Bayesian PCE as a Control Variate Method for Estimating Sobol' Sensitivity Indices

Hui (Alyssa) Duan

Florida State University

March 16, 2022

Global sensitivity analysis

- Provides a rigorous assessment of parameter sensitivity
- Variance-based method: Sobol' sensitivity index
- Application on applied math models and physical experiments

ANOVA decomposition

Consider a d-dimensional vector $\mathbf{x} = (x_1, \cdots, x_d)$, the ANOVA decomposition of function $f(\mathbf{x})$ is given by

$$f(\mathbf{x}) = \sum_{u \subseteq \mathcal{D}} f_u(\mathbf{x}^u)$$

- The expectation of f:

$$E[f(\mathbf{x})] = \int f(\mathbf{x}) d\mathbf{x} = \mu$$

- The variance of f:

$$\sigma^2 = \int f^2(\mathbf{x}) d\mathbf{x} - \mu^2$$

- The variance of f_u:

$$\sigma_u^2 = \int f_u^2(\mathbf{x}) d\mathbf{x} - (\int f_u(\mathbf{x}) d\mathbf{x})^2 = \int f_u^2(\mathbf{x}) d\mathbf{x}$$

if $u \neq \emptyset$

Sobol' sensitivity index

- Lower Sobol' sensitivity index:

$$\underline{S}_u = \frac{1}{\sigma^2} \sum_{v \subseteq u} \sigma_v^2 = \frac{\underline{\tau}_u}{\sigma^2}$$

- Upper Sobol' sensitivity index:

$$\overline{S}_u = \frac{1}{\sigma^2} \sum_{v \bigcap u \neq \emptyset} \sigma_v^2 = \frac{\overline{\tau}_u}{\sigma^2}$$

Theorem (Sobol' [1])

$$\underline{\tau}_u = \int f(\mathbf{x}^u, \mathbf{x}^{-u}) f(\mathbf{x}^u, \mathbf{z}^{-u}) d\mathbf{x} d\mathbf{z}^{-u} - \mu^2$$

$$\overline{\tau}_u = \frac{1}{2} \int [f(\mathbf{x}^u, \mathbf{x}^{-u}) - f(\mathbf{z}^u, \mathbf{x}^{-u})]^2 d\mathbf{x} d\mathbf{z}^u$$

Lower Sobol' estimator - MC approach

- Owen's estimator

$$\underline{S}_u^{\text{owen}} = \frac{\underline{\tau}_u^{\text{owen}}(\mathbf{x}, \mathbf{y}^u, \mathbf{z})}{\text{Var}(f(\mathbf{x}))} = \frac{\frac{1}{n}\sum_{i=1}^{n}(f(\mathbf{x}_i) - f(\mathbf{y}_i^u, \mathbf{x}_i^{-u}))(f(\mathbf{x}_i^u, \mathbf{z}_i^{-u}) - f(\mathbf{z}_i))}{\text{Var}(f(\mathbf{x}))} \quad (1)$$

- Janon/Monod's estimator

$$\underline{S}_u^{\text{janon}} = \frac{\frac{1}{N}\sum_{n=1}^{N} f(\mathbf{x}_i) f(\mathbf{x}_i^u, \mathbf{z}_i^{-u}) - f_0^2}{\frac{1}{N}\sum_{i=1}^{N} \frac{f(\mathbf{x}_i)^2 + f(\mathbf{x}_i^u, \mathbf{z}_i^{-u})^2}{2} - f_0^2} \quad (2)$$

where $f_0 = \frac{1}{N}\sum_{i=1}^{N} \frac{f(\mathbf{x}_i) + f(\mathbf{x}_i^u, \mathbf{z}_i^{-u})}{2}$

- Azzini & Rosati's estimator

$$\underline{S}_u^{\text{az}} = \frac{2\sum_{i=1}^{N}(f(\mathbf{x}_i^u, \mathbf{z}_i^{-u}) - f(\mathbf{z}_i))(f(\mathbf{x}_i) - f(\mathbf{z}_i^u, \mathbf{x}_i^{-u}))}{\sum_{i=1}^{N}(f(\mathbf{x}_i) - f(\mathbf{z}_i)))^2 + (f(\mathbf{x}_i^u, \mathbf{z}_i^{-u}) - f(\mathbf{z}_i^u, \mathbf{x}_i^{-u}))^2} \quad (3)$$

