Large-scale Bayesian Inversion for Geosciences Problems

Omar Ghattas
Oden Institute for Computational Engineering and Sciences
Departments of Geological Sciences and Mechanical Engineering
The University of Texas at Austin

Joint work with:
Tobin Isaac (Georgia Tech), Noémi Petra (UC-Merced), Georg Stadler (NYU)



SIAM Conference on Mathematical and Computational Issues in the Geosciences

Former PhD students who worked on geophysical inversion

Volkan Akcelik
ExxonMobil

Aysegul Askan
Middle East Tech U

Jifeng Xu
COMAC

George Biros
UT Austin

Hesheng Bao
Mgmt Sci Assoc

Ioannis Epanomeritakis
Univ Crete

Pearl Flath
Molex

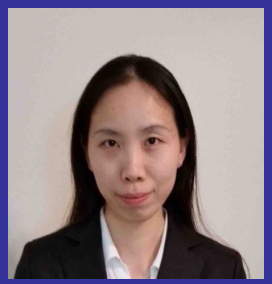
Judy Hill
ORNL

Nick Alger
UT Austin

James Martin

Ben Crestel
Element AI

Jenn Worthen
Emerson



Hongyu Zhu
UTRC

Johann Rudi
Argonne

Tiankai Tu
DE Shaw

Jamie Bramwell
LLNL

Tobin Isaac
Georgia Tech

Amal Alghamdi
UT Austin

Postdocs/research scientists in geophysical inversion

Carsten Burstedde
Univ Bonn

Lucas Wilcox
Naval Postgrad School

Georg Stadler
NYU

Tan Bui
UT Austin

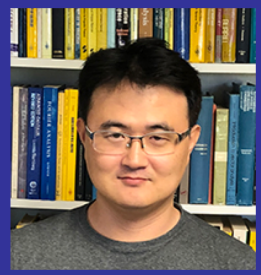
Noemi Petra
UC Merced

Alen Alexanderian
NC State

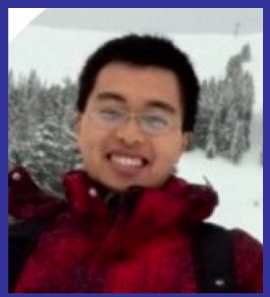
Hejun Zhu
UT Dallas

Loukas Kallivokas
UT Austin

Umberto Villa
Washington Univ



Jeonghun Lee
Baylor Univ



Peng Chen
UT Austin

Ilona Ambartsumyan
UT Austin

Eldar Khattatov
UT Austin

Hari Sundar
Univ Utah

Hossein Aghakhani
CGG US Imaging

The central role of inverse problems in the geosciences

- High fidelity predictive models of geophysical phenomena are often characterized by uncertainties in initial/boundary conditions, constitutive parameters, source terms, subgrid/closure models, etc.
- Rapidly growing abilities to instrument, sense, and remotely observe the Earth have give rise to expanding volumes of observational data that can be used to reduce uncertainties
- Sustained growth in high performance computing, networking, storage
- Advances in mathematical and computational inverse theory and algorithms
- SIAM Geosciences community has long been a fountain of inverse problems research

First SIAM Geosciences conference: Sept 1989 in Houston

Society for Industrial and Applied Mathematics

Preliminary Program

SIAM Conference on Mathematical and Computational Issues in Geophysical Fluid and Solid Mechanics

September 25 – 28, 1989

Stouffer Greenway Plaza Hotel
Houston, Texas

CONFERENCE THEMES

Reactive Flows in Porous Media
Adaptive Mesh Methods
Conservation Laws in Scientific Computing
Contaminant Transport Modeling
Fluid Flows Among Solid Obstacles
Domain Decomposition Methods
Mechanics of Geological Materials
Models of Fault Constitutive Properties

AND

SIAM Workshop on Geophysical Inversion

September 27 – 29, 1989

WORKSHOP THEMES

Inverse Scattering
Tomography
Wave Equation Migration
Velocity Estimation
Elastic Inversion
Multiparameter Inversion
Complex Wave Phenomena
Inverse Modeling

See page 10 for Workshop Program

CONFERENCE ORGANIZING COMMITTEE

William E. Fitzgibbon, Co-chair
University of Houston
Mary F. Wheeler, Co-chair
University of Houston
Richard E. Ewing
University of Wyoming
Roland Glowinski
University of Houston
Richard P. Kendall
J.S. Nolen & Associates, Houston
John MacBain
British Petroleum Exploration, Houston
George R. Sell
University of Minnesota
William W. Symes
Rice University
Alan Weiser
Exxon Production Research Corporation

WORKSHOP ORGANIZING COMMITTEE

J. Bee Bednar, Chair
Amerada Hess Corporation, Tulsa
Lawrence R. Lines
Amoco Production Company, Tulsa
R.H. Stolt
Conoco, Inc., Ponca City
A. B. Weglein
ARCO Exploration Company, Plano

Inverse problems within SIAM Geosciences community

SIAG/GS Career Prize winners active in inverse problems

- Mary Wheeler (2009)
- William Symes (2011)
- Clint Dawson (2013)
- Jérôme Jaffré (2015)
- Juan Restrepo (2017)

Inverse Problems at GS19

- IP3: Anatoly Baumstein
- IP5: Yunyue "Elita" Li
- IP6: Patrick Heimbach
- MS3/MS18: Extended-model Versus Reduced-model Strategies for Full Waveform Inversion
- MS4: Use of Adjoint Models in the Geosciences
- MS5/MS17: Scalable Methods for Coupling Water Resources Modeling & Parameter Estimation
- MS8: Scientific Machine Learning for Subsurface Geoscience
- MS13/26: Optimal Transport for Imaging in Geosciences
- MS25/MS35: Practical Aspects of Large-scale Sparsity-promoting Seismic Inversion
- MS30: Uncertainty Quantification in Subsurface Flow and Transport
- MS41: Advances in Bayesian Estimation Strategies in Subsurface Processes
- MS45: Uncertainty Quantification for Geophysical Inverse Problems

Outline

1. The inverse problem: Integrating data and models
2. Examples of Bayesian inverse problems
3. Target: Flow of the Antarctic ice sheet
4. Large-scale Bayesian inverse problems
5. Application to Antarctic ice sheet flow inverse problem

Anatomy of an inverse problem

Input parameters, computational model, and output observables

The forward problem

- Given model parameters m, solve forward model $\mathcal{F}$ to yield output observables d

$$\mathcal{F}(m) \longrightarrow d$$

- Well-posed: solution exists, is unique, and is stable to perturbations in inputs
- Causal: later-time solutions depend only on earlier time solutions
- Local: the forward operator includes derivatives that couple nearby solutions in space and time

Anatomy of an inverse problem

Input parameters, computational model, and output observables

The inverse problem

- Given output observations d_{obs} and forward model $\mathcal{F}$, infer model parameters m

$$m \longleftarrow \mathcal{F}^{-1} d_{\text{obs}}$$

- Ill-posed: observations are usually sparse and forward operator often smoothing; many different parameter values may be consistent with the data
- Non-causal: the inverse operator couples earlier time solutions with later time ones
- Global: the inverse operator couples solution values across all of space and time

Anatomy of an inverse problem

Input parameters, computational model, and output observables

Tikhonov's approach to ill-posedness

Employ regularization to penalize unwanted solution features, guarantee unique solution:

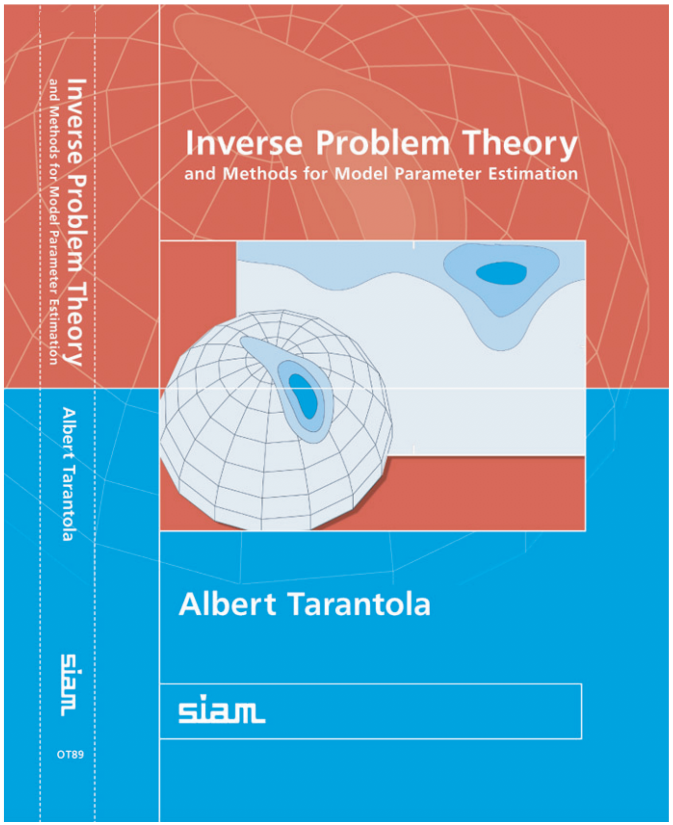
$$\min_m \ \frac{1}{2} \| \mathcal{F}(m) - d_{\text{obs}} \|_W^2 + \frac{\alpha}{2} \| m - m_{\text{ref}} \|_R^2$$

Bayesian approach to ill-posedness

Describe probability of all models that are consistent with the data and any prior knowledge of the parameters:

$$\pi(m|d_{\text{obs}}) \propto \exp\left(-\frac{1}{2} \| \mathcal{F}(m) - d_{\text{obs}} \|_{C_{\text{d}}^{-1}}^2 - \frac{1}{2} \| m - m_{\text{pr}} \|_{C_{\text{m}}^{-1}}^2 \right)$$

Tarantola's Bayesian inversion vision



Twin challenges of Bayesian inversion

- Geophysical forward problems often highly complex (nonlinear, wide range of spatiotemporal scales, complex geometry, heterogeneous/anisotropic, coupled, ...)
- Model parameters often infinite-dimensional fields (ICs, BCs, sources, material properties, geometry

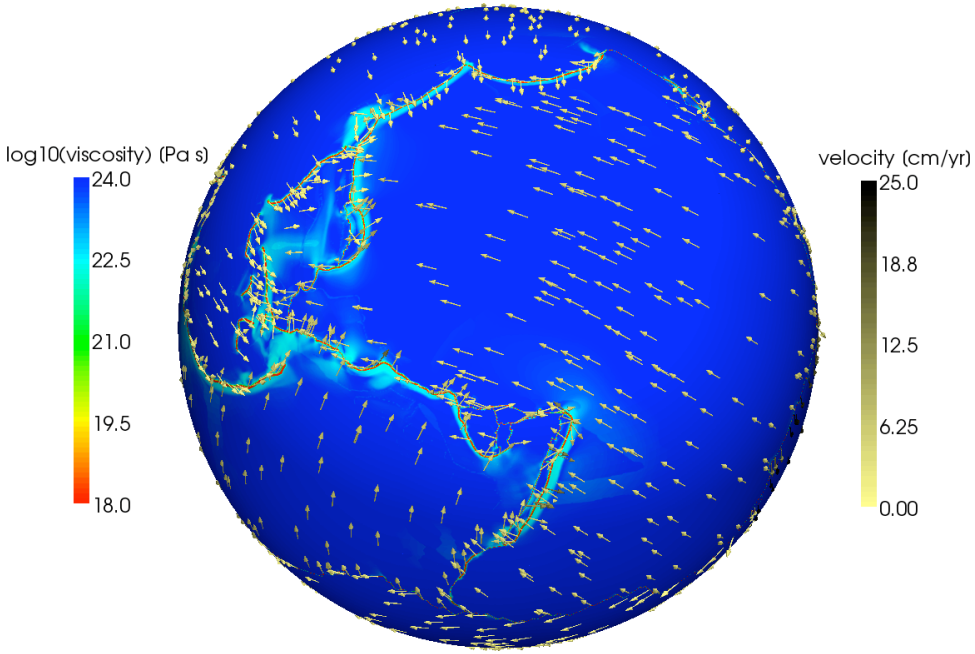
Outline

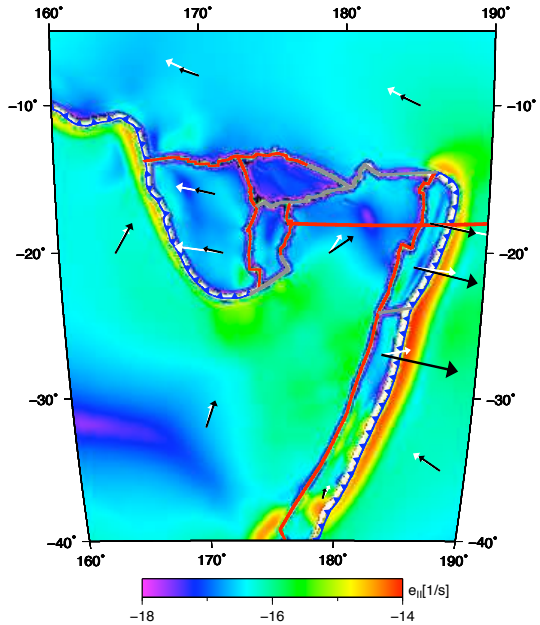
1. The inverse problem: Integrating data and models
2. Examples of Bayesian inverse problems
3. Target: Flow of the Antarctic ice sheet
4. Large-scale Bayesian inverse problems
5. Application to Antarctic ice sheet flow inverse problem

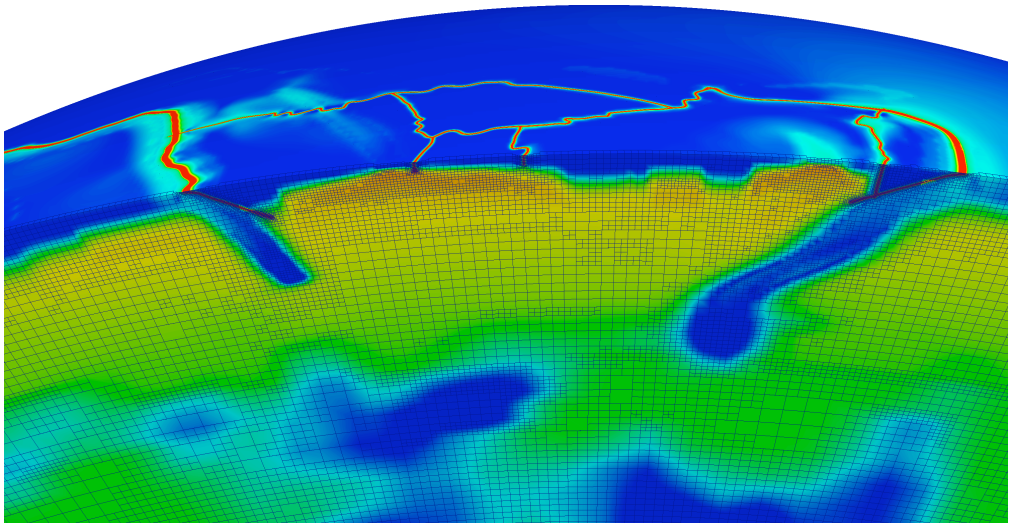
Some Bayesian inverse problems of interest

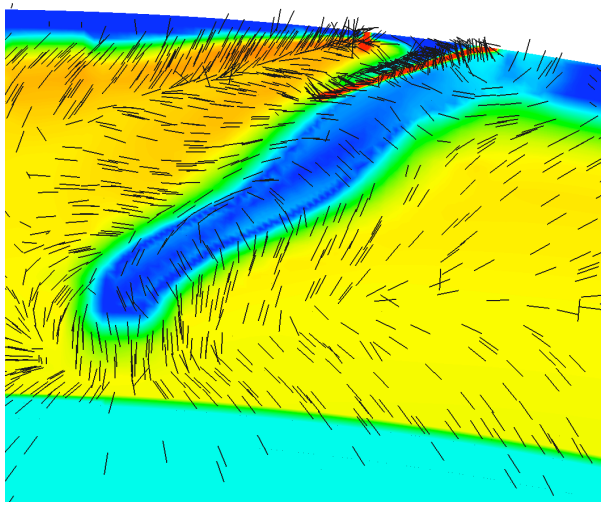
- Antarctic ice sheet flow (+ ocean dynamics)
  - Joint with Patrick Heimbach, Tom Hughes, Tobin Isaac (Georgia Tech), Tom O'Leary-Roseberry, Noemi Petra (UC-Merced), Georg Stadler (NYU), Umberto Villa (WashU), Alice Zhu (UTRC)
- Global and regional seismic inversion, joint seismic–EM inversion, inverse scattering
  - Joint with Hossein Aghakhani (CCG), Nick Alger, Tan Bui, Ben Crestel (Element AI), David Keyes (KAUST), George Turkiyyah (AUB), Georg Stadler (NYU), Umberto Villa (WashU)
- Global mantle convection
  - Joint w/ Mike Gurnis (Caltech), Johann Rudi (ANL), Georg Stadler (NYU)
- Poroelastic subsurface flow inversion and management of induced seismicity
  - Joint with Amal Alghamdi, Ann Chen, Marc Hesse, Georg Stadler (NYU), Umberto Villa (WashU), Karen Willcox
- Reservoir inversion
  - Joint with George Biros, Tan Bui, Clint Dawson, Sam Estes, John Lee (Baylor), Umberto Villa (WashU)

Inversion for mantle properties in global mantle convection

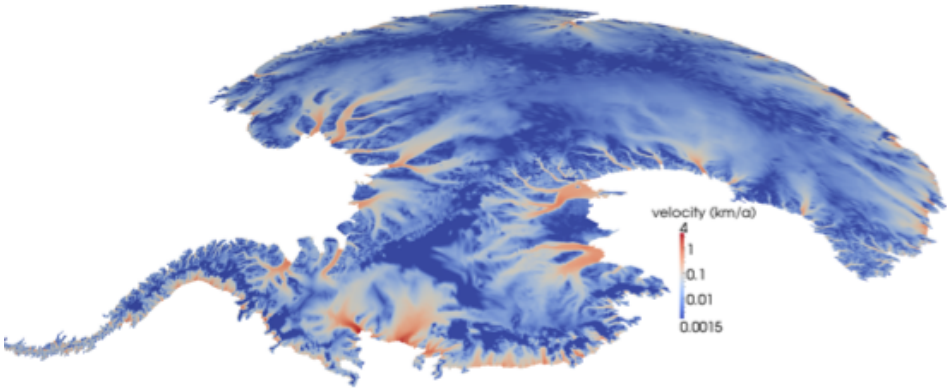




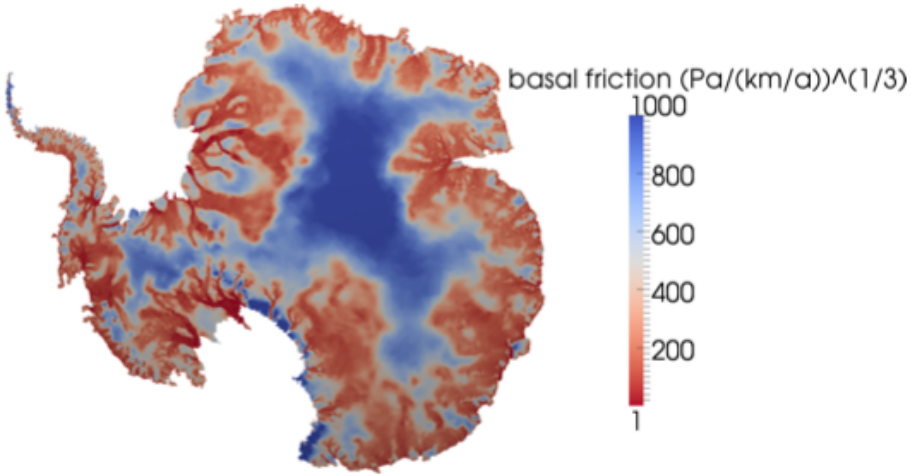




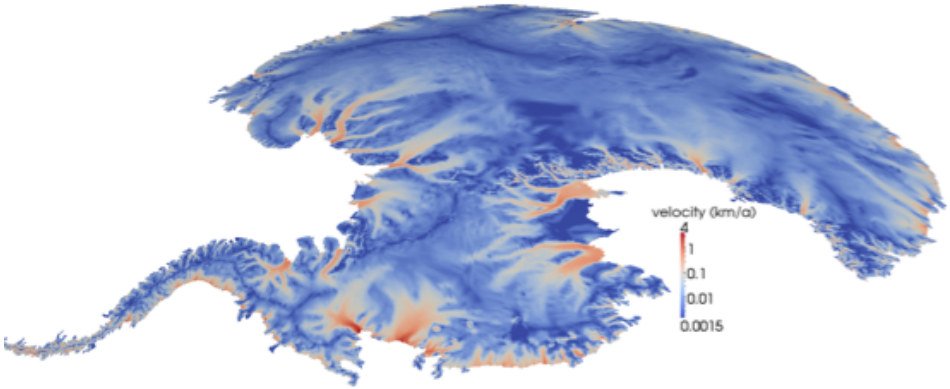
Bayesian inversion for basal friction field in Antarctica



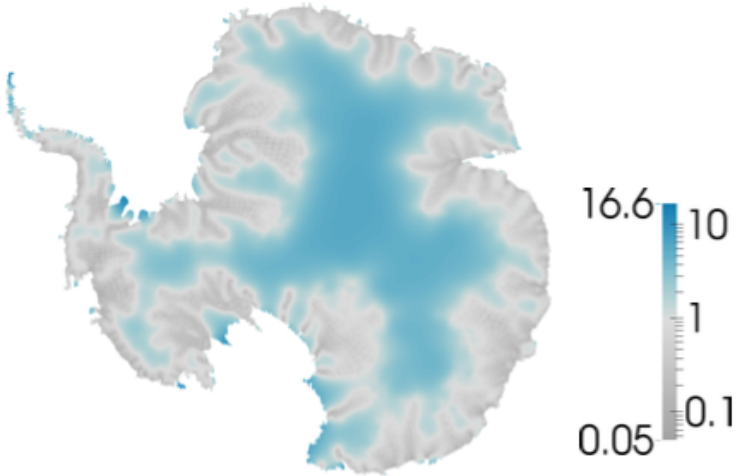
InSAR-based ice surface velocity observations



Inferred mean of basal friction field

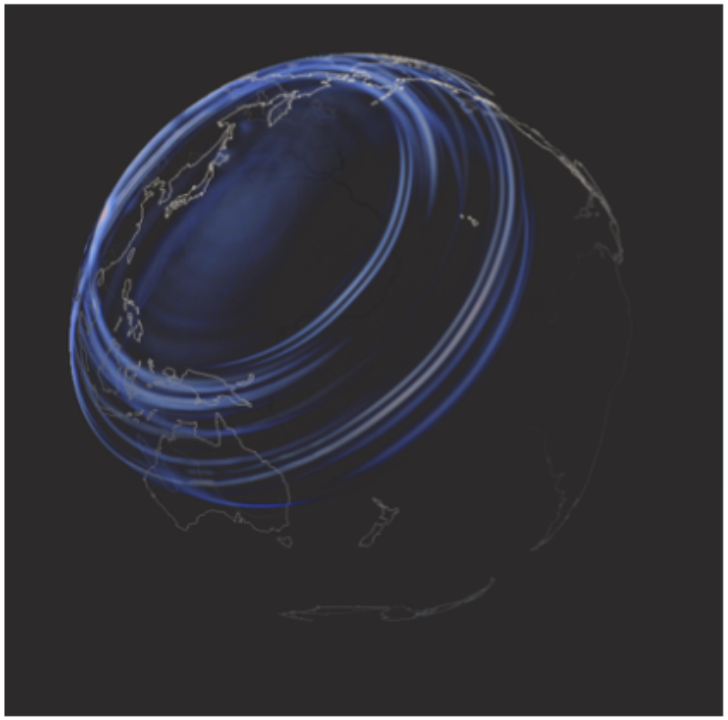


Reconstructed ice surface velocity field (based on inferred mean of basal friction field)

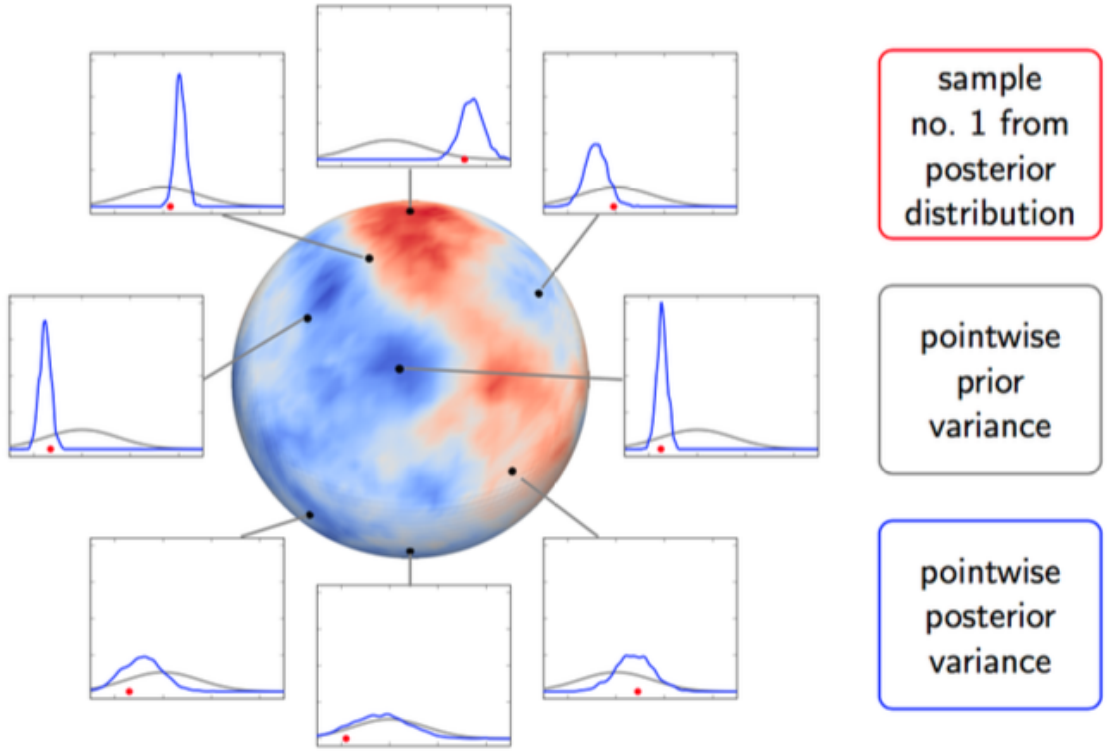


Inferred uncertainty in basal friction field (standard deviation of Gaussianized posterior of log basal friction)

