CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Civil Engineering

Ing. Anna Kučerová, Ph.D.

Inverse Problems in Modeling of Random Materials

Inverzní problémy v modelování náhodných materiálů

January, 2020

ACKNOWLEDGMENTS

TABLE OF CONTENTS

Czech Technical University in Prague

Abstract

Inverse Problems in Modeling of Random Materials

Anna Kučeroá

Reliable modeling of large-scale structures with the prediction of their durability is inevitably related to an adequate description of the materials used in the structure. Most of the building materials are heterogeneous with random morphology, which complicates namely the modeling of damage and application of traditional homogenization-based techniques. Therefore there are several strategies to handle this underlaying randomness of material and structural response, which are not directly connected to employed constitutive models, but rather to description of input material morphology and material parameters. This work is devoted to methods leading to the suitable characterization of the inputs to advanced constitutive and structural models, in other words to methods providing solutions to the related inverse problems. The emphasis is put to the classification of the methods and inverse problems, their appropriate formulation and namely to the clarification of several common misinterpretations of solutions to inverse problems. Finally, examples of three selected topics are discussed in more detail in particular chapters: (i) generating random microstructures with prescribed statistical properties, (ii) parameter identification for random material models and (iii) related approximation-based acceleration techniques.

Keywords: Random materials; Statistically equivalent periodic unit cell; Wang tiles, Statistical descriptors; Inverse problems, Parameter identification, Bayesian inference, Uncertainty quantification, Surrogate models, Polynomial chaos expansion, Artificial neural networks

České vysoké učení technické v Praze

Abstrakt

Inverzní problémy v modelování náhodných materiálů

Anna Kučeroá

Spolehlivé modelování stavebních konstrukcí je nevyhnutelně spojeno s adekvátním popisem chování použitých materiálů. Převážná většina stavebních materiálů má heterogenní strukturu s náhodnou morfologií, která představuje značnou komplikaci zejména v modelování poškození nebo při používání tradičních homogenizačních technik. Proto existuje celá řada postupů, které se náhodností v popisu materiálů zabývají. Tato práce se věnuje takovým postupům, které mají za cíl získat vhodný popis náhodných materiálů z různých typů dostupných měření, což vede k řešení různých inverzních úloh. Důraz je kladen na klasifikaci identifikačních postupů a inverzních problémů, jejich správnou formulaci a interpretaci výsledků. V práci jsou ukázány podrobněji tři vybraná témata v samostatných kapitolách: (i) generování náhodných mikrostruktur s předepsanými statistickými vlastnostmi, (ii) identifikace parametrů modelů náhodných materiálů a (iii) akcelerační techniky založené na konstrukci náhradních modelů.

Klíčová slova: Materiály s náhodnou morfologií; Statisticky equivalentní periodická jednotková buňka; Wangovo dláždění; Statistické deskriptory; Inverzní problémy; Identifikace parametrů; Bayesovská inference; Modelování nejistot; Aproximace; Polynomiální chaos; Umělé neuronové sítě

Chapter 1

INTRODUCTION

Random materials, which exhibit indeterministic behavior at macro-scale due to heterogeneities at fine-scales, are very common materials not only in civil engineering but in many other engineering fields, natural sciences, medicine etc., see Figure 1.1 for illustrative examples. Nevertheless, the best developed theoretical framework provided by homoge-
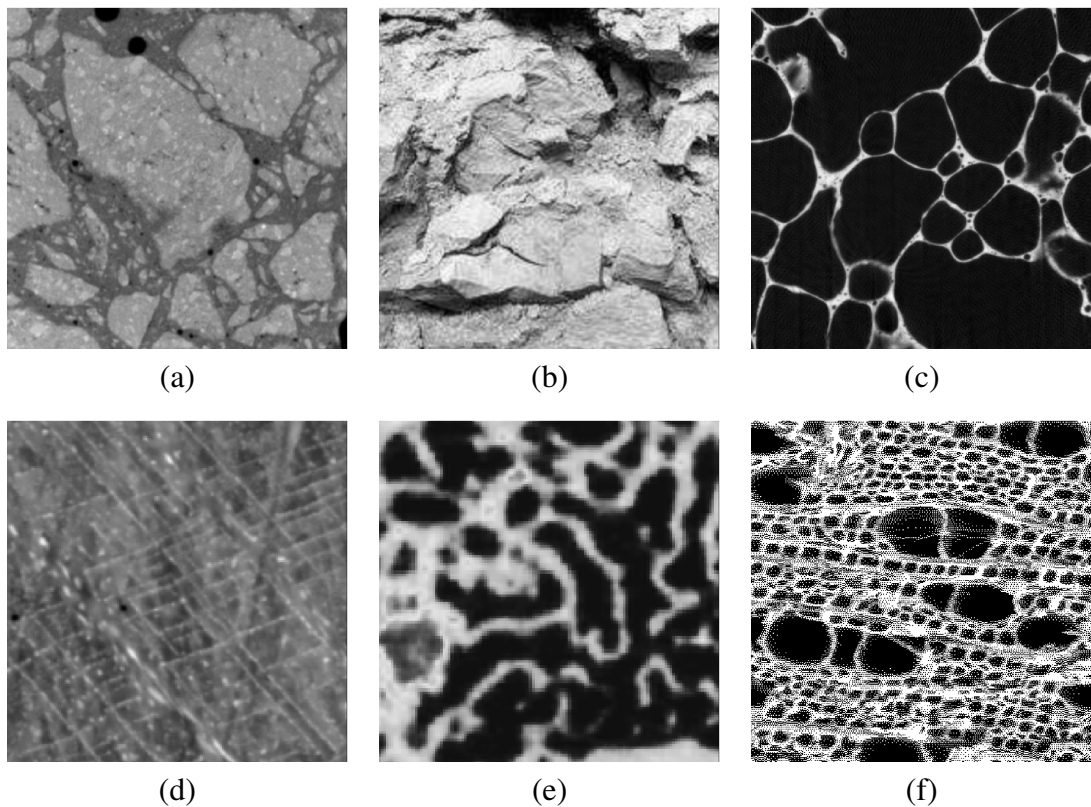
Figure 1.1: Examples of random materials: (a) Concrete, from Sitek, L., Bodnarová, L., Souček, K., Staš, L., and Gurková, L. (2015). Analysis of inner structure changes of concrete exposed to high temperatures using micro X-ray computed tomography. Acta Geodynamica et Geomaterialia, 12(1), 177. (b) Irregular filling (quarry masonry). (c) Metal foam. (d) Copper-silver alloy. (e) Trabecular bone, from Jiroušek, O., Zlámal, P., Kytýř, D., and Kroupa, M. (2011). Strain analysis of trabecular bone using time-resolved X-ray microtomography. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 633, S148-S151. (f) Birch.

nization theories, which aim at the replacement of the heterogeneous microstructure with an equivalent homogeneous material, is originally tailored for periodic or quasi-periodic ma-

terials, see Torquato (2006). Their application to random materials is driven namely by the lack of any better alternative and insufficient computational power limiting the development of probabilistic formulations. Only to briefly recall two main homogenization-based approaches, first are the effective media theories estimating material response analytically based on selected geometrical information. The computational requirements are very low, but their application is limited only to specific inclusion shapes, see e.g. Tanaka (1973); Vorel and Šejnoha (2009); Odegard et al. (2017); Tornabene et al. (2019). Second is the computational homogenization with more general applications, see E et al. (2007); Geers et al. (2010); Sýkora et al. (2012, 2013); Matouš et al. (2017). This numerical method is based on detailed finite element discretization of heterogeneous microstructure resulting in a detailed description of local response fields. The appropriate application is, however, conditioned by the existence of a representative volume element (RVE), which cannot be fulfilled in case of random materials.

The last decades witnessed an intense development of statistical and probabilistic reformulations of existing methods for modeling random materials and representation of their morphology, which can be for this work divided into three families briefly described below.

Advanced statistical cell-based representation:

The first group of methods is working with information about the material at the micro- or mesoscale level based on image analysis. These methods follow and extend the concept of RVE in two directions.

(i) The exactly periodic unit cell unavailable for a random material is replaced by the cell preserving spatial geometrical statistics also called as statistically equivalent periodic unit cell (SEPUC) or statistically similar representative volume element (SSRVE), see Zeman and Šejnoha (2007); Schröder et al. (2011); Niezgoda et al. (2010); Bostanabad et al. (2016). To this goal, different statistical descriptors are available such as the n-point probability function (Torquato and Stell (1982)), two-point cluster function (Doškář (2013)), lineal path function (Havelka et al. (2016)) or others, see Torquato (2006) or Chapter 2.2.

(ii) A single PUC or SEPUC is replaced by a set of statistically equivalent and geometrically compatible cells, so-called Wang tiles, allowing to represent an aperiodic random morphology without or with the controlled long-range correlations arising as artifacts of using a single PUC, see Novák et al. (2012, 2013); Doškář et al. (2014); Doškář and Novák (2016) or Chapter 2.1. As the tiling-based concept is relatively new, the further possible steps for tiling-based mechanical simulations inspired by homogenization or partition of unity method are discussed in Novák et al. (2013); Doškář and Novák (2016).

The principal attribute of these methods is preserving the sharp separation of particular phases at the microstructural level. The advantage consists in the fact that the material properties of these phases can be obtained separately from experiments on single-phase specimens. The key difficulty is thus shifted to a determination of particular geometry of statistically equivalent set of tiles based on the images of real material. This leads to a specific inverse problem discussed in more detail in Section 1.1. Another possibility to introduce the statistical description of heterogeneous material into homogenization-based framework is to

model the microstructural effective properties by random variables. This formulation leads to a group of the so-called stochastic multi-scale methods, see e.g. Babuška et al. (2014); Guan et al. (2015); Chu and Guilleminot (2019).

Lattice/particle-based discrete systems:

Another source of information available typically in the case of concrete is the knowledge of the constituents employed during the material preparation, such as granulometric distributions. This particular example leads to lattice/particle-based models of random materials developed in the scientific community studying mostly the failure phenomena in quasi-brittle materials on the meso- or microstructural levels, see Jirásek and Bažant (1995). Discrete models allow to realistically and relatively easily simulate the propagation of discontinuities or multiple cracks interaction in the heterogeneous materials. The internal structure is characterized by the finite number of coarse aggregate particles in accordance with granulometric distributions, which are randomly placed into the volume. According to physical laws describing the mechanical response of the system, the following three modeling strategies can be recognized: (i) lattice models, see Hrennikoff (1941); Ostoja-Starzewski (2002); Nikolić et al. (2018), (ii) particle models, see Cundall (1971); Bažant et al. (1990); De Schutter and Taerwe (1993); Wang et al. (2016), and (iii) lattice–particle models representing the synthesis of aforementioned formulations, see Grassl and Bažant (2009); Cusatis et al. (2011); Schauffert and Cusatis (2012); Marcon et al. (2017).

Probabilistic description based on random variables and/or fields:

The last category of methods considers no additional information about the material itself. In this situation, the estimation of its physical or mechanical properties, as well as their spatial distribution, needs to be identified from the macroscopic observations. A significant milestone in probabilistic modeling was establishing the theory for spectral stochastic finite element method by Ghanem and Spanos (1991), which started an intense development and elaboration of probabilistic reformulation of engineering material and structural models, just to cite a few Hamdia et al. (2017); Quarteroni et al. (2017); Beran et al. (2017); Muneo (2018). The so-called 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 due to insufficient knowledge of their inner characteristics, or both, cf. e.g. the reports Matthies (2007); Stefanou (2009); Xiu (2010); Gutiérrez and Krenk (2017). From a mathematical point of view, these systems can be characterized by stochastic ordinary/partial differential equations (SODEs/SPDEs), which can be solved by the stochastic finite element method (SFEM). SFEM is an extension of the classical deterministic finite element approach to the stochastic framework, where finite elements have random properties. Originally, Monte Carlo (MC) was widely used technique in simulating models driven by SODEs/SPDEs. MC simulations require thousands or millions of samples because of relatively slow convergence rate, thus the total cost of these numerical evaluations quickly becomes prohibitive. The surrogate models based on the polynomial chaos expansion (PCE), see Wiener (1938); Xiu and Karniadakis (2002), were developed as promising alternative. For more details, see Chapter 4.2, where different methods for the construction of PCE-based surrogates are compared on a simple civil engineering example of frame structure with uncertain parameters in loading and geometry. Moreover, to accelerate the simulations

for heterogeneous materials with spatially varying material properties characterized by random fields, the truncated Karhunen-Loève (KL) expansion is typically introduced to reduce the stochastic space dimension to a reasonable number of random variables Matthies (2007); Rosić and Matthies (2015). The inverse problem related to this type of random materials modeling thus leads to the estimation of random variables describing either the particular material properties or scaling the particular KL modes in the employed random fields. Such a probabilistic identification of random variables needs to be properly formulated to avoid misinterpretation of the results, as discussed in more detail in Section 1.2. One more elaborate example of probabilistic parameter identification of a coupled model for transport processes in a heterogeneous medium characterized by random fields is presented in Chapter 3.1.

A common feature of random fields-based and particle-based strategies is that in the further mechanical modeling the particular phases are not necessarily modeled with the sharp separated geometry. Particular material parameters considered in the FEM-based models are related either to homogenized or smoothly spatially varying medium (see Chapters 3.2 and 3.1, respectively, for detailed examples). The key inverse problem to be solved here is thus the identification of material parameters. As the meaning of these parameters is inevitably related to a particular constitutive model (and some of them are phenomenological), the estimation of their values needs to be based on the observations of the material response, which is generally in nonlinear relationship to particular material parameters to be identified. The deterministic formulation and numerical techniques available for parameter identification in the case of non-linear material models are thoroughly discussed in the doctoral thesis Kučerová (2007). One perspective on the probabilistic formulation of the parameter identification problem is presented in Rosić et al. (2013). Section 1.2 presents another classification of probabilistic parameter identification along with the indication of common misinterpretation of its results.

Note on computational acceleration:

The last topic related to a certain extent to all listed categories concerns the computational burden connected to probabilistic models of random materials or their calibration, which can be partially relieved by replacement of the material model by its cheaper surrogate. While the parameter identification community prefers for decades different types of artificial neural networks, see Kučerová (2007) for a review and other applications in Kučerová et al. (2009); Kučerová and Lepš (2014); Mareš et al. (2016), the uncertainty quantification community originally focused more on uncertainty propagation applies namely polynomial chaos expansion allowing for fast analytical evaluation of statistical moments of approximated quantities, see e.g. Kučerová et al. (2012); Rosić et al. (2013, 2016); Gutiérrez and Krenk (2017); Janouchová et al. (2018). For more details about both of these types of surrogates, we refer to Chapters 4.1 and 4.2, respectively. Nowadays, both communities start to blend and both types of surrogates are compared or combined, see e.g. the recent works Shahane et al. (2019); Schwab and Zech (2019); Zhang et al. (2019). Surrogate models for stochastic models, particularly, are discussed e.g. in Marrel et al. (2012); Azzi et al. (2020).

1.1   Search for statistically representative cells

Computational techniques suitable for the determination of statistically representative cells can be generally used both for the more traditional reconstruction of a single SEPUC or for the more recent design of a set of Wang tiles. A very thorough and recent review of techniques available in the literature is presented e.g. in Bostanabad et al. (2018). Here, more compressed classification into three categories follows as:

   (i) Optimization: Cells are optimized with respect to a chosen measure of statistical difference from the original medium. One set of measures is based on correlation functions such as two-point, lineal path or two-point cluster. The optimization is then governed by some of the stochastic algorithms based on the principles of simulated annealing (e.g. Yeong and Torquato (1998a,b); Kumar et al. (2008); Jiao et al. (2009); Collins et al. (2010); Novák et al. (2012, 2013); Chen et al. (2014); Havelka et al. (2016)). Another type of measures is physically based, for instance, the distance to nearest neighbor motivated by transport processes in medium with inclusions (Tewari and Gokhale (2004) or connectivity of the local stress/strain fields (Novák et al. (2013)).

  (ii) Random field: Here the random field-based description is not used to generate fields of smoothly varying material properties, but again to generate samples of cells with prescribed spatial correlation. Particular samples are created as random realizations of typically smooth Gaussian random field, which is then transformed into two-phase medium by means of chosen level-cutting, see Grigoriu (2003); Jiang et al. (2013). On the other hand, the correlation function can be also derived by image analysis directly from the scan of the microstructure and the advantages of this approach are discussed in Kučerová et al. (2014) and in Section 1.1.2 in more detail.

 (iii) Texture synthesis: These methods based on fusion of raster images were originally developed for computer graphics problems and later applied to the reconstruction of random media in (Wei and Levoy (2000); Efros and Freeman (2001); Doškář et al. (2014)). In particular, the image quilting algorithm seeks for a continuous path along which the desired pieces of microstructure are glued together, minimizing the sum of squared differences of pixel values.

Of course, not all of the related works fit one of the listed categories, for instance, an interesting work of Bostanabad et al. (2016) presents the microstructure reconstruction using supervised learning. To conclude, the methods based on random fields or texture synthesis are namely very fast, while the optimization-based strategies are more general and allow for better control of desired microstructural measures. Further search for more efficient optimization algorithms tailored specifically to this type of problem remains highly demanded. As an attempt to elaborate more on this topic, Section 1.1.1 is focused on the comparison of available optimization algorithms on the design of isotropic two-phase PUC for equal-sized hard-disk ensembles.

1.1.1   Optimization of an isotropic PUC

To compare available optimization algorithms suitable for seeking statistically representative cells, a simple example focused on the minimization of anisotropy is presented. In particular, the goal is to find an isotropic two-phase PUC consisting of equal-sized hard-disk particles, see Figure 1.2.
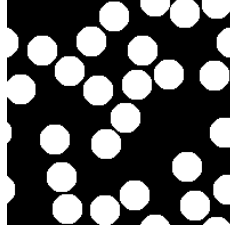
Figure 1.2: Example of PUC for two-phase medium with 25 equal-sized hard-disks repre-
senting 40 % of volume fraction.

1.1.1.1   An anisotropy measure formulation

To quantify the anisotropy of a given microstructure, the two-point probability function
seems to be convenient thanks to its possibly fast evaluation for binary discretized mi-
crostructures. More formally, the two-point probability function $S_{rs}(\mathbf{x_1}, \mathbf{x_2})$ quantifies the
probability of finding simultaneously the phase r at $\mathbf{x_1}$ and the phase s at $\mathbf{x_2}$ and can be
written in the form

$$S_{rs}(\mathbf{x_1}, \mathbf{x_2}) = \overline{\chi_r(\mathbf{x_1}, \alpha)\chi_s(\mathbf{x_2}, \alpha)}, \tag{1.1}$$

i.e. as the ensemble average of the product of characteristic functions $\chi_r(\mathbf{x_i}, \alpha)$, which are
equal to one when point $\mathbf{x_i}$ lies in the phase r in the sample α and equal to zero otherwise:

$$\chi_r(\mathbf{x_i}, \alpha) = \begin{cases} 1, & \text{if } \mathbf{x_i} \in D_r(\alpha) \\ 0, & \text{otherwise} \end{cases}. \tag{1.2}$$

In Equation (1.2), $D_r(\alpha)$ denotes the domain occupied by the r-th phase.

In general, the evaluation of these characteristics may prove to be prohibitively diffi-
cult. Fortunately, a simple method of attack can be adopted when accepting an assumption
regarding the material as statistically homogeneous, so that

$$S_{rs}(\mathbf{x_1}, \mathbf{x_2}) = S_{rs}(\mathbf{x_1} - \mathbf{x_2}) = S_{rs}(\mathbf{x}). \tag{1.3}$$

Last note concerns particularly the two-phase medium, where two-point probability func-
tions are related according to the following equation

$$S_{rr}(\mathbf{x}) = c_r - c_s + S_{ss}(\mathbf{x}), \tag{1.4}$$

where c_r and c_s are volume fractions of phases r and s, respectively. Thanks to this relation,
only one probability function needs to be determined to describe the two-phase medium.

To establish $S_{rr}(\mathbf{x})$ value for a particular binary image with dimensions $W \times H$, the
two-point probability functions can be re-written as a correlation of functions χ_r and χ_s.
By implementing the Discrete Fourier Transform (DFT), see Burrus and Parks (1985), the
two-point probability function S_{rs} becomes

$$S_{rs}(m, n) = \frac{1}{WH} \text{IDFT} \left\{ \text{DFT} \left\{ \chi_r(m, n) \right\} \overline{\text{DFT} \left\{ \chi_s(m, n) \right\}} \right\}, \tag{1.5}$$

where IDFT is the inverse DFT, $\bar{\cdot}$ stands for the complex conjugate and $\chi_r(m, n)$ denotes the
value of χ_r for the pixel located in the m-th row and n-th column of the binary image. This

method is very economical and its accuracy depends only on the selected resolution of the digitized medium. The Fast Fourier Transform, which needs only $O(WH \log(WH) + WH)$ operations, is called to carry out the numerical computations.

In the case of statistical isotropy assumption, the shape of two-point probability function should be invariant to rotation, i.e.

$$\forall l = \|\mathbf{x}\| : S_{rs}(\mathbf{x}) = S_{rs}(l) . \tag{1.6}$$

For an anisotropic medium, the spread of the two-point probability function values over the vectors of the same length is non-zero. For illustration, it is obvious that the microstructure shown in Figure 1.3a is not isotropic. The shape of the corresponding two-point probability function for white phase S_{ww} is shown in Figure 1.3b and its mean $\overline{S}_{ww}(l)$ and standard deviation $\text{STD}(S_{ww}(l))$ over the vectors of the same length are shown in Figures 1.3c. An

(a)

(b)

(c)
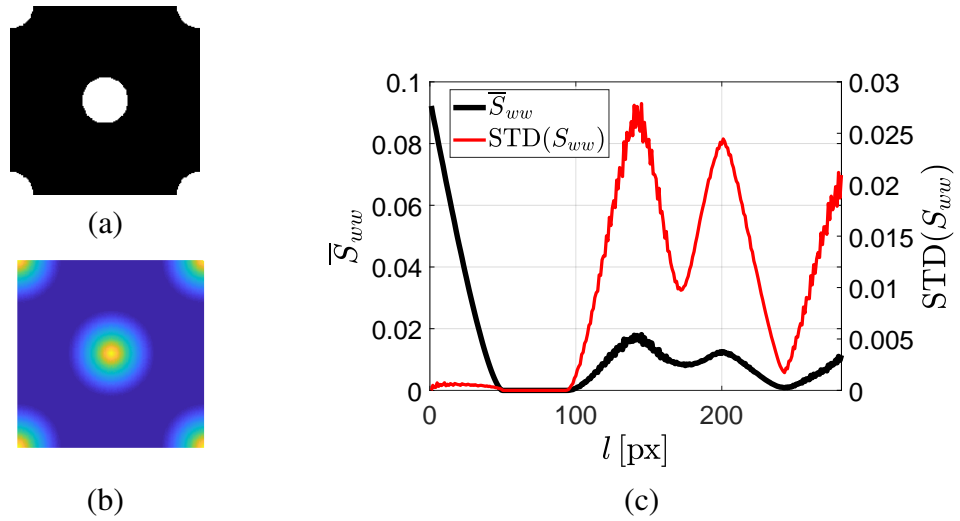
Figure 1.3: (a) Illustration of simple PUC with two equal-sized hard-disks. (b) Corresponding two-point correlation function. (c) Mean $\overline{S}_{ww}(l)$ and standard deviations $\text{STD}(S_{ww}(l))$ over vectors of same length.

anisotropy measure can be thus formulated as a sum of standard deviations in values of two-point probability function:

$$A = \sum_i \text{STD}(S_{ww}(l_i)) \tag{1.7}$$

for some choice of discrete vector lengths $l_i \in \{l_1, l_2, \ldots, l_a)\}$ and this value of anisotropy could be then considered as a cost function $A(\mathbf{c})$ for an optimization process, where center coordinates of hard-disks $\mathbf{c} = (x_1, y_1, x_2, y_2, \ldots, x_{N-1}, y_{N-1})$ represent optimized variables. To simplify the optimization process, the coordinates of N-th hard-disk are fixed to the origin of the coordinate system so that only $(N-1)$ hard-disks are optimized. For a microstructure with two hard-disks, the cost function has only two optimized variables and its shape is shown in Figure 1.4a. Figure 1.4c corresponds to the case of microstructure with 30 hard-disks, where all but one hard-disks are fixed to some possibly optimal values and the

<div align="center">(a)   (b)   (c)   (d)</div>
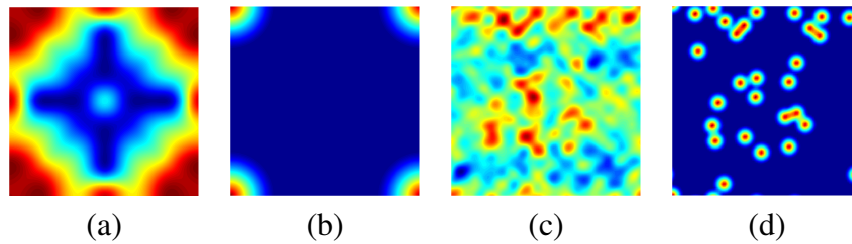
Figure 1.4: Illustration of terms in optimized cost function: (a) anisotropy and (b) penalty in cell with 2 disks, (c) anisotropy and (d) penalty in cell with 30 disks, all plotted in terms of coordinates of one selected disk, while other coordinates are fixed.

anisotropy value $A(\mathbf{c})$ is plotted for varying coordinates of only one hard-disk. This figure demonstrates high non-linearity and multi-modality of the anisotropy measure.

The optimization process is performed for different numbers of disks N and different values of their volume fraction V. Nevertheless, for some optimization algorithms, it can be more suitable to consider as feasible also solutions with smaller volume fraction due to some overlapping disks. Nevertheless, as these solutions are not preferred, they are penalized by the term added to the optimized cost function $F(\mathbf{c})$ as

$$F(\mathbf{c}) = A(\mathbf{c}) + \alpha |V - V(\mathbf{c})|, \tag{1.8}$$

where V is prescribed volume fraction to hard-disks, $V(\mathbf{c})$ is computed volume fraction of hard-disks in configuration $\mathbf{c}$ and α is a weighting factor fixed to $\alpha = 10^6$. The shape of the penalty term for the case of $V = 0.1$ is shown in Figures 1.4b and 1.4d for two and 30 hard-disks, respectively.

1.1.1.2 Optimization algorithms

The goal of this study is to test the performance of several available algorithms and namely of steady-state algorithms working with one solution such as Simulated Annealing and population-based genetic algorithms.

Metropolis algorithm (M)

is the simplest algorithm applied in Torquato (2006). This algorithm is used to generate Gibbs or equilibrium ensembles of interacting particles, spins, etc. Here a canonical ensemble is considered with a fixed number of N interacting hard-disks and fixed system volume V. The instantaneous particle positions at time t are denoted by $\mathbf{c}^t$. The simplest version of the Metropolis algorithm presented in Torquato (2006) and applied here starts with initial configuration of disks in a square array. Then a displacement is applied to each particle along each axis by random steps uniformly distributed in the interval $[-\delta, \delta]$, where δ is the maximum step size. The move is accepted if the particle on a new position does not overlap with any other particle. The cost function is not taken into account here and the algorithm is used only to generate some random configuration of hard-disks. A solution obtained by this algorithm is included just to justify the necessity of explicit optimization of the anisotropy and to quantify the contribution of optimization algorithms. The pseudocode of the Metropolis algorithm is written in Table 1.1(a).

```
1   Generate_initial_configuration();
2   for ( t=1; t<nsteps; t++ ){
3       for ( i=1; i<N; i++ ){
4           Move_particle(i, delta);
5           if Particle_not_overlap() {
6               Accept_move(); acm++; }
7           n++; }
8       if ((acm/n)>0.5) delta*=1.05;
9       else delta*=0.95; }
```
(a)

```
1    Generate_initial_configuration();
2    for ( t=1; t<nsteps; t++ ){
3        for ( i=1; i<N; i++ ){
4            Move_particle(i, delta);
5            if Particle_not_overlap() {
6                Evaluate_new_configuration();
7                if (F_new<F) {
8                    Accept_move(); acm++; }}
9            n++; }
10       if ((acm/n)>0.5) delta*=1.05;
11       else delta*=0.95; }
```
(b)

Table 1.1: Pseudocodes of (a) Metropolis algorithm and (b) Hill-climbing algorithm.

Hill-climbing algorithm (HC)

is a simple modification of the Metropolis algorithm, which takes into account a chosen cost function to be optimized. More specifically, the move of a particle is accepted only if the new position leads to a decrease in the cost function value. Its pseudocode is shown in Table 1.1(b), where the difference from the Metropolis algorithm consists of Lines 6 and 7. From the optimization point of view, such an algorithm is called the Hill-climbing algorithm as it enables only down-hill moves and is not able to escape from any local minimum.

Simulated Annealing (SA)

algorithm implemented here can be considered as a modification of the hill-climbing algorithm designed with the ability to escape from local extremes. The first version of this method was proposed by Kirkpatrick et al. (1983) and independently by Černý (1985). The core of SA is based on physical principles - on the analogy between the optimization problem and the annealing process of metals. In the physical process of annealing, the temperature of a solid is initially kept rather high and then decreases sufficiently slowly, so the individual crystals can attain the state with the minimal energy for a given constant temperature. As the temperature gradually decreases the energy of the whole body decreases as well and finally reaches the minimal value. SA works on the same principle - the initial solution is generated, some artificial parameter called temperature is set to initial value and the new solution is generated modifying the solution from the previous step. If the new solution is better than the preceding one in terms of the objective function, it is accepted automatically and replaces the original solution. However, if the new solution is worse than the preceding one, it has still a chance to replace the original solution enabling the solution to escape from a local minimum. The probability of accepting the worse solution depends on the difference of the cost function and actual temperature. The procedure is repeated several times for constant temperature and then the temperature is decreased until it reaches a certain prescribed minimum. The pseudocode of SA implemented here is presented in Table 1.2(a). Particularity of SA consists in Lines 2, $10 - 13$ and 17.

Simulated Re-Annealing

is again a modification of the preceding algorithm designed to even more increase its ability to escape from the local extreme. The original algorithm was proposed by Ingber (1989). The difference from SA consists of a faster cooling schedule and re-annealing phase, where

```
 1  Generate_initial_configuration();
 2  T = Tmax;
 3  for ( t=1; t<nsteps; t++ ){
 4    for ( i=1; i<N; i++ ){
 5      Move_particle(i, delta);
 6      if Particle_not_overlap() {
 7        Evaluate_new_configuration();
 8        if (F_new<F) {
 9          Accept_move(); acm++;
10        else
11          p=exp((F-F_new)/T);
12          if rand_num_uni(0,1) < p
13            Accept_move(); acm++; }}}
14    n++; }
15    if ((acm/n)>0.5) delta*=1.05;
16    else delta*=0.95;
17    T=T*pow(Tmin/Tmax,1/nsteps); }
```
(a)

```
 1  Generate_initial_configuration();
 2  T = Tmax;
 3  for ( t=1; t<nsteps; t++ ){
 4    for ( i=1; i<N; i++ ){
 5      Move_particle(i, delta);
 6      if Particle_not_overlap() {
 7        Evaluate_new_configuration();
 8        if (F_new<F) {
 9          Accept_move(); acm++;
10        else
11          p=exp((F-F_new)/T);
12          if rand_num_uni(0,1) < p
13            Accept_move(); acm++; }}}
14    n++; }
15    if ((acm/n)>0.5) delta*=1.05;
16    else delta*=0.95;
17    T=T*pow(Tmin/Tmax,10/nsteps);
18    if (T<Tmin) T=100*Tmax; }
```
(b)

Table 1.2: Pseudocodes of (a) Simulated Annealing and (b) Simulated Re-Annealing.

the temperature is increased again after achieving its minimum at the end of the cooling schedule. This step enables to escape from the local minimum, which is found here with higher probability due to faster cooling schedules. The particular version implemented here assumes six re-annealing phases within the maximum number of iterations. Its pseudocode in Table 1.2(b) contains its particularities in Lines 17 and 18.

Genetic algorithm (G)

is an algorithm operating with a population of feasible solutions simultaneously in contrast to all previously presented steady-state algorithms operating with just one individual. It is believed that population-based algorithm is usually more successful on multi-modal problems thanks to their ability to better explore whole domain space. When looking for such an algorithm, we have chosen the GRADE algorithm, which was successfully tested to solve highly multi-modal problems. Its detailed description, testing results, as well as the configuration of the algorithm control parameters, are presented in Kučerová (2007) or in Ibrahimbegović et al. (2004). For its brief introduction, it is a real-coded genetic algorithm applying one mutation and differential cross-over operator to create new individuals and inverse tournament selection operator for selecting a given number of solutions for the next population. One individual represents a vector of real-coded coordinates of particle centers $\mathbf{c}^i$. The genetic operators than create new individuals in the domain given for each variable by the size of microstructure volume. Nevertheless, genetic operators in their original form can create an individual with overlapping particles, therefore the penalization term is considered here in the cost function formulation. Pseudocode of this algorithm is shown in Table 1.3.

Genetic algorithm 2 (G2)

is the modification of the previous one motivated by its high percentage of penalized solutions. Particular modifications are listed below.

(i) The original GRADE algorithm starts with an initial population, where disks of each individual have coordinates randomly chosen from the prescribed domain. In such a

```
1   Generate_initial_population();
2   for ( i=0; i<n_generations; i++ ) {
3     Mutate();
4     Cross();
5     Evaluate();
6     Select(); }
```

Table 1.3: Pseudocode of GRADE algorithm

way, most of the configurations are penalized. Therefore, one individual with a square array configuration and no penalization is placed into the initial population. Thanks to the inverse tournament selection operator, this individual survives till some better solution without penalization is found.

(ii) The other problem concerns the identical configuration of disks obtained for different vectors of coordinates due to different ordering of disks inside the vector. Sorting of disks inside the vector of coordinates may facilitate the convergence because similar configurations having also the similar vector of coordinates can produce also similar offspring within the cross-over operator. Without this sorting, two parents with similar configurations may produce completely different offspring, which significantly slows down the convergence. Therefore, when two parental vectors are chosen[1], their coordinates are sorted to minimize their pair-wise distances, before applying the rule for generation of their offspring.

1.1.1.3 Comparison of algorithms in optimizing anisotropy of PUC

To compare the performance of the presented algorithms, the microstructure generation process was started 100 times for each algorithm. The number of considered hard-disks was prescribed to $N \in \{4, 9, 16, 25\}$ and their volume fraction to $V \in \{0.1, 0.2, \ldots, 0.7\}$. Each process was stopped after evaluating $2000 \times N$ solutions. Mean, maximum and minimum value of anisotropy measure at the end of the optimization process over 100 optimization runs are shown in Figure 1.5. The obtained results can be summarized in the following conclusions:

(i) The anisotropy of the unoptimized, just randomized solution obtained by the original Metropolis algorithm can be significantly improved by the optimization process and especially for cells with a higher number of hard-disks.

(ii) Despite many successful applications of GRADE algorithm presented in Kučerová (2007); Sýkora et al. (2013); Sýkora et al. (2018), it mostly failed in optimizing the hard-disks centers. Namely for higher volume fractions and higher number of disks, the GRADE algorithm resulted in average in worse solution than the one obtained by the Metropolis algorithm. Although the presented modifications to the GRADE algorithm were beneficial, they still did not sufficiently increase the number of offsprings, which are not penalized and the convergence of the algorithm remains significantly violated. In simplifying words, the algorithm is trying to make too big moves and

---

[1] In both the cross-over and mutation operators, the offspring is produced as a linear combination of two parental solutions. In cross-over, both parents are chosen from the population, while in mutation one parent comes from population and other is new randomly generated feasible solution.
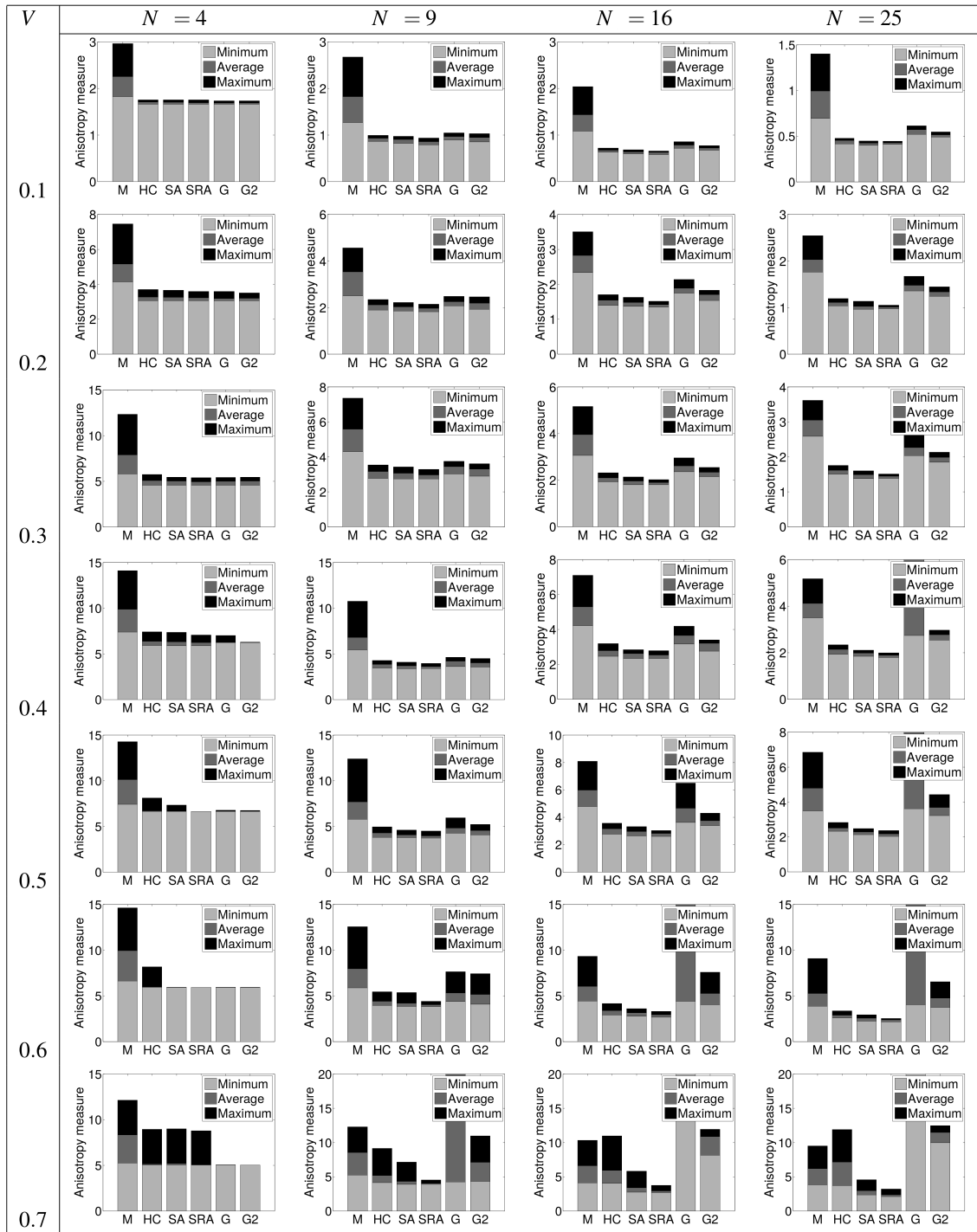
Figure 1.5: Statistics over anisotropy of obtained microstructures over 100 runs of each of the algorithms: M = Metropolis algorithm, HC = Hill-climbing algorithm, SA = Simulated Annealing, SRA = Simulated Re-Annealing, G = GRADE algorithm, G2 = modified GRADE algorithm, V = volume fraction of disks, N = number of disks.

most of them are penalized, which makes it being outperformed by all steady-state optimization algorithms.

(iii) Finally, all the consecutive improvements introduced to the Metropolis algorithm have proven their benefit and the Simulated Re-Annealing algorithm can be considered as the best suited for the elaborated problem.

1.1.2   Image-based covariance for microstructure reconstruction

A significantly faster method for reconstruction of material microstructure comparing to previously described optimization is based on random fields. Once having the characterization of microstructural spatial variations in terms of covariance function, the Karhunen-Loève expansion allows for a fast generation of samples with prescribed spatial covariance utilizing surprisingly few orthogonal modes from spectral decomposition of covariance matrix, see Adler and Taylor (2009). Jiang et al. (2013) use the method to reconstruct 3D microstructures from its 2D images. One significant drawback of their method is, however, construction of random field-based realizations for smooth Gaussian covariance function although the spatial evolution of correlation in the modeled microstructure is different. To address this inconvenience, they further refine the reconstructed image by optimizing the cluster function to corresponding better to the original image. Here, two strategies are proposed to derive more precise information about the spatial covariance directly from the image. Fist, the covariance function can be derived according to Lombardo et al. (2009) from the two-point probability function computed for the microstructure. Second, based on the erdogic assumption, the spectral decomposition of the image-based covariance can be derived from the samples of the microstructure employing principal component analysis (Jolliffe (2002)). Both these methods converge to the same solution and differ only numerically to the level of difference between the large piece of the original microstructure and its representation by a set of small segments. To quantify the effect of the considered spatial covariance, the proposed image-based covariances are compared with commonly used Gaussian and exponential functions in the following study.

1. Gaussian covariance kernel (GK) in two-dimensional space is given as

$$C(\mathbf{x_1}, \mathbf{x_2}) = \sigma_g^2 \exp\left(-\frac{(x_1 - x_2)}{2l_x} - \frac{(y_1 - y_2)}{2l_y}\right), \qquad (1.9)$$

where $\mathbf{x_1} = (x_1, y_1)$ and $\mathbf{x_2} = (x_1, y_2)$ are arbitrarily chosen two points, σ_g^2 is the variance of material property and $\mathbf{l} = (l_x, l_y)$ are the correlation lengths.

2. Exponential covariance kernel (EK) is defined as

$$C(\mathbf{x_1}, \mathbf{x_2}) = \sigma_g^2 \exp\left(-\left|\frac{x_1 - x_2}{l_x}\right| - \left|\frac{y_1 - y_2}{l_y}\right|\right). \qquad (1.10)$$

3. Image-based covariance kernel (IMGK) is derived from the two-point probability function $S_{rs}(\mathbf{x_1}, \mathbf{x_2})$ in a two-phase medium, where phases s and r have values of modeled material parameter prescribed to g_s and g_r, respectively. According to Lombardo et al. (2009), the covariance function is derived as

$$C(\mathbf{x_1}, \mathbf{x_2}) = \left(S_{ss}(\mathbf{x_1}, \mathbf{x_2}) - (c_s)^2\right)(g_s - g_r)^2, \qquad (1.11)$$

where c_s is the volume fraction of the phase s.

4. Image-based covariance based on sampling (IMGK-PCA) starts with a set of equal-sized image segments $\mathbf{g}_i$ arranged as rows of a matrix $\mathbf{G} \in \mathbb{R}^{n_s \times n_p}$, i.e. $\mathbf{G}^T = [\mathbf{g}_1 \ldots \mathbf{g}_{n_s}]$, where n_s is the number of segments and n_p is their length in pixels. The principal components analysis transforms them into the samples $\mathbf{Q} = \mathbf{GT}$, $\mathbf{Q} \in \mathbb{R}^{n_s \times n_p}$ in new coordinate system defined in fact by eigenmodes of the underlying covariance function of spatial fluctuations arranged as columns of transformation matrix $\mathbf{T} \in \mathbb{R}^{n_p \times n_p}$. According to the corresponding eigenvalues, only a subset of eigenmodes $\widetilde{\mathbf{T}} \in \mathbb{R}^{n_p \times n_m}$ allows to characterize the original segments with sufficient accuracy, i.e. truncated samples of principal components $\widetilde{\mathbf{Q}} = \mathbf{G}\widetilde{\mathbf{T}}$, $\widetilde{\mathbf{Q}} \in \mathbb{R}^{n_s \times n_m}$ approximate the original set of segments $\mathbf{G}$. Based on the samples $\widetilde{\mathbf{Q}}$ in the space of eigenmodes, it is possible to check also the distribution of the corresponding random variables. In the application presented here, the original samples of binomial variables transformed into samples distributed nearly normally. Therefore the application of Gaussian variables in the employed representation of random field was justified.

The first step of the presented study consists in the identification of correlation lengths used in the Gaussian and exponential covariance kernel functions, see Eq. (1.9) and Eq. (1.10), respectively. For a sake of clarity, a set of digitized images representing one-dimensional artificial structures 1×10^5 px is employed to illustrate the properties of particular correlation functions, see Fig. 1.6. The fitting process of the covariance lengths is relatively simple and intuitive, see also Kučerová et al. (2014) for more details. It starts from the calculation of the two-point probability function S_{ss} (according to Eq. (1.5)) of the original medium and subsequent derivation of the related covariance kernel, see Eq. (1.11). Then the correlation lengths in Gaussian and exponential kernels, see Eqs. (1.9) and (1.10), respectively, are optimized to fit these kernels to the one derived from the image. In our study, the in-house GRADE algorithm (Kučerová (2007); Kučerová et al. (2009) is utilized for such type of optimization. The computed results of optimized covariance lengths are listed in Tab. 1.4.

The results in Tab. 1.4 show that the computed correlation lengths are drastically smaller than the original dimensions of the investigated medium. Thanks to this fact, the dimension
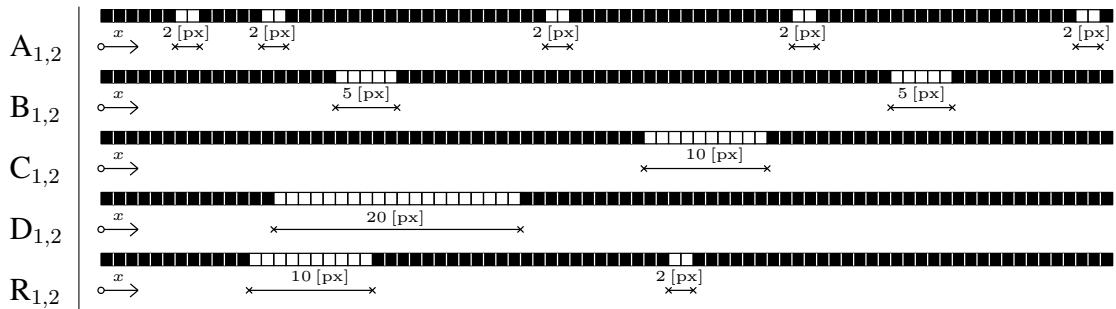
Figure 1.6: Reference media, size 1×10^5 px: A_1 - particles 2 px filling 10 % of vol.; A_2 - particles 1×2 px filling 50 % of vol.; B_1 - particles 5 px filling 10 % of vol.; B_2 - particles 5 px filling 50 % of vol.; C_1 - particles 1×10 px filling 10 % of vol.; C_2 - particles 10 px filling 50 % of vol.; D_1 - particles 1×20 px filling 10 % of vol.; D_2 - particles 20 px filling 50 % of vol.; R_1 - particles ranging from 2 px to 20 px filling 10 % of vol.; R_2 - particles ranging from 2 px to 20 px filling 50 % of vol.

		A_1	A_2	B_1	B_2	C_1	C_2	D_1	D_2	R_1	R_2
GK - l_x	[px]	0.75	0.50	1.91	1.10	3.78	2.33	7.66	4.77	4.91	2.48
EK - l_x	[px]	0.87	0.50	2.39	1.20	4.68	2.59	9.64	5.40	6.37	3.10

Table 1.4: Optimized covariance lengths calibrated for EK and GK.

of the problem can be reduced to $1 \times 100\,\mathrm{px}$, and thus the computational demands decrease to a reasonable level.

In the next step, $n_\mathrm{r} = 1000$ realizations of random fields with the length of $1 \times n_\mathrm{p} = 100\,\mathrm{px}$ were generated according to all four abovementioned representations of the covariance structure and for different number of employed eigenmodes. Each set of realizations was again arranged as rows of a matrix $\mathbf{R} \in \mathbb{R}^{n_\mathrm{r} \times n_\mathrm{p}}$, i.e. $\mathbf{R}^\mathrm{T} = [\mathbf{r}_1 \ldots \mathbf{r}_{n_\mathrm{r}}]$ and used to recompute their covariance matrix

$$\mathbf{C}_\mathrm{r} = \frac{1}{n_\mathrm{r} - 1} \mathbf{R}^\mathrm{T} \mathbf{R}\,. \tag{1.12}$$

The obtained covariance matrix $\mathbf{C}_\mathrm{r}$ is then compared with covariance matrix $\mathbf{C}_\mathrm{o}$ computed for 1000 randomly cutted segments from original medium. The comparison for the calculated covariances is shown in Fig. 1.7. The errors $e(\mathbf{C})$ computed as

$$e(\mathbf{C}) = \frac{\|\mathbf{C}_\mathrm{r} - \mathbf{C}_\mathrm{o}\|_{\mathrm{l}_2}}{\|\mathbf{C}_\mathrm{o}\|_{\mathrm{l}_2}} \tag{1.13}$$

are plotted as a function of the number of KL modes – M. There we can see the gain of image-based random fields construction and strong dependencies of errors for a small number of KL modes. These findings can be further explored in a real physical analysis propagating random fields.

1.2   Probabilistic identification of material model parameters

When dealing with models involving description of heterogeneous materials, the most important question concerns the real system, which is chosen to be modeled. Therefore, the following classification is introduced here to clearly distinguish two categories of systems, which are often confused when it comes to the identification of their parameters:

Deterministic system:

One object — specific laboratory specimen or more often specific real-world structure is called here as a deterministic system in a sense that it is one object with some specific material properties in every point of its spatial domain.

Stochastic system:

Set of objects — an ensemble or also group of objects having some common features defining the set, while some other their properties vary from one sample to another. For example, a
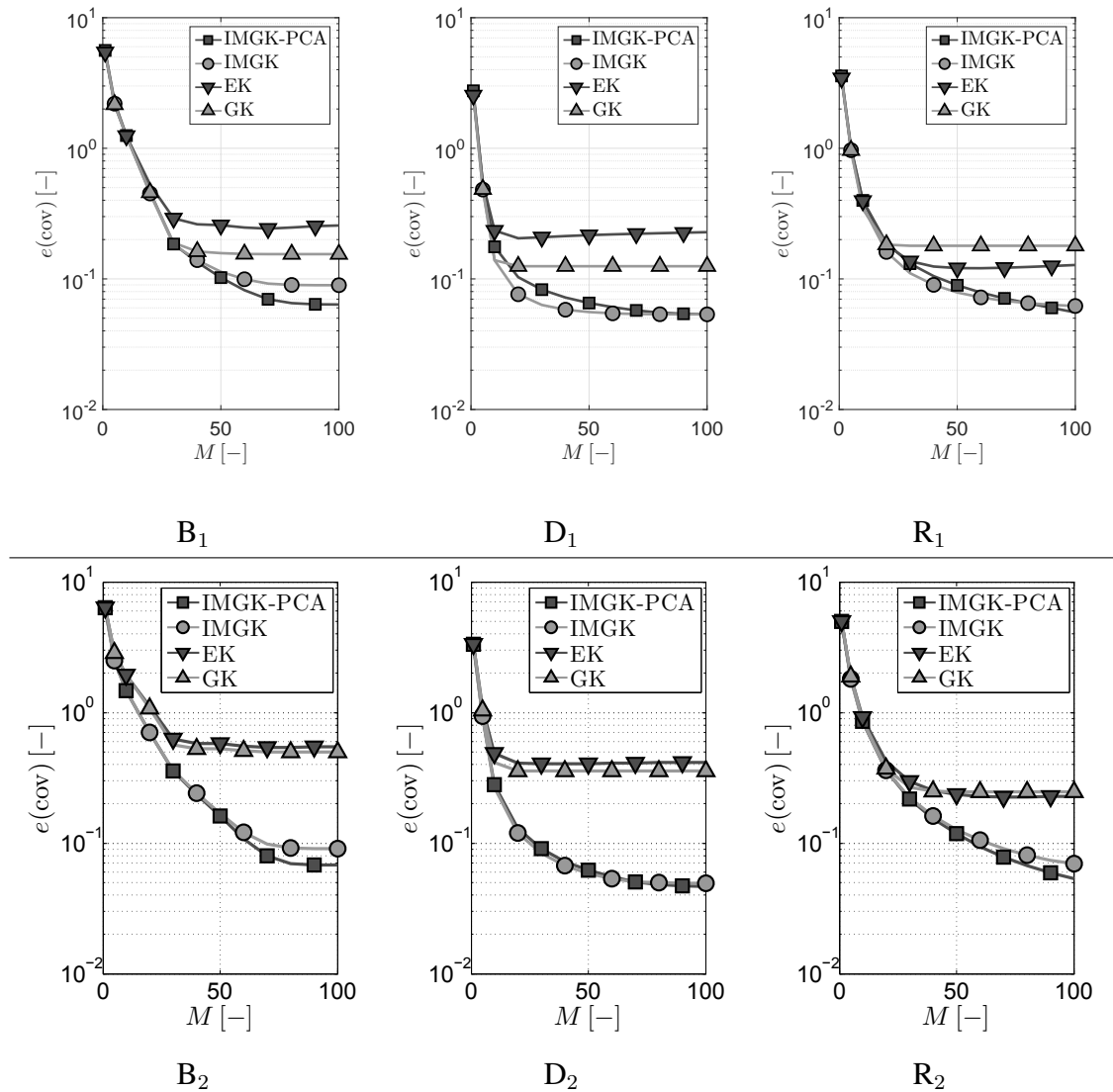
Figure 1.7: Relative error of covariance matrix as a function of KL terms – M.

set of laboratory specimens of heterogeneous concrete cylinders of the same size made of the same mix of the concrete, where the particular samples differ in their morphology. If the goal is to characterize whole system — a whole set of all possible cylinders belonging to the set, its material properties attain also a set of values, which can be modeled as random variables with some probability density function corresponding to the distribution of values among particular samples in the set.

The presented categories are not exactly the intrinsic properties of particular systems. *Modeling a system as deterministic or stochastic is, in fact, the result of a pragmatic choice of a modeler, whether he wants to model statistically a set of objects or one specific object, regardless of whether the object or set of objects are manufactured or just virtual*, see Figure 1.8.

In the field of uncertainty propagation through mechanical models, the careful rationale
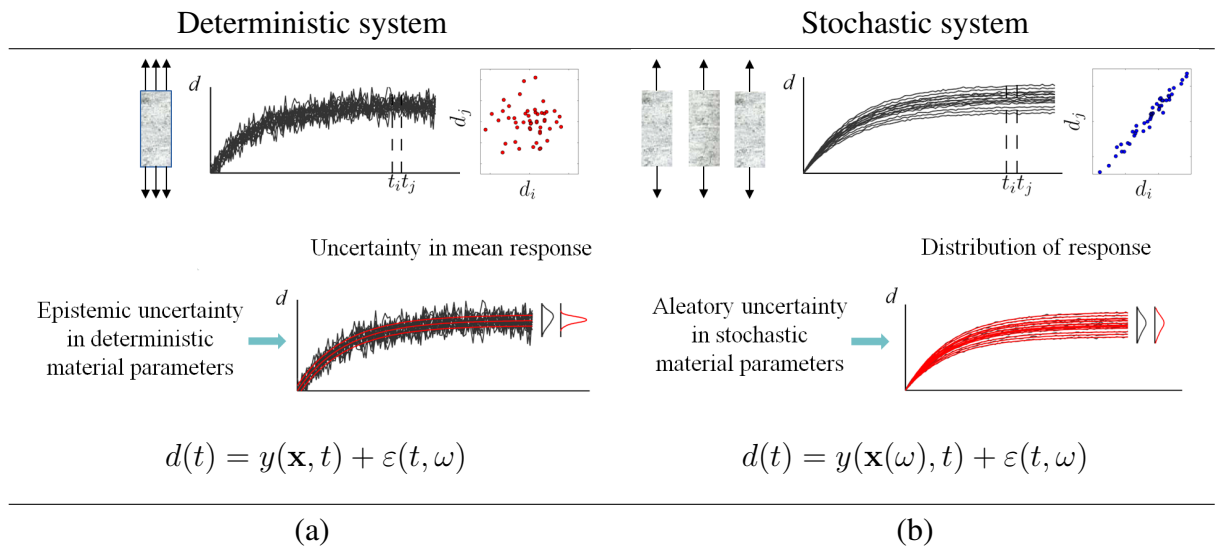
Figure 1.8: (a) One specimen with deterministic vector of material parameters $\mathbf{x}$ and (b) set of specimens with material parameters characterized by random variables $\mathbf{x}(\omega)$.

for modeling a system as stochastic or deterministic is, of course, also important, but it is also not subject to any misinterpretations. It becomes much more delicate when it comes to inverse analysis, or more specifically, to probabilistic identification of parameters of a modeled system trying to involve all relevant uncertainties. One of the widespread classifications of uncertainties in probabilistic parameter identification is the definition of the two following categories, which were often a subject of many discussions in past decades, see Der Kiureghian and Ditlevsen (2009).

> **Epistemic** uncertainty was often defined as reducible uncertainty caused by insufficient knowledge about the quantity of interest.
> **Aleatory** uncertainty was described as an irreducible, intrinsic property of the quantity of interest without a clear definition, what "intrinsic" means in this situation.

Considering uncertainties as epistemic or aleatory is very important, but it is again a choice of a modeler connected precisely to the chosen representation of the modeled system. When the system is modeled as stochastic, its properties varying among particular samples need to be modeled in some statistical manner, e.g. by a random variable. This randomness is then viewed within the inverse analysis as aleatory uncertainty, which is irreducible as it represents the random nature of the investigated properties among the considered set of objects. The goal of the modeler is then to find the correct distribution of these random quantities and not the reduce (sharpen or change in any sense) this distribution. On the other hand, epistemic uncertainties may be considered in modeling of both systems as these uncertainties reflect the lack of knowledge, which can be possibly reduced by additional observations, etc., see Figure 1.8. We refer an interested reader to Der Kiureghian and Ditlevsen (2009) for a brief review on this topic accompanied by a few clarifying examples.

Note on random inputs in homogenization-based models:

This work is focused especially on inverse problems connected to modeling heterogeneous materials with random morphology being hardly characterized by any periodic unit cell. Nevertheless, it does not implies that all models have to work with material properties described by random variables only. Even if a modeler aims to characterize a material as a stochastic system, he may decide to simplify the situation and work with some idealized averaged or statistically representative sample. This philosophy leads to homogenization-based models, where material properties attain again some deterministic values associated with the homogenized medium. Estimation of these homogenized or so-called effective properties may be again accompanied by some epistemic uncertainty about the correctness of the estimated deterministic values. Propagation of this uncertainty through a multi-scale model leads to stochastic multi-scale model, see e.g. Frank Xu (2007); Babuška et al. (2014); Guan et al. (2015); Chu and Guilleminot (2019). Characterization of this epistemic uncertainty may not be then interpreted as the variability of material properties within the heterogeneous materials as it is rather the modeling uncertainty reflecting our trust into the model itself.

Note on spatial variability and random fields:

It is possibly surprising that random fields are in the calibration of material models used mainly for characterizing deterministic systems, see e.g. Marzouk and Najm (2009); Kučerová and Matthies (2010); Kučerová and Sýkora (2013); Havelka et al. (2019); Yang et al. (2019); Uribe et al. (2020) or Chapter 3.1. In the listed applications, the random field is applied to characterize spatially varying material properties within one specific object. The goal of the parameter identification is to reduce the prior uncertainty about the deterministic values of material properties in each point of the modeled domain. Additional data brings a reduction of the resulting epistemic uncertainty, which is again related only to the correctness of the estimated deterministic values. As characterizing spatial variations within a continuous domain, the dimension of the inverse problem is, in fact, infinite as the number of points inside the domain.

Note on dimensionality reduction in estimating spatially varying properties:

To reduce the problem dimension, the spatial variation is commonly approximated by a truncated Karhunen-Loève expansion, where a limited number of random variables play a role of scaling factors for KL modes selected in the expansion. Particular modes are obtained as eigenfunctions of chosen covariance function, which is typically unknown and ad-hoc chosen to be smooth exponential or Gaussian function thus implying the assumption that the spatial variations are also smooth. The possibility to introduce also the non-smooth covariance functions based on the information about material morphology is already discussed in Section 1.1.2. If no information about the morphology is available, the introduction of smooth covariance is possible the only way to reduce the dimensionality of the inverse problem. It is interesting to point out that the same is valid not only for the random fields-based probabilistic formulation but also for the deterministic formulation of the problem, where the high problem dimensionality is commonly solved by the introduction of some regularization terms to the cost function. This regularization term reduces the oscillations in the optimized parameter field which has again the smoothing effect as consideration smooth covariance

function. Very recent and detailed discussion on the determination of spatial varying properties from boundary measurements in deterministic as well as in a probabilistic manner is presented in the doctoral thesis of Havelka (2019).

Note on stochastic models:

To cite an example of stochastic models, lattice/particle-based models aim to intrinsically model the stochasticity of heterogeneous material by stochastic lattice generation. In other words, their goal is to capture the aleatory uncertainty in material properties fully by randomized generation of micro- or mesostructural geometry based on the knowledge of the material constituents. Material properties to be calibrated can be then simplified again to deterministic material parameters being spatially constant. Their meaning is thus similar to homogenized or so-called effective properties. This category of models is called stochastic because for a given set of material parameters, different macroscopic response is obtained due to the random generation of the underlying geometry. The calibration of such models is then significantly complicated by this stochasticity. The calibration of this type of model is not much elaborated, but one example is discussed in very detail in Chapter 3.2.

1.2.1 *Quantification of aleatory uncertainty*

When no information about morphology or granularity of constituents is available, yet the goal is to model the material as a stochastic set (representing e.g. a type of concrete), the material parameters are no longer deterministic. Parameter identification turns into the quantification of aleatory uncertainty in material parameters described by random variables (or possibly fields), whose probability distribution characterizes the distribution of values of material properties within the set.

Quantification of aleatory uncertainties in parameters of a nonlinear physical model is significantly less elaborated than quantification of epistemic uncertainty in deterministic parameters mentioned above or in Chapter 3.1. Documented strategies to the quantification of aleatory uncertainty in parameters of nonlinear models can be divided into two following categories. Even though it is possible to characterize the aleatory uncertainty in material parameters by random fields, most of the works describe each material property by a random variable. Each sample of material is modeled with one sample of the material property being spatially constant with the domain of a specimen. The distribution of material property in a set of material samples can be thus viewed as a distribution of some type of effective (or homogenized) property among the set of samples.

1.2.1.1 *Deterministic fitting of aleatory uncertainty:*

The first category of methods is based on the parametrization of the probability density function (PDF) prescribed to each material parameter. The goal of the inverse analysis then turns into a search for the parameters of the prescribed PDF, which can be performed in a deterministic manner. Fonseca et al. (2005) compares the mean-centered first-order perturbation method and Monte Carlo-based maximum likelihood estimation (MLE) applied to determine statistical moments of the prescribed PDFs. The proposed methods have several crucial limitations. First, the assumption of a known type of statistical model is very limiting and prohibits the identification of higher statistical moments. The perturbation method (even

when extended by other nonlinear terms) is hardly applicable to nonlinear models and Monte Carlo-based MLE faces high computational requirements. Recently, Fang et al. (2014); Sepahvand and Marburg (2014) employed independently polynomial chaos (PC)-based surrogates to accelerate the identification of aleatory uncertainties formulated as a deterministic optimization problem. In particular, statistical moments or PC coefficients defining the statistical model of physical parameters are optimized so as to fit the corresponding moments of model response to the moments obtained from the experiments. While Fang et al. (2014) emphasizes that the PC-based surrogate of a model response provides an efficient way for computing sensitivities, Sepahvand and Marburg (2014) focuses on PC ability to effectively represent higher statistical moments and identify non-Gaussian parameters. Nevertheless, many related issues remain unsolved. For instance, quantification of related epistemic uncertainties is not considered at all. It means that the methodology is not able to reflect the number and the value of additional measurements. Also, the application of the proposed methods to the set of correlated observations, which are typically obtained by measuring e.g. load-displacement curves (see Figure 1.8) or by collecting the observations from a set of probes placed on each specimen can be problematic. In such a situation, the higher dimensionality of the observations leads to the increasing complexity of their joint distributions and thus to an increasing dimensionality or a number of the underlying optimization problems.

1.2.1.2 *Probabilistic estimation of aleatory uncertainty:*

Modeling aleatory uncertainty in material parameters together with related epistemic uncertainty in the correctness of estimated distributions is addressed in several recent works Arnst et al. (2010); Mehrez et al. (2012a,b); Debruyne et al. (2015). The authors use again the PC-based approximation of the distribution of material parameters, where the PC coefficients are obtained via the formulation of a likelihood function constructed for the measured data. Contrary to Fonseca et al. (2005), the PC coefficients are considered here as uncertain and identified in the Bayesian way by MCMC sampling. The introduction of new random variables for describing the PC coefficients thus allows us to quantify the epistemic uncertainty in the estimated distributions of aleatory material properties. Nevertheless, the method is derived only for a situation, where identified material parameters coincide with observed quantities. In the case of measuring a response of a nonlinear physical model, a set of deterministic inverse problems is solved first. Then quantification of aleatory uncertainty in estimated set of discrete parameter values follows in a separate step. Such separation is however not always feasible as the employed deterministic inverse problems are generally ill-posed. A similar concept was derived already much earlier for probabilistic estimation of statistical moments of PDFs prescribed to random variables Gelman et al. (2013); Šejnoha et al. (2017). Nevertheless, its generalization for estimation directly from observations of nonlinear response components was elaborated only recently in Behmanesh et al. (2015); Nagel and Sudret (2016), and its application to the estimation of viscoplastic material parameters was presented in Janouchová and Kučerová (2018). The principal limiting factor of this formulation for practical applications is the significant increase in problem dimensionality and computational effort with the number of material samples used to acquire the observations.

