Journal of Computational Physics 229 (2010) 4648–4663

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Numerical approach for quantification of epistemic uncertainty

John Jakeman [a], Michael Eldred [b,1], Dongbin Xiu [c,*,2]

[a] Department of Mathematics, Australian National University, Australia
[b] Sandia National Laboratories, Albuquerque, NM 87185, United States
[c] Department of Mathematics, Purdue University, West Lafayette, IN 47907, United States

ARTICLE INFO

Article history:
Received 7 October 2009
Received in revised form 8 February 2010
Accepted 3 March 2010
Available online 15 March 2010

Keywords:
Uncertainty quantification
Epistemic uncertainty
Generalized polynomial chaos
Stochastic collocation
Encapsulation problem

ABSTRACT

In the field of uncertainty quantification, uncertainty in the governing equations may assume two forms: aleatory uncertainty and epistemic uncertainty. Aleatory uncertainty can be characterised by known probability distributions whilst epistemic uncertainty arises from a lack of knowledge of probabilistic information. While extensive research efforts have been devoted to the numerical treatment of aleatory uncertainty, little attention has been given to the quantification of epistemic uncertainty. In this paper, we propose a numerical framework for quantification of epistemic uncertainty. The proposed methodology does not require any probabilistic information on uncertain input parameters. The method only necessitates an estimate of the range of the uncertain variables that encapsulates the true range of the input variables with overwhelming probability. To quantify the epistemic uncertainty, we solve an *encapsulation problem*, which is a solution to the original governing equations defined on the estimated range of the input variables. We discuss solution strategies for solving the encapsulation problem and the sufficient conditions under which the numerical solution can serve as a good estimator for capturing the effects of the epistemic uncertainty. In the case where probability distributions of the epistemic variables become known *a posteriori*, we can use the information to post-process the solution and evaluate solution statistics. Convergence results are also established for such cases, along with strategies for dealing with mixed aleatory and epistemic uncertainty. Several numerical examples are presented to demonstrate the procedure and properties of the proposed methodology.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Mathematical models are used to simulate a wide range of systems and processes in engineering, physics, biology, chemistry and environmental sciences. These systems are subject to a wide range of uncertainties. The effects of such uncertainty should be traced through the system thoroughly enough to allow one to evaluate their effects on the intended use of the model usually, but not always, related to prediction of model outputs.

There are two forms of model uncertainty: aleatory and epistemic. Aleatory uncertainty arises from the inherent variation associated with the system under consideration and is irreducible. Epistemic uncertainty represents any lack of knowledge

* Corresponding author. Tel.: +1 765 496 2846.
E-mail addresses: john.jakeman@anu.edu.au (J. Jakeman), mseldre@sandia.gov (M. Eldred), dxiu@purdue.edu (D. Xiu).
[1] Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.
[2] The work of D. Xiu was supported by AFOSR, DOE/NNSA, and NSF.

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.03.003

or information in any phase or activity of the modeling process [13] and is reducible through the introduction of additional information.

The sources of aleatory uncertainty are typically represented using a probabilistic framework under which the aleatory uncertainty can be represented by a finite number of random variables with some known distribution. The sources of aleatory uncertainty include both uncertainty in model coefficients (parametric uncertainty) and uncertainty in the sequence of possible events (stochastic uncertainty). Stochastic uncertainty is entirely aleatory by nature. Parametric uncertainty can also be completely aleatory if the complete distribution of all the model parameters are known *a priori*.

Frequently, strong statistical information such as probability distribution functions or high-order statistical moments is not available. Experimental data needed to construct this information is often expensive and consequently no data, or only a small collection of data points, may be obtainable. In these cases "expert opinion" is used in conjunction with the available data to produce weak inferential estimates of parametric characteristics, often in the form of lower and upper bounds. Other sources of epistemic uncertainty include limited understanding or misrepresentation of the modeled process, known commonly as "model form" uncertainty. Inclusion of "enough" additional information about either the model parameters or structure can lead to a reduction in the predicted uncertainty of a model output. Consequently, we can consider epistemic uncertainty as providing (conservative) bounds on an underlying aleatory uncertainty, where reduction and convergence to the true aleatory uncertainty (or, in some cases, a constant value) can be obtained given sufficient additional information.

