A new numerical scheme for simulating non-gaussian and non-stationary stochastic processes

Zhibao Zheng\textsuperscript{a,b,*}, Hongzhe Dai\textsuperscript{a,b}, Yuyin Wang\textsuperscript{a,b}, Wei Wang\textsuperscript{a,b}

\textsuperscript{a}Key Lab of Structures Dynamic Behavior and Control, Harbin Institute of Technology, Ministry of Education, Harbin 150090, China
\textsuperscript{b}School of Civil Engineering, Harbin Institute of Technology, Harbin 150090, China

Abstract

This paper presents a new numerical scheme for simulating stochastic processes specified by their marginal distribution functions and covariance functions. Stochastic samples are firstly generated to automatically satisfy target marginal distribution functions. An iterative algorithm is proposed to match the simulated covariance function of stochastic samples to the target covariance function, and only a few times iterations can converge to a required accuracy. Several explicit representations, based on Karhunen-Loève expansion and Polynomial Chaos expansion, are further developed to represent the obtained stochastic samples in series forms. Proposed methods can be applied to non-gaussian and non-stationary stochastic processes, and three examples illustrate their accuracies and efficiencies.

Keywords: Stochastic samples, Non-gaussian, Non-stationary, Karhunen-Loève expansion, Polynomial Chaos expansion

\*Corresponding author.
Email address: zhibaozheng@hit.edu.cn (Zhibao Zheng)

Preprint submitted to Journal August 11, 2020
1. Introduction

With widely developments of uncertainty quantification theories and methods, stochastic problems involving uncertainties commonly arise in various fields of engineering, such as computational mechanics [1], financial analysis [2] and biomedical science [3]. A large number of these problems involves uncertain quantities which should be modeled as random processes or fields. On the one hand, assumptions regarding probabilistic distributions are made due to the incomplete experimental data [4]. On the other hand, stochastic simulations are provided for sufficient observation data. Thus, applications of stochastic process and field theories to engineering problems have gained considerable interests. In general, stochastic processes are assumed to be gaussian because of simplicity and the Central Limit Theorem. Since the Gaussian stochastic processes can be completely described by their second-order statistics, [5], methods for simulating gaussian stochastic processes [6, 7, 8] have been quite well established. However, the gaussian assumption does not work due to the fact that some physical phenomena are obviously non-gaussian in some cases [9]. The difficulty for simulating non-gaussian stochastic processes is that all the joint distribution functions are needed to completely characterize the non-gaussian properties. The problem is even further complicated when the process is also non-stationary since the marginal distributions depend on time or space. With these motivations, the efficient simulation of non-gaussian stochastic processes are urgent because of practical and theoretical importance.

Spectral representation is the first widely developed method for the simulation of non-gaussian stochastic processes [10]. This method is implemented
in frequency domain and is initially developed for gaussian stochastic processes \[11\]. It has been extended to non-gaussian stochastic processes by combining the spectral representation method with non-linear transformations \[10\], i.e., transforming gaussian stochastic samples generated by the spectral representation method into the non-gaussian stochastic process and matching the target power spectral density function and non-gaussian marginal distribution function. Extensive studies based on this method can be found in \[12, 13, 14, 15\]. Different from the spectral representation method, Karhunen-Loève (KL) expansion \[1, 16\] is implemented in time or space domains, which is usually used in the simulation of stationary and non-stationary gaussian processes \[17, 18, 19, 20\]. Iterative algorithms for updating the non-gaussian expanded random variables are proposed in \[21, 22\] for the simulation of non-gaussian stochastic processes. The method can be applied to highly skewed non-gaussian marginal distribution functions. Hence, KL expansion provides a unified and powerful framework for the simulation of stochastic processes, which is potentially capable of providing a better fit to non-gaussian and non-translational data \[23\]. Another important technique, Polynomial Chaos (PC) expansion, has also been developed for simulation of non-gaussian and non-stationary stochastic processes and fields in \[24, 25\]. The method represents the target stochastic process and field as multidimensional Hermite polynomial chaos in a set of normalized gaussian random variables. The accuracy and efficiency of this method were further examined in \[26, 27\].

In this paper, we present numerical schemes for simulating non-gaussian and non-stationary stochastic processes that have been specified by their covariance functions and non-gaussian marginal distribution functions. The
basic idea is to firstly generate stochastic samples that satisfying target marginal distribution functions, and then match target covariance functions by developing delicate iterative algorithm. In this way, the simulation of both gaussian and non-gaussian stochastic processes can be implemented in an unified framework since marginal distribution functions are automatically satisfied by generated samples, and the accuracy and efficiency of the simulation are only dependent on matching the target covariance functions. Another advantage is that there are no differences in the application of the proposed iterative algorithm to stationary and non-stationary stochastic processes. Thus, the proposed method can be considered as a unified numerical scheme for simulating samples of stochastic processes. Further, it’s usually not convenient to apply stochastic samples in practical problems. In this paper, we exploit KL expansion for expanding the obtained stochastic samples since KL expansion is optimal among series expansion methods in the global mean square error with respect to the number of random variables in the representation. Thus, the proposed strategy is capable of representing stochastic processes with sufficient accuracy with as few random variables as possible. In order to meet the requirements of different practical problems, we also exploit PC expansion and KL-PC expansion (combination of KL expansion and PC expansion) to represent the obtained stochastic samples, whose methodology are similar to KL expansion but based on different expansions. The accuracies and efficiencies are demonstrated by several numerical examples. Proposed methods can be readily generalized to multi-dimensional random fields [25, 28, 29, 30, 31], but it’s beyond the scope of this article and will be studied in subsequent papers.
The paper is organized as follows: a new algorithm for simulating stochastic samples is presented in Section 2. Section 3 develops several numerical algorithms for representing the obtained stochastic samples and three illustrative examples are finally given in Section 4 to demonstrate the proposed algorithms.

