Comput. Methods Appl. Mech. Engrg. 271 (2014) 109–129

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

Numerical methods for the discretization of random fields by means of the Karhunen–Loève expansion

Wolfgang Betz *, Iason Papaioannou, Daniel Straub

Engineering Risk Analysis Group, Technische Universität München, Germany

ARTICLE INFO

Article history:
Received 12 June 2013
Received in revised form 15 October 2013
Accepted 20 December 2013
Available online 2 January 2014

Keywords:
Random field discretization
Karhunen–Loève expansion
Nyström method
Collocation method
Galerkin method
Finite cell method

ABSTRACT

The computational efficiency of random field representations with the Karhunen–Loève (KL) expansion relies on the solution of a Fredholm integral eigenvalue problem. This contribution compares different methods that solve this problem. Focus is put on methods that apply to arbitrary shaped domains and arbitrary autocovariance functions. These include the Nyström method as well as collocation and Galerkin projection methods. Among the Galerkin methods, we investigate the finite element method (FEM) and propose the application of the finite cell method (FCM). This method is based on an extension to the FEM but avoids mesh generation on domains of complex geometric shape. The FCM was originally presented in Parvizian et al. (2007) [17] for the solution of elliptic boundary value problems. As an alternative to the L^2-projection of the covariance function used in the Galerkin method, $H^{1/2}$-projection and discrete projection are investigated. It is shown that the expansion optimal linear estimation (EOLE) method proposed in Li and Der Kiureghian (1993) [18] constitutes a special case of the Nyström method. It is found that the EOLE method is most efficient for the numerical solution of the KL expansion. The FEM and the FCM are more efficient than the EOLE method in evaluating a realization of the random field and, therefore, are suitable for problems in which the time spent in the evaluation of random field realizations has a major contribution to the overall runtime – e.g., in finite element reliability analysis.

1. Introduction

Many engineering problems require the modeling of uncertain input parameters with inherent spatial variability. These include soil parameters and groundwater heights in geotechnical engineering, wind loads and snow loads in structural engineering, and the amount of precipitation and evaporation in hydrology. This type of uncertainty is modeled by means of random fields. A random field represents a random quantity at each point of a continuous domain, and, thus, consists of an infinite number of random variables. For computational purposes, the random field has to be expressed using a finite number of random variables. This step is referred to as random field discretization.

Various methods for random field discretization have been published in literature. A comprehensive overview of random field discretization methods is given in [1]. The efficiency of a method for random field discretization depends on its ability to approximate the original random field accurately with a minimum number of random variables. Accuracy is to be defined with respect to a certain error measure such as the mean square error.

Series expansion methods approximate the random field by a finite sum of products of deterministic spatial functions and random variables. The Karhunen–Loève (KL) expansion is optimal among series expansion methods in the global mean

* Corresponding author. Tel.: +49 89 289 23072.
E-mail addresses: wolfgang.betz@tum.de (W. Betz), iason.papaioannou@tum.de (I. Papaioannou), straub@tum.de (D. Straub).

0045-7825/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cma.2013.12.010

square error with respect to the number of random variables in the representation [2]. Consequently, it has received much attention in literature. The KL expansion was introduced in the engineering community by Spanos and Ghanem [3]. The expansion requires the solution of a Fredholm integral eigenvalue problem (IEVP), whose integral kernel is the autocovariance function of the field. Analytical solutions of the IEVP can be obtained only for specific types of autocovariance functions defined on rectangular domains. For random fields with arbitrary autocovariance functions defined on domains of complex geometrical shape, the solution of the IEVP needs to be approximated numerically. An overview of the numerical solution of Fredholm integral equations is given in [4].

For the numerical approximation of the KL expansion, approaches that are based on the Galerkin method are used most often in literature. On one-dimensional domains, the Galerkin scheme is often applied in a spectral sense, i.e., the basis functions are spanned over the entire domain. The convergence behavior of this approach is investigated in [5] for stationary and non-stationary problems and different covariance functions using polynomials with a degree up to ten as basis functions. Gutiérrez et al. [6] compares the Legendre polynomials with trigonometric basis functions. The wavelet-Galerkin method is an alternative Galerkin approach. It is applied in [7–9] for the discretization of the IEVP on one-dimensional domains. For random fields defined on two- and three-dimensional domains, the finite element method (FEM) is often used for the discretization of the IEVP. The use of the FEM for the approximate solution of the KL expansion was suggested by Ghanem and Spanos [2]. The FEM is a Galerkin approach that requires a spatial discretization of the domain, typically called finite element mesh. The method can cope with domains of arbitrary geometric shape. The convergence behavior of the FEM was investigated by Papaioannou [10] for two-dimensional domains. For the computation of the matrix eigenvalue problem in the FEM, a generalized fast multi-pole Krylov eigen-solver is used in [11,12]. On two- and especially three-dimensional random field domains, the computational costs of setting up the matrix eigenvalue problem and obtaining its solution become expensive. This problem was tackled in [13,14] by application of $\mathcal{H}$-matrices [15] for the decomposition of the covariance matrix. If the domain has a complex geometric shape, the generation of a finite element mesh is an involved task and can be time consuming. Therefore, approaches that avoid mesh generation might be preferred. Papaioannou [10] proposed a spectral Galerkin approach that deals with geometrically complex domains by embedding the actual domain in a domain of simple geometric shape and using the latter to solve the IEVP. Furthermore, a Galerkin approach employing meshless basis functions is proposed in [16]. Besides the Galerkin approach, the Nyström method [4] is an alternative technique for the discretization of the IEVP that also avoids mesh generation. In this method, the domain is represented by a set of points associated with weighting factors. In addition to the Galerkin approach and the Nyström method, the collocation method should be mentioned. This method is investigated in [6] for a one-dimensional domain.

This paper reviews different methods for the numerical solution of the IEVP in the KL expansion. Focus is put on methods that can cope with random field domains of arbitrary geometric shape and that do not require the autocovariance function to be isotropic or homogeneous. These include the Nyström method, collocation and Galerkin methods. For the Galerkin methods we focus on the FEM and propose the application of the finite cell method (FCM). In the FCM [17], the actual domain is embedded in a domain of simple geometric shape and the mesh is generated on the simple geometry. Higher-order basis functions ensure a fast rate of convergence of the approach. We present the principles of the FCM for the solution of the IEVP and discuss its implementation. Furthermore, we show that if equal weights are used in the Nyström method, the derived approximate KL expansion is equivalent to the expansion optimal linear estimation (EOLE) method, which is a series expansion method for discretization of random fields, proposed in [18].

The remainder of the paper is organized as follows: In Section 2, the KL expansion is presented along with its fundamental properties. The numerical solution of the IEVP is discussed in Section 3. Special emphasis is given to the description of the FCM and the treatment of the arising discontinuous integrands. Section 4 is devoted to numerical examples. In this section, the considered methods are compared with respect to their numerical efficiency. The efficiency is measured in terms of the computational efforts required to obtain a random field approximation, and in terms of the computational costs involved in the evaluation of a realization of the random field. The paper closes with the concluding remarks in Section 5.

2. Karhunen–Loève expansion

2.1. Definition of random fields

A continuous random field $H(\mathbf{x},\theta)$ may be loosely defined as a random function that describes a random quantity at each point $\mathbf{x}\in\Omega$ of a continuous domain $\Omega\subset\mathbb{R}^d, d\in\mathbb{N}_{>0}$. $\theta\in\Theta$ is a coordinate in the sample space Θ, and (Θ,F,P) is a complete probability space. If the random quantity attached to each point $\mathbf{x}$ is a random variable, the random field is said to be *univariate* or real-valued. If the random quantity is a random vector, the field is called *multivariate*. The dimension d of a random field is the dimension of its topological space Ω. One usually distinguishes between a *one-* and a *multidimensional* random field.

The field is said to be *Gaussian* if the distribution of $(H(\mathbf{x}_1,\theta),\dots,H(\mathbf{x}_n,\theta))$ is jointly Gaussian for any $(\mathbf{x}_1,\dots,\mathbf{x}_n)\in\Omega$ and any $n\in\mathbb{N}_{>0}$. A Gaussian field is completely defined by its mean function $\mu:\Omega\to\mathbb{R}$ and autocovariance function $\mathrm{Cov}:\Omega\times\Omega\to\mathbb{R}$. The autocovariance function can be expressed as $\mathrm{Cov}(\mathbf{x},\mathbf{x}')=\sigma(\mathbf{x})\cdot\sigma(\mathbf{x}')\cdot\rho(\mathbf{x},\mathbf{x}')$, where $\sigma:\Omega\to\mathbb{R}$ is the standard deviation function of the random field and $\rho:\Omega\times\Omega\to[-1,1]$ is its autocorrelation coefficient function.

The discussion in this work is restricted to univariate multidimensional random fields. We focus on Gaussian random fields and discuss a special case of non-Gaussian fields that can be expressed as functions of Gaussian fields.

2.2. Error measures for random field discretization

The approximation $\hat{H}(\cdot)$ of a continuous random field $H(\cdot)$ by a finite set of random variables $\{\chi_i, i = 1, \ldots, M\}$ with $M \in \mathbb{N}_{>0}$ is referred to as *random field discretization*. The approximation error $\varepsilon_H(\mathbf{x}, \theta)$ is defined as the difference between the original field and its approximation, i.e., $\varepsilon_H(\mathbf{x}, \theta) = H(\mathbf{x}, \theta) - \hat{H}(\mathbf{x}, \theta)$. The expectation of the squared approximation error is called the *mean square error*. Integration of the mean square error over the domain Ω gives the *global mean square error* [2]:

$$\overline{\varepsilon}_H^2 = \int_\Omega \mathrm{E}[(\varepsilon_H(\mathbf{x}, \theta))^2]\,\mathrm{d}\mathbf{x} \tag{1}$$

An alternative error measure for random field discretization is the normalized variance of the approximation error, denoted $\varepsilon_\sigma(\mathbf{x})$ [18]:

$$\varepsilon_\sigma(\mathbf{x}) = \frac{\mathrm{Var}[H(\mathbf{x}, \theta) - \hat{H}(\mathbf{x}, \theta)]}{\mathrm{Var}[H(\mathbf{x}, \theta)]} \tag{2}$$

$\varepsilon_\sigma(\mathbf{x})$ is called *error variance* in literature. The corresponding global error measure, namely the *mean error variance*, is defined as the weighted integral [1]:

$$\overline{\varepsilon}_\sigma = \frac{1}{|\Omega|} \int_\Omega \varepsilon_\sigma(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{3}$$

where $|\Omega| = \int_\Omega \mathrm{d}\mathbf{x}$. Another global error measure is the supremum norm of the error variance [18]: $\hat{\varepsilon}_\sigma = \sup_{\mathbf{x}\in\Omega} |\varepsilon_\sigma(\mathbf{x})|$. It reflects the maximum point-wise error in the domain. Global error measures are applied to compare random field discretization methods and to quantify the overall quality of a random field approximation. It was noted in [19,18] that different global error measures might favor different discretization methods. In this paper we consider only global error measures that average point-error measures over the domain. However, for some fields of application, global error measures that are based on the supremum norm might be of relevance as well.

It is convenient to assume that the mean of the random field can be represented exactly. In this case, the expectation of the approximation error is zero, and the expectation of the squared approximation error is equivalent to the variance of the approximation error, i.e., $\mathrm{E}[(\varepsilon_H(\mathbf{x}, \theta))^2] = \mathrm{Var}[\varepsilon_H(\mathbf{x}, \theta)]$. Consequently, the error variance is proportional to the mean square error. If the standard deviation of the field is constant on the domain Ω, i.e., $\sigma = \sigma(\mathbf{x})\,\forall \mathbf{x} \in \Omega$, the global mean square error can be expressed in terms of the mean error variance as:

$$\overline{\varepsilon}_H^2 = |\Omega| \cdot \sigma^2 \cdot \overline{\varepsilon}_\sigma \tag{4}$$

2.3. KL expansion of random fields

The Karhunen–Loève expansion is a series expansion method for the representation of the random field. The expansion is based on a spectral decomposition of the autocovariance function of the field. It states that a second-order random field can be represented exactly by the following expansion [20,21]:

$$H(\mathbf{x}, \theta) = \mu(\mathbf{x}) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\, \varphi_i(\mathbf{x})\, \xi_i(\theta) \tag{5}$$

where $\mu(\mathbf{x})$ is the mean function of the field, $\xi_i(\theta) : \Theta \to \mathbb{R}$ are standard uncorrelated random variables, and $\lambda_i \in [0, \infty), \varphi_i : \Omega \to \mathbb{R}$ are the eigenvalues and eigenfunctions of the autocovariance kernel obtained from solving the homogeneous Fredholm integral equation of the second kind:

$$\int_\Omega \mathrm{Cov}(\mathbf{x}, \mathbf{x}')\, \varphi_i(\mathbf{x}')\,\mathrm{d}\mathbf{x}' = \lambda_i \varphi_i(\mathbf{x}) \tag{6}$$

In this context, the autocovariance function $\mathrm{Cov}(\mathbf{x}, \mathbf{x}')$ is also referred to as kernel function. Any valid covariance function is a bounded, symmetric and positive semi-definite kernel [22]. Moreover, a continuous kernel function is assumed. Note that the kernel does not have to be stationary. According to Mercer's theorem, the eigenvalues λ_i are nonnegative, the eigenfunctions corresponding to positive eigenvalues are continuous and orthogonal to each other, and the kernel function can be written as the uniformly convergent expansion $\mathrm{Cov}(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i\, \varphi_i(\mathbf{x})\, \varphi_i(\mathbf{x}')$, where the eigenfunctions in the expression are normalized. Consequently, the eigenfunctions must be orthonormal to each other, i.e., $\int_\Omega \varphi_i(\mathbf{x})\, \varphi_j(\mathbf{x})\,\mathrm{d}\mathbf{x} = \delta_{ij}$, where δ_{ij} is one if $i = j$ and zero otherwise. Moreover, they form a complete basis of the space $L^2(\Omega)$ of square integrable functions on Ω.

