

Multi-fidelity sparse polynomial chaos expansion based on Gaussian process regression and least angle regression

Original

Multi-fidelity sparse polynomial chaos expansion based on Gaussian process regression and least angle regression / Xiao, D., Ferlauto, M., Song, L., Li, J.. - In: JOURNAL OF PHYSICS. CONFERENCE SERIES. - ISSN 1742-6588. - ELETTRONICO. - 1730:(2021), p. 012091. (9th International Conference on Mathematical Modeling in Physical Sciences, IC-MSQUARE 2020 grc 2020) [10.1088/1742-6596/1730/1/012091].

Availability:

This version is available at: 11583/2873614 since: 2021-03-08T20:58:12Z

Publisher:

IOP Publishing Ltd

Published

DOI:10.1088/1742-6596/1730/1/012091

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

PAPER • OPEN ACCESS

Multi-fidelity sparse polynomial chaos expansion based on Gaussian process regression and least angle regression

To cite this article: Dong Xiao *et al* 2021 *J. Phys.: Conf. Ser.* **1730** 012091

View the [article online](#) for updates and enhancements.



240th ECS Meeting ORLANDO, FL

Orange County Convention Center Oct 10-14, 2021



Abstract submission due: April 9

SUBMIT NOW

Multi-fidelity sparse polynomial chaos expansion based on Gaussian process regression and least angle regression

Dong Xiao^{1,*}, Michele Ferlauto², Liming Song¹ and Jun Li¹

¹ Xi'an Jiaotong University, Xianning West Road 28, 710049 Xi'an, China

² Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

*Corresponding author: leafree@stu.xjtu.edu.cn

Abstract. Polynomial chaos (PC) expansion meta-model has been widely employed and investigated in the field of uncertainty quantification (UQ) and sensitivity analysis (SA). However, the majority of the multi-fidelity polynomial chaos expansion (MF-PC) models in the literature are still focused on using high-fidelity (HF) PC model to correct low fidelity (LH) model directly, without cross-correlation between PC models of different fidelities. To address this shortcoming, a multi-fidelity sparse polynomial chaos expansion (MF-sPC) model is proposed based on least angle regression (LAR) and recursive Gaussian process regression (GPR) in this paper. From low to high degree of fidelity, the autoregressive scheme in MF GPR is employed to construct MF-sPC model, in which the sparse polynomial chaos (sPC) model of each fidelity is built iteratively coupling with GPR, LAR and cross validation (CV), as gradually expanding the design of experiment (DoE) to reach a given CV error. This recursive scheme finally yields a MF-sPC model with highest fidelity which takes advantage of all sPC models of the lower fidelities. And the proposed MF-sPC model is validated by a test example in detail, and the results reveal that this MF meta-model performs outstanding both in convergence speed and model accuracy.

1. Introduction

In engineering practice, there are data usually available with multiple fidelities. High-fidelity (HF) models are more accurate but more time-consuming, while low-fidelity (LF) ones are less accurate but more accessible. This has inspired the development of multi-fidelity (MF) models, which aim to build HF models using multi-fidelity data. The first MF model was the multi-fidelity Gaussian process regression (MF-GPR) by Kennedy and O'Hagan [1]. To apply the MF model to uncertainty quantification (UQ) and sensitivity analysis (SA), the multi-fidelity polynomial chaos expansion (MF-PC) model has been well developed in the last decade. Ng and Eldred [2] first put forward the MF-PC model, which was a fully expanded LF model based on stochastic configuration, with HF data to correct the one made from the LF data. On the basis of Ng's work, a few works have been carried out to further explore this corrected MF-PC model. First of all, Eldred [3] compared the MF-PC model constructed by stochastic configuration method and compressed sensing, and demonstrated three cases when the MF-PC model converges more rapidly than a single-fidelity PC. Palar [4] obtained this MF-PC model based on regression method and then applied it to sensitivity analysis. This MF model