2. Simulation of stochastic samples

Consider a stochastic process \( \omega(x, \theta) \), \( x \in D \) specified by its covariance function \( C(x_1, x_2) \) and marginal distribution function \( F(y; x) \). In order to obtain stochastic samples that satisfying the target covariance function and marginal distribution function, we discretize spatial domain as \( x = \{x_1, \cdots, x_n\} \) and generate random variables samples \( \{\{\eta_i(\theta_k)\}_{k=1}^N\}_{i=1}^n \) according to the target marginal distribution function \( F(y; x) \), where \( N \) is the number of random variables and \( \eta_i(\theta) = \omega(x_i, \theta), i = 1, \cdots, n \). Note that random variables \( \{\eta_i(\theta)\}_{i=1}^n \) automatically satisfy the target marginal distribution function, i.e., \( \eta_i(\theta) \sim F(y; x_i) \). However, the generated samples of random variables \( \{\eta_i(\theta)\}_{i=1}^n \) don’t match the target covariance functions \( C(x_1, x_2) \) since statistical correlations between samples \( \{\eta_i(\theta_k)\}_{k=1}^N \) and \( \{\eta_j(\theta_k)\}_{k=1}^N \) are no constraints. Thus, delicate algorithm is required to be developed to match the target covariance function \( C(x_1, x_2) \).

Statistical correlations between random variable \( \eta_i(\theta) \) and \( \eta_j(\theta) \) are given as

\[
T_{ij} = \frac{1}{N-1} \sum_{k=1}^{N} [\eta_i(\theta_k) - \bar{\eta}_i][\eta_j(\theta_k) - \bar{\eta}_j] \quad (1)
\]

where \( \bar{\eta}_i \) and \( \bar{\eta}_j \) are the mean of random variables \( \eta_i(\theta) \) and \( \eta_j(\theta) \), respec-
tively. Expanding Eq. (1) yields,

$$T_{ij} = \frac{1}{N - 1} \sum_{k=1}^{N} (\eta_i(\theta_k) \eta_j(\theta_k) - \bar{\eta}_i(\theta_k) \bar{\eta}_j(\theta_k) + \bar{\eta}_i \bar{\eta}_j)$$

$$= \frac{1}{N - 1} \left[ \frac{1}{N} \sum_{k=1}^{N} \eta_i(\theta_k) \eta_j(\theta_k) - \left( \frac{1}{N} \sum_{k=1}^{N} \eta_i(\theta_k) \right) \left( \frac{1}{N} \sum_{k=1}^{N} \eta_j(\theta_k) \right) \right]$$

$$= \frac{1}{N - 1} \left[ \frac{1}{N} \sum_{k=1}^{N} \eta_i(\theta_k) \eta_j(\theta_k) - \frac{1}{N} \left( \sum_{k=1}^{N} \eta_i(\theta_k) \right) \left( \sum_{k=1}^{N} \eta_j(\theta_k) \right) \right]$$

$$= \frac{1}{N - 1} \sum_{k=1}^{N} \eta_i(\theta_k) \eta_j(\theta_k) - \frac{1}{N(N - 1)} \left( \sum_{k=1}^{N} \eta_i(\theta_k) \right) \left( \sum_{k=1}^{N} \eta_j(\theta_k) \right)$$

(2)

By introducing matrix $Y$ and assembling random samples, we have

$$Y = \left[ \{\eta_1(\theta_k)\}_{k=1}^{N}, \cdots, \{\eta_n(\theta_k)\}_{k=1}^{N} \right] = \left[ \begin{array}{ccc} \eta_1(\theta_1) & \cdots & \eta_n(\theta_1) \\ \vdots & \ddots & \vdots \\ \eta_1(\theta_N) & \cdots & \eta_n(\theta_N) \end{array} \right]$$

(3)

Then Eq. (2) can be rewritten in matrix

$$T = \frac{Y^T Y}{N - 1} - \frac{Y^T U U^T Y}{N(N - 1)}$$

(4)

where $U = [1]_{N \times 1}$, and $T$ is the simulated covariance matrix of random variables samples $\left\{ \{\eta_i(\theta_k)\}_{k=1}^{N} \right\}_{i=1}^{n}$.

Both target covariance matrix $C$ and simulated covariance matrix $T$ are symmetric and positive definite, thus Cholesky decompositions can be used to performed on matrix $C$ and $T$ as

$$C = P^T P$$

(5)
and

\[ T = Q^T Q \]  \hspace{1cm} (6)

where \( Q \) and \( P \) are upper triangular matrices.

In order to match the target covariance matrix \( C \), a new random samples matrix \( Y' \) is introduced as

\[ Y' = YQ^{-1}P \]  \hspace{1cm} (7)

By the transformation in Eq.(7), simulated covariance matrix \( T' \) can match the target covariance matrix \( C \) and proof as follows,

**Proof.** Simulated covariance matrix \( T' \) of the new random samples \( Y' \) is

\[
T' = \frac{Y'^TY'}{N-1} - \frac{Y'^TTY'Q^{-1}P}{N(N-1)} \\
= \frac{P^TQ^{-TY'^TY}Q^{-1}P}{N-1} - \frac{P^TQ^{-TY'^TTYQ^{-1}P}}{N(N-1)} \\
= P^TQ^{-T} \left( \frac{Y^TYY}{N-1} - \frac{Y^TTYQ^{-1}P}{N(N-1)} \right) Q^{-1}P \\
= P^TQ^{-T}TQ^{-1}P \\
= C
\]

