



**HAL**  
open science

# Large scale random fields generation using localized Karhunen–Loève expansion

Alfonso Panunzio, Régis Cottureau, Guillaume Puel

► **To cite this version:**

Alfonso Panunzio, Régis Cottureau, Guillaume Puel. Large scale random fields generation using localized Karhunen–Loève expansion. *Advanced Modeling and Simulation in Engineering Sciences*, 2018, 5 (20), 10.1186/s40323-018-0114-7. hal-01861348

**HAL Id: hal-01861348**

**<https://hal.science/hal-01861348>**

Submitted on 13 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE

Open Access



# Large scale random fields generation using localized Karhunen–Loève expansion

Alfonso M. Panunzio<sup>1,2\*</sup>, Régis Cottereau<sup>2</sup> and Guillaume Puel<sup>2</sup>

\*Correspondence:

alfonso-  
michele.panunzio@ratp.fr  
<sup>1</sup>RATP, 54 Quai de la Rapée,  
75012 Paris, France  
Full list of author information is  
available at the end of the article

## Abstract

In this paper the generation of random fields when the domain is much larger than the characteristic correlation length is made using an adaptation of the Karhunen–Loève expansion (KLE). The KLE requires the computation of the eigen-functions and the eigen-values of the covariance operator for its modal representation. This step can be very expensive if the domain is much larger than the correlation length. To deal with this issue, the domain is split in sub-domains where this modal decomposition can be comfortably computed. The random coefficients of the KLE are conditioned in order to guarantee the continuity of the field and a proper representation of the covariance function on the whole domain. This technique can also be parallelized and applied for non-stationary random fields. Some numerical studies, with different correlation functions and lengths, are presented.

**Keywords:** Random fields generation, Karhunen–Loève expansion, Large scale random fields

## Introduction

The representation of fluctuating parameters by means of random fields is very common in many scientific domains. Samples of stationary random fields can be generated through a sum of harmonic functions with random uniform phase and amplitude depending on the spectral density [1, 2]. This kind of representation can be performed in the spectral domain [3–5], leading to the spectral representation method that can be efficiently computed using the FFT [6]. In case of multi-dimensional random fields, the spectral representation has been combined with the turning bands methods [7] for a more efficient computation [8]. In a huge domain the numerical cost can be a major issue. In [9], to deal with this problem, the domain is split in several small sub-domains in which the samples of the random fields are generated. Then, samples on the whole domain are obtained by using an overlapping technique.

Auto-regressive models, in which a state only depends linearly on its own previous values, can also be employed to represent random fields [10–13]. The linear dependency coefficients can be computed by maximizing a likelihood function or by solving a linear system involving the inversion of a matrix representing the discretized covariance.

Other methods use the direct decomposition of the covariance to simulate generation of the field. The Cholesky decomposition of the discretized covariance can be used to correlate a set of random variables representing the discretization of the random field

© The Author(s) 2018. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

[14]. In some methods, when the random generation is required on a large size domain, polynomial approximations of the square root of the covariance are computed [15, 16]. The covariance decomposition can be combined with a ARMA representation to improve the computational efficiency [17].

Another way for generating random fields is the Karhunen–Loève expansion (KLE, [18, 19]), which is based on the covariance kernel modal decomposition on a finite domain [20, 21]. The KLE has been extensively used for representing random fluctuating properties in different engineering problems [22–24]. While the spectral representation is optimal in mean-square sense on an infinitely large domain, the KLE is optimal on a finite domain. Moreover, one of the main advantages is that this expansion can also be directly applied for non-stationary processes. Different numerical methods exist for solving the covariance decomposition for the KLE [25, 26]. In the case of stationary covariances, the modal functions can be approximated by means of Fourier transforms [27–29] to reduce the computational complexity, but the application of the KLE in a very large domain compared to the correlation length still remains unaffordable. For a given mean-square error, The number of needed terms in the KLE grows as shown in [30], with the size of the domain. In some applications, for avoiding the KLE decomposition, known families of polynomials are used to parametrize the random field [31, 32], but they do not minimize the mean-square error as the KLE does.

The aim of this paper is to generate samples of a random field using the KLE, when the size of the domain is much larger than the correlation length and a direct KLE is not affordable because of the computational effort. The technique presented in [33], that deals with the representation of cross-correlated random fields using the KLE, is here adapted to overcome this issue. At first the whole domain is split in small sub-domains (with a size of few correlation lengths). The modal decomposition of the covariance operator is computed in a small sub-domain where the computational effort is easily affordable and a reduced number of terms are needed for the KLE. Then, independent random samples are generated in each sub-domain and, finally, the assembling is made by conditioning the KLE coefficients to obtain continuous samples of the random fields having the prescribed covariance function. In [33] the authors model a set of correlated random fields by imposing a correlation between the KLE coefficients of each random field. In this work, the same idea is used for correlating sets of KLE coefficients related to local regions of a large domain.

In this paper Gaussian random fields with different correlation structure are considered. Non-Gaussian random fields can be obtained by using the Rosenblatt transform [34] that allows to modify a Gaussian random field according to a chosen marginal first order probability density function (memoryless transformation). This transform also changes the correlation structure although, in most of the cases, one can deal with this issue by modifying the original correlation function as done in [35–37], where stationary fields are transformed into non-stationary fields. In other methods, non-Gaussian fields can be obtained through transformations with memory [38]. These aspects are not discussed in this paper, where only Gaussian fields are considered.

The proposed method is firstly presented for the case of 1-dimensional (1D) random processes (“Karhunen–Loève expansion for large scale 1D random processes” section) with an example of the application of the method. Some considerations about the continuity of the generated samples are discussed in “Continuity of the generated samples”

section. The method is then generalized to 2D and 3D random fields (“Generation of multi-dimensional random fields” section). Then, the generation of non-stationary fields is discussed in “Extension to non-stationary random fields” section. In this work only random fields with values in  $\mathbb{R}$  are considered. Some numerical applications are provided (“Numerical applications” section).

### **Karhunen–Loève expansion for large scale 1D random processes**

In this section the Karhunen–Loève expansion is adapted to generate samples of a large scale 1D stationary random process. In “Standard 1D Karhunen–Loève expansion” section, the generalities of the standard KLE, applied on a domain of size equal to  $L$ , are presented.

For simplicity, the method is firstly illustrated for a domain composed of just two sub-domains (“Principles of the generation method on a large domain” section). The continuity of the generated samples is investigated in “Continuity of the generated samples” section. Then the extension to a domain composed of an arbitrary number of sub-domains is discussed in “Extension of the expansion on an arbitrary large domain” section.

#### **Standard 1D Karhunen–Loève expansion**

Let  $(\Theta, \mathcal{F}, P)$  be a probability space and  $f(s, \theta)$  with  $\theta \in \Theta$  a centred stationary random field, indexed by the variable  $s$ , whose covariance function  $C(|s - t|)$  is equal to:

$$C(|s - t|) = \mathbb{E}[f(s, \theta)f(t, \theta)] \quad (1)$$

For simplicity, the case of a stationary random field is treated in this section. An extension to non-stationary fields is given in “Extension to non-stationary random fields” section.

For the application of the Karhunen–Loève expansion [18], the first step is an eigenvalue decomposition of the covariance operator:

$$\int_0^L C(|s - t|)\varphi_i(s)ds = \lambda_i\varphi_i(t) \quad (2)$$

The deterministic spatial functions  $\varphi_i(s)$  and the coefficients  $\lambda_i$  are the eigen-functions and the eigen-values of the covariance kernel operator  $C(|s - t|)$  on the domain  $[0, L]$ . Note that the eigen-values are real and non-negative because the covariance is semi-positive definite:

$$\int_0^L \int_0^L C(|s - t|)g(s)g(t)dsdt \geq 0 \quad (3)$$

for any function  $g(s)$  having finite  $L^2$  norm with  $s \in [0, L]$ . The eigen-functions  $\varphi_i(s)$  represent a complete orthonormal basis functions set:

$$\begin{aligned} \int_0^L \varphi_i(s)\varphi_j(s)ds &= \delta_{ij} \\ \sum_{i=0}^{\infty} \varphi_i(s)\varphi_i(t) &= \delta(s - t) \end{aligned} \quad (4)$$

where  $\delta_{ij}$  represents the Kronecker delta and  $\delta(s)$  is the Dirac distribution function. The second equation is due to the completeness of the eigen-functions set [39].

The random field  $f(s, \theta)$  and its covariance are therefore expressed as a truncated sum of  $N$  terms on the domain  $s \in [0, L]$ :

$$\begin{aligned} f(s, \theta) &\approx \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(s) \eta_i(\theta) \\ C(|s-t|) &\approx \sum_{i=1}^N \lambda_i \varphi_i(s) \varphi_i(t) \end{aligned} \quad (5)$$

where  $\eta_i(\theta)$  are random centred uncorrelated (Gaussian if the process is Gaussian) random variables with unit variance defined as the projection of the random process onto the KLE basis:

$$\begin{aligned} \eta_i(\theta) &= \frac{1}{\sqrt{\lambda_i}} \int_0^L f(s, \theta) \varphi_i(s) ds \\ \mathbb{E}[\eta_i(\theta) \eta_j(\theta)] &= \delta_{ij} \end{aligned} \quad (6)$$

In case of non-Gaussian processes, the probability distribution of the KLE coefficients can be obtained by projecting an available set of realisations of the fields onto the KLE basis [40] or by an iterative procedure [41,42].

When the random field is stationary, the KLE modes are alternatively symmetric or skew-symmetric as demonstrated in [27,43]:

$$\varphi_i(s) = \pm \varphi_i(L-s) \quad (7)$$

An odd (even)  $i$  corresponds to a symmetric (skew-symmetric) mode.

The mean-square truncation relative error  $\epsilon_{KL}^2$  is related to the sum of the eigen-values

$$\epsilon_{KL}^2 = \frac{\int_0^L \mathbb{E}[(f(s, \theta) - \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(s) \eta_i(\theta))^2] ds}{\int_0^L \mathbb{E}[f(s, \theta)^2] ds} = 1 - \frac{\sum_{i=1}^N \lambda_i}{\sum_{i=1}^{\infty} \lambda_i} \quad (8)$$

The mean square truncation error decreases monotonically with the number of terms retained in the expansion. This rate depends on the decay of the spectrum  $\mathfrak{S}(\omega)$  of covariance operator [44], where  $\omega$  indicates the frequency. The larger the rate of the spectral decay is (which means the more correlated the process is), the smaller the number of terms needed in the expansion for a given error.

When  $L/l_c \rightarrow \infty$  (where  $l_c$  is the correlation length) the KLE is equivalent to the spectral representation of random fields [30]. Equation (2) can be analytically solved in a few cases, such as rational spectra processes as detailed in [45] or in case of Slepian processes where the eigen-functions are finite trigonometric polynomial functions [46]. However, generally the problem has to be solved numerically. When the random field is discretized into  $n_s$  uniformly spaced points over a domain, Eq. 2 leads to a  $n_s \times n_s$  eigen-value problem. This corresponds to the optimal linear estimation method [25], which is used in this paper. When the domain is huge and a fine discretization of the field is needed the eigen-problem is very heavy to solve, having  $\mathcal{O}(n_s^3)$  complexity.

Other methods, described and compared in [26] can be used to approximate the eigenfunctions more efficiently, such as collocation and Galerkin integration [26]. In this case, the eigen-functions are approximated by a set of  $\tilde{n}_s < n_s$  basis functions, leading to  $\tilde{n}_s \times \tilde{n}_s$  matrix generalized eigen-problem whose complexity is  $\mathcal{O}(2\tilde{n}_s^3)$ . However, since  $\tilde{n}_s$  increases with the size of the domain, solving the eigen-problem still remains an obstacle for large scale random fields. In the frame of this work, any numerical method can be used of solving the KLE but, for simplicity, the optimal linear estimation method is employed.

### Principles of the generation method on a large domain

The general principles of the random fields generation method, which is the main object of this paper, are highlighted in this section through the explanation of a simple case.

In this section the eigen-functions and eigen-values calculated in Eq. (2), defined for  $s \in [0, L]$ , are used to generate a sample of the random process in a domain with  $s \in [0, 2L]$ . The size of the domain is thus doubled. The method can be straightforwardly extended to an arbitrary-sized domain by iterating the technique presented in this section. The general idea is to firstly generate two independent samples, each covering half of the domain, and then impose a correlation between the KLE coefficients of the two samples. At the end, continuous samples of the process on the whole domain with a respected correlation structure are obtained.

