Supplemental material for Shang *et al*., "An Sequential Experimental Design for Multivariate Sensitivity Analysis using Polynomial Chaos Expansion", *Engineering Optimization*, 2019.

Appendix

A. Proof 1

In the statistics, MLE is a method of estimating the parameters of a statistical model. Assuming the observation data are independent with each other, the log-likelihood function $L_F(\boldsymbol{b},\Lambda;Q)$ is given by

$$\begin{aligned} L_F(\boldsymbol{b},\Lambda;Q) &= \ln\rho(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_P) = \ln\prod_{i=1}^{P}\rho(\boldsymbol{y}_i) \\ &= -\sum_{i=1}^{P}\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)^{\mathrm{T}}\Lambda^{-1}\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)-\frac{Pn}{2}\ln(2\pi)-P\ln|\Lambda| \end{aligned}$$

where ρ denotes the probability density. The unknown parameters $\boldsymbol{b}$ and Λ can be obtained by maximizing the function $L_F(\boldsymbol{b},\Lambda;Q)$. To solve the maximization issue, the partial derivative of $L_F(\boldsymbol{b},\Lambda;Q)$ with respect to $\boldsymbol{b}$ and Λ are set to be zero, e.g.

$$\frac{\partial L_F(\boldsymbol{b},\Lambda;Q)}{\partial\boldsymbol{b}}=0,\frac{\partial L_F(\boldsymbol{b},\Lambda;Q)}{\partial\Lambda}=0$$

Since

$$\begin{aligned} \frac{\partial L_F(\boldsymbol{b},\Lambda;Q)}{\partial\boldsymbol{b}} &= \frac{\partial\sum_{i=1}^{P}\left[\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right]^{\mathrm{T}}\Lambda^{-1}\left[\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right]}{\partial\boldsymbol{b}} \\ &= \sum_{i=1}^{P}\frac{\partial\left[\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right]^{\mathrm{T}}\Lambda^{-1}\left[\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right]}{\partial\boldsymbol{b}} \\ &= \sum_{i=1}^{P}\frac{\partial\left[\boldsymbol{y}_i^{\mathrm{T}}\Lambda^{-1}\boldsymbol{y}_i+\boldsymbol{b}^{\mathrm{T}}\Phi(\boldsymbol{x}_i)\Lambda^{-1}\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}-2\boldsymbol{y}_i^{\mathrm{T}}\Lambda^{-1}\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right]}{\partial\boldsymbol{b}} \\ &= 2\sum_{i=1}^{P}\boldsymbol{b}^{\mathrm{T}}\Phi(\boldsymbol{x}_i)\Lambda^{-1}\Phi^{\mathrm{T}}(\boldsymbol{x}_i)-2\boldsymbol{y}_i^{\mathrm{T}}\Lambda^{-1}\Phi^{\mathrm{T}}(\boldsymbol{x}_i) \end{aligned}$$

Let $\frac{\partial L_F(\boldsymbol{b},\Lambda;Q)}{\partial\boldsymbol{b}}=0$, then the maximum likelihood estimator of $\boldsymbol{b}$ is given as

$$\hat{\boldsymbol{b}}=\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}\Psi^{\mathrm{T}}\Sigma^{-1}Q$$

where $\Psi=\left[\Phi(\boldsymbol{x}_1),\ldots,\Phi(\boldsymbol{x}_P)\right]^{\mathrm{T}}$ and $\Sigma^{-1}=I_P\otimes\Lambda^{-1}$.

Similarly, $\frac{\partial L_F(\boldsymbol{b},\Lambda;Q)}{\partial\Lambda}$ can written as

$$\begin{aligned}\frac{\partial L_F(\boldsymbol{b},\Lambda;Q)}{\partial\Lambda} &= \sum_{i=1}^{P}\frac{\partial\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)^{\mathrm{T}}\Lambda^{-1}\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)-P\ln|\Lambda|}{\partial\Lambda} \\ &= -\sum_{i=1}^{P}\Lambda^{-1}\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)^{\mathrm{T}}\Lambda^{-1}-P\Lambda^{-1} \\ &= -\left[\sum_{i=1}^{P}\Lambda^{-1}\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)^{\mathrm{T}}-P\right]\Lambda^{-1}\end{aligned}$$

