Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2023) 151-157

ECCOMAS
Proceedia

UNCECOMP 2023
5th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Science and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Athens, Greece, 12-14 June 2023

ON PHYSICALLY CONSTRAINED NON-INTRUSIVE POLYNOMIAL CHAOS EXPANSION

Lukáš Novák[1], **Himanshu Sharma**[2], **Michael D. Shields**[2]

[1]Brno University of Technology
Brno, Czech Republic
E-mail: novak.l@fce.vutbr.cz

[2]Johns Hopkins University
Baltimore, USA
E-mail: {hsharm10, michael.shields}@jhu.edu

Abstract. *Surrogate modeling of costly mathematical models representing physical systems represents a challenging task, since it is typically not possible to create a large experimental design and thus it is beneficial to restrict the approximation by additional information reflecting the physics of the original mathematical model. This paper presents a novel methodology for the construction of polynomial chaos expansion combining experimental design and additional information arising from the physics of an original model – physical constraints. Physical constraints investigated in this paper are represented by boundary conditions and given partial differential equations. A computationally efficient numerical algorithm based on least angle regression and Lagrange multipliers is proposed and compared to standard least angle regression. It is shown that the proposed algorithm leads to superior accuracy of the approximation and does not bring any significant computational burden.*

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

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2023.
doi: 10.7712/120223.10330.19857

1 INTRODUCTION

Mathematical models of real-life physical systems are typically highly computationally demanding and thus it is necessary to use surrogate models as computationally cheap approximations in order to perform an uncertainty quantification (UQ). The surrogate models often treat the original computational models as black-boxes and they are based only on several deterministic simulations for the given data points in the design domain and thus their practical employment is conditioned by a sufficient number of these data points covering the whole design space spanned by input random variables. However, it is beneficial to incorporate additional characteristics and constraints of analyzed physical systems in order to assure realistic behavior of a surrogate model. The paper is focused on Polynomial Chaos Expansion (PCE) reflecting given physical constraints of the original mathematical model. The proposed methodology is applied to reflect: a) boundary conditions in differential equations (Dirichlet, Neumann or mixed) and b) physics constraints described by ordinary differential equation (ODE) or partial differential equation (PDE), though it is generally possible to extend the methodology to various different types of constraints.

The exploitation of known boundary conditions or physical constraints in the construction of an approximation assures higher accuracy of the predictions, especially in parts of design space containing an insufficient number of data points describing the behavior of the original mathematical model. Generally, any numerical optimization technique can be employed for the construction of constrained PCE, since it can be seen as an estimation of deterministic coefficients for given set of basis functions. However this study is focused on equality constraints satisfying Karush-Kuhn-Tucker (KKT) conditions and thus it is possible to use well-known Lagrange multipliers and ordinary least squares for a computationally efficient solution.

2 POLYNOMIAL CHAOS EXPANSION

The PCE represents the quantity of interest (QoI) Y obtained as a result of the original mathematical model $\mathcal{M}(X)$ as a polynomial expansion of another random variable ξ called a germ with a known distribution. A set of polynomials, orthogonal with respect to the probability distribution of the germ, are used as a basis of a 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:

$$\langle \psi_j, \psi_k \rangle = \int \psi_j(\xi)\psi_k(\xi)p_\xi(\xi)\,\mathrm{d}\xi = 0. \tag{1}$$

Orthogonal polynomials ψ corresponding to common probability distributions p_ξ can be chosen according to Wiener-Askey scheme [10] or numerically constructed in case of arbitrary probability distributions [9]. In the case of $\boldsymbol{X}$ 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 QoI $Y = \mathcal{M}(\boldsymbol{X})$, can then be represented as [3]:

$$Y = \mathcal{M}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^M} \beta_{\boldsymbol{\alpha}}\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{2}$$

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. (2) 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

$$\mathcal{A}^{M,p} = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^M : |\boldsymbol{\alpha}| = \sum_{i=1}^{M} \alpha_i \leq p \right\}. \tag{3}$$

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 $\boldsymbol{X}$ 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

$$\boldsymbol{\beta} = \left(\boldsymbol{\Psi}^T \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\Psi}^T \mathcal{Y}. \tag{4}$$

The optimal size of ED is clearly given by the number of terms P dependent on M and p. Therefore, in case of a large stochastic model, the problem can become computationally highly demanding, and thus it is typically useful to find a sparse solution using advanced model selection algorithms such as Least Angle Regression (LAR) [2, 1], orthogonal matching pursuit [8] or Bayesian compressive sensing [4] to find an optimal set of PCE terms.

