

ADAPTIVE SEQUENTIAL SAMPLING FOR POLYNOMIAL CHAOS EXPANSION

Lukáš Novák¹, Miroslav Vořechovský¹, and Václav Sadílek¹

¹Brno University of Technology
Veveří 331/95, Brno 60200, Czech Republic
e-mail: {novak.l, vorechovsky.m, sadilek.v}@fce.vutbr.cz

Abstract. *The paper presents a sampling strategy created specifically for surrogate modeling via polynomial chaos expansion. The proposed method combines adaptivity of surrogate model and sequential sampling enabling one-by-one extension of an experimental design. The iteration process of sequential sampling selects from a large pool of candidate points by trying to cover the design domain proportionally to their local variance contribution. The criterion for the sample selection balances between exploitation of the surrogate model and exploration of the design domain. The obtained numerical results confirm its superiority over standard non-sequential approaches in terms of surrogate model accuracy and estimation of the output variance.*

Keywords: Polynomial Chaos Expansion, Adaptive Sampling, Sequential Sampling.

1 INTRODUCTION

Uncertainty quantification of mathematical model of physical system $Y = g(X)$ is attracting an increasing attention in the last decades. Since the quantity of interest (QoI) Y can be output of very computationally demanding model such as complex engineering structures solved by finite element method, it is often necessary to create an approximation of original mathematical model which significantly computationally cheaper. Moreover, it is beneficial to choose approximation which enables direct post-processing in order to obtain statistical moments and sensitivity indices without additional computational demands. Therefore, the paper is focused on a popular method: Polynomial Chaos Expansion (PCE). Although PCE is very accurate surrogate model, its accuracy is highly dependent on design of experiments (DOE). The paper presents a novel approach for sequential extension of the experimental design based on adaptively refined PCE. The proposed approach significantly reduces the number of samples in experimental design to achieve a good surrogate model and thus it reduces the necessary number of evaluations of the original mathematical model.

2 POLYNOMIAL CHAOS EXPANSION

Evaluation of mathematical model of QoI is often highly computationally demanding and thus it is necessary to create an efficient approximation. PCE is a method of representing the output variable Y as a function g^{PCE} of an another random variable ξ called the germ with given distribution

$$Y = g(X) \approx g^{PCE}(\xi), \quad (1)$$

and representing the function $g(X)$ via polynomial expansion. A set of polynomials, orthogonal with respect to the probability distribution of the germ, are used as a basis of the Hilbert space of all real-valued random variables of finite variance. The orthogonality condition for all $j \neq k$ is given by the inner product of the Hilbert space defined for any two functions ψ_j and ψ_k with respect to the weight function p_ξ (probability density function of ξ) as:

$$\langle \psi_j, \psi_k \rangle = \int \psi_j(\xi) \psi_k(\xi) p_\xi(\xi) d\xi = 0. \quad (2)$$

Orthogonal polynomials ψ corresponding to a selected probability distributions p_ξ can be chosen according to Wiener-Askey scheme [12]. For further processing, it is common to use normalized polynomials, where the inner product is equal to the Kronecker delta δ_{jk} , i.e. $\langle \psi_j, \psi_k \rangle = \delta_{jk}$, where $\delta_{jk} = 1$ if and only if $j = k$, and $\delta_{jk} = 0$ otherwise.

In the case of \mathbf{X} and $\boldsymbol{\xi}$ being vectors containing M random variables, the polynomial $\Psi(\boldsymbol{\xi})$ is multivariate and it is built up as a tensor product of univariate orthogonal polynomials. The quantity of interest (QoI), i.e. the response of the mathematical model $Y = g(\mathbf{X})$, can then be represented, according to Ghanem and Spanos [5], as

$$Y = g(\mathbf{X}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} \beta_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \quad (3)$$

where $\boldsymbol{\alpha} \in \mathbb{N}^M$ is a set of integers called the *multi-index*, $\beta_{\boldsymbol{\alpha}}$ are deterministic coefficients and $\Psi_{\boldsymbol{\alpha}}$ are multivariate orthogonal polynomials.