Upper Sobol' estimator - MC approach

- Jansen's estimator

$$\overline{S}_u^{\text{jansen}} = \frac{\overline{\tau}_u^{\text{jansen}}(\mathbf{x}, \mathbf{z}^u)}{\text{Var}(f(\mathbf{x}))} = \frac{\frac{1}{2n}\sum_{i=1}^{n}(f(\mathbf{x}_i) - f(\mathbf{z}_i^u, \mathbf{x}_i^{-u}))^2}{\text{Var}(f(\mathbf{x}))} \tag{4}$$

- Janon/Monod's estimator

$$\overline{S}_u^{\text{janon}} = 1 - \frac{\frac{1}{N}\sum_{n=1}^{N} f(\mathbf{x}_i)f(\mathbf{z}_i^u, \mathbf{x}_i^{-u}) - f_0'^2}{\frac{1}{N}\sum_{i=1}^{N} \frac{f(\mathbf{x}_i)^2 + f(\mathbf{z}_i^u, \mathbf{x}_i^{-u})^2}{2} - f_0'^2} \tag{5}$$

where $f_0' = \frac{1}{N}\sum_{i=1}^{N} \frac{f(\mathbf{x}_i) + f(\mathbf{z}_i^u, \mathbf{x}_i^{-u})}{2}$

- Azzini & Rosati's estimator

$$\overline{S}_u^{\text{az}} = \frac{\sum_{i=1}^{N}[f(\mathbf{z}_i) - f(\mathbf{x}_i^u, \mathbf{z}_i^{-u})]^2 + [f(\mathbf{x}_i) - f(\mathbf{z}_i^u, \mathbf{x}_i^{-u})]^2}{\sum_{i=1}^{N}[f(\mathbf{x}_i) - f(\mathbf{z}_i)]^2 + [f(\mathbf{x}_i^u, \mathbf{z}_i^{-u}) - f(\mathbf{z}_i^u, \mathbf{x}_i^{-u})]^2} \tag{6}$$

Polynomial chaos expansion

Consider a d-dimensional vector $\mathbf{x} = (x_1, x_2, \cdots, x_d)$ with probability density function $f_{\mathbf{x}}(x)$. Assume that the random output $Y = f(\mathbf{x})$ has finite variance, it can be written as the polynomial chaos expansion (PCE):

$$Y = f(\mathbf{x}) = \sum_{i=0}^{\infty} k_i \Psi_i(\mathbf{x})$$

- The truncated polynomial at order p is:

$$f_p(\mathbf{x}) = \sum_{i=0}^{P-1} k_i \Psi_i(\mathbf{x})$$

  where $P = \binom{p+d}{p}$
- The expectation of f_p: $\mathbb{E}[f_p] = k_0$
- The variance of f_p: $\sigma^2 = \sum_{i=1}^{P-1} k_i^2$

Estimating Sobol' sensitivity indices using PCE

- ANOVA component f_u:

$$f_u(\mathbf{x}_u) = \sum_{i \in \mathcal{A}_u} k_i \Psi_{\alpha^i}$$

$\mathcal{A}_u = \{i : \alpha^i_j > 0 \text{ for every } j \in u, \text{ and } \alpha^i_j = 0 \text{ otherwise}\}$

- Variance of the component function:

$$\sigma_u^2 = \sum_{i \in \mathcal{A}_u} k_i^2$$

- The lower Sobol' index:

$$\underline{S}_u^{\text{pce}} \approx \frac{\sum_{i \in \underline{\mathcal{A}}_{u,p}} \hat{k}_i^2}{\sigma^2} \tag{7}$$

$\underline{\mathcal{A}}_{u,p} = \{i : i < P, \text{and } \exists j \in u \text{ where } \alpha^i_j > 0 \ \alpha^i_j = 0 \text{ or every } j \in -u \ \}$

- The upper Sobol' index:

$$\overline{S}_u^{\text{pce}} \approx \frac{\sum_{i \in \overline{\mathcal{A}}_{u,p}} \hat{k}_i^2}{\sigma^2} \tag{8}$$

