Advanced Modeling and Simulation in Engineering Sciences

RESEARCH ARTICLE

Open Access

Bayesian stochastic multi-scale analysis via energy considerations

Muhammad S. Sarfaraz[1*], Bojana V. Rosić[2], Hermann G. Matthies[1] and Adnan Ibrahimbegović[3]

*Correspondence:
m.sarfaraz@tu-braunschweig.de
[1] Institute of Scientific
Computing, Technische
Universität Braunschweig,
Braunschweig 38106, Germany
Full list of author information is
available at the end of the article

Abstract

Multi-scale processes governed on each scale by separate principles for evolution or equilibrium are coupled by matching the stored energy and dissipation in line with the Hill-Mandel principle. We are interested in cementitious materials, and consider here the macro- and meso-scale behaviour of such a material. The accurate representations of stored energy and dissipation are essential for the depiction of irreversible material behaviour, and here a Bayesian approach is used to match these quantities on different scales. This is a probabilistic upscaling and as such allows to capture, among other things, the loss of resolution due to scale coarsening, possible model errors, localisation effects, and the geometric and material randomness of the meso-scale constituents in the upscaling. On the coarser (macro) scale, optimal material parameters are estimated probabilistically for certain possible behaviours from the class of generalised standard material models by employing a nonlinear approximation of Bayes's rule. To reduce the overall computational cost, a model reduction of the meso-scale simulation is achieved by combining unsupervised learning techniques based on a Bayesian copula variational inference with functional approximation forms.

Keywords: Upscaling, Bayesian updating, Gauss–Markov–Kalman filter

Introduction

The predictive modelling of highly nonlinear and damaging material behaviour such as occurs in cementitious-like materials which are heterogeneous over a large range of scales requires realistic mathematical models. As a detailed description on the desired macroscopic level is not computationally feasible for large-scale structures, some kind of multi-scale approach is called for. In this paper simple prototypical models of macro- and meso-scale descriptions of cementitious materials will be considered. However, smaller scales can be introduced as well in order to explicitly describe heterogeneities characterising the material structure of aggregates, the mortar matrix, or the interfacial zone.

Conceptually, the prevalent computational methods to tackle multi-scale problems can be classified into concurrent and non-concurrent approaches. Concurrent schemes consider both the macro- and meso-scales during the course of the simulation e.g. the FE^2 method [12,13], whereas the non-concurrent ones are based on a scale separation idea, by which the desired quantity of interest (QoI), e.g. average stresses or strains, are estimated given numerical experiments on a representative volume element (RVE), see [17]

for a recent overview on related techniques. Although the non-concurrent method has proven to work very well for elastic properties, i.e. homogenisation especially under the assumption of scale separation, problems appear when this is not satisfied, or the material is loaded into a range where irreversible material changes occur under complex loading paths, possibly with induced localisation effects, which are of crucial concern when dealing with material non-linearities such as plasticity and damage. In [22,41,44] a combination of concurrent and non-concurrent approaches is proposed. In case heterogeneity is present over a larger range of scales, the so-called size-effect problem appears and has to be resolved [9]. This is difficult to handle with standard local material models in the FE^2 method [12,13] as integration points loose any size information. In [22,36,37,44] this problem is approached by considering the *mesh in element approach* (MIEL) [24,25,41], in which the meso-scale structure is embedded in a macro-scale finite element. This allows the precise transfer of size-information between the scales.

In this paper we take the MIEL approach, and focus on the stochastic multiscale problem in which the meso-scale information is described by variations reflecting aleatoric uncertainties describing the geometry, the spatial distribution and the material properties of the individual meso-scale material constituents and their mutual interaction. The idea is to design an appropriate stochastic macro-scale model and to estimate its corresponding stochastic parameters such that the aleatoric meso-scale information is preserved.

In the literature stochastic homogenisation is usually performed on an ensemble of RVEs in order to extract the relevant statistical QoI [32–35,52,53,58–60]. For example, in [19,29], the moving-window approach is used to characterise the probabilistic uni-axial compressive strength of random micro-structures. Another active direction of research is to develop stochastic surrogate models for strain energy of random micro-structures as in [6–8,57]. The main goal is to mitigate the effect of the curse of dimensionality due to a large number of stochastic dimensions. With the rapid expansion of machine learning and data driven techniques, the current trend is to train neural network based approximate models [28,30,48,65,66] as a cheaper computational alternative in multiscale methods. Preserving mechanical invariance properties in such settings is a difficult task, which is why energy and dissipation based scale transfer methods were proposed [41]. Furthermore, to obtain a probabilistic description of macro-scale characteristic by incorporating micro-scale measurements, Bayesian methods have been applied to such problems with promising success, please see [5,14–16,26] for recent applications in the formulation of high dimensional probabilistic inverse problems generally, and estimating distribution of material parameters specifically. Moreover [11,27,50,54] demonstrate the application of the Bayesian framework to multi-scale problems.