Bayesian global seismic inversion

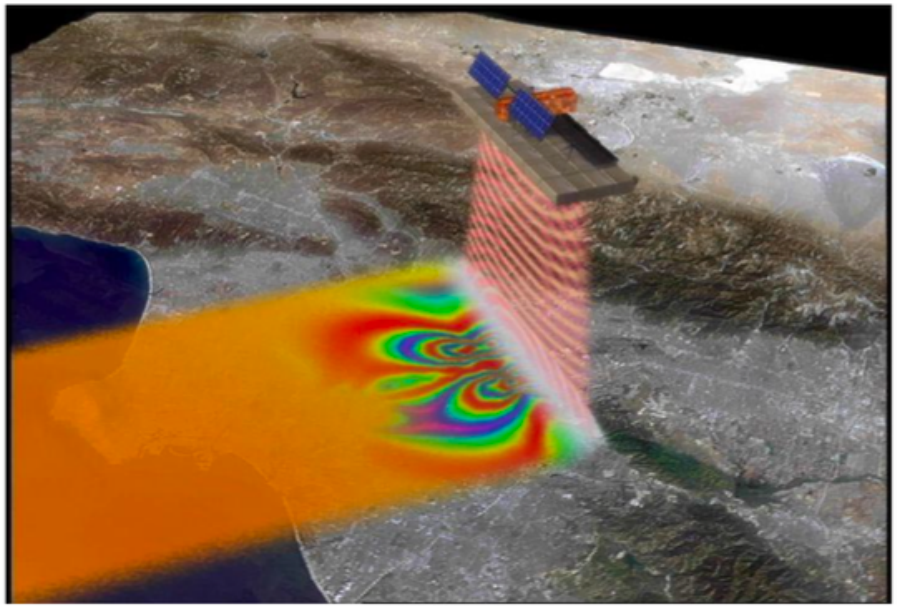


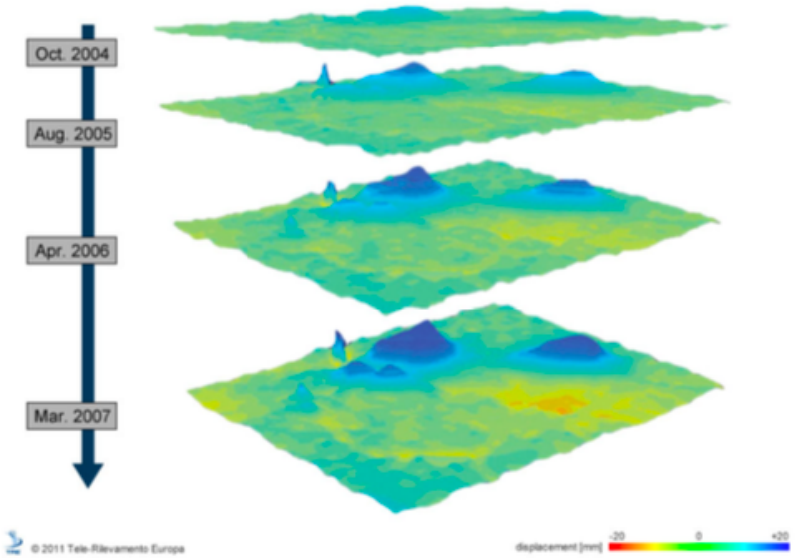
Parallel adaptive DG wave propagation



Prior and posterior seismic velocity marginals

Bayesian poroelastic inversion



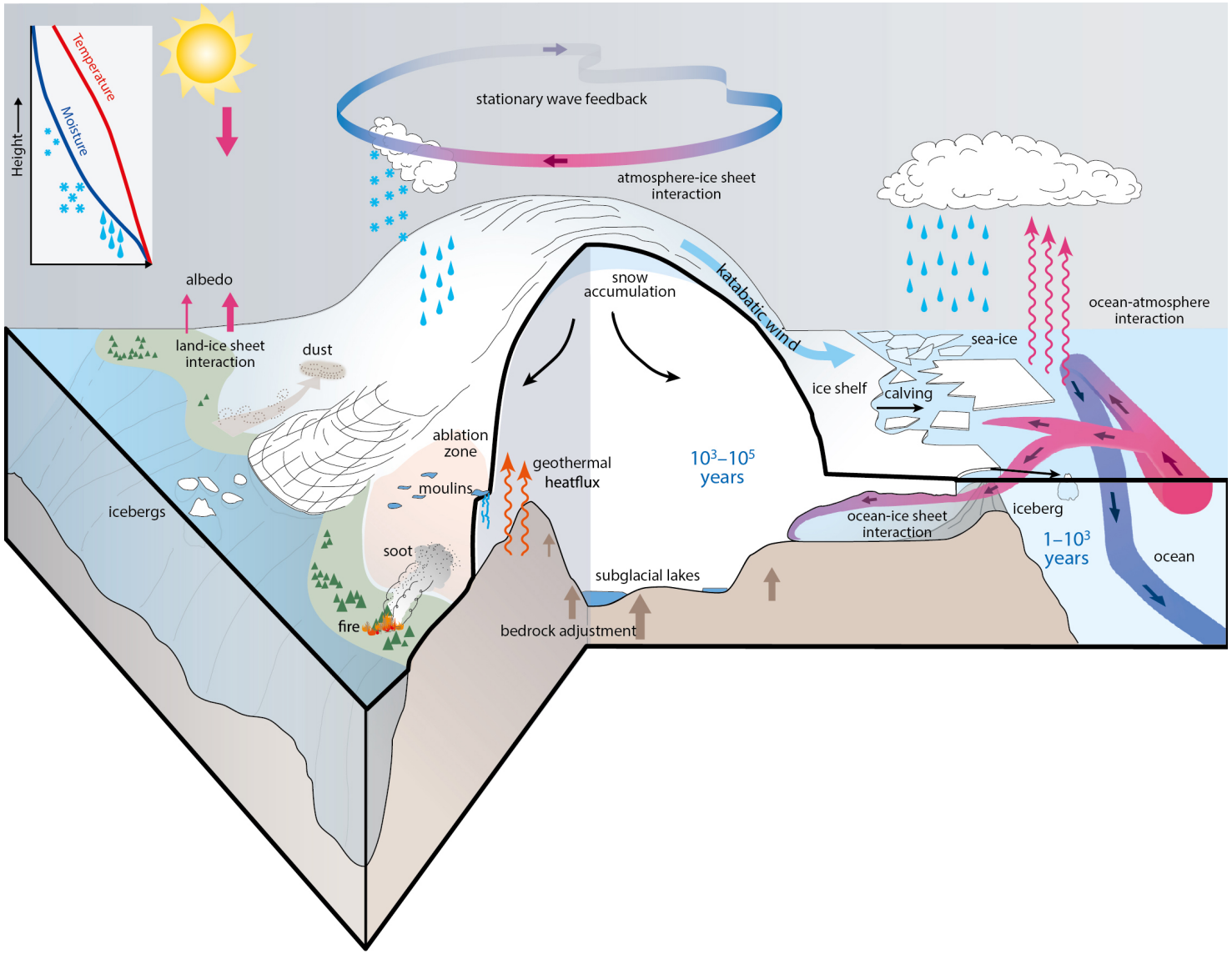


Use observations (InSAR, GPS) of surface deformation induced by CO2 or wastewater injection, in addition to well pressure measurements, to infer subsurface permeability and elastic properties. Forward predict and then ultimately optimize injection processes to avoid induced seismicity.

Outline

1. The inverse problem: Integrating data and models
2. Examples of Bayesian inverse problems
3. Target: Flow of the Antarctic ice sheet
4. Large-scale Bayesian inverse problems
5. Application to Antarctic ice sheet flow inverse problem

Motivation: Flow of the Antarctic ice sheet



Interactions between ice sheet, solid earth, ocean, and atmosphere

Motivation: Flow of the Antarctic ice sheet

- Ice flows from interior of polar ice sheets to ocean is primary contributor to sea level rise (200 billion tons/yr currently)
- Flow rates of outlet glaciers in Antarctica have been increasing over past several decades
- Thinning of ice shelves due to increased mixing in ocean driven by intensification of polar winds, bringing warmer water to surface;
- Ice shelf thinning leads to loss of buttressing effect and retreat of ice sheet
- 0.5 m sea level rise by 2070 estimated to jeopardize 136 largest port cities, with 150M inhabitants and $35 trillion in assets
- We need to develop predictive models, with quantified uncertainties, to better anticipate future sea level rise

Motivation: Flow of the Antarctic ice sheet II

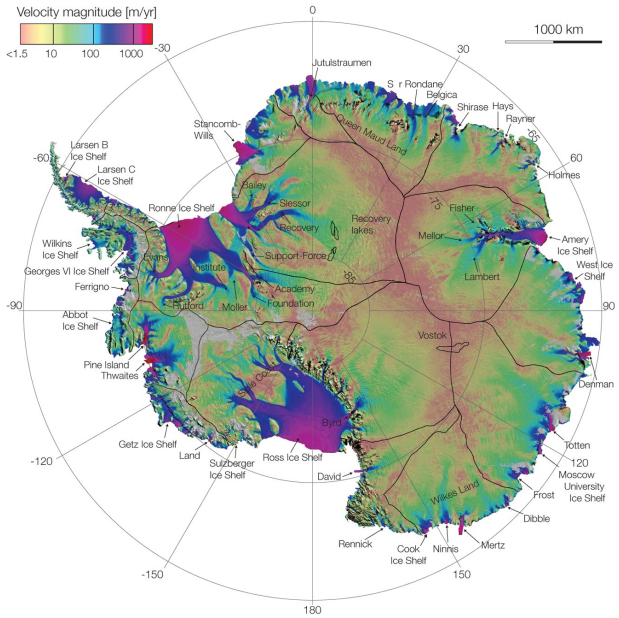
- Recent evidence suggests that sea level rose abruptly at the end of the last interglacial ($\sim$118 kyr ago) by $\sim$5-6m; the likely cause is catastrophic collapse of polar ice sheets
  - *Ice sheet collapse following a prolonged period of stable sea level during the last interglacial*, MJ O'Leary, PJ Hearty, WG Thompson, ME Raymo, JX Mitrovica, JM Webster, **Nature Geoscience**, 6, 796800, 2013.
- Recent work indicates that retreat of the Amundsen Sea Embayment (a portion of the West Antarctic ice sheet) is accelerating with no major bed obstacles to prevent draw down of the entire basin.
  - *Widespread, rapid grounding line retreat of Pine Island, Thwaites, Smith and Kohler glaciers, West Antarctica from 1992 to 2011*, E. Rignot, J. Mouginot, M. Morlighem, H. Seroussi, and B. Scheuchl, **Geophysical Research Letters**, 41(10):3502–3509, 2014.

Dynamics of the Antarctic ice sheet and sea level rise

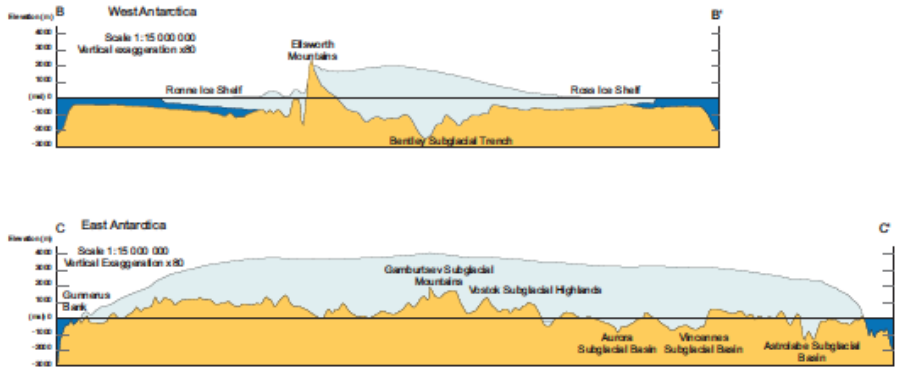
Glaciers flow thousands of miles from the continent's deep interior to its coast

Credit: NASA Goddard Space Flight Center/JPL-Caltech

Challenges in modeling ice sheet dynamics



Observed surface velocity (Rignot et al., 2011)



Fretwell et al., 2013

- strong nonlinearities and complex rheology
- highly ill-conditioned due to widely-varying coefficients
- complex, high aspect ratio geometry
- wide range of length scales: from $O(\text{m})$ to $O(10^3\ \text{km})$
- uncertainties in: basal boundary conditions, basal topography, rheology, geothermal heat flux
- diverse observational data (InSAR, laser altimetry, GRACE satellite, ice cores, ground-penetrating radar)

The Forward Ice Sheet Model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla}\cdot\overbrace{[2\eta(\boldsymbol{u},\theta)\,\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}-\boldsymbol{I}p]}^{\boldsymbol{\sigma}_{\boldsymbol{u}}}=\rho\boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}=\tfrac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u}+\boldsymbol{\nabla}\boldsymbol{u}^T)]$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}=0$$

$$\rho c\boldsymbol{u}\cdot\boldsymbol{\nabla}\theta-\boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}\theta)=\eta(\boldsymbol{u},\theta)\,\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)$$

Constitutive relation

$$\eta(\boldsymbol{u},\theta)=\frac{1}{2}\left\{A_0\exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}}\left(\tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)\right)^{\frac{1-n}{2n}}$$

Boundary conditions

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}=\boldsymbol{0} \qquad \theta=\theta_s \quad \text{on } \Gamma_t$$

$$\boldsymbol{u}\cdot\boldsymbol{n}=\frac{M}{\rho L_i} \qquad \boldsymbol{T}:=\boldsymbol{I}-\boldsymbol{n}\otimes\boldsymbol{n},\ \boldsymbol{T}\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}+\beta\boldsymbol{T}\boldsymbol{u}=\boldsymbol{0} \quad \text{on } \Gamma_b$$

$$K\boldsymbol{\nabla}\theta\cdot\boldsymbol{n}=G-M+\beta\boldsymbol{T}\boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}, \qquad \theta\le\theta_m,\ M\ge 0,\ M(\theta-\theta_m)=0 \quad \text{on } \Gamma_b$$

The Forward Ice Sheet Model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla}\cdot\overbrace{[2\eta(\boldsymbol{u},\theta)\,\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}-\boldsymbol{I}p]}^{\boldsymbol{\sigma}_{\boldsymbol{u}}}=\rho\boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}=\tfrac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u}+\boldsymbol{\nabla}\boldsymbol{u}^T)]$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}=0$$

$$\rho c\boldsymbol{u}\cdot\boldsymbol{\nabla}\theta-\boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}\theta)=\eta(\boldsymbol{u},\theta)\,\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)$$

Constitutive relation

rheology parameter $n=3$

$$\eta(\boldsymbol{u},\theta)=\frac{1}{2}\left\{A_0\exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}}\left(\tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)\right)^{\frac{1-n}{2n}}$$

Boundary conditions

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}=\boldsymbol{0} \qquad \theta=\theta_s \quad \text{on } \Gamma_t$$

$$\boldsymbol{u}\cdot\boldsymbol{n}=\frac{M}{\rho L_i} \qquad \boldsymbol{T}:=\boldsymbol{I}-\boldsymbol{n}\otimes\boldsymbol{n},\ \boldsymbol{T}\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}+\beta\boldsymbol{T}\boldsymbol{u}=\boldsymbol{0} \quad \text{on } \Gamma_b$$

$$K\boldsymbol{\nabla}\theta\cdot\boldsymbol{n}=G-M+\beta\boldsymbol{T}\boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}, \qquad \theta\le\theta_m,\ M\ge 0,\ M(\theta-\theta_m)=0 \quad \text{on } \Gamma_b$$

The Forward Ice Sheet Model

Balance of linear momentum, mass, and energy

$$-\nabla \cdot \overbrace{[2\eta(\boldsymbol{u},\theta)\,\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} - \boldsymbol{I}p]}^{\boldsymbol{\sigma}_{\boldsymbol{u}}} = \rho \boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} = \tfrac{1}{2}(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T)]$$

$$\nabla \cdot \boldsymbol{u} = 0$$

$$\rho c \boldsymbol{u} \cdot \nabla \theta - \nabla \cdot (K \nabla \theta) = \eta(\boldsymbol{u},\theta)\ \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)$$

Constitutive relation — Arrhenius-type thinning

$$\eta(\boldsymbol{u},\theta) = \frac{1}{2}\left\{A_0 \exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}} \left(\tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)\right)^{\frac{1-n}{2n}}$$

Boundary conditions

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n} = \boldsymbol{0} \qquad \theta = \theta_s \quad \text{on } \Gamma_t$$

$$\boldsymbol{u}\cdot\boldsymbol{n} = \frac{M}{\rho L_i} \qquad \boldsymbol{T} := \boldsymbol{I} - \boldsymbol{n}\otimes\boldsymbol{n},\ \boldsymbol{T}\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n} + \beta \boldsymbol{T}\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \Gamma_b$$

$$K\nabla\theta\cdot\boldsymbol{n} = G - M + \beta \boldsymbol{T}\boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}, \qquad \theta \le \theta_m,\ M \ge 0,\ M(\theta-\theta_m) = 0 \quad \text{on } \Gamma_b$$

The Forward Ice Sheet Model

Balance of linear momentum, mass, and energy

$$-\nabla \cdot \overbrace{[2\eta(\boldsymbol{u},\theta)\,\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} - \boldsymbol{I}p]}^{\boldsymbol{\sigma}_{\boldsymbol{u}}} = \rho \boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} = \tfrac{1}{2}(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T)]$$

$$\nabla \cdot \boldsymbol{u} = 0$$

$$\rho c \boldsymbol{u} \cdot \nabla \theta - \nabla \cdot (K \nabla \theta) = \eta(\boldsymbol{u},\theta)\ \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)$$

Constitutive relation — shear thinning with the second strain rate invariant

$$\eta(\boldsymbol{u},\theta) = \frac{1}{2}\left\{ A_0 \exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}} \left(\tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)\right)^{\frac{1-n}{2n}}$$

Boundary conditions

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n} = \boldsymbol{0} \qquad \theta = \theta_s \quad \text{on } \Gamma_t$$

$$\boldsymbol{u}\cdot\boldsymbol{n} = \frac{M}{\rho L_i} \qquad \boldsymbol{T} := \boldsymbol{I} - \boldsymbol{n}\otimes\boldsymbol{n},\ \boldsymbol{T}\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n} + \beta \boldsymbol{T}\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \Gamma_b$$

$$K\nabla\theta\cdot\boldsymbol{n} = G - M + \beta \boldsymbol{T}\boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}, \qquad \theta \le \theta_m,\ M \ge 0,\ M(\theta-\theta_m) = 0 \quad \text{on } \Gamma_b$$

The Forward Ice Sheet Model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla}\cdot\overbrace{[2\eta(\boldsymbol{u},\theta)\,\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}-\boldsymbol{I}p]}^{\boldsymbol{\sigma}_{\boldsymbol{u}}}=\rho\boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}=\tfrac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u}+\boldsymbol{\nabla}\boldsymbol{u}^T)]$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}=0$$

$$\rho c\boldsymbol{u}\cdot\boldsymbol{\nabla}\theta-\boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}\theta)=\eta(\boldsymbol{u},\theta)\,\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)$$

Constitutive relation

$$\eta(\boldsymbol{u},\theta)=\frac{1}{2}\left\{A_0\exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}}\left(\tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)\right)^{\frac{1-n}{2n}}$$

Boundary conditions

unknown parameter fields

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}=\boldsymbol{0} \qquad \theta=\theta_s \quad \text{on } \Gamma_t$$

$$\boldsymbol{u}\cdot\boldsymbol{n}=\frac{M}{\rho L_i} \qquad \boldsymbol{T}:=\boldsymbol{I}-\boldsymbol{n}\otimes\boldsymbol{n},\ \boldsymbol{T}\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}+\beta\boldsymbol{T}\boldsymbol{u}=\boldsymbol{0} \quad \text{on } \Gamma_b$$

$$K\boldsymbol{\nabla}\theta\cdot\boldsymbol{n}=G-M+\beta\boldsymbol{T}\boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}, \qquad \theta\le\theta_m,\ M\ge 0,\ M(\theta-\theta_m)=0 \quad \text{on } \Gamma_b$$

The Forward Ice Sheet Model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla}\cdot\overbrace{[2\eta(\boldsymbol{u},\theta)\,\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}-\boldsymbol{I}p]}^{\boldsymbol{\sigma}_{\boldsymbol{u}}}=\rho\boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}=\tfrac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u}+\boldsymbol{\nabla}\boldsymbol{u}^T)]$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}=0$$

$$\rho c\boldsymbol{u}\cdot\boldsymbol{\nabla}\theta-\boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}\theta)=\eta(\boldsymbol{u},\theta)\,\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)$$

Constitutive relation

$$\eta(\boldsymbol{u},\theta)=\frac{1}{2}\left\{A_0\exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}}\left(\tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}^2)\right)^{\frac{1-n}{2n}}$$

Boundary conditions

complementarity conditions

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}=\boldsymbol{0} \qquad \theta=\theta_s \quad \text{on } \Gamma_t$$

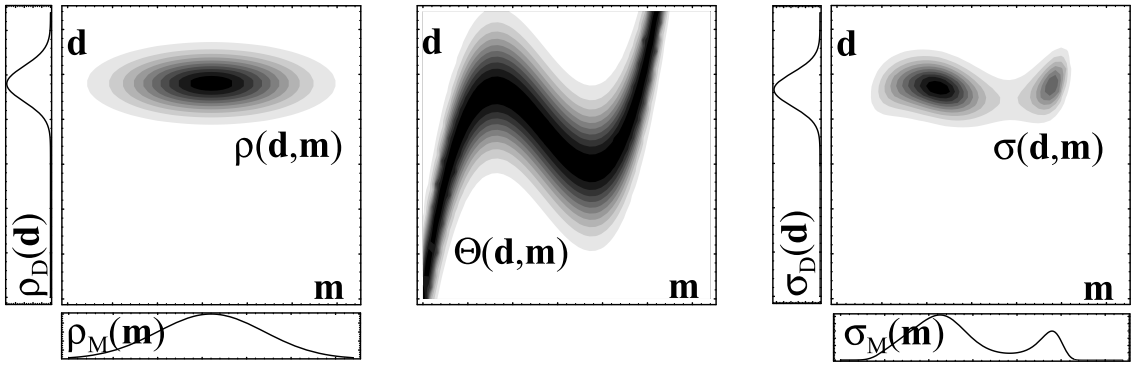
$$\boldsymbol{u}\cdot\boldsymbol{n}=\frac{M}{\rho L_i} \qquad \boldsymbol{T}:=\boldsymbol{I}-\boldsymbol{n}\otimes\boldsymbol{n},\ \boldsymbol{T}\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n}+\beta\boldsymbol{T}\boldsymbol{u}=\boldsymbol{0} \quad \text{on } \Gamma_b$$

$$K\boldsymbol{\nabla}\theta\cdot\boldsymbol{n}=G-M+\beta\boldsymbol{T}\boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}, \qquad \theta\le\theta_m,\ M\ge 0,\ M(\theta-\theta_m)=0 \quad \text{on } \Gamma_b$$

Outline

1. The inverse problem: Integrating data and models
2. Examples of Bayesian inverse problems
3. Target: Flow of the Antarctic ice sheet
4. Large-scale Bayesian inverse problems
5. Application to Antarctic ice sheet flow inverse problem

Bayesian inversion: Finite dim. w/Gaussian noise & prior



A. Tarantola, *Inverse Problem Theory*, SIAM, 2005

Given: prior pdf of model parameters $\boldsymbol{m}$: $\pi_{\text{pr}}(\boldsymbol{m}) = \exp(-\frac{1}{2} \parallel \boldsymbol{m} - \boldsymbol{m}_{\text{pr}} \parallel^2_{\boldsymbol{\Gamma}_{\text{pr}}^{-1}})$

prior pdf of the observables $\boldsymbol{d}$: $\pi_{\text{obs}}(\boldsymbol{d}) = \exp(-\frac{1}{2} \parallel \boldsymbol{d} - \boldsymbol{d}_{\text{obs}} \parallel^2_{\boldsymbol{\Gamma}_{\text{d}}^{-1}})$

conditional pdf relating $\boldsymbol{d}$ and $\boldsymbol{m}$: $\pi_{\text{model}}(\boldsymbol{d}|\boldsymbol{m}) = \exp(-\frac{1}{2} \parallel \boldsymbol{f}(\boldsymbol{m}) - \boldsymbol{d} \parallel^2_{\boldsymbol{\Gamma}_{\text{m}}^{-1}})$

Then the posterior pdf of model parameters (solution of Bayesian inverse problem) is given by:

$$\pi_{\text{post}}(\boldsymbol{m}) \stackrel{\text{def}}{=} \pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d}_{\text{obs}}) \propto \pi_{\text{pr}}(\boldsymbol{m}) \int_{\mathcal{D}} \pi_{\text{obs}}(\boldsymbol{d})\, \pi_{\text{model}}(\boldsymbol{d}|\boldsymbol{m})\, d\boldsymbol{d}$$

$$\propto \pi_{\text{pr}}(\boldsymbol{m})\, \pi_{\text{like}}(\boldsymbol{d}_{\text{obs}}|\boldsymbol{m})$$

Posterior pdf for Gaussian additive noise (with $\boldsymbol{\Gamma}_{\text{noise}} = \boldsymbol{\Gamma}_{\text{d}} + \boldsymbol{\Gamma}_{\text{m}}$) and Gaussian prior

$$\pi_{\text{post}}(\boldsymbol{m}) \propto \exp\left(-\frac{1}{2} \parallel \boldsymbol{f}(\boldsymbol{m}) - \boldsymbol{d}_{\text{obs}} \parallel^2_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}} -\frac{1}{2} \parallel \boldsymbol{m} - \boldsymbol{m}_{\text{pr}} \parallel^2_{\boldsymbol{\Gamma}_{\text{pr}}^{-1}} \right)$$

MAP estimate and linearized Bayesian formulation

- Maximum a posteriori (MAP) estimate given by maximizing posterior pdf, or equivalently minimizing $-\log \pi_{\text{post}}(\boldsymbol{m})$

$$\boldsymbol{m}_{\text{map}} := \arg\min_{\boldsymbol{m}\in\mathbb{R}^n} \ \tfrac{1}{2}\|\boldsymbol{f}(\boldsymbol{m}) - \boldsymbol{d}_{\text{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \tfrac{1}{2}\|\boldsymbol{m} - \boldsymbol{m}_{\text{pr}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{pr}}}$$

- Linearize the parameter-to-observable map about $\boldsymbol{m}_{\text{map}}$:

$$\boldsymbol{f}(\boldsymbol{m}) \approx \boldsymbol{f}(\boldsymbol{m}_{\text{map}}) + \boldsymbol{F}(\boldsymbol{m} - \boldsymbol{m}_{\text{map}})$$

- $\boldsymbol{F}$ is Fréchet derivative of $\boldsymbol{f}$ at $\boldsymbol{m}_{\text{map}}$; its action involves solution of linearized forward problem
- Posterior is Gaussian pdf $\mathcal{N}(\boldsymbol{m}_{\text{map}}, \boldsymbol{\Gamma}_{\text{post}})$ with

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

- $\boldsymbol{\Gamma}_{\text{post}}$ is the inverse of the Hessian $\boldsymbol{H}$ of the negative log posterior
- Cannot form Hessian; each column requires a (linearized) forward solve
- But action of Hessian requires a pair of linearized forward and adjoint solves

Employ infinite dimensional framework of A. Stuart (*Acta Numerica*, 2010)

- Gaussian random field prior, $\mu_0 := \mathcal{N}(m_0, \mathcal{A}^{-2})$ on $L^2(\Omega)$
- covariance operator is given by the inverse of differential operator $\mathcal{A}^2$, where $\mathcal{A}(m) := -\alpha \boldsymbol{\nabla} \cdot (\boldsymbol{\Theta} \boldsymbol{\nabla} m) + \alpha m$
- invert for $m \in E$, Cameron-Martin space with inner product $(\cdot, \cdot)_E := (\mathcal{A}\cdot, \mathcal{A}\cdot)$
- Bayesian solution of the inverse problem defined as conditional measure μ^d of m given the data $\boldsymbol{d}_{\text{obs}} \in \mathbb{R}^q$, where

$$\frac{d\mu^d}{d\mu_0} = \frac{1}{Z}\pi_{\text{like}}(\boldsymbol{d}_{\text{obs}}|m) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{f}(m) - \boldsymbol{d}_{\text{obs}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}}\right)$$

  is the Radon-Nikodym derivative of μ^d w.r.t. μ_0, $\boldsymbol{f}(m)$ is the parameter-to-observable map, and $\boldsymbol{\Gamma}_{\text{noise}}$ the noise covariance operator
- leads to well-posed Bayesian inverse problem; exploits fast elliptic solvers

Finite dimensional discretization

- Care must be taken when discretizing to ensure convergence to infinite dimensional posterior: since $m \in L^2$, finite dimensional approximation inherits L^2-inner product; model coefficient vector endowed with weighted inner product $(\cdot,\cdot)_{\boldsymbol{M}}$, where $\boldsymbol{M}$ is a mass matrix; this gives rise to mass matrix weightings when drawing samples, computing variances, etc.
- Also, adjoint of an operator is not the same as the transpose; covariance matrix not symmetric w.r.t. Euclidean inner product (but it is w.r.t. L^2-inner product and is therefore self-adjoint)
- Details in:
  - T. Bui-Thanh, O. Ghattas, J. Martin, G. Stadler, *A computational framework for infinite-dimensional Bayesian inverse problems. Part I: The linearized case, with applications to global seismic inversion*, **SIAM Journal on Scientific Computing**, 35(6):A2494–A2523, 2013.
  - N. Petra, J. Martin, G. Stadler, O. Ghattas, *A computational framework for infinite-dimensional Bayesian inverse problems: Part II. Stochastic Newton MCMC with application to ice sheet inverse problems*, **SIAM Journal on Scientific Computing**, 36(4):A1525–A1555, 2014.

