

PHYSICS-INFORMED POLYNOMIAL CHAOS EXPANSIONS: RECENT DEVELOPMENTS AND COMPARISONS

LUKÁŠ NOVÁK¹, QITIAN LU¹, HIMANSHU SHARMA², DIBAKAR ROY SARKAR²,
SOMDATTA GOSWAMI² AND MICHAEL D. SHIELDS²

¹ Brno University of Technology
Brno, Czechia
Lukas.Novak4@vut.cz

² Johns Hopkins University
Baltimore, USA

Key words: Scientific machine learning, Uncertainty quantification, Physics-informed Polynomial chaos expansion, Physics-informed deep operator networks, Statistical sampling

Abstract. This work presents recent developments in a constrained polynomial chaos expansion as a physics-informed machine learning technique. Specifically, an optimized numerical solver for straightforward updating of Lagrange multipliers and an improved statistical sampling method are compared to the original algorithm for estimating deterministic coefficients. Both techniques are applied to solve a heat equation with Neumann boundary conditions. A second study presents a preliminary numerical comparison of the constrained polynomial chaos expansion and physics-informed deep operator networks with respect to computational cost and achieved accuracy.

1 INTRODUCTION

Uncertainty quantification (UQ) of mathematical models representing physical systems is typically associated with a high computational burden. Consequently, there has been significant interest in the development of machine learning methods that offer both numerical efficiency and high accuracy. Surrogate modeling of complex mathematical models of physical systems is particularly challenging, as it requires satisfying physical constraints across the entire design domain while adhering to specific boundary conditions of the system under investigation. Additionally, constructing a large training set is often infeasible due to the immense computational cost.

One of the most widely used surrogate models for UQ is polynomial chaos expansion (PCE) [1], which generally offers a good balance between training set size and achieved accuracy, along with analytical post-processing capabilities for UQ [2]. However, standard PCE surrogate models may yield predictions that violate the underlying physical constraints of the system. To address this limitation, a recently proposed approach referenced as physics-informed polynomial chaos expansion (PC²) combines conventional experimental design with additional constraints derived from the system's physics [3, 4]. The PC² can be seen as a technique for operator learning. Therefore, a numerical comparison with a well-known operator learning method – physics-informed deep operator networks [5], is presented in this paper. Both techniques are applied to a heat equation with Neumann boundary

condition, and the obtained results are compared with respect to accuracy and numerical efficiency. In addition to the comparison, this contribution focuses on recent developments that aim to improve the scalability of the original PC² algorithm.

The PC² method, which employs an augmented system with Lagrange multipliers [3], demonstrates superior performance for linear equality constraints and relatively low-size experimental design. However, its computational cost increases rapidly with the stochastic dimension of the system and/or maximum polynomial order of PCE basis. In this paper, we investigate two directions for improvement of numerical efficiency. First of all, PC² algorithm can be improved by optimized algorithm for straightforward updating of Lagrange multipliers, which significantly reduces the computational time in comparison to naive implementation of constrained least squares. The computational cost is also influenced by the number of virtual and boundary samples. Optimizing the experimental design is thus crucial to further reduce the computational cost similarly to standard data-driven PCE [6, 7]. Both optimized numerical solver of linear equations together with D-optimal sampling of virtual points are applied to a heat equation and compared to the original PC² algorithm.

2 PHYSICS-INFORMED POLYNOMIAL CHAOS EXPANSION

Considering \mathcal{X} and $\boldsymbol{\xi}$ being vectors consisting of M independent random variables, the basis $\Psi(\boldsymbol{\xi})$ is in form of multivariate polynomials:

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \prod_{i=1}^M \psi_{\alpha_i}(\xi_i), \quad (1)$$

where $\boldsymbol{\alpha} \in \mathbb{N}^M$ is a the multi-index consisting of integers reflecting polynomial degrees associated to each ξ_i . Note that multivariate polynomials are simply constructed as a tensor product of univariate polynomials ξ_i . The quantity of interest (QoI), i.e. the response of the model $\mathcal{Y} = u(\mathcal{X})$, can then be represented by PCE as [8, 9]