1.2.1.3   Nonlinear transformation of random variables:

There are two principal drawbacks of the above-mentioned methods to quantify aleatory uncertainty. First is their high computational effort needed even in the case when some surrogate of material model is employed. Second is the necessity to prescribe some particular distribution family to each material parameter and identify its statistical moments. The later can be partially relaxed by describing each random material parameter by polynomial function and search for its coefficients thus allowing to find some more general probability distribution. The preliminary choice of the particular polynomial degree with the given number of polynomial coefficients is, in fact, very similar limitation to the choice of a distribution family with some given number of statistical moments. Driven by this motivation, another less general, but computationally more effective strategy is presented in the following text. Its simple idea is based on nonlinear transformation of random variables and its efficiency is demonstrated on two simple examples with tens of observed samples. Moreover, the second example is devoted to situations, where data are collected from two distinct groups of samples, each employed for another destructive experiment, where none of the experiments allow to fully identify all the model parameters.

Considering the nonlinear model mapping a vector of random inputs $\mathbf{x}$ into a vector of random outputs $\mathbf{y}$ as

$$\mathbf{y} = \mathbf{g}(\mathbf{x}). \tag{1.14}$$

The nonlinear transformation of probability distribution $f(\mathbf{x})$ into distribution $f(\mathbf{y})$ is then given as

$$f(\mathbf{y}) = f(\mathbf{g}(\mathbf{x})) \cdot |J_{\mathbf{y}}(\mathbf{x})|, \tag{1.15}$$

where $J_{\mathbf{y}}(\mathbf{x})$ denotes the Jacobian of $\mathbf{g}(\mathbf{x})$. Making this change of variables explicit imposes certain constraints on the model $\mathbf{g}(\mathbf{x})$. First, the inverse image support $\mathcal{X}$ of the distribution $f(\mathbf{x})$ must be contained within the range of $\mathbf{y}$, i.e. $\mathcal{Y} = \mathbf{g}^{-1}(\mathbf{x})$. Second, $\mathbf{g}(\mathbf{x})$ must be a differentiable transformation from $\mathcal{X}$ to $\mathcal{Y}$ with a differentiable inverse. The constraints do not be fully fulfilled by the original material model, but they can be easily valid for its polynomial approximation employed for the acceleration of the transformation process. In such a case, the Jacobian can be easily obtained analytically. Moreover, the second constraint is not a great liability, when the transformation is performed by means of suitable sampling methods such as Markov chain Monte Carlo, see Marzouk et al. (2007).

The crucial condition of the method is that the mapping (1.14) must be bijective. In practice, laboratory experiments often results in some load-deflection curves discretized into a set of points corresponding to observed vector $\mathbf{d}$. The dimension n_{d} of $\mathbf{d}$ is thus typically much higher than the dimension n_{x} of $\mathbf{x}$ and particular components of $\mathbf{d}$ are highly correlated, see Figure 1.8b. Once the data from n_{r} repetitions of the experiments are collected into the matrix $\mathbf{D} = (\mathbf{d}_1, \dots, \mathbf{d}_{n_{\mathrm{r}}}) \in \mathbb{R}^{n_{\mathrm{d}} \times n_{\mathrm{r}}}$, the correlations between couples of particular response components can be computed directly from data. By means of principal component analysis (PCA) they can be easily transformed into a set of uncorrelated quantities and ordered according to their particular variance Jolliffe (2002). The first few components account for most of the statistical variability in all of the original data. PCA thus allows to reduce the dimensionality of observed data $\mathbf{D}$ to the set $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_{n_{\mathrm{r}}}) \in \mathbb{R}^{n_{\mathrm{x}} \times n_{\mathrm{r}}}$ of first n_{x} most important components without any significant information loss. The uncorrelation of the principal components, however, do not imply their statistical independence. Therefore, the construction of their joint probability density function $f(\mathbf{y})$ is again not trivial and one has

to keep in mind that its decomposition into a product of independent marginal distributions $f(\mathbf{y}) = \prod f(y_i)$ may contain certain error.

1.2.1.4 *Estimation of parameter distributions of cyclic loading test*

This example was a part of the project governed by European Space Agency, where the results are confidential and thus all the data are scaled. A set of fifty curves were simulated to represent pseudo-experimental data from the cyclic loading test on a heterogeneous viscoplastic material, see Figure 1.9.
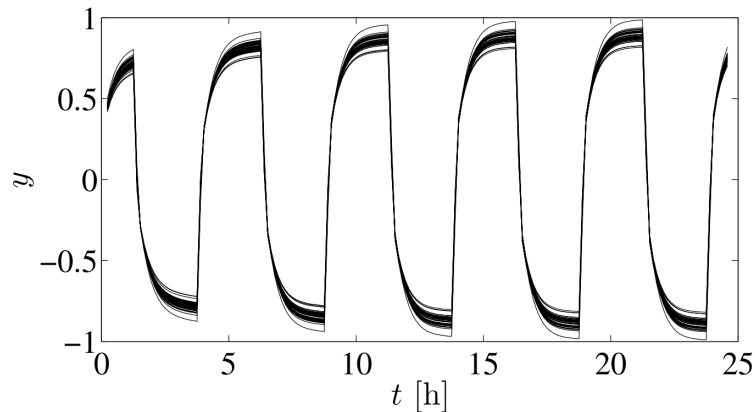
Figure 1.9: Fifty pseudo-experimental curves obtained from cyclic loading test.

The corresponding model has six material parameters and the model itself is considered as a black box. The given data set is produced for inputs where two of the parameters are strongly correlated and one parameter is insignificant the elaborated experiment. The goal is to identify the relevant parameters within the given loading test and their probability distributions corresponding to the distribution of the pseudo-experimental data.

The data obtained by discretization of the curves are first transformed into six most important principal components. Corresponding response components of the material model are approximated by sixth order Legendre polynomials constructed as optimal within the feasible intervals prescribed to each material parameter. The marginal distribution of each principal component is formulated as normal distribution with mean and standard deviation estimated from the samples and their joint distribution is obtained as a product of marginals under the simplifying assumption of their mutual independence. Joint distribution of material parameters is then obtained by MCMC sampling of the (1.15). The obtained probability density functions of material parameters are shown in Figure 1.10. The results confirm that the proposed method is able to determine sufficiently precisely the distributions of five relevant material parameters including the strong correlation between the first two of them. The incorrect distribution is obtained only for sixth parameter, which is insignificant in the simulated experiment and the data contain no information about it. Prediction of the model for identified distribution of random material inputs are compared with the prescribed pseudo-experimental data in Figure 1.11.
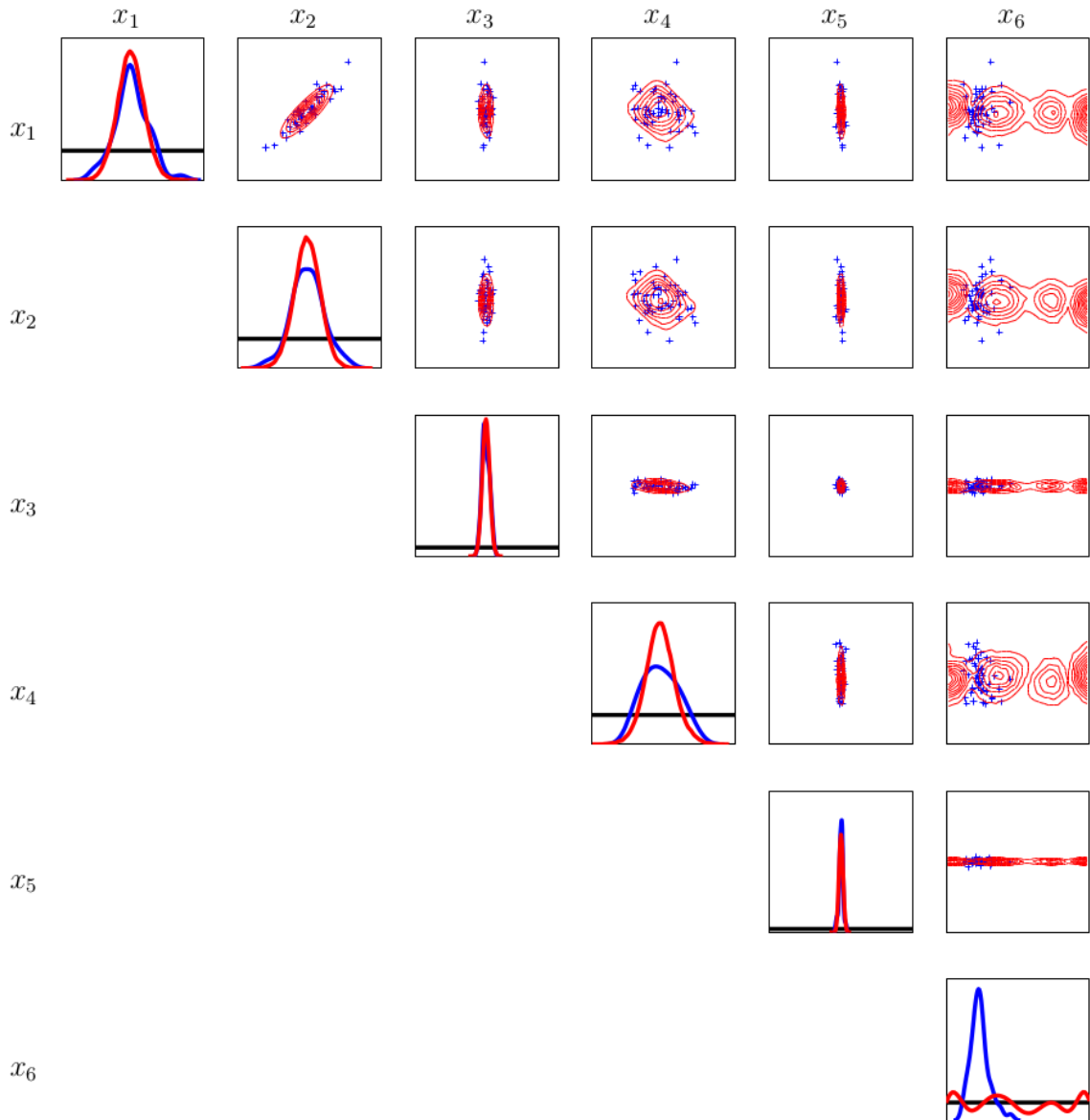
Figure 1.10: Identified 1D and 2D marginal PDFs of material parameters (red) and original samples used for simulating pseudo-experimental data (blue).

1.2.1.5 *Estimation of parameter distributions from independent tests*

This example concentrates on identification of distribution of two damage parameters from two distinct type of destructive experiments. The damage model is governed by the Landgraf - Morrow equation, see Landgraf (1970), where the relation between the strain range $\Delta \varepsilon$ and the number of cycles to a failure of the specimen N_f is given as

$$N_f = \left(\frac{\Delta \varepsilon}{S} \right)^{-s}. \tag{1.16}$$

Figure 1.11: Comparison of pseudo-experimental data and model predictions for identified distributions of material parameters.

The material parameters to be identified are the fatigue ductility coefficient S [-] and fatigue ductility exponent s [-]. According to the expert knowledge, the parameters are limited to feasible intervals $S \in [0.1; 1.2]$ and $s \in [1.0; 2.8]$.

In order to identify the damage parameters, two types of destructive experiments are considered and only pseudo-experimental data are used in this study. The first experiment is a tensile test with 50 repetitions, where the measured quantity is the strain at rupture and directly corresponds to the identified parameter S. In the second experiment with 30 realizations, the number of cycles to a failure of the specimen N_f is measured under cyclic loading with the strain range $\Delta\varepsilon = 0.03$. The histograms of the pseudo-experimental data and the corresponding marginal distributions are depicted in Figure 1.12.
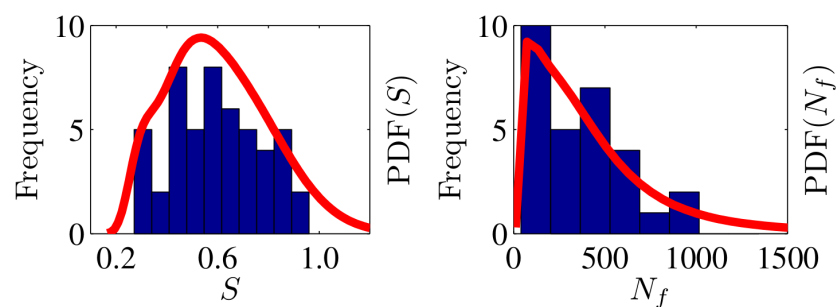
Figure 1.12: Distributions of pseudo-experimental observations.

Recording data from two distinct types of experiments on different specimens leads to missing information about the inherent correlation in the data. Therefore this correlation is estimated approximately based on model simulations for samples of material parameters generated from prescribed feasible intervals. The estimated correlation is combined with the estimated marginal distributions of data based on a normal kernel function to generate new synthetic data with desired correlation. These are used to built the approximation of

their joint probability density function, which is applied for the transformation of variables according to (1.15) to obtain the identified distribution of material parameters depicted in Figure 1.13.

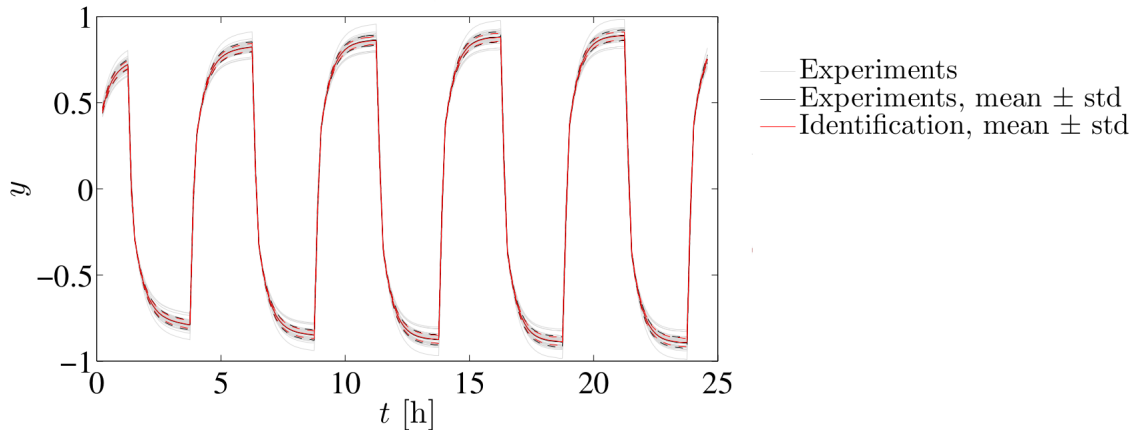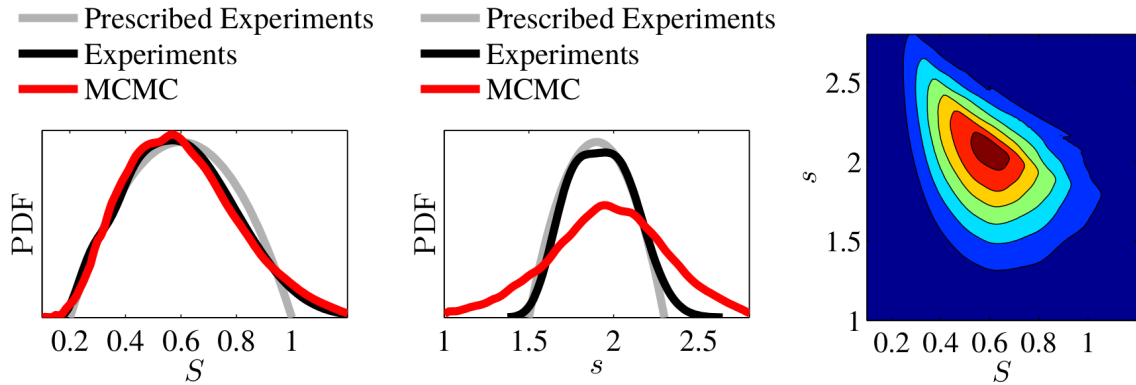Figure 1.13: Identified distribution of damage parameters - marginals and isolines of joint PDF.
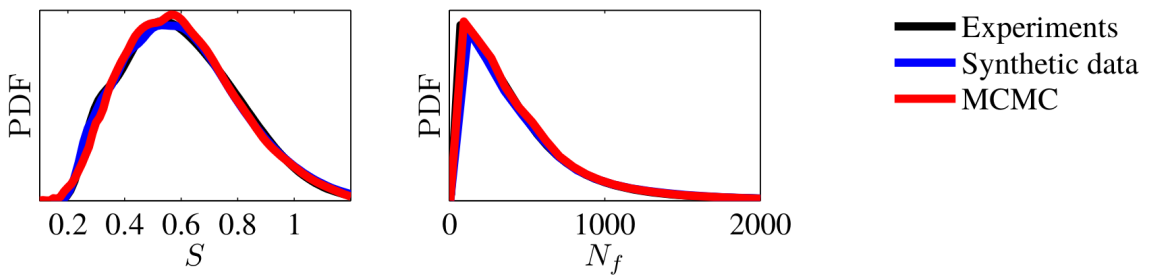
Figure 1.14: Comparison of pseudo-experimental data and model predictions for identified distributions of damage parameters.

The identification of the fatigue ductility coefficient S is essentially perfect, whereas variance of the fatigue ductility exponent s is not identified well. The obtained difference in the variance of s can be caused by inaccurate estimation of the joint distribution of both data components S and N_f, which is here principally unknown. Therefore is this joint distribution approximated at two levels. First, both components are modeled as simply linearly dependent with estimated correlation. Second, their correlation is only roughly estimated based on other simulated data sets obtained for parameters distributed in feasible intervals. Nevertheless, according to Figure 1.14 the known marginal distributions of both data components are fitted with high accuracy. As a possible simple extension of this work, the applied identification steps can be sequentially repeated. New estimate of data correlation can be obtained based on new simulations from identified distribution of material parameters instead of original feasible intervals and further applied to gain new updated estimation of parameter distributions.

References

Adler, R. J. and Taylor, J. E. (2009). *Random fields and geometry*. Springer Science & Business Media.

Arnst, M., Ghanem, R., and Soize, C. (2010). Identification of Bayesian posteriors for coefficients of chaos expansions. *Journal of Computational Physics*, 229(9):3134–3154.

Azzi, S., Sudret, B., and Wiart, J. (2020). Sensitivity analysis for stochastic simulators using differential entropy. *International Journal for Uncertainty Quantification*. Forthcoming Article.

Babuška, I., Motamed, M., and Tempone, R. (2014). A stochastic multiscale method for the elastodynamic wave equation arising from fiber composites. *Computer Methods in Applied Mechanics and Engineering*, 276:190 – 211.

Bažant, Z., Tabbara, M. R., Kazemi, M. T., and Pijaudier-Cabot, G. (1990). Random particle model for fracture of aggregate or fiber composites. *Journal of Engineering Mechanics*, 116(8):1686–1705.

Behmanesh, I., Moaveni, B., Lombaert, G., and Papadimitriou, C. (2015). Hierarchical Bayesian model updating for structural identification. *Mechanical Systems and Signal Processing*, 64–65:360–376.

Beran, P., Stanford, B., and Schrock, C. (2017). Uncertainty quantification in aeroelasticity. *Annual Review of Fluid Mechanics*, 49:361–386.

Bostanabad, R., Bui, A. T., Xie, W., Apley, D. W., and Chen, W. (2016). Stochastic microstructure characterization and reconstruction via supervised learning. *Acta Materialia*, 103:89–102.

Bostanabad, R., Zhang, Y., Li, X., Kearney, T., Brinson, L. C., Apley, D. W., Liu, W. K., and Chen, W. (2018). Computational microstructure characterization and reconstruction: Review of the state-of-the-art techniques. *Progress in Materials Science*, 95:1–41.

Burrus, C. S. and Parks, T. W. (1985). *DFT/FFT and Convolution Algorithms and Implementation*. John Wiley & Sons.

Chen, D., Teng, Q., He, X., Xu, Z., and Li, Z. (2014). Stable-phase method for hierarchical annealing in the reconstruction of porous media images. *Physical Review E*, 89(1):013305.

Chu, S. and Guilleminot, J. (2019). Stochastic multiscale modeling with random fields of material properties defined on nonconvex domains. *Mechanics Research Communications*, 97:39–45.

Collins, B. C., Matouš, K., and Rypl, D. (2010). Three-dimensional reconstruction of statistically optimal unit cells of multimodal particulate composites. *International Journal for Multiscale Computational Engineering*, 8(5):489–507.

Cundall, P. A. (1971). A computer model for simulating progressive, large-scale movement in blocky rock system. In *Proceedings of the International Symposium on Rock Mechanics*. Society for Rock Mechanics (ISRM), France.

Cusatis, G., Pelessone, D., and Mencarelli, A. (2011). Lattice discrete particle model (LDPM) for failure behavior of concrete. I: Theory. *Cement and Concrete Composites*, 33(9):881–890.

Černý, J. (1985). Thermodynamical approach to the traveling salesmanproblem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45:41–51.

De Schutter, G. and Taerwe, L. (1993). Random particle model for concrete based on Delaunay triangulation. *Materials and Structures*, 26(2):67–73.

Debruyne, S., Vandepitte, D., and Moens, D. (2015). Identification of design parameter variability of honeycomb sandwich beams from a study of limited available experimental dynamic structural response data. *Computers & Structures*, 146:197–213.

Der Kiureghian, A. and Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, 31(2):105 – 112.

Doškář, M. (2013). Wang tilings for real world material systems. Master's thesis, CTU in Prague.

Doškář, M. and Novák, J. (2016). A jigsaw puzzle framework for homogenization of high porosity foams. *Computers & Structures*, 166:33–41.

Doškář, M., Novák, J., and Zeman, J. (2014). Aperiodic compression and reconstruction of real-world material systems based on wang tiles. *Physical Review E*, 90(6):062118.

E, W., Engquist, B., Li, X., Ren, W., and Vanden-Eijnden, E. (2007). Heterogeneous multiscale methods: a review. *Communications in Computational Physics*, 2(3):367–450.

Efros, A. A. and Freeman, W. T. (2001). Image quilting for texture synthesis and transfer. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH'01, pages 341–346, New York, NY, USA. Association for Computing Machinery.

Fang, S.-E., Zhang, Q.-H., and Ren, W.-X. (2014). Parameter variability estimation using stochastic response surface model updating. *Mechanical Systems and Signal Processing*, 49(1– 2):249–263.

Fonseca, J. R., Friswell, M. I., Mottershead, J. E., and Lees, A. W. (2005). Uncertainty identification by the maximum likelihood method. *Journal of Sound and Vibration*, 288(3):587–599.

Frank Xu, X. (2007). A multiscale stochastic finite element method on elliptic problems involving uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 196(25–28):2723–2736.

Geers, M. G. D., Kouznetsova, V. G., and Brekelmans, W. A. M. (2010). Multi-scale computational homogenization: Trends and challenges. *Journal of Computational and Applied Mathematics*, 234(7):2175–2182.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC, third edition.

Ghanem, R. and Spanos, P. D. (1991). *Stochastic finite elements: A spectral approach*. Springer-Verlag, New York, first edition.

Grassl, P. and Bažant, Z. (2009). Random lattice-particle simulation of statistical size effect in quasi-brittle structures failing at crack initiation. *Journal of Engineering Mechanics*, 135(2):85–92.

Grigoriu, M. (2003). Random field models for two-phase microstructures. *Journal of Applied Physics*, 94:3762.

Guan, X., Liu, X., Jia, X., Yuan, Y., Cui, J., and Mang, H. A. (2015). A stochastic multiscale model for predicting mechanical properties of fiber reinforced concrete. *International Journal of Solids and Structures*, 56–57:280–289.

Gutiérrez, M. A. and Krenk, S. (2017). *Stochastic Finite Element Methods*, pages 1–25. American Cancer Society.

Hamdia, K. M., Silani, M., Zhuang, X., He, P., and Rabczuk, T. (2017). Stochastic analysis of the fracture toughness of polymeric nanoparticle composites using polynomial chaos expansions. *International Journal of Fracture*, 206(2):215–227.

Havelka, J., Kučerová, A., and Sýkora, J. (2016). Compression and reconstruction of random microstructures using accelerated lineal path function. *Computational Materials Science*, 122:102–117.

Havelka, J., Kučerová, A., and Sýkora, J. (2019). Dimensionality reduction in thermal tomography. *Computational Materials Science*, 122:102–117.

Havelka, M. (2019). *Application of Boundary Inverse Methods in Civil Engineering*. PhD thesis, CTU in Prague.

Hrennikoff, A. (1941). Solution of problems of elasticity by the framework method. *Journal of Applied Mechanics*, 8:A619–A715.

Ibrahimbegović, A., Knopf-Lenoir, C., Kučerová, A., and Villon, P. (2004). Optimal design and optimal control of structures undergoing finite rotations and elastic deformations. *International Journal for Numerical Methods in Engineering*, 61(14):2428–2460.

Ingber, L. (1989). Very fast simulated re-annealing. *Mathematical and Computer Modelling*, 12(8):967–973.

Janouchová, E. and Kučerová, A. (2018). Bayesian inference of heterogeneous viscoplastic material parameters. In Padevět, P., editor, *9th annual Conference NANO & MACRO MECHANICS 2018*, volume 15 of *Acta Polytechnica CTU Proceedings*, pages 41–45.

Janouchová, E., Sýkora, J., and Kučerová, A. (2018). Polynomial chaos in evaluating failure probability: A comparative study. *Applications of Mathematics*, 63(6):713–737.

Jiang, Z., Chen, W., and Burkhart, C. (2013). Efficient 3d porous microstructure reconstruction via gaussian random field and hybrid optimization. *Journal of Microscopy*, 525(2):135–148.

Jiao, Y., Stillinger, F. H., and Torquato, S. (2009). A superior descriptor of random textures and its predictive capacity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(42):17634–17639.

Jirásek, M. and Bažant, Z. P. (1995). Particle model for quasibrittle fracture and application to sea ice. *Journal of Engineering Mechanics*, 121(9):1016–1025.

Jolliffe, I. (2002). *Principal component analysis*. Wiley Online Library.

Kirkpatrick, S., Gelatt, J. C., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220:671–680.

Kučerová, A., Sýkora, J., Rosić, B., and Matthies, H. G. (2012). Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials. *Journal of Computational and Applied Mathematics*, 236(18):4862–4872.

Kučerová, A., Sýkora, J., and Zeman, J. (2014). Stochastictic modelling of heterogeneous materials based on image analysis. In *Engineering Mechanics 2014*, pages 371–374.

Kučerová, A. (2007). *Identification of nonlinear mechanical model parameters based on softcomputing methods*. PhD thesis, Ecole Normale Supérieure de Cachan, Laboratoire de Mécanique et Technologie.

Kučerová, A., Brancherie, D., Ibrahimbegović, A., Zeman, J., and Bittnar, Z. (2009). 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):128–144.

Kučerová, A. and Lepš, M. (2014). Soft computing-based calibration of microplane M4 model parameters: Methodology and validation. *Advances in Engineering Software*, 72:226–235.

Kučerová, A. and Matthies, H. G. (2010). Uncertainty updating in the description of heterogeneous materials. *Technische Mechanik*, 30(1–3):211–226.

Kučerová, A. and Sýkora, J. (2013). Uncertainty updating in the description of coupled heat and moisture transport in heterogeneous materials. *Applied Mathematics and Computation*, 219(13):7252—-7261.

Kumar, N. C., Matouš, K., and Geubelle, P. H. (2008). Reconstruction of periodic unit cells of multimodal random particulate composites using genetic algorithms. *Computational Materials Science*, 42:352–367.

Landgraf, R. W. (1970). The resistance of metals to cyclic deformation. In *Achievement of high fatigue resistance in metals and alloys*. ASTM International.

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

Marcon, M., Vorel, J., Ninčević, K., and Wan-Wendner, R. (2017). Modeling adhesive anchors in a discrete element framework. *Materials*, 10(8):917.

Mareš, T., Janouchová, E., and Kučerová, A. (2016). Artificial neural networks in calibration of nonlinear mechanical models. *Advances in Engineering Software*, 95:68–81.

Marrel, A., Iooss, B., Veiga, S. D., and Ribatet, M. (2012). Global sensitivity analysis of stochastic computer models with joint metamodels. *Statistics and Computing*, 22:833–847.

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

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

Matouš, K., Geers, M. G. D., Kouznetsova, V. G., and Gillman, A. (2017). A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *Journal of Computational Physics*, 330:192 – 220.

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

Mehrez, L., Doostan, A., Moens, D., and Vandepitte, D. (2012a). Stochastic identification of composite material properties from limited experimental databases, Part II: Uncertainty modelling. *Mechanical Systems and Signal Processing*, 27:484–498.

Mehrez, L., Moens, D., and Vandepitte, D. (2012b). Stochastic identification of composite material properties from limited experimental databases, part I: Experimental database construction. *Mechanical Systems and Signal Processing*, 27:471–483.

Muneo, H. (2018). *Introduction to computational earthquake engineering*. World Scientific.

Nagel, J. B. and Sudret, B. (2016). A unified framework for multilevel uncertainty quantification in Bayesian inverse problems. *Probabilistic Engineering Mechanics*, 43:68–84.

Niezgoda, S. R., Turner, D. M., Fullwood, D. T., and Kalidindi, S. R. (2010). Optimized structure based representative volume element sets reflecting the ensemble-averaged 2-point statistics. *Acta Materialia*, 58(13):4432–4445.

Nikolić, M., Karavelić, E., Ibrahimbegovic, A., and Miščević, P. (2018). Lattice element models and their peculiarities. *Archives of Computational Methods in Engineering*, 25(3):753–784.

Novák, J., Kučerová, A., and Zeman, J. (2012). Compressing random microstructures via stochastic wang tilings. *Physical Review E*, 86(4):040104.

Novák, J., Kučerová, A., and Zeman, J. (2013). Microstructural enrichment functions based on stochastic wang tilings. *Modelling and Simulation in Materials Science and Engineering*, 21(2):025014.

Odegard, G., Clancy, T., and Gates, T. (2017). Modeling of the mechanical properties of nanoparticle/polymer composites. In *Characterization of Nanocomposites*, pages 319–342. Jenny Stanford Publishing.

Ostoja-Starzewski, M. (2002). Lattice models in micromechanics. *Applied Mechanics Reviews*, 55(1):35–60.

Quarteroni, A., Manzoni, A., and Vergara, C. (2017). The cardiovascular system: Mathematical modelling, numerical algorithms and clinical applications. *Acta Numerica*, 26:365–590.

Rosić, B., Kučerová, A., Sýkora, J., Pajonk, O., Litvinenko, A., and Matthies, H. G. (2013). Parameter identification in a probabilistic setting. *Engineering Structures*, 50:179–196.

Rosić, B. and Matthies, H. G. (2015). Variational theory and computations in stochastic plasticity. *Archives of Computational Methods in Engineering*, 22:457–509.

Rosić, B., Sýkora, J., Pajonk, O., Kučerová, A., and Matthies, H. G. (2016). Comparison of numerical approaches to Bayesian updating. In *Computational Methods for Solids and Fluids*, pages 427–461. Springer.

Schauffert, E. A. and Cusatis, G. (2012). Lattice discrete particle model for fiber-reinforced concrete. I: Theory. *Journal of Engineering Mechanics*, 138(7):826–833.

Schröder, J., Balzani, D., and Brands, D. (2011). Approximation of random microstructures by periodic statistically similar representative volume elements based on lineal-path functions. *Archive of Applied Mechanics*, 81(7):975–997.

Schwab, C. and Zech, J. (2019). Deep learning in high dimension: Neural network expression rates for generalized polynomial chaos expansions in uq. *Analysis and Applications*, 17(1):19–55.

Sepahvand, K. and Marburg, S. (2014). Identification of composite uncertain material parameters from experimental modal data. *Probabilistic Engineering Mechanics*, 37:148–153.

Shahane, S., Aluru, N. R., and Vanka, S. P. (2019). Uncertainty quantification in three dimensional natural convection using polynomial chaos expansion and deep neural networks. *International Journal of Heat and Mass Transfer*, 139:613–631.

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

Sýkora, J., Jarušková, D., Šejnoha, M., and Šejnoha, J. (2018). Fire risk analysis focused on damage of the tunnel lining. *Fire Safety Journal*, 95:51–65.

Sýkora, J., Krejčí, T., Kruis, J., and Šejnoha, M. (2012). Computational homogenization of non-stationary transport processes in masonry structures. *Journal of Computational and Applied Mathematics*, 18:4745—-4755.

Sýkora, J., Šejnoha, M., and Šejnoha, J. (2013). Homogenization of coupled heat and moisture transport in masonry structures including interfaces. *Applied Mathematics and Computation*, 219(13):7275–7285.

Šejnoha, M., Janda, T., Melzerová, L., Nežerka, V., and Šejnoha, J. (2017). Modeling glulams in linear range with parameters updated using Bayesian inference. *Engineering Structures*, 138:293–307.

Tanaka, T. M. K. (1973). Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta Metallurgica*, 21(5):571–574.

Tewari, A. and Gokhale, A. M. (2004). Nearest-neighbor distances between particles of finite size in three-dimensional uniform random microstructures. *Materials Science and Engineering A*, 385(1–2):332–341.

Tornabene, F., Bacciocchi, M., Fantuzzi, N., and Reddy, J. N. (2019). Multiscale approach for three-phase CNT/polymer/fiber laminated nanocomposite structures. *Polymer Composites*, 40(S1):E102–E126.

Torquato, S. (2006). *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*. Springer; Corrected edition.

Torquato, S. and Stell, G. (1982). Microstructure of two-phase random media. I. The n-point probability functions. *The Journal of Chemical Physics*, 77(4):2071–2077.

Uribe, F., Papaioannou, I., Betz, W., and Straub, D. (2020). Bayesian inference of random fields represented with the Karhunen–Loève expansion. *Computer Methods in Applied Mechanics and Engineering*, 358:112632.

Vorel, J. and Šejnoha, M. (2009). Evaluation of homogenized thermal conductivities of imperfect carbon-carbon textile composites using the Mori-Tanaka method. *Structural Engineering and Mechanics*, 33:429–446.

Wang, X., Zhang, M., and Jivkov, A. P. (2016). Computational technology for analysis of 3d meso-structure effects on damage and failure of concrete. *International Journal of Solids and Structures*, 80:310–333.

Wei, L.-Y. and Levoy, M. (2000). Fast texture synthesis using tree-structured vector quantization. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH'00, pages 479–488, USA. ACM Press/Addison-Wesley Publishing Co.

Wiener, N. (1938). The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936.

Xiu, D. (2010). *Numerical methods for stochastic computations: a spectral method approach*. Princeton University Press.

Xiu, D. and Karniadakis, G. E. (2002). The wiener–askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644.

Yang, H.-Q., Zhang, L., Xue, J., Zhang, J., and Li, X. (2019). Unsaturated soil slope characterization with karhunen–loève and polynomial chaos via Bayesian approach. *Engineering with Computers*, 35(1):337–350.

Yeong, C. L. E. and Torquato, S. (1998a). Reconstructing random media. *Physical Review E*, 57(1):495.

Yeong, C. L. E. and Torquato, S. (1998b). Reconstructing random media. II. Three-dimensional media from two-dimensional cuts. *Physical Review E*, 58(1):224.

Zeman, J. and Šejnoha, M. (2007). From random microstructures to representative volume elements. *Modelling and Simulation in Materials Science and Engineering*, 15(4):325–335.

Zhang, D., Lu, L., Guo, L., and Karniadakis, G. E. (2019). Quantifying total uncertainty in physics-informed neural networks for solving forward and inverse stochastic problems. *Journal of Computational Physics*, 397:108850.

Chapter 2

REPRESENTATION OF RANDOM MATERIALS MORPHOLOGY

This chapter contains two parts devoted to advanced techniques for characterization and reconstruction of morphology in random materials. While the first technique concentrates on reducing long-range correlations arising in models based on a single representative cell, the second part elaborates numerical methods for evaluating higher-order statistical descriptors for morphology characterization. To be more specific, the first part introduces a new concept for representing material morphology by a set of representative cells – Wang tiles – allowing to reconstruct an aperiodic medium. The second part analyzes the properties of lineal path function, its benefits, and drawbacks in comparison with widely applied two-point probability function. To make the comparison computational feasible, the evaluation of lineal path function is accelerated by porting part of its code to a graphics processing unit using the CUDA (Compute Unified Device Architecture) programming environment.

List of selected journal papers and book chapters

Novák, J., Kučerová, A., and Zeman, J. (2012). Compressing random microstructures via stochastic wang tilings. *Physical Review E*, 86(4):040104.

Havelka, J., Kučerová, A., and Sýkora, J. (2016). Compression and reconstruction of random microstructures using accelerated lineal path function. *Computational Materials Science*, 122:102–117.

List of related journal papers and book chapters

Novák, J., Kučerová, A., and Zeman, J. (2013). Microstructural enrichment functions based on stochastic wang tilings. *Modelling and Simulation in Materials Science and Engineering*, 21(2):025014.

Compressing Random Microstructures via Stochastic Wang Tilings

Jan Novák

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Anna Kučerová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Jan Zeman

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Abstract

This paper presents a stochastic Wang tiling based technique to compress or reconstruct disordered microstructures on the basis of given spatial statistics. Unlike the existing approaches based on a single unit cell, it utilizes a finite set of tiles assembled by a stochastic tiling algorithm, thereby allowing to accurately reproduce long-range orientation orders in a computationally efficient manner. Although the basic features of the method are demonstrated for a two-dimensional particulate suspension, the present framework is fully extensible to generic multi-dimensional media.

Keywords: Microstructure compression, Reconstructing algorithms, Wang tiles, Aperiodic tilings

In 1961, Hao Wang introduced a tiling concept based on square dominoes with colored edges permitting their mutual assembly in a geometrically compatible (hard) manner [1]. Since then, his tiles have been the subject of studies in discrete mathematics [2, 3, 4] and found an extensive use in computer graphics [5], game industry [6], theory of quasicrystals [7] or biology [8]. From the perspective of this paper, the appealing feature of Wang tilings is that they can compress and reproduce naturally looking planar patterns or three-dimensional surfaces by employing only a small number of distinct tiles [9, 3, 10]. Motivated by this observation, we further explore the potential of Wang tiles to represent long-range spatial correlations in disordered microstructures; a problem common to materials science [11], geostatistics [12] or image analysis [13].

In this regard, two closely related applications can be distinguished, namely the *microstructure reconstruction* [14, 15, 16] based on given spatial statistics and *microstructure*

Email addresses: `novakj@cml.fsv.cvut.cz` (Jan Novák), `anicka@cml.fsv.cvut.cz` (Anna Kučerová), `zemanj@cml.fsv.cvut.cz` (Jan Zeman)

compression [17, 18, 19] aiming at efficient representation of materials structure in multi-scale computations [20]. Our focus is on the latter, since these procedures usually have the microstructure reconstruction techniques at heart, hereby covering the common features of both.

To the best of our knowledge, compression algorithms reported to date use a single cell (PUC) that is periodically extended to tile the plane in a deterministic manner [18]. Such structures then inevitably manifest strong long-range correlations with a period of the PUC dimensions. We shall demonstrate that these artifacts can be effectively controlled when utilizing small Wang tile sets [10, 3], carefully designed to capture morphological trademarks of compressed media, combined with fast stochastic Cohen-Shade-Hiller-Deussen (CSHD) tiling algorithm [5] for real-time texture generation. A potential of this approach will be demonstrated for equilibrium two-dimensional particulate suspensions consisting of equi-sized disks of radius ρ uniformly distributed in a homogeneous matrix, cf. [21].

To this goal, consider a two-dimensional microstructured domain $\mathcal{D}$ discretized by a regular square lattice. Each lattice cell contains specific morphological patterns that are compatible on contiguous boundaries, Fig. 1(b). If there are no missing cells in $\mathcal{D}$, the discretization is called a *valid tiling*, and a single cell is referred to as the *Wang tile* [1], Fig. 1(a). The tiles have different codes on their edges, lower-case Greek symbols in Fig. 1(a), and are not allowed to rotate when tiling a plane. The number of distinct tiles within $\mathcal{D}$ is limited, though arranged in such a fashion that none of them or any of their sub-sequence periodically repeats. The gathered distinct tiles are referred to as the *tile set*, Fig. 1(a). Sets that enable uncountably many, always aperiodic, tilings are called *aperiodic* [3, 4]. In real world applications, the assumption of strict aperiodicity of the tile sets is relaxed to aperiodicity of tilings, ensured e.g. by the CSHD algorithm [5] introduced next.
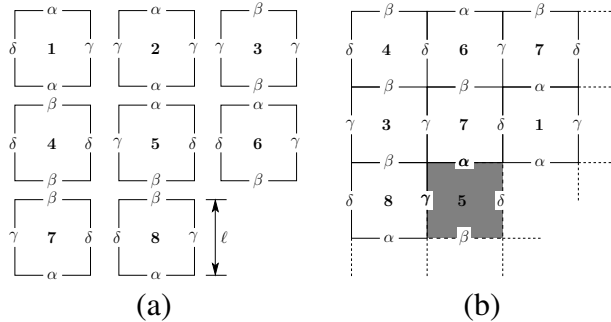
Figure 1: (a) A Wang tile set W8/2–2 with edge length ℓ and codes $\{\alpha, \beta, \gamma, \delta\}$. (b) Example of an aperiodic valid tiling.

Intuitively, the ability of a tile set to control long-range order effects arises from tile and edge code diversities, Fig. 1(a). Both factors are related, so that while the number of edge codes n_i^{c} in i–th spatial direction can be chosen arbitrarily, the number of tiles n^{t} must satisfy

$$n^{\mathrm{t}} = n^{\mathrm{NW}}\sqrt{n^{\mathrm{cs}}}, \tag{1}$$

where $n^{\mathrm{cs}} = (n_1^{\mathrm{c}} n_2^{\mathrm{c}})^2$ is the number of tiles in the complete set and $n^{\mathrm{NW}} = 2, \ldots, \sqrt{n^{\mathrm{cs}}}$ stands for the optional number of tiles with identical arrangement of north-western (NW) edge codes. The complete set of n^{cs} tiles is obtained by permuting the chosen codes c_i. In valid tilings, the south-eastern edge codes must match those assigned to NW edges, Fig. 1(b).

Thus, the tiles in the complete set are sorted according to existing NW edge code combinations and the desired number of tiles in a user defined set is formed by selecting n^{NW} of unique tiles from each group. Such a set is referred to as $\mathrm{W}n^{\mathrm{t}}/n_1^{\mathrm{c}}\text{--}n_2^{\mathrm{c}}$ in what follows. Notice that the W1/1−1 set corresponds to the PUC setting.

In the stochastic tiling algorithm, the index of a new tile to be placed is selected randomly with the uniform probability from an appropriate NW group compatible with the eastern code of the tile previously placed and the southern code of the tile above the one to be placed (edges γ and α in bold adjacent to shaded cell in Fig. 1(b)). Aperiodicity of the resulting tiling is guaranteed provided that the random generator never returns a periodic sequence of numbers and that each NW group contains at least two distinct tiles [5].

Analogously to the existing works on reconstruction and compression of random media, the tile morphologies are designed by an optimization procedure expressed in terms of suitable statistical descriptors. As our focus is to control long-range artifacts, we limit the exposition to the two-point probability function $S_2(\mathbf{x})$ [11]. For statistically uniform ergodic media, it provides the probability that two arbitrary points from $\mathcal{D}$, separated by $\mathbf{x}$, are simultaneously found in the particle phase. The function satisfies $S_2(\mathbf{0}) = \phi$, where ϕ is the volume fraction of particles, and $S_2(\mathbf{x}) \approx \phi^2$ for $\|\mathbf{x}\| > \lambda$ indicates the absence of long-range orders at the characteristic length λ, Fig. 2(b).

In the current setting, the Wang tiling compression consists of a set of n^{t} tiles of the edge length ℓ, in which we distribute n^{d} disks. The configuration of particles is determined by the parameter vector $[t_d, \xi_{1,d}, \xi_{2,d}]_{d=1}^{n^{\mathrm{d}}}$, where $t_d \in \{1, \ldots, n^{\mathrm{t}}\}$ specifies the parent tile index of the d-th disk and $\xi_{j,d} \in [0, \ell]$ the local position of the disk at j-th direction. To determine the two-point probability function $\widetilde{S}_2$ for a given configuration, we assemble a tiling that covers the domain of the same size as the representative sample $\mathcal{D}$, Fig. 2(a). Notice that such tiling corresponds to a realization of a statistically homogeneous material, since the tiles are selected from NW edge groups with the uniform probability. The proximity of the tile-based morphology to the original sample is quantified by an objective function

$$E = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} \left(S_2(\mathbf{x}) - \widetilde{S}_2(\mathbf{x}) \right)^2 \, \mathrm{d}\mathbf{x} \qquad (2)$$

which can be efficiently evaluated using the Fast Fourier Transform techniques, e.g. [11]. The minimization of (2) is carried out by the Simulated Re-Annealing method with compu-
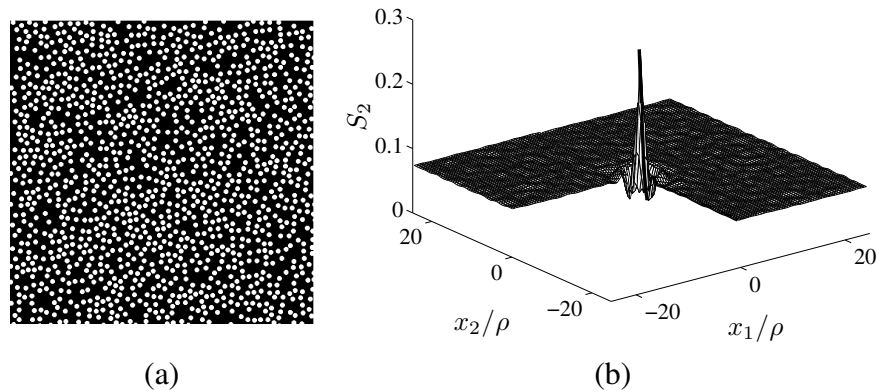
Figure 2: (a) Reference two-phase medium of size $174\rho \times 174\rho$ formed by equilibrium distribution of $1,300$ equi-sized disks of volume fraction 26.8% and (b) its two-point probability function S_2; ρ is the disk radius.

tational cost similar to existing PUC design strategies [22]. The algorithm ensures that the tiles in the set satisfy the corner constraint [5], requiring that the tile corners are not occupied by a disk, and determines the number of disks n^{d} and the cell size ℓ such that the local volume fractions associated with edges (grey disks in Fig. 5) and tile interiors (white disks in Fig. 5) are as close to the target value ϕ as possible.
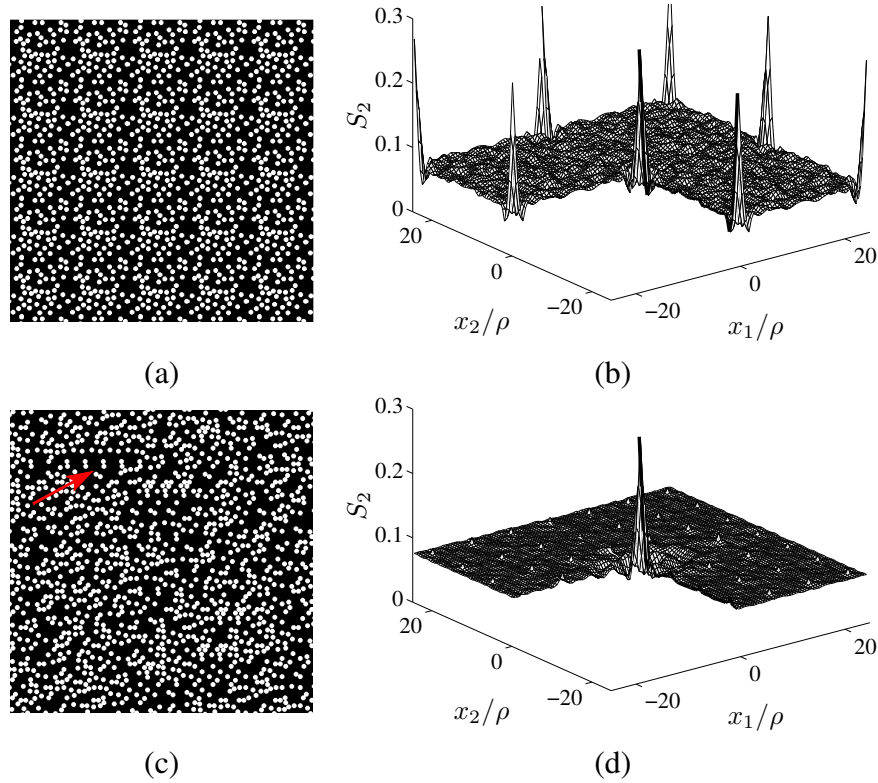
Figure 3: Optimized microstructures and two point probability functions $\widetilde{S}_2$ for PUC (a,b) and set W18/3–3 (c,d). The arrow in (c) denotes periodic region due to local character of tile placement in CSHD tiling algorithm.

An example of optimal approximations of the target microstructure from Fig. 2 in terms of a PUC and the Wang tile set W18/3–3 is shown in Fig. 3. The representations are based on $n^{\mathrm{d}} = 49$ particles and tile sizes $\ell = 24.5\rho$ and $\ell = 7.5\rho$, respectively. Evidently, both heterogeneity patterns carry long-range order effects with the period of ℓ, manifested as the local peaks

$$\widehat{S}_2 = \max_{\boldsymbol{k}\backslash\{\boldsymbol{0}\}} \widetilde{S}_2(\boldsymbol{k}\ell) \tag{3}$$

in the two-point probability functions, Fig. 3(b,d). Notice that $\widehat{S}_2$ is always equal to ϕ for the PUC approach, whereas the Wang tiles are capable of adjusting these artifacts by the proper morphology design. This is also reflected in visual regularity of the generated suspensions, compare Fig. 2(a) with Fig. 3(a,c). Also observe the locally periodic region in Fig. 3(c), arising from the local character of CSHD algorithm and from the lowest number of tiles in groups of admissible NW edge code combinations, $n^{\mathrm{NW}} = 2$. Such phenomenon is thus less likely when increasing parameter n^{NW}, however, at the expense of increasing set sizes, especially for higher edge code diversities, recall Eq. (1).

The principal features of Wang tile-based compressions are further illustrated in Fig. 5. Instead of relying on a single cell containing the complete morphological information (see
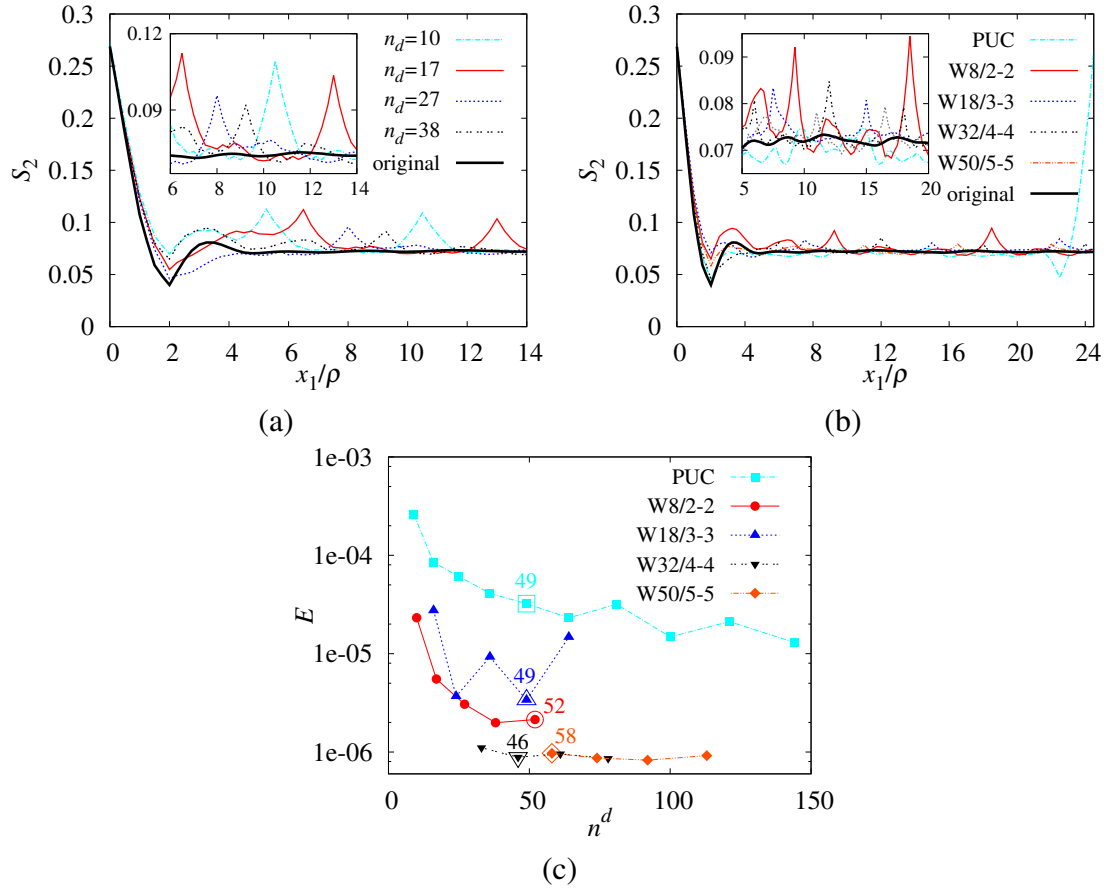
Figure 4: (Color online) Comparison of $\widetilde{S}_2$ in x_1–S_2 plane for (a) tile set W8/2-2 with respect to number of disks n^{d} involved, (b) different sets. (c) Comparison of different sets in terms of objective function (2) and with respect to number of disks n^{d} involved. The curves in (b) are plotted for particular values of n^{d} highlighted in (c).

Fig. 5(a)), the tiling-based approach utilizes substantially simpler building blocks, Fig. 5(b), assembled to comply with edge constraints (grey disks in Fig. 5). This, however, restricts the space of admissible disk configurations in Wang tiles compared to the single PUC design.

Now, we are in the position to quantify to which extent is the quality of reconstructed suspensions determined by the tile set diversity and the morphology design itself. This aspect is examined first in Fig. 4 by means of sections of the two-point probabilities $S_2(x_1, 0)$ and the objective function E, revealing that two effects govern the amplitude and period of the local extremes $\widehat{S}_2$. First, for a fixed tile set, increasing the number of disks increases the tile edge dimension ℓ (and thus the period) and slightly decreases the amplitude, Fig. 4(a). On the other hand, increasing the number of tiles decreases the period as well as the magnitude of local extremes, Fig. 4(b). Also notice that the quality of the W18/3–3 set in terms of the objective function (2) is systematically inferior to W8/2-2, Fig. 4(c). This is caused by an inaccurate representation of disk volume fraction for the former set, which pollutes the shape of $\widehat{S}_2$ statistics, Fig. 4(b) [1]. It further follows from Fig. 4(c) that increasing the

[1]The best resulting disk volume fraction for W18/3–3 and 24 disks was by 1.3% higher than the prescribed value. The remaining sets, however, never resulted in a worse scatter than 0.2%.
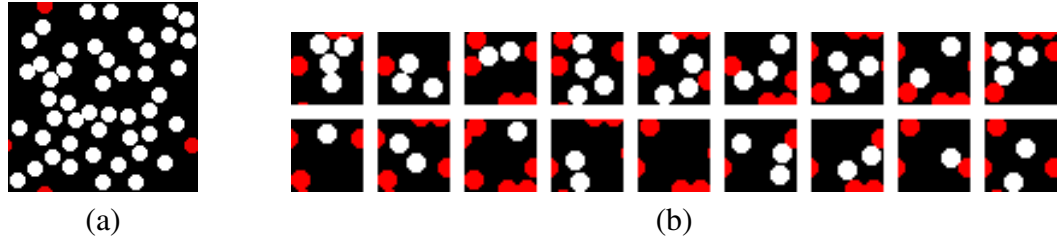
Figure 5: Building blocks of microstructure compression based on (a) PUC and (b) tile set W18/3–3 with 49 disks assigned to tile edges (grey) and interiors (white).

tile set diversity is much more efficient; for sets containing more than 32 tiles, the error is almost independent of the number of disks. This saturation value reflects the inaccurate representation of short-range values of S_2, caused by the particular form of the objective function (2). If needed, the local details can be incorporated in terms of higher-order statistics or specifically tailored descriptors [14, 15, 24, 25].

It is now clear that the local extremes can be attributed to a limited number of tiles used in a repetitive, although random fashion. Actually, two components repeat when tiling the plane: tile edges and interiors. To study the local artifacts analytically, we consider user defined sets with tiles selected so that their edges incorporate each code c_i at least once. Assuming that tiles and edges repeat independently, the maximum local extremes can be estimated as

$$\widehat{S}_2^{\text{p}} \approx \frac{\phi^{\text{t}}}{n^{\text{t}}} \left[\phi + (n^{\text{t}} - 1) \phi^2 \right] + \max_i \left\{ \frac{\phi^{\text{e}}}{n_i^{\text{c}}} \left[\phi + (n_i^{\text{c}} - 1)\phi^2 \right] \right\} \tag{4}$$

where $\phi^{\text{t}} = (\ell - 4\rho)^2/\ell^2$ and $\phi^{\text{e}} = 1 - \phi^{\text{t}}$ denote the volume fractions of tile interiors (occupied by white disks in Fig. 5) and edges (occupied by grey disks in Fig. 5), respectively [2].

In Fig. 6, we compare the actual values of $\widehat{S}_2$ with theoretical predictions (4) for several values of ϕ^{e}. Apart from the limit cases, $\phi^{\text{e}} \in \{0, 1\}$, $\widehat{S}_2^{\text{p}}$ was also explored for $\phi^{\text{e}} = 0.2$ (average value from all considered tile sets). We observe that an almost exact match is obtained for the lower bound with $\phi^{\text{e}} = 0$, red solid curve in Fig. 6(a), demonstrating that the long-range artifacts are carried mainly by the tile interiors. This is rather surprising, since all considered tile sets satisfy $n^{\text{t}} \gg n_1^{\text{c}} = n_2^{\text{c}}$, so that edges repeat more often than the tile interiors. Moreover, the magnitude of spatial artifacts converges rapidly to the limit value ϕ^2. Altogether, this leads us to the conclusion that artifacts due to discrete nature of Wang tilings can be almost eliminated by a proper morphology optimization. Also note that the accuracy of the estimate (4) appears to be reasonable, both for the average value of $\phi^{\text{e}} = 0.2$, blue double dotted curve in Fig. 6(a), as well as for values corresponding to individual tile sets, Fig. 6(b). It may thus serve as a basis for the a-priori selection of the tile set parameters n_i^{c} and n^{t}.

---

[2]Observe that the estimate (4) contains contributions from tile interiors and edges. In addition, the tile interior part arises from two complementary events. If the two adjacent tiles are identical, the probability of simultaneously locating two disks distant by ℓ amounts to $\phi^{\text{t}}\phi/n^{\text{t}}$. Otherwise, we consider the disks in both tiles as independent which gives rise to the term $\phi^{\text{t}}(1 - 1/n^{\text{t}})\phi^2$. The contribution of repeated tile edges related to i–th spatial direction is established analogously, by estimating the probability of simultaneously matching two edges distant by ℓ as $1/n_i^{\text{c}}$.
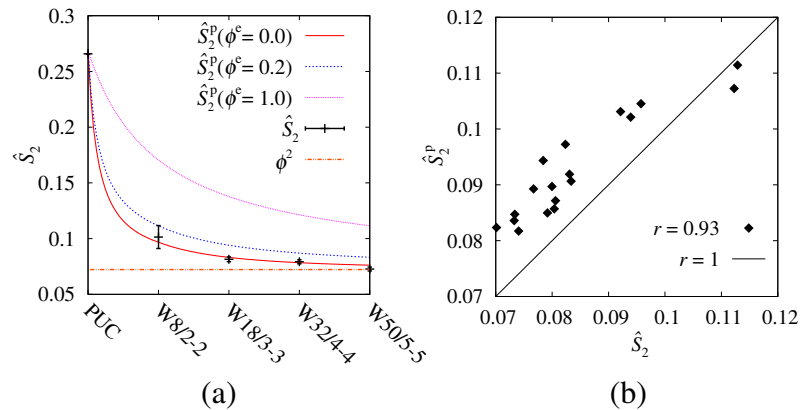
Figure 6: (Color online) (a) Dependence of local extremes $\widehat{S}_2$ on the particular tile set. (b) Correlation of local peaks obtained from two-point probabilities of optimized microstructures and their predictions given by Eq. (4); r in (b) denotes the Pearson correlation coefficient.

Summary.

A new compression/reconstruction technique based on Wang tilings has been proposed and applied to two-dimensional microstructures of disordered particulate media. The technique is extensible to generic three-dimensional microstructures, adopting the frameworks of Wang cubes [27, 28] and image synthesis [5]; it substantially generalizes the periodic unit cell concept by making use of multiple tiles instead of a single cell, preserves long range spatial features, and is computationally efficient. A formula for estimates of long-range order spatial artifacts has also been proposed and verified for the studied material system.

Acknowledgement.

[1] H. Wang, Bell. Syst. Tech. J. **40**, 1 (1961).

[2] R. M. Robinson, Invent. Math. **12**, 177 (1971).

[3] B. Grünbaum and G. Shephard, *Tilings and patterns* (WH Freeman & Co., 1986).

[4] K. Culik, Discret. Math. **160**, 245 (1996).

[5] M. Cohen, J. Shade, S. Hiller, and O. Deussen, ACM Trans. Graph. **22**, 287 (2003).

[6] E. Demaine and M. Demaine, Graphs Comb. **23**, 195 (2007).

[7] D. Aristoff and C. Radin, J. Phys. A – Math. Theor. **44**, 255001 (2011).

[8] E. Winfree, F. Liu, L. Wenzler, and N. Seeman, Nature **394**, 539 (1998).

[9] A. Glassner, IEEE Comput. Graph. Appl. **24**, 86 (2004).

[10] A. Lagae and P. Dutré, Comput. Graph. Forum **27**, 114 (2008).

[11] S. Torquato, *Random heterogenous materials* (Springer-Verag, New York, 2002).

[12] J. Chilés and P. Delfiner, *Geostatistics* (John Wiley & Sons, 1999).

[13] J. Serra, *Image analysis and mathematical morphology* (Academic Press, London, 1982).

[14] C. Yeong and S. Torquato, Phys. Rev. E **57**, 495 (1998).

[15] P. Čapek, V. Hejtmánek, L. Brabec, A. Zikánová, and M. Kočiřík, Transp. Porous Media **76**, 179 (2009).

[16] M. Davis, S. Walsh, and M. Saar, Phys. Rev. E **83**, 026706 (2011).

[17] G. L. Povirk, Acta Metall. Mater. **43**, 3199 (1995).

[18] J. Zeman and M. Šejnoha, Model. Simul. Mater. Sci. Eng. **15**, S325 (2007).

[19] H. Lee, M. Brandyberry, A. Tudor, and K. Matouš, Phys. Rev. E **80**, 061301 (2009).

[20] D. T. Fullwood, S. R. Niezgoda, B. L. Adams, and S. R. Kalidindi, Prog. Mater. Sci. **55**, 477 (2010).

[21] M. D. Rintoul and S. Torquato, J. Colloid Interface Sci. **186**, 467 (1997).

[22] J. Novák, A. Kučerová, and J. Zeman, "Microstructural enrichment functions based on stochastic Wang tilings," available at `http://arxiv.org/abs/1110.4183` (2011).

[23] The best resulting disk volume fraction for W18/3–3 and 24 disks was by 1.3% higher than the prescribed value. The remaining sets, however, never resulted in a worse scatter than 0.2%.

[24] Y. Jiao, F. Stillinger, and S. Torquato, Proc. Natl. Acad. Sci. U. S. A. (2009).

[25] R. Piasecki and A. Plastino, Physica A **389**, 397 (2010).

[26] Observe that the estimate (4) contains contributions from tile interiors and edges. In addition, the tile interior part arises from two complementary events. If the two adjacent tiles are identical, the probability of simultaneously locating two disks distant by ℓ amounts to $\phi^{\mathrm{t}}\phi/n^{\mathrm{t}}$. Otherwise, we consider the disks in both tiles as independent which gives rise to the term $\phi^{\mathrm{t}}(1 - 1/n^{\mathrm{t}})\phi^2$. The contribution of repeated tile edges related to i–th spatial direction is established analogously, by estimating the probability of simultaneously matching two edges distant by ℓ as $1/n_i^{\mathrm{c}}$.

[27] K. Culik and J. Kari, J. Univers. Comput. Sci. **1**, 675 (1995).

[28] A. Lu, D. S. Ebert, W. Qiao, M. Kraus, and B. Mora, ACM Trans. Graph. **26** (2007).

Compression and Reconstruction of Random Microstructures Using Accelerated Lineal Path Function

Jan Havelka

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Anna Kučerová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Jan Sýkora

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Abstract

Microstructure reconstruction and compression techniques are designed to identify microstructures with desired properties. While a microstructure reconstruction involves searching for a microstructure with prescribed statistical properties, a microstructure compression focuses on efficient representation of material morphology for the purpose of multiscale modelling. Successful application of these techniques, nevertheless, requires proper understanding of the underlying statistical descriptors quantifying morphology of a material. In this paper, we focus on a lineal path function designed to capture short-range effects and phase connectedness, which can hardly be handled by the commonly used two-point probability function. Usage of the lineal path function is, however, significantly limited because of huge computational requirements. So as to examine the properties of the lineal path function during computationally exhaustive compression and reconstruction processes, we start with an acceleration of the lineal path evaluation, namely by porting part of its code to a graphics processing unit using the CUDA (Compute Unified Device Architecture) programming environment. This allows us to present a unique comparison of the entire lineal path function with the commonly used rough approximation based on the Monte Carlo and/or sampling template. Moreover, this accelerated version of the lineal path function is then compared to the two-point probability function during the compression and reconstruction of two-phase morphologies. Their significant features are discussed and illustrated using a set of artificial periodic as well as real-world random microstructures.

