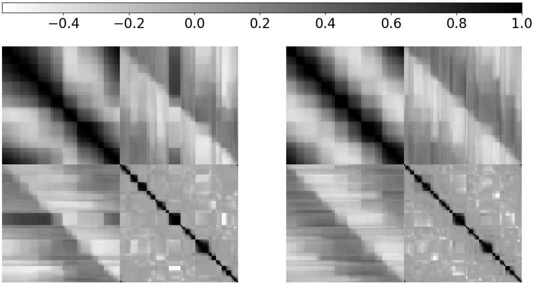
$$\pi_{\text{like}}(\boldsymbol{d}|\beta) \propto \exp\left\{-\frac{1}{2}\left\|\mathcal{F}(\beta) - \boldsymbol{d} + \boldsymbol{\nu}_*\right\|^2_{\boldsymbol{\Gamma}^{-1}_{\boldsymbol{\nu}}}\right\}$$
$$\boldsymbol{\Gamma}_{\text{post}} = (\mathbf{F}^T(\beta_{\text{MAP}})\boldsymbol{\Gamma}^{-1}_{\boldsymbol{\nu}}\mathbf{F}(\beta_{\text{MAP}}) + \boldsymbol{\Gamma}^{-1}_{\text{prior}})^{-1}$$

Low-dimensional structure for the model error

ISMIP-HOM benchmark problem



The spectrum of the approximation errors covariance matrix $\mathbf{\Gamma}_\varepsilon$ for small (orange) and large (blue) approximation errors and for increasing samples sizes.

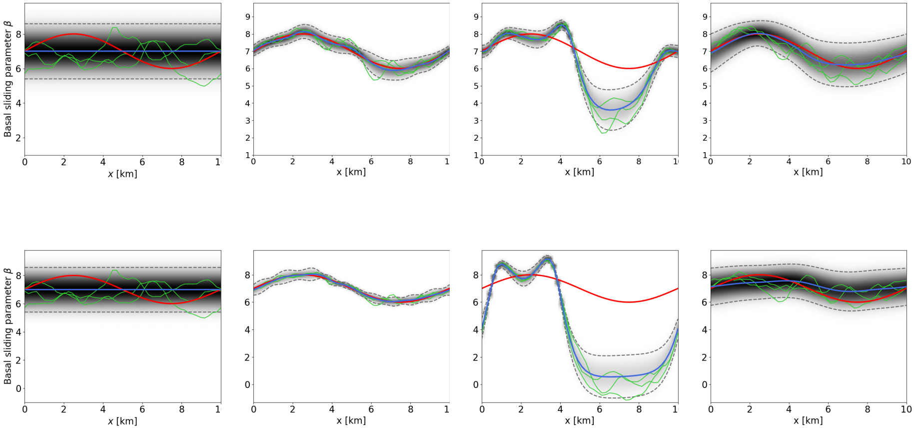


;

The correlation matrix of the approximation error for small (left) and large (right) error cases. The 2×2 block structure is due to velocity measurements in the x- and y-directions.

The “truth” may not be supported by the posterior

ISMIP-HOM benchmark problem



(prior) (REF) (CEM) (BAE)

red: true β, blue: the mean, green: samples, gray shading: $\pm$ 2 standard deviation

hIPPYlib: **I**nverse **P**roblems **PY**thon **lib**rary

- Supported by NSF-SSI2: *Integrating Data with Complex Predictive Models under Uncertainty: An Extensible Software Framework for Large-Scale Bayesian Inversion* (Co-PIs: O. Ghattas (UT Austin), Y. Marzouk (MIT), N. Petra (UC Merced), M. Parno (CRREL) and U. Villa (Washington University in St. Louis))
- `hIPPYlib` contains the implementation of state-of-the-art scalable adjoint-based algorithms (and their extensions) for PDE-based deterministic and Bayesian inverse problems.
- `hIPPYlib` builds on FEniCS (a parallel finite element element library) for the discretization of the PDEs, and on PETSc (**P**ortable, **E**xtensible **T**oolkit for **S**cientific **C**omputation) for scalable and efficient linear algebra operations and solvers.
- hIPPYlib is implemented in a mixture of C++ and Python and has been released under the GNU General Public License version.

- `hIPPYlib` 3.0. can be downloaded from:

  http://hippylib.github.io

Details in: U. Villa, N. Petra, and O. Ghattas. "An Extensible Software Framework for Large-Scale Deterministic and Linearized Bayesian Inverse Problems". ACM Transactions on Mathematical Software, Vol. 47, No. 2, 2021.

Summary and Conclusions

- We considered the problem of inferring the basal sliding coefficient field for **uncertain** Stokes ice sheet forward model from (synthetic) surface velocity measurements under additional uncertainty.
- To account for the associated model uncertainties (error) we employed the **Bayesian Approximation Error (BAE)** approach to approximately premarginalize over both the noise in measurements and uncertainty in the forward model.
- Our findings suggest that accounting for secondary uncertainties in the inference is crucial.
- The results also suggest that hierarchical off-diagonal low-rank approximation and H-matrix compression combined with model reduction techniques may be a path forward towards solving large-scale Bayesian (ice sheet) inverse problems.