If the random field $H(\mathbf{x}, \theta)$ is Gaussian, then $\xi_i(\theta)$ are independent standard normal random variables [2]. In any other case, the joint distribution of $\xi_i(\theta)$ is almost impossible to obtain. Hence, the KL expansion is mainly applicable to the discretization of Gaussian fields.

The direct modeling of non-Gaussian random fields by means of the KL expansion was discussed by Phoon et al. [23]. The authors proposed an iterative framework to simulate non-stationary non-Gaussian processes. The procedure was refined in

[24] for highly skewed non-Gaussian processes. Moreover, non-Gaussian fields are commonly modeled by combining the KL expansion with the polynomial chaos expansion. Ghanem [25] proposed a general framework in which the non-Gaussian field is projected onto an orthogonal polynomial basis with argument an underlying Gaussian field that is then discretized by the KL expansion. Matthies and Keese [26] proposed to perform the KL expansion of the non-Gaussian field and project the random variables involved in the expansion to an underlying independent Gaussian random variable space. In Section 2.5, we discuss the treatment of a special case of non-Gaussian random fields within the context of the KL expansion.

2.4. Truncated KL expansion

The KL expansion can be approximated by sorting the eigenvalues λ_i and the corresponding eigenfunctions $\varphi_i(\mathbf{x})$ in a descending order and truncating the expansion after M terms:

$$\tilde{H}(\mathbf{x},\theta) = \mu(\mathbf{x}) + \sum_{i=1}^{M} \sqrt{\lambda_i}\,\varphi_i(\mathbf{x})\,\xi_i(\theta) \tag{7}$$

For fixed M, the resulting random field approximation $\tilde{H}(\mathbf{x},\theta)$ is optimal among series expansion methods with respect to the global mean square error (Eq. (1)) [2]. The variance of $\tilde{H}(\mathbf{x},\theta)$ is given as

$$\mathrm{Var}[\tilde{H}(\mathbf{x},\theta)] = \sum_{i=1}^{M} \lambda_i\,\varphi_i^2(\mathbf{x}) \tag{8}$$

In case of the truncated KL expansion, the error variance introduced in Eq. (2) can be expressed as [1]:

$$\varepsilon_{\sigma,\mathrm{KL}}(\mathbf{x}) = 1 - \frac{\sum_{i=1}^{M} \lambda_i\,\varphi_i^2(\mathbf{x})}{\sigma^2(\mathbf{x})} \tag{9}$$

wherein the numerator in the fraction represents the variance of the truncated field. Eq. (9) can be derived by expressing $H(\mathbf{x},\theta)$ by its KL expansion and using the orthonormality of the random variables ξ_i. From Eq. (9) it can be deduced that the truncated KL expansion always underestimates the true variability of the original random field. This property of the KL expansion was also discussed in [1]. Moreover, as pointed out in [9], the truncated KL expansion of homogeneous random fields is only approximately homogeneous, since the standard deviation function of the truncated field will always vary in space. The mean error variance is given as:

$$\overline{\varepsilon}_{\sigma,\mathrm{KL}} = 1 - \frac{1}{|\Omega|} \sum_{i=1}^{M} \lambda_i \int_{\Omega} \frac{\varphi_i^2(\mathbf{x})}{\sigma^2(\mathbf{x})}\,\mathrm{d}\mathbf{x} \tag{10}$$

The equation of the mean square error can be transformed to $\mathrm{E}[(\varepsilon_H(\mathbf{x},\theta))^2] = \sigma^2(\mathbf{x}) - \sum_{i=1}^{M} \lambda_i\,\varphi_i^2(\mathbf{x})$ using the orthonormality of the random variables ξ_i. The global mean square error reads [1]:

$$\overline{\varepsilon}^2_{H,\mathrm{KL}} = \int_{\Omega} \sigma^2(\mathbf{x})\,\mathrm{d}\mathbf{x} - \sum_{i=1}^{M} \lambda_i \tag{11}$$

If the standard deviation of the field is constant, the equation for the mean error variance, Eq. (10), reduces to [1]:

$$\overline{\varepsilon}_{\sigma,\mathrm{KL}} = 1 - \frac{1}{|\Omega|}\,\frac{1}{\sigma^2} \sum_{i=1}^{M} \lambda_i \tag{12}$$

For this special case, the truncated KL expansion is also optimal with respect to the mean error variance.

2.5. Non-Gaussian translation random fields

General non-Gaussian random fields are not suitable to be expressed by means of Gaussian random fields. If a non-Gaussian random field belongs to the class of translation fields, it can be expressed in terms of a Gaussian random field through a nonlinear mapping of the form $H_{\mathrm{transl}.}(\mathbf{x},\theta) = \mathrm{g}(H(\mathbf{x},\theta))$, where $H_{\mathrm{transl}.}(\mathbf{x},\theta)$ represents the non-Gaussian random field defined in terms of the Gaussian field $H(\mathbf{x},\theta)$ and the strictly increasing nonlinear mapping $\mathrm{g}:\mathbb{R}\to\mathbb{R}$ [27]. Discretization of the field $H_{\mathrm{transl}.}(\mathbf{x},\theta)$ is achieved by replacing $H(\mathbf{x},\theta)$ by its KL expansion $\tilde{H}(\mathbf{x},\theta)$ and applying $\tilde{H}_{\mathrm{transl}.}(\mathbf{x},\theta) = \mathrm{g}(\tilde{H}(\mathbf{x},\theta))$. However, it cannot be confirmed that the transformed field $\tilde{H}_{\mathrm{transl}.}(\mathbf{x},\theta)$ inherits the optimality property that the Gaussian random field approximation $\tilde{H}(\mathbf{x},\theta)$ may possess [18]. All random field models used in reliability analysis and probabilistic mechanics belong essentially to the class of translation fields.

A subclass of translation random fields constitute fields where the Nataf multivariate distribution [28] is applied to perform the nonlinear mapping $\mathrm{g}(\cdot)$. For this class of translation random fields, the underlying Gaussian field has zero mean and unit variance. Its autocorrelation coefficient function $\rho(\mathbf{x},\mathbf{x}')$ is linked to the target autocorrelation coefficient function $\rho_{\mathrm{transl}.}(\mathbf{x},\mathbf{x}')$ of the desired non-Gaussian field through an integral equation [18]. However, not for all $\rho_{\mathrm{transl}.}(\mathbf{x},\mathbf{x}')$ a corresponding $\rho(\mathbf{x},\mathbf{x}')$ can be found [29]. Moreover, it is computationally demanding to evaluate a $\rho(\mathbf{x},\mathbf{x}')$ that is associated with

a given $\rho_{\text{transl}}\cdot(\mathbf{x},\mathbf{x}')$, because of their implicit relationship in form of an integral equation. Therefore, it is often simpler to estimate the autocorrelation coefficient function of the underlying Gaussian random field $\rho(\mathbf{x},\mathbf{x}')$ directly. This can be achieved by transforming available data to Gaussian data using the inverse mapping $g^{-1}:\mathbb{R}\to\mathbb{R}$. It should be noted that a direct estimation of $\rho(\mathbf{x},\mathbf{x}')$ will result in a different autocorrelation function of $H(\mathbf{x},\theta)$ than the one arising from the solution of the integral equation according to translation field theory. However, such a direct estimation will overcome the problem that often occurs when the solution of the integral equation does not result in an autocorrelation function that is nonnegative definite.

3. Numerical methods to solve the KL expansion

3.1. Introduction

Integral eigenvalue problems of the type given in Eq. (6) are difficult to solve analytically except for a few autocovariance functions defined on domains Ω of simple geometric shape. Analytical solutions for exponential and triangular kernels are discussed in [2] for one-dimensional domains. Extensions to multidimensional rectangular domains can be derived assuming a separable covariance structure (e.g., see [1]). In general, the integral eigenvalue problem is solved numerically. The random field approximation of the truncated KL expansion given in Eq. (7) is approximated as:

$$\hat{H}(\mathbf{x},\theta)=\mu(\mathbf{x})+\sum_{i=1}^{M}\sqrt{\hat{\lambda}_i}\,\hat{\varphi}_i(\mathbf{x})\,\hat{\xi}_i(\theta) \tag{13}$$

where $\hat{\lambda}_i$ and $\hat{\varphi}_i$ are approximations to the true eigenvalues λ_i and eigenfunctions φ_i. $\hat{\xi}_i(\theta)$ are standard uncorrelated random variables, i.e., $\mathrm{E}[\hat{\xi}_i(\theta)\,\hat{\xi}_j(\theta)]=\delta_{ij}\ \forall i,j\leqslant M$. Note that due to the approximate character of the numerical solution, the random variables $\hat{\xi}_i$ are not necessarily orthogonal to the random variables ξ_i used in the representation of Eq. (5). This means that the expression for the error variance given in Eq. (9) cannot be derived from Eq. (2) for the numerical approximation of the KL expansion. Therefore, the error measures listed in Section 2.4 are not strictly valid for the approximated truncated KL expansion. Moreover, it is important to note that the random field approximation given in Eq. (13) does no longer possess the optimality property of the truncated analytical KL expansion.

Numerical algorithms for the solution of Fredholm integral eigenvalue problems approximate the eigenfunctions by a set of functions $h_j:\Omega\to\mathbb{R}$ as:

$$\varphi_i(\mathbf{x})\approx\hat{\varphi}_i(\mathbf{x})=\sum_{j=1}^{N}d_j^i\,h_j(\mathbf{x}) \tag{14}$$

where the coefficients $d_j^i\in\mathbb{R}$ have to be determined. In general, all algorithms can be categorized into three main categories: degenerate kernel methods, Nyström methods, and projection methods. Projection methods can be further subdivided into collocation methods and Galerkin methods.

3.2. Nyström method

In the Nyström method [4], the integral in the eigenvalue problem of Eq. (6) is approximated by a numerical integration scheme. Applications to integral eigenvalue problems published in literature include [30–32]. Numerical algorithms are discussed in [33,34]. The problem is approximated as:

$$\sum_{j=1}^{N}w_j\,\mathrm{Cov}(\mathbf{x},\mathbf{x}_j)\,\hat{\varphi}_i(\mathbf{x}_j)=\hat{\lambda}_i\hat{\varphi}_i(\mathbf{x}) \tag{15}$$

where $\mathbf{x}_j\in\Omega$ with $j\in\{1,\ldots,N\}$, $N\in\mathbb{N}$ represent a finite set of integration points, and w_j is the integration weight associated with each $\mathbf{x}_j$. For a given N, the distribution of the integration points $\mathbf{x}_j$ and the value of the integration weights w_j depend on the applied numerical integration scheme. Special integration techniques exist for kernels that are non-differentiable on the diagonal, see [33,34]. It is assumed that for the applied numerical integration scheme, the solution of Eq. (15) converges against the analytical solution with increasing N.

In the Nyström method, Eq. (15) is solved at the integration points, i.e.:

$$\sum_{j=1}^{N}w_j\,\mathrm{Cov}(\mathbf{x}_n,\mathbf{x}_j)\,\hat{\varphi}_i(\mathbf{x}_j)=\hat{\lambda}_i\hat{\varphi}_i(\mathbf{x}_n),\quad n=1,\ldots,N \tag{16}$$

The above system of equations can be formulated in matrix notation as

$$\mathbf{C}\mathbf{W}\mathbf{y}_i=\hat{\lambda}_i\mathbf{y}_i \tag{17}$$

where $\mathbf{C}$ is a symmetric positive semi-definite $N\times N$ matrix with elements $c_{nj}=\mathrm{Cov}(\mathbf{x}_n,\mathbf{x}_j)$, $\mathbf{W}$ is a diagonal matrix of size N with nonnegative diagonal entries $\mathbf{W}_{jj}=w_j$, and $\mathbf{y}_i$ is a N-dimensional vector whose nth entry is $y_{i,n}=\hat{\varphi}_i(\mathbf{x}_n)$. Since the

integration weights w_j are nonnegative, the matrix $\mathbf{W}$ is symmetric and positive semi-definite. The problem in Eq. (17) is a matrix eigenvalue problem. This matrix eigenvalue problem can be reformulated to an equivalent matrix eigenvalue problem $\mathbf{B}\mathbf{y}_i^* = \hat{\lambda}_i \mathbf{y}_i^*$, where the matrix $\mathbf{B}$ is defined as $\mathbf{B} = \mathbf{W}^{\frac{1}{2}}\mathbf{C}\mathbf{W}^{\frac{1}{2}}$, where $\mathbf{W}^{\frac{1}{2}}$ is a diagonal matrix with entries $(\mathbf{W}^{\frac{1}{2}})_{jj} = \sqrt{w_j}$. The matrix $\mathbf{B}$ is a symmetric positive semi-definite matrix and, thus, the eigenvalues $\hat{\lambda}_i$ are nonnegative real numbers and the eigenvectors $\mathbf{y}_i^*$ are orthogonal to each other. The eigenvectors $\mathbf{y}_i$ can be obtain as $\mathbf{y}_i = \mathbf{W}^{-\frac{1}{2}}\mathbf{y}_i^*$, where $\mathbf{W}^{-\frac{1}{2}}$ denotes the inverse of the matrix $\mathbf{W}^{\frac{1}{2}}$.

