Extending strategies from modular Bayesian inference to solve inverse problems with uncertainty

Authors: Giri Gopalan (Los Alamos National Lab; ggopalan@lanl.gov), corresponding – Devin Francom (Los Alamos National Lab; dfrancom@lanl.gov)

Topics: Probabilistic approaches for solving inverse problems, Uncertainty-aware hybrid modeling for solving inverse problems, Scalable algorithms for inverse problems

Challenges: Computer simulations of complex physical systems often couple together multiple submodels, typically implemented with numerical codes, that each constitute an important piece of physics driving the system of interest. Examples of this modeling strategy abound for both problems relevant to national security and fundamental science. Moreover, such multiphysics systems have experimental data for each submodel that can be used in order to solve inverse problems for each of the system components. Rooted in a statistical foundation, we advocate for solving such inverse problems as modifications of Bayesian parameter inference, which provide a way to produce probabilistic assessment of uncertainty. Specifically, this entails breaking up the problem with conditional probability distributions as in Bayesian hierarchical or Bayesian calibration approaches.

An important issue with such approaches is accounting for model error, misspecification, or discrepancy, which is crucial for making precise and accurate extrapolations that properly mitigate uncertainty, as is aptly demonstrated in [1]. The essential argument is that accurate prior knowledge of discrepancy is required in order to mitigate bias for physical parameter estimates. Improperly accounting for or neglecting model error can lead to a situation where collecting larger experimental data sets will end up producing very small uncertainties for parameters but at wrong values, which leads to a false sense of confidence in the parameter estimates and consequent predictions. Simultaneously, a practical challenge is that without having significant prior knowledge of discrepancy, an overly flexible prior on discrepancy will conflate with the parameter estimates. Hence, our first challenge is to account for model error and/or misspecification when there is no reliable prior information with which to inform discrepancy, when using a Bayesian inference approach.

A second consideration in the multiphysics setup described above is that the solution of inverse problems is now often aided with the help of a surrogate model or emulator for each of the core pieces of physics. In modern settings it is very common to use a neural-network based emulator (for instance a fine-tuned pretrained model or a physics-informed neural network). However, Bayesian neural networks are notoriously difficult to fit especially for modern architectures, and it is not resolved how to incorporate uncertainties into the neural network while solving an inverse problem in a Bayesian way [2].

Opportunities: We believe modeling and computational strategies from modular, generalized, and cut-Bayesian inference [3-6] provide important, uncharted research avenues to the two major challenges highlighted above, namely: 1) accounting for model error/misspecification and 2) incorporating modern neural-network surrogates with uncertainty into a Bayesian framework for inverse problems. The essence of cut-Bayesian inference is to infer a probability distribution for certain parameters using trusted/high-fidelity numerical models and experiments. The uncertainty of these distributions is propagated to other submodels or integrated models when making inference for additional physics parameters with Bayesian inference. But in contrast to usual Bayesian inference, the benefit of this strategy is that the high-fidelity parameter inferences are not corrupted by the low-fidelity physical models. (e.g., **Figure 1** of [3], where Z is specified by a high-fidelity model and Y is specified by a misspecified/low-fidelity model).

In the case of Bayesian neural networks that are trained to emulate simulations, a prodigious number of parameters can be intractable for inference. A point estimate for weights ignores uncertainty in the neural network and therefore potentially underestimates uncertainties. We believe cut-Bayes provides one way to handle this situation. Inspired by literature on deep echo-state networks in a Bayesian hierarchical framework [7], the majority of neural network parameters can be randomly sampled from a cut-distribution, and conditional on these sampled weights, Bayesian inference can be performed on a small subset of the large number of neural network weights – this is inherently a cut-Bayesian strategy.

Innovation: While cut-Bayesian computational methods were proposed and developed for multiphysics problems in [5], work needs to be done to hybridize this approach with discrepancy modeling [1]. This is because cut-Bayes posteriors can avoid parameter biases for only those entering into high-fidelity physics models, but not those that appear in low-fidelity, discrepant ones. A hybrid cut-Bayes/discrepancy model handles realistic scenarios with varying model fidelities. Along with this, developing scalable, computational strategies that involve both cut Bayes and discrepancy functions is a technical innovation that has not been developed, though [5] provides a start with sequential Monte Carlo algorithms with rigorous finite-sample guarantees that don't depend on the parameter set dimensions.

The second technical innovation that will be needed is to employ a computational procedure for cut-Bayesian inference when there is a prodigious number of neural network parameters that have cut distributions. To the best of our knowledge, technical cut-Bayesian or modular Bayesian computational solutions do not exist in the general neural-network setting. Some progress has been made for this problem with deep echo state networks with several Gibbs sampling steps for each sample of randomly selected neural network weights [7]. Several Gibbs sampling steps may not scale well for high dimensional parameter settings. A starting point to perform cut-Bayesian computation could be the sequential Monte Carlo approach of [5] and the variational inference approach of [6], which are known to be computationally efficient. Nonetheless, these methods don't address neural networks, suggesting that genuine innovation is needed to adapt cut-Bayes computation for modern AI neural networks.

References:

[1] Brynjarsdóttir, J., & O'Hagan, A. (2014). Learning about physical parameters: The importance of model discrepancy. *Inverse problems*, *30*(11), 114007.
[2] Papamarkou, T., Skoularidou, M., Palla, K., Aitchison, L., Arbel, J., Dunson, D., ... & Zhang, R. (2024). Position: Bayesian deep learning is needed in the age of large-scale AI. *arXiv preprint arXiv:2402.00809*.
[3] Plummer, M. (2015). Cuts in Bayesian graphical models. *Statistics and Computing*, *25*, 37-43.
[4] Frazier, D. T., & Nott, D. J. (2024). Cutting feedback and modularized analyses in generalized Bayesian inference. *Bayesian Analysis*, *1*(1), 1-29.
[5] Mathews, J., Gopalan, G., Gattiker, J., Smith, S., & Francom, D. (2024). Sequential Monte Carlo for cut-Bayesian posterior computation. *Computational Statistics*, 1-31.
[6] Chakraborty, A., Nott, D. J., Drovandi, C. C., Frazier, D. T., & Sisson, S. A. (2023). Modularized Bayesian analyses and cutting feedback in likelihood-free inference. *Statistics and Computing*, *33*(1), 33.
[7] McDermott, P. L., & Wikle, C. K. (2019). Deep echo state networks with uncertainty quantification for spatio‐temporal forecasting. *Environmetrics*, *30*(3), e2553.