## 12. Large-scale algorithms for Bayesian inversion, with application to flow of the Antarctic ice sheet (UT-Austin, NYU, UC Merced)

Research thrust: *Advanced methods for inference*
Research sub-thrusts: *Goal-oriented, multifidelity, and multiscale methods for inference; Randomized approaches for inference*

The question of how one infers unknown parameters characterizing a given model of a physical system from observations of the outputs of that model is fundamentally an inverse problem. In order to address the intrinsic ill-posedness of many inverse problems, regularization is often invoked to render the inverse solution unique. The Bayesian formulation of the inverse problem, on the other hand, seeks to infer the probability of all models that are consistent with the data and any prior knowledge. The central challenge in Bayesian inversion is to explore this "posterior" probability, which is the solution of the inverse problem. This task is intractable when the parameter dimension is large and the forward problem is expensive to solve. We have developed efficient and scalable algorithms for Laplace approximation of large-scale/high-dimensional Bayesian inverse problems, and instantiated them in the context of modeling the flow of the Antarctic ice sheet and its effect on loss of grounded ice to the ocean [62, 63, 114, 147]. The ice is modeled as a viscous, incompressible, creeping, shear-thinning fluid. The observational data come from satellite measurements of surface ice flow velocity, and the uncertain parameter field to be inferred is the basal friction, represented by a heterogeneous coefficient in a Robin boundary condition at the base of the ice sheet. The prediction quantity of interest is the present-day ice mass flux from the Antarctic continent to the ocean. We have shown that the work required to solve this inverse problem—measured by number of (linearized) forward and adjoint ice sheet model solves—is independent of the state dimension, parameter dimension, and data dimension. The key to achieving this dimension independence is to exploit the fact that, despite their large size, the observational data typically provide only sparse information on model parameters. This property can be exploited to construct a low rank approximation of the linearized parameter-to-observable map via randomized SVD methods and adjoint-based actions of Hessians of the data misfit functional. Fig. 16 illustrates the Bayesian inverse solution for 1.19 million parameters. The low-rank approximation of the parameter-to-observable map resulted in effective parameter dimensionality reduction of over two orders of magnitude (from 1.19M to 4000). The combination of very high parameter dimension with highly nonlinear forward problem results in the most challenging Bayesian inversion problem ever solved, to the best of our knowledge.
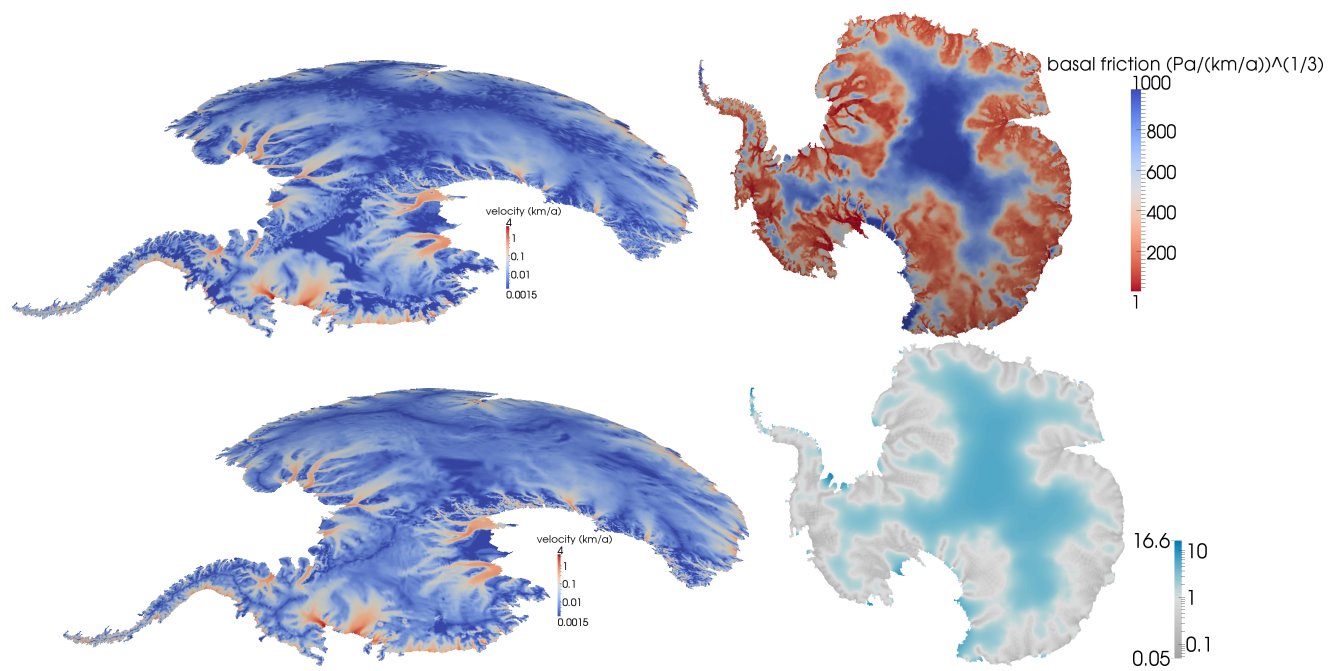
Figure 16: Top left: InSAR-based surface velocity observations of flow of Antarctic ice sheet. Top right: Inferred mean of basal friction field. Bottom left: Reconstructed ice surface velocity field (based on inferred basal friction field). Bottom right: Inferred uncertainty in basal friction field (standard deviation of Gaussianized posterior of log basal friction).