Uncertainty updating in the description of heterogeneous materials

A. Kučerová, H. G. Matthies

At a macroscopic scale, the details of mechanical behaviour are often uncertain, due to incomplete knowledge of details at small scales; this is especially acute if the materials have to be described before actually having been manufactured, such as in the case of concrete. Here the uncertainties are described by probabilistic methods, using recent numerical techniques for random fields based on white noise analysis. Numerical procedures are then developed to change/update the description once the materials have been manufactured, in order to take into account additional information, obtained for example from measurements. This results in an improved description of the uncertainties of the material behaviour.

1 Introduction

There are many phenomena of high engineering importance which cannot be analyzed straightforwardly due to insufficient data on material properties. This happens especially in the description of heterogeneous materials due to incomplete knowledge of details at small scales.

In recent decades, the development of multi-scale simulation tools has mainly been driven by the fundamental research in applied mathematics and theoretical mechanics with emphasis given to rigorous foundations of the homogenization theories for materials with well-defined geometry and components described by simple constitutive laws, see Šejnoha and Zeman (2008). Nevertheless, there are many engineering materials where the description of geometry remains non-trivial as in the case of concrete (Wriggers and Moftah (2006)), asphalt (Valenta et al. (2009)), or irregular masonry (Šejnoha et al. (2008)).

Another more elaborate possibility is casting the description of heterogeneous material in probabilistic framework. Here, a convenient way for describing uncertain material properties in time and/or space is through stochastic processes and fields, see Keese (2003); Matthies (2007). It means that instead of assuming material properties to be constant in a homogenized medium, their more realistic random nature can be conserved describing them by a finite number of random fields, see Ghanem and Kruger (1996), Matthies and Keese (2005), or Stefanou (2009).

The rigorous approach to quantify the probabilistic description of the material properties should include all possible sources of the information. The most augmented way to the estimation of material properties is based on fitting the response of the numerical model of the system into the results of real experiment, see Kučerová et al. (2007, 2009). Nevertheless, this approach usually completely omits the preliminary knowledge about the material properties, which can be represented for example by the physical meaning of particular properties (e.g. Young's modulus must be positive, Poisson's ratio must lay within the -1 and 0.5), or by our more detailed knowledge of the material constituents, its fractions and properties etc.

Bayesian inference is the statistical inference in which experimental observations are not only used as a source of information, they are also used to update the preliminary probability description of system - the so-called prior information - to deliver the posterior distribution (Kennedy and O'Hagan (2001)). In realistic applications, observations may be noisy, uncertain and limited in number relative to the dimension or complexity of the model space. Also, the forward model may have limitations on its predictive value because of its imprecision, filtering or smoothing effects. Taking into account all pertinent uncertainties, the process of material properties estimation cannot lead to a single 'optimal' parameter set, but one has to find a probability distribution of parameters that represents the knowledge about parameter values. The Bayesian setting for inverse problems offers a rigorous foundation for inference from noisy data and uncertain forward models, a natural mechanism for incorporating prior information, and a quantitative assessment of uncertainty in the inferred results summarising all available

information about the unknown quantity (Tarantola (2005)). In addition, unlike other techniques that aim to regularise the ill-posed inverse problem to achieve a point estimate, the Bayesian method treats the inverse problem as a well-posed problem in an expanded stochastic space.

The primary computational challenge remains one of extracting information from the posterior density (Mosegaard and Tarantola (2002)). Most estimates take the form of integrals over the posterior, which may be computed with asymptotic methods, deterministic methods, or sampling. The deterministic quadrature or cubature (Evans and Swartz (1995)) may be attractive alternatives to the Monte Carlo method at low to moderate dimensions, but the Markov chain Monte Carlo method (MCMC) (Tierney (1994), Gilks et al. (1995)) remains the most general and flexible method for complex and high-dimensional distributions. All of these methods, however, require the evaluation of the likelihood or posterior at many values of the model parameters. In this setting, evaluating the likelihood requires solving the forward problem. With complex forward models, such as those described by partial differential equations, each single evaluation can be computationally expensive (Higdon et al. (2003)). As Monte Carlo simulations require thousands or millions of samples, the total cost of these forward evaluations quickly becomes prohibitive. The efficient forward propagation of uncertainty - i.e. from model parameters to model predictions - is a central challenge of uncertainty quantification. This will be done here by expressing uncertainty through generalised stochastic (e.g. polynomial chaos) expansions to give a fast forward model surrogate. One example of this is also presented in Marzouk et al. (2007); Marzouk and Najm (2009).

Probabilistic or stochastic mechanics deals with mechanical systems, which are either subject to random external influences - a random or uncertain environment, or are themselves uncertain, or both, see e.g. the reports of Gutiérrez and Krenk (2004); Keese (2003); Matthies (2007). A powerful tool in computational stochastic mechanics is the stochastic finite element method (SFEM). SFEM is an extension of the classical deterministic FE approach to the stochastic framework, i.e. to the solution of stochastic (static and dynamic) problems involving finite elements whose properties are random, see Stefanou (2009). From a mathematical point of view, SFEM can be seen as a powerful tool for the solution of stochastic partial differential equations (SPDEs). One of the main variants of SFEM in the literature is the spectral stochastic finite element method SSFEM, which has been elaborated by other authors: Chen and Soares (2008); Ghanem and Kruger (1996); Matthies and Keese (2005); Stefanou (2009). In the initial version of the method presented in Chen and Soares (2008), the random properties of a structure are described by a Gaussian stochastic field which is represented using the Karhunen-Loève (KLE) expansion. Since then, much work has been done to model non-Gaussian fields (Colliat et al. (2007); Matthies et al. (1997); Mosegaard and Tarantola (2002)). Other possibilities to represent random properties are discussed e.g. in Lombardo et al. (2009). The resulting structural response may then be described using polynomial chaos expansion (PCE).

In this contribution we will limit our attention to the Bayesian update of a material property of an arbitrary heterogeneous material described by a Gaussian random field. For the sake of simplicity, different aspects of applied methods are studied in the framework of a simple conductivity problem. Particularly, the reduction of random variables via a Karhunen-Loève expansion and the construction of a surrogate model to a forward model based on a polynomial chaos is examined.

2 Setup

To introduce the reader into the subject of this paper, we start by defining the real *system* S under study, where system parameters m, loading f and the system response u are related as

$$\mathsf{S}(\mathsf{u};\mathsf{m}) = \mathsf{f}\,. \tag{1}$$

Then, one can derive an *explicit operator* V defining the system response u as a function of system parameters m and loading f

$$\mathsf{u} = \mathsf{V}(\mathsf{m};\mathsf{f}) \tag{2}$$

In the field of inverse problems, system parameters m are to be determined from some observed quantities d. Therefore, we define the operator F relating the complete model response u to the set of particular observations d

$$\mathsf{F}(\mathsf{m};\mathsf{u}) = \mathsf{F}(\mathsf{m};\mathsf{V}(\mathsf{m},\mathsf{f})) = \mathsf{d} \tag{3}$$

and finally we obtain an *observation operator* A mapping the response of a system given system parameters m and loading f to the observed quantities d

$$\mathsf{A}(\mathsf{m},\mathsf{f}) = \mathsf{d}\,. \tag{4}$$

The subject of this paper is concerned with the description of heterogeneous materials and therefore the system parameters m are related here to the material behaviour, i.e. they are material parameters. Moreover, the material parameters are usually determined in the context of a chosen experimental setup where the loading conditions are fixed. Hence, the loading f is assumed to be constant in the following text.

3 Bayesian inference

The Bayesian approaches to inverse problems have received much recent interest, since increasing performance of modern computers and clusters enables exhausting Monte Carlo computations. Among recent applications one can cite applications in environmental modelling (e.g. Hosack and Eldridge (2009), Rao (2007) or Yee et al. (2008)), hydrology (Fu and Gómez-Hernández (2009)), neuroscience (Quiros et al. (2010)) or heat transfer (Parthasarathy and Balaji (2008)). We review this approach briefly later; for more extensive introductions see Tarantola (2005) or Mosegaard and Tarantola (2002).

The main principle of Bayesian inference is to cast the inverse problem in a probabilistic setting, where the material parameters m as well as the observations d and also the response of observation operator $A(\cdot)$ are considered as random variables or random fields. Therefore, we introduce the following notation: We consider a set Ω of random elementary events ω together with a σ-algebra $\mathscr{S}$ to which a real number in the interval $[0, 1]$ may be assigned, the probability of occurrence - mathematically a measure $\mathbb{P}$.

In the Bayesian setting we assume three sources of information and uncertainties which should be taken into account. The first one is our prior knowledge about the model/material parameters $m(\omega)$, which is represented by defining the prior density function $p_{\mathsf{m}}(\mathsf{m})$. Prior models may embody simple constraints on m, such as a range of feasible values, or may reflect more detailed knowledge about the parameters, such as correlations or smoothness.