$\tilde{f}_1(s, \theta_1)$  and  $\tilde{f}_2(s, \theta_2)$ , with  $\theta_1, \theta_2 \in \Theta$ , are two independent samples sets of the random field  $f(s, \theta)$  introduced in the previous section (see the black curves in Fig. 1):

$$\begin{aligned}\tilde{f}_1(s, \theta_1) &= \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(s) \eta_i(\theta_1) \quad \text{with } s \in [0, L] \\ \tilde{f}_2(s, \theta_2) &= \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(s - L) \eta_i(\theta_2) \quad \text{with } s \in [L, 2L]\end{aligned}\tag{9}$$

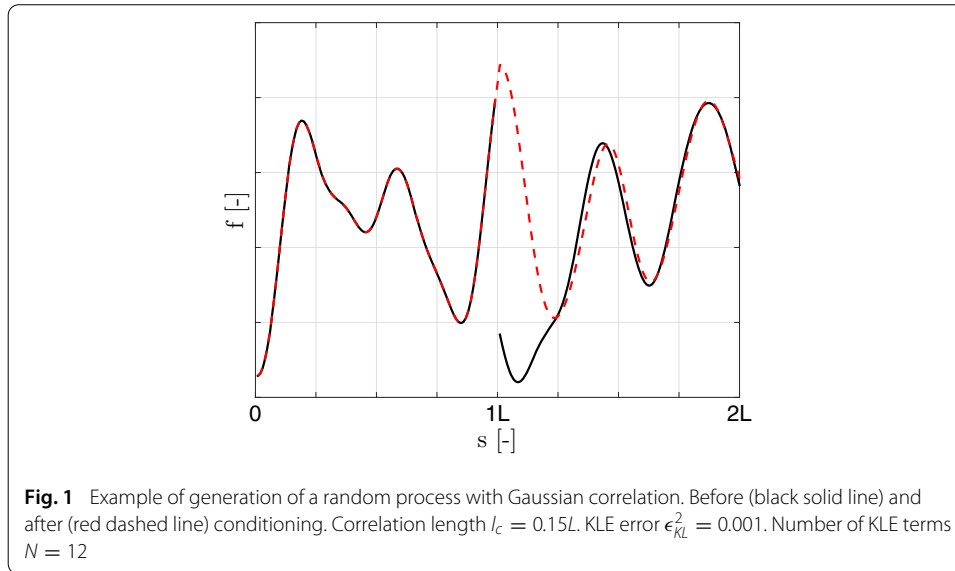
Each one of the two KLE coefficients sets  $\eta_i(\theta_1)$  and  $\eta_i(\theta_2)$  is composed of normalized and uncorrelated random variables as stated in Eq. (6). Since the two sets are independently generated, they are uncorrelated, implying the decorrelation of the random fields:

$$\mathbb{E}[\eta_i(\theta_1) \eta_j(\theta_2)] = 0 \quad \forall i, j \implies \mathbb{E}[\tilde{f}_1(s, \theta_1) \tilde{f}_2(t, \theta_2)] = 0\tag{10}$$

The expression of the covariance function  $C(|s - t|)$  of the field  $f(s, \theta)$  is known for  $0 \leq s, t \leq 2L$ .  $\mathbf{K}$  denotes the  $N \times N$  matrix, here called coupling matrix, whose elements are given by the projection of  $C(|s - t|)$ , in the domain  $s \in [0, L]$  and  $t \in [L, 2L]$ , onto the basis  $\varphi_i(s)$ :

$$\begin{aligned}K_{ij} &= \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_{s=0}^L \int_{\tilde{t}=L}^{2L} C(|s - \tilde{t}|) \varphi_i(s) \varphi_j(\tilde{t} - L) d\tilde{t} ds \\ &= \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_{s=0}^L \int_{t=0}^L C(|s - t - L|) \varphi_i(s) \varphi_j(t) dt ds\end{aligned}\tag{11}$$

The matrix  $\mathbf{K}$  represents the correlation that the two KLE coefficients sets (completely uncorrelated since independently sampled) should have in order to represent a correlated process on the complete domain. This matrix will be therefore used to impose a correlation structure between the two KLE coefficients set.



$\mathbf{L}$  denotes the lower triangular matrices defined through Cholesky decomposition as:

$$\mathbf{I} - \mathbf{K}^T \mathbf{K} = \mathbf{L} \mathbf{L}^T \quad (12)$$

where  $\mathbf{I}$  is the  $N \times N$  identity matrix. Note that the matrices  $\mathbf{K}$  and  $\mathbf{L}$  are defined and used in [33] for generating sets of correlated random fields. In this work they represent the coupling between two sub-domains belonging to a large domain.

The positive definiteness of the matrix  $\mathbf{I} - \mathbf{K}^T \mathbf{K}$  is demonstrated in B. Note that the numerical cost to perform this Cholesky decomposition is not related to the size of the whole domain: it is related to the size of the sub-domain  $L$  and the KLE truncation error  $\epsilon_{KL}^2$  (which determines the number of terms  $N$ ).

$\mathbf{H}$  is a  $N$ -dimensional column vector gathering all the KLE coefficients  $\eta_i(\theta)$ . Therefore the matrices  $\mathbf{K}$  and  $\mathbf{L}$  are used to condition the second KLE coefficients generation set:

$$\tilde{\mathbf{H}}(\theta_2, \theta_1) = \mathbf{K}^T \mathbf{H}(\theta_1) + \mathbf{L} \mathbf{H}(\theta_2) \quad (13)$$

This new set of coefficients  $\tilde{\eta}_i(\theta_2, \theta_1)$ , gathered in the vector  $\tilde{\mathbf{H}}(\theta_2, \theta_1)$ , is then used to generate samples of the random fields  $\tilde{f}(s, \theta_1, \theta_2)$  defined on the domain  $[0, 2L]$ :

$$\tilde{f}(s, \theta_1, \theta_2) = \begin{cases} \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(s) \eta_i(\theta_1), & \text{if } s \in [0, L] \\ \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(s - L) \tilde{\eta}_i(\theta_2, \theta_1), & \text{if } s \in (L, 2L] \end{cases} \quad (14)$$

An example of the generation by doubling the size of the domain is shown in Fig. 1 for a random process characterized by a Gaussian correlation function with the correlation length equal to  $0.15L$  and a KLE truncation error set to  $0.001$  (which leads to 12 retained terms). As shown, in the second part of the domain the realization is modified according to the generation in the first part of the domain. Note that the modification effect is mostly localised near the breaking point  $s = L$ , while it is weak far from this point.

### Correlation structure across the coupled sub-domains

In Eq. (14), the samples are piecewise generated in the two sub-domains. The KLE coefficient set related to the second sub-domain has been correlated to the first one in Eq. (13).

This correlation imposed to the KLE coefficients sets determines the cross-covariance between the two sub-samples:

$$\begin{aligned} \mathbb{E}[\eta_i(\theta_1)\tilde{\eta}_j(\theta_2, \theta_1)] &= K_{ij} \implies \\ \mathbb{E}[\tilde{f}(s, \theta_1, \theta_2)\tilde{f}(t, \theta_1, \theta_2)] &= \sum_{i=1}^N \sum_{j=1}^N K_{ij} \sqrt{\lambda_i \lambda_j} \varphi_i(s) \varphi_j(t-L) \\ &= \sum_{i=1}^N \sum_{j=1}^N \left( \int_0^L \int_L^{2L} C(|s' - t'|) \varphi_i(s') \varphi_j(t' - L) dt' ds' \right) \varphi_i(s) \varphi_j(t-L) \end{aligned} \quad (15)$$

for  $s \in [0, L]$  and  $t \in [L, 2L]$ . In practice, the correlation structure across the two sub-domains is approximated under its projection onto the KLE basis  $\varphi_i(s)$ . This basis, as explained in “Standard 1D Karhunen–Loève expansion” section, is optimal for the representation of the covariance  $C(|s - t|)$  in one sub-domain (Eq. 2). As said the basis terms are selected according to a truncation error decaying as the spectrum  $\mathfrak{S}(\omega)$ , which is the Fourier transform of the covariance  $C(|s - t|)$ .

When  $N \rightarrow \infty$ , because of the basis completeness (Eq. 4), the expression in Eq. (15) is exactly equal to  $C(|s - t|)$ . Otherwise, in the same way as the standard KLE, the cross-covariance is approximated. The cross-spectrum of the two sub-samples is equal to  $\mathfrak{S}(\omega)e^{-i\omega L}$ . The more similar the decays of  $\mathfrak{S}(\omega)$  and cross-spectrum are, the fewer extra terms are needed for a good approximation of the correlation structure.

While the matrix  $\mathbf{K}$  imposes a correlation between two sets of KLE coefficients (coupling effect), the matrix  $\mathbf{L}$  guarantees that the new KLE coefficients  $\tilde{\mathbf{H}}(\theta_2, \theta_1)$  are uncorrelated between them, i.e. Eq. (6) is still satisfied. This means that the KLE representation of the correlation structure, in the second sub-domain, is preserved:

$$\begin{aligned} \mathbb{E}[\tilde{\mathbf{H}}(\theta_2, \theta_1)\tilde{\mathbf{H}}(\theta_2, \theta_1)^T] &= \mathbf{I} \implies \\ \mathbb{E}[\tilde{f}(s, \theta_1, \theta_2)\tilde{f}(t, \theta_1, \theta_2)] &= \sum_{i=1}^N \lambda_i \varphi_i(s-L) \varphi_i(t-L) \end{aligned} \quad (16)$$

for  $s \in [L, 2L]$  and  $t \in [L, 2L]$ .

### Continuity of the generated samples

Let us suppose that the process  $f(s, \theta)$ , defined in “Standard 1D Karhunen–Loève expansion” section, is almost surely continuous, i.e. almost all its sample paths are continuous [47], for  $s \in [0, L]$ :

$$\mathbb{P} \left[ \bigcap_{t \in [0, L]} \left\{ \lim_{s \rightarrow t^-} f(s, \theta) = \lim_{s \rightarrow t^+} f(s, \theta) \right\} \right] = 1 \quad (17)$$

If this condition is satisfied, then the KLE eigen-functions  $\varphi_i(s)$  are continuous on their domain. Note that the almost-sure continuity condition is supposed (in the sub-domain)



for simplicity. When this condition is not fulfilled, the KLE can still be used. Concerning the samples piecewise-generated in Eq. (14), the continuity is not automatically ensured at the breaking point ( $s = L$ ). In this section, the continuity of the sample paths is investigated in this location.

Let us introduce  $\tilde{f}(s, \theta_1, \theta_2)$  as the two random variables corresponding to the left and right limits of the random fields calculated at the breaking point locations ( $s = L$ ):

$$\begin{aligned} l_l(\theta_1) &= \lim_{s \rightarrow L^-} \tilde{f}(s, \theta_1, \theta_2) = \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(L) \eta_i(\theta_1) \\ l_r(\theta_1, \theta_2) &= \lim_{s \rightarrow L^+} \tilde{f}(s, \theta_1, \theta_2) = \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(0) \tilde{\eta}_i(\theta_2, \theta_1) \end{aligned} \quad (18)$$

The following continuity error  $\epsilon_c$  is defined as:

$$\begin{aligned} \epsilon_c &= \frac{\mathbb{E}[(l_l(\theta_1) - l_r(\theta_1, \theta_2))^2]}{2\mathbb{E}[l_l(\theta_1)^2]} = 1 - \frac{\mathbb{E}[l_l(\theta_1)l_r(\theta_1, \theta_2)]}{\mathbb{E}[l_l(\theta_1)^2]} \\ &= 1 - \frac{\sum_{i=1}^N \sum_{j=1}^N K_{ij} \sqrt{\lambda_i \lambda_j} \varphi_i(L) \varphi_j(0)}{\sum_{i=1}^N \lambda_i \varphi_i(L)^2} \end{aligned} \quad (19)$$

where the expectations are calculated using Eqs. (5) and (15). This error compares the variance of the difference between the two-sided limits to the variance of the limits. In other words, if the covariance of the two limits is equal to their variance, the two limits are equal. When this error is small enough, one can assume that the discontinuity jump, at the breaking point, is small compared to the variance of the process.

Note that, because of the completeness of the KLE basis (Eq. 4), when  $N \rightarrow \infty$  the error tends to zero, making equal to one the probability that the left and the right limits take the same value (satisfying the continuity at the breaking location  $s = L$ ).

In numerical applications, the random process is discretized in  $n_s$  steps with discretization step equal to  $\Delta l = \frac{L}{n_s}$ . The continuity error  $\epsilon_c$  can still be evaluated as:

$$\bar{\epsilon}_c = 1 - \frac{\mathbb{E}[\tilde{f}(L, \theta_1, \theta_2) \tilde{f}(L + \Delta l, \theta_1, \theta_2)]}{\mathbb{E}[\tilde{f}(L, \theta_1, \theta_2) \tilde{f}(L - \Delta l, \theta_1, \theta_2)]} = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^N K_{ij} \sqrt{\lambda_i \lambda_j} \varphi_i(L) \varphi_j(\Delta l)}{\sum_{i=1}^N \lambda_i \varphi_i(L) \varphi_i(L - \Delta l)} \quad (20)$$

In practice, the covariance evaluated across the breaking point is compared to the covariance at the lag equal to  $\Delta l$  for  $s = L$  (at the border of the first sub-domain). For instance, when the correlation at the lag equal to  $\Delta l$  is very weak, just a few terms in the sum will make the error  $\bar{\epsilon}_c$  small. The decay of the continuity error is numerically evaluated in ‘‘Inuence of the correlation kernel’’ section for different correlation structures.

Note that an overlapping method for representing large scale random fields is proposed in [9]. This overlapping strategy can be applied to any random fields generation methods (KLE included). By overlapping the sub-domains the continuity issues are avoided, but an error on the correlation representation is introduced. In this paper, the correlation across the sub-domains is imposed. Then the continuity error that can be reduced by adding more terms in the expansion. Since the modal decomposition is affordably performed in one sub-domain, adding more terms in the expansion is not a numerically expensive task.

