Parameter Reduction for Nonlinear Models Based On Hankel Singular Values and Sensitivity Analysis

Chuili Sun and Juergen Hahn

Dept. of Chemical Engineering, Texas A&M University, College Station, TX 77843-3122

1. Introduction

Chemical processes usually contain a large number of parameters, all of which are only known to a certain degree. However, it is not always necessary to identify the values of all of the parameters from data for building a model to be used for monitoring and control. Instead, a reduction of the parameter set can be performed such that only a limited number of them need to be estimated.

This paper addresses this point and presents new techniques for reducing the parameter set of fundamental models. These novel methodologies reduce the set of parameters to be considered for a model via a technique derived from balanced model reduction [1] and further reduces this set of parameters via sensitivity analysis [2]. In the first step, the contribution that a parameter has to the input-output behavior of a system is identified based on the Hankel singular values. In a second step, the interactions between parameters are investigated via singular value decomposition of a sensitivity covariance matrix. The result can be a significantly smaller parameter set to be retained in the model.

As the first step can be very conservative, while the second step looses the physical interpretation of the parameters, this combination allows that many parameters can still retain their physical interpretation, while at the same time a reduction of the parameters comparable to that of the second method can be achieved. However, each step can be an independent method for parameter reduction, therefore, this paper presents the comparison of these two independent methods and the combination methods in the case study.

2. Parameter reduction for nonlinear systems

This section presents three new approaches for reducing the parameter set of models derived from first principles. The advantages and drawbacks of each technique are discussed in detail.

2.1. Parameter reduction based on Hankel singular values

The state space realization of a nonlinear system with parameters can be expressed as

$$\begin{aligned} \dot{x} &= f(x,\theta,u) \\ y &= g(x,\theta,u) \end{aligned}, \text{ with } x(0)=x_0 \qquad (1)$$

where $x\in\Re^n$ refers to a vector of the states of the system, $u\in\Re^l$ is a vector of inputs, $y\in\Re^m$ represents a vector containing the measured variables, and $\theta\in\Re^s$ refers to the parameters of the system. Without loss of generality suppose that all variables have been normalized.

The effect of parameters to the system can be identified by determining the contributions of parameters as if they are regarded as inputs. The contribution of the original inputs *u* can be determined by the controllability covariance matrix [3] $W_{C,1}$. Similarly, the controllability covariance matrix $W_{C,i}$, *i* = 2, …, *s*+1, can be computed for each

parameter. $W_{C,1}$ corresponds to the controllability covariance matrix representing the input-to-state behavior for excitation with the inputs *u*. $W_{C,2}$ to $W_{C,s+1}$ on the other hand contain information about the input-to-state behavior if the system is only excited by variations in one of the parameters at a time. Since the observability gramian W_O, does not depend upon the input or the parameters, the observability covariance matrix of the original system is the same as that of each subsystem. Therefore, the observability covariance matrix only needs to be computed once for this investigation.

The Hankel singular values corresponding to the inputs or each parameter can be computed based on $W_{C,i}$,*i* = 1, …, *s*+1, and W_O. These Hankel singular values can be used to determine the contribution of individual states to the input-output behavior. Let $\sigma_{i,u}$ refer to the *i*-th Hankel singular value corresponding to the inputs, and where σ_{i,θ_j} refers to the *i*-th Hankel singular value corresponding to each parameter. Based on the error bound for balanced truncation of linear system [4], a parameter θ_j can be safely assumed to be constant if

$$\sigma_{1,u} >> 2\sum_{i=1}^{n}\sigma_{i,\theta_j} \tag{2}$$

The reason behind this argument is that the maximal error that is incurred by setting a parameter to its constant value is significantly smaller than the smallest error that can result from neglecting the effect that changes in the inputs have on the system. It should be noted that this is a very conservative condition. In most cases, the parameters that satisfy $\sigma_{1,u} \approx 2\sum_{i=1}^{n}\sigma_{i,\theta_j}$ can also be reduced.