Solving Eq. (15) for $\hat{\varphi}_i(\mathbf{x})$, we obtain the so-called Nyström interpolation formula of the eigenfunction $\hat{\varphi}_i(\mathbf{x})$. Taking into account that $\hat{\varphi}_i(\mathbf{x}_j) = \frac{1}{\sqrt{w_j}} y_{i,j}^*$, this results in:

$$\hat{\varphi}_i(\mathbf{x}) = \frac{1}{\hat{\lambda}_i} \sum_{j=1}^{N} \sqrt{w_j}\, y_{i,j}^* \operatorname{Cov}(\mathbf{x}, \mathbf{x}_j) \tag{18}$$

where $y_{i,j}^*$ is the jth element of the eigenvector $\mathbf{y}_i^*$.

The eigenfunctions have to be normalized such that $\int_\Omega (\hat{\varphi}_i(\mathbf{x}))^2 \,\mathrm{d}\mathbf{x} = 1$. Applying a numerical integration scheme, the inner product $\int_\Omega \hat{\varphi}_i(\mathbf{x})\, \hat{\varphi}_j(\mathbf{x})\, \mathrm{d}\mathbf{x}$ can be approximated as $\sum_{n=1}^{N} w_n\, \hat{\varphi}_i(\mathbf{x}_n)\, \hat{\varphi}_j(\mathbf{x}_n)$. Using the same integration points and integration weights as the ones used in Eq. (16), the approximation of the inner product can be simplified to $\int_\Omega \hat{\varphi}_i(\mathbf{x})\, \hat{\varphi}_j(\mathbf{x})\, \mathrm{d}\mathbf{x} \approx \left(\mathbf{y}_i^*\right)^T \mathbf{y}_j^*$. Therefore, the approximate eigenfunctions are orthonormal if and only if the eigenvectors are orthonormal.

Equivalence of the EOLE method with the Nyström method. The expansion optimal linear estimation (EOLE) method is a series expansion method for discretization of random fields that was developed in [18] based on linear estimation theory. Here we show that the EOLE method with a uniform distribution of points over the domain can be considered a special case of the Nyström method.

Assume that points $\mathbf{x}_j, j = 1, \ldots, N$ uniformly distributed over the domain Ω are available. The points $\mathbf{x}_j$ can be chosen either at random by sampling the uniform distribution over Ω or by application of the rectangle quadrature using the nodes of an equispaced structured grid. If the domain Ω does not have a simple shape, the integration procedure can be performed on a geometrically simpler domain Ω^* that contain Ω, i.e., $\Omega \subseteq \Omega^*$. In this case, points outside of Ω are not taken into account. If the points $\mathbf{x}_j, j = 1, \ldots, N$ are selected with one of the above procedures, then all the integration weights w_j in the integration scheme in Eq. (15) will be the same, i.e., $w_j = w\ \forall j = 1, \ldots, N$. Consequently, matrix $\mathbf{W}$ in Eq. (17) can be written as $\mathbf{W} = w\mathbf{I}$, where $\mathbf{I}$ is the identity matrix and $w = |\Omega|/N$. In this special case, the matrix eigenvalue problem of Eq. (17) can be reformulated as:

$$\mathbf{C}\mathbf{y}_i = \hat{\lambda}_i^* \mathbf{y}_i \tag{19}$$

where $\hat{\lambda}_i^*$ is related to $\hat{\lambda}_i$ in Eq. (17) as $\hat{\lambda}_i^* = \frac{N}{|\Omega|}\hat{\lambda}_i$. $\hat{\lambda}_i^*$ and $\mathbf{y}_i$ are the eigenvalues and eigenfunctions of the covariance matrix $\mathbf{C}$, respectively. Assuming normalized eigenvectors $\mathbf{y}_i$, i.e., $\|\mathbf{y}_i\| = 1$ for all i, gives after some algebra the following approximate truncated KL expansion:

$$\hat{H}(\mathbf{x}, \theta) = \mu(\mathbf{x}) + \sum_{i=1}^{M} \frac{\hat{\xi}_i(\theta)}{\sqrt{\hat{\lambda}_i^*}} \sum_{j=1}^{N} y_{i,j} \operatorname{Cov}(\mathbf{x}, \mathbf{x}_j) \tag{20}$$

where $y_{i,j}$ is the jth element of $\mathbf{y}_i$.

The matrix eigenvalue problem of Eq. (19) is the problem that needs to be solved for the EOLE method, and the expansion in Eq. (20) is equivalent to the one obtained in the EOLE method. Consequently, the EOLE method is equivalent to an approximate KL expansion, whereby the IEVP is solved by the Nyström method with a uniform distribution of integration points.

3.3. Projection methods

Due to the approximation of the eigenfunctions stated in Eq. (14), the Fredholm integral equation of Eq. (6) can be solved only approximately. Inserting Eq. (14) into Eq. (6) gives the residual $r_{\mathrm{IEVP}}(\mathbf{x})$:

$$r_{\mathrm{IEVP}}(\mathbf{x}) = \sum_{j=1}^{N} d_j^i \left(\int_\Omega \operatorname{Cov}(\mathbf{x}, \mathbf{x}') h_j(\mathbf{x}')\, \mathrm{d}\mathbf{x}' - \hat{\lambda}_i h_j(\mathbf{x}) \right) \tag{21}$$

In projection methods, the coefficients $d_1^i, \ldots, d_N^i$ are determined such that this residual is minimized in some sense. In the following, we discuss two different categories of projection methods; collocation and Galerkin methods.

3.4. Collocation methods

In collocation methods, the residual r_{IEVP} defined in Eq. (21) is minimized pointwise for a given set of points $\{\mathbf{x}_l\}_{l=1}^{P}$. Let $\mathbf{A}$ and $\mathbf{N}$ be matrices of size $P \times N$. The elements of $\mathbf{A}$ and $\mathbf{N}$ are $a_{lj} = \int_\Omega \operatorname{Cov}(\mathbf{x}_l, \mathbf{x})\, h_j(\mathbf{x})\, \mathrm{d}\mathbf{x}$ and $n_{lj} = h_j(\mathbf{x}_l)$, respectively. Moreover, let $\mathbf{d}_i$ be a vector of dimension N with elements $(\mathbf{d}_i)_j = d_j^i$. For the special case of $P = N$, the problem can be formulated as $r_{\mathrm{IEVP}}(\mathbf{x}_l) = 0,\ \forall l = 1, \ldots, P$, which is equivalent to the following generalized matrix eigenvalue problem:

$$\mathbf{A}\mathbf{d}_i = \hat{\lambda}_i \mathbf{N}\mathbf{d}_i \tag{22}$$

Note that $P = N$ is a rather restrictive case, especially if higher-order basis functions are applied: For higher-order basis functions, a good approximation of the eigenfunctions is usually achieved with a relatively small number of basis functions N, whereas a large P is typically required to minimize Eq. (21) in a global sense.

For the case $P > N$, the condition $r_{\mathrm{IEVP}}(\mathbf{x}_l) = 0,\ \forall l = 1,\dots,P$ cannot be satisfied at all points $\mathbf{x}_l$. In this case, the coefficients d_j^i can be determined by means of a least squares minimization, i.e., $\arg\min_{\mathbf{d}_i} \sum_{l=1}^{P} (r_{\mathrm{IEVP}}(\mathbf{x}_l))^2,\ \forall i = 1,\dots,N$. This leads to the following nonlinear matrix eigenvalue problem:

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{d}_i - \hat{\lambda}_i \mathbf{A}^{\mathrm{T}}\mathbf{N}\mathbf{d}_i - \hat{\lambda}_i \mathbf{N}^{\mathrm{T}}\mathbf{A}\mathbf{d}_i + \hat{\lambda}_i^2 \mathbf{N}^{\mathrm{T}}\mathbf{N}\mathbf{d}_i = 0 \tag{23}$$

This kind of eigenvalue problem is also known as quadratic eigenvalue problem [35], and can be transformed to the following generalized matrix eigenvalue problem:

$$\begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{d}_i \\ \mathbf{c}_i \end{pmatrix} = \hat{\lambda}_i \begin{pmatrix} \mathbf{N}^{\mathrm{T}}\mathbf{A} + \mathbf{A}^{\mathrm{T}}\mathbf{N} & -\mathbf{N}^{\mathrm{T}}\mathbf{N} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{d}_i \\ \mathbf{c}_i \end{pmatrix} \tag{24}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{c}_i = \hat{\lambda}_i \mathbf{d}_i$. According to [35] the eigenvalues $\hat{\lambda}_i$ are real and positive, and the eigenvectors $\mathbf{d}_i$ are linearly independent. The numerical solution of quadratic eigenvalue problems is discussed in [35].

3.5. Galerkin methods

3.5.1. Introduction

In Galerkin methods, the coefficients d_j^i are chosen such that the residual r_{IEVP} becomes orthogonal to the subspace of $L^2(\Omega)$ spanned by the basis functions $\{h_j\}_{j=1}^{N}$. Hence, the following problem is solved:

$$\int_\Omega r_{\mathrm{IEVP}}(\mathbf{x})\, h_j(\mathbf{x})\, \mathrm{d}\mathbf{x} = 0 \quad \forall j = 1,\dots,N \tag{25}$$

Eq. (25) can be expressed in matrix notation as the generalized matrix eigenvalue problem

$$\mathbf{B}\mathbf{d}_i = \hat{\lambda}_i \mathbf{M}\mathbf{d}_i \tag{26}$$

where $\mathbf{B}$ is a symmetric positive semi-definite $N \times N$ matrix whose elements are defined as

$$b_{ln} = \int_\Omega h_l(\mathbf{x}) \int_\Omega \mathrm{Cov}(\mathbf{x},\mathbf{x}')\, h_n(\mathbf{x}')\, \mathrm{d}\mathbf{x}'\, \mathrm{d}\mathbf{x} \tag{27}$$

and $\mathbf{M}$ is a symmetric positive definite $N \times N$ matrix with elements

$$m_{ln} = \int_\Omega h_l(\mathbf{x})\, h_n(\mathbf{x})\, \mathrm{d}\mathbf{x} \tag{28}$$

On three-dimensional domains, Eq. (27) constitutes a six-folded integral. This renders the assembly of matrix $\mathbf{B}$ computationally expensive. To speed up the assembly of $\mathbf{B}$, it is suggested in [13,14] to approximate $\mathbf{B}$ as a hierarchical matrix ($\mathcal{H}$-matrix). If $\mathbf{B}$ is expressed as a hierarchical matrix, sub-blocks of $\mathbf{B}$ that have small contribution to the overall system are approximated as low-rank matrices. Therefore, not all coefficients b_{ln} have to be computed explicitly. Moreover, matrix operations on hierarchic matrices can be performed with almost linear complexity [15].

The advantage of Galerkin-based approaches compared to the other investigated numerical methods is that the error is minimized over the entire domain and not only at certain points within the domain. The convergence behavior of the Galerkin method with high-order basis functions is investigated in [5].

3.5.2. Finite element method

In practical problems, the domain Ω often has a complex geometrical shape. One solution is to partition the domain Ω into multiple elements and construct a functional space such that each function is non-zero only on a small number of neighboring elements. This allows for the definition of local basis functions defined on the elemental domains. This approach is termed finite element method (FEM) and the partition of Ω is called finite element mesh (e.g., [36]). The FEM was applied to the solution of the IEVP of the KL expansion in [2]. If the random field is needed as input for a finite element simulation, the FEM is straightforward to implement because a finite element mesh is readily available. In the FEM, the local basis functions are typically chosen as piecewise linear polynomials. However, it is noted that the use of higher-order basis functions can greatly improve the convergence behavior of the method [37]. The selection of appropriate high-order basis functions is discussed in Section 3.5.4. Contrary to other random field discretization methods, as for example, the midpoint method [38], very fine meshes do not lead to numerical problems.

However, if a meshing of the domain is solely required for random field discretization, meshless approaches might be preferred in order to overcome the obstacles of mesh generation. A meshless approach for defining the basis functions in the Galerkin method was investigated in [16,39]. However, it should be noted that this method requires the evaluation of

the integrals over Ω defined in Eqs. (27) and (28) and, therefore, the shape of the domain Ω has to be considered for integration.

3.5.3. Finite cell method

Another quasi meshless approach is based on so-called finite cells. The finite cell method (FCM) [17] was originally developed as an extension to the FEM for the solution of linear elasticity problems. The FCM is a fictitious domain method of higher approximation order. The principal idea of the method is independent of the applied element basis functions and so far has been successfully tested for integrated Legendre polynomials [17,40,41], as known from the p-version of the FEM, as well as for B-splines [41] and NURBS [42]. The FCM is extended here for application to the solution of the IEVP in the KL expansion.

Let $\Omega \subset \mathbb{R}^d$ be the domain of interest and $\Omega^* \subset \mathbb{R}^d$ a geometrically simpler domain with $\Omega \subseteq \Omega^*$. The geometrically simpler domain Ω^* is called *primitive domain*, and the original domain Ω is referred to as *physical domain*. In the FCM, the primitive domain is meshed instead of the physical domain, and, therefore, mesh generation is a trivial task. The elements of the mesh of the primitive domain are referred to as cells and the cell domain is denoted Ω_e. The finite cell approach is illustrated in Fig. 1.