For two reasons the response of an observation operator $A(\cdot)$ is usually not exactly equal to the observed quantities: first, the measurements can be noisy including some measuring error $\eta(\bar{\omega})$, second, the observation operator $A(\cdot)$ can be imperfect, when for example the description of the real system S does not include all important phenomena, and therefore the observation operator response $A(\cdot, \bar{\bar{\omega}})$ can be assumed as uncertain. The probabilistic formulation of Eq. (4) is

$$\mathsf{d} = \mathsf{A}(\mathsf{m}(\omega), \bar{\bar{\omega}}) + \eta(\bar{\omega}) \,. \tag{5}$$

If the modelling uncertainties $\bar{\bar{\omega}}$ cannot be neglected, they can be described by a conditional probability density $p(\mathsf{d}|\mathsf{m})$ for predicted data d given model parameters m. If these uncertainties can be neglected, only model parameters $m(\omega)$ and observation $\mathsf{d}(\bar{\omega})$ remain uncertain. In practise, it is sometimes difficult to distinguish the imperfection of the system description S from measurement error η. Hence modelling uncertainties $\bar{\bar{\omega}}$ can be hidden in measuring error $\eta(\bar{\omega})$. Finally, for noisy measurements we define the last probability density $p_{\mathsf{d}}(\mathsf{d})$.

To update our prior knowledge about the model parameters we must include measurements with our theoretical knowledge. The Bayesian update is based on the idea of Bayes' rule defined for probabilities:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}, \tag{6}$$

where $P(\cdot)$ is the probability of particular events X and Y.

The definition of Bayes' rule for continuous distribution is, however, more problematic and hence Tarantola (2005) derives the posterior state of information $\pi(\mathsf{m}, \mathsf{d})$ as a conjunction of all information at hand

$$\pi(\mathsf{m}, \mathsf{d}) = \kappa p_{\mathsf{m}}(\mathsf{m}) p_{\mathsf{d}}(\mathsf{d}) p(\mathsf{d}|\mathsf{m}) \,, \tag{7}$$

where κ is a normalisation constant.

The posterior state of information defined in the space of model parameters m is given by the marginal probability density

$$\pi_{\mathsf{m}}(\mathsf{m}) = \mathbb{E}_{\bar{\omega}}\left[\pi(\mathsf{m}, \mathsf{d})\right] = \kappa p_{\mathsf{m}}(\mathsf{m}) \int_{\bar{\Omega}} p(\mathsf{d}|\mathsf{m}) p_{\mathsf{d}}(\mathsf{d}) \, \mathbb{P}(\mathrm{d}\bar{\omega}) = \kappa p_{\mathsf{m}}(\mathsf{m}) \, L(\mathsf{m}), \tag{8}$$

where measured data d enter through the *likelihood function* $L(\mathsf{m})$, which gives a measure of how well an observation operator $A(\mathsf{m})$ explains the data d. Here $\mathbb{E}_{\bar{\omega}}$ is the expectation operator averaging over $\bar{\Omega}$.

To keep the presentation of different numerical aspects of particular methods clear and transparent, we focus here on a quite common and simple case where modelling-uncertainties are neglected and measurement errors are assumed to be Gaussian. Then the likelihood function takes the form

$$L(\mathsf{m}) = \kappa \exp\left(-\frac{1}{2}\left(\mathsf{A}(\mathsf{m}) - \mathsf{d}\right)^{\mathrm{T}} C_{\mathrm{obs}}^{-1}\left(\mathsf{A}(\mathsf{m}) - \mathsf{d}\right)\right), \tag{9}$$

where C_{obs} is a covariance among measurements d.

The Bayesian estimation typically gives rise to integrals over the posterior density. Several sophisticated means of estimating such integrals are discussed, e.g., in Marzouk et al. (2007) and it is shown that Markov chain Monte Carlo method (MCMC) encompasses a wide class of methods that simulate drawing samples from the posterior.

4 Representation of heterogeneity and reduction of random variables

When we are modelling heterogeneous material, some material parameters are not constants, but can be described as random fields. It means that the uncertainty in the particular material parameter m is modelled by defining $m(\boldsymbol{x})$ for each $\boldsymbol{x} \in \mathcal{G}$ as a random variable $m(\boldsymbol{x}) : \Omega \to \mathbb{R}$ on a suitable probability space $(\Omega, \mathscr{S}, \mathbb{P})$ in the bounded admissible region $\mathcal{G} \subset \mathbb{R}^d$. As a consequence, $m : \mathcal{G} \times \Omega \to \mathbb{R}$ is a random field and one may identify Ω with the set of all possible values of m or with the space of all real-valued functions on $\mathcal{G}$. Alternatively, $m(\boldsymbol{x}, \omega)$ can be seen as a collection of real-valued random variables indexed by $\boldsymbol{x} \in \mathcal{G}$.

Assuming the random field $m(\boldsymbol{x}, \omega)$ to be Gaussian, it is defined by its mean

$$\mu_m(\boldsymbol{x}) = \mathbb{E}[m(\boldsymbol{x}, \omega)] = \int_\Omega m(\boldsymbol{x}, \omega)\,\mathbb{P}(\mathrm{d}\omega) \tag{10}$$

and its covariance

$$C_m(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(m(\boldsymbol{x}, \omega) - \mu_m(\boldsymbol{x}))(m(\boldsymbol{x}', \omega) - \mu_m(\boldsymbol{x}'))] = \int_\Omega (m(\boldsymbol{x}, \omega) - \mu_m(\boldsymbol{x}))(m(\boldsymbol{x}', \omega) - \mu_m(\boldsymbol{x}'))\,\mathbb{P}(\mathrm{d}\omega). \tag{11}$$

Some non-Gaussian fields may be synthesised as functions of Gaussian fields (see e.g. Grigoriu (1995) or Matthies (2007)), that is $m(\boldsymbol{x}, \omega) = \Xi(\boldsymbol{x}, g(\boldsymbol{x}, \omega))$, where Ξ is a pointwise - usually nonlinear - transformation, and $g(\boldsymbol{x}, \omega)$ is a Gaussian field. Therefore in the following text we focus on Gaussian fields.

In a computational setting, this field and the forward model must be discretised. If a parameter field $m(\boldsymbol{x})$ can be adequately represented on a finite collection of points $\boldsymbol{x}_{i\,i=1}^{n} \in \mathbb{R}$, then we can write both the prior and posterior densities in terms of $\boldsymbol{m} = (m(\boldsymbol{x}_1), \ldots, m(\boldsymbol{x}_n))$, where $m_i = m(\boldsymbol{x}_i)$ are random variables usually correlated among each other. The vector $\boldsymbol{m}$, however, will likely be high-dimensional and it renders an MCMC exploration of the posterior more challenging and costly. Therefore, the Karhunen−Loève expansion can be applied for dimensionality reduction (Marzouk and Najm (2009)).

Karhunen−Loève expansion (KLE) is an extremely useful tool for the concise representation of the stochastic processes. Based on the spectral decomposition of covariance function $C_m(\boldsymbol{x}, \boldsymbol{x}')$ and the orthogonality of eigenfunctions ϕ_i, the random field $m(\boldsymbol{x}, \omega)$ can be written as

$$m(\boldsymbol{x}, \omega) = \mu_m(\boldsymbol{x}) + \sum_{i=0}^{\infty} \sqrt{\lambda_i}\xi_i(\omega)\phi_i(\boldsymbol{x}), \tag{12}$$

where $\boldsymbol{\xi}(\omega)$ is a set of uncorrelated random variables of zero mean and unit variance. In case of several stochastically dependent fields (several different but dependent material parameters), a joint KLE has to be computed. Otherwise each field may be expanded by itself.

The spatial KLE functions $\phi_i(\boldsymbol{x})$ are the eigenfunctions of the Fredholm integral equation with the covariance function as the integral kernel

$$\int_{\mathcal{A}} C_m(\boldsymbol{x}, \boldsymbol{x}')\phi_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x}' = \lambda_i\phi_i(\boldsymbol{x}), \tag{13}$$

where λ_i are positive eigenvalues ordered in a descending order.

Since the covariance is symmetric and positive definite, it can thus be expanded in the series

$$C_m(\boldsymbol{x}, \boldsymbol{x}') = \sum_{i=1}^{\infty} \lambda_i \phi_i(\boldsymbol{x}) \phi(\boldsymbol{x}') \,. \tag{14}$$

However, computing the eigenfunctions analytically is usually not feasible. Therefore, one discretises the covariance spatially according to chosen grid points (usually corresponding to a finite element mesh). The resulting covariance matrix $\mathbf{C}_m$ is again symmetric and positive definite, and Eq. (13) becomes a symmetric matrix eigenvalue problem (Matthies and Keese (2005)), where the eigenfunctions $\phi_i(\boldsymbol{x})$ are replaced by eigenvectors ϕ_i. The eigenvalue problem is usually solved by a Krylov subspace method with a sparse matrix approximation. For large eigenvalue problems, Khoromskij and Litvinenko (2008) propose the efficient low-rank and data sparse hierarchical matrix technique.

For practical implementation, the series (12) and (14) are truncated after N_ξ terms, yielding the approximations

$$\hat{\boldsymbol{m}}(\omega) \quad \approx \quad \mu_m + \sum_{i=1}^{N_\xi} \sqrt{\lambda_i} \xi_i(\omega) \phi_i \,, \tag{15}$$

$$\hat{\mathbf{C}}_m \quad \approx \quad \sum_{i=1}^{N_\xi} \lambda_i \phi_i^T \phi_i \,, \tag{16}$$

Such spatial semi-discretisation is optimal in the sense that the mean square error resulting from a truncation after the N_ξ-th term is minimised.