The strong points of this type of procedure for reducing the set of parameters are that

- The physical meaning of the parameters is retained during the procedure
- Explicit error bounds can be given for the effect that neglecting changes in the parameters have on the input-output behavior of a linear system
- the computed error bounds are not just valid for the steady-state behavior, but will also hold if the parameters vary with time within their uncertainty range.

A drawback of this method is that it can be very conservative, especially if the condition given by Eq. (2) is used for determining which parameters to neglect.

2.2. Parameter reduction based on sensitivity analysis

Define the sensitivity functions for the parameters: $x_\theta = \frac{\partial x}{\partial \theta} \in \Re^{n\times s}$, $u_\theta = \frac{\partial u}{\partial \theta} \in \Re^{l\times s}$, $y_\theta = \frac{\partial y}{\partial \theta} \in \Re^{m\times s}$. By differentiating Eq. (1), the parameter variation system can be obtained

$$\dot{x}_\theta = \frac{d}{dt}\left(\frac{\partial x}{\partial \theta}\right) = \frac{\partial}{\partial \theta}\left(\frac{dx}{dt}\right) = \frac{\partial}{\partial \theta}\left(\dot{x}\right) = \frac{\partial}{\partial \theta}\left(f(x,\theta,u)\right) = J_{f,x}x_\theta + J_{f,u}u_\theta + J_{f,\theta} \tag{3}$$

$$y_\theta = \frac{\partial y}{\partial \theta} = \frac{\partial}{\partial \theta}(g(x,\theta,u)) = J_{g,x}x_\theta + J_{g,u}u_\theta + J_{g,\theta} \tag{4}$$

where $J_{f,x}$, $J_{f,u}$, $J_{g,x}$, $J_{g,u}$ are Jacobian matrices of the vector functions *f*, *g* and with respect to the state and input vectors, $J_{f,p}$ $J_{g,p}$ are Jacobian matrices of the vector functions *f*, *g* with respect to the parameter vector θ.

Since all Jacobian matrices are functions of *x*, *u* and θ, it is necessary for Eq. (1) and (3) to be solved simultaneously. The set of equations need to be integrated over the time interval [0 *T*], with initial conditions $x(0) = x_0$ and $x_\theta(0) = 0$ and the input held at zero.

The sensitivity covariance matrix can then be computed from

$$\mathrm{SCM} = \int_0^\infty \bar{y}_\theta^T \bar{y}_\theta dt = \int_0^\infty (y_\theta - y_{\theta,ss})^T (y_\theta - y_{\theta,ss}) dt \tag{5}$$

where $y_{\theta,ss}$ is the final steady state value of y_θ.

The information contained in the SCM allows to extract the principle directions in the parameter space that will give the largest contribution to the parameter-output behavior. Extracting the most important directions in the parameter space, allows to take the interactions in the parameters into account and can result in a significantly reduced parameter set. This section describes how the information contained in the SCM can be extracted and interpreted.

Since the sensitivity covariance matrix defined in Eq. (5) is a positive semi-definite matrix, all eigenvalues are non-negative real numbers. Furthermore, due to its symmetry, this matrix can be diagonalized by an orthogonal matrix even if multiple eigenvalues exist [5]. Therefore, singular value decomposition can be applied to this sensitivity covariance matrix

$$\mathrm{SCM} = T^T \Lambda T \tag{6}$$

where $T \in \Re^{s \times s}$, an orthogonal matrix; superscript ^{T} indicates the transpose; and $\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ are eigenvalues of SCM.

Then, the linear transformation relationship $\bar{\theta} = T\theta$ is introduced, such that the system in the new parameter space can be given by

$$\begin{aligned} \dot{x} &= f(x, T^{-1}\bar{\theta}, u) \\ y &= g(x, T^{-1}\bar{\theta}, u) \end{aligned} \tag{7}$$

Since the diagonal entries of $\overline{\mathrm{SCM}}$ provide a measure of the importance of the corresponding parameters in the new parameter space, the parameters can be classified as belonging to an important and a less important category. The transformed parameter vector $\bar{\theta}$ can be partitioned into two parts, $\bar{\theta}_1$ and $\bar{\theta}_2$, as shown in Eq. (8). $\bar{\theta}_1$ represents

the more important parameters, which should be retained during parameter reduction and $\bar{\theta}_2$ are of lesser importance and will be reduced.