$\overline{\mathcal{A}}_{u,p} = \{i : i < P, \text{and } \exists j \in u \text{ where } \alpha^i_j > 0 \ \}$

Polynomial Chaos Expansion

PCE coefficients:

$$k_i = \int_S \Psi_i(\mathbf{x}) f(\mathbf{x}) w(\mathbf{x}) d\mathbf{x}$$

- Monte Carlo method:

$$\hat{k}_i = \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) \Psi_i(\mathbf{x}_n)$$

- Ordinary least-square minimization:

$$\hat{k} = \text{argmin}_k ||\Psi(x)k - Y||_2$$

- Sparse PCE: an expansion for most coefficients are zero
- Sparse solver: Bayesian PCE

PCE v.s. Bayesian PCE

Problem Setup: a model to surrogate $Y = f(\mathbf{x})$

- Test functions: Ishigami function, Morris function
- Input $\mathbf{x}$: Sobol' sequence
- Goal: estimate $\int f(\mathbf{x})d\mathbf{x}$
- Compare function estimations between PCE and Bayesian PCE using

$$f_p(\mathbf{x}) = \sum_{i=0}^{P-1} k_i \Psi(\mathbf{x}_i)$$

- Error estimation:

$$\epsilon = \frac{\mathbb{E}[f(\mathbf{x}) - f_p(\mathbf{x})]^2}{Var[f(\mathbf{x})]}$$

- **Ishigami function** (dim = 3, $p = 6$)

$$f(\mathbf{x}) = \sin(x_1) + a\sin^2(x_2) + bx_3^4\sin(x_1)$$

where $\mathbf{x} = (x_1, x_2, x_3) \sim U[-\pi, \pi]^3$ and $a = 7,\ b = 0.1$

PCE v.s. Bayesian PCE

- **Morris function** (dim $= 20$, $p = 3$)

$$Y = \beta_0 + \sum_{i=1}^{20} \beta_i X_i + \sum_{i<j}^{20} \beta_{ij} X_i X_j + \sum_{i<j<k}^{20} \beta_{ijk} X_i X_j X_k + \sum_{i<j<k<l}^{20} \beta_{ijkl} X_i X_j X_k X_l$$

where

$$X_j = \begin{cases} 2(1.1x_i/(x_i+0.1) - 0.5) & \text{if } i = 3,5,7 \\ 2(x_i - 0.5) & \text{otherwise} \end{cases}$$

$x_i \sim U[0,1]^{20}$. The coefficients β_i are assigned as follows:

$$\begin{cases} \beta_i = 20 & \text{for } i = 2,\cdots,10 \\ \beta_{ij} = -15 & \text{for } i,j = 1,\cdots,6 \\ \beta_{ijk} = -10 & \text{for } i,j,k = 1,\cdots,5 \\ \beta_{ijkl} = 5 & \text{for } i,j,k,l = 1,\cdots,4 \end{cases}$$

The remaning coefficients are defined by $\beta_0 = 0$, $\beta_i = (-1)^i$ and $\beta_{ij} = (-1)^{i+j}$

PCE v.s. Bayesian PCE

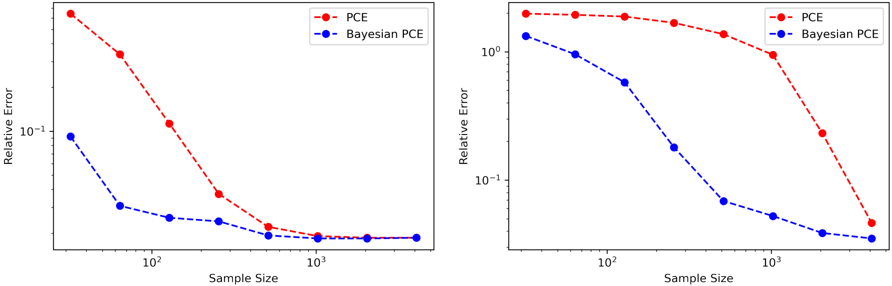


Figure 1: Relative error of PCE and Bayesian PCE using Sobol' sequence (left: Ishigami function, right: Morris function)