### Extension of the expansion on an arbitrary large domain

The sequential extension on an arbitrary long domain is straightforward.  $ML$  is the length of the domain composed of  $M$  sub-domains.  $M$  independent realizations are generated in each sub-domain. Then, the  $M$  independently generated random coefficients sets are sequentially correlated:

$$\tilde{\mathbf{H}}^{(m)}(\theta_m, \theta_{m-1}, \dots, \theta_1) = \mathbf{K}^T \tilde{\mathbf{H}}^{(m-1)}(\theta_{m-1}, \dots, \theta_1) + \mathbf{L}\mathbf{H}^{(m)}(\theta_m) \quad (21)$$

with  $m = (2, \dots, M)$  and  $\tilde{\mathbf{H}}^{(1)} = \mathbf{H}^{(1)}(\theta_1)$ . This generation implies that, in each sub-domain, the random process  $m$  is conditioned by the  $m - 1$  previous parts:

$$\tilde{f}_m(s, \theta_m, \theta_{m-1}, \dots, \theta_1) = \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(s - (m-1)L) \tilde{\eta}_i(\theta_m, \theta_{m-1}, \dots, \theta_1) \quad (22)$$

with  $s \in ((m-1)L, mL]$ . By assembling all the parts  $\tilde{f}_m(s, \theta_m, \theta_{m-1}, \dots, \theta_1)$ , with  $\tilde{f}_1(s, \theta_1) = f_1(s, \theta_1)$  generated as in Eq. (9), samples of the process  $\tilde{f}(s, \theta_1, \dots, \theta_M)$  are generated on the whole domain.

In this section the generation in each sub-domain is performed sequentially. Some aspects concerning the parallelisation are discussed in “Parallel computing of the random field generation” section.

### Generation of multi-dimensional random fields

In this section, the random fields generation method presented in “Karhunen–Loève expansion for large scale 1D random processes” section for 1D processes is generalized to 2D and 3D random fields. Note that the term “multi-dimensional” refers to the dimension of the indexing variable of the fields. In this work only random fields with values in  $\mathbb{R}$  are considered. The principle of the method is essentially the same as what has been presented in “Standard 1D Karhunen–Loève expansion” section. The only difference is that, in case of multi-dimensional random fields, one sub-domain must be conditioned with more sub-domains along the different directions (and not only one direction as for 1D processes).

In this section the generation method is described for a general case. If the tensorization is possible, the computational cost for the modal decomposition of the covariance and the random field generation can be reduced. The case generation with tensorizable correlation is reported in A.

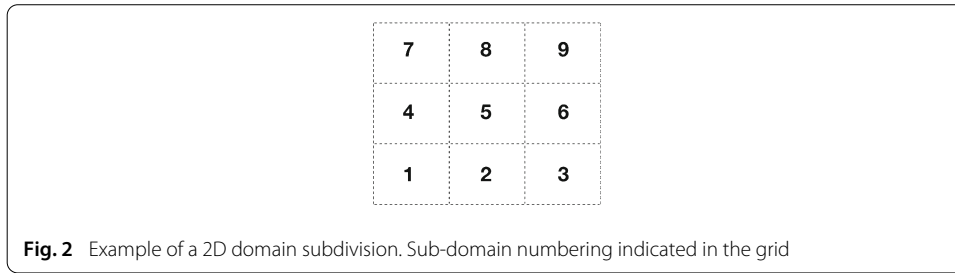
$f(\mathbf{s}, \theta)$  is centred random field, indexed by the variable  $\mathbf{s} \in \mathbb{R}^d$  (with  $d$  being the dimensionality), with covariance function equal to  $C(\mathbf{s}, \mathbf{t})$ .

For the application of the KLE, over the domain  $\mathbf{D} = [0, L]^d$ , the modal decomposition of the covariance operator is performed as in Eq. (2) with multi-dimensional eigen-functions. Then,  $N$  eigen-values and eigen-functions are retained in the expansion to obtain realizations of the field on the domain  $\mathbf{D}$  as in Eq. (5).

In this section a larger domain of size  $\hat{\mathbf{D}} = [0, ML]^d$  is considered. The steps of the method are here summarized. For clarity purposes, a 2D example is illustrated step-by-step.

- **Step 1: Domain subdivision**

The first step consists in the domain subdivision into  $M^d$  sub-domains  $\mathbf{D}_k$  (with



$k = 1, \dots, M^d$ ), each one  $L^d$  sized. An example of the domain subdivision in a 2D case with  $M = 3$  is shown in Fig. 2.

- **Step 2: Determination of the coupling matrices**

The next step is the computation of the coupling matrices. Each sub-domain is connected with the surrounding sub-domains.  $\mathbf{K}^{(pq)}$  indicates the coupling matrix concerning the sub-domains  $\mathbf{D}_p$  and  $\mathbf{D}_q$ . Its elements are calculated as:

$$K_{ij}^{(pq)} = \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_{\mathbf{s} \in \mathbf{D}_p} \int_{\mathbf{t} \in \mathbf{D}_q} C(\mathbf{s}, \mathbf{t}) \varphi_i(\mathbf{s} - \mathbf{o}_p) \varphi_j(\mathbf{t} - \mathbf{o}_q) d\mathbf{s} d\mathbf{t} \quad (23)$$

with  $\mathbf{o}_k = [\min s_1, \dots, \min s_d] \mid \mathbf{s} \in \mathbf{D}_k$ . Note that, when the field is stationary,  $\mathbf{K}^{(pq)} = \mathbf{K}^{(qp)\text{T}}$ . Moreover, two coupling matrices are equal if the relative position between their respective sub-domains is the same in stationary conditions. For instance, with respect to Fig. 2,  $\mathbf{K}^{(12)} = \mathbf{K}^{(23)}$ ,  $\mathbf{K}^{(14)} = \mathbf{K}^{(47)}$ ,  $\mathbf{K}^{(15)} = \mathbf{K}^{(59)}$ , and so on. In this way the number of coupling matrices, and the consecutive Cholesky decompositions, is reduced.

- **Step 3: KLE coefficients conditioning**

The third step is the conditioning of the KLE coefficients sets in each sub-domain with respect to its neighbour sub-domain. At first, the order in which the KLE coefficients sets are conditioned is chosen. Different strategies can be adopted. For example, referring to Fig. 2, the order can be chosen by simply using the domain numbering. Then, each set is generated and conditioned with the sets of all its neighbour sub-domains that have been already generated.

$\mathbf{D}_k$  is the sub-domain which is connected with  $n_k$  sub-domains already generated whose indices are gathered in the set  $\mathcal{I}_k$ . For the sub-domains already generated the following equation holds:

$$\mathbb{E}[\tilde{\mathbf{H}}^{(p)}(\boldsymbol{\theta}_p) \tilde{\mathbf{H}}^{(q)\text{T}}(\boldsymbol{\theta}_q)] = \mathbf{K}^{(pq)} \quad (24)$$

where  $p, q \in \mathcal{I}_k$  and  $\tilde{\mathbf{H}}^{(p)}(\boldsymbol{\theta}_p)$  is a vector gathering the KLE coefficients in the sub-domain  $\mathbf{D}_p$ .

A linear system of  $n_k$  coupled matrix equations is defined such that each equation takes the form:

$$\mathbf{X}_q + \sum_{p \in \mathcal{I}_k \setminus \{q\}} \mathbf{K}^{(pq)} \mathbf{X}_p = \mathbf{K}^{(qk)} \quad (25)$$

where  $q \in \mathcal{I}_{\mathbf{k}}$  and  $\mathbf{X}_q$  are the  $n_{\mathbf{k}}$  ( $N \times N$  sized) matrix unknowns. There are  $n_{\mathbf{k}}$  coupled matrix equations, each one  $N \times N$  sized. This leads to  $N$  linear systems of equations having the same coefficient matrix of size  $n_{\mathbf{k}}N \times n_{\mathbf{k}}N$ . This coefficient matrix is symmetric positive definite.

After its resolution, the matrix  $\mathbf{L}^{(k)}$  is defined through Cholesky decomposition:

$$\mathbf{I} - \sum_{q \in \mathcal{I}_{\mathbf{k}}} \mathbf{K}^{(kq)} \mathbf{X}_q = \mathbf{L}^{(k)} \mathbf{L}^{(k)T} \quad (26)$$

Finally the set of KLE coefficients  $\tilde{\mathbf{H}}^{(k)}(\boldsymbol{\theta}_{\mathbf{k}})$  in the domain  $\mathbf{D}_{\mathbf{k}}$  is generated as:

$$\tilde{\mathbf{H}}^{(k)}(\boldsymbol{\theta}_{\mathbf{k}}) = \sum_{q \in \mathcal{I}_{\mathbf{k}}} \mathbf{X}_q^T \mathbf{H}^{(q)}(\boldsymbol{\theta}_q) + \mathbf{L} \mathbf{H}^{(k)}(\boldsymbol{\theta}_{\mathbf{k}}) \quad (27)$$

where  $\boldsymbol{\theta}_{\mathbf{k}} = [\theta_1, \dots, \theta_{\mathbf{k}}]$ .

The structure of the linear system ensures that the cross-correlation between the neighbour sets is taken into account: by using the expectation of Eq. (24) and the definition of the system in Eq. (25) it can be proven that:  $\mathbb{E}[\tilde{\mathbf{H}}^{(k)}(\boldsymbol{\theta}_{\mathbf{k}}) \tilde{\mathbf{H}}^{(q)T}(\boldsymbol{\theta}_q)] = \mathbf{K}^{(kq)}$ . This ensures the respect of the correlation structure in the same way as for the 1D case (Eq. 15).

As for the computation of the coupling matrix described in the previous step, if the field is stationary, the linear system solution and the Cholesky decomposition are the same if two sub-domains take the same relative position with respect to their neighbour sub-domains previously generated. For example, with respect to the sequential generation shown in Fig. 2, this situation occurs for the sub-domains 2 and 3, 4 and 7, 5 and 8, and so on.

#### • Step 4: Random field generation

The last step is the generation of the field in each sub-domain:

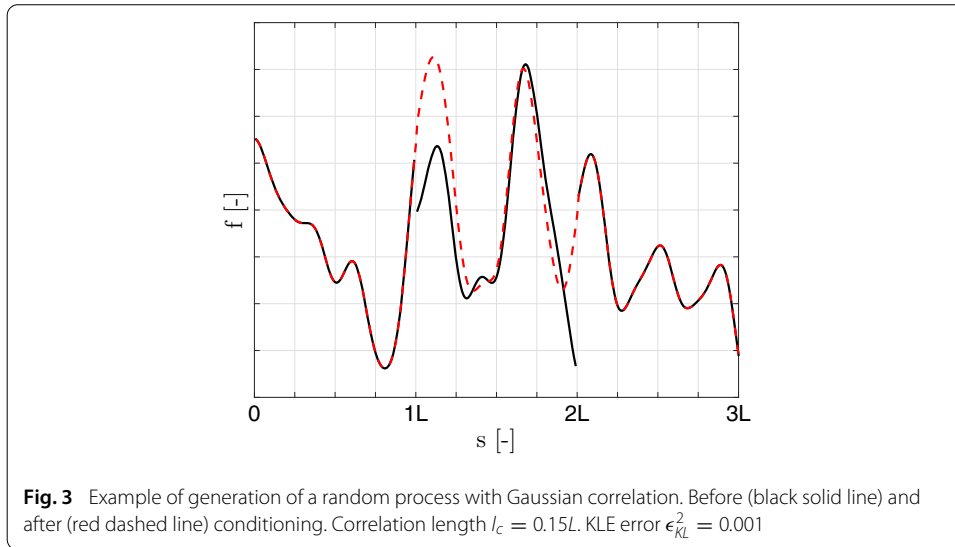
$$f(\mathbf{s}, \boldsymbol{\theta}_{\mathbf{k}}) = \sum_{i=1}^N \sqrt{\lambda_i} \varphi_i(\mathbf{s}) \tilde{\eta}_i^{(k)}(\boldsymbol{\theta}_{\mathbf{k}}) \quad \text{with } \mathbf{s} \in \mathbf{D}_{\mathbf{k}} \quad (28)$$

### Parallel computing of the random field generation

In this section the strategies to adopt for parallelizing the generation method are discussed. The sequential conditioning presented above is simply applicable but, since each part is conditioned by the previous ones iteratively, this technique is not parallelizable. When one wants to use several processors to generate a very large sample of the field another strategy is more advisable.

The first part of this section discusses the parallelization of the 1D processes generation technique. Then, the parallelization strategy for general multi-dimensional fields generation is described.

In this section the parallelization is performed with respect to the indexing variable of the random field (computation of several sub-domains at the same time). However, it is always possible to run distributed computations along the statistical axis, if one needs to sample several realizations of the random fields.