Making large-scale Hessians tractable

- $\boldsymbol{H}$ is sum of Hessian of data misfit, which is often a compact operator, and the inverse of a prior, which we take as an elliptic differential operator:

$$\boldsymbol{H} = \overbrace{\boldsymbol{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{F}}^{\text{Hessian of data misfit}} + \overbrace{\boldsymbol{\Gamma}_{\text{pr}}^{-1}}^{\text{Hessian of prior}}$$

- Factor out $\boldsymbol{\Gamma}_{\text{pr}}^{1/2}$ to expose prior-preconditioned data misfit Hessian, $\tilde{\boldsymbol{H}}_{\text{data}}$:

$$\boldsymbol{H} = \boldsymbol{\Gamma}_{\text{pr}}^{-1/2} (\underbrace{\boldsymbol{\Gamma}_{\text{pr}}^{1/2} \boldsymbol{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{F} \, \boldsymbol{\Gamma}_{\text{pr}}^{1/2}}_{\tilde{\boldsymbol{H}}_{\text{data}}} + \boldsymbol{I}) \boldsymbol{\Gamma}_{\text{pr}}^{-1/2}$$

- For many ill-posed inverse problems, eigenvalues of $\tilde{\boldsymbol{H}}_{\text{data}}$ decay rapidly: data have limited ability to inform parameter field, due to:
  - limited number of observations (rank of $\tilde{\boldsymbol{H}}_{\text{data}} \leq q$)
  - smoothness of parameter-to-observable map
- Construct *low rank* approximation of $\tilde{\boldsymbol{H}}_{\text{data}}$
- Cannot form $\tilde{\boldsymbol{H}}_{\text{data}}$; must use *matrix-free* methods such as Lanczos or randomized SVD that require only actions of $\tilde{\boldsymbol{H}}_{\text{data}}$
  - each $\tilde{\boldsymbol{H}}_{\text{data}}$ action entails a pair of Poisson solves (with operator $\mathcal{A}$) and a pair of linearized forward/adjoint solves (with $\boldsymbol{F}$ and $\boldsymbol{F}^*$)

Randomized SVD for low rank approximation of $\tilde{\boldsymbol{H}}_{\text{data}}$

- Computing $\tilde{\boldsymbol{H}}_{\text{data}}$ is intractable using conventional algorithms
- Randomized SVD constructs low rank approximation of $\tilde{\boldsymbol{H}}_{\text{data}}$ with matrix-free algorithm: cost is $2r$ products with random vectors (r is numerical rank of $\tilde{\boldsymbol{H}}_{\text{data}}$)
- Resulting cost is $2r$ incremental forward/adjoint solves and $4r$ Poisson solves (Steps 2 and 4)

Randomized SVD (double pass algorithm)

1. Generate i.i.d. Gaussian matrix $\boldsymbol{R} \in \mathbb{R}^{n \times r}$ with $r =$ numerical rank of $\tilde{\boldsymbol{H}}_{\text{data}}$ ($r \ll n$)
2. Form $\boldsymbol{Y} = \tilde{\boldsymbol{H}}_{\text{data}} \boldsymbol{R}$
3. Compute $\boldsymbol{Q} =$ orthonormal basis for $\boldsymbol{Y}$
4. Define $\boldsymbol{B} \in \mathbb{R}^{r \times r} := \boldsymbol{Q}^T \tilde{\boldsymbol{H}}_{\text{data}} \boldsymbol{Q}$
5. Decompose $\boldsymbol{B} = \boldsymbol{Z} \boldsymbol{\Lambda} \boldsymbol{Z}^T$
6. Low-rank approximation: $\tilde{\boldsymbol{H}}_{\text{data}} \approx \boldsymbol{V} \boldsymbol{\Lambda} \boldsymbol{V}^T$, where $\boldsymbol{V} \in \mathbb{R}^{n \times r} := \boldsymbol{Q} \boldsymbol{Z}$

- Prior operator is compact and Hessian is often so (consequence of smoothing of p2o map: limited ability to infer parameter modes from data)
- Thus r is finite dimensional and $r \ll n$
- **Randomized SVD can construct accurate low rank approximation of prior-preconditioned Hessian at cost measured in (linearized) PDE solves that is independent of parameter dimension, and depends only on number of parameter modes that are informed by the data**

Randomized SVD for low rank approximation of $\tilde{\boldsymbol{H}}_{\text{data}}$

N. Halko, P.-G. Martinsson, and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, **SIAM Review**, 53(2)217–288, 2011

- Note that for deterministic SVD, $\left(\sum_{j=r+1}^{n} \lambda_j\right)^{1/2}$ is the minimal Frobenius norm error in approximating $\boldsymbol{A}$ with a rank-r matrix
- Result for randomized SVD (with Gaussian test matrix R): oversample with p random vectors:

$$\mathbb{E} \parallel \tilde{\boldsymbol{H}}_{\text{data}} - \tilde{\boldsymbol{H}}_{\text{data}}^{r+p} \parallel_F \leq \left(1 + \frac{r}{p-1}\right)^{1/2} \left(\sum_{j=r+1}^{n} \lambda_j\right)^{1/2}$$

- p need not be too large to produce expected error within a small constant factor of optimal
- Concentration of measure for random matrices:
  $\longrightarrow$ Variance of $\|\tilde{\boldsymbol{H}}_{\text{data}} - \tilde{\boldsymbol{H}}_{\text{data}}^{r+p}\|$ is small

Low rank approximation of data misfit Hessian

- Use Sherman-Morrison-Woodbury to invert/factor:

$$
\begin{aligned}
\boldsymbol{\Gamma}_{\text{post}} &= \boldsymbol{H}^{-1} \\
&= \left(\boldsymbol{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{F} + \boldsymbol{\Gamma}_{\text{pr}}^{-1}\right)^{-1} \\
&= \boldsymbol{\Gamma}_{\text{pr}}^{1/2} \Big(\underbrace{\boldsymbol{\Gamma}_{\text{pr}}^{1/2} \boldsymbol{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{F}\, \boldsymbol{\Gamma}_{\text{pr}}^{1/2}}_{\approx \boldsymbol{V}_r \boldsymbol{\Lambda}_r \boldsymbol{V}_r^*} + \boldsymbol{I} \Big)^{-1} \boldsymbol{\Gamma}_{\text{pr}}^{1/2} \\
&= \boldsymbol{\Gamma}_{\text{pr}}^{1/2} \left[\boldsymbol{I} - \boldsymbol{V}_r \boldsymbol{D}_r \boldsymbol{V}_r^* + \mathcal{O}\left(\sum_{i=r+1}^{n} \frac{\lambda_i}{\lambda_i + 1} \right) \right] \boldsymbol{\Gamma}_{\text{pr}}^{1/2} \\
&\approx \boldsymbol{\Gamma}_{\text{pr}} - \boldsymbol{\Gamma}_{\text{pr}}^{1/2} \boldsymbol{V}_r \boldsymbol{D}_r \boldsymbol{V}_r^* \boldsymbol{\Gamma}_{\text{pr}}^{1/2}
\end{aligned}
$$

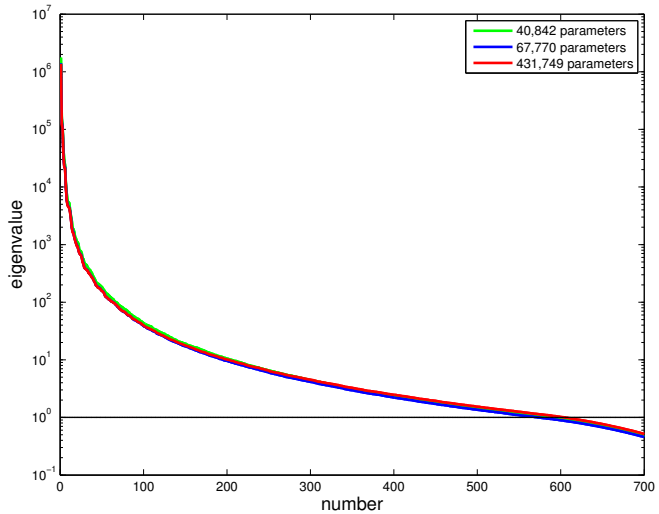
  where $\boldsymbol{V}_r, \boldsymbol{\Lambda}_r$ are truncated eigenvector/values of prior-preconditioned data misfit Hessian, and $\boldsymbol{D}_r = \text{diag}(\lambda_i/(\lambda_i + 1))$
- **Optimality of the low rank Hessian approximation:**
  A. Spantini, et al., SISC 2015
- Can bound error between exact and approximate $\boldsymbol{\Gamma}_{\text{post}}$ independent of mesh size for specific model problems (P. Flath dissertation, 2013)

Low-rank-based posterior covariance

Posterior covariance is given by prior covariance less information gained from data; it is a low rank update to the prior:

$$\boldsymbol{\Gamma}_{\text{post}} \approx \underbrace{\boldsymbol{\Gamma}_{\text{pr}}}_{\text{rank } n} - \underbrace{\boldsymbol{\Gamma}_{\text{pr}}^{1/2} \boldsymbol{V}_r \boldsymbol{D}_r \boldsymbol{V}_r^* \boldsymbol{\Gamma}_{\text{pr}}^{1/2}}_{\text{rank } r}$$

The data term is the only place where the forward PDE is involved



$$\lambda_i(\boldsymbol{\Gamma}_{\text{pr}}^{1/2} \boldsymbol{F}^* \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{F}\, \boldsymbol{\Gamma}_{\text{pr}}^{1/2})$$

Spectra of prior-preconditioned data misfit spectrum
for a large-scale problem using three different mesh refinement levels

Dimension-independence (scalability) of entire process

Never necessary to form dense operators:

$$\boldsymbol{H} = \boldsymbol{\Gamma}_{\text{pr}}^{-1/2}\left[\boldsymbol{V}_r\boldsymbol{\Lambda}_r\boldsymbol{V}_r^* + \boldsymbol{I}\right]\boldsymbol{\Gamma}_{\text{pr}}^{-1/2}$$

$$\boldsymbol{H}^{-1}\boldsymbol{g} = \boldsymbol{\Gamma}_{\text{pr}}^{1/2}\left\{\boldsymbol{V}_r\left[(\boldsymbol{\Lambda}_r + \boldsymbol{I}_r)^{-1} - \boldsymbol{I}_r\right]\boldsymbol{V}_r^* + \boldsymbol{I}\right\}\boldsymbol{\Gamma}_{\text{pr}}^{1/2}\boldsymbol{g} \quad \text{(Newton step)}$$

$$\boldsymbol{H}^{-1/2}\boldsymbol{x} = \boldsymbol{\Gamma}_{\text{pr}}^{1/2}\left\{\boldsymbol{V}_r\left[(\boldsymbol{\Lambda}_r + \boldsymbol{I}_r)^{-1/2} - \boldsymbol{I}_r\right]\boldsymbol{V}_r^* + \boldsymbol{I}\right\}\boldsymbol{M}^{-1/2}\boldsymbol{x} \quad \text{(Sample generation)}$$

$$c^{\text{post}}(\boldsymbol{x},\boldsymbol{x}) = c^{\text{prior}}(\boldsymbol{x},\boldsymbol{x}) - \sum_{k=1}^{r} d_k \left(\left[\boldsymbol{\Gamma}_{\text{pr}}^{1/2}\boldsymbol{v}_k\right](\boldsymbol{x})\right)^2 \quad \text{(Pointwise variance field)}$$

Complexity of these operations is scalable (i.e. requires a number of forward PDE solves that is independent of the parameter and data dimensions) when:

- prior-preconditioned data misfit Hessian is compact with mesh/data independent dominant spectrum (ill-posed inverse problem)
- dominant spectrum captured in $O(r)$ matvecs; use randomized SVD
- Matrix-free adjoint-based Hessian-vector products $\Rightarrow$ two (linearized) PDE solves
- fast $O(n)$ elliptic solvers used for $\boldsymbol{\Gamma}_{\text{pr}}^{1/2}\boldsymbol{z}$ (e.g., multigrid)
- note $\boldsymbol{M}$ is a well-behaved operator (with spectrum close to the identity), so $\boldsymbol{M}^{-1/2}\boldsymbol{x}$ can be computed in $O(n)$ using iterative techniques (Y. Saad, et al. SISC 2011)

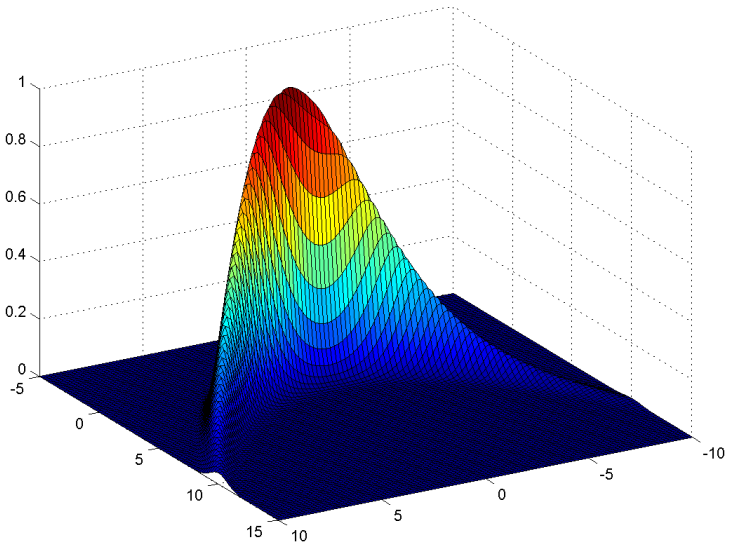
What about non-Gaussian posteriors? Hessian still plays an important role

- Importance sampling: generate samples from Hessian-based Gaussian at MAP, re-weight using true posterior values
- Stochastic Newton: use Hessian-based Gaussian as proposal for MCMC (Martin, Wilcox, Burstedde, & Ghattas, SISC 2012)
- Dimension-Independent Likelihood-Informed (DILI) method (Cui, Law, & Marzouk, arXiv 2014)
- Optimal transport (Marzouk et al.)
- Riemannian manifold Langevin, Hamiltonian Monte Carlo (Girolami & Calderhead, JRSS 2011; Bui-Thanh & Girolami, Inverse Problems 2014)
- Randomized Maximum Likelihood (Oliver, 2008), Randomized-then-Optimize (Bardsley, SISC 2012)
- Inf-dim Geometric MCMC (Beskos, Girolami, Lan, Farrell, Stuart, JCP 2017)

Markov chain Monte Carlo sampling of $\pi_{\text{post}}(\boldsymbol{m})$

For Gaussian additive noise $\sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_{\text{noise}})$ and Gaussian prior $\sim \mathcal{N}(\boldsymbol{m}_{\text{pr}}, \boldsymbol{\Gamma}_{\text{pr}})$

$$\pi_{\text{post}}(\boldsymbol{m}) \propto \exp\left(-\frac{1}{2} \parallel \boldsymbol{f}(\boldsymbol{m}) - \boldsymbol{d}_{\text{obs}} \parallel^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} - \frac{1}{2} \parallel \boldsymbol{m} - \boldsymbol{m}_{\text{pr}} \parallel^2_{\boldsymbol{\Gamma}^{-1}_{\text{pr}}}\right)$$



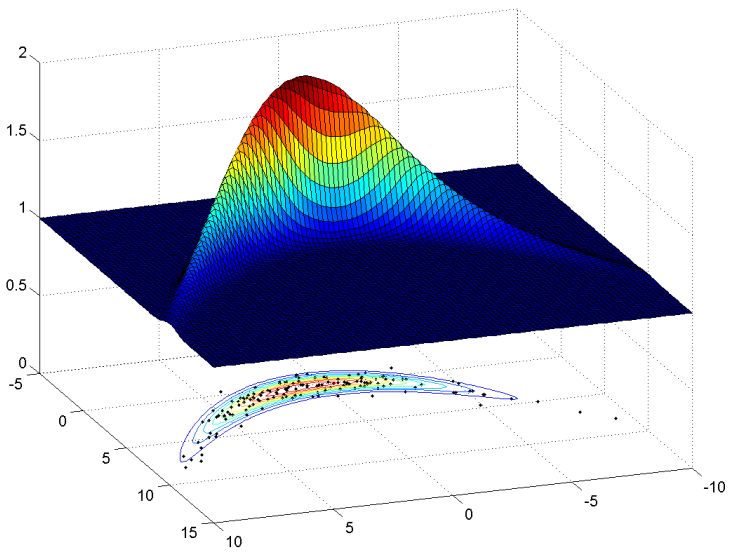
example probability density

- How do we explore this distribution?
- Often high dimensional
- Computationally expensive

Markov chain Monte Carlo sampling of $\pi_{\text{post}}(\boldsymbol{m})$

For Gaussian additive noise $\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\text{noise}})$ and Gaussian prior $\sim \mathcal{N}(\boldsymbol{m}_{\text{pr}}, \boldsymbol{\Gamma}_{\text{pr}})$

$$\pi_{\text{post}}(\boldsymbol{m}) \propto \exp\left(-\frac{1}{2} \parallel \boldsymbol{f}(\boldsymbol{m}) - \boldsymbol{d}_{\text{obs}} \parallel^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} -\frac{1}{2} \parallel \boldsymbol{m} - \boldsymbol{m}_{\text{pr}} \parallel^2_{\boldsymbol{\Gamma}^{-1}_{\text{pr}}}\right)$$



sampled probability density

The MCMC approach

- Replace $\pi(\boldsymbol{m})$ by sample chain $\{\boldsymbol{m}_k\}$
- Compute using ergodic averages

$$\mathbb{E}[f(M)] = \int_{\mathbb{R}^n} f(m)\pi(dm)$$

$$\approx \frac{1}{N}\sum_{j=1}^{N} f(m_k)$$

Markov Chain Monte Carlo

Metropolis-Hastings MCMC

Initialize $\boldsymbol{m}_0$, $\pi_{\text{post}}(\boldsymbol{m}_0)$

for $k = 0, \ldots$ **do**

Draw sample $\boldsymbol{y}$ from the proposal density $q(\boldsymbol{m}_k, \cdot)$

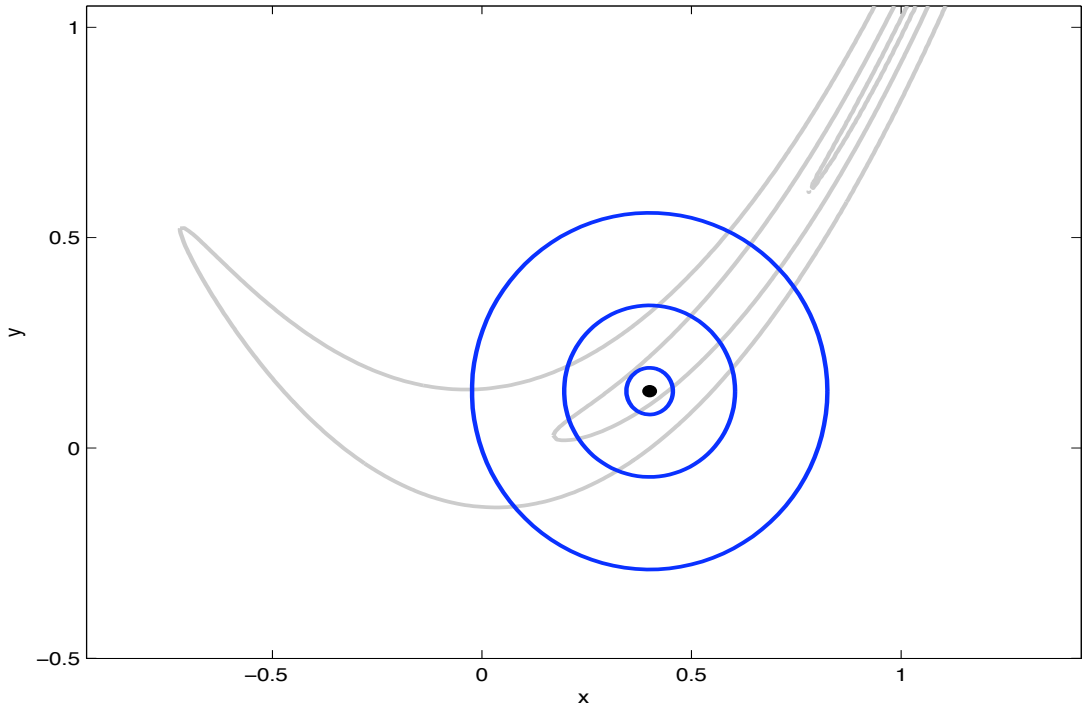
Compute the acceptance probability:

$$\alpha(\boldsymbol{m}_k, \boldsymbol{y}) = \min\left\{1, \frac{\pi_{\text{post}}(\boldsymbol{y})\; q(\boldsymbol{y}, \boldsymbol{m}_k)}{\pi_{\text{post}}(\boldsymbol{m}_k)\; q(\boldsymbol{m}_k, \boldsymbol{y})}\right\}$$

Accept/Reject

end for

Isotropic Gaussian proposal: Gaussian random walk

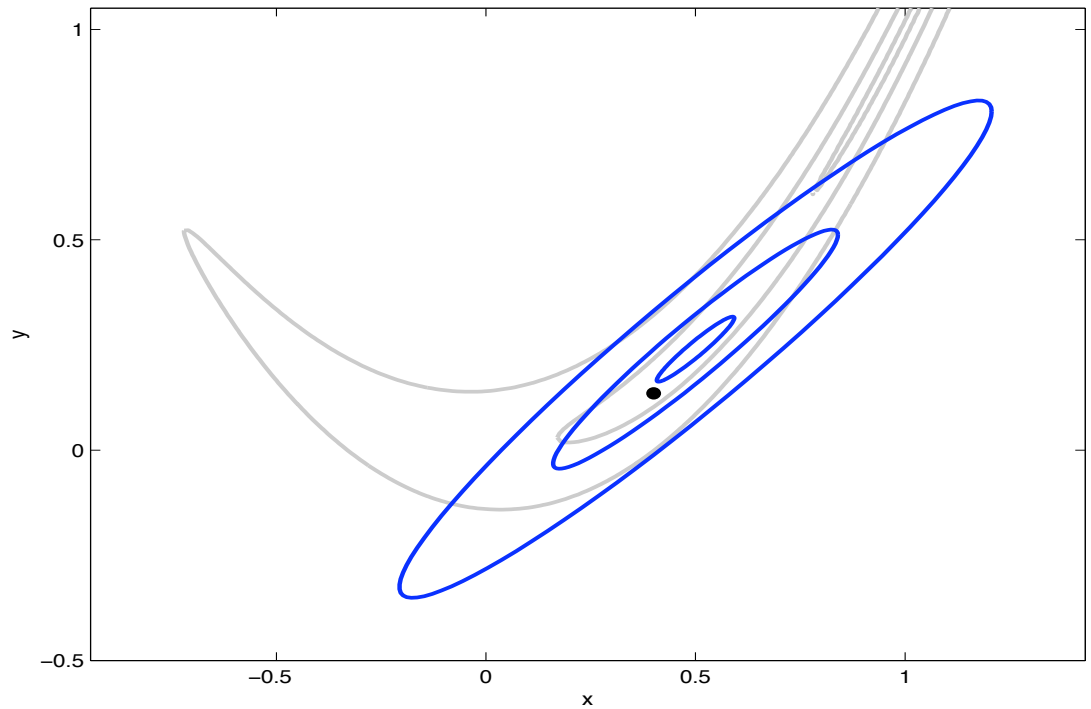


- an isotropic Gaussian gives the *random walk method*:

$$q(\boldsymbol{m}_k, \boldsymbol{y}) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp[-(\boldsymbol{y} - \boldsymbol{m}_k)^T (\boldsymbol{y} - \boldsymbol{m}_k)/(2\sigma^2)]$$