$$\bar{\theta} = \begin{bmatrix} \bar{\theta}_1 \\ \bar{\theta}_2 \end{bmatrix} \tag{8}$$

The last step is to truncate the less important states $\bar{\theta}_2$ via a projection matrix *P* leading to the reduced system

$$\begin{aligned} \dot{x} &= f(x, PT^{-1}\bar{\theta}, u) \\ y &= g(x, PT^{-1}\bar{\theta}, u) \\ \bar{\theta}_2 &= \bar{\theta}_{2,ss} \end{aligned} \tag{9}$$

where $P = [I_{k\times k} \quad 0_{k\times(s-k)}]$, *k* is the number of parameters retained.

The strong points of this second technique are that

- it is possible to take interactions between the parameters into account during the reduction procedure
- the reduced set of parameter can be significantly smaller than the original set.

The main drawback is that the physical meaning of the parameters is lost during this procedure.

2.3. Combination of above two methods

As Methods I and II both have their advantages and drawbacks, it is possible to combine them to avoid some of the disadvantages. A third technique is proposed here which performs the following steps:

(1) the sensitivity of the outputs with respect to changes in the parameters is investigated via the Hankel singular value-based technique. This investigation will keep all the parameters at their nominal value if changes in them will not result in significantly different behavior of the outputs.

(2) Compute the SCM for the reduced set of parameters from (1) and perform parameter space reduction.

The advantages that this combination of the techniques offers are that, the physical meaning of all reduced parameters in step (1) are retained, while it is possible to substantially reduce the number of parameters due to the reduction from step (2).

3. Performance evaluation

Evaluating the performance of systems where the parameter set has been reduced is an important step of the parameter reduction procedure. A technique based on Monte Carlo simulation method is applied to evaluate the performance in time domain.

This method is based upon varying the inputs and the parameters of the system, computing the trajectories of the outputs in the time domain, and comparing the results of the systems with reduced parameter sets to the behavior of the original process. A variety of different inputs and changes in the parameters has to be considered for this technique. This is achieved by randomly varying the inputs and the parameters within their bounds

and computing a large number of trajectories (usually greater than 1000). The average error introduced by reduction of the parameter set can then be computed from this run.

It is important, however, to also determine confidence intervals for the average error introduced by the reduction procedure. This is done by repeating the Monte Carlo simulations several times and determining a mean of the average error as well as a standard deviation of the average error:

$$Err = \frac{\sigma}{Mean(|y|)} \tag{10}$$

where $\sigma = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}[(y(k)-y_r(k))]^2}$ is the standard deviation. $Mean(|y|) = \frac{1}{N}\sum_{k=1}^{N}|y(k)|$ is the mean of absolute value of the responses of the original system. y represents the output of the original system and y_r refers to that of the reduced system. N is the sample length for one simulation.

Choosing the number of simulations used for each Monte Carlo simulation as well as the number of repeated Monte Carlo simulations is a key component for this evaluation method. In general, it can be said that if both numbers are sufficiently large then the standard deviation will be small.

4. Case study

A system of continuously stirred tank reactors in series is applied to illustrate the applicability of these techniques, which was studied in [6]. This system has one input, the coolant flow rate and one output, the effluent concentration from the second tank. This model contains four states, which are the temperatures and concentrations of the products in each tank. And, the system contains 14 parameters,

$$\theta = [q, C_{Af}, T_f, T_{cf}, V_1, V_2, \rho, \rho_c, C_p, C_{pc}, -\Delta H, \frac{E}{R}, hA_1, hA_2]^T \tag{11}$$

some of which are physical constants, which depend upon the properties of the streams while some others are equipment related or even related to the operating conditions.