Let $\frac{\partial L_F(\boldsymbol{b},\Lambda;Q)}{\partial\Lambda}=0$, then the maximum likelihood estimator of Λ is

$$\hat{\Lambda}=\frac{1}{P}\sum_{i=1}^{P}\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)^{\mathrm{T}}$$

B. Proof 2

Since

$$\hat{\boldsymbol{b}}=\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}\Psi^{\mathrm{T}}\Sigma^{-1}Q$$

Then the expected value of $\hat{\boldsymbol{b}}$ is

$$E[\hat{\boldsymbol{b}}]=E\left[\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}\Psi^{\mathrm{T}}\Sigma^{-1}Q\right]=\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\boldsymbol{b}=\boldsymbol{b}$$

$\hat{\boldsymbol{b}}$ is an unbiased estimator of $\boldsymbol{b}$. Besides, the covariance matrix of $\hat{\boldsymbol{b}}$ is derived:

$$C[\hat{\boldsymbol{b}}]=\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}\Psi^{\mathrm{T}}\Sigma^{-1}C(Q)\Sigma^{-1}\Psi\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}=\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}$$

Therefore $\hat{\boldsymbol{b}}\sim N\left(\boldsymbol{b},\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1}\right)$.

For the covariance matrix estimator $\hat{\Lambda}$, since

$$\hat{\Lambda}=\frac{1}{P}\sum_{i=1}^{P}\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)^{\mathrm{T}}$$

then, the expected value of $\hat{\Lambda}$ is

$$E\left[\hat{\Lambda}\right]=\frac{1}{P}\sum_{i=1}^{P}E\left[\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)\left(\boldsymbol{y}_i-\Phi^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{b}\right)^{\mathrm{T}}\right]=\frac{1}{P}\sum_{i=1}^{P}E\left[\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{T}}\right]=\Lambda$$

Therefore, $\hat{\Lambda}$ is an unbiased estimator of Λ.

C. Proof 3

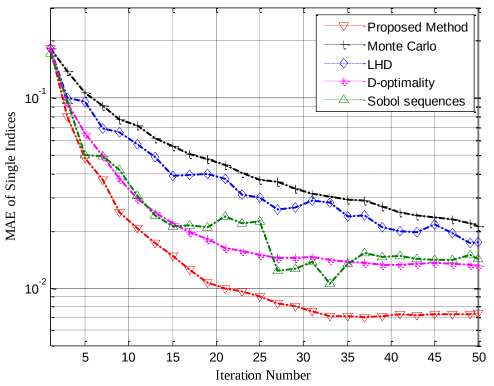
Since $\Psi = \left[\Phi(\boldsymbol{x}_1), \Phi(\boldsymbol{x}_2), ..., \Phi(\boldsymbol{x}_P)\right]^{\mathrm{T}}$, $\Phi^{\mathrm{T}}(\boldsymbol{x}_i) = I_n \otimes \boldsymbol{\varphi}^{\mathrm{T}}(\boldsymbol{x}_i), i = 1, ..., P$ and $\Sigma^{-1} = I_P \otimes \Lambda^{-1}$, then

$$\begin{aligned}
\Psi^{\mathrm{T}}\Sigma^{-1}\Psi &= \left[\Phi(\boldsymbol{x}_1), \Phi(\boldsymbol{x}_2), ..., \Phi(\boldsymbol{x}_P)\right] \begin{bmatrix} \Lambda^{-1} & 0 & ... & 0 \\ 0 & \Lambda^{-1} & ... & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & ... & \Lambda^{-1} \end{bmatrix} \begin{bmatrix} \Phi^{\mathrm{T}}(\boldsymbol{x}_1) \\ \Phi^{\mathrm{T}}(\boldsymbol{x}_2) \\ \vdots \\ \Phi^{\mathrm{T}}(\boldsymbol{x}_P) \end{bmatrix} \\
&= \sum_{i=1}^{P} \Phi(\boldsymbol{x}_i)\Lambda^{-1}\Phi^{\mathrm{T}}(\boldsymbol{x}_i) \\
&= \sum_{i=1}^{P} \left[I_n \otimes \phi(\boldsymbol{x}_i)\right]\Lambda^{-1}\left[I_n \otimes \phi^{\mathrm{T}}(\boldsymbol{x}_i)\right] \\
&= \Lambda^{-1} \otimes \sum_{i=1}^{P} \phi(\boldsymbol{x}_i)\phi^{\mathrm{T}}(\boldsymbol{x}_i)
\end{aligned}$$