- challenge: devise a proposal density $q(\boldsymbol{m}_k, \boldsymbol{y})$ that is cheap, easy to sample, and a good representation of the underlying posterior—in high dimensions and for expensive-to-evaluate likelihoods
- must exploit structure of parameter-to-observable map

Hessian-based Gaussian proposal for MCMC

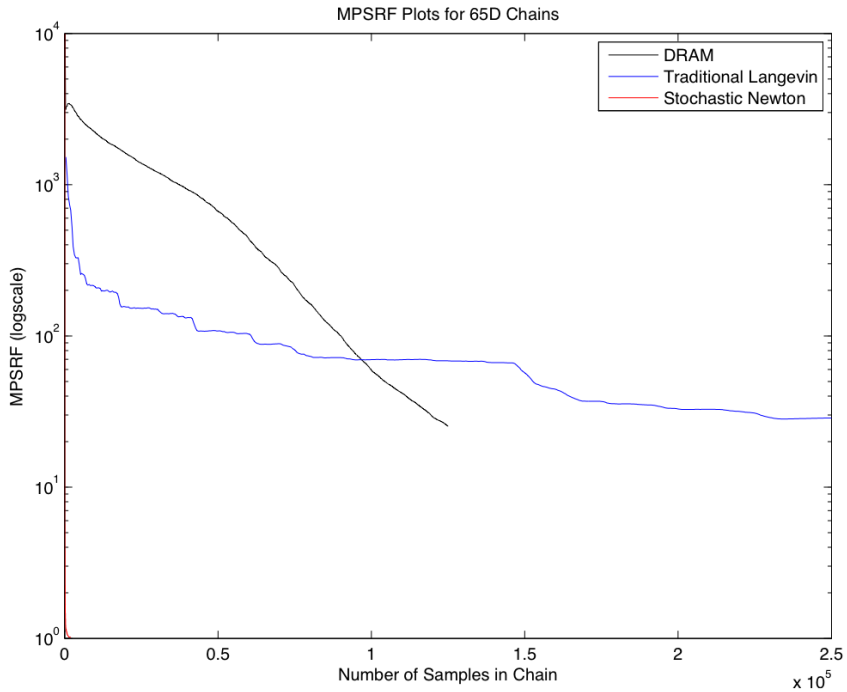


- **Stochastic Newton method:** proposal is a local Gaussian approximation based on local gradient $\boldsymbol{g}$ & Hessian $\boldsymbol{H}$ of negative log posterior:

$$q(\boldsymbol{m}_k, \boldsymbol{y}) = \frac{\det \boldsymbol{H}^{1/2}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\left(\boldsymbol{y} - \boldsymbol{m}_k + \boldsymbol{H}^{-1}\boldsymbol{g}\right)^T \boldsymbol{H}\left(\boldsymbol{y} - \boldsymbol{m}_k + \boldsymbol{H}^{-1}\boldsymbol{g}\right)\right)$$

- Can also be derived via inverse Hessian-preconditioned (Metropolized) Langevin dynamics: $\boldsymbol{m}_{k+1}^{\text{prop}} = \boldsymbol{m}_k - \boldsymbol{H}^{-1}\nabla_{\boldsymbol{m}}(-\log\pi) + \mathcal{N}(\boldsymbol{0}, \boldsymbol{H}^{-1})$

Convergence comparison for 65-parameter problem



Multivariate potential scale reduction factor convergence statistic
unpreconditioned Langevin vs. stochastic Newton vs. Adaptive Metropolis

J. Martin, L.C. Wilcox, C. Burstedde, and O. Ghattas, *A Stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion*, **SIAM Journal on Scientific Computing**, 34(3):A1460-A1487, 2012.

See hIPPYlib library: `https://hippylib.github.io`

Outline

1. The inverse problem: Integrating data and models
2. Examples of Bayesian inverse problems
3. Target: Flow of the Antarctic ice sheet
4. Large-scale Bayesian inverse problems
5. Application to Antarctic ice sheet flow inverse problem

Newton-CG for finding MAP points: Gradients & Hessians

Minimize regularized data misfit

$$\min_{\beta} \mathcal{J}(\beta) := \frac{1}{2}\int_{\Gamma_t} (\mathcal{B}\boldsymbol{u}(\beta) - \boldsymbol{d}_{\text{obs}})^2 \, d\boldsymbol{s} + \frac{\alpha}{2}\int_{\Gamma_b} \nabla_\Gamma \beta \cdot \nabla_\Gamma \beta \, d\boldsymbol{s}$$

Solve Newton system using CG, inexactly and operator-free

$$\mathcal{H}(\beta)\tilde{\beta} = -\mathcal{G}(\beta), \qquad \beta_{\text{new}} = \beta + \gamma\tilde{\beta}$$

Gradient given by:

$$\mathcal{G}(\beta) := \exp(\beta)\, \boldsymbol{u}_\Gamma(\beta) \cdot \boldsymbol{v}_\Gamma(\beta) + \alpha\Delta_\Gamma\beta \quad \text{on } \Gamma_b$$

- $\boldsymbol{u}$: velocity, $\boldsymbol{v}$: adjoint velocity, β: log basal sliding coefficient field
- $\boldsymbol{d}_{\text{obs}}$: observed surface velocity, α: regularization parameter
- $\mathcal{B}$: observation operator, $\mathcal{G}$: gradient operator, $\mathcal{H}$: Hessian operator
- Γ: indicates surface tangential component or surface operator

Gradient computation: adjoint (linear) Stokes equation

u and p satisfy the *forward (nonlinear) Stokes equations*

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

v and q satisfy the *adjoint Stokes equations*

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{v} &= 0 && \text{in } \Omega \\
-\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}_{\boldsymbol{v}} &= \boldsymbol{0} && \text{in } \Omega \\
\boldsymbol{\sigma}_{\boldsymbol{v}} \boldsymbol{n} &= -\mathcal{B}^*(\mathcal{B}\boldsymbol{u} - \boldsymbol{d}_{\text{obs}}) && \text{on } \Gamma_t \\
\boldsymbol{v} \cdot \boldsymbol{n} = 0, \quad (\boldsymbol{\sigma}_{\boldsymbol{v}} \boldsymbol{n})_\Gamma + \exp(\beta) \boldsymbol{v}_\Gamma &= \boldsymbol{0} && \text{on } \Gamma_b
\end{aligned}
$$

where the adjoint stress $\boldsymbol{\sigma}_{\boldsymbol{v}}$ is

$$
\boldsymbol{\sigma}_{\boldsymbol{v}} := 2\eta(\boldsymbol{u}) \left(\mathsf{I} + \frac{1-n}{n} \frac{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} \otimes \dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}}{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}} \cdot \dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}} \right) \dot{\boldsymbol{\varepsilon}}_{\boldsymbol{v}} - \boldsymbol{I} q
$$

Hessian action: 2 additional (linearized) Stokes-like eqns

Action of Hessian operator in direction $\hat{\beta}$ evaluated at β

$$\mathcal{H}(\beta)\hat{\beta} := \exp(\beta)\left(\hat{\beta}\,\boldsymbol{u}_\Gamma \cdot \boldsymbol{v}_\Gamma + \hat{\boldsymbol{u}}_\Gamma \cdot \boldsymbol{v}_\Gamma + \boldsymbol{u}_\Gamma \cdot \hat{\boldsymbol{v}}_\Gamma\right) + \alpha\Delta_\Gamma\hat{\beta}$$

where $\hat{u}$ and $\hat{p}$ satisfy the *incremental forward equations*

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\hat{\boldsymbol{u}} &= 0 && \text{in } \Omega\\
-\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_{\hat{\boldsymbol{u}}} &= \boldsymbol{0} && \text{in } \Omega\\
\boldsymbol{\sigma}_{\hat{\boldsymbol{u}}}\boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma_t\\
\hat{\boldsymbol{u}}\cdot\boldsymbol{n} = 0, \quad (\boldsymbol{\sigma}_{\hat{\boldsymbol{u}}}\boldsymbol{n})_\Gamma + \exp(\beta)\hat{\boldsymbol{u}}_\Gamma &= -\hat{\beta}\exp(\beta)\boldsymbol{u}_\Gamma && \text{on } \Gamma_b
\end{aligned}$$

with $\boldsymbol{\sigma}_{\hat{\boldsymbol{u}}} := 2\eta(\boldsymbol{u})\left(\mathsf{I} + \frac{1-n}{n}\frac{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\otimes\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}}{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\cdot\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}}\right)\dot{\boldsymbol{\varepsilon}}_{\hat{\boldsymbol{u}}} - \boldsymbol{I}\hat{p}$

N. Petra, H. Zhu, G. Stadler, T.J.R. Hughes, O. Ghattas, *A scalable adjoint-based inexact Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model*, Journal of Glaciology, 58(211):889903, 2012.

Hessian action: 2 additional (linearized) Stokes-like eqns

Action of Hessian operator in direction $\hat{\beta}$ evaluated at β

$$\mathcal{H}(\beta)\hat{\beta} := \exp(\beta)\left(\hat{\beta}\,\boldsymbol{u}_\Gamma \cdot \boldsymbol{v}_\Gamma + \hat{\boldsymbol{u}}_\Gamma \cdot \boldsymbol{v}_\Gamma + \boldsymbol{u}_\Gamma \cdot \hat{\boldsymbol{v}}_\Gamma\right) + \alpha\Delta_\Gamma\hat{\beta}$$

where $\tilde{v}$, $\tilde{q}$ satisfy the *incremental adjoint equations*

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\hat{\boldsymbol{v}} &= 0 && \text{in } \Omega\\
-\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}_{\hat{\boldsymbol{v}}} &= -\boldsymbol{\nabla}\cdot\boldsymbol{\tau}_{\hat{\boldsymbol{u}}} && \text{in } \Omega\\
\boldsymbol{\sigma}_{\hat{\boldsymbol{v}}}\boldsymbol{n} &= -\mathcal{B}^*\mathcal{B}\hat{\boldsymbol{u}} - \boldsymbol{\tau}_{\hat{\boldsymbol{u}}}\boldsymbol{n} && \text{on } \Gamma_t\\
\hat{\boldsymbol{v}}\cdot\boldsymbol{n} = 0, \quad (\boldsymbol{\sigma}_{\hat{\boldsymbol{v}}}\boldsymbol{n})_\Gamma + \exp(\beta)\hat{\boldsymbol{v}}_\Gamma &= -(\boldsymbol{\tau}_{\hat{\boldsymbol{u}}}\boldsymbol{n})_\Gamma && \text{on } \Gamma_b
\end{aligned}$$

with $\boldsymbol{\sigma}_{\hat{\boldsymbol{v}}} := 2\eta(\boldsymbol{u})\left(\mathsf{I} + \frac{1-n}{n}\frac{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\otimes\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}}{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\cdot\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}}\right)\dot{\boldsymbol{\varepsilon}}_{\hat{\boldsymbol{v}}} - \boldsymbol{I}\hat{q}$, and $\boldsymbol{\tau}_{\hat{\boldsymbol{u}}} = 2\eta(\boldsymbol{u})\Psi\dot{\boldsymbol{\varepsilon}}_{\hat{\boldsymbol{u}}}$, where

$$\Psi = (1 + \frac{1-n}{n}\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\cdot\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}})\mathsf{I} + \frac{1-n}{n}\left[\frac{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\otimes\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}}{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\cdot\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}} + 2\frac{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\otimes\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{v}}}{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\cdot\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}} + \frac{1-3n}{n}\frac{\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\otimes\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}}{(\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}}\cdot\dot{\boldsymbol{\varepsilon}}_{\boldsymbol{u}})^2}\right]$$

N. Petra, H. Zhu, G. Stadler, T.J.R. Hughes, O. Ghattas, *A scalable adjoint-based inexact Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model*, Journal of Glaciology, 58(211):889903, 2012.

Scalability of inexact Newton-CG

Pine Island Glacier region

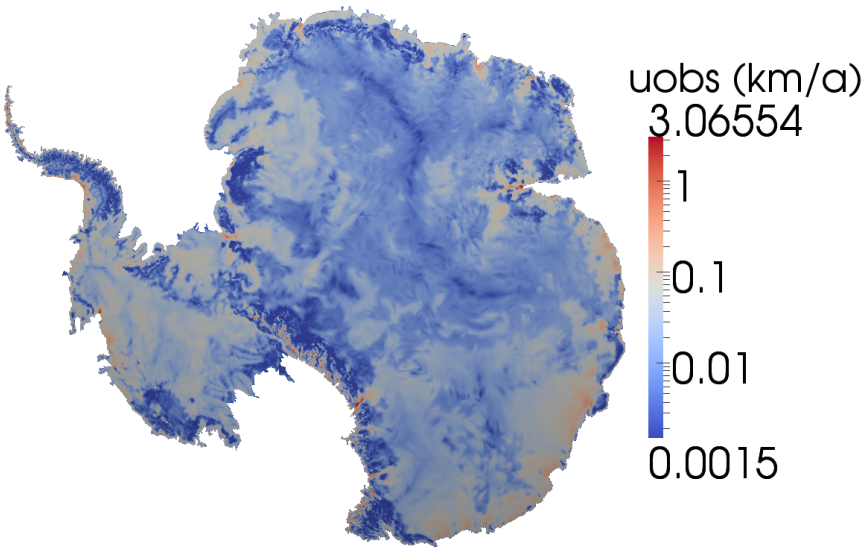
#sdof	#pdof	#N	#CG	avgCG	#Stokes
95,796	10,371	42	2718	65	7031
233,834	25,295	39	2342	60	6440
848,850	91,787	39	2577	66	6856
3,372,707	364,649	39	2211	57	6193
22,570,303	1,456,225	40	1923	48	5376

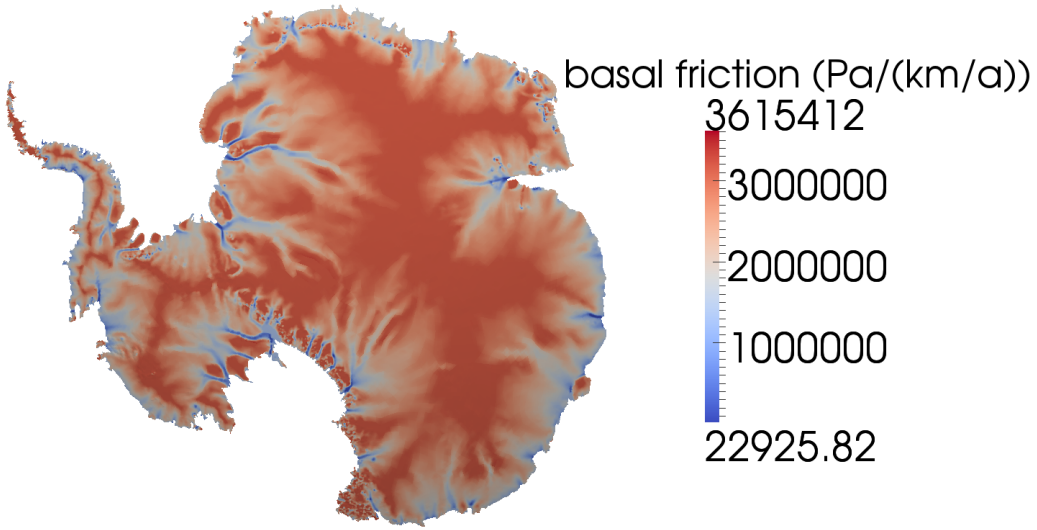
- **#sdof** : number of degrees of freedom for the state variables
- **#pdof** : number of degrees of freedom for the inversion parameter field
- **#N** : number of Newton iterations for the inverse problem
- **#CG** : total number of CG iterations for the inverse problem
- **#avgCG** : average number of CG iterations per Newton iteration
- **#Stokes** : total number of linear(ized) Stokes solves
- refinements obtained by decreasing max area of an element by a factor of 4
- convergence = reduction of gradient by factor of 10^5
- cost (measured by # of forward solves) is independent of parameter dimension and data dimension (CG performance consequence of preconditioned Hessian operator of form compact perturbation of identity)

Antarctic synthetic inversion for β: setup & performance

- Spatial discretization (all on same mesh):
  - velocity (state, adjoint, incremental state, incremental adjoint): $\mathbb{Q}_2$
  - pressure (state, adjoint, incremental state, incremental adjoint): $\mathbb{Q}_0^{\text{disc}}$
  - log basal sliding coefficient and its increment: $\mathbb{Q}_2$ (biquadratic)
- inexact Newton-CG, preconditioned by L-BFGS
- # state parameters: 4,085,841
- # inversion parameters: 409,545
- # elements: 99,984
- # of processor cores: 1024
- noise level used to synthesize data: 1%
- reduction in norm of gradient: 10^{-6}
- # of Gauss-Newton iterations: 18
- average # of CG iterations per Gauss-Newton iteration: 85
- total # of (linearized) Stokes solves: 2600

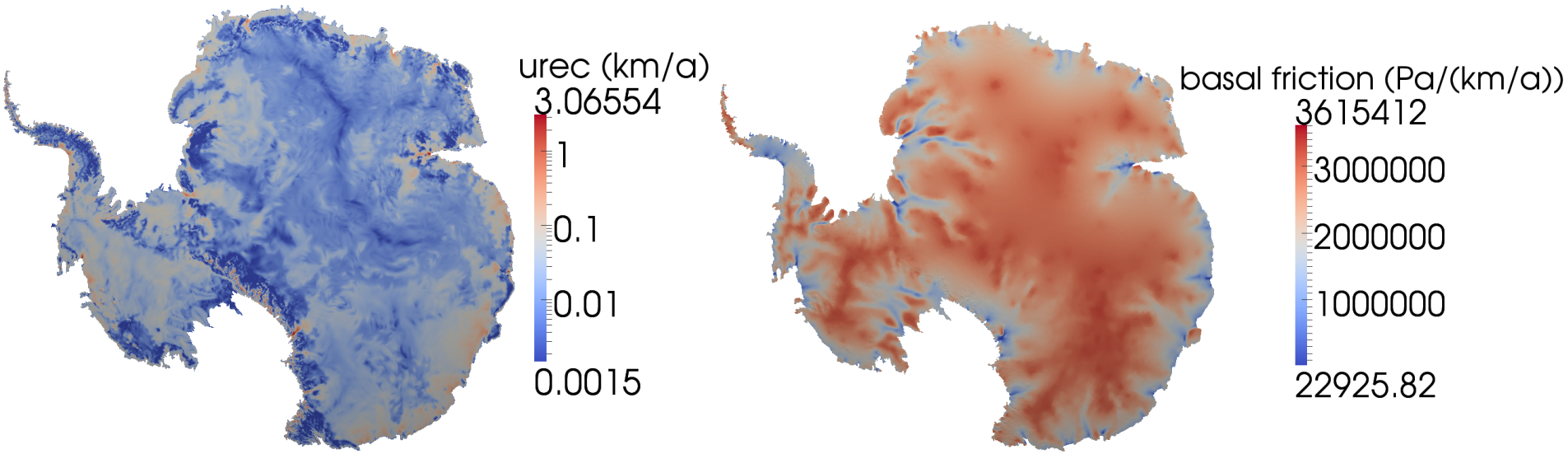
Antarctic ice sheet inversion for basal sliding field: Synthetic data





Left: Synthetic surface velocity observations
Right: "Truth" basal sliding field

Antarctic ice sheet inversion for basal sliding field: Synthetic data

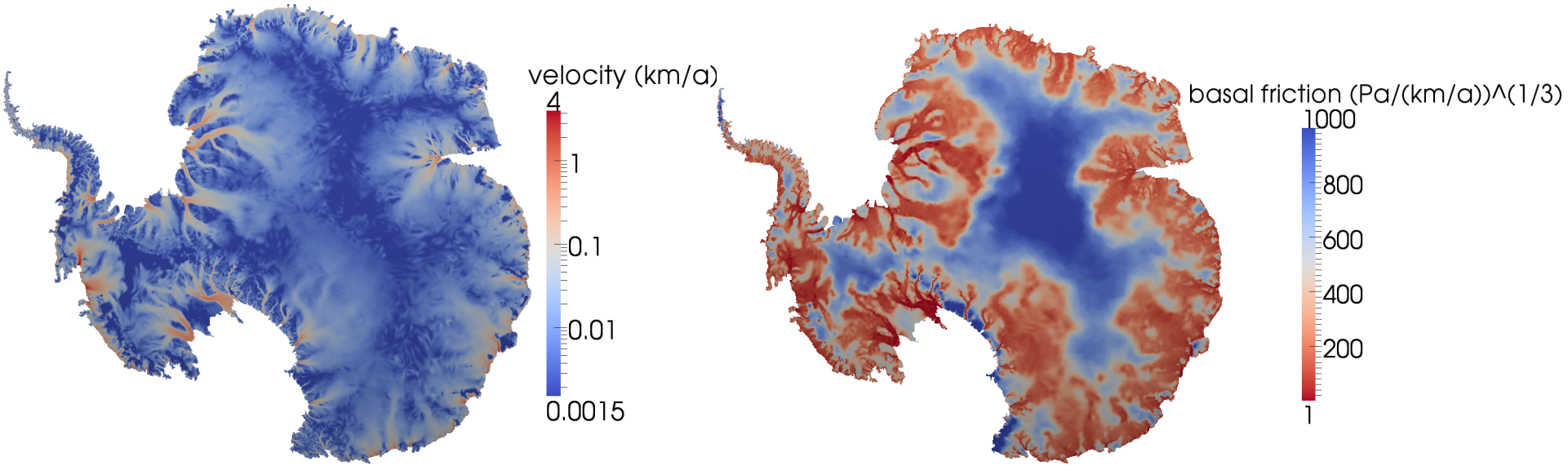


Left: Reconstructed surface velocity field
Right: Inferred basal sliding field

InSAR observed surface ice velocity dataset

- E. Rignot, J. Mouginot, and B. Scheuchl, Ice Flow of the Antarctic Ice Sheet, *Science*, 333(6048):1427–1430, 2011.
- Available from `http://nsidc.org/data/nsidc-0484.html`
- Assembled from multiple satellite interferometric synthetic-aperture radar data from 2007–2009
- Data set integrates 900 satellite tracks and more than 3,000 orbits

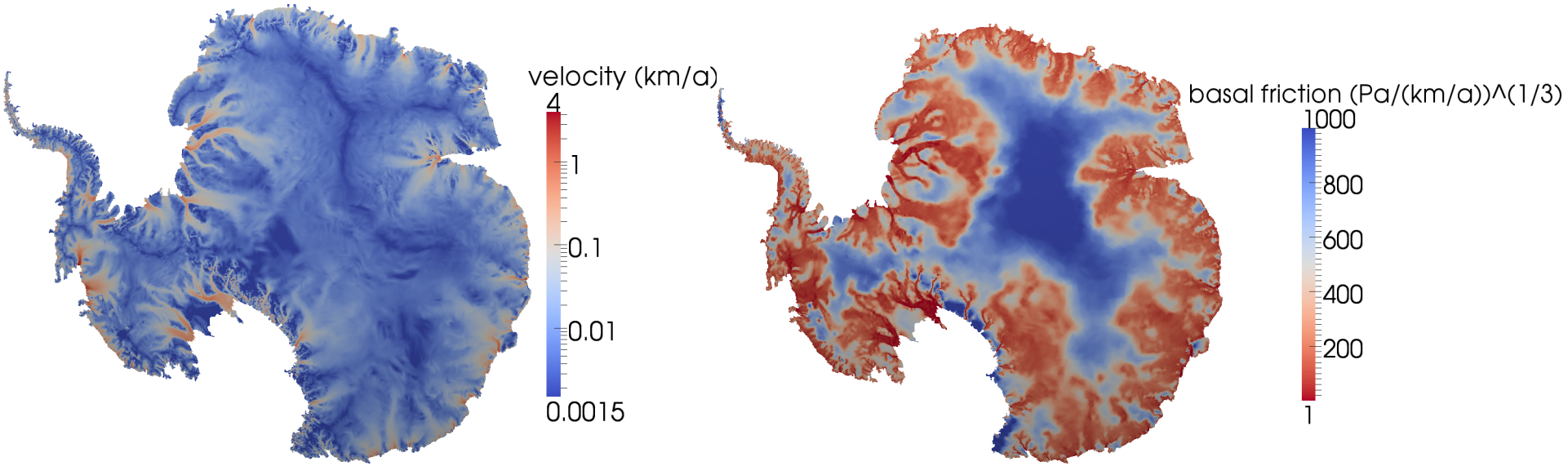
Antarctic ice sheet inversion for basal sliding field: InSAR data



Left: InSAR-based Antarctica ice surface velocity observations
Right: Inferred basal sliding field

More details: 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*, **Journal of Computational Physics**, 296(1):348–368, 2015.