Keywords: Lineal path function, Two-point probability function, Statistically equivalent periodic unit cell, Microstructure reconstruction, Microstructure compression, Graphics processing unit, Compute Unified Device Architecture

Email addresses: `jan.havelka.1@fsv.cvut.cz` (Jan Havelka), `anicka@cml.fsv.cvut.cz` (Anna Kučerová), `jan.sykora.1@fsv.cvut.cz` (Jan Sýkora)

1. Introduction

Computational modelling of random heterogeneous materials is a nontrivial multi-disciplinary problem with a wide range of relevant engineering applications. FE^2-methods have been developed as promising techniques for material modelling and have been used to derive effective models at various scales of interest. A unifying theoretical framework has been provided by homogenization theories aiming to replace heterogeneous microstructures with equivalent homogeneous materials, see [1]. Currently, two main approaches are available: (i) computational homogenization and (ii) effective media theories.

The latter approach estimates material response analytically on the basis of limited geometrical information (e.g. volume fractions of constituents) about an analysed medium. Structural imperfections are introduced in the cumulative sense using one of the averaging schemes, e.g. the Mori-Tanaka method [2]. The computational requirements are very low, however, because such an analytical solution is available only for a limited spectrum of microstructural geometries, such as media with specific inclusion shapes.

Methods based on computational homogenization have more general applications, see [3, 4]. They study the distribution of local fields within a typical heterogeneity pattern using a numerical method. It is generally accepted that detailed discretisation techniques and the finite element method, in particular, remain the most powerful and flexible tools available. Despite tedious computational times, these provide us with details about local fields, see e.g. [5, 6]. The principal requirement of these techniques is finding a representative volume element (RVE). This can be intriguing when considering real-world random microstructures. Recent studies suggest that structure preserving spatial geometrical statistics, such as a statistically equivalent periodic unit cell (SEPUC) [7], a statistically similar representative volume element (SSRVE) [8] and/or RVE Sets introduced in [9], is computationally very efficient when compared to classical RVE concept.

A relatively new concept in microstructure modelling is based on the production of a set of structures morphologically similar to the original media, so called Wang tiles, see [10, 11]. This approach allows us to obtain aperiodic local fields in heterogenous media with a small set of statistically representative tiles. The main advantages of stochastic Wang tilings are their computational efficiency and long range spatial correlations, neglected in classical homogenization techniques, see [10]. In some cases tiles can be produced by a computationally efficient image quilting algorithm [12] or generally by optimising a chosen statistical descriptor.

This paper is devoted to statistical descriptors defining statistically/mor-phologically similar material structures (cells or tiles). Such structures are generally obtained using a process of microstructure reconstruction [13, 14, 15] or compression [10, 16] in order to represent the microstructure as accurately as possible in terms of a selected statistical descriptor. In particular, we focus on two commonly used descriptors, the two-point probability function [17, 18, 19] and the lineal path function [8, 20, 21, 22, 23]. To gain further insight into n-point statistics and the other types of spatial correlation functions, one may consult monographs [24] and [1], respectively.

The goal of this paper is to investigate the properties and differences of these two descriptors within the compression and reconstruction process in more detail. To achieve this goal, we concentrate on the calculation of the entire lineal path function instead of its often used rough discretisation with a sampling template evaluated approximately using a Monte Carlo-based procedure, see [25]. Since the evaluation of the entire lineal path function can be

computationally extremely exhaustive, we introduce certain acceleration steps for the algorithmic as well as implementation sides, where significant acceleration is achieved, namely by porting the algorithm to the graphics processing unit (GPU) using a CUDA environment.

This article is organised as follows: Section 2 describes the theoretical formulation for both descriptors. Section 3 is devoted to describing the acceleration of the lineal path function and presents the resulting improvement in speed obtained at the GPU in comparison with the sequential central processing unit (CPU) formulation. Section 4 briefly introduces the optimisation algorithm employed for microstructure compression and reconstruction, which are discussed in Sections 5 and 6, respectively. A summary of the essential findings is provided in Section 7.

2. Statistical description of random media

The morphology quantification for random heterogeneous materials starts with the introduction of the concept of an ensemble established by Kröner [26] and Beran [27]. Proposed mathematical formulations are considered as one of the milestones in statistical physics and the basic idea is that macroscopic observables can be calculated by performing averages over the systems in an ensemble. In other words, the ensemble represents a collection of geometrical systems having different microstructures but being completely identical from the macroscopic point of view [25].

A variety of statistical descriptors have been developed to describe the morphology of a multi-phase random heterogenous material [7, 19, 24, 25] based on the concept of an ensemble. Here, the two-point probability function and the lineal path function are investigated as frequently used descriptors. Therefore, this section provides their brief analytical description and classical numerical implementation.

As a preamble, throughout this paper we consider an ensemble of a two-phase medium consisting of a black and white phase labelled by superscripts $i, j \in \{\mathrm{b}, \mathrm{w}\}$. We also model the medium only as a two-dimensional system, where the position of an arbitrary point $\boldsymbol{x}_a$ is defined by the Cartesian coordinates $\boldsymbol{x}_a = (x_a, y_a)$. Nevertheless, extension into three-dimensional systems could be very straightforward.

2.1. Two-point probability function

More formally, the two-point probability function $S_2^{ij}(\boldsymbol{x}_1, \boldsymbol{x}_2)$[1] quantifies the probability of finding simultaneously the phase i and the phase j at two arbitrarily chosen points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively, and can be written in the form, see [25, 28],

$$S_2^{ij}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \langle \chi^i(\boldsymbol{x}_1, \alpha)\chi^j(\boldsymbol{x}_2, \alpha) \rangle, \qquad (1)$$

where the symbol $\langle \cdot \rangle$ denotes the ensemble average of the product of characteristic functions $\chi^i(\boldsymbol{x}_a, \alpha)$ which are equal to one when the point $\boldsymbol{x}_a$ lies in the phase i in the sample α and equal to zero otherwise:

$$\chi^i(\boldsymbol{x}_a, \alpha) = \left\{ \begin{array}{ll} 1, & \text{if } \boldsymbol{x}_a \in D^i(\alpha) \\ 0, & \text{otherwise} \end{array} \right. . \qquad (2)$$

In Eq. (2), $D^i(\alpha)$ denotes the domain occupied by the i-th phase. In general, the evaluation

---

[1]with S_2^i abbreviating S_2^{ii}

Figure 1: Illustration of the two-point probability function: (a) Example of a homogeneous system, size 500×500 [px]; (b) S_2^{w}-function; (c) S_2^{b}-function; (d) Comparison of S_2-functions in cut 1-1.

of these characteristics may prove to be prohibitively difficult. Fortunately for homogeneous systems, S_2^i depends only on the relative position of the two points $\boldsymbol{x} = \boldsymbol{x}_2 - \boldsymbol{x}_1$ and has following asymptotic properties, see [13],

$$S_2^i(|\boldsymbol{x}| = 0) = \phi^i, \tag{3}$$

$$\lim_{|\boldsymbol{x}| \to \infty} S_2^i(\boldsymbol{x}) = (\phi^i)^2, \tag{4}$$

where ϕ^i is the volume fraction of the i-th phase. Eq. (3) follows from definition (1) and means that the probability of a randomly thrown point (i.e. vector of zero length) falling into phase i is equal to the volume fraction of phase i. On the other hand, Eq. (4) assumes that the system has no long-range correlations and thus, falling of the two distant points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ into phase i are independent events, each having a probability equal to ϕ^i, see Fig. 1 as an illustrative example of such a system.

Even though we aim to characterize generally non-periodic media with SEPUC, whose boundaries are constructed as periodic, it has been demonstrated in [29] that the assumption of periodic boundaries does not introduce systematic bias into the values of statistical descriptors. On the other hand, the assumption of periodicity simplifies computation of the two-point probability function because we do not need to consider all possible orientations of the vector $\boldsymbol{x}$. As shown in Fig. 2a, four differently oriented vectors are actually connecting identical points and obviously have the same value as the two-point probability function. As a consequence, the evaluation of the two-point probability function for vectors oriented into the first quadrant includes information about all other vector orientations, see Fig. 2b.

The last note concerns a two-phase medium where the two-point probability functions of particular phases are related according to the following equation

$$S_2^i(\boldsymbol{x}) = (\phi^i)^2 - (\phi^j)^2 + S_2^j(\boldsymbol{x}), \tag{5}$$
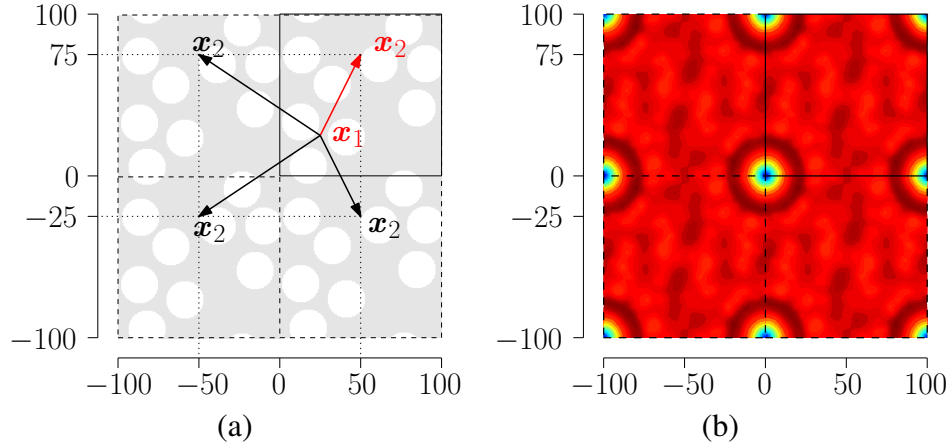
Figure 2: Vectors connecting identical points in a periodical setting (a) and the corresponding identical values of the periodic two-point probability function (b)

i.e. they differ only by a constant, also visible in Fig. 1d. Since the constant is given by known volume fractions of particular phases, only one two-point probability function needs to be determined to describe the two-phase medium. For this reason, we may drop the superscript of $S_2^i(\boldsymbol{x})$ and write the two-point probability function as $S_2(\boldsymbol{x})$.

Implementation of the two-point probability function is based on the assumption of a discrete description of the system being studied, primarily binary images in our case. The general and simple Monte Carlo-based evaluation strategy throws two points randomly into the investigated medium and counts successful "hits" of both points into phase i. This approach is, however, not only approximate but is also very computationally demanding. Therefore, another practical method was introduced on the basis of rewriting the two-point probability function as an autocorrelation of the characteristic function χ^i for a periodic medium as, see [25],

$$S_2^i(x,y) = \frac{1}{WH} \sum_{x_1=0}^{W-1} \sum_{y_1=0}^{H-1} \chi^i(x_1,y_1)\chi^i((x_1+x)\%W,(y_1+y)\%H), \tag{6}$$

where the symbol $\%$ is the modulo; $\chi^i(x_1,y_1)$ denotes the value of χ^i for the pixel located in the y_1-th row and the x_1-th column of the digitised image with dimensions $W \times H$; and where x and y are the vertical and horizontal distances between two pixels, see Fig. 3. According to [28], Eq. (6) can be computed in an efficient way using the Fast Fourier Transform. Applying this, the reformulation of the two-point probability function S_2^i for a periodic medium can be written as

$$S_2^i(x,y) = \frac{1}{WH} \text{IDFT} \left\{ \text{DFT}\left\{\chi^i(x,y)\right\} \overline{\text{DFT}\left\{\chi^i(x,y)\right\}} \right\}, \tag{7}$$

where IDFT is the inverse Discrete Fourier Transform (DFT) and where the symbol $^{-}$ stands for the complex conjugate. This method is very efficient and its accuracy depends only on the selected resolution of the digitised medium, see [25, 7]. The Fast Fourier Transform, which needs only $\mathcal{O}(WH \log(WH) + WH)$ operations, is used to perform the numerical computations presented below.
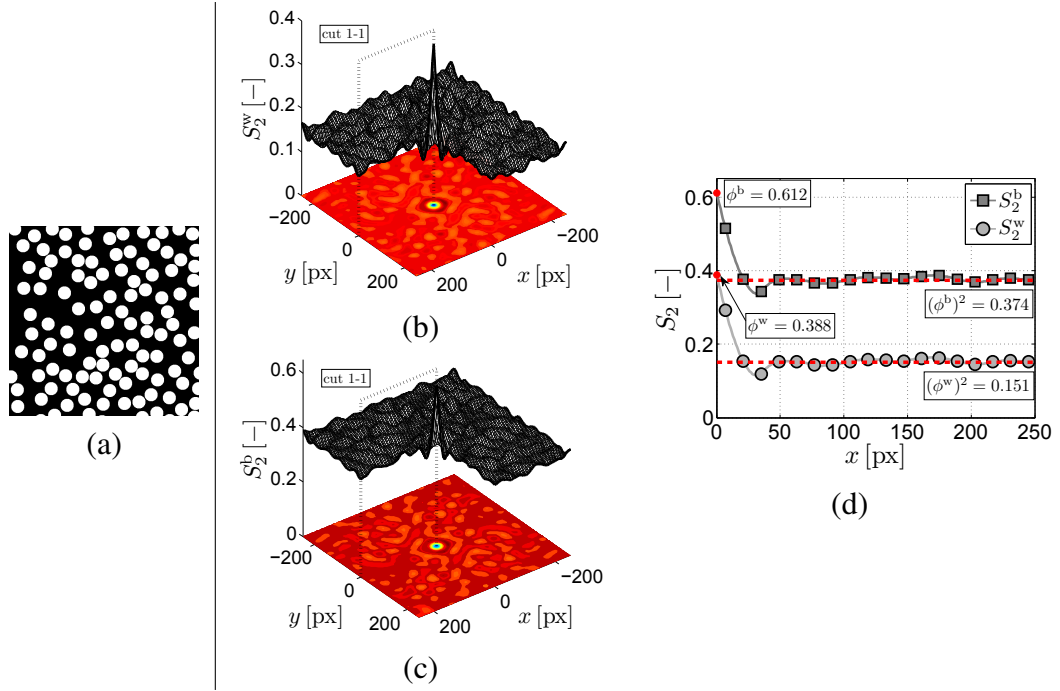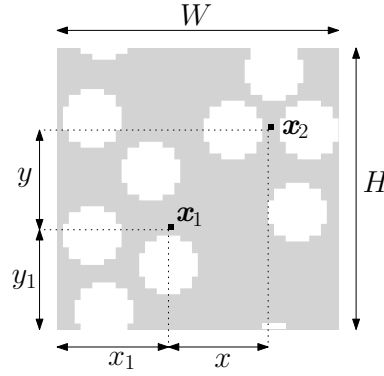
Figure 3: Illustration of a digitised image
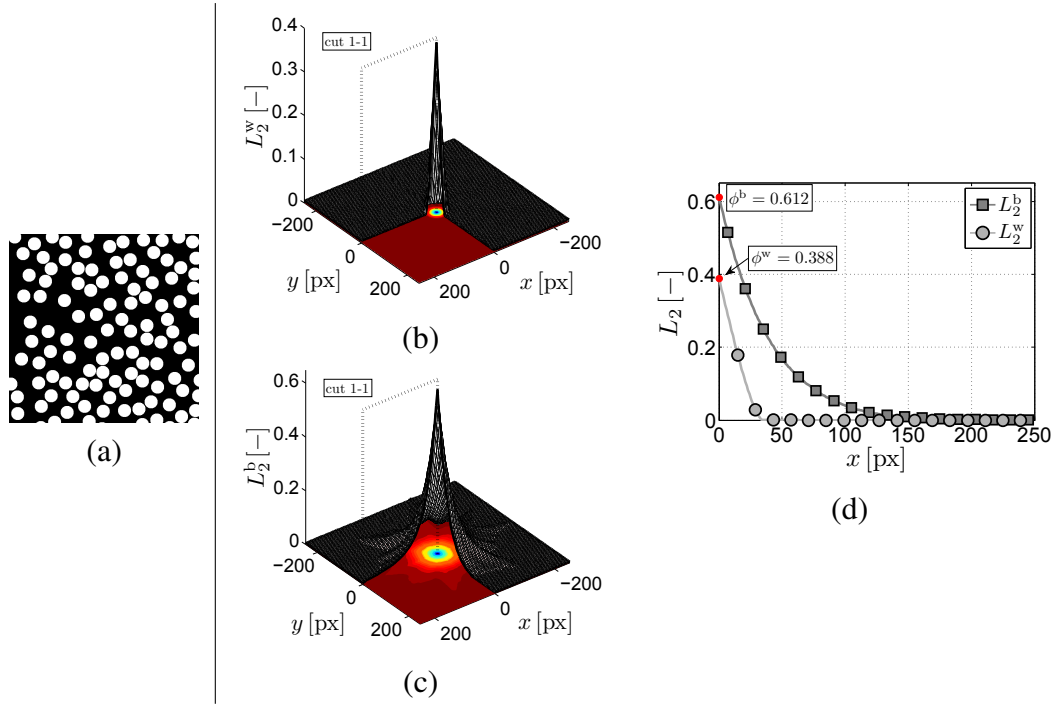
Figure 4: Illustration of the two-point probability function: (a) Example of a homogeneous system, size 500×500 [px]; (b) L_2^{w}-function; (c) L_2^{b}-function; (d) Comparison of L_2-functions in cut 1-1.

2.2. Lineal path function

Another frequently used statistical descriptor for microstructural morphology quantification is the lineal path function $L_2^i(\boldsymbol{x}_1, \boldsymbol{x}_2)$, originally introduced in [20] and further elaborated in [13, 25]. It is defined as a low-order descriptor based on the more complex fundamental function λ^i which is able to describe certain information about phase connectedness and put more emphasis on short-range correlations because its value quickly vanishes to zero with increasing $|\boldsymbol{x}|$. The fundamental function λ^i is defined as

$$\lambda^i(\boldsymbol{x}_1, \boldsymbol{x}_2, \alpha) = \begin{cases} 1, & \text{if } \boldsymbol{x}_1 \boldsymbol{x}_2 \subset D^i(\alpha), \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

i.e. a function which is equal to 1 when the segment $\boldsymbol{x}_1 \boldsymbol{x}_2$ is contained in phase i for sample α, and 0 otherwise. The lineal path function is defined as the probability that the line segment
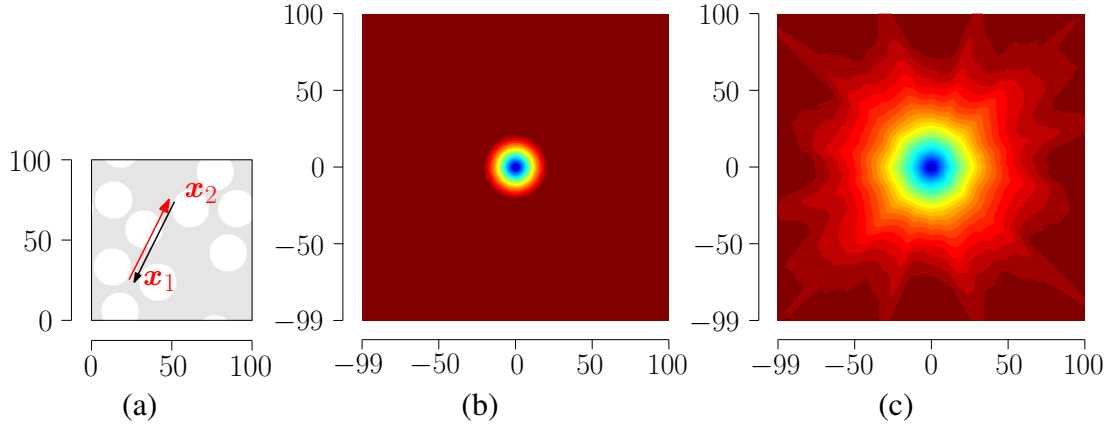
Figure 5: Vectors corresponding to identical segments (a), point symmetry of the lineal path function of black (b) and white (c) phases

$x_1 x_2$ lies entirely in phase i and can be written as the ensemble averaging fundamental function given as

$$L_2^i(\boldsymbol{x}_1, \boldsymbol{x}_2) = \langle \lambda^i(\boldsymbol{x}_1, \boldsymbol{x}_2, \alpha) \rangle. \tag{9}$$

As mentioned above, under the assumption of statistical homogeneity [25], the function can again be simplified to $L_2^i(\boldsymbol{x}_1, \boldsymbol{x}_2) = L_2^i(\boldsymbol{x})$ with $\boldsymbol{x} = \boldsymbol{x}_2 - \boldsymbol{x}_1$ and yielding

$$L_2^i(|\boldsymbol{x}| = 0) = \phi^i \tag{10}$$
$$\lim_{|\boldsymbol{x}| \to \infty} L_2^i(\boldsymbol{x}) = 0. \tag{11}$$

Here again, Eq. 11 assumes no long-range correlations and thus the probability that the line segment $x_1 x_2$ lies entirely in phase i vanishes to zero as its length increases, see Fig. 4 for an illustration of such a homogeneous system.

For the sake of consistency regarding formulation and computation of the two-point probability function, we again introduce the assumption of periodicity in our numerical implementation. However, there are no computational benefits, since all vectors in Fig. 2a connect the same points via different paths. Nevertheless, we need to keep in mind that line segment $x_1 x_2$ is identical to line segment $x_2 x_1$ and thus

$$L_2^i(\boldsymbol{x}) = L_2^i(-\boldsymbol{x}), \tag{12}$$

which means that the lineal path possesses point symmetry, see Fig. 5. Hence, we need to compute the lineal path function only for half of all possible orientations of vector x and the rest is obtained by symmetry.

In contrast to the evaluation of S_2^i, see Eq. (5), the lineal path function computed for one phase does not include all the information about the lineal path function regarding the other phase, which thus needs to be computed separately, see Fig. 5. This yields additional information about structural morphology but also has higher computational demands.

With this in mind, the standard numerical implementation of a sequential version of the entire L_2^i starts from the definition of line segments connecting two pixels x_1 and x_2 within an image with dimensions $W \times H$. The set of pixels representing a segment starting in $x_1 = (0,0)$ and ending in $x_2 = (x,y)$ is specified by an algorithm originally proposed by Bresenham [30] defining a unique solution for any positions for boundary pixels x_1, x_2. Due
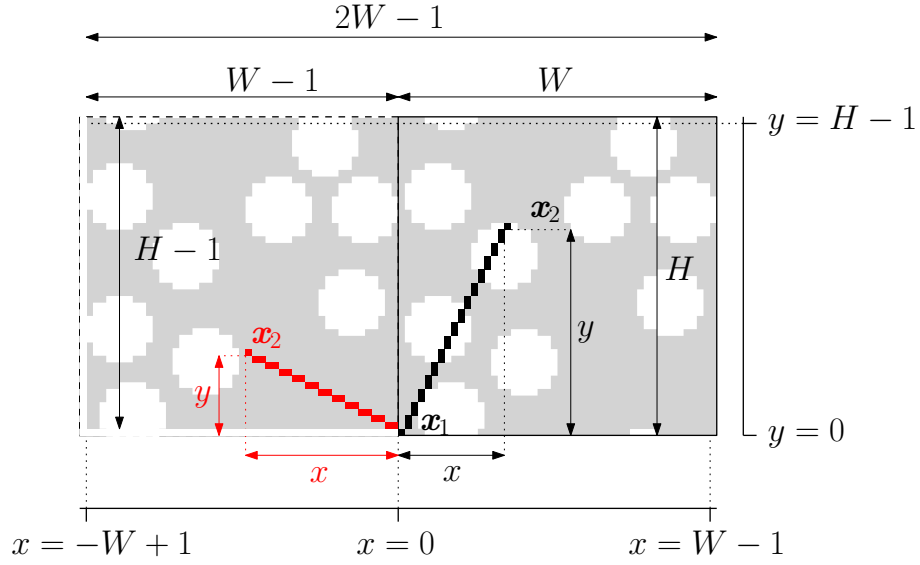
Figure 6: Illustration of line segments

to point symmetry for the lineal path function, all orientations of a line segment necessary for its computation are produced by moving point $\boldsymbol{x}_2$ within domain $\mathbb{D}$ which is given by two rectangles specifying the left bottom corner of a pixel, i.e.

$$\mathbb{D} := [-W+1; -1] \times [1; H-1] + [0; W-1] \times [0; H-1] \tag{13}$$

see Fig. 6. The number of segments defining the lineal path function is thus given as a cardinality of domain $\mathbb{D}$, which is $|\mathbb{D}| = 2HW - H - W + 1$. With defined segments, the computation of the lineal path function involves simple translations of each segment throughout the image and a comparison for examining if all pixels of the segment at a given position correspond to image pixels with the value of an investigated phase. Such an intuitive description represents, however, a computationally costly procedure leading to $\mathcal{O}(H^3W^2)$ operations for periodic media with $W \leq H$. For our purposes, this simplifies to $\mathcal{O}(W^5)$ for $W = H$, i.e. a square shape of SEPUC/SSRVE. In order to reduce the computational demands, several algorithmic and hardware acceleration steps are introduced and described in the following section.

3. Numerical implementation of L_2

In order to avoid the tremendous computational requirements for an entire L_2 evaluation, some authors (see e.g. [25]) compute only its approximation using the Monte Carlo-based procedure. In such cases, the line segments are not compared with the image at all available positions equal to a number of all pixels in the image, but only for a limited number of randomly selected positions. The level of error produced by such approximation is illustrated in Figure 7 for three different types of microstructures as a function of the number N of selected positions. The values on the vertical axis are the least square errors between the exact lineal path function L_2 and its approximation $\widetilde{L}_2$ given as

$$e(L_2^i) = \sum_{p \in \mathbb{D}} (L_2^i(\boldsymbol{x}_p) - \widetilde{L}_2^i(\boldsymbol{x}_p))^2, \tag{14}$$
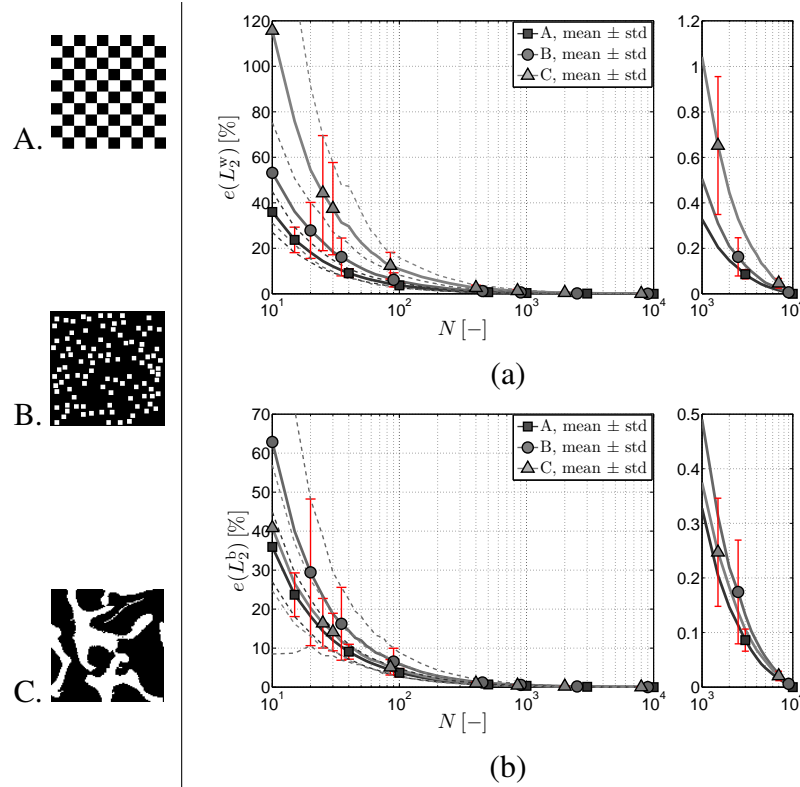
Figure 7: Convergence analysis of MC-based approximation of lineal path functions computed for (a) white phase and (b) black phase

where superscript i denotes the phase and subscript p covers all the oriented segments defining the lineal path function given by (13). Dimensions of all three microstructural images are 100×100 pixels. Since the lineal path function is evaluated at N positions obtained as a random N-combination of set $|\mathbb{D}|$ without repetition, error converges to zero for $N = |\mathbb{D}|$.

3.1. Parallelisation on GPU using a CUDA environment

The key idea described here is porting part of the code to a GPU device. Parallel computations on GPUs have become very popular within the last decade thanks to their high performance at a relatively low financial cost. Moreover, the programming environment CUDA (Compute Unified Device Architecture) simplifies GPU-based software development because it uses standard C/C++ language, see [31]. In order to clearly describe GPU parallelism, we start with the algorithmic structure of the L_2 evaluation consisting of several computational steps:

1. generating line segments for given input dimensions,
2. allocating the inputs (e.g. each input representing a binary image),
3. calculating the lineal path function based on translations of each segment and its comparison with an image.

Regarding the computational requirements for particular steps, one needs to keep in mind that L_2 is supposed to be called repeatedly within an optimisation process for new feasible solutions (i.e. new binary images) of the same dimensions $W \times H$. This means that the
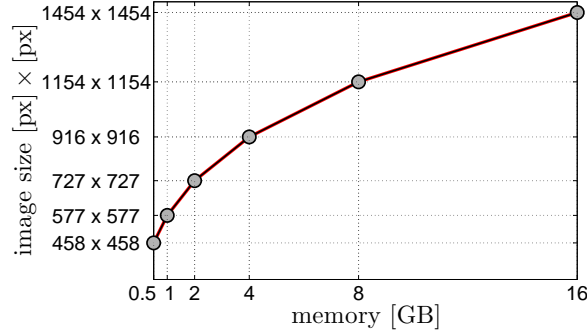
Figure 8: Illustration of memory requirements for line segment storage.

definition of line segments remains the same during the entire optimisation process, thus enabling step 1 to run only once at the beginning of the optimisation, while steps $2-3$ need to be repeatedly called. This also implies that step 1 is critical namely from the memory usage point of view, while steps $2-3$ need to be optimised with respect to computational time.

Before starting an implementation of Bresenham's algorithm for defining line segments, one needs to carefully consider line segment coding. While the definition of a particular pixel according to its (x, y) coordinates is very intuitive, it is excessively memory intensive. It is much more efficient to index all pixels in the image by only one integer value from 0 to $WH - 1$. Then the number of integer values required for definition of all line segments is given as

$$M = \sum_{i=1}^{W} \sum_{j=1}^{H} \max(i, j) + \sum_{i=2}^{W} \sum_{j=2}^{H} \max(i, j), \tag{15}$$

which leads to

$$M = \frac{W(3H^2 + 3H + W^2 - 1)}{3} - \frac{W(W + 1)}{2} - \frac{H(H + 1)}{2} + 1 \tag{16}$$

for $W \leq H$ and to

$$M = \frac{4W^3 - 4W + 3}{3} \tag{17}$$

for $H = W$. Figure 8 shows the dimensions of square images which can be handled by cards with a given memory size, assuming that one integer requires 4 bytes.

As mentioned above, the consecutive steps 2 and 3 are supposed to be called repeatedly within the optimisation process and thus represent the principal demands on computational time, see Alg. 1 for a more detail algorithmic structure for step 3. The image enters the algorithm as matrix A twice wider and twice higher than the original because it is periodically copied onto a grid 2×2 to enable simpler translation of the segments starting within the image and to ensure that segments never end outside an image. Translation is thus defined by moving the starting point of a segment within one quadrant of the entering image. To facilitate repeatedly called computations, the indices of moves within one quadrant are precomputed and stored in separate matrix C.

The structure of Alg. 1 suggests several possibilities for parallelisation. One method involves parallelisation over particular segments (line 1), which would, however, lead to

---

Data:

$A \ldots$ a binary image defined as an integer vector with size $2W \cdot 2H$

$B \ldots$ an irregular 2D integer matrix defining the pixels of Bresenham's line segments of size $S \times$ segment size, where $S = 2WH - W - H + 1$

$C \ldots$ an integer vector of size WH defining a translation within an image $W \times H$ mapped onto a periodically copied image of size $2W \times 2H$

$D \ldots$ an integer vector of size S defining the size of particular segments $phase \ldots$ an integer defining phase, for which L_2 is evaluated

Result: $L \ldots$ an integer vector of size S defining L_2

```
 1  for seg=0 to S − 1 do
 2  │   for transl=0 to WH − 1 do
 3  │   │   for pix=0 to D[seg] do
 4  │   │   │   if A[B[seg][pix]+C[transl]] ≠ phase then
 5  │   │   │   │   break;
 6  │   │   │   end
 7  │   │   end
 8  │   │   if pix = D[seg] then
 9  │   │   │   L[seg] = L[seg] + 1;
10  │   │   end
11  │   end
12  end
```

Algorithm 1: Algorithmic structure of implementation designed for a CPU device; seg, $transl$ and pix represent integer variables used to govern the corresponding items for loops.

extremely asynchronous computation due to large differences in the lengths of segments. Parallelisation over translations (line 2) is not completely synchronous because its inner cycle over pixels of the segment (line 3) is stopped when it proceeds to a pixel which does not lie in a given phase, which is dependent on the morphology of a particular image. Nevertheless, computation has at least a chance to be more synchronous than the surely highly asynchronous parallelisation over segments.

Fortunately, the algorithm clearly consists of a very large number of extremely simple logic and arithmetic operations and is thus well-suited for parallelisation on a GPU because of the following reasons:

(i) It allows for a nearly synchronous parallelisation scheme, thus respecting the basic GPU programming rule – memory coalescence;

(ii) It corresponds to the SIMD (single instruction, multiple data) architecture: a single instruction is an index of a segment to be compared with an image at all possible positions, which thus represents the multiple data;

(iii) Most of the memory transfer corresponding to copying of images, translations, line segments and their sizes is done only once and in large chunks, thus reducing related system overhead.
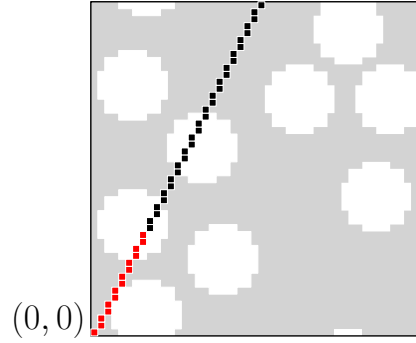
Figure 9: Illustration of overlapping line segments

The parallel algorithmic structure proposed to increase numerical efficiency of L_2 computation is given in Alg. 2. The crucial step for implementation efficiency involves line 9, where all $n_t = WH$ translations are distributed into available multiprocessors (MP). Since particular GPU architectures differ slightly from each other, we concentrate here on a Fermi compute architecture [31], where each MP has 32 single-precision CUDA cores. This means that each MP can simultaneously solve up to 32 tasks – so-called threads – defining one warp. Besides currently computing threads, the MP can already load and prepare other threads up to maximally 1536 threads = 48 warps. The tasks are sent to the MP in blocks, where each particular translation is assigned to a particular thread automatically according to its position within a block. Storing the translations in a 1D vector instead of a 2D matrix thus allows for more even distribution of translations among the MPs. Each MP can handle maximally 8 blocks at the same time. The size of a block can be chosen by a programmer, but finding the optimum size is not a straightforward process. To maximise the occupancy of MPs, it is convenient to define the size of block B as

$$
B = \begin{cases} 1 \times 6 \, \text{warps} = 1 \times 192 \, \text{threads}, & \text{if } \left\lceil \frac{n_t}{n_{mp}} \right\rceil \geq 1536, \\ 1 \times \left\lceil \frac{n_t}{8 n_{mp}} \right\rceil \text{threads}, & \text{otherwise}, \end{cases} \tag{18}
$$

where $\lceil \cdot \rceil$ denotes a round-up operation to the nearest integer and n_{mp} is the number of available MPs. Nevertheless, other aspects related to shared memory and registers [31] may shift the preferences towards bigger blocks. A more detailed study of optimal block size is beyond the scope of this paper. In our computations, we focuse on occupancy maximisation only and the block size was set according to Eq. (18).

3.2. Algorithmic acceleration of L_2 evaluation

Besides parallelisation, we also propose one simple algorithmic acceleration of lineal path computation. The idea stems from the discrete nature of segments and the fact that some shorter segments overlap with some longer segments. See Fig. 9, where all segments start at $x_1 = (0,0)$ and those ending in red pixels are overlapped by segments ending in black pixels. If a segment never falls entirely in a given phase and its L_2 value is zero, it is obvious that all longer overlapping segments will have a value of zero as well. This simple logic brings additional significant time savings in L_2 evaluation. One only needs to precalculate a vector containing the indices of the longest shorter overlapping segments (LSOS). Such a precalculation is computationally expensive but is done only once at the

```
1  CPU:          calculating line segments: indices in B and sizes in D;
2  CPU→GPU: copying B and D into GPU;
3  CPU:          loading and copying binary image onto grid 2 × 2 saved
4                    into A, defining translations C ;
5  CPU→GPU: copying C and phase into GPU;
6  CPU→GPU: copying A into GPU;
7  for seg = 0 to S − 1 do
8      CPU→GPU: copying seg into GPU;
9      GPU calls threads: for transl = 0 to WH − 1 do
10         L[seg] = 0;
11         for pix=0 to D[seg] do
12             if A[B[seg][pix]+C[transl]] ≠ phase then
13                 break;
14             end
15         end
16         if pix = D[seg] then
17             L[seg] = L[seg] + 1;
18         end
19     end
20 end
21 CPU←GPU: copying L to CPU;
```

Algorithm 2: Simplified algorithmic structure of implementation designed for a single GPU device; All variables are defined in Alg. 1.

beginning of the algorithm. Keeping in mind that steps 6 to 21 in Alg. 2 are supposed to be called repeatedly within an optimisation process, this precalculation should take place before step 3. Then one simple if-condition is added before translating and comparing the segments with the image. If the LSOS corresponding to the current segment has a zero value for L_2, then the L_2 value of the current segment is also automatically given a zero value and the translating and comparing phase is skipped. To be more specific, there are two possibilities for solving this crucial if-condition. An intuitive solution involves solving the if-condition on the CPU, i.e. before line 8 of the Alg. 2, so as to skip the calling of the GPU entirely. However, in such a case, the CPU needs to have knowledge about previously computed segments, which means that the value of the lineal path function has to be sent to the CPU for every segment separately inside the for-loop before line 18. Our computations, however, revealed that repeated sending of one integer from GPU to CPU is time-consuming and it is faster to repeatedly call the GPU, solve the if-condition there (i.e. before line 10) and store all computed values of the lineal path function on the GPU only until the last segment is computed. Then sending the whole vector of the lineal path values saves significant time. We implemented this latter variant and refer to it as *enhanced*, while the original version of the algorithm without any algorithmic acceleration is called *standard*.

The performance of GPU parallelism is demonstrated by evaluating the L_2-function on three different microstructures: (i) a chess-type morphology with dimensions of squares 10×10 [px], (ii) a particulate suspension consisting of equal-sized squares with dimensions 4×4 [px] and (iii) a metal foam taken from [32]. Tab. 1 compares the amount of time necessary
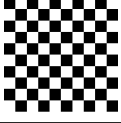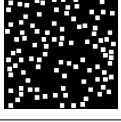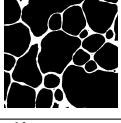
	Standard			Enhanced			
dim.	GPU	CPU	S	GPU	CPU	S	S^*
[px]	[s]	[s]	[−]	[s]	[s]	[−]	[−]
10×10	$2.4 \cdot 10^{-3}$	$0.59 \cdot 10^{-3}$	**0.2×**	$2.4 \cdot 10^{-3}$	$0.39 \cdot 10^{-3}$	**0.2×**	**0.2×**
20×20	$11.9 \cdot 10^{-3}$	$11.9 \cdot 10^{-3}$	**1.0×**	$11.1 \cdot 10^{-3}$	$7.8 \cdot 10^{-3}$	**0.7×**	**1.1×**
50×50	0.15	0.55	**3.7×**	0.10	0.24	**2.5×**	**5.8×**
100×100	2.1	7.0	**3.3×**	0.26	0.27	**1.0×**	**27.0×**
200×200	32.6	110.9	**3.4×**	2.0	1.0	**0.5×**	**53.9×**
500×500	318.4	1071.1	**3.4×**	16.3	6.4	**0.4×**	**65.8×**

	Standard			Enhanced			
dim.	GPU	CPU	S	GPU	CPU	S	S^*
[px]	[s]	[s]	[−]	[s]	[s]	[−]	[−]
10×10	$2.5 \cdot 10^{-3}$	$0.71 \cdot 10^{-3}$	**0.3×**	$2.5 \cdot 10^{-3}$	$0.38 \cdot 10^{-3}$	**0.2×**	**0.3×**
20×20	$12.9 \cdot 10^{-3}$	$16.0 \cdot 10^{-3}$	**1.2×**	$12.5 \cdot 10^{-3}$	$8.2 \cdot 10^{-3}$	**0.7×**	**1.3×**
50×50	0.16	0.81	**5.0×**	0.13	0.36	**2.7×**	**6.0×**
100×100	2.5	16.1	**6.4×**	1.7	7.2	**4.1×**	**9.3×**
200×200	41.9	256.3	**6.1×**	20.8	78.0	**3.7×**	**12.3×**
500×500	411.6	2544.0	**6.2×**	205.1	789.1	**3.9×**	**12.4×**

	Standard			Enhanced			
dim.	GPU	CPU	S	GPU	CPU	S	S^*
[px]	[s]	[s]	[−]	[s]	[s]	[−]	[−]
10×10	$2.4 \cdot 10^{-3}$	$0.73 \cdot 10^{-3}$	**0.3×**	$2.4 \cdot 10^{-3}$	$0.65 \cdot 10^{-3}$	**0.3×**	**0.3×**
20×20	$12.0 \cdot 10^{-3}$	$13.8 \cdot 10^{-3}$	**1.1×**	$11.9 \cdot 10^{-3}$	$9.7 \cdot 10^{-3}$	**0.8×**	**1.2×**
50×50	0.17	1.31	**7.6×**	0.15	0.67	**4.4×**	**8.7×**
100×100	2.8	31.9	**11.5×**	2.22	17.2	**7.7×**	**14.3×**
200×200	48.3	577.9	**12.0×**	32.9	241.7	**7.3×**	**17.5×**
500×500	542.1	7911.7	**14.6×**	445.2.1	3884.2	**8.7×**	**17.7×**

Table 1: Comparison of CPU and GPU performance averaged over five evaluations (S stands for speedup and S^* represents overall speedup obtained by hardware and software accelaration)

averaged over five evaluations of the lineal path function for both phases on single CPU or GPU devices depending on image size and chosen variant of the algorithm. Computational times correspond to a part of the lineal path function computation, which is called repeatedly within the optimisation process, i.e. evaluation of lines 1 to 5 in Alg. 2 is excluded. We show that for very small images, the use of CPU outperforms use of the GPU because of the additional time spent by communicating with the GPU. Nevertheless, for images of 50×50 [px], the GPU achieves an evident acceleration which mostly increases as the dimensions of the image increase. An exception is the chess-type microstructure, where a specific phase distribution limits the length of most of the line segments to 10 [px]. This significantly

elevates the gain in speed obtained for the enhanced variant of the algorithm and even the CPU version is so fast that communication with the GPU leads again to deceleration, which increases according to the image dimensions.

The computations presented in Tab. 1 were performed on a $2\times$ INTEL Xeon $E5-2620$ @ 2.0 GHz, 96 GB RAM, $2\times$ GPU - NVIDIA QUADRO 4000 with a Micrsosoft Windows 7 64-bit operating system and CUDA v. 6.5. Furthermore, the algorithm is also designed for dual GPUs. However, scalability towards multiple GPU devices is not considered here. The logical step for implementing a dual GPU algorithm would be to uniformly distribute the generated segments so that each device holds only a certain amount of them. This improvement would thus result in lower memory requirements.

4. Optimisation procedure

Before proceeding to a comparative study of the lineal path and two-point probability function, we briefly describe the optimisation procedures employed in our computations. Here, we used the framework first introduced by Yeong and Torquato [13] for digitised media and further elaborated, for instance, in [33, 34, 19, 23]. The algorithm is based on a simulated annealing method independently developed by Kirkpatrick et al. [35] and Černý [36]. It starts with a randomly generated microstructure and quantification of its quality according to a chosen statistical descriptor. The microstructure is then modified with a chosen operator and its new quality is evaluated. The acceptance of the proposed modification is governed by the Metropolis rule, which allows, with certain probability, acceptance of a worse solution and thus avoids a local extreme. Such a generic optimisation scheme opens up the possibility for defining the modification operator suitable for a given microstructure. For instance, a particulate suspension consisting of equal-sized discs can be modified by moving the centre of an arbitrarily chosen disc, see e.g. [10, 11]. Such a move affects the whole set of pixels and allows preservation of the known shape of particles, thus accelerating the optimisation procedure. Most of the microstructures, however, do not consist of particles having a specific known shape. In such cases, the simplest modification operator is based on interchanging two randomly chosen pixels from different phases, which at least allows for preservation of their volume fractions [13]. The very simple acceleration employed in our implementation consists of a random selection of interfacial pixels which leads to a significant increase in accepted modifications, as presented in [37].

Since the proposed way of porting lineal path evaluation onto the GPU counts with copying the whole image from CPU to GPU for any new proposed modification, the modification operator can be designed in any convenient way. Nevertheless, our further computations only use the interchange of two pixels. The particular structure of the employed optimisation algorithm is given in Alg. 3. First of all, a random digitised image P is created with the same volume fractions of phases as the original morphology. Its statistical similarity to the original image is then evaluated using the chosen statistical descriptor SD as the least square error:

$$e(\text{SD}^i) = \sum_{p \in \mathbb{D}} (\text{SD}^i_{\text{original}}(\boldsymbol{x}_p) - \text{SD}^i(\boldsymbol{x}_p))^2, \tag{19}$$

where superscript i denotes the phase for which the SD is evaluated and subscript p corresponds to the component of a discretised descriptor. If superscript i is missing in the following text, SD is evaluated for both phases. Note that the least square error in Eq. (19)

Data: binary image with dimensions $W \times H$
Result: optimised SEPUC corresponding to given image
1   $create\_random\_image(P)$;
2   $SDP = evaluate(P)$;
3   $T = T_{\max}$;
4   $T_{\mathrm{mult}} = (T_{\min}/T_{\max})^{(succ_{\max}/N_{\mathrm{step}})}$;
5   **while** $c < N_{\mathrm{step}}$ **do**
6      $c = s = 0$;
7   **while** $c < c_{\max}$ & $s < s_{\max}$ **do**
8         $c = c + 1$;
9         $Q = modify(P)$;
10        $SDQ = evaluate(Q)$;
11   **if** *random_number U[0,1]* $< \exp((SDQ - SDP)/T)$ **then**
12            $s = s + 1$;
13            $P = Q$;
14            $SDP = SDQ$;
15   **end**
16   **end**
17      $T = T \cdot T_{\mathrm{mult}}$;
18 **end**

Algorithm 3: Algorithmic structure of simulated annealing

also consists of a large number of simple arithmetic operations, which are again efficiently evaluated in parallel on the GPU.

Control parameters of the algorithm were set to following values: The value of $T_{\max}$

$$N_{\mathrm{step}} = 4 \cdot 10^6 \qquad c_{\max} = 0.1 N_{\mathrm{step}}$$
$$T_{\min} = 0.01 T_{\max} \quad s_{\max} = 0.01 N_{\mathrm{step}}$$

Table 2: Control parameters of simulated annealing method

was manually changed for every particular computation so as to achieve approximately the ratio $s/c = 0.5$ within the first few steps of the algorithm. Some other recommendations for setting these parameters can be found e.g. in [38].

5. Microstructure reconstruction

Reconstruction of a microstructure from its statistical description is an inverse problem addressed by several authors in different ways, see [37, 39, 40] and the references therein. Here we follow the concept proposed in [13], where the discretised randomly generated microstructure is optimised in relation to the prescribed statistical descriptor. The authors in [41] presented numerical evidence that a periodic medium discretised into pixels is completely specified by its two-point correlation function, up to a translation and, in some cases,

---

[3] Similarly as in the introduction part, superscripts $\cdot^{\mathrm{b}}$ and $\cdot^{\mathrm{w}}$ represent again black and white phases, respectively.

inversion. This conclusion implies that the reconstruction process based on the discretised two-point probability function has a unique solution. For many microstructural morphologies, the same holds also for reconstruction from the lineal path function. For instance, the chess-type morphology is fully defined not only by the lineal path function computed for both phases, but only one of the phases is fully sufficient to completely define the morphology, see Figure 10. Nevertheless, generic evidence for a unique solution of the lineal path function-based reconstruction is missing and is suggested only by findings concerning orientation-dependent chord length distributions in continuous domains [42]. On the contrary, we can demonstrate that when employing Bresenham's algorithm for the definition of the line segments, the lineal path function does not define a unique solution for a reconstruction process based on a discretised medium.

Fig. 11 shows an example of two different periodic cells with dimensions 2×8 of pixels. Due to the same volume fraction of both phases, the two-point probability functions obtained for both phases in Fig. 11c-d are identical, which is in agreement with Eq. 5. However, both functions differ from the corresponding ones obtained for the other cell in Fig. 11i-j. The lineal path functions are, on the other hand, identical for both phases in Fig. 11e-f as well as for both phases obtained for the second cell in Fig. 11k-l. This proves a non-unique solution for the reconstruction process of the chosen highly rough discretisation. Of course, for a higher resolution, the difference between the lineal path functions obtained for both cells can be revealed again. As a conclusion, a reconstruction process based on a discretised medium always has a unique solution for a two-point probability function and mostly for a lineal path function where the differences among a potential set of solutions decrease with increasing resolution.

Based on this conclusion, we can proceed to comparing the entire lineal path function L_2 with its Monte Carlo-based approximation $\tilde{L}_2$ within the reconstruction process. Considering the microstructures depicted in Fig. 7, we may assume that their reconstruction based on the entire lineal path function will lead to almost the same microstructures as the original ones in the case of the microstructures B and C, and to exactly the same one in the case of A. To decrease the computational demands of the comparison, we reduced the dimensions of the microstructures to 50×50 pixels. The reconstruction process is driven again as the minimisation of the least square error given in Eq. 14. The Monte Carlo-based approximation
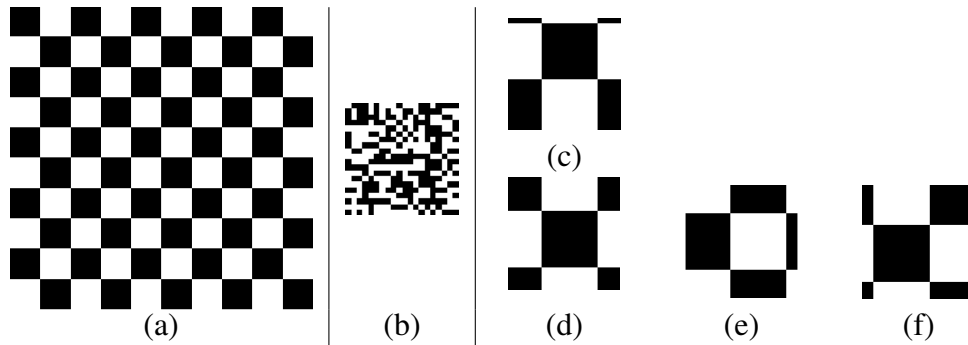
Figure 10: Chess microstructure: (a) Original medium with size 100×100 [px] and characteristic lengths 20×20 [px]; (b) Random initial structure, size 20×20 [px]; (c) S_2-based reconstructed image; (d) L_2-based reconstructed image; (e) L_2^{b}-based reconstructed image; (f) L_2^{w}-based reconstructed image; obtained within less than $5 \cdot 10^5$ iterations[3].

described in Sec. 3 is applied here to evaluate both the lineal path functions of the original as well as of the reconstructed image, respectively. In order to investigate the influence of the approximation quality, we have considered three levels corresponding to a Monte Carlo evaluation based on $N = 10, 100$ and 1000 samples. The results are compared with the reconstruction based on the entire lineal path function, where the number of samples is identical with the number of pixels in the images, i.e. $N = 2500$.

The relative errors of the final reconstructed L_2-based images related to the entire lineal path function L_2 of original images are displayed in Fig. 12. The displayed values reveal that the reconstruction procedure based on lineal path approximation converges very slowly with the number of evaluated samples and the reconstruction process thus leads to images according to the lineal path function which are very different from the original image.
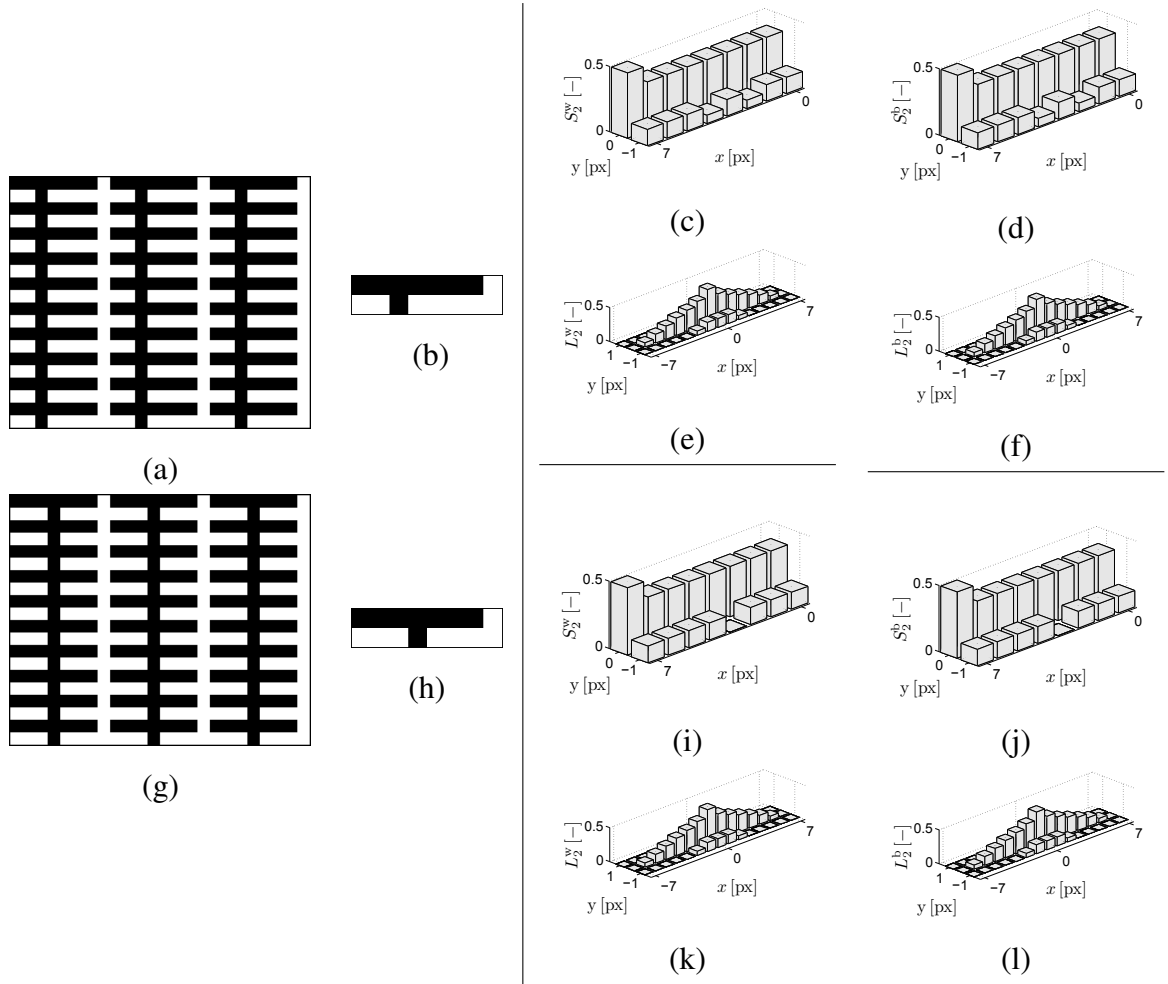
Figure 11: (a) Original medium, size 20×24 [px]; (b) Periodic unit cell (PUC), size 2×8 [px]; (c) S_2^{w}-function of PUC; (d) S_2^{b}-function of PUC; (e) L_2^{w}-function of PUC; (f) L_2^{b}-function of PUC; (g) Original medium, size 20×24 [px]; (h) Periodic unit cell, size 2×8 [px]; (i) S_2^{w}-function of PUC; (j) S_2^{b}-function of PUC; (k) L_2^{w}-function of PUC; (l) L_2^{b}-function of PUC;
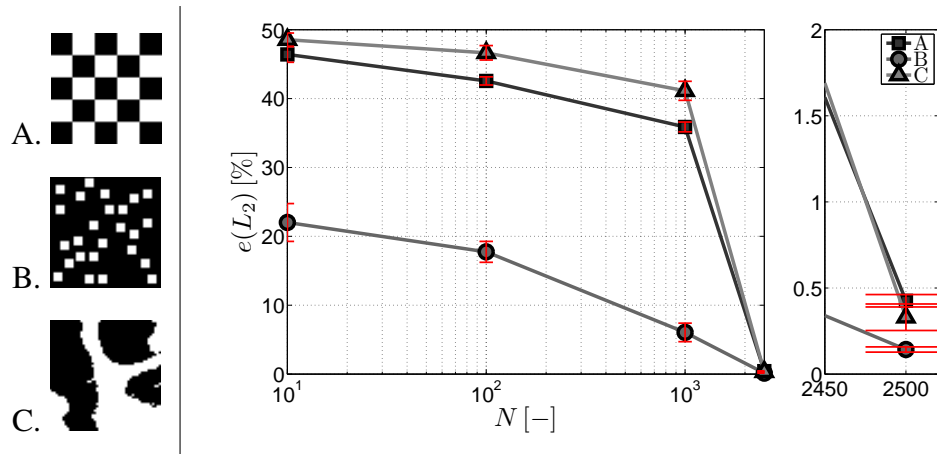
Figure 12: Comparison of the entire L_2 and its approximation $\tilde{L}_2$ used in microstructure reconstruction for different values of N

6. Microstructure compression

While the reconstruction process aims at rediscovering a microstructure with dimensions and spatial statistics defined by a given descriptor, the compression process tries to reduce the information content of the given morphology and searches for its compressed representation using a small statistically similar periodic cell [7, 43] or a set of compatible cells [10]. After evaluating a chosen statistical descriptor over the whole available domain of the original medium, one needs to determine the cells' dimensions and accordingly cut the dimensions of the descriptor. Then the compression process proceeds in exactly the same manner as the microstructure reconstruction.

As our numerical implementations of the two-point probability function and the lineal path function are both based on the assumption of periodicity, they will not provide precise results when applied to original random and non-periodic microstructures. Nevertheless, as already mentioned, in [29], we have shown that the assumption of periodicity does not introduce systematic bias in the values of the descriptors.

6.1. Particulate suspension

The first example comparing properties of S_2- and L_2-based compression concerns an artificially created particulate suspension consisting of equal-sized white squares randomly distributed within a black matrix, see Fig. 13. The shape of particles is a very significant property in such microstructures that can easily be preserved by modifying the optimisation algorithm so as to start with randomly distributed particles and then to move their centres during the optimisation process. Nevertheless, here we tested the descriptors according to their ability for capturing such an important property within the compression process.

Fig. 13c shows that S_2-based compression leads to significant deterioration of the shape of particles. This is caused by the very small ratio of particles over the size of the PUC, i.e. 4×4 [px] vs. 50×50 [px]. The information regarding shape is thus saved on a small portion of the descriptor's domain which corresponds to short-range correlations (see Fig. 14a), while most of the domain defines long-range correlations corresponding to the mutual distances between the particles. The lineal path function allows us to separate the
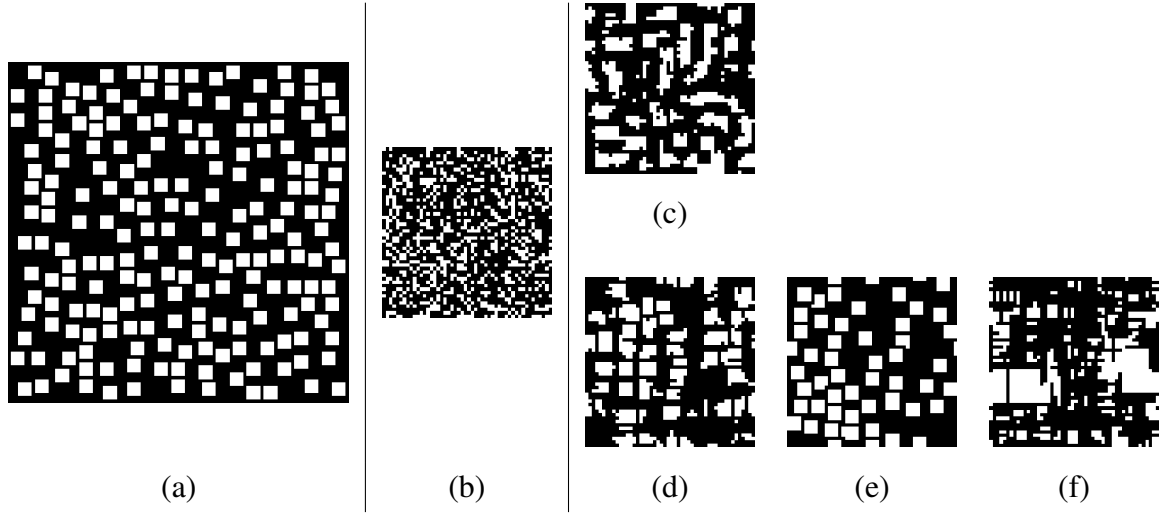
Figure 13: Particulate suspension: (a) Original medium, size 100×100 [px], particles 4×4 [px]; (b) Random initial structure, size 50×50 [px]; (c) Compressed S_2-based image, size 50×50 [px]; (d) Compressed L_2^{b} and L_2^{w}-based image, size 50×50 [px]; (e) Compressed L_2^{w}-based image, size 50×50 [px]; (f) Compressed L_2^{b}-based image, size 50×50 [px]
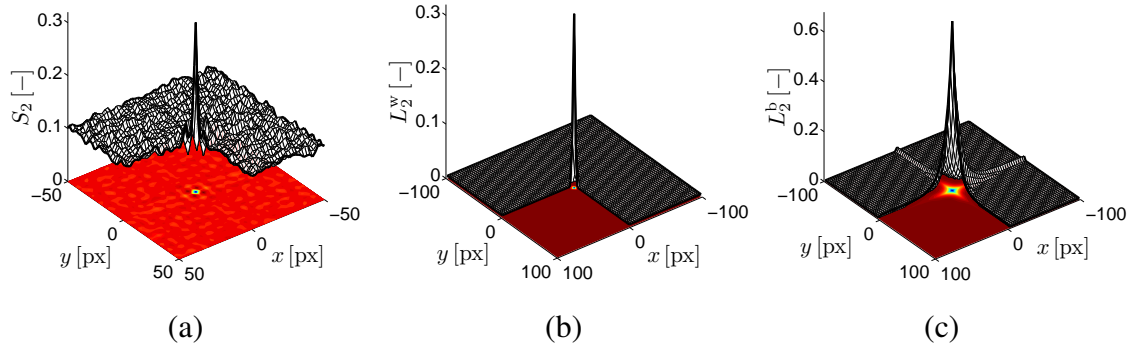
Figure 14: Particulate suspension: (a) Original medium, size 100×100 [px], particles 4×4 [px]; (b) S_2-function; (c) L_2^{w}-function; (d) L_2^{b}-function

information about the shape of particles from their mutual positions, since L_2^{w} in Fig. 14b only contains information about shape, while L_2^{b} in Fig. 14c defines position. Accordingly, L_2^{w}-based compression as in Fig. 13e obviously leads to well compressed particle shapes, while L_2^{b}-based compression as in Fig. 13f does not capture the shape of particles at all and L_2-based compression as in Fig. 13d provides a compromise solution. It is hard to evaluate the quality of obtained structures in an objective manner, but we can conclude that the L_2 function allows a user to emphasise short-range effects as needed.

Another interesting aspect concerns mutual comparison of the compressed microstructures and the corresponding errors in describing original media according to Eq. (14), listed in Tab. 3. While the optimisation of S_2 leads to a microstructure which is relatively good also in relation to L_2 and comparable with microstructures obtained for L_2^{w}- or L_2^{w}-based optimi-

sation, the opposite is not true. The microstructures optimised w.r.t. one or both phases of L_2 manifest very poor correlations, which are comparable or even worse than those obtained for a random image. The lineal path function cannot therefore be applied to correlations-based compression.

Compressed medium	$e(S_2)$	$e(L_2)$
Random Fig. 13b	$3.12 \cdot 10^{-1}$	$1.68 \cdot 10^{1}$
S_2-based Fig. 13c	$9.20 \cdot 10^{-3}$	$2.35 \cdot 10^{0}$
L_2-based Fig. 13d	$3.09 \cdot 10^{-1}$	$2.73 \cdot 10^{-2}$
L_2^{w}-based Fig. 13e	$3.53 \cdot 10^{-1}$	$1.08 \cdot 10^{0}$
L_2^{b}-based Fig. 13f	$6.52 \cdot 10^{-1}$	$9.60 \cdot 10^{-1}$

Table 3: Mutual comparison of compressed microstructures.