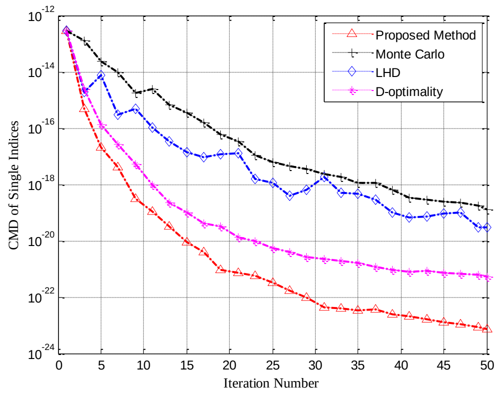
Therefore,

$$\left(\Psi^{\mathrm{T}}\Sigma^{-1}\Psi\right)^{-1} = \Lambda \otimes \left[\sum_{i=1}^{P}\left(\phi(\boldsymbol{x}_i)\phi^{\mathrm{T}}(\boldsymbol{x}_i)\right)\right]^{-1}$$

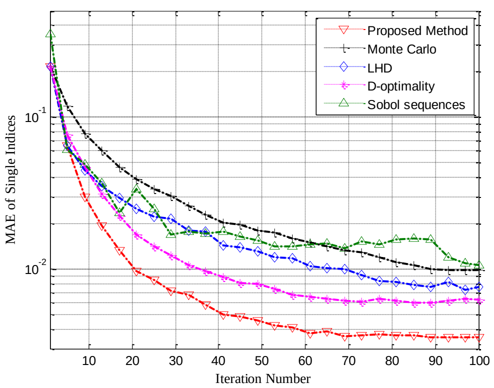
D. Comparison of performance metrics



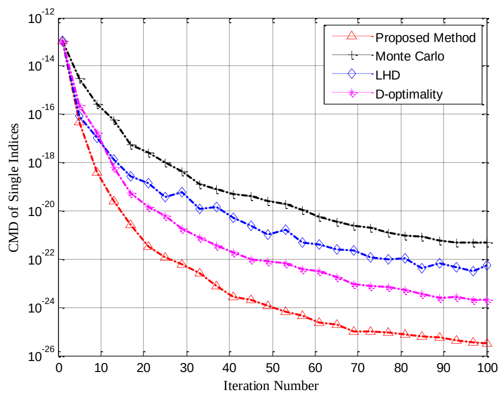
a) MAE with *p*=4



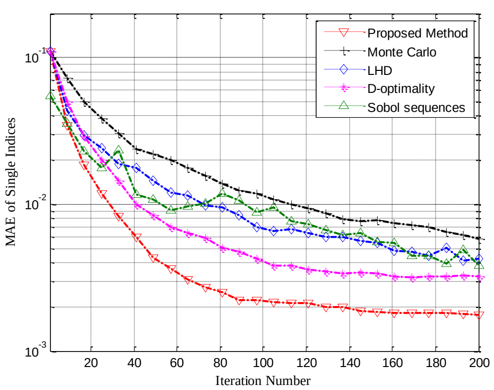
b) CMD with *p*=4



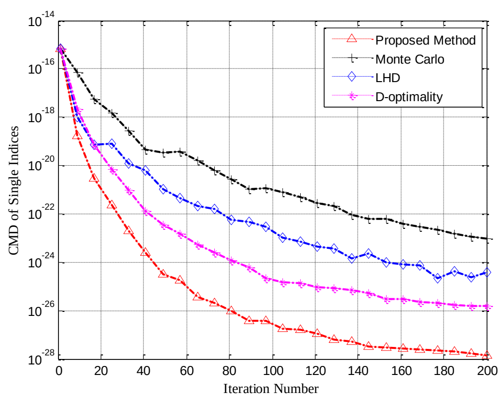
c) MAE with *p*=6



d) CMD with *p*=6

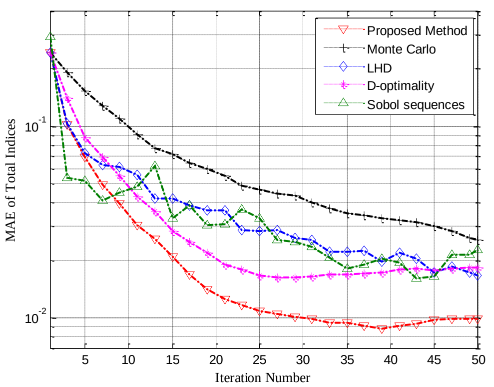


e) MAE with *p*=9

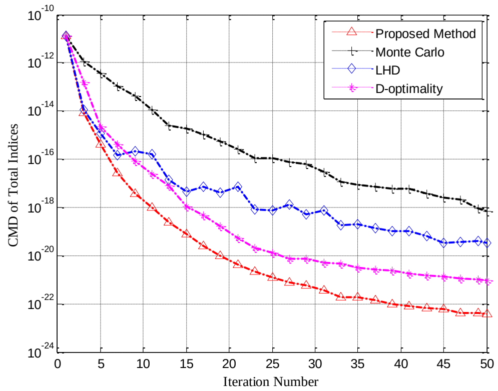


f) CMD with *p*=9

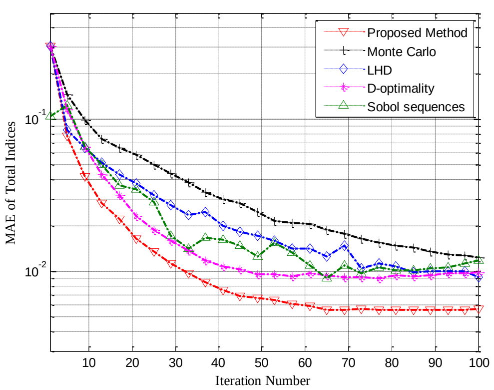
Figure 1 MAE and CMD of multivariate single effect sensitivity indices for Campbell function



