Sensitivity Analysis for Model with Dependent Inputs Using Sparse Polynomial Chaos Expansions

(Article begins on next page)

14 May 2024

SENSITIVITY ANALYSIS FOR MODEL WITH DEPENDENT INPUTS USING SPARSE POLYNOMIAL CHAOS EXPANSIONS

DONG XIAO[1], MICHELE FERLAUTO[2], LIMING SONG[1] AND JUN LI[1]

[1]Xi'an Jiaotong University
Xianning West Road 28, CHN-710049 Xi'an
leafree@stu.xjtu.edu.cn

[2]Politecnico di Torino
Corso Duca degli Abruzzi 24, IT-10129 Torino
michele.ferlauto@polito.it

Key words: Sensitivity Analysis, Dependent Inputs, Polynomial Chaos Expansions, Least Angle Regression, Weighted Leja Sequences, Cross Validation.

Abstract. Polynomial chaos expansions (PCE) meta-model has been wildly used and investigated in the last decades in sensitivity analysis (SA), which adopts a variety of orthogonal polynomials to approximate the system response and calculates sensitivity indices directly from the polynomial coefficients. The Sobol' index is one of prevalent sensitivity indices for model with independent inputs and can be easily obtained after constructing generalized polynomial chaos (gPC). But for dependent inputs, a typical approach is based on the procedure of transforming the dependent inputs into independent inputs according to the literature. This paper demonstrates a global sensitivity analysis (GSA) approach for dependent inputs, in which Gram-Schmidt orthogonalization (GSO) numerically computes the orthonormal polynomials for PCE. The especial procedure for dependent inputs to obtain sensitivity indices lies in the linearly independent polynomials basis for GSO must be in an intended order. Besides, to alleviate the curse of dimensionality, the sparse polynomial chaos (sPC) is built coupling with least angle regression (LAR) and a nested experimental design called weighted Leja sequences (wLS). Then cross validation (CV) determines the best truncated set for sPC with the suitable size of experimental design in use. In the end, this proposed approach is validated on a benchmark function with dependent inputs. The results reveal that the proposed approach performs well to calculate sensitivity indices for model with dependent inputs.

1 INTRODUCTION

This paper aims to proposes a global sensitivity analysis (GSA) approach for dependent inputs based on sparse polynomial chaos (sPC). To implement this method, a data-driven method to computes the orthonormal polynomials for PCE model, a generalization of the Sobol' indices for dependent inputs and an algorithm for construct sPC to alleviate the curse of dimensionality are adopted in succession.

As we know, generalized polynomial chaos (gPC) based on Askey scheme[1] only suits for

independent inputs which requires the full knowledge of the distributions of inputs to select the proper polynomials. But for dependent inputs, a typical method is mapping dependent random inputs to independent ones, such as such as Nataf[2-3] and Rosenblatt[4] transformation. Nevertheless, mapping methods can significantly degrade performance since the Jacobian of the map must be approximated. Another known method is Gram-Schmidt orthogonalization (GSO), which can directly compute orthonormal polynomials for arbitrary inputs from the discrete input data. And it was proved that the polynomials for independent variables generated by GSO using monic polynomials accord with Askey scheme, which assures the accuracy of GSO to some extent.

When orthonormal polynomials are generated by GSO, the sensitivity indices can be calculated by algebraic operations of polynomial coefficients of PCE. However, due to the correlation between input variables, Sobol' indices are not applicable to dependent inputs. Kucherenko[5], Caniouhave[6] and Mara[7-8] proposed three different generalizations of the Sobol' indices. Kucherenko directly decomposed the variance and evaluated sensitivity indices with a double loop Monte Carlo estimation. Caniou have adopted a High Dimensional Model Representation (HDMR) as structural substitute of the actual model and regarded input variables as independent, but for computing sensitivity indices, the interactive and correlative effects of inputs had been taken into consideration. Mara defined the full sensitivity indices and the uncorrelated sensitivity indices for dependent inputs, and those sensitivity indices could be solved by reverse Rosenblatt transformation when conditional probability density functions of the inputs are known. Based on Mara's definition of sensitivity indices for dependent inputs, Liu[9] applied GSO to generate orthogonal polynomials in an intended order based on observed data of model inputs and outputs, which is also adopted in this paper.

Usually, the projection[10] and regression[11] methods are two main approaches computing expansion coefficients of PCE, both suffering from curse of dimensionality badly. To alleviate this symptom, Blatman and Sudret[12] adopted regression method and firstly attempted to deal with the efficient sparse representation by a hyperbolic scheme and least angle regression (LAR) algorithm. Based on Askey scheme, they successfully applied sPC to computing sensitivity indices for independent inputs[13]. Inspired by their pioneering work and a nested sampling strategy called weighted Leja sequences (wLS)[14], we combine these two techniques and propose an algorithm to construct sPC gradually. The key ideal of this algorithm is sampling continuously until stop-sampling conditions about error of PCE surrogate are reached, in which cross validation (CV) is adopted to appraise the accuracy of PCE surrogates.