6.2. Epithelial cells

Epithalial cells provide a typical example of morphology characterised by very thin and continuous walls, see Fig. 15a. Their volume fraction is very small, only $4.97\,[\%]$ and thickness is mostly equal to only 1 pixel. Assembling continuous walls from a random initial arrangement is rather unattainable. As can be expected, the two-point probability function fails completely in this task, see Fig. 15c. Nevertheless, the assumption that the continuity of the white walls can be captured by a lineal path computed for the white phase is incorrect. As a matter of fact, the nonlinear walls are composed of a set of short line segments
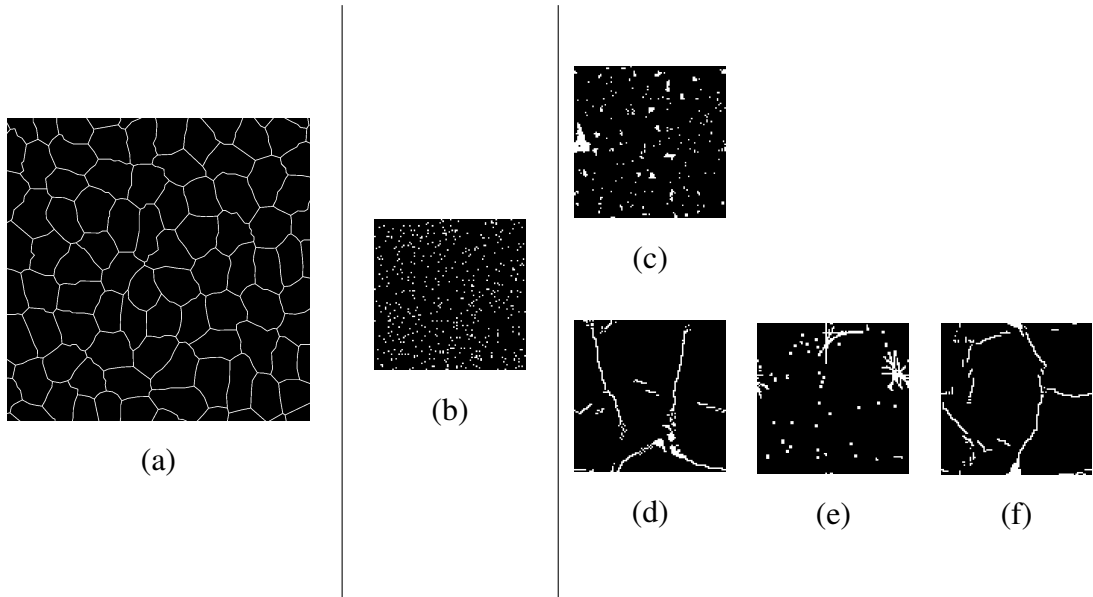
Figure 15: Epithelial cells: (a) Original medium, size 510×510 [px]; (b) Random initial structure, size 100×100 [px]; (c) Compressed S_2-based image, size 100×100 [px]; (d) Compressed L_2^{b} and L_2^{w}-based image, size 100×100 [px]; (e) Compressed L_2^{w}-based image, size 100×100 [px]; (f) Compressed L_2^{b}-based image, size 100×100 [px]

Figure 16: Epithalial cells: (a) Original medium, size 100×100 [px]; (b) S_2-function; (c) L_2^{w}-function; (d) L_2^{b}-function

and L_2^{w}-based compression thus leads to a random arrangement resembling stars as visible in Fig. 15e. The continuity of walls is actually closely related to cells, whose limited size requires a continuity in the surrounding medium. As a consequence, the information about the continuity of walls is surprisingly hidden in L_2^{b}; see results of L_2^{b}-based compression in Fig. 15f. Due the small volume fraction of the white phase, its influence on L_2-based compression is rather small and the results are principally similar to those from L_2^{b}-based compression, cf. Figs. 15d and 15f. The remaining discontinuities are very difficult to improve with the proposed optimisation strategy based on random interchanging of two pixels. We can only assume that full connectivity of walls can be obtained using a more sophisticated modification operator.
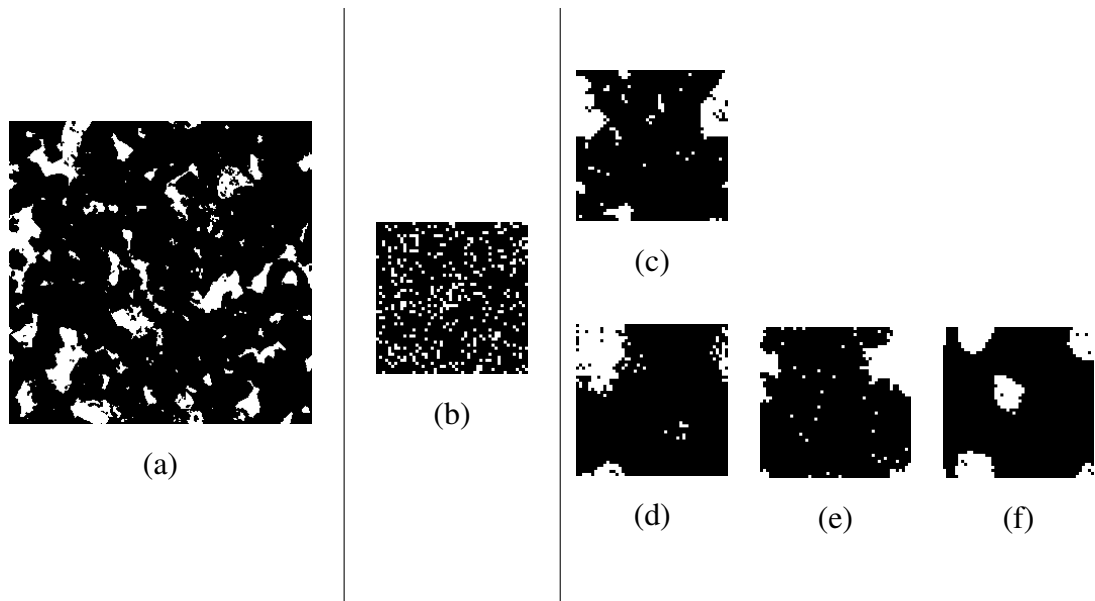
Figure 17: Sandstone: (a) Original medium, size 300×300 [px]; (b) Random initial structure, size 50×50 [px]; (c) Compressed S_2-based image, size 50×50 [px]; (d) Compressed L_2^{b} and L_2^{w}-based image, size 50×50 [px]; (e) Compressed L_2^{w}-based image, size 50×50 [px]; (f) Compressed L_2^{b}-based image, size 50×50 [px]

6.3. Sandstone

The next example of microstructure compression is devoted to the morphology of sandstone, which is characterized by irregular shapes of its porous system, corresponding here to the white phase. A slice of digitised sandstone 300×300 [px] is depicted in Fig. 17a.

As in previous examples, it is worthwhile to mutually compare again the final compressed structures. The corresponding errors are listed in Tab. 4 and it can be seen that computed results differ from those calculated for the regular shape of inclusions presented in Sec. 6.1. L_2-based compression also yields relatively good results for the two-point probability function, but not vice versa. This confirms that the level of connectedness information starts to play a role and it is not captured by S_2-based optimisation.

Compressed medium	$e(S_2)$	$e(L_2)$
Random Fig. 17b	$1.14 \cdot 10^0$	$1.48 \cdot 10^3$
S_2-based Fig. 17c	$3.97 \cdot 10^{-1}$	$2.22 \cdot 10^2$
L_2-based Fig. 17d	$6.91 \cdot 10^{-1}$	$9.45 \cdot 10^{-1}$
L_2^{w}-based Fig. 17e	$5.61 \cdot 10^{-1}$	$1.24 \cdot 10^2$
L_2^{b}-based Fig. 17f	$7.63 \cdot 10^{-1}$	$1.77 \cdot 10^0$

Table 4: Mutual comparison of compressed microstructures.

6.4. Trabecular bone

The last example concerns trabecular bone, which represents a medium with approximately equal volume fractions of both phases creating continuous irregular branches. Our original microstructural specimen was a $100 \times 100 \times 100$ [px] three-dimensional image obtained with micro Computed Tomography [44]. We divide this data into an ensemble of 100 two-dimensional cuts 100×100 [px] and, by employing the assumption of ergodicity, the statistical descriptors are computed as an average over the ensemble. The computational effort for the L_2-based reconstruction is enormous. Although part of the L_2 calculation was ported to a GPU, the entire compression process for image 100×100 [px] lasted days, recall Tab. 1 for the time requirements for single L_2 evaluation. According to the relations in Sec. 2, the overall number of operations for a single standard calculation[4] of the L_2^i-function is $6.716 \cdot 10^9$ comparing to $1.021 \cdot 10^5$ operations of S_2 evaluation.

The final compressed S_2 and L_2-based structures are shown in Figs. 18c and d, respectively. The S_2- and L_2-functions of original and new microstructures are then summarised in Fig. 19. It is clearly visible that the optimised functions coincide very well with the prescribed ones, see Fig. 19b,d,f. However, the same does not hold for a microstructure optimised w.r.t. one descriptor but then evaluated w.r.t. to another. The L_2-based compressed microstructure manifests much stronger short-range correlations, see Fig. 19b, while the S_2-based compressed microstructure underestimates the connectivity and, especially in the black phase, consists of a smaller number of continuous line segments, see Fig. 19f.

---

[4]The number of operations related to an enhanced version of L_2^i evaluation cannot be determined exactly because of missing knowledge about the zero segments. However, there are approximately 97 percent fewer operations for dimensions 100×100 [px], i.e. approximately $2.0 \cdot 10^8$ operations.

(a)

(b)   (c)   (d)
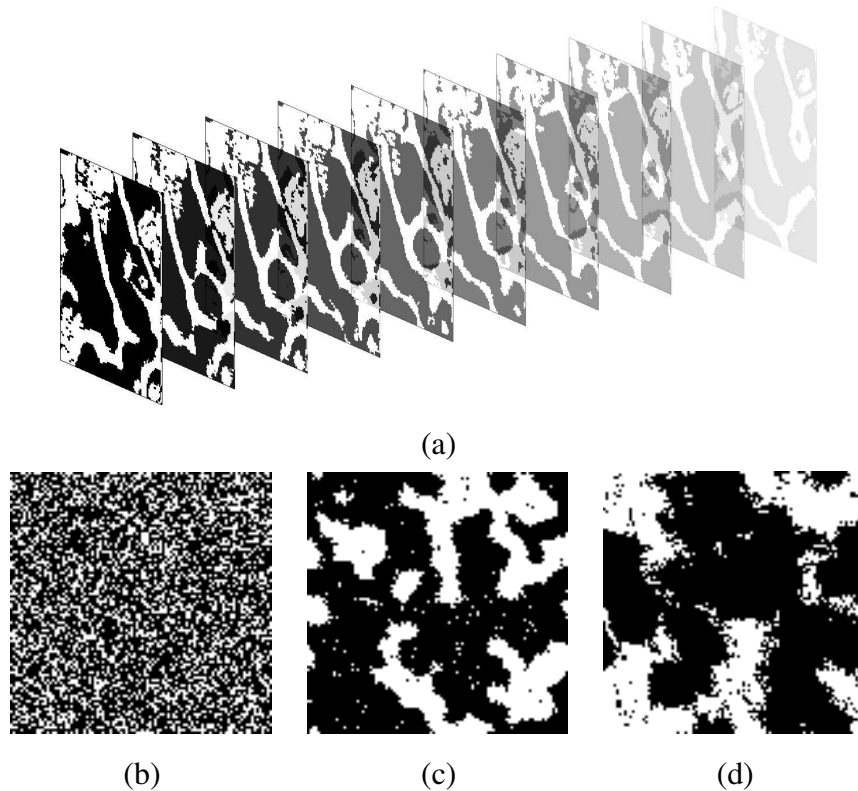
Figure 18: Trabecular bone microstructure obtained by micro Computed Tomography [44]: (a) 3-D cuts of original structure, 100×100 [px]; (b) Initial random morphology corresponding to volume fraction ϕ^{b} and ϕ^{w} of original medium, 100×100 [px]; (c) Compressed S_2-based structure, 100×100 [px]; (d) Compressed L_2-based structure, 100×100 [px]

Nevertheless, we can conclude that both descriptors provide visually well compressed microstructures and even the two-point probability function enables continuous regions similar to the original medium to be obtained.

7. Conclusions

This paper is devoted to comparing the lineal path function and the two-point probability function in reconstruction and compression of two-phase microstructures. To investigate the properties of the descriptors in a sufficient detail and to avoid some misleading conclusions based on rough discretisation of the lineal path by an approximately evaluated sampling template, an accelerated version of the entire lineal path function was proposed. Acceleration involves namely reformulation of the sequential C/C++ code for the repeatedly called part of the lineal path function into the parallel C/C++ code with CUDA extensions enabling the use of computational potential for the NVIDIA graphics processing unit (GPU). Even though the algorithm requires copying relatively large data structures to the GPU, we have shown that the principal limitations reside in the computational time required within the compression or reconstruction process, where the lineal path function often needs to be called, indeed, more than a million times. Despite parallel evaluation of the lineal path function on the GPU, the evaluation of the two-point probability function remains faster even on a single CPU thanks to an accelerated formulation based on the Fast Fourier Transform.
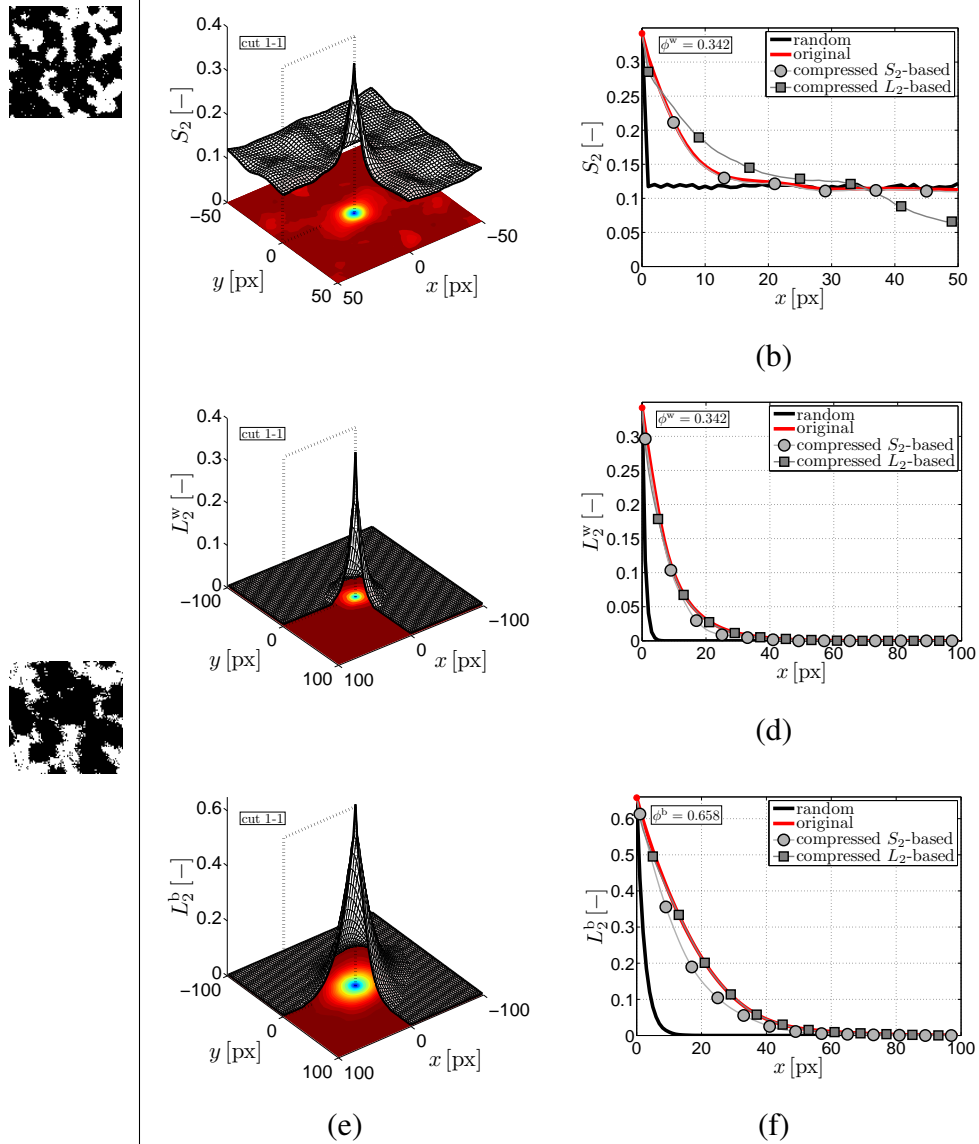
Figure 19: Results of compressed trabecular bone microstructure: (a) S_2-function of compressed S_2-based medium; (b) Comparison of S_2-functions in cut 1-1; (c) L_2^{w}-function of compressed L_2-based system; (d) Comparison of L_2^{w}-functions in cut 1-1; (e) L_2^{b}-function of compressed L_2-based system; (f) Comparison of L_2^{b}-functions in cut 1-1;

The accelerated discrete versions of both descriptors were successfully employed for microstructure reconstruction and compression processes governed by the simulated annealing algorithm based on interchanging two interfacial pixels belonging to opposite phases. We demonstrated that unlike the two-point probability function, the discrete version of the lineal path function based on line segments defined by Bresenham's algorithm does not generally ensure a unique solution for a reconstruction process. Nevertheless, the difference among feasible solutions is small and decreases with increasing resolution. On the other hand, many different morphologies could be fully defined by the lineal path computed for only one continuous phase.

Three particular microstructures were employed for illustrating typical features of both descriptors. A particulate suspension consisting of equal sized squares revealed the S_2 function was unable to capture the shape of particles. This can be emphasised by the L_2 function. The example of epithelial cells demonstrated that the S_2 function also cannot capture very thin walls and that the computation of L_2 corresponding to the phase of the walls is surprisingly not needed to achieve the most connected walls in the compressed cell. Trabecular bone, on the other hand, represents an example of mutually penetrating phases of comparable volume fractions, where both descriptors provided visually well developed microstructures.

We may conclude that despite the proposed acceleration steps, the lineal path function is still a computationally expensive descriptor which can, however, be essential when compressing morphologies consisting of specific formations such as particles or thin walls.

Acknowledgment

References

[1] S. Torquato, Random Heterogeneous Materials: Microstructure and Macroscopic Properties, Springer; Corrected edition, 2006.

[2] J. Vorel, M. Šejnoha, Evaluation of homogenized thermal conductivities of imperfect carbon-carbon textile composites using the Mori-Tanaka method, Structural Engineering and Mechanics 33 (2009) 429–446.

[3] W. E, B. Engquist, X. Li, W. Ren, E. Vanden-Eijnden, Heterogeneous multiscale methods: a review, Communications in Computational Physics 2 (3) (2007) 367–450.

[4] M. G. D. Geers, V. G. Kouznetsova, W. A. M. Brekelmans, Multi-scale computational homogenization: Trends and challenges, Journal of Computational and Applied Mathematics 234 (7) (2010) 2175–2182.

[5] J. Sýkora, T. Krejčí, J. Kruis, M. Šejnoha, Computational homogenization of nonstationary transport processes in masonry structures, Journal of Computational and Applied Mathematics 18 (2012) 4745–4755.

[6] J. Sýkora, M. Šejnoha, J. Šejnoha, Homogenization of coupled heat and moisture transport in masonry structures including interfaces, Applied Mathematics and Computation 219 (13) (2013) 7275–7285.

[7] J. Zeman, M. Šejnoha, From random microstructures to representative volume elements, Modelling and Simulation in Materials Science and Engineering 15 (4) (2007) 325–335.

[8] J. Schröder, D. Balzani, D. Brands, Approximation of random microstructures by periodic statistically similar representative volume elements based on lineal-path functions, Archive of Applied Mechanics 81 (7) (2011) 975–997.

[9] S. R. Niezgoda, D. M. Turner, D. T. Fullwood, S. R. Kalidindi, Optimized structure based representative volume element sets reflecting the ensemble-averaged 2-point statistics, Acta Materialia 58 (13) (2010) 4432–4445.

[10] J. Novák, A. Kučerová, J. Zeman, Compressing random microstructures via stochastic wang tilings, Physical Review E 86 (4) (2012) 040104.

[11] J. Novák, A. Kučerová, J. Zeman, Microstructural enrichment functions based on stochastic wang tilings, Modelling and Simulation in Materials Science and Engineering 21 (2) (2013) 025014.

[12] M. Doškář, J. Novák, J. Zeman, Aperiodic compression and reconstruction of real-world material systems based on Wang tiles, Physical Review E 90 (2014) 062118.

[13] C. L. Y. Yeong, S. Torquato, Reconstructing random media, Physical Review E 57 (1998) 495–506.

[14] D. T. Fullwood, S. R. Niezgoda, S. R. Kalidindi, Microstructure reconstructions from 2-point statistics using phase-recovery algorithms, Acta Materialia 56 (5) (2008) 942–948.

[15] D. T. Fullwood, S. R. Kalidindi, S. R. Niezgoda, A. Fast, N. Hampson, Gradient-based microstructure reconstructions from distributions using fast fourier transforms, Materials Science and Engineering A 494 (1-2) (2008) 68–72.

[16] G. L. Povirk, Incorporation of microstructural information into models of two-phase materials, Acta Metallurgica et Materialia 43 (8) (1995) 3199 – 3206.

[17] B. L. Adams, P. R. Morris, T. T. Wang, K. S. Willden, S. I. Wright, Description of orientation coherence in polycrystalline materials, Acta Metallurgica 35 (12) (1987) 2935–2946.

[18] S. R. Niezgoda, D. T. Fullwood, S. R. Kalidindi, Delineation of the space of 2-point correlations in a composite material system, Acta Materialia 56 (18) (2008) 5285–5292.

[19] M. V. Karsanina, K. M. Gerke, E. B. Skvortsova, D. Mallants, Universal spatial correlation functions for describing and reconstructing soil microstructure, PLoS ONE 10 (5) (2015) 1–26.

[20] B. Lu, S. Torquato, Lineal-path function for random heterogeneous materials, Physical Review A 45 (2) (1992) 922–929.

[21] H. Singh, A. M. Gokhale, S. I. Lieberman, S. Tamirisakandala, Image based computations of lineal path probability distributions for microstructure representation, Materials Science and Engineering A 474 (1-2) (2008) 104–111.

[22] D. S. Li, M. A. Tschopp, M. Khaleel, X. Sun, Comparison of reconstructed spatial microstructure images using different statistical descriptors, Computational Materials Science 51 (1) (2012) 437–444.

[23] K. M. Gerke, M. V. Karsanina, D. Mallants, Universal stochastic multiscale image fusion: An example application for shale rock, Scientific Reports 5 (2015) 15880.

[24] D. T. Fullwood, S. R. Niezgoda, B. L. Adams, S. R. Kalidindi, Microstructure sensitive design for performance optimization, Progress in Materials Science 55 (6) (2010) 477–562.

[25] J. Zeman, Analysis of composite materials with random microstructure, Ph.D. thesis, Klokner Institute, Czech Technical University in Prague (2003).

[26] E. Kröner, Bounds for effective elastic moduli of disordered materials, Journal of the Mechanics and Physics of Solid 25 (1977) 137–155.

[27] M. J. Beran, Statistical continuum theories, Monographs in Statistical Physics, Interscience Publishers, 1968.

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

[29] J. Gajdošík, J. Zeman, M. Šejnoha, Qualitative analysis of fiber composite microstructure: Influence of boundary conditions, Probabilistic Engineering Mechanics 21 (4) (2006) 317–329.

[30] J. E. Bresenham, Algorithm for computer control of a digital plotter, IBM System journal 4 (1) (1965) 25–30.

[31] NVIDIA corporation (www), `http://developer.nvidia.com/page/home.html`, accessed: 2015-02-23.

[32] O. Jiroušek, T. Doktor, D. Kytýř, P. Zlámal, T. Fíla, P. Koudelka, I. Jandejsek, X-ray and finite element analysis of deformation response of closed-cell metal foam subjected to compressive loading, Journal of Instrumentation 8 (2) (2013) 2012–2016.

[33] C. L. Y. Yeong, S. Torquato, Reconstructing random media. II. Three-dimensional media from two-dimensional cuts, Physical Review E 58 (1998) 224–233.

[34] M. S. Talukdar, O. Torsaeter, M. A. Ioannidis, J. J. Howard, Stochastic reconstruction of chalk from 2D images, Transport in Porous Media 48 (2002) 101–123.

[35] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.

[36] V. Černý, Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm, Journal of Optimization Theory and Applications 45 (1) (1985) 41–51.

[37] M. G. Rozman, M. Utz, Efficient reconstruction of multiphase morphologies from correlation functions, Physical Review E 63 (6) (2001) 066701.

[38] M. Lepš, M. Šejnoha, New approach to optimization of reinforced concrete beams., Computers & Structures 81 (18–19) (2003) 1957–1966.

[39] P. Čapek, V. Hejtmánek, L. Brabec, A. Zikánová, M. Kočiřík, Stochastic reconstruction of particulate media using simulated annealing: Improving pore connectivity, Transport in Porous Media 76 (2) (2009) 178–198.

[40] M. A. Davis, S. D. C. Walsh, M. O. Saar, Statistically reconstructing continuous isotropic and anisotropic two-phase media while preserving macroscopic material properties, Physical Review E 83 (2011) 026706.

[41] M. G. Rozman, M. Utz, Uniqueness of reconstruction of multiphase morphologies from two-point correlation functions, Physical Review Letters 89 (13) (2002) 135501.

[42] W. Nagel, Orientation-dependent chord length distributions characterize convex polygons, Journal of Applied Probability 30 (3) (1993) 730–736.

[43] H. Lee, M. Brandyberry, A. Tudor, K. Matouš, Three-dimensional reconstruction of statistically optimal unit cells of polydisperse particulate composites from microtomography, Physical Review E 80 (2009) 061301.

[44] O. Jiroušek, P. Zlámal, D. Kytýř, M. Kroupa, Strain analysis of trabecular bone using time-resolved X-ray microtomography, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 663 (2011) S148–S151.

Chapter 3

CALIBRATION OF RANDOM MATERIAL MODELS

This chapter elaborates on the methods for identifying properties of random materials from the macroscopic behavior of specimens in a laboratory or real-world structures. The identification method applied here is based on Bayesian inference combining the expert knowledge about the realistic values of material properties with the information contained in the limited and noisy experimental data. Its principal advantage is the well-posed probabilistic formulation allowing to quantify the uncertainty in the estimated values of material properties. To preserve the robustness of the method, the probabilistic description of the estimated material parameters is obtained via Markov chain Monte Carlo sampling. Its crucial drawback consists of high computational requirements connected to repeated simulations of a finite element-based nonlinear model of the experiments. To overcome this inconvenience, the simulations by the structural model are replaced by evaluations of polynomial chaos-based approximation of simulated response components. The robustness and efficiency of the method are demonstrated in two examples:(i) identification of the spatial distribution of material transport properties within one rectangular 2D sample based on measurements of temperature and moisture modeled by highly nonlinear Künzel's coupled heat and moisture transport model and (ii) identification of homogenized mechanical properties of stochastic lattice discrete particle model for concrete from a set of stress-strain diagrams obtained in unconfined compression cube test and notched three-point bending test.

List of selected journal papers and book chapters

Kučerová, A., Sýkora, J., Rosić, B., and Matthies, H. G. (2012). Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials. *Journal of Computational and Applied Mathematics*, 236(18):4862–4872.

Janouchová, E., Kučerová, A., Sýkora, J., Vorel, J., and Wan-Wedner, R. (2020). Robust probabilistic calibration of stochastic lattice discrete particle model for concrete. *Cement and Conrete Composites.*, Submitted.

List of related journal papers and book chapters

Kučerová, A. and Sýkora, J. (2013). Uncertainty updating in the description of coupled heat and moisture transport in heterogeneous materials. *Applied Mathematics and Computation*, 219(13):7252–7261.

Rosić, B. V., Kučerová, A., Sýkora, J., Pajonk, O., Litvinenko, A., and Matthies, H. G. (2013). Parameter identification in a probabilistic setting. *Engineering Structures*, 50:179–196.

Acceleration of Uncertainty Updating in the Description of Transport Processes in Heterogeneous Materials

Anna Kučerová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Jan Sýkora

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Bojana Rosić

Institute of Scientific Computing, Technische Universität Braunschweig, Germany

Hermann G. Matthies

Institute of Scientific Computing, Technische Universität Braunschweig, Germany

Abstract

The prediction of thermo-mechanical behaviour of heterogeneous materials such as heat and moisture transport is strongly influenced by the uncertainty in parameters. Such materials occur e.g. in historic buildings, and the durability assessment of these therefore needs a reliable and probabilistic simulation of transport processes, which is related to the suitable identification of material parameters. In order to include expert knowledge as well as experimental results, one can employ an updating procedure such as Bayesian inference. The classical probabilistic setting of the identification process in Bayes's form requires the solution of a stochastic forward problem via computationally expensive sampling techniques, which makes the method almost impractical. In this paper novel stochastic computational techniques such as the stochastic Galerkin method are applied in order to accelerate the updating procedure. The idea is to replace the computationally expensive forward simulation via the conventional finite element (FE) method by the evaluation of a polynomial chaos expansion (PCE). Such an approximation of the FE model for the forward simulation perfectly suits the Bayesian updating. The presented uncertainty updating techniques are applied to the numerical model of coupled heat and moisture transport in heterogeneous materials with spatially varying coefficients defined by random fields.

Keywords: Uncertainty updating, Bayesian inference, Heterogeneous materials, Coupled heat and moisture transport, Künzel's model, Stochastic finite elements, Galerkin methods, Polynomial chaos expansion, Karhunen-Loève expansion

Email addresses: `anicka@cml.fsv.cvut.cz` (Anna Kučerová), `jan.sykora.1@fsv.cvut.cz` (Jan Sýkora), `bojana.rosic@tu-bs.de` (Bojana Rosić), `wire@tu-bs.de` (Hermann G. Matthies)

1. Introduction

Durability of structures is influenced by moisture damage processes. High moisture levels cause metal corrosion, wood decay and other structural degradation. Thermal expansion and contraction, on the other hand, can induce large displacements and extensive damage to structural materials with differing coefficients ,e.g. masonry. The Charles Bridge in Prague, currently the subject of rehabilitation works, is a typical example, see [1]. A study of the coupled heat and moisture transport behaviour is thus essential in order to improve the building materials' performance. So far, a vast number of models have been introduced for the description of transport phenomena in porous media. An extensive overview of transport models can be found in [2]. In this work we focus on the model by Künzel [3], since the predicted results comply well with the results of experimental measurements [4], once the relevant material parameters are well estimated.

Material properties are usually determined from experimental measurements via an identification procedure, see e.g. [5]. However, the experimental measurements as well as the identification methods involve some inevitable errors. Bayesian updating, employed within this study, provides a general framework for inference from noisy and limited data. It enables mutually involving both expert knowledge of the material, such as limit values of physical parameters, and information from experimental observations and measurements. In other words, it uses experimental data to update the so-called a priori uncertainty in the material description and results in a posterior probabilistic description of material performance [6]. In addition, unlike traditional identification techniques that aim to regularise the ill-posed inverse problem to achieve a point estimate, the Bayesian identification process leads to a well-posed problem in an expanded stochastic space.

The main disadvantage of Bayesian updating lies in the significant computational effort that results from the sampling-based estimation of posterior densities [7]. While deterministic quadrature or cubature may be attractive alternatives to Monte Carlo at low to moderate dimensions [8], computationally exhaustive Markov chain Monte Carlo (MCMC) remains the most general and flexible method for complex and high-dimensional distributions [9, 10]. In a sampling-based procedure, the posterior distribution must be evaluated for any sample generated from the prior one in order to decide, whether the sample is admissible or not. The computation of the posterior involves the evaluation of the computational model—the FE discretisation of a non-linear partial differential equation (PDE)—relating model (i.e. material) parameters and observable quantities (i.e. model outputs). Hence, complex and time-consuming models can make the sampling procedure practically unfeasible.

Bayesian updating of uncertainty in the description of the parameters of Künzel's model is thoroughly described in [11] for the case of heterogeneous material, where material parameters are described by random fields (RFs). It was shown that Bayesian updating is applicable even for such a complex and nonlinear model as Künzel's model. However, the demonstrated results were performed for a sample with a coarse FE, thereby rendering the evaluation of the numerical model computationally relatively cheap. A higher complexity of modelled structure and its FE-based numerical model lead to time-consuming simulations and are prohibitive for the sampling procedure. In such a case one may construct an approximation of the model response and evaluate this within the sampling procedure in order to render the updating procedure feasible [12, 13].

The efficient forward propagation of uncertainty, which may describe material properties, the geometry of the domain, external loading etc., from model parameters to model

outputs is a main topic of stochastic mechanics. The recently developed polynomial chaos (PC) variant of the stochastic finite element method (SFEM)—the spectral SFEM (SSFEM) [14, 15, 16, 17, 18]—has become one of the promising techniques in this area. Some of the uncertainties in the model are represented as random fields/processes. Here one often employed technique in SFEM computations is the use of a truncated Karhunen-Loève expansion (KLE) to represent the RFs in a computationally efficient manner by means of a minimal set of random variables (RVs) [19, 16, 20, 21], via an eigenvalue decomposition of the covariance. This approach involves the introduction of an orthogonal—hence uncorrelated—basis in a space of RVs. These are projections of the RF onto the orthogonal KL eigenfunctions, and in the case of Gaussian RFs consists of Gaussian RVs. In that case they are not only uncorrelated but independent—a computationally very important property [22]. However, the material properties very often cannot be modelled as Gaussian due to crucial constraints such as positive definiteness, boundedness in some interval, etc. In such a case, one has to adopt non-Gaussian models and their corresponding approximations, see [17, 18, 23], often as a non-linear transformation of a Gaussian RF. The orthogonal or uncorrelated RVs alluded to above are not Gaussian in that case, and hence not independent. One then may adopt a pure PC representation of the RF in terms of polynomials of independent Gaussian RVs [14], or—to take advantage of the dimension reduction inherent in the KLE truncation—one uses the PC representation for the orthogonal/uncorrelated non-Gaussian RVs from the KLE [16, 20].

In this paper, we focus on Künzel's model [24, 3], defined by uncertain positive-definite material parameters, modelled as log-normal RFs according to the maximum entropy principle. Since these RFs are non-Gaussian, their spectral decomposition (KLE) gives a set of uncorrelated but not necessarily independent RVs. To address this problem, we project the RVs onto a PC basis constructed from Hermite polynomials in independent Gaussian RVs as alluded to in the previous paragraph. Such a combined expansion (KL/PC) is then used to represent the RFs as inputs to the FE discretisation of the nonlinear Künzel model. The solution procedure of Galerkin type for this SPDE is chosen in an "intrusive" manner based on analytic computations in the PC/Hermite algebra [20, 25, 26]. This brings huge computational savings in case of small and moderate problem dimensions, but it requires complete knowledge of the model (the FEM system can not be used in black-box fashion).

Once such a representation is propagated through the physical model, one obtains a description of all desired output quantities in terms of simply evaluable functions—in this case polynomials—of known independent Gaussian RVs. This is often called a surrogate model or a response surface.

The paper is organised in the following way. The next Section 2 reviews Künzel's model. Section 3 is focused on the probabilistic description of heterogeneous material properties where particular material parameters are not spatially constant. Intrusive stochastic Galerkin method for computing coefficients of the PC-based surrogate of outputs of Künzel's model is developed in Section 4 and the related outcomes a presented in Section 5. Finally, Section 6 presents the Bayesian updating procedure on Künzel's model with the results summarized in Section 7, and Section 8 concludes.

2. Coupled heat and moisture transfer

Künzel [24, 3] derived balance equations describing coupled heat and moisture transport through porous media using the concepts of Krischer and Kiessl. Krischer [27] identified two

transport mechanisms for material moisture, one being the vapour diffusion and the other being described as capillary water movement. In other words, he introduced the gradient of partial pressure in air as a driving force for the water vapour transport and the gradient of liquid moisture content as the driving force for the water transport. This model is then extended by Kiessl [28] who introduced the so-called moisture potential Φ used for unification of the description of moisture transport in the hygroscopic $\varphi \leq 0.9$ and over-hygroscopic $\varphi > 0.9$ range (where φ is relative humidity). The introduction of the moisture potential brings several advantages, especially very simple expressions for the moisture transport across the interface. On the other hand, the definition of the moisture potential in the over-hygroscopic range was too artificial, and Kiessl introduced it without any theoretical background, see [2].

For the description of simultaneous water and water vapour transport Künzel chose the relative humidity φ as the only moisture potential for both the hygroscopic and the over-hygroscopic range. He also divided the over-hygroscopic region into two sub-ranges—the capillary water region and supersaturated region—where different conditions for water and water vapour transport are considered. In comparison with Kiessl's or Krischer's model Künzel's model brings certain simplifications. Nevertheless, the proposed model describes all substantial phenomena and the predicted results comply well with experimentally obtained data [4]. Therefore, it was chosen as a physical basis for the formulation of the probabilistic framework.

Künzel's model is described by the energy balance equation

$$\frac{\mathrm{d}H}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} = \boldsymbol{\nabla}^{\mathrm{T}}[\lambda(\varphi)\boldsymbol{\nabla}\theta] + h_{\mathrm{v}}(\theta)\boldsymbol{\nabla}^{\mathrm{T}}[\delta_{\mathrm{p}}(\theta)\boldsymbol{\nabla}\{\varphi p_{\mathrm{sat}}(\theta)\}] \tag{1}$$

and the conservation of mass equation

$$\frac{\mathrm{d}w}{\mathrm{d}\varphi}\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \boldsymbol{\nabla}^{\mathrm{T}}[D_{\varphi}(\varphi)\boldsymbol{\nabla}\varphi] + \boldsymbol{\nabla}^{\mathrm{T}}[\delta_{\mathrm{p}}(\theta)\boldsymbol{\nabla}\{\varphi p_{\mathrm{sat}}(\theta)\}], \tag{2}$$

where the transport coefficients defining the material behaviour are nonlinear functions of structural responses, i.e. the temperature $\theta[^\circ\mathrm{C}]$ and moisture $\varphi[-]$ fields. We briefly recall their particular relations [3]:

- Thermal conductivity $[\mathrm{Wm}^{-1}\mathrm{K}^{-1}]$:

$$\lambda = \lambda_0 \left(1 + \frac{b_{\mathrm{tcs}}w_{\mathrm{f}}(b-1)\varphi}{\rho_{\mathrm{s}}(b-\varphi)} \right). \tag{3}$$

- Evaporation enthalpy of water $[\mathrm{Jkg}^{-1}]$:

$$h_{\mathrm{v}} = 2.5008 \cdot 10^6 \left(\frac{273.15}{\theta + 273.15} \right)^{(0.267 + 3.67 \cdot 10^{-4}\theta)}. \tag{4}$$

- Water vapour permeability $[\mathrm{kgm}^{-1}\mathrm{s}^{-1}\mathrm{Pa}^{-1}]$:

$$\delta_{\mathrm{p}} = \frac{1.9446 \cdot 10^{-12}}{\mu} \cdot (\theta + 273.15)^{0.81}. \tag{5}$$

- Water vapour saturation pressure [Pa]:

$$p_{\text{sat}} = 611 \exp\left(\frac{17.08\theta}{234.18 + \theta}\right). \tag{6}$$

- Liquid conduction coefficient [$\text{kgm}^{-1}\text{s}^{-1}$]:

$$D_\varphi = 3.8 \frac{a^2}{w_{\text{f}}} \cdot 10^{\frac{3w_{\text{f}}(b-1)\varphi}{(b-\varphi)(w_{\text{f}}-1)}} \cdot \frac{b(b-1)}{(b-\varphi)^2}. \tag{7}$$

- Total enthalpy of building material [Jm^{-3}]:

$$H = \rho_{\text{s}} c_{\text{s}} \theta. \tag{8}$$

- Water content [kgm^{-3}]:

$$w = w_{\text{f}} \frac{(b-1)\varphi}{b - \varphi}. \tag{9}$$

A more detailed discussion on the transport coefficients can be found in [3, 29]. Some of them defined by Eqs. (3)–(8) depend on a subset of the material parameters listed in Tab. 1. The approximation factor b appearing in Eqs. (3) and (7) can be determined from the relation:

$$b = \frac{0.8(w_{80} - w_{\text{f}})}{w_{80} - 0.8w_{\text{f}}}, \tag{10}$$

where w_{80} is the equilibrium water content at $0.8\,[-]$ relative humidity. Moreover, the free water saturation w_{f} must always be greater than w_{80}. Therefore we introduce the water content increment $\mathrm{d}w_{\text{f}} > 0$ and define the free water saturation as

$$w_{\text{f}} = w_{80} + \mathrm{d}w_{\text{f}}. \tag{11}$$

Consequently, w_{80} and $\mathrm{d}w_{\text{f}}$ substitute b and w_{f} as material parameters to be identified within the updating procedure. Tab. 1 presents the resulting list of $W = 8$ material parameters to be identified. As an outcome of such a substitution, all identified parameters should be positive and thus described by log-normal RFs (a priori information) with second order statistics (mean values μ_q and standard deviations σ_q) given in Tab. 1. Those particular values are chosen to correspond to materials used in masonry [30].

The partial differential equations (1) and (2) are discretised in space by standard finite elements. This also goes well with the use of the stochastic Galerkin method for the discretisation in the stochastic space. Performing first only the spatial discretisation, the temperature and moisture fields are spatially approximated as

$$\theta(\boldsymbol{x}) = \sum_{n=1}^{N} \phi_n(\boldsymbol{x}) u_{\theta,n}, \qquad \varphi(\boldsymbol{x}) = \sum_{n=1}^{N} \phi_n(\boldsymbol{x}) u_{\varphi,n} \tag{12}$$

where N is the number of nodes in FE discretisation, $\phi_n(\boldsymbol{x})$ are the shape functions (according to the type of used elements) and $u_{\theta,n}$ and $u_{\varphi,n}$ are the nodal values of temperature field θ and moisture field φ, respectively.

Parameter			μ_q	σ_q
$\mathrm{d}w_{\mathrm{f}}$	$[\mathrm{kgm}^{-3}]$	water content increment	100	20
w_{80}	$[\mathrm{kgm}^{-3}]$	water content at $0.8\,[-]$ relative humidity	50	10
λ_0	$[\mathrm{Wm}^{-1}\mathrm{K}^{-1}]$	thermal conductivity of dry material	0.3	0.1
b_{tcs}	$[-]$	thermal conductivity supplement	10	2
μ	$[-]$	water vapour diffusion resistance factor	12	5
a	$[\mathrm{kgm}^{-2}\mathrm{s}^{-0.5}]$	water absorption coefficient	0.6	0.2
c_{s}	$[\mathrm{Jkg}^{-1}\mathrm{K}^{-1}]$	specific heat capacity	900	100
ρ_{s}	$[\mathrm{kgm}^{-3}]$	bulk density of building material	1650	50

Table 1: Mean values and standard deviations of material parameters

Using the approximations Eq. (12) and Eqs. (1), (2), we obtain a set of first order differential equations

$$\boldsymbol{K}(\boldsymbol{u})\boldsymbol{u} + \boldsymbol{C}(\boldsymbol{u})\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \boldsymbol{F}, \tag{13}$$

where $\boldsymbol{K}(\boldsymbol{u})$ is the conductivity matrix, $\boldsymbol{C}(\boldsymbol{u})$ is the capacity matrix, $\boldsymbol{u}^{\mathrm{T}} = (u_{\theta,1}, \ldots, u_{\theta,N}, u_{\varphi,1}, \ldots, u_{\varphi,N})$ is the vector of nodal values, and $\boldsymbol{F}$ is the vector of prescribed fluxes transformed into nodes. For a detailed formulation of the matrices $\boldsymbol{K}(\boldsymbol{u})$ and $\boldsymbol{C}(\boldsymbol{u})$ and the vector $\boldsymbol{F}$, we refer the interested reader to the doctoral thesis [31, Chapter 3.1].

The numerical solution of the system Eq. (13) is based on a simple temporal finite difference discretisation. If we use time steps $\Delta\tau$ and denote the quantities at time step i with a corresponding superscript, the time-stepping equation is

$$\boldsymbol{u}^{i+1} = \boldsymbol{u}^i + \Delta\tau[(1-\gamma)\dot{\boldsymbol{u}}^i + \gamma\dot{\boldsymbol{u}}^{i+1}], \tag{14}$$

where γ is a generalised midpoint integration rule parameter. In the results presented in this paper the Crank-Nicolson (trapezoidal rule) integration scheme with $\gamma = 0.5$ was used. Expressing $\dot{\boldsymbol{u}}^{i+1}$ from Eq. (14) and substituting into the Eq. (13), one obtains a system of non-linear equations:

$$(\gamma\Delta\tau\boldsymbol{K}^{i+1} + \boldsymbol{C}^{i+1})\boldsymbol{u}^{i+1} = \gamma\Delta\tau\boldsymbol{F}^{i+1} + \boldsymbol{C}^{i+1}[\boldsymbol{u}^i + \Delta\tau(1-\gamma)\dot{\boldsymbol{u}}^i], \tag{15}$$

which can be solved by some iterative method such as Newton-Raphson. For clarification and easier reading, we rewrite Eq. (15) using the symbols $\boldsymbol{A}^{i+1}(\boldsymbol{u}^{i+1}) := \gamma\Delta\tau\boldsymbol{K}^{i+1}(\boldsymbol{u}^{i+1}) + \boldsymbol{C}^{i+1}(\boldsymbol{u}^{i+1})$ and $\boldsymbol{f}^{i+1}(\boldsymbol{u}^{i+1}) := \gamma\Delta\tau\boldsymbol{F}^{i+1} + \boldsymbol{C}^{i+1}(\boldsymbol{u}^{i+1})[\boldsymbol{u}^i + \Delta\tau(1-\gamma)\dot{\boldsymbol{u}}^i]$ in the following form

$$\boldsymbol{A}^{i+1}(\boldsymbol{u}^{i+1})\boldsymbol{u}^{i+1} = \boldsymbol{f}^{i+1}(\boldsymbol{u}^{i+1}). \tag{16}$$

3. Uncertain properties of heterogeneous materials

When dealing with heterogeneous material, some material parameters can vary spatially in an uncertain fashion and therefore RFs are suitable for their description. This means that the uncertainty in a particular material parameter q is modelled by defining $q(\boldsymbol{x})$ for each $\boldsymbol{x} \in \mathcal{G}$ as a RV $q(\boldsymbol{x}) : \Omega \to \mathbb{R}$ on a suitable probability space $(\Omega, \mathscr{S}, \mathbb{P})$ in some bounded admissible region $\mathcal{G} \subset \mathbb{R}^d$. As a consequence, $q : \mathcal{G} \times \Omega \to \mathbb{R}$ is a RF and one may identify Ω with the set of all possible realisations of q. Alternatively, $q(\boldsymbol{x}, \omega)$ can be seen as a collection of real-valued RVs indexed by $\boldsymbol{x} \in \mathcal{G}$.

The description of log-normal RFs given in Tab. 1 can be derived from a Gaussian RF $g(\boldsymbol{x}, \omega)$, which is defined by its mean

$$\mu_g(\boldsymbol{x}) = \mathbb{E}[g(\boldsymbol{x}, \omega)] = \int_\Omega g(\boldsymbol{x}, \omega)\,\mathbb{P}(\mathrm{d}\omega)\,, \tag{17}$$

and its covariance

$$\begin{aligned} C_g(\boldsymbol{x}, \boldsymbol{x}') &= \mathbb{E}[(g(\boldsymbol{x}, \omega) - \mu_g(\boldsymbol{x}))(g(\boldsymbol{x}', \omega) - \mu_g(\boldsymbol{x}'))] \\ &= \int_\Omega (g(\boldsymbol{x}, \omega) - \mu_g(\boldsymbol{x}))(g(\boldsymbol{x}', \omega) - \mu_g(\boldsymbol{x}'))\,\mathbb{P}(\mathrm{d}\omega)\,. \end{aligned} \tag{18}$$

The log-normal RF $q(\boldsymbol{x}, \omega)$ can be then obtained by a nonlinear transformation of a zero-mean unit-variance Gaussian RF $g(\boldsymbol{x}, \omega)$ [20, 26] as

$$q(\boldsymbol{x}, \omega) = \exp(\mu_g(\boldsymbol{x}) + \sigma_g g(\boldsymbol{x}, \omega))\,. \tag{19}$$

The statistical moments μ_g and σ_g of the Gaussian field can be obtained from the statistical moments μ_q and σ_q given for the log-normally distributed material property according to the following relations [26]:

$$\sigma_g^2 = \ln\left(1 + \left(\frac{\sigma_q}{\mu_q}\right)^2\right)\,, \quad \mu_g = \ln\mu_q - \frac{1}{2}\sigma_g^2\,. \tag{20}$$

In numerical computation random fields are first spatially discretised by finite element method (see Eq. (12)) into a finite collection of points $\{\boldsymbol{x}_{i=1}^n\} \in \mathcal{G}$. Further, the semi-discretised RF are described by a finite—but probably very large—number of RVs $\boldsymbol{q}(\omega) = (q(\boldsymbol{x}_1, \omega), \ldots, q(\boldsymbol{x}_n, \omega))$, which are usually highly correlated. Large number of RVs is, however, very challenging for the efficient numerical implementation of forward problem, as well as for MCMC identification. As already alluded to previously, the number of RVs can be reduced by the approximation $\hat{\boldsymbol{g}}(\omega)$ of a RF $\boldsymbol{g}(\omega)$ based on a truncated KLE including much smaller number of RVs [20, 13]. Here we use the KLE on the underlying Gaussian field $\boldsymbol{g}(\omega)$, and hence the RVs in the KLE are independent Gaussian RVs, as already indicated above.

The spatial discretisation of a given RF concerns also the discretisation of corresponding covariance function $C_g(\boldsymbol{x}, \boldsymbol{x}')$ into the covariance matrix $\boldsymbol{C}_g$ which is symmetric and positive definite [16, 20]. The KLE is based on the spectral decomposition of the covariance matrix $\boldsymbol{C}_g$ leading to the solution of a symmetric matrix eigenvalue problem

$$\boldsymbol{C}_g \boldsymbol{\psi}_i = \varsigma_i^2 \boldsymbol{\psi}_i\,, \tag{21}$$

where $\boldsymbol{\psi}_i$ are orthogonal eigenvectors and ς_i^2 are positive eigenvalues ordered in a descending order. The KLE approximation $\hat{\boldsymbol{g}}(\omega)$ of a RF $\boldsymbol{g}(\omega)$ can then be written as

$$\hat{\boldsymbol{g}}(\omega) = \boldsymbol{\mu}_g + \sum_{i=0}^{M} \varsigma_i\,\xi_i(\omega)\boldsymbol{\psi}_i, \tag{22}$$

where $\xi_i(\omega) = \boldsymbol{\psi}_i^T(\boldsymbol{g}(\omega) - \boldsymbol{\mu}_g)/\varsigma_i$ are uncorrelated RVs of zero mean and unit variance, and in case that $g(\boldsymbol{x}, \omega)$ and hence $\boldsymbol{g}(\omega)$ are Gaussian, then $x_i(\omega)$ are Gaussian and independent. The number $M \le n$—the number of points used for the discretisation of the

spatial domain—is chosen such that Eq. (22) gives a good approximation, i.e. captures a high proportion of the total variance. Higher values of M lead to better description of a RF, smaller values imply faster exploration by MCMC. The eigenvalue problem Eq. (21) is usually solved by a Krylov subspace method with a sparse matrix approximation. For large eigenvalue problems, the authors in [32] propose efficient low-rank and data sparse hierarchical matrix techniques. The approximation of a non-Gaussian RF can be then obtained by a nonlinear transformation of the KLE obtained for a Gaussian RF such as in our particular case, where the approximation of a given RF $\hat{\boldsymbol{q}}(\omega)$ is obtained from the Eq. (19) by the substitution of the Gaussian RF $\boldsymbol{g}(\omega)$ by its KLE $\hat{\boldsymbol{g}}(\omega)$.

We assume full spatial correlation among material properties, i.e. spatial fluctuations for all parameters differ only in magnitude. Taking into account a log-normal distribution of the parameters, the final formulation of the RF describing the parameter q then becomes

$$\hat{\boldsymbol{q}}(\omega) = \exp\left(\boldsymbol{\mu}_g + \sigma_g \sum_{i=1}^{M} \sqrt{\varsigma_i}\xi_i(\omega)\boldsymbol{\psi}_i\right) , \tag{23}$$

where the exponential is to be used at each spatial point, i.e. for each component of the vector inside the parentheses. The statistical moments $\boldsymbol{\mu}_g$ and σ_g are derived from the prior mean $\boldsymbol{\mu}_q$ and standard deviation σ_q for each material parameter according to Eq. (20). The eigenvectors $\boldsymbol{\psi}_i$ are obtained for the a priori exponential covariance function

$$C(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\frac{|r_1|}{l_{x_1}} - \frac{|r_2|}{l_{x_2}}\right) , \tag{24}$$

where $\boldsymbol{r} = (r_1, r_2) = \boldsymbol{x} - \boldsymbol{x}'$, and $l_{x_1} = 0.1$ [m] and $l_{x_2} = 0.04$ [m] are a priori covariance lengths. Determination of correlation lengths is generally not obvious. In material modelling, one possible way is based on image analysis as described in [33]. A numerical study for a differing number of modes M included in the KLE is presented in [11].

4. Surrogate of Künzel's model

While the KLE can be efficiently applied to reduce the number of RVs and thus to accelerate the exploration of the MCMC method in terms of the number of samples, construction of a surrogate of the computational model can be used for a significant acceleration of each sample evaluation. In [12, 13] methods were introduced for accelerating Bayesian inference in this context through the use of stochastic spectral methods to propagate the prior uncertainty through the forward problem. Here we employ the stochastic Galerkin method [15, 16] to construct the surrogate of Künzel's model based on polynomial chaos expansion (PCE).

According to Eq. (23), all model parameters are characterised by M independent standard Gaussian RVs $\boldsymbol{\xi}(\omega) = [\xi_1(\omega), \ldots, \xi_M(\omega)]$. Hence, the discretised model response $\boldsymbol{u}(\boldsymbol{\xi}(\omega)) = (\ldots, u_i(\boldsymbol{\xi}(\omega)) \ldots)^{\mathrm{T}}$ is a random vector which can be expressed in terms of the same RVs $\boldsymbol{\xi}(\omega)$. Since $\boldsymbol{\xi}(\omega)$ are independent standard Gaussian RVs, Wiener's PCE based on multivariate Hermite polynomials—orthogonal in the Gaussian measure—$\{H_\alpha(\boldsymbol{\xi}(\omega))\}_{\alpha \in \mathcal{J}}$ (see [16, 20] for the notation) is the most suitable choice for the approximation $\tilde{\boldsymbol{u}}(\boldsymbol{\xi}(\omega))$ of the model response $\boldsymbol{u}(\boldsymbol{\xi}(\omega))$ [34], and it can be written as

$$\tilde{\boldsymbol{u}}(\boldsymbol{\xi}(\omega)) = \sum_{\alpha \in \mathcal{J}} \boldsymbol{u}_\alpha H_\alpha(\boldsymbol{\xi}(\omega)) \tag{25}$$

where $\boldsymbol{u}_\alpha$ is a vector of PC coefficients and the index set $\mathcal{J} \subset \mathbb{N}_0^{(\mathbb{N})}$ is a finite set of non-negative integer sequences with only finitely many non-zero terms, i.e. multi-indices, with cardinality $|\mathcal{J}| = R$. We collect all the PC coefficients in $\mathbf{u} := [\dots, \boldsymbol{u}_\alpha, \dots]_{\alpha \in \mathcal{J}}$. Assuming the uncertainty in all material parameters listed in Tab. 1 and consequently in the model response, Eq. (16) can be rewritten as

$$\boldsymbol{A}^{i+1}\left(\boldsymbol{\xi}; \boldsymbol{u}^{i+1}(\boldsymbol{\xi})\right) \boldsymbol{u}^{i+1}(\boldsymbol{\xi}) = \boldsymbol{f}^{i+1}\left(\boldsymbol{\xi}; \boldsymbol{u}^{i+1}(\boldsymbol{\xi})\right). \tag{26}$$

Substituting the model response $\boldsymbol{u}^{i+1}(\boldsymbol{\xi})$ by its PC approximation $\tilde{\boldsymbol{u}}^{i+1}(\boldsymbol{\xi})$ given in Eq. (25) and applying a Bubnov-Galerkin projection, one requires that the weighted residuals vanish:

$$\forall \beta \in \mathcal{J}: \quad \mathbb{E}([\boldsymbol{f}^{i+1}\left(\boldsymbol{\xi}; \tilde{\boldsymbol{u}}^{i+1}(\boldsymbol{\xi})\right) - \tilde{\boldsymbol{A}}^{i+1}(\boldsymbol{\xi})\tilde{\boldsymbol{u}}^{i+1}(\boldsymbol{\xi})]H_\beta(\boldsymbol{\xi})) = 0\,, \tag{27}$$

where $\tilde{\boldsymbol{A}}^{i+1}(\boldsymbol{\xi}) := \boldsymbol{A}^{i+1}\left(\boldsymbol{\xi}; \tilde{\boldsymbol{u}}^{i+1}(\boldsymbol{\xi})\right)$. Eq. (27) together with Eq. (25) leads to

$$\forall \beta \in \mathcal{J}: \quad \sum_{\alpha \in \mathcal{J}} \mathbb{E}\left(H_\beta(\boldsymbol{\xi})\tilde{\boldsymbol{A}}^{i+1}(\boldsymbol{\xi})H_\alpha(\boldsymbol{\xi})\right) \boldsymbol{u}_\alpha^{i+1} = \mathbb{E}(\boldsymbol{f}^{i+1}(\boldsymbol{\xi})H_\beta(\boldsymbol{\xi}))\,, \tag{28}$$

which is a non-linear system of equations of size $N \times R$.

The approximation $\tilde{\boldsymbol{u}}^{i+1}(\boldsymbol{\xi})$ can be represented through its PC coefficients $\mathbf{u}^{i+1}$, and similarly for all other quantities. Denoting the block-matrix $\mathbf{A}^{i+1}(\mathbf{u}^{i+1}) := (\mathbb{E}\left(H_\beta(\boldsymbol{\xi})\boldsymbol{A}^{i+1}(\boldsymbol{\xi})\right.$ $\left. H_\alpha(\boldsymbol{\xi})\right))_{\beta,\alpha \in \mathcal{J}}$, and the right hand side $\mathbf{f}^{i+1} := (\mathbb{E}(\boldsymbol{f}^{i+1}(\boldsymbol{\xi})H_\beta(\boldsymbol{\xi})))_{\beta \in \mathcal{J}}$, the system Eq. (28) may succinctly be written as

$$\mathbf{A}^{i+1}(\mathbf{u}^{i+1})\mathbf{u}^{i+1} = \mathbf{f}^{i+1}. \tag{29}$$

The matrix $\mathbf{A}^{i+1}$ has more structure than is displayed here, but this is outside the scope of this paper; see [16, 20] for details and possible computational procedures.

The evaluation of expected values in Eq. (28) can often be performed analytically in intrusive Galerkin procedures—that is their advantage—using the Hermite algebra [20]. In case they are to be computed numerically, they may be approximated by a weighted sum of samples drawn from the prior distributions. To that purpose, one can apply some integration technique: the Monte Carlo (MC) method, the quasi-Monte Carlo (QMC) method, or some quadrature rule, see [20] for a recent review. The latter ones allow to take advantage of a possibly regular behaviour in the stochastic variables and consequently reduce the number of samples. Since the system of equations Eq. (28) can be quite large, the evaluation of the left hand side for each sample of $\boldsymbol{\xi}$ becomes costly. Here we apply a sparse-grid Smolyak quadrature rule [35, 22, 16, 20], sometimes also named hyperbolic cross integration method, which is an efficient alternative for integration over Gaussian RVs.

After solving the system Eq. (29), one has via Eq. (25) a surrogate representation of the model outputs. This model approximation may be evaluated orders of magnitude more quickly than the evaluation containing the full FE simulation.

5. Numerical results for the uncertainty propagation

For an illustration of the described method, we employ the same simple example as in [11] with the two-dimensional rectangular domain discretised by an FE mesh into $N = 80$ nodes and 120 triangular elements. Its geometry together with the specific loading conditions are shown in Fig. 1. The initial temperature is $\theta_{\text{in}} = 14$ [°C] and the moisture $\varphi_{\text{in}} = 0.5$ [-]
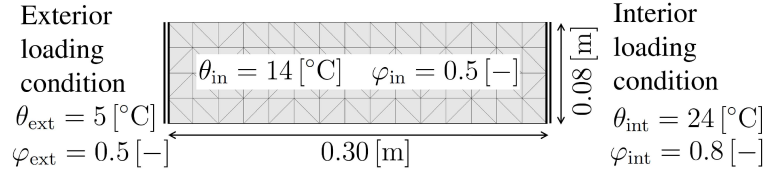
Figure 1: Experimental setup

in the whole domain. One side of the domain is submitted to exterior loading conditions $\theta_{\text{ext}} = 5$ [°C] and $\varphi_{\text{ext}} = 0.5$ [-], while the opposite side is submitted to interior loading conditions $\theta_{\text{int}} = 24$ [°C] and $\varphi_{\text{int}} = 0.8$ [-]. The solution of the time-dependent problem in Eq. (29) also involves a discretisation of the time domain $\mathcal{T}$ into $T = 151$ time steps and hence the PCE-based surrogate model consist of $N \times T = 12,080$ PCEs for the temperature, and the same for the moisture.

In order to describe the accuracy of such a surrogate model, let us define the MC estimate of the error expectation $\varepsilon(\boldsymbol{u})$ as a relative difference between two response fields $\boldsymbol{u}^a$ and $\boldsymbol{u}^b$ over the discretised spatial and time domain as

$$\varepsilon(\boldsymbol{u}) := \mathbb{E}_{\Omega} \left(\sum_{n=1}^{N} \sum_{t=1}^{T} \frac{\left| u_{n,t}^a - u_{n,t}^b \right|}{u_{n,t}^a} \right) . \tag{30}$$

The quality of a PC-based surrogate model depends on the number M of eigenmodes involved in KLE describing the fields of material properties as well as on the degree of polynomials P used in the expansion Eq. (25)[1]. Figure 2 shows the error estimate $\varepsilon(\boldsymbol{u})$ computed
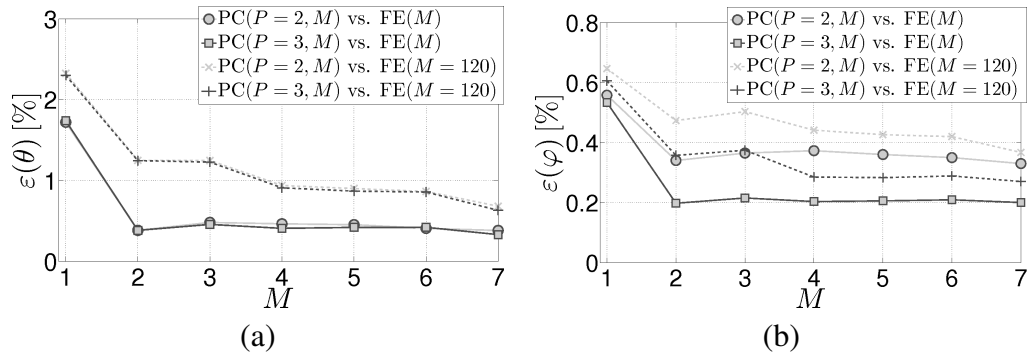
Figure 2: Errors in approximation of the temperature (a) and the moisture (b) field induced by PCE and KLE as a function of number of eigenmodes.

for different numbers of eigenmodes M and for the polynomial order $P = 2$ and $P = 3$. Here, the response fields $\boldsymbol{u}^a$ are computed by the FEM based on one realization of the KLE of the parameter fields (further shortly called FE simulations) and the response fields $\boldsymbol{u}^b$ are obtained by evaluation of the constructed PCE in the same sample point. In order to distinguish the portion of error induced by the KL approximation of the parameter fields, the estimate $\varepsilon(\boldsymbol{u})$ is computed once for the FE simulations using all $M = 120$ (dashed lines),

---

[1]We assume the full PC expansion, where number of polynomials R is fully determined by the degree of polynomials P and number of eigenmodes M according to the well-known relation $R = (M + P)!/(M!P!)$.

and once for the FE simulation using the same number of eigenmodes as in the constructed PCE (solid lines). In other words, the solid lines represent the error induced by PC approximation and the difference between the solid and corresponding dashed line quantifies the error induced by the KL approximation of the parameter fields.
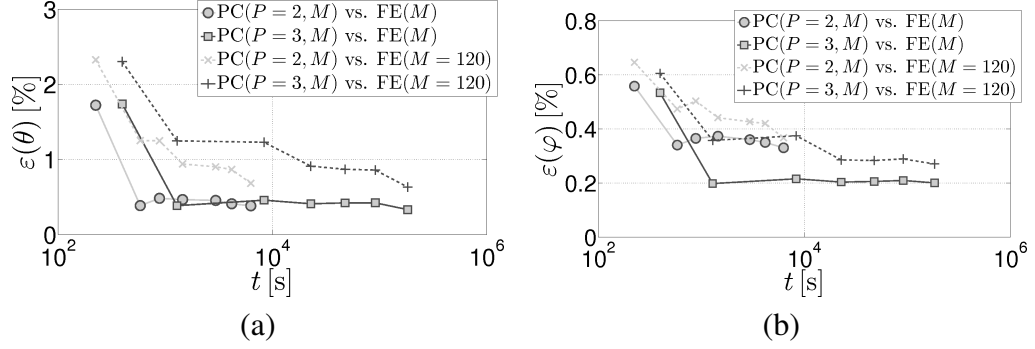
Figure 3: Errors in approximation of the temperature (a) and the moisture (b) field induced by PCE and KLE as a function of computational time needed for a PCE construction.