5 Surrogate of forward model

Although the KLE is a very effective tool to limit the dimensionality of the inverse problem at hand, MCMC exploration still requires repeated solutions of the forward model, once for each proposed move of the Markov chain. Therefore, it is desirable to avoid repeated forward solutions altogether. Marzouk and Najm (2009); Marzouk et al. (2007) introduced methods for accelerating Bayesian inference in this context by using stochastic spectral methods to propagate prior uncertainty through the forward problem. Polynomial chaos expansion (PCE) is a suitable tool able to circumvent the described problem.

According to the previous section, the vector of model parameters $\boldsymbol{m}(\omega)$ represents now the discretised spatial field $m(\boldsymbol{x}, \omega)$. Since its nature is random, it is described by a limited number N_ξ of uncorrelated random variables $\boldsymbol{\xi}(\omega) = \xi_1(\omega) \ldots \xi_{N_\xi}(\omega)$ according to Eq. (15). Hence, the discretised system response $\boldsymbol{u}(\omega) = (\ldots, u_i \ldots)^{\mathrm{T}} = \boldsymbol{V}(\boldsymbol{m}(\omega))$ has also the random nature and can be expressed in terms of the same random variables $\boldsymbol{\xi}(\omega)$. If $\boldsymbol{\xi}(\omega)$ are standard Gaussian random variables, they can be utilised to construct the surrogate $\tilde{\boldsymbol{u}}(\boldsymbol{\xi}(\omega))$ of the model response $\boldsymbol{u}(\boldsymbol{\xi}(\omega))$ as the Wiener polynomial chaos expansion

$$\tilde{\boldsymbol{u}}(\boldsymbol{\xi}) = \sum_{\alpha} \boldsymbol{\beta}_\alpha H_\alpha(\boldsymbol{\xi}(\omega)) \tag{17}$$

where $\boldsymbol{\beta}_\alpha = (\ldots, \beta_{\alpha,i} \ldots)^{\mathrm{T}}$ is a vector of PC coefficients $\beta_{\alpha,i}$ corresponding to a particular component of system response $u_i(\omega)$. $H_\alpha(\boldsymbol{\xi}(\omega))$ are multivariate Hermite polynomials

$$H_\alpha(\boldsymbol{\xi}(\omega)) = \prod_{j=1}^{\infty} h_{\alpha_j}(\xi_j(\omega)) \,, \tag{18}$$

and h_α are the usual univariate Hermite polynomials. The expansion (17) is usually truncated to the limited number of terms N, which is very often related to the dimension N_ξ of vector $\boldsymbol{\xi}(\omega)$ and to the maximal degree of polynomials N_p according to the relation

$$N = \frac{(N_p + N_\xi)!}{N_p! N_\xi!} \,. \tag{19}$$

The Equation (17) can also be rewritten using the matrix notation

$$\tilde{\boldsymbol{u}}(\boldsymbol{\xi}) = (\mathbf{I} \otimes \boldsymbol{H}^*(\boldsymbol{\xi})) \cdot \boldsymbol{\beta} \,, \tag{20}$$

where $\mathbf{I} \in \mathbb{R}^{M,M}$ is the unity matrix, $\boldsymbol{H}(\boldsymbol{\xi})$ is an N-dimensional vector of Hermite polynomials, and $\boldsymbol{H}^*(\boldsymbol{\xi})$ is its dual, $\boldsymbol{\beta} = \left(\dots, \boldsymbol{\beta}_\alpha^{\mathrm{T}}, \dots \right)^{\mathrm{T}} \in \mathbb{R}^{M \cdot N}$ is a vector of PC coefficients, M is the number of chosen grid points (e.g. nodes in FE mesh) discretising the spatial domain $\mathcal{G}$, and $\otimes$ is a Kronecker product.

After an appropriate discretisation of Eq. (1), the model of the system is usually given by a system of equations

$$\mathbf{S}(\boldsymbol{u}, \boldsymbol{m}) = \boldsymbol{f} \,. \tag{21}$$

For the sake of simplicity, we focus now on the linear case, i.e. $D_u \mathbf{S}(\boldsymbol{u}, \boldsymbol{m}) = \mathbf{K}$ then

$$\mathbf{K}(\boldsymbol{m})\boldsymbol{u} = \boldsymbol{f} \,. \tag{22}$$

Substituting the model response $\boldsymbol{u}$ by its PC approximation $\tilde{\boldsymbol{u}}$ given in Eq. (20) and applying Galerkin conditions we can compute PC coefficients:

$$\int_\Omega \boldsymbol{H}(\boldsymbol{\xi}) \otimes \mathbf{K}(\boldsymbol{m}(\boldsymbol{\xi})) \otimes \boldsymbol{H}^*(\boldsymbol{\xi}) \, \mathrm{d}\mathbb{P}(\omega) \cdot \boldsymbol{\beta} - \int_\Omega \boldsymbol{H}(\boldsymbol{\xi}) \, \mathrm{d}\mathbb{P}(\omega) \otimes \boldsymbol{f} = \mathbf{0} \,. \tag{23}$$

If the operator $\mathbf{K}(\boldsymbol{m})$ is linear also with respect to model parameters $\boldsymbol{m}$, it can be expanded according to Eq. (15) as

$$\tilde{\mathbf{K}}(\tilde{\boldsymbol{m}}(\boldsymbol{\xi})) \approx \bar{\mathbf{K}} + \sum_{i=1}^{N_\xi} \sqrt{\lambda_i} \xi_i \mathbf{K}_i \tag{24}$$

and then substituted in Eq. (23):

$$\left[\bar{\mathbf{K}} \otimes \int_\Omega \boldsymbol{H}(\boldsymbol{\xi}) \otimes \boldsymbol{H}^*(\boldsymbol{\xi}) \, \mathrm{d}\mathbb{P}(\omega) + \sum_{i=1}^{N_\xi} \sqrt{\lambda_i} \mathbf{K}_i \otimes \int_\Omega \xi_i \cdot \boldsymbol{H}(\boldsymbol{\xi}) \otimes \boldsymbol{H}^*(\boldsymbol{\xi}) \, \mathrm{d}\mathbb{P}(\omega) \right] \cdot \boldsymbol{\beta} - \int_\Omega \boldsymbol{H}(\boldsymbol{\xi}) \, \mathrm{d}\mathbb{P}(\omega) \otimes \boldsymbol{f} = \mathbf{0} \,. \tag{25}$$

The integrals in Eq. (25) can be solved analytically or numerically depending on their complexity. Eq. (25) then results in a system of $M \cdot N$ equations for PC coefficients $\boldsymbol{\beta}$. The computation can be slightly simplified by normalising the Hermite polynomials. Then $\int_\Omega \boldsymbol{H}(\boldsymbol{\xi}) \otimes \boldsymbol{H}^* \mathrm{d}\mathbb{P}(\boldsymbol{\xi})$ is the identity matrix, and $\int_\Omega \boldsymbol{H}(\boldsymbol{\xi}) \mathrm{d}\mathbb{P}(\boldsymbol{\xi})$ is the first unit vector.

It is clear that the complexity of PC expansion increases quite fast with the increasing degree of polynomials N_p and/or dimensions N_ξ of $\boldsymbol{\xi}(\omega)$ according to Eq. (19). The number of random variables is already reduced by transformation from vector of correlated random variables $\boldsymbol{m}(\omega)$ into uncorrelated variables $\boldsymbol{\xi}(\omega)$. These variables are related to eigenvalues of different magnitude expressing the importance of particular eigenmodes. One possibility to decrease N is to neglect terms in PC expansion, where less important variables ξ_i are in higher orders (Matthies (2007)). The maximal degree of polynomials p can be also reduced e.g. by a division of modelled domain $\mathcal{G}$ as presented e.g. by Marzouk et al. (2007). Other computational savings can be achieved by formulating Eq. (17) in tensorial notation instead of matrix one (Matthies (2007)).

The derivation was done for the Hermite basis. But the development is completely general and holds also for any other basis - except maybe for the orthonormality relation; e.g. such as the Wiener-Askey generalised polynomial chaos (Xiu and Karniadakis (2002)).

These methods effectively create a 'surrogate' containing PC representations of the forward model outputs. This model approximation may be evaluated in orders of magnitude more quickly than the 'direct' evaluation containing the full forward problem.

6   Numerical example

For the sake of simplicity we consider the model of heat transport under steady state conditions. To be more specific, the description of heat transfer is governed by the following energy balance equation,

$$0 = \nabla \left[k \nabla T \right], \tag{26}$$

where k is the thermal conductivity and T is the temperature. For the spatial discretisation of the partial differential equations, a finite element method is preferred here to the finite volume technique (Sýkora et al. (2009)). Using the weighted residual statement and applying Green's theorem then yields

$$0 = -\int_{\mathcal{G}} \nabla \delta w_T \left[k \nabla T \right] \mathrm{d}\mathcal{G} + \int_{\Gamma_T^{\bar{q}}} \delta w_T \, \bar{q}_\nu \mathrm{d}\Gamma, \tag{27}$$

where δw_T is the weighting function such as $\delta w_T=0$ on Γ_T, $\bar{q}_\nu$ is the prescribed heat flux perpendicular to the boundary Γ_T^q. On Γ_T the temperature T is equal to its prescribed value T. Vector ν stores the components of the unit outward normal. Using approximation $T(\boldsymbol{x}) = \boldsymbol{N}(\boldsymbol{x})\,\boldsymbol{T}$ converts Eq. (27) into the form

$$\mathbf{K}_T(k)\,\boldsymbol{T} = \boldsymbol{q}_{\mathrm{ext}}. \tag{28}$$

Geometry and FE mesh (80 nodes and 120 elements) of a chosen experiment together with the specific loading conditions are shown in Figure 1.
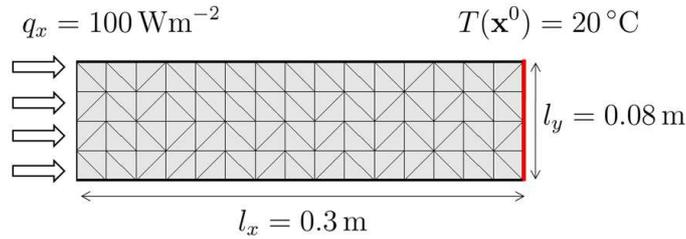
Figure 1: Experimental setup.

