Journal of Computational Physics 255 (2013) 521–539

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Improved statistical models for limited datasets in uncertainty quantification using stochastic collocation

Aravind Alwan, N.R. Aluru *

Department of Mechanical Science and Engineering, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign, 405 N. Mathews Avenue, Urbana, IL 61801, United States

ARTICLE INFO

Article history:
Received 18 March 2013
Received in revised form 11 July 2013
Accepted 14 August 2013
Available online 27 August 2013

Keywords:
Uncertainty quantification
Stochastic collocation
Density estimation
Moment matching
Reproducing kernel Hilbert space (RKHS)
Microelectromechanical systems (MEMS)

ABSTRACT

This paper presents a data-driven framework for performing uncertainty quantification (UQ) by choosing a stochastic model that accurately describes the sources of uncertainty in a system. This model is propagated through an appropriate response surface function that approximates the behavior of this system using stochastic collocation. Given a sample of data describing the uncertainty in the inputs, our goal is to estimate a probability density function (PDF) using the kernel moment matching (KMM) method so that this PDF can be used to accurately reproduce statistics like mean and variance of the response surface function. Instead of constraining the PDF to be optimal for a particular response function, we show that we can use the properties of stochastic collocation to make the estimated PDF optimal for a wide variety of response functions. We contrast this method with other traditional procedures that rely on the Maximum Likelihood approach, like kernel density estimation (KDE) and its adaptive modification (AKDE). We argue that this modified KMM method tries to preserve what is known from the given data and is the better approach when the available data is limited in quantity. We test the performance of these methods for both univariate and multivariate density estimation by sampling random datasets from known PDFs and then measuring the accuracy of the estimated PDFs, using the known PDF as a reference. Comparing the output mean and variance estimated with the empirical moments using the raw data sample as well as the actual moments using the known PDF, we show that the KMM method performs better than KDE and AKDE in predicting these moments with greater accuracy. This improvement in accuracy is also demonstrated for the case of UQ in electrostatic and electrothermomechanical microactuators. We show how our framework results in the accurate computation of statistics in micromechanical systems.

1. Introduction

Uncertainty quantification (UQ) has become a necessary step in the design of most modern engineering systems due to the need to create robust devices that can tolerate variations in the manufacturing process or in the operating environment. These variations or uncertainties can be represented by stochastic variables which perturb the deterministic behavior of the device about the nominal value for which it was designed. The UQ process consists of identifying the relevant uncertain parameters, assigning appropriate stochastic models to them and quantifying their effect on the final performance of the device. By leveraging fast and accurate deterministic solvers [1–5], it is possible to perform UQ on the computer so as to be able to predict the sensitivity of a given design to an applied set of uncertainties. In this paper, we restrict our focus

* Corresponding author.
E-mail address: aluru@illinois.edu (N.R. Aluru).
URL: http://www.illinois.edu/~aluru (N.R. Aluru).

0021-9991/$ – see front matter
http://dx.doi.org/10.1016/j.jcp.2013.08.024

to a broad category of devices that are collectively called microelectromechanical systems (MEMS). These devices have dimensions that are of the order of micrometers and are particularly sensitive to uncertainties that arise due to an inability to precisely control manufacturing tolerances. Performing UQ for these devices is further complicated by the fact that the characterization data, which describes the variation in the design parameters, is very often unavailable or sparse in quantity. Hence, our goal in this paper is to perform UQ in micromechanical devices by developing a framework that can produce reliable predictions given the limited amount of data describing the uncertainties.