Figure 3 represents the same errors $\varepsilon(\boldsymbol{u})$ as Fig. 2, but this time with respect to the computational effort needed for the computation of PC coefficients. Regarding the obtained results, we focus our following computations on the KL approximation of the material parameters including $M = 7$ eigenmodes and a PCE of order $P = 2$ providing, at reasonable time, sufficiently good approximation of the model response, namely of the temperature field where the errors are more significant.

For a more detailed presentation of the PCE accuracy, Fig. 4 compares the model response in one node of FE mesh (the node No. 1 at Fig. 5) at the time $t = 400[\mathrm{h}]$ obtained by the FE simulation and by the PCE as a function of the first stochastic variable ξ_1.
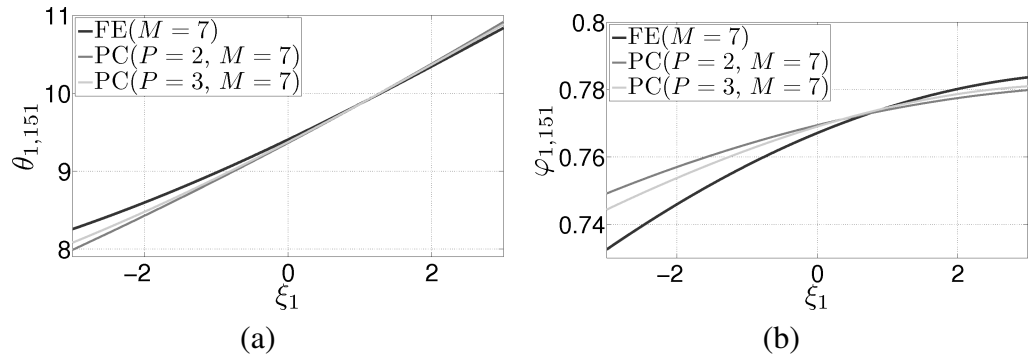
Figure 4: Detailed comparison of the temperature (a) and moisture (b) with their PC approximation as functions of the first stochastic variable ξ_1.

6. Bayesian updating procedure

In the Bayesian approach to parameter identification, we assume three sources of information and uncertainties which should be taken into account. The first one is the prior knowledge about the model/material parameters $q_i(\omega)$ defining the prior density functions. In our particular case, we know that all the identified parameters are positive-definite and

the log-normal random fields with the statistical moments given in Tab. 1 are suitable for a description of the prior information. In fact, they are maximum entropy distributions for this case. We describe the material parameters using the KLE which is fully defined by a finite set of standard Gaussian variables $\boldsymbol{\xi} = [\xi_1(\omega), \ldots, \xi_M(\omega)]$ with the probability density function (pdf) $p_{\boldsymbol{\xi}}(\boldsymbol{\xi})$ and thus, the updating procedure can be performed in terms of $\boldsymbol{\xi}$ turning them into non-Gaussian variables.

Other source of information comes from measurements, which are violated by uncertain experimental errors $\epsilon(\bar{\omega})$. Last uncertainty $\bar{\bar{\omega}}$ arises from imperfection of the numerical model, when for example the description of a real system does not include all important phenomena. However, it is a common situation that the imperfection of the system description cannot be distinguished from measurement error ϵ and the modelling uncertainties $\bar{\bar{\omega}}$ can be hidden inside the measuring error $\epsilon(\bar{\omega})$. Then we can define the pdf $p_{\boldsymbol{z}}(\boldsymbol{z})$ for noisy measurements $\boldsymbol{z}(\bar{\omega})$.

Bayesian update is based on the idea of Bayes' rule defined for probabilities. Definition of Bayes' rule for continuous distribution is, however, more problematic and hence [6, Chapter 1.5] derived the posterior state of information $\pi(\boldsymbol{\xi}, \boldsymbol{z})$ as a conjunction of all information at hand

$$\pi(\boldsymbol{\xi}, \boldsymbol{z}) = \kappa p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) p_{\boldsymbol{z}}(\boldsymbol{z}) p(\boldsymbol{z}|\boldsymbol{\xi}) \,, \tag{31}$$

where κ is a normalization constant.

The posterior state of information defined in the space of model parameters $\boldsymbol{\xi}$ is given by the marginal pdf

$$\pi_{\boldsymbol{\xi}}(\boldsymbol{\xi}) = \mathbb{E}_{\bar{\Omega}}\left[\pi(\boldsymbol{\xi}, \boldsymbol{z})\right] = \kappa p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) \int_{\bar{\Omega}} p(\boldsymbol{z}|\boldsymbol{\xi}) p_{\boldsymbol{z}}(\boldsymbol{z}) \, \mathbb{P}(\mathrm{d}\bar{\omega}) = \kappa p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) \, L(\boldsymbol{\xi}), \tag{32}$$

where $\bar{\Omega}$ is a set of random elementary events $\bar{\omega}$ and measured data $\boldsymbol{z}$ enters through the *likelihood function* $L(\boldsymbol{\xi})$, which gives a measure of how good a numerical model is in explaining the data $\boldsymbol{z}$.

The most general way of extracting the information from the posterior density $\pi_{\boldsymbol{\xi}}(\boldsymbol{\xi})$ is based on sampling procedure governed by MCMC method. For more details about this approach to Bayesian updating of uncertainty in description of couple heat and moisture transport we refer to [11]. In this paper, we focus on the comparison of the posterior information obtained from the sampling procedure using directly the computationally exhaustive numerical model (16) on one hand and using the PC approximation of the model (25) on the other hand.

7. Numerical results for the Bayesian update

Due to the lack of experimental data, we prepared a virtual experiment using a FE simulation based on parameter fields obtained by the KLE with 7 eigenmodes so as to avoid the error induced by KLE, which is mainly the subject of the work presented in [11]. A related set of random variables $\boldsymbol{\xi}$ is drawn randomly from the prior distribution and stored for a purpose of latter comparison with the prior and the posterior state of knowledge. The resulting temperature and moisture fields considered as a so-called "true state" or simply the "truth" are shown in Fig. 5. According to [11] the values of temperature and moisture are measured in 14 points (see Figs. 5 (a) and (c)), and at three distinct times (see Figs. 5 (b) and (d)). Hence, the observations $\boldsymbol{z}$ consist of 84 values. To keep the presentation of the different
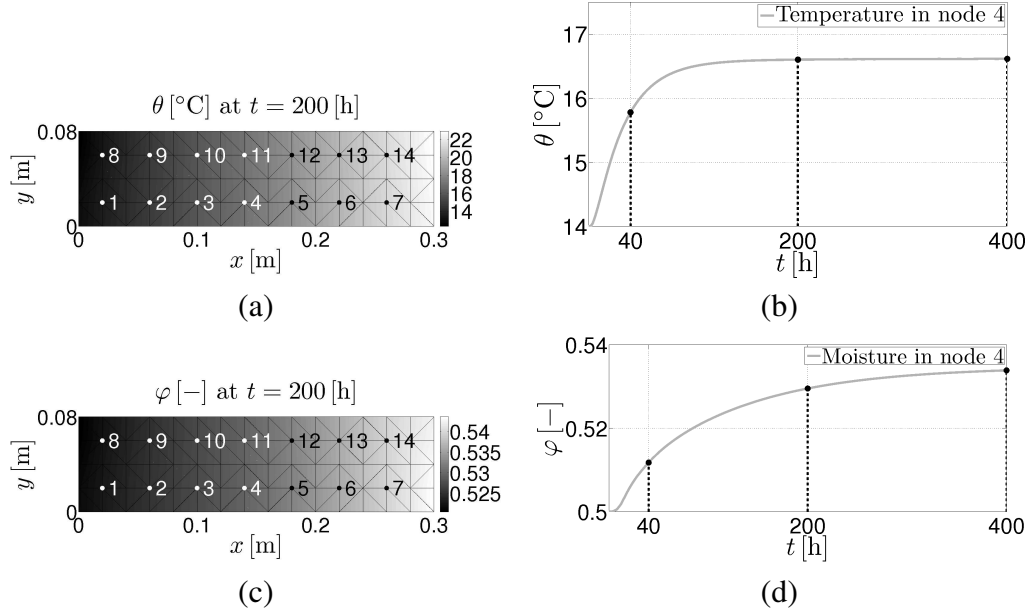
Figure 5: Virtual observations: (a) and (c) spatial arrangement of probes; (b) and (d) temporal organization of measurements

numerical aspects of the presented 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

$$L(\boldsymbol{\xi}) = \kappa \exp\left(-\frac{1}{2}\left(\boldsymbol{Y}(\boldsymbol{\xi}) - \boldsymbol{z}\right)^{\mathrm{T}} \mathbf{C}_{\mathrm{obs}}^{-1} \left(\boldsymbol{Y}(\boldsymbol{\xi}) - \boldsymbol{z}\right)\right), \tag{33}$$

where $\boldsymbol{Y}(\boldsymbol{\xi})$ is an observation operator mapping the model response $\boldsymbol{u}$ given parameters $\boldsymbol{\xi}$ and loading $\boldsymbol{f}$ to observed quantities $\boldsymbol{z}$. $\mathbf{C}_{\mathrm{obs}}$ is a covariance matrix representing the uncertainty in experimental error, which is obtained by perturbing the virtual observations by Gaussian noise with standard deviation for temperature $\sigma_\theta = 0.2\ [^\circ\mathrm{C}]$ and for moisture $\sigma_\varphi = 0.02\ [-]$ so as to get 100 virtual as an input for the covariance matrix evaluation. In order to be able to compare the posterior state with the true state also in terms of model parameters $\boldsymbol{\xi}$, we assume an artificial situation where the observed quantities $\boldsymbol{z}$ correspond exactly to the true state of temperature and moisture.

The Bayesian update was performed using Metropolis-Hasting algorithm and $100,000$ samples were generated in order to sample the posterior density (31) over the variables $\boldsymbol{\xi} = (\xi_1 \ldots \xi_{M=7})$. The truth state, prior and posterior pdfs obtained by the FE simulations and using the PCE are plotted in Fig. 6. One can see that the error induced by PC surrogate of model response are negligible in terms of the resulting posterior densities. Figure 6 also demonstrates the fact that the variables $\boldsymbol{\xi}$ being a priori standard Gaussian should not be a posteriori Gaussian.

During the sampling procedure, we stored also the corresponding values of parameter fields and response fields in order to obtain their posterior state of information. As a result, Fig. 7 shows the comparison of the truth, and prior and posterior pdfs for two material parameters λ_0 and μ in the top-right corner FE element, and similarly, Fig. 8 presents pdfs for the temperature and moisture in FE node 7 at $400[\mathrm{h}]$ (i.e. at the 151-th time step).

We should note that the similarity of the prior and the posterior pdfs for moisture in
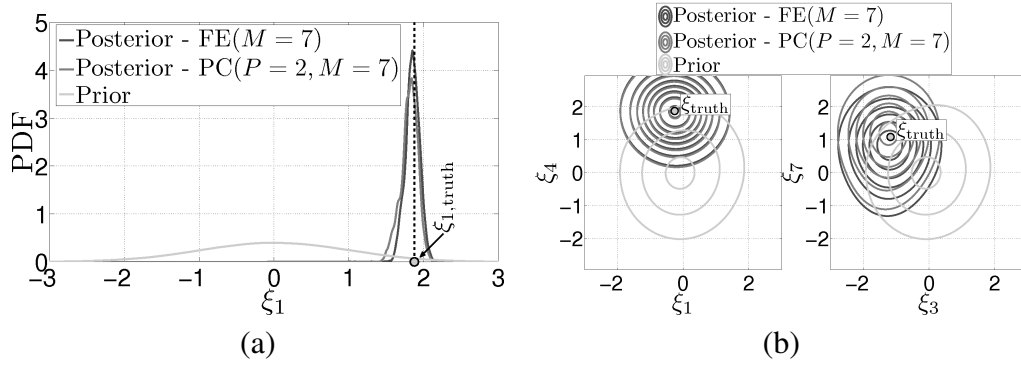
Figure 6: Comparison of pdfs (a) for the separate variable ξ_1 and (b) for the pairs of variables.
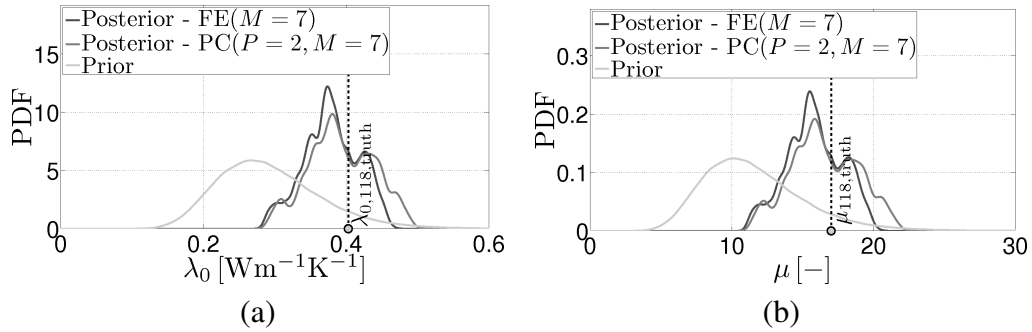
Figure 7: Comparison of pdfs for material properties in FE node 7: (a) the thermal conductivity of dry material λ_0 and (b) water vapour diffusion resistance factor μ.
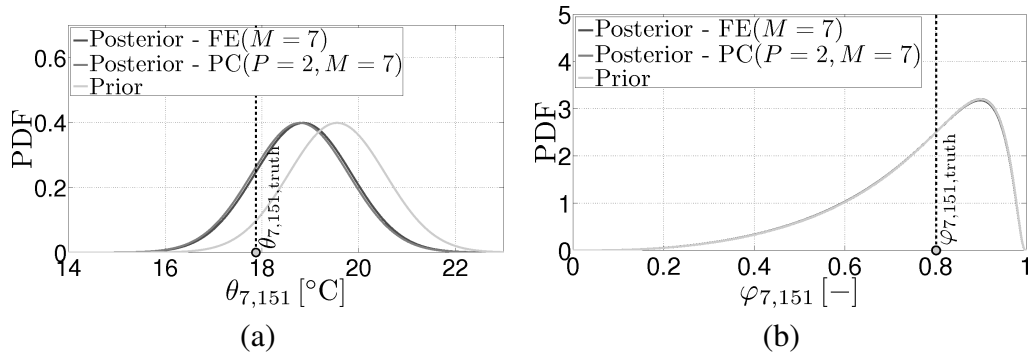
Figure 8: comparison of pdfs (a) for the temperature and (b) for the moisture in FE node 7 at 151-th time step.

Fig. 8 is probably caused by the very slight influence of the studied material parameters to the moisture value or more precisely, the prior standard deviations were very small.

Beside the comparison of the PCE accuracy, we also compared the time necessary to generate the samples. In case of PCE, the total time also includes the time of PC coefficients computation. Particular comparison of computational time needed by FE simulations and by PCE evaluations is demonstrated in Fig. 9.
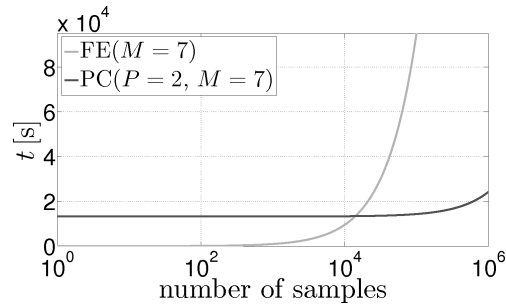
Figure 9: Comparison of time necessary for evaluation of samples.

8. Conclusions

The presented paper presents an efficient approach to propagation and updating of uncertainties in description of coupled heat and moisture transport in heterogeneous material. In particular, we employed the Künzel's model, which is sufficiently robust to describe real-world materials, but which is also highly nonlinear, time-dependent and is defined by 8 material parameters difficult to be estimated from measurements. The updating procedure starts with the prior information about the parameters' properties such as positive-definitness and second order statistics. Heterogeneity of the material under the study is taken into account by describing the material properties by random fields, which are for a simplicity considered as fully correlated. Then, the corresponding correlation lengths are assume to be known as another a priori information. In order to limit the number of random variables necessary to describe the material, the random fields are approximated by Karhunen-Loève expansion and hence, all the remaining uncertainties are described by a set of standard Gaussian variables whose number is given by the number of eigenmodes involved in KLE.

These uncertainties are then propagated through the numerical model so as to provide a probabilistic characterization of the model response, here the moisture and temperature fields. Simultaneously, the other information including uncertainties coming from the experimental measurements is used to update the prior uncertainties in the model parameters. In order to imitate the experimental measurements, a virtual experiment is prepared together with the relating uncertainties given by a covariance matrix. The Markov Chain Monte Carlo method is then employed so as to sample the posterior state of information.

The primary objective of the presented paper is to accelerate the sampling procedure. To this goal, a polynomial chaos-based approximation of the model response is constructed in order to replace computationally expensive FE simulations by fast evaluations of the PCE during the sampling. In particular, the PC coefficients are obtained by an intrusive stochastic Galerkin method. It is shown that the resulting approximations exhibit high accuracy and the related posterior probability density functions are sufficiently precise as well. Finally, the comparison of the computational effort confirmed the large savings in case of PC evaluations.

While the acceleration obtained by the presented procedure is significant, it can be still unfeasible for very large problems. Our future work will be focused on the elimination of the MCMC sampling procedure itself by the update directly in terms of parameters of probability density functions as proposed in [36].

[1] J. Zeman, J. Novák, M. Šejnoha, J. Šejnoha, Pragmatic multi-scale and multi-physics analysis of Charles Bridge in Prague, Engineering Structures 30 (11) (2008) 3365–3376.

[2] R. Černý, P. Rovnaníková, Transport Processes in Concrete, London: Spon Press, 2002.

[3] H. Künzel, K. Kiessl, Calculation of heat and moisture transfer in exposed building components, International Journal of Heat Mass Transfer 40 (1997) 159–167.

[4] J. Sýkora, M. Šejnoha, J. Šejnoha, Homogenization of coupled heat and moisture transport in masonry structures including interfaces, Appl. Math. Comput.doi:10.1016/j.amc.2011.02.05.

[5] A. Kučerová, Identification of nonlinear mechanical model parameters based on soft-computing methods, Ph.D. thesis, Ecole Normale Supérieure de Cachan, Laboratoire de Mécanique et Technologie (2007).

[6] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, 2005.

[7] K. Mosegaard, A. Tarantola, Probabilistic Approach to Inverse Problems, Academic Press, 2002, pp. 237–265.

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

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

[10] W. R. Gilks, S. Richardson, D. Spiegelhalter, Markov Chain Monte Carlo in Practice, Chapman & Hall/CRC, 1995.

[11] A. Kučerová, J. Sýkora, Uncertainty updating in the description of coupled heat and moisture transport in heterogeneous materials, Appl. Math. Comput.doi:10.1016/j.amc.2011.02.078.

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

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

[14] R. Ghanem, P. D. Spanos, Stochastic finite elements: A spectral approach, second revised Edition, Dover Publications, Mineola, New York, 2003.

[15] I. Babuška, R. Tempone, G. E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM Journal on Numerical Analysis 42 (2) (2004) 800–825.

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

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

[18] A. Keese, Numerical solution of systems with stochastic uncertainties, Ph.D. thesis, Institute of Scientific Computing, Department of Mathematics and Computer Science, Technische Universität Braunschweig (2004).

[19] R. Ghanem, Analysis of stochastic systems with discrete elements, Ph.D. thesis, Rice University (1989).

[20] H. G. Matthies, Encyclopedia of Computational Mechanics, John Wiley & Sons, Ltd., 2007, Ch. Uncertainty Quantification with Stochastic Finite Elements.

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

[22] A. Keese, H. G. Matthies, Numerical methods and Smolyak quadrature for nonlinear stochastic partial differential equations, Informatikbericht 2003-5, Institute of Scientific Computing, Department of Mathematics and Computer Science, Technische Universität Braunschweig, Brunswick (2003).

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

[24] H. M. Künzel, Simultaneous heat and moisture transport in building components, Tech. rep., Fraunhofer IRB Verlag Stuttgart (1995).

[25] B. V. Rosić, H. G. Matthies, M. Živkoviić, A. Ibrahimbegović, Stochastic plasticity- a variational and functional approximation approach i: The small strain case, Tech. rep., Institute of Scientific Computing, TU Braunschweig (2011).

[26] B. Rosić, H. G. Matthies, Computational approaches to inelastic media with uncertain parameters, Journal of the Serbian Society for Computational Mechanics 2 (1) (2008) 28–43.

[27] O. Krischer, W. Kast, Die wissenschaftlichen Grundlagen der Trocknungstechnik, Dritte Auflage, Berlin: Springer, 1978.

[28] K. Kiessl, Kapillarer und dampfförmiger Feuchtetransport in mehrschichtlichen Bauteilen, Ph.D. thesis, Universität in Essen (1983).

[29] R. Černý, J. Maděra, J. Kočí, E. Vejmelková, Heat and moisture transport in porous materials involving cyclic wetting and drying, in: Computational Methods and Experimental Measurements XIV, Vol. 48 of WIT Transactions on Modelling and Simulation, 2009, pp. 3–12.

[30] Z. Pavlík, J. Mihulka, J. Žumár, R. Černý, Experimental monitoring of moisture transfer across interfaces in brick masonry, in: Structural Faults and Repair, 2010.

[31] J. Sýkora, Multiscale modeling of transport processes in masonry structures, Ph.D. thesis, Czech Technical University in Prague, Faculty of Civil Engineering, Department of Mechanics (2010).

[32] B. N. Khoromskij, A. Litvinenko, Data sparse computation of the Karhunen-Loève expansion, in: Proceedings of International Conference on Numerical Analysis and Applied Mathematics 2008, Vol. 1048, 2008, pp. 311–314.

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

[34] D. Xiu, G. E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.

[35] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Math. Dokl. 4 (1963) 240–243.

[36] B. Rosić, A. Litvinenko, O. Pajonk, H. G. Matthies, Direct bayesian update of polynomial chaos representations, Journal of Computational PhysicsSubmitted for publication.

Robust Probabilistic Calibration of Stochastic Lattice Discrete Particle Model for Concrete

Eliška Janouchová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Anna Kučerová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Jan Sýkora

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Jan Vorel

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Roman Wan-Wedner

Magnel Laboratory for Concrete Research, Ghent University, Gent, Belgium

Abstract

Numerical modelling of quasi-brittle materials arising from lattice or particle formulations is based on a priori discretisation of a medium according to an idealization of its granularity. This paper concentrates on the so-called Lattice Discrete Particle Model (LDPM) which provides more accurate modelling of damage initiation and crack propagation at various length scales. However, its simulations are very computationally demanding. In this paper, we propose an automated identification procedure allowing for its widespread utilisation without the necessity of the deep expert knowledge of the model details. Such an automated procedure is complicated namely due to stochasticity of the LDPM related to the random generation of the particle configuration. The particle size distribution is generated to be statistically corresponding to prescribed concrete granulometric distributions, but each realisation of particle configuration is created at random. The robustness and stability of the proposed identification procedure are obtained by probabilistic Bayesian formulation, where the time requirements are kept reasonable thanks to employed polynomial chaos approximation of the LDPM. The Bayesian formulation solves such an inverse problem as well-posed and provides a quantitative assessment of the underlying uncertainty in the identified values of material properties. In particular, the procedure is applied to identify seven material parameters from an unconfined compression cube test and notched three-point bending test. Its efficiency is verified on the synthetic data with known parameter values to quantify the accuracy of the estimates and it is also validated on experimental data to prove its robustness.

1. Introduction

Lattice or particle formulations of models are suitable for the description of the macroscopic behaviour of quasi-brittle materials such as concrete due to their potential to capture the material morphology and its impact on material macroscopic behaviour. Material morphology or granularity is defined by a set of geometrical parameters (or granulometric distributions) which govern the generation of the lattice and thus equip this type of formulations with inherent characteristic lengths. This allows more realistic simulations of a damage initiation and crack propagation at various length scales, but the related computational demands of model simulations are very high (in order of hours for standard experimental setup, i.e. three-point bending test or cube compression test), which prohibits an easy hand-fitting-based calibration of the model from experimental data. In particular, in this paper, we employ the so-called Lattice Discrete Particle Model (LDPM) proposed by [1], which was calibrated and validated against both quasi-static and dynamic loading conditions and it was demonstrated to possess superior predictive capability, see [2]. The stochastic response of the model caused by the random generation of the particle configuration is, however, another burden for the calibration process. Calibration procedure employed for model validation in [2] was based on hand-fitting and deep expert knowledge of the model. But the necessity of such knowledge makes the further practical applications of the model almost impossible or at least impractical.

Combination of trial-and-error method and automated calibration of the model is presented in [3]. The main drawback of this identification strategy, however, consists in consideration of a fixed position of grains thus disabling the model stochasticity. Such simplification significantly facilitates and accelerates the identification process, but it cannot be applied in a general manner. The authors in [3] justify this simplification by a selection of a specific experimental setup with a deep notch, where the strong stress concentration minimizes the effect of model stochasticity. Oppositely to this approach, the goal of this paper is to present an identification strategy allowing an automatisation, while preserving the applicability to any commonly used experimental setups, such as cube compression test employed herein. In order to reduce the effect of model stochasticity and to accelerate identification process, we introduce the model approximation based on polynomial chaos (PC) expansion [4, 5, 6, 7] with an averaging effect. In other words, we aim to construct the PC surrogate using a limited set of model simulations for selected values of material parameters as well as for various randomly generated meso-structures. While the material parameters enter the PC expansion as input variables, randomness in meso-structure geometry is smoothed out by keeping low polynomial order of the approximation to avoid its random geometry-induced oscillations.

The authors in [8] developed an artificial neural network trained to mimic the recommendations in building codes for predicting three fracture and shear strength parameters of LDPM based on three concrete macroscopic properties: compressive strength, maximum size of aggregates and initial fracture energy. The results obtained within validation on experimental data from uniaxial compression (UC) and notched three-point bending (TPB) tests show good performance of the proposed procedure, however, it reveals its high sensitivity to input values of initial fracture energy and its non-trivial estimation. Our aim is

to propose the identification procedure, which does not require any deep insight into the model detailed formulation or knowledge of some other macroscopic properties, which are not easy to determine. Therefore, we want to consider the measured load-deflection curves as input values without the necessity of any other complicated preprocessing of the data before the use of the identification method. So as to fulfil this goal, the procedure needs to be sufficiently robust and numerically stable for a large variety of concretes characterised by reasonably large feasible intervals for values of material parameters. Such requirements lead us to Bayesian inference [9, 10, 11, 12], which formulates the inverse problem as well-posed not only for limited and noisy experimental data but also in this case of stochastic model response. This general probabilistic formulation of an inverse problem combines the available information about the parameter values by means of the related uncertainty to each information source. Here, we can thus consider three relevant sources of uncertainty: (i) in the prior expert knowledge and recommendation about the realistic values of parameters, (ii) in the model randomness leading to uncertainty about the assessment of the mean model response for a given set of values of parameters and (iii) uncertainty in the experimental observations caused by measurement errors. By combining these three sources of information and related uncertainty we obtain a joint posterior probability distribution of material parameters quantifying our resulting uncertainty in estimated values of parameters. The probabilistic description of estimates of parameters thus allow to quantify, whether the information contained in the selected experimental data is sufficient to determine the parameters of such a stochastic model with a satisfactory accuracy.

In order to demonstrate the advantages and drawbacks of the proposed identification strategy, we apply it to the LDPM calibration utilising notched three-point bending tests and cube compression tests. In particular, seven material parameters are to be identified: normal modulus, shear-normal coupling, tensile strength, tensile characteristic length, softening exponent, shear/strength ratio and initial friction. The strategy is verified on ten synthetic data sets, where the true values of parameters are known and then validated on one set of real experimental data, where each both experiments were performed in three repetitions on six samples of the same material. The material properties are thus assumed to attain the same values in all samples.

Particular formulation of an expert knowledge about the identified values of material parameters in terms of prior probability density functions along with the description of the both considered laboratory experiments is presented in the following section. The identification strategy based on Bayesian inference is described in Section 3. The acceleration of the identification process based on the polynomial chaos expansion is described in Subsection 3.1 along with the formulation of estimated theoretical uncertainty in the modelling of the mean LDPM response. Resulting formulation of the posterior uncertainty in the estimates of material parameters combining prior knowledge, theoretical uncertainty due to model stochasticity and the uncertainty in data caused by measurement errors is presented in Subsection 3.2. Verification of the proposed identification strategy is presented in Section 4 for synthetic data simulated for ten different concrete mixes with known material parameters. The proposed LDPM calibration is finally validated on real experimental data in Section 5 and the concluding remarks are summarized in Section 6.

2. Concrete behaviour

This section is focused on the material model description and introduction of the employed experimental tests which are typically performed to characterize material properties of concrete.

2.1. Lattice discrete particle model

The examined numerical model, which is capable of an accurate description of the macroscopic behaviour of concrete in different loading scenarios, falls into the category of models which are based on lattice or particle formulations in which materials are discretized "a priori" according to an idealization of their internal structure. Particle size and size of the contact area among particles, for particle models, as well as lattice spacing and cross-sectional area, for lattice models, introduce inherent characteristic lengths into this type of formulations. Therefore, lattice/particle models are able to simulate the geometrical features of material internal structure more authentically.

Earlier attempts to formulate particle and lattice models for fracture are reported in [13, 14, 15, 16, 17, 18]. A comprehensive discrete formulation for concrete was proposed by Cusatis and coworkers [1, 2] who formulated the so-called Lattice Discrete Particle Model (LDPM). This model describes a granular microstructure through a system of polyhedral particles connected through a three-dimensional lattice. Such particles are placed randomly across the volume following a prescribed grain size distribution, thus enabling the direct representation of a heterogeneous system of grains surrounded by a bonding agent. Particle contact behaviour represents the mechanical interaction among adjacent aggregate particles through the embedding mortar. More specifically, the used constitutive law is based on the strains and stresses defined in terms of one normal and two shear components on each facet, the interested reader is referred to [1] for detailed formulation. LDPM was calibrated and validated against a large variety of loading conditions in both quasi-static and dynamic loading conditions [19, 20, 21]. The material parameters of the models can be divided into two groups: (i) parameters characterising the concrete mix which are utilised for the generation of concrete mesostructure; (ii) parameters used for the description of material behaviour (e.g. elasticity, fracture, etc.).

In the present study, the basic properties of the tested concrete mix are fixed at values listed in Table 1. Note that in a laboratory the concrete mix is prepared according to these

Table 1: Values of parameters governing the generation of concrete meso-structure

Material property		Unit	Value
Minimum particle size	d_0	mm	4
Maximum particle size	d_a	mm	16
Cement content	c	kg/m^3	240
Water to cement ratio	w/c	-	0.7
Aggregate to cement	a/c	-	8.83
Fuller coefficient	n_F	-	0.5
Concrete density	ρ	kg/m^3	2400

prescribed parameters and their values can be used further for the generation of a corresponding concrete meso-scale structure. The seed number governing the sampling of cumulative

distribution function of concrete granulometric distributions by a pseudo-random number generator is kept pseudo-random. The model response is thus depending on the specific distribution of particles which causes stochasticity of the model. As experimental tests employed for the model calibration, an unconfined compression test and notched three-point bending test are considered as the most commonly used laboratory tests for testing material properties of concrete. In particular, seven material parameters are identified, two additional parameters (compressive strength and transitional stress) are prescribed to be proportional to the tensile strength and other seven material parameters of LDPM are prescribed to values presented in [2]. To properly identify the latter parameters, more advanced experimental test data would be needed (hydrostatic compression, triaxial compression, etc.). These tests are, however, not trivial, financially expensive and are not typically available. Therefore, two aforementioned concrete tests are employed in this study. All material parameters of the model along with their prescribed values or prior uncertain knowledge in terms of feasible intervals (interpreted further as uniform probability density functions) are summarised in Table 2. Note that the intervals were defined based on the authors' experience corresponding to the normal strength concrete and according to the parameters presented in [2].

2.2. Experimental tests

Two types of standardized tests for concrete are utilised in the paper, i.e. cube compression and three-point bending tests. These two tests are usually available and carried out when the material properties of concrete are in demand. More specifically, the standard cube compression test (with a side length of $150\,\text{mm}$), tested according to [22], is employed. All experiments were performed with a rate of loading equal to $0.008\,\text{mm/s}$. The three-point bending test was performed on notched prisms with dimensions of $100 \times 100 \times 400\,\text{mm}$, a relative notch depth of 30% and a span of $300\,\text{mm}$. These measurements were performed using

Table 2: Values of material model parameters used in the numerical simulations

Material property		Unit	Value/Prior bounds
Normal modulus	E_0	MPa	$20000 - 70000$
Shear-normal coupling	α	-	$0.2 - 0.3$
Tensile strength	σ_t	MPa	$1.5 - 5$
Tensile characteristic length	l_t	mm	$50 - 300$
Softening exponent	n_t	-	$0.1 - 1$
Shear/strength ratio	σ_s/σ_t	-	$1.5 - 8$
Initial friction	μ_0	-	$0.001 - 0.5$
Compressive strength	σ_{c0}	MPa	$\sigma_{c0} = 40\sigma_t$
Transitional stress	σ_{N0}	MPa	$\sigma_{N0} = 240\sigma_t$
Initial hardening modulus ratio	$H_{c0}/E0$	-	0.4
Transitional strain ratio	κ_{c0}	-	4
Deviatoric strain threshold ratio	κ_{c1}	-	1.0
Deviatoric damage parameter	κ_{c2}	-	5.0
Asymptotic friction	μ_∞	-	0.0
Densification ratio	E_d/E_0	-	1.0
Volumetric-deviatoric coupling	β	-	0

the servo-hydraulic testing machine Walter & Bai LFV-63 kN. To ensure stable tests and especially post-peak response, the tests were controlled by crack mouth opening displacement (CMOD). A constant loading rate equal to 0.0001 mm/s was applied using an extensometer of type Epsilon 3542-050M-025-HT2. Experimental data are available for three repetitions of both experimental types.

For later use, the measured experimental data are post-processed and nominal values are obtained. As mentioned above, the typical compression test is performed on cubes of $a = 150$ mm edge length with the nominal stress σ_{cN} and nominal strain ε_{cN} defined as

$$\sigma_{cN} = \frac{F}{a^2} \;\; \text{and} \;\; \varepsilon_{cN} = \frac{u}{a}, \tag{1}$$

where F is the applied load and u denotes the load point displacement. In three-point bending test the nominal values of stress and strain take the form

$$\sigma_{bN} = \frac{3Fl}{dh^2} \;\; \text{and} \;\; \varepsilon_{bN} = \frac{CMOD}{h}, \tag{2}$$

where l stands for the beam span and h, d are the beam height and width, respectively. $CMOD$ is the crack mouth opening displacement measured over the notch.

Moreover, to reduce the measurement error caused by the testing machine stiffness, the elastic part of the stress-strain diagram is eliminated from the observed data in both loading tests and the inelastic strains are then obtained according to

$$\varepsilon_{\xi N}^{\text{inel}} = \varepsilon_{\xi N} - \sigma_{\xi N}/K_{\xi N}, \tag{3}$$

where $\xi = c, b$ for compression and three-point bending tests, respectively. $K_{\xi N}$ is the elastic stiffness estimated as a slope of elastic part of the stress-strain diagram. Furthermore, the value of elastic stiffness K_{cN}, obtained from the compression test, is then employed as an additional observation. This value is less perturbated by measurement error than an elastic part of the three-point bending test. The slope is obtained by using the linear regression on the initial part of the stress-strain diagram corresponding to the range from 20 to 60 % of the peak load.

3. Bayesian inference

Since the computational capacity and stochastic modelling are sufficiently developed to be used for parameter identification in a probabilistic way, many researchers in different fields have taken into account uncertainties to obtain more information about the parameter estimates. Probabilistic parameter estimation, allowing an update of the initial knowledge about the parameter values by processing the experimental data, is based on Bayes' formula [9, 10, 11]. The Bayesian identification method consists of modelling unknown parameters $\boldsymbol{x}$ as well as observations $\boldsymbol{z}$ as random variables. The combination of the prior probability distribution $p(\boldsymbol{x})$ and the likelihood function $p(\boldsymbol{z}|\boldsymbol{x})$ results in the posterior probability distribution of parameters $p(\boldsymbol{x}|\boldsymbol{z})$ expressing uncertainties about the identified values as

$$p(\boldsymbol{x}|\boldsymbol{z}) = \frac{p(\boldsymbol{z}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{z})} = \frac{p(\boldsymbol{z}|\boldsymbol{x})p(\boldsymbol{x})}{\int_{\boldsymbol{x}} p(\boldsymbol{z}|\boldsymbol{x})p(\boldsymbol{x})\mathrm{d}\boldsymbol{x}}. \tag{4}$$

A Bayesian framework for the model calibration is well-suited namely for situations where only a few experimental observations are available. In case of stress-strain diagrams observed in terms of many discrete points, it is questionable, whether each observed point can be considered as an independent measurement or how to build a likelihood function corresponding to the joint probability distribution of observed and possibly correlated values. Underestimating of the correlation among particular data leads to the overestimation of their information content and vice versa. More information in data is connected to their higher significance and more pronounced reduction of the resulting uncertainty in parameter estimates. Considered data correlation has of course only the magnifying effect to the posterior uncertainty reduction. For example, if the data contain no information allowing to distinguish two model parameters, this lack of information in the data cannot be compensated by the underestimated correlation of the data. On the other hand, high overestimation of the data correlation can in the extreme situation lead to significant information loss. From the computational point of view, considering a high number of observations corresponding to very dense discrete points of stress-strain diagram may simply lead to increased computational costs or numerical difficulties connected to highly sharp likelihood function representing very low remaining uncertainty caused by measurement errors. Unfortunately, there is no strict and clear rule, how to properly model the uncertainty and information in data for each particular situation. Moreover, particular selection among the observed data points and neglecting some observations can be also driven by the expert knowledge about the significance of particular parts of the stress-strain diagram. Considering the higher density of data points has then similar effect as considering smaller uncertainty caused by measurement error.

Here we decided to reduce the experimental data, consisting of $8,500 - 10,500$ discrete points in case of compression tests and $14,000 - 18,000$ discrete points in case of three-point bending tests, to stresses corresponding to $n_\sigma = 200$ equisized strain steps in each experiment obtained by a linear interpolation among the original data points. We are convinced that the resulting data points are still very dense and we do not lose any relevant information, while a smaller number of observations significantly facilitates our implementation of model approximation described in the following section. More specifically, the strain range from 0 to 0.0293 is used for the compression test and 0 to 0.0265 for the three-point bending test. The simulated stress-strain diagrams are then also linearly interpolated to match the experimental points. Overall, the considered model response consists thus of $n_y = 401$ components and it is defined as

$$\boldsymbol{y}(\boldsymbol{x}) = \{\boldsymbol{\sigma}_{cN}(\boldsymbol{x}), \boldsymbol{\sigma}_{bN}(\boldsymbol{x}), K_{cN}(\boldsymbol{x})\}, \tag{5}$$

where $\boldsymbol{\sigma}_{cN}(\boldsymbol{x})$ is a vector of n_σ stress values obtained in compression test and $\boldsymbol{\sigma}_{bN}(\boldsymbol{x})$ is a vector of n_σ stress values obtained in three-point bending test. $\boldsymbol{x}$ is a vector of $n_x = 7$ material parameters, specifically, $\boldsymbol{x} = \{E_0, \alpha, \sigma_t, l_t, n_t, \sigma_s/\sigma_t, \mu_0\}$.

3.1. Surrogate model

As mentioned in previous sections, the LDPM simulations of compression and three-point bending tests are highly time-demanding. We run the LDPM simulations on a Vienna Scientific Cluster[1] using one node with 16 cores, where the running time is approx. 3 hours

---

for a single compression test and approx. 3-4 hours for a single three-point bending test. Such computational cost is caused by both the high number of particles and corresponding facets (approx. 280,000 for uniaxial compression and 320,000 for three-point bending tests, respectively) used for the evaluation of the mechanical interaction, and small time steps used by the explicit solver.

To make the probabilistic identification process feasible, some surrogate model of the LDPM have to be employed. In order to consider the uncertainty in material parameters $\boldsymbol{x}$, we model them as random variables defined over some probability space $(\Omega, \mathfrak{A}, \mathbb{P})$, where Ω is the basic probability set of elementary events, $\mathfrak{A}$ a σ-algebra of subsets of Ω, and $\mathbb{P}$ a probability measure. Because the prior knowledge of the values of searched parameters is expressed in terms of feasible intervals given in Table 2, we model them as uniformly distributed independent random variables. Therefore, we choose the approximation of model responses $\boldsymbol{y}(\boldsymbol{x})$ based on the polynomial chaos expansion (PCE) [4, 5, 10] and in particular on the expansion of Legendre polynomials allowing an easy computation of integrals over the uniform probability space [23]. To keep the surrogate complexity (namely its dimensionality, non-linearity and possible multi-modality) as low as possible, we build one independent PC expansion $\tilde{y}_i(\boldsymbol{x})$ for approximating each response component $y_i(\boldsymbol{x})$ separately as

$$\tilde{y}_i(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}} \boldsymbol{\beta}_{\boldsymbol{\alpha},i}\psi_{\boldsymbol{\alpha}}(\boldsymbol{x}), \tag{6}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{n_x})$ is a vector of n_x non-negative integer components that indicate degrees of multivariate polynomial $\psi_{\boldsymbol{\alpha}}(x_1, \ldots, x_{n_x}) = \psi_{\alpha_1}(x_1) \cdot \ldots \cdot \psi_{\alpha_{n_x}}(x_{n_x})$ with $\psi_{\alpha_j}(x_j)$ being univariate polynomial with a degree α_j. The vector $\boldsymbol{\beta}_{\boldsymbol{\alpha},i}$ is a vector of PC coefficients corresponding to a response component y_i.

Due to the time limitations and a high computational cost, we decided to built the PC-based surrogates using only $n_s = 400$ LDPM simulations of both loading tests. Therefore, the input sets of material parameters are selected by means of stratified Latin hypercube sampling (LHS) which is able to select samples according to the prescribed probability distributions. More specifically, we derive a particular LHS from the Halton sequence, which provides an LHS design which is reasonable for joint statistical representation of the prescribed prior probability distributions of material properties including their mutual independence/non-correlation, as demonstrated e.g. in [24].

Using the obtained LDPM simulations, the PC coefficients $\boldsymbol{\beta}_{\boldsymbol{\alpha},i}$ are computed by means of a linear regression based on the least-squares method (LSM) [6, 7]. Scheme of such a surrogate construction is sketched in Figure 1. Every of the total $n_y = 401$ PC expansions
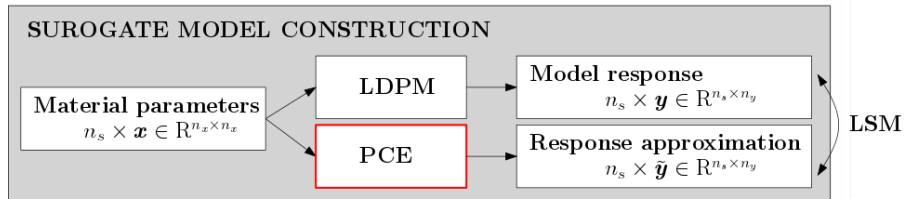
Figure 1: Scheme of surrogate model construction.

is prescribed to have the same degree. The choice of the appropriate degree of the polynomials arises from the performed leave-one-out cross-validation. The Table 3 presents the

corresponding cross-validation errors for different degrees of the polynomials p computed as mean absolute error relative to the span of simulated values obtained for a given response component, i.e.

$$\varepsilon_p = \frac{1}{n_s} \sum_{j=1}^{n_s} \frac{1}{n_y} \sum_{i=1}^{n_y} \frac{|y_i(\boldsymbol{x}_j) - \tilde{y}_i(\boldsymbol{x}_j)|}{\max(\boldsymbol{y}(\boldsymbol{x}_j)) - \min(\boldsymbol{y}(\boldsymbol{x}_j))}, \tag{7}$$

where $\boldsymbol{y}(\boldsymbol{x}_j) = (\ldots, y_i(\boldsymbol{x}_j), \ldots)$ is a vector of n_y response components obtained from the j-th LDPM simulation for values of material parameters $\boldsymbol{x}_j$ and $\tilde{\boldsymbol{y}}(\boldsymbol{x}_j)$ is a vector of their corresponding PC-based approximations. The maximal degree of polynomials is limited by the number of the available model simulations n_s which has to be at least equal to the number of polynomial terms n_p to make the least-squares system of equations solvable. The number of terms n_p in polynomial of the p-th degree and n_x variables is given as follows:

$$n_p = \frac{(p + n_x)!}{p! n_x!}. \tag{8}$$

Of course, the situation when $n_s = n_p$ leads to polynomial exact interpolation among the simulated response values, which is not desired here as we do not want to interpolate the response oscillations due to the particle placement randomicity. Based on this knowledge, we rather prefer a lower degree of the polynomial degree to reduce the undesired oscillations. According to results in Table 3, the second and third degree provide very similar values of the cross-validation error and thus, we decided to choose the lower degree, i.e. we employ the second degree in all our further computations.

Table 3: Cross-validation error and required minimal number of simulations corresponding to the first five degrees of polynomial approximations.

p	1	2	3	4	5
ε_p [%]	4.5217	2.6459	2.5573	7.0772	-
n_p	8	36	120	330	792

Now let us recall that the random generation of particle positions does not allow to exactly represent the real material morphology of one specific example. However, it is assumed that it represents well, from a statistical point of view, a set of concrete specimens. Once we build the model surrogate to eliminate the response variance arising from the randomness of particle positions, we may view the surrogate as an imperfect model unable to capture effects connected to material morphology. Due to the lack of another knowledge, we may consider the difference between the model response $\boldsymbol{y}$ and its approximation $\tilde{\boldsymbol{y}}$ as caused namely by this omitted effect. We thus consider these discrepancies in particular response components as a vector of theoretical model errors

$$\boldsymbol{\varepsilon}_t = \boldsymbol{y} - \tilde{\boldsymbol{y}}, \tag{9}$$

which can be modelled as a vector of correlated random variables. As the surrogate is built to minimise the mean error, see Eq. (7), we may assume $\boldsymbol{\varepsilon}_t$ to have zero mean. In order to keep the computational effort feasible, we estimate their covariance matrix $\mathbf{C}_t$ by using the same n_s samples employed for the surrogate construction. By arranging the sampled responses $\boldsymbol{y}(\boldsymbol{x}_j)$ as the columns of a matrix $\mathbf{Y} = [\boldsymbol{y}(\boldsymbol{x}_1)\ \boldsymbol{y}(\boldsymbol{x}_2)\ \cdots\ \boldsymbol{y}(\boldsymbol{x}_{n_s})]$ with n_s columns and n_y

rows and in the same way their corresponding approximations $\tilde{\boldsymbol{y}}(\boldsymbol{x}_j)$ into the matrix $\tilde{\mathbf{Y}}$, the sample covariance matrix of $\boldsymbol{\varepsilon}_t$ is given as

$$\mathbf{C}_t = \frac{1}{n_s - 1}(\mathbf{Y} - \tilde{\mathbf{Y}})(\mathbf{Y} - \tilde{\mathbf{Y}})^{\mathrm{T}}. \tag{10}$$

We consider this statistical description of theoretical model errors $\boldsymbol{\varepsilon}_t$ to be our best estimate of the response variance and/or the uncertainty arising from the material morphology and we thus employ this uncertainty also in the quantification of the posterior uncertainty of estimates of material parameters as described in the following section.

3.2. Formulation of likelihood function

The likelihood function is defined as a probability distribution of the difference between the available experimental data $\boldsymbol{z}$ and the numerical model response $\boldsymbol{y}$. In our computations, we replace the model response $\boldsymbol{y}$ by its approximation $\tilde{\boldsymbol{y}}$. Then the difference between the experimental data and response approximation written as

$$\boldsymbol{z} = \tilde{\boldsymbol{y}}(\boldsymbol{x}) + \boldsymbol{\varepsilon}_t + \boldsymbol{\varepsilon}_m \tag{11}$$

has two sources of error: (i) the theoretical error $\boldsymbol{\varepsilon}_t$ arising from the inability of the model approximation $\tilde{\boldsymbol{y}}$ to reproduce the effects of particular material morphology and (ii) the measurement errors $\boldsymbol{\varepsilon}_m$ caused by the inaccuracy of experimental measurements. According to the expert knowledge, the measurement errors $\boldsymbol{\varepsilon}_m$ are considered as independent normally distributed random variables with zeros mean values and prescribed standard deviations $\boldsymbol{\sigma}_m$, i.e. $\varepsilon_{m,i} \sim N(0, \sigma_{m,i})$. Particularly, we consider to the best of our knowledge the measurement error in all stress values obtained from the compression test to have the variance $\sigma_{m,cN}^2 = 1.3^2\,\mathrm{MPa}^2$, the measurement error variance in the stress for the three-point bending test to be $\sigma_{m,bN}^2 = 0.3^2\,\mathrm{MPa}^2$ and the measurement error variance in the elastic stiffness K to be $\sigma_{m,K}^2 = 150^2\,\mathrm{MPa}^2$. By arranging these measurement error variances onto the diagonal of an $n_y \times n_y$ matrix in an order corresponding to response components organised in vector $\boldsymbol{y}(\boldsymbol{x})$ and by placing zeros on all off-diagonal terms, we obtain the covariance matrix for measurement errors $\mathbf{C}_m$. The joint probability distribution of measurement errors is thus $\boldsymbol{\varepsilon}_m \sim N(\mathbf{0}, \mathbf{C}_m)$.

If we consider the theoretical errors $\boldsymbol{\varepsilon}_t$ to be also normally distributed according to $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \mathbf{C}_t)$, the likelihood function can be then simply formulated again as multivariate normal distribution as

$$p(\boldsymbol{z}|\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^{n_y}|\mathbf{C}|}}e^{\frac{1}{2}(\tilde{\boldsymbol{y}}(\boldsymbol{x})-\boldsymbol{z})^T\mathbf{C}^{-1}(\tilde{\boldsymbol{y}}(\boldsymbol{x})-\boldsymbol{z})}, \tag{12}$$

where $\mathbf{C} = \mathbf{C}_t + \mathbf{C}_m$. To justify this simplification for the presented example, we plot 1D and 2D marginal probability density functions of theoretical errors for selected response components in Figure 2. Note that the density functions are relatively rough as they are constructed using only 400 error samples obtained from computed LDPM simulations ($n_s = 400$ in current study). Nevertheless, we can conclude that the assumption of the normal distribution of errors is in agreement with these data. It is also clearly visible that the theoretical errors in neighbouring response components are highly correlated and thus this correlation needs to be taken into account.

The updated joint probability distribution of estimates $p(\boldsymbol{x}|\boldsymbol{z})$ is formulated as a product of the likelihood function $p(\boldsymbol{z}|\boldsymbol{x})$ and prescribed uniform prior distribution $p(\boldsymbol{x})$ with bounds
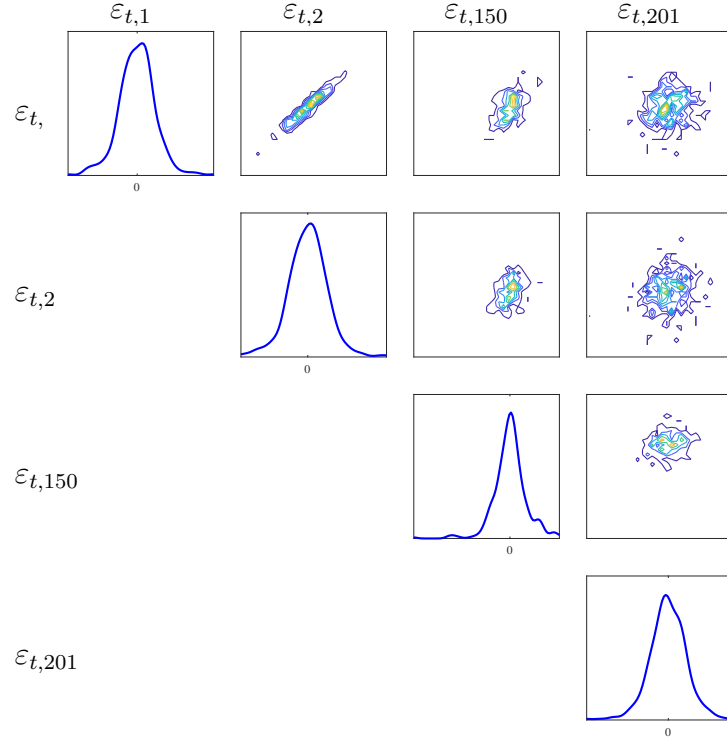
Figure 2: 1D and 2D marginal pdfs of theoretical errors in selected response components: $\varepsilon_{t,1}$, $\varepsilon_{t,2}$ and $\varepsilon_{t,150}$ correspond to the first, second and 150th response components obtained from the compression test and $\varepsilon_{t,201}$ corresponds to the first response component obtained from the three-point bending test.

given in Table 2. Due to the complicated formulation of the posterior distribution, the most general method to obtain its characteristics in terms of statistical moments is Markov chain Monte Carlo (MCMC) sampling [25], where the denominator in Eq. (4) does not need to be evaluated. The MCMC method requires a large number of the posterior probability evaluations involving the model simulations, which is a general obstacle of sampling methods. Nevertheless, we overcome this problem here by substituting the LDPM simulations by fast evaluations of its PCE-based surrogate.

To be more specific, we compute a million of Markov chain steps by using the Metropolis algorithm with symmetrical proposal distribution [26]. A random walk is governed by joint Gaussian distribution of independent steps with zero means and standard deviations equal to 1.75 % of the prior parameters ranges in order to attain acceptance rate between 20 % and 40 %. According to [27], the acceptance rate is recommended to be between 10 % and 60 % for providing good mixing properties of the generated Markov chain. In other words, the proposal standard deviation can not be too high because large steps are frequently rejected and the chain is often stuck, but on the other side too small random move slows down exploring the domain. The appropriate setting of proposal distribution is crucial task for optimising the convergence rate of this sampling method. The starting point of the algorithm is always prescribed in the middle of the prescribed feasible domain and the initial part of the chain, called a burn-in period, is considered in a length of 5.000 samples which are discarded.

4. Verification

To verify the proposed identification method, synthetic experimental data are used, where true values of material parameters are known and the identification method can be evaluated not only based on the quality of the predicted data fit but also in terms of estimated values of material parameters, see Figure 3 for a brief summary of the verification process. Red rectangles highlight the outcomes of such an identification process used for its verification.
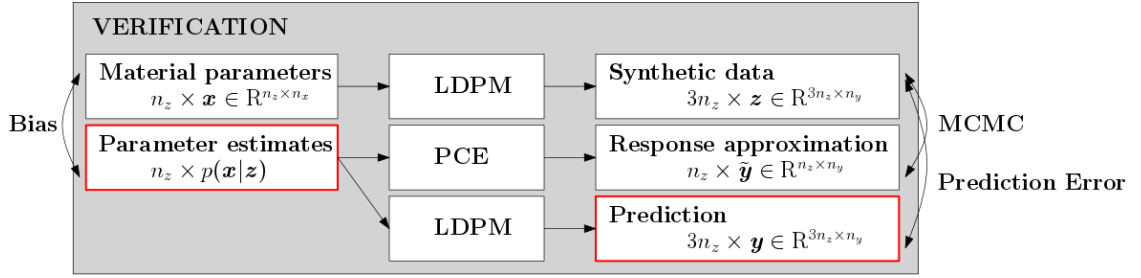
Figure 3: Scheme of verification of the proposed parameter identification strategy on synthetic data.

In particular, ten data sets of model parameters ($n_z = 10$) are utilised to obtain the numerical responses for three different meso-structures (different random particle configurations) of each test. We selected the parameter sets for these simulations again with the help of LHS design derived from the Halton sequence. The obtained synthetic data for both experiments are depicted in Figure 4. The identification procedure is performed for each data set independently and the obtained results, in terms of predicted stress-strain diagrams based on estimated parameter values, are shown in Figure 5. In both Figures 4 and 5, the stress-strain diagrams obtained by the LDPM with the purpose to construct the PC-based model surrogate ($n_s = 400$) are drawn by grey colour in order to demonstrate the range of feasible diagrams corresponding to prescribed prior bounds of material parameters. The data corresponding to one data set are displayed by one colour in both experiments. The solid lines show the synthetic experimental data and the filled areas in Figure 5 correspond to the identification results. More precisely, the displayed areas are centered around a stress-strain diagram $\widehat{\boldsymbol{y}}(\widehat{\boldsymbol{x}})$
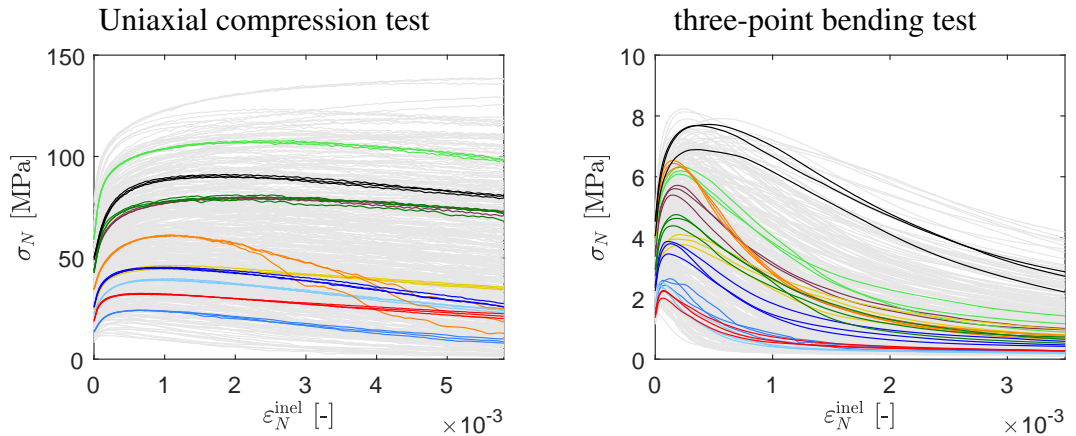
Figure 4: Gray stress-strain diagrams employed for construction of LDPM surrogate vs. synthetic experimental data sets in colour. Each colour corresponds to one verification data set.
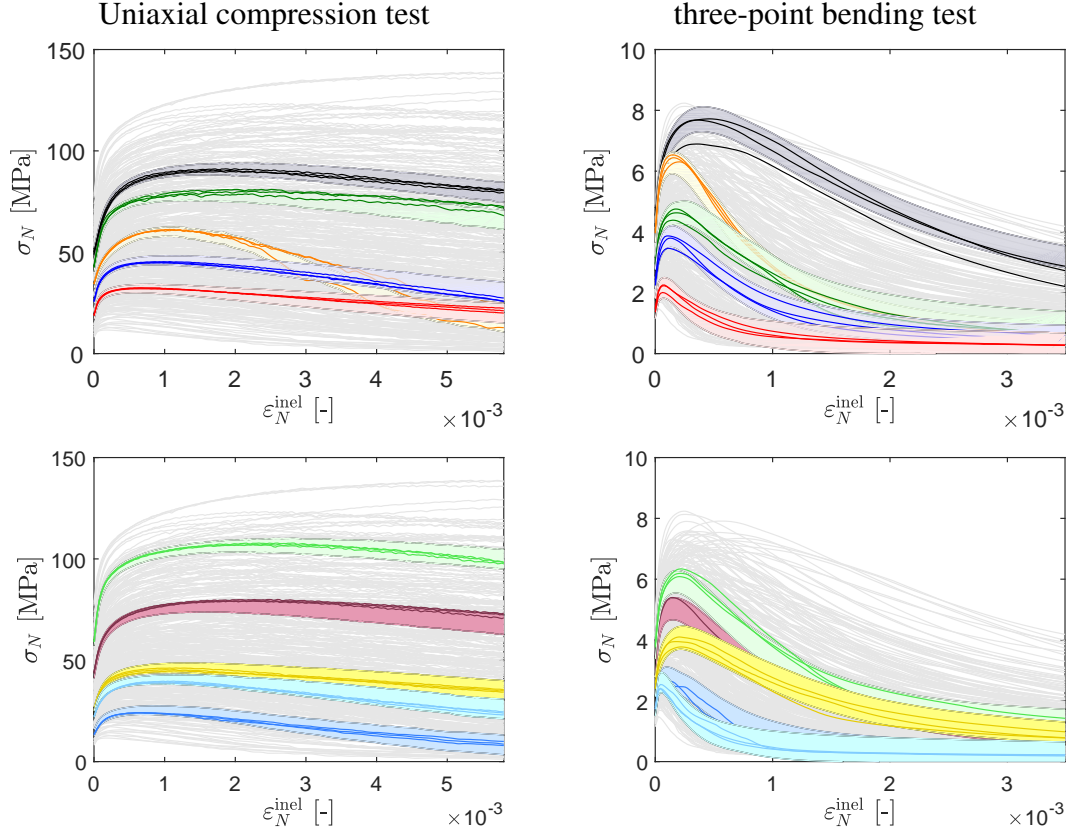
Figure 5: Gray stress-strain diagrams employed for construction of LDPM surrogate vs. results of verification in colour. Solid lines are synthetic experimental data while areas in colour represent identified regions of high prediction probability. Each colour corresponds to one verification data set.

obtained as a mean of the three diagrams simulated by LDPM for three randomly generated particle configurations, but for material parameters fixed to the posterior mean values $\widehat{\boldsymbol{x}}$. Since the measurement and theoretical errors cause a significant randomness of the model response, the areas in colour aim to reflect this relevant randomness. In particular, it corresponds to the interval $[\widehat{y}_i - \sigma_{m,i} - \sigma_{t,i}, \widehat{y}_i + \sigma_{m,i} + \sigma_{t,i}]$ for i-th response component. Under the assumptions that $\widehat{\boldsymbol{y}}(\widehat{\boldsymbol{x}})$ is the estimate of the mean predicted LPDM response[2] and the normal distributions of both measurement ($\boldsymbol{\varepsilon}_m$) and theoretical ($\boldsymbol{\varepsilon}_t$) errors, the highlighted area covers the region where the LPDM response, including both considered errors, can be expected with 68.2% probability. We should also note that the area thus reflects the measurement error with modelled randomness of material morphology, but it does not include the uncertainty in the parameter estimates. We excluded this uncertainty as it can be significantly reduced by the amount of the data involved in the identification process, while the included errors are expected to be always present to a similar extent. In order to illustrate also particular contributions of both errors ($\boldsymbol{\varepsilon}_m$ and $\boldsymbol{\varepsilon}_t$), we plotted their standard deviations in particular response components in Figure 6. It can be thus clearly noticed that measurement error is considered to be equal at all loading steps, while the influence of the randomly generated

---

[2]Note that it is only very rough estimate of the mean response as it is based on average values from only three LDPM simulations.
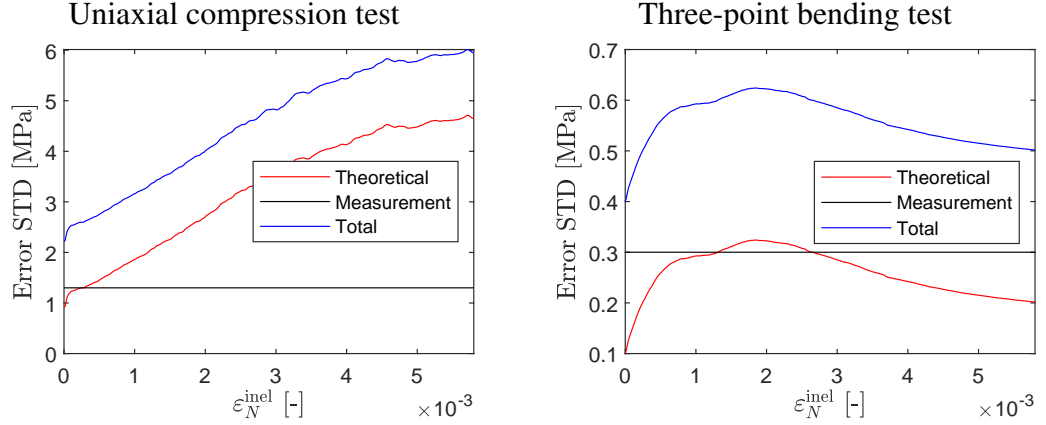
Figure 6: Standard deviations of measurement and theoretical errors.

meso-structure significantly differs and its impact is most pronounced at the later post-peak of the compression test and around the maximal stress achieved in three-point bending test. Based on the results shown in Figure 5, we thus conclude that the proposed identification strategy is sufficiently robust and was satisfactorily verified against ten synthetic data sets.

4.1. Discussion on posterior distribution of material parameters for one data set

Besides the ability of the identification strategy to determine LDPM material parameters allowing satisfactory reproduction of the synthetic data, we are also interested in the information contained in the elaborated experiments about the material parameters and the precision of the parameter estimates. In order words, we are interested in the remaining uncertainty in the estimates of material parameters, which is expressed in terms of posterior marginal 1D and 2D probability distributions plotted in Figure 7. This particular posterior distribution is obtained from an inverse analysis of one synthetic data set depicted by purple colour in Figures 4 and 5.