successfully passed the tests of several analytic functions and engineering examples and the results revealed that when the correlation R^2 between LF and HF model is large (larger than 0.9), the accuracy of MF model is higher than that of single-fidelity model. In addition, the accuracy of the MF model also depended on the magnitude of the corrections. When the absolute error between the high and low fidelity models was large, the single-fidelity model could be more accurate than MF model. In the meantime, several researchers have constructed MF-PC models in their unique ways, and successfully applied it to different fields, such as inverse problems [5] and aeroelastic flutter analysis [6]. Further on, considering previous corrected MF-PC models ignores the mutual correlation between high and low-fidelity model, Cheng [7] established a MF-PC model based on recursive Gaussian process regression (GPR) to correlate different fidelity PC models, in which an iterative algorithm was proposed to obtain sparse polynomial chaos expansion (sPC) of different levels of fidelity in ascending order. However, this iterative algorithm did not build a high-precision sPC model for each fidelity, and it took lots of iterations to converge. To improve the accuracy of this MF-PC and accelerate convergence, in this paper, an adaptive approach based on least angle regression (LAR) algorithm is added into this MF-PC method to build an effective sPC representation. This adaptive method was first proposed by Blatman and Sudret [8], showing excellent sparse reconstruction of PC model.

Above all, this paper is devoted to develop a multi-fidelity sparse polynomials chaos expansion (MF-sPC) meta-model for arbitrary inputs, in which LAR constructs sPC model of each fidelity, GPR correlates different fidelity sPC models, and cross validation (CV) judges the accuracy of different sPC models and estimates hyper-parameters in the process. The remainder of this paper is organized as follows. The detailed methods to construct MF-sPC model based on GPR, LAR are summarized in section 2. Then section 3 displays a benchmark test in detail to illustrate the outstanding performances of proposed MF-sPC model. In the end, this paper is concluded in section 4.

2. Multi-fidelity sparse polynomial chaos expansions (MF-sPC)

In this section, the corrected multi-fidelity polynomial chaos expansions meta-model is firstly introduced, which is a common way to construct MF-PC. Then we present the sparse polynomial chaos expansions based on GPR and LAR, in which PC models of different fidelities are cross-correlated.

2.1. Corrected multi-fidelity polynomial chaos expansions

In Ng and Eldred's MF-PC, the HF model can be regarded as a combination of the LF model and a correction model

$$y_h = \rho y_l + y_c \quad (1)$$

where y_h , y_l and y_c represent HF, LF and correction term respectively, ρ indicates the scaling factor. In other words, the key point of this method is using HF data to correct LF model. Here we insist that the polynomials in correction model is a subset of the LF PC, because it is convinced the polynomials used in LF PC can approximate the HF model in the same way. The polynomial expansions model can be expressed as

$$y_h = \sum_{i \in \Lambda} (\rho a_i + a_{c_i}) \Phi_i(X) \quad (2)$$

where the multi-index Λ is determined by the hyperbolic truncation of the lowest fidelity sPC model.

2.2. Sparse polynomial chaos expansions based on GPR and LAR

To deduce a MF-PC based on GPR, one must infer a sPC model based on GP at first. And to this end, a covariance function by the inner product of PC basis functions is defined as

$$\text{Cov}(x_i, x_j) = \sum_{\alpha \in \Lambda} \omega_\alpha^2 \Phi_\alpha(x_i) \Phi_\alpha(x_j) \quad (3)$$

where ω_α are weight factors. Then the response is regarded as a Gaussian process (GP) $Z(X) \sim N(\mu(X), Cov(X, X'))$, where $\mu(X)$ and $Cov(X, X')$ is the mean and covariance function of $Z(X)$ respectively. For the sake of simplicity, here we set the mean as zero and the weight factors in covariance function as one. Next given a data set $X_1 = \{x_1, x_2, \dots, x_m\}$ and its evaluation $Y_1 = \{y_1, y_2, \dots, y_m\}$, we can infer that the posterior distribution of $Z(X)$ given X_1, Y_1 is still Gaussian distribution $N(\tilde{\mu}(x), C\tilde{ov}(x, x'))$, in which

$$\begin{aligned}\tilde{\mu}(x) &= Cov(x, X_1)Cov(X_1, X_1)^{-1}Y_1 \\ C\tilde{ov}(x, x') &= Cov(x, x') - Cov(x, X_1)Cov(X_1, X_1)^{-1}Cov(x', X_1)^T\end{aligned}\quad (4)$$