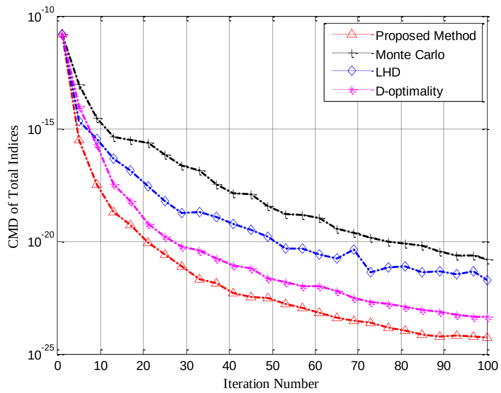
a) MAE with *p*=4



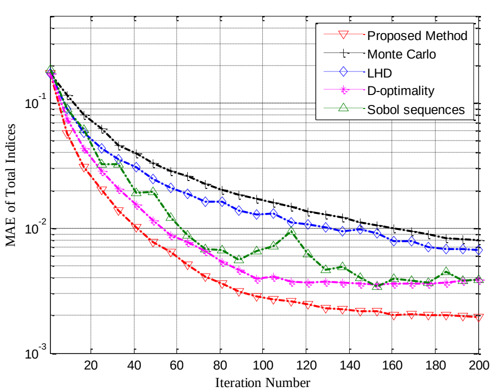
b) CMD with *p*=4



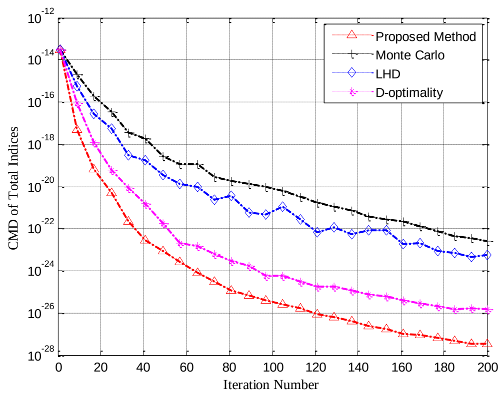
c) MAE with *p*=6



d) CMD with *p*=6

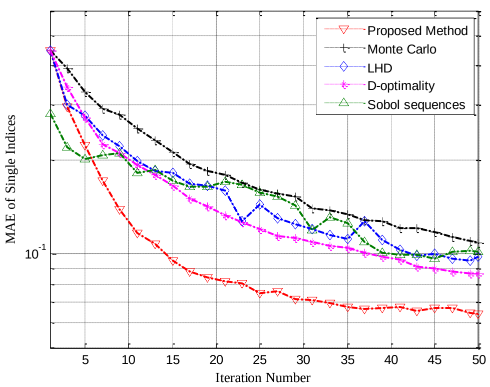


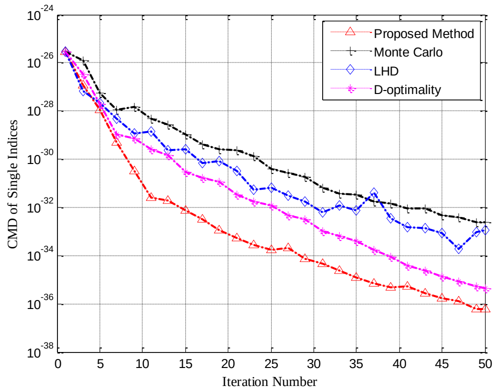
e) MAE with *p*=9



f) CMD with *p*=9

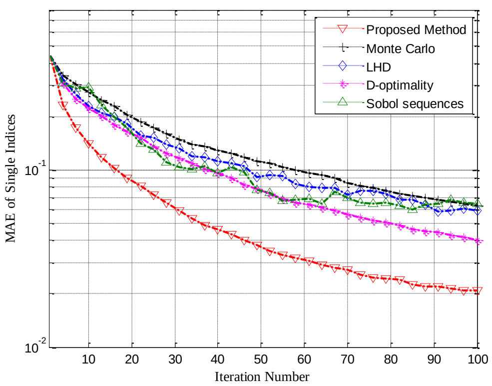
Figure 2 MAE and CMD of multivariate total effect sensitivity indices for Campbell function