In this paper we assume that the stochastic macro-scale continuum model can be described as a stochastic generalised standard material model [18,20,24], characterised by the specification of the stored energy and the dissipation. By using physics based principles and Bayesian inference, we employ these two thermodynamics functions as a coupling constituent between the stochastic meso- and macro-scales as previously suggested in [25,41,49–51] for one specific realisation of a representative volume element (RVE). Hence, the stochastic stored energy and dissipation on the meso-scale are captured, and further used as measurements in the Bayesian estimation of the stochastic macro-scale properties of the generalised model. In this regard, the focus of the paper is two-fold: on one hand we suggest the Bayesian approach for upscaling the uncertain meso-

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.* (2020)7:50

Page 3 of 35

scale information to the macro-scale, and on the other we suggest a novel approach for reducing the dimension of the stochastic meso-scale measurement. By repeating experiments on several RVEs we collect the meso-scale measurement samples, i.e. instances of measurement data that represent aleatoric uncertainty. Furthermore, we assume that the distribution of measurement uncertainty is not known, and search for the minimally parametrized model that represents the stochastic meso-scale data via an unsupervised learning algorithm. The goal is to represent the meso-scale measurement as a nonlinear function of simpler independent random variables (e.g. Gaussian random variables), and hence indirectly determine the distribution of the meso-scale data. Here this is achieved by employing a copula based Bayesian variational inference on a generalised mixture model. Once the measurement data are quantified we use them in a Bayesian upscaling procedure to estimate the stochastic macro-scale properties.

The paper is organised as follows: a generalised model problem is presented in "Abstract model problem" section, and the research questions are discussed. "Bayesian upscaling of random meso-structures" section describes a Bayesian framework for upscaling random meso-scale information with a particular focus on the approximate posterior estimation. From a computational point of view this is done through unsupervised learning. The energy-like observations used to map the information between the scale is described in concrete terms using an example material model in "Bayesian upscaling of random meso-structures" section and its algorithm performance is analysed on two different numerical examples in "Numerical results" section. "Conclusion" section concludes the discussion on the proposed approach.

Abstract model problem

Building on earlier work [22,25,41,47,49–51], a stochastic multi-scale formulation is developed which allows stochastic upscaling in order to have simpler macro-scale material models and thus save expensive meso-scale evaluations. We concentrate here on isothermal and quasi-static but irreversible material behaviour at small strains. In the linear (elastic) case when the meso- and macro-scale are well separated, homogenisation approaches can be successfully used, in which the macro-scale properties such as the stiffness matrix are evaluated, given the meso-scale structural response of a representative material element. If the scale separation criterion is not satisfied, the upscaling becomes more difficult. This is even more pronounced in the nonlinear and irreversible case, as the upscaling of meso-scale information to the macro-scale is not as straightforward. The so-called size-effect problem appears [9], as the size relation of meso-scale features to the macro-scale has an influence on the macro-scale response. To overcome this issue, in this paper the focus is put on the so-called the *mesh in element approach* (MIEL) [24,25,41], in which the meso-scale structure is embedded in a macro-scale finite element.

As will become clear later, the meso-scale model actually does not have to be a continuum model, e.g. [22,23], but for the macro-scale we assume a continuum model, which obeys the usual quasi-static iso-thermal equilibrium equations

$$ -\operatorname{div} \boldsymbol{\sigma}(\boldsymbol{u}) = \boldsymbol{f} \text{ in } \mathcal{G}, \quad \boldsymbol{u} = \boldsymbol{u}_0 \text{ on } \Gamma_d \subset \partial\mathcal{G}, \quad \boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{g} \text{ on } \Gamma_n \subset \partial\mathcal{G}, \tag{1} $$

plus appropriate essential boundary conditions $\boldsymbol{u}_0$ for the displacement $\boldsymbol{u}$ of the macro-scale body on the Dirichlet boundary $\Gamma_d \subset \partial\mathcal{G}$, where $\mathcal{G} \subset \mathbb{R}^d$ is the domain occupied

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 4 of 35

by said body, under volume load f and boundary tractions g on the Neumann boundary $\Gamma_n \subset \partial\mathcal{G}$, equilibrated by the Cauchy stress $\sigma(u)$ in the body, which depends on the displacement through constitutive relations, which we turn to next.