where $Cov(x, x)$, $Cov(x, X_1)$ and $Cov(X_1, X_1)$ are the covariance matrix for given samples in brackets. Further deduction shows that the posterior mean value is a PC model. We hold

$$\tilde{\mu}(x) = Cov(x, X_1)\beta = \sum_{i=1}^m \beta_i \Phi(x_i) \Phi(X) = \alpha \cdot \Phi(X) = \sum_{\lambda \in \Lambda} a_\lambda \Phi_\lambda(X) \quad (5)$$

in which m is the sample size, $\beta = [\beta_1, \beta_2, \dots, \beta_m]^T = Cov(X_1, X_1)^{-1}Y_1$, $\alpha = \sum_{i=1}^m \beta_i \Phi(x_i)$ and $\Phi(X)$ are the vectors of coefficients and orthogonal polynomials respectively.

Cheng in ref [7] adopted an iterative algorithm to update weight factors of covariance function and finally yield the sPC model. However, this iterative algorithm did not build a high-precision sPC meta-model for each fidelity, and it took lots of iterations to converge. To improve the accuracy of MF-PC and accelerate convergence, an adaptive approach based on LAR algorithm is employed to build an effective sparse representation and then inspire the posterior distribution of $Z(X)$. The whole algorithm for constructing sPC model based on LAR and GPR is given in pseudocode below.

Algorithm 1. sPC model based on LAR

- 1: **Initialization.** $K = 10, m_0 = 10d, \Delta m = \text{int}(0.25m_0 \sim 0.3m_0), \bar{\varepsilon}_{CV} = 0.001, m = m_0$
- 2: **while** stop-sampling conditions are not met **do**
- 3: **for** $p = 1, 2, \dots, p_{\max}$ **do**
- 4: **for** $q = 0.1, 0.2, \dots, 1$ **do**
- 5: Run LAR for a set of PCE model and compute coefficients by least square method, calculate ε_{CV} for each PCE and find the one with minimal ε_{CV}^{\min}
- 6: Discard some (p, q) combinations with large ε_{CV} consecutively
- 7: $m = m + \Delta m$
- 8: Record (p, q) , polynomials coefficients α and ε_{CV} for the best model selected

Algorithm 2. posterior distribution of $Z(X)$ given X_1, Y_1 , with mean of final sPC model

- 9: $k = 1, u = |\text{sign}(\alpha)|, Cov^{(1)}(x_i, x_j) = u \cdot \Phi(x_j) \cdot u \cdot \Phi(x_j), \beta = Cov^{(1)}(X_1, X_1)^{-1}Y_1,$
 $\alpha^{(1)} = \sum_{i=1}^m \beta_i \Phi(x_i)$ and $l_1^{(1)} = \|\alpha^{(1)}\|, \delta = 1$
 - 10: **while** $\delta > 0.001$ **do**
 - 11: $k = k + 1, u = \prod_{i=1}^{k-1} a^{(i)}, Cov^{(k)}(x_i, x_j) = u \cdot \Phi(x_j) \cdot u \cdot \Phi(x_j), \beta = Cov^{(k)}(X_1, X_1)^{-1}Y_1$
 $\alpha^{(k)} = \sum_{i=1}^m \beta_i u^2 \Phi(x_i), l_1^{(k)} = \|\alpha^{(k)}\|, \delta = |l_1^{(k)} - l_1^{(k-1)}| / l_1^{(k-1)}$
 - 12: **Output.** $\alpha = \alpha^{(k)}, Cov(x_i, x_j) = Cov^{(k)}(x_i, x_j) = u \cdot \Phi(x_j) \cdot u \cdot \Phi(x_j), \tilde{\mu}(x), C\tilde{ov}(x, x')$
-

