

Finite dimensional models for random functions

M. Grigoriu

Cornell University, Ithaca, NY 14853-3501, USA



ARTICLE INFO

Article history:

Received 19 July 2018

Received in revised form 11 September 2018

Accepted 13 September 2018

Available online 18 September 2018

Keywords:

Convergence of random series

Karhunen–Loève series

Optimization

Polynomial chaos

Translation vectors

ABSTRACT

Truncated Karhunen–Loève (KL) representations are used to construct finite dimensional (FD) models for non-Gaussian functions with finite variances. The second moment specification of the random coefficients of these representations are enhanced to full probabilistic characterization by using translation, polynomial chaos, and translated polynomial chaos models, referred to as T, PC, and PCT models. Following theoretical considerations on KL representations and T, PC, and PCT models, three numerical examples are presented to illustrate the implementation and performance of these models. The PCT models inherit the desirable features of both T and PC models. It approximates accurately all quantities of interest considered in these examples.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

The states of mechanical, economic, environmental, physical, and other systems satisfy equations with random coefficients and/or inputs and end conditions, referred to as stochastic equations. For example, consider the one-dimensional stochastic transport equation $d(\mathcal{A}(x)U'(x))/dx = Y(x)$, $x \in D \subset \mathbb{R}$, accompanied by deterministic or random boundary conditions, where the random field $\mathcal{A}(x)$ characterizes material properties and $Y(x)$ denotes a source term. Its solution $U(x)$ is a functional of the material properties $\mathcal{A}(x)$. Since $\mathcal{A}(x)$ is an uncountable family of random variables indexed by $x \in D$, its stochastic dimension is infinity. For numerical solutions of this equation, $\mathcal{A}(x)$ has to be approximated by a finite dimensional (FD) model $A(x) = h(x, Z)$, where h is a deterministic function of the spatial argument $x \in D$ and a random vector Z . The stochastic dimension of $A(x)$ is finite and equal to the dimension of Z . It is common to construct FD models of the input random fields to stochastic equations from truncated Karhunen–Loève (KL) series representations of these fields [3,5,12,20,22,27,35,36]. The resulting FD models are linear forms of deterministic functions of space and/or time with random coefficients, e.g., the components of Z for the FD model $A(x) = h(x, Z)$ of $\mathcal{A}(x)$ for the above transport equation. These coefficients are uncorrelated irrespective of the law of $\mathcal{A}(x)$. They are independent and Gaussian for Gaussian functions but dependent and non-Gaussian for non-Gaussian functions.

The probabilistic characterization of the solutions of stochastic equations and quantities of interest derived from these solutions requires full probabilistic characterization of the random coefficients in the definition of FD models, e.g., the joint distribution of Z is needed to characterize the extreme $\max_{x \in D} \{U(x)\}$ of the solution of the above transport equation with $A(x, Z)$ in place of $\mathcal{A}(x)$. To simplify calculations, it is frequently assumed that the random coefficients of the FD models of non-Gaussian functions are independent [2–4,9,11,12,34]. This assumption, referred to as the finite dimensional noise assumption, can significantly affect the solution quality since the resulting FD models are approximately Gaussian irrespective of the law of the target non-Gaussian field [19].

E-mail address: mdg12@cornell.edu.

Our objective is to develop FD models $A(x)$ which capture the essential features of the laws of target non-Gaussian functions $\mathcal{A}(x)$. A two-phase approach is proposed. First, truncated KL representations are constructed for $\mathcal{A}(x)$ such that they provide accurate representations of the second moment target properties. The random coefficients of these models, i.e., the components of the random vectors Z , are defined up to their first two moments. The construction of KL representations can be based on the multigrid finite element method in [37] which provides efficient and accurate algorithms for solving Fredholm integral equations. Second, the probabilistic characterization of Z is completed. We represent Z by translation vectors $\tilde{Z}$ (T models), truncated polynomials chaos Z_n (PC models), where n denotes the truncation level of the PC representation of Z , and translation of PC models $\tilde{Z}_n$ (PCT models), which are derived from PC models by memoryless transformations. The computation times to implement PC and PCT models are similar but much larger than the time to implement T models. The PC and PCT models are of great interest in applications since, in contrast to T models which cannot capture the dependence between the random coefficients of KL expansions, they quantify the dependence between these coefficients.

The available information consists of independent samples of Z which can be calculated from samples of the target random function $\mathcal{A}(x)$ and the expression of the FD model $A(x)$ of this function. It can be argued that there is no need to develop models for Z since (1) the joint distribution of Z can be estimated from samples of this vector provided the sample size is sufficiently large and that (2) the Rosenblatt transformation [32] can be used to construct a mapping $U \mapsto Z$, where U is a random vector with independent $U(0, 1)$ components and the same dimension of Z . Samples of U , which can be generated by elementary techniques, can be mapped into samples of Z and $A(x)$ by the transformation $U \mapsto Z$ and the definition of the FD model $A(x)$. In principle, this is a valid argument. In practice, the implementation of the Rosenblatt transformation is computationally prohibitive when dealing with large dimensional random vectors since the construction of the mapping $U \mapsto Z$ requires to (i) estimate the distribution F_1 of the first component Z_1 of Z , (ii) estimate the distributions $F_{i|i-1, \dots, 1}$ of the conditional random variables $Z_i | Z_{i-1} \dots Z_1$, $i \geq 2$, and (iii) invert the distributions F_1 and $F_{i|i-1, \dots, 1}$. Moreover, the accuracy of the resulting Rosenblatt mapping $U \mapsto Z$ depends on the sample size and the law of Z .

The paper is organized as follows. Section 2 illustrates the construction of FD models for real- and vector-valued random functions from truncated KL representations (Sects. 2.1 and 2.2) and give formulas for calculating samples of the random coefficients of these models (Sect. 2.3). The T, PC, and PCT models and the corresponding FD-based models are constructed and examined in Sects. 3, 4, and 5. The relationships between T, PC, and PCT models are explored in Sect. 6. Applications are presented in Sect. 7. They include a two-dimensional vector for which the mapping $U \mapsto Z$ is known, a real-valued non-Gaussian field, and Gaussian and Gamma diffusion processes. Conclusions are in Sect. 8.

2. Second-moment FD models

Let $\mathcal{A}(x)$, $x \in D$, be an $\mathbb{R}^d$ -valued random function defined on a subset D of $\mathbb{R}^q$. The FD models of $\mathcal{A}(x)$ have the form $A(x) = h(x, Z)$, where h is a deterministic function of $x \in D$ and a random vector Z . The model has finite stochastic dimensions which is equal to the dimension of Z . If Z is defined on a probabilistic space $(\Omega, \mathcal{F}, P)$ and h is measurable, then $A(x, Z)$ is a random element on this space. The law of $A(x)$ is completely defined by h and the distribution of Z . In many applications, h is a linear function of Z .

We limit our discussion to random functions $\mathcal{A}(x)$ which are defined on bounded subsets $D \subset \mathbb{R}^q$, have finite variances, and continuous correlation functions so that they admit KL series representations [23] (Theorem 6.2.1). Truncated versions of these series are examples of FD models of $\mathcal{A}(x)$ which are linear in the components of Z . This class of FD models is considered here. The function h is a sum of the top eigenfunctions of the correlation functions of $\mathcal{A}(x)$ multiplied by random coefficients which are collected in the random vector Z . The random vector Z of KL representations are partially specified by its first two moments unless $\mathcal{A}(x)$ is a Gaussian function in which case Z is a Gaussian vector.

The following two subsections define and summarize essential properties of KL-based FD models for real- and vector-valued random functions. Additional information can be found in [15]. The subsequent subsection calculates samples of Z from samples of $\mathcal{A}(x)$ by projection. These samples are then used to construct probabilistic models for the law of Z .

2.1. Real-valued random functions

Suppose $\mathcal{A}(x)$, $x \in D$, is a real-valued random function with mean $E[\mathcal{A}(x)] = 0$, variance $E[\mathcal{A}(x)^2] < \infty$, and continuous correlation function $r(x, y) = E[\mathcal{A}(x)\mathcal{A}(y)]$, $x, y \in D$, where $D \subset \mathbb{R}^q$ is a bounded subset. The KL representation of this function has the form

$$\mathcal{A}(x) = \sum_{k=1}^{\infty} Z_k \varphi_k(x), \quad x \in D, \quad (1)$$

where $Z_k = \lambda_k^{1/2} \xi_k$, $\{\xi_k\}$ are uncorrelated random variables with zero means and unit variances so that $\{Z_k\}$ are uncorrelated variables with zero means and variances $\{\lambda_k\}$, and $\{\lambda_k\}$ and $\{\varphi_k\}$ are the eigenvalues and eigenfunctions of $\int_D r(x, y) \varphi(y) dy = \lambda \varphi(x)$, $x \in D$. The set of eigenvalues $\{\lambda_k\}$ is at most countable, the eigenvalues $\{\lambda_k\}$ are strictly positive, and the eigenfunctions corresponding to distinct eigenvalues are orthogonal, i.e., $\int_D \varphi_k(x) \varphi_l(x) dx = 0$ for $\lambda_k \neq \lambda_l$, $k \neq l$. The above series converges in the m.s. sense and its first two moments coincide with those of $\mathcal{A}(x)$ [23] (Chap. 6).

Truncated versions,

$$A^{(m)}(x) = \sum_{k=1}^m Z_k \varphi_k(x), \quad m = 1, 2, \dots, \quad (2)$$

of the KL series in Eq. (1) are FD models of $\mathcal{A}(x)$ which are used extensively in applications [5,19,26,28,34]. The number $m \geq 1$ of terms retain from the representation in Eq. (1) is called truncation level and gives the stochastic dimension of $A^{(m)}(x)$. The implementation of these models is relatively simple and requires knowledge of only the first two moments of the target random function $\mathcal{A}(x)$.

We note that (1) the representations of Eqs. (1) and (2) hold for the infinite family of random functions which have the same mean and correlation functions of $\mathcal{A}(x)$, (2) the distribution of the $\mathbb{R}^m$ -valued random variable $Z = (Z_1, \dots, Z_m)$ is unknown and so is the law of $A^{(m)}(x)$, (3) the random variables $\{Z_k\}$ are independent $\{N(0, \lambda_k)\}$ if $\mathcal{A}(x)$ is a Gaussian random function, and (4) the eigenfunctions $\{\varphi_k\}$ provide a useful basis for the approximate representation of samples $\mathcal{A}(x)$. The following properties summarize essential features of the FD models in Eq. (2).

Property 1. If the correlation function $r(x, y)$ is square integrable and continuous in $D \times D$, then (i) the series representation $r(x, y) = \sum_{k=1}^{\infty} \lambda_k \varphi_k(x) \varphi_k(y)$ converges absolutely and uniformly in $D \times D$ and (ii) the sequence of random fields $\{A^{(m)}(x), m = 1, 2, \dots\}$ converges in mean square (m.s.) to $\mathcal{A}(x)$, i.e., $E[(\mathcal{A}(x) - A^{(m)}(x))^2] \rightarrow 0$ as $m \rightarrow \infty$. The truncation error is

$$E[(\mathcal{A}(x) - A^{(m)}(x))^2] = \sum_{k=m+1}^{\infty} \lambda_k \varphi_k(x)^2. \quad (3)$$

Proof. This property follows from the Mercer theorem ([1], Sect. 3.3, [8], Appendix 2 and Sect. 6–4, and [23], Sect. 6.2). Since $E[(\mathcal{A}(x) - A^{(m)}(x))^2] \rightarrow 0$, the truncation error can be made as small as desired. $\square$

Property 2. Augmented versions of $A^{(m)}(x)$ obtained by assuming that the random variables $\{\xi_k\}$ are independent and have known distributions [2–4,9,11,12,34] are Gaussian as $n \rightarrow \infty$.

Proof. See [19]. This implies that the laws of $A^{(m)}(x)$ and $\mathcal{A}(x)$ may differ significantly for non-Gaussian random fields. If $\mathcal{A}(x)$ is a Gaussian function, then $\{Z_k\}$ are independent Gaussian variables so that $A^{(m)}(x)$ is a Gaussian function for any $m \geq 1$. The FD model $A^{(m)}(x)$ matches the finite dimensional distributions of $\mathcal{A}(x)$ asymptotically as $m \rightarrow \infty$. $\square$

Property 3. Generally, FD models $A^{(m)}(x)$, $m < \infty$, of homogeneous target fields $\mathcal{A}(x)$ are inhomogeneous.

Proof. Suppose that $\mathcal{A}(x)$ is a weakly homogeneous random field and assume that $A^{(m)}(x)$ is also homogeneous for all $m \geq 1$. Then

$$E[(\sum_{k=1}^{m+1} Z_k \varphi_k(x))^2] = E[A^{(m)}(x)^2] + \lambda_{m+1} \varphi_{m+1}(x)^2$$

by Mercer's theorem so that $\varphi_{m+1}(x)$ must be invariant with x , which is a contradiction. If it is assumed that $\{Z_k\}$ are independent random variables, then the characteristic functions of $A^{(m)}(x)$ for the truncation levels m and $m+1$ are related by

$$E[\exp(iu \sum_{k=1}^{m+1} Z_k \varphi_k(x))] = E[\exp(iu A^{(m)}(x))] E[\exp(iu Z_{m+1} \varphi_{m+1}(x))].$$

Since the characteristic function of $A^{(m)}(x)$ is x -invariant by assumption, then the expectation $E[\exp(iu Z_{m+1} \varphi_{m+1}(x))]$ should not depend on x , a contradiction. Numerical illustrations can be found in [10]. $\square$

Property 4. Almost all samples of $A^{(m)}(x)$ are in $L^2(D)$.