In conclusion, the main idea of this paper is constructing sPC model generated by GSO, LAR and wLS, while CV determines an appropriate size of experimental design and the optimal sPC model. After the high-precision surrogates are built, the generalized sensitivity indices defined in [7] and can be obtained by algebraic operations of polynomial coefficients. The remainder of this paper is organized as follows. Section 2 introduces a generalization of the Sobol' indices for arbitrary inputs. The GSO algorithm to compute orthogonal polynomials and its corresponding order for calculating sensitivity indices for dependent inputs are summarized in Section 3. Section 4 proposes an adaptive algorithm to obtain accurate sparse representation of PCE based on LAR, wLS and CV. Section 5 displays a benchmark tests with dependent inputs for our proposed approach. In the end this paper is concluded in Section 6.

2 SENSITIVITY INDICES

2.1. Probabilistic formulation of the problem

Define a probability space $(\Omega, \mathcal{F}, P)$ with sample space Ω, σ-algebra $\mathcal{F}$ and probability measure P. Let f: be a function with a d-variate random variable $X = (X_1, X_2 ... X_d)$, where $f_X(X)$ denotes the joint probability density function of the d-variate random input variables. For an input vector $x = (x_1, x_2 ... x_m)$, their realizations are denoted by $y = (y_1, y_2 ... y_m)$. Note m that represents the number of samples in the random input vector. And remember that the capital letter of X, Y denotes random variables and the response, while lowercase of x, y indicates input vector and output vector.

2.2. A generalization of the Sobol' indices

The original idea behind the Sobol' indices is to represent the model as a sum of component functions with increasing dimensionality, in which those component functions satisfy a few specific properties. For independent variables X and a square-integrable model M with finite variance there exists a unique decomposition

$$Y = M(X) = M_0 + \sum_{i=1}^{d} M_i(X_i) + \sum_{1 \le i \le j \le d} M_{ij}(X_i, X_j) + ... + \sum M_{12...d}(X) \tag{1}$$

The uniqueness of the decomposition is ensured by those two properties:

$$M_0 = \int_{\Omega_X} M(X) f_X(X) dx \tag{2}$$

$$\int_{\Omega_{X_k}} M_{i_1,...i_s}(X_{i_1}, ..., X_{i_s}) f_{X_k}(X_k) dx_k = 0, \ 1 \le i_1 < \cdots < i_s \le d, k \in \{i_1, ..., i_s\} \tag{3}$$

where Ω_X denotes the support of the random variables X while Ω_{X_k} is the support with respect to variables in set k.

The variance of model M can be deduced from above unique decomposition

$$Var(Y) = \int_{\Omega_X} M(X) f_X(X) dx - M_0^2 = \sum_{u \subseteq \{1,2,...,d\}} D_u(Y) \tag{4}$$

where

$$D_u(Y) = Cov(M_u(X_u), Y) \tag{5}$$

Other than the unique decomposition of independent inputs, the component functions of dependent inputs are not unique and vary from the order of inputs variables. In addition, independent inputs hold $D_i(Y) = Cov(M_i(X_i), Y) = Var(M_i(X_i))$, while one can divide the $D_i(Y)$ into two parts when correlations between inputs exist:

$$D_i(Y) = Cov(M_i(X_i), Y) = Var(M_i(X_i)) + Cov(M_i(X_i), Y - M_i(X_i)) \tag{6}$$

To calculate the sensitivity indices, just divide the equation by model variance $Var(Y)$

$$S_i = \frac{Cov(M_i(X_i), Y)}{Var(Y)} = \frac{Var(M_i(X_i))}{Var(Y)} + \frac{Cov(M_i(X_i), Y - M_i(X_i))}{Var(Y)} = S_i^u + S_i^c \tag{7}$$

The second term in above equation isn't equal to zero when some variables are dependent with X_i. These two summands are known as uncorrelated (or structural) and correlative sensitivity indices of X_i, represented by S_i^u, S_i^c. In [7], the sensitivity indices for a given

variable X_i are defined by the first-order full sensitivity indices S_i, total full sensitivity indices ST_i, first-order uncorrelated sensitivity indices S_i^u and total uncorrelated sensitivity indices ST_i^u. Just as the S_i, S_i^u given in above equation, the ST_i, ST_i^u also can be interpreted as

$$ST_i = \sum_{i \in u} \frac{Cov(M_u(X_u),Y)}{Var(Y)} = \sum_{i \in u} \frac{Var(M_u(X_u))}{Var(Y)} + \sum_{i \in u} \frac{Cov(M_u(X_u),Y-M_u(X_u))}{Var(Y)} = ST_i^u + ST_i^c \tag{8}$$

in the same way. It's worth mentioning that the uncorrelated sensitivity indices are same as Sobol' indices for independent inputs.