It has to be noted that sample realizations \( \{\eta_i(\theta_k)\}_{k=1}^N \) in each column of \( Y' \) in Eq.(7) are different from those in each column of \( Y \) in Eq.(3) due to the factor matrix \( Q^{-1}P \). The realizations in each column of \( Y' \) change with updated factor matrix \( Q^{-1}P \) and do not match the marginal distribution function \( F(y;x_i) \). A fact that re-ordering the sample realizations will
not change distributions of random variables but change the statistical correlations, i.e., simulated covariance matrix, is enlightened. Hence, we use the strategy that re-ordering the sample realizations \( \{ \eta_i (\theta_k) \}_{k=1}^N \) in each column of \( Y \) in Eq.(3) follows the ranking of the realizations in each column of \( Y' \) in Eq.(7). The target covariance matrix \( C \) is matched by repeating the procedure of Eq.(4), Eq.(6), Eq.(7) and re-ordering stochastic samples.

The resulting procedure for simulating stochastic process samples is summarized in Algorithm 1 as follows,

**Algorithm 1** Algorithm for simulating stochastic process samples
1: Discretize spatial domain \( x = \{ x_1, \ldots, x_i, \ldots, x_n \} \) and generate random variables samples \( \{ \eta_i (\theta_k) \}_{k=1}^N \sim F(y; x_i), i = 1, \ldots, n \).
2: Compute the upper triangular matrix \( P \) by use of a Cholesky decomposition in Eq.(5).
3: repeat
4: Compute upper triangular matrix \( Q^{(k-1)} \) by Eq.(6) based on simulated covariance matrix \( T^{(k-1)} \) obtained by Eq.(4).
5: Compute \( Y^{(k-1)} \) through Eq.(7) and re-order samples in each column of \( Y \) following the ranking of the realizations in each column of \( Y' \).
6: Compute simulated covariance matrix \( T^{(k)} \) by Eq.(4).
7: until \( \| T^{(k)} - C \| / \| C \| < \varepsilon \)

Algorithm 1 provides a simple and efficient framework to simulate stochastic processes samples. Spatial discrete points and initial random variables samples are generated in Step 1. The number of random variables is equal to the that of spatial points and it’s not necessary to discretize excessive
spatial points. A Cholesky decomposition of target covariance matrix $C$ is performed in Step 2. The computational cost can be neglected since only one time decomposition needs to be computed for matrix $C$. The Step 3 to Step 7 includes a loop iteration procedure to match the target covariance matrix $C$, and the computational cost in these steps is low since only Cholesky decompositions and re-ordering samples are involved. The convergence error in Step 7 can adopt 2-norm or infinite-norm (here same to 1-norm) and we adopt 2-norm in this paper. Note that, Algorithm 1 can be applied to non-gaussian and non-stationary stochastic processes and can be readily generalized to high-dimensional random fields.

3. Expansions of stochastic processes

Algorithm 1 provides an efficient procedure to simulate samples of stochastic processes. However, sample-based descriptions of stochastic processes are not suitable for subsequent applications in some cases [1, 4] and it’s necessary to develop methods to represent the obtained samples of stochastic process. In general, a stochastic process $\omega(x, \theta)$ can be expressed as

$$\omega(x, \theta) = \sum_{i=0}^{\infty} \xi_i(\theta) f_i(x) \quad (8)$$

where $\{\xi_i(\theta)\}_{i=0}^{\infty}$ and $\{f_i(x)\}_{i=0}^{\infty}$ are a set of random variables and deterministic functions, respectively. In practical, Eq.(8) can be truncated at the term $M$ as

$$\omega(x, \theta) = \sum_{i=0}^{M} \xi_i(\theta) f_i(x) \quad (9)$$

There exist three unknown ‘variables’ in Eq.(9), i.e., stochastic process $\omega(x, \theta)$, random variables $\{\xi_i(\theta)\}_{i=0}^{M}$ and deterministic functions $\{f_i(x)\}_{i=0}^{M}$. 
Eq. (9) can be determined if any two variables are available. In this context, only a set of random variables \( \{\xi_i(\theta)\}_{i=0}^M \) or deterministic functions \( \{f_i(x)\}_{i=0}^M \) are required to be determined since samples of the stochastic process \( \omega(x, \theta) \) have been obtained by Algorithm 1. Three strategies are developed to simulate stochastic processes through Eq. (9): (i) select a set of deterministic basis \( \{f_i(x)\}_{i=0}^M \), then compute random variables \( \{\xi_i(\theta)\}_{i=0}^M \); (ii) select a set of random variables \( \{\xi_i(\theta)\}_{i=0}^M \), then compute deterministic functions \( \{f_i(x)\}_{i=0}^M \); (iii) select a set of deterministic basis \( \{f_i(x)\}_{i=0}^M \) and random variables \( \{\xi_i(\theta)\}_{i=0}^M \), then compute unknown projection coefficients. We’ll develop corresponding algorithms based on above strategies in Section 3.1, Section 3.2 and Section 3.3 respectively.