Proof. We have

$$E[\|A^{(m)}(\cdot, \omega)\|_{L^2(D)}^2] = E\left[\int_D A^{(m)}(x, \omega)^2 dx\right] = \sum_{k=1}^m \lambda_k \leq \sum_{k=1}^{\infty} \lambda_k, \quad m \geq 1,$$

and $\sum_{k=1}^{\infty} \lambda_k < \infty$ by properties of the eigenvalues. Since $E[\|A^{(m)}(\cdot, \omega)\|_{L^2(D)}^2]$ is finite and $P(\|A^{(m)}(\cdot, \omega)\|_{L^2(D)}^2 > a) \leq (1/a) \sum_{k=1}^m \lambda_k \rightarrow 0$, $a \rightarrow \infty$, by Markov's inequality [31] (Sect. 5.2.3), we conclude that almost all samples of $A^{(m)}(x)$, $m \geq 1$, are square integrable on D , i.e., are elements of $L^2(D)$. $\square$

The truncation level m will be selected such that $A^{(m)}(x)$ approximates the target random function $\mathcal{A}(x)$ satisfactorily. Once selected, the truncation level m is kept fixed. For simplicity, we drop the superscript m and will refer to $A^{(m)}(x)$ as $A(x)$.

2.2. Vector-valued random functions

Suppose $\mathcal{A}(x)$, $x \in D$, is an $\mathbb{R}^d$ -valued random function whose components $\mathcal{A}_i(x)$, $i = 1, \dots, d$, have means $E[\mathcal{A}_i(x)] = 0$, variances $E[\mathcal{A}_i(x)^2] < \infty$, and continuous correlation functions $r_{ij}(x, y) = E[\mathcal{A}_i(x)\mathcal{A}_j(y)]$, $i, j = 1, \dots, d$, $x, y \in D$. We present two methods to construct FD models of $\mathcal{A}(x)$.

The standard method uses truncated KL series representations

$$A^{(m)}(x) = \sum_{k=1}^m Z_k \psi_k(x) \in \mathbb{R}^d, \quad x \in D, \quad (4)$$

of $\mathcal{A}(x)$, where m denotes the truncation level, $\{Z_k = \mu_k^{1/2} \eta_k \in \mathbb{R}\}$, $\{\eta_k\}$ are uncorrelated random variables with zero means and unit variances, $\{\mu_k, \psi_k(x) \in \mathbb{R}^d\}$ are the eigenvalues and eigenfunctions of $\int_D r(x, y) \psi(y) dy = \mu \psi(x)$, $x \in D$, and $r(x, y) = \{r_{ij}(x, y)\}$. Since the selection of the truncation level m is based on the m.s. error of $A^{(m)}(x)$ and the dominant components of $\mathcal{A}(x)$ control this error, the components of the resulting FD model may have different accuracies.

An alternative method can be used to construct FD models of $\mathcal{A}(x)$ whose components have prescribed accuracies. We develop FD models of $\mathcal{A}(x)$ component-by-component. Generally, they have different truncation levels to satisfy required accuracies, e.g., the FD model of $\mathcal{A}_i(x)$ for the truncation level m_i has the form

$$A_i^{(m_i)}(x) = \sum_{k=1}^{m_i} Z_{i,k} \varphi_{i,k}(x), \quad m_i = 1, 2, \dots, \quad i = 1, \dots, d, \quad (5)$$

where $Z_{i,k} = \lambda_{i,k}^{1/2} \xi_{i,k}$, $\{\xi_{i,k}\}$, $k = 1, \dots, m_i$, are uncorrelated random variables with zero means and unit variances so that $\{Z_{i,k}\}$ are uncorrelated variables with zero means and variances $\{\lambda_{i,k}\}$, and $\{\lambda_{i,k}\}$ and $\{\varphi_{i,k}\}$ are the top eigenvalues and the eigenfunctions of the correlation function $r_{ii}(x, y) = E[\mathcal{A}_i(x)\mathcal{A}_i(y)]$, $x, y \in D$, of the real-valued random function $\mathcal{A}_i(x)$. The properties of the individual FD models $A_i^{(m_i)}(x)$ follow directly from those of the real-valued random function in the previous subsection.

We now summarize properties of the FD model $A^{(m_1, \dots, m_d)}(x) = (A_1^{(m_1)}(x), \dots, A_d^{(m_d)}(x))$ of the vector-valued random function $\mathcal{A}(x)$ which are relevant to our discussion.

Property 5. The set of functions $\{\theta_{ij,kl} = \varphi_{i,k} \varphi_{j,l}\}$, $i, j = 1, 2, \dots, d$, $k, l = 1, 2, \dots$ is an orthonormal basis of $L^2(D \times D)$.

Proof. See [15]. This shows that the basis functions for vector-valued random functions can be obtained from those of their components by elementary calculations. The resulting representations have components with prescribed accuracies. $\square$

Property 6. The correlation functions of distinct components of $\mathcal{A}(x)$ admits the representation

$$r_{ij}(x, y) = \sum_{k,l=1}^{\infty} \langle r_{ij}, \theta_{ij,kl} \rangle \theta_{ij,kl}(x, y), \quad (x, y) \in D \times D, \quad (6)$$

where the series convergent in $L^2(D \times D)$.

Proof. This holds since $\{\theta_{ij,kl}\}$ is an orthonormal system of the Hilbert space $L^2(D \times D)$ and the sequence $\{r_{ij}^{(m_i, m_j)}\}$ in Eq. (7) is Cauchy in $L^2(D \times D)$ which is a complete space [14] (Theorem 11.2, p. 25). $\square$

Property 7. The projection

$$r_{ij}^{(m_i, m_j)}(x, y) = \sum_{k=1}^{m_i} \sum_{l=1}^{m_j} \langle r_{ij}, \theta_{ij,kl} \rangle \theta_{ij,kl}(x, y), \quad (x, y) \in D \times D, \quad (7)$$

of r_{ij} on the subspace spanned by $\{\theta_{ij,kl}\}$, $k = 1, \dots, m_i$, $l = 1, \dots, m_j$, is the best m.s. approximation for the truncation levels.

Proof. Direct calculations show that the projections $\langle r_{ij} - r_{ij}^{(m_i, m_j)}, \theta_{ij, kl} \rangle$ of the truncation error are zero, which shows that $r_{ij}^{(m_i, m_j)}$ is the best m.s. approximation of r_{ij} . $\square$

Property 8. The Fourier coefficients $\langle r_{ij}, \theta_{ij, kl} \rangle$ of the representation in Eq. (7) can be calculated from

$$\langle r_{ij}, \theta_{ij, kl} \rangle = \lambda_{i,k} \delta_{ij} + (1 - \delta_{ij}) \int_{D \times D} r_{ij}(x, y) \varphi_{i,k}(x) \varphi_{j,l}(y) dx dy. \quad (8)$$

Proof. The term for $i = j$ is that for real-valued functions in the previous section and the term for $i \neq j$ is the expression of the Fourier coefficients for r_{ij} . $\square$

As for real-valued random function, we assume that the truncation levels $\{m_i\}$ of the FD models of the components of $\mathcal{A}(x)$ have been selected such that $A^{(m_1, \dots, m_d)}(x)$ approximates satisfactorily the target random function $\mathcal{A}(x)$. Once selected, the truncation levels $\{m_i\}$ are kept fixed. For simplicity, we will denote $A^{(m_1, \dots, m_d)}(x)$ by $A(x)$.

2.3. Samples of random coefficients

Let m and $\{m_i\}$ denote the selected truncation levels for the KL representations of real- and vector-valued random function $\mathcal{A}(x)$ (see Eq. (2) and (5)). The corresponding dimensions of the vectors of random coefficients of the resulting FD models are m and $m = \sum_{i=1}^d m_i$, respectively. Samples of these coefficients can be obtained from samples of $\mathcal{A}(x)$ by projection as indicated by the following property.

Property 9. Samples $\{Z_k(\omega)\}$ and $\{Z_{i,k}(\omega)\}$ of the random coefficients of $A(x)$ are given by

$$\begin{aligned} \langle \mathcal{A}(\cdot, \omega), \varphi_p \rangle &= \sum_{k=1}^{\infty} Z_k(\omega) \langle \varphi_k, \varphi_p \rangle = Z_p(\omega), \quad p = 1, \dots, m, \quad \text{and} \\ \langle \mathcal{A}_i(\cdot, \omega), \varphi_{i,p} \rangle &= \sum_{k=1}^{\infty} Z_{i,k}(\omega) \langle \varphi_{i,k}, \varphi_{i,p} \rangle = Z_{i,p}(\omega), \quad p = 1, \dots, m_i, \quad i = 1, \dots, d \end{aligned} \quad (9)$$

for real- and vector-valued target random functions $\mathcal{A}(x)$.

Proof. Suppose $\mathcal{A}(x)$ is a real-valued random function. Consider its FD model $A(x) := A^{(m)}(x)$ and the increasing sequence of real-valued non-negative random variables

$$\|A(\cdot, \omega)\|_{L^2(D)}^2 = \int_D A(x, \omega)^2 dx = \sum_{k=1}^m Z_k(\omega)^2 \leq \sum_{k=1}^{\infty} Z_k(\omega)^2, \quad m = 1, 2, \dots,$$

which holds since $\{\varphi_k\}$ is an orthonormal system in $L^2(D)$. The expectations

$$E[\|A(\cdot, \omega)\|_{L^2(D)}^2] = \sum_{k=1}^m \lambda_k \leq \sum_{k=1}^{\infty} \lambda_k, \quad m \geq 1,$$

of these random variables are finite by properties of the eigenvalues of the correlation function of $\mathcal{A}(x)$. Markov's inequality gives $P(\|A(\cdot, \omega)\|_{L^2(D)}^2 > a) \leq a^{-1} \sum_{k=1}^m \lambda_k \rightarrow 0, a \rightarrow \infty$, so that almost all samples of $A(x)$, $m \geq 1$, are in $L^2(D)$ and can be expanded in any orthonormal basis of this space. $\square$

The components $\{Z_k, k = 1, \dots, m\}$ of Z for real-valued random functions are uncorrelated. For vector-valued random functions, the components $\{Z_{i,k}, k = 1, \dots, m_i\}$ of Z corresponding to individual components of $\mathcal{A}(x)$ are uncorrelated but the random vectors $\{Z_{i,k}\}$ and $\{Z_{j,l}\}$, $i \neq j$, may be correlated, i.e., $E[Z_{i,k} Z_{i,l}] = 0, k \neq l$, and, generally, $E[Z_{i,k} Z_{j,l}] \neq 0, i \neq j$. Since $E[\mathcal{A}(x)] = E[A(x)] = 0$ by assumption, the components of Z have zero means.

The FD models of $A(x)$ of $\mathcal{A}(x)$ defined by Eqs. (2) and (5) have the form

$$A(x) = \varphi(x) Z, \quad (10)$$

where $\varphi(x)$ and Z are matrices whose entries are basis functions and random coefficients. For real-valued random functions, $\varphi(x)$ is the row vector $(\varphi_1(x), \dots, \varphi_m(x))$ for each $x \in D$ and $Z = (Z_1, \dots, Z_m)'$. For vector-valued functions, $\varphi(x)$ is an (d, m) -matrix with non-zero entries $(\varphi_{i,1}(x), \dots, \varphi_{i,m_i}(x))$ in row $i = 1, \dots, d$ and columns $(i d + 1, \dots, i d + m_i)$ and Z is

a column vector with components $(Z_{1,1}, \dots, Z_{1,m_1}, \dots, Z_{d,1}, \dots, Z_{1,m_d})$. The covariance matrix of Z in Eq. (10) is diagonal with entries $\{\lambda_k\}$ for real-valued random functions $\mathcal{A}(x)$. This matrix has diagonal blocks along the main diagonal of size (m_i, m_i) , $i = 1, \dots, d$, with entries $\{\lambda_{i,k}, k = 1, \dots, m_i\}$ and off-diagonal blocks of size (m_i, m_j) , $i, j = 1, \dots, d$, $i \neq j$, with entries $\{(\lambda_{i,k} \lambda_{j,l})^{1/2}, k = 1, \dots, m_i, l = 1, \dots, m_j\}$.

Our objective is to develop FD models of $\mathcal{A}(x)$ of the type in Eq. (10). A two-step approach is used to achieve this objective. First, we develop probabilistic models of the random vector Z . Three models, referred to as T, PC, and PCT models, are defined and examined. The T and PC models are available in the literature and have been studied extensively [13,17]. The PCT models attempt to retain the desirable features of T and PC models and overcome their limitations. Second, the properties of the FD models $A(x)$ of $\mathcal{A}(x)$ are inferred from those of the T, PC, and PCT models of Z and the definition of the FD models $A(x)$ in Eq. (10).

3. T models

Let $\{F_i\}$ denote the marginal distributions of the components $\{Z_i\}$, $i = 1, \dots, m$, of Z . The components of Z are uncorrelated for real-valued functions. For vector-valued functions, the components of Z are uncorrelated if they correspond to the same component of $A(x)$ and correlated otherwise. Denote by $\{\zeta_{ij}\}$ the correlation coefficients of Z . The distributions $\{F_i\}$ and the correlations $\{\zeta_{ij}\}$ can be estimated from samples of Z . Alternative correlations can be used to construct translation models, e.g., Spearman's correlation [7]. It is assumed that the set of samples of Z is sufficiently large such that the statistical uncertainty in these estimates can be disregarded. We construct translation models for the random vector Z and use these models to characterize $A(x)$.