3 POLYNOMIAL CHAOS EXPANSIONS FOR DEPENDENT INPUTS

3.1. Gram-Schmidt orthogonalization

For Gram-Schmidt orthogonalization to compute orthogonal polynomials, first one needs to determine a set of linearly independent basis polynomials. To assure GSO generates the same orthogonal polynomials with Askey scheme, the set of monic polynomials are adopted

$$\varphi_j(X) = \prod_{i=1}^{d} X_i^{\alpha_i^j}, \ j = 0,1...N, \ \sum_{i=1}^{d} \alpha_i^j \le p \tag{9}$$

in which d is the dimension of random space, p is the maximal order of monic polynomials and α_i^j is the multi-index. Then the orthogonal polynomials are numerically constructed according to the Gram-Schmidt algorithm:

$$\Phi_0(X) = 1$$
$$\Phi_j(X) = \varphi_j(X) - \sum_{k=0}^{j-1} c_{jk} \Phi_k(X), 1 \le j \le P \tag{10}$$

the P denotes the number of basis functions determined by the number of monic polynomials, the coefficients c_{jk} are defined by

$$c_{jk} = \left\langle \varphi_j(X), \Phi_k(X) \right\rangle / \left\langle \Phi_k^2(X) \right\rangle \tag{11}$$

where the $\langle \ \rangle$ means inner product.

3.2. Polynomial chaos representation

While orthogonal polynomials with regard to input random variables X are computed by GSO, the model response Y can be represented by a linear combination of these polynomials, which is called PCE meta-model. In practice, the PCE should be truncated into a finite set of polynomials when output random variable is a second-order variable which satisfies $E(Y^2) < +\infty$. This truncation set is defined by a multi-index $\Lambda = N_o^d$ and the truncated PCE can be expressed as

$$Y = \sum_{\lambda \in \Lambda} a_\lambda \Phi_\lambda(X) \tag{12}$$

A simple and commonly applied truncation approach is hyperbolic truncation [51] based on $q-norm$:

$$\Lambda = \Lambda_{p,q}^d = \left\{ \lambda, \|\lambda\|_q \le p \right\}, \ \|\lambda\|_q = \left(\sum_{i=1}^{d} \lambda_i^q \right)^{1/q} \tag{13}$$

4

$0 < q \leq 1$ is a tuning parameter and p is the maximal total degree of the polynomials. A decreasing q leads to a smaller set of polynomials as well as a smaller number of interactive polynomials. When $q = 1$ results in a total-degree space where $\Lambda = \Lambda_{p,1}^{d} = C_{p+d}^{p}$. For any Λ, there exists a corresponding PCE and the choice of Λ determines the required number of samples for given PCE accuracy.

3.3. Computation of the polynomial chaos coefficients

This paper adopts regression method for solving a system of linear functions by least-square minimization. For an input vector $x = (x_1, x_2 \ldots x_m)$, their realizations are denoted by $y = (y_1, y_2 \ldots y_m)$, the expansion coefficients are calculated by minimizing the expectation of the least-squares residual:

$$\hat{a} = \arg \min_{a \in \mathbb{R}^{|\Lambda|}} \frac{1}{m} \sum_{i=1}^{m} (y_i - \sum_{\lambda \in \Lambda} a_\lambda \Phi_\lambda(X)) \tag{14}$$

The coefficients $\hat{a}$ are the least-square solution to the linear system

$$\hat{a} = (\phi^T \phi)^{-1} \phi^T y, \; \phi_{ij} = \Phi_j(x_i) \tag{15}$$

where ϕ is a matrix with m rows and P columns, with m denoting number of samples and P number of basis function. The ϕ_{ij} corresponds to the value of the i-th point on the j-th basis function.

3.4. Estimators of accuracy for the polynomial chaos approximations

In terms of PCE meta-model tends to overfit when PCE contains high order polynomials for a simple objective function with a limited number of samples, the cross validation (CV) [53] has been come up with to estimate the generalization error. Suppose $B = \phi^T \phi$, the relative K-fold CV error is defined as

$$\varepsilon_{CV} = \sqrt{\sum_{k=1}^{K} (B_{[k]} a_{[-k]} - y_{[k]})^2} \, / \, \|y\|_2 \tag{16}$$

where $B_{[k]}$ represents B contains the corresponding rows in set k, $a_{[-k]}$ denotes the coefficient computed without considering set k, while $y_{[k]}$ indicates the vector y without corresponding response value of set k, the $\|y\|_2$ represents the 2-*norm* of vector y.

3.5. Calculation of generalized sensitivity indices for dependent inputs

In this part, the orders of monic polynomials for GSO are introduced to compute full and uncorrelated sensitivity indices. As mentioned in [7], suppose the initial input variables are ordered as $(X_i, X_{i+1}, \ldots X_d, X_1, X_2, \ldots, X_{i-1})$, and their transformation independent variables are $(\bar{X}_i, \bar{X}_{i+1}, \ldots \bar{X}_d, \bar{X}_1, \bar{X}_2, \ldots, \bar{X}_{i-1})$, then the full sensitivity indices $(S_i = S_{\bar{i}}, ST_i = ST_{\bar{i}})$ and the uncorrelated sensitivity indices $(S_i^u = S_{\overline{i-1}}^u, ST_i = ST_{\overline{i-1}})$ are given. It inspires us to design an intended order of linearly independent functions for GSO to calculate sensitivity indices.