Until recently, most uncertainty analysis has focused on aleatory uncertainty. Numerous methods have been developed that provide accurate and efficient estimates of this form of uncertainty. In particular, stochastic Galerkin (SG) [2,9,28] and stochastic collocation (SC) [1,5,8,16,21,25,27] methods provide accurate representations of aleatory uncertainty and have the ability to deal with steep non-linear dependence of the solution on random model data. For a detailed review on the methods, see [26].

In comparison to the quantification of aleatory uncertainty, the analysis of epistemic uncertainty has proved more challenging. Probabilistic representations of epistemic uncertainty are inappropriate, since the characterization of input epistemic uncertainty through well-defined probability distributions imposes a large amount of unjustified structure on the influence of the inputs on the model predictions. This can result in stronger inferences than are justified by the available data. Evidence theory [12], possibility theory [4] and interval analysis [11,19] have been proposed as more appropriate alternatives, where they are listed in descending order based on the amount of imposed input structure.

Of the aforementioned methods, evidence theory is the most closely related to probability theory. Evidence theory starts from basic probability assignments on the inputs, propagates these descriptions through a model using standard sampling techniques, and produces estimates of the lowest and highest probabilities of the model observables. These estimates define cumulative belief and cumulative plausibility functions that represent the uncertainty in the output metrics, where belief provides a measure of the amount of information that supports an event being true and plausibility measures the absence of information that supports the event being false. The evidence theory representation of uncertainty approaches the probabilistic representation as the amount of information about the input data increases [12].

Possibility theory is closely related to fuzzy set theory and, similar to evidence theory, utilizes two descriptions of likelihood, necessity and possibility. These two measures are based upon the properties of individual elements of the universal set of events, unlike plausibility and belief which are derived from the properties of subsets of the universal set. For more details, see [10].

Evidence and possibility theory require aggregation of data from multiple sources into a format consistent with the chosen technique. In practice, this can be difficult and time consuming. Interval analysis [17], on the other hand, only requires upper and lower bounds on the uncertain input data. Sampling and/or optimization [5,19] is then used to generate upper and lower bounds (intervals) on the model outputs from predefined intervals on the input data.

The application of evidence theory, possibility theory and interval analysis to non-linear and complex problems often requires a prohibitively large number of samples and typically underestimates the output extrema. Global surrogate models have been used in an attempt to alleviate this problem [10]; however, the performance of these approaches is highly dependent on the accuracy of the surrogate model and construction costs can be high when global accuracy is required and convergence rates are not exponentially fast. In more recent work, surrogates with adaptive refinement strategies have been combined with stochastic collocation methods [5,6,19,20] in order to segregate aleatory quantification with stochastic expansions from epistemic quantification using optimization-based interval estimation.

The choice of the aforementioned methods depends on the amount of available information which can be utilized to characterize the input uncertainty. Consequently this choice is highly problem dependent. Here we propose a new and more general framework to numerically quantify epistemic uncertainty. This proposed method can deal with varying amounts of information on the input data from simple bounds to full probabilistic descriptions, and thus can seamlessly handle the problems with both epistemic and aleatory uncertainties. Furthermore the proposed approach utilizes the classical approximation theory in multi-dimensional space and achieves high efficiency than the methods currently available.

Unlike many existing numerical methods for quantifying epistemic uncertainty, the proposed method requires only an approximation of the ranges of the input data that encapsulates the "true" bounds of the input data. We then propose solving an "encapsulation problem" which generates a solution to the governing equations in a domain that encloses the true (and unknown) probability space. Here a multi-dimensional polynomial expansion can be employed to approximate the solution on the larger encapsulation space. We show that if such a representation converges in the encapsulation space then this method will also converge in the true probability space. Furthermore, convergence is maintained even in the presence of

Download English Version:

https://daneshyari.com/en/article/521309

Download Persian Version:

https://daneshyari.com/article/521309

Daneshyari.com