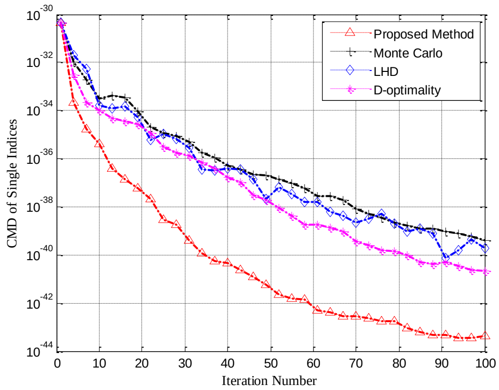




a) MAE with p=3

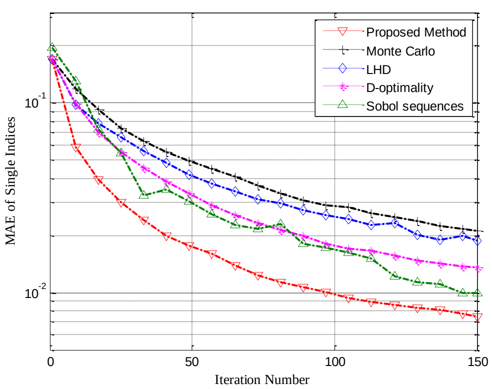
b) CMD with p=3

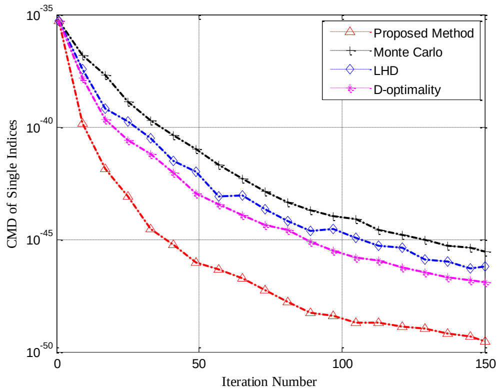




c) MAE with p=5

d) CMD with p=5

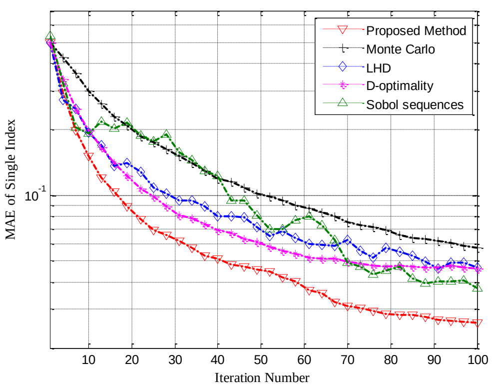




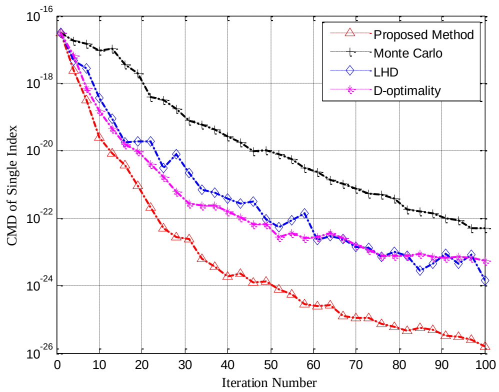
e) MAE with p=8

f) CMD with p=8

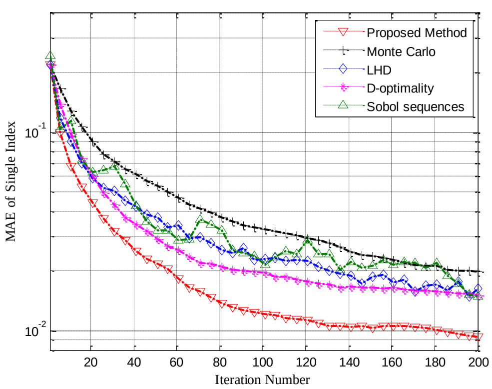