3 CONSTRAINED POLYNOMIAL CHAOS EXPANSION

Boundary conditions are prescribed by their type (Dirichlet, Neumann or mixed), coordinates $\boldsymbol{\xi}_b$ and their values $\mathbf{c}$. Similarly, ODE/PDE is numerically evaluated for given coordinates $\boldsymbol{\xi}_c$ containing typically realizations from ED but it can be also extended by virtual points $\boldsymbol{\xi}_v$ similarly as in physics-informed neural networks [7]. Note that PCE and its derivatives have same coefficients $\boldsymbol{\beta}$ but different basis functions. Therefore instead of Eq.4, the system of linear equations reflecting physical constrained solved by OLS is in the following form:

$$\underbrace{\begin{bmatrix} \boldsymbol{\Psi}^T\boldsymbol{\Psi} & \boldsymbol{A}^T \\ \boldsymbol{A} & \mathbf{0} \end{bmatrix}}_{\text{KKT matrix}} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}^T\mathcal{Y} \\ \boldsymbol{c} \end{bmatrix} \tag{5}$$

In case of boundary conditions, sub-matrix $\mathbf{A}$ contains basis functions of PCE (Dirichlet conditions) or their derivatives (von Neumann conditions) evaluated for $\boldsymbol{\xi}_b$, and vector $\mathbf{c}$ contains the associated numerical value of a boundary condition at $\boldsymbol{\xi}_b$. In case of PDE/ODE with virtual samples, vector $\mathbf{c}$ contains numerical values of derivatives obtained from PDE/ODE at $\boldsymbol{\xi}_c$ using the original model or at virtual samples. Naturally, the numerical values of derivatives satisfying PDE/ODE at virtual samples are not known at the start of the algorithm and thus it is necessary to construct PCE (estimate coefficients $\boldsymbol{\beta}$) without virtual samples in the first step and iteratively improve the estimation by evaluation of PDE/ODE for virtual samples. Since the whole process is computationally efficient, it can be easily combined with sparse solvers such as LAR. For the sake of clarity, the whole algorithm combining LAR, Lagrange multipliers and virtual samples is summarized in Algorithm 1.

Algorithm 1 Numerical algorithm of LAR-KKT with virtual samples

Input: p, $\mathbf{X}$ and corresponding $\mathcal{Y}$, coordinates of boundary conditions $\boldsymbol{\xi}_b$, their numerical values $\mathbf{c}$ and types of constraints, ODE/PDE function evaluated for ED, virtual samples $\boldsymbol{\xi}_v$, n_{iter}

construct set of P candidate basis functions – $\mathcal{A}^{M,p}$
start LAR algorithm
while $\epsilon \geq \epsilon_{target}$ **do**
    identify the most important basis function and add it to active $\mathcal{A}$
    construct data matrix Ψ
    construct matrix $\mathbf{A}$
    construct vector $\mathbf{c}$
    solve the system of linear equations by a method of Lagrange multipliers (Eq. 5)
    construct PCE function
    for 1 to n_{iter} **do**
        evaluate ODE/PDE at $\boldsymbol{\xi}_v$
        solve the system of linear equations by a method of Lagrange multipliers (Eq. 5)
        construct PCE function
    end for
    get the error ϵ of PCE
end while

Output: set of basis functions $\mathcal{A}$ and corresponding $\boldsymbol{\beta}$

4 NUMERICAL EXAMPLE

The numerical example is represented by the following 1D simple first-order non-linear ODE (Logistic differential equation) modeling population growth:

$$\frac{df(x)}{dx} = f(x)\left(1 - f(x)\right) \tag{6}$$

with boundary condition $f(0) = 0.5$ and $X \sim R[-5, 5]$. The analytical solution is