3.5.1. Full sensitivity indices

To compute full sensitivity indices of X_i, we place all the linearly independent functions

$\varphi_j(X), j \in \{1, 2, ..., N\}$ in equation 9 relating to X_i in the front, of which $\alpha_i^j > 0$. The order the linearly independent function $\varphi_j(X)$ can be represented as

$$S = (St_0, St_{11}, St_1 - St_{11}, St_2, St_3, ..., St_d) \tag{17}$$

where $St_0 = \varphi_0(X) = 1$. The definition of St_{11} and St_i are as follow:

$$St_{11} = \left\{ \varphi_j(X), 0 < \alpha_1^j \le p, \alpha_{i \ne 1}^j = 0 \right\}$$
$$St_i = (S - \bigcup_{j=0}^{i-1} St_j) \bigcap \left\{ \varphi_j(X), \alpha_i^j > 0, \alpha_i^j \le p \right\} \tag{18}$$

where St_{11} is compose of functions which only relate to X_1, and $St_1 \supseteq St_{11}$ is the functions which as long as involve X_1, while $St_i, i \ge 2$ contains all the functions only relating to X_i and the interaction terms between $(X_i, X_{i+1}, ..., X_d)$.

Based on this, the corresponding PCE can be acquired with GSO, the sparse representation of PCE and its coefficients a of orthogonal polynomials are obtained with LAR and wLS. Therefore, the first-order full sensitivity indices S_1 and the total full sensitivity indices ST_1 are defined as

$$S_1 = \frac{\sum_{j \in St_{11}} a_j^2}{Var(Y)}, \quad ST_1 = \frac{\sum_{j \in St_1} a_j^2}{Var(Y)} \tag{19}$$

By permuting the order of the inputs, such as to calculate S_i, ST_i, the inputs can be adjusted into $(X_i, X_{i+1}, ..., X_d, X_1, X_2, ..., X_{i-1})$, the different sensitivity indices for each variable can be further calculated.

3.5.2. Uncorrelated sensitivity indices

In order to calculate the uncorrelated sensitivity indices of X_i, we place the linearly independent functions relating to X_i in the end. The order the linearly independent function for calculating uncorrelated sensitivity indices can be defined as

$$S = (St_0, St_{-1}, St_{11}, St_1 - St_{11}) \tag{20}$$

where $St_0 = \varphi_0(X) = 1$, St_{-1} refers to all the functions that are irrelevant with X_1, St_{11} and St_1 are defined hereinabove.

Therefore, the first-order uncorrelated sensitivity indices S_1^u and total uncorrelated sensitivity indices ST_1^u can be computed by

$$S_1^u = \frac{\sum_{j \in St_{11}} a_j^2}{Var(Y)}, \quad ST_1^u = \frac{\sum_{j \in St_1} a_j^2}{Var(Y)} \tag{21}$$

In a same way, the uncorrelated sensitivity indices variable X_i can be calculated by permuting the order of the inputs from $(X_1, X_2, ..., X_d)$ to $(X_i, X_{i+1}, ..., X_d, X_1, X_2, ..., X_{i-1})$.

Note that uncorrelated sensitivity indices are not always less than corresponding full sensitivity indices since some interactions of inputs are likely to counteract their effects on the model variance.

4 SPARSE POLYNOMIAL CHAOS BASED ON LEAST ANGLE REGRESSION

4.1. Least angle regression

This paper adopts hyperbolic truncation to generate non-empty finite subset Λ of $\mathbb{N}^d$, corresponding to truncated polynomial chaos. We aim at selecting those predictors which influence the model response most. Least angle regression (LAR) is an efficient approach for feature selection. It generates a sequence of PC representations, in a such way the i-th metamodel includes predictors from 1 to i. The LAR algorithm is given below.

(i) Initialize the coefficients to $a_{\lambda_0}, a_{\lambda_1} ... a_{\lambda_{P-1}}$. Set the initial residual equal to the vector of observations.

(ii) Find the vector Φ_{λ_i} which is most correlated with the current residual.

(iii) Move a_{λ_i} from 0 toward the least-square coefficient of the current residual on Φ_{λ_i}, until some other predictor Φ_{λ_j} has as much correlation with the current residual as does Φ_{λ_i}.

(iv) Move jointly $(a_{\lambda_i}, a_{\lambda_j})$ in the direction defined by their joint least-square coefficient of the current residual on $(\Phi_{\lambda_i}, \Phi_{\lambda_j})$, until some other predictor Φ_{λ_k} has as much correlation with the current residual.

(v) Continue this way until $N = \min (P, m-1)$ predictors have been entered.