Figure 3 MAE and CMD of multivariate single effect sensitivity indices for Sobol function



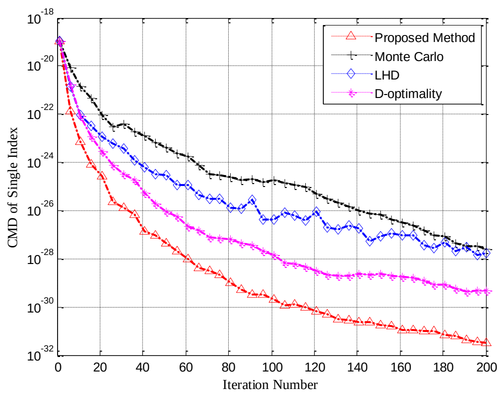
a) MAE with *p*=4



b) CMD with *p*=4

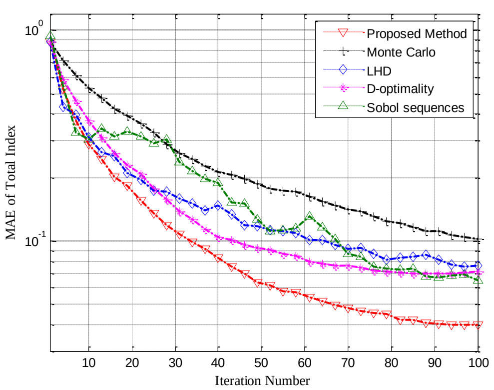


c) MAE with *p*=7

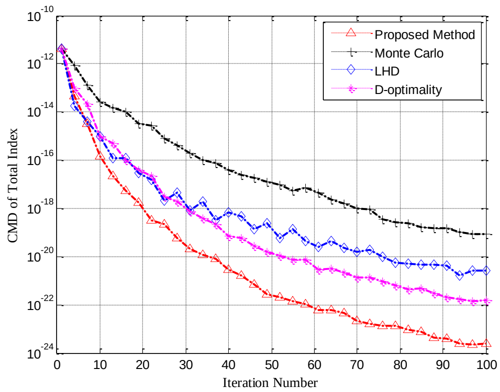


d) CMD with *p*=7

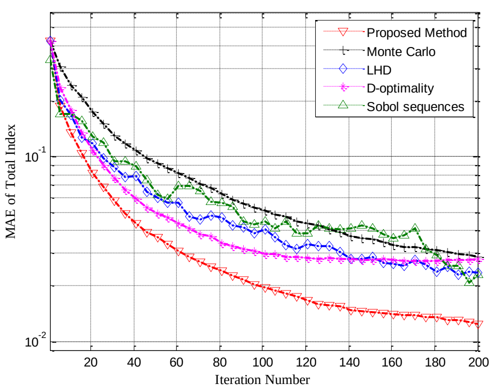
Figure 4 MAE and CMD of multivariate single effect sensitivity indices for HYMOD model



