Journal of Computational Physics 324 (2016) 24–43

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

An information theoretic approach to use high-fidelity codes to calibrate low-fidelity codes

Allison Lewis [a,*], Ralph Smith [a], Brian Williams [b], Victor Figueroa [c]

[a] Department of Mathematics, North Carolina State University, Raleigh, NC 27695, United States
[b] Los Alamos National Laboratory, Los Alamos, NM 87545, United States
[c] Sandia National Laboratories, Albuquerque, NM 87185, United States

ARTICLE INFO

Article history:
Received 21 May 2015
Received in revised form 31 July 2016
Accepted 1 August 2016
Available online 5 August 2016

Keywords:
Model calibration
Bayesian experimental design
Optimal evaluation
Mutual information

ABSTRACT

For many simulation models, it can be prohibitively expensive or physically infeasible to obtain a complete set of experimental data to calibrate model parameters. In such cases, one can alternatively employ validated higher-fidelity codes to generate simulated data, which can be used to calibrate the lower-fidelity code. In this paper, we employ an information-theoretic framework to determine the reduction in parameter uncertainty that is obtained by evaluating the high-fidelity code at a specific set of design conditions. These conditions are chosen sequentially, based on the amount of information that they contribute to the low-fidelity model parameters. The goal is to employ Bayesian experimental design techniques to minimize the number of high-fidelity code evaluations required to accurately calibrate the low-fidelity model. We illustrate the performance of this framework using heat and diffusion examples, a 1-D kinetic neutron diffusion equation, and a particle transport model, and include initial results from the integration of the high-fidelity thermal-hydraulics code Hydra-TH with a low-fidelity exponential model for the friction correlation factor.

Introduction

Most complex simulation models have inputs comprised of parameters – such as those in closure relations, initial conditions, boundary conditions, or exogenous forces, which must be calibrated to ensure that the model accurately quantifies the considered physical system. Ideally, one would employ experimental data to calibrate the model. However, there are numerous settings for which this is prohibitively expensive or physically infeasible.

In the context of nuclear power plant design, this can be illustrated by the difficulties associated with measuring CRUD deposits. In the ideal scenario, this measurement requires the complete shutdown of the plant and removal of the fuel rods. After data acquisition, engineers are left with only a low-resolution image of the deposit from which the data must be digitized and used for least-squares inference. The process is further complicated by the fact that thermal contraction of cladding during cooling can cause CRUD to break off rods, thus distorting measurements. Furthermore, cold CRUD does not accurately reflect boron levels during operation where hot CRUD serves as a boron absorber.

For some applications, one can alternatively employ validated higher-fidelity codes to generate synthetic data, which can be used to calibrate lower-fidelity codes. Such high-fidelity models and codes can be based on, for example, conservation

* Corresponding author.
E-mail address: lewis.allison10@gmail.com (A. Lewis).

http://dx.doi.org/10.1016/j.jcp.2016.08.001
0021-9991/

laws or physics that are highly resolved in the regimes for which predictions are sought. When employing high-fidelity codes to generate synthetic data in this manner, it is critical that simulations be restricted to the validation regimes for which the high-fidelity code is statistically determined to be accurate.

Due to their complexity, high-fidelity codes are typically computationally expensive and, in some cases, can take hours or days to run. Hence a critical issue concerning their use for generating synthetic data centers on evaluation strategies that optimize the information content provided by the simulated data. The goal is to reduce the uncertainty associated with calibrated low-fidelity inputs using a minimal number of high-fidelity model evaluations.

In the context of nuclear transport, a number of methods have recently been proposed to address the integration of high- and low-fidelity codes to predict future observations in an efficient manner. As described fully in [12] and introduced earlier in [14], multi-scale frameworks have been proposed for use in many nuclear transport models. In these models, the full domain is split into multiple sub-domains on which a micro-scale approach is applied. A fine-scale transport equation is solved over each sub-domain characterized by appropriate boundary conditions, and cross-sections over angular direction, energy, and space are determined in a manner that preserves neutron reaction rates. A coarse model is then applied to the domain as a whole; a low-order approximation of the transport equation is solved over the entire spatial domain, but without the dependence on the angular direction that was removed via the fine-scale method. This allows for a more efficient evaluation of the model. However, there are weaknesses to this approach, the most significant of which is that the boundary conditions used in the micro-scale model to preserve neutron reaction rates are not necessarily refined based on the coarse-scale solution. Schaefer et al. [12] discuss an approach that addresses the updating of the fine-scale boundary conditions to meet the requirements of the coarse-scale model.