As a priori knowledge we assume that k can be described as a Gaussian random field with the mean value $\mu_k = 2\,\mathrm{Wm}^{-1}\mathrm{K}^{-1}$ and *exponential* two-dimensional covariance kernel

$$C(\boldsymbol{x}, \boldsymbol{x}\prime) = \sigma_k^2 \mathrm{e}^{-\left|\frac{x-x\prime}{L_x}\right| - \left|\frac{y-y\prime}{L_y}\right|}, \tag{29}$$

where the standard deviation $\sigma_k = 0.3\ \mathrm{Wm}^{-1}\mathrm{K}^{-1}$ and correlation lengths are $L_x = 0.1$ m and $L_y = 0.04$ m, if not specified otherwise. This choice of correlation lengths corresponds to material with a 'coarse' microstructure. Some effects of different correlation lengths are also examined in the following text.

As mentioned in the previous section, the complexity of the surrogate of forward model based on PC expansion is quickly increasing with the number of random variables ξ_i used for the description of the input random field. Therefore, it is convenient to reasonably limit the number of terms in the KL expansion (15), but attention should be paid to the related error in the description of the input field itself as well as in the subsequent response of the model. The comparison of a realisation of the k-field using different numbers of terms in KLE can be seen in Figure 2. Figure 2a shows the k-field constructed using all $N_\xi = 120$ eigenmodes of covariance matrix, while its approximation $\hat{k}$ built up with the same values of random variables ξ, but involving only 6 first eigenmodes, is depicted in Figure 2b. The corresponding pointwise error $|k(\boldsymbol{x}) - \hat{k}^{(N_\xi=6)}(\boldsymbol{x})|$ is plotted in Figure 3a. The error is then summarised over the whole discretised domain $\mathcal{G}$ and averaged over 100 random choices of random vectors $\boldsymbol{\xi}$ according to

$$E(\boldsymbol{k}, \hat{\boldsymbol{k}}) = \frac{1}{100} \sum_{j}^{100} \sum_{i=1}^{M} |k_i(\boldsymbol{\xi}_j) - \hat{k}_i^{(N_\xi)}(\boldsymbol{\xi}_j)|. \tag{30}$$

Figure 3b shows the evolution of error $E(\boldsymbol{k}, \hat{\boldsymbol{k}})$ as a function of a number of modes $N_\xi = 1 \ldots 120$ for several different choices of correlation lengths. One can see that smaller correlation length leads to higher error in approximation and this error also decreases slower with increasing number of eigenmodes.

Since the covariance kernel given in Eq. (29) is stationary, i.e. its value does not depend on the particular position of points $\boldsymbol{x}$ and $\boldsymbol{x}'$, but on their relative position, the covariance can be easily visualised in terms of the difference between points $\boldsymbol{x}$ and $\boldsymbol{x}'$, see Figure 4a. For comparison, an approximation of the covariance $\hat{C}$ based on 6 eigenmodes is shown in Figure 4b.

For a better examination of the influence of a thermal conductivity field description to a model response, we extract the average temperature field $\bar{T}(\boldsymbol{x})$ (see Figure 5a) obtained for the mean value of field $k(\boldsymbol{x}) = \mu_k(\boldsymbol{x}) =$
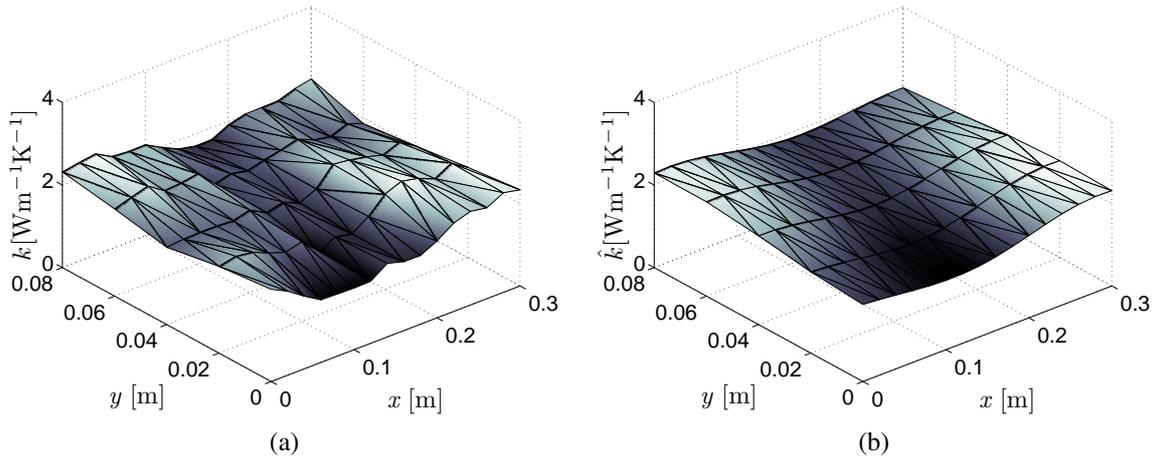
Figure 2: Realisation of k random field using (a) all eigenmodes or (b) only 6 eigenmodes.
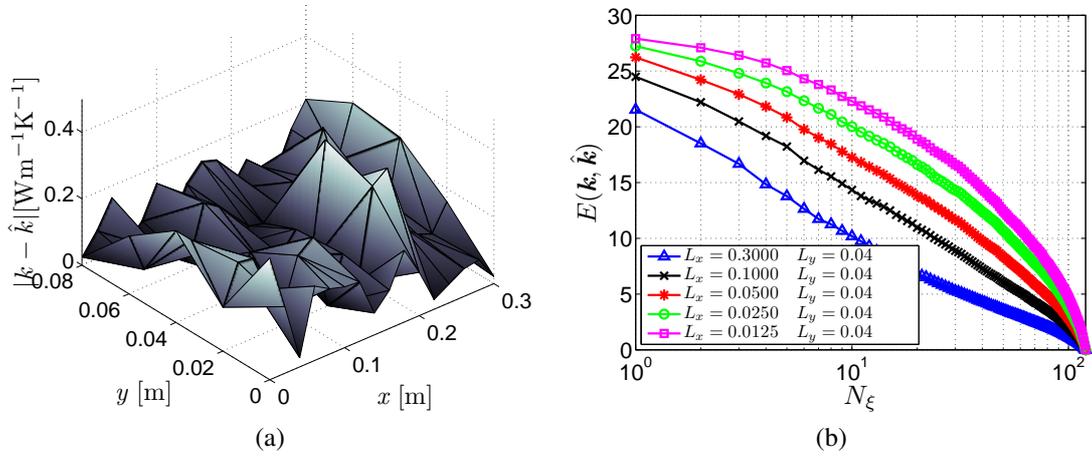
Figure 3: Error of random field approximation: (a) pointwise error $|k(\boldsymbol{x}) - \hat{k}^{(N_\xi=6)}(\boldsymbol{x})|$; (b) evolution of summarised and averaged pointwise error $E(\boldsymbol{k}, \hat{\boldsymbol{k}})$ as a function of number of eigenvectors involved.
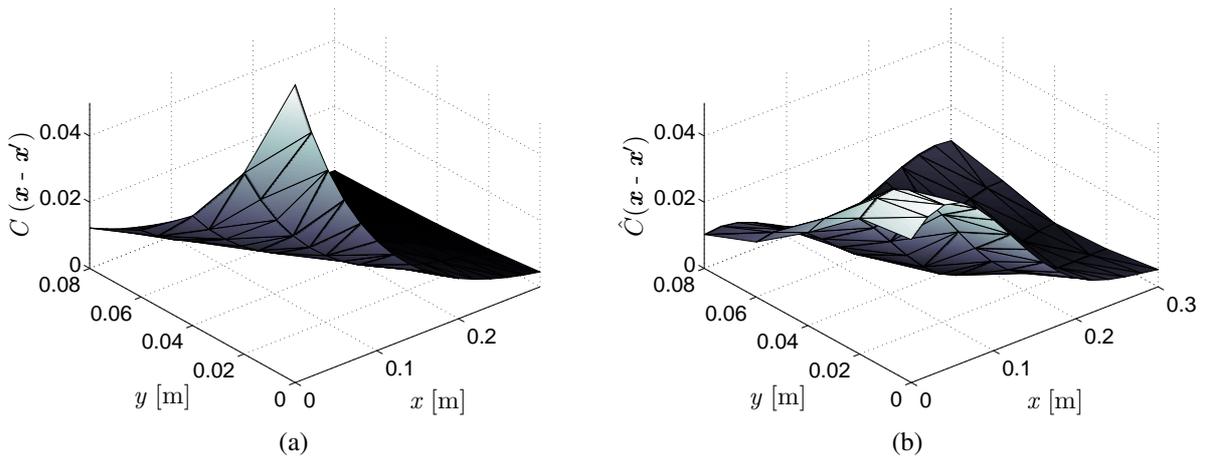
Figure 4: Covariance as a function of difference between points $(\boldsymbol{x} - \boldsymbol{x}\prime)$: (a) full, i.e. involving $N_\xi = 120$ eigenmodes; (b) reduced for $N_\xi = 6$ eigenmodes.