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{7}$$

The following methods are compared in this study: a) standard PCE using LAR implemented in UQPy [6], b) PCE using LAR and Lagrange multipliers (LAR-KKT) and c) extension of LAR-KKT with virtual samples (Virtual). For correct comparison, all methods have set identical p and ED. Typical realization with significant differences in all methods can be seen in Fig. 1. Results corresponds to $p = 7$ together with 8 simulations in ED. Standard LAR approach clearly leads to over-fitting, since it has no additional information about the mathematical model (ODE). This can be seen also in derivative of PCE compared to the original model (second row in figure). The LAR-KKT takes into account ODE evaluated only in ED and thus both approximations of the original model and its derivative are more accurate. However, the approximation is not perfect, since there is not enough samples in ED. The extension of LAR-KKT by virtual samples ($\boldsymbol{\xi}_v$ contains 10 samples) then lead to superior results, since the shape of derivative must be fulfilled in additional locations. Note that numerical values of derivative in $\boldsymbol{\xi}_v$ are obtained from Eq. 6 using PCE instead of the original function and thus there is not any additional computational burden associated to virtual points.

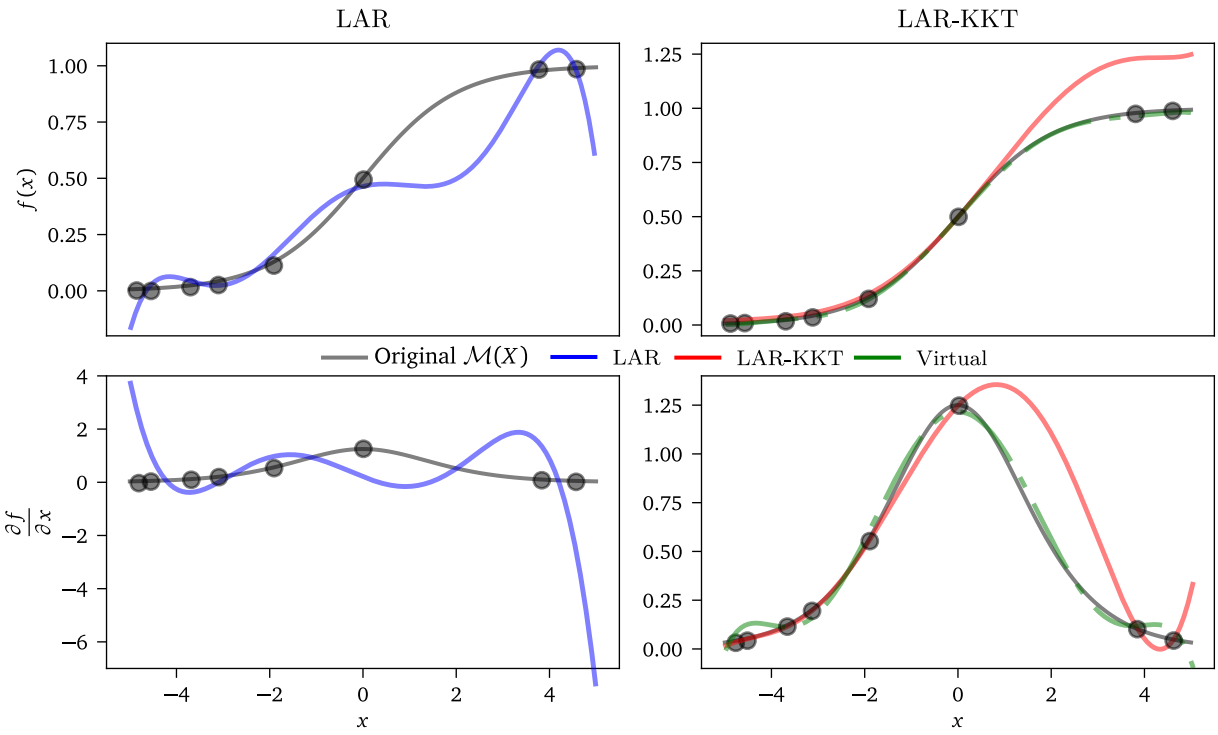

Figure 1: Comparison of the proposed method to standard LAR approach. Figure shows typical realization of the comparison with $p = 7$ and 8 simulations in ED. Top row shows approximations of the mathematical model and bottom row corresponds to derivative of the function.