4.2. Weighted Leja sequences

Weighted Leja sequences is a stepwise search method which acquires sample one by one among thousands of given candidates. Suppose an input vector compose by given candidates $x = (x_1, x_2 ... x_N)$, N is the number of sampling candidates. We aim to find a solution to linear system $\phi a = y$, where $a = (a_1, a_2 ... a_P)$ is vector the coefficients of PCE and ϕ is the matrix satisfying $\phi_{ij} = \Phi_j(x_i)$, Φ_j is the j-th orthogonal polynomials. Assure the unisolvence of $\phi a = y$, multiplying a non-vanishing weight function in the linear system still has a unique solution.

$$\phi a = y \Leftrightarrow V\phi a = Vy, V_{ii} = v(x_i) \tag{22}$$

V is a diagonal matrix and v is the weight function. Inspired by the introduction of weight function, a sampling strategy is proposed by building sampling sequentially to maximize the determinant of $V\phi$. Then a nested sequence can be created by seeking sample among candidates to maximize determinant of matrix $V\phi$ gradually, that is

$$x^i = \arg\max \left| \det \left[V\phi(:,1:i) \right] \right|, \ i \geq 1 \tag{23}$$

Greedy algorithms can be employed to solve such maximization. In fact, the choice of weight v is vitally important and determining special properties of the sequence. When joint probability density $f_X(X)$ is known, adopting $v(X) = \sqrt{f_X(X)}$ as weight constructs a asymptotically optimal wLS. While $f_X(X)$ is unknown, an alternative showed in [2] is using the root inverse of the Christoffel function $v(X) = 1/\sqrt{\sum_{\lambda \in \Lambda} \Phi_\lambda^2(X)}$, where Christoffel function is the root of sum of squares of the whole basis functions.

4.3. Cross validation criterion for selecting the optimal LAR metamodel

With this sequential experimental design (wLS), We aim to gradually add samples into experimental design in use until satisfying the stop-sampling criteria. The K-fold cross validation is adopted to appraise the accuracy of PCE surrogate and select the best PCE model with the minimal K-fold cross validation error.

At first, we must set some parameters for this algorithm.

• For K-fold cross validation, a small K tends to low variance and high bias, while a large K tends to low bias and high variance as well as higher computational costs since it needs to solve more instances of the reduced problem. This paper we set $K = 10$ since the experimental design can be added in by multiples of 10 easily.

• A reasonable initial number of samples $m_0 = 10d$, and the increment for next step Δm take an integer divisible by 10 in interval $[0.25m_0, 0.35m_0]$

• Stop-sampling criteria. When the K-fold CV error converges, the algorithm comes to an end. We adopt the gradient of K-fold CV error as evaluation indicator

$$\eta_k = (\varepsilon_{CV}^{k-1} - \varepsilon_{CV}^k)/\varepsilon_{CV}^{k-1} \tag{24}$$

k is the times for adding samples to the experimental design. The algorithm is terminated when three or four successive η_k less than hundredth. Three or four depends on the increment Δm, we aim to assure a convergence over consecutive sample's increment (greater than m_0). In addition, in case of the K-fold CV error drops drastically without convergence, we set a threshold of $\varepsilon_{CV}^{Threshold}$ as thousandth. Once relative K-fold CV error less than the threshold value, terminate the whole algorithm. Besides, the maximal size of experimental design is given on account of the limited computational resources.

The pseudocode for the overall method is presented as follow.

Algorithm for selecting the optimal LAR metamodel based on CV

1: Set parameters. Select CV fold size K; initial sample size m_0; sample increment size Δm;
   stop-sampling parameters
2: $m = m_0$
3: **while** stop-sampling conditions are not met
4:    **for** $p = 1, 2, ..., p_{max}$ **do**
5:      **for** $q = 0.1, 0.2, ..., 1$ **do**
6:        Run GSO and LAR for a set of PCE model and compute coefficients by least
         square method, calculate ε for each PCE and find the one with minimal ε_{min}
7:      **end for**
8:    **end for**
9:    Discard some (p, q) combinations which have poor performance consecutively
10:   **if** stop-sampling conditions are met **then**
11:     Record (p, q), polynomials coefficients and ε for the best model selected
12:   **else**
13:     $m = m + \Delta m$
14:   **end if**

15: **end while**

In practice, it is no need to traverse all q from 0 to 1. Because for a high dimension problem, when q equals to a large value like 1, plenty of orthogonal polynomials ($P = C_d^{p+d}$) result in high computational costs. One prefers increasing p_{max} rather than q to expand the polynomial candidates. Empirically, set $q \in [0.3, 0.6]$ is a reasonable choice. Besides, the LAR is carried on the whole data available, we just assign the experimental design into K folds to calculate K-fold error instead of running K times of LAR for different folds.