3.1. Random coefficients

The components of the translation model $\tilde{Z}$ of Z are defined by

$$\tilde{Z}_i = F_i^{-1} \circ \Phi(G_i) = \tilde{h}_i(G_i), \quad i = 1, \dots, m, \quad (11)$$

where F_i^{-1} is the inverse of F_i and Φ denotes the distribution of $N(0, 1)$. For simplicity, it is assumed that the distributions F_i are continuous with no flat spots. The random variables $\{\tilde{Z}_i\}$ are nonlinear functions of the components of a standard Gaussian vector $G = (G_1, \dots, G_m)$ with correlation matrix $\{\rho_{ij} = E[G_i G_j]\}$. Generally, $\{\rho_{ij}\}$ is selected such that the covariance matrix of $\tilde{Z}$ is as close as possible to that of Z [17] (Chap. 3). Optimization algorithms can be used to find the correlations $\{\rho_{ij}\}$ of the Gaussian image G of $\tilde{Z}$. The law of the random vector $\tilde{Z}$ is completely defined by the correlations $\{\rho_{ij}\}$ and the marginal distributions $\{F_i\}$. The random vector $\tilde{Z}$ is defined on the same probability space as G and has the following properties.

- (1) $\tilde{Z}_i \stackrel{d}{=} Z_i$, $i = 1, \dots, m$, irrespective of the correlations $\{\rho_{ij}\}$ of their Gaussian images by construction.
- (2) The discrepancy between the joint distributions of $\tilde{Z}$ and its version with independent components can be bounded by

$$\left| P\left(\bigcap_{i=1}^m \{\tilde{Z}_i \leq x_i\}\right) - \prod_{i=1}^m F_i(x_i) \right| \leq c \sum_{1 \leq i < j \leq m} \exp\left(-\frac{(u_i^2 + u_j^2)/2}{1 + |\rho_{ij}|}\right), \quad (12)$$

where $c > 0$ is a constant and $u_i = \tilde{h}_i^{-1}(x_i)$, $i = 1, \dots, m$. The bound results from the normal comparison lemma by a change of variables [25] (Theorem 4.2.1).

- (3) The components of $\tilde{Z}$ corresponding to uncorrelated components of Z are independent since, if Z_i and Z_j are uncorrelated, ρ_{ij} has to be zero [17] (Sect. 3.1.1) which means that the components G_i and G_j are independent and so are their images $\tilde{Z}_i$ and $\tilde{Z}_j$.
- (4) The components of $\tilde{Z}$ have independent tails, i.e., $P(\tilde{Z}_i > x \mid \tilde{Z}_j > x) \rightarrow 0$ as $x \rightarrow \infty$ for $i \neq j$, so that $\tilde{Z}_i$ and $\tilde{Z}_j$ are unlikely to be simultaneously large [30] (Sect. 5.2).

3.2. FD models for random functions

The translation models $\tilde{A}(x)$ of $A(x)$ are given by Eq. (10) with Z replaced by $\tilde{Z}$ in Eq. (11). Simplicity and computation efficiency are the main features of translation FD models. However, the random functions $\tilde{A}(x)$ have some limitations.

- (1) The random variable $\tilde{A}(x) = \sum_{k=1}^m Z_k \varphi_k(x) =: \sum_{k=1}^m U_k(x)$ for a fixed arbitrary $x \in D$ is Gaussian as the KL truncation level $m \rightarrow \infty$, if $\int_{|z|>\alpha} z^2 dF_{U_k}(z) \sim O(m^{-(1+\alpha)})$, $\alpha > 0$, and $\mathcal{A}(x)$ is real-valued. The proof uses the Lindeberg-Feller condition ([31], Theorem 9.8.1) and can be found in [19].
- (2) Under the above conditions, $\tilde{A}(x)$, $x \in D$, becomes a Gaussian function as $m \rightarrow \infty$. It is sufficient to show that $\sum_{r=1}^R \tilde{A}(x_r) = \sum_{r=1}^R \sum_{k=1}^m \beta_k U_k(x_r)$ is asymptotically Gaussian as $m \rightarrow \infty$ for any $R \geq 1$ and $x_1, \dots, x_R \in D$ [19].

- (3) If $\mathcal{A}(x)$ is a vector-valued random function and the conditions in the first item are satisfied for the components $\{\mathcal{A}_i(x)\}$ of this function, the FD translation models $\{\tilde{A}_i(x)\}$ are asymptotically Gaussian as the KL truncation levels m_i increase indefinitely. Under additional rather mild conditions, it can be shown that the vector-valued function $\tilde{A}(x)$ become Gaussian as $m_i \rightarrow \infty$, $i = 1, \dots, d$.

These properties pose difficulties on the selection of the KL truncation level for T models. For real-valued random functions $\mathcal{A}(x)$, the m.s. error of the translation model $\tilde{A}(x)$ decreases with the KL truncation level m but its distribution approaches the Gaussian distribution as m increases. This means $\tilde{A}(x)$ and $\mathcal{A}(x)$ have similar first two moments but different distributions for sufficiently large truncation levels m . In addition, the weak tail dependence between the components of $\tilde{Z}$ can result in unsatisfactory approximations of quantities of interest involving extremes of $\mathcal{A}(x)$.

We note that the models $\tilde{A}(x)$ in this section are not the standard translation models $\tilde{A}_0(x)$. For real-valued stationary functions $\mathcal{A}(x)$, these models are defined by

$$\tilde{A}_0(x) = F_0^{-1} \circ \Phi(G(x)), \quad x \in D,$$

where $G(x)$ is a zero-mean, unit-variance, stationary Gaussian function with correlation function $\rho(\cdot)$ and F_0 denotes the marginal distribution of $\mathcal{A}(x)$. The correlation function of $G(x)$ is selected to minimize differences between the correlation functions of $\tilde{A}_0(x)$ and $\mathcal{A}(x)$ [17] (Chap. 3). The random functions $\tilde{A}_0(x)$ and $\mathcal{A}(x)$ are stationary and have the same and similar marginal distributions and correlation functions. A notable limitation of $\tilde{A}_0(x)$ if used to model extremes of $\mathcal{A}(x)$ is the tail independence, i.e., $P(\tilde{A}_0(x) > u \mid \tilde{A}_0(y) > u) \rightarrow 0$, $x \neq y$, as $u \rightarrow \infty$.

4. PC models

We outline the construction and essential properties of PC models for random vectors and derive properties of the resulting FD models of $\mathcal{A}(x)$. Most of the content of this section is available in the literature, an introduction on PC models can be found in [13].

4.1. Random coefficients

The PC model Z_n of $Z \in \mathbb{R}^m$ for truncation level n has the form

$$Z_n = \sum_{k_1, \dots, k_m \geq 0, \sum_{r=1}^m k_r = 1, \dots, n} a_{k_1 \dots k_m} \prod_{r=1}^m H_{k_r}(N_r), \quad (13)$$

where $\{a_{k_1 \dots k_m}\}$ are $\mathbb{R}^m$ -valued coefficients, $\{H_{k_r}\}$ are Hermite, Legendre, or other polynomials, and $N = (N_1, \dots, N_m) \in \mathbb{R}^m$ has independent components with standard Gauss, uniform, or other distributions, respectively [38]. The PC model has the following asymptotic properties.

- (1) $Z_{n,i} \xrightarrow{L_2} Z_i$, $i = 1, \dots, m$, which implies the convergence in probability and distribution by properties of PC models for real-valued random variables [24].
- (2) $Z_n \xrightarrow{L_2} Z$ by the L_2 convergence of the components of Z_n to those of Z .
- (3) $Z_n \xrightarrow{d} Z$ since $E[\|Z_n - Z\|^2] \leq E[\|Z_n - Z\|^2] \rightarrow 0$ by the Cauchy–Schwarz inequality so that $P(\|Z_n - Z\| > \varepsilon) \leq E[\|Z_n - Z\|]/\varepsilon$ by the Chebyshev inequality. This gives the convergence $Z_n \xrightarrow{p} Z$, $n \rightarrow \infty$, which implies the convergence in distribution of Z_n to Z .
- (4) If the marginal distributions $\{F_i\}$ of Z are continuous and have no flat spots, the marginal distributions $\{F_{n,i}\}$ of Z_n converge uniformly to $\{F_i\}$, i.e., $\sup_{x \in \mathbb{R}} |F_{n,i}(x) - F_i(x)| \rightarrow 0$, $n \rightarrow \infty$. This holds by properties of monotone functions. We have $\sup_{x \in [a,b]} |F_{n,i}(x) - F_i(x)| \rightarrow 0$, $n \rightarrow \infty$, for arbitrary bounded intervals $[a, b]$ [29] (Proposition 2.1) and the observation that a and b can be taken sufficiently small and large, respectively, such that the differences $|F_{n,i}(x) - F_i(x)|$ can be made as small as desired for $x < a$ and $x > b$.

The above asymptotic properties are of limited value in many applications since (i) the truncation level n has to be kept relatively low to limit the computation demand, (ii) the rate of convergence of Z_n to Z is slow, and (iii) the discrepancy between Z and Z_n may not decrease monotonically with the truncation level n . These observations suggest that the distributions of Z and Z_n are likely to differ significantly for large dimensional vectors Z since the truncation levels n must be low to assure that the implementation of Z_n is feasible. A notable feature of the PC models Z_n , $n > 1$, is that their components are dependent as functions of the random vector $N = (N_1, \dots, N_m)$ (see Eq. (13)).

The implementation of the PC models Z_n requires to calculate the coefficients $\{a_{k_1 \dots k_m}\}$ in Eq. (13). The number of these coefficients increases rapidly with the dimension m of Z and the truncation level n . The coefficients $\{a_{k_1 \dots k_m}\}$ can be obtained efficiently by projection if Z is a known function of the random vector N in the definition of Z_n . Otherwise, they

can be obtained by matching moments and/or other statistics of Z and Z_n via optimization algorithms, e.g., MATLAB genetic algorithms. The algorithms can be supplemented with constraint involving the first two moments of Z and Z_n . We have $E[Z_n] = 0$ irrespective of the values of $\{a_{k_1 \dots k_m}\}$ since $E[\prod_{r=1}^m H_{k_r}(N_r)] = \prod_{r=1}^m E[H_{k_r}(N_r)] = 0$. The covariances of Z_n with Hermite polynomials $\{H_{k_r}\}$ have the expressions

$$\begin{aligned} E[Z_{n,p} Z_{n,q}] &= \sum_{k_1, \dots, k_m} \sum_{l_1, \dots, l_m} a_{k_1 \dots k_m}^{(p)} a_{l_1 \dots l_m}^{(q)} E\left[\prod_{r=1}^m H_{k_r}(N_r) \prod_{s=1}^m H_{l_s}(N_s)\right] \\ &= \sum_{k_1, \dots, k_m} \sum_{l_1, \dots, l_m} \prod_{i=1}^m a_{k_1 \dots k_m}^{(p)} a_{l_1 \dots l_m}^{(q)} E[H_{k_i}(N_i) H_{l_i}(N_i)] \\ &= \sum_{k_1, \dots, k_m} \sum_{l_1, \dots, l_m} \prod_{i=1}^m a_{k_1 \dots k_m}^{(p)} a_{l_1 \dots l_m}^{(q)} \prod_{i=1}^m k_i!, \quad p, q = 1, \dots, m, \end{aligned}$$

by properties of Hermite polynomials, i.e., $E[H_\alpha(X) H_\beta(Y)] = \alpha! \rho^\alpha \delta_{\alpha\beta}$ for X, Y correlated standard Gaussian variables with correlation coefficient ρ . The indices of summations in the above expression of $E[Z_{n,p} Z_{n,q}]$ have the range of those in Eq. (13) and $a_{k_1 \dots k_m}^{(p)}$ denotes the component $p = 1, \dots, m$ of $a_{k_1 \dots k_m}$. The possible values of the coefficients $\{a_{k_1 \dots k_m}\}$ can be constrained by limiting differences between the covariances of Z and the covariances of Z_n , which depend on the coefficients $\{a_{k_1 \dots k_m}\}$. The constraints define a subset of $\mathbb{R}^m$ which contains the optimal coefficients so that the search is performed in a subset of $\mathbb{R}^m$ rather than the entire $\mathbb{R}^m$. We will specify the objective functions and the constraints used to implement Z_n in the numerical examples of Sect. 7 on Applications.

4.2. FD models for random functions

The PC model $A_n(x)$ of $A(x)$ is the random function $A(x)$ in Eq. (10) with Z replaced by Z_n in Eq. (13), i.e.,

$$A_n(x) = \varphi(x) Z_n, \quad x \in D. \quad (14)$$

The properties of $A_n(x)$ follow directly from those of Z_n and the functional form of $A(x)$. For example, $A_n(x)$ has the following properties for large truncation levels n .

