Journal of Computational Physics 307 (2016) 94–109

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Enhancing sparsity of Hermite polynomial expansions by iterative rotations

Xiu Yang [a,1], Huan Lei [a,1], Nathan A. Baker [b], Guang Lin [c,d,*]

[a] *Advanced Computing, Mathematics, and Data Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA*
[b] *Computational and Statistical Analytics Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA*
[c] *Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA*
[d] *School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA*

ARTICLE INFO

Article history:
Received 13 June 2015
Received in revised form 10 November 2015
Accepted 19 November 2015
Available online 30 November 2015

Keywords:
Uncertainty quantification
Generalized polynomial chaos
Compressive sensing
Iterative rotations
Active subspace
High dimensions

ABSTRACT

Compressive sensing has become a powerful addition to uncertainty quantification in recent years. This paper identifies new bases for random variables through linear mappings such that the representation of the quantity of interest is more sparse with new basis functions associated with the new random variables. This sparsity increases both the efficiency and accuracy of the compressive sensing-based uncertainty quantification method. Specifically, we consider rotation-based linear mappings which are determined iteratively for Hermite polynomial expansions. We demonstrate the effectiveness of the new method with applications in solving stochastic partial differential equations and high-dimensional ($\mathcal{O}(100)$) problems.

1. Introduction

Uncertainty quantification (UQ) plays an important role in constructing computational models as it helps to understand the influence of uncertainties on the quantity of interest. In this paper, we study parametric uncertainty, which treats some of the parameters as random variables. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, where Ω is the event space and P is a probability measure on the σ-field $\mathcal{F}$. We consider a system depending on a d-dimensional random vector $\boldsymbol{\xi}(\omega) = (\xi_1(\omega), \xi_2(\omega), \cdots, \xi_d(\omega))^T$, where ω is an event in Ω. For simplicity, we denote $\xi_i(\omega)$ as ξ_i. We aim to approximate the quantity of interest $u(\boldsymbol{\xi})$ with a generalized polynomial chaos (gPC) expansion [1,2]:

$$u(\boldsymbol{\xi}) = \sum_{n=1}^{N} c_n \psi_n(\boldsymbol{\xi}) + \varepsilon(\boldsymbol{\xi}), \tag{1.1}$$

where ε is the truncation error, N is a positive integer, c_n are coefficients, ψ_n are multivariate polynomials which are orthonormal with respect to the distribution of $\boldsymbol{\xi}$:

$$\int_{\mathbb{R}^d} \psi_i(\boldsymbol{\xi}) \psi_j(\boldsymbol{\xi}) \rho(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} = \delta_{ij}, \tag{1.2}$$

* Corresponding author at: Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA.
E-mail address: guanglin@purdue.edu (G. Lin).

[1] The first two authors made equal contributions to this manuscript.

http://dx.doi.org/10.1016/j.jcp.2015.11.038
0021-9991/

where $\rho(\boldsymbol{\xi})$ is the probability distribution function (PDF) of $\boldsymbol{\xi}$ and δ_{ij} is the Kronecker delta. The approximation converges in the L_2 sense as N increases if u is in the Hilbert space associated with the measure of $\boldsymbol{\xi}$ (i.e., the weight of the inner product is the PDF of $\boldsymbol{\xi}$) [2–4]. Stochastic Galerkin and probabilistic collocation are two popular methods [1,2,5–8] used to approximate the gPC coefficients $\boldsymbol{c} = (c_1, c_2, \cdots, c_N)^T$. Stochastic collocation starts by generating samples of input $\boldsymbol{\xi}^q, q = 1, 2, \cdots, M$ based on $\rho(\boldsymbol{\xi})$. Next, the computational model is calculated for each $\boldsymbol{\xi}^q$ to obtain corresponding samples of the output $u^q = u(\boldsymbol{\xi}^q)$. Finally, coefficients $\boldsymbol{c}$ are approximated based on u^q and $\boldsymbol{\xi}^q$. Note that in many practical problems, it is very costly to obtain u^q and, due to the limited computational sources, we will often have $M < N$ or even $M \ll N$. The smaller number of samples than basis functions implies that the following linear system is under-determined:

$$\boldsymbol{\Psi}\boldsymbol{c} = \boldsymbol{u} + \boldsymbol{\varepsilon}, \tag{1.3}$$

where $\boldsymbol{u} = (u^1, u^2, \cdots, u^M)^T$ is the vector of output samples, $\boldsymbol{\Psi}$ is an $M \times N$ matrix with $\Psi_{ij} = \psi_j(\boldsymbol{\xi}^i)$ and $\boldsymbol{\varepsilon} = (\varepsilon^1, \varepsilon^2, \cdots, \varepsilon^M)^T$ is a vector of error samples with $\varepsilon^q = \varepsilon(\boldsymbol{\xi}^q)$. The compressive sensing method is effective at solving this type of under-determined problem when $\boldsymbol{c}$ is sparse [9–12] and recent studies have applied this approach to uncertainty quantification (UQ) problems [13–24].