### 1D processes generation parallel computing

Let us consider that  $M$  is odd, without loss of generality. At first,  $M$  sets of KLE coefficients are independently generated ( $\mathbf{H}(\theta_1), \dots, \mathbf{H}(\theta_M)$ ). Then each part corresponding to an even  $m$  is conditioned by the parts at the left and the right:

$$\tilde{\mathbf{H}}^{(m)}(\theta_m, \theta_{m-1}, \theta_{m+1}) = \mathbf{K}^T \mathbf{H}^{(m-1)}(\theta_{m-1}) + \mathbf{K} \mathbf{H}^{(m+1)}(\theta_{m+1}) + \mathbf{R} \mathbf{H}^{(m)}(\theta_m) \quad (29)$$

with  $m = 2 \times (1, \dots, (M-1)/2)$  and  $\mathbf{R}$  being the lower triangular matrix such that:

$$\mathbf{I} - \mathbf{K}^T \mathbf{K} - \mathbf{K} \mathbf{K}^T = \mathbf{R} \mathbf{R}^T \quad (30)$$

The correlation structure, as in the case of the one-sided conditioning (“Principles of the generation method on a large domain” section), is ensured by the correlation between the KLE coefficients sets: because of the conditioning in Eq. (29), it follows that

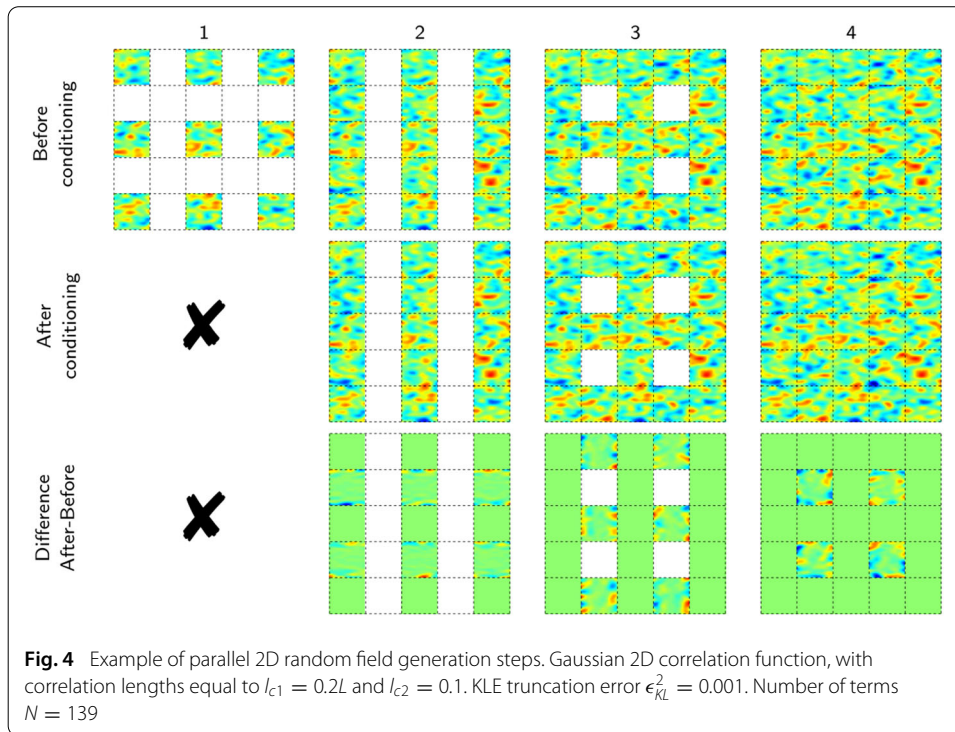
$$\begin{aligned} \mathbb{E}[\tilde{\mathbf{H}}^{(m)}(\theta_m, \theta_{m-1}, \theta_{m+1}) \mathbf{H}^{(m-1)}(\theta_{m-1})] &= \mathbf{K}^T, \\ \mathbb{E}[\tilde{\mathbf{H}}^{(m)}(\theta_m, \theta_{m-1}, \theta_{m+1}) \mathbf{H}^{(m+1)}(\theta_{m+1})] &= \mathbf{K} \end{aligned} \quad (31)$$

In this way, each even part can be parallely conditioned and, then, the KLE coefficients sets are used to generate the sample in each sub-domain. The coefficients corresponding to the odd parts are not changed. Therefore only  $M-1$  coefficients sets are conditioned. An example to show how this technique works is presented in Fig. 3 with  $M=3$ .

Note that when  $M$  is even, the only difference is that the last sub-domain ( $m=M$ ) is only conditioned from the left side (as done in “Principles of the generation method on a large domain” section).

### Multi-dimensional fields generation parallel computing

For parallelizing the multi-dimensional fields generation method presented in “Generation of multi-dimensional random fields” section, the only difference with the sequential conditioning technique is the sub-domains ordering. Indeed, if two sub-domains do not interact, i.e. they are enough far to consider their correlation equal to zero, they can be



processed at the same time. Then the equations for the conditioning are the same as in “Generation of multi-dimensional random fields” section.

One possible strategy is to condition, sequentially, as many not connected sub-domains as available. Subdividing the domain in  $(ML)^d$ , with  $M$  odd, a total number of  $2^d$  steps, in which more sub-domains are parallelly processed, are needed. For each  $r$  going from 0 to  $d$  there will be  $\binom{d}{r}$  steps in which the number of running processes is equal to:

$$n_{\text{proc}}(r) = \left(\frac{M+1}{2}\right)^{d-r} \left(\frac{M-1}{2}\right)^r \quad \text{for } \binom{d}{r} \text{ steps} \quad (32)$$

For instance, with respect to Fig. 2 where  $d = 2$  and  $M = 3$ , the sub-domains 1–3–7–9 are processed parallelly and independently. Then the sub-domains 4 and 6 are parallelly computed, followed by the numbers 2 and 8 (computed at the same time). At the end the sub-domain number 5 is processed.

A total number of 4 steps are needed for generating a 2D random field. This sequential generation with parallel computing is illustrated in Fig. 4 for a case when  $M = 5$ . The generation is performed in 4 steps. The top row of the figure shows the generation without conditioning the KLE random variable, while the middle row shows the generation after the conditioning. Note that the steps 2 and 3 cannot be combined in a single step because the sub-domains generated in step 3 are conditioned by the sub-domains generated in step 2, just around their corners.

### Extension to non-stationary random fields

All the methods cited above concern stationary random fields. The representation of non-stationary random fields can be achieved by modifying a stationary field (already generated) using one of the methods described in the introduction or the method proposed in

this paper. A stationary process can be multiplied by a deterministic slowly-varying function for reproducing a non-stationary effect as in [48,49]. In [3,50] the spectral representation is extended to non-stationary processes with evolutionary power spectrum [51], i.e. when the power spectral density can be modulated by a deterministic function depending on the frequency and the support variable, with the possibility to improve the computation by using the FFT [52]. Concerning the ARMA method, it has been extended to non-stationary processes by introducing time-dependent coefficients [53].

Conversely, the application of the KLE does not require that the random field is stationary.  $C(s, t)$  is the covariance of a non-stationary random field. After splitting the domain of size  $ML$  in  $M$  parts, to apply the method described in “Karhunen–Loève expansion for large scale 1D random processes” section, the covariance decomposition has to be performed in each sub-domain since the covariance kernel depends on the indexing variable:

$$\int_{(m-1)L}^{mL} C(s, t) \varphi_i^{(m)}(t) dt = \lambda_i^{(m)} \varphi_i^{(m)}(s) \quad (33)$$

with  $m = 1, \dots, M$  and  $s \in [(m-1)L, mL]$ . The coupling matrix, defined in Eq. 11, also depends on the domain part:

$$K_{ij}^{(m)} = \frac{1}{\sqrt{\lambda_i^{(m)} \lambda_j^{(m+1)}}} \int_{(m-1)L}^{mL} \int_{(m-1)L}^{mL} C(s, t+L) \varphi_i^{(m)}(s) \varphi_j^{(m+1)}(t) ds dt \quad (34)$$

with  $s, t \in [(m-1)L, mL]$ .

When the field is non-stationary and the variance varies along the indexing variable, it is not possible to condition the KLE coefficients from the left and the right side, i.e. the decomposition in Eq. 30 is not possible. For this reason the conditioning is sequential:

$$\tilde{\mathbf{H}}(\theta_m, \theta_{m-1}, \dots, \theta_1) = \mathbf{K}^{(m)\top} \tilde{\mathbf{H}}(\theta_{m-1}, \dots, \theta_1) + \mathbf{L}^{(m)} \mathbf{H}(\theta_m) \quad (35)$$

with:

$$\mathbf{I} - \mathbf{K}^{(m)\top} \mathbf{K}^{(m)} = \mathbf{L}^{(m)} \mathbf{L}^{(m)\top} \quad (36)$$

Examples of generation, with  $M = 3$  and covariance shown in Fig. 5, are illustrated in Fig. 6.

Note that, because of covariance non-stationarity, the sets of eigen-functions (and also their number)  $\varphi_i^{(m)}(s)$  and  $\varphi_i^{(m+1)}(s)$  related to two sub-domains can be different. For this reason, the matrices  $\mathbf{K}^{(m)}$  and  $\mathbf{L}^{(m)}$  are not necessarily square.

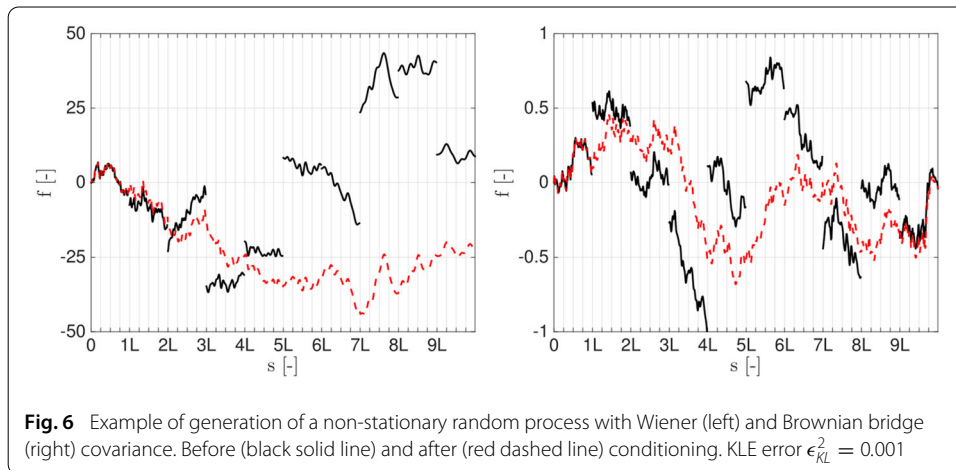
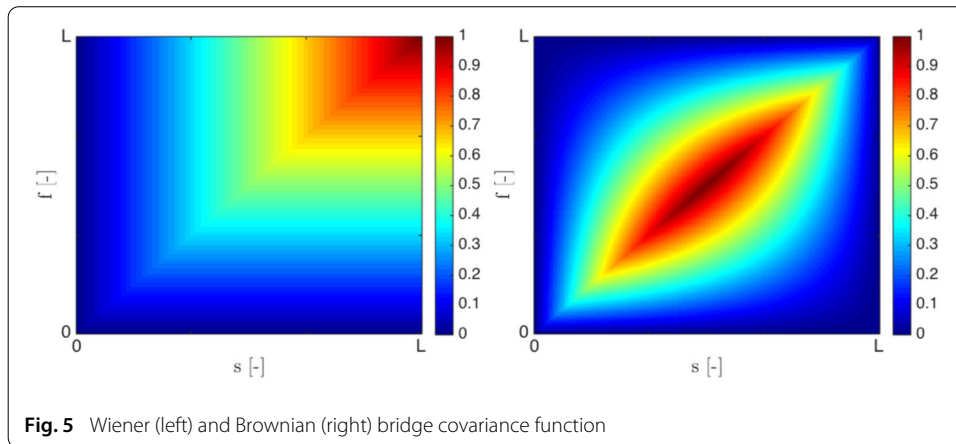
## Numerical applications

### Example of generation on a large domain

In this section the generation of a sample of a 2D random field on the domain  $\mathbf{D} \in [0, ML]^2$ , with  $M = 21$  is presented. The random field is characterized by the following (anisotropic and non-tensorizable) correlation function:

$$C(\mathbf{s}, \mathbf{t}) = \exp \left( -\sqrt{\left( \frac{s_1 - t_1}{l_{c1}} \right)^2 + \left( \frac{s_2 - t_2}{l_{c2}} \right)^2} \right) \quad (37)$$

with  $l_{c1} = 0.2L$  and  $l_{c2} = 0.1L$ .



The field is discretized into  $(n_s M)^2$  steps, with  $n_s = 100$ . The KLE truncation error  $\epsilon_{KL}^2$  is set to 0.001, giving a number of retained terms in the expansion  $N$  equal to 9600. Using directly the standard KLE on such a large domain as the one here considered is unaffordable: using the optimal linear estimation method [25] would require an eigen-decomposition of a  $(n_s M)^2 \times (n_s M)^2$ , i.e.  $4,410,000 \times 4,410,000$ , sized matrix.

The method described in “Generation of multi-dimensional random fields” section is employed in this section to generate the sample. No parallel computing is performed: this field is generated sequentially in each sub-domain.

The computational time for the generation is indicated in Table 1. Note that the computation of the coupling matrices is the most step of the process. This fact is due to the number of retained KLE terms, which determines the size of the coupling matrices and therefore the complexity of the linear systems (solved using the Cholesky decomposition of the coefficient matrix) formulated in “Generation of multi-dimensional random fields” section. A sample of the random field thus generated is shown in Fig. 7.