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

The coefficients β that minimize the error can be further estimated simply by ordinary least squares (OLS). In order to estimate β by OLS, it is necessary to create a set of realizations used as a training data set referenced also as the experimental design (ED). PCE can be seen as a simple linear regression model and thus the vector β can be simply obtained as

$$\beta = (\Psi^T \Psi)^{-1} \Psi^T Y, \quad (3)$$

where Y contains numerical values of QoI, Ψ is the data matrix. Once the PCE is constructed, the first statistical moment of QoI is simply given by the first estimated coefficient of the PCE $\mu_Y = \langle \mathcal{Y}^1 \rangle = \beta_0$ and the variance $\sigma_Y^2 = \langle \mathcal{Y}^2 \rangle - \mu_Y^2$ can be obtained from the all squared coefficients excluding the intercept.

Extension of the standard data-driven PCE in form of PC² was designed to solve any stochastic partial differential equation, generally described by the following formulas:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, t, \mathcal{X}(\omega); u(\mathbf{x}, t, \mathcal{X}(\omega))) &= f(\mathbf{x}, t, \mathcal{X}(\omega)), \quad \forall \mathbf{x} \in \mathcal{D}, t \in \mathcal{T}, \omega \in \Omega \\ \mathcal{B}(\mathbf{x}, t, \mathcal{X}(\omega); u(\mathbf{x}, t, \mathcal{X}(\omega))) &= g(\mathbf{x}, t, \mathcal{X}(\omega)), \quad \forall \mathbf{x} \in \partial\mathcal{D}, t \in \mathcal{T}, \omega \in \Omega \end{aligned} \quad (4)$$

where QoI (solution of the stochastic partial differential equation) is $u(\cdot)$, $\mathcal{T} \subset \mathbb{R}$, $\mathcal{D} \subset \mathbb{R}^3$ having a boundary $\partial\mathcal{D}$, differential operator is represented by \mathcal{L} , boundary operator \mathcal{B} , source terms are f, g , and $\boldsymbol{\mathcal{X}}(\omega) \in \mathbb{R}^d$ is a random vector containing d random variables having a sample space Ω . The objective function of a regression problem constrained by Eq. (4) can be formulated as

$$\begin{aligned} \mathcal{M}(\boldsymbol{\beta}) &= \min_{\boldsymbol{\beta}} \sum_{j=1}^{n_{\text{sim}}} [Y^j - u^{\text{PCE}}(\boldsymbol{x}^j, t^j, \mathbf{X}^j)]^2 = \min_{\boldsymbol{\beta}} \|Y - \boldsymbol{\Psi}\boldsymbol{\beta}\|^2 \\ \text{s.t. } \mathcal{L}(\boldsymbol{x}_V, t_V, \boldsymbol{\mathcal{X}}(\omega); u(\boldsymbol{x}_V, t_V, \boldsymbol{\mathcal{X}}(\omega))) &= f(\boldsymbol{x}_V, t_V, \boldsymbol{\mathcal{X}}(\omega)), \\ \mathcal{B}(\boldsymbol{x}_{\text{BC}}, t_{\text{BC}}, \boldsymbol{\mathcal{X}}(\omega); u(\boldsymbol{x}_{\text{BC}}, t_{\text{BC}}, \boldsymbol{\mathcal{X}}(\omega))) &= g(\boldsymbol{x}_{\text{BC}}, t_{\text{BC}}, \boldsymbol{\mathcal{X}}(\omega)) \end{aligned} \quad (5)$$

Note that two new sets of points in input space are defined extending standard ED: points enforcing boundary conditions $(\boldsymbol{x}_{\text{BC}}, t_{\text{BC}}, \mathbf{X})$ and n_V virtual points used for evaluation of the prescribed differential equations $(\boldsymbol{x}_V, t_V, \mathbf{X})$.