Figure 7 clearly reveals the very significant correlation between the estimates of E_0 and α which attains the value -0.963. Remarkable correlation can be then noticed also namely between l_t and n_t, between l_t and σ_t, and between l_t and E_0 whose absolute values lie in the range [0.55, 0.75]. Eight other parameter couples achieve absolute correlation values in the interval [0.4, 0.55] and only nine parameter couples have the absolute correlation smaller. We thus conclude that the overall correlation in parameter estimates is a high and different types of additional experimental data can be significantly beneficial for more accurate parameter estimation and namely more clear separation of the effect of particular parameters on the measured data.

Regarding 1D marginal distributions of particular parameters, we can see large remaining uncertainty in the estimates of parameters α, l_t, and n_t. To clearly quantify the posterior uncertainty, limits in all plots in Figure 7 are set to the prescribed prior bounds of parameter values. We can thus easily evaluate the change from the uniform prior in these prescribed bounds to the resulting posterior distributions and simultaneously the information contained in the data responsible for this distribution change. We have to also emphasise that the sharpness of all distributions depends strongly on the number of observations, i.e. the more sharp distributions (and also 1D marginals) can be achieved by employing a higher number of measured points considered to define the stress-strain diagram. Nevertheless, from the parameter
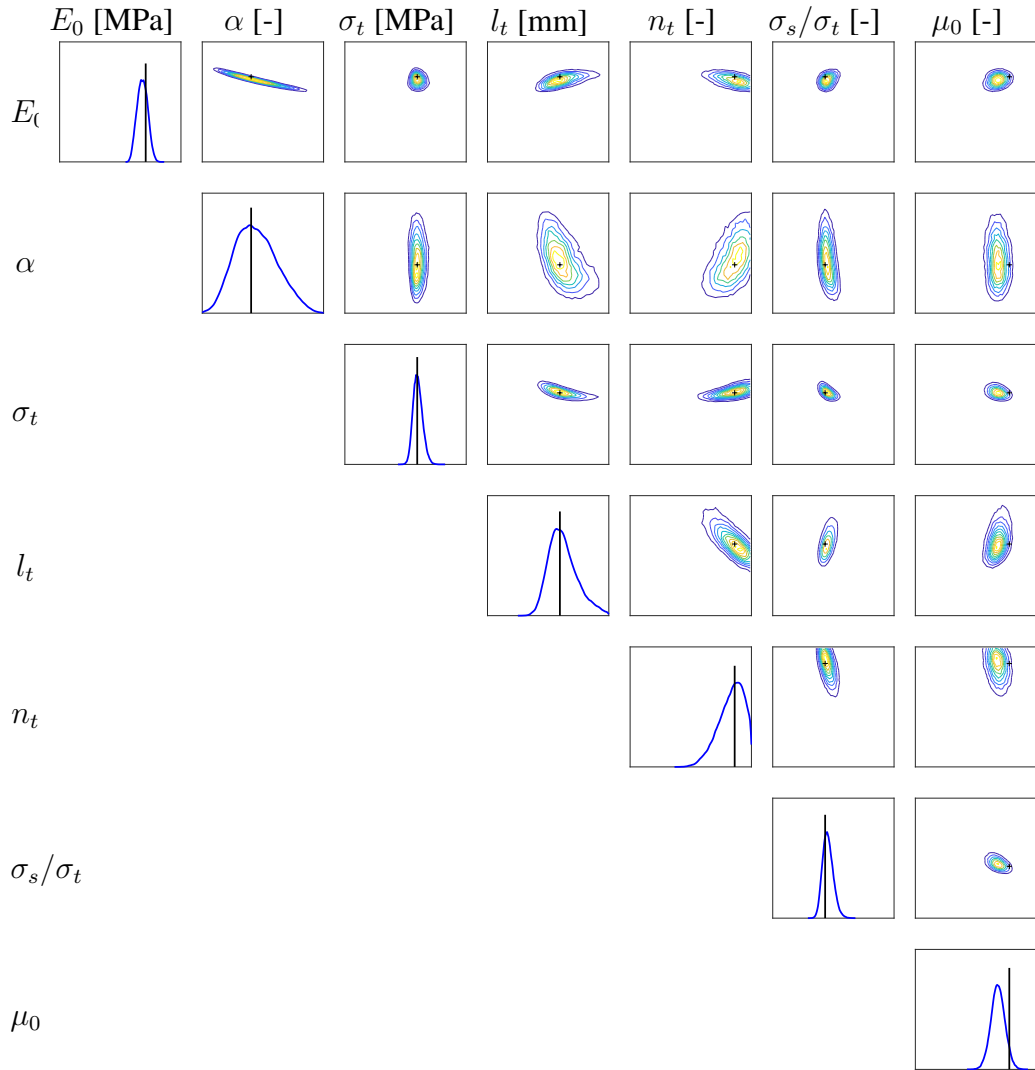
Figure 7: 1D and 2D marginal posterior distributions of model parameters corresponding to purple data set, black line in 1D marginals and black '+' symbol in 2D marginals correspond to the true value of parameter used to simulate the data.

identification point of view, it is more interesting to compare particular 1D marginals mutually. This leads us to the conclusion that additional experimental data should be selected with the emphasis namely to parameters α, l_t and n_t.

We should also mention that the maxima of 1D marginals do not always correspond to the true value of parameter depicted by the black line in Figure 7. It can be noticed especially for the parameter μ_0. This can be explained again by the model stochasticity. Even though the synthetic data were obtained for the known true values of parameters, the particular three synthetic experimental curves corresponding to one parameter set were obtained each for randomly generated particle configuration. These three curves are thus only a very rough representation of the material morphology and their mean response contains still a significant bias from the true mean material response corresponding to the used set of material parameters. Also, the employed PCE-based surrogate of the LDPM is only the approximation of the LDPM mean response, where the corresponding approximation error is roughly

quantified by the theoretical error depicted in Figure 6. Both these errors in the estimates of the mean measured and modelled responses may lead the identification process to a slightly different set of parameters than the true one. Note that the quantification of bias in the mean of synthetic experimental response would require much more LDPM simulations performed for the true values of material parameters, which will be again highly computationally expensive. Moreover, in the case of model calibration from real experimental data, a high number of stress-strain diagrams from both experimental types will be too expensive in both the financial and man-power aspects. Therefore, we consider as a reasonable to demonstrate the functionality of the proposed identification strategy within the limits of its further practical applications, where again only a very limited number of stress-strain diagrams will be available.

4.2. Quantification of bias in posterior mean estimates

On the other hand, in the case of synthetic experimental data obtained for known values of material parameters, we can easily evaluate the bias in resulting parameter estimates. To that purpose, the estimated posterior mean $\widehat{x}_{i,j}$ of i-th parameter for j-th synthetic data set is compared to its known true values $x^*_{i,j}$ and the obtained bias is normalised with respect to the prior range of parameter values (see Table 2) according to

$$\varepsilon_{i,j} = \frac{|x^*_{i,j} - \widehat{x}_{i,j}|}{\max(x_i) - \min(x_i)} \ . \tag{13}$$

The mean, minimal and maximal bias obtained over all $n_z = 10$ synthetic data sets are then summarised in Table 4. By comparing these results with posterior 1D marginal distributions

Table 4: Normalised bias in estimated mean of material parameters.

	E_0	α	σ_t	l_t	n_t	σ_s/σ_t	μ_0
$\bar{\varepsilon}_i$ [%]	4.81	24.60	3.04	12.62	14.86	4.82	7.84
$\min \varepsilon_i$ [%]	0.33	0.63	0.24	1.66	0.27	0.00	0.53
$\max \varepsilon_i$ [%]	12.31	67.87	8.61	28.02	29.58	12.44	21.90

in Figure 7, it can be noticed that the parameters with a higher amount of the posterior uncertainty in their estimates (namely α, l_t and n_t) have also the higher value of bias in their posterior mean. These values of biases clearly suggest that remaining posterior uncertainty cannot be fully solved by increasing number of considered measured points of stress-strain diagrams, but there is a principal lack of information in the two elaborated experiments and some other type of experimental data is necessary for determination and separation of the particular material parameters.

4.3. Global sensitivity analysis

Another inexpensive way to quantify the information about material parameters hidden in particular data points is offered by the so-called global sensitivity analysis. Once having a constructed PCE-based surrogate of the LDPM, its coefficients can be easily used to compute analytically the Sobol' sensitivity indices as described in [28]. The idea of this sensitivity analysis is based on analysis of variance, where the variance of each response component is decomposed into fractions which can be attributed to particular material parameters. The

first-order Sobol' indices include the variance induced solely by one parameter. Total Sobol' indices, on the other hand, include the effect of all polynomial terms depending on a given parameter including the effect of mixed polynomial terms depending on more parameters simultaneously. Attributing the effect of mixed terms is questionable as there is no possibility to clearly separate the role of individual parameters. Here we introduce modified Sobol' indices, where variance of a mixed polynomial term is equally distributed among all involved parameters in that term. This modified Sobol' index S_{Mi} corresponding to i-th material parameter is then given as

$$S_{Mi} = \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{I}_{Mi}} {}^{1}/n_{\boldsymbol{\alpha}} \beta_{\boldsymbol{\alpha}}^2 \mathbb{E}[\psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\xi})]}{\sum_{\boldsymbol{\alpha} \in \mathcal{I}} \beta_{\boldsymbol{\alpha}}^2 \mathbb{E}[\psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\xi})]}, \qquad (14)$$

where $n_{\boldsymbol{\alpha}}$ is a number of variables included in the polynomial term given by the multi-index $\boldsymbol{\alpha}$, $\mathcal{I}$ is a set of multi-indices or n_x-tuples $\boldsymbol{\alpha}$ defining all polynomial terms present in the constructed PCE and $\mathcal{I}_{Mi} \subset \mathcal{I}$ is a subset of terms depending on the i-th material parameter x_i. The advantage of the proposed sensitivity index S_{Mi} is simply that its sum over all material parameters equals to one, i.e. $\sum_{i=1}^{n_x} S_{Mi} = 1$ and the role of individual parameters can be compared also mutually among particular response components. The results of such sensitivity analysis for the elastic stiffness K_{cN} and discrete stress values in the uniaxial compression test and three-point bending test are shown in Figure 8.
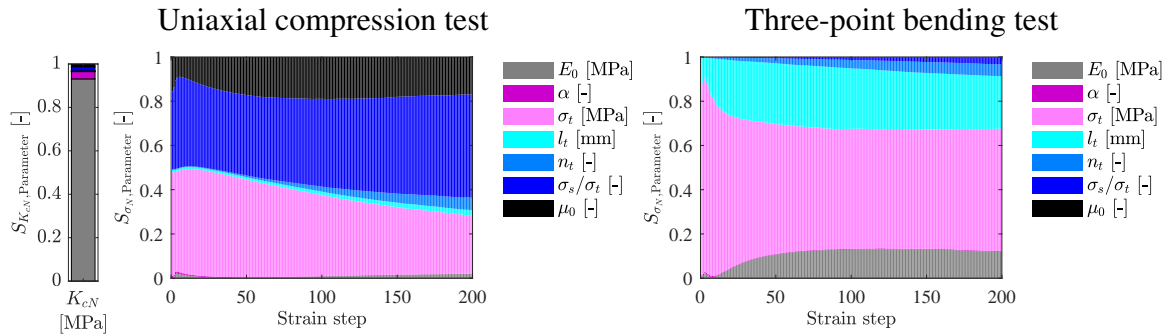
Figure 8: Sensitivity analysis - Sobol' indices computed from PCE coefficients.

The sensitivity analysis shows, in accordance with the posterior 1D marginal distributions in Figure 7 and biases in estimates in Table 4, that the material parameters α and n_t are really insignificant for all the response components. σ_t is highly significant for both experiments, E_0 is extremely significant for elastic stiffness K_{cN} and σ_s/σ_t is highly significant in uniaxial compression. μ_0 and l_t parameters seem to be also relatively significant at least for one of the two experiments, but their effect can be diminished by the other more significant parameters and the underlying approximation errors. Moreover, the unsatisfactory estimation of l_t can be also caused by its high posterior correlation with other significant parameters such as σ_t, whose correlation is -0.67, or with E_0 and correlation 0.55.

To verify these results, we performed one inverse analysis also separately for both experiment. In particular, one analysis was performed by using the elastic stiffness K_{cN} and a response from uniaxial compression, other analysis for response from three-point bending test. The posterior 1D marginal distributions depicted in Figure 9 clearly confirm that σ_t can be determined from both experiments even separately, while E_0 can, as expected, be well
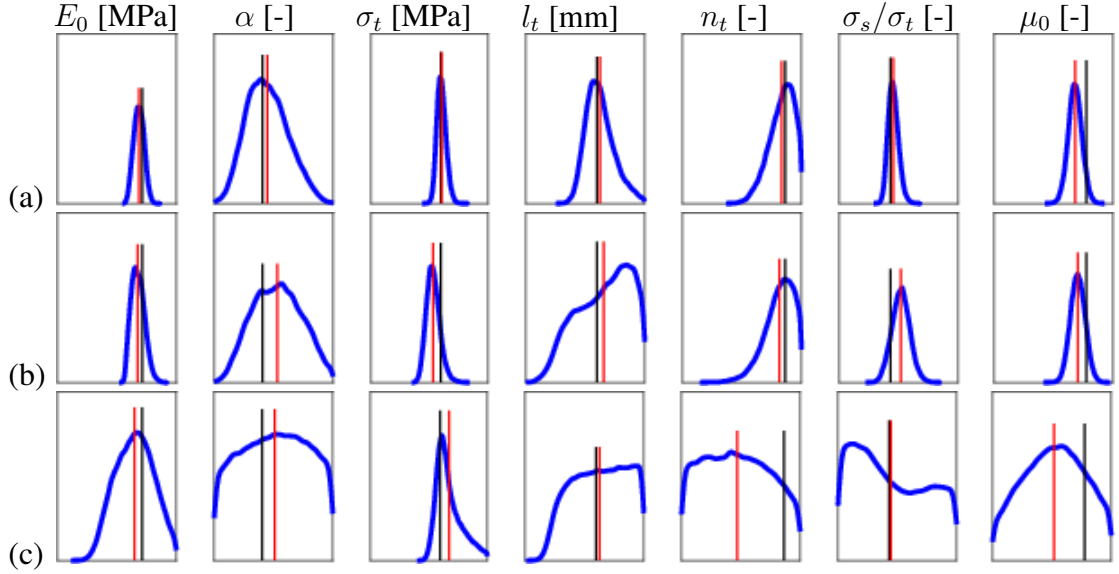
Figure 9: Identified 1D marginal pdfs of model parameters corresponding to experimental data from both tests (a), uniaxial compression test including elastic stiffness K_{cN} (b), and three point-bending test (c). True values are marked by black and identified values by red colour.

identified namely from elastic stiffness K_{cN} (probably even without the responses from uniaxial compression) and σ_s/σ_t or μ_0 can be well identified from uniaxial compression test. Regarding the results of l_t, its separate identification from any of the experiments leads to much worse results, which is consistent with our assumption of its disturbing correlation with other parameters which is reduced by combining both experiments leading to its more precise estimation.

The presented sensitivity analysis thus provides a useful tool allowing to distinguish the role of particular material parameters to particular response components. Once a PCE-based surrogate to the material model is constructed, its evaluation is analytical and very computationally inexpensive. Its results can be used for planning of additional observations with emphasis on material parameters which are insignificant for the simulation of experimental tests employed in this paper.
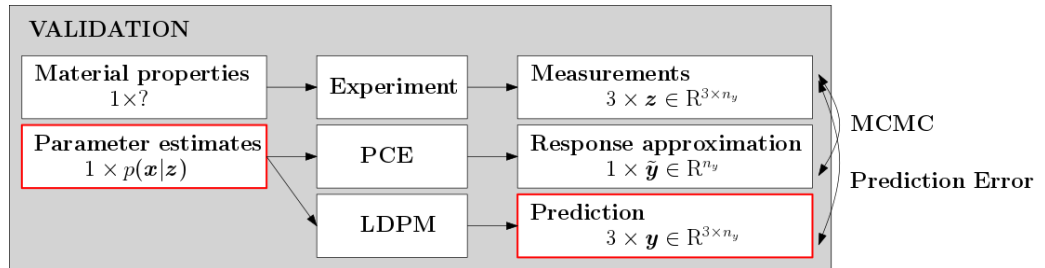
Figure 10: Scheme of validation of the proposed parameter identification strategy on experimental data.

5. Validation

The verified identification approach is then validated on experimental data, where the LDPM is calibrated by the proposed strategy to fit these data according to the scheme in

Table 5: Identified mean values and standard deviations of parameters.

	E_0 [MPa]	α [-]	σ_t [MPa]	l_t [mm]	n_t [-]	σ_s/σ_t [-]	μ_0[-]
MIN	20000	0.2	1.5	50	0.1	1.5	0.001
MAX	70000	0.3	5	300	1	8	0.5
MEAN	36781	0.2723	2.7436	164.99	0.9472	1.6475	0.1742
STD	1471	0.0224	0.1910	28.94	0.0505	0.1307	0.0358

Figure 10. As already mentioned in Section 2.2, the real experimental data of normal strength concrete, available for three repetitions of both experimental types, are utilised.

Table 5 contains the posterior means and standard deviations of parameter estimates along with the prior bounds. The posterior distribution is also depicted in Figure 11 in the form of marginal 1D and 2D probability distributions of the material parameters. By
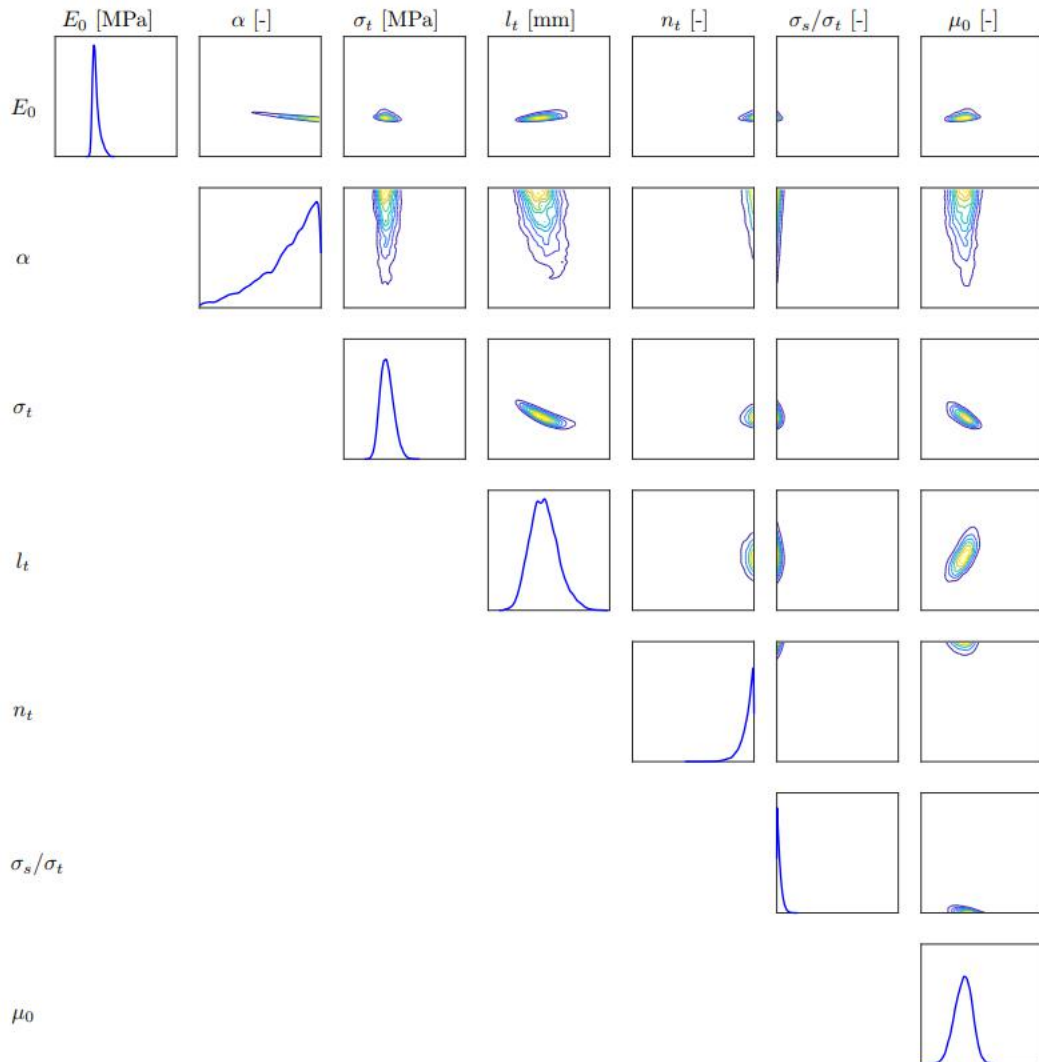
Figure 11: 1D and 2D marginal posterior distributions of model parameters corresponding to experimental data, black line in 1D marginals and black '+' symbol in 2D marginals correspond to the true values of parameters used to simulate the data.

comparing the posterior means with the prior bounds, we can highlight that the estimated means of n_t and σ_s/σ_t lie very close to the prior bounds. Also, the value of α approaches remarkably its prior maximum, see its posterior distribution in Figure 11. This phenomenon suggest that the prior bounds are not sufficiently wide and the elaborated experimental data lie beyond these bounds. Luckily based on the results of verification we know that the parameters α and n_t are not much significant and their inappropriate prior limitations probably do not significantly harm the results of the validation. However, the opposite is true for the σ_s/σ_t parameter, which is highly significant for uniaxial compression. Its incorrect estimation due to inappropriate prior bounds may cause namely unsatisfactory fit of the measured curves in uniaxial compression, as can be seen in Figure 12, where the experimental data are plotted along with the LDPM response obtained for the posterior parameter means. We can also conclude that the measured data obtained from three-point bending test are predicted well for the estimated parameters and in accordance with the results of verification, values of E_0 and σ_t are determined with very low posterior uncertainty and μ_0 with l_t seems to be identified also with good accuracy. In order to achieve higher accuracy in μ_0, l_t and namely in α and n_t, other types of experiments needs to be designed and incorporated in the LDPM calibration process.
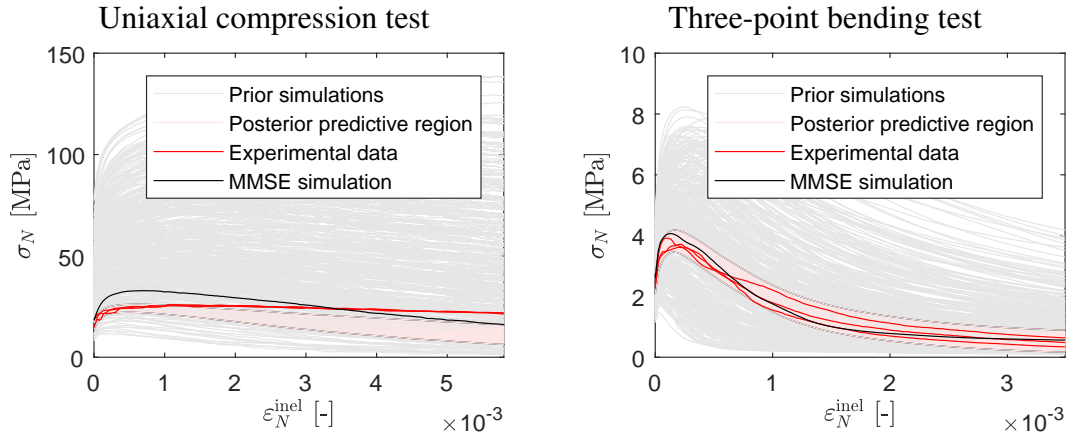
Figure 12: Gray model output curves for construction of approximation vs. results of the validation in colour. Solid lines are experimental data while filled areas represent identified regions of high prediction probability. Black lines represent the closest prior simulation to experimental data, i.e. prior simulation with minimum mean square error (MMSE).

6. Conclusions

The paper is focused on the probabilistic parameter identification of lattice discrete particle model for concrete. The essential benefit of the presented identification method is that it is general, computationally feasible, and automated (i.e. relatively easy-to-use) thus allowing for wide model utilisation without necessity of deep understanding of its sophisticated formulation. The core of the model calibration procedure is the Bayes' theorem for probabilistic parameter identification, which allows to take into account uncertainties connected to the available experimental data as well as the model inaccuracy. This probabilistic formulation requires, however, high number of LDPM simulations. To reduce the related computational cost, the LDPM is replaced by its polynomial approximation. This substitution is

not convenient only for reducing the computational time, but it namely solves the obstacle of the LDPM stochasticity as the polynomial chaos expansion can be considered as a rough approximation of the mean LDPM response. The model calibration is thus turned into fitting the mean LDPM response to the mean observed stress-strain diagram for a concrete mix with a given material parameters. The Bayesian inference results in posterior probability distribution of deterministic values of material parameters corresponding to given concrete mix and this distribution expresses the uncertainty due to the insufficient information contained in the employed experiments, measurement error and estimated model stochasticity.

The effect of model stochasticity is estimated here only very roughly due to the limited number of LDPM simulations employed for construction of the surrogate model. Runtime of these simulations is the principal computational burden of the whole proposed calibration method. Once the surrogate model is constructed, the Markov chain Monte Carlo-based estimation of posterior distribution of material parameters is already relatively fast. From another point of view, this type of surrogate has a significant potential in a repeated use for any new data observed within the given framework, i.e. the prescribed experimental setups (geometry of specimens, boundary conditions and loading).

The following conclusions can be drawn from the results presented in this study:

- the proposed identification strategy is sufficiently robust and precise to fit the synthetic stress-strain diagrams representing ten different concrete mixes (Figure 5);

- the marginal posterior distributions (Figure 7) of material parameters for one selected concrete mix suggests that many parameters play in the employed experiments similar role and there is not enough independent data to clearly separate their effects, i.e. the information contained in the employed experiments is insufficient to precisely determine all the identified parameters;

- lack of the information in the experiments is also projected to the remaining uncertainty in the estimates of individual parameters (Figure 7) as well as in the bias in the estimated posterior mean of the particular parameters (Table 2), where the biggest uncertainty and/or error remains namely in the estimates of α, l_t and n_t. This conclusion is also supported by the results of global sensitivity analysis (Figure 8) revealing no observations significantly sensitive to these parameters;

- the prior lower/upper bounds has a substantial impact on the final results and must be appropriately chosen before the surrogate model construction.

Overall we conclude that the paper presents an identification strategy for material parameters of LDPM, which is robust, computationally feasible and allows for an automatisation thus enabling wider usage of the LDPM within the engineering community. As a recommendation we suggest to be very careful with the prescribing prior bounds of the particular material parameters, as their further change requires the nontrivial reconstruction of the employed surrogate model. Moreover, the proposed procedure allows a comfortable incorporation of any type of experimental setup for which the surrogate model can be created based on the LDPM simulations and arbitrary extension to all relevant material parameters used within LDPM.

Acknowledgment

References

[1] G. Cusatis, D. Pelessone, A. Mencarelli, Lattice discrete particle model (LDPM) for failure behavior of concrete. I: Theory, Cement and Concrete Composites 33 (9) (2011) 881–890.

[2] G. Cusatis, A. Mencarelli, D. Pelessone, J. Baylot, Lattice discrete particle model (LDPM) for failure behavior of concrete. II: Calibration and validation, Cement and Concrete Composites 33 (9) (2011) 891–905.

[3] J. Eliáš, M. Vořechovský, J. Skoček, Z. P. Bažant, Stochastic discrete meso-scale simulations of concrete fracture: Comparison to experimental data, Engineering Fracture Mechanics 135 (2015) 1–16.

[4] R. Ghanem, P. D. Spanos, Stochastic finite elements: A spectral approach, second revised Edition, Dover Publications, Mineola, New York, 2003.

[5] H. G. Matthies, Encyclopedia of Computational Mechanics, John Wiley & Sons, Ltd., 2007, Ch. Uncertainty Quantification with Stochastic Finite Elements.

[6] G. Blatman, B. Sudret, An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis, Probabilistic Engineering Mechanics 25 (2) (2010) 183–197.

[7] E. Janouchová, J. Sýkora, A. Kučerová, Polynomial chaos in evaluating failure probability: A comparative study, Applications of Mathematics 63 (2018) 713–737.

[8] M. Alnaggar, N. Bhanot, A machine learning approach for the identification of the lattice discrete particle model parameters, Engineering Fracture Mechanics 197 (2018) 160–175.

[9] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, 2005.

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

[11] A. Kučerová, J. Sýkora, B. Rosić, H. G. Matthies, Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials, Journal of Computational and Applied Mathematics 236 (18) (2012) 4862–4872.

[12] E. Janouchová, A. Kučerová, Bayesian inference of heterogeneous viscoplastic material parameters, Acta Polytechnica CTU Proceedings 15 (2018) 41–45.

[13] G. Cusatis, Z. P. Bažant, L. Cedolin, Confinement-shear lattice model for concrete damage in tension and compression: I. Theory, Journal of Engineering Mechanics 129 (2003) 1439.

[14] G. Cusatis, Z. P. Bažant, L. Cedolin, Confinement-shear lattice CSL model for fracture propagation in concrete, Computer Methods in Applied Mechanics and Engineering 195 (52) (2006) 7154–7171.

[15] G. Cusatis, Strain-rate effects on concrete behavior, International Journal of Impact Engineering 38 (4) (2011) 162–170.

[16] Z. P. Bažant, M. R. Tabbara, M. T. Kazemi, G. Pijaudier-Cabot, Random particle model for fracture of aggregate or fiber composites, Journal of Engineering Mechanics 116 (8) (1990) 1686–1705.

[17] M. Yip, Z. Li, B. Liao, J. E. Bolander, Irregular lattice models of fracture of multiphase particulate materials, International journal of fracture 140 (2006) 113–124.

[18] G. Cusatis, H. Nakamura, Discrete modeling of concrete materials and structures, Cement and Concrete Composites 33 (9) (2011) 865 – 866.

[19] E. A. Schauffert, G. Cusatis, D. Pelessone, J. O'Daniel, J. Baylot, Lattice discrete particle model for fiber-reinforced concrete. II: Tensile fracture and multiaxial loading behavior, Journal of Engineering Mechanics 138 (7) (2012) 834–841.

[20] M. Marcon, J. Vorel, K. Ninčević, R. Wan-Wendner, Modeling adhesive anchors in a discrete element framework, Materials 10 (8) (2017) 917.

[21] J. Vorel, M. Marcon, G. Cusatis, F. Caner, G. Di Luzio, R. Wan-Wendner, Constitutive modeling of concrete - comparison of the state of the art models, Computers & Structures. Submitted.

[22] British Standard Institute, BS EN 12390-3:2009 - Testing hardened concrete. Compressive strength of test specimens.

[23] D. Xiu, G. E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM Journal on Scientific Computing 24 (2) (2002) 619–644.

[24] M. Lepš, E. Myšáková, Fast LHS design of experiments and its multicriteria comparison, in: UNCECOMP 2019. Athens: Institute of Structural Analysis and Antiseismic Research, NTUA, 2019.

[25] W. R. Gilks, Markov chain Monte Carlo, Wiley Online Library, 2005.

[26] J. C. Spall, Estimation via Markov chain Monte Carlo, IEEE Control Systems Magazine 23 (2) (2003) 34–45. doi:10.1109/MCS.2003.1188770.

[27] J. S. Rosenthal, Handbook of Markov Chain Monte Carlo, Chapman & Hall/CRC, Boca Raton, Fla., 2011, Ch. Optimal Proposal Distributions and Adaptive MCMC, pp. 93–140.

[28] G. Blatman, B. Sudret, Efficient computation of global sensitivity indices using sparse polynomial chaos expansions, Reliability Engineering and System Safety 95 (2010) 1216–1229.

Chapter 4

ACCELERATION OF MODEL CALIBRATION BASED ON SURROGATES

This chapter concentrates on accelerating the calibration of nonlinear material and/or structural models by constructing their computationally efficient surrogates. Two parts of the chapter are devoted to two widely used types of surrogate models: (i) layered feed-forward artificial neural network and (ii) polynomial chaos expansion. The first part elaborates and discusses different strategies for applying surrogates in the calibration of nonlinear material models. The second part is focused on the comparison of different numerical methods for constructing polynomial chaos-based surrogates.

List of selected journal papers and book chapters

Mareš, T., Janouchová, E., and Kučerová, A. (2016). Artificial neural networks in calibration of nonlinear mechanical models. *Advances in Engineering Software*, 95:68–81.

Janouchová, E., Sýkora, J., and Kučerová, A. (2018). Polynomial chaos in evaluating failure probability: A comparative study. *Applications of Mathematics*, 63(6):713–737.

List of related journal papers and book chapters

Janouchová, E. and Kučerová, A. (2013). Competitive comparison of optimal designs of experiments for sampling-based sensitivity analysis. *Computers & Structures*, 124:47–60.

Artificial neural networks in calibration of nonlinear mechanical models

Tomáš Mareš

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Eliška Janouchová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Anna Kučerová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Abstract

Last decades witness rapid development in numerical modelling of materials and the complexity of new models increases quickly together with their computational demands. Despite the growing performance of modern computers and clusters, calibration of such models from noisy experimental data remains a nontrivial and often computational exhaustive task. The layered neural networks thus represent a robust and efficient technique to overcome the time-consuming simulations of a calibrated model. The potential of neural networks consists in simple implementation and high versatility in approximating nonlinear relationships. Therefore, there were several approaches proposed to accelerate the calibration of nonlinear models by neural networks. This contribution reviews and compares three possible strategies based on approximating (i) model response, (ii) inverse relationship between the model response and its parameters and (iii) error function quantifying how well the model fits the data. The advantages and drawbacks of particular strategies are demonstrated on calibration of four parameters of affinity hydration model from simulated data as well as from experimental measurements. This model is highly nonlinear, but computationally cheap thus allowing its calibration without any approximation and better quantification of results obtained by the examined calibration strategies.

Keywords: Artificial neural network, Multi-layer perceptron, Parameter identification, Principal component analysis, Sensitivity analysis, Affinity hydration model

Email addresses: `marestom87@gmail.com` (Tomáš Mareš),
`eliska.janouchova@fsv.cvut.cz` (Eliška Janouchová), `anicka@cml.fsv.cvut.cz` (Anna Kučerová)

1. Introduction

Development in numerical modelling provides the possibility to describe a lot of complex phenomena in material or structural behaviour. The resulting models are, however, often highly nonlinear and defined by many parameters, which have to be estimated so as to properly describe the investigated system and its behaviour. The aim of the model calibration is thus to rediscover unknown parameters knowing the experimentally obtained response of a system to the given excitations. The principal difficulty of model calibration is related to the fact that while the numerical model of an experiment represents a well-defined mapping from input (model, material, structural, or other parameters) to output (structural response), there is no guarantee that the inverse relation even exists.

The most broadly used approach to parameter identification is usually done by means of an error minimisation technique, where the distance between parameterised model predictions and observed data is minimised [1]. Since the inverse relation (mapping of model outputs to its inputs) is often ill-posed, the error minimisation technique leads to a difficult optimisation problem, which is highly nonlinear and multi-modal. Therefore, the choice of an appropriate identification strategy is not trivial.

Another approach intensively developed during the last decade is based on Bayesian updating of uncertainty in parameters' description [2, 3]. The uncertainty in observations is expressed by corresponding probability distribution and employed for estimation of the so-called posterior probabilistic description of identified parameters together with the prior expert knowledge about the parameter values [4, 5]. The unknown parameters are thus modelled as random variables originally endowed with prior expert-based probability density functions which are then updated using the observations to the posterior density functions. While the error minimisation techniques lead to a single point estimate of parameters' values, the result of Bayesian inference is a probability distribution that summarizes all available information about the parameters. Another very important advantage of Bayesian inference consists in treating the inverse problem as a well-posed problem in an expanded stochastic space.

Despite the progress in uncertainty quantification methods [6, 7], more information provided by Bayesian inference is generally related to more time-consuming computations. In many situations, the single point estimate approach remains the only feasible one and development of efficient tools suitable for this strategy is still an actual topic. Within the several last decades, a lot of attention was paid to the so-called intelligent methods of information processing and among them especially to soft computing methods such as artificial neural networks (ANNs), evolutionary algorithms or fuzzy systems [8]. A review of soft computing methods for parameter identification can be found e.g. in [9]. In this paper, we focus on applications of ANNs in the single point approach to parameter identification.

2. Artificial neural network

Artificial neural networks (ANNs) [10, 11] are powerful computational systems consisting of many simple processing elements - so-called neurons - connected together to perform tasks analogously to biological brains. Their main feature is the ability to change their behaviour based on external information that flows through the ANN during the learning (training) phase.

A particular type of ANN is the so-called feedforward neural network, which consists of neurons organized into layers where outputs from one layer are used as inputs into the following layer, see Figure 1. There are no cycles or loops in the network, no feed-back connections. Most frequently used example is the multi-layer perceptron (MLP) with a sigmoid transfer function and a gradient descent method of training called the back-propagation learning algorithm. In practical usage, the MLPs are known for their ability to approximate nonlinear relations and therefore, when speaking about an ANN, the MLP is considered in the following text.
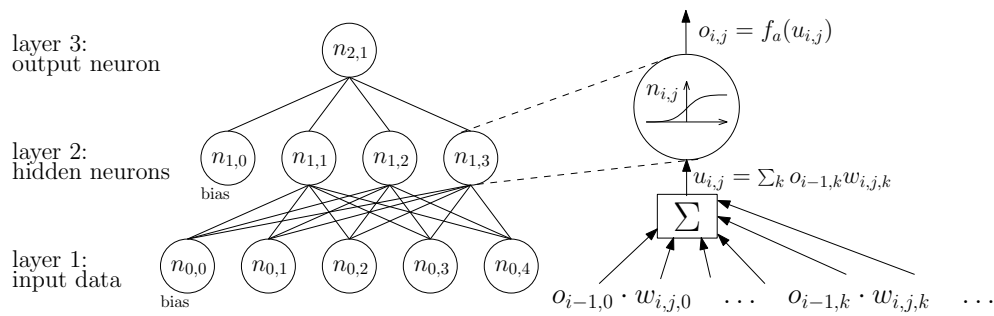
Figure 1: Architecture of multi-layer perceptron

The input layer represents a vector of input parameters which are directly the outputs of the input layer. The outputs $o_{i-1,k}$ of the $(i-1)$-th layer are multiplied by a vector of constants $w_{i,j,k}$, the so-called synaptic weights, summarized and used as inputs $u_{i,j}$ into the following i-th layer. Elements in the hidden and output layers - neurons - are defined by an activation function $f_a(u_{i,j})$, which is applied on the input and produces the output value of the j-th neuron in the i-th layer, i.e.

$$o_{i,j} = f_a\left(u_{i,j}\right) \qquad \text{where} \qquad u_{i,j} = \sum_k \left(o_{i-1,k} w_{i,j,k}\right) . \tag{1}$$

The synaptic weights $w_{i,j,k}$ are parameters of an ANN to be determined during the training process. The type of the activation function is usually chosen in accordance with the type of a function to be approximated. In the case of continuous problems, the sigmoid activation function given as

$$o_{i,j} = f_a\left(u_{i,j}\right) = \frac{1}{1 + e^{-u_{i,j}}} \tag{2}$$

is the most common choice.

One bias neuron is usually added into the input and hidden layers. It does not contain an activation function, but only a constant value. Its role is to enable to shift the value of a sum over the outputs of his neighbouring neurons before this sum enters as the input into the neurons in the following layer. The value of biases is determined by training process together with the synaptic weights.

Despite of ANN's popularity there are only few recommendations for the choice of ANN's architecture. The authors, e.g. in [12, 13], show that the ANN with any of a wide variety of continuous nonlinear hidden-layer activation functions and one hidden layer with an arbitrarily large number of units suffices for the "universal approximation" property. Therefore, we limit our numerical experiments to such case. The number of units in the input

and the output layer is usually given by the studied problem itself, but there is no theory yet specifying the number of units in the hidden layer.

To overcome this problem, some model selection technique [14] has to be applied in order to perform a guided choice of the ANN's topology. The most general strategy is the cross-validation, because it does not involve any probabilistic assumptions or dependencies on a identification problem. The idea of cross-validation is based on repeated ANN prediction error evaluation for a chosen subset of training data and selection of the ANN with the smallest averaged prediction errors. Comparing to the well-known model validation on some independent set of data, the advantage of cross-validation consists in smaller requirements on data amount [15].

3. Strategies for application of ANN in model calibration

In model calibration, the goal is to find a set of model parameters minimising the difference between the model response and experimental measurements, see Figure 2. An in-
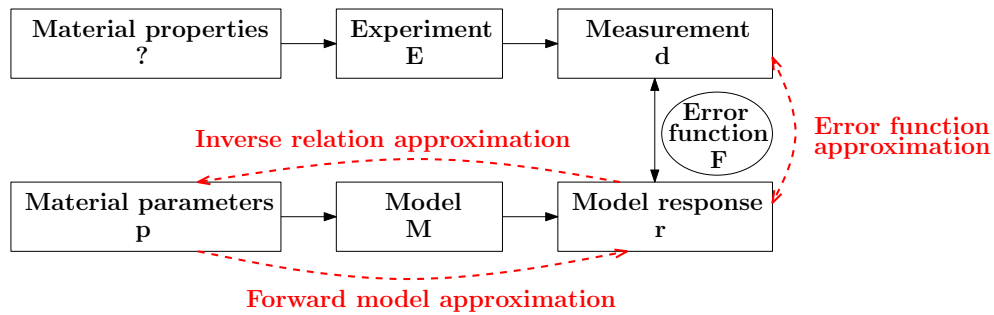
Figure 2: Scheme of model calibration procedure.

tuitive way of solving calibration problem is to formulate an error function quantifying this difference and to minimise the error function using some optimisation algorithm. The most common error functions are given as

$$F_1 = \sum_{i=1}^{N_R} (r_i - d_i)^2, \tag{3}$$

$$F_2 = \sum_{i=1}^{N_R} |r_i - d_i|, \tag{4}$$

where r_i is the i-th component of model response corresponding to the i-th measured quantity d_i and N_R is a number of measured quantities. The difficulty arises from the nonlinear relation between the model response and model parameters often causing complexity of the error function such as multi-modality or non-differentiability. Therefore, the computationally efficient methods based on analytically or numerically obtained gradient can be applied only in specific cases.

A more general possibility is to apply some evolutionary algorithm which can handle the multi-modality once furnished by a sufficient number of function evaluations. However, one evaluation of an error function always involves a simulation of the model. Even for the relatively fast model simulation, the optimisation can become easily unfeasible because of

the huge number of function evaluations commonly needed by evolutionary algorithms, even though they usually need less simulations than uncertainty based methods mentioned in the introductory part of the paper.

One way of reducing the number of model simulations is to construct a **forward model approximation** based e.g. on an ANN. The error function minimisation then becomes a minimisation of distance between the ANN's predictions and experimental data. The efficiency of this strategy relies on the evaluation of the trained ANN to be significantly much faster than the full model simulation. The advantage of this strategy is that the ANN is used to approximate a known mapping which certainly exists and is well-posed. Computational costs of this strategy are separated in two parts of a similar size: (i) the ANN training - optimisation of synaptic weights and (ii) the minimisation of an error in the ANN prediction for experimental data - optimisation of ANN inputs (i.e. determination of investigated model parameters). An important shortcoming of this method is that this ill-posed optimisation problem needs to be solved repeatedly for any new experimental measurement. This way of ANN application to the parameter identification was presented e.g. in [16], where an ANN is used for predicting load-deflection curves and the conjugate directions algorithm is then applied for optimisation of ductile damage and fracture parameters. Authors in [17] train an ANN to approximate the results of FE simulations of jet-grouted columns and optimise the column radius and a cement content of the columns by a genetic algorithm. Principally same methods are used for identification of elasto-plastic parameters in [18].

One more difficulty of the forward model approximation concerns the number of parameters and response components. It is very common that the experimental observations are represented by a discretized curves or surfaces in time or space dimensions being defined as a vectors with a large number of components. A forward model then represents a mapping from usually low-dimensional parameter space to high-dimensional response space. Although this mapping is well-posed, the surrogate model must have a large number of outputs or the time and/or space dimensions have to be included among the model inputs.

Another way of avoiding the mapping to a large number of outputs is to construct the **error function approximation**, where the model parameters are mapped onto only one scalar value. One important inconvenience of such strategy is of course the complexity of the error function, which can be, as mentioned above, highly nonlinear, multi-modal and/or non-smooth. Higher complexity of the approximated relation leads to a higher number of simulations needed for the approximation construction. This concerns another problem of estimation and choice of an appropriate design of experiments, i.e. sets of parameters, to perform the simulations which will enable to build up the surrogate with a relatively small error. This problem can be reduced by adaptive addition of design points, i.e. new simulations, close to the minimum of the error function approximation. Result of the new simulation is then used for an improvement of the surrogate and a new optimisation process is run again. Such an approach is usually well suited for surrogates based on kriging or radial basis function networks [19, 20]. In this paper, we limit our attention to application of feedforward layered neural networks and thus, we investigated their ability to approximate the error function with a limited number of simulations in non-adaptive fashion.

While the strategy of the forward model approximation involves a new optimisation process for any new data, the strategy of the error function approximation involves not only the optimisation process, but also the surrogate model construction. Regarding this aspect, the most convenient strategy is the **inverse relation approximation**, which needs only one eval-

uation to furnish the parameters corresponding to new observations. Of course, by the new observations we mean observations of the system with different properties but performed under the same external conditions (e.g. a different material, but the same specimen geometry and loading conditions). The strategy of the inverse relation approximation assumes the existence of an inverse relationship between the outputs and the inputs of the calibrated model. If such a relationship exists at least on a specified domain of parameters' values, it can be approximated by an ANN. Here the ANN training process represents the whole computational costs and a solution of the ill-posed problem. This way of ANN application to parameter identification was presented e.g. in [21] or recently in [22] for identification of mechanical material parameters, in [23] for estimation of elastic modulus of the interface tissue on dental implants surfaces, in [24] for identification of interfacial heat transfer coefficient or in [25] for determination of geometrical parameters of circular arches.

In order to illustrate the advantages and disadvantages of the outlined strategies of ANN's application to model calibration, we have chosen computationally simple but nonlinear affinity hydration model briefly described in the following section. The model was successfully validated on Portland cements in [26] and thus allows us to also validate the described identification strategies on experimental data as summarized in Section 9.

4. Affinity hydration model

Affinity hydration models provide a framework for accommodating all stages of cement hydration. We consider hydrating cement under isothermal temperature $25°C$. At this temperature, the rate of hydration can be expressed by the *chemical affinity* $\tilde{A}_{25}(\alpha)$ under isothermal $25°C$

$$\frac{d\alpha}{dt} = \tilde{A}_{25}(\alpha), \tag{5}$$

where the chemical affinity has a dimension of time^{-1} and α stands for the degree of hydration.

The affinity for isothermal temperature can be obtained experimentally; isothermal calorimetry measures a heat flow $q(t)$ which gives the hydration heat $Q(t)$ after integration. The approximation is given

$$\frac{Q(t)}{Q_{pot}} \approx \alpha, \tag{6}$$

$$\frac{1}{Q_{pot}}\frac{dQ(t)}{dt} = \frac{q(t)}{Q_{pot}} \approx \frac{d\alpha}{dt} = \tilde{A}_{25}(\alpha), \tag{7}$$

where Q_{pot} is expressed in J/g of cement paste. Hence the normalized heat flow $\frac{q(t)}{Q_{pot}}$ under isothermal $25°C$ equals to chemical affinity $\tilde{A}_{25}(\alpha)$.

Cervera et al. [27] proposed an analytical form of the normalized affinity which was refined in [28]. A slightly modified formulation is proposed here

$$\tilde{A}_{25}(\alpha) = B_1\left(\frac{B_2}{\alpha_\infty} + \alpha\right)(\alpha_\infty - \alpha)\exp\left(-\bar{\eta}\frac{\alpha}{\alpha_\infty}\right), \tag{8}$$

where B_1, B_2 are coefficients related to chemical composition, α_∞ is the ultimate hydration degree and $\bar{\eta}$ represents microdiffusion of free water through formed hydrates. The parameters in (8) express isothermal hydration at $25°C$.

Parameter	Minimum	Maximum	Relation
$B_1\,[h^{-1}]$	0.1	1	$p_1 = (B_1 - 0.1)/0.9$
$B_2\,[-]$	10^{-6}	10^{-3}	$p_2 = (\log B_2 + 6)/3$
$\bar{\eta}\,[-]$	2	12	$p_3 = (\bar{\eta} - 2)/10$
$\alpha_\infty\,[-]$	0.7	1.0	$p_4 = (\alpha_\infty - 0.7)/0.3$

Table 1: Bounds for affinity model parameters.

When hydration proceeds under varying temperature, maturity principle expressed via Arrhenius equation scales the affinity to arbitrary temperature T

$$\tilde{A}_T = \tilde{A}_{25} \exp\left[\frac{E_a}{R}\left(\frac{1}{273.15 + 25} - \frac{1}{T}\right)\right], \tag{9}$$

where R is the universal gas constant ($8.314\ \mathrm{Jmol^{-1}K^{-1}}$). For example, simulating isothermal hydration at $35°C$ means scaling $\tilde{A}_{25}$ with a factor of 1.651 at a given time. This means that hydrating concrete for 10 hours at $35°C$ releases the same amount of heat as concrete hydrating for 16.51 hours under $25°C$. Note that setting $E_a = 0$ ignores the effect of temperature and proceeds the hydration under $25°C$.

The integration of (7) needs to be carried out numerically. The predictor-corrector scheme is implemented as follows

$$\alpha_0^{(1)} = \alpha^{(0)} + \tilde{A}_T(\alpha^{(0)}) \cdot \Delta t, \tag{10}$$

$$\alpha_{n+1}^{(1)} = \alpha^{(0)} + \left[\tilde{A}_T(\alpha^{(0)}) - \tilde{A}_T(\alpha_n^{(1)})\right] \cdot \frac{\Delta t}{2}. \tag{11}$$

5. Sensitivity analysis

Since the ANN training needs a preparation of a set of training data, it is also worthy to use these data for a sampling-based sensitivity analysis [29, 30] and obtain some information about importance of particular observations or significance of each parameter for a system behaviour. To achieve some reliable information from sensitivity analysis as well as a good approximation by an ANN, one has to choose the training data carefully according to a suitable design of experiments, see e.g. [31] for a competitive comparison of several experimental designs.

Since the bounds for model parameters vary in orders, one can employ the expert knowledge about the parameter meanings and before preparation of the training data transform them into standardised parameters $p_i \in [0; 1]$. The bounds for the affinity model parameters together with their relations to the standardised parameters p_i are given in Table 1.

The affinity hydration model was chosen not only for its nonlinearity, but especially for its relatively simple interpretation and computationally fast simulation. Hence, we assume that the model is eligible to illustrate typical features of particular identification strategies. In order to understand the influence of the model parameters to its response more deeply, Figure 3 demonstrates the changes of the response induced by changes in a chosen parameter while the other parameters are fixed. On the other hand, to illustrate the spread of the model response corresponding to the parameters varying within the given domain, we prepare a design of experiments (DoE) having $N_{\mathrm{DoE}} = 100$ samples in the space of standardised
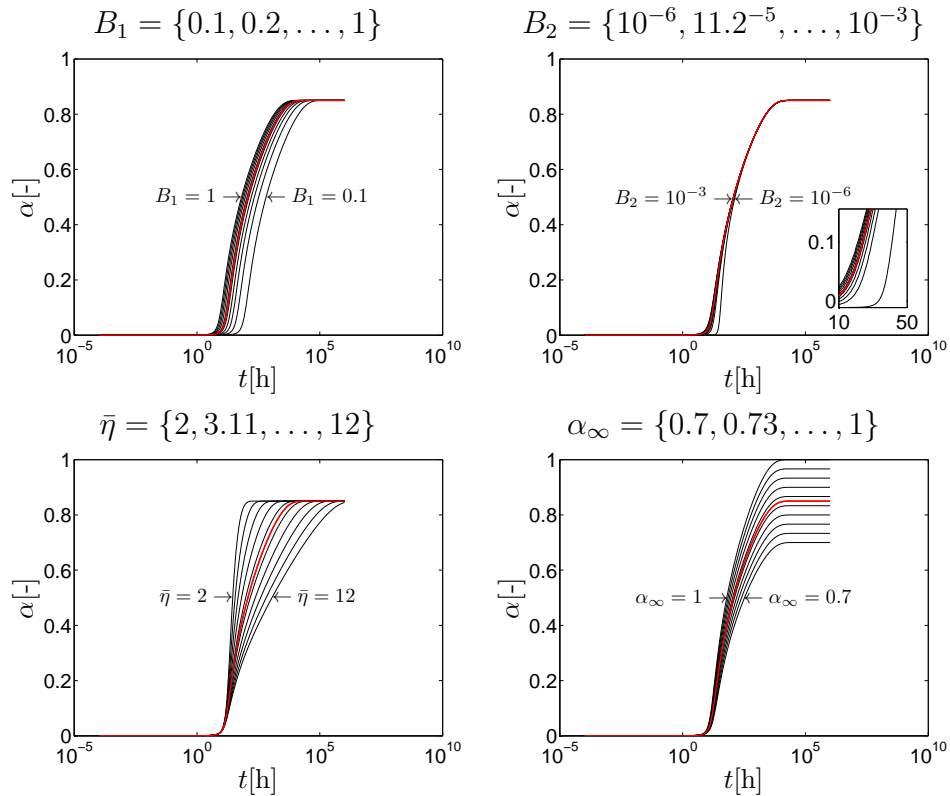
Figure 3: Influence of model parameters to model response α.

parameters. The DoE is generated as Latin Hypercube Sampling optimised with respect to the modified L_2 discrepancy. In [31] it is shown that such an experimental design has a good space-filling property and is nearly orthogonal. For each design point we perform a model simulation to obtain a bundle of N_{DoE} curves for the degree of hydration $\alpha(t)$, see Figure 4a.
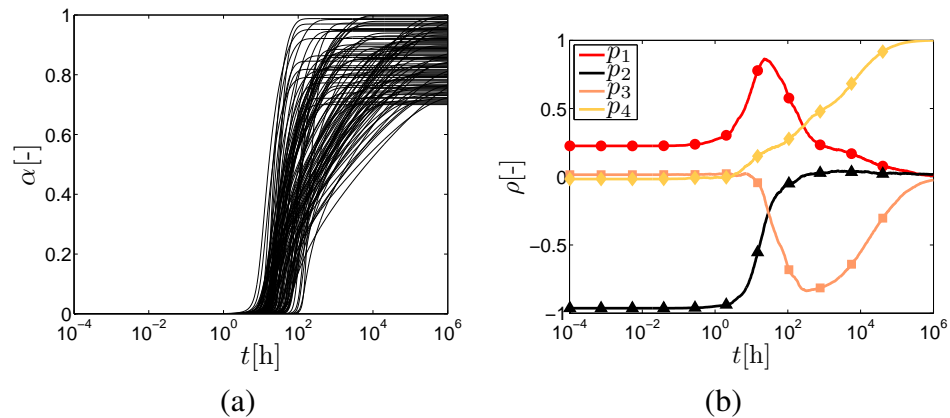
Figure 4: Bundle of degree of hydration curves obtained for design points (a) and sensitivity analysis for input-output pairs (b).

Since the model response is represented by the degree of hydration being a function of the time, the time domain is discretised into 1161 steps uniformly distributed with the logarithm of the time. Hence, the model input vector $\boldsymbol{p} = (p_1, p_2, p_3, p_4)$ consists of 4 parameters

and the output vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{N_{\text{time}}})$ consists of $N_{\text{time}} = 1161$ components. In order to quantify the influence of the model parameters to particular response components, we evaluate Spearman's rank correlation coefficient ρ for each $(\boldsymbol{p}_i, \boldsymbol{\alpha}_i)$ pair using all the $i \in \{1, \ldots, N_{\text{DoE}}\}$ simulations. The results of such a sampling-based sensitivity analysis [29] are plotted in Figure 4b.

In the inverse mode of identification, the model output vector $\boldsymbol{\alpha}$ consisting of $N_{\text{time}} = 1161$ components is too large for usage as an input vector for the ANN. Hence, we performed the principal component analysis (PCA) in order to reduce this number to $N_{\text{PCA}} = 100$ components $\bar{\boldsymbol{\alpha}} = (\bar{\alpha}_1, \ldots, \bar{\alpha}_2)$ with non-zero variance (this number is related to the number of simulations involved in PCA, i.e. $N_{\text{PCA}} = N_{\text{DoE}}$). The components are ordered according to their relative variance, see Figure 5a for the nine most important ones. Resulting
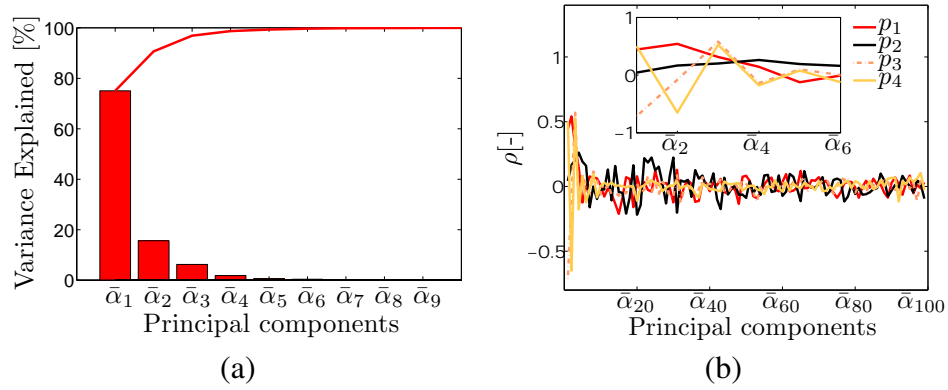
Figure 5: Variance explained by the first nine principal components (a) and sensitivity analysis for model inputs $\boldsymbol{p}_i$ - principal components $\bar{\boldsymbol{\alpha}}_i$ (b).

principal components are technically new quantities obtained by a linear combination of the original model outputs $\bar{\boldsymbol{\alpha}} = \bar{A}(\boldsymbol{\alpha})$. This transformation has of course an influence to sensitivity analysis and thus we computed correlations between the model inputs $\boldsymbol{p}_i$ and principal components $\bar{\boldsymbol{\alpha}}_i$, see Figure 5b.

6. Implementation of approximation strategies

Results of the described simulations are also used as training simulations for ANNs, i.e. $\mathcal{D}_{\text{train}} = \{(\boldsymbol{p}_i, \boldsymbol{\alpha}_i) \mid i \in \{1, 2, \ldots, N_{\text{train}}\}, N_{\text{train}} = N_{\text{DoE}} = 100\}$. Particular approximation strategies, however, process the training simulations in a different way.

The strategy of the forward model approximation can be formulated in two ways, which differ in handling the high dimensionality of the model output $\boldsymbol{\alpha}$. In the first formulation, we can consider the time step t_k as the fifth model parameter (i.e. the fifth model input) and thus the model output reduces into only one scalar value of the hydration degree α_k corresponding to the given time t_k. As the objective of the ANN is thus to span the parameter as well as the time space, we called this strategy as **Forward Complex** (ForwComp). In such a configuration, the results of N_{train} training simulations turns into $N_{\text{train}} \times N_{\text{time}} = 116, 100$ training samples. Evaluation of so many samples at every iteration of ANN training process is, however, very time-consuming. Therefore, only every m-th time step is included for ANN training and thus the training set is given as $\mathcal{D}_{\text{train}}^{\text{ForwComp}} = \{((\boldsymbol{p}_i, t_k), \alpha_{i,k}) \mid i \in \{1, 2, \ldots, N_{\text{train}}\}, k \in \{1, 1+m, 1+2m, \ldots, N_{\text{time}}\}\}$. In our particular implementation, we

selected $m = 10$ leading to $|\mathcal{D}_{\text{train}}^{\text{ForwComp}}| = 11,700$ samples. Note that in all other strategies, the number of training samples equals the number of training simulations, see Table 2, where the significant parameters of particular approximation strategies are briefly summarised.

Strategy	N_{ANN}	Inputs	Outputs	$|\mathcal{D}_{\text{train}}|$
Forward Complex	1	p_1, p_2, p_3, p_4, t_k	$\alpha_k \mid k \in \{1, 11, \ldots, 1161\}$	11700
Forward Split	9	p_1, p_2, p_3, p_4	$\alpha_{300}; \alpha_{400}; \ldots; \alpha_{1100}$	100
Forward Split II	22	p_1, p_2, p_3, p_4	$\alpha_{100}; \alpha_{150}; \ldots; \alpha_{1150}$	100
Forward Split III	43	p_1, p_2, p_3, p_4	$\alpha_{100}; \alpha_{130}; \alpha_{150}; \alpha_{170}; \ldots; \alpha_{1150}$	100
Error F_1	1	p_1, p_2, p_3, p_4	F_1	100
Error F_2	1	p_1, p_2, p_3, p_4	F_2	100
Inverse Expert	4	$\alpha_{300}, \alpha_{400}, \ldots, \alpha_{1100}$	$p_1; p_2; p_3; p_4$	100
Inverse Expert II	4	$\alpha_{200}, \alpha_{300}, \ldots, \alpha_{1100}$	$p_1; p_2; p_3; p_4$	100
Inverse PCA	4	$\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_9$	$p_1; p_2; p_3; p_4$	100

Table 2: Parameters of approximation strategies

The second way of the model output approximation is based on training an independent ANN for every time step t_k. Here, the particular ANN approximates simpler relation and span only the parameter space. A training data set for ANN approximating the response component α_k is thus given as $\mathcal{D}_{\text{train}}^{\text{ForwSpli},\alpha_k} = \{(\boldsymbol{p}_i, \alpha_{i,k}) \mid i \in \{1, 2, \ldots, 100\}\}$ having only $|\mathcal{D}_{\text{train}}^{\text{ForwSpli},\alpha_k}| = 100$ samples. A disadvantage of such an approach consists in training a large number N_{ANN} of smaller ANNs. As training of $N_{\text{ANN}} = N_{\text{time}} = 1161$ different ANNs can be almost unfeasible, we select only a few of the time steps, where the approximation is constructed and thus, the model output approximation is more rough. The choice of the important time steps and their number can be driven by the expert knowledge or results of the sensitivity analysis. Hence, we present three different choices so as to illustrate its influence, see Table 2. We further call these strategies as **Forward Split** (ForwSpli), **Forward Split II** (ForwSpliII) and **Forward Split III** (ForwSpliIII).

The error function approximation is the only strategy where the high dimensionality of the model output does not impose any complications. The model output is used for evaluation of the error function and the ANN is trained to approximate the mapping from the parameter space to a single scalar value of the error function, i.e. $\mathcal{D}_{\text{train}}^{\text{Error},F_a} = \{(\boldsymbol{p}_i, F_a) \mid i \in \{1, 2, \ldots, N_{\text{train}}\}\}$ and $|\mathcal{D}_{\text{train}}^{\text{Error},F_a}| = 100$, where F_a stands for a chosen error function. As we already mentioned in Section 3, there are two very common error functions given by Eqs. (3) and (4) and thus we investigate both considering the two strategies further called as **Error** F_1 and **Error** F_2, respectively.

In case of the inverse relation approximation, the high dimensionality of the model output needs again some special treatment so as to keep the number of ANN inputs and thus the ANN complexity reasonable. An intuitive approach is a simple selection of a limited number of output values $\boldsymbol{a} = A(\boldsymbol{\alpha})$. Here, one ANN is trained to predict one model parameter p_j and thus $\mathcal{D}_{\text{train}}^{\text{InvExp},p_j} = \{(\boldsymbol{a}_i, p_{i,j}) \mid i \in \{1, 2, \ldots, N_{\text{train}}\}\}$ and $|\mathcal{D}_{\text{train}}^{\text{InvExp},p_j}| = 100$. A particular choice of components in the vector $\boldsymbol{a}_i$ defined by the operator A should take into account not only the results of sensitivity analysis, but also a possible measurement error in experimental data as well as any other expert knowledge. Hence we present again two different choices in order to illustrate its influence, see Table 2 and we further call these configurations as **Inverse Expert** (InvExp) and **Inverse Expert II** (InvExpII).

In order to reduce the influence of the expert choice, the principal components $\bar{\alpha}$ computed as described in the previous section can be used as ANN inputs and one has to choose only their number. To compare the information contained in the same number of inputs selected by an expert, we have chosen the same number of principal components as the number of inputs in the Inverse Expert configuration and thus $\mathcal{D}_{\text{train}}^{\text{InvPCA},p_j} = \{((\bar{\alpha}_{i,1}, \ldots, \bar{\alpha}_{i,9}), p_{i,j}) \mid i \in \{1, 2, \ldots, N_{\text{train}}\}\}$ and $|\mathcal{D}_{\text{train}}^{\text{InvPCA},p_j}| = 100$. The principal components based strategy is further called **Inverse PCA** (InvPCA).

Then, the last preparatory step concerns the generation of testing data for a final assessment of the resulting ANNs consisting of $N_{\text{test}} = 50$ simulations for randomly generated sets of input parameters. The obtained data are then processed by particular approximation strategies in the same way as the training data described above.