$2\,\mathrm{Wm^{-1}K^{-1}}$. Fluctuations of temperature field $T(\boldsymbol{x})$ computed for a particular realisation of conductivity field from Figure 2a based on a complete set of KLE terms are shown in Figure 5b.

Again, the pointwise error between temperature field $T(\boldsymbol{x})$ and its surrogate $\hat{T}(\boldsymbol{x})$ computed for the approximated conductivity field $\hat{k}(\boldsymbol{x})$ from Figure 2b is depicted in Figure 6a. The error can be also summarised similarly to
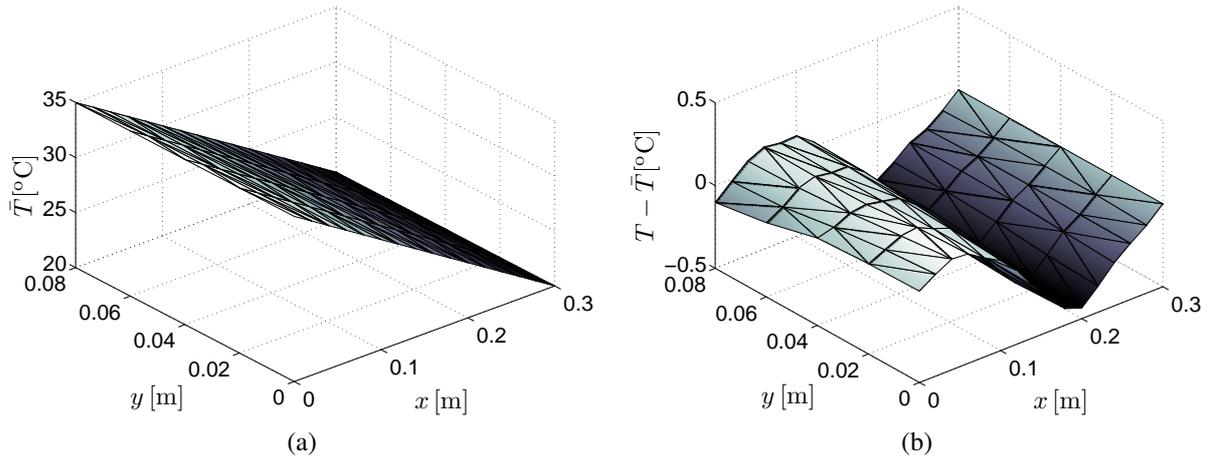
Figure 5: (a) Temperature field $\bar{T}(\boldsymbol{x})$ obtained for an averaged field $k(\boldsymbol{x}) = 2\,\mathrm{Wm^{-1}K^{-1}}$; (b) pointwise difference between temperature field $\bar{T}(\boldsymbol{x})$ and temperature field $T(\boldsymbol{x})$ computed for $k(\boldsymbol{x})$ shown in Figure 2a.

Eq. (30), but in terms of discrete values of temperature fields. The resulting error $E(\boldsymbol{T}, \hat{\boldsymbol{T}})$ evaluated for different correlation lengths and different numbers N_ξ of eigenmodes involved is shown in Figure 6b.
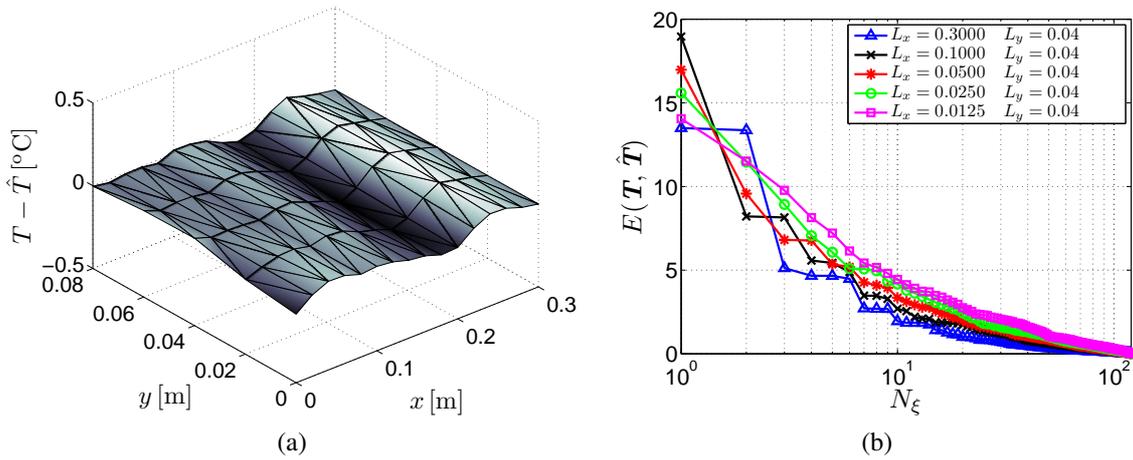
Figure 6: Error of random field approximation propagated into the response of the model (a) pointwise error $|T(\boldsymbol{x}) - \hat{T}^{(N_\xi=6)}(\boldsymbol{x})|$; (b) evolution of summarised and averaged pointwise error $E(\boldsymbol{T}, \hat{\boldsymbol{T}})$ as a function of number of eigenvectors involved.

One should notice that in terms of forward model response, smaller correlation length does not mean higher error in the response field. However, this phenomenon cannot be generalised and particularly it will not be true in case of models describing some localised effects as, e.g., damage models.

Note that surrogate $\hat{T}(\boldsymbol{x})$ is computed for the approximation of conductivity field $\hat{k}(\boldsymbol{x})$ using full model $\hat{\boldsymbol{T}} = \mathbf{V}(\hat{\boldsymbol{k}}, \cdot)$. When the full model is replaced by its PC approximation, another surrogate $\tilde{T}(\boldsymbol{x})$ is constructed. The error given in Eq. (30) can be now expressed in terms of discretised temperature field $\boldsymbol{T}$ and its new discretised surrogate $\tilde{\boldsymbol{T}}$ as $E(\boldsymbol{T}, \tilde{\boldsymbol{T}})$, and its value can be compared for different maximal degrees of Hermite polynomials and again for different eigenmodes involved as well as for different choices of correlation length in Figure 7.

Beside the accuracy also the time necessary for the computation of polynomial chaos coefficients and the computing of samples is very important. Such a comparison for computing 100.000 samples via direct evaluation of a full FE model and via PCE involving PC coefficients computation is presented in Figure 8 for different numbers of eigenmodes taken into account and for different maximal degrees of polynomials. These results were obtained using the MATLAB/Octave toolbox for stochastic Galerkin methods (Zander (2009)).

Also the time evolution with respect to the increasing number of samples is interesting. Such a comparison of direct evaluation of forward problem and approximated forward model by PC expansion of 2nd order based on