Estimation of deterministic coefficients can be efficiently done by constrained least squares using Lagrange multipliers. The boundary conditions are represented by \mathcal{B} , evaluated at n_{BC} points $\boldsymbol{\xi}_{\text{BC}}$, and the corresponding vector \mathbf{c}_{BC} consisting of n_{BC} rows determined by $c_b = g(\boldsymbol{\xi}_{\text{BC}}^{(b)})$. Constraints in form of differential equations are represented by the differential operator \mathcal{L} , evaluated at n_V virtual points $\boldsymbol{\xi}_V$, and the corresponding vector \mathbf{c}_V consisting of n_V rows given by $c_v = u(\boldsymbol{\xi}_V^{(v)})$.

The prescribed constraints are then assembled into a matrix \mathbf{A} containing the first n_{BC} rows containing $a_b = \left\{ a_b^j = \mathcal{B}(\Psi_j(\boldsymbol{\xi}_{\text{BC}}^{(b)})) \right\}$, $j = 0, \dots, P-1$ and followed by n_V rows containing $a_v = \left\{ a_v^j = \mathcal{L}(\Psi_j(\boldsymbol{\xi}_V^{(v)})) \right\}$, $j = 0, \dots, P-1$. Finally it is possible to construct a system of linear equations solved by the constrained OLS:

$$\underbrace{\begin{bmatrix} \boldsymbol{\Psi}^T \boldsymbol{\Psi} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix}}_{\text{KKT matrix}} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}^T Y \\ \mathbf{c} \end{bmatrix}. \quad (6)$$

The input vector of PC² determined by its coefficients obtained by constrained OLS contains mixture of M deterministic and random variables $\boldsymbol{\xi} = (\boldsymbol{x}, t, \boldsymbol{\mathcal{X}})$. Deterministic part of this vector (\boldsymbol{x}, t) contains spatial variables and time, therefore it is necessary to filter-out their influence on statistical moments of QoI [3]. In other words, we need to estimate statistical moments only from a reduced basis set containing a subset of the input vector, that is, $\boldsymbol{\mathcal{X}} \subset \{\boldsymbol{x}, t, \boldsymbol{\mathcal{X}}\}$.

3 EFFICIENT SOLVERS AND STATISTICAL SAMPLING

The computational cost of naive PC² implemented in UQPy package [10] grows rapidly with size of KKT matrix and thus it is necessary to use advanced algorithms to efficiently solve large linear least squares problems, such as straightforward updating approach based on Lagrange multipliers (SULM) algorithm summarized in Algorithm 1 [11].

To compare the accuracy and efficiency of the KKT and SULM, both with and without the D-optimality, we consider the following stochastic problem related to the heat equation with Neumann

Algorithm 1 Straightforward updating approach based on Lagrange multipliers

- 1: Solve the unconstrained least squares problem $\|\Psi\tilde{\beta} - Y\|^2$
- 2: Solve $\mathbf{H}\mathbf{J} = -\mathbf{A}^T$
- 3: Set $\mathcal{Y} = \mathbf{A}\mathbf{J}$
- 4: Solve $\mathcal{Y}\lambda = \mathbf{c} - \mathbf{A}\tilde{\beta}$ for λ
- 5: Set $\beta = \tilde{\beta} + \mathbf{J}\lambda$

Output: deterministic coefficients β of PC²

boundary and uniformly distributed coefficient of thermal diffusivity \mathcal{D} :

$$\frac{\partial u(x, y, t)}{\partial t} = \mathcal{D} \left(\frac{\partial^2 u(x, y, t)}{\partial x^2} + \frac{\partial^2 u(x, y, t)}{\partial y^2} \right), \quad x, y \in [0, 1], \quad t \in [0, 1], \quad \mathcal{D} \sim \mathcal{U}[0.001, 0.1] \quad (7)$$

$$u(0, y, t) = u(1, y, t) = u(x, 0, t) = u(x, 1, t) = 0, \quad u(x, 0) = \sin(\pi x)$$

where u represents the temperature field in a 2D spatial domain and 1D time, and \mathcal{D} is the uncertain thermal diffusion coefficient of the medium. For the reference solution, we solve this problem numerically using the finite element method implemented in FEniCS library [12].