A comparison of Hankel singular values based upon covariance matrices computed for the system excited by the input and the parameters shown in Fig. 1-a. In this figure, the sequence of parameters is consistent with that of the vector θ. Based on this figure, it can be concluded that the relative importance of the parameters is

$$\frac{E}{R} > T_f > T_{cf} > C_{Af} > -\Delta H > q > \rho = C_p > \rho_c = C_{pc} > V_1 > hA_1 > V_2 > hA_2 \tag{12}$$

It also can be seen that there are three parameters, hA_1, V_2 and hA_2, whose Hankel singular values are very small and can be neglected.

The sensitivity covariance matrix is obtained based on the system trajectories. The eigenvalues of the sensitivity covariance matrix in descending order are shown in Fig. 1-b.

For the combination method, the initial screening step, based on the Hankel singular values, determined that only hA_1, V_2 and hA_2 are to be reduced. Parameter space reduction is then performed on the 11 remaining parameters.

Model performance evaluation for this nonlinear example is conducted in the time domain by use of Monte Carlo simulations. Each Monte Carlo experiment included simulating the system with variations in the inputs and the parameters 5,000 times. The variations of the inputs and parameters are generated randomly in their ranges based on a uniform distribution. The errors between the original system and reduced systems can be computed as defined in Eq. (10). Based upon the results for 5000 simulations, the percentage of the errors below threshold values of 0.0005, 0.001, 0.005, 0.01 and 0.05 is recorded. While this Monte Carlo run determined value, it is important to repeat this experiment several times to statistically evaluate the results. For this work, the experiment has been repeated 10 times and standard deviations of the mean value of the error bound have been obtained. For example, for each experiment, the percentages of the error below 0.01 for the system reduced by the combination method were determined to be 0.9936, 0.9938, 0.9936, 0.9928, 0.9936, 0.9926, 0.9932, 0.9950, 0.9938 and 0.9932, respectively. Then the estimation of this percentage is 0.9935 ± 0.00066. For each threshold value and each method, such an estimation is generated. Fig. 2 shows the comparison of the statistics results for three methods in four cases: where 4, 5, 6 and 7 parameters are retained.

From these figures, it can be concluded that all three methods achieve a satisfactory performance for the reduced system. Additionally, it can be seen that if the same number of parameters needs to be retained, that the sensitivity analysis method and the combination method have a better performance than Method I (Hankel singular value-based technique). This point is also illustrated by Fig. 3, in which the relative errors for three methods are plotted for one Monte Carlo experiment. It can also be concluded for this example that Method II (sensitivity analysis method) and Method III (combination method) have comparable performance. When the number of retained parameters is more than 5, then the percentages of the models that Method III produces do not exceed the threshold are slightly larger than those for Method II. On the other hand, when the number of retained parameters is less or equal to 5, the Method II shows a slightly better performance. Another comparison is shown in Table 1. The percentages in this table represent the number of cases where the relative error resulting from Method III is smaller than the one for Method II. It can be seen Method III slightly outperforms Method II for this study, even though some information is lost in the initial screening step for this technique

5. Conclusions

Techniques for reducing the parameter set of fundamental models are presented in this paper. The first method focuses on determining the relative importance of parameters for the system behavior based on Hankel singular values computed for each parameter. The second technique is based on parameter space reduction, which allows reduction of a significant number of parameters while retaining most of the system behavior. Considering the advantages and drawbacks of these two methods, a combination of these two methods is also developed in this work. An initial screening step similar to analysis via Hankel singular values is applied followed by a reduction of the parameter space spanned by the remaining parameters.

All three techniques are illustrated via a CSTR example. While each method can result in a good approximation, the number of parameters reduced via Method II and III is significantly larger than for the first technique if comparable performance is to be achieved. Both Method II and the combination method exhibited comparable performance in the

example, where the combination technique has the advantage that the physical interpretation of some parameters is retained in the model.

References

[1] B.C. Moore, Principal component analysis in linear systems: controllability, observability, and model reduction, IEEE Transactions on Automatic Control 26 (1) (1981), 17-32.

[2] A. Saltelli, K. Chan, E. M. Scott, Sensitivity Analysis, Wiley Series in Probability and Statistics, Chichester, New York, Wiley, 2000.