To provide a context for the process of uncertainty quantification, we consider an example that illustrates the issues involved. Microelectromechanical devices are typically batch-fabricated on silicon wafers through a repeated process of etching or deposition after masking or exposing selective areas of the wafer. As a result of process variations during batch processing, there may be significant difference in material and geometric properties of devices fabricated on different parts of a wafer or even across different wafers. Uncertainty quantification is a valuable tool used by the designer to predict the effect of these variations on the performance of a nominal device by considering a numerical model of the device where some of the parameters are uncertain. UQ can be performed by either assuming stochastic models for random parameters or by fitting models to experimentally measured characterization data. For instance, a set of measurements of the elastic modulus of the material that forms the structural layer in a microactuator, can be used to estimate a stochastic model for elastic modulus, which can then be propagated through a numerical model of the device. The designer can then predict the effect of a particular stochastic parameter on some output quantity of interest, say the mean displacement of the actuator under an applied stimulus.

There are many different ways of propagating uncertainties through a numerical model of a device. Although traditionally, sampling-based methods were popular for design optimization in microsystems [6–8], they have been replaced by faster methods like generalized polynomial chaos [9–11] and more recently, by stochastic collocation [12,13] that do not rely on statistical sampling. Adaptive improvements to these methods [11,14–16] have further improved the efficiency by adaptively concentrating the computational effort in the areas where the error is large. With the help of these methods, we can carry out the step of propagating uncertainties to any arbitrary precision in an efficient manner.

With the advent of these efficient propagation methods, the error in predictions is no longer limited by the precision of the propagation methods, but rather, by the accuracy of the stochastic models that represent the input uncertainties. For many sources of randomness, the nature of uncertainty is not known in advance and in the absence of a physical justification, it is not possible to apply a standard distribution function to represent the uncertainty. Instead we resort to a data-driven approach, where we estimate a probabilistic model using experimentally measured characterization data. Although density estimation methods have been popular for a few decades with the advent of kernel density estimation (KDE) [17–20] and its adaptive modifications [21], these methods are mostly designed for asymptotic convergence when the size of the data set is large. On the other hand, for the purpose of UQ in MEMS, it is more appropriate to pick methods that try to preserve what is known from the limited data. We look at moment matching methods [22–24] that try to estimate PDFs that reproduce the statistics of either the raw data, or more interestingly, the statistics of the system response with respect to the data.

In this paper, we consider a method called kernel moment matching (KMM) [25], which can be used to estimate PDFs that accurately reproduce the moments of all functions that lie in a reproducing kernel Hilbert space (RKHS). Since we use stochastic collocation to construct approximations to the system response function, we leverage this property to identify a suitable RKHS for this process. We argue that the choice of this RKHS allows the resulting PDF to optimally reproduce the output statistics of any system modeled using stochastic collocation. This effectively decouples the density estimation process from the propagation step and allows for the estimation of PDFs without evaluating the actual system response function. Finally, we demonstrate through numerical examples that this process improves the accuracy of UQ in micromechanical systems like electrostatic and electrothermomechanical actuators. We propose that this method can be used in other engineering systems as well, where the amount of data that is available to characterize uncertainties is limited.

The rest of this paper is organized as follows: Section 2 gives a brief overview of uncertainty propagation, with a specific focus on stochastic collocation, in order to introduce some notation that will be used in the later section. The main contribution of this paper, which is the construction of stochastic models with limited data, is presented in Section 3, where we modify existing density estimation methods to suit this goal. Section 4 presents the numerical results, where we consider density estimation in univariate and multivariate cases in order to demonstrate the advantage of our framework over traditional approaches. This section also includes MEMS examples to which we applied synthetic uncertainties as well as models estimated from actual data. We finally present the conclusions in Section 5.

2. Propagation of uncertainties

Before we discuss density estimation, which is the primary focus of this paper, we introduce some concepts related to uncertainty propagation and the computation of statistics, since we will use these ideas later on to guide the density estimation process. The most important and computationally intensive step of uncertainty quantification is the process of propagating uncertainties through a numerical model of a device, which is nothing but a deterministic solver that can compute the output parameter of interest for a given set of device parameters. The UQ process starts with the identification of random parameters that act as sources of uncertainty in the model. Once we identify the uncertain parameters that are relevant, we can call the deterministic solver multiple times with different sets of input parameters in order to get an idea