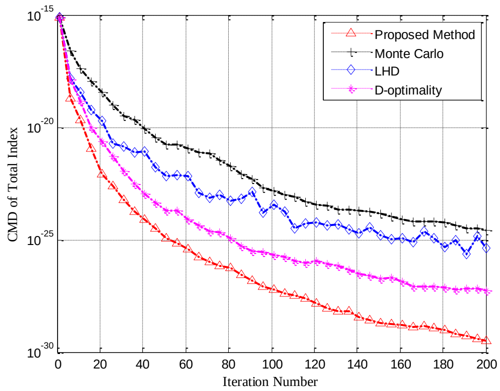
a) MAE with *p*=4



b) CMD with *p*=4

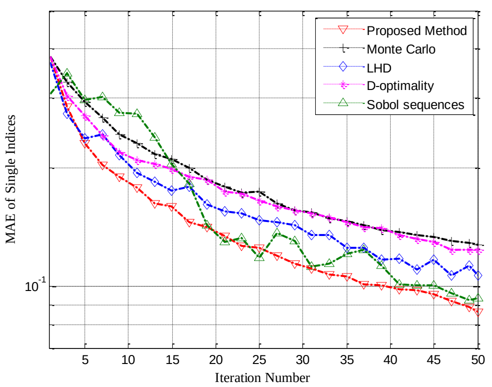


c) MAE with *p*=7

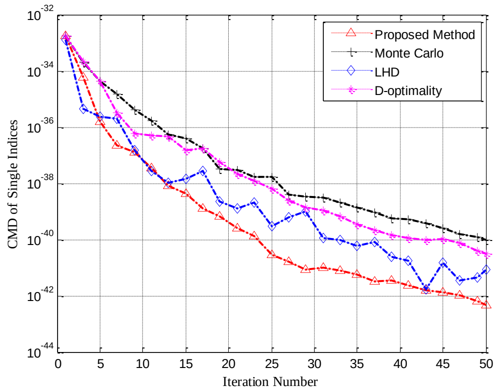


d) CMD with *p*=7

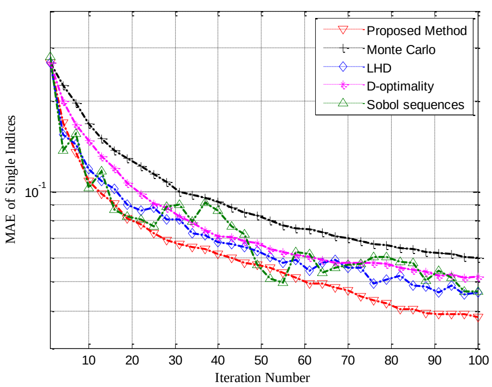
Figure 5 MAE and CMD of multivariate total effect sensitivity indices for HYMOD model



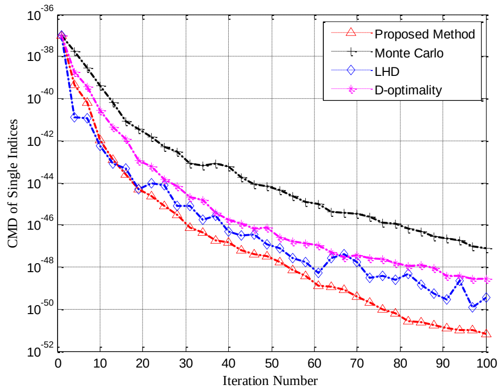
a) MAE with *p*=3



b) CMD with *p*=3



c) MAE with *p*=4



d) CMD with *p*=4

Figure 6 MAE and CMD of multivariate single effect sensitivity indices for truss structure model