Control variate method

Let $Y = f(\mathbf{x})$ be an estimator for θ, i.e. $\theta = \mathbb{E}[Y]$. The unbiased control variate estimator is given by

$$Y(\beta) = Y - \beta(C - \mathbb{E}[C])$$

where C is a random variable called a control variate for the estimator Y with known mean μ_C and correlated with Y. β is some real number and it is picked by minimizing the variance of $Y(\beta)$

$$\beta^* = \frac{\text{Cov}(Y, C)}{\text{Var}(C)}$$

The control variate estimator is given by

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^{N} Y_i - \hat{\beta}_N^*(C_i - \mathbb{E}[C])$$

Bayesian PCE as a control variate method

The control variate estimator with truncated PCE as a control is given by

$$\hat{f}_{cv} = \frac{1}{N}\sum_{i=1}^{N} f(\mathbf{x}_i) - \beta(f_p(\mathbf{x}_i) - \mathbb{E}[f_p]) \quad (9)$$

where f_p is estimated using Bayesian PCE and optimal coefficient β is estimated by

$$\beta^* = \frac{\mathrm{Cov}(f(\mathbf{x}), f_p(\mathbf{x}))}{\mathrm{Var}(f_p(\mathbf{x}))}$$

Control variate method 1 (cv1)

The lower Sobol' index:

$$\underline{S}_u^{cv1} = \frac{\underline{\tau}_u^{cv1}(\mathbf{x}, \mathbf{y}^u, \mathbf{z})}{\mathrm{Var}(f(\mathbf{x}))} \tag{10}$$

where

$$\underline{\tau}_u^{cv1}(\mathbf{x}, \mathbf{y}^u, \mathbf{z}) = \underline{\tau}_u^{\text{owen}}(\mathbf{x}, \mathbf{y}^u, \mathbf{z}) - \underline{\beta}^*(\underline{\tau}_{u,p}^{\text{owen}}(\mathbf{x}, \mathbf{y}^u, \mathbf{z}) - \mathbb{E}[\underline{\tau}_{u,p}^{\text{owen}}])$$

The upper Sobol' index:

$$\overline{S}_u^{cv1} = \frac{\overline{\tau}_u^{cv1}(\mathbf{x}, \mathbf{z}^u)}{\mathrm{Var}(f(\mathbf{x}))} \tag{11}$$

$$\overline{\tau}_u^{cv1}(\mathbf{x}, \mathbf{z}^u) = \overline{\tau}_u^{\text{jansen}}(\mathbf{x}, \mathbf{z}^u) - \overline{\beta}^*(\overline{\tau}_{u,p}^{\text{jansen}}(\mathbf{x}, \mathbf{z}^u) - \mathbb{E}[\overline{\tau}_{u,p}^{\text{jansen}}])$$

Control variate method 2 (cv2)

The lower Sobol' index:

$$\underline{S}_u^{cv2} = \frac{\underline{\tau}_{u,p}^{cv2}(\mathbf{x}, \mathbf{y}^u, \mathbf{z})}{\mathrm{Var}(f(\mathbf{x}))} \tag{12}$$

where

$$\begin{aligned}\underline{\tau}_{u,p}^{cv2}(\mathbf{x}, \mathbf{y}^u, \mathbf{z}) = {} & (f(\mathbf{x}) - \hat{f}_p(\mathbf{x}) - f(\mathbf{y}^u, \mathbf{x}^{-u}) + \hat{f}_p(\mathbf{y}^u, \mathbf{x}^{-u})) \\ & (f(\mathbf{x}^u, \mathbf{z}^{-u}) - \hat{f}_p(\mathbf{x}^u, \mathbf{z}^{-u}) - f(\mathbf{z}) + \hat{f}_p(\mathbf{z})) + \mathbb{E}[\underline{\tau}_{u,p}]\end{aligned}$$