2.3. Multi-fidelity sparse polynomial chaos expansions (MF-sPC)

Consider the autoregressive model $Z_{t+1}(X) = \rho_t Z_t(X) + g_{t+1}(X)$ in MF GPR, and give multiple data sets $X_1, X_2 \dots X_s$ and their realization $Y_1, Y_2 \dots Y_s$, initialize the GP $g_{t+1}(X)$ with zero mean and its weight factors ω_k is equal to 1, then it can be deduced that the HF model $Z_2(X) \sim N(\mu_2(X), Cov_2(X, X'))$ is a GP, and its inferior mean and covariance are expressed as

$$\begin{aligned}\mu_2(x) &= \rho_1 \tilde{\mu}_1(x) \\ Cov_2(x, x') &= \rho_1^2 C\tilde{ov}_1(x, x') + Cov_{g_2}(x, x')\end{aligned}\quad (6)$$

These two equations work to map the mean and covariance from model of low to high fidelity. Given training data set X_2 and its realization Y_2 , the posterior distribution of $Z_2(X)$ is still a GP, whose posterior mean and covariance are expressed as

$$\begin{aligned}\tilde{\mu}_2(x) &= \rho_1 \tilde{\mu}_1(x) + Cov_2(x, X_2)Cov_2(X_2, X_2)^{-1}(Y_2 - \rho_1 \tilde{\mu}_1(x)) \\ C\tilde{ov}_2(x, x') &= Cov_2(x, x') - Cov_2(x, X_2)Cov_2(X_2, X_2)^{-1}Cov_2(x', X_2)^T\end{aligned}\quad (7)$$

Note the weight factors of GP $g_{t+1}(X)$ is updated by the sPC algorithm given above, and the posterior mean $\tilde{\mu}_2(x)$ in Eq. (7) of final HF model $Z_2(X)$ is still a sPC model. Following this procedure, a sparse MF-PC meta-model can be constructed recursively from a low to high level of fidelity.

Another key point of this algorithm is how to estimate the hyper-parameters. For a problem with s fidelities, one must determine the scaling factor ρ_i and maximal degree of polynomials p_i for each fidelity. In sum, the hyper-parameters $p_1, \rho_1, p_2, \rho_2 \dots p_{s-1}, \rho_{s-1}, p_s$ with a number of $2s-1$ need to be estimated gradually. As we said before, the polynomials in HF model is a subset of that in the LH PC, thus this inequality $p_1 \geq p_2 \geq \dots \geq p_{s-1} \geq p_s$ is valid. In the present study, CV is employed to estimate the optimal hyper-parameters, with which the current MF-PC model has the minimal CV error.

3. Benchmark tests

In this section, an eight-dimensional borehole function is employed to examine the capabilities of the multi-fidelity method in detail, from the perspective of convergence and accuracy.

The borehole function models water flow through a borehole [4], and the HF and LF functions are expressed as

$$f_h(x) = \frac{2\pi T_u(H_u - H_l)}{\ln(r/r_\omega) \left(1 + \frac{2LT_u}{\ln(r/r_\omega)r_\omega^2 K_\omega} + \frac{T_u}{T_l}\right)}\quad (8)$$

$$f_l(x) = \frac{5T_u(H_u - H_l)}{\ln(r/r_\omega) \left(1.5 + \frac{2LT_u}{\ln(r/r_\omega)r_\omega^2 K_\omega} + \frac{T_u}{T_l}\right)}\quad (9)$$

respectively, where the various parameters and their distributions can be found in ref. [4]. To analyse the similarity and discrepancy of those two-fidelity functions, the R^2 correlation with a value of 0.9999 indicates that these functions are highly correlated, while the mean absolute relative error $MARE$ is equal to 0.204 means that the LF function is extremely inaccurate.

To dealing with this problem, the first thing is to construct an accurate LH sPC model, in which the number of samples required and the optimal truncation parameters p, q need to be estimated.

Execute Algorithm 1 to construct sPC, we obtain the convergence curves of error as increasing the number of samples in figure 1.