For practical computation, PCE expressed in Eq. (3) must be truncated to a finite number of terms P . The truncation is commonly achieved by retaining only terms whose total degree $|\boldsymbol{\alpha}|$

is less than or equal to a given p . Therefore, the truncated set of PCE terms is then defined as

$$\mathcal{A}^{M,p} = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^M : |\boldsymbol{\alpha}| = \sum_{i=1}^M \alpha_i \leq p \right\}. \quad (4)$$

Additional reduction of the truncated set was proposed by Blatman and Sudret [2] as a “hyperbolic” truncation scheme. Such an approach leads to a dramatic reduction in the cardinality of the truncated set for high total polynomial orders p .

From a statistical point of view, truncated PCE is a simple linear regression model with intercept. Therefore, it is possible to use ordinary least square (OLS) regression to minimize the error ε . In order to use OLS for $\boldsymbol{\beta}$ estimation, it is necessary to first sample n_{sim} realizations of the input random vector \mathbf{X} and the corresponding results of the original mathematical model \mathcal{Y} , together called the experimental design (ED). Then, the vector of deterministic coefficients $\boldsymbol{\beta}$ is calculated using data matrix $\boldsymbol{\Psi}$ as

$$\boldsymbol{\beta} = (\boldsymbol{\Psi}^T \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^T \mathcal{Y}. \quad (5)$$

The number of terms P is dependent on the number of input random variables M and the maximum total degree of polynomials p . Therefore, in case of a large stochastic model, the problem can become computationally highly demanding. The solution can utilize advanced model selection algorithms such as Least Angle Regression (LAR) [3] to find an optimal set of PCE terms as proposed by Blatman and Sudret [2]. Note that, similar techniques such as orthogonal matching pursuit [11] or Bayesian compressive sensing [7] achieve similar numerical results. The sparse set of basis functions obtained by any adaptive algorithm is further denoted for the sake of clarity as \mathcal{A} .

3 ADAPTIVE SEQUENTIAL SAMPLING

The accuracy of PCE is unfortunately highly dependent on given experimental design similarly as in case of any surrogate model. Although there are many sampling schemes suitable for PCE, the recent study [4] shows an advantage of sequential approach. Therefore this paper is focused on iterative selection of the new sampling points according to specific criteria created particularly for PCE. Note that there are two different strategies for sequential sampling. The first is to enrich the initial ED according to a space-filling criterion (exploration) without assuming any knowledge of the mathematical model or PCE form. The second strategy works with the structure of the PCE in order to identify an optimal sample. Unfortunately, in situations when the initial screening overlooks a globally important region, the exploitation criterion may continue refinement of some other, locally important region that was detected, and there is a risk of never discovering a globally important region. Therefore, it is important to include a balance between both criteria in search for the best candidate. Note that, such approach was employed already in a different context [10]: a criterion motivated by the Koksma-Hlawka inequality [8] was proposed and coupled with stratified sampling in order to improve the efficiency of statistical integration. Beside sequential sampling, the ideal algorithm should be able to adaptively reconstruct the PCE using model selection algorithms in order to identify a sparse set of basis functions \mathcal{A} in each iteration.

We propose an adaptive sequential sampling strategy accompanied by a criterion designed for non-intrusive PCE. Once a pool of candidates containing n_{pool} realizations of the random vector $\boldsymbol{\xi}$ generated by any sampling technique is available, it is necessary to construct a criterion

Θ for the selection of the best candidate balancing between the exploitation and exploration of the design domain:

$$\Theta(\xi^{(c)}) \equiv \sqrt{\sigma_A^2(\xi^{(c)}) \cdot \sigma_A^2(\xi^{(s)})} l_{c,s}^M. \quad (6)$$

The criterion is a product of two terms: the exploitation term and the exploration term. The exploration aspect is maintained by accounting for the distance $l_{c,s}$ between a candidate $\xi^{(c)}$ and its nearest neighboring point from the existing ED, $\xi^{(s)}$. For the distance term we select the Euclidean distance between the candidate and its nearest neighbor as

$$l_{c,s} = \sqrt{\sum_{i=1}^M |\xi_i^{(c)} - \xi_i^{(s)}|^2}. \quad (7)$$

The exploitation in candidate selection is motivated by our desire to uniformly cover local contributions to the total variance, σ_Y^2 . The variance can be thought of as an integral of local contributions $\sigma_A^2(\xi)$ over the design domain indexed by coordinates ξ . Once the PCE has been established at any given stage of the algorithm, the local variance is computationally cheap to evaluate for any location ξ as

$$\sigma_A^2(\xi) = \left[\sum_{\substack{\alpha \in \mathcal{A} \\ \alpha \neq 0}} \beta_\alpha \Psi_\alpha(\xi) \right]^2 p_\xi(\xi). \quad (8)$$