Consider a set of basis functions $h_j^*(\mathbf{x}) \in L^2(\Omega^*)$ that form a basis of a subspace in $L^2(\Omega^*)$. In this regard, the approximation of the eigenfunctions given in Eq. (14) is redefined as

$$\varphi_i(\mathbf{x}) \approx \hat{\varphi}_i(\mathbf{x}) = \sum_{j=1}^{N} d_j^i \, h_j^*(\mathbf{x}) \tag{29}$$

That is, the solution of the integral equation defined on the physical domain Ω is approximated with basis functions spanned over the primitive domain Ω^*. This is achieved by performing a Galerkin projection on the space $L^2(\Omega^*)$, while requiring that the residual of (21) is minimized over the domain Ω, i.e.,

$$\int_\Omega r_{\mathrm{IEVP}}(\mathbf{x}) \, h_j^*(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 0 \quad \forall \, j = 1, \ldots, N \tag{30}$$

The above leads to a generalized matrix eigenvalue problem identical to (25), whereby the elements of matrices $\mathbf{B}$ and $\mathbf{M}$ can be written as the following integrals over the primitive domain Ω^*:

$$b_{ln} = \int_{\Omega^*} \alpha(\mathbf{x}) h_l^*(\mathbf{x}) \int_{\Omega^*} \mathrm{Cov}(\mathbf{x}, \mathbf{x}') h_n^*(\mathbf{x}') \alpha(\mathbf{x}') \, \mathrm{d}\mathbf{x}' \, \mathrm{d}\mathbf{x} \tag{31}$$

$$m_{ln} = \int_{\Omega^*} \alpha(\mathbf{x}) \, h_l^*(\mathbf{x}) \, h_n^*(\mathbf{x}) \, \mathrm{d}\mathbf{x} \tag{32}$$

The mapping $\alpha : \Omega^* \to \{0, 1\}$ is one for $\mathbf{x} \in \Omega$ and zero otherwise. Hence, only the region of Ω^* that is within the physical domain Ω is considered in the integration. However, the approximation of the eigenfunctions $\hat{\varphi}_i(\mathbf{x})$ extends to the region outside of the physical domain, since the basis functions are defined over the entire domain Ω^*. Therefore, the eigenfunctions need to be normalized on the physical domain Ω, i.e., $\int_{\Omega^*} \alpha(\mathbf{x}) \, \hat{\varphi}_i(\mathbf{x})^2 \, \mathrm{d}\mathbf{x} = 1$.

Higher-order basis functions are of crucial importance for the applicability of the method because they yield a fast rate of convergence [17]. The selection of appropriate high order polynomial basis functions for this problem is discussed in Section 3.5.4.

Integration technique for the FCM. The FCM shifts the problem from mesh generation on geometrically complex domains to the integration of discontinuous functions. An accurate evaluation of the integrals is essential for the effectiveness of the method. However, the evaluation of the double integral over Ω^* in Eq. (31) may involve considerable computational costs. This is mainly due to the discontinuity that appears at the boundary of the physical domain, denoted $\partial\Omega$. The discontinuity is caused by the indicator function $\alpha(\mathbf{x})$. If a cell Ω_e is cut by the boundary $\partial\Omega$, standard Gaussian quadrature exhibits a slow rate of convergence since here the integrand cannot be approximated well by a polynomial function.

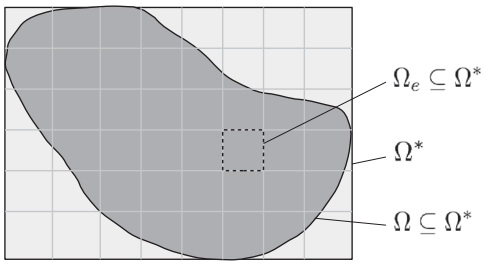

Fig. 1. The finite cell method – idea and notation.

In order to overcome this obstacle, it is suggested to apply a so-called staggered Gaussian quadrature scheme, proposed in [40]. The idea of this numerical integration scheme is illustrated in Fig. 2 for a finite cell that is cut by the boundary $\partial\Omega$: The domain Ω_e of a cell cut by $\partial\Omega$ is gradually divided into sub-cells. For one-, two- and three-dimensional elements, a binary-, quad- and oct-tree is used, respectively. A sub-cell is further refined if it is cut by the boundary $\partial\Omega$ and if the level of the sub-cell is smaller than the specified maximum tree-depth. The level of the original finite cell is defined to be zero. In the example depicted in Fig. 2, the maximum tree-depth is four.

Standard Gaussian integration is performed on each leaf sub-cell of the resulting grid. In the context of this work, the number of Gauss-points used to integrate on a leaf sub-cell of a certain level is decreased with an increasing level of the sub-cell. This is contrary to the approach presented in [17,40,41], where all sub-cells are integrated with the full number of Gauss-points. However, the smaller a sub-cell becomes, the better the function to integrate can be approximated by lower order polynomials. Moreover, the contribution of a sub-cell to the total integral decreases with an increasing level. In this work, the number of Gauss-points in each coordinate direction is halved at the third level and again for all levels larger or equal than five. However, at least two Gauss-points in each coordinate direction are used for all sub-cells with a level smaller than the maximum tree-depth. For the sub-cells that are cut by $\partial\Omega$, only a single Gauss-point located in the barycenter of the cell is used. The computation of the barycenters of all cut sub-cells is performed once prior to the integration. The time needed to compute the barycenters does not contribute much to the total time required to solve the integral in Eq. (31).

The reduction of Gauss-points speeds up the evaluation of the double integral in Eq. (31) significantly, compared to a numerical integration using the full number of Gauss-points at each sub-cell. Additionally, the use of a single integration point in the barycenter of a sub-cell that is cut by $\partial\Omega$ leads to a smooth convergence behavior with respect to an increasing maximum tree-depth. The resolution of the boundary $\partial\Omega$ can be refined by increasing the maximum tree-depth, but not by increasing the total number of Gauss-points of a finite cell. Note that if a cell is located completely outside of the physical domain, i.e., $\Omega_e \cap \Omega = \emptyset$, it does not have to be integrated since the integral is zero.

Some kernels $\mathrm{Cov}(\mathbf{x}, \mathbf{x}')$ are non-differentiable on the diagonal. Consequently, the integrand of the inner integral in Eq. (31) is non-differentiable for $\mathbf{x} = \mathbf{x}'$, additionally to the discontinuity at $\partial\Omega$. Moreover, integration over cells that are not cut by $\partial\Omega$ as well as over elements in the classical FEM is also affected by non-differentiable kernels. This additional difficulty can be dealt with by a slight modification of the integration scheme proposed above: Instead of refining only the region around $\partial\Omega$ with a high resolution, the entire domain of all cells (or elements) is refined with at least a predefined minimum resolution. However, a refinement is only required if both $\mathbf{x}$ and $\mathbf{x}'$ are located in the same cell (or element). In the h-version of the FEM, a finite element mesh consists usually of many elements. Fixing the current integration point $\mathbf{x}$ of the outer integral, only the element containing $\mathbf{x}$ needs to be refined with at least a minimum resolution. Therefore, for the FEM, the loss in efficiency due to non-differentiable kernels is almost negligible. For the FCM, this does not hold since a small number of cells is typically used due to the application of higher-order basis functions. Consequently, for the proposed integration scheme and non-differentiable kernels it is advantageous to refine Ω^* using multiple cells instead of just a single cell - to speed up the integration.

3.5.4. Selection of the basis functions

Different types of basis functions can be used in conjunction with a Galerkin-based approach. For an optimal choice of basis functions, matrix $\mathbf{M}$ would become a diagonal matrix and, hence, the generalized matrix eigenvalue problem of Eq. (26) would reduce to a standard matrix eigenvalue problem. If the domain Ω has a rectangular shape, the basis functions can be selected such that they are defined globally on Ω, leading to a spectral Galerkin discretization. In this case, matrix $\mathbf{M}$ would become a diagonal matrix if the classical Legendre polynomials are selected as basis functions. However, in a spectral approach, the only way to improve the approximation is to increase the polynomial order of the involved polynomials.

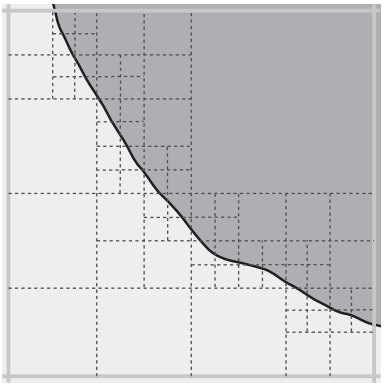

Fig. 2. Staggered Gaussian integration: mesh for integration on a finite cell cut by $\partial\Omega$.

This can lead to numerical problems if the correlation length of the random field becomes too small, since for such cases the required polynomial order for an accurate approximation is typically large.

A more flexible approach is to define the basis functions locally on the element domains, as is done in the FEM and the FCM. In this case, a hierarchic basis is of advantage, since it allows a straightforward coupling of neighboring elements. Also, in hierarchic functional spaces existing basis functions do not change when the basis is extended, e.g., when the polynomial order of the basis functions is increased.

In the following, we propose a new hierarchic polynomial basis defined in terms of the Gegenbauer polynomials [43]. For the problem at hand, this hierarchic basis leads to matrices with a better condition number than the hierarchic basis introduced by Szabó and Babuška [44]. Moreover, the proposed hierarchic basis leads to a sparse structure of $\mathbf{M}$ for one-dimensional problems and two-/three-dimensional problems with structured Cartesian meshes. The construction of a hierarchic basis defined in terms of the integrated Legendre polynomials [44] is discussed in [37].

Hierarchic Gegenbauer polynomials. In order to ensure continuity between the elemental domains, the first two one-dimensional local basis functions of the hierarchic basis are chosen as the following piecewise linear polynomials:

$$h_1^{1D}(\eta) = \frac{1}{2}(1-\eta) \tag{33}$$

$$h_2^{1D}(\eta) = \frac{1}{2}(1+\eta) \tag{34}$$

where η is the local coordinate that runs on the standard interval $[-1,1]$. All basis functions $h_i^{1D}(\eta)$ with $i>2$ should satisfy the conditions

$$h_i^{1D}(-1) = 0 \ \wedge \ h_i^{1D}(1) = 0 \quad \forall\, i>2 \tag{35}$$

For the problem at hand, we require

$$\int_{-1}^{1} h_i^{1D}(\eta) h_j^{1D}(\eta)\,\mathrm{d}\eta = \delta_{ij} \quad \forall\, i,j>2 \tag{36}$$

which ensures that the matrix $\mathbf{M}$ is sparse for one-dimensional problems. The polynomial degree of h_i^{1D} for $i>2$ is equal to $i-1$. The two boundary conditions defined in Eq. (35) can be enforced by defining the basis functions as

$$h_i^{1D}(\eta) = (1-\eta^2)\phi_{i-3}(\eta) \quad \forall\, i>2 \tag{37}$$

where $\phi_j(\eta)$ with $j \geqslant 0$ are polynomials of order j. The requirement of Eq. (36) is fulfilled if the polynomials ϕ_j are chosen as the normalized Gegenbauer polynomials with $\alpha = 2.5$. The Gegenbauer polynomials C_j^{α} are orthogonal to the weight function $(1-\eta^2)^{\alpha-\frac{1}{2}}$. For the special case where $\alpha = 2.5$, they can be defined by means of the recursive formula:

$$C_0^{2.5}(\eta) = 1$$

$$C_1^{2.5}(\eta) = 5\eta$$

$$C_i^{2.5}(\eta) = \frac{1}{i}\left[2\eta(i+1.5)C_{i-1}^{2.5}(\eta) - (i+3)C_{i-2}^{2.5}(\eta)\right]$$

The corresponding normalized Gegenbauer polynomials read:

$$\phi_i(\eta) = a_i \cdot C_i^{2.5}(\eta) \quad \forall\, i \geqslant 0$$

where the positive and non-zero scaling factor a_i is defined as

$$a_i = \sqrt{\frac{i!(i+2.5)[\Gamma(2.5)]^2}{\pi 2^{-4}\Gamma(i+5)}} \quad \forall\, i \geqslant 0$$

with Γ the Euler gamma function. The first eight basis functions of the one-dimensional hierarchic Gegenbauer basis are sketched in Fig. 3. A hierarchic basis for two- or three-dimensional elements can be constructed from the set of one-dimensional basis functions. Often this is done using either the *tensor product space* that comprises all combinations of one-dimensional basis functions or the *trunk space* that excludes some higher-order combinations [37].

Note that among all one-dimensional hierarchic bases that comply with Eqs. (33)–(35), the sparsity of the matrix $\mathbf{M}$ is maximized for the choice of the hierarchic Gegenbauer polynomials. This holds also for rectangular two-dimensional elements and three-dimensional elements that have the shape of a rectangular cuboid. The structure of $\mathbf{M}$ for a single one-dimensional element with higher-order basis functions is depicted in Fig. 4. A strictly diagonal shape cannot be achieved for $\mathbf{M}$ with a hierarchic basis. This is because $h_3^{1D}(\eta) = a_0(1-\eta^2)$ and, consequently, $\int_{-1}^{1} h_i^{1D}(\eta) h_3^{1D}(\eta)\,\mathrm{d}\eta > 0$ if $i=1$ or $i=2$. For unstructured meshes in two- or three-dimensional problems or for problems that are solved with the FCM, the hierarchic Gegenbauer basis loses its optimality, i.e., the sparsity of $\mathbf{M}$ decreases. In the FCM, this effect is caused by the cells that are cut by the boundary $\partial\Omega$. It is noted, however, this problem is present whenever a hierarchic basis is used in the context of the FEM or FCM. In fact, the hierarchic Gegenbauer basis results in a relatively small condition number of $\mathbf{M}$ compared to other hierarchic bases, such as the integrated Legendre polynomials that are typically applied in the p-version of the FEM.