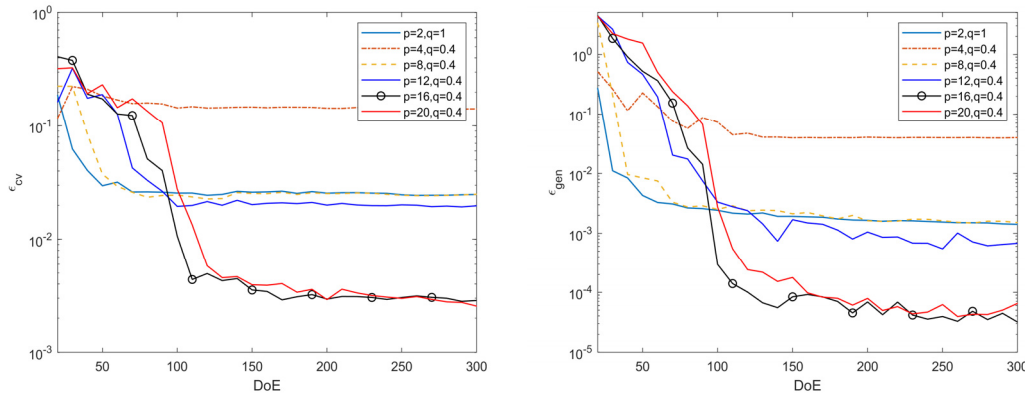


Figure 1: Relative K-fold CV error and relative generalization error for different truncated sets

As expanding the size of design of experiment (DoE), the relative K-fold CV error ϵ_{CV} and relative generalization error ϵ_{gen} for all truncated sets begin decreasing sharply as first, until the sample number m_1 reaches 100, then errors declines slowly and the algorithm converges. The polynomial truncated set generated by truncated parameters $p=16, q=0.4$ proves to construct the most accurate sPC model among those candidate sets when $m \geq 100$, which is adopted ultimately in this paper. For the truncated parameters $p=16, q=0.4$, then LAR algorithm meets its convergence conditions when $m_1 = 140$, at this point $\epsilon_{CV} = 0.0045$, $\epsilon_{gen} = 5.5089e-05$.

Furthermore, the trends of the K-fold CV error ϵ_{CV} are consistent with that of generalization error ϵ_{gen} for all truncated sets while increasing samples, which indicates CV is an efficient approach to select the optimal sPC model and truncation parameters. We must emphasize that the relative generalization error ϵ_{gen} cannot be acquired in the engineering practice due to lack of information about the true model. In the case that ϵ_{gen} is unobtainable, it is proved that ϵ_{CV} can work as excellent metrics of the model accuracy in this algorithm.

When the best PC model is found, then run Algorithm 1 to obtain the GP $Z_1(X)$, as well as its ultimate weight factors and the covariance function. Next we construct the GP $Z_2(X)$ by the autoregressive model $Z_2(X) = \rho_1 Z_1(X) + g_2(X)$, in which the hyper-parameters including scaling factor ρ_1 and maximal degree of polynomial p_2 are estimated by CV, as shown in figure 2.

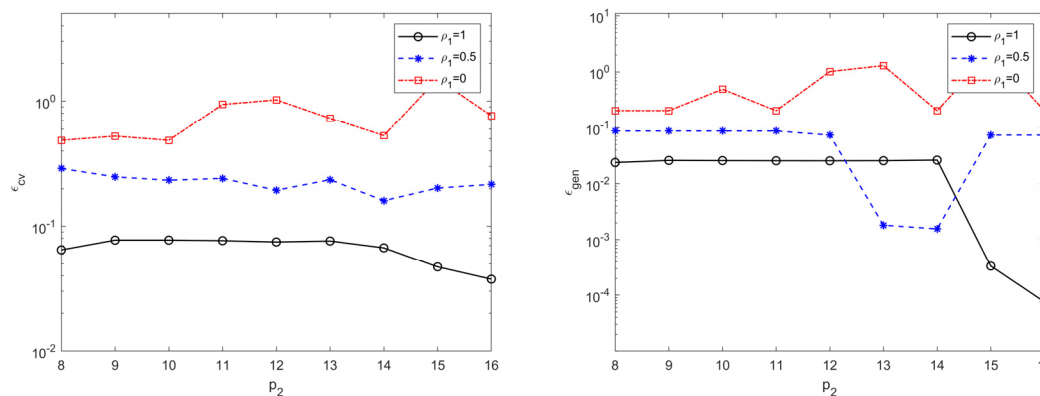


Figure 2: Relative K-fold CV error and relative generalization error for different hyper-parameters