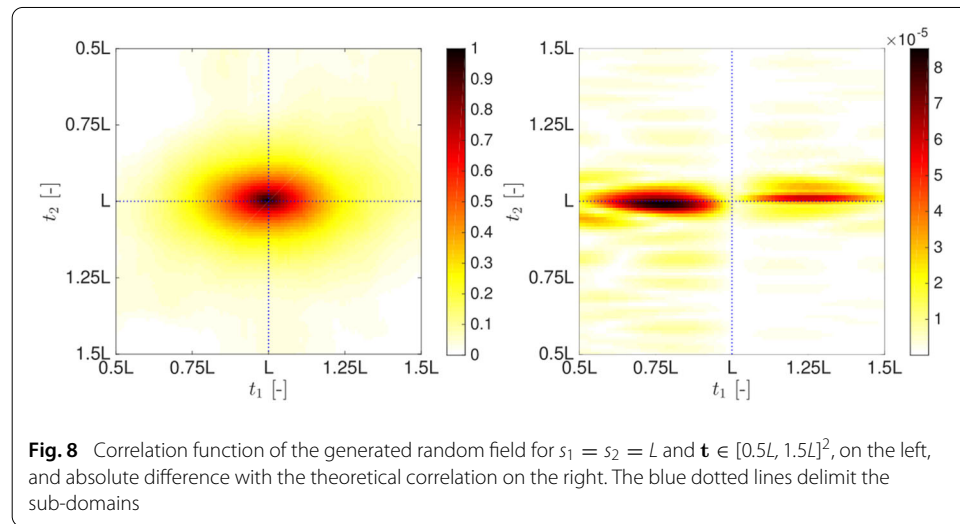
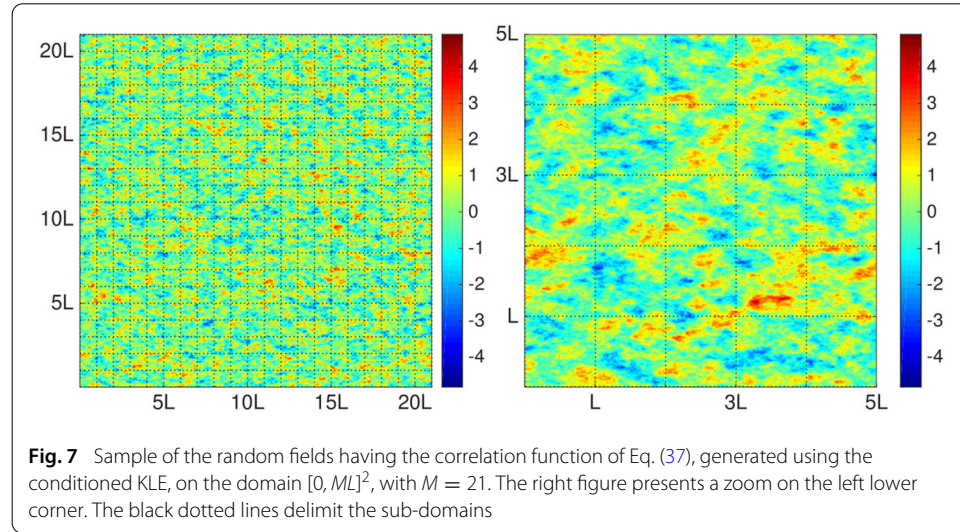
The correlation function of the generated random field is shown in Fig. 8 for two different locations. The correlation structure is well respected in the proximity of the junction point ( $\mathbf{s} = [L, L]$ ). Note that in this location the error is lower than the chosen truncation error.



**Table 1** Computational time for generating a 2D random field with correlation function as in Eq. (37) on an extended domain  $[0, ML]^2$ , with  $M = 21$  using the conditioned KLE

Operation	Elapsed time (s)
Kernel modal decomposition	56.31
Conditioning matrices computation	2220.17
Random field sampling	1351.95

Computed on Intel® Xeon® CPU E5-2695 v3 @2.30 GHz. 1 processor used



### Computational complexity

In this section the computation time of the standard KLE is compared with the generation method proposed in this work in the case of tensorizable and non-tensorizable correlation functions. A 2D random field defined on the domain  $\mathbf{s} \in [0, ML]^d$ , with  $d$  being the dimension. The modal decomposition is solved using the optimal linear estimation method [25] in which the domain is uniformly discretized. Let us indicate as  $n_s$  the number of discretization steps of a segment of length  $L$  along one of the directions. The domain is thus discretized in  $(n_s M)^d$  parts.

**Table 2 Numerical complexity of the standard KLE and the conditioned KLE**

Operation	Standard KLE	Conditioned KLE
Kernel modal decomposition	$\mathcal{O}((n_s M)^{3d})$	$\mathcal{O}(n_s^{3d})$
Conditioning matrices computation	–	$\mathcal{O}(N^3(3^d - 1)^3)$
Random field sampling	$\mathcal{O}(n_s^d M^{2d} N)$	$\mathcal{O}((n_s M)^d N)$

Non-tensorizable covariance kernel.  $n_s$  number of discretization steps of a segment of length  $L$ ,  $M$  number of prolongations in one direction,  $d$  dimension,  $N$  total number of retained KLE terms for  $M = 1$

**Table 3 Numerical complexity of the standard KLE and the conditioned KLE**

Operation	Standard KLE	Conditioned KLE
Kernel modal decomposition	$\mathcal{O}((n_s M)^3 d)$	$\mathcal{O}(n_s^3 d)$
Conditioning matrices computation	–	$\mathcal{O}(\tilde{N}^3 d)$
Random field sampling	$\mathcal{O}(n_s^d M^{d+1} \tilde{N})$	$\mathcal{O}((n_s M)^d \tilde{N})$

Tensorizable covariance kernel.  $n_s$  number of discretization steps of a segment of length  $L$ ,  $M$  number of prolongations in one direction,  $d$  dimension,  $\tilde{N}$  maximal number of retained KLE terms among all the directions (dimensions) for  $M = 1$

In the first part of this section, an example concerning a tensorizable 2D random field is presented. In the second part, the numerical complexity of the standard and the conditioned KLE are compared.

In Tables 2 and 3 the complexity of the standard KLE and the conditioned KLE proposed in this work are compared for the case of, respectively, non-tensorizable and tensorizable kernel covariance. In these tables  $N$  indicates the total number of retained KLE terms, while  $\tilde{N}$  indicates the maximal number of retained KLE terms among all the directions (dimensions) for  $M = 1$ .

#### **Kernel modal decomposition**

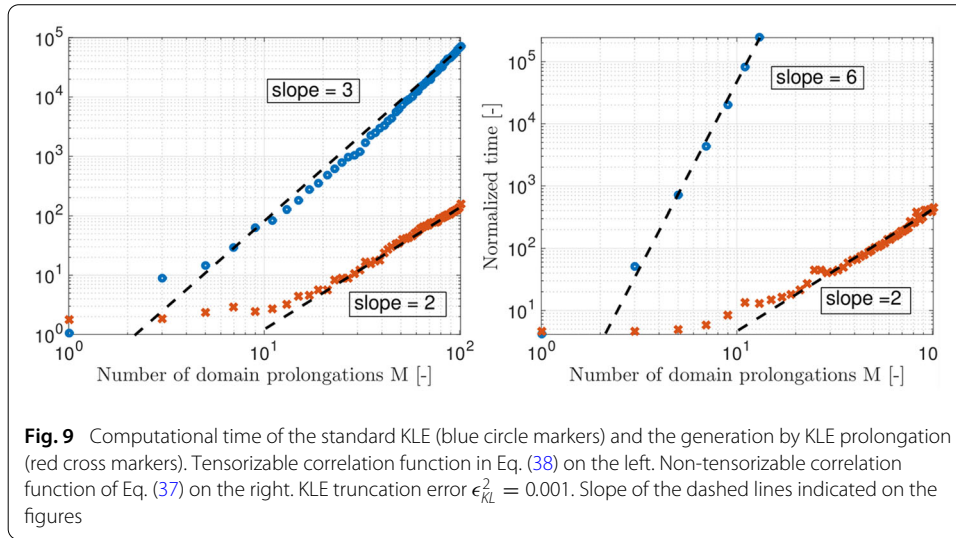
The modal decomposition complexity does not depend on  $M$  when the conditioned KLE is employed. For the standard KLE, this complexity grows with  $\mathcal{O}(M^3)$ , when the covariance kernel is tensorizable, and  $\mathcal{O}(M^{3d})$  when not. This is the main limit for directly using the KLE on the whole domain.

#### **Conditioning matrices computation**

The computation of the matrices used for conditioning the KLE coefficients, described in “Generation of multi-dimensional random fields” section, requires some Cholesky decomposition operations. Its complexity does not depend on  $M$ , but on the number of KLE terms (depending on the truncation error) and the sequential prolongation strategy. In fact, the size of the linear system in Eq. (25), that is solved with Cholesky decomposition of the coefficient matrix, depends on the number of connections of the considered sub-domain. The complexity, indicated in Table 2, for the non-tensorizable kernel case is referred to the resolution of the largest linear system (the one that is associated with the sub-domain having the largest number of neighbours).

#### **Random field sampling**

The sample generation is the only operation growing with  $M$  when the conditioned KLE is used: in this case the complexity grows with  $\mathcal{O}(M^d)$ . The use of the standard KLE requires a complexity, for this stage, of  $\mathcal{O}(M^{2d})$ , in case of non-tensorizable kernel and  $\mathcal{O}(M^{d+1})$ , in case of a tensorizable kernel. Another advantage of using the method proposed in this work, is the possibility, due to the domain splitting, of storing separately each part of the



sample corresponding to a sub-domain. This can prevent memory issues when the total size of the domain is huge.

As example, the evolution of computational costs for the random field generation when the numbers of sub-domains increases (using the standard and the conditioned KLE) is analysed for two cases: non-tensorizable correlation of Eq. (37) and tensorizable correlation defined as:

$$C(\mathbf{s}, \mathbf{t}) = \exp\left(-\left(\frac{|s_1 - t_1|}{l_{c1}}\right)^2\right) \exp\left(-\left(\frac{|s_2 - t_2|}{l_{c2}}\right)^2\right) \quad (38)$$

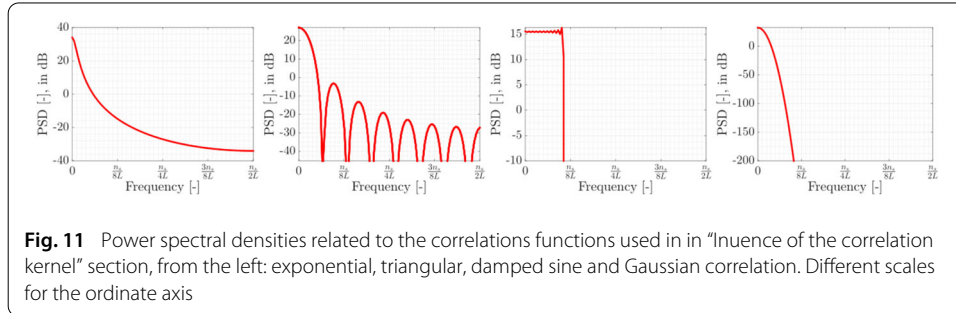
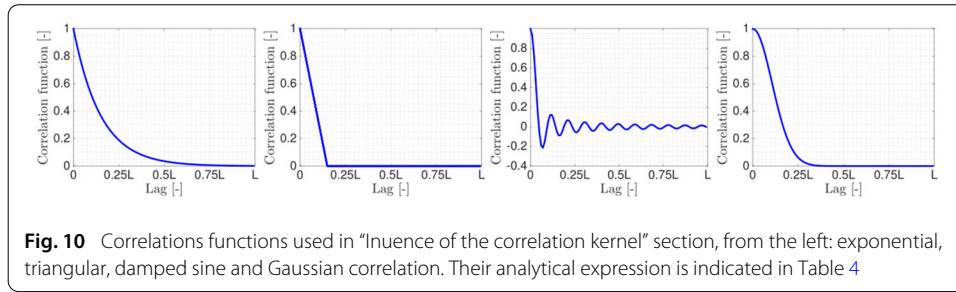
with  $l_{c1}$  and  $l_{c2}$  respectively equal to  $0.3L$  and  $0.2L$ . The KLE truncation error is chosen to be equal to 0.001 (corresponding to a number of terms  $N = 54$  when the size of the domain is  $[0, L]^2$ ).

The evolution of the computation cost is shown in Fig. 9. The slopes indicated on the figures are coherent with the complexities (with respect to the numbers of sub-domains  $M$ ) indicated in Tables 2 and 3.

Since the modal decomposition is the most expensive stage, it is convenient to partition the domain in smaller sub-domains, even though, in this case, the random field sampling stage will be more expensive. However, the correlation structure should be well represented in one sub-domain, i.e. the correlation should tend to zero at a lag equal to  $L$ . This is because, with the method proposed in this paper, the correlation across two sub-domains which are not neighbour is neglected.