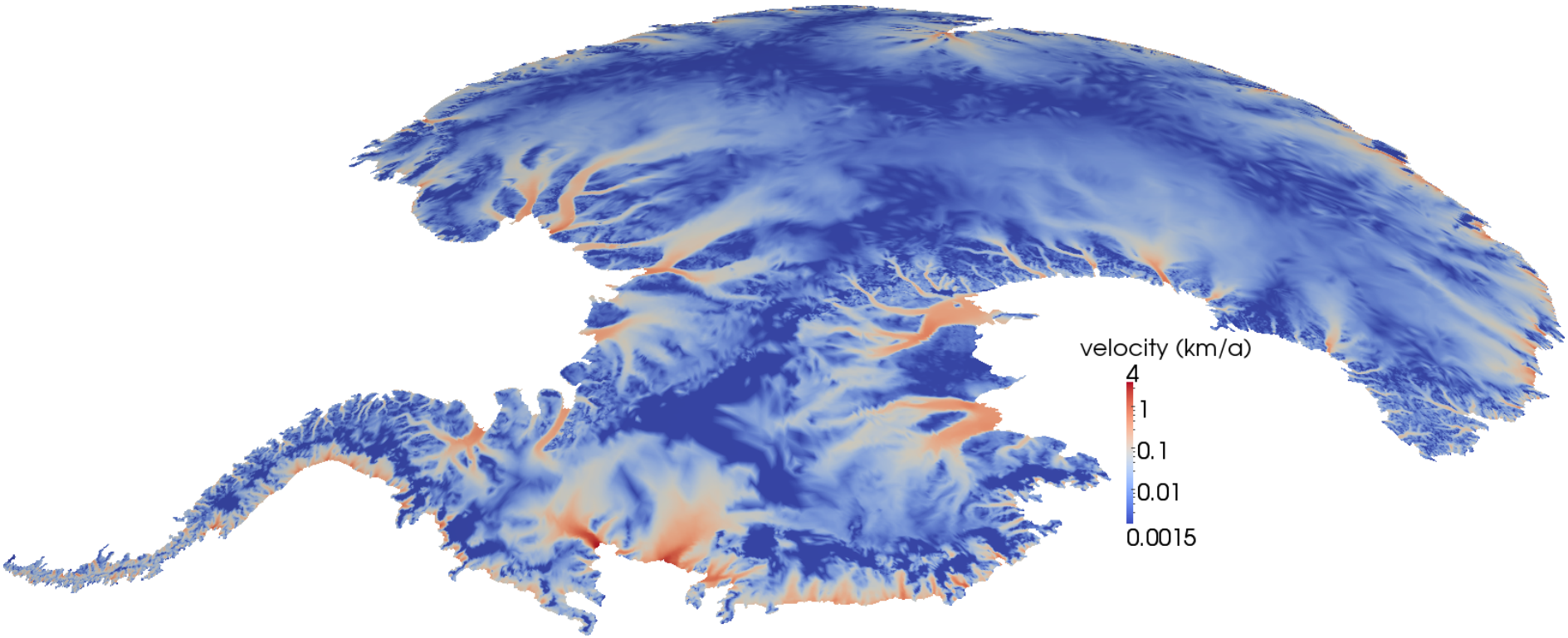
Antarctic ice sheet inversion for basal sliding field: InSAR data



Left: Reconstructed ice surface velocity field
Right: Inferred basal sliding field

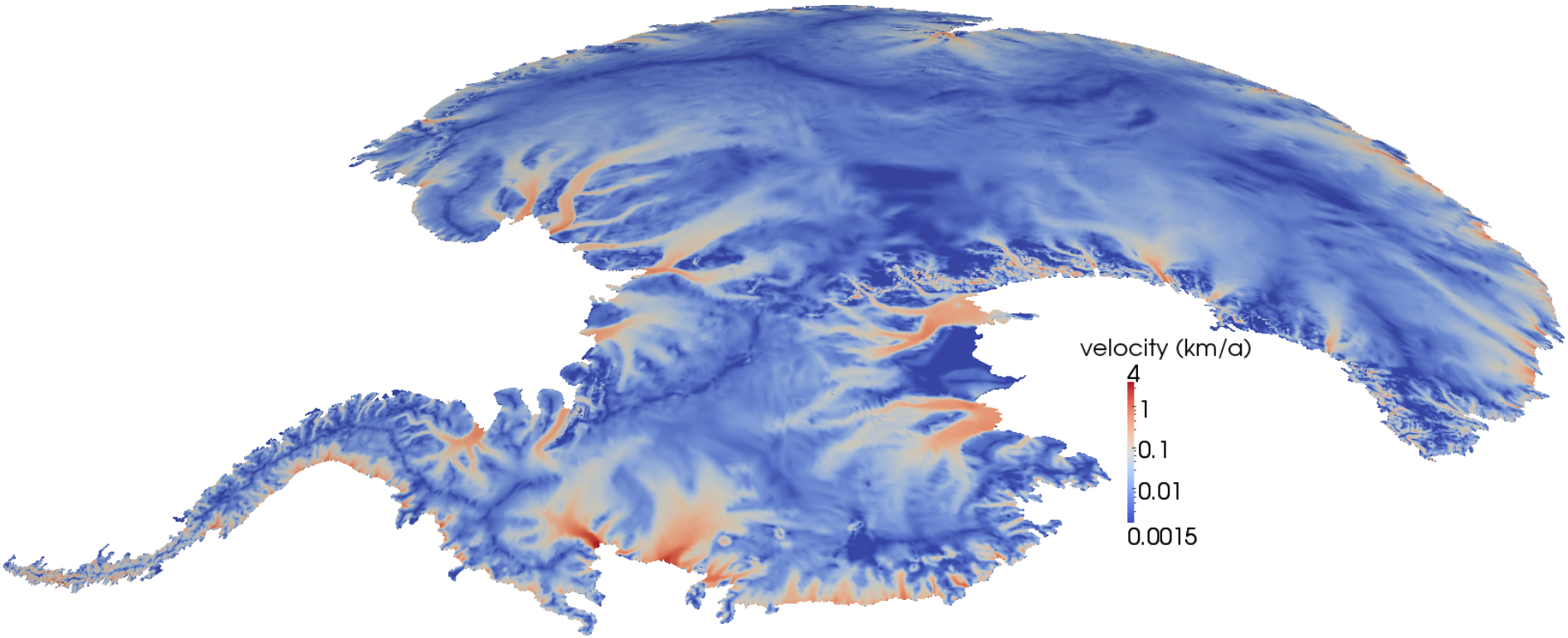
More details: T. Isaac, N. Petra, G. Stadler, O. Ghattas, *Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet*, **Journal of Computational Physics**, 296(1):348–368, 2015.

Antarctic ice sheet inversion for basal sliding field: InSAR data



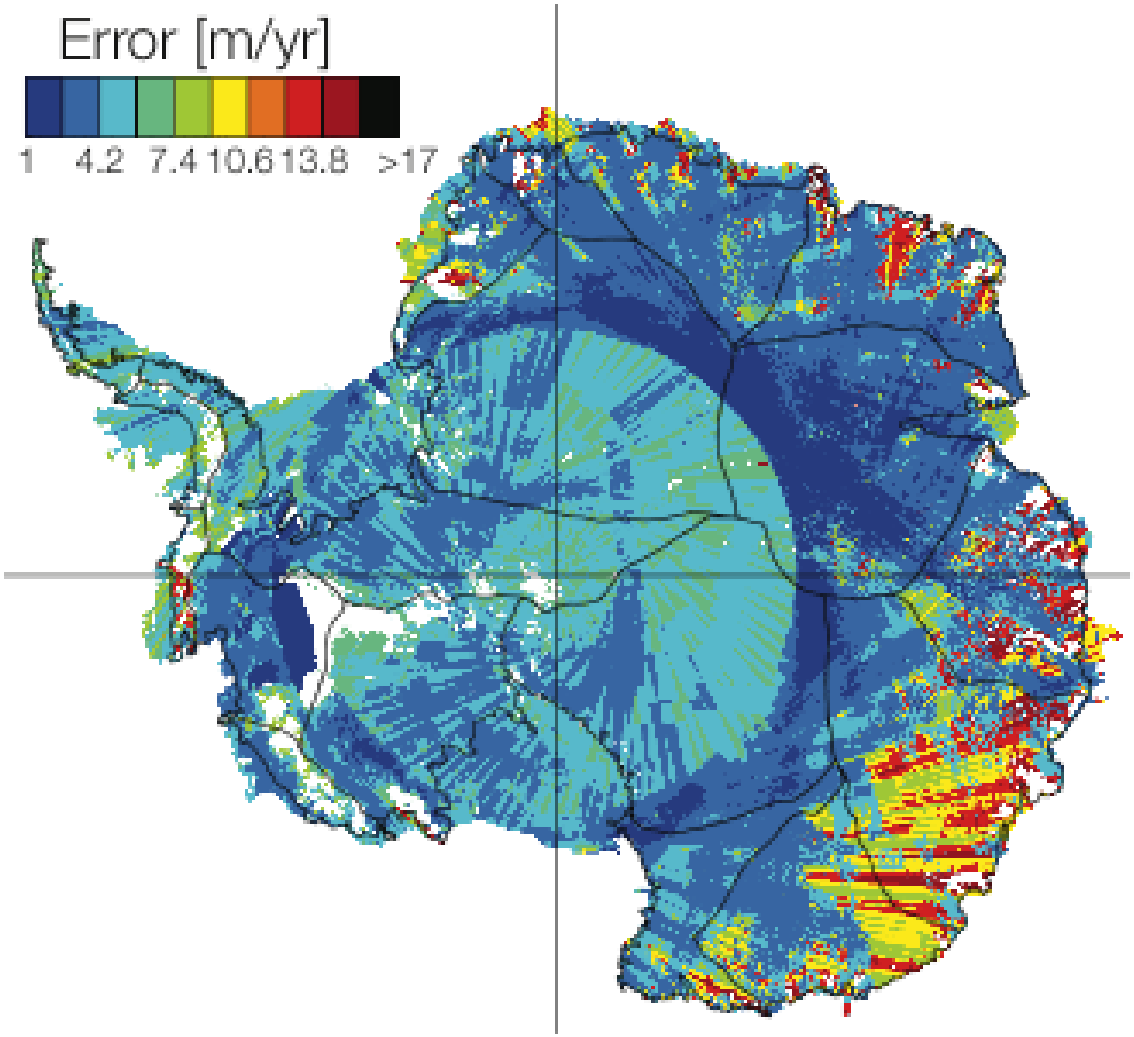
InSAR-based Antarctica ice surface velocity observations

Antarctic ice sheet inversion for basal sliding field: InSAR data



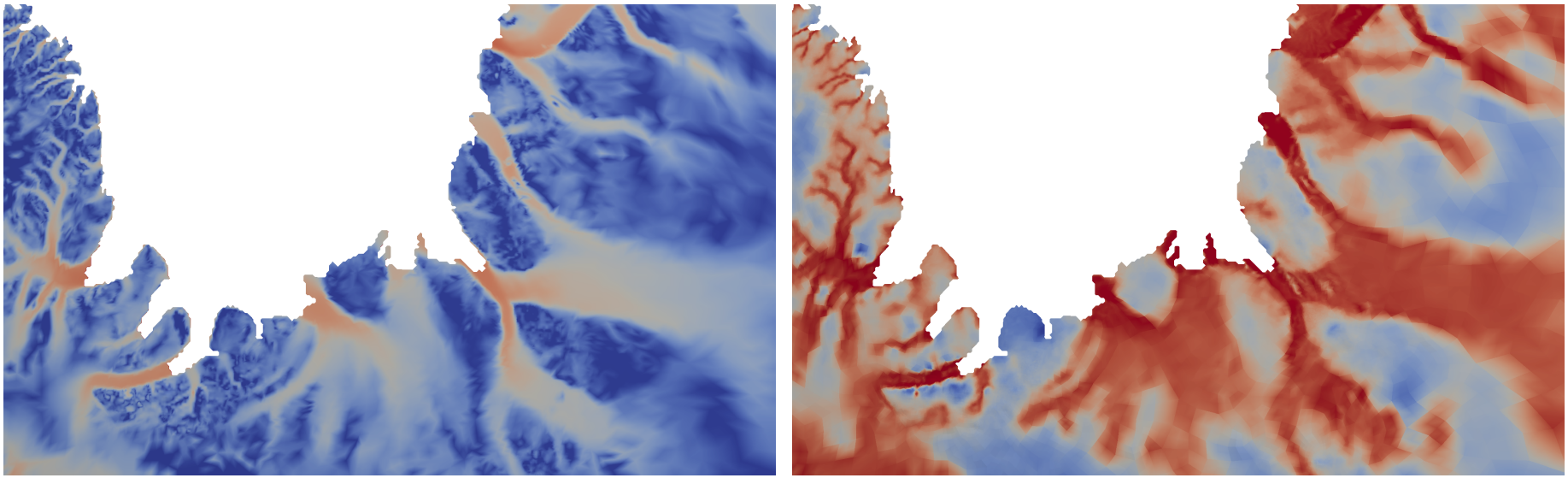
Reconstructed ice surface velocity field

Antarctic ice sheet inversion for basal sliding field: InSAR data



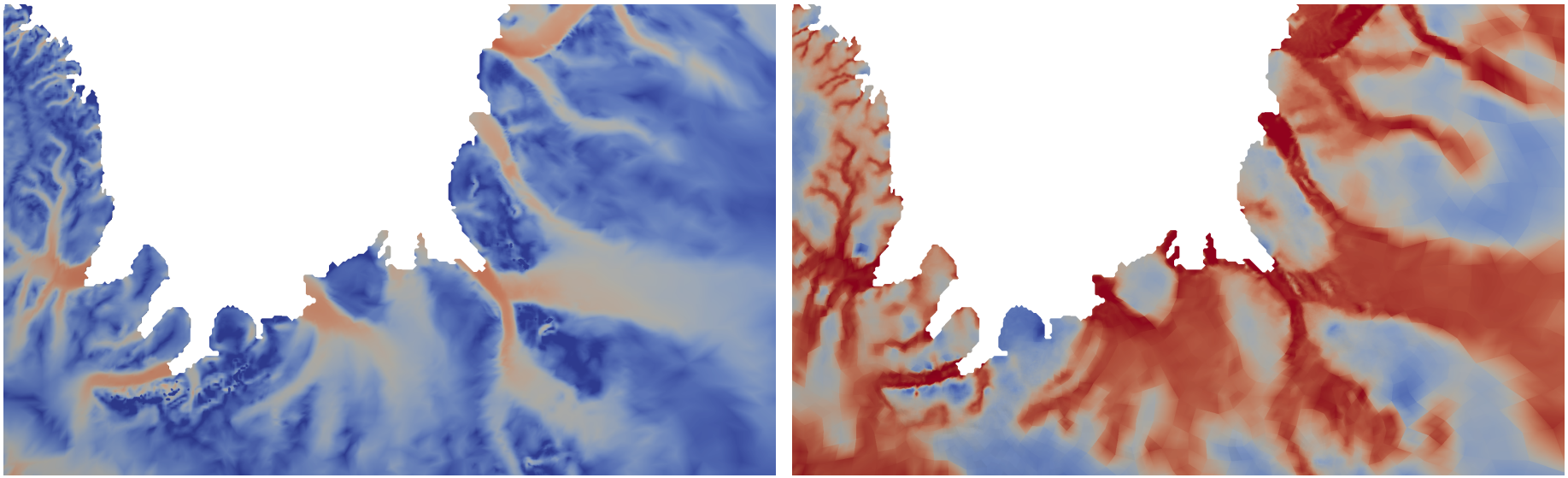
Error in velocity observations

Antarctic ice sheet inversion for basal sliding field: InSAR data (Ronne ice shelf region)



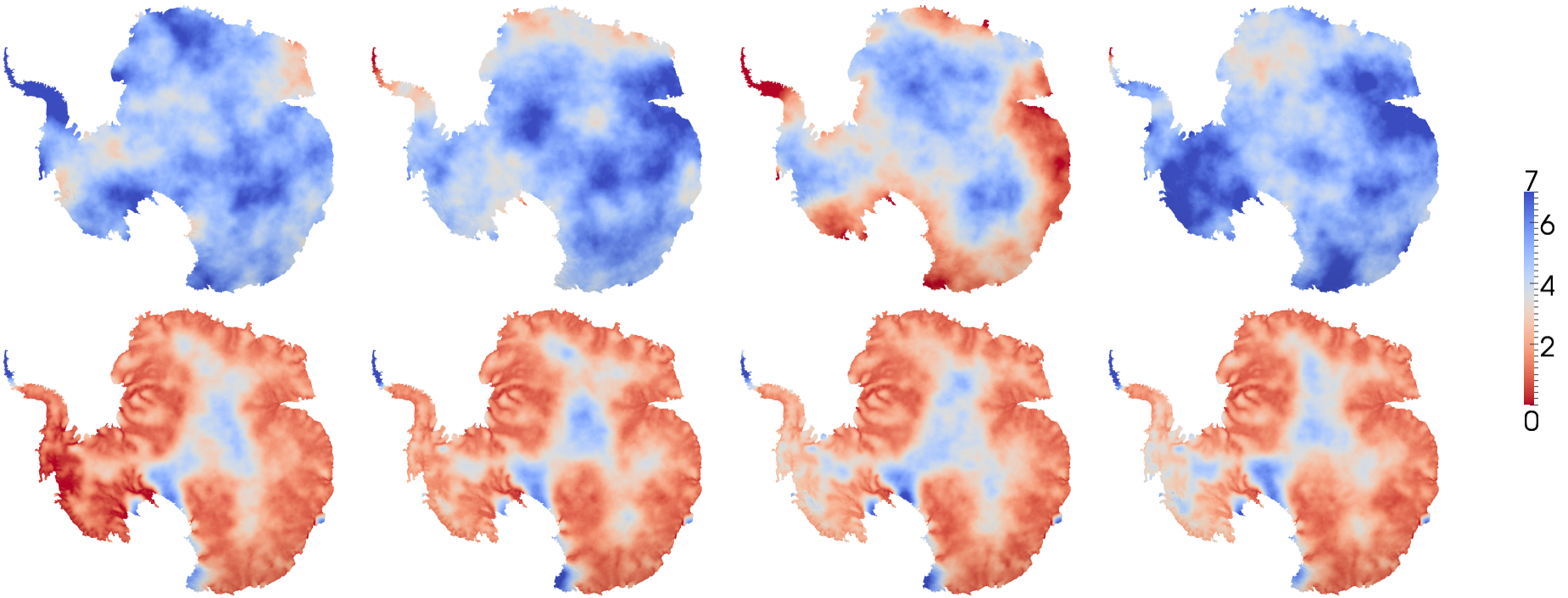
Left: InSAR-based Antarctica ice surface velocity observations
Right: Inferred basal sliding field

Antarctic ice sheet inversion for basal sliding field: InSAR data (Ronne ice shelf region)



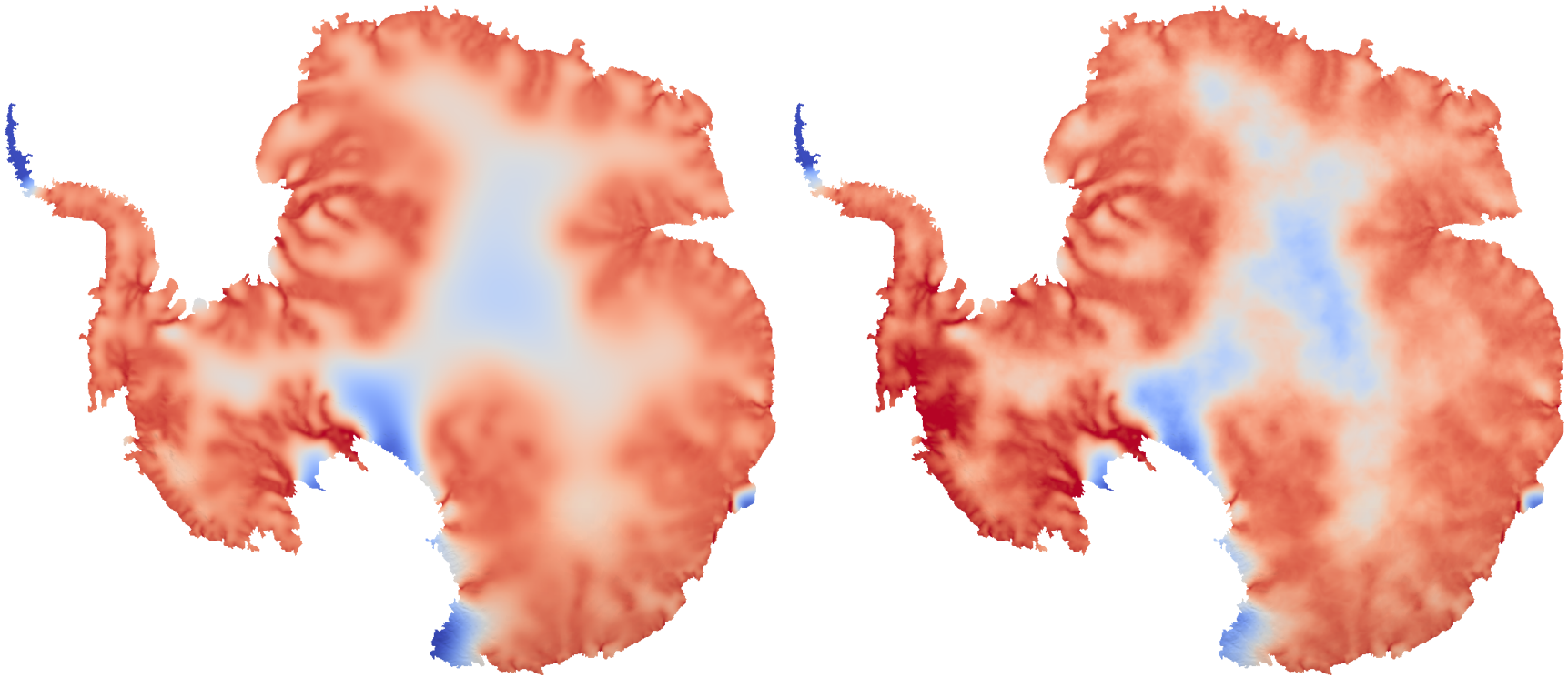
Left: Reconstructed ice surface velocity observations
Right: Inferred basal sliding field

Samples from prior and (Gaussianized) posterior



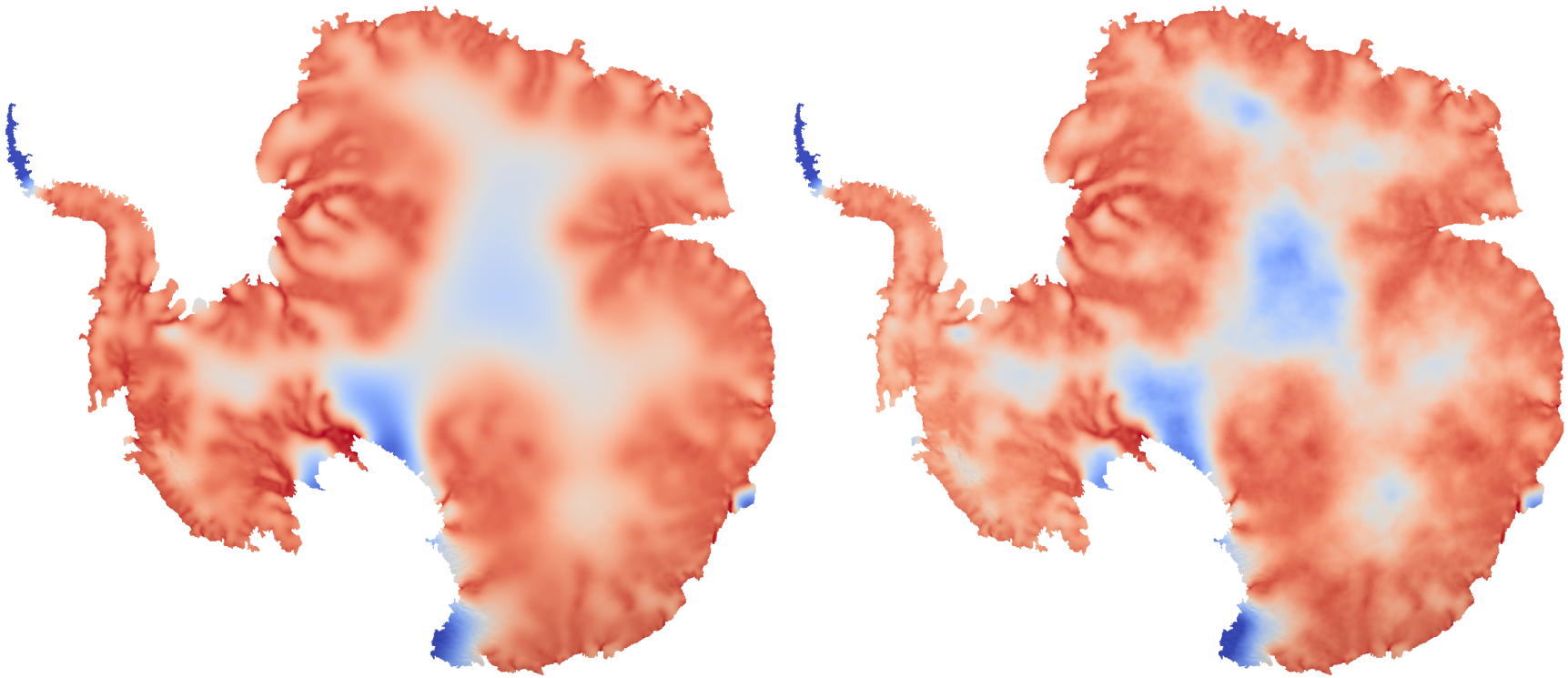
- Top: samples from prior
- Bottom: samples from posterior
- Difference between top and bottom reflects information gained (and uncertainty reduction) from data

Posterior mean and samples for Antarctic inversion



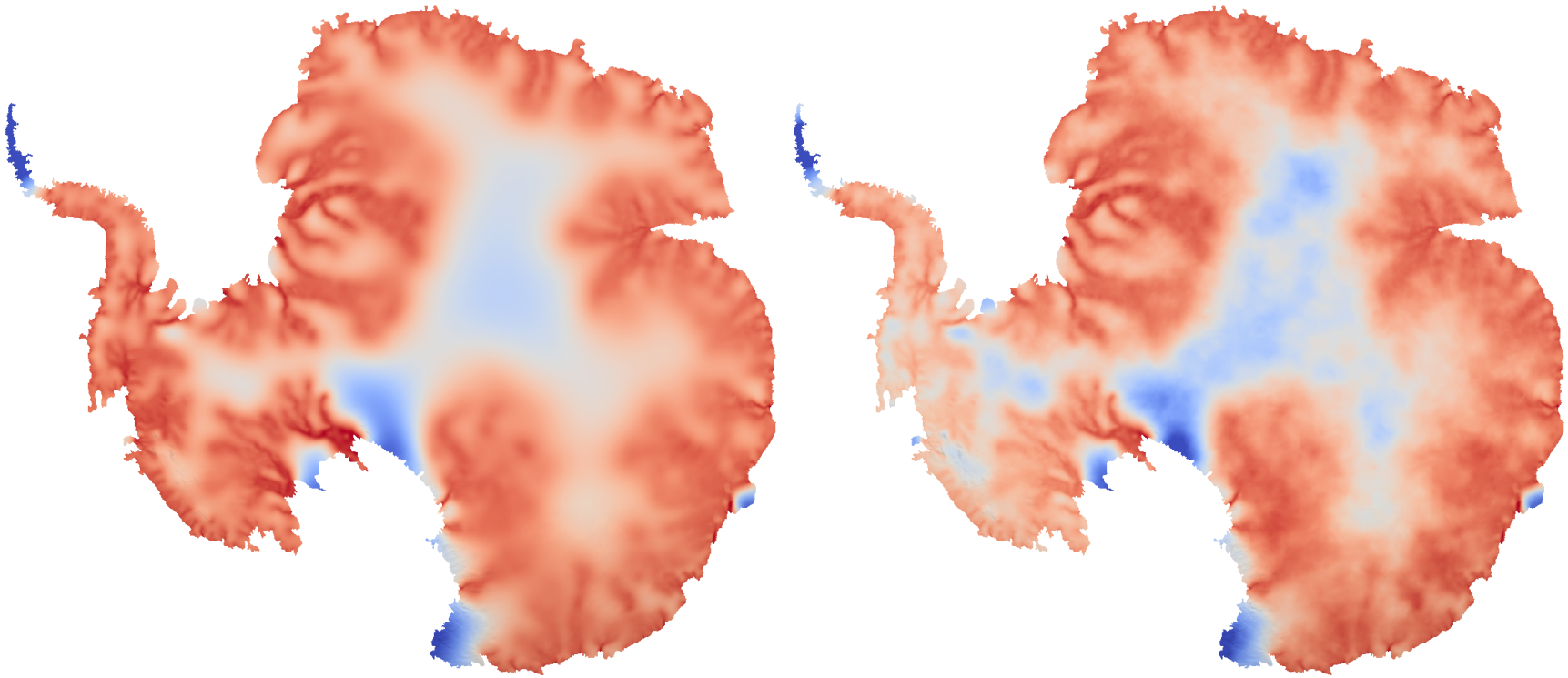
mean (left); sample from the posterior pdf (right)