Naturally in order to compare all three approaches, it is necessary to perform statistical analysis of the obtained accuracy. Fig. 2 presents statistical results for $p = 6$ (top), $p = 8$ (middle) and $p = 11$ (bottom) with 50 repetitions. Note that two graphs are presented for each setting: boxplots with median and outliers (left), and mean $\pm sigma$ interval (right). The obtained errors are measured by mean squared errors, i.e.

$$\epsilon(\boldsymbol{X}) := \mathbb{E}\left[\left(f(\boldsymbol{X}) - \tilde{f}(\boldsymbol{X})\right)^2\right]. \tag{8}$$

It can be seen from statistical results that LAR-KKT and its Virtual modification has significantly lower variance of obtained accuracy for both $p = 6$ and $p = 8$. The low $p = 6$ is limiting the possible accuracy of PCE and thus mean values are almost identical for all compared methods. However, once the polynomial order is increased to $p = 8$, it is clear that also mean (and median) errors obtained by Virtual modification of LAR-KTT are significantly lower than standard LAR approach. This trend is further highlighted with increasing $p = 11$, where the difference in error is up to two orders of magnitude. It can be concluded that LAR-KKT has higher influence on accuracy of PCE for higher p, since it brings additional information and thus restricts set of possible solutions.

5 CONCLUSIONS

The novel methodology for a construction of non-intrusive PCE referenced as LAR-KKT was presented in this paper. The algorithm is based on least angle regression employed as a sparse solver (selection of suitable basis functions) combined with Lagrange multipliers. The presented approach is created for a specific task – UQ of costly mathematical models of phys-