219

(a) $L_x = 0.3$ m (b) $L_x = 0.1$ m

(c) $L_x = 0.05$ m (d) $L_x = 0.025$ m
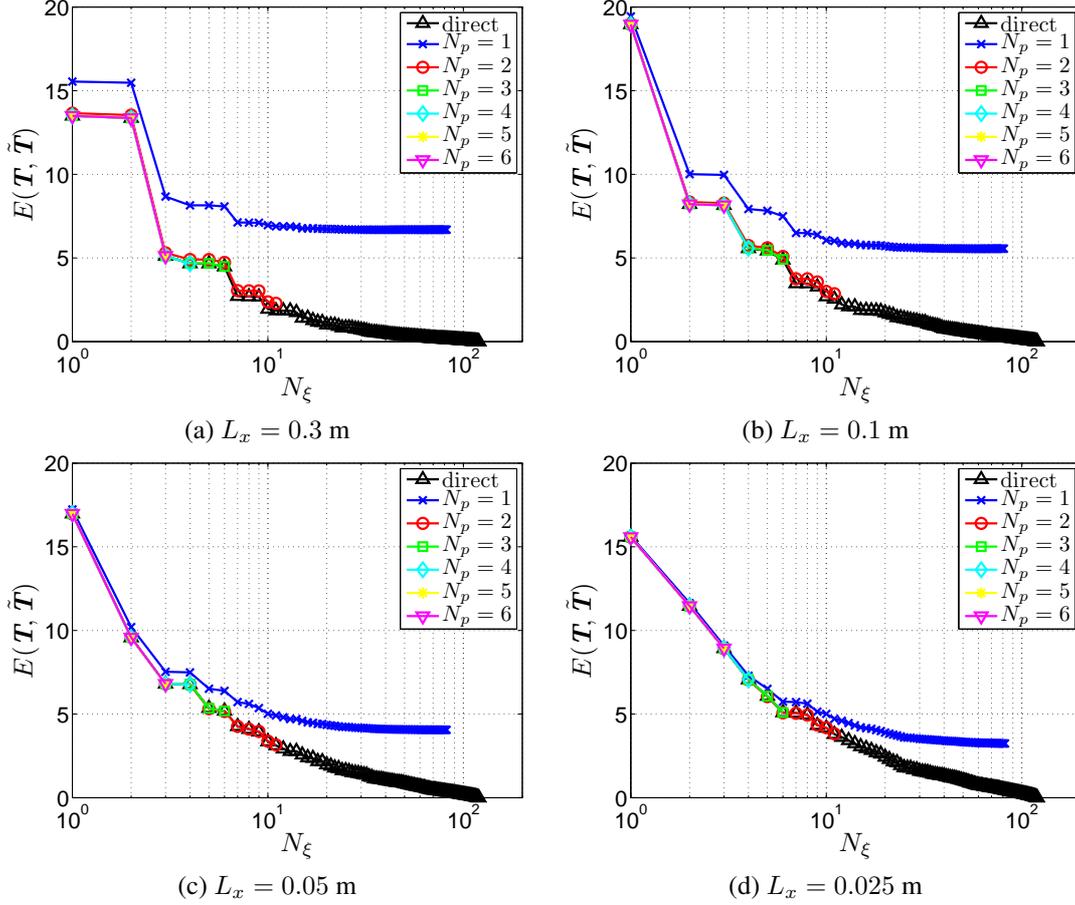
Figure 7: Summarised and averaged pointwise error $E(\boldsymbol{T}, \hat{\boldsymbol{T}})$ as a function of number N_ξ of eigenvectors involved for different choices of maximal degrees of polynomials N_p and for correlation lengths $L_y = 0.04$ and (a) $L_x = 0.3$, (b) $L_x = 0.1$, (c) $L_x = 0.05$, (d) $L_x = 0.025$.
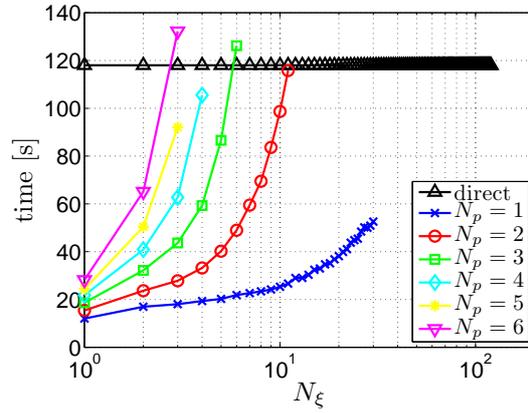
Figure 8: Comparison of time necessary for the computation of PC coefficients and 100,000 random samples of the model response via direct evolution of FE code and via its PC approximation with different degrees of polynomials.

input random field built up with 6 eigenmodes is shown in Figure 9.

Of course, the presented model is very simple (it leads to the system of only M = 80 linear equations), hence, even the direct simulation via full FE model is very fast, but once we use for example finer FE mesh, FE simulation becomes slower, but evaluation of PC expansion remains the same and becomes much more effective (cf. Figures 9a and 9b). So, for more "complex" models, acceleration via PC expansion becomes much more attractive.

To study the properties of the presented methods in Bayesian identification, we must start with the preparation of
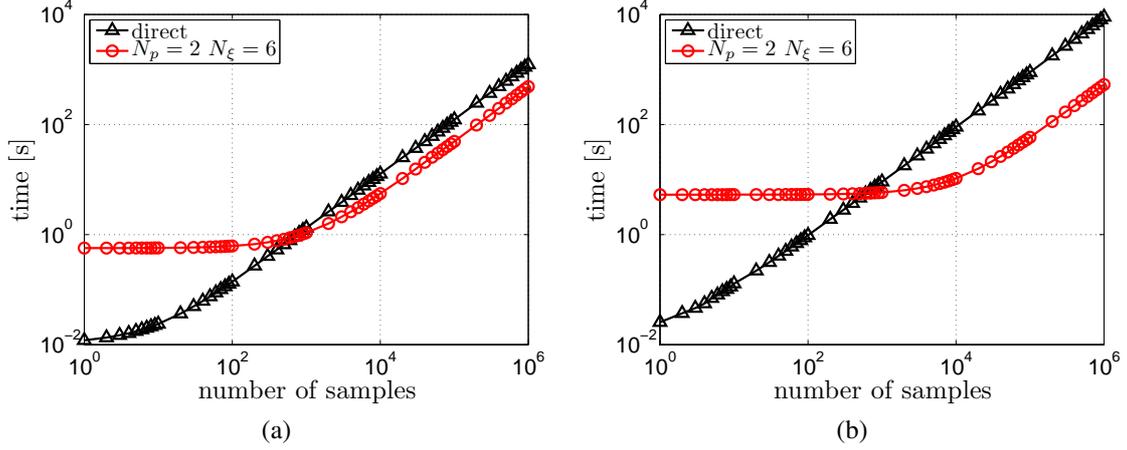
Figure 9: Time necessary for computing the response of a FE model and PCE as a function of number of samples for a FE mesh with (a) 120 elements and (b) 480 elements.

"artificial" measurements. We have generated a random vector $\boldsymbol{\xi} = \xi_1 \ldots \xi_{120}$, and the corresponding realisation of input field $k(\boldsymbol{x})$, which we consider as a known "true" shape, is shown in Figure 10a. Using the subset of the first 6 values $\xi_1 \ldots \xi_6$, we construct the "true approximation" $\hat{k}(\boldsymbol{x})$ of random field $k(\boldsymbol{x})$. The shape of such an approximation is shown in Figure 10b.
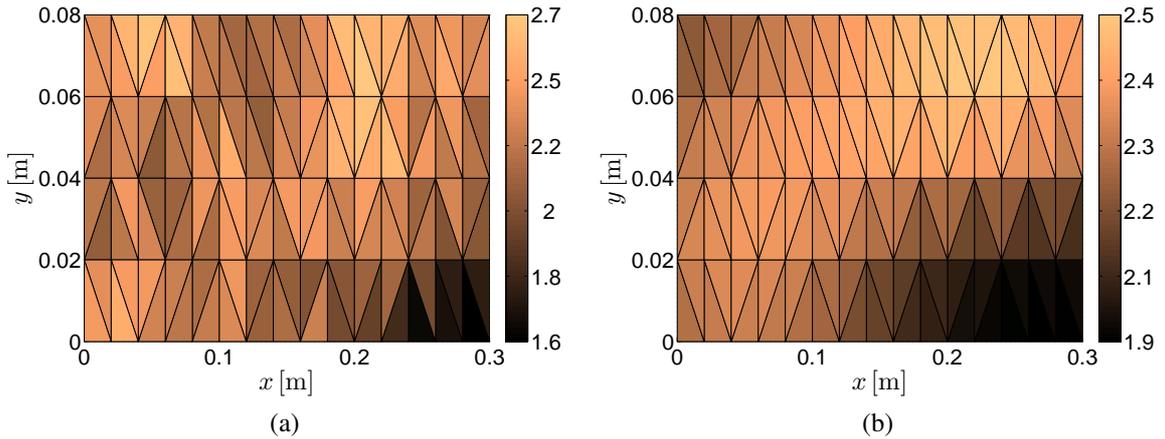
Figure 10: Shape of a particular realisation of k field used for the computation of an "artificial" measurement. The shape is obtained using KL expansion based on (a) all 120 eigenmodes and (b) only the first 6 eigenmodes.

The true shape of the input field $k(\boldsymbol{x})$ is then used to compute the forward model response considered as an artificial measurement. Adding the Gaussian noise with zero mean and standard deviation $\sigma_T = 0.1$ (if not specified otherwise) we obtain 100 measurements, and using them the covariance matrix $\mathbf{C}_{\mathrm{obs}}$ is computed to characterise the reliability of measuring. This covariance matrix then takes place in the expression of a likelihood function (9).

Our a priori knowledge about the random variables ξ_i is that they are standard Gaussian $\sim N(0, 1)$. Then we construct the polynomial basis for $N_\xi = 6$ random variables and with the maximal degree $N_p = 2$, so the basis consists of $N = 28$ polynomials. We compute Galerkin projection and obtain PC coefficients for our surrogate of FE model solving the system of $80 \cdot 28 = 2240$ linear equations. Finally, we apply the Metropolis algorithm to choose $100,000$ samples from the a posteriori distribution of $\xi_1 \ldots \xi_6$. For each sample $\boldsymbol{\xi}_j = \xi_{1,j} \ldots \xi_{6,j}$ we construct the field of $\hat{k}_j(\boldsymbol{x})$ and from those fields the average field is computed (see Figure 11a). In the same way we also compute the standard deviation of those fields, which is plotted in Figure 11b. It is visible that from the identification point of view, the average field from our posterior samples well corresponds to the "true approximation" of the reference field shown in Figure 10b.