Figure 2 depicts the trend of relative K-fold CV error ε_{CV} and relative generalization error ε_{gen} for various scaling factor ρ_1 and maximal degree of polynomial p_2 when number of LF and HF samples are set as $m_l = 140, m_h = 40$ respectively. Here we discover that the general trend of ε_{CV} is in line with the trend of ε_{gen} for different ρ_1 while varying p_2 . When ε_{CV} reaches its least value, the ε_{gen} is exactly at its minimum. And in this test, the optimal hyper-parameters are found to be $\rho_1 = 1, p_2 = 16$, by this time ε_{gen} is equal to $7.5492e-05$.

Hereto, the MF-sPC model has been constructed successfully, on the strength of LAR and GPR. To further analyse the effects of LAR and GPR, firstly a comparative study of convergence for two methods named MF-PC with LAR (LAR-MF) and original MF-PC without LAR (Ori-MF) is carried out. Figure 3 illustrates the convergence for Algorithm 2 for sPC model of each fidelity when given $m_l = 400, m_h = 80$.

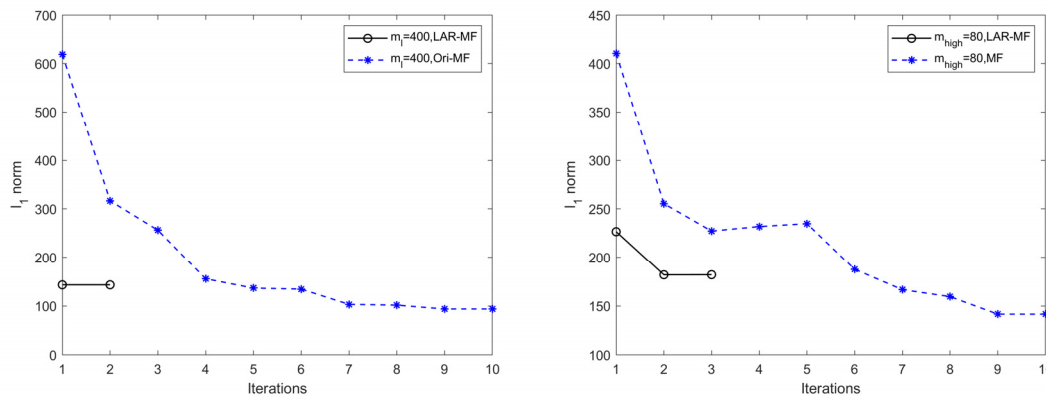


Figure 3: Convergence of Algorithm 1 for PC model of each fidelity

It is obvious that LAR accelerates the convergence of the whole algorithm, the convergent iterations decrease from 10 to 2 and 3 for low and HF sPC model respectively. And the various convergent values of l_1 norm indicate model accuracies differ for LAR-MF and Ori-MF PC model. Here figure 4 gives the accuracy of the sPC model generated by three methods as increasing m_h for given $m_l = 400$: LAR-MF, LAR and Ori-MF, in which LAR means using LAR construct sPC solely.

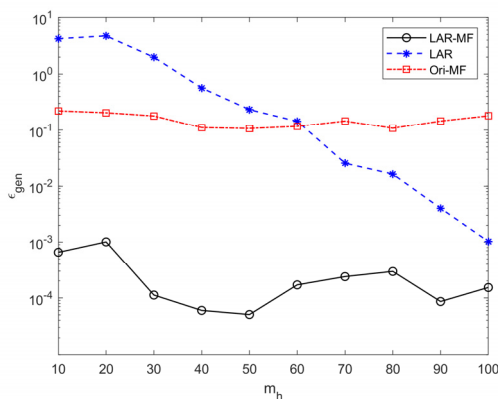


Figure 4: Error convergence for three different sPC methods as increasing m_h when fixing m_l