[3] J. Hahn, T.F. Edgar, W. Marquardt, Controllability and observability covariance matrices for the analysis and order reduction of stable nonlinear systems, Journal of Process Control 13 (2003), 115-127.

[4] Glover, K., 1984. All optimal Hankel-norm approximations of linear multivariable systems and their L-errors bounds, International Journal of Control 39 (6), 1115-1193.

[5] Fairman, F.W., 1998. Linear Control Theory: The State Space Approach, John Wiley & Sons, New York.

[6] Henson, M.A., Seborg, D.E., 1990. Input-output linearization of general nonlinear processes, AIChE Journal, 36 (11), 1753-1757.

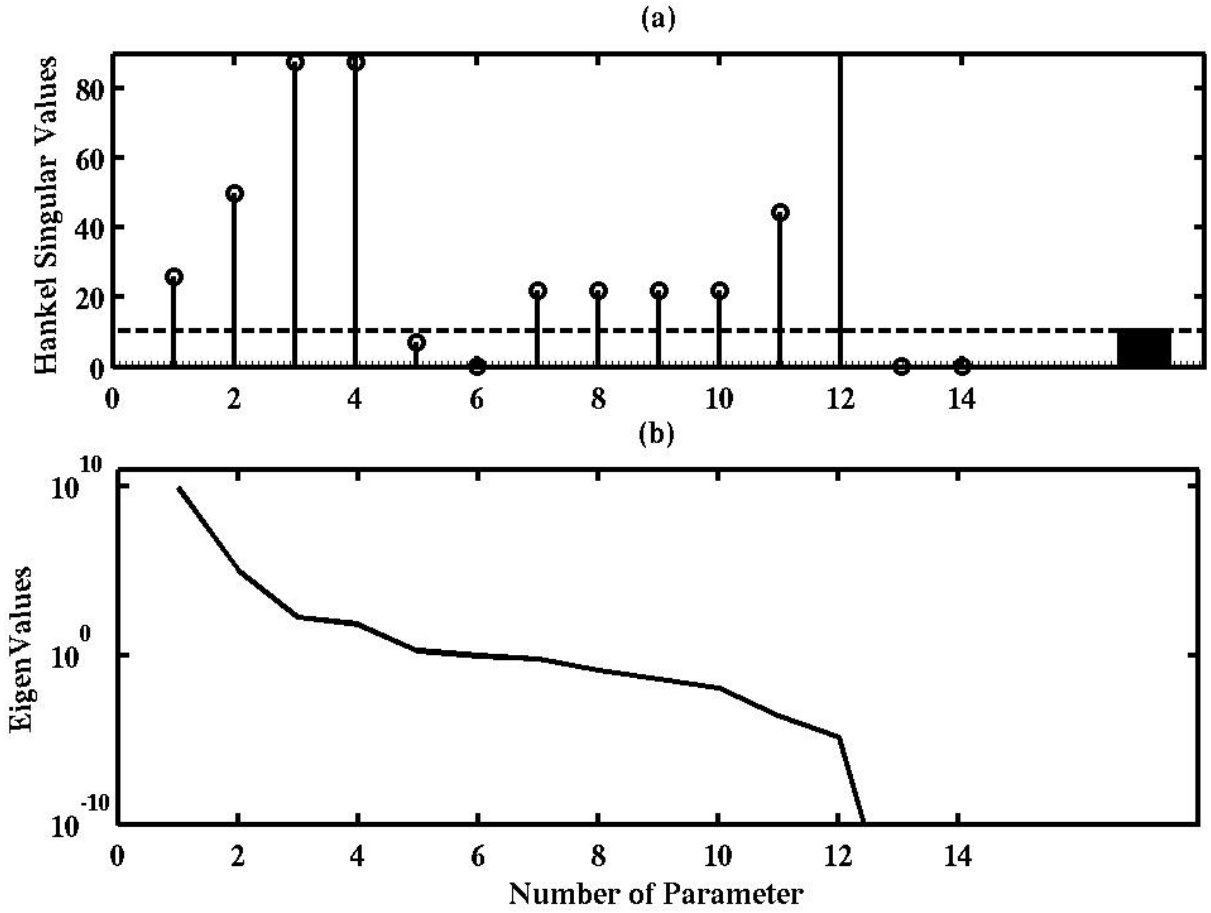

Fig. 1. Hankel singular values for parameters (Method I) and eigenvalues of sensitivity covariance matrix (Method II), for example 2.