The upper Sobol' index:

$$\overline{S}_u^{cv2} = \frac{\overline{\tau}_{u,p}^{cv2}(\mathbf{x}, \mathbf{z}^u)}{\mathrm{Var}(f(\mathbf{x}))} \tag{13}$$

where

$$\overline{\tau}_{u,p}^{cv2}(\mathbf{x}, \mathbf{z}^u) = \frac{1}{2}(f(\mathbf{x}) - \hat{f}_p(\mathbf{x}) - f(\mathbf{z}^u, \mathbf{x}^{-u}) + \hat{f}_p(\mathbf{z}^u, \mathbf{x}^{-u}))^2 + \mathbb{E}[\overline{\tau}_{u,p}]$$

Numerical results

- Problem: estimate the Sobol' sensitivity index
- Estimators: cv1, cv2, MC estimators, PCE
- Input: $N = 10,000$ pseudorandom numbers generated by Mersenne Twister
- Repeat the process $K = 50$ times independently to obtain the root mean square error(RMSE)

$$\mathrm{RMSE}_j = [\frac{1}{K}\sum_{k=1}^{K}(\hat{S}_j^{(k)} - S_j)^2]^{1/2}$$

- Compute sample standard deviation σ of the estimates with 50 independent estimates and record the computing time t to compare the efficiency:

$$E = \sigma^2 \times t$$

Case 1

The problem is inexpensive to evaluate and the "low" truncation can not approximate the problem accurately

Ishigami function (dim = 3, $p = 2$)

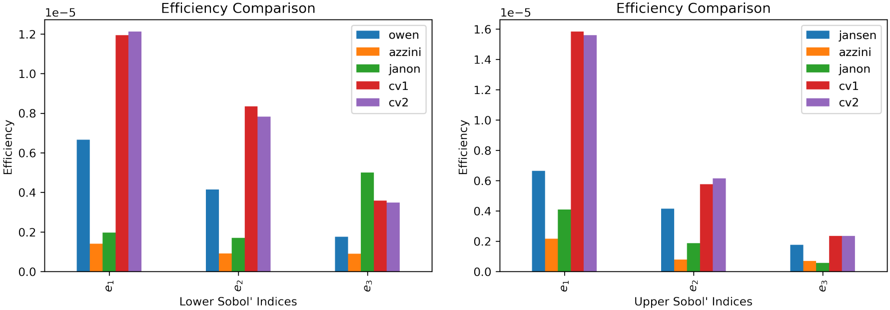
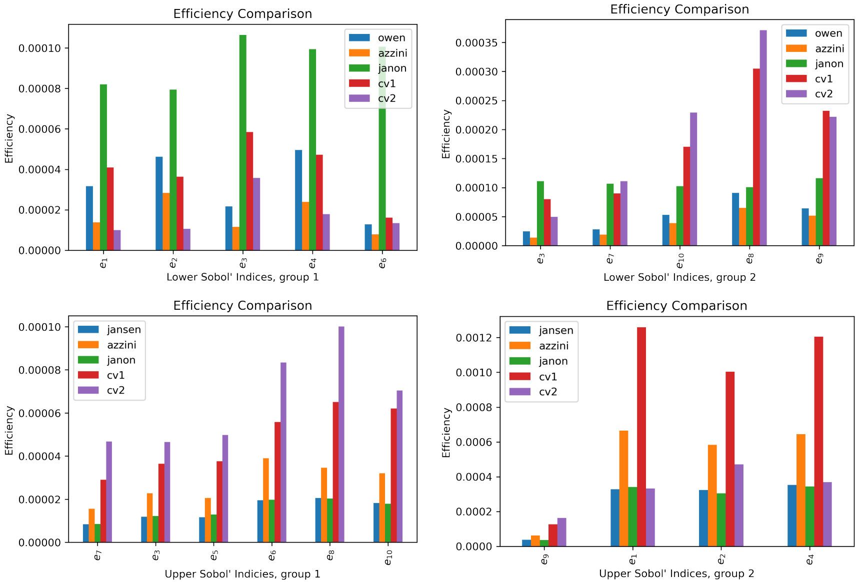


Figure 2: Efficiency comparison for Ishigami function

Case 2

The problem is inexpensive to evaluate and the "low" truncation can approximate the problem accurately

Morris function (dim = 20, $p = 2$)



Case 3

The problem is expensive to evaluate but the "low" truncation can not approximate the function well

SIR model (dim $= 4$, $p = 2$)