5 BENCHMARK TEST WITH DEPENDENT INPUTS

The Ishigami function is adopted as test function with dependent inputs and defined as

$$y = \sin(x_1) + 7\sin^2(x_2) + 0.1x_3^4 \sin(x_1) \tag{25}$$

where $x \sim U(-\pi, \pi)$ and we set there exist correlation ρ_{13} between variables X_1, X_3. It's easy to find the linear combination which satisfies above requirements.

$$x_{1-2} \sim U_1(-\pi, -\pi)$$
$$x_3 = sign(\rho_{13})hx_1 + (1-h)U_3(-\pi, -\pi), h = 1/(\sqrt{(1-\rho_{13}^2)/\rho_{13}^2} + 1) \tag{26}$$

After computing the orthogonal polynomials for different truncated set (different values of p, q), the selection of the optimal PCE metamodel based on LAR and CV is carried out and given below in figure 1.
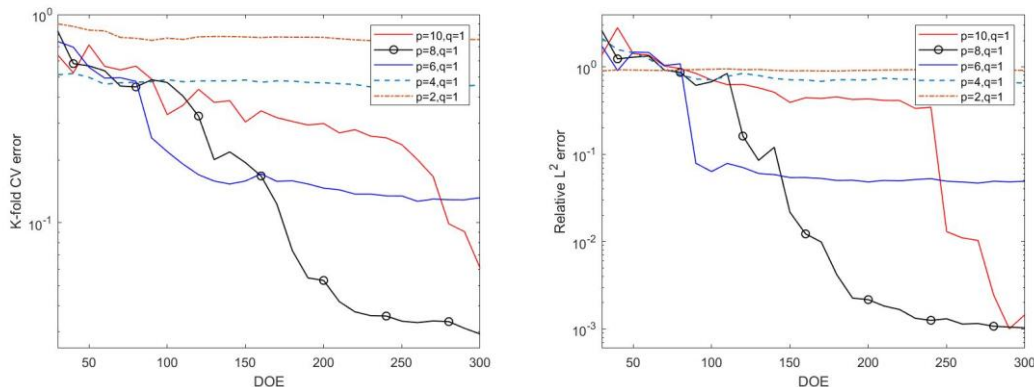

Figure 1: K-fold CV error and relative L^2 error for different truncated set for Ishigami function

Figure 1 depicts the evolution of K-fold CV error and relative L^2 error for different truncated set when $\rho_{13} = 0.5$. It's almost the same that the truncated set $p = 8, q = 1$ behaves best for all $\rho_{13} \in [-0.9, 0.9]$, therefore all sensitivity indices computations are based on this truncated set with same experimental design.

And all sensitivity indices are given in figure 2. Just see into the Ishigami function, the function items with regard to X_1, X_2, X_3 are nothing but $\{\sin x_1, \sin x_2, x_3^4\}$, in which variable X_2 is independent with variables X_1 and X_3. We can easily conclude that the variance of variable X_2 ($V_2 = Var(7\sin^2(x_2))$) is fixed. In other words, the equation $S_2 = S_2^u = ST_2^u$ is satisfied all the time. And one can deduce the variance of variables X_1, X_3 as

$$V_{1,3} = Var(\sin(x_1) + 0.1x_3^4 \sin(x_1)) = Var(\sin(x_1)) + Var(0.1x_3^4 \sin(x_1)) + 2Cov(\sin(x_1), 0.1x_3^4 \sin(x_1))$$

$$= Var(\sin(U_1)) + 0.1^2 \cdot Var\left[(\pm h U_1 + (1-h)U_3)^4 \cdot \sin U_1 \right] + 2E(0.1x_3^4 \sin^2(x_1)), h > 0$$

$$= Var(\sin(U_1)) + 0.1^2 \cdot Var\left[(c_1 U_1^4 \pm c_2 U_1^3 U_3^1 + c_3 U_1^2 U_3^2 \pm c_4 U_1^1 U_3^3 + c_5 U_3^4) \cdot \sin U_1 \right]$$

$$+ 0.2 \cdot E(c_1 U_1^4 + c_3 U_1^2 U_3^2 + c_5 U_3^4), c_{i \in \{1,\dots,5\}} > 0$$
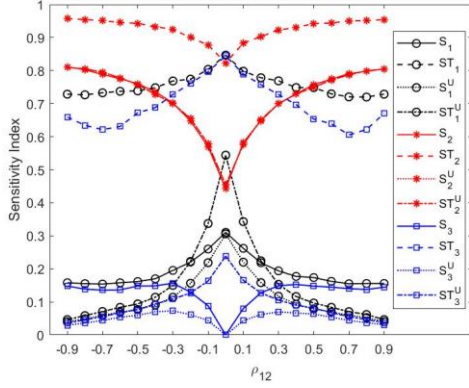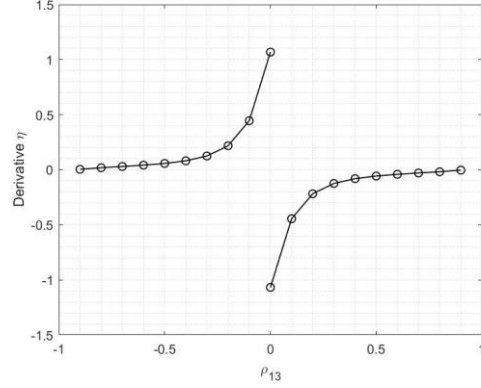
(27)

Figure 2: Sensitivity indices $\qquad$ **Figure 3**: Derivative of variance $V_{1,3}$ for different ρ_{13}

based on the variance formula $Var(A+B) = Var(A) + Var(B) + 2Cov(A,B)$.