- (1) For arbitrary fixed $x \in D$, the sequence of random variables $A_n(x)$ converges in m.s. to $A(x)$ as $n \rightarrow \infty$, and, therefore, in probability and distribution, since

$$E[(A(x) - A_n(x))^2] = \varphi(x) E[(Z - Z_n)(Z - Z_n)'] \varphi(x)',$$

and

$$|E[(Z_i - Z_{n,i})(Z_j - Z_{n,j})]| \leq (E[(Z_i - Z_{n,i})^2] E[(Z_j - Z_{n,j})^2])^{1/2}$$

by the Cauchy-Schwarz inequality, and the m.s. convergence of Z_n to Z as $n \rightarrow \infty$.

- (2) For arbitrary $\bar{n}$ and $x_1, \dots, x_{\bar{n}} \in D$, the vector with components $(A_n(x_1), \dots, A_n(x_{\bar{n}}))$ converges in m.s. to $(A(x_1), \dots, A(x_{\bar{n}}))$ as $n \rightarrow \infty$ by the previous property since these vectors are linear in Z . This implies the convergence in probability and distributions of these vectors, so that the finite dimensional distributions of $A_n(x)$ converge to those of $A(x)$.

As previously stated, the attractive asymptotic properties of Z_n and $A_n(x)$ are of limited value in many applications since, generally, the rate of convergence of Z_n to Z is slow, the computational effort for using large truncation levels is excessive when dealing with large dimensional vectors, and the accuracy of Z_n and, therefore $A_n(x)$, does not increase monotonically with the truncation level n .

5. PCT models

We have seen that the translation models $\tilde{Z}$ match the marginal distributions of Z but cannot capture its dependence structure and that the PC models Z_n have dependent components but their distributions differ from those of Z for finite truncation levels n . The PCT models are intended to retain the desirable features of T and PC models and overcome most of their limitations. The models have the functional form in Eq. (10) with Z replaced by PCT models $\tilde{Z}_n$ of Z .

5.1. Random coefficients

The components of the *PCT models* for random vectors Z are defined by

$$\tilde{Z}_{n,i} = F_i^{-1} \circ F_{n,i}(Z_{n,i}) = h_{n,i}(Z_{n,i}), \quad i = 1, \dots, d, \quad n = 0, 1, \dots, \quad (15)$$

where $\{F_i^{-1}\}$ are the inverses of the marginal distributions $\{F_i\}$, $Z_n = \{Z_{n,i}\}$ denote the PC models in Eq. (13), and $\{F_{n,i}\}$ are the marginal distributions of $\{Z_n\}$. The distributions $\{F_{n,i}\}$ depend on the coefficients $\{a_{k_1 \dots k_m}\}$ which, generally, differ for the PC and PCT models of Z . The PCT models have the following properties.

- (1) $\tilde{Z}_{n,i} \xrightarrow{d} Z_i$, $i = 1, \dots, m$, $n = 1, 2, \dots$, by construction since $F_{n,i}(Z_{n,i}) \sim U(0, 1)$ irrespective of the truncation level n and the values of the coefficients $\{a_{k_1 \dots k_m}\}$.
- (2) $h_{n,i}$ converges to the identity function uniformly for each $i = 1, \dots, m$.
This follows from (i) the uniform continuity of F_i^{-1} as a continuous function on a bounded interval, i.e., for given $\varepsilon > 0$, there exists $\delta > 0$ such that $|u - v| < \delta$ implies $|F_i^{-1}(u) - F_i^{-1}(v)| < \varepsilon$ and (ii) the uniform convergence $F_{n,i} \rightarrow F_i$, i.e., for any $\delta > 0$ there is an index n_δ such that $|F_{n,i}(x) - F_i(x)| < \delta$ for all $x \in \mathbb{R}$ and $n \geq n_\delta$. In summary, for an arbitrary $\varepsilon > 0$, we can find $\delta > 0$ and n_δ such that $|F_i^{-1}(F_{n,i}(x)) - F_i^{-1}(F_i(x))| = |h_{n,i}(x) - x| < \varepsilon$ for $x \in \mathbb{R}$ and $n \geq n_\delta$.
- (3) $\tilde{Z}_{n,i} \xrightarrow{L_2} Z_i$, $i = 1, \dots, m$, as $n \rightarrow \infty$ which implies the convergence in probability and distribution of $\tilde{Z}_{n,i}$ to Z_i .
We have $(Z_i - \tilde{Z}_{n,i})^2 \leq 2(Z_i - Z_{n,i})^2 + 2(Z_{n,i} - \tilde{Z}_{n,i})^2$, $E[(Z_i - \tilde{Z}_{n,i})^2] \rightarrow 0$ by the m.s. convergence of $Z_{n,i}$ to Z_i , and $E[(Z_{n,i} - \tilde{Z}_{n,i})^2] = \int (x - h_{n,i}(x))^2 dF_{n,i}(x)$ which converges to zero by the previous item.
- (4) $\tilde{Z}_n \xrightarrow{L_2} Z$ as $n \rightarrow \infty$ by the convergence $\tilde{Z}_{n,i} \xrightarrow{L_2} Z_i$, $i = 1, \dots, m$. This implies the convergence $Z_n \rightarrow Z$ in probability and distribution so that the joint distributions of $\tilde{Z}_n$ converge to the joint distribution F of Z .

These properties show that the PCT model $\{\tilde{Z}_n\}$ in Eq. (15) preserves the desirable properties of T and PC models, $\tilde{Z}$ and Z_n . Like $\tilde{Z}$, it matches the marginal distributions of Z exactly for any n and $\{a_{k_1 \dots k_m}\}$. In contrast to $\tilde{Z}$ whose components have independent tails, the components of $\tilde{Z}_n$, $n > 1$, are dependent as functions of the same random vector $N = (N_1, \dots, N_d)$. The latter features is inherited from PC models.

We reiterate that the coefficients $\{a_{k_1 \dots k_m}\}$ of the PC models Z_n in Eq. (13) and those of the PCT model $\tilde{Z}_n$ in Eq. (15), generally, differ. These coefficients are obtained by optimization algorithms if the mapping $N \mapsto Z$ is not known. The objective functions and the constraints used to implement Z_n and $\tilde{Z}_n$ will be specified in the numerical examples of Sect. 7 on Applications.

5.2. FD models for random functions

The random function $A(x)$ in Eq. (10) with $\tilde{Z}_n$ given by Eq. (15) in place of Z yields the associated *PCT model* $\tilde{A}_n(x)$ of $A(x)$, i.e.,

$$\tilde{A}_n(x) = \varphi(x) \tilde{Z}_n, \quad x \in D. \quad (16)$$

The properties of $\tilde{A}_n(x)$ result from those of $\tilde{Z}_n$ and the above expression of $\tilde{A}_n(x)$. They parallel the properties of the PC model $A_n(x)$.

- (1) For arbitrary fixed $x \in D$, the sequence of random variables $\tilde{A}_n(x)$ converges in m.s. to $A(x)$ as $n \rightarrow \infty$, and, therefore, in probability and distribution since

$$E[(A(x) - \tilde{A}_n(x))^2] = \varphi(x) E[(Z - \tilde{Z}_n)(Z - \tilde{Z}_n)'] \varphi(x)',$$

and $\tilde{Z}_n$ converged in the m.s. to Z as $n \rightarrow \infty$.

- (2) For arbitrary $\bar{n}$ and $x_1, \dots, x_{\bar{n}} \in D$, the vector with components $(\tilde{A}_n(x_1), \dots, \tilde{A}_n(x_{\bar{n}}))$ converges in m.s. to $(A(x_1), \dots, A(x_{\bar{n}}))$ as $n \rightarrow \infty$ by the previous property. This implies the convergence of in probability and distribution so that the finite dimensional distributions of $\tilde{A}_n(x)$ converge to those of $A(x)$.

6. Relations between T, PC, and PCT models

The models proposed for $A(x)$ have the form of $A(x)$ in Eq. (10), i.e.,

$$\begin{aligned} A_n(x) &= \varphi(x) Z_n, \\ \tilde{A}(x) &= \varphi(x) \tilde{Z}, \quad \text{and} \\ \tilde{A}_n(x) &= \varphi(x) \tilde{Z}_n, \end{aligned} \quad (17)$$

where Z_n , $\tilde{Z}$, and $\tilde{Z}_n$ are the random vectors in Eqs. (13), (11), and (15). The samples of the models in Eq. (17) are elements of the vector space spanned by the deterministic basis functions $\varphi(x)$, which are common to all models. The differences between these models relate to differences between the random coefficients Z_n , $\tilde{Z}$, and $\tilde{Z}_n$. We now assess these differences and their implications on the models of $A(x)$ for given truncation level $n < \infty$ of the PC series representation of Z .

6.1. T and PCT models

The random vectors $\tilde{Z}_n$ and $\tilde{Z}$ match exactly the marginal distributions of Z by construction. For $\tilde{Z}_n$, this holds irrespective of the truncation level n and the values of the coefficients $\{a_{k_1 \dots k_m}\}$ of the underlying PC model of Z .

The components of $\tilde{Z}_n$ are dependent/independent if the components of Z are dependent/independent. The accuracy of the joint distribution of $\tilde{Z}_n$ depends on the truncation level n and the values of $\{a_{k_1 \dots k_m}\}$. In contrast, $\tilde{Z}$ cannot capture the dependence structure of Z . For example, if $A(x)$ is a real-valued random function, the components of $\tilde{Z}$ are independent since the components of Z are uncorrelated.

The properties of $\tilde{Z}$ imply that $\tilde{A}(x) = \varphi(x) \tilde{Z}$ is approximately Gaussian so that its law may differ significantly from that of $A(x)$. Approximations of, e.g., extremes $\max_{x \in D} |A(x)|$ of $A(x)$ based on $\tilde{A}(x)$ are likely to be unsatisfactory. On the other hand, essential features of the dependence structure of Z are captured by $\tilde{Z}_n$ provided that the coefficients $\{a_{k_1 \dots k_m}\}$ minimize the discrepancy between joint statistics of Z and $\tilde{Z}_n$.

6.2. PC and PCT models

There are notable similarities between PC and PCT models and some differences which may affect their performance significantly. The properties of $\tilde{Z}_n$ and Z_n are transferred to the FD models $\tilde{A}_n(x)$ and $A_n(x)$ of $A(x)$.

The similarities of $\tilde{Z}_n$ and Z_n result from their definitions in Eqs. (13) and (15) which show that the PCT models are constructed component-by-component by translating the components of PC models. Following are similarities between these models.

- For given truncation level, PC and PCT models depend on the same number of parameters, the coefficients $\{a_{k_1, \dots, k_m}\}$.
- Optimization algorithms are used to select the coefficients $\{a_{k_1, \dots, k_m}\}$. The algorithms have similar complexity and they require similar computation time to find the values of these coefficients.
- The random vectors $\tilde{Z}_n$ and Z_n have similar asymptotic properties. They converge to Z in L^2 , probability, and distribution as the truncation level n increases indefinitely.
- The dependence structures of Z_n and $\tilde{Z}_n$ are closely related in the sense that the probabilities that two or more components of these vectors are simultaneously large are similar, provided they have the same truncation level n and coefficients $\{a_{k_1 \dots k_m}\}$. For example, for a two-dimensional vector Z with identically distributed components, this means that the probabilities $P(Z_{n,1} > z, Z_{n,2} > z)$ and $P(\tilde{Z}_{n,1} > \tilde{z}, \tilde{Z}_{n,2} > \tilde{z})$ are similar, where z and $\tilde{z}$ are large thresholds at the scales of the distributions of $Z_{n,1}$ and $\tilde{Z}_{n,1}$. Generally, $P(Z_{n,1} > z, Z_{n,2} > z)$ is an unsatisfactory approximation of $P(Z_1 > z, Z_2 > z)$ since the distributions of Z_n and Z differ for relatively low truncation levels.

The differences between $\tilde{Z}_n$ and Z_n relate to the construction of PCT models which retain the essential property of T models. For example,

- The coefficients $\{a_{k_1, \dots, k_m}\}$ define completely the law of Z_n . The law of $\tilde{Z}_n$ is defined by these coefficients and the marginal distributions of Z .
- PCT models capture exactly the marginal distributions of Z irrespective of the truncation level n and the values of the coefficients $\{a_{k_1, \dots, k_m}\}$. In contrast, the marginal distributions of Z_n and Z differ for $n < \infty$ unless, e.g., the components of Z are polynomials of $N = (N_1, \dots, N_m)$ in Eq. (13) and n is larger than the degrees of these polynomials.
- The optimal coefficients $\{a_{k_1, \dots, k_m}\}$ of $\tilde{Z}_n$ and Z_n differ. Generally, the objective functions for PC models quantify differences between moments, distributions, and other properties of Z_n and Z . If the objective functions include only differences between statistics of some quantities of interest, e.g., $\max_{x \in D} \{A(x)\}$ for real-valued random functions, the properties of resulting random vectors Z_n may differ significantly from those of Z . PCT models do not exhibit this behavior since they match the marginal distributions of Z exactly. The coefficients $\{a_{k_1 \dots k_m}\}$ can be tuned to best capture the dependence structure of Z and/or various quantities of interest.

7. Applications

The construction of T, PC, and PCT based FD models for random functions involves optimization algorithms which require various information on the target random vector Z , e.g., marginal distributions and correlations for T models and joint moments and/or histograms for PC and PCT models.

The optimization algorithm to construct T models identifies the correlation coefficients $\{\rho_{ij}\}$ in the Gaussian space which minimize the discrepancy between the target correlations $\{\xi_{ij}\}$ and the correlations $\{\tilde{\xi}_{ij}\}$ induced by $\{\rho_{ij}\}$ (see Eq. (11)). We

measure the discrepancy between these correlations by the norm of $\{\rho_{ij} - \xi_{ij}\}$. The selection of the optimal $\{\rho_{ij}\}$ is efficient since, generally, $\rho_{ij} \simeq \xi_{ij}$ [17] (Chap. 3).