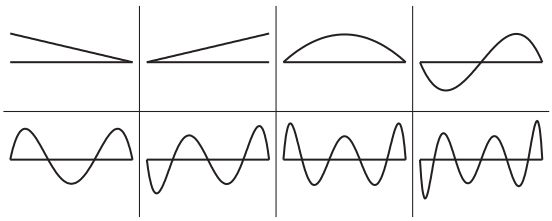

Fig. 3. The first eight hierarchic Gegenbauer polynomials.

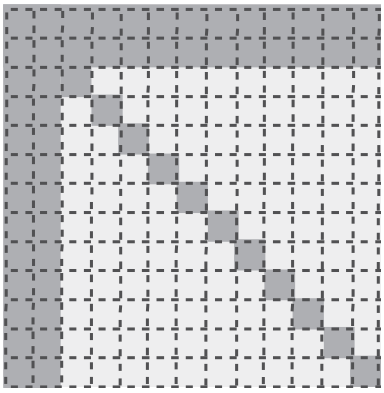

Fig. 4. Structure of the local element mass matrix **M** obtained with the hierarchic Gegenbauer polynomials for a one-dimensional problem.

If a hierarchic basis is not enforced, a diagonal structure of **M** can be obtained by means of Gram-Schmidt orthogonalization [45]. However, the problem of finding a diagonal structure becomes involved if the mesh is unstructured or the FCM is applied.

3.5.5. Projection of the autocovariance function

The matrix **B** is computationally expensive to assemble. The computational cost becomes especially high on three-dimensional domains, since for this case Eq. (27) (or Eq. (31) for the FCM) consists of a six-folded integral. An alternative to the assembly of matrix **B** according to Eq. (27) (or Eq. (31))) is to project the autocovariance function onto the space spanned by the basis functions $\{h_j\}_{j=1}^N$ (or $\left\{h_j^*\right\}_{j=1}^N$ for the FCM):

$$\mathrm{Cov}(\mathbf{x},\mathbf{x}') \approx \sum_{j=1}^{N}\sum_{n=1}^{N} k_{jn}\, h_j(\mathbf{x})\, h_n(\mathbf{x}') \tag{38}$$

This approach is similar to the technique applied in the degenerate kernel methods (Section 3.6). Note that the projection of the autocovariance function decreases the accuracy of the random field approximation, due to the approximation of the autocovariance function. This does not hold for the special case of L^2-projection.

Inserting Eq. (38) into Eq. (27) (or Eq. (31)) leads to the formulation $\mathbf{B} = \mathbf{MKM}$, where the matrix **K** is a symmetric positive semi-definite $N \times N$ matrix whose elements k_{jn} need to be determined. In [46], the coefficients k_{jn} were defined as the covariance between nodes j and n of the finite element mesh and a linear interpolation between the nodes was used as a projection approach. We investigate other approaches to perform the projection, namely L^2-projection, $H^{1/2}$-projection and discrete projection. The projection methods aim at minimizing the residual $r_{\mathrm{Cov}}(\mathbf{x},\mathbf{x}')$:

$$r_{\mathrm{Cov}}(\mathbf{x},\mathbf{x}') = \mathrm{Cov}(\mathbf{x},\mathbf{x}') - \sum_{j=1}^{N}\sum_{n=1}^{N} k_{jn}\, h_j(\mathbf{x})\, h_n(\mathbf{x}') \tag{39}$$

L^2-projection In L^2-projection, the residual is minimized with respect to

$$\int_\Omega h_l(\mathbf{x}) \int_\Omega r_{\mathrm{Cov}}(\mathbf{x},\mathbf{x}')\, h_m(\mathbf{x}')\, \mathrm{d}\mathbf{x}'\, \mathrm{d}\mathbf{x} = 0 \quad \forall\, l,m = 1,\ldots,N \tag{40}$$

Inserting Eq. (39) into Eq. (40) and comparing the obtained expression to Eq. (27) (or Eq. (31)) shows that the assembly of matrix **B** according to Eq. (27) (or Eq. (31)) is equivalent to a L^2-projection of the autocovariance function. Moreover, there is no gain in computational efficiency compared to solving Eq. (27) (or Eq. (31)) directly.

$H^{1/2}$-projection $H^{1/2}$-projection was developed for an efficient estimation of coefficients of higher-order basis functions for application of boundary conditions [47]. We discuss its application to the projection of the covariance kernel on a set of basis functions in conjunction with the FEM and FCM approaches. In $H^{1/2}$-projection, the problem is not solved in one global step as in L^2-projection, but in a number of local steps. This approach is only applicable if hierarchic basis functions are used at the elemental domains. In a first step, the autocovariance function is evaluated at the nodes of the mesh. Linear basis functions are used to interpolate between the nodes. In the subsequent steps, the modes that have already been determined are subtracted from Eq. (39 series of local L^2-projections is applied to determine the remaining modes. That is, in the second stage, the edge-modes are determined for each edge separately. Thereafter, for two- and three-dimensional elements, the face modes are determined for each face separately in a third step. Finally, for three-dimensional elements, the internal modes have to be approximated for each element separately in a last step. Compared to the L^2-projection, the problem is solved by means of several local L^2-projections, which may lead to significant gain in computational cost. The disadvantage of the $H^{1/2}$-projection is that its implementation in numerical codes is complex - even for one-dimensional elements.

If only linear basis functions are applied, the problem reduces to the trivial case of computing the coefficients as $k_{jn} = \mathrm{Cov}(\mathbf{x}_j, \mathbf{x}_n)$, where $\mathbf{x}_j$ and $\mathbf{x}_n$ denote the coordinates of the nodes of the mesh. The projection becomes equivalent to a linear interpolation of the covariance at the nodes of the mesh. This technique has been applied in [46]; it is referred to as linear projection in the following.

Discrete projection. This projection method minimizes the residual $r_{\mathrm{Cov}}(\mathbf{x}, \mathbf{x}')$ pointwise at a given set of points $\{\mathbf{x}_j\}_1^P$, where P denotes the number of points in the set, and $P \geqslant N$. Using a linear least squares minimization with $\arg\min_{\mathbf{k}} \sum_k^P \sum_l^P r_{\mathrm{Cov}} r(\mathbf{x}_k, \mathbf{x}_l)^2$, the problem can be formulated as:

$$\overline{\mathbf{N}}^{\mathrm{T}} \overline{\mathbf{N}} \overline{\mathbf{k}} = \overline{\mathbf{N}}^{\mathrm{T}} \overline{\mathbf{c}} \tag{41}$$

where $\overline{\mathbf{N}}$ is a $P^2 \times N^2$ matrix with coefficients $\overline{n}_{(kl),(jm)} = h_j(\mathbf{x}_k)\, h_m(\mathbf{x}_l)$, $\overline{\mathbf{k}}$ is a vector of size N^2 that contains all coefficients k_{jm}, and $\overline{\mathbf{c}}$ is a vector of size P^2 whose coefficients are defined as $\overline{c}_{(kl)} = \mathrm{Cov}(\mathbf{x}_k, \mathbf{x}_l)$. Due to symmetry of the solution vector $\overline{\mathbf{k}}$, i.e., $\overline{k}_{(ij)} = \overline{k}_{(ji)}$, the size of the problem to solve can be condensed from N^2 to $\frac{N^2+N}{2}$.

Contrary to L^2- and $H^{1/2}$-projection, for the discrete projection approach no explicit integration needs to be performed. Therefore, the quality of the numerical solution depends, for a given set of basis functions, solely on the number of points used in the projection. The drawback of this projection method is the size of the matrix $\overline{\mathbf{N}}$ in the linear system of Eq. (41).

3.6. Degenerate kernel methods

In degenerate kernel methods [4], the kernel $\mathrm{Cov}(\mathbf{x}, \mathbf{x}')$ is approximated as:

$$\mathrm{Cov}(\mathbf{x}, \mathbf{x}') \approx \hat{K}(\mathbf{x}, \mathbf{x}') = \sum_{j=1}^{N} \sum_{n=1}^{N} k_{jn}\, \alpha_j(\mathbf{x})\, \beta_n(\mathbf{x}') \tag{42}$$

with coefficients $k_{jn} \in \mathbb{R}$, and linear independent functions $\alpha_1(\mathbf{x}), \ldots, \alpha_N(\mathbf{x})$ and $\beta_1(\mathbf{x}'), \ldots, \beta_N(\mathbf{x}')$. The approximation of (42) offers many alternatives on how to solve the IEVP of Eq. (6), some of which lead to the methods already presented in the previous sections. Therefore, the degenerate kernel approach is not considered explicitly here.

4. Numerical studies

4.1. Error measures

Computable expressions for the mean error variance or the global mean square error cannot be derived for all of the numerical algorithms presented in Section 3. Therefore, two error measures are introduced that can be evaluated numerically for all investigated methods:

$$\varepsilon_{\mathrm{Var}} = \frac{1}{|\Omega|} \int_\Omega \frac{|\mathrm{Var}[H(\mathbf{x}, \theta)] - \mathrm{Var}[\hat{H}(\mathbf{x}, \theta)]|}{\mathrm{Var}[H(\mathbf{x}, \theta)]}\, \mathrm{d}\mathbf{x} \tag{43}$$

$$\varepsilon_{\mathrm{Cov}} = \frac{1}{|\Omega|} \int_\Omega \frac{\int_{\mathbf{x}' \in \Omega} |\Gamma_{\mathbf{x}\mathbf{x}'}(\mathbf{x}, \mathbf{x}') - \hat{\Gamma}_{\mathbf{x}\mathbf{x}'}(\mathbf{x}, \mathbf{x}')|\, \mathrm{d}\mathbf{x}'}{\int_{\mathbf{x}' \in \Omega} |\Gamma_{\mathbf{x}\mathbf{x}'}(\mathbf{x}, \mathbf{x}')|\, \mathrm{d}\mathbf{x}'}\, \mathrm{d}\mathbf{x} \tag{44}$$

where $\Gamma_{\mathbf{x}\mathbf{x}'}(\mathbf{x}, \mathbf{x}') = \mathrm{Cov}[H(\mathbf{x}, \theta), H(\mathbf{x}', \theta)]$ and $\hat{\Gamma}_{\mathbf{x}\mathbf{x}'}(\mathbf{x}, \mathbf{x}') = \mathrm{Cov}[\hat{H}(\mathbf{x}, \theta), \hat{H}(\mathbf{x}', \theta)]$.

For the analytical solution of the IEVP, the error measure $\varepsilon_{\mathrm{Var}}$ is equivalent to the mean error variance defined in Eq. (3), since $\mathrm{Var}[H(\mathbf{x}, \theta)] \geqslant \mathrm{Var}[\tilde{H}(\mathbf{x}, \theta)] \forall \mathbf{x} \in \Omega$ [2]. In practice, the latter also holds for $\tilde{H}(\mathbf{x}, \theta) = \hat{H}(\mathbf{x}, \theta)$, except for cases where all numerically computed eigenpairs are used in the expansion, i.e., when $M = N$. Therefore, $\varepsilon_{\mathrm{Var}}$ can be obtained by means of Eq. (10) or Eq. (12), even if the approximated eigenvalues and eigenfunctions are not equivalent to the analytical ones. However, it should be noted that $\varepsilon_{\mathrm{Var}}$ is not identical to the mean error variance since $\mathrm{E}[\hat{\xi}_i(\theta)\, \xi_j(\theta)] \neq \delta_{ij}$ (see Section 3.1).

The error measure $\varepsilon_{\mathrm{Cov}}$ can be considered a more general measure than $\varepsilon_{\mathrm{Var}}$ [48], since $\varepsilon_{\mathrm{Var}}$ accounts only for the quality of the approximation of the variance, whereas $\varepsilon_{\mathrm{Cov}}$ quantifies the quality of the entire approximated covariance structure. It should be noted that $\varepsilon_{\mathrm{Var}} = 0$ does not imply that $\varepsilon_{\mathrm{Cov}} = 0$ as well.

In order to compare the rate of convergence of the numerical truncated KL expansion $\hat{H}(\mathbf{x},\theta)$ to the analytical KL expansion $\tilde{H}(\mathbf{x},\theta)$ for a fixed M, it is more appropriate to look at relative errors

$$\varepsilon_{\mathrm{Var,rel}} = \frac{|\varepsilon_{\mathrm{Var}} - \varepsilon_{\mathrm{Var,ref}}|}{\varepsilon_{\mathrm{Var,ref}}} \tag{45}$$

$$\varepsilon_{\mathrm{Cov,rel}} = \frac{|\varepsilon_{\mathrm{Cov}} - \varepsilon_{\mathrm{Cov,ref}}|}{\varepsilon_{\mathrm{Cov,ref}}} \tag{46}$$

where $\varepsilon_{\mathrm{Var}}$ and $\varepsilon_{\mathrm{Cov}}$ are defined according to Eqs. (43) and (44), respectively, and $\varepsilon_{\mathrm{Var}}, \mathrm{ref}$ and $\varepsilon_{\mathrm{Cov}}, \mathrm{ref}$ are the errors obtained for the corresponding analytical solution $\tilde{H}(\mathbf{x},\theta)$, i.e., substituting $\tilde{H}(\mathbf{x},\theta)$ for $\hat{H}(\mathbf{x},\theta)$ in Eqs. (43) and (44).