At first, we aim at proving the $\pm$ which depends on the sign of ρ_{13} doesn't affect the variance $V_{1,3}$. The expansion items of x_3^4 can be divided into group $G_1 = \{c_1 U_1^4, c_3 U_1^2 U_3^2, c_5 U_3^4\}$ and group $G_2 = \{\pm c_2 U_1^3 U_3^1, \pm c_4 U_1^1 U_3^3\}$ based on whether involves $\pm$, $c_i, i = 1,\dots,5$ is abbreviation of coefficient relating to h. And it's demonstrated that the elements between group G_1 and group G_2 is uncorrelated mutually, which means $Cov(G_1^i, G_2^j) = 0$, $i = 1,2,3$ and $j = 1,2$ are the i-th and j-th element in group G_1 and group G_2.

$$Cov(G_1^i, G_2^j) = E\left[(G_1^i - E(G_1^i)) \cdot (G_2^j - E(G_2^i)) \right] = E\left[G_1^i \cdot G_2^j \right], E(G_2^j) = 0$$

$$= E\left[c_{ij} U_1^{k_1} U_3^{k_2} \right] = 0, \quad k_1, k_2 = 2n+1, n \geq 1$$

(28)

As a result, one can simplify the second item of $V_{1,3}$ into

$$Var\left[(c_1 U_1^4 \pm c_2 U_1^3 U_3^1 + c_3 U_1^2 U_3^2 \pm c_4 U_1^1 U_3^3 + c_5 U_3^4) \cdot \sin U_1 \right], c_{i \in \{1,\dots,5\}} > 0$$

$$= Var[(c_1 U_1^4 + c_3 U_1^2 U_3^2 + c_5 U_3^4) \cdot \sin U_1] + Var[(\pm c_2 U_1^3 U_3^1 \pm c_4 U_1^1 U_3^3) \cdot \sin U_1]$$

(29)

Then go on to deduce the second item with $\pm$ in above formula

$$Var[(\pm c_2 U_1^3 U_3^1 \pm c_4 U_1^1 U_3^3) \cdot \sin U_1]$$

$$= E[(\pm c_2 U_1^3 U_3^1 \pm c_4 U_1^1 U_3^3)^2 \cdot \sin^2 U_1], \quad E[(\pm c_2 U_1^3 U_3^1 \pm c_4 U_1^1 U_3^3) \cdot \sin U_1] = 0$$

$$= E[(c_2 U_1^3 U_3^1 + c_4 U_1^1 U_3^3)^2 \cdot \sin^2 U_1]$$

(30)

and we eliminate the $\pm$ induced by the sign of ρ_{13}. One knows that $h = 1/(\sqrt{(1-\rho_{13}^2)/\rho_{13}^2} + 1)$ is an even function, so that the variance $V_{1,3}$ and all sensitivity indices should be even functions. The results of our proposed method, as shown in figure 2, accord with the analytical prove that all sensitivity indices are symmetric with regard to $\rho_{13} = 0$.

In the following, we take derivative on $V_{1,3}$ for parameter h, As the below formula shown.

$$\eta_h = \frac{dV_{1,3}}{dh} = \frac{d[Var(0.1x_3^4 \sin(x_1))]}{dh} = 0.1^2 \frac{d[E(x_3^8 \sin^2(x_1))]}{dh}$$

$$= 0.1^2 \frac{d[E((hU_1+(1-h)U_2)^8 \sin^2(U_1))]}{dh} = \frac{0.1^2}{(2\pi)^2} \frac{d[\iint_{-\pi}^{\pi}((hU_1+(1-h)U_2)^8 \sin^2(U_1))du_2du_1]}{dh} \quad (31)$$

The derivative η_h is not difficult to obtain and have an analytical solution. Then we draw the relation between the derivative $\eta = dV_{1,3}/\rho_{13}$ and the correlation ρ_{13} in figure 3. The derivative of variance $V_{1,3}$ is less than 0 for ρ_{13} in the range of $(0,0.9]$ and greater than 0 for ρ_{13} in $[-0.9,0)$. In other words, the correlation ρ_{13} reduces the total variance $V_{1,3}$ of variables X_1, X_3. In the figure 2, correspondingly, the sensitivity indices S_2, ST_2, S_2^u, ST_2^u go up as $|\rho_{13}|$ increasing while model variance V_y decreases. For variable X_1 and X_3, as the correlation $|\rho_{13}|$ increasing, the total full sensitivity indices (ST_1, ST_3) and the total uncorrelated sensitivity indices (ST_1^u, ST_3^u) go down. However, their first-order sensitivity indices (S_1, S_3) and first-order uncorrelated sensitivity indices (S_1^u, S_3^u) is more concerned about the functional relation ($x_3^4 \sin(x_1)$) and is difficult to explain their variations. But the results produced by our proposed approach tells us the first-order sensitivity indices for X_1 (S_1^u, S_1) diminish while S_3^u, S_3 increase.