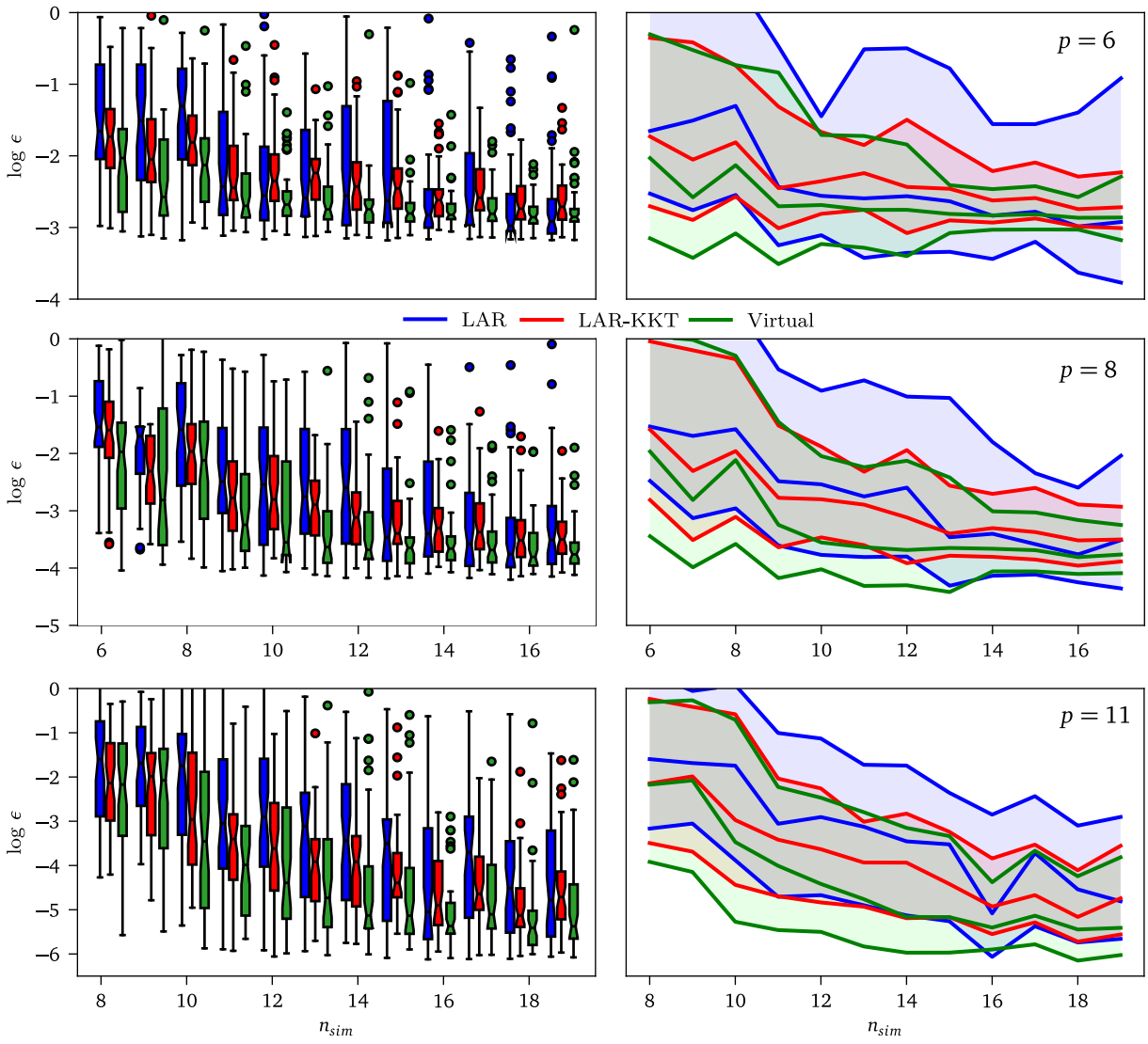

Figure 2: Statistical results of 50 replications of obtained results by the proposed method and standard LAR approach. The first row shows results for fixed $p = 6$ and the second row corresponds to $p = 8$ and the last row to $p = 11$. The left column shows standard boxplots with outliers and the right column shows mean $\pm\sigma$ interval.

ical systems using low-sized ED. From obtained numerical results presented in this very first study of LAR-KKT, it is clear that there is a significant improvement of the accuracy of PCE for low-size to mid-size ED without any additional computational burden. Naturally, the presented method further converges with increasing n_{sim} to the identical accuracy as standard LAR approach, though the variance of the accuracy obtained by LAR-KKT is still significantly lower.

However, there are still several open questions that should be investigated in further research. First of all, the presented numerical example is simple 1D non-linear ODE and thus it must be applied to more complicated examples to confirm its efficiency and scalability. Further, Lagrange multipliers will be combined with active learning techniques [5], which could significantly improve the performance of the presented methodology.

ACKNOWLEDGMENT

The first author would like to express his thanks for the support provided by the Czech Science Foundation under Project No. 23-04712S.

REFERENCES

[1] Géraud Blatman and Bruno Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6):2345–2367, 2011.

[2] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. *The Annals of Statistics*, 32(2):407–451, 2004.

[3] Roger G. Ghanem and Pol D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer New York, 1991.

[4] Shihao Ji, Ya Xue, and Lawrence Carin. Bayesian compressive sensing. *IEEE Transactions on Signal Processing*, 56(6):2346–2356, 2008.

[5] Lukáš Novák, Miroslav Vořechovský, Václav Sadílek, and Michael D. Shields. Variance-based adaptive sequential sampling for polynomial chaos expansion. *Computer Methods in Applied Mechanics and Engineering*, 386:114105, 2021.

[6] Audrey Olivier, Dimitrios Giovanis, BS Aakash, Mohit Chauhan, Lohit Vandanapu, and Michael D Shields. Uqpy: A general purpose python package and development environment for uncertainty quantification. *Journal of Computational Science*, 47:101204, 2020.

[7] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[8] 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.

[9] Xiaoliang Wan and George Em Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *Journal of Computational Physics*, 209(2):617–642, 2005.

[10] 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.