Several useful approaches have been developed to enhance the efficiency of solving Eq. (1.3) in UQ applications. First, re-weighted ℓ_1 minimization assigns a weight to each c_n and solves a weighted ℓ_1 minimization problem to enhance the sparsity [25]. The weights can be estimated in *a priori* [18,26] or, for more general cases, can be obtained iteratively [15,17]. Second, better sampling strategies can be used, such as minimizing the mutual coherence [27,20]. Third, Bayesian compressive sensing method provides the posterior distribution of the coefficients [23,16]. Finally, adaptive basis selection selects basis functions to enhance the efficiency instead of fixing the basis functions at the beginning [22]. Recently, we propose an approach [17] to enhance the sparsity of $\boldsymbol{c}$ through the rotation of the random vector $\boldsymbol{\xi}$ to a new random vector $\boldsymbol{\eta}$, where the rotation operator is determined by the sorted variability directions of the quantity of interest u based on the active subspace method [28].

In this work, we aim to extend our previous work [17] and consider the specific case where the system depends on i.i.d. Gaussian random variables; i.e., $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ where $\boldsymbol{0}$ is a d-dimensional zero vector and $\boldsymbol{I}$ is a $d \times d$ identity matrix. This assumption appears in a wide range of physics and engineering problems. We aim to find a mapping $g : \mathbb{R}^d \mapsto \mathbb{R}^d$ which maps $\boldsymbol{\xi}$ to a new set of i.i.d. Gaussian random variables $\boldsymbol{\eta} = (\eta_1, \eta_2, \cdots, \eta_d)^T$ such that the gPC expansion of u with respect to $\boldsymbol{\eta}$ is sparser. In other words,

$$u(\boldsymbol{\xi}) \approx \sum_{n=1}^{N} c_n \psi_n(\boldsymbol{\xi}) = \sum_{n=1}^{N} \tilde{c}_n \tilde{\psi}_n(\boldsymbol{\eta}(\boldsymbol{\xi})) \approx u(\boldsymbol{\eta}(\boldsymbol{\xi})), \tag{1.4}$$

where $\tilde{\psi}_n$ are orthonormal polynomials associated with the new random vector $\boldsymbol{\eta}$ and $\tilde{c}_n$ are the corresponding coefficients. Note that $\psi_n = \tilde{\psi}_n$ since $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. We intend to find the set $\tilde{\boldsymbol{c}} = (\tilde{c}_1, \tilde{c}_2, \cdots, \tilde{c}_N)^T$ which is sparser than $\boldsymbol{c}$ while preserving the properties of matrix $\tilde{\boldsymbol{\Psi}}$ (with $\tilde{\Psi}_{ij} = \tilde{\psi}_j(\boldsymbol{\eta}^i)$) close to those of $\boldsymbol{\Psi}$ to improve the efficiency of the compressive sensing method. To accomplish this, we will use a linear mapping, based on the idea of active subspaces [28], to obtain $\boldsymbol{\eta}$ as first proposed in [17]. Unlike our previous work, we build this mapping iteratively in order to obtain a sparser $\tilde{\boldsymbol{c}}$ and improve the efficiency of the gPC approximation by compressive sensing. We also provide the analytical form of the "gradient matrix" (see Eq. (3.3)) to avoid estimating it with Monte Carlo methods. Our method is applicable for both ℓ_0 and ℓ_1 minimization problems. Especially, for the latter, we can also integrate the present method with re-weighted ℓ_1 minimization method to further reduce the error. We demonstrate that, compared with the standard compressive sensing methods, our approach reduces the relative L_2 error of the gPC approximation.

2. Brief review of the compressive sensing-based gPC method

2.1. Hermite polynomial chaos expansions

In this paper we study systems relying on d-dimensional Gaussian random vector $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. Therefore, the gPC basis functions are constructed by tensor products of univariate orthonormal Hermite polynomials. For a multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_d), \alpha_i \in \mathbb{N} \cup \{0\}$, we set

$$\psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \psi_{\alpha_1}(\xi_1)\psi_{\alpha_2}(\xi_2)\cdots\psi_{\alpha_d}(\xi_d). \tag{2.1}$$

For two different multi-indices $\boldsymbol{\alpha}_i = ((\alpha_i)_1, (\alpha_i)_2, \cdots, (\alpha_i)_d)$ and $\boldsymbol{\alpha}_j = ((\alpha_j)_1, (\alpha_j)_2, \cdots, (\alpha_j)_d)$, we have the property

$$\int_{\mathbb{R}^d} \psi_{\boldsymbol{\alpha}_i}(\boldsymbol{\xi})\psi_{\boldsymbol{\alpha}_j}(\boldsymbol{\xi})\rho(\boldsymbol{\xi})\mathrm{d}\xi = \delta_{\boldsymbol{\alpha}_i\boldsymbol{\alpha}_j} = \delta_{(\alpha_i)_1(\alpha_j)_1}\delta_{(\alpha_i)_2(\alpha_j)_2}\cdots\delta_{(\alpha_i)_d(\alpha_j)_d}, \tag{2.2}$$

where

$$\rho(\boldsymbol{\xi}) = \left(\frac{1}{\sqrt{2\pi}}\right)^d \exp\left(-\frac{\xi_1^2 + \xi_2^2 + \cdots + \xi_d^2}{2}\right). \tag{2.3}$$

For simplicity, we denote $\psi_{\boldsymbol{\alpha}_i}(\boldsymbol{\xi})$ as $\psi_i(\boldsymbol{\xi})$.