Posterior mean and samples for Antarctic inversion



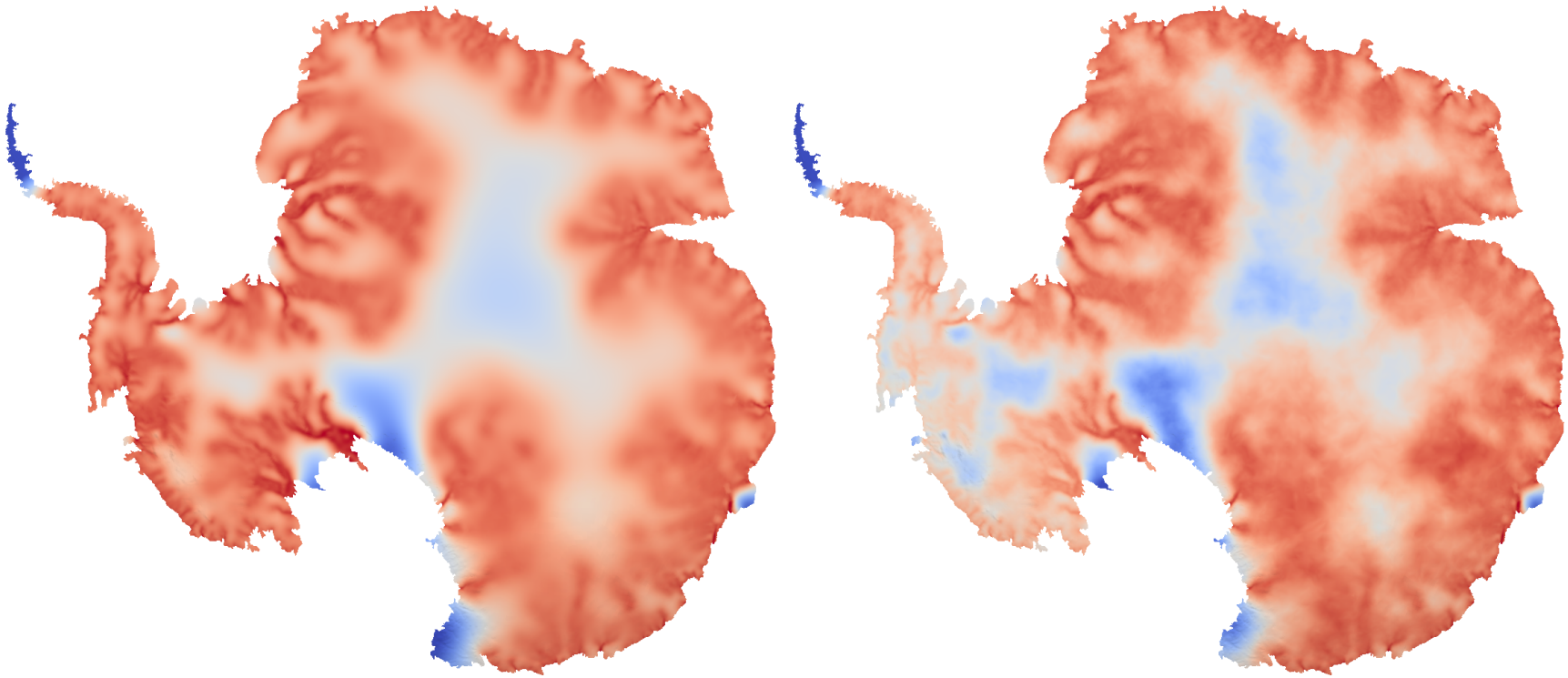
mean (left); sample from the posterior pdf (right)

Posterior mean and samples for Antarctic inversion



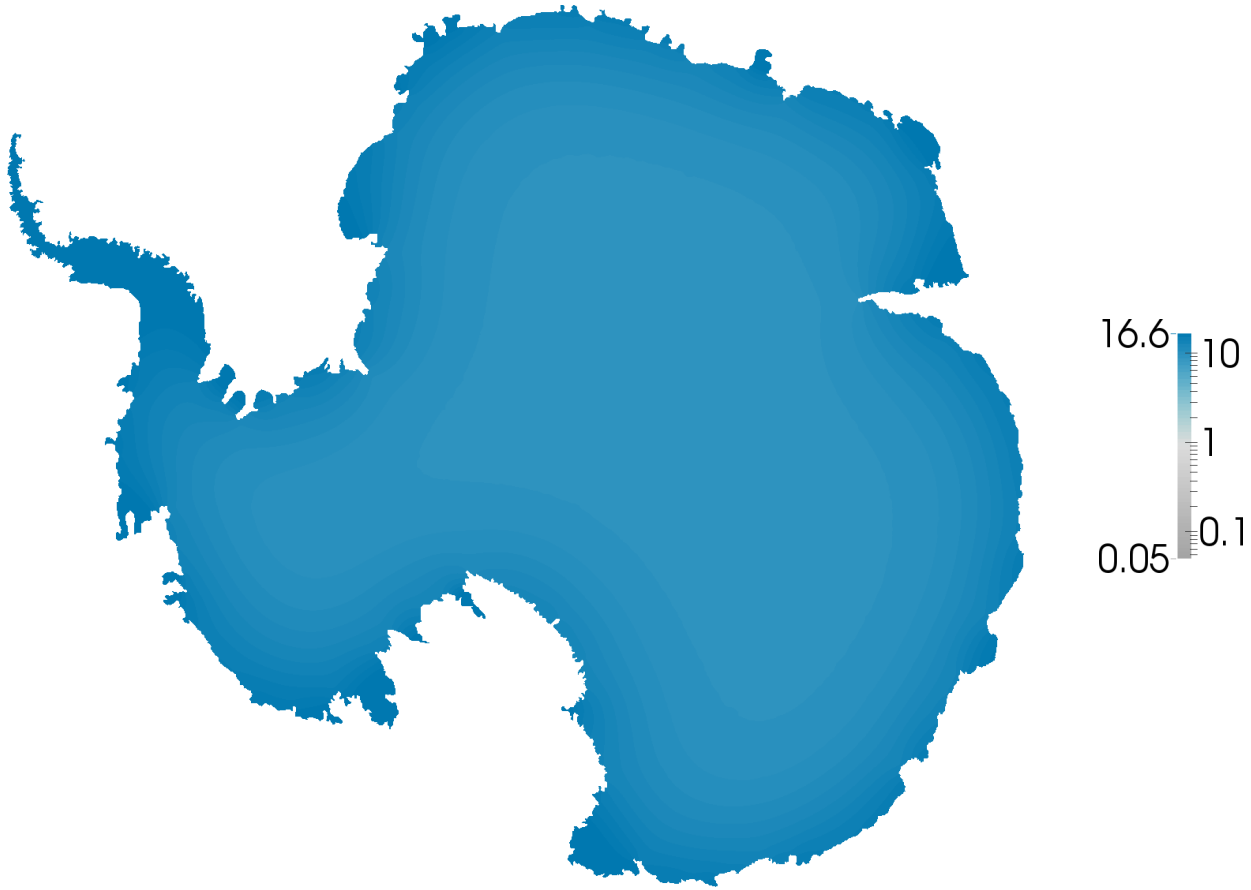
mean (left); sample from the posterior pdf (right)

Posterior mean and samples for Antarctic inversion



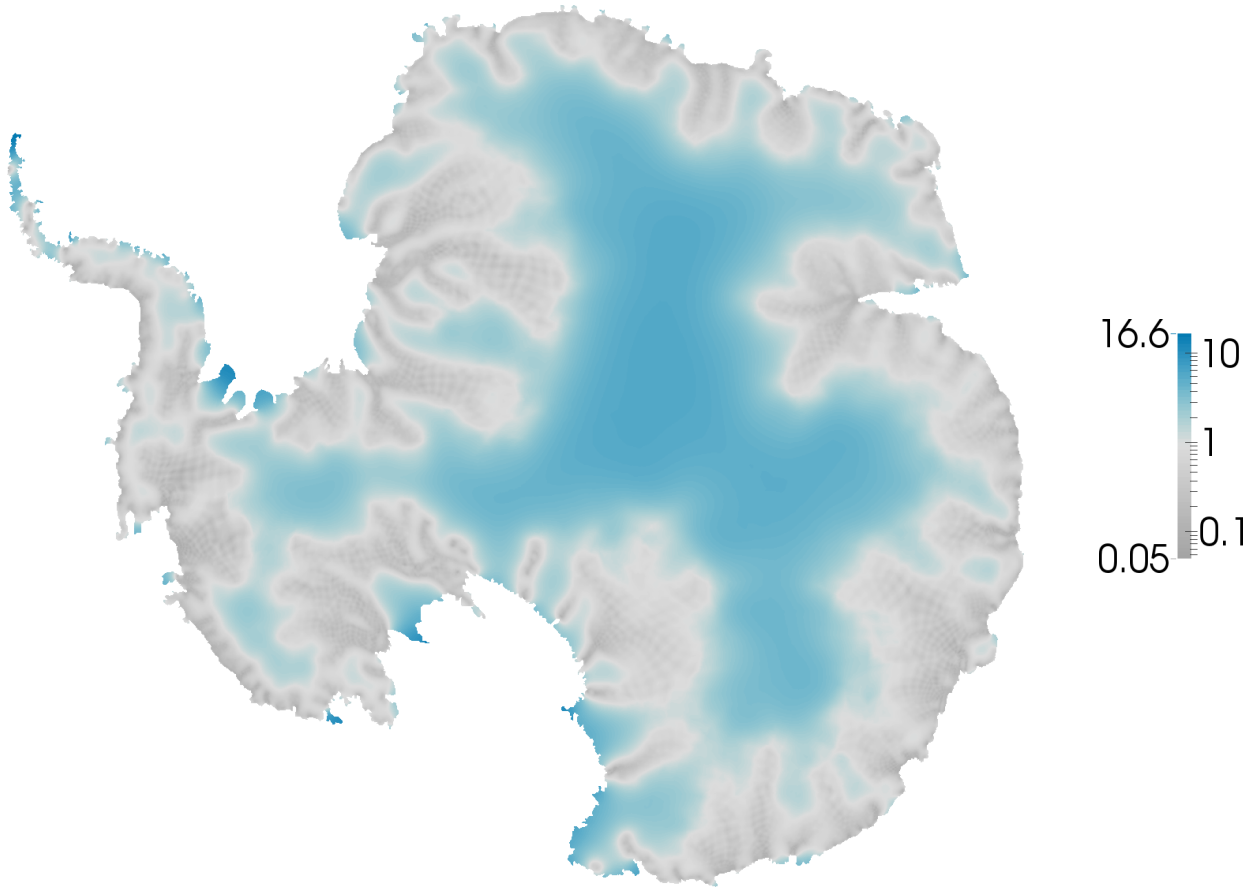
mean (left); sample from the posterior pdf (right)

Prior and posterior standard deviation of parameter field



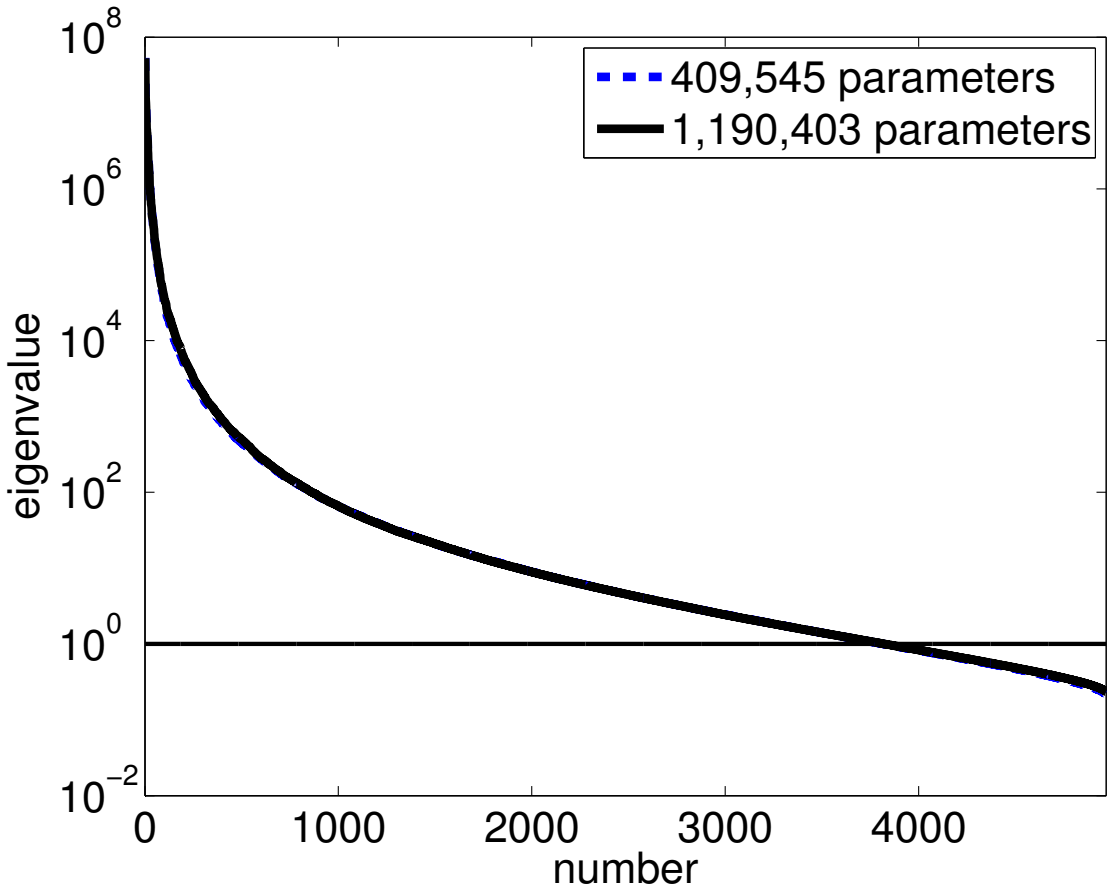
Prior standard deviation of log basal sliding parameter

Prior and posterior standard deviation of parameter field



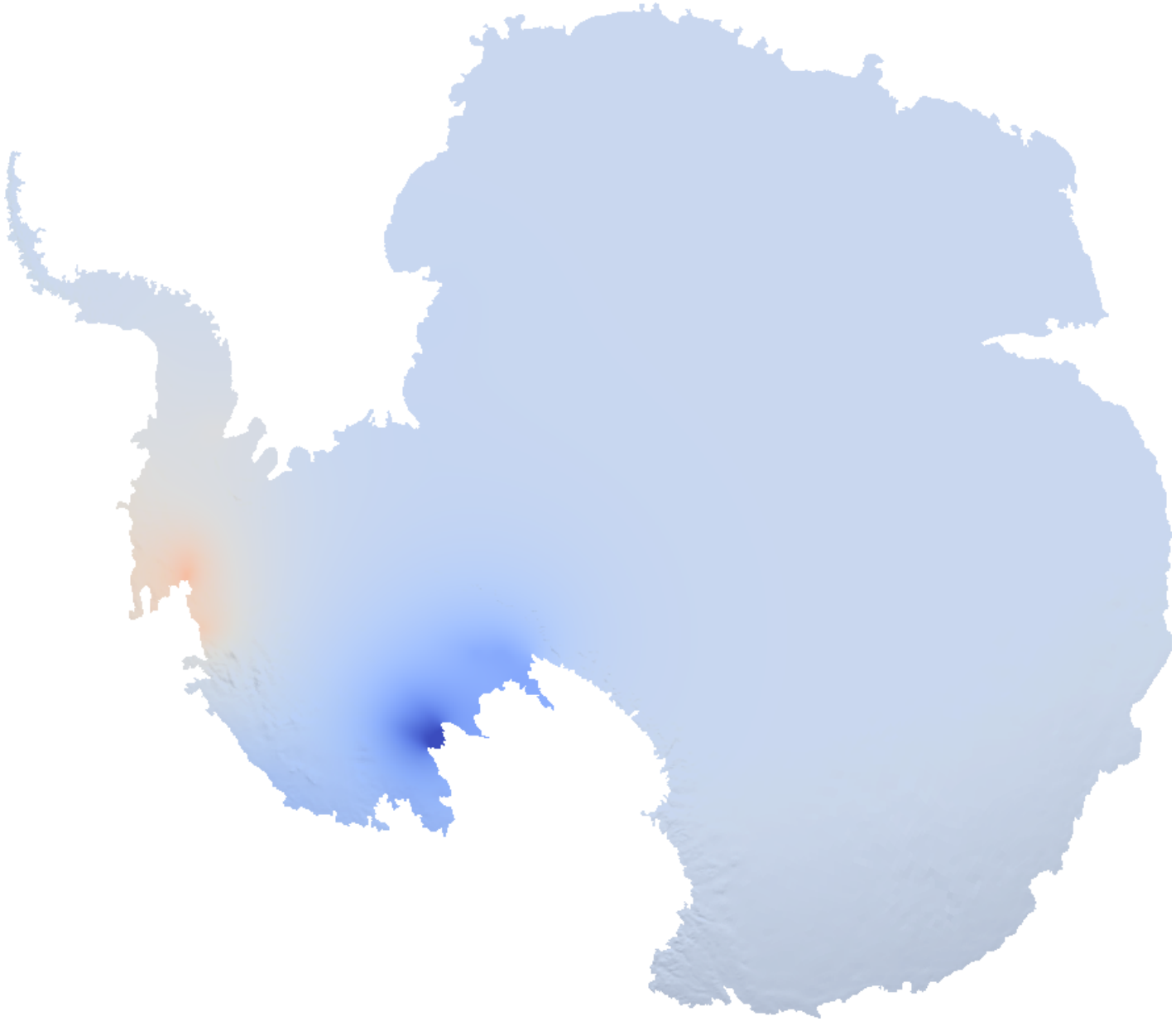
Posterior standard deviation of log basal sliding parameter

Spectrum of the prior-preconditioned data misfit Hessian

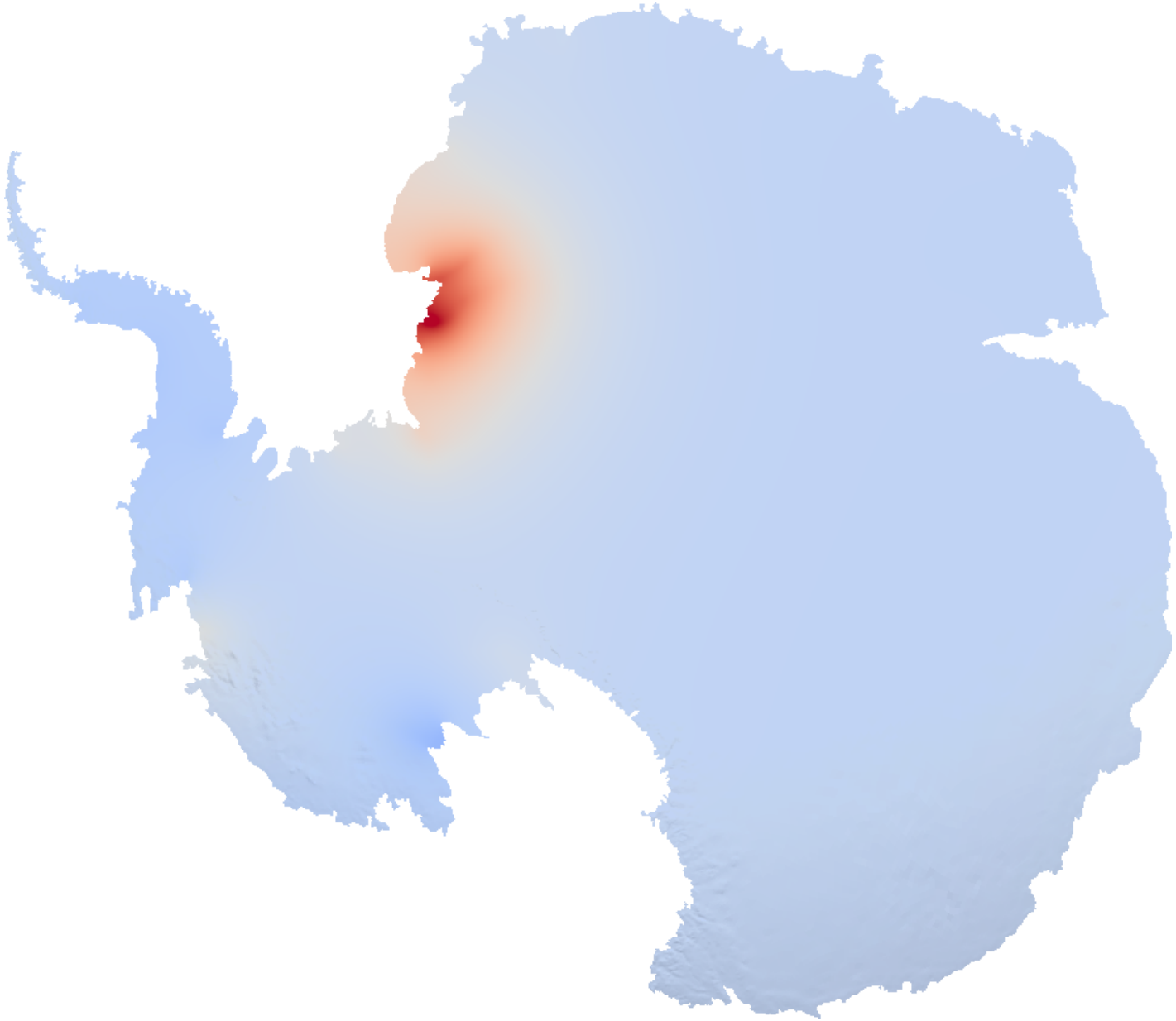


- Spectrum of $\boldsymbol{\Gamma}_{\text{pr}}^{1/2} \boldsymbol{F}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{F} \boldsymbol{\Gamma}_{\text{pr}}^{1/2}$ for Antarctica inverse problem with 410K and 1.19M basal sliding parameters (observed to decay like i^{-3})
- 4000 dominant modes, independent of parameter and data dimension
- intrinsic problem dimension depends on information content of data

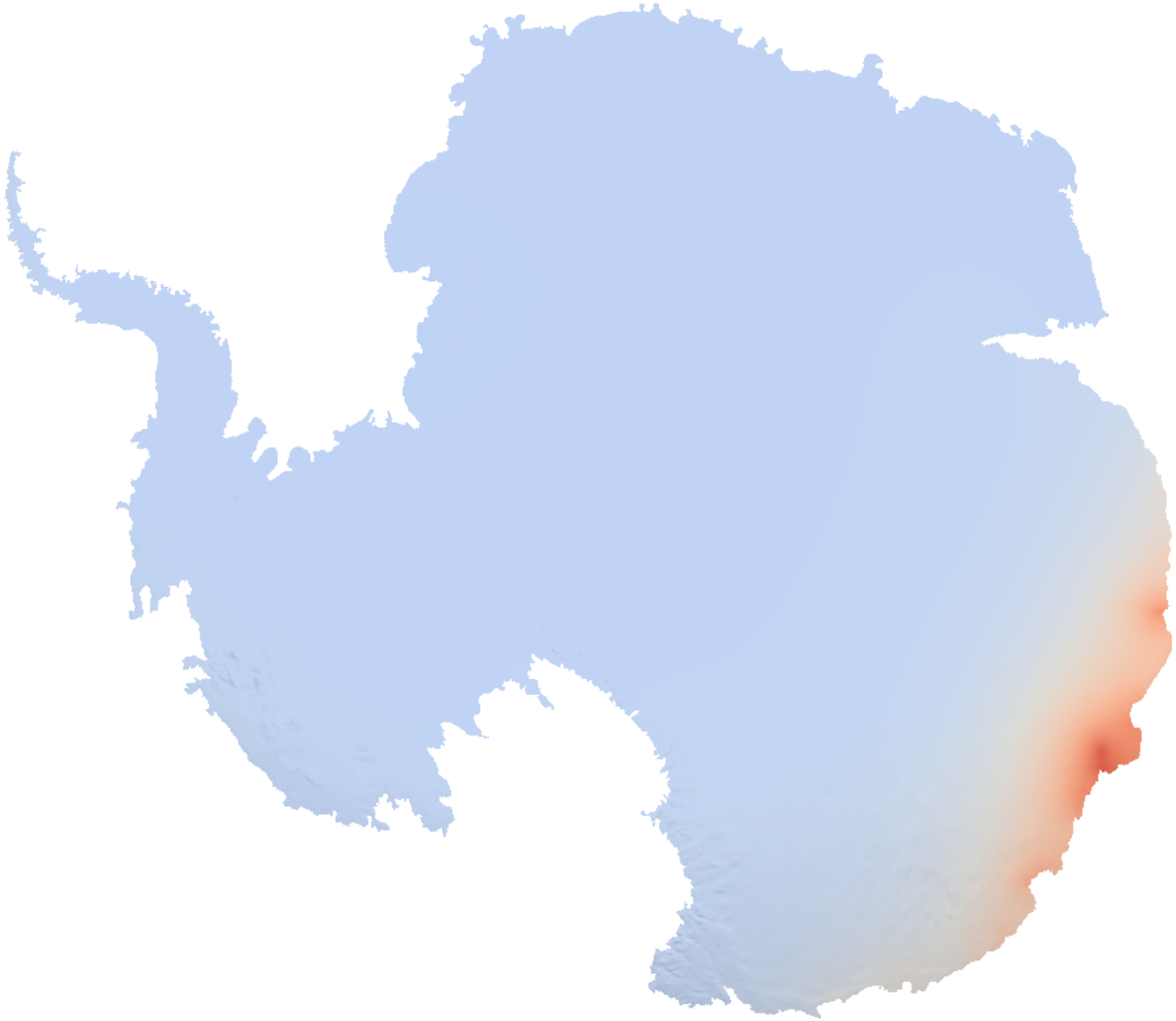
Eigenvector 1 of prior-preconditioned data misfit Hessian