7. Neural network training algorithm and topology choice

The quality of the ANN-based approximation estimated on a given data set $\mathcal{D}$ can be expressed as the mean relative prediction error $\varepsilon^{\text{MRP}}(\mathcal{D})$ given as

$$\varepsilon^{\text{MRP}}(\mathcal{D}) = \frac{\sum_{i=1}^{|\mathcal{D}|} |O_i - T_{i,\mathcal{D}}|}{|\mathcal{D}|(T_{\max,\mathcal{D}_{\text{train}}} - T_{\min,\mathcal{D}_{\text{train}}})} , \tag{12}$$

where O_i is the ANN output corresponding to the target value $T_{i,\mathcal{D}}$ contained in the data set $\mathcal{D}$, which consists of $|\mathcal{D}|$ samples. $T_{\max,\mathcal{D}_{\text{train}}}$ and $T_{\min,\mathcal{D}_{\text{train}}}$ are the maximal and minimal target values in the training data set $\mathcal{D}_{\text{train}}$, so the error $\varepsilon^{\text{MRP}}(\mathcal{D})$ is always scaled by the same factor for any chosen data set $\mathcal{D}$ and this factor corresponds to the range of the training data.

The conjugate gradient-based method [32] was applied as a training algorithm for synaptic weights computation and the cross-validation method was employed to determine the number of hidden neurons. In V-fold cross-validation we break the training data set $\mathcal{D}_{\text{train}}$ into V approximatelly equisized subsets $\mathcal{D}_{\text{train}} = \mathcal{D}_{\text{train},1} \cup \mathcal{D}_{\text{train},2} \cup \cdots \cup \mathcal{D}_{\text{train},V}$ and then we perform V training processes, each time leaving out one of the subsets $\mathcal{D}_{\text{train},i}$ and using the rest of the training data set $\mathcal{D}_{\text{train}} \setminus \mathcal{D}_{\text{train,i}}$.

The criterion for stopping the training process is governed by the prediction errors ratio r_k^{PE} computed at the k-th iteration of the training algorithm given as

$$r_k^{\text{PE}}(\mathcal{D}_{\text{train}} \setminus \mathcal{D}_{\text{train,i}}) = \frac{\sum_{j=k-J}^{k} \varepsilon_j^{\text{MRP}}(\mathcal{D}_{\text{train}} \setminus \mathcal{D}_{\text{train,i}})}{\sum_{j=k-2J}^{k-J-1} \varepsilon_j^{\text{MRP}}(\mathcal{D}_{\text{train}} \setminus \mathcal{D}_{\text{train,i}})} , \tag{13}$$

where $\varepsilon_j^{\text{MRP}}(\mathcal{D}_{\text{train}} \setminus \mathcal{D}_{\text{train,i}})$ is the mean relative prediction error obtained at the j-th iteration of the training algorithm obtained on the training data set without its i-th partition. J is the chosen number of iterations considered for computing the ratio r_k^{PE} for its smoothing effect on r_k^{PE}. The training process is stopped either when the number of iterations achieves its chosen maximal value K or if the prediction errors ratio r_k^{PE} exceeds a chosen critical value $r_{\max}^{\text{PE}}$.

Once the training process is completed, the ANN is evaluated on the remaining part of the training data $\mathcal{D}_{\text{train,i}}$, which was not used in the training process. The quality of the ANN with a particular number of hidden neurons h is assessed by the cross-validation error $\varepsilon_h^{\text{CV}}$,

which is computed as a mean of the errors obtained for the ANNs trained on the subsets $\mathcal{D}_{\text{train}} \setminus \mathcal{D}_{\text{train,i}}$ and then evaluated on the remaining subset $\mathcal{D}_{\text{train,i}}$, i.e.

$$\varepsilon_h^{\text{CV}} = \frac{1}{V} \sum_{i=1}^{V} \varepsilon^{\text{MRP}}(\mathcal{D}_{\text{train,i}}) . \tag{14}$$

We start with an ANN having $h_{\min}$ hidden neurons and we compute the corresponding cross-validation error. Then, one hidden neuron is added and after all the training processes on training data subsets, the new cross-validation error is evaluated. We compute the cross-validation error ratio r_h^{CVE} as

$$r_h^{\text{CVE}} = \varepsilon_h^{\text{CV}} / \varepsilon_{h-1}^{\text{CV}} . \tag{15}$$

We count the situations when the ratio r_h^{CVE} exceeds a chosen critical value $r_{\max}^{\text{CVE}}$. If this happened W times, the addition of hidden neurons is stopped. Then we choose the architecture having the smallest cross-validation error $\varepsilon_h^{\text{CV}}$ and the particular ANN with the synaptic weights having the smallest training error ε^{MRP}.

Number of subsets in cross-validation	V	10
Number of iteration considered in r_k^{PE}	J	100
Maximal number of training iterations	K	5000
Maximal value of prediction errors ratio	$r_{\max}^{\text{PE}}$	0.999
Starting value of hidden neurons	$h_{\min}$	1
Maximal value of cross-validation error ratio	$r_{\max}^{\text{CVE}}$	0.99
Maximal value of $r_{\max}^{\text{CVE}}$ exceeding	W	3

Table 3: Parameters of ANN training algorithm and cross-validation method

The resulting ANNs are tested on an independent testing data set $\mathcal{D}_{\text{test}}$. Since some of the approximation strategies consist of a high number of ANNs, the resulting number of hidden neurons and achieved errors on training and testing data for all the trained ANNs are listed in Appendix A. Brief summary of these results is presented in Table 4[1].

Strategy	h	$\varepsilon^{\text{MRP}}(\mathcal{D}_{\text{train}})[\%]$	$\varepsilon^{\text{MRP}}(\mathcal{D}_{\text{test}})[\%]$
Forward Complex	7	2.03	2.67
Forward Split	3 to 10	0.06 to 1.06	0.06 to 1.27
Forward Split II	4 to 13	0.06 to 1.42	0.07 to 2.04
Forward Split III	3 to 13	0.03 to 1.50	0.03 to 1.98
Error F_1	10	0.40 to 0.54	0.57 to 0.74
Error F_2	9 to 11	0.78 to 1.36	0.96 to 1.56
Inverse Expert	5 to 8	1.14 to 5.74	1.31 to 6.43
Inverse Expert II	4 to 6	1.38 to 5.79	1.36 to 6.52
Inverse PCA	4 to 8	0.28 to 10.50	0.33 to 16.73

Table 4: Architecture of particular ANNs in inverse strategies and their errors on training and testing data.

---

[1]The error function approximation strategies are intrinsically related to particular experimental curve. The results here are obtained for experimental "Mokra" data described in Section 9 in more detail.

Regarding the number of hidden neurons, the results point to higher complexity of the error function relationships. Nevertheless, the differences in hidden neurons among particular strategies are relatively small.

The quality of the resulting ANNs in approximation of the given relationships is measured by the obtained errors on all the training $\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{train}})$ and testing $\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{test}})$ data. Small differences between the training and testing errors refer to well-trained ANNs and to the good quality of the training method as well as the method for topology estimation. Note that overtrained ANNs usually lead to significantly higher errors on testing data.

Comparing the approximation quality of the particular strategies, we can point out good results of the forward model approximation and error function approximation, where the errors did not exceed the value of 3 %. The good approximation of the forward model is not surprising since the relationship is well-defined, smooth and relatively simple. The good results of the error function approximation are more unexpected, because the relationship here is probably more nonlinear and complex. One possible explanation is a large spread of error function values on the training data, which is used to scale the errors (see Eq. (12)). While the error functions converge to zero near the optimal parameter values, they quickly rise to extremely high values for parameter values more distant from the optimum. Hence, we presume that the small errors obtained in the error function approximation do not promise comparably good results in the final parameter identification.

The results of the inverse relation approximation are not very good, but it was foreseen due to unknown and probably ill-posed relationship. Nevertheless, the obtained errors are actually the final errors of the whole identification process for the training and testing data, since there is no other following step concerning any optimisation as in the case of other identification strategies. Hence, further comments on these results are presented in the following section concerning verification of the overall identification strategies on the testing data.

8. Verification of model calibration

Since the errors in Table 4 represent only the quality of the constructed ANNs, we have to also investigate the quality of the identification procedures. This section is devoted to verification of model calibration, where the goal is to predict the model parameters' values corresponding to the simulated data, which are not perturbated by any noise. The advantage of verification is that we also know the true values of the parameters and thus, we can easily evaluate the quality of their estimation by each strategy. In particular, the calibration strategies were applied to estimate the parameters' values for all the training and testing simulations.

As mentioned, in case of the inverse relation approximation, the outputs of ANNs are directly the predicted values of the identified parameters $\widehat{\boldsymbol{p}}$. In case of the forward model approximation, we have to run a subsequent optimisation process. Here, the evolutionary algorithm GRADE, see [9] for details about this method[2], is applied to find a set of parameters' values $\widehat{\boldsymbol{p}}$ minimising the square distance δ between components of the model response

---

[2]The parameters of GRADE algorithm were set to pool_rate = 4, radioactivity = 0.33 and cross_limit = 0.1. The algorithm was stopped after 10000 cost function evaluations.

α_k and their corresponding ANN-based approximated counterparts $\widetilde{\alpha}_k$, i.e.

$$\delta = \sum_k (\alpha_k - \widetilde{\alpha}_k)^2 \,, \tag{16}$$

where k corresponds to selected approximated components defined for particular identification strategies in Table 2. In such a way, the parameters $\widehat{p}$ are predicted for all the training as well as testing data. As the true values of parameters p are known in the verification process, the mean prediction errors $\widehat{\varepsilon}$ are computed relatively to the spread of the training data, i.e.

$$\widehat{\epsilon}(\widehat{p}_j) = \frac{\sum_{i=1}^{|\mathcal{D}|} |p_{i,j} - \widehat{p}_{i,j}|}{|\mathcal{D}|(p_{\max(\mathcal{D}_{\text{train}}),j} - p_{\min(\mathcal{D}_{\text{train}}),j})} \,, \tag{17}$$

and the obtained errors for particular identification strategies are listed in Table 5. In applica-

	$\widehat{\varepsilon}(\widehat{p}_1)$		$\widehat{\varepsilon}(\widehat{p}_2)$		$\widehat{\varepsilon}(\widehat{p}_3)$		$\widehat{\varepsilon}(\widehat{p}_4)$		$\widehat{\varepsilon}(\widehat{\alpha})$	
	train	test	train	test	train	test	train	test	train	test
Forward Complex	16.78	17.09	52.20	47.91	6.06	5.45	3.67	2.69	1.08	1.09
Forward Split	9.48	11.62	30.18	38.45	3.14	4.65	1.17	3.10	0.31	0.37
Forward Split II	5.09	6.47	13.34	15.03	1.69	2.60	0.67	1.02	0.14	0.21
Forward Split III	4.12	4.84	10.73	10.65	1.49	1.63	0.57	0.64	0.12	0.16
Inverse Expert	5.74	6.43	5.15	6.21	1.99	2.16	1.14	1.31	0.49	0.49
Inverse Expert II	5.79	6.23	5.60	6.52	2.60	3.18	1.38	1.36	0.44	0.53
Inverse PCA	3.86	5.10	10.50	16.73	1.25	1.89	0.28	0.33	0.38	1.21

Table 5: Results of verification of particular identification strategies in terms of mean relative prediction errors $\widehat{\varepsilon}$ [%].

tion of identification strategy to real experimental data, the parameter values are not known, but the success of the identification process is quantified by quality of fitting the data by the model response obtained for the identified parameters. Hence, the model simulations were performed for all the identified parameter sets and prediction errors $\widetilde{\varepsilon}$ in terms of predicted responses $\widetilde{\alpha}$ are computed analogously to the Eq. (17). Their values averaged also over all the response components are then listed in Table 5.

The results for strategies based on an approximation of the error function are missing here, because they require to build a particular ANN for every curve of the hydration degree and for each require to run an additional minimisation procedure. This is overwhelming and thus these strategies are only validated on experimental data as described in the following section.

One can see that among the forward strategies, the complex variant provided the worst results in the training process as well as in the final identification. The complex relationship covering the time domain cause apparently certain difficulties to the training process. We can conclude that training of a set of neural networks means more work, but offers significantly better quality of the model approximation. We can also point out the large differences in errors of particular parameters, which correspond to influence of particular parameters to the model response. As demonstrated in Figure 3, the largest spread of the model response is related namely to change in the parameters p_4 and p_3, while the parameter p_1 and even more p_2 seem to be almost negligible. The sensitivity analysis illustrated in Figure 4b shows

very high sensitivity of the model response to the parameter p_2 at early stage of hydration, nevertheless, at this stage the spread of the model response is almost negligible and even a very small error in the response approximation can be fatal for identification of the parameter p_2. On the other hand, it is not surprising that the identification accuracy is significantly improved with an increasing number of approximated response components, i.e. increasing number of trained ANNs.

Despite the worse results in training of ANNs, the inverse strategies achieved comparably good results with the forward strategies in parameter identification and also in fitted measurements. More precisely, the results of measurements fitting are slightly worse, but the errors in parameter prediction are smaller. Especially the Inverse Expert strategies provided surprisingly small errors in p_2 prediction and the errors in parameters are generally more balanced. This phenomenon can be possibly explained by fact that each ANN is trained to predict each parameter separately, thus automatically selecting and emphasizing the combinations of model response critical for the parameter. In the strategy Inverse Expert II, the usage of one additional input at the early stage of hydration caused no improvement of the resulting prediction, which is probably caused again by fact that the responses at this stage have negligible spread and almost no predictive value. Last interesting result concerns the application of principal component analysis. The Inverse PCA strategy provided again significantly different errors in prediction of particular parameters, similarly to the forward strategies. The reason resides possibly in fact that PCA emphasize the most important components, while it can mix the effects of the less significant parameters. Nevertheless, when compared with strategies Forward Split and Inverse Expert using the same number of response components, the Inverse PCA provided the best results in prediction of all the parameters except p_2. Its quality of measurement fitting is, however, the worst among those strategies.

From this thorough comparison we may conclude that all the inverse strategies provide very good results, which makes them highly promising considering their very simple implementation which does not include any additional optimisation process except the only training of ANNs. Moreover, the Inverse Expert strategies can be especially recommended for identification of less significant parameters.

9. Validation of model calibration

The previous section was focused on mutual comparison of the presented identification strategies on simulated data. However, a complete comparison has to include their validation on experimental data. To that purpose we used the four experimental data obtained by isothermal calorimetry: one for cement "**Mokra**" CEM I 42.5 R taken directly from Heidelberg cement group's kiln in Mokrá, Czech Republic [33] and three others from the following literature: "**Boumiz**" [34], "**Hua**" [35] and "**Princigallo**" [36].

In parameter identification from experimental data, one often face to difficulties related to (i) experimental errors and (ii) model imperfections. Especially in case of models with parameters having a specific physical meaning – like affinity hydration model – it happens that the experimental data seems to lie beyond the physically meaningful values of the model parameters. This is exactly what we face in case of the four experimental curves depicted in Figure 6. The grey curves represent the training samples generated in an optimised fashion so as to maximally cover the parameter space. Nevertheless, it is visible that all the experimental curves lie out of the bundle of the training samples. Applying the identification strategies to
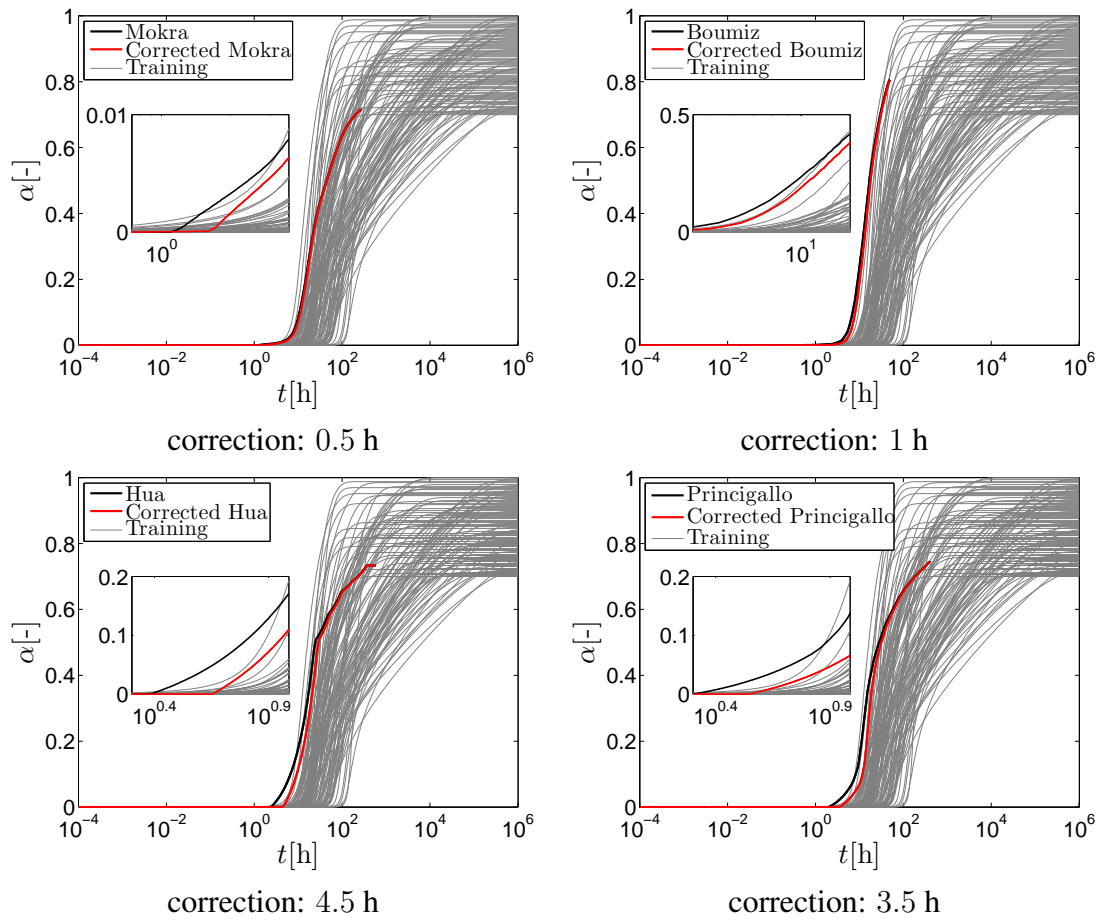
Figure 6: Corrections of experimental curves.

these data will require the ANNs to extrapolate and it will probably lead to unphysical and wrong predictions of the model parameters. Such results were presented for "Mokra" in [37]. Looking in more detail on the experimental curves, one can see that the difference between the experimental data and simulations can be explained by wrong estimation of the origin of hydration. Correction of the starting time moves the curves into the bundle of response simulations. As a matter of fact, the correction in orders of hours is negligible comparing to the duration of the whole hydration process lasting often days or weeks. Moreover, the goal of this paper is not to argue against the correctness of the model or data, but to demonstrate the properties of particular identification strategies which can be better illustrated in situation, where the observed data are not outliers w.r.t. sampled parameter domain. For an interested reader about the identification of outliers we refer to [37].

In general, validation does not allow for a comparison in terms of parameters' values, because these are not known a priori. Nevertheless, the simplicity and the fast simulation of the affinity hydration model permit a direct optimisation of the model parameters so as to fit the measured data without any incorporated approximation. The resulting optimal solutions can be then compared with the results obtained using the ANN approximations. To that purpose, we employ again the error functions given in Eqs. (3) and (4) and the GRADE algorithm with the same setting as in the previous section to minimise the both error functions. The obtained results are referred to as **Direct1** and **Direct2**, respectively.

Subsequently, the identification strategies were applied to the experimental data using the prepared ANNs. Since the ANNs are constructed for specific time steps of the hydration degree, the experimental curves are interpolated to the time steps required by the particular ANNs. If necessary, the data are extrapolated beyond the last measured time step assuming the further progress of hydration to be constant at the last measured value. The identified parameters together with the parameters' values obtained by the direct optimisation are written in Tables 6 and 7. The identified parameters were used as inputs for simulations, whose

Method	"Mokra"					"Boumiz"				
	p_1	p_2	p_3	p_4	$\widehat{\varepsilon}(\widehat{\alpha})$	p_1	p_2	p_3	p_4	$\widehat{\varepsilon}(\widehat{\alpha})$
Direct1	0.84	0.99	0.18	0.05	0.70	0.93	1.00	0.02	0.36	2.37
Direct2	0.82	0.98	0.18	0.05	0.65	0.93	1.00	0.02	0.35	2.70
Forward Complex	0.81	1.00	0.18	0.03	1.35	1.00	0.61	0.08	0.36	12.67
Forward Split	0.82	1.00	0.19	0.05	1.15	0.96	1.00	0.08	0.35	5.44
Forward Split II	0.78	1.01	0.18	0.05	0.83	1.00	1.00	0.08	0.35	4.11
Forward Split III	0.80	1.00	0.19	0.05	0.91	0.98	1.00	0.05	0.35	3.03
Error F_1	0.78	0.73	0.09	0.07	3.89	-	-	-	-	-
Error F_2	1.00	1.19	0.15	-0.06	2.73	-	-	-	-	-
Inverse Expert	1.16	-0.18	0.29	0.03	6.83	0.78	-0.24	0.22	0.30	35.11
Inverse Expert II	1.21	-0.06	0.19	0.16	4.68	1.27	-0.14	0.20	0.13	25.94
Inverse PCA	0.75	0.83	0.18	0.06	1.82	0.78	0.87	0.02	0.35	10.82

Table 6: Results of identification strategies obtained for "Mokra" and "Boumiz": identified values of model parameters and mean relative error in degree of hydration $\widehat{\varepsilon}(\widehat{\alpha})$ [%].

Method	"Hua"					"Princigallo"				
	p_1	p_2	p_3	p_4	$\widehat{\varepsilon}(\widehat{\alpha})$	p_1	p_2	p_3	p_4	$\widehat{\varepsilon}(\widehat{\alpha})$
Direct1	1.00	0.94	0.20	0.11	2.24	1.00	0.85	0.19	0.14	3.46
Direct2	0.99	0.96	0.21	0.11	2.46	1.00	0.88	0.21	0.15	3.27
Forward Complex	1.00	0.64	0.22	0.08	4.10	1.00	0.58	0.23	0.14	6.21
Forward Split	0.87	1.00	0.19	0.11	2.84	0.78	0.98	0.18	0.15	4.39
Forward Split II	0.93	0.96	0.21	0.11	2.92	0.92	0.82	0.20	0.14	4.44
Forward Split III	0.87	1.01	0.18	0.10	2.71	0.89	0.92	0.18	0.14	3.75
Inverse Expert	0.94	-0.29	0.26	0.12	10.64	1.07	-0.16	0.22	0.15	9.02
Inverse Expert II	1.26	-0.27	0.19	0.02	6.23	1.52	-1.38	0.13	-0.24	15.05
Inverse PCA	1.00	0.89	0.15	0.12	2.41	1.13	0.74	0.19	0.15	3.62

Table 7: Results of identification strategies obtained for "Hua" and "Princigallo": identified values of model parameters and mean relative error in degree of hydration $\widehat{\varepsilon}(\widehat{\alpha})$ [%].

results are compared with the experimental data in Figures 7 and 8. To quantify the quality of obtained fits, Tables 6 and 7 contain also the mean relative error $\widehat{\varepsilon}(\widehat{\alpha})$ [%] computed in the same manner as in Table 5 for an easy comparison of the verification and validation results.

The strategies based on the error function approximation are illustrated on parameter identification from "Mokra" data, which are used to define the error functions, which are approximated by ANNs. Trained ANNs are then optimised by GRADE algorithm so as to provide the optimal set of identified parameters. As we presumed, the identification results
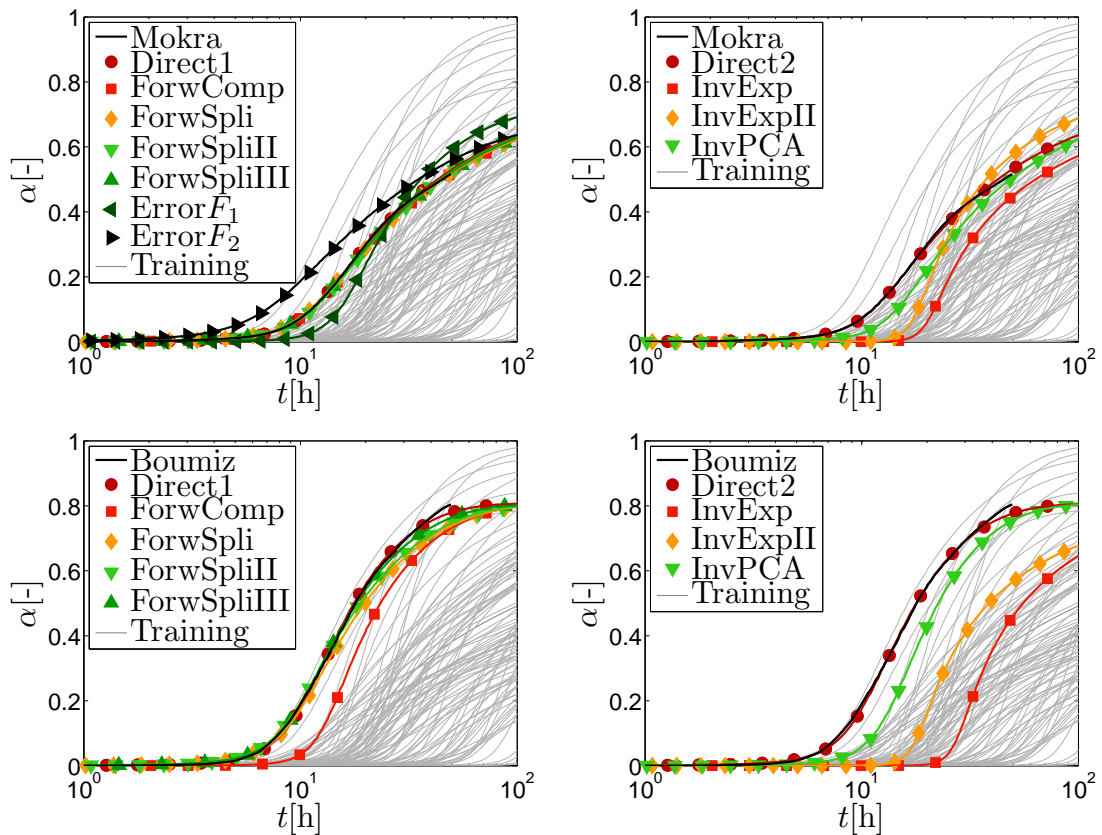
Figure 7: Comparison of corrected experimental data "Mokra" and "Boumiz" and corresponding results of calibration strategies.

are not satisfactory despite ostensibly very good results of the ANNs' training processes, see Table 4. The training and testing errors are small relatively to the spread of error functions' values, which increase quickly with the distance from the optimal solution. The strategy, however, requires high precision of the ANN's approximation near the optimal solution, which can be hardly achieved due the overall complex shape of the error functions.

The worst results on all the experimental curves were obtained by the inverse strategies based on selected components of the model response used as ANNs' inputs. The results pointed out the high sensitivity of this strategy to measurement noise and to specific choice of the inputs. Both drawbacks are overcome by employing principal component analysis, which allows to employ high number of response components and filter the measurement noise out of the several first principal components. The Inverse PCA strategy thus achieved significantly better results.

The forward strategies provided generally the best results consistent with the results of the verification on simulated data. These strategies thus proved to be rather immune to the noise in experimental data.

10. Conclusions

The presented paper reviews and compares several possible applications of artificial neural networks in calibration of numerical models. In particular, the feedforward layered neural network is employed in three basic schemes to surrogate: (i) response of a model, (ii) inverse
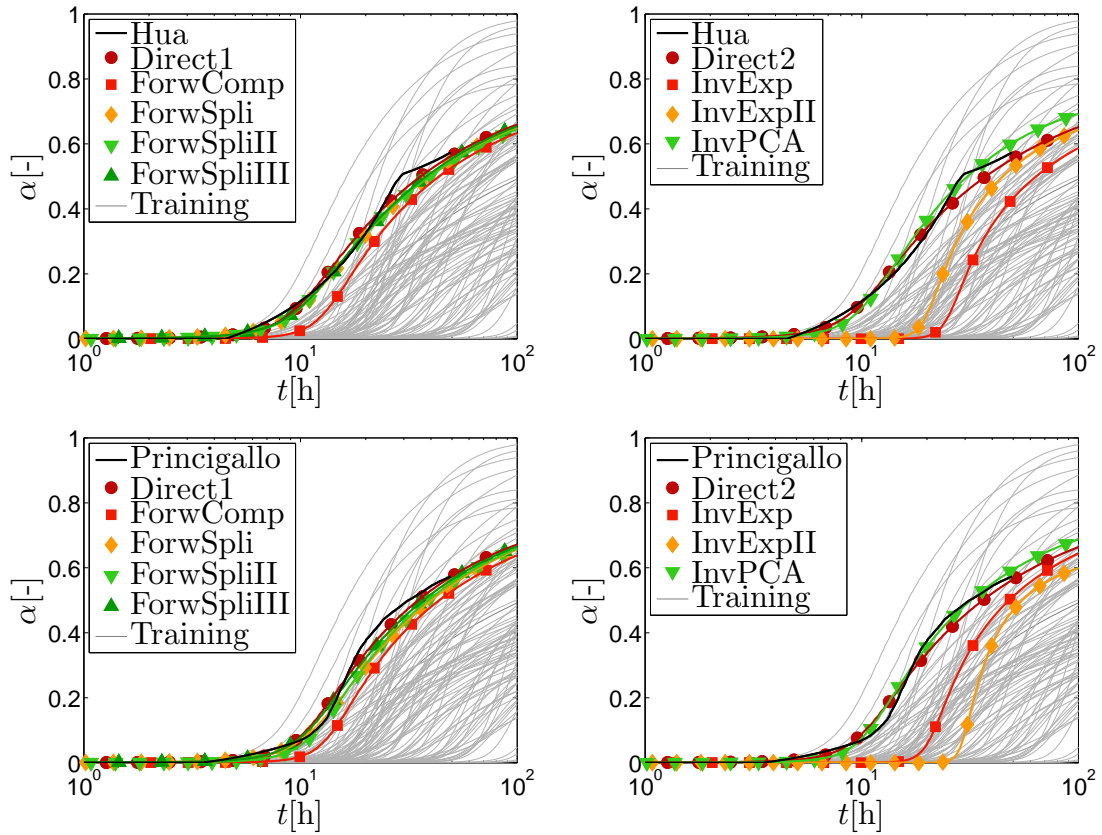
Figure 8: Comparison of corrected experimental data "Hua" and "Princigallo" and corresponding results of calibration strategies.

relationship of model parameters and model response and (iii) error function quantifying how well the model response fits the experimental data. Their advantages and drawbacks are illustrated on calibration of four parameters of the affinity hydration model. The model is chosen for its nonlinearity, difference in sensitivities to particular parameters on one hand and simplicity and very fast numerical evaluation on the other. The later allow for model calibration based on stochastic evolutionary algorithm without any involved approximation and thus better quantification of calibration results provided by particular strategies. Investigated calibration strategies are verified on 50 simulated curves of hydration degree and validated on four experimental ones.

Strategy	N_{ANN}	optimisation	new data	errors
Forward Complex	1	yes	optimisation	middle
Forward Split	N_α	yes	optimisation	low
Error F	1	yes	traininng + optimisation	high
Inverse Expert	N_{p}	no	-	high
Inverse PCA	N_{p}	no	-	middle

Table 8: Simplified summary of calibration strategies. N_α stands for a number of approximated components of model response, N_{p} is a number of model parameters.

Simplified summary of the obtained results is written in Table 8. One of the simplest strategies from the implementation point of view is based on approximation of the error

function (Error F), where only one neural network needs to be trained for the prediction of the error function values. This simplicity, however, does not hold in case of multiple experimental measurements, where the whole identification process including the neural network training as well as its optimisation needs to be done all over again for any new experiment. Moreover, the presented examples revealed that the complexity of the error function may cause difficulties for neural network training resulting in high errors in identified parameters. The potential of the neural network is wasted on approximating the whole domain, while the accurate predictions are required only in the vicinity of the optimal values of parameters. Hence, this strategy is more suited for surrogate models based on radial basis function networks or kriging, which can be trained along with the optimisation of the error function thus allowing to improve the precision in the promising area, see e.g. [20].

Equally simple strategy is based on the approximation of the model response, where time or space variables are included among the neural network inputs (Forward Complex). This strategy is better suited for layered neural networks, which is trained only once and then can be used repeatedly for any new observations. The effort invested into the approximation of the whole domain is thus not wasted. The application to new data requires only one new optimisation process. The results obtained by this strategy were not excellent, but can be considered as satisfactory solution at a low price.

The best results were achieved by separate approximations of particular response components, where a higher number of neural networks is trained to approximate rather simple relationship defined by the calibrated model (Forward Split). This procedure requires more work on networks preparation, which is compensated by high accuracy of the obtained results. The accuracy is proportionally increasing with the number of approximated response components and can be thus influenced by work invested to the surrogate construction. Moreover, the constructed approximations can be then used again for any new data, where only the optimisation of model parameters needs to be repeated.

The worst results were obtained by strategy approximating the inverse mapping from the response components to the model parameters (Inverse Expert). Such relationship does not have to exist and can be hardly approximated. Moreover, if the inputs for neural network are not properly selected and thus highly sensitive to measurement error, the procedure provide unsatisfactory results. Nevertheless, using an expert knowledge for proper selection of inputs as presented in [22], this strategy gives good results at a very low price, since neither training nor optimisation process, but only simple evaluation of the trained networks is needed for parameter identification from new data.

The necessity of the expert knowledge and sensitivity to measurement error can be easily circumvented by employing principal component analysis on model response components (Inverse PCA). Then only the number of components entering as inputs in the neural network needs to be selected. The strategy thus represents a compromise solution providing satisfactory results at a low price especially in repeated application to new observed data.

References

[1] G. Stavroulakis, G. Bolzon, Z. Waszczyszyn, L. Ziemianski, 3.13 - inverse analysis, in: E. in Chief: I. Milne, R. O. Ritchie, , B. Karihaloo (Eds.), Comprehensive Structural Integrity, Pergamon, Oxford, 2003, pp. 685 – 718.

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

[3] A. Kučerová, J. Sýkora, B. Rosić, H. G. Matthies, Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials, Journal of Computational and Applied Mathematics 236 (18) (2012) 4862–4872.

[4] E. T. Jaynes., Probability Theory, The Logic of Science, Cambridge University Press, Cambridge, 2003.

[5] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, 2005.

[6] H. G. Matthies, Encyclopedia of Computational Mechanics, John Wiley & Sons, Ltd., 2007, Ch. Uncertainty Quantification with Stochastic Finite Elements.

[7] B. V. Rosić, A. Kučerová, J. Sýkora, O. Pajonk, A. Litvinenko, H. G. Matthies, Parameter identification in a probabilistic setting, Engineering Structures 50 (2013) 179–196.

[8] J.-S. R. Jang, C.-T. Sun, Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1997.

[9] A. Kučerová, Identification of nonlinear mechanical model parameters based on soft-computing methods, Ph.D. thesis, Ecole Normale Supérieure de Cachan, Laboratoire de Mécanique et Technologie (2007).

[10] K. N. Gurney, An introduction to neural networks, UCL Press, London, 2002.

[11] S. S. Haykin, Neural networks and learning machines, 3rd Edition, Prentice Hall/Pearson, New York, 2009.

[12] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989) 359–366.

[13] K. Hornik, Some new results on neural network approximation, Neural Networks 6 (1993) 1069–1072.

[14] U. Anders, O. Korn, Model selection in neural networks, Neural Networks 12 (2) (1999) 309–323.

[15] J. Moody, Prediction risk and architecture selection for neural networks, From Statistics to Neural Networks, NATO ASI Series 136 (1994) 147–165.

[16] M. Abendroth, M. Kuna, Identification of ductile damage and fracture parameters from the small punch test using neural networks, Engineering Fracture Mechanics 73 (2006) 710–725.

[17] B. Pichler, R. Lackner, H. A. Mang, Back analysis of model parameters in geotechnical engineering by means of soft computing, International journal for numerical methods in engineering 57 (2003) 1943–1978.

[18] H. Aguir, H. BelHadjSalah, R. Hambli, Parameter identification of an elasto-plastic behaviour using artificial neural networks–genetic algorithm method, Materials and Design 32 (2011) 48–53.

[19] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, P. K. Tucker, Surrogate-based analysis and optimization, Progress in Aerospace Sciences 41 (2005) 1–28.

[20] A. Kučerová, D. Brancherie, A. Ibrahimbegović, J. Zeman, Z. Bittnar, 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.

[21] D. Novák, D. Lehký, ANN inverse analysis based on stochastic small-sample training set simulation, Engineering Applications of Artificial Intelligence 19 (2006) 731–740.

[22] A. Kučerová, M. Lepš, Soft computing-based calibration of microplane M4 model parameters: Methodology and validation, Advances in Engineering Software 72 (2014) 226–235.

[23] K. Zaw, G. R. Liu, B. Deng, K. B. C. Tan, Rapid identification of elastic modulus of the interface tissue on dental implants surfaces using reduced-basis method and a neural network, Journal of Biomechanics 42 (2009) 634–641.

[24] L. Zhang, L. Li, H. Ju, B. Zhu, Inverse identification of interfacial heat transfer coefficient between the casting and metal mold using neural network, Energy Conversion and Management 51 (2010) 1898–1904.

[25] M. Klos, Z. Waszczyszyn, Modal analysis and modified cascade neural networks in identification of geometrical parameters of circular arches, Computers and Structures 89 (2011) 581–589.

[26] W. R. L. da Silva, V. Šmilauer, Fuzzy affinity hydration model, Journal of Intelligent & Fuzzy Systems 28 (2015) 127–139.

[27] M. Cervera, J. Oliver, T. Prato, Thermo-chemo-mechanical model for concrete. i: Hydration and aging, Journal of Engineering Mechanics ASCE 125 (9) (1999) 1018–1027.

[28] D. Gawin, F. Pesavento, B. Schrefler, Hygro-thermo-chemo-mechanical modelling of concrete at early ages and beyond. part i: Hydration and hygro-thermal phenomena, International Journal for Numerical Methods in Engineering 67 (3) (2006) 299–331.

[29] J. Helton, J. Johnson, C. Sallaberry, C. Storlie, Survey of sampling-based methods for uncertainty and sensitivity analysis, Reliability Engineering & System Safety 91 (2006) 1175–1209.

[30] A. Saltelli, K. Chan, E. M. Scott, Sensitivity analysis, NY:Wiley, New York, 2000.

[31] E. Janouchová, A. Kučerová, Competitive comparison of optimal designs of experiments for sampling-based sensitivity analysis, Computers & Structures 124 (2013) 47–60.

[32] J. R. Shewchuk, An introduction to the conjugate gradient method without the agonizing pain, Tech. rep., Pittsburgh, PA, USA (1994).

[33] V. Šmilauer, Multiscale Hierarchical Modeling of Hydrating Concrete, Saxe-Coburg Publications, Stirling, 2015.

[34] A. Boumiz, C. Vernet, F. C. Tenoudji, Mechanical properties of cement pastes and mortars at early ages, Advanced Cement Based Materials 3 (3–4) (1996) 94–106.

[35] C. Hua, P. Acker, A. Ehrlacher, Analyses and models of the autogenous shrinkage of hardening cement paste. I. modelling at macroscopic scale, Cement and Concrete Research 25 (7) (1995) 1457–1468.

[36] A. Princigallo, P. Lura, K. van Breugel, G. Levita, Early development of properties in a cement paste: A numerical and experimental study, Cement and Concrete Research 33 (7) (2003) 1013–1020.

[37] T. Mareš, E. Janouchová, A. Kučerová, Application of artificial neural networks in identification of affinity hydration model parameters, in: Proceedings of the Eighth International Conference on Engineering Computational Technology, Vol. 100, Civil-Comp Press Ltd, Stirling, 2012.

Appendix A. Configurations and results of particular neural networks

The particular choice of ANN inputs and outputs are presented in Tables A.9 and A.11 for forward and inverse mode strategies, respectivelly.

Strategy	Inputs	h	Output	$\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{train}})[\%]$	$\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{test}})[\%]$
Forward Complex	p_1, p_2, p_3, p_4, t_k	7	α_k	2.03	2.67
Forward Split	p_1, p_2, p_3, p_4	7	α_{300}	0.06	0.06
	p_1, p_2, p_3, p_4	8	α_{400}	0.07	0.12
	p_1, p_2, p_3, p_4	7	α_{500}	0.08	0.11
	p_1, p_2, p_3, p_4	2	α_{600}	0.62	1.01
	p_1, p_2, p_3, p_4	7	α_{700}	0.79	1.01
	p_1, p_2, p_3, p_4	6	α_{800}	1.06	1.27
	p_1, p_2, p_3, p_4	8	α_{900}	0.28	0.32
	p_1, p_2, p_3, p_4	10	α_{1000}	0.22	0.27
	p_1, p_2, p_3, p_4	9	α_{1100}	0.21	0.30
Forward Split II	p_1, p_2, p_3, p_4	7	α_{100}	0.52	0.88
	p_1, p_2, p_3, p_4	4	α_{150}	0.86	1.39
	p_1, p_2, p_3, p_4	4	α_{200}	0.08	0.11
	p_1, p_2, p_3, p_4	8	α_{250}	0.68	0.80
	p_1, p_2, p_3, p_4	4	α_{300}	0.44	0.60
	p_1, p_2, p_3, p_4	5	α_{350}	0.48	0.95
	p_1, p_2, p_3, p_4	4	α_{400}	0.06	0.07
	p_1, p_2, p_3, p_4	6	α_{450}	0.07	0.10
	p_1, p_2, p_3, p_4	6	α_{500}	0.07	0.13
	p_1, p_2, p_3, p_4	9	α_{550}	0.15	0.22
	p_1, p_2, p_3, p_4	6	α_{600}	1.42	2.04
	p_1, p_2, p_3, p_4	5	α_{650}	0.84	1.19
	p_1, p_2, p_3, p_4	6	α_{700}	0.55	0.73
	p_1, p_2, p_3, p_4	8	α_{750}	0.60	1.18
	p_1, p_2, p_3, p_4	7	α_{800}	0.46	0.62
	p_1, p_2, p_3, p_4	9	α_{850}	0.75	1.09
	p_1, p_2, p_3, p_4	7	α_{900}	0.20	0.23
	p_1, p_2, p_3, p_4	13	α_{950}	0.28	0.43
	p_1, p_2, p_3, p_4	9	α_{1000}	0.73	1.16
	p_1, p_2, p_3, p_4	6	α_{1050}	0.10	0.17
	p_1, p_2, p_3, p_4	9	α_{1100}	0.10	0.19
	p_1, p_2, p_3, p_4	7	α_{1150}	0.08	0.14
Error F_1	p_1, p_2, p_3, p_4	10	F_1 for Mokra	0.54	0.74
	p_1, p_2, p_3, p_4	10	F_1 for shifted Mokra	0.40	0.57
Error F_2	p_1, p_2, p_3, p_4	9	F_2 for Mokra	0.78	0.96
	p_1, p_2, p_3, p_4	9	F_2 for shifted Mokra	1.36	1.56

Table A.9: Architecture of particular ANNs constructed in forward strategies and their errors on training and testing data.

Strategy	Inputs	h	Output	$\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{train}})[\%]$	$\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{test}})[\%]$
Forward Split III	p_1, p_2, p_3, p_4	7	α_{100}	0.07	0.08
	p_1, p_2, p_3, p_4	4	α_{130}	0.56	0.45
	p_1, p_2, p_3, p_4	4	α_{150}	1.03	0.95
	p_1, p_2, p_3, p_4	8	α_{170}	0.08	0.06
	p_1, p_2, p_3, p_4	4	α_{200}	0.76	0.74
	p_1, p_2, p_3, p_4	5	α_{230}	0.44	0.41
	p_1, p_2, p_3, p_4	4	α_{250}	0.49	0.45
	p_1, p_2, p_3, p_4	6	α_{270}	0.07	0.06
	p_1, p_2, p_3, p_4	6	α_{300}	0.07	0.05
	p_1, p_2, p_3, p_4	9	α_{330}	0.07	0.11
	p_1, p_2, p_3, p_4	6	α_{350}	0.16	0.37
	p_1, p_2, p_3, p_4	5	α_{370}	1.47	1.90
	p_1, p_2, p_3, p_4	6	α_{400}	0.84	1.06
	p_1, p_2, p_3, p_4	8	α_{430}	0.59	0.96
	p_1, p_2, p_3, p_4	7	α_{450}	0.71	0.92
	p_1, p_2, p_3, p_4	9	α_{470}	0.54	0.55
	p_1, p_2, p_3, p_4	7	α_{500}	0.89	0.98
	p_1, p_2, p_3, p_4	13	α_{530}	0.23	0.40
	p_1, p_2, p_3, p_4	9	α_{550}	0.30	0.44
	p_1, p_2, p_3, p_4	6	α_{570}	0.73	0.63
	p_1, p_2, p_3, p_4	9	α_{600}	0.12	0.20
	p_1, p_2, p_3, p_4	7	α_{630}	0.11	0.18
	p_1, p_2, p_3, p_4	7	α_{650}	0.07	0.08
	p_1, p_2, p_3, p_4	4	α_{670}	0.55	0.49
	p_1, p_2, p_3, p_4	6	α_{700}	0.07	0.09
	p_1, p_2, p_3, p_4	8	α_{730}	0.06	0.06
	p_1, p_2, p_3, p_4	9	α_{750}	0.07	0.06
	p_1, p_2, p_3, p_4	8	α_{770}	0.03	0.03
	p_1, p_2, p_3, p_4	8	α_{800}	0.05	0.04
	p_1, p_2, p_3, p_4	5	α_{830}	0.09	0.10
	p_1, p_2, p_3, p_4	5	α_{850}	0.77	0.42
	p_1, p_2, p_3, p_4	3	α_{870}	0.23	0.27
	p_1, p_2, p_3, p_4	6	α_{900}	1.06	0.99
	p_1, p_2, p_3, p_4	7	α_{930}	1.50	1.88
	p_1, p_2, p_3, p_4	8	α_{950}	0.37	0.49
	p_1, p_2, p_3, p_4	7	α_{970}	1.38	1.98
	p_1, p_2, p_3, p_4	7	α_{1000}	0.93	1.05
	p_1, p_2, p_3, p_4	8	α_{1030}	0.26	0.35
	p_1, p_2, p_3, p_4	7	α_{1050}	0.83	0.87
	p_1, p_2, p_3, p_4	6	α_{1070}	1.12	1.04
	p_1, p_2, p_3, p_4	8	α_{1100}	0.31	0.36
	p_1, p_2, p_3, p_4	11	α_{1130}	0.13	0.20
	p_1, p_2, p_3, p_4	7	α_{1150}	0.14	0.20

Table A.10: Architecture of particular ANNs constructed in forward strategies and their errors on training and testing data.

Strategy	Inputs	h	Output	$\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{train}})[\%]$	$\varepsilon^{\mathrm{MRP}}(\mathcal{D}_{\mathrm{test}})[\%]$
Inverse Expert	9 values: $\alpha_{300}, \alpha_{400}, \ldots, \alpha_{1100}$	5	p_1	5.74	6.43
	9 values: $\alpha_{300}, \alpha_{400}, \ldots, \alpha_{1100}$	7	p_2	5.15	6.21
	9 values: $\alpha_{300}, \alpha_{400}, \ldots, \alpha_{1100}$	8	p_3	1.99	2.16
	9 values: $\alpha_{300}, \alpha_{400}, \ldots, \alpha_{1100}$	5	p_4	1.14	1.31
Inverse Expert II	10 values: $\alpha_{200}, \alpha_{300}, \ldots, \alpha_{1100}$	5	p_1	5.79	6.23
	10 values: $\alpha_{200}, \alpha_{300}, \ldots, \alpha_{1100}$	4	p_2	5.60	6.52
	10 values: $\alpha_{200}, \alpha_{300}, \ldots, \alpha_{1100}$	6	p_3	2.60	3.18
	10 values: $\alpha_{200}, \alpha_{300}, \ldots, \alpha_{1100}$	5	p_4	1.38	1.36
Inverse PCA	9 values: $\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_9$	6	p_1	3.86	5.10
	9 values: $\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_9$	4	p_2	10.50	16.73
	9 values: $\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_9$	8	p_3	1.25	1.89
	9 values: $\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_9$	8	p_4	0.28	0.33

Table A.11: Architecture of particular ANNs in inverse strategies and their errors on training and testing data.

Polynomial chaos in evaluating failure probability: A comparative study

Eliška Janouchová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Jan Sýkora

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Anna Kučerová

Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Abstract

Recent developments in the field of stochastic mechanics and particularly regarding the stochastic finite element method allow to model uncertain behaviours for more complex engineering structures. In reliability analysis, polynomial chaos expansion is a useful tool because it helps to avoid thousands of time-consuming finite element model simulations for structures with uncertain parameters. The aim of this paper is to review and compare available techniques for both the construction of polynomial chaos and its use in computing failure probability. In particular, we compare results for the stochastic Galerkin method, stochastic collocation, and regression method based on Latin hypercube sampling with predictions obtained by crude Monte Carlo sampling. As an illustrative engineering example, we consider a simple frame structure with uncertain parameters in loading and geometry with prescribed distributions defined by realistic histograms.

Keywords: Uncertainty quantification, Reliability analysis;, Probability of failure, Safety margin, Polynomial chaos expansion, Regression method, Stochastic collocation method, Stochastic Galerkin method, Monte Carlo method

1. Introduction

Reliability analysis and modelling of structures in general need to take into account all relevant information as well as any uncertainties in environmental conditions, loading, or structural properties. Input uncertainties influence the behaviour of an investigated system, which thus also becomes uncertain. Description of this phenomenon is provided with an uncertainty quantification process. Extensive development of efficient methods for stochastic modelling enables uncertainty quantification, even for complex models.

Methods quantifying uncertainties can be classified into two groups: (i) reliability analysis methods such as the first- and second-order reliability method (FORM/SORM [8]) for computing the probability of failure related to limit states; (ii) higher moment analysis focused on estimation of higher-order statistical moments of structural response such as stochastic finite element methods (SFEM), see [24, 30, 14] for a review. SFEM is a powerful tool in computational stochastic mechanics extending the classical deterministic finite element method (FEM) to the stochastic framework involving finite elements whose properties are random [13].

In this contribution, we concentrate on SFEM based on polynomial chaos expansion (PCE) [31]. PCE is used to accelerate reliability analysis by replacing time-consuming FEM simulations within the Monte Carlo (MC) sampling of failure probability [20, 21, 27, 17]. To this end, PCE can be employed in two ways: (i) as an approximation of the model response – typically displacements (subsequently referred to **Variant A**) or (ii) as an approximation of the resulting safety margin (**Variant B**). While statistical moments for any approximated quantity can then be computed analytically from the PCE coefficients, failure probability still needs to be estimated from MC simulations. The acceleration of the latter case comes from replacement of an FEM simulation with rapid evaluation of the constructed PCE. The efficiency of SFEM thus depends on the computational requirements of PCE construction and its consequent accuracy.

There are several methods for constructing PCE-based surrogates: the regression method [4, 5, 7], stochastic collocation methods [3, 34], and the stochastic Galerkin method [13, 2, 23, 9]. The principal differences among these methods are outlined. The regression method constructs the polynomial approximation of a response by using the least squares method. It is a stochastic method based on a set of model simulations performed for a stochastic design of experiments, usually obtained using Latin Hypercube Sampling. PCE coefficients are then obtained by regression of the model outputs at the design points. This leads to a solution of a system of equations. In contrast, the stochastic collocation method is a deterministic method involving a set of model simulations on a sparse grid constructed for a chosen level of accuracy. The computation of PCE coefficients is based on an explicit formula. The stochastic Galerkin method leads to a solution of a large system of deterministic equations and requires an intrusive modification of the numerical model itself [11, 29]. These methods were compared within the uncertainty quantification of stiff systems in [6]. The aim of this paper is to extend the previous work presented in [19] devoted to a comparison of these methods in the prediction of failure probability in reliability analysis. In particular, we compare the three methods in terms of computational requirements and resulting accuracy for failure probability of a simple frame structure, where uncertain parameters occur in the geometry of a structure and its loading [12]. The underlying random variables are described by discrete histograms to illustrate a common situation in engineering practice.

Methods for construction of PCE-based surrogates are briefly recalled in the following section. The selected example of a frame structure with uncertain parameters is described in Section 3, followed by numerical study in Section 4. The obtained results are then summarized in Section 5.

2. Polynomial chaos expansion

In order to accelerate the sampling procedure for an uncertainty quantification process, the evaluations of a numerical model

$$r = g(m), \tag{1}$$

where $r = (r_1, \ldots, r_{n_r})^{\mathrm{T}}$ is a vector of model responses and $m = (m_1, \ldots, m_{n_m})^{\mathrm{T}}$ is a vector of random model input parameters, can be replaced with evaluations of a model surrogate. In the stochastic model problem, we assume the model parameters m to be random variables defined over some probability space $(\Omega, \mathfrak{A}, \mathbb{P})$, where Ω is the basic probability set of elementary events, $\mathfrak{A}$ a σ-algebra of subsets of Ω, and $\mathbb{P}$ a probability measure.

In particular, we search for an approximation of response r using polynomial chaos expansion (PCE) [24, 30]. PCE can be used to approximate response with respect to the probability distribution of random variables. For example, Hermite polynomials are associated with Gaussian distribution, Legendre polynomials with uniform distribution, and so on.

In the case of model variables m having another distribution, new standard random variables ξ with the appropriate distribution defined by joint probability density function w_ξ must be introduced. Once we have expressed model parameters m as functions of standard variables $\xi = (\xi_1, \ldots, \xi_{n_\xi})^{\mathrm{T}}$, the model response also becomes a function of these variables. In this paper, we assume particular components of m as well as ξ as independent random variables. When we use PCE-based approximation such that each model input is expressed as a polynomial with one standard variable (i.e. m_j is a univariate function of ξ_j), then the number of newly introduced standard variables n_ξ equals n_m. Let the random model output r be approximated by a PCE $\tilde{r}$ whose polynomials are orthogonal with respect to the probability density function of the distribution of ξ. We write

$$\tilde{r}(\xi) = \sum_\alpha \beta_\alpha \psi_\alpha(\xi), \tag{2}$$

where $\alpha = (\alpha_1, \ldots, \alpha_{n_\xi})$ is a vector of n_ξ non-negative integer components that indicates degrees of multivariate ponynomial $\psi_\alpha(\xi_1, \ldots, \xi_{n_\xi}) = \psi_{\alpha_1}(\xi_1) \cdot \ldots \cdot \psi_{\alpha_{n_\xi}}(\xi_{n_\xi})$ with $\psi_{\alpha_j}(\xi_j)$ being univariate polynomials with a degree α_j. The vector β_α is a vector of PCE coefficients $\beta_{\alpha,i}$ corresponding to a particular component of system response r_i.

Expansion (2) is usually truncated to a limited number of terms, often related to n_ξ and n_p, the number of random variables and the maximal degree of polynomials, respectively [34]. Denoting $|\alpha| = \sum_{j=1}^{n_\xi} \alpha_j$ and considering $|\alpha| \leq n_{\mathrm{p}}$, the number of all terms is n_β given as follows:

$$n_\beta = \frac{(n_{\mathrm{p}} + n_\xi)!}{n_{\mathrm{p}}! n_\xi!}. \tag{3}$$

Polynomial chaos-based surrogate modelling enables computation of statistical moments for an approximated model response $\tilde{r}_i$ analytically from the PCE coefficients [35]. In particular, the mean value can be computed as

$$\mu_{\tilde{r}_i} = \mathbb{E}[\tilde{r}_i] = \underbrace{\int \ldots \int}_{n_\xi} \sum_{|\alpha| \leq n_{\mathrm{p}}} \beta_{\alpha,i} \psi_\alpha(x) w_\xi(x) \mathrm{d}x_1 \ldots \mathrm{d}x_{n_\xi} = \beta_{0,i} \tag{4}$$

and the standard deviation as

$$\sigma_{\tilde{r}_i} = \sqrt{\mathbb{E}[(\tilde{r}_i - \mu_{\tilde{r}_i})^2]} = \sqrt{\sum_{0 < |\boldsymbol{\alpha}| \leq n_p} \mathbb{E}[\psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\xi})] \beta_{\boldsymbol{\alpha},i}^2}, \tag{5}$$

where

$$\mathbb{E}[\psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\xi})] = \underbrace{\int \ldots \int}_{n_\xi} \prod_{j=1}^{n_\xi} (\psi_{\alpha_j}^2(x_j)) w_{\boldsymbol{\xi}}(\boldsymbol{x}) \mathrm{d}x_1 \ldots \mathrm{d}x_{n_\xi}. \tag{6}$$

Specifically, the expected value of the product of Hermite polynomials, which are employed in this paper, is

$$\mathbb{E}[\psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\xi})] = \prod_{j=1}^{n_\xi} \alpha_j!, \tag{7}$$

where α_j is a polynomial degree of variable ξ_j in a polynomial $\psi_{\boldsymbol{\alpha}}$.

The efficiency of this method thus depends mainly on the computational demands of the PCE construction and its accuracy, likewise connected with the method chosen for the construction of the surrogate model [28, 25, 1].

2.1. Regression method

A very general method of computing PC coefficients in Eq. (2) is a well-known regression method [4]. The underlying assumption of this method is that surrogate $\tilde{r}$ is a linear combination of multivariate polynomials $\psi_{\boldsymbol{\alpha}}$, but does not have to be linear in the independent variables $\boldsymbol{\xi}$. The application is based on the following three steps:

   i Preparation of data $\mathbf{X} \in \mathbb{R}^{n_\xi \times n_d}$ obtained as n_d samples $\{\boldsymbol{x}_k, k = 1, \ldots, n_d\}$ of parameter vector $\boldsymbol{\xi}$

   ii Evaluation of the model for samples $\{\boldsymbol{x}_k, k = 1, \ldots, n_d\}$ resulting in response samples $\{\boldsymbol{r}_k, k = 1, \ldots, n_d\}$ organised into matrix $\mathbf{R} \in \mathbb{R}^{n_r \times n_d}$, where n_r is the number of response components, and

   iii Computation of PCE coefficients $\boldsymbol{\beta}_{\boldsymbol{\alpha}}$ organised into matrix $\mathbf{B} \in \mathbb{R}^{n_r \times n_\beta}$ using e.g. the ordinary least square method.

Since the most time-consuming part of this method consists in evaluating the model for samples of random variables, the choice of these samples represents a crucial task with the highest impact on computational time requirements. The simplest way is to choose samples using the Monte Carlo method, i.e. to draw them randomly from a prescribed probability distribution. However, the accuracy of the resulting surrogate depends on the quality with which the samples cover the defined domain [16]. The same quality can be achieved with a smaller number of samples when drawn according to a stratified procedure called design of experiments (DoE). Latin hypercube sampling (LHS) is a well-known DoE able to respect the prescribed probability distributions. There also exist more enhanced ways of optimising LHS (see e.g. [18]), but these are out of scope for this paper. Here, the simplest version of unoptimised LHS is employed. Each computation of a response sample $\boldsymbol{r}_k$ then includes an evaluation of the transformations between model variables $\boldsymbol{m}$ and standard variables $\boldsymbol{\xi}$ and evaluation of model (1).

The computation of PCE coefficients $\mathbf{B}$ starts with an evaluation of all polynomial terms $\boldsymbol{\psi}$ for all samples $\{\boldsymbol{x}_k, k = 1, \ldots, n_d\}$ and saving them in matrix $\mathbf{Z} \in \mathbb{R}^{n_d \times n_\beta}$. The ordinary least square method then leads to

$$\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\mathbf{B}^{\mathrm{T}} = \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}, \tag{8}$$

which is n_r linear systems of n_β equations.

2.2. Stochastic collocation

The stochastic collocation method [33, 26, 22, 10] is based on an explicit expression of the PCE coefficients of orthogonal polynomials with respect to the probability distribution:

$$\beta_{\boldsymbol{\alpha},i} = \frac{1}{\mathbb{E}[\psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\xi})]} \underbrace{\int \ldots \int}_{n_\xi} r_i(\boldsymbol{x})\psi_{\boldsymbol{\alpha}}(\boldsymbol{x})\, w_{\boldsymbol{\xi}}(\boldsymbol{x}) \mathrm{d}x_1 \ldots \mathrm{d}x_{n_\xi}, \tag{9}$$

which can be calculated numerically using an appropriate integration rule (quadrature) on $\mathbb{R}^{n_\xi}$. Eq. (9) then becomes

$$\beta_{\boldsymbol{\alpha},i} \approx \frac{1}{\mathbb{E}[\psi_{\boldsymbol{\alpha}}^2(\boldsymbol{\xi})]} \sum_{l=1}^{n_d} r_i(\boldsymbol{x}_l)\psi_{\boldsymbol{\alpha}}(\boldsymbol{x}_l)v_l, \tag{10}$$

where $\boldsymbol{x}_l$ stands for an integration node and v_l is a corresponding weight. Here we employ versions of the Smolyak quadrature rule, in particular quadratures with Gaussian rules (GQN) and nested Kronrod-Patterson quadrature rules (KPN) derived for normal distribution, see [15].

It is clear that the stochastic collocation method is similar to the regression method, because in both cases the evaluation of a set of model simulations requires the most computational effort. The principal difference can be seen in sample generation, where the stochastic collocation method uses preoptimised sparse grids while the regression method is based on stochastic LHS.

2.3. Stochastic Galerkin

The stochastic Galerkin method is principally different than the previous ones based on a set of independent model simulations. This method spreads the classical finite element method into stochastic space given by equation

$$\mathbf{K}(\boldsymbol{m})\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{m}), \tag{11}$$

where $\mathbf{K}$ is a $n_r \times n_r$ stiffness matrix, $\boldsymbol{f}$ is a $n_r \times 1$ loading vector, and $\boldsymbol{r}$ is an $n_r \times 1$ unknown displacement vector. The stochastic Galerkin method is an intrusive method, i.e. it requires reformulation of this governing equation. To this end, we rewrite Eq. (2) using a matrix notation

$$\tilde{\boldsymbol{r}}(\boldsymbol{\xi}) = (\mathbf{I} \otimes \boldsymbol{\psi}^{\mathrm{T}}(\boldsymbol{\xi}))\boldsymbol{\beta}, \tag{12}$$

where $\mathbf{I} \in \mathbb{R}^{n_r \times n_r}$ is the identity matrix, $\otimes$ is the Kronecker product, $\boldsymbol{\psi}(\boldsymbol{\xi})$ is a n_β-dimensional column vector of polynomials, and $\boldsymbol{\beta}$ is a $(n_\beta \cdot n_r)$-dimensional column vector of PCE coefficients organised here as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\beta}_i^{\mathrm{T}}, \ldots, \boldsymbol{\beta}_{n_r}^{\mathrm{T}})^{\mathrm{T}}$, where $\boldsymbol{\beta}_i = (\ldots, \beta_{\boldsymbol{\alpha},i}, \ldots)^{\mathrm{T}}$ consists of PCE coefficients corresponding to the i-th response component.

Substituting model response $\boldsymbol{r}$ in Eq. (11) with its PCE approximation $\tilde{\boldsymbol{r}}$ given in Eq. (12) and applying the Galerkin projection, we obtain

$$\underbrace{\int \ldots \int}_{n_\xi} \boldsymbol{\psi}(\boldsymbol{x}) \otimes \mathbf{K}(\boldsymbol{x}) \otimes \boldsymbol{\psi}^{\mathrm{T}}(\boldsymbol{x})\, w_{\boldsymbol{\xi}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}\,\boldsymbol{\beta} = \underbrace{\int \ldots \int}_{n_\xi} \boldsymbol{\psi}(\boldsymbol{x}) \otimes \boldsymbol{f}(\boldsymbol{x})\, w_{\boldsymbol{\xi}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}\,, \quad (13)$$

which is a linear system of $(n_\beta \cdot n_\mathrm{r})$ equations with unknowns in the vector $\boldsymbol{\beta}$. Integration can be performed numerically or analytically. The analytical solution is available e.g. when all terms in the stiffness matrix and in the loading vector are polynomials with respect to $\boldsymbol{\xi}$. In such cases, the method is called *fully intrusive*. In other cases, a numerical integration leading to a *semi-intrusive* Galerkin method is inevitable and can be again solved with the help of the Smolyak integration rule, namely GQN [15].

3. Description of a frame structure with uncertain parameters

The goal of the work presented here is to compare the methods described for approximating model response and accelerating Monte Carlo (MC) sampling performed in order to estimate the probability distribution of the safety margin and the probability of structural failure.

In order to demonstrate the performance of the methods described on an engineering structure, we have chosen a simple frame with two beams (cross-section HEB 100) and a column (cross-section HEB 120) presented in [12]. To keep the comparison study clear, the geometry, loading conditions, input material parameters, random variables, and their corresponding notations are preserved as in [12]. As an illustration, the initial geometry and loading conditions are displayed in Figure 1.
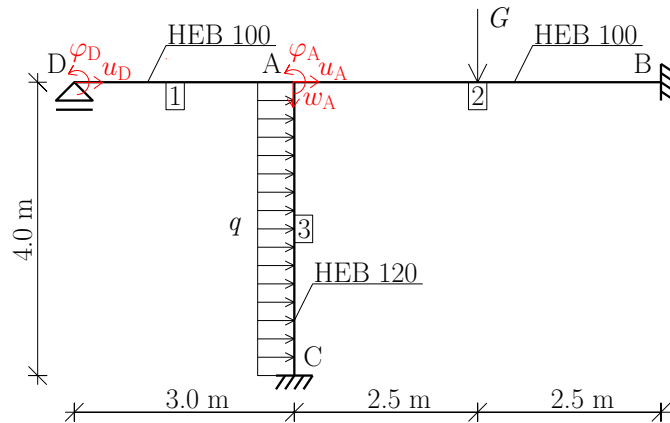
Figure 1: Scheme of the frame structure.