#### Influence of the correlation kernel

In this section the influence of the correlation kernel is discussed when the conditioned KLE is employed. Four 1D random processes, indexed by the variable  $s \in [0, ML]$ , are considered in this section. For the numerical representation, the processes are discretized into  $n_s M$  steps, such that  $\Delta l = \frac{L}{n_s} = 0.01L$ . Their correlation functions and their power spectral densities (PSDs) are respectively shown in Figs. 10 and 11. Their analytical



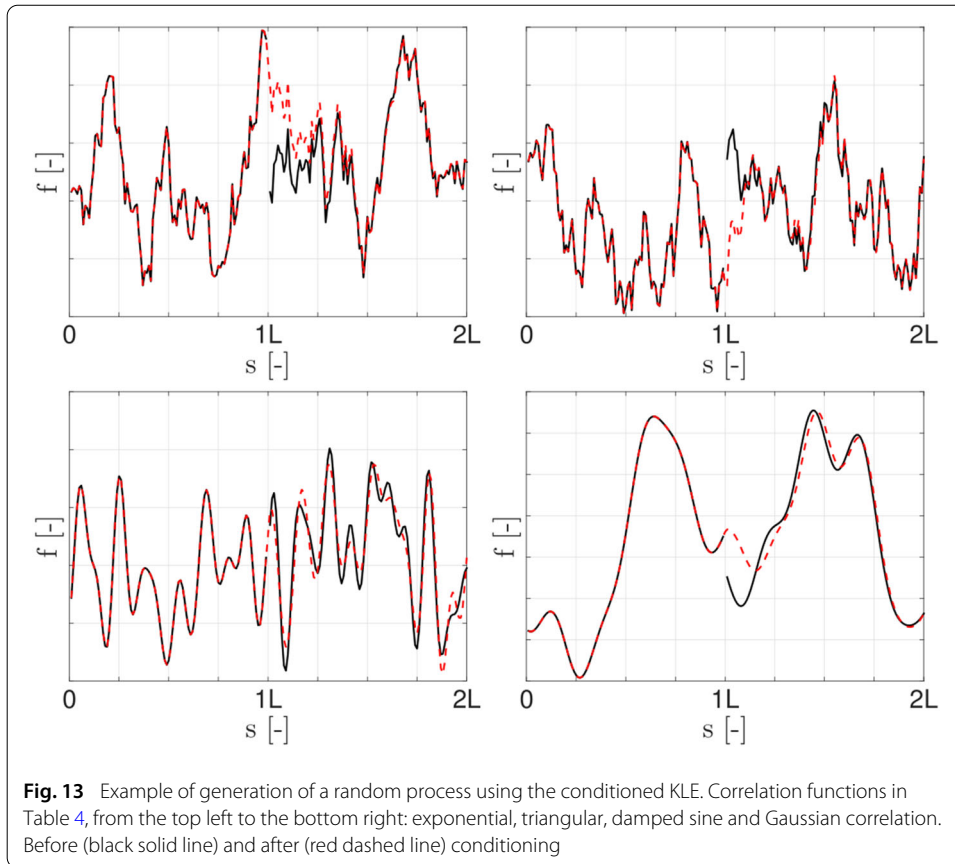
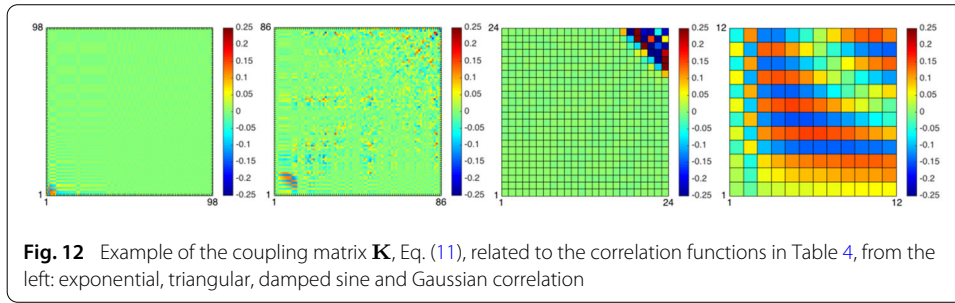
**Table 4** Correlation functions used in “Influence of the correlation kernel” section and relative truncation errors, number of retained KLE terms and continuity errors, with  $l_c = 0.15L$  in all the cases and  $\tau = |s - t|$

Kernel name	Function	Truncation error $\epsilon_{KL}^2$	Retained terms $N$	Continuity error $\epsilon_c$
Exponential	$\exp\left(-\frac{\tau}{l_c}\right)$	0.001	98	$1.7 \times 10^{-6}$
Triangular	$\max\left(1 - \frac{\tau}{l_c}, 0\right)$	0.001	86	$1.8 \times 10^{-3}$
Damped sine	$\frac{l_c}{10\tau} \sin\left(\frac{10\tau}{l_c}\right)$	0.001	24	$1.9 \times 10^{-2}$
Gaussian	$\exp\left(-\frac{\tau}{l_c}\right)^2$	0.001	12	$9.2 \times 10^{-3}$

expression is indicated in Table 4. The parameter  $l_c$  (here called correlation length) is equal to  $0.15L$  for all the cases.

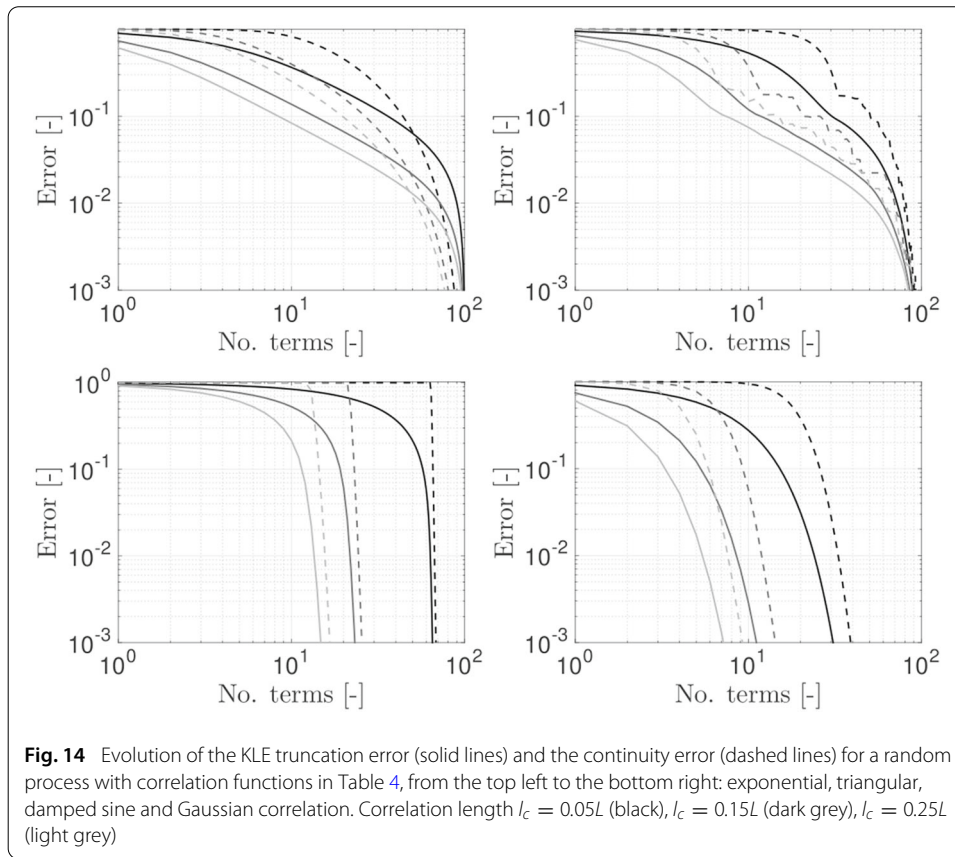
The KLE truncation error  $\epsilon_{KL}^2$  (Eq. 8) is set to 0.001. The choice of this truncation error determines the number of retained KLE terms  $N$  and the continuity error (defined in Eq. 19), that are indicated in Table 4. Note how the number of terms increases as the spectral density decay (Fig. 11) is slower. For instance, the exponential correlated processed requires more than 8 times more terms than the Gaussian correlated process.

The coupling matrices  $\mathbf{K}$  of each random process are displayed in Fig. 12. Note that, even though the number of required KLE terms to guarantee the same error is the largest one for the exponential correlated process, only few first modes are significant for the coupling. Conversely, for the sine damped correlated process, only the last modes are important for the coupling. In fact, the last (less energetical) modes represent the oscillation (not totally damped for  $s = L$ ) of this correlation function. This oscillation is fundamental for the coupling. Concerning the triangular correlation function, all the modes interact for the coupling.



Examples of generated samples, when the size of the domain is split into two sub-domains ( $M = 2$ ), are shown in Fig. 13. Note that, in all the cases except the damped sine correlated process, the corrections due to the conditioning of the KLE terms only concern the region around the breaking points ( $s = L$ ), while the process samples are weakly modified far from this location. In case of the damped sine correlation function, the correction modifies the sample in the whole second sub-domain. This is due to the oscillating behaviour of the correlation function (Fig. 10).

The evolution of the continuity error defined in Eq. (19) is shown in Fig 14 and compared with the KLE truncation error (Eq. 8) for correlation kernels considered in this section (Table 4) with three different correlation lengths. Note how the continuity error, for the damped sine correlation, is firstly constant before suddenly decaying after adding more terms. For the triangular correlation, this error decays not continually.



Note that the sample-path continuity of the field is ensured only if the number of terms in the KLE tends to infinity (use of a complete orthonormal basis). In the other cases, the sample-path continuity is not recovered in the breaking points locations, but the discontinuity jump can be reduced in order to be enough small for the numerical applications.

## Conclusion

Solving the KLE modal decomposition, when the domain is much larger than the correlation length and a small discretization step is needed, represents a computational issue that can become unaffordable. To deal with this issue, a “conditioned” Karhunen–Loève expansion is proposed. The domain is subdivided in sub-domains where the modal decomposition can be comfortably computed. Then, parts of the field are generated in each sub-domain and conditioned with their neighbours in order to ensure the continuity and the correlation structure.

This generation method is applicable to multi-dimensional fields with a strong simplification in case of tensorizable covariance kernels. The capability of the KLE for generating non-stationary random fields is also preserved with the proposed generation method. In every case the computational cost is largely reduced. Another important advantage is that the proposed technique can be easily parallelized to further reduce the computational time.

Moreover, another advantage of using the method proposed in this work is the possibility, due to the domain splitting, of working locally on each part of the sample corresponding to a sub-domain. This can prevent memory issues when the total size of the domain is huge.

The method presented in this paper concerns Gaussian centred random fields. Non-Gaussian fields can be obtained by transforming the generated Gaussian fields by the application of the Rosenblatt transformation to obtain the prescribed first order marginal probability function.

#### Authors' contributions

AMP, RC and GP conceived the presented technique and carried out the mathematical description. AMP wrote this manuscript with support from RC and GP. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup>RATP, 54 Quai de la Rapée, 75012 Paris, France, <sup>2</sup>Laboratoire MSSMat, UMR CNRS 8579, CentraleSupélec, Université Paris-Saclay, 8-10 Rue Joliot Curie, 91190 Gif-sur-Yvette, France.

#### Acknowledgements

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Availability of data and materials

Data will be available on demand.

#### Consent for publication

All authors consent for publication.

#### Ethics approval and consent to participate

Not applicable.

#### Funding

Not applicable.

## Appendix A: Tensorizable correlation function

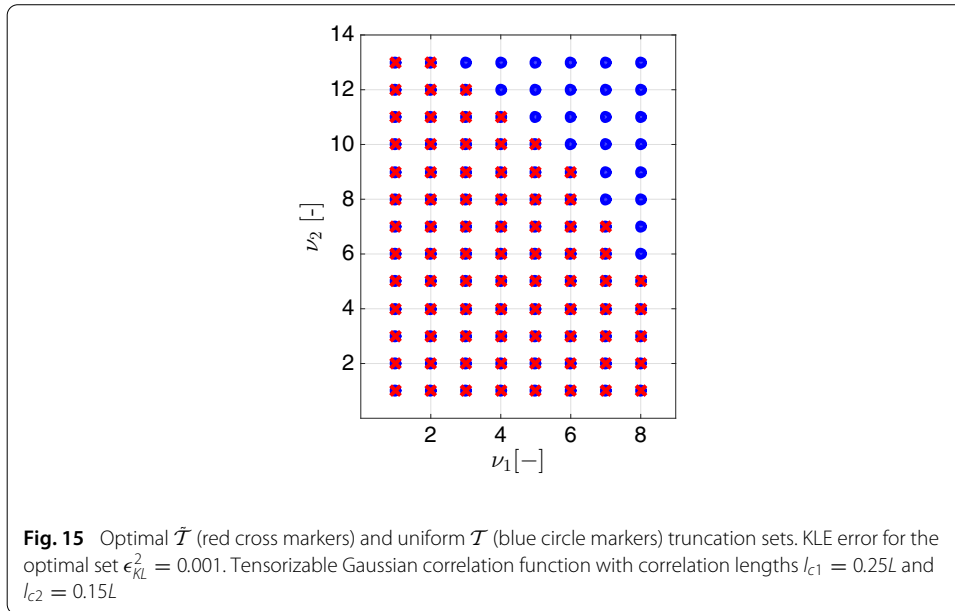
In this section the general random fields generation method, described above, is simplified in case of tensorizable covariance kernel. Indeed, in this situation, one does not need to solve the modal decomposition in a multi-dimensional space. By combining the 1D random processes generation (presented in “Karhunen–Loève expansion for large scale 1D random processes” section) along each dimension, a multi-dimensional realization is obtained.

$\nu = [\nu_1, \dots, \nu_d]$  denotes a  $d$ -dimensional index with  $\nu_1, \dots, \nu_d \in \mathbb{N}^+$ . A tensorizable correlation function can be written in the form:

$$C(\mathbf{s}, \mathbf{t}) = \prod_{l=1}^d C_l(|s_l - t_l|) \quad (39)$$

The KLE decomposition is simplified. For the application of the KLE, over the domain  $\mathbf{D} = [0, L]^d$ , the modal decomposition of the covariance operator can be performed separately along each dimension. The multi-dimensional KLE solutions are obtained by the tensor product of the eigen-functions and the product of the eigen-values related the the 1D KLE in each dimension:

$$\lambda_\nu = \prod_{l=1}^d \mu_{\nu_l} \quad \text{and} \quad \varphi_\nu(\mathbf{s}) = \prod_{l=1}^d \psi_{\nu_l}(s_l) \quad (40)$$



By sorting the eigen-values  $\lambda_\nu$  by decreasing order, a truncation  $d$ -dimensional set  $\tilde{\mathcal{T}}$  (containing the indices of the retained eigen-values) is chosen for a given KLE truncation error (Eq. 8). This multi-dimensional set, which is optimal in the sense of the  $\ell^2$  error, is composed of  $d$  uni-dimensional sets  $\mathcal{T}_l = \{1, \dots, N_l\}$ . For each multi-index  $\nu \in \mathcal{T}$  there are  $d$  indices  $\nu_l \in \mathcal{T}_l$ . Choosing a KLE truncation error the set  $\tilde{\mathcal{T}}$  and the corresponding sets  $\mathcal{T}_l$  are automatically defined.