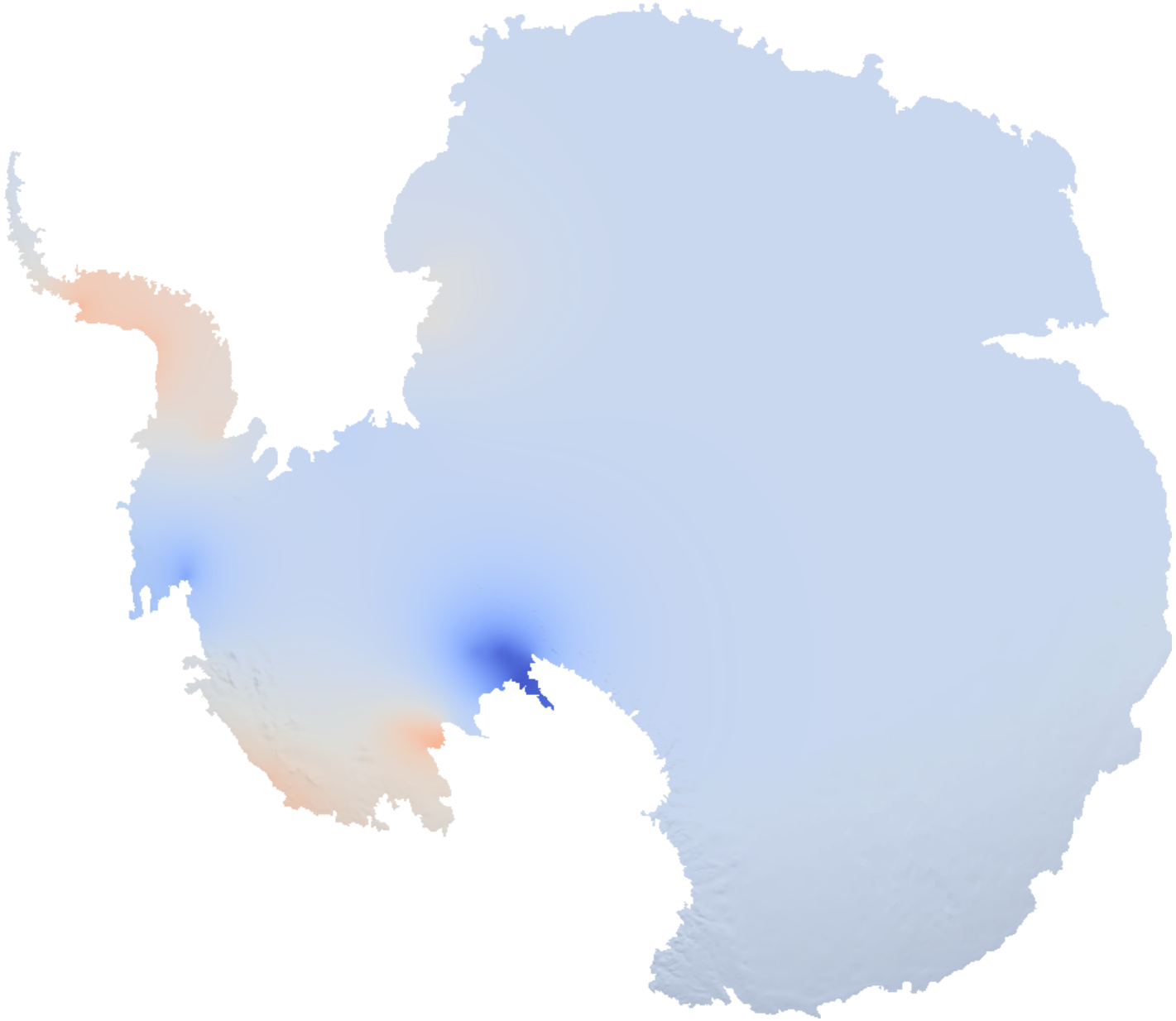
Eigenvector 2 of prior-preconditioned data misfit Hessian



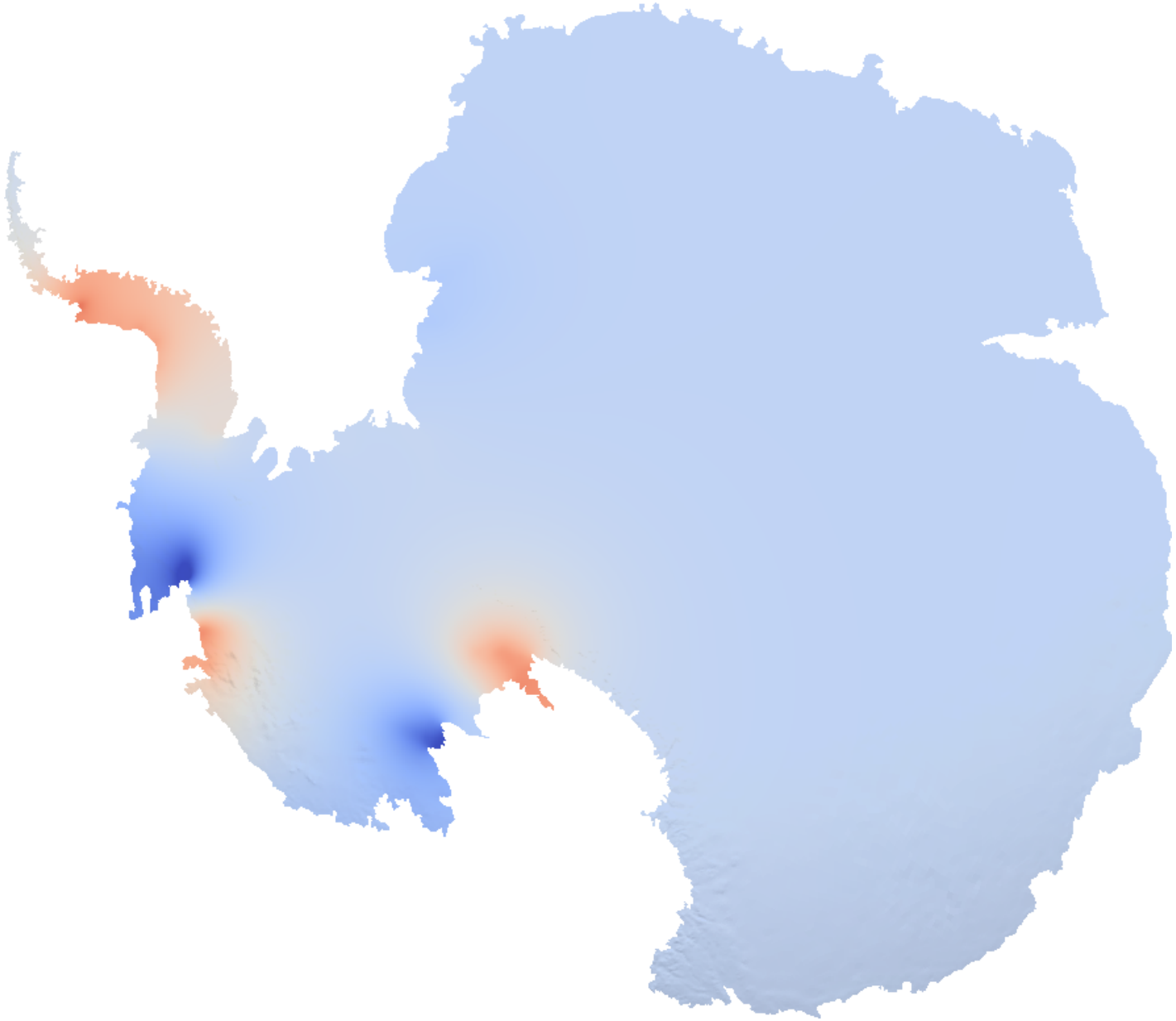
Eigenvector 3 of prior-preconditioned data misfit Hessian



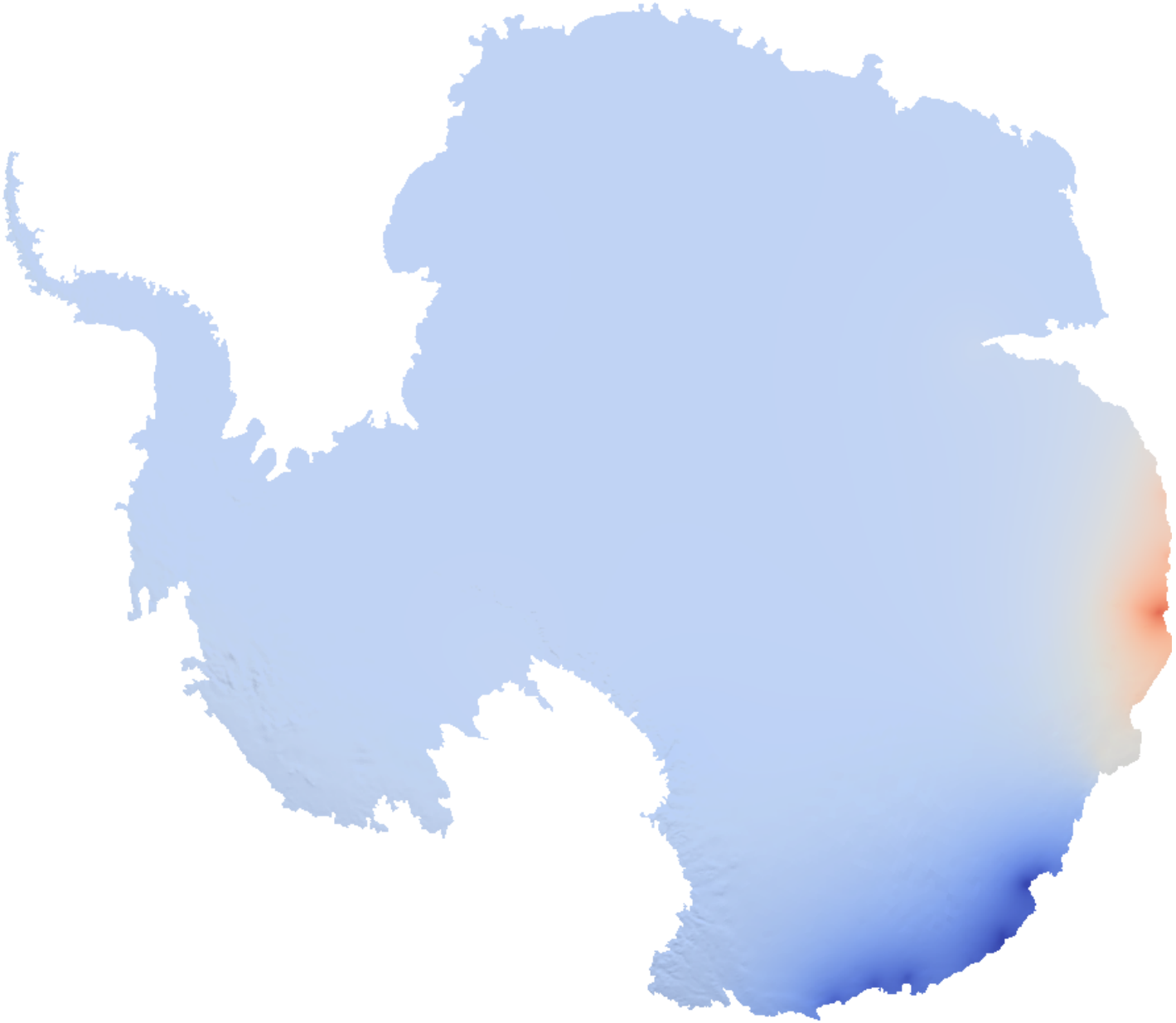
Eigenvector 4 of prior-preconditioned data misfit Hessian



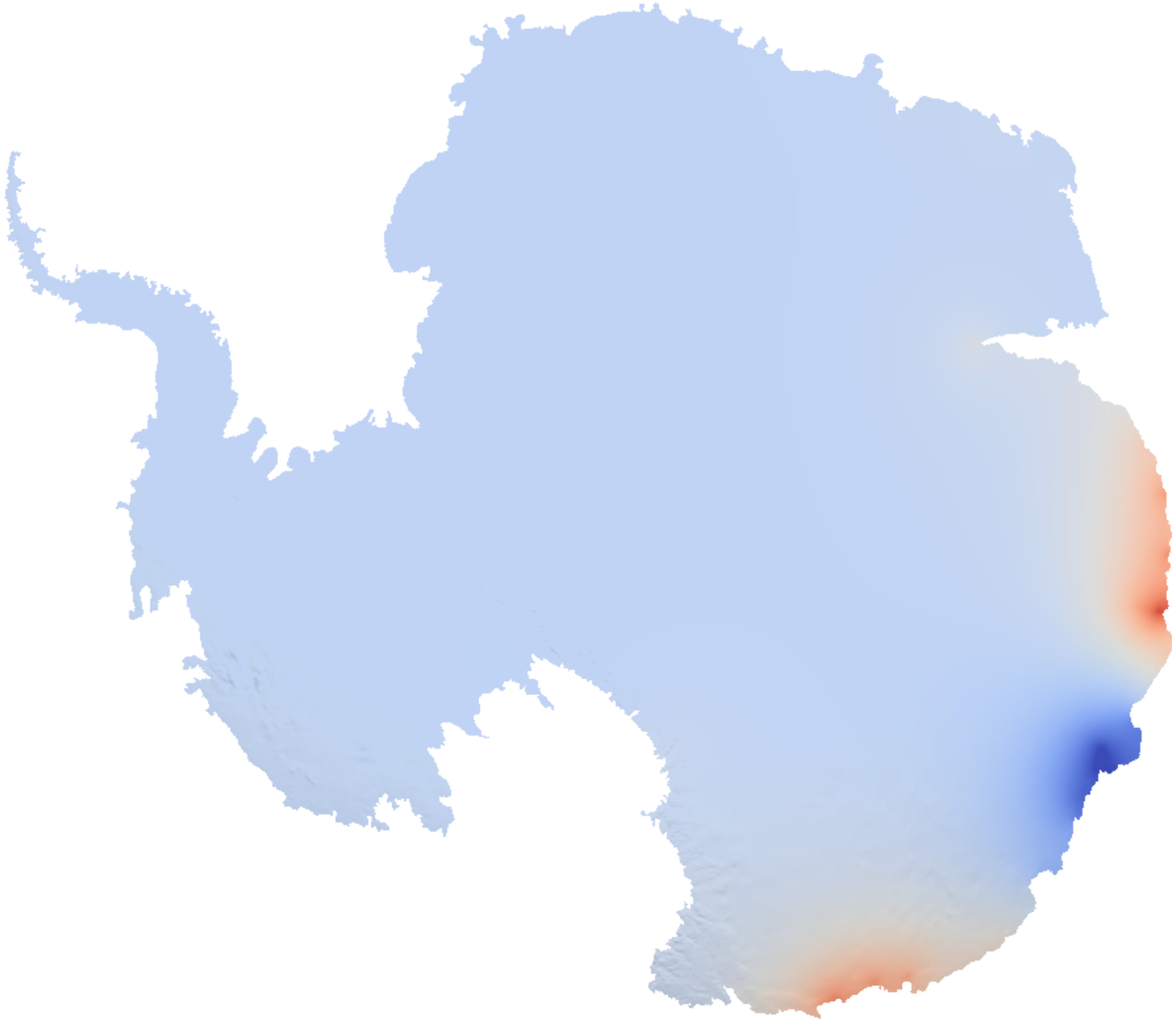
Eigenvector 5 of prior-preconditioned data misfit Hessian



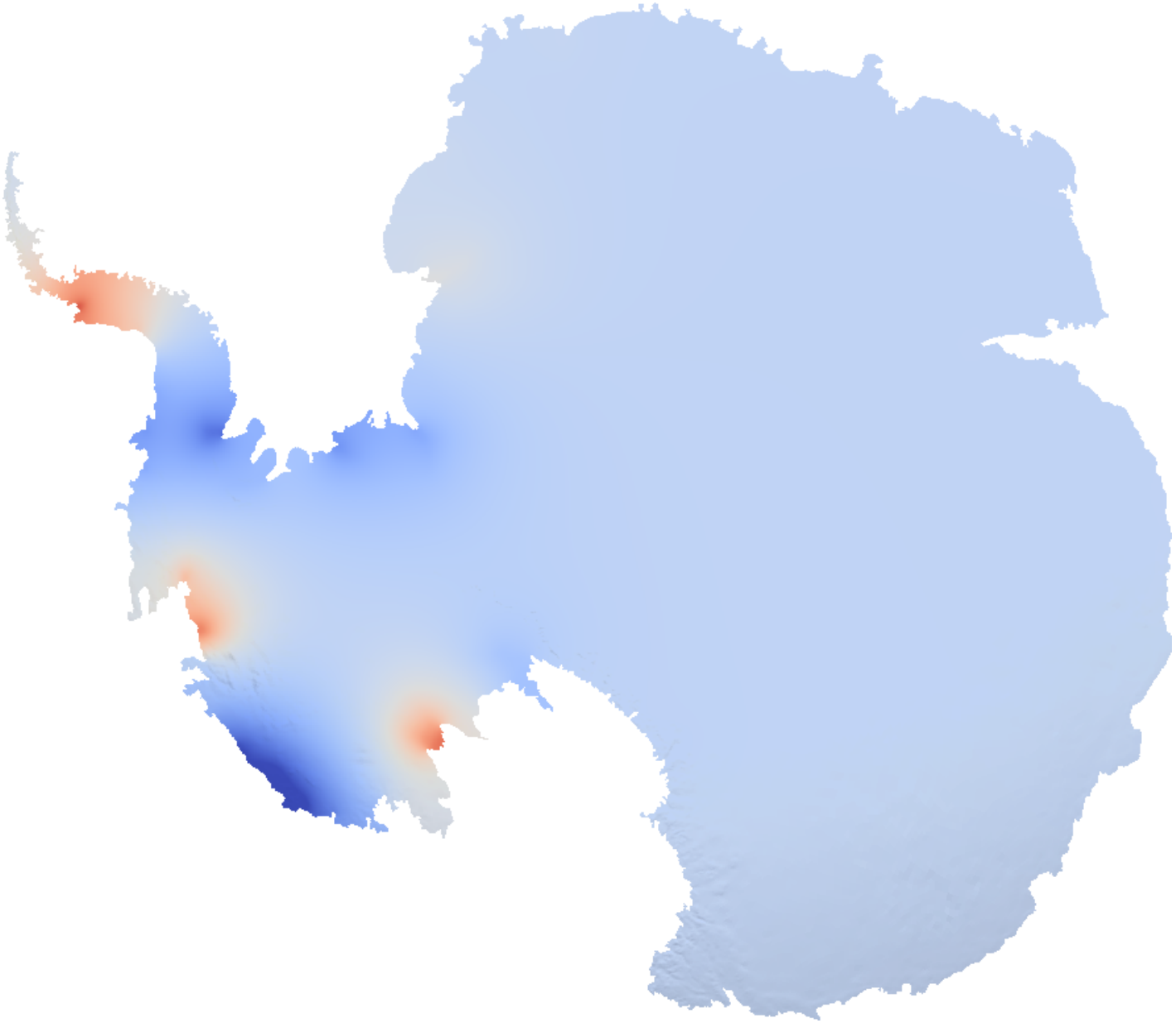
Eigenvector 6 of prior-preconditioned data misfit Hessian



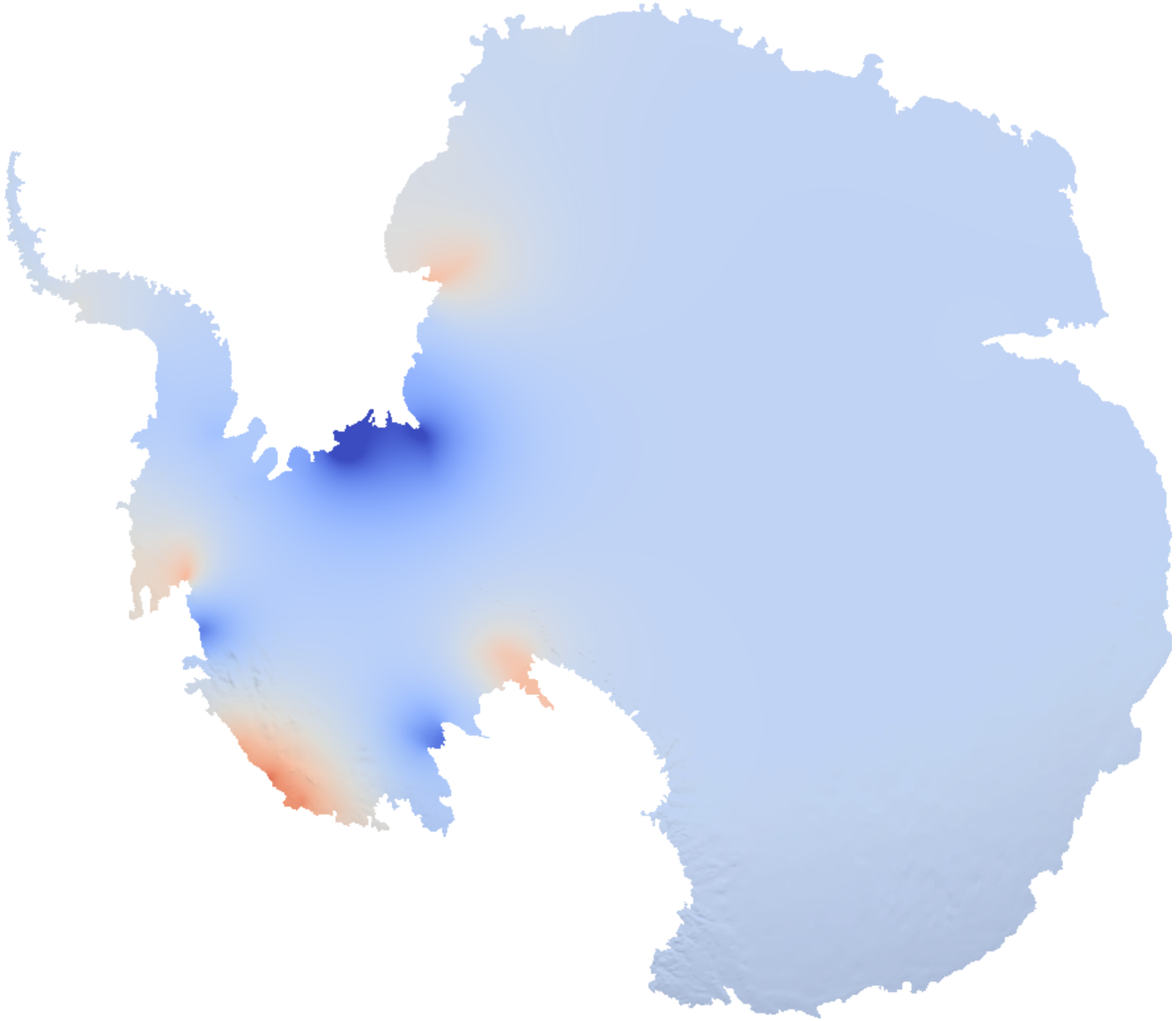
Eigenvector 7 of prior-preconditioned data misfit Hessian



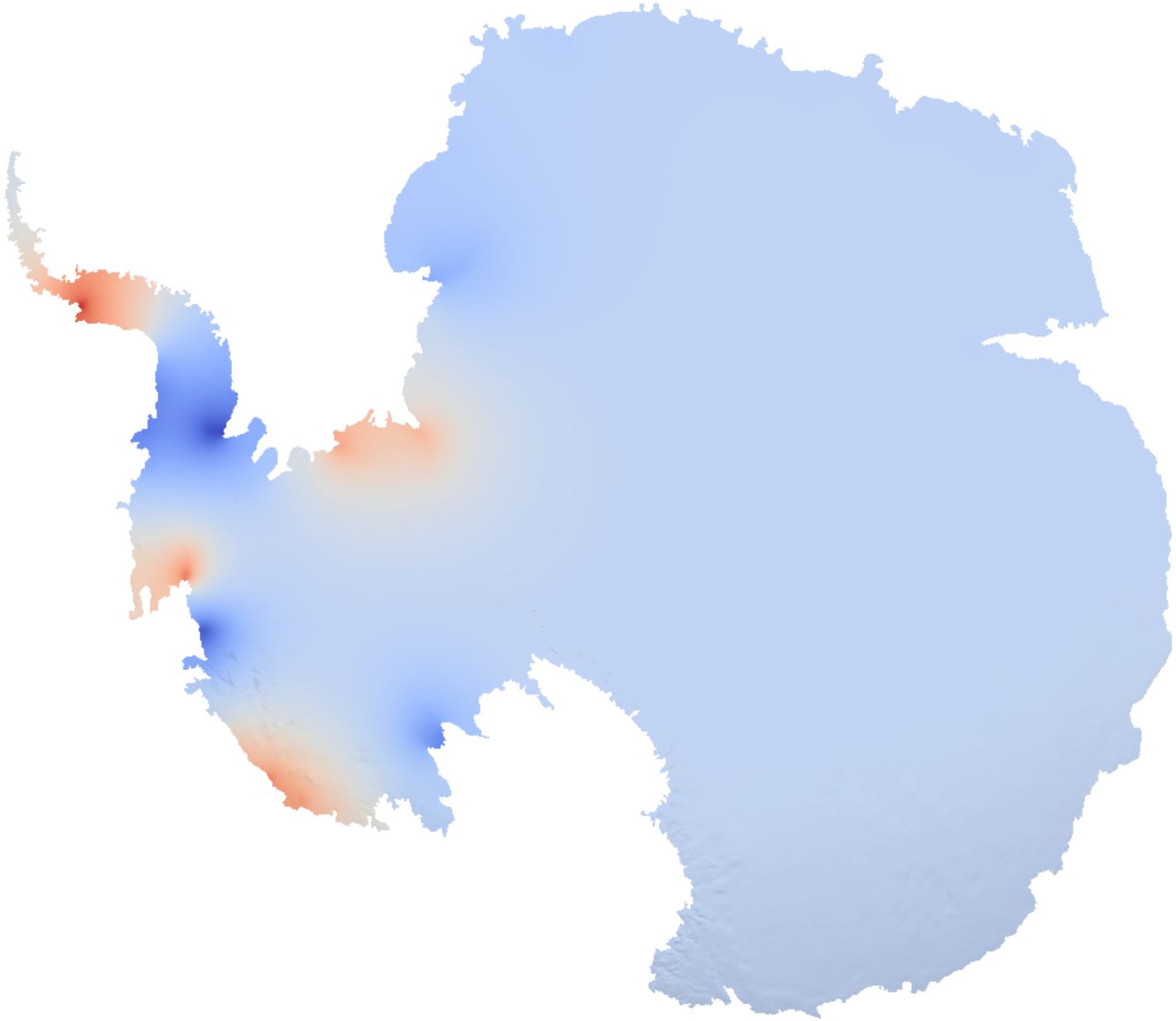
Eigenvector 8 of prior-preconditioned data misfit Hessian



