13. Bayesian nonlinear optimal experimental design for systems governed by PDEs (UT Austin, MIT)

Research thrusts: *Optimization under uncertainty; Advanced methods for inference*
Research sub-thrusts: *Optimal experimental design; Dimensionality reduction*

Motivated by recent advances in theory and numerical algorithms for large-scale Bayesian inverse problems, we have begun working on the "outer" problem of optimal experimental design (OED) for such problems. Specifically, we consider Bayesian inverse problems governed by PDEs whose solution is the posterior probability law for a parameter field. In this context, OED aims to design an observing system (for example, locations of sensors) such that the inference of the model parameters is optimal. What constitutes *optimal* depends on the choice of the design criterion. Here we have pursued two parallel but complementary approaches: one based on a information theoretic design criterion—the expected Shannon information gain in the parameters—making no assumptions of linearity or Gaussianity; and a second that generalizes classical alphabetic A-optimality to the weakly nonlinear and non-Gaussian case. The first approach has emphasized design optimality in relatively low-dimensional settings where nonlinearity and non-Gaussianity are significant, while the second has emphasized scalability to expensive models and infinite-dimensional parameter spaces under the Laplace approximation.

The key computational bottleneck of the first approach is the evaluation of the expected information gain (EIG) for any candidate design. Though several approaches have been proposed to mitigate this cost—e.g., surrogate models [60] or Gaussian approximations—they typically introduce errors that can become arbitrarily large for complex nonlinear problems, or that cannot be reduced through numerical refinement. We have instead developed a new *consistent* estimator of EIG based on layered multiple importance sampling [48]. At the heart of EIG is the need to estimate log-normalizing constants of the Bayesian posterior under multiple realizations of the data. Our algorithm makes careful use of past model evaluations to sequentially construct better and better biasing distributions for the estimation of these normalizing constants, based on initially crude estimates of posterior moments. This scheme is both structure-exploiting and asymptotically unbiased. Compared to previous schemes, our new estimator can achieve up to five orders of magnitude smaller mean-square error for a given computational effort.

In the second approach, we target designs that minimize the average variance of the inversion parameters, which is known as *A-optimal design*. For linear inverse problems, this is equivalent to minimizing the Bayes risk of the MAP estimator, even in infinite dimensions [1]. Here, our goal is to construct theory and algorithms for OED that scale to large-scale problems. We have developed a scalable method for computing A-optimal designs for infinite-dimensional Bayesian *linear* inverse problems, and applied it to the problem of locating sensors to best estimate the state of an atmospheric contaminant [2]. By approximating the parameter-to-observable map offline using a low-rank randomized SVD, our method does not require PDE solutions during the optimal design process—and requires only a fixed number of PDE solves in the offline phase. Moreover, we incorporate sparsity controls given by a sequence of penalty functions that successively approximate the ℓ_0-norm. The result is an OED method that scales independently of the parameter dimension, data dimension, and observation dimension.

Nonlinear inverse problems present additional difficulties for OED, since for these the design criterion (the variance) depends on the actual data, which are unknown at the time the experimental design is computed. Here, we have developed a scalable method for computing A-optimal designs that minimizes the average variance of a Gaussian approximation to the inversion parameters at the posterior mode [3]. Our formulation consists of a bi-level optimization problem which includes as constraints the optimality conditions defining the solution of the inverse problem, as well as the PDEs describing the uncertainty in the inverse solution. Our scheme integrates efficient forward solvers, scalable methods for Bayesian inversion, and scalable OED for linear inverse problems. We have applied the method to a subsurface flow problem [3], and employed it for optimal source compression in inverse scattering problems [31].