We employ a Bayesian information-theoretic framework to specify evaluation strategies for high-fidelity codes that optimize the information required to calibrate low-fidelity code inputs and reduce associated uncertainties. The goal is to accurately calibrate low-fidelity model parameters using as few high-fidelity model evaluations as possible. By measuring the mutual information between potential designs and parameter distributions, we can select the design that will most significantly reduce the amount of uncertainty in the parameters. We utilize a sequential design setting in which each specific design is selected based on its optimal ability to reduce parameter uncertainty. Once selected, the corresponding high-fidelity simulation is run and the newly acquired data is used to recalibrate the low-fidelity model parameters. The posterior distribution resulting from this calibration becomes the prior distribution for the next cycle. Mutual information between parameters and designs is computed again, and the next most profitable experiment or high-fidelity simulation is chosen. Once a point is reached where information gain is no longer significant, the process is terminated and the cost of additional high-fidelity model evaluations or expensive experimental data acquisition is avoided. We consider two methods for estimating the mutual information between random variables or distributions. The first is based on Monte Carlo evaluation whereas the second utilizes a kth-nearest neighbor (kNN) algorithm to approximate the mutual information.

Of the two, we focus primarily on this kNN approach, since it is generally more efficient than Monte Carlo sampling, particularly in cases with moderate-to-high dimensionality. Within the kNN approach, we investigate two methods for identifying the kth nearest neighbor to a query point. The first, utilized by Kraskov et al. [9] in their initial presentation of this estimate, uses a brute force search, requiring computation times on the order of $O(n^2)$ for n data points. In addition, we investigate the performance of an approximate nearest neighbor (ANN) search algorithm, proposed by Arya et al. [1], in the context of mutual information estimation. This relaxation on the requirement to identify the strict nearest neighbor allows for significant computational savings, requiring computation time on the order of $O(n \log n)$, and the substitution of these approximate nearest neighbors into the mutual information estimate does not affect the order in which we select design conditions for high-fidelity model evaluation.

The use of high-fidelity simulation models to calibrate lower-fidelity models is closely related to the Method of Manufactured Universes (MMU) proposed in [16]. In this framework, the researcher defines the laws of their manufactured "universe" and simulates experimental data—coinciding with the high-fidelity synthetic data in this investigation—that follows these laws and may contain measurement error. These "experiments" are then simulated via the proposed model—the low-fidelity model—and the differences between simulation and manufactured reality are determined. Once the input uncertainties are quantified, the simulated model may be used to predict future observations with a corresponding level of uncertainty. The goal is to provide a framework to test proposed uncertainty quantification methods and assess the predictive capabilities of proposed models. These goals can be similarly achieved using the high-to-low methodology employed here.

The mutual information approach of using high-fidelity codes to calibrate lower-fidelity codes is based on the methodology reported in [3,17] and more generally in [10]. In this paper, we extend these results in three ways. The first centers on the use of the Delayed Rejection Adaptive Metropolis (DRAM) algorithm to construct posterior densities. This yields mutual information algorithms that are robust for models exhibiting highly nonlinear parameter dependencies or correlated parameters. Secondly, we provide a verification framework in which direct Monte Carlo evaluation can be used to assess the accuracy of the computationally more efficient kNN and ANN estimates used to approximate the mutual information. Finally, we establish the relation between the employed high-to-low framework and the Method of Manufactured Universes, which has been used to assess the accuracy and feasibility of uncertainty quantification methods. Specifically, we demonstrate how the method can be employed to construct prediction intervals for the high-fidelity model.

In Section 1, we outline our design algorithm and proceed with descriptions of the kNN and ANN estimates for mutual information in Section 2. In Section 3, we define the Monte Carlo method of mutual information estimation as a means