In the criterion we take the geometric mean of two local variance contributions representing average variance contribution of the region between the candidate and its nearest neighbor. When this geometric mean is multiplied by the M th power of the distance between the two points, $l_{c,s}^M$, the volume (variance contribution) of an area between them is estimated. The proposed criterion maintains a balance between exploration and exploitation, since a candidate which is close to an existing point can only be selected if the corresponding variance density is significant. Similarly, when a region with low contribution is being detected by the PCE, candidates from such regions are ignored. Maximization of the proposed criterion leads to the best candidate, which is added to active ED. The pool of candidates can be generated by commonly known LHS and the proposed criterion is employed for the selection of the best candidate at every iteration of sequential sampling.

4 NUMERICAL EXAMPLE

The pilot numerical study is represented by Ishigami function [6]. The function is strongly nonlinear, non-monotonic and presents strong interactions. We set the coefficients as in [9]. Let $\mathbf{X} \sim \mathcal{U}[-\pi, \pi]^3$ and the mathematical model

$$Y = \sin(X_1) + 7 \sin^2(X_2) + 0.1 X_3^4 \sin(X_1). \quad [\sigma_Y \approx 13.844587940] \quad (9)$$

The setup of PCE was as follows: PCE is solved by non-intrusive OLS, a sparse set of the basis functions \mathcal{A} is obtained by LAR with maximum total polynomial order $p = 10$ and $p = 20$. The initial ED for the PCE construction before the first step of the proposed iterative algorithm is generated by LHS and it contains an initial screening design with $n_{\text{sim}} = 10$ realizations of the input random vector. The results are compared in terms of the relative error in variance of QoI

$$\epsilon = \frac{|\sigma - \sigma_Y|}{\sigma_Y}, \quad (10)$$

defined as the absolute deviation of the estimated variance σ from the exact value σ_Y divided by the exact variance; and commonly used leave-one-out error of PCE approximation Q^2 [1]. The calculations were repeated 100 times and the orders of errors (\log_{10}) are depicted in Fig.1. Solid lines represent mean values and the scatters represent $\pm \sigma$ confidence intervals.