Of course, this correspondence depends on how much we trust to our a priori knowledge on one hand, and how much we trust measurements on the other hand. This fact is demonstrated in Figure 12. Figure 12a shows boxplot of a priori distributions of $\xi_1 \ldots \xi_6 \sim N(0, 1)$. Figure 12b then shows the a posteriori distributions corresponding
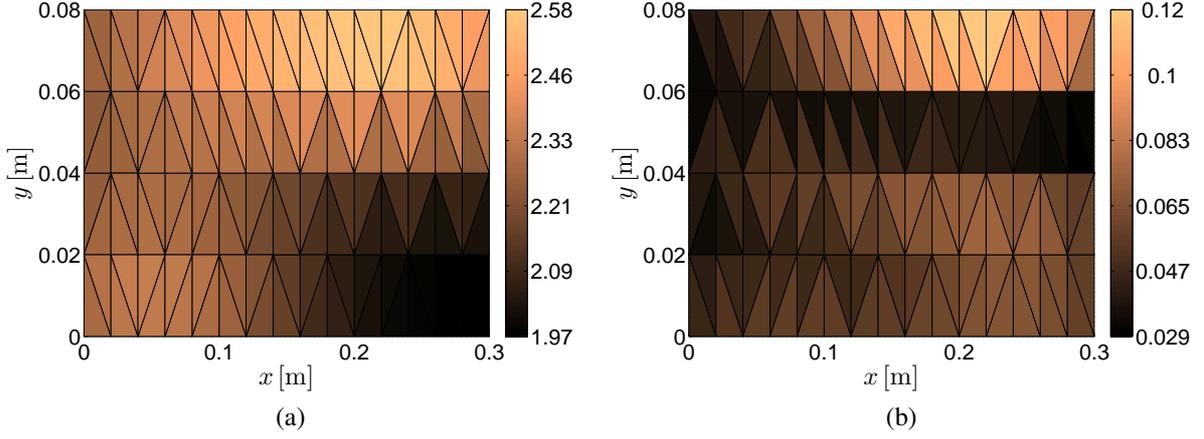
Figure 11: A posteriori prediction of a particular realisation of k field (a) mean and (b) standard deviation.

to measurements with the error $\sigma_T = 1$, and Figure 12c shows again the a posteriori distribution, but for $\sigma_T = 0.1$. The later case represents a situation when we have more precise measurements or we trust them more. Therefore, the mean values in Figure 12b are closer to zero (the a priori value), and in Figure 12c they are more distant from zero and more determined by measurements. Also standard deviations are smaller in the later case.
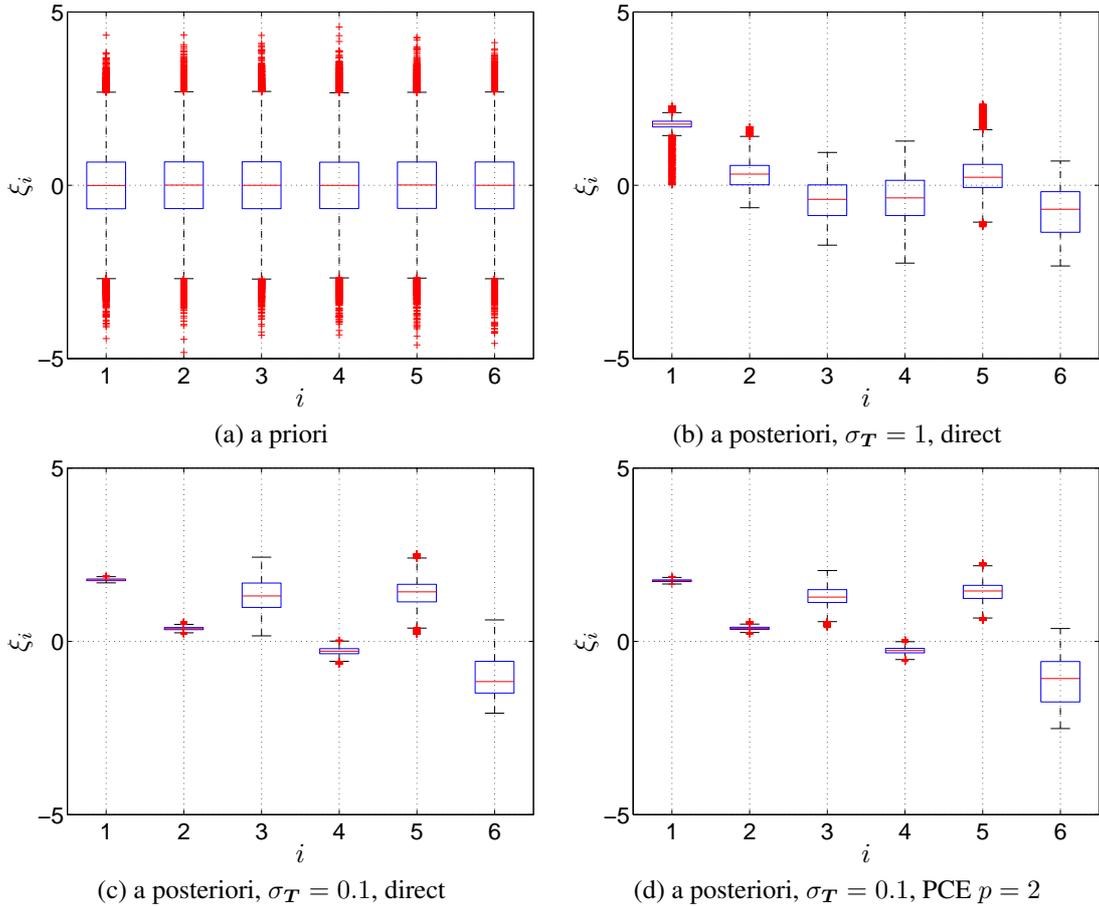
Figure 12: Boxplot of distribution of particular random variables $\xi_1 \ldots \xi_6$: (a) a priori, (b) a posteriori for error of thermometer $\sigma_T = 1$ and direct evaluation of FE model, (c) a posteriori for error of thermometer $\sigma_T = 0.1$ and direct evaluation of FE model, (d) a posteriori for error of thermometer $\sigma_T = 0.1$ and PC approximation with $N_p = 2$.

Effect of imprecise PC approximation of FE model is visible by comparing Figure 12c with Figure 12d, where the a posteriori is computed using PC expansion instead of direct evaluation of FE model. We conclude that

the error introduced by PC approximation seems to be negligible. More detailed comparison of a priori and a posteriori distribution computed via direct evaluation of direct problem or PC expansion of degree $N_p = 2$ is shown in Figure 13 for each particular variable ξ_i.
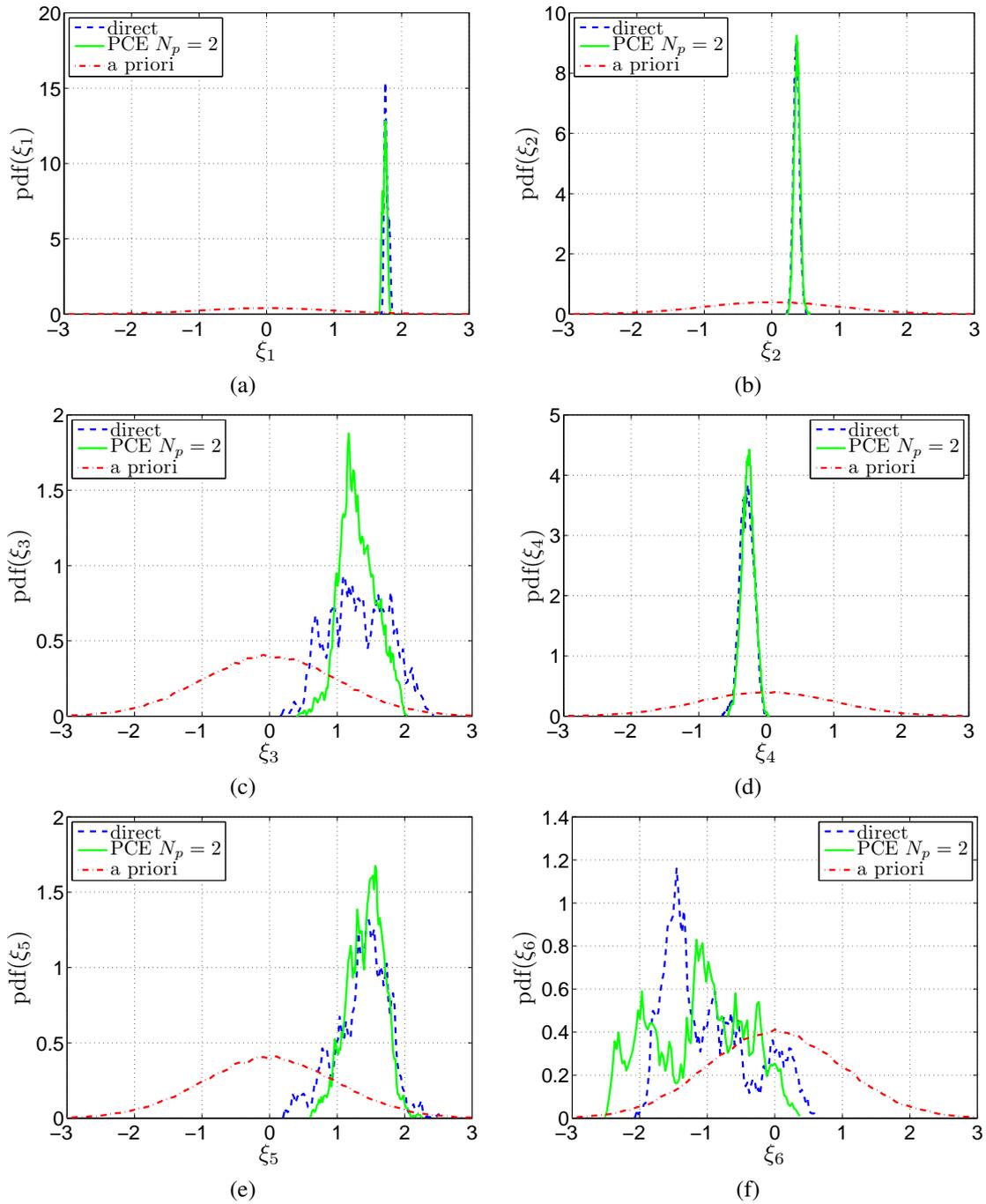
Figure 13: A posteriori probability density functions for (a) $\xi_1 \ldots$ (f) ξ_6.