When $\rho_{13} = 0$, it's the situation for test of Ishigami function with independent. The uncorrelated sensitivity indices are equivalent to the Sobol' indices for independent test $(S_i^u = S_i, ST_i^u = ST_i)$. Just as shown in table 1, sPC obtains sensitivity indices accurately with lower computational costs in contrast to Monte Carlo (MC) and reference values.

Table 1: Sobol' indices of Ishigami function with independent inputs

Sensitivity indices	Reference	Sparse PCE $\varepsilon_{CV} = 0.035$ $\varepsilon_{L^2} = 0.020$	Sparse PCE $\varepsilon_{CV} = 0.008$ $\varepsilon_{L^2} = 0.001$	Crude MC
S_1	0.31	0.2639	0.3040	0.3245
S_2	0.44	0.4540	0.4544	0.4430
S_3	0	0.0000	0.0000	0.0098
ST_1	0.56	0.5387	0.5444	0.5572
ST_2	0.44	0.4544	0.4545	0.4330
ST_3	0.24	0.2816	0.2431	0.2425
Evaluations		60	100	200000

6 CONCLUSIONS

In this paper, a global sensitivity analysis (GSA) approach for dependent inputs has been proposed based on sPC. Firstly, a generalization of the Sobol' indices is introduced. Then PCE model is constructed by Gram-Schmidt algorithm directly. To reduce the computational costs, The LAR algorithm is employed to produce a set of sparse PCE models as a feature selection method, while CV appraise those models and find the best one with minimal CV

error. An adaptive algorithm is designed to determine best truncated set for sparse representation of PCE and the suitable size of the nested experimental design called wLS. After the accurate sparse PCE has been established, sensitivity indices for dependent inputs can be calculated by algebraic operations on the coefficients of the polynomial expansions.

The benchmark test illustrates that the adaptive algorithm based on LAR, CV and wLS works well for constructing sparse representation of PCE. And it's proved that our proposed method is effective and accurate to compute sensitivity indices for dependent inputs, with no need to fit variable data to any distribution, with no need to preset or fix the size of experimental design, and with no need to transform dependent inputs into independent inputs.

ACKNOWLEDGEMENT

REFERENCES

[1] Xiu D, Karniadakis G E. The Wiener--Askey polynomial chaos for stochastic differential equations[J]. SIAM journal on scientific computing, 2002, 24(2): 619-644.

[2] Jakeman J D, Franzelin F, et al. Polynomial chaos expansions for dependent random variables[J]. Computer Methods in Applied Mechanics and Engineering, 2019, 351: 643-666.

[3] Wang G, Xin H, et al. Data-driven Arbitrary Polynomial Chaos-based Probabilistic Load Flow Considering Correlated Uncertainties[J]. IEEE Transactions on Power Systems, 2019.

[4] Rosenblatt M. Remarks on a multivariate transformation[J]. The annals of mathematical statistics, 1952, 23(3): 470-472.

[5] Kucherenko S, Tarantola S, Annoni P. Estimation of global sensitivity indices for models with dependent variables[J]. Computer physics communications, 2012, 183(4): 937-946.

[6] Caniou Y. Global sensitivity analysis for nested and multiscale models[D]. Ph. D. thesis, Université Blaise Pascal, Clermont-Ferrand, 2012.

[7] Mara T A, Tarantola S. Variance-based sensitivity indices for models with dependent inputs[J]. Reliability Engineering & System Safety, 2012, 107: 115-121.

[8] Mara T A, Tarantola S. Non-parametric methods for global sensitivity analysis of model output with dependent inputs[J]. Environmental modelling & software, 2015, 72: 173-183.

[9] Liu Z, Choe Y. Data-driven sensitivity indices for models with dependent inputs using the polynomial chaos expansion[J]. arXiv preprint arXiv:1803.10978, 2018.

[10] Ghanem R G, Spanos P D. Stochastic finite elements: a spectral approach[M]. Courier Corporation, 2003.

[11] Xiu D, Hesthaven J S. High-order collocation methods for differential equations with random inputs[J]. SIAM Journal on Scientific Computing, 2005, 27(3): 1118-1139.

[12] Blatman G, Sudret B. Adaptive sparse polynomial chaos expansion based on least angle regression[J]. Journal of Computational Physics, 2011, 230(6): 2345-2367.

[13] Blatman G, Sudret B. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions[J]. Reliability Engineering & System Safety, 2010, 95(11): 1216-1229.

[14] Reichel L. Newton interpolation at Leja points[J]. BIT, 1990, 30(2): 332-346.