Commun. Math. Res.
doi: 10.4208/cmr.2020-0010

Vol. **36**, No. 2, pp. 128-153
May 2020

Sparse Approximation of Data-Driven Polynomial Chaos Expansions: An Induced Sampling Approach

Ling Guo[1], Akil Narayan[2], Yongle Liu[3] and Tao Zhou[4,*]

[1] *Department of Mathematics, Shanghai Normal University, Shanghai 200234, China.*

[2] *Department of Mathematics, and Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, UT 84112, USA.*

[3] *Department of Mathematics, Southern University of Science and Technology, Shenzhen 518055, China.*

[4] *LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.*

Received 20 January 2020; Accepted 25 April 2020

Abstract. One of the open problems in the field of forward uncertainty quantification (UQ) is the ability to form accurate assessments of uncertainty having only incomplete information about the distribution of random inputs. Another challenge is to efficiently make use of limited training data for UQ predictions of complex engineering problems, particularly with high dimensional random parameters. We address these challenges by combining data-driven polynomial chaos expansions with a recently developed preconditioned sparse approximation approach for UQ problems. The first task in this two-step process is to employ the procedure developed in [1] to construct an "arbitrary" polynomial chaos expansion basis using a finite number of statistical moments of the random inputs. The second step is a novel procedure to effect sparse approximation via ℓ^1 minimization in order to quantify the forward uncertainty. To enhance the performance of the preconditioned ℓ^1 minimization problem, we sample from the so-called induced distribution, instead of using Monte Carlo

*Corresponding author. *Email addresses:* `lguo@shnu.edu.cn` (L. Guo), `akil@sci.utah.edu` (A. Narayan), `11749318@mail.sustc.edu.cn` (Y. Liu), `tzhou@lsec.cc.ac.cn` (T. Zhou)

(MC) sampling from the original, unknown probability measure. We demonstrate on test problems that induced sampling is a competitive and often better choice compared with sampling from asymptotically optimal measures (such as the equilibrium measure) when we have incomplete information about the distribution. We demonstrate the capacity of the proposed induced sampling algorithm via sparse representation with limited data on test functions, and on a Kirchoff plating bending problem with random Young's modulus.

AMS subject classifications: 41A10, 65D15, 62E17

Key words: Uncertainty quantification, data-driven polynomial chaos expansions, sparse approximation, equilibrium measure, induced measure.

1 Introduction

The main purpose of uncertainty quantification (UQ) is to quantify the effect of various sources of randomness on output model predictions. An effective and popular approach is to build an approximation for the map between the random input and the model output, i.e., the quantities of interest (QoI). Many techniques have been developed to construct such approximations. When the distribution of the random input variables are known, the generalized Polynomial Chaos (gPC) [33, 37] expansion is a popular approach. The basic idea is to build a polynomial approximation of the QoI and then the goal is to compute the polynomial expansion coefficients. There are intrusive and non-intrusive approaches to compute the unknown gPC coefficients. In this paper, we focus on the non-intrusive stochastic collocation method, which constructs a global polynomial approximations after running the model with a sample set of the random variables. We refer to [24, 31] and references therein for recent developments for stochastic collocation methods.

However, in real-world scientific applications, one of the main challenges for UQ is the lack of explicit knowledge of the random inputs. Recently, extension of the gPC method to more general input distributions or arbitrary distributions have been investigated. These methods include Multi-element generalized polynomial chaos (ME-gPC) using local polynomial expansion [27,32], Global polynomial expansions based on Gram-Schmidt orthogonalization [34,35], and moment match method to deal with arbitrary distributions (termed aPC) [26]. There has also been extensive work on constructing polynomial approximation for dependent random variables or arbitrary distributions, we refer to the recent paper [19] and references therein.