(a) Comparison of Hankel singular values in Method I

Legend: Twice of the sum of Hankel singular values for parameters, The 1st Hankel singular value for inputs, A threshold line with the value equal to $\sigma_{1,u}$, A threshold line with the value equal to $0.1\sigma_{1,u}$

(b) Eigenvalues of sensitivity covariance matrix in Method II

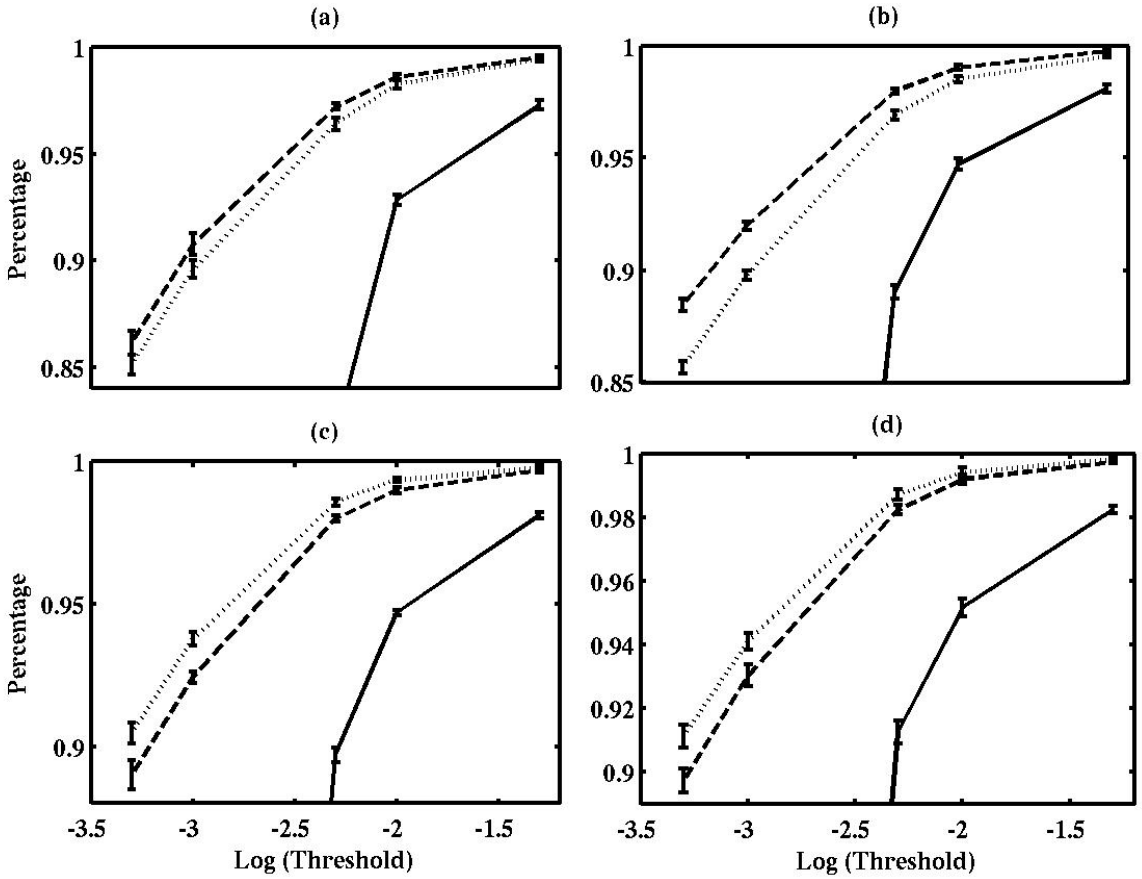

Fig. 2. Comparison of statistics results for three methods: (a) 4 parameters retained; (b) 5 parameters retained; (c) 6 parameters retained; (d) 7 parameters retained.