3.1. Algorithm based on stochastic samples and KL expansion

In order to obtain deterministic basis \( \{f_i(x)\}_{i=0}^M \) in Strategy (i), we adopt Karhunen-Loève (KL) expansion due to its minimum mean-square error property. The KL expansion is a special case of Eq. (8) and is based on a spectral decomposition of the covariance function \( C(x_1, x_2) \) of the stochastic process \( \omega(x, \theta) \) with the form

\[
\omega(x, \theta) = \bar{\omega}(x) + \sum_{i=1}^{M} \xi_i(\theta) \sqrt{\lambda_i} f_i(x)
\]

(10)

where \( \bar{\omega}(x) \) is the mean function of the stochastic process \( \omega(x, \theta) \), \( M \) is the number of terms of KL expansion and \( \{\xi_i(\theta)\}_{i=1}^M \) is a set of uncorrelated random variables with zero mean and unit variance, i.e.,

\[
E\{\xi_i(\theta)\} = 0, \ E\{\xi_i(\theta)\xi_j(\theta)\} = \delta_{ij}
\]

(11)
and given by

$$\xi_i(\theta) = \frac{1}{\sqrt{\lambda_i}} \int_D [\omega(x, \theta) - \bar{\omega}(x)] f_i(x) \, dx \quad (12)$$

where \(\{\lambda_i\}\) and \(\{f_i(x)\}\) are the eigenvalues and eigenfunctions of the covariance function \(C(x_1, x_2)\), obtained from solving the following homogeneous Fredholm integral equation of the second kind

$$\int_D C(x_1, x_2) f_i(x_1) \, dx_1 = \lambda_i f_i(x_2) \quad (13)$$

which satisfies

$$\int_D f_i(x) f_j(x) \, dx = \delta_{ij} \quad (14)$$

The solution of Eq.(13) can be determined numerically for problems of practical interests. It is known that, for fixed \(M\), the resulting random process approximation \(\omega(x, \theta)\) is optimal among series expansion methods with respect to the global mean square error [1]. If the stochastic process \(\omega(x, \theta)\) is gaussian, then \(\{\xi_i(\theta)\}^{M}_{i=1}\) are independent standard gaussian random variables. But for non-gaussian stochastic processes, \(\{\xi_i(\theta)\}^{M}_{i=1}\) are generally non-gaussian and there are no general methods for determining them. In order to determine their distributions, Eq.(12) has to be solved. Here we adopt a sample-based method to compute the samples of random variables \(\{\xi_i(\theta)\}^{M}_{i=1}\) as

$$\xi_i(\theta_k) = \frac{1}{\sqrt{\lambda_i}} \int_D [\omega(x, \theta_k) - \bar{\omega}(x)] f_i(x) \, dx, \ k = 1, \cdots, N \quad (15)$$

which needs \(N\) times deterministic integral and has very low computational costs. The above method is summarized in Algorithm 2 as
Algorithm 2 Algorithm based on stochastic samples and KL expansion
1: Generate samples of the stochastic process by use of Algorithm 1.
2: Solve a set of deterministic basis \(\{f_i(x)\}_{i=1}^M\) by Eq. (13).
3: Compute random variables’ samples \(\{\xi_i(\theta_k)\}_{k=1}^N\) by Eq. (15).

Algorithm 2 provides an efficient method to expand stochastic samples obtained from Algorithm 1. Deterministic basis \(\{f_i(x)\}_{i=1}^M\) are solved by use of KL expansion, thus guaranteeing the minimum numbers of expansion items \(M\). Random variables \(\{\xi_i(\theta)\}_{i=1}^M\) are described by samples and can be computed with very low computational costs.

3.2. Algorithm based on stochastic samples and PC expansion

For the Strategy (ii), it’s not easy to select random variables \(\{\xi_i(\theta)\}_{i=1}^M\) in Eq. (9) directly, here we utilize Polynomial Chaos (PC) expansion to represent the random variables \(\{\xi_i(\theta)\}_{i=1}^M\) and then compute deterministic coefficient functions \(\{f_i(x)\}_{i=1}^M\). A general representation of the stochastic process \(\omega(x, \theta)\) in PC expansion has the following form

\[
\omega(x, \theta) = \bar{\omega}(x) + \sum_{i=1}^P \Gamma_i(\theta) f_i(x) \tag{16}
\]

where \(\bar{\omega}(x)\) is the mean function of the stochastic process \(\omega(x, \theta)\), \(\{\Gamma_i(\theta)\}_{i=1}^P\) are Polynomial Chaos basis and can be generated by Rodriguez formula [32].

The deterministic coefficient functions \(\{f_i(x)\}_{i=1}^M\) can be obtained by

\[
f_i(x) = \frac{E \{[\omega(x, \theta) - \bar{\omega}(x)] \Gamma_i(\theta)\}}{E \{\Gamma_i^2(\theta)\}} \tag{17}
\]

where \(E \{\cdot\}\) is the expectation operator and the above method is summarized in Algorithm 3 as
Algorithm 3 Algorithm based on stochastic samples and PC expansion

1: Generate samples of the stochastic process by use of Algorithm 1.
2: Choose standard random variables \( \{ \gamma_i (\theta) \}_{i=1}^{M_\gamma} \).
3: Generate PC basis \( \{ \Gamma_i (\theta) \}_{i=1}^P \) of random variables \( \{ \gamma_i (\theta) \}_{i=1}^{M_\gamma} \).
4: Compute \( \{ f_i (x) \}_{i=1}^P \) by Eq. (17).

There are various choices for specifying random variables \( \gamma_i (\theta) \) in Step 2 such as gaussian random variables and uniform random variables, and the corresponding PC basis should be adopted as Hermite Polynomial Chaos basis [1] and Generalized Polynomial Chaos basis [1, 32], respectively. Further, Algorithm 3 provides a suitable method to simulate stochastic processes if Eq. (13) is not easy to solve. However the computational efficiency decreases if the number \( M_\gamma \) of random variables \( \gamma_i (\theta) \) or the order of PC basis is too large.