One can notice that the a posteriori distribution of ξ_1, ξ_2 and ξ_4 is narrower than the others. It means that these variables are more determined and probably more important for description of model response. Interesting is also the fact, the from this point of view, ξ_3 seems to be less important than ξ_4, while according to KLE it should be in contrariwise.

7 Conclusion

This paper presents a Bayesian updating/identification of a heterogeneous material property which can be described by a random field. Since the Bayesian inference consists of a posterior distribution sampling, it is profitable to replace the expensive simulation of a computation model by its cheaper approximation. Polynomial chaos is used here as an appropriate tool to overcome this problem.

The cost of PCE evaluation is, however, quickly increasing with the degree of polynomials N_p and the number of random variables N_ξ. Therefore, the description of a random field via simple spatial discretisation is inconvenient, since a chosen spatial grid needs to be dense enough to describe the field properly. Moreover, resulting random variables are usually highly correlated among each other. A more suitable characterisation is e.g. via truncated Karhunen-Loève expansion describing the random field by a limited set of independent random variables. The degree of polynomials can also be reduced, e.g., by the division of modelled domain as presented by Marzouk et al. (2007).

The whole concept is demonstrated on a very simple example of heat conduction, where interesting phenomena can be clearly understood. The particular conclusion can be listed as follows:

i Karhunen-Loève expansion can significantly reduce the number of random variables necessary for the description of a random field with a satisfactory accuracy. The error in the description of the random field itself increases with smaller correlation lengths, but it should not be true for the model response, where the random field is used as input.

ii Polynomial chaos expansion is a powerful approximation tool. In the presented application, the 2nd degree polynomial is sufficient to describe the nonlinear relation of model response $u(m(\xi)) = (K(m(\xi)))^{-1}f$ with sufficient precision.

iii Both the KLE and the PCE enables a more efficient Bayesian update of model parameters. Savings in computational time increases with the complexity of the model, or especially with finer FE discretisation.

Our future work will be concerned with the development of the same updating procedure for a nonlinear model.

References

Chen, N. Z.; Soares, C. G.: Spectral stochastic finite element analysis for laminated composite plates. *Computer Methods in Applied Mechanics and Engineering*, 197, 51–52, (2008), 4830–4839.

Colliat, J.-B.; Hautefeuille, M.; Ibrahimbegović, A.; Matthies, H. G.: Stochastic approach to size effect in quasi-brittle materials. *Comptes Rendus Mécanique*, 335, 8, (2007), 430–435.

Evans, M.; Swartz, T.: Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science*, 10, 3, (1995), 254–272.

Fu, J.; Gómez-Hernández, J. J.: Uncertainty assessment and data worth in groundwater flow and mass transport modeling using a blocking markov chain monte carlo method. *Journal of Hydrology*, 364, 3–4, (2009), 328–341.

Ghanem, R. G.; Kruger, R. M.: Numerical solution of spectral stochastic finite element systems. *Computer Methods in Applied Mechanics and Engineering*, 129, 3, (1996), 289–303.

Gilks, W. R.; Richardson, S.; Spiegelhalter, D.: *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC (December 1995).

Grigoriu, M.: *Applied non-Gaussian Processes.*. Prentice Hall: Englewood Cliffs (1995).

Gutiérrez, M.; Krenk, S.: *Encyclopedia of Computational Mechanics*, chap. Stochastic finite element methods. John Wiley & Sons, Ltd. (2004).

Higdon, D.; Lee, H.; Holloman, C.: Markov chain Monte Carlo-based approaches for inference in computationally intensive inverse problems. *Bayesian Statistics*, 7, (2003), 181–197.

Hosack, G. R.; Eldridge, P. M.: Do microbial processes regulate the stability of a coral atolls enclosed pelagic ecosystem? *Ecological Modelling*, 220, 20, (2009), 2665–2682.

Keese, A.: A review of recent developments in the numerical solution of stochastic partial differential equations (stochastic finite elements). Tech. rep., Institute of Scientific Computing, Technical University Braunschweig, Brunswick, Germany (2003).

Kennedy, M. C.; O'Hagan, A.: Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 3, (2001), 425–464.

Khoromskij, B. N.; Litvinenko, A.: Data sparse computation of the karhunen-loève expansion. In: *Proceedings of International Conference on Numerical Analysis and Applied Mathematics 2008*, vol. 1048, pages 311–314 (2008).

Kučerová, A.; Brancherie, D.; Ibrahimbegović, A.; Zeman, J.; Bittnar, Z.: Novel anisotropic continuum-discrete damage model capable of representing localized failure of massive structures. part II: identification from tests under heterogeneous stress field. *Engineering Computations*, 26, 1/2, (2009), 128–144.

Kučerová, A.; Lepš, M.; Zeman, J.: Back analysis of microplane model parameters using soft computing methods. *CAMES: Computer Assisted Mechanics and Engineering Sciences*, 14, 2, (2007), 219–242.

Lombardo, M.; Zeman, J.; Šejnoha, M.; Falsone, G.: Stochastic modeling of chaotic masonry via mesostructural characterization. *International Journal for Multiscale Computational Engineering*, 7, 2, (2009), 171–185.

Marzouk, Y.; Najm, H.: Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics*, 228, 6, (2009), 1862–1902.

Marzouk, Y.; Najm, H.; Rahn, L.: Stochastic spectral methods for efficient Bayesian solution of inverse problems. *Journal of Computational Physics*, 224, 2, (2007), 560–586.

Matthies, H.; Brenner, C.; Bucher, C.; Soares, C. G.: Uncertainties in probabilistic numerical analysis of structures and solids - stochastic finite elements. *Structural Safety*, 19, 3, (1997), 283–336.

Matthies, H.; Keese, A.: Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 194, 12-16, (2005), 1295–1331.

Matthies, H. G.: *Encyclopedia of Computational Mechanics*, chap. Uncertainty Quantification with Stochastic Finite Elements. John Wiley & Sons, Ltd. (2007).

Mosegaard, K.; Tarantola, A.: *Probabilistic Approach to Inverse Problems*, pages 237–265. Part A, Academic Press (2002).

Parthasarathy, S.; Balaji, C.: Estimation of parameters in multi-mode heat transfer problems using bayesian inference effect of noise and a priori. *International Journal of Heat and Mass Transfer*, 51, 9–10, (2008), 2313–2334.

Quiros, A.; Diez, R. M.; Gamerman, D.: Bayesian spatiotemporal model of fMRI data. *NeuroImage*, 49, 1, (2010), 442–456.

Rao, K. S.: Source estimation methods for atmospheric dispersion. *Atmospheric Environment*, 41, 33, (2007), 6964–6973.

Stefanou, G.: The stochastic finite element method: Past, present and future. *Computer Methods in Applied Mechanics and Engineering*, 198, 9–12, (2009), 1031–1051.

Sýkora, J.; Vorel, J.; Krejčí, T.; Šejnoha, M.; Šejnoha, J.: Analysis of coupled heat and moisture transfer in masonry structures. *Materials and Structures*, 42, 8, (2009), 1153–1167.

Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics (2005).

Tierney, L.: Markov chains for exploring posterior distributions. *Annals of Statistics*, 22, 4, (1994), 1701–1728.

Valenta, R.; Šejnoha, M.; Zeman, J.: Macroscopic constitutive law for mastic asphalt mixtures from multiscale modeling. *International Journal for Multiscale Computational Engineering*, accepted for publication.

Šejnoha, J.; Šejnoha, M.; Zeman, J.; Sýkora, J.; Vorel, J.: Mesoscopic study on historic masonry. *Structural Engineering and Mechanics*, 30, 1, (2008), 99–117.

Šejnoha, M.; Zeman, J.: Micromechanical modeling of imperfect textile composites. *International Journal of Engineering Science*, 46, 6, (2008), 513–526.

Wriggers, P.; Moftah, S.: Mesoscale models for concrete: Homogenisation and damage behaviour. *Finite Elements in Analysis and Design*, 42, 7, (2006), 623–636.

Xiu, D.; Karniadakis, G. E.: The wiener–askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24, 2, (2002), 619–644.

Yee, E.; Lien, F.-S.; Keats, A.; D'Amours, R.: Bayesian inversion of concentration data: Source reconstruction in the adjoint representation of atmospheric diffusion. *Journal of Wind Engineering and Industrial Aerodynamics*, 96, 10–11, (2008), 1805–1816.

Zander, E.: SGLIB. `http://www.wire.tu-bs.de/mitarbeiter/ezander/e_index.html` (2009).

Address: Ing. Anna Kučerová, Ph.D., and Prof. Hermann Matthies, Ph.D., Institut für Wissenschaftliches Rechnen, Technische Universität Braunschweig, D-38092 Braunschweig
email: `a.kucerova@tu-bs.de; wire@tu-bs.de`