To simulate the scenario of scarce or expensive sampling, we do not introduce any additional training set; instead, we increase the number of virtual points n_V to capture information beyond the Neumann boundary and initial conditions in order to investigate convergence of algorithms. The parameters of PCE itself remain unchanged, with the polynomial order ($p = 15$), number of boundary condition points ($n_{BC} = 2000$) and number of initial condition points ($n_{init} = 2000$). Note that, methods don't use any p -adaptivity, q -adaptivity or best model selection algorithms [13].

The absolute errors (AE) and mean squared errors (MSE) for four combinations of algorithms and D-optimal sampling [7] are presented in Fig. 1. When the number of constraints is significantly insufficient, the initial AEs and MSEs of the KKT and KKT-D are lower than those of the SULM and SULM-D. During this under-determined stage, the AE and MSE for all methods initially exhibit an upward trend. As the number of virtual points approaches the fully constrained scenario (well-determined system of linear equations), both AE and MSE begin to decrease sharply (starting from $n_V = 1500$). Notably, for the KKT and KKT-D methods, when the number of virtual points approaches the theoretical requirement for a fully constrained scenario, the AE and MSE basically stabilize near their minimum values. In contrast, the AE and MSE for the SULM and SULM-D methods continue to decline with the increasing number of virtual points, ultimately achieving a more accurate fitted solution. However, as the problem gradually transitions from a fully constrained toward an over-constrained scenario, both AE and MSE obtained by KKT and KKT-D start to diverge.

We also discuss the computational efficiency of the four algorithms. As shown in Fig. 2 (left), for both KKT and SULM methods, the introduction of D-optimality obviously increases the required computation cost, and this increase becomes more pronounced as the number of virtual points increases. Furthermore, compared to the basic and advanced KKT methods, the corresponding SULM methods demonstrate a significant advantage in computational efficiency. In this case, even with the introduction of D-optimality, the computational cost of SULM-D only surpasses that of the standard KKT (without D-optimality) when the number of virtual points exceeds 8000.

If the error requirements are not as strict, as typical for sensitivity analysis (e.g. $\text{MSE} < 10^{-2}$), the KKT and KKT-D methods are more computationally efficient near the fully constrained scenario,

as can be seen in Fig. 2 (right). However, if it is important to obtain higher accuracy, we can obtain more accurate results with less additional computational cost by SULM with appropriately increasing the number of virtual points.

4 COMPARISON TO DEEP OPERATOR NETWORKS

PC^2 can be seen as a physics-informed machine learning (PIML) method together with well-known physics-informed neural networks [14], physics-informed deep operator networks (PI-DON) [15] or physics-informed Gaussian process [16]. In this section, a comparison of PI-DON and PC^2 is presented with respect to the computational efficiency and obtained accuracy. Both techniques are applied to the same numerical example as in the previous section – 2D heat equation with Neumann boundary conditions and random uncertain thermal diffusion coefficient of the medium.