3.3. Algorithm based on stochastic samples and KL-PC expansion

A natural idea is to combine methods in Section 3.1 and Section 3.2 for the Strategy (iii). Hence, we choose deterministic functions \( \{ f_i (x) \}_{i=1}^M \) obtained by Eq. (13) and express random variable \( \xi_i (\theta) \) in Eq. (10) by use of PC expansion as

\[
\xi_i (\theta) = \sum_{j=1}^P c_{ij} \Gamma_j (\theta)
\]

Substituting Eq. (18) into Eq. (10) yields

\[
\omega (x, \theta) = \bar{\omega} (x) + \sum_{i=1}^M \sum_{j=1}^P c_{ij} \Gamma_j (\theta) \sqrt{\lambda_i} f_i (x)
\]
Hence, unknown projection coefficients $c_{ij}$ can be computed by

$$c_{ij} = \frac{1}{\sqrt{\lambda_i}} E \{ \Gamma_j^2 (\theta) \} \int_{D} E \{ [\omega (x, \theta) - \bar{\omega} (x)] \Gamma_j (\theta) \} f_i (x) \, dx \quad (20)$$

The above method is summarized in Algorithm 4 as

Algorithm 4 Algorithm based on stochastic samples and KL-PC expansion

1: Generate samples of the stochastic process by use of Algorithm 1.
2: Solve a set of deterministic basis $\{ f_i (x) \}_{i=1}^{M}$ by Eq. (13).
3: Choose standard random variables $\{ \gamma_i (\theta) \}_{i=1}^{M_{\gamma}}$.
4: Generate PC basis $\{ \Gamma_i (\theta) \}_{i=1}^{P}$ of random variables $\{ \gamma_i (\theta) \}_{i=1}^{M_{\gamma}}$.
5: Compute $\{ c_{ij} \}_{i,j=1}^{i,j}$ by Eq. (20).

Algorithm 4 combines KL expansion and PC expansion since PC expansion provides an explicit representation for random variables $\{ \xi_i (\theta) \}_{i=1}^{M}$ obtained in Algorithm 2. Algorithm 4 is a good choice if sample-based descriptions are not convenient for some practical problems. Further, stochastic processes may be described directly by real samples instead of being specified by their covariance functions and marginal probability distributions. We can use real samples instead of stochastic samples in Step 1 of Algorithm 2. Algorithm 3 and Algorithm 4, thus Algorithm 1 is no longer necessary.

This section provides three numerical algorithms to expand the stochastic samples obtained by Algorithm 1. Different algorithms can be chosen according to properties of practical problems and the performances of proposed algorithms are shown in next section.
4. Numerical examples

Three numerical examples, including a non-gaussian and stationary stochastic process, a non-gaussian and non-stationary stochastic process and a strongly non-gaussian and non-stationary stochastic process, are used to demonstrate the accuracies and efficiencies of the proposed algorithms. The performances of Algorithm 2, Algorithm 3, and Algorithm 4 are indicated in Example 4.1, and only Algorithm 2 is used to represent stochastic samples in Example 4.2 and Example 4.3. Since the simulated marginal distribution function automatically matches the target one, the accuracies of the proposed methods are validated by comparing the simulated covariance function with the exact one. For all examples, spatial domains are discretized as 50 points, convergence errors are set as $5 \times 10^{-3}$ and $1 \times 10^{4}$ initial samples for each random variable are generated.

4.1. Example 1: non-gaussian and stationary stochastic process

Consider a stochastic process with beat marginal distribution function

$$F(y; p, q) = \frac{\Gamma(p + q)}{\Gamma(p) \Gamma(q)} \int_{0}^{u} z^{p-1}(1 - z)^{q-1} dz$$

(21)

and squared exponential covariance function

$$C(x_1, x_2) = e^{-(x_1-x_2)^2}$$

(22)

In Eq.(21), $\Gamma(\cdot)$ is the gamma function and parameters are given by

$$u = \frac{y - y_{\min}}{y_{\max} - y_{\min}}, \ p = 4, \ q = 2$$

(23)
The expectation function and variance function of the beat distribution in Eq.(21) are

\[
\begin{align*}
\mu_F(x) &= (y_{\text{max}} - y_{\text{min}}) \frac{p}{p+q} + y_{\text{min}} \\
\sigma^2_F(x) &= (y_{\text{max}} - y_{\text{min}})^2 \frac{pq}{(p+q)^2(p+q+1)}
\end{align*}
\] (24)

According to Eq.(22), we can obtain \(\sigma^2_F(x) = 1\). Letting \(\mu_F(x) = 0\) and solving Eq.(24) yield \(y_{\text{min}} = -\sqrt{14} \approx -3.74, y_{\text{max}} = \sqrt{3.5} \approx 1.87\).

Fig.1 shows iterations of sample covariance functions \(\{T^{(k)}\}_{k=0}^{5}\), where \(T^{(0)}\) is the covariance function of the initial stochastic samples. It indicates that a rough approximation of the target covariance function can be obtained with only one round of iteration. The convergence error in each iteration is shown in Fig.2, which demonstrates Algorithm 1 has a good convergence and a high accuracy.

![Figure 1: Iterations of simulated covariance functions \(\{T^{(k)}\}_{k=0}^{5}\).](image-url)
In order to verify the performances of the proposed methods, we utilize Algorithm 2, Algorithm 3, and Algorithm 4 to represent obtained stochastic samples, respectively. For Algorithm 3, Fig. 3 shows first four eigenfunctions and eigenvalues of $C(x_1, x_2)$ obtained by Eq. (13).