4.2. Autocorrelation coefficient functions

The behavior of the presented random field discretization methods is investigated for the isotropic exponential, squared exponential and sine autocorrelation coefficient functions, defined respectively by ([18])

$$\rho_{\mathrm{A}}(|\mathbf{x}-\mathbf{x}'|) = \exp\left(-\frac{|\mathbf{x}-\mathbf{x}'|}{l_{\mathrm{A}}}\right) \tag{47}$$

$$\rho_{\mathrm{B}}(|\mathbf{x}-\mathbf{x}'|) = \exp\left[-\left(\frac{|\mathbf{x}-\mathbf{x}'|}{l_{\mathrm{B}}}\right)^2\right] \tag{48}$$

$$\rho_{\mathrm{C}}(|\mathbf{x}-\mathbf{x}'|) = \frac{\sin\left(\frac{|\mathbf{x}-\mathbf{x}'|}{l_{\mathrm{C}}}\right)}{\frac{|\mathbf{x}-\mathbf{x}'|}{l_{\mathrm{C}}}} \tag{49}$$

where $|\mathbf{x}-\mathbf{x}'|$ denotes the distance between $\mathbf{x}$ and $\mathbf{x}'$, and $l_{\mathrm{A}}, l_{\mathrm{B}}$ and l_{C} are the correlation lengths of the corresponding correlation model.

Note that although the autocorrelation coefficient functions selected here are isotropic, the methods discussed here can cope with arbitrary types of autocorrelation coefficient functions (including non-isotropic and non-homogeneous types). The numerical procedures that were used to compute the examples presented in this paper can be applied without modification for problems with non-homogeneous types of autocorrelation coefficient functions. The error measures given in Eqs. (43) and (44) are already formulated for non-homogeneous problems.

4.3. 1D example

As a first example, a one-dimensional random field defined on a domain Ω of length one, i.e., $\Omega = [0,1]$, is investigated. The random field is assumed to have a constant standard deviation of $\sigma = 1$. The number of random variables in the expansion is fixed to $M = 10$.

An autocorrelation coefficient function of exponential type (type A in Eq. (47)) is selected. For this kind of problem, the Fredholm IEVP of Eq. (6) can be solved analytically [2]. The correlation length l_{A} is chosen such that the error $\varepsilon_{\mathrm{Var}}, \mathrm{ref}$ is approximately 5% ($l_{\mathrm{A}} = 0.42385$). It should be noted that the exponential correlation model is difficult to treat numerically because the kernel $\rho_{\mathrm{A}}(|\mathbf{x}-\mathbf{x}'|)$ is not differentiable at the diagonal, i.e., for $\mathbf{x} = \mathbf{x}'$. Therefore, numerical integration must be handled carefully. For the Nyström method, a non-differentiable kernel implies that the approximated eigenfunctions will also be non-differentiable.

In a first study, L^2-projection is compared to $H^{1/2}$-projection in the Galerkin-based approach. It is reminded that $H^{1/2}$ consists of a series of localized L^2-projections (compare Section 3.5.5). For L^2-projection, the matrices in Eq. (26) are assembled according to Eqs. (27) and (28). The isolines of the error measure $\varepsilon_{\mathrm{Var}}, \mathrm{rel}$ are depicted in Fig. 5 for different numbers of elements N_{el} and various maximum polynomial orders p. It can be seen that piecewise linear basis functions exhibit a slower rate of convergence than higher-order basis functions. Furthermore, L^2-projection leads to an error $\varepsilon_{\mathrm{Var}}, \mathrm{rel}$ that is orders of magnitude smaller than the error obtained with $H^{1/2}$-projection. Considering additionally the complexity of the $H^{1/2}$-approach, L^2-projection is clearly the one to favor.

Fig. 6 shows the same study as Fig. 5 using the error measure $\varepsilon_{\mathrm{Cov}}, \mathrm{rel}$ instead of $\varepsilon_{\mathrm{Var}}, \mathrm{rel}$. The reference error $\varepsilon_{\mathrm{Cov}}, \mathrm{ref}$ required in Eq. (46) was obtained by means of a numerical integration of Eq. (44) using the analytical eigenvalues and eigenfunctions. Comparing both plots, it can be seen that convergence with respect to $\varepsilon_{\mathrm{Cov}}, \mathrm{rel}$ is slower than for $\varepsilon_{\mathrm{Var}}, \mathrm{rel}$. However, the same conclusions can be drawn from Fig. 6 as were drawn from Fig. 5. In Fig. 6, the isoline corresponding to $\varepsilon_{\mathrm{Cov}}, \mathrm{rel} = 10^{-3}$ presents significant fluctuations between 4 and 15 elements. This implies that for seven elements and a polynomial order of three $\varepsilon_{\mathrm{Cov}}, \mathrm{rel} < 10^{-3}$, whereas for ten elements and a polynomial order of three $\varepsilon_{\mathrm{Cov}}, \mathrm{rel} > 10^{-3}$. Such fluctuations can occur for increasing N_{el} if p is fixed, because the basis functions of the case $N_{el} = 7 \wedge p = 3$ are not a subset of the basis functions of the case $N_{el} = 10 \wedge p = 3$. It is noted, however, that for fixed N_{el} and increasing p convergence is smooth, because the basis functions are hierarchic. Moreover, the fluctuations do not occur in the convergence plot of $\varepsilon_{\mathrm{Var}}, \mathrm{rel}$ (Fig. 5), which is minimized by the KL expansion.

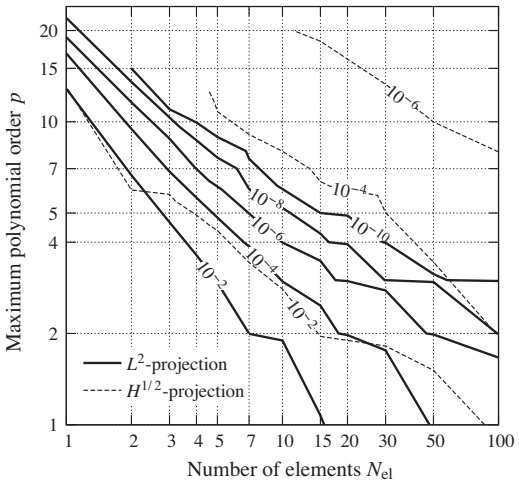

Fig. 5. Isolines of the relative error $\varepsilon_{\mathrm{Var,rel}}$ for a one-dimensional random field of length one with $\sigma = 1$ and correlation structure $\rho_A, l_A = 0.42385$; $M = 10$.

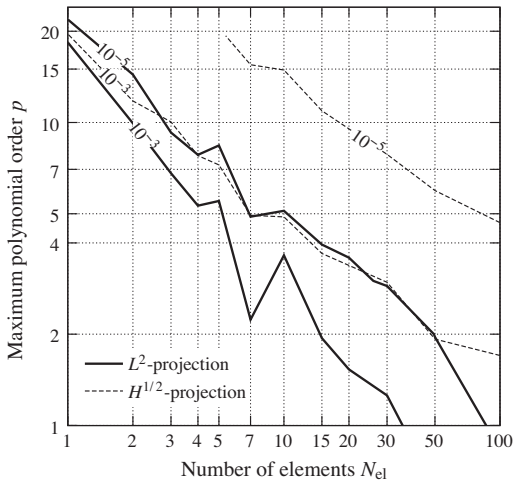

Fig. 6. Isolines of the relative error $\varepsilon_{\mathrm{Cov,rel}}$ for a one-dimensional random field of length one with $\sigma = 1$ and correlation structure $\rho_A, l_A = 0.42385$; $M = 10$.

Looking at the results presented in Figs. 5 and 6, it seems that the best approach would be to use only a single element and to employ basis functions with a large maximum polynomial order p. However, this approach is not recommended in most cases since a very large p can lead to numerical problems in the solution of the matrix eigenvalue problem. Therefore, we suggest to restrict the maximum polynomial order of the basis functions to $p \leq 12$.

The comparison of the projection methods, presented in Figs. 5 and 6, does not include the discrete projection method, because its error depends on the number of projection points used for given N_{el} and p. For an infinite number of points, discrete projection is equivalent to L^2-projection. In Fig. 7 the discrete projection (DP) method is compared with the Nyström method and the collocation method for an increasing number of points P. For the Nyström method only the case with equal weights is investigated, which is equivalent to the EOLE method. For the collocation method only piecewise linear basis functions are used. The number of basis functions is chosen equivalent to the number of points, i.e., $P = N$. For the discrete projection approach, two different sets of basis functions are applied: For the first set 20 elements with a maximum polynomial order of five are used, and for the second set 100 elements with piecewise linear basis functions are employed. Both sets have the same number of basis functions.

For the specified problem, the EOLE method, the collocation method and the discrete projection approach exhibit approximately the same rate of convergence. For the discrete projection approach it is important to note that the chosen set of basis functions must be capable of representing the eigenfunctions with the required accuracy. The set using piecewise linear basis functions lacks this quality for $\varepsilon_{\mathrm{Var}}, \mathrm{rel} \leqslant 5 \cdot 10^{-5}$. The advantage of the EOLE method and the discrete projection approach over the collocation method is that they do not require an explicit integration. Moreover, the problem can be solved faster with the EOLE method than with the discrete projection approach since in the EOLE method no multiplication

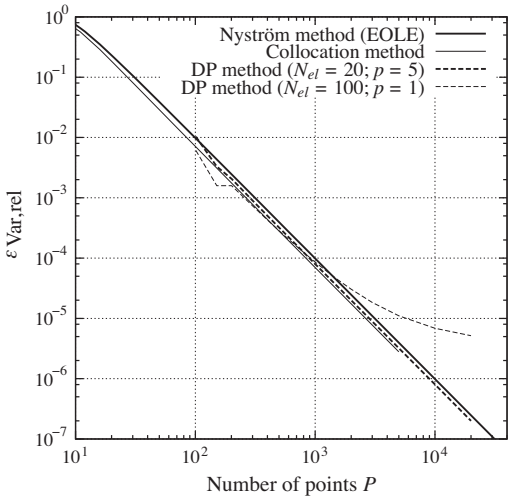

Fig. 7. Convergence behavior of the Nyström method, collocation method and discrete projection (DP) method for a random field with correlation structure $\rho_{\mathrm{A}}, l_{\mathrm{A}} = 0.42385$ and $\sigma = 1; M = 10$.

of matrices needs to be performed additionally to the solution of the matrix eigenvalue problem. Therefore, this method is easy to implement and due to its simplicity very efficient in obtaining an approximate solution.

Next, we examine the performance of the methods for the differentiable kernels ρ_{B} and ρ_{C} defined in Eqs. (48) and (49), respectively. As in the previous example, we take $M = 10$ random variables in the expansion and the correlation lengths l_{B} and l_{C} are chosen such that the error $\varepsilon_{\mathrm{Var}}$, ref is approximately 5% ($l_{\mathrm{B}} = 0.093065, l_{\mathrm{C}} = 0.031143$). It is noted that with differentiable kernels we can represent smaller correlation lengths than with non-differentiable kernels, while keeping the same number of random variables in the expansion. The reference error $\varepsilon_{\mathrm{Var}}$, ref was obtained numerically using L^2-projection with 1000 elements and a maximum polynomial order of ten. Special care was taken to ensure a good quality of the numerical integrals involved by choosing the number of Gauss-points so that convergence is achieved.

The isolines of the error measure $\varepsilon_{\mathrm{Var}}$, rel are depicted for the two correlation models in Figs. 8 and 9 for different numbers of elements N_{el} and various maximum polynomial orders p. Comparing both plots reveals that the two kernels ρ_{B} and ρ_{C} behave analogously. Moreover, the convergence behavior of L^2-projection is similar to the one shown in Fig. 5 for the exponential model. However, for the two differentiable correlation structures, $H^{1/2}$-projection performs better than for the non-differentiable structure of exponential type. This is because the differentiable kernels ρ_{B} and ρ_{C} can be approximated better by a projection onto the basis functions than the kernel ρ_{A}. Nevertheless, L^2-projection still converges much faster than $H^{1/2}$-projection.

For one-dimensional problems, the time needed to obtain a random field approximation is usually not an issue. Therefore, in principle any of the investigated methods can be applied for random field discretization. However, in most cases the EOLE

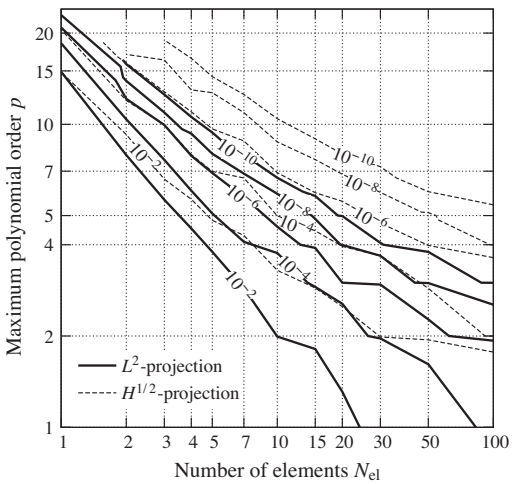

Fig. 8. Isolines of the relative error $\varepsilon_{\mathrm{Var,rel}}$ for a one-dimensional random field of length one with $\sigma = 1$ and correlation structure $\rho_{\mathrm{B}}, l_{\mathrm{B}} = 0.093065$; $M = 10$.