The optimization algorithms to construct the PC and PCT models, Z_n and $\tilde{Z}_n$, of Z identify the coefficients $\{a_{k_1 \dots k_m}\}$ of the expansions in Eq. (13) and (15). The algorithms are computationally more intensive than for T models since (1) the number of coefficients $\{a_{k_1 \dots k_m}\}$ increases rapidly with the dimension of Z and the truncation level n and (2) the objective functions involve model statistics which need to be recalculated for each proposed values of the coefficients $\{a_{k_1 \dots k_m}\}$. For example, let $\mu(r_1, \dots, r_m) = E[\prod_{j=1}^m Z_j^{r_j}]$ and $\mu_n(r_1, \dots, r_m; \{a_{k_1 \dots k_m}\}) = E[\prod_{j=1}^m Z_{n,j}^{r_j} | \{a_{k_1 \dots k_m}\}]$ be moments of order $(r_1, \dots, r_m)$ of Z and Z_n or $\tilde{Z}_n$ and $\{h_{ij}\}$ and $\{h_{n,ij}(\{a_{k_1 \dots k_m}\})\}$, $i, j = 1, \dots, m$, $i \leq j$, be histograms of pairs of target and model components of Z and Z_n or $\tilde{Z}_n$. We select the coefficients $\{a_{k_1 \dots k_m}\}$ to minimize objective functions which quantify differences between target and model moments and/or two-dimensional histograms, e.g., the metrics

$$\alpha \|\mu(r_1, \dots, r_m) - \mu_n(r_1, \dots, r_m; \{a_{k_1 \dots k_m}\})\| + \beta \|\{h_{ij}\} - \{h_{n,ij}(\{a_{k_1 \dots k_m}\})\}\|, \quad (18)$$

where $\alpha, \beta > 0$ are weights. The moments $\{\mu_n(r_1, \dots, r_m; \{a_{k_1 \dots k_m}\})\}$ and the histograms $\{h_{n,ij}(\{a_{k_1 \dots k_m}\})\}$ have to be recalculated for every new proposed values of the coefficients $\{a_{k_1 \dots k_m}\}$ which slows computation. The computation times to calculate these statistics of PC and PCT models are similar.

The following subsections implement T, PC, and PCT models for a two-dimensional non-Gaussian vector and two types of random functions and assess their performance and relative accuracy.

7.1. Two-dimensional non-Gaussian vector

Let $Z = (Z_1, Z_2)$ be a random vector with the joint density

$$f(z_1, z_2) = 6z_2 \mathbf{1}(z_1 \geq 0, z_2 \geq 0, z_1 + z_2 \leq 1), \quad z = (z_1, z_2) \in \mathbb{R}^2, \quad (19)$$

and marginal distributions

$$\begin{aligned} F_1(z_1) &= 1 - (1 - z_1)^3, \quad z_1 \in [0, 1], \\ F_2(z_2) &= 3z_2^2 - 2z_2^3, \quad z_2 \in [0, 1]. \end{aligned} \quad (20)$$

It can be shown that

$$\begin{aligned} Z_1 &\stackrel{d}{=} 1 - U_1^{1/3} \\ Z_2 &\stackrel{d}{=} U_1^{1/3} U_2^{1/2} \end{aligned} \quad (21)$$

where U_1 and U_2 are independent $U(0, 1)$ variables [33] (Sect. 3.5.1).

Suppose that the quantity of interest is the joint distribution of Z . We construct approximations of these quantities based on T, PC, and PCT models, $\tilde{Z}$, Z_n , and $\tilde{Z}_n$, of Z . Since the target vector Z is given as a function of the $U(0, 1)$ random variables $\{U_i\}$, Legendre rather Hermite polynomials are used to construct PC models.

T models: The vector $\tilde{Z}$ is the image of a standard Gaussian vector $G = (G_1, G_2)$ with correlated components. Simple calculations show that $\rho_{12} = -0.5900$ yields the correlation coefficient -0.5674 of $\tilde{Z}$, which is very close to the correlation coefficient -0.5713 of Z . The model $\tilde{Z}$ is completely defined by the correlation coefficients ρ_{12} and the marginal distributions in Eq. (20). The components of these models are

$$\begin{aligned} \tilde{Z}_1 &= F_1^{-1} \circ \Phi(G_1) \\ \tilde{Z}_2 &= F_2^{-1} \circ \Phi(G_2). \end{aligned} \quad (22)$$

PC and PCT models: The general form of the PC models of Z is

$$Z_n = \sum_{k_1, k_2 \geq 0; \leq k_1 + k_2 \leq n} a_{k_1 k_2} L_{k_1}(U_1) L_{k_2}(U_2) \quad (23)$$

where $\{L_{k_i}\}$ are Legendre polynomials and $\{U_i\}$, $i = 1, 2$, denote independent $U(0, 1)$ variables. The coefficients $\{a_{k_1 k_2} \in \mathbb{R}^2\}$ can be obtained by projecting Z on the Legendre basis $\{L_{k_1}(U_1) L_{k_2}(U_2)\}$ and using properties of the Legendre polynomials (see Eq. (9)), and have the expressions

$$a_{k_1 k_2} = \frac{1}{(2k_1 + 1)(2k_2 + 1)} E[Z L_{k_1}(U_1) L_{k_2}(U_2)]. \quad (24)$$

The expectations in their expressions can be calculated or estimated from samples of (U_1, U_2, Z) since Z is a known function of (U_1, U_2) .

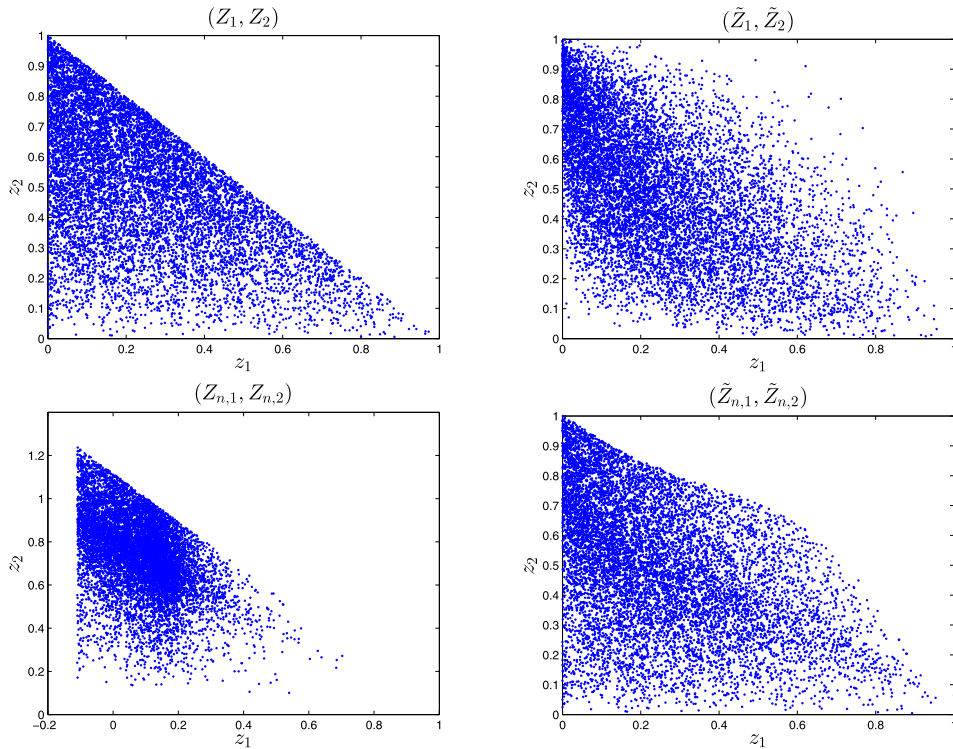


Fig. 1. Scatter plots of Z (top left panel), $\tilde{Z}$ (top right panel), Z_n (bottom left panel), and $\tilde{Z}_n$ (bottom right panel).

The components of the PCT models $\tilde{Z}_n$ result from Eq. (22) with Z_n in Eq. (23) in place of $G = (G_1, G_2)$ and the distributions of Z_n in place of Φ . The coefficients $\{a_{k_1 k_2}\}$ in the expression of Z_n can be selected to minimize specified objective functions of the type in Eq. (18). Generally, the coefficients $\{a_{k_1 k_2}\}$ of PC and PCT coefficients differ.

Two approaches have been considered to implement PC and PCT models. The first constructs the mapping $U \mapsto Z_n$ in Eq. (23) by using Eq. (24). The second uses the structure of this problem to reduce calculations, i.e., the fact Z_1 depends only on U_1 and Z_2 is a product of the functions $U_1^{1/3}$ and $U_2^{1/2}$ of U_1 and U_2 . We construct two independent PC models, one for $U_1^{1/3}$ and one for $U_2^{1/2}$ and use them to define a PC model of $Z = (Z_1, Z_2)$. The PC models of $U_1^{1/3}$ and $U_2^{1/2}$ have the form

$$Y_{n_i, i} = \sum_{r=0}^{n_i} \alpha_{i,r} L_r(U_i), \quad i = 1, 2, \quad (25)$$

where $\{n_i\}$ denote truncation levels. The corresponding PC models of the components of Z are $Z_{n_1,1} = 1 - \sum_{r=0}^{n_1} \alpha_{1,r} L_r(U_1)$ and $Z_{n_1, n_2, 2} = (\sum_{r=0}^{n_1} \alpha_{1,r} L_r(U_1)) (\sum_{r=0}^{n_2} \alpha_{2,r} L_r(U_2))$. The coefficients of these PC models can be obtained by projections since the components of Z are known functions of $U = (U_1, U_2)$. We continue to call Z_n the resulting PC model $(Z_{n_1,1}, Z_{n_1, n_2, 2})$ by abuse of notations.

The following results are for $n_1 = n_2 = 4$. The scatter plots in Fig. 1 have been obtained from 10,000 independent samples of Z (top left panel), $\tilde{Z}$ (top right panel), Z_n (bottom left panel), and $\tilde{Z}_n$ (bottom right panel). The PCT model $\tilde{Z}_n$ in this figure has been obtained from the PC model Z_n by translation so that it has the same coefficients $\{a_{k_1 k_2}\}$ as Z_n . The scatter plot of $\tilde{Z}$ and of the target random vector Z differs significantly although these two random vectors have the same marginal distributions. As expected, the scatter plot of $\tilde{Z}$ resembles that of negatively correlated Gaussian variables. The scatter plot of the PC model Z_n is qualitatively similar to the scatter plot of Z . However, the range of values of Z_n and their probabilities differ from those of Z significantly. The scatter plot of $\tilde{Z}_n$ (bottom right panel) is rather similar to that of Z although its coefficients $\{a_{k_1 k_2}\}$ are those of the PC model. This PCT model is just a translation of the PC model.

The left panel of Fig. 2 is the scatter plots of Z in the top left panel of Fig. 1. The right panel is the scatter of the PCT model $\tilde{Z}_n$ in Eq. (7) whose coefficients have been selected to minimize the discrepancy between its two-dimensional histogram and the histogram of Z , which can be quantified by the objective function in Eq. (18) with $\alpha = 0$ and $\beta = 1$. The support and the distribution of samples of Z and $\tilde{Z}_n$ nearly coincide. The PCT model $\tilde{Z}_n$ with its optimal coefficients provides an accurate representation of Z and is a significant improvement over the T and PC models.

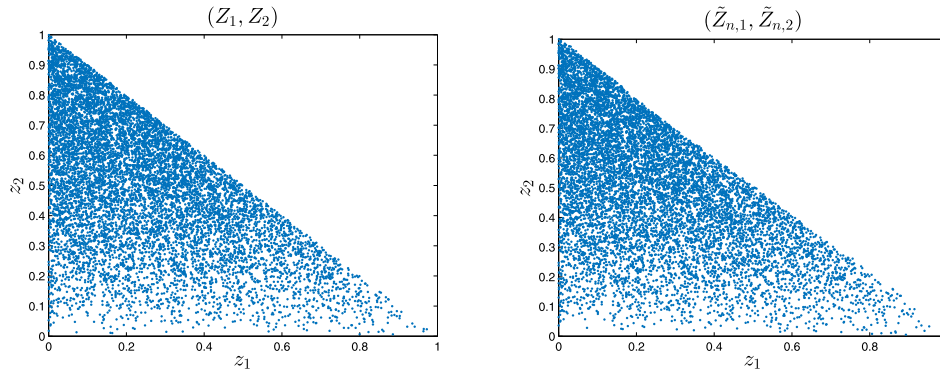


Fig. 2. Scatter plots of Z and the PCT model $\tilde{Z}_n$ (left and right panels).