Legend: ▬ Method I; ▬ ▬ Method II; ····· Method III

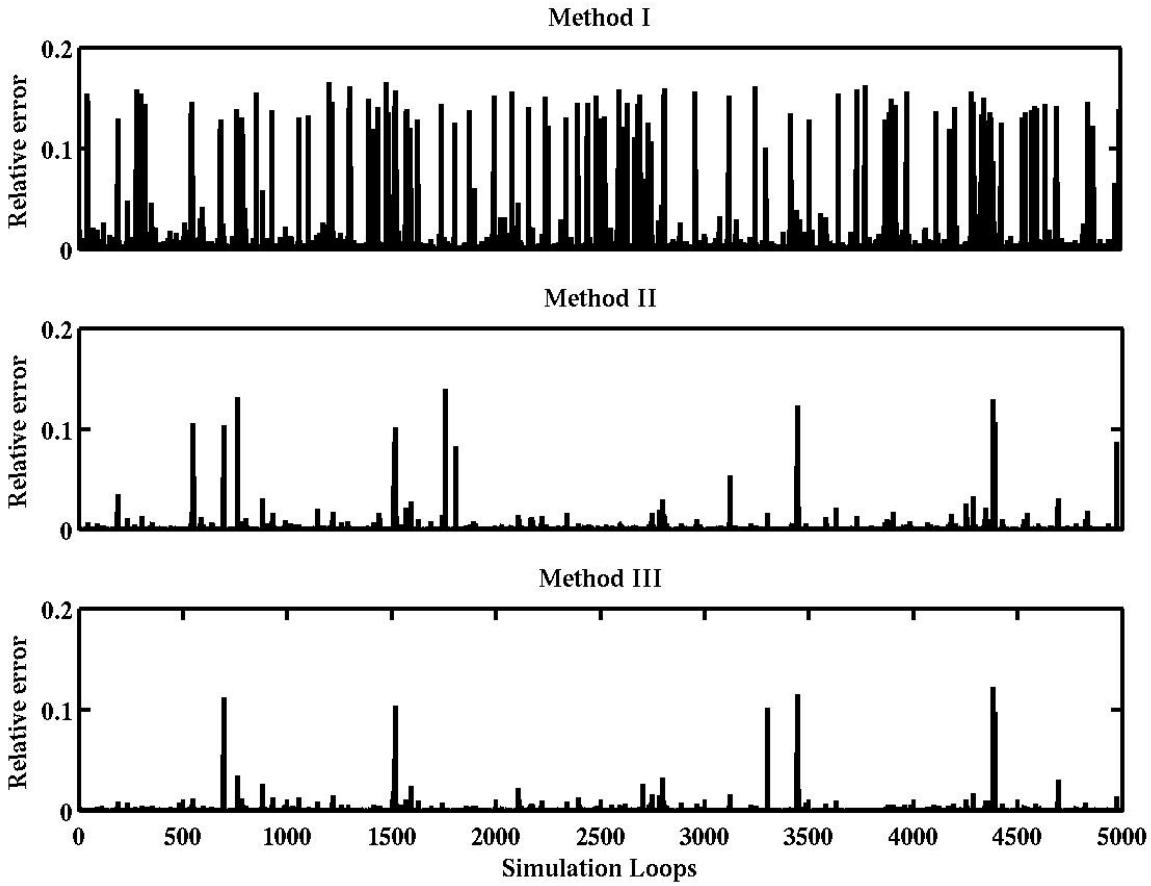

Fig. 3. The relative error set in one Monte Carlo simulation.

Table 1 The percentage of cases in which Method III is better than Method II.

	4 retained	5 retained	6 retained
Percentage	52.69±0.61%	39.43±0.59%	57.74±0.57%
	7 retained	8 retained	9 retained
Percentage	60.09±0.87%	59.52±0.37%	70.91±0.62%