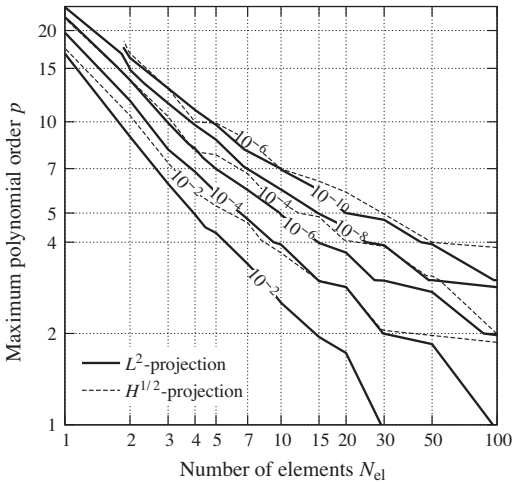

Fig. 9. Isolines of the relative error $\varepsilon_{\mathrm{Var,rel}}$ for a one-dimensional random field of length one with $\sigma = 1$ and correlation structure $\rho_{\mathrm{C}}, l_{\mathrm{C}} = 0.031143$; $M = 10$.

method is to be recommended, since it is relatively simple to implement and solves the problem within a relatively short time - without the need to perform an integration. On the other hand, the Galerkin approach with L^2-projection provides an elegant solution that exhibits a fast rate of convergence, especially in combination with higher-order basis functions.

4.4. 2D example (complex domain)

4.4.1. Squared exponential kernel

Mesh generation on one-dimensional domains is usually a straightforward and trivial task. The same cannot be said for two- or three-dimensional problems. In this example, the rectangular domain with a hole depicted in Fig. 10 is investigated. The autocorrelation coefficient function of the random field is assumed to be of squared exponential type ρ_{B}. The number of random variables in the expansion is fixed to $M = 30$. The random field is modeled with a constant standard deviation and the correlation length l_{B} was chosen such that the error $\varepsilon_{\mathrm{Var}}$, ref becomes approximately 5% ($l_{\mathrm{B}} = 0.77$). Three different approaches for random field discretization are compared: the FEM with bilinear basis functions, the FCM and the Nyström method with equal integration weights (i.e., the EOLE method). For the FCM, trunk space was used to construct the hierarchic two-dimensional basis functions.

In the linear FEM, 4-node quadrilateral elements are applied to mesh the physical domain Ω. The domain Ω is meshed by means of an unstructured mesh generator [49]; an exemplary finite element mesh is depicted in Fig. 11. The more elements are used, the better Ω is represented by the mesh. We investigate the convergence behavior of the random field discretization with respect to an increasing number of elements. Additional to the conventional L^2-projection, linear projection is applied as a special case of $H^{1/2}$-projection.

In the FCM, the physical domain Ω is embedded in a so-called primitive domain Ω^* of simple geometric shape. Three different structured meshes are studied for the discretization of Ω^*: $1 \times 1, 2 \times 2$ and 4×4 cells (Fig. 10). For each structured mesh, the convergence behavior of the random field discretization is investigated with respect to an increasing polynomial order of the basis functions of the cells.

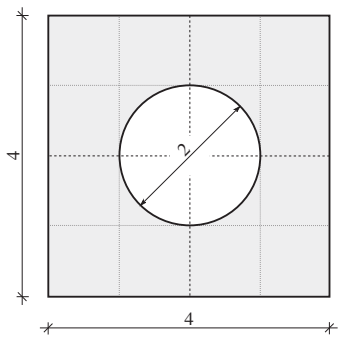

Fig. 10. Shape of the domain used in the 2D example.

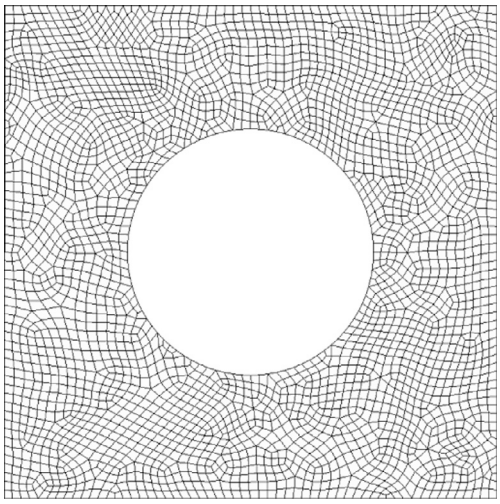

Fig. 11. Finite elment mesh with 2893 elements ($N = 3052$).

The EOLE method can be considered a truly meshless method. In this study, the integration points are distributed uniformly over the primitive domain Ω^* of simple geometric shape. Points located outside of the physical domain Ω are not considered in the analysis. The convergence behavior of the random field discretization is investigated with respect to an increasing number of points.

For each of the random field discretization methods, a matrix eigenvalue problem must be solved. In a first study, the convergence behavior of the relative error $\varepsilon_{\mathrm{Var},\mathrm{rel}}$ is investigated with respect to the size of the matrix eigenvalue problem. The reference error $\varepsilon_{\mathrm{Var},\mathrm{ref}}$ was obtained numerically as 0.049931 using the FCM with 4×4 cells and a maximum polynomial order of 15. The results of the analysis are depicted in Fig. 12. The finite cell discretization scheme exhibits an exponential rate of convergence. The influence of the coarseness of the finite cell mesh is small. A smaller number of cells results in a faster convergence. However, to reach the same relative error, a larger maximum polynomial order is required on a coarser mesh. Contrary to the exponential rate of convergence of the FCM, the EOLE method and the linear FEM show only a linear rate of convergence in the log–log plot in Fig. 12. The much faster convergence of the FCM compared to the FEM is due to the use of higher-order basis functions in the FCM. For the FEM, it is observed that linear projection converges slower than L^2-projection on linear basis functions. Therefore, going through the complex steps of implementing $H^{1/2}$-projection for higher-order basis functions appears to be dispensable. Since the size of the matrix eigenvalue problem can become rather large (i.e., $N \gg M$) for the FEM and the EOLE method, it is often considerable faster to compute only the M largest eigenvalues and corresponding eigenvectors instead of finding all the N eigenvalues and eigenvectors. This can be achieved by means of Lanczos algorithms for the solution of the matrix eigenvalue problem.

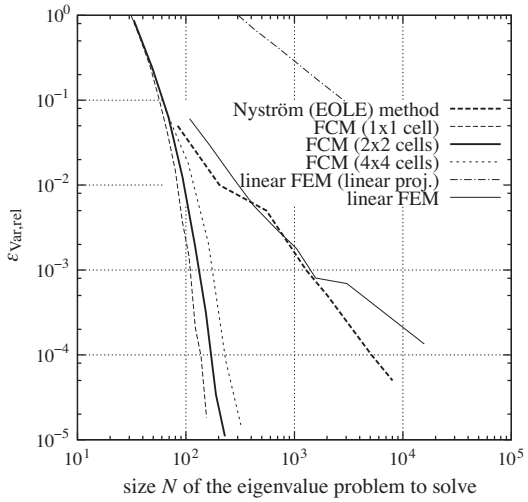

Fig. 12. Convergence of the relative error $\varepsilon_{\mathrm{Var,rel}}$ with respect to the size of the matrix eigenvalue problem to solve. (squared exponential kernel ρ_{B}, $M = 30$).

The size of the matrix eigenvalue problem to solve is only one property of the overall solution process. Another factor that has considerable influence on the overall computational cost is the assembly of the matrices, which differs significantly between the different methods. Therefore, for practical purposes, a comparison of the computational time needed to obtain an approximation of the KL expansion of the random field is of interest. This approximation is referred to as random field approximation in the following. The computational time to obtain a random field approximation for given relative errors ε_{Var}, rel is plotted in Fig. 13. The study was performed on a Intel® Core™ i7-3770 running at 3.40 GHz. All investigated methods were implemented in *C++*. Some effort was put into optimizing the implementations. In this regard, the EOLE method is most efficient. Except for errors smaller than $1 \cdot 10^{-3}$, the linear FEM is faster than the FCM. This is due to the time needed for the integration of discontinuous functions in the FCM; the faster rate of convergence due to the use of higher-order basis functions can compensate this only for errors smaller than $1 \cdot 10^{-3}$. Contrary to the finding of Fig. 12, it is better to use more than a single cell. Compared to the other investigated methods, the linear projection is very inefficient and, therefore, not of interest for practical applications. The efficiency of the linear FEM might be increased further by application of the $\mathcal{H}$-matrix approach proposed in [13,14].

If the random field is required as input for a non-intrusive finite element reliability analysis [10,50], in each run of the finite element method a realization of the random field needs to be evaluated at every Gauss-point. For this type of problem, a significant part of the overall runtime is spent in evaluating realizations of the random field. Therefore, it is relevant to compare the random field discretization methods with respect to the time needed to evaluate a single realization of the random field. In this regard, no other method is more efficient than the linear FEM. This is because independent of the accuracy of the random field approximation, the number of basis functions to evaluate for a realization of the field remains constant: for a one-, two- and three-dimensional element two, four and eight basis functions must be evaluated, respectively. On a Intel® Core™ i7-3770 running at 3.40 GHz this time is approximately $t_{\text{hFEM}} = 1.13 \cdot 10^{-6}$ s on a two-dimensional element. For all other methods, the time needed to obtain a realization was weighted with t_{hFEM}. The results are presented in Fig. 14. The computational costs of both the FCM and the EOLE method increase with a decreasing relative error ε_{Var}, rel. However, the EOLE method is more than an order of magnitude slower than the FCM. This is due to the fact that in the EOLE method (and for Nyström methods in general) the autocorrelation coefficient functions must be computed at every integration point, whereas in the FCM only the basis functions local to the cell must be evaluated.

Therefore, if the time spent in the evaluation of random field realizations has a major contribution to the overall runtime, the EOLE method is not the best choice for random field discretization. If a finite element mesh is available (e.g., as in non-intrusive finite element reliability analysis), the use of the FEM for the discretization of the random field is recommended. In this case, the same mesh that is used for the finite element reliability analysis should be used for the random field discretization as well. To understand the reason for this one needs to consider the fact that in FE reliability analysis, realizations of the random field need to be computed at the Gauss-points. If the same mesh is used, this is a trivial task, because the shape functions that have to be evaluated are known. If two different meshes are used, the element of the random field mesh in which the Gauss-point of interest is located has to be determined beforehand. Especially if unstructured meshes are used for the random field discretization, this is not a straightforward task. For very small correlation lengths, the existing finite element mesh might be too coarse to approximate the solution of the IEVP well with linear basis functions. In this case higher-order basis functions can be used to improve the quality of the random field approximation. The use of higher-order basis

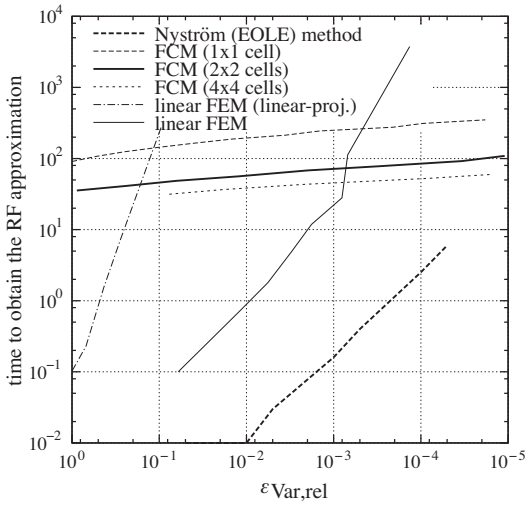

Fig. 13. Time needed to obtain an approximation of the KL expansion of the random field for a certain relative error $\varepsilon_{\text{Var,rel}}$. (squared exponential kernel ρ_{B}, $M = 30$).

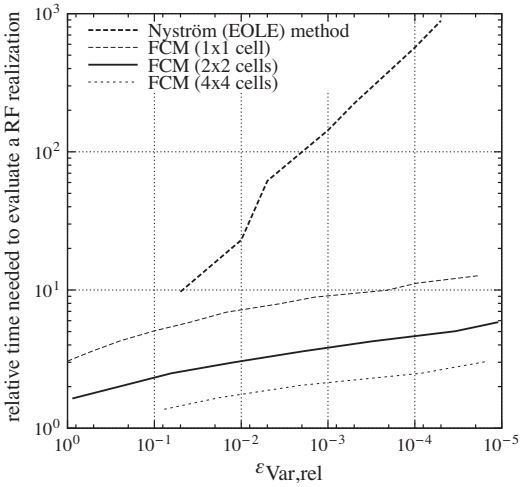

Fig. 14. Relative time needed to evaluate a realization of the random field for a certain relative error $\varepsilon_{\mathrm{Var,rel}}$. The results are relative to t_{hFEM}. (kernel ρ_{B}, $M = 30$).

functions is also recommended to check if the approximation has already converged with sufficient accuracy. If a mesh is not available, the FCM is the method of choice.