![Figure 3: Eigenfunctions (left) and eigenvalues (right) of covariance $C(x_1, x_2)$.](image)

Fig. 4 shows cumulative distribution functions (CDFs) of random vari-
ables \( \{\xi_i(\theta)\}_{i=1}^4 \) computed by Eq.(15) and Table.1 indicates that they are uncorrelated random variables, which is consistent with the theory of KL expansion.

![CDFs of random variables \( \{\xi_i(\theta)\}_{i=1}^4 \).](image)

**Figure 4:** CDFs of random variables \( \{\xi_i(\theta)\}_{i=1}^4 \).

| \( E\{\xi_i\xi_j\} \) | \( \xi_1(\theta) \) | \( \xi_2(\theta) \) | \( \xi_3(\theta) \) | \( \xi_4(\theta) \) |
|-----------------|--------------|--------------|--------------|--------------|
| \( \xi_1(\theta) \) | 1.0005       |              |              |              |
| \( \xi_2(\theta) \) | 0.0004       | 0.9995       |              |              |
| \( \xi_3(\theta) \) | 0.0042       | 0.0016       | 0.9927       |              |
| \( \xi_4(\theta) \) | 0.0272       | 0.0087       | -0.0082      | 1.0248       |

**Table 1:** Statistical correlations between \( \xi_i(\theta) \) and \( \xi_j(\theta) \)

Fig.5 shows comparisons between target (top left), sample (top mid) and KL-simulated (top right) covariance. The relative errors between target and
sample covariance (bottom left), target and KL-simulated covariance (bottom mid) and sample and KL-simulated covariance (bottom right) demonstrate the high accuracy of the proposed Algorithm 2.

Figure 5: Comparisons between target, sample and KL-simulated covariance.

For Algorithm 3, we adopt 2-order Hermite Polynomial Chaos basis, the number of total terms $M$ is 14. Fig. 6 shows $\{f_i(x)\}_{i=1}^{P}$ obtained by Eq. (17).
Comparisons between target, sample and PC-simulated covariance (the corresponding legends of relative errors are the same as Fig.5) are shown in Fig.7 which demonstrates the high accuracy of the proposed Algorithm 3.

Figure 6: Coefficient functions $\{f_i(x)\}_{i=1}^{14}$ of PC expansion.

Figure 7: Comparisons between target, sample and PC-simulated covariance.
For Algorithm 4, Table 2 shows projection coefficients $c_{ij}$ obtained by Eq. (20) and comparisons between target, sample and KL-PC-simulated covariance (the corresponding legends of relative errors are the same as Fig. 5) are shown as Fig. 8 which demonstrates the high accuracy of the proposed Algorithm 4.

\begin{table}
\centering
\caption{Coefficients $c_{ij}$ of the KL-PC expansion}
\begin{tabular}{cccccccc}
\hline
$c_{ij}$ & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
\hline
1 & -0.2273 & -0.3286 & -0.6603 & -0.4586 & 0.0511 & 0.5482 & 0.3417 \\
2 & -0.3109 & 0.8586 & 0.2774 & 0.5514 & -0.9518 & -0.6500 & -0.6917 \\
3 & 0.3333 & -0.0232 & 1.1992 & -0.1781 & 0.3611 & 0.0587 & -0.3919 \\
4 & -0.3514 & -1.0478 & 0.7115 & -0.1398 & 1.4399 & 1.0510 & 0.5405 \\
\hline
$c_{ij}$ & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\
\hline
1 & 0.2638 & -0.0346 & -0.0014 & 0.3181 & -0.3325 & -0.5259 & 0.7197 \\
2 & 0.4569 & 0.0296 & 0.7990 & -0.3317 & -0.2799 & -0.0816 & 0.2825 \\
3 & -1.0618 & 0.7456 & 0.3166 & -0.6234 & 0.2455 & -0.1098 & -0.4125 \\
4 & -1.3680 & -0.2531 & -0.2757 & -0.3412 & 0.0206 & -0.1233 & -0.9378 \\
\hline
\end{tabular}
\end{table}
4.2. Example 2: non-gaussian and non-stationary stochastic process

Consider the stochastic process with beat marginal distribution function in Eq.(21) and Brown-Bridge covariance function [21]

\[ C(x_1, x_2) = \min(x_1, x_2) - x_1 x_2 \]  

(25)

According to Eq.(25), we can obtain \( \sigma_F^2(x) = C(x, x) = x - x^2 \). Letting \( \mu_F(x) = 0 \) and solving Eq.(24) yields,

\[ y_{\min} = -\sqrt{14(x - x^2)}, \quad y_{\max} = \sqrt{3.5(x - x^2)} \]  

(26)

Fig.9 shows iterations of sample covariance functions \( \{T^{(k)}\}_{k=0}^4 \) and Fig.10 shows convergence error in each iteration, which once again demonstrate the good convergence and the high accuracy of Algorithm [1].
Example 4.1 validate Algorithm 2, Algorithm 3 and Algorithm 4 here we only consider using Algorithm 2 expand the obtained stochastic samples from Algorithm 1. According to Eq.(13), analytical eigenfunctions and eigenvalues of $C(x_1, x_2)$ can be obtained as

$$f_k(x) = \sqrt{2} \sin k\pi x, \quad \lambda_k = \frac{1}{k^2\pi^2}, \quad k = 1, 2, \ldots$$

First six eigenfunctions and eigenvalues are shown in Fig. 11.
Figure 11: Eigenfunctions (left) and eigenvalues (right) of covariance $C(x_1, x_2)$.