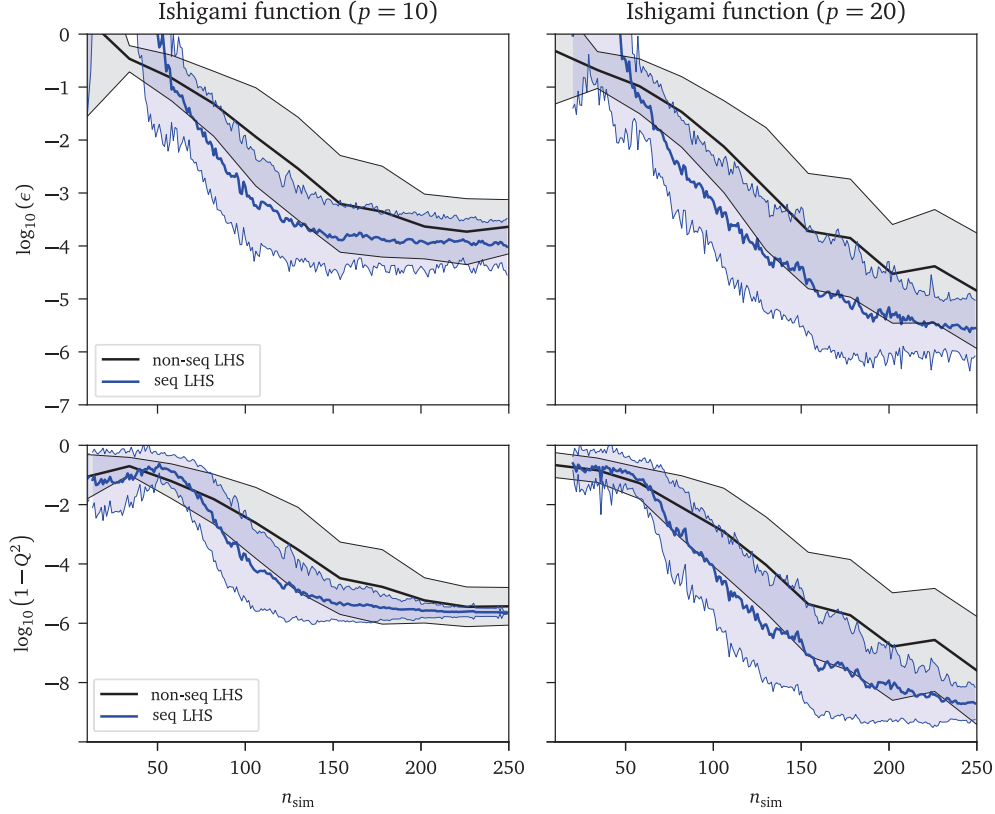


Figure 1: Obtained results for the Ishigami function. The first two row represents the accuracy measured by ϵ and the second row shows Leave-one-out error Q^2 .

5 DISCUSSION AND CONCLUSION

The paper presented innovative approach of adaptive sequential sampling for polynomial chaos expansion. The proposed method combines adaptivity of a PCE and sequential sampling. The sequential sampling is based on one-by-one extension of existing experimental design by a selection of the best candidate from large pool. The best sample candidate is identified by proposed criterion consisting of two parts: average local variance between a candidate “c” and its nearest neighbor “s” and the Euclidean distance between them. Both parts of the criterion together maintain the balance between exploration of the design domain and exploitation of current form of PCE. The presented method can be easily coupled with any existing sampling method such as LHS, which was employed in numerical example. From obtained results, one can see significant improvement in accuracy of PCE using sequential sampling in comparison to standard non-sequential LHS. The improvement is especially clearly visible for mid-size ED. Moreover, the benefit of the sequential sampling is higher in case of $p = 20$ as can be seen in Fig.1 (right) since the maximum polynomial order does not limit the convergence of PCE. Further work will be focused on combination of the sequential approach with advanced sampling schemes.

ACKNOWLEDGMENT

The authors acknowledge financial support provided by the Ministry of Education, Youth and Sports of the Czech Republic under project no. LTAUSA19058. Additionally the first author is supported by BUT internal grant project FAST-J-21-7209.

REFERENCES

- [1] Géraud Blatman and Bruno Sudret. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. *Comptes Rendus Mécanique*, 336(6):518–523, 2008.
- [2] Géraud Blatman and Bruno Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6):2345–2367, 2011.
- [3] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. *The Annals of Statistics*, 32(2):407–451, 2004.
- [4] Noura Fajraoui, Stefano Marelli, and Bruno Sudret. Sequential design of experiment for sparse polynomial chaos expansions. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):1061–1085, 2017.
- [5] Roger G. Ghanem and Pol D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer New York, 1991.
- [6] T. Ishigami and T. Homma. An importance quantification technique in uncertainty analysis for computer models. In *Proceedings. First International Symposium on Uncertainty Modeling and Analysis*. IEEE Comput. Soc. Press, 1990.
- [7] Shihao Ji, Ya Xue, and Lawrence Carin. Bayesian compressive sensing. *IEEE Transactions on Signal Processing*, 56(6):2346–2356, 2008.
- [8] J. F. Koksma. Een algemeene stelling uit de theorie der gelijkmatige verdeeling modulo 1. *Mathematica B*, 11:7–11, 1942/1943.
- [9] Amandine Marrel, Bertrand Iooss, Béatrice Laurent, and Olivier Roustant. Calculations of sobol indices for the gaussian process metamodel. *Reliability Engineering & System Safety*, 94(3):742–751, 2009.
- [10] Michael D. Shields. Adaptive Monte Carlo analysis for strongly nonlinear stochastic systems. *Reliability Engineering & System Safety*, 175:207–224, July 2018.
- [11] Joel A. Tropp and Anna C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, 2007.
- [12] Dongbin Xiu and George Em Karniadakis. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644, 2002.