4.4.2. Exponential kernel

The squared exponential kernel ρ_{B} that was used in the previous example is a differentiable kernel. In practical applications often the non-differentiable exponential kernel ρ_{A} is used instead [19]. In this study, the same problem as in Section 4.4.1 is investigated, however, instead of the kernel ρ_{B} the kernel ρ_{A} is applied. Again, the correlation length is chosen such that the reference error is approximately 5% (i.e., $l_{\mathrm{A}} = 4.2$). The kernel ρ_{A} is non-differentiable at the diagonal, i.e., at $\mathbf{x} = \mathbf{x}'$. As is mentioned in Section 3.5.3, this introduces yet another difficulty in the integration of Eqs. (27) and (31). The reference error $\varepsilon_{\mathrm{Var}}$, ref was obtained numerically as 0.049954 using the FCM with 4×4 cells and a maximum polynomial order of 20. Special care was taken to ensure a good quality of the numerical integrals involved by choosing the number of Gauss-points so that convergence is achieved.

The time needed to obtain the random field approximation is shown in Fig. 15. Comparing Fig. 15 with the results obtained for the squared exponential kernel (Fig. 13), the most noticeable difference is the efficiency of the EOLE method with respect to the FEM. For the case with an exponential correlation structure the two methods behave similar, and for the case with the exponential squared correlation structure the EOLE method is clearly more efficient than the FEM. This is because the EOLE method approximates the eigenfunctions using the kernel $\mathrm{Cov}(\mathbf{x}, \mathbf{x}')$. If the kernel is not differentiable,

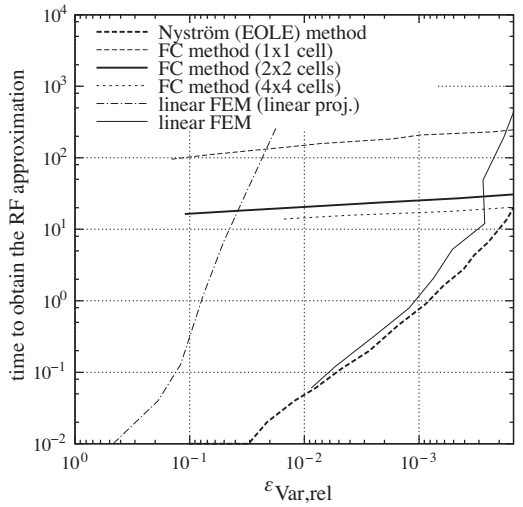

Fig. 15. Time needed to obtain an approximation of the KL expansion of the random field for a certain relative error $\varepsilon_{\mathrm{Var,rel}}$. (exponential kernel ρ_{A}, $M = 30$).

the theoretically smooth eigenfunctions are approximated by a linear combination of non-differentiable functions. For the exponential kernel, the relatively large time difference observed between using a single cell and using four cells in the FCM can be explained by a more efficient numerical treatment of the integrals in the later case (compare Section 3.5.3). Apart from that, the behavior of the investigated methods is similar to the one observed for the exponential squared correlation structure.

5. Conclusion

This paper compared methods for the numerical solution of the integral eigenvalue problem in the KL expansion in terms of the computational costs for obtaining a random field approximation and for evaluating a realization of the random field. Moreover, a novel approach for discretization of random fields based on the KL expansion was proposed, namely the use of the finite cell method (FCM). For the implementation of the FCM, a special integration technique as well as a novel hierarchic polynomial basis were introduced.

Based on the assessment, the EOLE method (as a special case of the Nyström method) is shown to be particularly efficient in obtaining a random field approximation. In our opinion, the EOLE method is the only true meshless method for the discretization of the KL expansion. Furthermore, the EOLE method is straightforward to implement - mainly because no integration is required to assemble the matrix eigenvalue problem. This is contrary to the Galerkin-based approaches (i.e., FEM and FCM), where the coefficients of the matrix eigenvalue problem are obtained from a two-folded integration over the domain of the random field. Additionally, the implementation of the FCM is involved, because the integrands become discontinuous and higher-order basis functions are essential to the method.

The advantage of the Galerkin-based procedures is an efficient evaluation of realizations of the field, whereby the FEM is faster than the FCM. In this regard, the FEM and the FCM clearly outperform the Nyström method. This is of interest if the time spent in the evaluation of random field realizations has a major contribution to the overall runtime (e.g., in finite element reliability analysis). If a finite element mesh is readily available, the use of the FEM is recommended. If for complex geometries a mesh is not available, the FCM is suggested. The proposed FCM can be considered a quasi-meshless method, because a domain of simple geometric shape is meshed and, therefore, the shape of the actual physical domain does not have to be taken into account for mesh generation. The use of higher-order basis functions in the FEM and the FCM provides fast convergence against the theoretical solution of the KL expansion for a given mesh. However, for most meshes generated for application of the linear FEM, linear basis functions provide already a good approximation on the available mesh.

The collocation method as well as the $H^{1/2}$-projection or the discrete projection approach were found to have no particular advantage compared to the other investigated methods.

Acknowledgments

With the support of the Technische Universität München – Institute for Advanced Study, funded by the German Excellence Initiative.

References

[1] B. Sudret, A. Der Kiureghian, Stochastic Finite Element Methods and Reliability – A State-of-the-Art Report, Tech. Rep. UCB/SEMM-2000/08, Department of Civil & Environmental Engineering, Univ. of California, Berkeley, 2000.
[2] R.-G. Ghanem, P.-D. Spanos, Stochastic Finite Elements – A Spectral Approach, Springer, New York, 1991.
[3] P.D. Spanos, R. Ghanem, Stochastic finite element expansion for random media, J. Eng. Mech. ASCE 115 (5) (1989) 1035–1053.
[4] K. Atkinson, The Numerical Solution of Integral Equations of the Second Kind, Cambridge University Press, Cambridge, 1997.
[5] S. Huang, S. Quek, K. Phoon, Convergence study of the truncated Karhunen–Loève expansion for simulation of stochastic processes, Int. J. Numer. Methods Eng. 52 (9) (2001) 1029–1043.
[6] R. Gutiérrez, J.C. Ruiz, M.J. Valderrama, On the numerical expansion of a second order stochastic process, Appl. Stoch. Models Data Anal. 8 (2) (1992) 67–77.
[7] K. Phoon, S. Huang, S. Quek, Implementation of Karhunen–Loève expansion for simulation using a wavelet-Galerkin scheme, Probab. Eng. Mech. 17 (3) (2002) 293–303.
[8] K. Phoon, H. Huang, S. Quek, Comparison between Karhunen–Loève and wavelet expansions for simulation of Gaussian processes, Comput. Struct. 82 (13) (2004) 985–991.
[9] G. Stefanou, M. Papadrakakis, Assessment of spectral representation and Karhunen–Loève expansion methods for the simulation of Gaussian stochastic fields, Comput. Methods Appl. Mech. Eng. 196 (21) (2007) 2465–2477.
[10] I. Papaioannou, Non-intrusive Finite Element Reliability Analysis, SVH-Verlag, Saarbrücken, 2013.
[11] P. Frauenfelder, C. Schwab, R.A. Todor, Finite elements for elliptic problems with stochastic coefficients, Comput. Methods Appl. Mech. Eng. 194 (2) (2005) 205–228.
[12] C. Schwab, R.A. Todor, Karhunen–Loève approximation of random fields by generalized fast multipole methods, J. Comput. Phys. 217 (1) (2006) 100–122.
[13] B.N. Khoromskij, A. Litvinenko, H.G. Matthies, Application of hierarchical matrices for computing the Karhunen–Loève expansion, Computing 84 (1–2) (2009) 49–67.
[14] D.L. Allaix, V.I. Carbone, Karhunen–Loève decomposition of random fields based on a hierarchical matrix approach, Int. J. Numer. Methods Eng.
[15] W. Hackbusch, A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices, Computing 62 (2) (1999) 89–108.
[16] S. Rahman, H. Xu, A meshless method for computational stochastic mechanics, Int. J. Comput. Methods Eng. Sci. Mech. 6 (1) (2005) 41–58.
[17] J. Parvizian, A. Düster, E. Rank, Finite cell method, Comput. Mech. 41 (2007) 121–133.
[18] C.-C. Li, A. Der Kiureghian, Optimal discretization of random fields, J. Eng. Mech. ASCE 119 (6) (1993) 1136–1154.

[19] B. Sudret, Uncertainity Propagation and Sensitivity Analysis in Mechanical Models – Contributions to Strucutral Reliability and Stochastic Spectral Methods, Université Blaise Pascal – Clermont II, Habilitation, 2007.
[20] K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, Ann. Acad. Sci. Fenn. 37 (1947) 3–79.
[21] M. Loève, Fonctions aleatoire du second ordre, supplement to P. Levy, in: Processus Stochastic et Mouvement Brownien, Gauthier-Villars, Paris, 1948.
[22] E. Vanmarcke, Random Fields: Analysis and Synthesis, second ed., World Scientific Publishing, Singapore, 2010.
[23] K. Phoon, S. Huang, S. Quek, Simulation of second-order processes using Karhunen–Loève expansion, Comput. Struct. 80 (12) (2002) 1049–1060.
[24] K. Phoon, H. Huang, S. Quek, Simulation of strongly non-Gaussian processes using Karhunen–Loève expansion, Probab. Eng. Mech. 20 (2) (2005) 188–198.
[25] R. Ghanem, Ingredients for a general purpose stochastic finite elements implementation, Comput. Methods Appl. Mech. Eng. 168 (1) (1999) 19–34.
[26] H.G. Matthies, A. Keese, Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations, Comput. Methods Appl. Mech. Eng. 194 (12) (2005) 1295–1331.
[27] M. Grigoriu, Crossing of non-Gaussian translation processes, J. Eng. Mech. ASCE 110 (41) (1984) 610–620.
[28] A. Der Kiureghian, P.-L. Liu, Structural reliability under incomplete probability information, J. Eng. Mech. ASCE 112 (1) (1986) 85–104.
[29] M. Grigoriu, Simulation of stationary non-Gaussian translation processes, J. Eng. Mech. ASCE 124 (2) (1998) 121–126.
[30] J. Hurtado, Analysis of one-dimensional stochastic finite elements using neural networks, Probab. Eng. Mech. 17 (1) (2002) 35–44.
[31] X. Wan, G. Karniadakis, A sharp error estimate for the fast Gauss transform, J. Comput. Phys. 219 (1) (2006) 7–12.
[32] H. Zhu, X. Zeng, W. Cai, J. Xue, D. Zhou, A sparse grid based spectral stochastic collocation method for variations-aware capacitance extraction of interconnects under nanometer process technology, in: Design, Automation & Test in Europe Conference & Exhibition, 2007. DATE'07, IEEE, 2007, pp. 1–6.
[33] W. Press, S. Teukolsky, W. Vetterling, B. Flannery, Numerical Recipes, The Art of Scientific Computing, third ed., Cambridge University Press, 2007.
[34] K. Atkinson, L. Shampine, Algorithm 876: solving Fredholm integral equations of the second kind in Matlab, ACM Trans. Math. Softw. (TOMS) 34 (4) (2008) 21.
[35] F. Tisseur, K. Meerbergen, The quadratic eigenvalue problem, SIAM Rev. 43 (2) (2001) 235–286.
[36] T.J. Hughes, The Finite Element Method: Linear Static and Dynamic Finite Element Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1987.
[37] B. Szabó, A. Düster, E. Rank, Encyclopedia of Computational Mechanics, Chap. The p-Version of the Finite Element Method, Wiley, 2004.
[38] A. Der Kiureghian, J. Ke, The stochastic finite element method in structural reliability, Probab. Eng. Mech. 3 (2) (1988) 83–91.
[39] C. Arun, B. Rao, S. Srinivasan, Stochastic meshfree method for elasto-plastic damage analysis, Comput. Methods Appl. Mech. Eng. 199 (37) (2010) 2590–2606.
[40] A. Düster, J. Parvizian, Z. Yang, E. Rank, The finite cell method for three-dimensional problems of solid mechanics, Comput. Methods Appl. Mech. Eng. 197 (2008) 3768–3782.
[41] D. Schillinger, M. Ruess, N. Zander, Y. Bazilevs, A. Düster, E. Rank, Small and large deformation analysis with the p- and B-spline versions of the finite cell method, Comput. Mech. 50 (4) (2012) 445–478.
[42] M. Ruess, D. Schillinger, Y. Bazilevs, V. Varduhn, E. Rank, Weakly enforced essential boundary conditions for NURBS-embedded and trimmed NURBS geometries on the basis of the finite cell method, Int. J. Numer. Methods Eng. (2013) (accepted for publication).
[43] M. Abramowitz, I.A. Stegun, Handbook of mathematical functions: with formulas, graphs, and mathematical tables, in: National Bureau of Standards, Applied Mathematics Series, 55, 1964.
[44] B. Szabó, I. Babuška, Finite Element Analysis, Wiley-Interscience, 1991.
[45] X. Zheng, S. Dong, An eigen-based high-order expansion basis for structured spectral elements, J. Comput. Phys. 230 (23) (2011) 8573–8602.
[46] A. Keese, A general purpose framework for stochastic finite elements (Ph.D. thesis), schoolTU Braunschweig, 2004.
[47] G. Karniadakis, S. Sherwin, Spectral/hp Element Methods for Computational Fluid Dynamics, Oxford University Press, Oxford, 2005.
[48] J. Zhang, B. Ellingwood, Orthogonal series expansions of random fields in reliability analysis, J. Eng. Mech. ASCE 120 (12) (1994) 2660–2677.
[49] C. Sorger, F. Frischmann, S. Kollmannsberger, E. Rank, TUM. GeoFrame: Automated high-order hexahedral mesh generation for shell-like structures, Eng. Comput. (2012) 1–16.
[50] G. Stefanou, The stochastic finite element method: past, present and future, Comput. Methods Appl. Mech. Eng. 198 (9) (2009) 1031–1051.