7.2. Non-Gaussian random field

Consider the real-valued FD random field

$$A(x) = (G_1 \cos(\nu x) + G_2 \sin(\nu x))^r, \quad x \in [0, l], \quad (26)$$

where $r > 1$ and $\nu > 0$ are a specified integer and real and (G_1, G_2) denote independent $N(0, 1)$ variables. The random field $A(x)$ is non-Gaussian as a power of the Gaussian field $G_1 \cos(\nu x) + G_2 \sin(\nu x)$. It has the expression

$$A(x) = \sum_{r_1, r_2 \geq 0; r_1 + r_2 = r} \frac{r!}{r_1! r_2!} G_1^{r_1} G_2^{r_2} \cos^{r_1}(\nu x) \sin^{r_2}(\nu x), \quad x \in [0, l], \quad (27)$$

which, for $r = 4$, takes the form

$$A(x) = \sum_{k=1}^5 Z_k \varphi_k(x), \quad x \in [0, l], \quad (28)$$

where the components of the 5-dimensional random vector Z are $Z_1 = G_1^4$, $Z_2 = 4 G_1^3 G_2$, $Z_3 = 6 G_1^2 G_2^2$, $Z_4 = 4 G_1 G_2^3$, and $Z_5 = G_2^4$ and the deterministic functions $\{\varphi_k(x)\}$ have the expressions $\varphi_1(x) = \cos^4(\nu x)$, $\varphi_2(x) = \cos^3(\nu x) \sin(\nu x)$, $\varphi_3(x) = \cos^2(\nu x) \sin^2(\nu x)$, $\varphi_4(x) = \cos(\nu x) \sin^3(\nu x)$, and $\varphi_5(x) = \sin^4(\nu x)$.

The quantity of interest is the probability $P(\max_{0 \leq x \leq l} \{A(x)\} > a)$. Approximations of this probability are constructed by replacing Z in Eq. (28) with T, PC, and PCT models, $\tilde{Z}$, Z_n , and $\tilde{Z}_n$, of this random vector. We note that the PC models of Z are exact for truncation levels $n \geq 4$ and approximate for $n < 4$ and that the components of Z are correlated. The following numerical results are for $r = 4$, $n = 3$, $\nu = 2\pi/l$, and $l = 1$.

T model: Samples of Z can be obtained by elementary calculations from samples of the standard Gaussian vector (G_1, G_2) with independent components and the mapping $G = (G_1, G_2) \mapsto Z = (Z_1, \dots, Z_5)$. Estimates of the correlation coefficients $\{\zeta_{ij}\}$ of the components (Z_i, Z_j) , $i, j = 1, \dots, 5$, of Z are $\zeta_{12} = -0.0595$, $\zeta_{13} = 0.4317$, $\zeta_{14} = -0.0177$, $\zeta_{15} = 0.0018$, $\zeta_{23} = -0.0566$, $\zeta_{24} = 0.6043$, $\zeta_{25} = -0.0076$, $\zeta_{34} = -0.0293$, $\zeta_{35} = 0.4233$, and $\zeta_{45} = -0.0046$. The Gaussian image of Z is a standard 5-dimensional Gaussian vector whose correlations $\{\rho_{ij}\}$ are very close to the target correlations $\{\zeta_{ij}\}$. We have set $\{\rho_{ij} = \zeta_{ij}\}$. These correlations and estimates of the marginal distributions of Z define completely the T model of Z (see Eq. (11)).

PC and PCT models: The PC model of Z has the form

$$Z_n = E[Z] + \sum_{k_1, k_2 \geq 0; 1 \leq k_1 + k_2 \leq n} a_{k_1 k_2} H_{k_1}(G_1) H_{k_2}(G_2) \quad (29)$$

for truncation level n (see Eq. (13)). The coefficients $\{a_{k_1 k_2} \in \mathbb{R}^5\}$ of these representations can be calculated from

$$a_{k_1 k_2} = \frac{1}{k_1! k_2!} E[(Z - E[Z]) H_{k_1}(G_1) H_{k_2}(G_2)], \quad (30)$$

which results by projecting Z on the Hermite basis $\{H_{k_1}(G_1) H_{k_2}(G_2)\}$ (see Eq. (9)). The expectations in Eq. (30) can be calculated directly by using properties of Gaussian variables since the components of Z are polynomials of G . We estimated these expectations from independent samples $(G(\omega), Z(\omega))$ of (G, Z) . The PC model of $A(x)$ results by replacing Z in Eq. (28) with Z_n in Eq. (29) with coefficients $\{a_{k_1 k_2}\}$ obtained from Eq. (30).

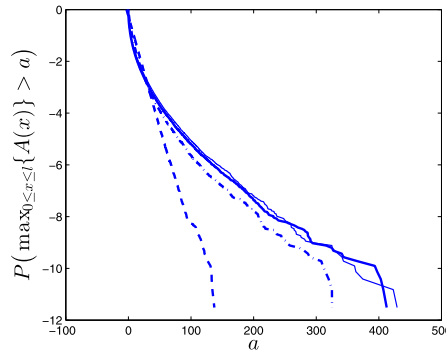


Fig. 3. Estimates of $P(\max_{0 \leq x \leq l} \{A(x)\} > a)$: target (thin solid line) and T, PC, and PCT models (heavy dash-dotted, dotted, and solid lines).

The PCT models $\tilde{Z}_n$ of Z are translations of PC models Z_n of the type in Eq. (29) such that the marginal distributions of the resulting vectors match the marginal distributions of Z irrespective of the values of the coefficients $\{a_{k_1 k_2}\}$ and the truncation level n (see Eq. (15)). The coefficients $\{a_{k_1 k_2}\}$ are selected to minimize specified objective functions which may or may not depend on quantities of interest. Similar optimization algorithms, rather than projection as commonly used in applications, can be used to calculate the coefficients $\{a_{k_1 k_2}\}$ of PC models. We use the discrepancy between all pairs of 2D histograms of $\tilde{Z}_n$ of Z to find the optimal PCT coefficients $\{a_{k_1 k_2}\}$, i.e., the objective function in Eq. (18) with $\alpha = 0$ and $\beta = 1$.

The probability $P(\max_{0 \leq x \leq l} \{A(x)\} > a)$ is illustrated by the thin solid line in Fig. 3. It is based on 100,000 independent samples of Z and the expression of $A(x)$ in Eq. (28). The heavy dash-dotted, dotted, and solid lines in the figure are estimates of this probability obtained from 100,000 independent samples of the T, PC, and PCT models, $\tilde{A}(x)$, $A_n(x)$, and $\tilde{A}_n(x)$, of $A(x)$ which are obtained from Eq. (28) with $\tilde{Z}$, Z_n , and $\tilde{Z}_n$ in place of Z . The PC based estimate of $P(\max_{0 \leq x \leq l} \{A(x)\} > a)$ is inaccurate at moderate and large thresholds. The estimate of $P(\max_{0 \leq x \leq l} \{A(x)\} > a)$ improves significantly if based on the T model $\tilde{Z}$ of Z . If Z is approximated by a PCT model $\tilde{Z}_n$, the resulting approximation of $P(\max_{0 \leq x \leq l} \{A(x)\} > a)$ traces closely this distribution. This is rather impressive since the PCT model has been constructed by translating the PC model Z_n to have the target marginal distributions, i.e., the coefficients $\{a_{k_1 k_2}\}$ of $\tilde{Z}_n$ are those of Z_n .

There is a close relationship between the T and PCT models. The models are images of Gaussian and non-Gaussian random vectors via the inverses of the marginal distributions $\{F_i^{-1}\}$ of Z . Moreover, Z_n is Gaussian for $n = 1$. If the dimension of the Gaussian vector G in the definition of Z_n is larger than or equal to that of Z , the T model is a special case of the PCT model $\tilde{Z}_n$, $n > 1$, so that $\tilde{Z}_n$ is at least as accurate as $\tilde{Z}$. The relative performance of T and PCT models is difficult to quantify in general since it is problem-dependent.

7.3. Gaussian & non-Gaussian diffusion processes

Let $\mathcal{X}(t)$ be a real-valued process defined by the stochastic differential equation

$$d\mathcal{X}(t) = -\lambda \mathcal{X}(t) dt + b(\mathcal{X}(t)) dB(t), \quad t \geq 0, \quad (31)$$

with drift and diffusion coefficients $-\lambda \mathcal{X}(t)$, $\lambda > 0$, and $b(\mathcal{X}(t))$, where $dB(t)$ denotes the increment of the standard Brownian motion process $B(t)$, $t \geq 0$, during an infinitesimal time interval dt .

The stationary solution of Eq. (31) exists [18] (Sect. 4.7). Its mean and correlation functions are $E[\mathcal{X}(t)] = 0$ and $r(s, t) = E[\mathcal{X}(s)\mathcal{X}(t)] = \sigma^2 \exp(-\lambda|s-t|)$ irrespective of the form of the diffusion coefficient, where $\sigma^2 = r(t, t)$ denotes the variance of $\mathcal{X}(t)$. The diffusion coefficient $b(\mathcal{X}(t))$ can be chosen such that the marginal distribution of the stationary state $\mathcal{X}(t)$ has a specified form [6,16]. For example, the stationary marginal distribution $f(x)$ of $\mathcal{X}(t)$ is the standard Gaussian and shifted Gamma with shape, scale, and shift $k > 0$, $\theta > 0$, and $-k\theta$, i.e.,

$$f(x) = \phi(x) = \exp(-x^2/2)/\sqrt{2\pi} \quad \text{and} \\ f(x) = \frac{(x+k\theta)^{k-1} \exp(-(x/\theta+k))}{\theta^k \Gamma(k)}, \quad x > -k\theta, \quad (32)$$

for the diffusion coefficients $b(x)^2 = 2\lambda$ and $b(x)^2 = 2\lambda\theta^2(x/\theta+k)$, $k, \theta > 0$.

The following numerical results are for $\lambda = 0.3$, a time interval $[0, \tau]$, $\tau = 5$, $k = 2$, and $\theta = 3$ and stationary Gaussian and shifted Gamma diffusion processes $\mathcal{X}(t)$ in Eq. (31). The processes $\mathcal{X}(t)$ are scaled by their standard deviations so that they have unit variances and exponential correlations with decay parameter λ . For simplicity, the stationary solutions of Eq. (31) continued to be denoted by $\mathcal{X}(t)$.

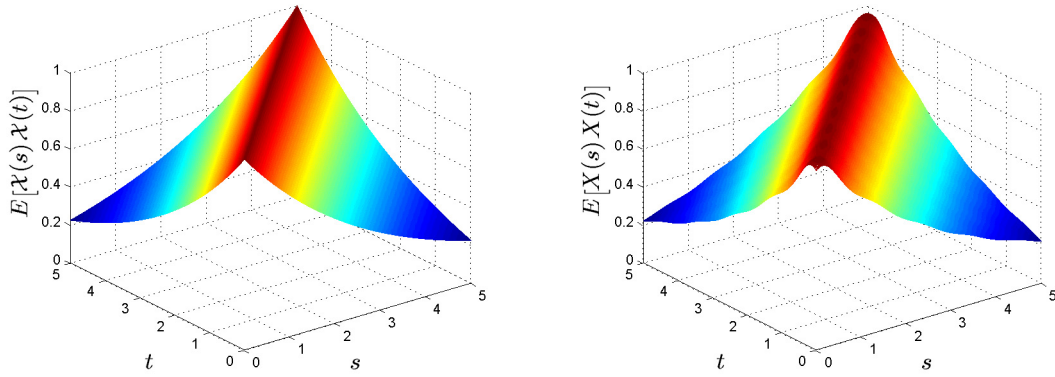


Fig. 4. Covariance functions of $\mathcal{X}(t)$ and $X(t)$ (left and right panels).

We construct FD models for these diffusion processes by using the methodology in Sect. 2, in which $\mathcal{A}(x)$ is replaced with $\mathcal{X}(t)$. First, the top eigenfunctions of the correlation functions of $\mathcal{X}(t)$ are used to define the functional form of the FD models $X(t)$ of $\mathcal{X}(t)$ (see Eq. (2)). Second, T, PC, and PCT models, $\tilde{Z}$, Z_n , and $\tilde{Z}_n$, are developed for the random coefficients $Z = (Z_1, \dots, Z_m)$ in the expression of $X(t)$, where n is the truncation level of the PC model. Third, T, PC, and PCT models, $\tilde{X}(t)$, $X_n(t)$, and $\tilde{X}_n(t)$, are constructed from $\tilde{Z}$, Z_n , and $\tilde{Z}_n$ and the functional form of $X(t)$.

7.3.1. Functional form of $X(t)$

Since the Gaussian and Gamma diffusion processes $\mathcal{X}(t)$ have zero means and scaled correlation functions $r(s, t) = \exp(-\lambda|s - t|)$, their KL series representations have the same form. The truncated versions of the common KL series of $\mathcal{X}(t)$ is

$$X(t) = \sum_{k=1}^m Z_k \varphi_k(t), \quad t \in [0, \tau], \quad (33)$$

where the basis functions $\{\varphi_k\}$ are the top m eigenfunctions of the exponential correlation function $r(s, t)$ (see Eqs. (2) and (10)), i.e., the eigenfunctions associated with the m largest eigenvalues of this correlation function.

The FD model $X(t)$ with $m = 10$ is approximately equal to $\mathcal{X}(t)$ in the second moment sense. Its mean and standard deviation are $E[X(t)] \simeq 0$ and $\text{Std}[X(t)] = 0.99$. The correlation functions of $\mathcal{X}(t)$ and $X(t)$ are similar, see the left and right panels of Fig. 4. These observations suggest that the truncation level $m = 10$ is sufficiently large to capture accurately the first two moments of the target process. The following subsection compares sample properties of $\mathcal{X}(t)$, $X(t)$, and Z for this truncation level.