PC^2 is constructed using the standard KKT algorithm with $p = 18$, $n_{BC} = 2 \times 1000$, $n_{initial} = 8000$ and $n_V = 4000$. The architecture of PI-DON is as follows: branch net contains 1, 128, 128, 128, 100 (layers); trunk net contains 3, 128, 128, 128, 100 (layers); $n_{initial} = 20 \times 20$; boundary conditions points $n_{BC} = 20 \times 20$; collocation points $n_V = 10000$; and 120000 epochs with batch size 75. Note that we do not use any data points of real solution obtained by FEM for training. The codes were tested on an NVIDIA A100 GPU. The total computational cost for PC^2 is 284 s and PI-DON 868 s. The

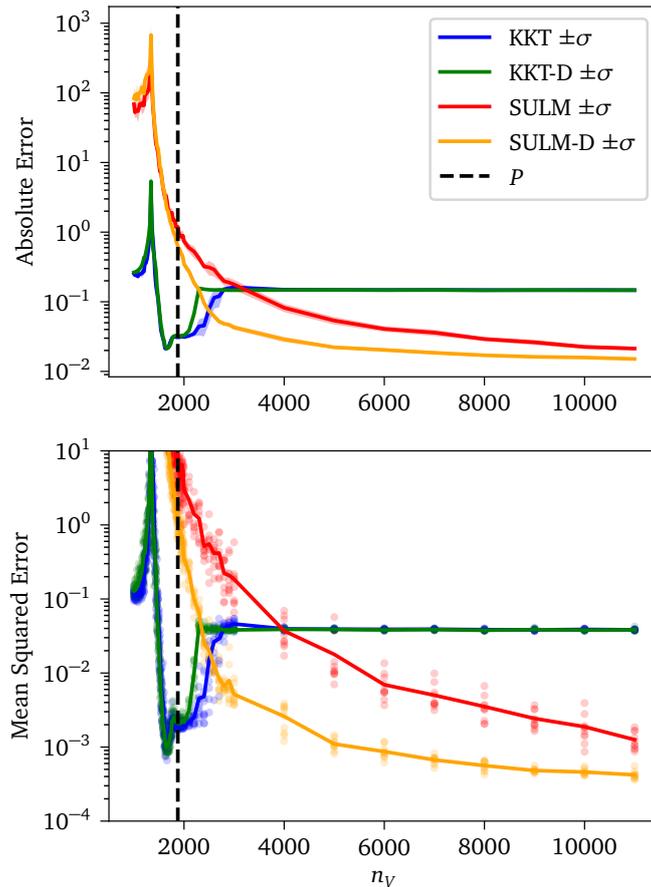


Figure 1: Convergence of the selected algorithms measured by absolute errors (top); and mean squared errors (bottom) for increasing number of virtual points. Results are obtained from 20 statistical replications of the algorithms.

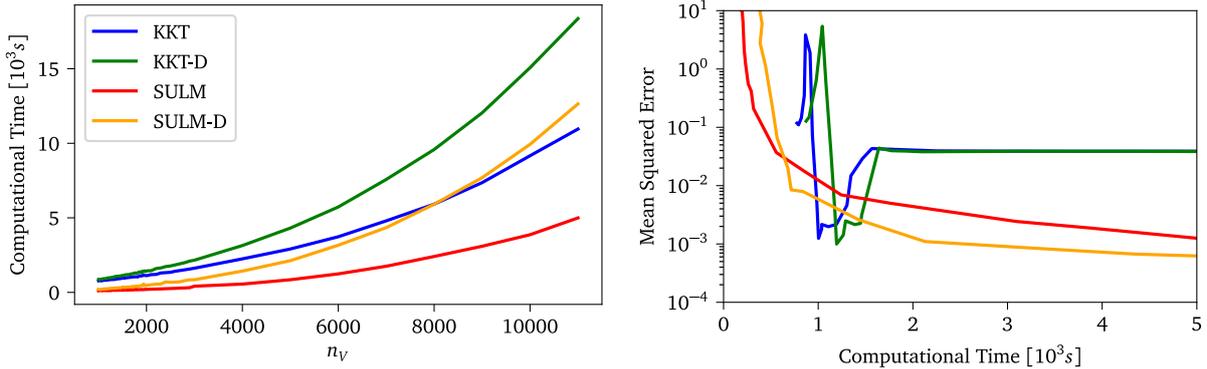


Figure 2: Computational cost of the selected algorithms: computational time for increasing number of virtual points (left); computational time and obtained accuracy (right).