However, for keeping the notation simple, without loss of generality, in this section a uniform multi-dimensional grid  $\mathcal{T}$  is chosen as truncation set. This means that the set  $\mathcal{T}$  is obtained by tensorizing the sets  $\mathcal{T}_l$  (that are derived from the set  $\tilde{\mathcal{T}}$ ) and  $\tilde{\mathcal{T}} \subseteq \mathcal{T}$ . The sets  $\tilde{\mathcal{T}}$  and  $\mathcal{T}$  are shown in Fig. 15 for a case involving a tensorizable Gaussian correlation function with correlation lengths  $l_{c1} = 0.25L$  and  $l_{c2} = 0.15L$  and choosing a KLE error for the optimal set  $\epsilon_{KL}^2 = 0.001$ .

Therefore, in the domain  $\mathbf{D}$ , the field is generated as:

$$f(\mathbf{s}, \theta) = \sum_{\nu \in \mathcal{T}} \sqrt{\lambda_\nu} \varphi_\nu(\mathbf{s}) \eta_\nu(\theta) = \prod_{l=1}^d \left( \sum_{\nu_l \in \mathcal{T}_l} \sqrt{\mu_{\nu_l}} \psi_{\nu_l}(s_l) \right) \eta_\nu(\theta) \quad (41)$$

For the random fields generation on a large domain of size  $[0, ML]^d$ , the 1D prolongation technique (“Principles of the generation method on a large domain” section) is used in each dimension. The steps are here reported.

- **Step 1: Domain subdivision**

The first step is the same as for the non-tensorizable correlation case:  $(ML)^d$  sub-domains are obtained.

- **Step 2: Determination of the conditioning matrices**

In this step the matrices used for conditioning the KLE coefficients are computed. Differently from the non-tensorizable case, where the matrices are computed for all possible multi-dimensional directions, here for each uni-dimensional direction there



is one coupling matrix  $\mathbf{K}^{(l)}$  and one triangular matrix  $\mathbf{L}^{(l)}$  computed as in Eqs. (11) and (12).

- **Step 3: Iterative 1D random generation**

As suggested from Eq. (41), the random generation can be performed in a drop-down way starting from the dimension  $d$  and arriving to the first dimension.

- *Sub-step a*

Let us consider the 1D random process  $f^{(d)}(s_d, \theta)$ , with  $s_d \in [0, ML]$ , whose correlation function is equal to  $C_d(|s_d - t_d|)$  and represented through KLE (on the domain  $[0, L]$ ) as:

$$f^{(d)}(s_d, \theta) = \sum_{v_d=1}^{N_d} \sqrt{\mu_{v_d}} \psi_{v_d}(s_d) \eta_{v_d}(\theta) \quad (42)$$

Using the method described in “Extension of the expansion on an arbitrary large domain” section, its realizations are extended on the domain  $s_d \in [0, ML]$  by using the matrices  $\mathbf{D}^{(d)}$  and  $\mathbf{L}^{(d)}$  computed at the previous step.

- *Sub-step b*

Let us consider the random field  $f^{(d-1)}(s_{d-1}, s_d, \theta)$  whose correlation function is equal to  $C_{d-1}(|s_{d-1} - t_{d-1}|)C_d(|s_d - t_d|)$  and represented on the domain  $s_{d-1} \in [0, L]$  and  $s_d \in [0, ML]$ :

$$f^{(d-1)}(s_{d-1}, s_d, \theta) = \sum_{v_{d-1}=1}^{N_{d-1}} \sqrt{\mu_{v_{d-1}}} \psi_{v_{d-1}}(s_{d-1}) f_{v_{d-1}}^{(d)}(s_d, \theta) \quad (43)$$

where  $f_{v_{d-1}}^{(d)}(s_d, \theta)$  are random processes sampled as in Eq. (42). Its realizations are extended along the direction  $d - 1$  on the domain  $s_{d-1} \in [0, ML]$  by using the method described in “Extension of the expansion on an arbitrary large domain” section, regardless the dimension  $s_d$  ( $\forall s_d \in [0, ML]$ ).

By recursively iterating the sub-step b, until the first dimension, a complete realization of  $f(\mathbf{s}, \theta)$  is obtained on the whole domain  $\mathbf{s} \in [0, ML]^d$ .

## Appendix B: Positive definiteness condition

In this section some considerations about the positive definiteness condition needed to perform the Cholesky decomposition of Eq. (12) are discussed. It is shown that the Cholesky decomposition is always possible.

The matrix  $\mathbf{K}^T \mathbf{K}$  is symmetric, real and semi-positive definite, because it is a product of a real matrix and its transpose. It follows that the matrix  $\mathbf{I} - \mathbf{K}^T \mathbf{K}$  is real and symmetric.

The Sylvester’s criterion asserts that a real and symmetric matrix is positive definite if and only if all its leading principal minors are positive. The application of this criterion to Eq. (12) gives the following condition for the required positive definiteness:

$$\det(\mathbf{I}_r - (\mathbf{K}^T \mathbf{K})_r) > 0, \quad \text{with } r = 1, \dots, N \quad (44)$$

where the subscript  $r$  indicates the  $r \times r$  top left corner sub-matrix. The above equation represents a necessary and sufficient condition for the positive definiteness of the matrix  $\mathbf{I} - \mathbf{K}^T \mathbf{K}$ .

Since the matrix  $\mathbf{K}^T \mathbf{K}$  is semi-positive definite, all its leading minors are non-negative. Using the properties of the determinants, it follows that:

$$\det(\mathbf{I}_r - (\mathbf{K}^T \mathbf{K})_r) \geq 1 \pm \det((\mathbf{K}^T \mathbf{K})_r) \geq 1 - \det((\mathbf{K}^T \mathbf{K})_r) \quad (45)$$

An upper bound for the leading minors of the matrix  $\mathbf{K}^T \mathbf{K}$  is given by:

$$\sqrt[r]{\det((\mathbf{K}^T \mathbf{K})_r)} \leq \frac{1}{r} \text{tr}((\mathbf{K}^T \mathbf{K})_r) = \frac{1}{r} \sum_{i=1}^r \sum_{j=1}^N K_{ij}^2 \quad (46)$$

By replacing the expression of the elements of the matrix  $\mathbf{K}$ , given in Eq. (11), it follows that:

$$\sum_{j=1}^N K_{ij}^2 = \sum_{j=1}^N \left| \int_0^L \int_0^L C(s, t + L) \frac{\varphi_i(s)\varphi_j(t)}{\sqrt{\lambda_i \lambda_j}} ds dt \right|^2, \quad \text{with } i = 1, \dots, N \quad (47)$$

By introducing the power spectral density  $\mathfrak{S}(\omega)$  of the considered random process in Eq. (2), where the modal decomposition of the covariance kernel operator is performed, the following equation is obtained:

$$\begin{aligned} \int_0^L C(s, t)\varphi_i(s)ds &= \int_0^L \int_{-\infty}^{\infty} \mathfrak{S}(\omega)e^{i\omega(s-t)}\varphi_i(s)d\omega ds \\ &= \int_{-\infty}^{\infty} \mathfrak{S}(\omega)\bar{\Phi}_i(\omega)e^{-i\omega t}d\omega = \lambda_i\varphi_i(t) \end{aligned} \quad (48)$$

where  $\Phi_i(\omega) = \int_0^L \varphi_i(s)e^{-i\omega s}ds$  and the over bar represents the complex conjugate. From the above equation, the following property is derived:

$$\int_{-\infty}^{\infty} \mathfrak{S}(\omega)\bar{\Phi}_i(\omega)e^{-i\omega t}d\omega = \lambda_i\varphi_i(t) = \lambda_i \int_{-\infty}^{\infty} \bar{\Phi}_i(\omega)e^{-i\omega t}d\omega \implies \mathfrak{S}(\omega)\bar{\Phi}_i(\omega) = \lambda_i\bar{\Phi}_i(\omega) \quad (49)$$

By the same way it can be derived that  $\mathfrak{S}(\omega)\Phi_i(\omega) = \lambda_i\Phi_i(\omega)$ .

By using the orthonormality and the completeness of functions  $\varphi_i(s)$  basis set (Eq. (4)), the following property is obtained:

$$\int_{-\infty}^{\infty} \mathfrak{S}(\omega)\Phi_j(\omega)\bar{\Phi}_i(\omega)d\omega = \delta_{ij}\lambda_i \quad (50)$$

Note that the functions  $\Phi_i(\omega)$  are a complete set of orthonormal basis functions equipped with the following inner product:

$$\int_{-\infty}^{\infty} \Phi_i(\omega)\bar{\Phi}_j(\omega)d\omega = \delta_{ij} \quad (51)$$

$\mathfrak{S}(\omega)$  denotes the power spectral density that can be introduced in Eq. (47). By using the properties derived above, one can assert that:

$$\begin{aligned} K_{ij}^2 &= \left| \int_0^L \int_0^L C(s, t + L) \frac{\varphi_i(s)\varphi_j(t)}{\sqrt{\lambda_i\lambda_j}} ds dt \right|^2 \\ &= \left| \int_0^L \int_0^L \int_{-\infty}^{\infty} \mathfrak{S}(\omega) e^{i\omega(s-t-L)} \frac{\varphi_i(s)\varphi_j(t)}{\sqrt{\lambda_i\lambda_j}} d\omega ds dt \right|^2 \\ &= \left| \int_{-\infty}^{\infty} \mathfrak{S}(\omega) \frac{\Phi_j(\omega)\bar{\Phi}_i(\omega)}{\sqrt{\lambda_i\lambda_j}} e^{-i\omega L} d\omega \right|^2 = \left| \int_{-\infty}^{\infty} \Phi_j(\omega)\bar{\Phi}_i(\omega) e^{-i\omega L} d\omega \right|^2 \end{aligned} \quad (52)$$

By the application of the Bessel's inequality, the expression in Eq. (47) is upper bounded for each  $i = 1, \dots, N$ :

$$\sum_{j=1}^N K_{ij}^2 = \sum_{j=1}^N \left| \int_{-\infty}^{\infty} \bar{\Phi}_i(\omega) e^{-i\omega L} \Phi_j(\omega) d\omega \right|^2 \leq \int_{-\infty}^{\infty} \bar{\Phi}_i(\omega) e^{-i\omega L} e^{i\omega L} \Phi_i(\omega) d\omega = 1 \quad (53)$$

Note that the above inequality becomes an equality when  $N \rightarrow \infty$  leading to the Parseval's theorem.

Finally, by applying this bound to Eq. (46), it follows that, for every considered correlation function:

$$\frac{1}{r} \sum_{i=1}^r \sum_{j=1}^N K_{ij}^2 \leq 1, \quad \text{with } r = 1, \dots, N \quad (54)$$

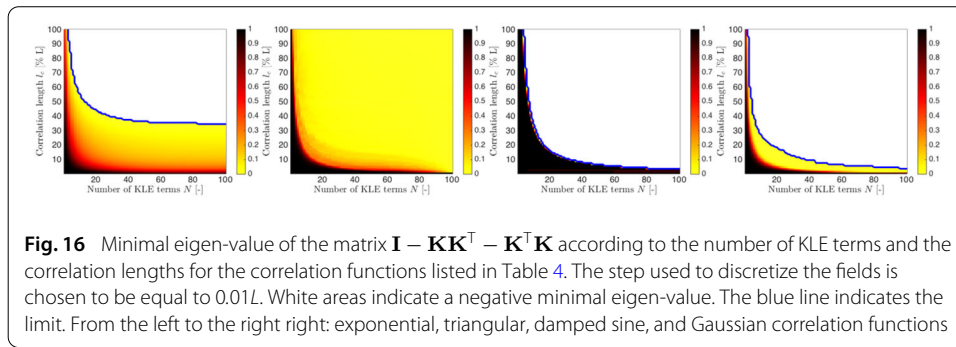
Note that this upper bound, derived through the application of the Bessel's inequality in Eq. (53), is strictly lower than one for every  $i = 1, \dots, N$  except when  $N \rightarrow \infty$  or when the correlation function is constant. In this latter particular case the matrix  $\mathbf{K}^T \mathbf{K}$  has one only diagonal element equal to one and all the other elements equal to zero, leading to a null determinant for the matrix  $\mathbf{K}^T \mathbf{K}$  (in this case there is not a unique Cholesky decomposition). In practice, a constant correlation function means that each sample of the random field is constant in the whole domain.

For all the other correlation functions, the application of the bound of Eq. (54) in Eq. (45), with a strict inequality, states that all the leading minors of the matrix  $\mathbf{I} - \mathbf{K}^T \mathbf{K}$  are greater than zero, satisfying the Sylvester's criterion (Eq. 44) and meaning the positive definiteness of that matrix.

### Appendix C: Influence of the sub-domain size

The Cholesky decomposition in Eqs. (12)–(30) implies a positive definiteness of the matrix to decompose. After imposing a correlation between two KLE sets using the matrix  $\mathbf{K}$  (Eq. 11), the matrices  $\mathbf{L}$  and  $\mathbf{R}$  guarantee that, in each sub-domain, the KLE coefficients are normalized and uncorrelated, i.e. that the generated random field has a prescribed correlation structure.