Fig. 12 shows CDFs of random variables $\{\xi_i (\theta)\}_{i=1}^{6}$ and they are uncorrelated as shown as Table 3, which is once again consistent with the theory of KL expansion. Fig. 13 shows comparisons between target, sample and KL-simulated covariance (the corresponding legends of relative errors are the same as Fig. 5), which verify the applicability of the proposed Algorithm 2 to non-gaussian and non-stationary stochastic processes.
Figure 12: CDFs of random variables $\{\xi_i(\theta)\}_{i=1}^6$.

Table 3: Statistical correlations between $\xi_i(\theta)$ and $\xi_j(\theta)$

| $E\{\xi_i\xi_j\}$ | $\xi_1(\theta)$ | $\xi_2(\theta)$ | $\xi_3(\theta)$ | $\xi_4(\theta)$ | $\xi_5(\theta)$ | $\xi_6(\theta)$ |
|---------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| $\xi_1(\theta)$    | 0.9991          |                 |                 |                 |                 |                 |
| $\xi_2(\theta)$    | -0.0010         | 1.0006          |                 |                 |                 |                 |
| $\xi_3(\theta)$    | -0.0025         | 0.0015          | 1.0005          |                 |                 |                 |
| $\xi_4(\theta)$    | -0.0011         | 0.0017          | 0.0004          | 1.0001          |                 |                 |
| $\xi_5(\theta)$    | -0.0004         | 0.0006          | -0.0010         | -0.0006         | 1.0000          |                 |
| $\xi_6(\theta)$    | 0.0026          | -0.0009         | -0.0003         | 0.0009          | 0.0005          | 0.9999          |
Consider a stochastic process with shifted lognormal marginal distribution proposed in [21]

\[ F(y; \mu, \delta, \sigma) = \Phi\left(\frac{\ln(y - \delta(x)) - \mu(x)}{\sigma}\right) \]  

(28)

and covariance function

\[ C(x_1, x_2) = e^{-(x_1 + x_2) - |x_1 - x_2|} \]  

(29)

The expectation function and variance function of the shifted lognormal distribution in Eq.(28) are

\[
\begin{align*}
\mu_F(x) &= \delta(x) + e^{\mu(x)} + \frac{1}{2} \sigma^2 \\
\sigma^2_F(x) &= \left(e^{\sigma^2} - 1\right)e^{2\mu(x) + \sigma^2}
\end{align*}
\]

(30)

According to Eq.(29), the variance function is \( \sigma^2_F(x) = C(x, x) = e^{-2x} \). Letting \( \mu_F(x) = 0, \sigma = 1 \) and solving Eq.(30) yield \( \mu(x) = -x - \ln\sqrt{e(e - 1)} \approx -x - 0.7707 \) and \( \delta(x) = -\frac{e^{-x}}{\sqrt{e - 1}} \approx -0.7629e^{-x} \).
Fig. 14 shows iterations of sample covariance functions $\{T^{(k)}\}_{k=0}^6$ and the corresponding convergence error of each iteration is shown as Fig. 15. The good convergence of Algorithm 2 for non-gaussian and non-stationary stochastic processes is demonstrated.

Figure 14: Iterations of sample covariance functions $\{T^{(k)}\}_{k=0}^6$. 
According to Eq. (13), first six eigenfunctions and eigenvalues of $C(x_1, x_2)$ are shown in Fig. 16.

Fig. 17 shows CDFs of random variables $\{\xi_i(\theta)\}_{i=1}^{6}$ and their uncorrelated properties are shown as Table 4. Fig. 18 shows comparisons between
target, sample and KL-simulated covariance (the corresponding legends of relative errors are the same as Fig.5), which verify the applicability of the proposed Algorithm 2 to strongly non-gaussian and non-stationary stochastic processes.

![Figure 17: CDFs of random variables \( \{\xi_i(\theta)\}_{i=1}^6 \).](image)

Table 4: Statistical correlations between \( \xi_i(\theta) \) and \( \xi_j(\theta) \)

| \( E \{\xi_i\xi_j\} \) | \( \xi_1(\theta) \) | \( \xi_2(\theta) \) | \( \xi_3(\theta) \) | \( \xi_4(\theta) \) | \( \xi_5(\theta) \) | \( \xi_6(\theta) \) |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| \( \xi_1(\theta) \) | 1.0051 |
| \( \xi_2(\theta) \) | -0.0015 0.9979 |
| \( \xi_3(\theta) \) | -0.0148 0.0014 1.0049 sym. |
| \( \xi_4(\theta) \) | 0.0141 0.0004 -0.0028 0.9988 |
| \( \xi_5(\theta) \) | -0.0050 0.0066 0.0077 0.0038 0.9997 |
| \( \xi_6(\theta) \) | -0.0093 0.0069 0.0029 -0.0076 0.0061 1.0042 |
5. Conclusion

In this paper, efficient numerical schemes have been presented for simulating non-gaussian and non-stationary stochastic processes specified by covariance functions and marginal distribution functions. In order to simulate samples of the target stochastic process, stochastic samples automatically matching the target marginal distribution function are firstly generated, and an iterative algorithm is proposed to match the target covariance function by transform the order of initial stochastic samples. Three numerical examples demonstrate the fast convergence and the high accuracy of the proposed algorithm. In order to overcome the difficulty that sample-descriptions are not convenient to applied to subsequent stochastic analysis, three numerical algorithms are developed to represent the obtained stochastic samples based on KL expansion and PC expansion. Different algorithms can be used for different problems of practical interests and the performances of the developed
algorithms are indicated by numerical examples. All proposed algorithms can be readily extended to multi-dimensional random fields and will be shown in subsequent researches.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Project 11972009). This support is gratefully acknowledged.

References

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

[2] P. Tankov, Financial modelling with jump processes, Chapman and Hall/CRC, 2003.