The frame is made of steel with Young's modulus $E = 210\,[\mathrm{GPa}]$ and uncertain yield stress f_y obtained by the product of the nominal value $f_{y,\mu}$ and dimensionless variation $f_{y,\sigma}$ defined by a prescribed histogram (see Figure 2). The geometrical parameters of particular beams are considered to be uncertain and defined as products of the corresponding nominal

Variable	Nominal value	Dimensionless variation
Yield stress	$f_{y,\mu} = 235\,[\text{GPa}]$	$f_{y,\sigma}\,[-]$
Moment of inertia	$I_1 = 449.5\,[\text{cm}^4]$	$I_{\sigma 1}\,[-]$
Cross-sectional area	$A_1 = 26.04\,[\text{cm}^2]$	$I_{\sigma 1}\,[-]$
Moment of inertia	$I_2 = 449.5\,[\text{cm}^4]$	$I_{\sigma 2}\,[-]$
Cross-sectional area	$A_2 = 26.04\,[\text{cm}^2]$	$I_{\sigma 2}\,[-]$
Moment of inertia	$I_3 = 864.4\,[\text{cm}^4]$	$I_{\sigma 3}\,[-]$
Cross-sectional area	$A_3 = 34.01\,[\text{cm}^2]$	$I_{\sigma 3}\,[-]$
Elastic section modulus	$W_3 = 144.1\,[\text{cm}^3]$	$I_{\sigma 3}\,[-]$
Length	$l_1 = 3.0\,[\text{m}]$	$l_\sigma\,[-]$
Length	$l_2 = 5.0\,[\text{m}]$	$l_\sigma\,[-]$
Length	$l_3 = 4.0\,[\text{m}]$	$l_\sigma\,[-]$
Dead load	$D_1 = 11\,[\text{kNm}^{-1}]$	$D_{\sigma 1}\,[-]$
Short-lasting load	$S_1 = 9\,[\text{kNm}^{-1}]$	$S_{\sigma 1}\,[-]$
Long-lasting load	$L_1 = 5.5\,[\text{kNm}^{-1}]$	$L_{\sigma 1}\,[-]$
Dead load	$D_2 = 3.5\,[\text{kN}]$	$D_{\sigma 2}\,[-]$
Short-lasting load	$S_2 = 2.2\,[\text{kN}]$	$S_{\sigma 2}\,[-]$
Long-lasting load	$L_2 = 1.7\,[\text{kN}]$	$L_{\sigma 2}\,[-]$

Table 1: Material, geometrical, and loading data with corresponding variations.

values and dimensionless variations given in [12] and listed in Table 1. Particular histograms are also depicted in Figure 2.

The prescribed loading conditions are linear combinations of dead, long-lasting, and short-lasting load given as:

$$q = D_1 D_{\sigma 1} + S_1 S_{\sigma 1} + L_1 L_{\sigma 1}\ [\text{kNm}^{-1}], \tag{14}$$

and

$$G = D_2 D_{\sigma 2} + S_2 S_{\sigma 2} + L_2 L_{\sigma 2}\ [\text{kN}], \tag{15}$$

where particular loads are statistically independent and described by random variables with nominal values $(D_1, S_1, L_1, D_2, S_2, L_2)$ and variations $(D_{\sigma 1}, S_{\sigma 1}, L_{\sigma 1}, D_{\sigma 2},\ S_{\sigma 2}, L_{\sigma 2})$ defined in Table 1 and by the histograms depicted in Figure 2.

Let us simplify the notation and denote all random variables as m_i,

$$\boldsymbol{m} = (m_1, \ldots, m_{n_m})^{\mathrm{T}} = (I_{\sigma 1}, I_{\sigma 2}, I_{\sigma 3}, l_\sigma, D_{\sigma 1}, S_{\sigma 1}, L_{\sigma 1}, D_{\sigma 2}, S_{\sigma 2}, L_{\sigma 2})^{\mathrm{T}}. \tag{16}$$

None of these variables has a continuous probability density function (PDF), which is necessary for constructing PCE-based approximation, but their distribution is described by discrete histograms. For this reason, we introduce new standard random variables $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{n_\xi})^{\mathrm{T}}$, $n_\xi = n_m$, with a continuous PDF. The variables m_j can be then expressed by transformation functions t_j with variables ξ_j, i.e.

$$m_j = t_j(\xi_j). \tag{17}$$
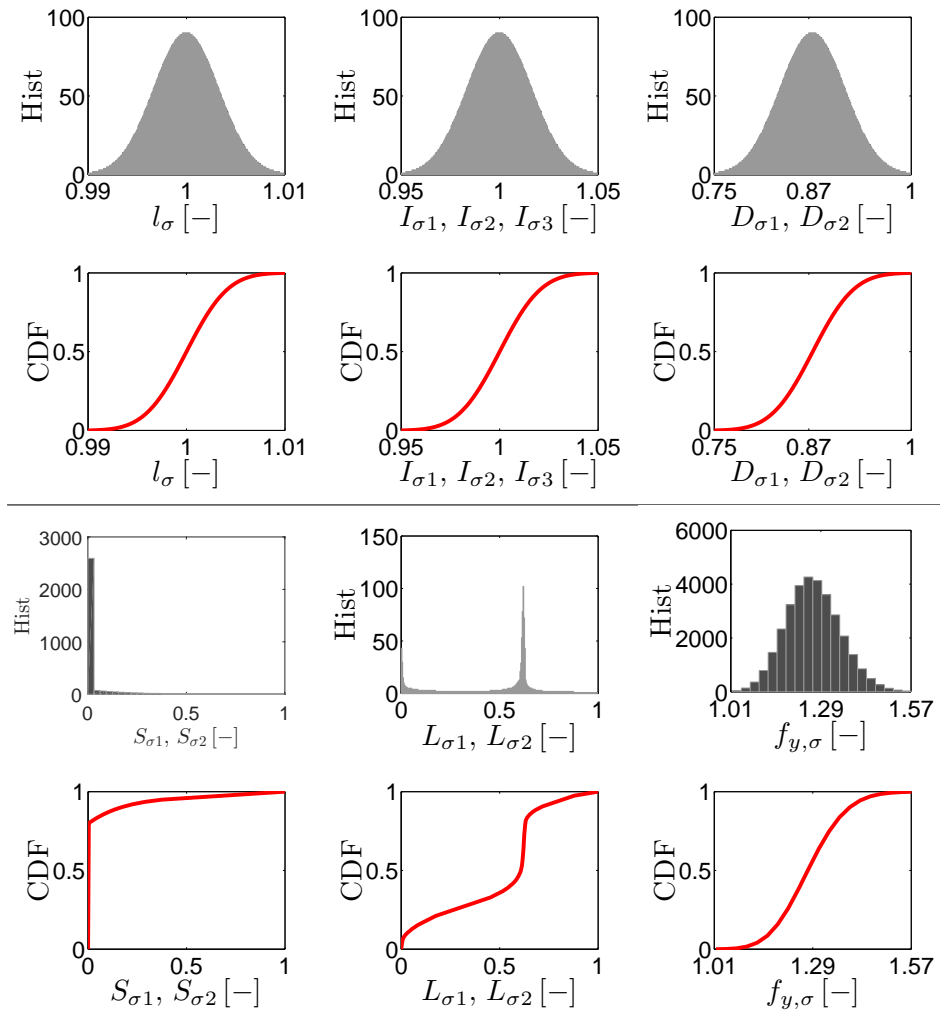
Figure 2: Histograms of uncertain parameters and corresponding cumulative density functions.

For discrete histograms, the transformation functions are non-smooth. Particular examples of transformation functions will be discussed in Section 4.

Since the axial force in the column does not achieve critical intensity, instability does not play any role and thus the column has only one failure mode determined by material yield stress. The maximal internal forces will appear in the column at support 'C' and can be computed from the displacement and rotation of joint 'A'. The unknown displacements $r = (u_D, \varphi_D, u_A, w_A, \varphi_A)$ can be – for the geometrical and material linearity considered here – computed using the finite element method or displacement method, both of which are very well-known. Hence, we start directly with the latter method with a discretised form of the equilibrium equations given in Eq. (11), which – after applying the boundary conditions – comprises a system of five linear equations for unknown vector $r = (u_D, \varphi_D, u_A, w_A, \varphi_A)$.

The stiffness matrix is given as

$$
\mathbf{K} = \frac{12E}{l_\sigma} \cdot \begin{bmatrix}
\frac{A_1 I_{\sigma 1}}{12 l_1} & 0 & -\frac{A_1 I_{\sigma 1}}{12 l_1} & 0 & 0 \\
0 & \frac{I_1 I_{\sigma 1}}{3 l_1} & 0 & \frac{I_1 I_{\sigma 1}}{2 l_1^2 l_\sigma} & \frac{I_1 I_{\sigma 1}}{6 l_1} \\
-\frac{A_1 I_{\sigma 1}}{12 l_1} & 0 & k_{33} & 0 & \frac{I_3 I_{\sigma 3}}{2 l_3^2 l_\sigma} \\
0 & \frac{I_1 I_{\sigma 1}}{2 l_1^2 l_\sigma} & 0 & k_{44} & \frac{I_1 I_{\sigma 1}}{2 l_1^2 l_\sigma} - \frac{I_2 I_{\sigma 2}}{2 l_2^2 l_\sigma} \\
0 & \frac{I_1 I_{\sigma 1}}{6 l_1} & \frac{I_3 I_{\sigma 3}}{2 l_3^2 l_\sigma} & \frac{I_1 I_{\sigma 1}}{2 l_1^2 l_\sigma} - \frac{I_2 I_{\sigma 2}}{2 l_2^2 l_\sigma} & k_{55}
\end{bmatrix},
\tag{18}
$$

where $k_{33} = \frac{A_1 I_{\sigma 1}}{12 l_1} + \frac{A_2 I_{\sigma 2}}{12 l_2} + \frac{I_3 I_{\sigma 3}}{l_3^3 l_\sigma^2}$, $k_{44} = \frac{I_1 I_{\sigma 1}}{l_1^3 l_\sigma^2} + \frac{I_2 I_{\sigma 2}}{l_2^3 l_\sigma^2} + \frac{A_3 I_{\sigma 3}}{12 l_3}$, $k_{55} = \frac{I_1 I_{\sigma 1}}{3 l_1} + \frac{I_2 I_{\sigma 2}}{3 l_2} + + \frac{I_3 I_{\sigma 3}}{3 l_3}$
and the loading vector as

$$
\boldsymbol{f} = \begin{pmatrix}
0 \\
0 \\
\frac{(D_1 D_{\sigma 1} + S_1 S_{\sigma 1} + L_1 L_{\sigma 1})) l_3 l_\sigma}{2} \\
\frac{D_2 D_{\sigma 2} + S_2 S_{\sigma 2} + L_2 L_{\sigma 2}}{2} \\
-\frac{(D_2 D_{\sigma 2} + S_2 S_{\sigma 2} + L_2 L_{\sigma 2}) l_2 l_\sigma}{8} + \frac{(D_1 D_{\sigma 1} + S_1 S_{\sigma 1} + L_1 L_{\sigma 1})(l_3 l_\sigma)^3}{12}
\end{pmatrix}.
\tag{19}
$$

Safety margin M of the column is the difference between yield stress f_y and stress σ produced by external load and given as

$$
\sigma = -w_\mathrm{A} E + \left(\frac{q l_3^2 l_\sigma^2}{12 I_{\sigma 3}} + \frac{2 E I_3}{l_3 l_\sigma} \varphi_\mathrm{A} + \frac{6 E I_3}{l_3 l_\sigma} u_\mathrm{A} \right) / W_3.
\tag{20}
$$

Failure F occurs when σ exceeds f_y. The probability of failure $\Pr(F)$ is then estimated to be the number of failures divided by the total number of performed simulations n:

$$
\Pr(F) = \frac{1}{n} \sum_{s=1}^{n} I[f_y - \sigma_s \le 0],
\tag{21}
$$

where $I[f_y - \sigma \le 0]$ is an indicator function with a value one if $f_y - \sigma \le 0$, and zero otherwise.

In this numerical study, two variants of a model response are considered. For the first variant, the safety margin is calculated from the vector $\tilde{\boldsymbol{r}}$ that approximates the displacements $\boldsymbol{r}$ (**Variant A**). In the second case, the approximated random output is directly safety margin M (**Variant B**). In the first variant, the stochastic Galerkin method can be applied and compared with the other two methods described in Section 2.

The results for constructed surrogate models are compared with reference results obtained by MC sampling with $n = 10^7$ samples. For **Variant A**, the relative errors [%] in an estimation of the mean of particular response components are defined as

$$
\varepsilon_{\mu_i} = \frac{\left| \mu_{i,\mathrm{PCE}} - \mu_{i,\mathrm{MC}} \right|}{\mu_{i,\mathrm{MC}}} \cdot 100, \, i = 1, \dots, n_r,
\tag{22}
$$

where $\mu_{i,\mathrm{MC}}$ is the MC-based estimation of mean and $\mu_{i,\mathrm{PCE}}$ is the mean computed for a particular surrogate model. Analogical expression is also used to evaluate errors in estimating standard deviation.

In both variants of model response, the relative errors [%] in predicting the safety margin are obtained with

$$\varepsilon_M = \frac{1}{n} \sum_{s=1}^{n} \frac{|M_{\mathrm{PCE},s} - M_{\mathrm{MC},s}|}{\max M_{\mathrm{MC}} - \min M_{\mathrm{MC}}} \cdot 100, \tag{23}$$

where M_{MC} stands for the samples of the safety margin estimated using the MC method. $M_{\mathrm{MC},s}$ resp. $M_{\mathrm{PCE},s}$ stands for a particular sample of the safety margin obtained by the MC method with the full numerical model or a chosen surrogate model respectively.

4. Numerical study

We assume that the components of $\boldsymbol{\xi}$ are standard Gaussian variables and thus we employ Hermite polynomials for the model surrogate. The study involves several variants of distributions for model parameters $\boldsymbol{m}$.

At first, each model parameter m_j is distributed according to the prescribed histogram and the corresponding transformation to the standard variable ξ_j consists of two steps. The variable ξ_j is transformed using the cumulative distribution function of the standard normal distribution

$$\Phi(\xi_j) = \int_{-\infty}^{\xi_j} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \mathrm{d}t \tag{24}$$

to the uniformly distributed variables on the interval $\langle 0; 1 \rangle$. Then, the final transformation step is based on piecewise linear inverse cumulative distribution functions arising from the prescribed histogram.

For transformation functions, the stochastic Galerkin method can be applied in its semi-intrusive form. In our particular example, we multiply the governing Eq. (11) involving expressions in Eq. (18) and (19) by $(l_\sigma)^3$ so as to obtain polynomials in terms of model parameters $\boldsymbol{m}$. However, we will not obtain polynomials in terms of $\boldsymbol{\xi}$ due to non-smooth transformations (Eq. (17)) produced by the discrete nature of the histograms prescribed to $\boldsymbol{m}$.

Table 2 contains results of uncertainty quantification for response components in terms of their means and standard deviations. These results are obtained for **Variant A**, where one PCE is used to approximate each response component. The presented methods are compared here in this specific form: the regression method (LHS), the stochastic collocation method in two variants (KPN, GQN), and semi-intrusive Galerkin method based on GQN quadrature rules (GM GQN) for four polynomial degrees n_{p}. The results show good predictions obtained using the regression method, while stochastic collocation based on KPN rules leads to significant errors in estimating standard deviations, with the method based on GQN rules appearing to be even divergent. The semi-intrusive Galerkin method achieves better results in estimating standard deviations than stochastic collocation based on the same quadrature rules. Estimated PDFs for displacement u_{A} depicted in Figure 3 are not sufficient, even for the regression method.

Moreover, the PCE-based approximations are compared to the polynomial approximation (PA) without orthogonality of the polynomial basis with respect to the input distribution. In this case, the polynomials approximate directly the relations between model parameters $\boldsymbol{m}$ and model responses $\boldsymbol{r}$. The PA coefficients are computed by the regression method based on LHS design respecting the prescribed histograms of the model parameters. The benefit

Method	n_{p}	n_{d}	Time [s]	$\mu_{u_{\mathrm{A}}}$ [mm]	$\sigma_{u_{\mathrm{A}}}$ [mm]	$\mu_{w_{\mathrm{A}}}$ [mm]	$\sigma_{w_{\mathrm{A}}}$ [mm]	$\mu_{\varphi_{\mathrm{A}}}$ [mrad]	$\sigma_{\varphi_{\mathrm{A}}}$ [mrad]
MC	$-$	10^7	22191	0.207	0.033	0.009	0.002	4.090	0.795
				$\varepsilon_{\mu_{u_{\mathrm{A}}}}$ [%]	$\varepsilon_{\sigma_{u_{\mathrm{A}}}}$ [%]	$\varepsilon_{\mu_{w_{\mathrm{A}}}}$ [%]	$\varepsilon_{\sigma_{w_{\mathrm{A}}}}$ [%]	$\varepsilon_{\mu_{\varphi_{\mathrm{A}}}}$ [%]	$\varepsilon_{\sigma_{\varphi_{\mathrm{A}}}}$ [%]
LHS	1	21	0	0.18	11.96	0.32	60.63	0.20	14.20
	2	201	0	0.26	4.76	0.36	2.49	0.34	4.41
	3	1201	3	0.08	1.36	0.01	0.02	0.10	1.54
	4	5301	19	0.02	1.31	0.09	0.98	0.04	1.21
KPN	1	21	0	4.81	9.34	4.26	8.42	4.90	9.38
	2	201	0	4.81	5.54	4.26	4.32	4.90	5.53
	3	1201	3	2.26	7.37	1.98	3.73	2.31	5.18
	4	5301	13	0.30	11.36	0.29	6.20	0.30	7.96
GQN	1	21	0	6.68	22.99	6.06	15.52	6.78	23.00
	2	221	0	4.81	73.94	4.25	50.93	4.90	58.21
	3	1581	4	3.11	59.73	2.81	37.07	3.16	46.92
	4	8761	21	1.13	187.92	1.10	129.41	1.14	147.83
GM GQN	1	21	1	0.37	17.50	1.28	16.47	0.29	19.82
	2	221	0	0.37	1.81	1.28	6.67	0.29	7.90
	3	1581	3	0.37	1.24	1.28	4.97	0.29	4.13
	4	8761	41	0.37	6.51	1.28	5.44	0.29	1.29
PA	1	21	302	0.01	0.41	0.17	1.32	$3 \cdot 10^{-3}$	0.36
	2	201	377	$3 \cdot 10^{-4}$	0.01	$1 \cdot 10^{-4}$	$5 \cdot 10^{-3}$	$2 \cdot 10^{-4}$	$2 \cdot 10^{-3}$
	3	1201	540	$5 \cdot 10^{-6}$	$2 \cdot 10^{-4}$	$2 \cdot 10^{-5}$	$2 \cdot 10^{-4}$	$1 \cdot 10^{-6}$	$3 \cdot 10^{-5}$
	4	5301	1117	$4 \cdot 10^{-6}$	$9 \cdot 10^{-5}$	$1 \cdot 10^{-6}$	$7 \cdot 10^{-5}$	$2 \cdot 10^{-6}$	$1 \cdot 10^{-5}$

Table 2: Time requirements and errors in predicting means and standard deviations of displacement vector components for the **prescribed histograms** with model parameters m.

$n_{\mathrm{p}} = 1$ $n_{\mathrm{p}} = 4$
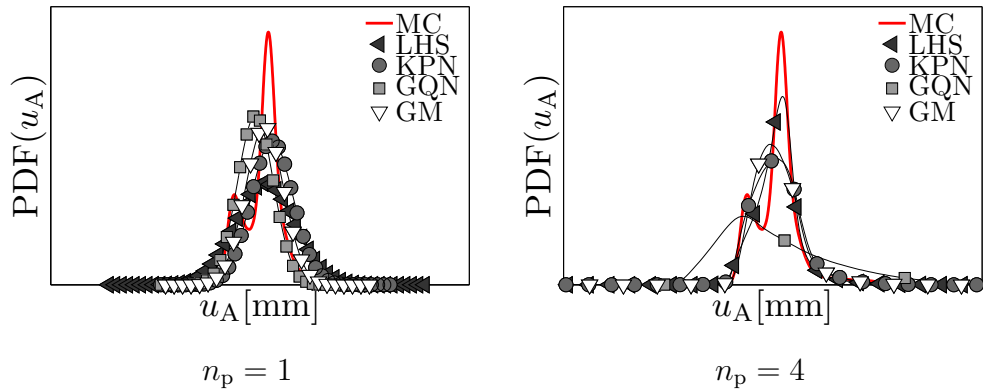
Figure 3: Probability density functions of displacement u_{A} for different n_{p} for the **prescribed histograms** with model parameters m.

of this approach is elimination of the nonlinear transformation functions from the approximated relations which enables to reach a higher accuracy with the same polynomial degree in comparison with PCE, see Table 2. On the other side, there are disadvantages associated

with the computational aspects. No restrictions to the model parameters' distribution can lead to ill-conditioned system of equations in the least square method as it is in this investigated case. On top of that, the polynomial coefficients cannot be used to compute statistical moments of the approximated model response analytically, but for this purpose the Monte Carlo samples of PA have to be employed which increases the computation time.

Predictions of safety margin M and failure probability $\Pr(F)$ are summarised in Table 3 for **Variant A** as well as for **Variant B**, where PCE is used as a direct surrogate for safety margin M. One can see that the estimations of failure probability are unsatisfactory for all methods examined. Predictions of the safety margin seems better in both variants, but **Variant A** significantly outperforms **Variant B**, see Figure 4. The number of random variables involved in the PCEs constructed appears to be a crucial factor here. The total number of random variables for **Variant B** is eleven, including uncertain yield stress f_y, missing from **Variant A** PCEs except during sampling of failure probability. Therefore, even though **Variant B** requires construction of only one PCE, the additional variable causes an enormous increase in complexity because of an increasing polynomial degree, which also quickly increases computational times.

			Variant B				Variant A		
Method	n_{p}	n_{d}	Time [s]	$\Pr(F)\,[-]$	ε_M [%]	n_{d}	Time [s]	$\Pr(F)\,[-]$	ε_M [%]
MC	$-$	10^7	23825	$7\cdot10^{-5}$	$-$	10^7	22191	$7\cdot10^{-5}$	$-$
LHS	1	23	32	$7\cdot10^{-7}$	4.20	21	473	$3\cdot10^{-5}$	1.33
	2	243	164	$16\cdot10^{-5}$	2.66	201	549	$6\cdot10^{-5}$	0.63
	3	1607	757	$157\cdot10^{-5}$	2.12	1201	772	$18\cdot10^{-5}$	0.44
	4	7767	2702	$138\cdot10^{-5}$	1.79	5281	1416	$18\cdot10^{-5}$	0.37
KPN	1	23	31	$1\cdot10^{-6}$	4.31	21	487	$2\cdot10^{-5}$	0.88
	2	243	164	$49\cdot10^{-5}$	3.11	201	547	$9\cdot10^{-5}$	0.64
	3	1607	760	$113\cdot10^{-5}$	2.98	1201	801	$13\cdot10^{-5}$	0.59
	4	7767	2701	$84\cdot10^{-5}$	3.43	5281	1416	$11\cdot10^{-5}$	0.65
GQN	1	23	31	0	3.51	21	475	$4\cdot10^{-6}$	0.71
	2	265	164	$5\cdot10^{-5}$	9.75	221	549	$4\cdot10^{-5}$	1.78
	3	2069	764	$719\cdot10^{-5}$	6.98	1581	759	$62\cdot10^{-5}$	1.29
	4	12453	2713	$3229\cdot10^{-5}$	15.74	8761	1404	$93\cdot10^{-5}$	2.75
GM GQN	1	$-$	$-$	$-$	$-$	21	500	$1\cdot10^{-5}$	0.67
	2	$-$	$-$	$-$	$-$	221	586	$5\cdot10^{-5}$	0.48
	3	$-$	$-$	$-$	$-$	1581	807	$15\cdot10^{-5}$	0.38
	4	$-$	$-$	$-$	$-$	8761	1461	$15\cdot10^{-5}$	0.37
PA	1	23	467	$4\cdot10^{-5}$	0.14	21	328	$7\cdot10^{-5}$	0.03
	2	243	583	$7\cdot10^{-5}$	$2\cdot10^{-3}$	201	382	$8\cdot10^{-5}$	$2\cdot10^{-4}$
	3	1607	916	$7\cdot10^{-5}$	$4\cdot10^{-5}$	1201	713	$8\cdot10^{-5}$	$3\cdot10^{-6}$
	4	7767	5604	$7\cdot10^{-5}$	$6\cdot10^{-5}$	5281	1420	$8\cdot10^{-5}$	$6\cdot10^{-6}$

Table 3: Time requirements, probability of failure and errors in predicting safety margin for the **prescribed histograms** with model parameters $\boldsymbol{m}$.

The overall unsatisfactory results of PCE approximations are a result of a highly nonlinear transformation (Eq. (17)). This hypothesis is supported by the good results of PA.
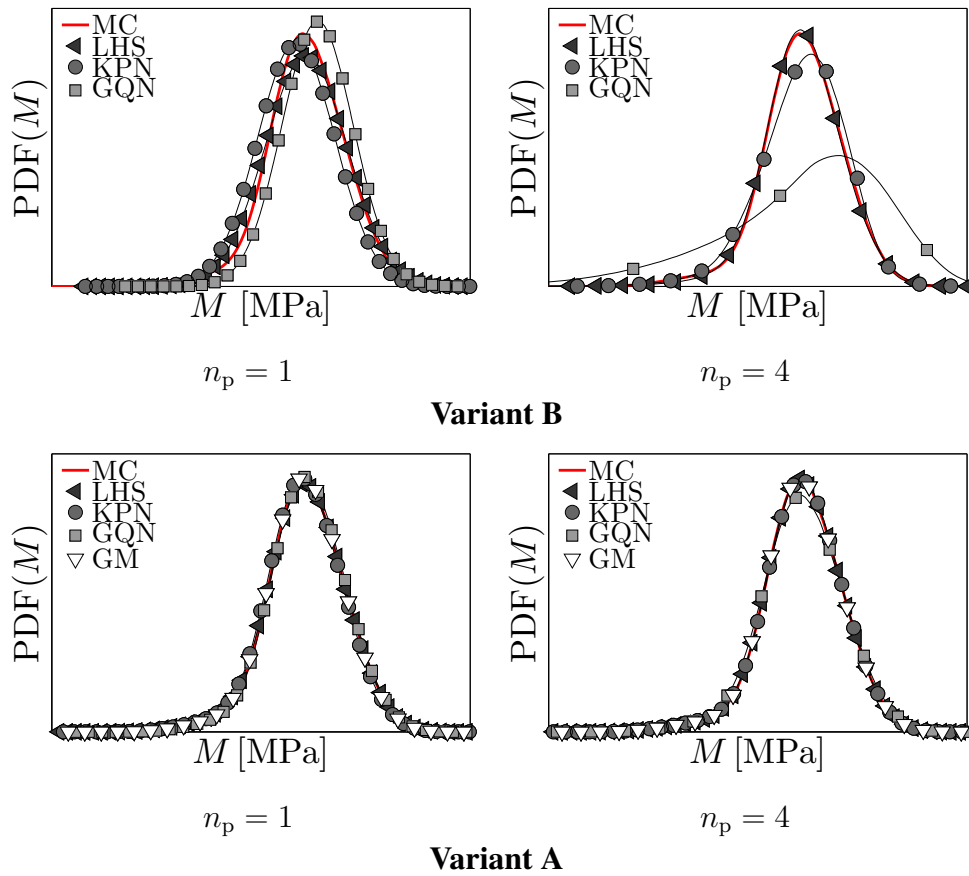
Figure 4: Probability density functions of safety margin M for different n_{p} for the **prescribed histograms** with model parameters $\boldsymbol{m}$.

The most problematic relations are likely transformation of parameters with the prescribed histograms $S_{\sigma 1}$, $S_{\sigma 2}$ and $L_{\sigma 1}$, $L_{\sigma 2}$, respectively, to standard normal variables, as shown in Figure 5. In order to test this assumption, we have replaced these two prescribed histograms with new ones closer to normal distribution, see Figure 6. The new histograms respect the
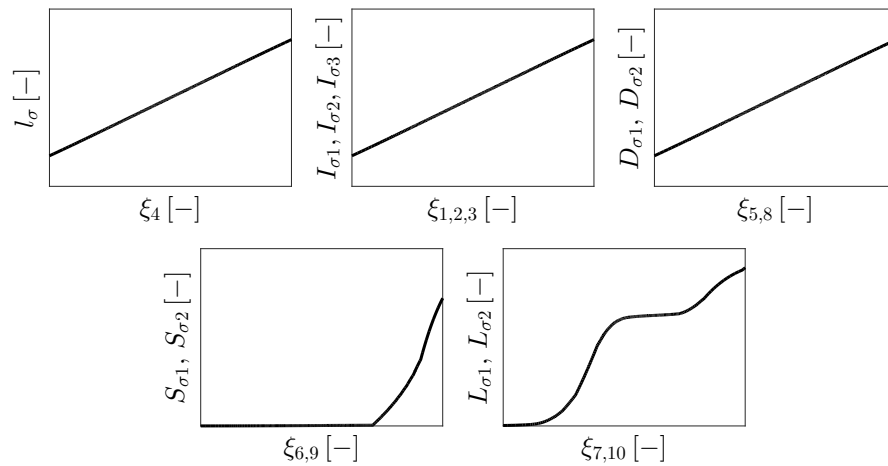
Figure 5: Transformation relations for prescribed histograms.
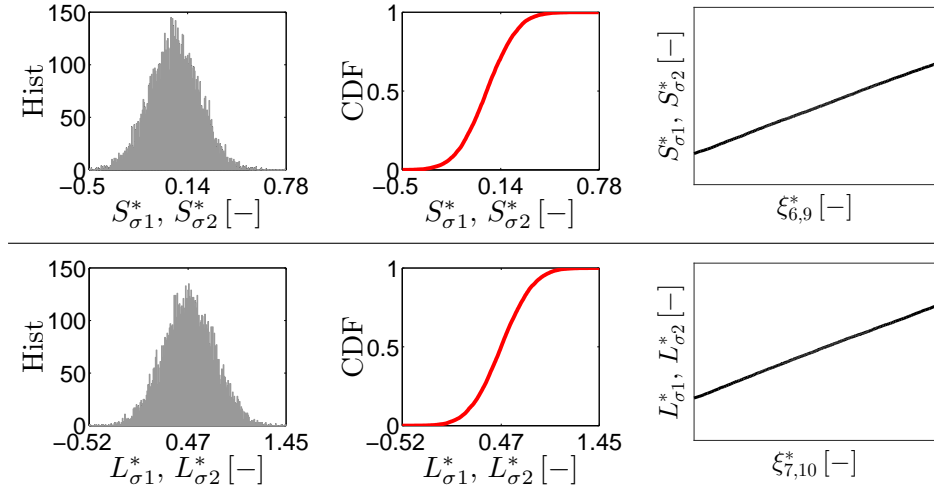
Figure 6: New histograms of model parameters with corresponding cumulative density functions and transformation relations.

initial values of means and standard deviations from the prescribed histograms.

Results obtained for the case of the new histograms and **Variant A** are listed in Table 4. One can see that the replacement of the two histograms led to a significant improvement in results achieved for all methods. GQN based collocation yields the worst results and the method still suffers from problems with convergence. The regression method LHS provides

| Method	n_{p}	n_{d}	Time [s]	$\mu_{u_{\mathrm{A}}}$ [mm]	$\sigma_{u_{\mathrm{A}}}$ [mm]	$\mu_{w_{\mathrm{A}}}$ [mm]	$\sigma_{w_{\mathrm{A}}}$ [mm]	$\mu_{\varphi_{\mathrm{A}}}$ [mrad]	$\sigma_{\varphi_{\mathrm{A}}}$ [mrad]
MC	−	10^7	21874	0.206	0.033	0.009	0.002	4.056	0.789
				$\varepsilon_{\mu_{u_{\mathrm{A}}}}$ [%]	$\varepsilon_{\sigma_{u_{\mathrm{A}}}}$ [%]	$\varepsilon_{\mu_{w_{\mathrm{A}}}}$ [%]	$\varepsilon_{\sigma_{w_{\mathrm{A}}}}$ [%]	$\varepsilon_{\mu_{\varphi_{\mathrm{A}}}}$ [%]	$\varepsilon_{\sigma_{\varphi_{\mathrm{A}}}}$ [%]
LHS	1	21	0	0.01	0.91	0.13	3.06	0.02	0.02
	2	201	0	0.01	0.08	0.02	0.16	0.01	0.09
	3	1201	3	0.00	0.13	0.03	0.34	0.00	0.13
	4	5301	19	0.01	0.00	0.01	0.06	0.01	0.00
KPN	1	21	0	0.05	0.13	0.04	0.06	0.05	0.11
	2	201	0	0.05	0.09	0.04	0.00	0.05	0.09
	3	1201	3	0.01	0.27	0.01	0.19	0.01	0.26
	4	5301	13	0.01	0.13	0.01	0.13	0.01	0.13
GQN	1	21	0	0.08	0.21	0.06	0.20	0.08	0.20
	2	221	0	0.05	0.07	0.04	0.01	0.05	0.07
	3	1581	4	0.03	0.47	0.03	0.30	0.03	0.47
	4	8761	21	0.01	0.20	0.01	0.21	0.01	0.20
GM GQN	1	21	1	0.07	0.31	0.01	0.09	0.01	0.32
	2	221	0	0.07	0.35	0.01	0.15	0.01	0.28
	3	1581	3	0.07	0.35	0.01	0.16	0.01	0.21
	4	8761	41	0.07	0.36	0.01	0.17	0.01	0.07

Table 4: Time requirements and errors in predicting means and standard deviations for displacement vector components for the **new histograms** with model parameters m.

the worst estimation for polynomials of the first degree, but the error here decreases with increasing polynomial degree. The behaviour of the semi-intrusive Galerkin method and stochastic collocation is very similar due to numerical integration based on GQN rules.

The same improvement can be seen also in the prediction of the whole probability density function for safety margin M. The corresponding errors in predicting safety margin and failure probability are listed in Table 5.

We can also notice that GQN-based collocation provides the worst results for the response corresponding to safety margin. The semi-intrusive Galerkin method delivers the worst prediction. The resulting estimation of failure probability is now satisfactory; in **Variant A**, it is excellent.

In both previous examples, the discrete nature of prescribed histograms led to the necessity of numerical integration in the stochastic Galerkin method resulting in its semi-intrusive variant. In order to investigate the performance of the fully intrusive stochastic Galerkin method avoiding all the numerical approximations, we have changed the prescribed distributions for model parameters once more. This time, we assume all the parameters to be normally distributed with the original values of means and standard deviations. In such cases, the transformation (17) becomes a 1st order polynomial and hence, analytical integration is available.

Figure 7 shows the functional dependence of safety margin M for considered types of probability distribution prescribed to model parameters. Figure 7(a) shows that the relation between M and model parameters $\boldsymbol{m}$ is linear, while high nonlinearity appears in the relation to standard variables $\boldsymbol{\xi}$ for the prescribed histograms, see Figure 7(b). Replacement of the

Method	n_p	Variant B				Variant A			
		n_d	Time [s]	$\Pr(F)\,[-]$	$\varepsilon_M\,[\%]$	n_d	Time [s]	$\Pr(F)\,[-]$	$\varepsilon_M\,[\%]$
MC	–	10^7	23833	$5 \cdot 10^{-7}$	–	10^7	21874	$5 \cdot 10^{-7}$	–
LHS	1	23	32	$9 \cdot 10^{-7}$	$2.86 \cdot 10^{-1}$	21	479	$5 \cdot 10^{-7}$	$3.20 \cdot 10^{-2}$
	2	243	165	$6 \cdot 10^{-7}$	$7.98 \cdot 10^{-2}$	201	537	$5 \cdot 10^{-7}$	$1.42 \cdot 10^{-2}$
	3	1607	753	$3 \cdot 10^{-7}$	$8.26 \cdot 10^{-2}$	1201	782	$5 \cdot 10^{-7}$	$1.43 \cdot 10^{-2}$
	4	7767	2701	$15 \cdot 10^{-7}$	$2.00 \cdot 10^{-1}$	5281	1437	$5 \cdot 10^{-7}$	$2.75 \cdot 10^{-2}$
KPN	1	23	31	$10 \cdot 10^{-7}$	$2.59 \cdot 10^{-1}$	21	475	$5 \cdot 10^{-7}$	$2.02 \cdot 10^{-2}$
	2	243	164	$6 \cdot 10^{-7}$	$8.21 \cdot 10^{-2}$	201	551	$5 \cdot 10^{-7}$	$1.32 \cdot 10^{-2}$
	3	1607	758	$1 \cdot 10^{-7}$	$1.66 \cdot 10^{-1}$	1201	778	$5 \cdot 10^{-7}$	$2.41 \cdot 10^{-2}$
	4	7767	2786	$8 \cdot 10^{-7}$	$1.30 \cdot 10^{-1}$	5281	1425	$5 \cdot 10^{-7}$	$2.20 \cdot 10^{-2}$
GQN	1	23	31	$10 \cdot 10^{-7}$	$2.41 \cdot 10^{-1}$	21	484	$5 \cdot 10^{-7}$	$2.12 \cdot 10^{-2}$
	2	265	164	$6 \cdot 10^{-7}$	$1.47 \cdot 10^{-1}$	221	560	$5 \cdot 10^{-7}$	$2.35 \cdot 10^{-2}$
	3	2069	763	$1 \cdot 10^{-7}$	$1.98 \cdot 10^{-1}$	1581	781	$5 \cdot 10^{-7}$	$3.39 \cdot 10^{-2}$
	4	12453	2714	$2 \cdot 10^{-7}$	$2.94 \cdot 10^{-1}$	8761	1418	$5 \cdot 10^{-7}$	$3.48 \cdot 10^{-2}$
GM GQN	1	–	–	–	–	21	491	$5 \cdot 10^{-7}$	$1.99 \cdot 10^{-2}$
	2	–	–	–	–	221	575	$5 \cdot 10^{-7}$	$2.67 \cdot 10^{-2}$
	3	–	–	–	–	1581	782	$5 \cdot 10^{-7}$	$4.03 \cdot 10^{-2}$
	4	–	–	–	–	8761	1446	$6 \cdot 10^{-7}$	$4.45 \cdot 10^{-2}$

Table 5: Time requirements, probability of failure and errors in predicting safety margin for the **new histograms** with model parameters $\boldsymbol{m}$.

two histograms $S_{\sigma 1}$, $S_{\sigma 2}$ and $L_{\sigma 1}$, $L_{\sigma 2}$, respectively, with new ones more similar to normal distributions leads to an almost linear $M - \xi$ relation, namely in the high probability region, see Figure 7(c). Finally, prescription of the normal distribution to model parameters leads to a linear $M - \xi$ relation as shown in Figure 7(d).

Results for normally distributed model parameters $\boldsymbol{m}$ and **Variant A** are shown in Table 6. The results prove that the $u_{\mathrm{A}} - \xi$ relation is now linear and thus the 1st order polynomials are sufficient for constructing an excellent surrogate. The differences among the various methods here are negligible in terms of accuracy and time requirements.

Figure 8 compares achieved accuracy in estimating the mean of displacement u_{A} for all the variants of the presented distribution of the parameters. The graphs show convergence of the mean estimation with help of the Monte Carlo method with 10^2 to 10^7 simulations for the full numerical model compared to the estimations obtained with coefficients of polynomials of the 4th degree.

The most accurate estimation in the variant of the prescribed histograms (Figure 8(a)) is obtained with a surrogate model based on the regression method, while stochastic collocation based on GQN rules yields the worst result. For the new histograms (Figure 8(b)), all methods except the semi-intrusive Galerkin method provide very accurate results. The last graph in Figure 8(c) shows excellent estimations for all methods investigated.

The errors in prediction of safety margin and failure probability are given in Table 7. The
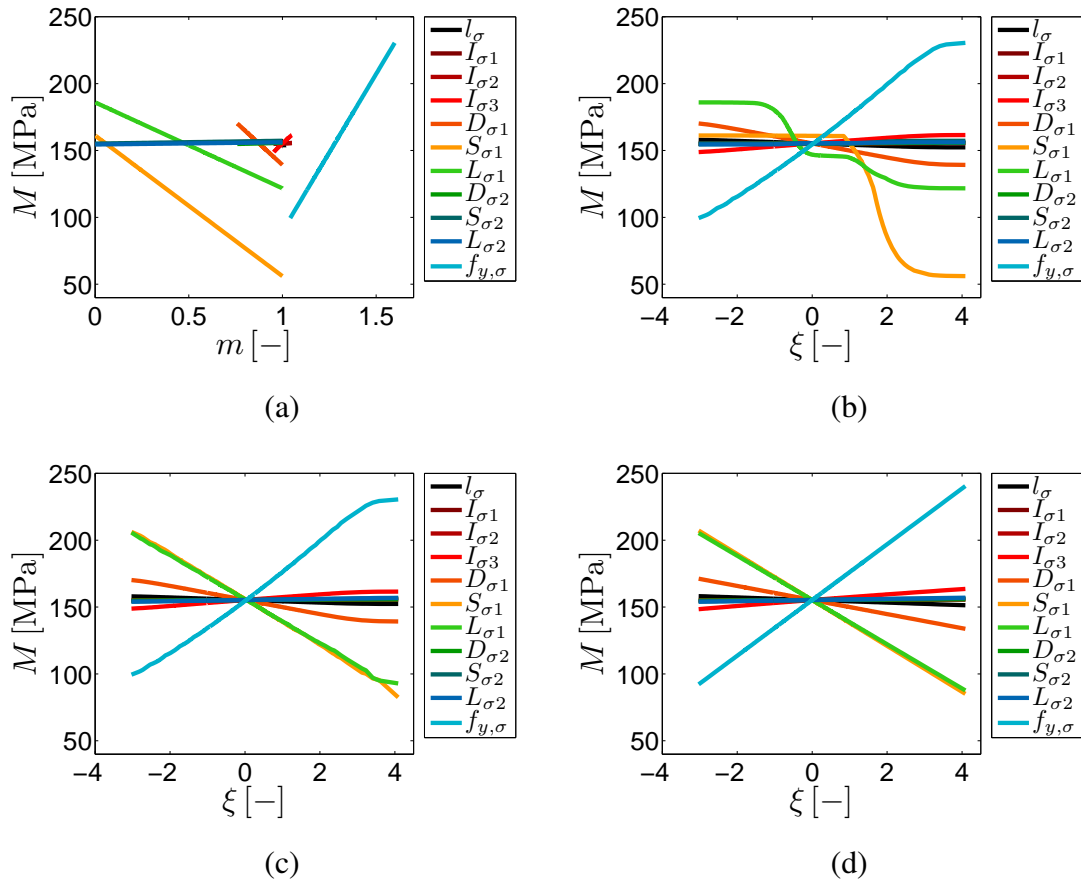

Figure 7: Functional dependence of safety margin M on model parameters $\boldsymbol{m}$ (a), on standard variables $\boldsymbol{\xi}$ for the **prescribed histograms** (b), **new histograms** (c), and **normal distribution** (d).

Method	n_p	n_d	Time [s]	μ_{u_A} [mm]	σ_{u_A} [mm]	μ_{w_A} [mm]	σ_{w_A} [mm]	μ_{φ_A} [mrad]	$\sigma_{\varphi_\mathrm{A}}$ [mrad]
MC	−	10^7	3692	0.207	0.033	0.009	0.002	4.090	0.795
				$\varepsilon_{\mu_{u_\mathrm{A}}}$ [%]	$\varepsilon_{\sigma_{u_\mathrm{A}}}$ [%]	$\varepsilon_{\mu_{w_\mathrm{A}}}$ [%]	$\varepsilon_{\sigma_{w_\mathrm{A}}}$ [%]	$\varepsilon_{\mu_{\varphi_\mathrm{A}}}$ [%]	$\varepsilon_{\sigma_{\varphi_\mathrm{A}}}$ [%]
LHS	1	21	0	$5.2 \cdot 10^{-2}$	0.2800	$5.8 \cdot 10^{-2}$	0.5300	$6.6 \cdot 10^{-2}$	0.5800
	2	201	0	$7.9 \cdot 10^{-3}$	0.0140	$6.8 \cdot 10^{-3}$	0.0042	$7.5 \cdot 10^{-3}$	0.0059
	3	1201	2	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
	4	5301	10	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
KPN	1	21	0	$7.9 \cdot 10^{-3}$	0.0490	$5.5 \cdot 10^{-3}$	0.0530	$7.9 \cdot 10^{-3}$	0.0260
	2	201	0	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
	3	1201	0	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
	4	5301	3	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
GQN	1	21	0	$7.9 \cdot 10^{-3}$	0.0050	$5.5 \cdot 10^{-3}$	0.0530	$7.9 \cdot 10^{-3}$	0.0260
	2	221	0	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0031	$7.9 \cdot 10^{-3}$	0.0029
	3	1581	1	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
	4	8761	5	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
GM	1	−	0	$7.9 \cdot 10^{-3}$	0.0045	$5.5 \cdot 10^{-3}$	0.0500	$7.9 \cdot 10^{-3}$	0.0170
	2	−	0	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0031	$7.9 \cdot 10^{-3}$	0.0029
	3	−	3	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029
	4	−	45	$7.9 \cdot 10^{-3}$	0.0069	$5.5 \cdot 10^{-3}$	0.0032	$7.9 \cdot 10^{-3}$	0.0029

Table 6: Time requirements and errors in predicting mean and standard deviation of displacement vector components for **normal distribution** with model parameters m.
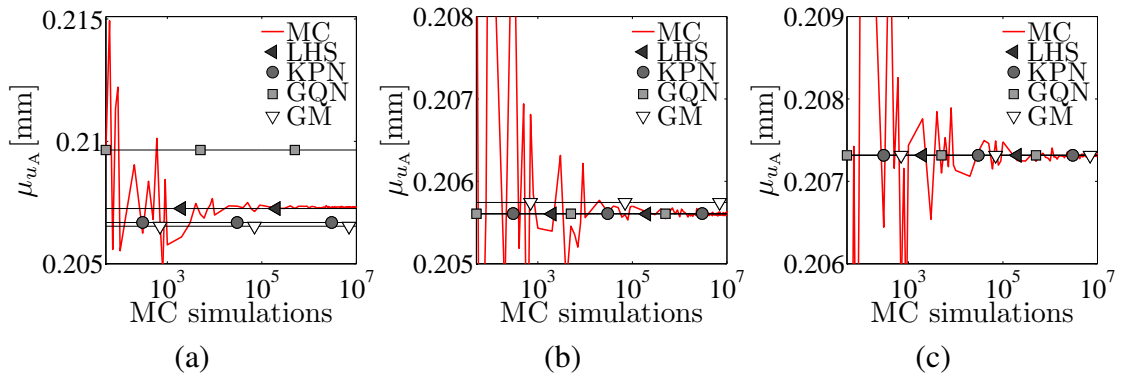

Figure 8: Comparison of predicted mean of displacement u_A based on PCE and MC with different numbers of simulations for the **prescribed histograms** (a), **new histograms** (b), and **normal distribution** (c).

$M - \xi$ relation is now linear and thus the 1st order polynomials are sufficient for constructing an excellent surrogate using all the particular methods.

5. Conclusion

The paper presents a survey and comparison of three methods for construction of a polynomial chaos-based surrogate for a numerical model assuming random model parameters. The methods investigated include the regression method based on Latin Hypercube Sam-

pling, stochastic collocation, and the stochastic Galerkin method. Particular features of these methods are discussed in the paper. The quality of obtained surrogates in terms of accuracy and time requirements is demonstrated using a comparison to the traditional Monte Carlo method with a frame structure serving as a simple illustrative example.

To obtain a PCE-based surrogate model, specific orthogonal polynomials corresponding to the probability distribution of the underlying variables must be used. The orthogonality enables to compute statistical moments of approximated model response analytically so PCE can be used very efficiently in uncertainty quantification as well as in sensitivity analysis. In this study, Hermite polynomials are employed to approximate a model response as a function of standard normal variables. The regression method provides the most accurate PCE-based approximation and thus the best results for uncertainty quantification. The stochastic collocation method has a problem with convergence. The semi-intrusive stochastic Galerkin method behaves similarly to the collocation method because the same quadrature rules are used for numerical integration in both methods. The stochastic Galerkin method is also employed in its fully intrusive form, where all numerical estimations are eliminated and integration is performed analytically. Results of this form of the stochastic Galerkin method are good, but it is worth mentioning that the application of this method is more complicated than for the other methods because reformulation of the full numerical model is required.

In terms of time requirements, all the methods investigated are comparable and, in comparison to the Monte Carlo method, they are significantly less time-consuming. The goal of the example presented is to predict the probability of failure of a simple engineering structure. For this, two variants of PCE applications are analyzed: **A)** approximation of selected structural displacements and rotations, **B)** approximation of safety margin. It is worth men-

Method	n_p	Variant B				Variant A			
		n_d	Time [s]	$\mathrm{Pr}(F)\,[-]$	$\varepsilon_M\,[\%]$	n_d	Time [s]	$\mathrm{Pr}(F)\,[-]$	$\varepsilon_M\,[\%]$
MC	–	10^7	3819	$12\cdot 10^{-7}$	–	10^7	3773	$12\cdot 10^{-7}$	–
LHS	1	23	32	$12\cdot 10^{-7}$	$1.14\cdot 10^{-1}$	19	439	$12\cdot 10^{-7}$	$2.49\cdot 10^{-2}$
	2	243	179	$12\cdot 10^{-7}$	$2.10\cdot 10^{-3}$	163	515	$12\cdot 10^{-7}$	$2.50\cdot 10^{-4}$
	3	1607	802	$12\cdot 10^{-7}$	$4.17\cdot 10^{-5}$	871	738	$12\cdot 10^{-7}$	$3.05\cdot 10^{-6}$
	4	7789	2987	$12\cdot 10^{-7}$	$1.35\cdot 10^{-6}$	3481	1374	$12\cdot 10^{-7}$	$4.88\cdot 10^{-8}$
KPN	1	23	38	$12\cdot 10^{-7}$	$7.42\cdot 10^{-2}$	19	447	$12\cdot 10^{-7}$	$1.34\cdot 10^{-2}$
	2	243	214	$12\cdot 10^{-7}$	$1.30\cdot 10^{-3}$	163	521	$12\cdot 10^{-7}$	$1.49\cdot 10^{-4}$
	3	1607	875	$12\cdot 10^{-7}$	$3.32\cdot 10^{-5}$	871	725	$12\cdot 10^{-7}$	$2.21\cdot 10^{-6}$
	4	7789	2997	$12\cdot 10^{-7}$	$1.00\cdot 10^{-6}$	3481	1362	$12\cdot 10^{-7}$	$4.08\cdot 10^{-8}$
GQN	1	23	31	$12\cdot 10^{-7}$	$7.42\cdot 10^{-2}$	19	444	$12\cdot 10^{-7}$	$1.34\cdot 10^{-2}$
	2	265	212	$12\cdot 10^{-7}$	$1.30\cdot 10^{-3}$	181	520	$12\cdot 10^{-7}$	$1.49\cdot 10^{-4}$
	3	2069	848	$12\cdot 10^{-7}$	$3.32\cdot 10^{-5}$	1177	734	$12\cdot 10^{-7}$	$2.21\cdot 10^{-6}$
	4	12453	2796	$12\cdot 10^{-7}$	$9.98\cdot 10^{-6}$	5965	1376	$12\cdot 10^{-7}$	$4.08\cdot 10^{-8}$
GM	1	–	–	–	–	–	456	$12\cdot 10^{-7}$	$1.34\cdot 10^{-2}$
	2	–	–	–	–	–	551	$12\cdot 10^{-7}$	$1.49\cdot 10^{-4}$
	3	–	–	–	–	–	746	$12\cdot 10^{-7}$	$2.21\cdot 10^{-6}$
	4	–	–	–	–	–	1425	$12\cdot 10^{-7}$	$4.07\cdot 10^{-8}$

Table 7: Time requirements, probability of failure and errors in predicting safety margin for **normal distribution** with model parameters m.

tioning that for **Variant B**, where PCE approximates the safety margin, PCE involves one additional random variable - yield stress - and the dimension of PCE is thus by one higher than for **Variant A**. On the other hand, for **Variant A**, we approximate a set of five displacements and rotations by constructing five PCEs. From Tables 2, 4 and 6 we can conclude that the time required for construction of the PCEs (including the evaluation of model simulations for LHS or stochastic collocation) is negligible and the most computational time is needed for repeated evaluations of PCE within the sampling of failure probability. Computational time grows exponentially with the number of variables and polynomial order. We can thus point out that evaluating of five ten-dimensional PCEs is faster than the evaluation of one eleven-dimensional PCE up to the second order. With the third order, the computational time needed for evaluation of one eleven-dimensional PCE becomes increasingly more demanding.

The paper also demonstrates the practical aspects of PCE application related to nonlinearity of the approximated relationship. Results in Table 7 correspond to a utopian situation, where the approximated relationship is linear (see Fig. 7d). The approximation is thus exact even in the case of the first order PCE for both variants **A** and **B**. This leads to significant time savings.

In order to benefit from the orthogonality of polynomial basis w.r.t. distribution of random variables, we have to involve some transformation from some chosen standard distribution in the case of the random variables defined by histograms. The nonlinearity of the approximated relationship thus consists not only of nonlinearity of the relationship between model responses and random inputs, but also the transformation from standard random variables. Results in Tables 4 and 5 correspond to such a situation, with random inputs defined by histograms very close to normal distribution and a nearly linear underlying transformation, see Fig. 7c. Nevertheless, the predictions of failure probability are remarkably worse for **Variant B**; for **Variant A**, they are still precise. This is because of the yield stress involved in the **Variant B** PCE. Its transformation is nonlinear only in the low probability region, but this small nonlinearity is important due to high sensitivity of the safety margin for this input (according to Fig. 7c, it belongs among the three most important inputs). We can conclude that nonlinearity, even in only low probability regions, is significant for predicting failure probability.

Finally, an even more significant difference between both variants is demonstrated in Table 3 for the prescribed histograms taken from the literature. Fig. 2 reveals that the histograms prescribed to short- and long-lasting loads are far from exhibiting normal distributions, and the corresponding transformation from normal variables depicted in Fig. 5 is highly nonlinear. This nonlinearity is thus remarkable, also in the approximated relationship of the safety margin as visible in Fig. 7b. This nonlinearity is present equally in both variants, but predictions for **Variant B** worsened more significantly (even in terms of orders), although the difference in both variants did not change and consists only of an additional variable - yield stress - for **Variant B**. Therefore, significantly worse results for **Variant B** compared to **Variant A** are caused only by a slightly higher dimension for the approximated relationship.

We thus conclude that failure probability is extremely sensitive to approximation errors. This is related namely to nonlinearity and dimensionality of the approximated relationship. The results presented show that - regarding the computational time and accuracy of predictions - it is more efficient to construct a set of five ten-dimensional approximations than one

approximation with eleven dimensions. Moreover, when applying orthogonal polynomial chaos, it is important to be careful when introducing nonlinear transformations to standard random variables.

[1] F. Augustin, A. Gilg, M. Paffrath, P. Rentrop, M. Villegas, U. Wever, "An accuracy comparison of polynomial chaos type methods for the propagation of uncertainties", Journal of Mathematics in Industry, Journal of Mathematics in Industry, 3(1), 2013.

[2] I. Babuska, R. Tempone, G. E. Zouraris, "Galerkin Finite Element Approximations of Stochastic Elliptic Partial Differential Equations", SIAM Journal on Numerical Analysis, 42(2), 800–825, 2004.

[3] I. Babuska, F. Nobile, R. Tempone, "A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data", SIAM Journal on Numerical Analysis, 45(3), 1005–1034, 2007.

[4] G. Blatman, B. Sudret, "An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis", Probabilistic Engineering Mechanics, 25(2), 183–197, 2010.

[5] G. Blatman, B. Sudret, "Adaptive sparse polynomial chaos expansion based on least angle regression", Journal of Computational Physics, 230(6), 2345-2367, 2011.

[6] H. Cheng, A. Sandu, "Efficient uncertainty quantification with the polynomial chaos method for stiff systems", Mathematics and Computers in Simulation, 79(11), 3278–3295, 2009.

[7] S.-K. Choi, R. V. Grandhi, R. A. Canfield, C. L. Pettit, "Polynomial Chaos Expansion with Latin Hypercube Sampling for Estimating Response Variability", AIAA Journal, 42(6), 1191–1198, 2004.

[8] O. Ditlevsen, H. O. Madsen, "Structural Reliability Methods", John Wiley & Sons Ltd, Chichester, England, 1996.

[9] M. Eigel, C. J. Gittelson, C. Schwab, E. Zander, "Adaptive stochastic Galerkin FEM", Computer Methods in Applied Mechanics and Engineering, 270(1), 247–269, 2014.

[10] M. S. Eldred, J. Burkardt, "Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification", American Institute of Aeronautics and Astronautics, 976.2009, 1–20, 2009.

[11] H. C. Elman, C. W. Miller, E. T. Phipps, R. S. Tuminaro, "Assessment of collocation and galerkin approaches to linear diffusion equations with random data", International Journal for Uncertainty Quantification, 1(1), 19–33, 2011.

[12] A. Fülöp, M. Iványi, "Safety of a column in a frame", In P. Marek, J. Brozzetti & M. Guštar & P. Tikalsky, editors, Probabilistic assessment of structures using Monte Carlo simulation, Institute of Theoretical and Applied Mechanics, Academy of Sciences of the Czech Republic, Prague, second edition, 2003.

[13] R. G. Ghanem, P. D. Spanos, "Stochastic Finite Elements: A Spectral Approach", Dover Publications, Revised edition, 2012.

[14] M. Gutiérrez, S. Krenk, " Encyclopedia of Computational Mechanics, chap. Stochastic finite element methods", John Wiley & Sons, Ltd., 2004.

[15] F. Heiss, V. Winschel, "Likelihood approximation by numerical integration on sparse grids", Journal of Econometrics, 144(1), 62–80, 2008.

[16] S. Hosder, R. W. Walters, M. Balch, "Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables", Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, AIAA paper, Vol. 1939, 2007.

[17] C. Hu, B. D. Youn, "Adaptive-sparse polynomial chaos expansion for reliability analysis and design of complex engineering systems", Structural and Multidisciplinary Optimization, 43(3), 419-442, 2011.

[18] E. Janouchová, A. Kučerová, "Competitive Comparison of Optimal Designs of Experiments for Sampling-based Sensitivity Analysis", Computers & Structures, 124, 47–60, 2013.

[19] E. Janouchová, A. Kučerová, J. Sýkora, "Polynomial Chaos Construction for Structural Reliability Analysis", in Y. Tsompanakis, J. Kruis, B.H.V. Topping, (Editors), "Proceedings of the Fourth International Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering", Civil-Comp Press, Stirlingshire, UK, Paper 9, 2015. doi:10.4203/ccp.109.9.

[20] J. Li, D. Xiu, "Evaluation of failure probability via surrogate models", Journal of Computational Physics, 229(23), 8966–8980, 2010.

[21] J. Li, J. Li, D. Xiu, "An efficient surrogate-based method for computing rare failure probability", Journal of Computational Physics, 230(24), 8683–8697, 2011.

[22] X. Ma, N. Zabaras, "An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations", Journal of Computational Physics, 228(8), 3084-3113, 2009.

[23] H. G. Matthies, A. Keese, "Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations", Computer Methods in Applied Mechanics and Engineering, 194(12–16), 1295–1331, 2005.

[24] H. G. Matthies, "Uncertainty quantification with stochastic finite elements", In Erwin Stein, René de Borst, and Thomas J. R. Hughes, editors, Encyclopaedia of Computational Mechanics, John Wiley & Sons, Chichester, 2007.

[25] H. N. Najm, "Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics", Annual Review of Fluid Mechanics, 41(1), 35-52, 2009.

[26] F. Nobile, R. Tempone, C. G. Webster, "A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data", SIAM Journal on Numerical Analysis, 46(5), 2309–2345, 2008.

[27] M. Paffrath, U. Wever, "Adapted polynomial chaos expansion for failure detection. Journal of Computational Physics", 226(1), 263–281, 2007.

[28] M. P. Pettersson, G. Iaccarino, J. Nordstrm, "Polynomial Chaos Methods for Hyperbolic Partial Differential Equations, chap. Polynomial Chaos Methods", Mathematical Engineering, Springer International Publishing, 23–29, 2015.

[29] R. Pulch, "Stochastic collocation and stochastic Galerkin methods for linear differential algebraic equations", Journal of Computational and Applied Mathematics, 262, 281–291, 2014.

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

[31] N. Wiener, "The Homogeneous Chaos", American Journal of Mathematics, 60(4), 897–936, 1938.

[32] D. Xiu, G. E. Karniadakis, "The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations", SIAM Journal on Scientific Computing, 24(2), 619–644, 2002.

[33] D. Xiu, J. S. Hesthaven, "High-Order Collocation Methods for Differential Equations with Random Inputs", SIAM Journal on Scientific Computing, 27(3), 1118–1139, 2005.

[34] D. Xiu, "Fast Numerical Methods for Stochastic Computations: A Review", Communications in Computational Physics, 5(2–4), 242–272, 2009.

[35] D. Xiu, "Numerical Methods for Stochastic Computations: A Spectral Method approach", Princeton University Press, 2010.

LIST OF AUTHOR'S PUBLICATIONS

Theses

Kučerová, A. (2003). Optimisation de forme et contrôle de chargement des structures elastique soumis de rotation finis en utilisant les algorithmes génétiques. Master's thesis.

Kučerová, A. (2007). *Identification of nonlinear mechanical model parameters based on softcomputing methods*. PhD thesis, Ecole Normale Supérieure de Cachan, Laboratoire de Mécanique et Technologie.

WOS journal papers

Havelka, J., Kučerová, A., and Sýkora, J. (2019). Dimensionality reduction in thermal tomography. *Computers & Mathematics with Applications*, 78(9):3077–3089.

Janouchová, E., Sýkora, J., and Kučerová, A. (2018). Polynomial chaos in evaluating failure probability: A comparative study. *Applications of Mathematics*, 63(6):713–737.

Jarušková, D. and Kučerová, A. (2017). Estimation of thermophysical parameters revisited from the point of view of nonlinear regression with random parameters. *International Journal of Heat and Mass Transfer*, 106:135–141.

Havelka, J., Kučerová, A., and Sýkora, J. (2016). Compression and reconstruction of random microstructures using accelerated lineal path function. *Computational Materials Science*, 122:102–117.

Mareš, T., Janouchová, E., and Kučerová, A. (2016). Artificial neural networks in calibration of nonlinear mechanical models. *Advances in Engineering Software*, 95:68–81.

Kučerová, A. and Lepš, M. (2014). Soft computing-based calibration of microplane M4 model parameters: Methodology and validation. *Advances in Engineering Software*, 72:226–235.

Janouchová, E. and Kučerová, A. (2013). Competitive comparison of optimal designs of experiments for sampling-based sensitivity analysis. *Computers & Structures*, 124:47–60.

Novák, J., Kučerová, A., and Zeman, J. (2013). Microstructural enrichment functions based on stochastic wang tilings. *Modelling and Simulation in Materials Science and Engineering*, 21(2):025014.

Rosić, B. V., Kučerová, A., Sýkora, J., Pajonk, O., Litvinenko, A., and Matthies, H. G. (2013). Parameter identification in a probabilistic setting. *Engineering Structures*, 50:179–196

Kučerová, A. and Sýkora, J. (2013). Uncertainty updating in the description of coupled heat and moisture transport in heterogeneous materials. *Applied Mathematics and Computation*, 219(13):7252–7261.

Novák, J., Kučerová, A., and Zeman, J. (2012). Compressing random microstructures via stochastic wang tilings. *Physical Review E*, 86(4):040104.

Kučerová, A., Sýkora, J., Rosić, B., and Matthies, H. G. (2012). Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials. *Journal of Computational and Applied Mathematics*, 236(18):4862–4872.

Kučerová, A., Brancherie, D., Ibrahimbegović, A., Zeman, J., and Bittnar, Z. (2009). 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):128–144.

Ibrahimbegović, A., Knopf-Lenoir, C., Kučerová, A., and Villon, P. (2004). Optimal design and optimal control of structures undergoing finite rotations and elastic deformations. *International Journal for Numerical Methods in Engineering*, 61(14):2428–2460.

Hrstka, O. and Kučerová, A. (2004). Improvements of real coded genetic algorithms based on differential operators preventing the premature convergence. *Advances in Engineering Software*, 35(3–4):237–246.

Hrstka, O., Kučerová, A., Lepš, M., and Zeman, J. (2003). A competitive comparison of different types of evolutionary algorithms. *Computers & Structures*, 81(18–19):1979–1990.

Non-WOS journal papers

Kučerová, A. and Matthies, H. G. (2010). Uncertainty updating in the description of heterogeneous materials. *Technische Mechanik*, 30(1–3):211–226.

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

Book chapters

Rosić, B., Sýkora, J., Pajonk, O., Kučerová, A., and Matthies, H. G. (2016). Comparison of numerical approaches to bayesian updating. In *Computational Methods for Solids and Fluids*, pages 427–461. Springer.

Selected proceedings papers

Jarušková, D. and Kučerová, A. (2017). Different views on additional random parameters in experiment design for thermophysical parameters estimation. In *UNCECOMP 2017 - Proceedings of the 2nd International Conference on Uncertainty Quantification in Computational Sciences and Engineering*, pages 543–553.

Kučerová, A., Sýkora, J., Janouchová, E., Jarušková, D., and Chleboun, J. (2016). Acceleration of robust experiment design using sobol indices and polynomial chaos expansion. In *Proceedings of 7th International Workshop on Reliable Engineering Computing*, pages 411–425.