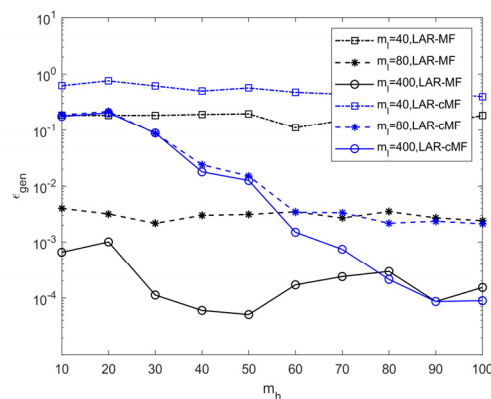


Figure 5: Error Convergence for two multi-fidelity models and different m_l as increasing m_h

According to figure 4, when constructing sPC model, it is revealed that Ori-MF method cannot reduce the ε_{gen} observably, while only using LAR algorithm based on HF samples decreases ε_{gen} significantly

as increasing m_h for given m_l . However, Ori-MF behaves better when m_h is small because this method takes advantage of the accurate LH sPC model while LAR does not. Therefore, LAR-MF combines the advantages of LAR and Ori-MF effectively and performs outstanding for all the different values of m_h , which is exactly what we want. And we can conclude that LAR speeds up the decline of ε_{gen} as increasing m_h . Another phenomenon is ε_{gen} fluctuates along with m_h for LAR-MF method within a pretty small value, this volatility may stem from diverse hyper-parameters ρ_1 and p_2 for different m_h .

Finally, to illustrate the impact of GPR, in other words, to show the effect of the cross-correlation between sPC of different fidelity, the performances of proposed MF-PC based on LAR and GPR (LAR-MF) with corrected MF-PC based on LAR(LAR-cMF) are compared for three different m_l as increasing m_h in figure 5. The three different $m_l = 40, 80, 400$ represent three LF sPC models whose ε_{gen} are equal to 0.4614, 0.5457e-2, 0.4892e-4 respectively. It can be found that even when the LF sPC models are different, our proposed LAR-MF method is able to construct a more accurate MF-PC model than LAR-cMF, especially when m_h is small. And the accuracy of LF sPC model influences the accuracy of MF-PC model for both methods, as we can see, a high-accuracy LF sPC model determines a high-precision MF-PC model.

4. Conclusions

In this paper, a multi-fidelity sparse polynomial chaos expansion meta-model based on GPR, LAR and CV has been proposed for dealing with UQ problems, in which GPR guarantees the cross-correlation between sPC models of different levels of fidelity, LAR accelerates the algorithm convergence and improves model accuracy, meanwhile CV estimates the accuracy of sPC model and determines the hyper-parameters in the algorithm. The body of this algorithm is to construct sPC model by regarding it as GP, then the MF-sPC meta-model can be obtained by implementing autoregressive model from low to high fidelity.

The benchmark test illustrates that the proposed method is able to construct high-accuracy MF-PC meta-model, even when the HF samples are scarce. Compared with corrected MF-PC method, the proposed method constructs more accurate MF-PC meta-model due to the cross-correlation between sPC models of different fidelities. In comparison with original MF-PC method based on GPR, the introduction of LAR algorithm successfully promotes the algorithm convergence as well as model accuracy.

5. References

- [1] Kennedy M C and O'Hagan A 2000 *Biometrika* **87** 1-13
- [2] Ng L W T, Eldred M 2012 *53rd AIAA/ASME/SAE Struct., Struct. Dyn., Mater. Conf.* 1852
- [3] Eldred M S, Ng L W T, Barone M F and Domino S P 2015 No. SAND2015-1803B SNL-NM
- [4] Palar P S, Tsuchiya T and Parks G T 2016 *Comput. Meth. Appl. Mech. Eng.* **305** 579-606
- [5] Yan L, Zhou T 2019 *J. Comput. Phys.* **381** 110-28.
- [6] Rumpfkeil M P, Bryson D E and Beran P S 2019 *AIAA Scitech forum* 1998
- [7] Cheng K, Lu Z and Zhen Y 2019 *Comput. Meth. Appl. Mech. Eng.* **349** 360-77
- [8] Blatman G and Sudret B 2011 *J. Comput. Phys.* **230** 2345-67

Acknowledgments

This work was supported by the Graduate Mobility Program of Xi'an Jiaotong University.