7.3.2. Samples of $\mathcal{X}(t)$, $X(t)$, and Z

The Euler method was used to generate samples of $\mathcal{X}(t)$, i.e., the iteration formula

$$\mathcal{X}(t + \Delta t) \simeq (1 - \lambda \Delta t) \mathcal{X}(t) + b(\mathcal{X}(t)) \Delta B(t), \quad (34)$$

where $\Delta t > 0$ is the integration time step and $\Delta B(t) = B(t + \Delta t) - B(t)$ denotes the increment of the Brownian motion in the time interval $(t, t + \Delta t)$. Since $\mathcal{X}(t + \Delta t) | \mathcal{X}(t)$ is a Gaussian variable with mean $(1 - \lambda \Delta t) \mathcal{X}(t)$ and variance $b(\mathcal{X}(t))^2 \Delta t$, then $\mathcal{X}(t + \Delta t)$ takes values outside the support of $f(x)$ for the Gamma diffusion processes with non-zero probability. To mitigate the occurrence of these events, the integration time step Δt is reduced when $\mathcal{X}(t)$ enters a small vicinity of the ends of the support of $f(x)$ and/or the process is restarted at an interior point of the support of $f(x)$ close to its boundary.

Samples of the FD model $X(t)$ are constructed from samples of Z and its expression in Eq. (33). The left and right panels of Fig. 5 show with dashed and solid lines five samples of the Gauss and Gamma diffusion process $\mathcal{X}(t)$ and the corresponding samples of $X(t)$ for $m = 10$. The samples of $X(t)$ capture the essential features of the samples of $\mathcal{X}(t)$ but miss high frequency fluctuations of these samples since they are not included in the basis functions $\{\varphi_k\}$, $k = 1, \dots, m$. The sample of $X(t)$ can capture additional high frequency details of the samples of $\mathcal{X}(t)$ for higher truncation levels [21].

The top and bottom panels of Fig. 6 show target densities $f(x)$ and marginal histograms of $\mathcal{X}(t)$ and $X(t)$ for the Gauss and Gamma diffusion processes (left and right panels). The histograms are based on 10,000 independent samples and match the marginal densities of the Gaussian and Gamma processes $\mathcal{X}(t)$. The plots of the bottom panels show that $X(t)$ and $\mathcal{X}(t)$ have similar marginal distributions. We also note that the estimates of the marginal skewness and kurtosis coefficients of $\mathcal{X}(t)$ and $X(t)$ are similar. They are approximately 0 and 3 for Gauss and 1.43 and 6.23 for Gamma diffusion processes. Since $X(t)$ and $\mathcal{X}(t)$ have similar first two moments, the samples of $X(t)$ capture essential features of the samples of $\mathcal{X}(t)$,

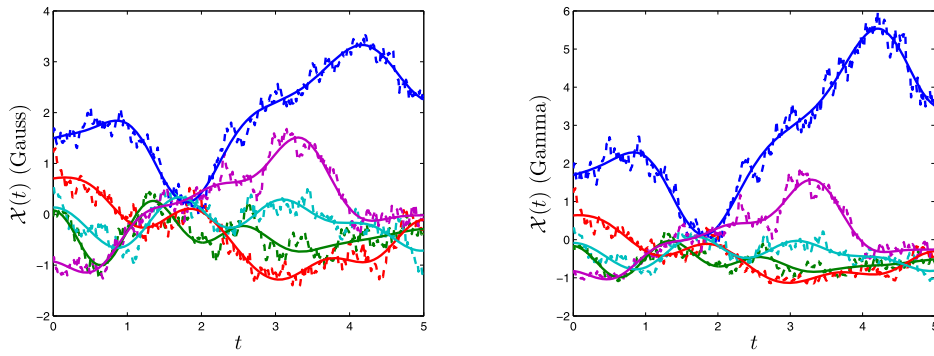


Fig. 5. Five samples of the Gauss and Gamma diffusion process $\mathcal{X}(t)$ and the corresponding samples of $X(t)$ (left and right panels) for $m = 10$.

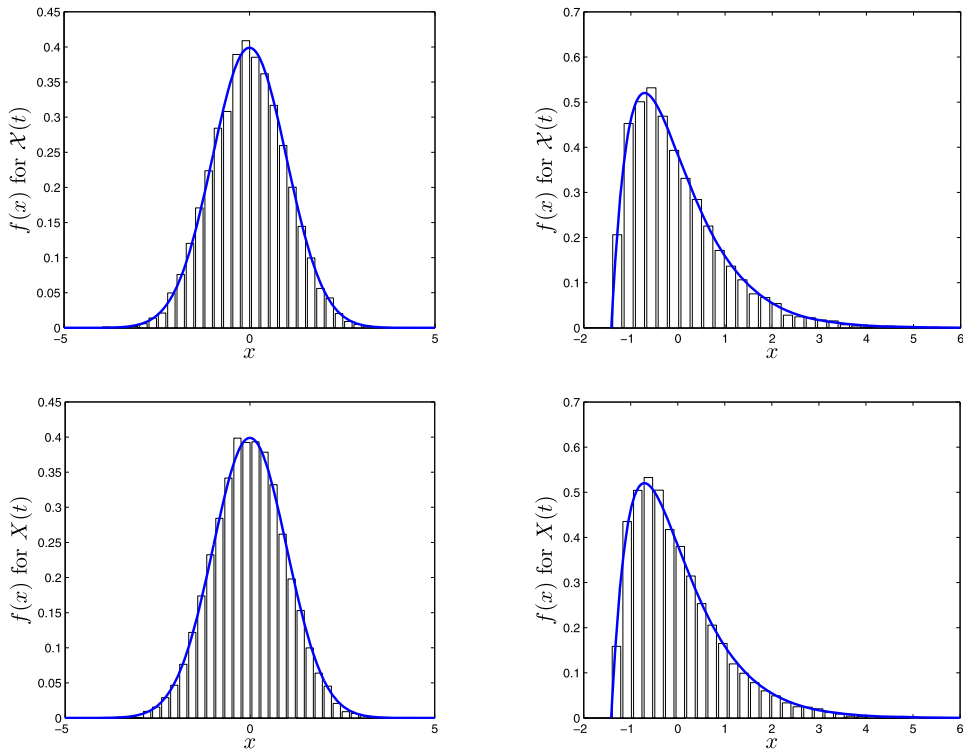


Fig. 6. Stationary marginal distributions and histograms of $\mathcal{X}(t)$ and $X(t)$, $m = 10$, (top and bottom panels) for Gauss and Gamma diffusion processes (left and right panels).

and the marginal histograms of $X(t)$ match the target densities $f(x)$, it is concluded that the truncation level $m = 10$ is adequate. The FD process $X(t)$ with $m = 10$ is taken as target for our subsequent developments.

Samples of the m -dimensional random vector Z result from samples of $\mathcal{X}(t)$ by projection via Eq. (9). The components of Z are uncorrelated for both processes $\mathcal{X}(t)$ (see Sect. 2.3). If $\mathcal{X}(t)$ Gaussian, the components $\{Z_k\}$ of Z are independent and Gaussian so that the law of Z is completely defined by the means and variances of its components.

If $\mathcal{X}(t)$ is the Gamma diffusion process, the components of Z are dependent and non-Gaussian so that the joint distribution of this vector is needed to characterize the FD model $X(t)$. Properties of samples of Z can and should be used to simplify the implementation of its PC and PCT models, a comment similar to that in the first example. The left and right panels of Fig. 7 show scatter plots of the components (Z_1, Z_2) and (Z_1, Z_4) of Z obtained from 10,000 samples of this vector. They show that these random variables are strongly dependent. Similar scatter plots are for all pairs $\{Z_1, Z_k\}$, $k \geq 2$, of components of Z . The left and right panels of Fig. 8 show scatter plots of (Z_2, Z_4) obtained from 10,000 samples of Z and from 10,000 samples of Z_2 and Z_4 assumed to be independent variable following the marginal distributions of the second and fourth components of Z . Since the scatter plots are similar, we conclude that the dependence between these component of Z is weak. Similar scatter plots have been obtained for the other pair (Z_k, Z_l) , $k, l \geq 2$, $k \neq l$ of Z .

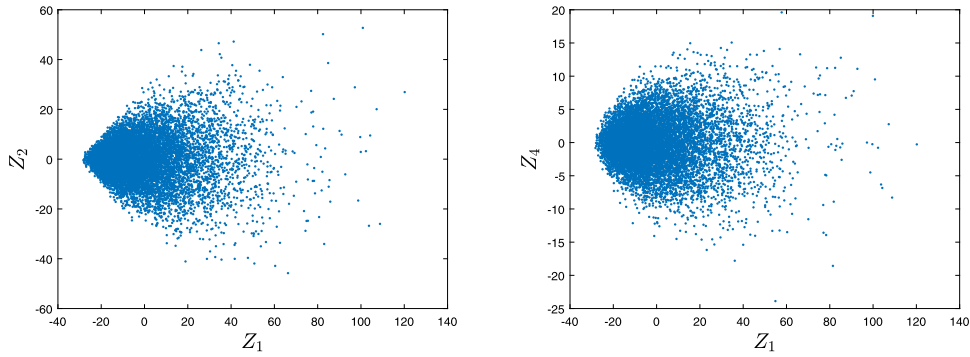


Fig. 7. Scatter plots of (Z_1, Z_2) and (Z_1, Z_4) (left and right panels).

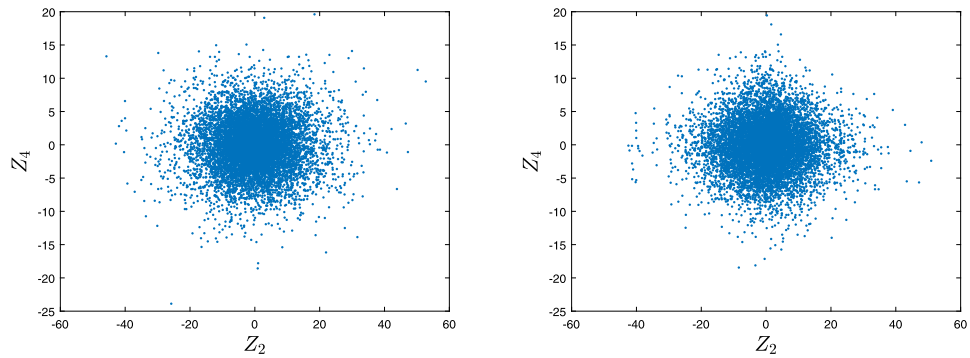


Fig. 8. Scatter plots of (Z_2, Z_4) and of (Z_2, Z_4) under the independence assumption (left and right panels).

7.3.3. T, PC, and PCT models of Z

Samples of Z are used to construct T, PC, and PCT models, $\tilde{Z}$, Z_n , and $\tilde{Z}_n$, for this random vector. These models are defined by Eqs. (13), (11), and (15).

T model: The implementation of the model $\tilde{Z}$ of Z in Eq. (11) requires to specify the marginal distributions $\{F_i\}$ of Z and the correlation coefficients $\{\rho_{ij}\}$ of the Gaussian image of this vector. Estimates of $\{F_i\}$ are obtained from samples of Z . The correlations coefficients $\{\rho_{ij}\}$, $i \neq j$, are zero since and the components of Z are uncorrelated. Samples of the FD model $X(t)$ result from its definition in Eq. (33) and samples of $\tilde{Z}$ which can be generated by standard algorithms for random variables since $\tilde{Z}$ has independent components.

PC and PCT models: The general formulation in Eqs. (13) and (15) can be applied to find the optimal coefficients $\{a_{k1\dots km} \in \mathbb{R}^m\}$ of the PC and PCT models, Z_n and $\tilde{Z}_n$, of Z , i.e., the coefficients which minimize objective functions quantifying discrepancies between target and model statistics (see the objective function in Eq. (18)).

We consider a simplified version of this general formulation which uses the properties of Z illustrated in Figs. 7 and 8 to reduce calculations. First, PC and PCT models, $Z_{n,1}$ and $\tilde{Z}_{n,1}$ are constructed for Z_1 . The PC model has the form

$$Z_{n,1} = \sum_{k=1}^{n_1} a_k H_k(G_1), \quad (35)$$

where the coefficients $\{a_k\}$ can be obtained by minimizing the discrepancy between moments of $Z_{n,1}$ and Z_1 , which is measured by objective functions of the type in Eq. (18) with $(\alpha, \beta) = (1, 0)$. The PCT model $\tilde{Z}_{n,1}$ has the same distribution as Z_1 irrespective of the truncation level n_1 and the values of the coefficients $\{a_k\}$. Second, to capture the dependences between the pairs (Z_1, Z_k) , $k \geq 2$, we construct PC and PCT models for the conditional random variables $\{Z_k | Z_1\}$, $k = 2, \dots, m$, an idea which was proposed previously [7]. The PC model of the conditional random variables $Z_k | Z_1$ has the form

$$Z_k | Z_1 \simeq Z_{m_k,k} = E[Z_k | Z_1] + \sum_{s=1}^{m_k} \beta_{k,s} H_s(G_k), \quad k = 2, \dots, m, \quad (36)$$

where m_k denotes the truncation level. For $m_k = 1$, $Z_{m_k,k}$ is a Gaussian variable with mean $\mu_k(Z_1) = E[Z_k | Z_1]$ and standard deviation $\sigma_k(Z_1) = \text{Std}[Z_k | Z_1]$.