obtained accuracy measured by MSEs evaluated at 100 000 randomly selected points are 4.69×10^{-4} for PC^2 and 4.21×10^{-4} for PI-DON. The solution obtained by both techniques for $\mathcal{D} = 0.009$ can be seen in Fig. 3 for the selected time steps $t = [0.03, 0.33, 0.66, 0.99]$. It can be seen that the errors are similar for both methods except for the boundaries ($t \rightarrow 0$), where polynomial bases of PC^2 lead to greater inaccuracy.

5 DISCUSSION & CONCLUSION

The paper presents recent developments in physics-informed polynomial chaos expansion constructed by constrained least squares. It is shown that computational efficiency can be significantly improved by advanced solvers (e.g. SULM employed in this paper) as well as by optimized statistical sampling of virtual points. Moreover, from the numerical results it can be concluded that optimization of statistical sampling is particularly important in combination with the SULM algorithm in order to accelerate the convergence rate. The convergence graph of the KKT algorithm shows, that it is crucial to create a well-determined system of equations, otherwise the obtained accuracy is significantly lower. Therefore, with respect to computational cost and the obtained accuracy, it can be concluded that standard KKT with a well-determined number of virtual points ($n_V = 2000$ in the numerical example) leads to almost identical results as SULM-D with a significantly higher number of virtual points ($n_V = 4000$), or SULM with $n_V = 11000$. In other words, optimization of number and position of virtual points is crucial for both algorithms SULM and KKT. However, in case of SULM offering low computational cost and stable convergence, it is beneficial to create an overdetermined system of equations contrary to KKT. The obtained numerical results clearly demonstrated the importance of numerical efficiency, and thus future research will be focused on optimization of experimental design including active learning algorithms [17] and the use of advanced numerical solvers.

The second study presents a pilot comparison of PC^2 algorithm to PI-DON in the selected example – 2D heat equation with Neumann boundary conditions and random uncertain thermal diffusion coefficient of the medium. The obtained accuracy measured by MSE, shows that both methods are comparable, with PI-DON yielding slightly higher accuracy. Although the computational cost of PC^2 is $3 \times$ lower in comparison to PI-DON and it also offers an analytical post-processing for UQ [3], PC^2 based on constrained least-squares can only approximate solutions of linear PDEs while PI-DON can also solve non-linear PDEs. Further comparisons will focus on the UQ capabilities of both methods.