In "Appendix B", it is shown that the Cholesky decomposition is always possible when the sub-domain is only conditioned from one side. When trying to correlate the KLE coefficients with the left and the right sets, as in "1D processes generation parallel computing"



**Fig. 16** Minimal eigen-value of the matrix  $\mathbf{I} - \mathbf{K}\mathbf{K}^T - \mathbf{K}^T\mathbf{K}$  according to the number of KLE terms and the correlation lengths for the correlation functions listed in Table 4. The step used to discretize the fields is chosen to be equal to  $0.01L$ . White areas indicate a negative minimal eigen-value. The blue line indicates the limit. From the left to the right right: exponential, triangular, damped sine, and Gaussian correlation functions

section, the positive definiteness of the matrix  $\mathbf{I} - \mathbf{K}\mathbf{K}^T - \mathbf{K}^T\mathbf{K}$  needs to be checked. In B it has been shown that  $\sum_{j=1}^N K_{ij}^2 \leq 1$  for every  $i$  (Eq. 53). When imposing a correlation from two sides, the same constraint is twice imposed. The latter sum must be lower than 0.5 to ensure the positive definiteness condition:

$$2 \sum_{j=1}^N K_{ij}^2 \leq 1 \tag{55}$$

Adding more terms in the KLE increases the value of the sum that can exceed the value of 0.5. In fact, when more KLE terms are added for representing the random field, although the correlation structure is better represented (the KLE truncation error is reduces), a larger number of constraints act on the KLE coefficients set.

The correlation length is another limitation. For a given correlation structure and truncation error, the more correlated (lower ratio  $\frac{L}{l_c}$ ) the process is, the fewer number of KLE terms are needed [30]. This means that, for highly correlated processes, few terms are needed to represent the correlation and therefore the left and right constraints conditioning constraint: the condition in Eq. (55) can be exceeded with few terms.

The minimal eigen-value of the matrix  $\mathbf{I} - \mathbf{K}\mathbf{K}^T - \mathbf{K}^T\mathbf{K}$  is shown in Fig. 16 according to the number of KLE terms and the correlation lengths (parameter  $l_c$ ) for the correlation functions listed in Table 4. The step used to discretize the fields is chosen to be equal to  $0.01L$ . For shorter correlation lengths the conditioning operation is possible since the eigen-values are positive. For the damped sine correlation a negative eigen-value appears suddenly when the number of terms is increased. This is due to the periodic behaviour of the correlation function. Note that, for the triangular correlation, the conditioning is always possible, even when the correlation length is equal to the size of the sub-domain.

In order to prevent this problem, a larger size of the sub-domain can be chosen (increasing the ratio  $\frac{L}{l_c}$ ).

**Publisher’s Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 3 April 2018 Accepted: 24 July 2018

Published online: 02 August 2018

**References**

1. Borgman LE. Ocean wave simulation for engineering design. Tech. Rep. HEL-9-13. Berkeley: California University Berkley Hydraulic Engineering Laboratory; 1967.

2. Shinozuka M. Simulation of multivariate and multidimensional random processes. *J Acoust Soc Am.* 1970;19(1):357–68. <https://doi.org/10.1121/1.1912338>.
3. Shinozuka M, Jan CM. Digital simulation of random processes and its applications. *J Sound Vibrat.* 1972;25(1):111–28. [https://doi.org/10.1016/0022-460X\(72\)90600-1](https://doi.org/10.1016/0022-460X(72)90600-1).
4. Shinozuka M, Deodatis G. Simulation of stochastic processes by spectral representation. *Appl Mech Rev.* 1991;44(4):191–204. <https://doi.org/10.1115/1.3119501>.
5. Grigoriu M. On the spectral representation method in simulation. *Probab Eng Mech.* 1993;8:75–90. [https://doi.org/10.1016/0266-8920\(93\)90002-D](https://doi.org/10.1016/0266-8920(93)90002-D).
6. Yang JN. On the normality and accuracy of simulated random processes. *J Sound Vibrat.* 1973;26(3):417–28. [https://doi.org/10.1016/S0022-460X\(73\)80196-8](https://doi.org/10.1016/S0022-460X(73)80196-8).
7. Matheron G. The intrinsic random functions and their applications. *Adv Appl Probab.* 1973;5(3):439–68. <https://doi.org/10.2307/1425829>.
8. Mantoglou A, Wilson JL. The turning bands method for simulation of random-fields using line generation by a spectral method. *Water Resour Res.* 1982;18(5):1379–94. <https://doi.org/10.1029/WR018i005p01379>.
9. De Carvalho Paludo L, Bouvier V, Cottreau R, Clouteau D. Efficient parallel generation of random field of mechanical properties for geophysical application. In: 12e Colloque national en calcul des structures. Giens: CSMA; 2015. p. 1–4.
10. Whittle P. On stationary processes on the plane. *Biometrika.* 1954;41:434–49. <https://doi.org/10.2307/2332724>.
11. Box GEP, Jenkins G. *Time series analysis: forecasting & control.* San Francisco: Holden-Day; 1970.
12. Gersch W, Yonemoto J. Synthesis of multivariate random vibration systems: a two-stage least squares AR-MA model approach. *J Sound Vibrat.* 1977;52(4):553–65. [https://doi.org/10.1016/0022-460X\(77\)90370-4](https://doi.org/10.1016/0022-460X(77)90370-4).
13. Samaras E, Shinozuka M, Tsurui A. ARMA representation of random processes. *J Eng Mech.* 1985;111:449–61. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1985\)111:3\(449\)](https://doi.org/10.1061/(ASCE)0733-9399(1985)111:3(449)).
14. Rue H. Fast sampling of Gaussian Markov random fields. *J R Stat Soc B.* 2001;63(2):325–38. <https://doi.org/10.1111/1467-9868.00288>.
15. Aune E, Eidsvik J, Pokern Y. Iterative numerical methods for sampling from high dimensional Gaussian distributions. *Stat Comput.* 2013;23(4):501–21. <https://doi.org/10.1007/s11222-012-9326-8>.
16. Chow E, Saad Y. Preconditioned Krylov subspace methods for sampling multivariate Gaussian distributions. *SIAM J Sci Comput.* 2014;36(2):A588–608. <https://doi.org/10.1137/130920587>.
17. Fang J, Tacher L. An efficient and accurate algorithm for generating spatially-correlated random fields. *Commun Numer Methods Eng.* 2003;19(10):801–8. <https://doi.org/10.1002/cnm.621>.
18. Loève M. *Probab Theory.* 4th ed. Berlin: Springer; 1977.
19. Fukunaga K, Koontz WLG. Representation of random processes using the finite Karhunen–Loève expansion. *Inf Control.* 1970;16(1):85–101. [https://doi.org/10.1016/S0019-9958\(70\)80043-2](https://doi.org/10.1016/S0019-9958(70)80043-2).
20. Delhomme JF. Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach. *Water Resour Res.* 1979;15(2):269–80. <https://doi.org/10.1029/WR015i002p00269>.
21. Yamazaki F, Shinozuka M. Simulation of Stochastic fields by statistical preconditioning. *J Eng Mech.* 1990;116:268–87. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1990\)116:2\(268\)](https://doi.org/10.1061/(ASCE)0733-9399(1990)116:2(268)).
22. Spanos PD, Ghanem R. Stochastic finite element expansion for random media. *J Eng Mech.* 1989;115(5):1035–53.
23. Ghanem R, Spanos PD. Polynomial chaos in stochastic finite elements. *J Appl Mech.* 1990;57(1):197. <https://doi.org/10.1115/1.2888303>.
24. Ghanem R, Spanos PD. A stochastic Galerkin expansion for nonlinear random vibrations analysis. *Probab Eng Mech.* 1993;8:255–64.
25. Li CC, Der Kiureghian A. Optimal discretization of random fields. *J Eng Mech.* 1993;119:1136–54. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1993\)119:6\(1136\)](https://doi.org/10.1061/(ASCE)0733-9399(1993)119:6(1136)).
26. Betz W, Papaioannou I, Straub D. Numerical methods for the discretization of random fields by means of the Karhunen–Loève expansion. *Comput Methods Appl Mech Eng.* 2014;271:109–29. <https://doi.org/10.1016/j.cma.2013.12.010>.
27. Unser M. On the approximation of the discrete Karhunen–Loeve transform for stationary processes. *Signal Proces.* 1984;7:231–49. [https://doi.org/10.1016/0165-1684\(84\)90002-1](https://doi.org/10.1016/0165-1684(84)90002-1).
28. Dietrich CR, Newsam GN. Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM J Sci Comput.* 1997;18(4):1088–107. <https://doi.org/10.1137/S1064827592240555>.
29. Fancourt CL, Principe JC. On the relationship between the karhunen-loeve transform and the prolate spheroidal wave functions. In: 2000 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No.00CH37100), vol. 1. 2000. p. 261–264. <https://doi.org/10.1109/ICASSP.2000.861937>
30. Huang SP, Quek ST, Phoon KK. Convergence study of the truncated Karhunen–Loeve expansion for simulation of stochastic processes. *Int J Numer Methods Eng.* 2001;52(9):1029–43. <https://doi.org/10.1002/nme.255>.
31. Zhang J, Ellingwood B. Orthogonal series expansions of random fields in reliability analysis. *J Eng Mech.* 1994;120(12):2660–77. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1994\)120:12\(2660\)](https://doi.org/10.1061/(ASCE)0733-9399(1994)120:12(2660)).
32. Grigoriu M. Parametric models of nonstationary Gaussian processes. *Probab Eng Mech.* 1995;10(95):95–102. [https://doi.org/10.1016/0266-8920\(95\)00008-M](https://doi.org/10.1016/0266-8920(95)00008-M).
33. Cho H, Venturi D, Karniadakis GE. Karhunen–Loève expansion for multi-correlated stochastic processes. *Probab Eng Mech.* 2013;34:157–67. <https://doi.org/10.1016/j.proengmech.2013.09.004>.
34. Rosemblatt M. Remarks on a multivariate transformation. *Ann Math Stat.* 1952;23:470–2. <https://doi.org/10.1214/aoms/1177729394>.
35. Grigoriu M. Simulation of stationary non-Gaussian translation processes. *J Eng Mech.* 1998;124:121–6. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1998\)124:2\(121\)](https://doi.org/10.1061/(ASCE)0733-9399(1998)124:2(121)).
36. Grigoriu M. Non-Gaussian models for stochastic mechanics. *Probab Eng Mech.* 2000;15:15–23. [https://doi.org/10.1016/S0266-8920\(99\)00005-3](https://doi.org/10.1016/S0266-8920(99)00005-3).
37. Puig B, Poirion F, Soize C. Non-Gaussian simulation using Hermite polynomial expansion: convergences and algorithms. *Probab Eng Mech.* 2002;17(3):256–64. [https://doi.org/10.1016/S0266-8920\(02\)00010-3](https://doi.org/10.1016/S0266-8920(02)00010-3).

38. Guillemot J, Soize C. Itô SDE-based generator for a class of non-Gaussian vector-valued random fields in uncertainty quantification. *SIAM J Sci Comput.* 2014;36(6):A2763–86. <https://doi.org/10.1137/130948586>.
39. Davis HT, Thomson KT. *Linear algebra and linear operators in engineering: with applications in mathematica.* San Diego: Academic Press; 2000.
40. Perrin G. *Random fields and associated statistical inverse problems for uncertainty quantification.* Doctoral thesis, Université Paris-Est; 2013.
41. Phoon KK, Huang HW, Quek ST. Simulation of strongly non-Gaussian processes using Karhunen–Loève expansion. *Probab Eng Mech.* 2005;20(2):188–98. <https://doi.org/10.1016/j.probengmech.2005.05.007>.
42. Li LB, Phoon KK, Quek ST. Comparison between Karhunen–Loève expansion and translation-based simulation of non-Gaussian processes. *Comput Struct.* 2007;85(5–6):264–76. <https://doi.org/10.1016/j.compstruc.2006.10.010>.
43. Makhoul J. On the eigenvectors of symmetric Toeplitz matrices. *IEEE Trans Acoustics Speech Signal Proces.* 1981;29(4):868–72. <https://doi.org/10.1109/TASSP.1981.1163635>.
44. Van Trees H, Bell K. *Detection, estimation, and modulation theory part I.* New York: Wiley; 1968.
45. Le Maître OP, Knio OM. *Spectral methods for uncertainty quantification*, chap. 2. Berlin: Springer; 2010. p. 21.
46. Hasofer AM, Ghahreman S. On the slepian process of a random Gaussian trigonometric polynomial. *IEEE Trans Inf Theory.* 1989;35(4):868–73. <https://doi.org/10.1109/18.32163>.
47. Kloeden PE, Platen E. *Numerical solution of stochastic differential equations.* Berlin: Springer; 1992. <https://doi.org/10.2977/prims/1166642153>.
48. Goto H, Toki K. *Structural response to nonstationary random excitation.* In: 4th world conference on earthquake engineering. Santiago: Chile; 1969. p. 130–44.
49. Ramadan O, Novak M. Simulation of spatially incoherent random ground motions. *J Eng Mech.* 1993;119(5):997–1016. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1993\)119:5\(997\)](https://doi.org/10.1061/(ASCE)0733-9399(1993)119:5(997)).
50. Deodatis G. Non-stationary stochastic vector processes: seismic ground motion applications. *Probab Eng Mech.* 1996;11(3):149–67. [https://doi.org/10.1016/0266-8920\(96\)00007-0](https://doi.org/10.1016/0266-8920(96)00007-0).
51. Priestley MB. Evolutionary spectra and non-stationary processes. *J R Stat S B (Methodological).* 1965;27(2):204–37.
52. Li Y, Kareem A. Simulation of multivariate nonstationary random processes by FFT. *J Eng Mech.* 1991;117(5):1037–58. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1991\)117:5\(1037\)](https://doi.org/10.1061/(ASCE)0733-9399(1991)117:5(1037)).
53. Deodatis G, Shinozuka M. Auto-regressive model for nonstationary stochastic processes. *J Eng Mech.* 1988;114(11):1995–2012. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1988\)114:11\(1995\)](https://doi.org/10.1061/(ASCE)0733-9399(1988)114:11(1995)).

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---