[3] M. Aboy, O. W. Márquez, J. McNames, R. Hornero, T. Trong, B. Goldstein, Adaptive modeling and spectral estimation of nonstationary biomedical signals based on kalman filtering, IEEE Transactions on Biomedical Engineering 52 (2005) 1485–1489.

[4] G. Stefanou, The stochastic finite element method: past, present and future, Computer methods in applied mechanics and engineering 198 (2009) 1031–1051.

[5] A. Papoulis, S. U. Pillai, Probability, random variables, and stochastic processes, Tata McGraw-Hill Education, 2002.
[6] M. Grigoriu, Evaluation of karhunen–löève, spectral, and sampling representations for stochastic processes, Journal of engineering mechanics 132 (2006) 179–189.

[7] C. E. Rasmussen, Gaussian processes in machine learning, in: Summer School on Machine Learning, Springer, 2003, pp. 63–71.

[8] K. Phoon, H. Huang, S. Quek, Comparison between karhunen–loève and wavelet expansions for simulation of gaussian processes, Computers & structures 82 (2004) 985–991.

[9] G. Deodatis, R. C. Micaletti, Simulation of highly skewed non-gaussian stochastic processes, Journal of engineering mechanics 127 (2001) 1284–1295.

[10] F. Yamazaki, M. Shinozuka, Digital generation of non-Gaussian stochastic fields, Journal of Engineering Mechanics 114 (1988) 1183–1197.

[11] M. Shinozuka, G. Deodatis, Simulation of stochastic processes by spectral representation, Applied Mechanics Reviews 44 (1991) 191–204.

[12] I. Elishakoff, Y. J. Ren, M. Shinozuka, Conditional simulation of non-gaussian random fields, Engineering Structures 16 (1994) 558–563.

[13] R. Popescu, G. Deodatis, J. H. Prevost, Simulation of homogeneous nongaussian stochastic vector fields, Probabilistic Engineering Mechanics 13 (1998) 1–13.

[14] M. Grigoriu, Simulation of stationary non-Gaussian translation processes, Journal of engineering mechanics 124 (1998) 121–126.
[15] Z. Liu, W. Liu, Y. Peng, Random function based spectral representation of stationary and non-stationary stochastic processes, Probabilistic Engineering Mechanics 45 (2016) 115–126.

[16] S. Huang, S. Quek, K. Phoon, Convergence study of the truncated karhunen–loève expansion for simulation of stochastic processes, International journal for numerical methods in engineering 52 (2001) 1029–1043.

[17] B. Sudret, A. Der Kiureghian, Stochastic finite elements and reliability: a state-of-the-art report, University of California, Berkeley (2000) 114–120.

[18] F. Poirion, I. Zentner, Non-gaussian non-stationary models for natural hazard modeling, Applied Mathematical Modelling 37 (2013) 5938–5950.

[19] H. Kim, M. D. Shields, Modeling strongly non-gaussian non-stationary stochastic processes using the Iterative Translation Approximation Method and Karhunen–Loève expansion, Computers & Structures 161 (2015) 31–42.

[20] Z. Liu, Z. Liu, Y. Peng, Dimension reduction of Karhunen-Loève expansion for simulation of stochastic processes, Journal of Sound and Vibration 408 (2017) 168–189.

[21] K. Phoon, S. Huang, S. Quek, Simulation of second-order processes using Karhunen–Loève expansion, Computers & structures 80 (2002) 1049–1060.
[22] K. Phoon, H. Huang, S. Quek, Simulation of strongly non-Gaussian processes using Karhunen–Loeve expansion, Probabilistic Engineering Mechanics 20 (2005) 188–198.

[23] H. Dai, Z. Zheng, H. Ma, An explicit method for simulating non-gaussian and non-stationary stochastic processes by Karhunen–Loève and polynomial chaos expansion, Mechanical Systems and Signal Processing 115 (2019) 1–13.

[24] S. Sakamoto, R. Ghanem, Polynomial chaos decomposition for the simulation of non-Gaussian nonstationary stochastic processes, Journal of engineering mechanics 128 (2002) 190–201.

[25] S. Sakamoto, R. Ghanem, Simulation of multi-dimensional non-gaussian non-stationary random fields, Probabilistic Engineering Mechanics 17 (2002) 167–176.

[26] B. Puig, F. Poirion, C. Soize, Non-Gaussian simulation using hermite polynomial expansion: convergences and algorithms, Probabilistic Engineering Mechanics 17 (2002) 253–264.

[27] R. Field Jr, M. Grigoriu, On the accuracy of the polynomial chaos approximation, Probabilistic Engineering Mechanics 19 (2004) 65–80.

[28] N. D. Lagaros, G. Stefanou, M. Papadrakakis, An enhanced hybrid method for the simulation of highly skewed non-gaussian stochastic fields, Computer Methods in Applied Mechanics and Engineering 194 (2005) 4824–4844.
[29] Z. Zheng, H. Dai, Simulation of multi-dimensional random fields by Karhunen–Loève expansion, Computer Methods in Applied Mechanics and Engineering 324 (2017) 221–247.

[30] W. Betz, I. Papaioannou, D. Straub, Numerical methods for the discretization of random fields by means of the Karhunen–Loève expansion, Computer Methods in Applied Mechanics and Engineering 271 (2014) 109–129.

[31] V. Christou, P. Bocchini, M. J. Miranda, Optimal representation of multi-dimensional random fields with a moderate number of samples: Application to stochastic mechanics, Probabilistic Engineering Mechanics 44 (2016) 53–65.

[32] D. Xiu, Numerical methods for stochastic computations: a spectral method approach, Princeton university press, 2010.