- **SIR model:**

$$\frac{dS}{dt} = \delta N - \delta S - \gamma k I S \ , \ S(0) = S_0$$

$$\frac{dI}{dt} = \gamma k I S - (r + \gamma) I \ , \ I(0) = I_0$$

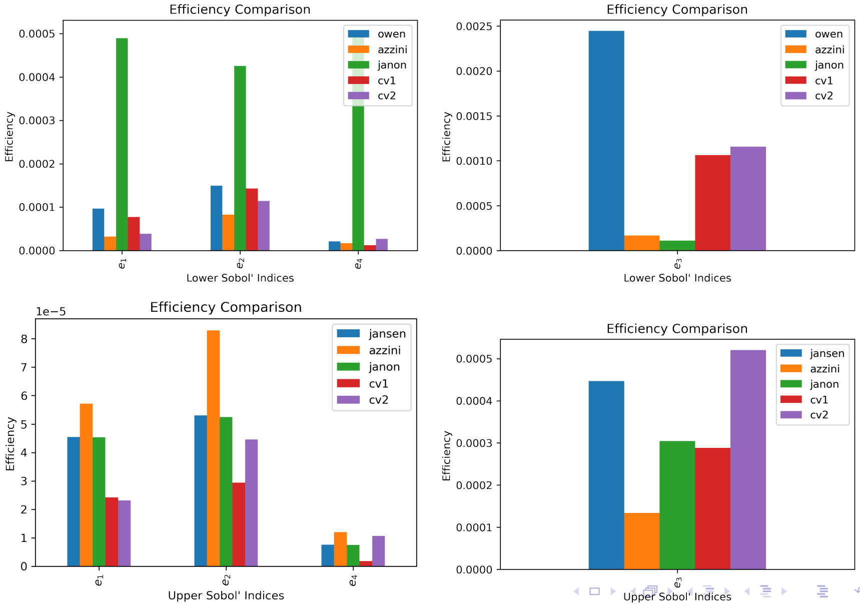
$$\frac{dR}{dt} = rI - \gamma R \ , \ R(0) = R_0$$

Consider the scalar response:

$$Y = \int_0^1 R(t, \theta) dt$$

where $S_0 = 900$, $I_0 = 100$, $R_0 = 0$ and input parameters $\theta = [\gamma, k, r, \delta] \sim U[0, 1]^4$. We use 3,300,000 model evaluations to estimate the exact values for both lower Sobol' and upper Sobol' indices

SIR model (dim = 4, $p = 2$)



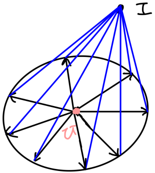
Stiefel canonical distance model (dim $= 7$, $p = 2$)

• **Stiefel canonical distance model:**
Denote the Stiefel manifold, the set of k ordered orthonormal vectors in $\mathbb{R}^n$ with

$$\mathbf{St}_{n,k} = \{U \in \mathbb{R}^{n \times k} : U^T U = I_k\}$$

with $n = 5, k = 2$

- Goal: The expectation of distance from an area to a point
- Input: $\mathbf{x} \sim U[-1, 1]^7$
- Use 4,950,000 model evaluations to estimate the exact values for both lower Sobol' and upper Sobol' indices



Stiefel canonical distance model (dim $= 7$, $p = 2$)

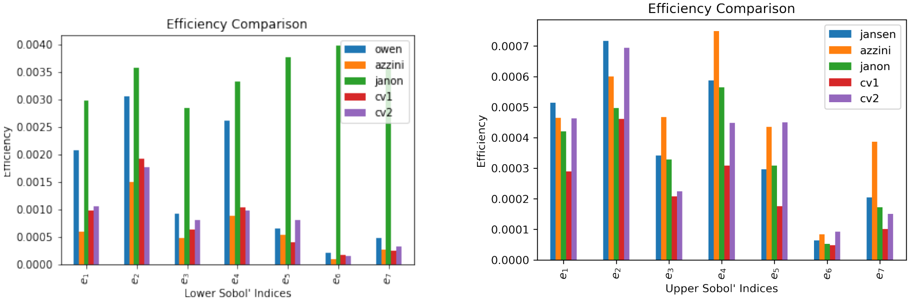


Figure 5: Efficiency comparison for Stiefel canonical distance model

Control variate estimator v.s. PCE

- Estimators: cv1 and PCE
- Models: SIR model, Stiefel canonical distance model
- Comparison: RMSE and efficiency

cv1 v.s. PCE: SIR Model, lower Sobol'

p	S_1	S_2	S_3	S_4
6	0.0033528	0.0022629	0.01265	0.0016435
7	**0.0023681**	**0.0015596**	**0.0093005**	**0.0014156**
cv1	**0.0025475**	**0.0034642**	**0.0094597**	**0.0010009**