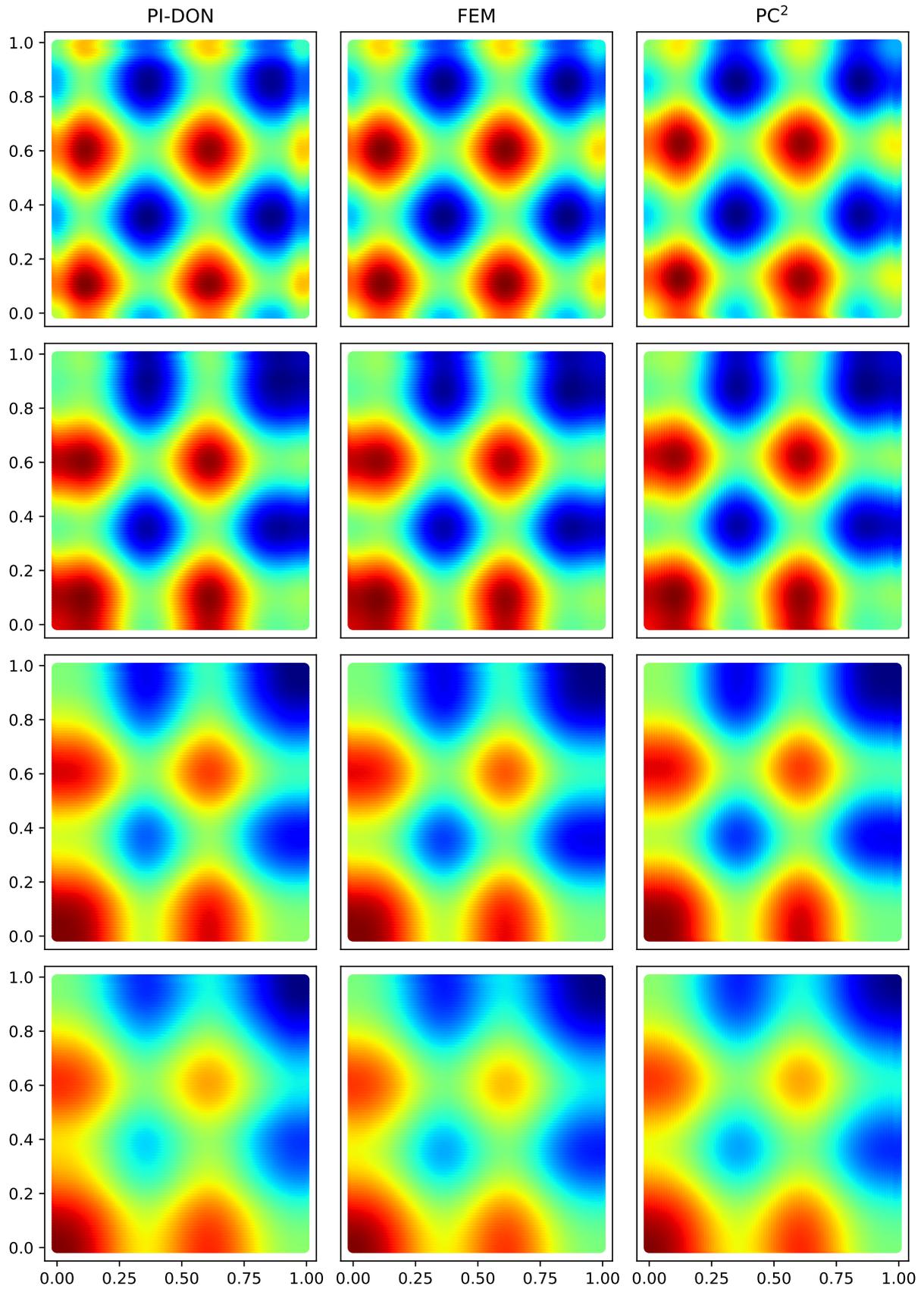


Figure 3: Comparison of PC^2 and PI-DON for the selected time steps $t = [0.03, 0.33, 0.66, 0.99]$

ACKNOWLEDGMENT

Authors gratefully acknowledge the support of the Czech Science Foundation under project No. 24-11845S. The international collaboration was supported by the Ministry of Education, Youth and Sports of the Czech Republic under project No. LUAUS24260.