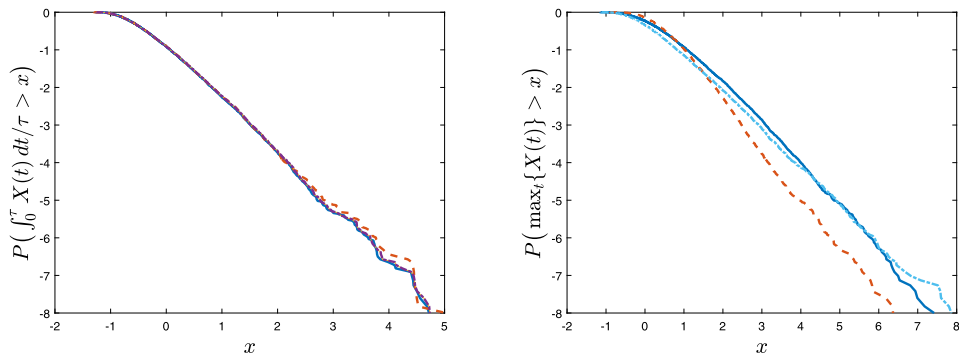


Fig. 9. Distributions of $\int_0^\tau X(t) dt/\tau$ and $\max_{0 \leq t \leq \tau} \{X(t)\}$ (left and right panels); the solid lines are target distributions and the dashed, and dash-dotted are T based and PCT based approximations.

The following numerical results are for truncation levels $m_k = 1$ so that the PC models $\{Z_{m_k,k}\}$ are the Gaussian variables $N(\mu_k(Z_1), \sigma_k(Z_1)^2)$. We have $\tilde{Z}_{n,1} \stackrel{d}{=} Z_1$ irrespective of the truncation level n of the underlying PC model $Z_{n,1}$ of Z_1 . The generation of samples of this model of Z is sequential. First, generate a sample of Z_1 . Second, generate the corresponding samples of the conditional PC models $\{Z_{m_k,k}\}$ corresponding to this sample of Z_1 . Third, use the resulting sample of Z to calculate the corresponding sample of the FD model $X(t)$ in Eq. (33).

Two quantities of interest are considered, the temporal average $\int_0^\tau X(t) dt/\tau$ and the extreme $\max_{0 \leq t \leq \tau} \{X(t)\}$. The left and right panels in Fig. 9 show the distributions of $\int_0^\tau X(t) dt/\tau$ and $\max_{0 \leq t \leq \tau} \{X(t)\}$. The solid lines in the figure are the target probabilities $P(\int_0^\tau X(t) dt/\tau > x)$ and $P(\max_{0 \leq t \leq \tau} \{X(t)\} > x)$. The dashed and dash-dotted lines are approximations of these distributions based on T and PCT models. Both models approximate the distribution of $\int_0^\tau X(t) dt/\tau$ satisfactorily. Since this quantity of interest is the sum $\sum_{k=1}^m Z_k \int_0^\tau \varphi_k(t) dt/\tau$ of the random variables $\{Z_k\}$ weighted by the deterministic coefficients $\{\int_0^\tau \varphi_k(t) dt/\tau\}$, detail features of the joint distributions of $\tilde{Z}$ and $\tilde{Z}_n$ have a minor influence.

This is not the case with the quantity of interest $\max_{0 \leq t \leq \tau} \{X(t)\}$. The T based approximation underestimates the right tail of the distribution of this quantity of interest. This unsatisfactory performance is likely to be caused by the fact marginal distribution of the T-based FD model $X(t)$ is approximately Gaussian as a sum of independent random variables, the components of $\tilde{Z}$ weighted by the deterministic functions $\{\varphi_k(t)\}$ (see Eq. (33)). The PCT model approximates accurately the distribution of $\max_{0 \leq t \leq \tau} \{X(t)\}$ although it is based on a simplified representation of Z . In contrast to the T model which ignores the dependence between the components of Z , the PCT model accounts explicitly for this dependence via the underlying PC model.

The PC model nearly coincides with the PCT model in this example since the PC model $Z_{n,1}$ in Eq. (35) with $n = 4$ matches closely the distribution of Z_1 . The coefficients $\{a_k\}$ of this representation have been obtained by minimizing discrepancies between moments of Z_n and Z which can be quantified by the objective function in Eq. (18) with $(\alpha, \beta) = (1, 0)$. The PC-based approximations of $P(\int_0^\tau X(t) dt/\tau > x)$ and $P(\max_{0 \leq t \leq \tau} \{X(t)\} > x)$ are indistinguishable from those of PCT models at the scale of Fig. 9.

Results are not shown for the Gaussian diffusion process $\mathcal{X}(t)$ since the components of Z are independent Gaussian variables so that $\tilde{Z} \stackrel{d}{=} Z_n \stackrel{d}{=} \tilde{Z}_n \stackrel{d}{=} Z$. Accordingly, the distributions of the quantities of interest $\int_0^\tau X(t) dt/\tau$ and $\max_{0 \leq t \leq \tau} \{X(t)\}$ based on the T, PC, and PCT models coincide and match exactly the target distributions.

8. Conclusions

The random entries of stochastic equations, i.e., equations with random coefficients, source terms, and/or end conditions, are random functions, i.e., uncountable families of random variables indexed by space and/or time arguments. Numerical solutions of these equations require to approximate these random functions by finite dimensional (FD) models, i.e., deterministic functions of space/time arguments which depend on random vectors Z .

Three models, referred to as T, PC, and PCT models, $\tilde{Z}$, Z_n and $\tilde{Z}_n$, have been developed for Z . The implementation of T models involves the least calculations. The models match exactly the marginal distributions of Z but cannot capture the dependence between its components. The implementation of PC and PCT models is computationally more demanding. Optimization algorithms in spaces which, generally, have relatively large dimensions are used to find the unspecified parameters of these models. The PC models Z_n capture the dependence structure of Z but provide less satisfactory for the joint distributions of this vector. The PCT models $\tilde{Z}_n$ inherit the desired features of T and PC models. Like T models, they match exactly the marginal distributions of Z and, like PC models, capture essential features of the dependence between the components of this vector.

Following the construction of FD models in Sect. 2, the presentation of properties of T, PC, and PCT models (Sects. 3, 4 and 5) and the discussion on the relationships between these models (Sect. 6), three numerical examples have been presented to illustrate the implementation and the relative performance of T, PC, and PCT models. The examples include a two-dimensional vector Z whose components are known functions of two independent $U(0, 1)$ variables, a real-valued non-Gaussian random field, and Gaussian and Gamma diffusion processes. The T model $\tilde{Z}$ was found to be accurate for quantities of interest which are controlled by the marginal distributions of Z rather than its dependence structure. The PC model Z_n was capable to capture the essential features of the dependence structure of Z but delivered unsatisfactory approximations of the distributions of Z in some cases.

The PCT model $\tilde{Z}_n$ approximated accurately all quantities of interest considered in this study. Their satisfactory performance is attributed to the fact that the coefficients $\{a_{k_1 \dots k_m}\}$ of the underlying PC model Z_n are only used to minimize discrepancies between the dependences of $\tilde{Z}_n$ and Z since $\tilde{Z}_n$ matches exactly the marginal distributions of Z irrespective of the truncation level n and the values of $\{a_{k_1 \dots k_m}\}$. In contrast, the optimal coefficients $\{a_{k_1 \dots k_m}\}$ of PC models are required to minimize differences between all statistics of Z_n and Z .

Acknowledgements

The work reported in this paper has been partially supported by the National Science Foundation under grants CMMI-1265511 and CMMI-1639669. This support is gratefully acknowledged.

References

- [1] R.J. Adler, *The Geometry of Random Fields*, John Wiley & Sons, New York, 1981.
- [2] I.M. Babuška, K-M. Liu, On solving stochastic initial-value differential equations, *Math. Models Methods Appl. Sci.* 13 (5) (2003) 715–745.
- [3] I.M. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, *SIAM J. Numer. Anal.* 45 (3) (2007) 1005–1034.
- [4] I.M. Babuška, R. Tempone, E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, *SIAM J. Numer. Anal.* 42 (2) (2004) 800–825.
- [5] I.M. Babuška, R. Tempone, E. Zouraris, Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation, *Comput. Methods Appl. Mech. Eng.* 194 (12–14) (2005) 1251–1294.
- [6] G.Q. Cai, Y.K. Lin, Generation of non-Gaussian stationary stochastic processes, *Phys. Rev. E* 54 (1) (1995) 299–303.
- [7] S. Das, R. Ghanem, S. Finette, Polynomial chaos representation of spatio-temporal random fields from experimental measurements, *J. Comput. Phys.* 228 (2009) 8726–8751, <https://doi.org/10.1016/j.jcp.2009.08.025>.
- [8] W.B. Davenport, W.L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill Book Company, Inc., New York, 1958.
- [9] H.C. Elman, O.G. Ernst, D.P. O’Leary, M. Stewart, Efficient iterative algorithms for stochastic finite element method with applications to acoustic scattering, *Comput. Methods Appl. Mech. Eng.* 194 (2005) 1037–1055.
- [10] R.V. Field Jr., M. Grigoriu, On the accuracy of the polynomial chaos approximation, *Probab. Eng. Mech.* 19 (1–2) (2004) 65–80.
- [11] J. Foo, X. Wan, E. Karniadakis, The multi-element probabilistic collocation method (ME-PCM): error analysis and applications, *J. Comput. Phys.* 227 (2008) 9572–9595.
- [12] P. Frauenfelder, C. Schwab, R.A. Todor, Finite element for elliptic problems with stochastic coefficients, *Comput. Methods Appl. Mech. Eng.* 194 (2005) 205–228.
- [13] R.G. Ghanem, P.D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.
- [14] I. Gohberg, S. Goldberg, *Basic Operator Theory*, Birkhäuser, Boston, 1980.
- [15] M. Grigoriu, Monte Carlo algorithm for vector-valued Gaussian functions with preset component accuracies, *Monte Carlo Methods Appl.* 23 (3) (2017) 165–188.
- [16] M. Grigoriu, Translation models revisited, *Probab. Eng. Mech.* 48 (2017) 68–75.
- [17] M. Grigoriu, *Applied Non-Gaussian Processes: Examples, Theory, Simulation, Linear Random Vibration, and MATLAB Solutions*, Prentice Hall, Englewood Cliffs, NJ, 1995.
- [18] M. Grigoriu, *Stochastic Calculus. Applications in Science and Engineering*, Birkhäuser, Boston, 2002.
- [19] M. Grigoriu, Probabilistic models for stochastic elliptic partial differential equations, *J. Comput. Phys.* 229 (2010) 8406–8429, <https://doi.org/10.1016/j.jcp.2010.07.023>.
- [20] M. Grigoriu, A method for solving stochastic equations by reduced order models and local approximations, *J. Comput. Phys.* 231 (19) (2012) 6495–6513.
- [21] M. Grigoriu, Models for space–time random functions, *Probab. Eng. Mech.* 43 (2016) 5–16.
- [22] H.H. Cho, D. Venturi, G.E. Karniadakis, Karhunen–Loève expansion for multi-correlated stochastic processes, *Probab. Eng. Mech.* 34 (2013) 157–167.
- [23] D.B. Hernández, *Lectures on Probability and Second Order Random Fields*, World Scientific, London, 1995.
- [24] Hui-Hsiung Kuo, *Introduction to Stochastic Integration*, Springer, New York, 2006.
- [25] M.R. Leadbetter, G. Lindgren, H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York, 1983.
- [26] X. Ma, A.F. Vakakis, L.A. Bregman, Karhunen–Loève models of a truss: transient response reconstruction and experimental verification, *AIAA J.* 39 (4) (2001) 687–696.
- [27] Y.M. Marzouk, H.N. Najm, L.A. Rahn, Stochastic spectral methods for efficient Bayesian solution of inverse problems, *J. Comput. Phys.* 224 (2007) 560–586.
- [28] G. Perrin, C. Soize, D. Duhamel, C. Funfschilling, Karhunen–Loève expansion revisited for vector-valued random fields: scaling, errors and optimal basis, *J. Comput. Phys.* (2013) 607–622.
- [29] S. Resnick, *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer, New York, 2007.
- [30] S.I. Resnick, *Extreme Values, Regular Variation, and Point Processes*, Springer Series in Operations Research and Financial Engineering, Springer, New York, 1987.
- [31] S.I. Resnick, *A Probability Path*, Birkhäuser, Boston, 1998.
- [32] M. Rosenblatt, Remarks on a multivariate transformation, *Ann. Math. Stat.* 23 (3) (1952) 470–472.
- [33] R. Rubinstein, *Simulation and the Monte Carlo Method*, John Wiley & Sons, New York, NY, 1981.
- [34] C. Schwab, R.A. Todor, Karhunen–Loève approximation of random fields by generalized fast multipole methods, *J. Comput. Phys.* 217 (2006) 100–122.

- [35] B. Sudret, A. Der Kiureghian, Comparisons of finite element reliability methods, *Probab. Eng. Mech.* 17 (2002) 337–348.
- [36] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos methods for stochastic differential equations, *J. Comput. Phys.* 209 (2) (2005) 617–642, <https://doi.org/10.1016/j.jcp.2005.03.023> C822.
- [37] H. Xie, T. Zhou, A multilevel finite element method for Fredholm integral eigenvalue problems, *J. Comput. Phys.* 303 (2015) 173–184.
- [38] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.* 24 (2) (2002) 619–644, <https://doi.org/10.1137/S1064827501387826>.