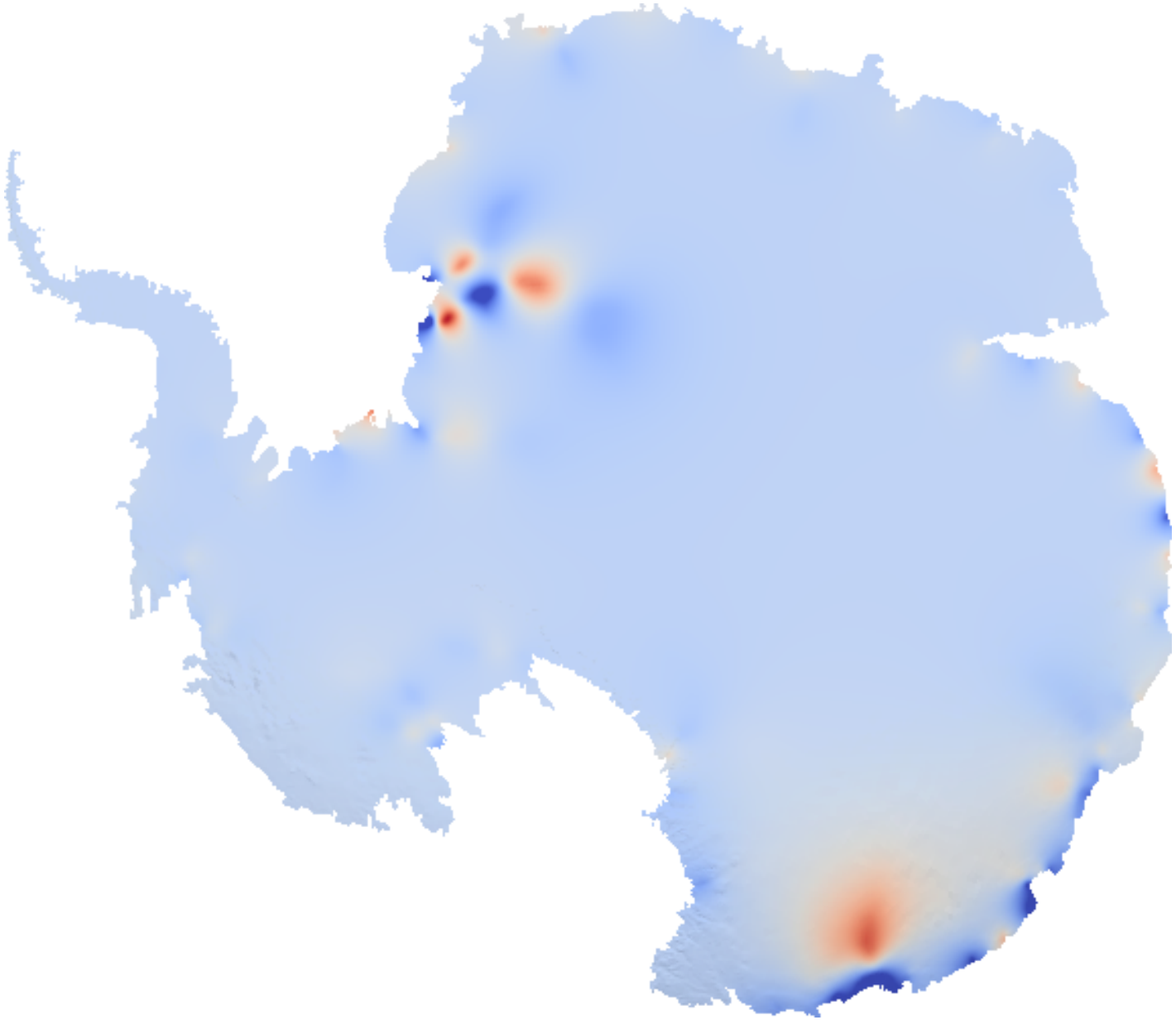
Eigenvector 9 of prior-preconditioned data misfit Hessian



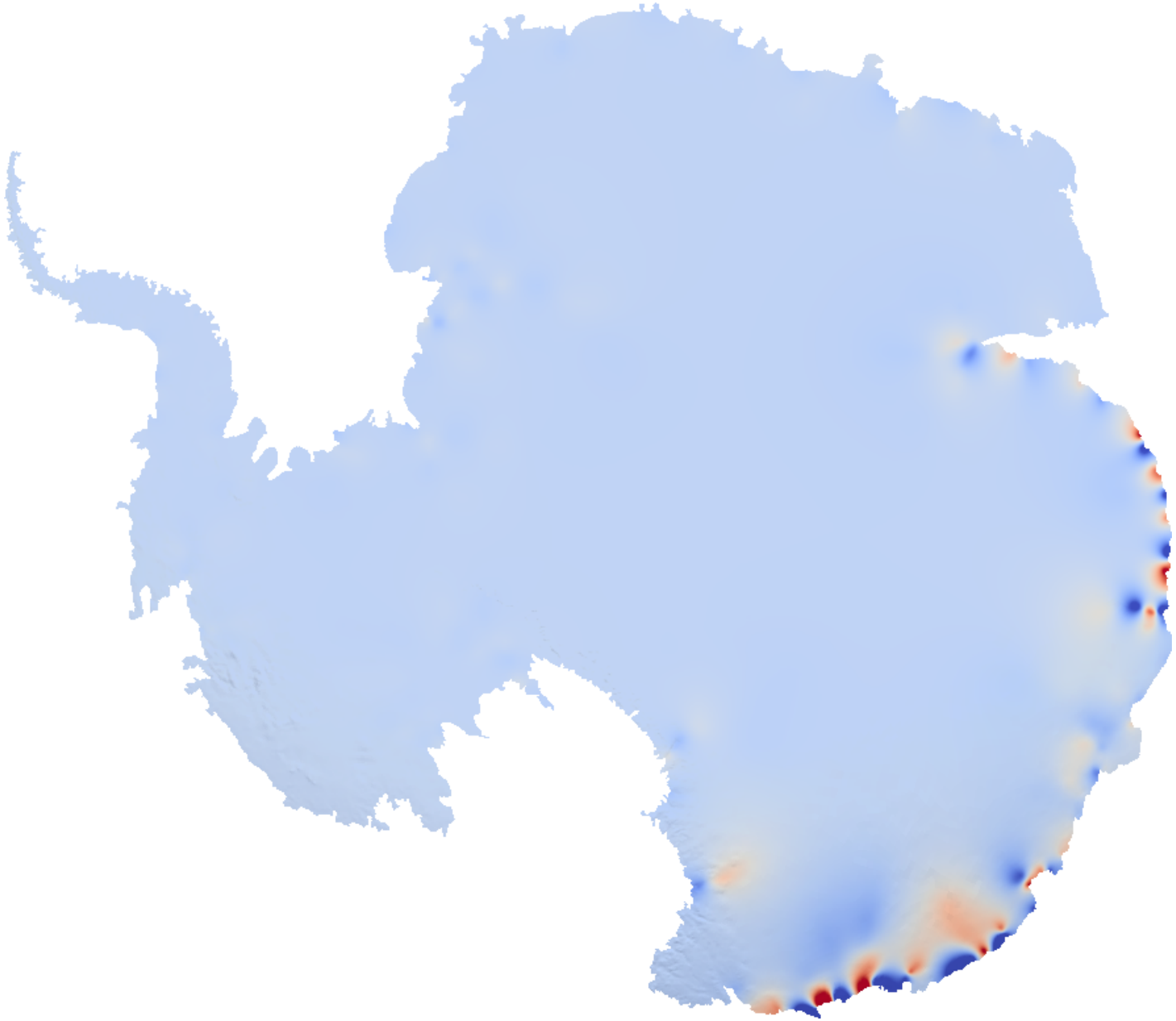
Eigenvector 10 of prior-preconditioned data misfit Hessian



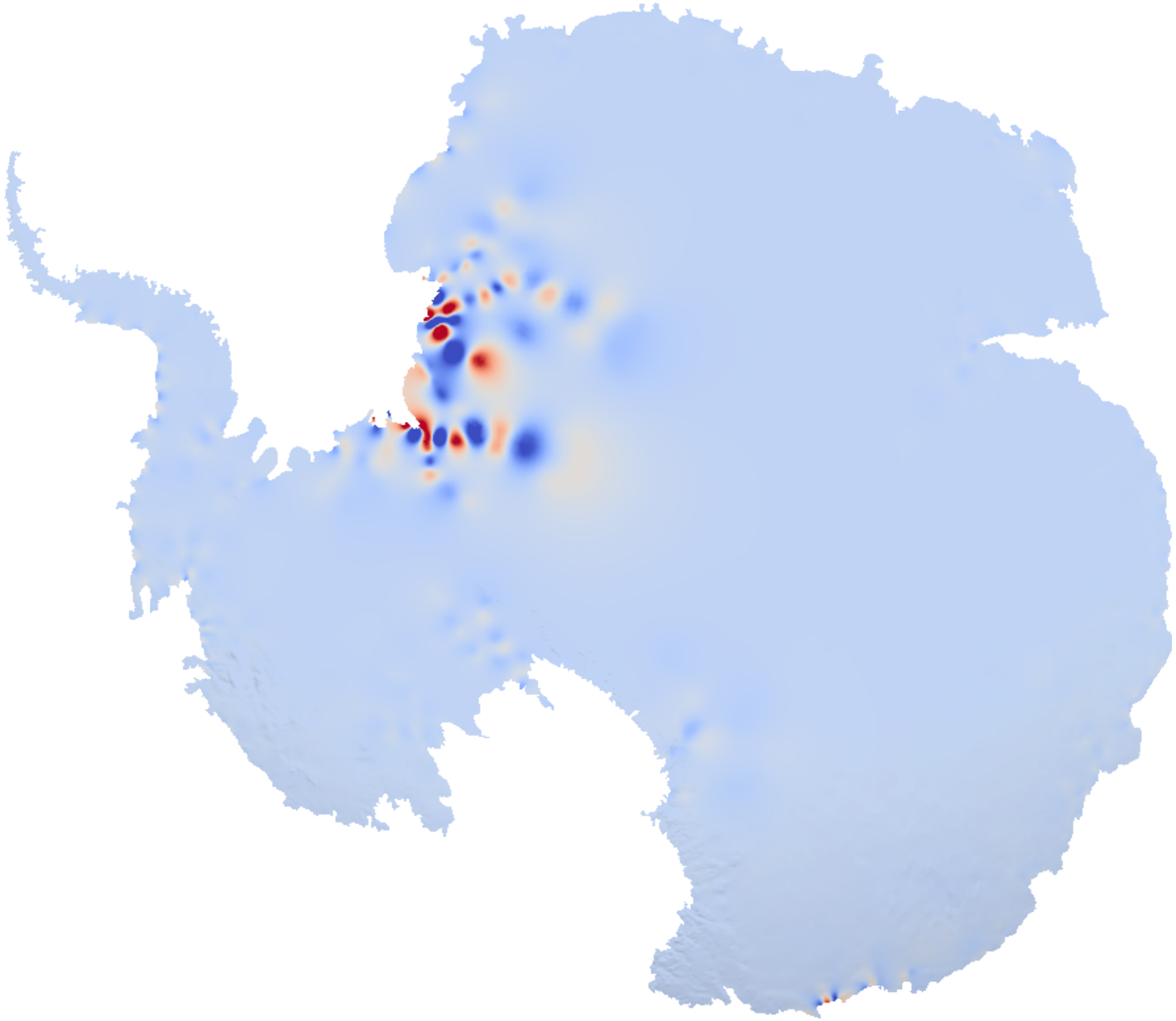
Eigenvector 100 of prior-preconditioned data misfit Hessian



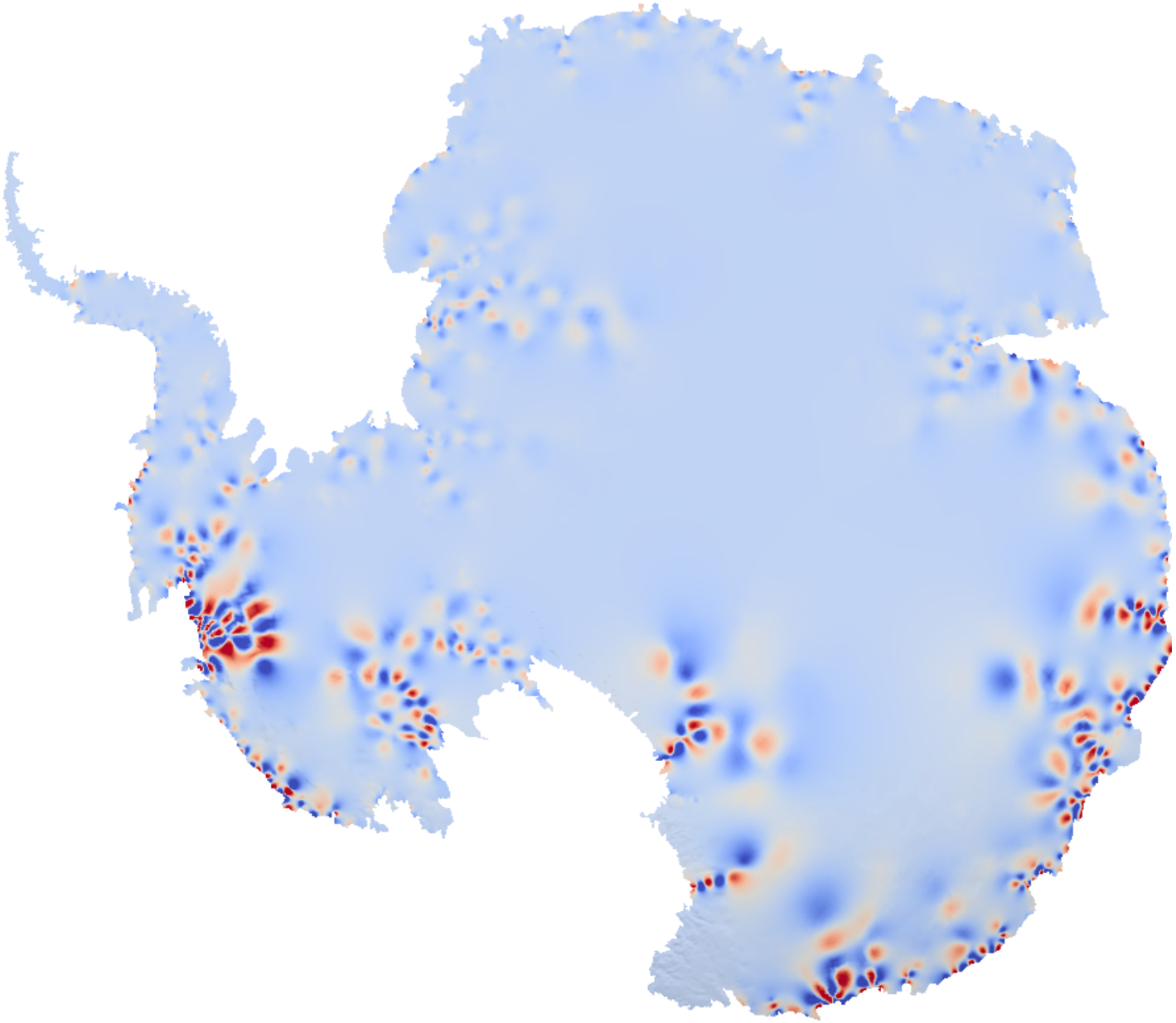
Eigenvector 200 of prior-preconditioned data misfit Hessian



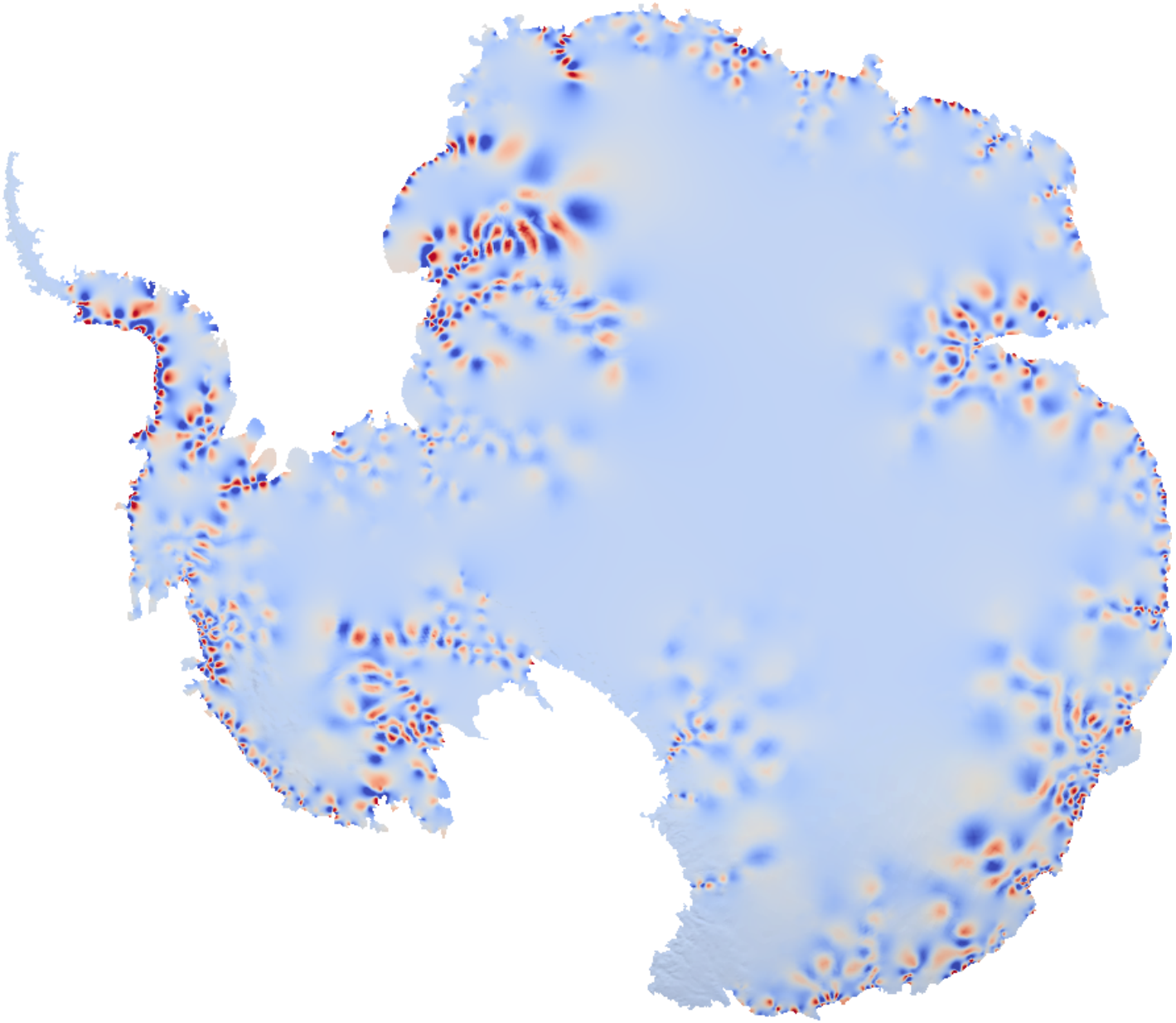
Eigenvector 500 of prior-preconditioned data misfit Hessian



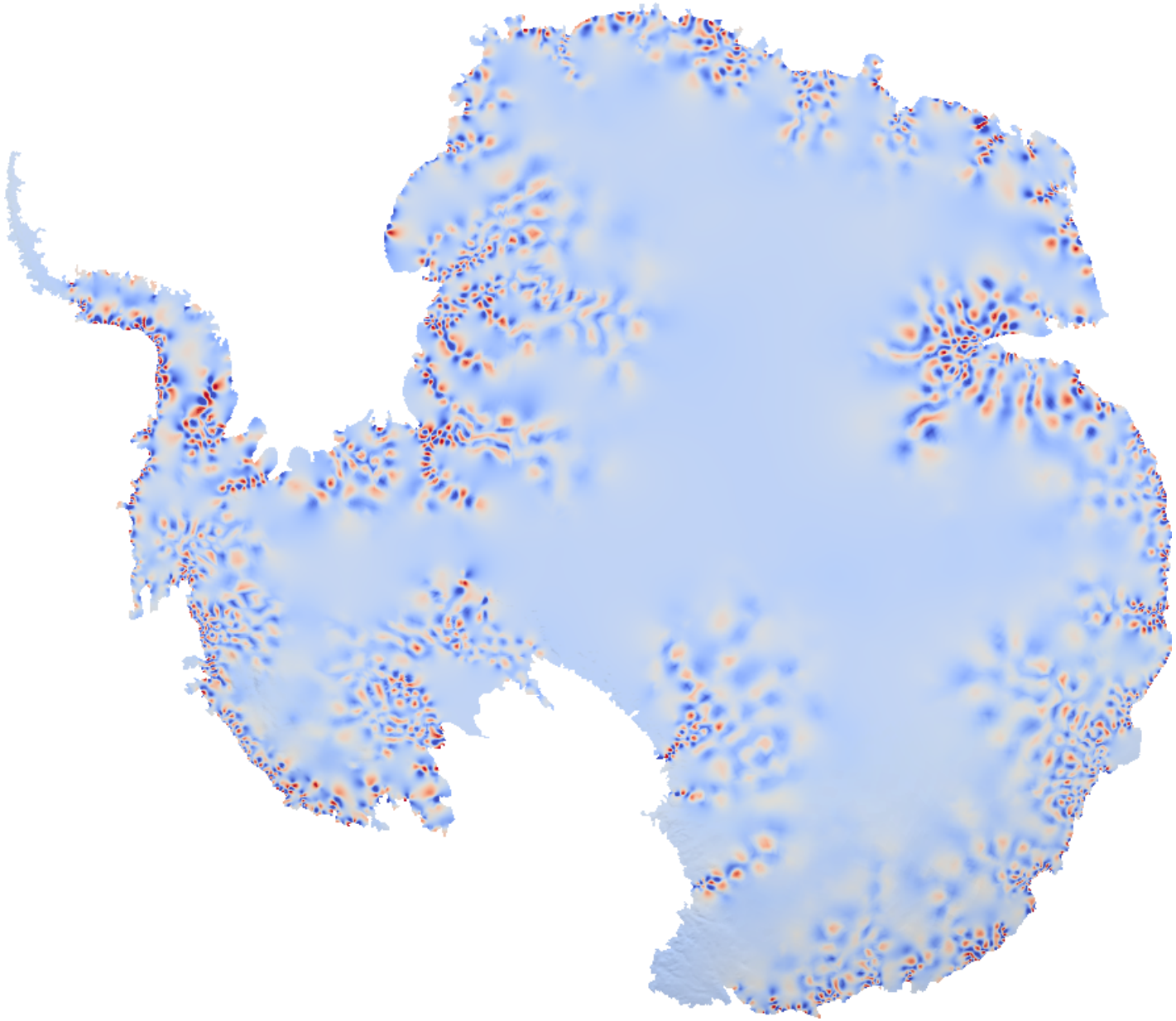
E-vector 1000 of prior-preconditioned data misfit Hessian



E-vector 2000 of prior-preconditioned data misfit Hessian



E-vector 4000 of prior-preconditioned data misfit Hessian



Summary

- Regularized inversions of surface velocity InSAR data for basal sliding field point to weaknesses that stretch deep into interior of Antarctic ice sheet
- But we need to quantify the uncertainty in these inversions, which leads to a large-scale Bayesian inverse problem governed by nonlinear, anisotropic, heterogeneous ice sheet flow model
- Linearization of the parameter-to-observable map leads to a Gaussianized approximation that builds on large-scale optimization algorithms
- Hessian manipulations can be made tractable by low-rank approximation of (prior-preconditioned) Hessian of the data misfit term, leading to several orders of magnitude effective dimensionality reduction
- Cost (in terms of forward/adjoint PDE solves) is independent of parameter dimension, data dimension, and state dimension
- Preliminary Gaussianized Bayesian inversions for Antarctica point to lower uncertainties in large velocity gradient regions
- Even with Gaussianized approximation of posterior and scalable algorithms, 117,578 linearized forward/adjoint Stokes solves are required

Open questions

- Can compactness of the Hessian of the data-misfit be proven for broader classes of problems?
- Can convergence of the discretized posterior to the infinite dimensional measure be proven for broader classes of problems?
- What to do when the data misfit Hessian does not admit a low-rank approximation?
- Freezing the Hessian at the MAP point vs. approximately recomputing the Hessian at each sample point
- What is the meaning of adaptive mesh refinement for inverse problems?
- How to be exploit hierarchies of available forward models (multifidelity methods)?
- How to best account for forward model error?
- Full MCMC sampling of Antarctic posterior
- Mechanistic sub-basal models? Glacial and sub-glacial hydrology?
- Two-way coupling with ocean models: continuum coupling?

Acknowledgments

- Many thanks to prize committee, past and present students and postdocs, collaborators, UT Austin colleagues, and SIAM Geosciences community
- Research supported by:
  - DOE grants DE-SC0009286 (MMICCs), DE-SC0010518 (ExaUQ), DE-SC0002710 (SciDAC)
  - NSF grants ARC-0941678 (CDI), CMMI-1028889 (CDI)
- Resources on ORNL Jaguar and Titan systems provided through ALCC award at ORNL Leadership Computing Facility
- Resources on TACC Lonestar, Longhorn, and Stampede systems provided through awards from TACC and XSEDE