References

- [1] Soize, Christian and Ghanem, Roger. Physical Systems with Random Uncertainties: Chaos Representations with Arbitrary Probability Measure. *SIAM Journal on Scientific Computing* 26.2 (2004), pp. 395–410. DOI: 10.1137/s1064827503424505.
- [2] Novák, Lukáš. On distribution-based global sensitivity analysis by polynomial chaos expansion. *Computers & Structures* 267 (2022), p. 106808. ISSN: 0045-7949. DOI: 10.1016/j.compstruc.2022.106808.
- [3] Novák, Lukáš, Sharma, Himanshu, and Shields, Michael D. Physics-informed polynomial chaos expansions. *Journal of Computational Physics* 506 (2024), p. 112926. ISSN: 0021-9991.
- [4] Sharma, Himanshu, Novák, Lukáš, and Shields, Michael. Physics-constrained polynomial chaos expansion for scientific machine learning and uncertainty quantification. *Computer Methods in Applied Mechanics and Engineering* 431 (2024), p. 117314. ISSN: 0045-7825.
- [5] Lu, Lu, Jin, Pengzhan, Pang, Guofei, Zhang, Zhongqiang, and Karniadakis, George Em. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence* 3.3 (2021). ISSN: ISSN 2522-5839.
- [6] Borkowski, John J and Valeroso, Elsie S. Comparison of Design Optimality Criteria of Reduced Models for Response Surface Designs in the Hypercube. *Technometrics* 43.4 (2001), pp. 468–477. DOI: 10.1198/00401700152672564.
- [7] Hadigol, Mohammad and Doostan, Alireza. Least squares polynomial chaos expansion: A review of sampling strategies. *Computer methods in Applied Mechanics and Engineering* 332 (2018), pp. 382–407. ISSN: 0045-7825. DOI: 10.1016/j.cma.2017.12.019.
- [8] Ghanem, Roger G. and Spanos, Pol D. *Stochastic Finite Elements: A Spectral Approach*. Springer New York, 1991. ISBN: 0387974563. DOI: 10.1007/978-1-4612-3094-6.
- [9] Xiu, Dongbin and Karniadakis, George Em. The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM Journal on Scientific Computing* 24.2 (2002), pp. 619–644. DOI: 10.1137/s1064827501387826.
- [10] Tsapetis, Dimitrios, Shields, Michael D., Giovanis, Dimitris G., Olivier, Audrey, Novak, Lukas, Chakroborty, Promit, Sharma, Himanshu, Chauhan, Mohit, Kontolati, Katiana, Vandanapu, Lohit, Loukrezis, Dimitrios, and Gardner, Michael. UQpy v4.1: Uncertainty quantification with Python. *SoftwareX* 24 (2023), p. 101561. ISSN: 2352-7110.

- [11] Scott, Jennifer and Tůma, Miroslav. Solving large linear least squares problems with linear equality constraints. *BIT Numerical Mathematics* 62 (2022), pp. 1765–1787. ISSN: 1572-9125.
- [12] Alnæs, Martin, Blechta, Jan, Hake, Johan, Johansson, August, Kehlet, Benjamin, Logg, Anders, Richardson, Chris, Ring, Johannes, Rognes, Marie E, and Wells, Garth N. The FEniCS Project Version 1.5. *Archive of Numerical Software* 3.100 (2015), pp. 9–23. DOI: 10.11588/ANS.2015.100.20553.
- [13] Blatman, Géraud and Sudret, Bruno. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics* 230.6 (2011), pp. 2345–2367. DOI: 10.1016/j.jcp.2010.12.021.
- [14] Raissi, M., Perdikaris, P., and Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378 (2019), pp. 686–707. ISSN: 0021-9991. DOI: <https://doi.org/10.1016/j.jcp.2018.10.045>.
- [15] Goswami, Somdatta, Bora, Aniruddha, Yu, Yue, and Karniadakis, George Em. Physics-Informed Deep Neural Operator Networks. *Machine Learning in Modeling and Simulation: Methods and Applications*. Ed. by Timon Rabczuk and Klaus-Jürgen Bathe. Cham: Springer International Publishing, 2023, pp. 219–254. ISBN: 978-3-031-36644-4. DOI: 10.1007/978-3-031-36644-4_6. URL: https://doi.org/10.1007/978-3-031-36644-4_6.
- [16] Pang, Guofei and Karniadakis, George Em. Physics-informed learning machines for partial differential equations: Gaussian processes versus neural networks. *Emerging Frontiers in Non-linear Science* (2020), pp. 323–343.
- [17] Novák, Lukáš, Vořechovský, Miroslav, Sadílek, Václav, and Shields, Michael D. Variance-based adaptive sequential sampling for Polynomial Chaos Expansion. *Computer Methods in Applied Mechanics and Engineering* 386 (2021), p. 114105. ISSN: 0045-7825. DOI: <https://doi.org/10.1016/j.cma.2021.114105>.