Table 1: RMSE comparison of PCE with cv1 for SIR model, lower Sobol'

p	S_1	S_2	S_3	S_4
7	**1.2310e-04**	**5.3400e-05**	**1.8993e-03**	**4.4000e-05**
cv1	**6.6220e-06**	**1.2245e-05**	**9.1312e-05**	**1.0220e-06**

Table 2: Efficiency comparison of PCE with cv1 for SIR model, lower Sobol'

cv1 v.s. PCE: SIR model, upper Sobol'

p	S_1	S_2	S_3	S_4
9	0.0047618	0.0045820	0.0026967	0.00074129
10	**0.0039348**	**0.0037069**	**0.0019462**	**0.00058521**
cv1	**0.0027806**	**0.0030625**	**0.0095914**	**0.0007628**

Table 3: RMSE comparison of PCE with cv1 for SIR model, upper Sobol'

p	S_1	S_2	S_3	S_4
10	**9.2195e-03**	**8.1822e-03**	**2.2554e-03**	**2.0390e-04**
cv1	**7.8900e-06**	**9.5700e-06**	**9.3872e-05**	**5.9400e-07**

Table 4: Efficiency comparison of PCE with cv1 for SIR model, upper Sobol'

cv1 v.s. PCE: Stiefel canonical distance, lower Sobol'

p	S_1	S_2	S_3	S_4	S_5	S_6	S_7
4	0.0044	0.0041	0.0017	0.0041	0.0024	0.0009	0.0015
5	**0.0028**	**0.0025**	**0.0013**	**0.0025**	**0.0016**	**0.0007**	**0.0011**
cv1	**0.0037**	**0.0053**	**0.003**	**0.0039**	**0.0024**	**0.0016**	**0.0019**

Table 5: RMSE comparison of PCE with cv1 for distance model, lower Sobol'

p	S_1	S_2	S_3	S_4	S_5	S_6	S_7
5	**0.0020**	**0.0016**	**0.0004**	**0.0016**	**0.0006**	**0.0001**	**0.0003**
cv1	**0.0010**	**0.0019**	**0.0006**	**0.0010**	**0.0004**	**0.0002**	**0.0003**

Table 6: Efficiency comparison of PCE with cv1 for distance model, lower Sobol'

cv1 v.s. PCE: Stiefel canonical distance, upper Sobol'

p	S_1	S_2	S_3	S_4	S_5	S_6	S_7
6	0.0084	0.0095	0.01	0.0103	0.0094	0.0073	0.0093
7	**0.0063**	**0.0073**	**0.0073**	**0.0077**	**0.0069**	**0.0062**	**0.0071**
cv1	**0.0030**	**0.0044**	**0.0028**	**0.0036**	**0.0028**	**0.0015**	**0.0024**

Table 7: RMSE comparison of PCE with cv1 for distance model, upper Sobol'

p	S_1	S_2	S_3	S_4	S_5	S_6	S_7
7	**0.0722**	**0.0961**	**0.0968**	**0.1091**	**0.0877**	**0.0703**	**0.0908**
cv1	**0.0003**	**0.0005**	**0.0002**	**0.0003**	**0.0002**	**0**	**0.0001**

Table 8: Efficiency comparison of PCE with cv1 for distance model, upper Sobol'

Conclusions

- Control variate estimator (cv1) has better performance for Sobol' sensitivity indices estimation when the problem becomes expensive and the "low" truncation order can not approximate the function well
- cv1 yields best efficiency among all MC estimators and beats the PCE around 2-3 magnitudes, especially for estimating upper Sobol' indices

Thank You