As this is to be a model for possibly more complex behaviour of the meso-scale, we assume that the macro-scale continuum model can be described as a generalised standard material model [18,20,24]. These materials obey the maximum dissipation hypothesis, and are thus in a sense *optimal* in fulfilling the requirements of the second law of thermodynamics. They have the additional advantage that these materials are completely characterised by the specification of two scalar functions, the stored energy resp. Helmholtz free energy density W, and the dissipation pseudo-potential density F. In our view this description is also a key for the connection with the meso-scale behaviour. No matter how the physical and mathematical/computational description on the meso-scale has been chosen, in all cases where the description is based on physical principles it will be possible to define the stored (Helmholtz free) energy and the dissipation (entropy production). These two thermodynamic functions will thus be employed as "measurements" resp. coupling quantitites in the Bayesian inference used to identify the macro-scale model parameters given the meso-scale response stored energy and dissipation. This approach is also a good start for computational procedures (e.g. [10,24,40,55]), and has for some specific subset of materials been given a fully variational formulation in a Hilbert space context [21]. This description has been subsequently extended to much more general cases for the here interesting case of rate-independent behaviour [43].

To summarise the generalised standard material model in a nutshell [18,20,24,40,42], for an isothermal small-strain situation with strain ε, from the Clausius–Duhem inequality it follows that the stress σ at some material point $x \in \mathcal{G}$ is

$$\sigma(x) = \mathrm{D}_{\varepsilon} W(\varepsilon, w, Q, x), \text{ with } \varepsilon(x) = \nabla^s u(x), \tag{2}$$

where ∇^s is the symmetric part of the gradient, w are the internal phenomenological variables [18,20,24,40,42] describing possibly irreversible changes in the material, D_{ε} is the partial derivative w.r.t. ε, the collection Q of tensors of even order describes the specific material (and has to be identified later), and $\zeta = -\mathrm{D}_w W(\varepsilon, w, Q)$ are "thermodynamic forces" conjugate to the "thermodynamic fluxes" $v := \dot{w}$ —the inner product $\langle \zeta, \dot{w} \rangle$ is a rate of (dissipated) energy.

The evolution of w—i.e. $v = \dot{w}$—is then [18,20,24,40,42] defined by the dissipation pseudo-potential F through a variational inclusion:

$$\zeta \in \partial_v F(\varepsilon, w, \dot{w}, Q) \quad \Leftrightarrow \quad \dot{w} \in \partial_{\zeta} F^*(\varepsilon, w, \zeta, Q), \tag{3}$$

where $\partial_v F$ is the subdifferential of F w.r.t v, which is by definition equivalent with the variational inequalities

$$\forall v: \ F(\varepsilon, w, v, Q) \geq F(\varepsilon, w, \dot{w}, Q) + \langle \zeta, v - \dot{w} \rangle \tag{4}$$

$$\Leftrightarrow \quad \forall \xi: \ F^*(\varepsilon, w, \xi, Q) \geq F^*(\varepsilon, w, \zeta, Q) + \langle \xi - \zeta, \dot{w} \rangle; \tag{5}$$

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 5 of 35

here F^* is the Legendre–Fenchel dual of the dissipation potential $F(\boldsymbol{\varepsilon}, \boldsymbol{w}, \boldsymbol{v}, \boldsymbol{Q})$ w.r.t. $\boldsymbol{v} = \dot{\boldsymbol{w}}$, where F is assumed as a convex and lower semi-continuous function w.r.t. that variable. Convex analysis and variational inequalities enter, as for rate-independent material behaviour—which we are interested in here—the dissipation pseudo-potential F *cannot* be smooth [20,24,40,42,43], in fact it is positively homogeneous of the first degree in $\boldsymbol{v} = \dot{\boldsymbol{w}}$; thus it is appropriate to use the *subdifferential* $\partial_{\boldsymbol{v}}$ w.r.t. $\boldsymbol{v} = \dot{\boldsymbol{w}}$ resp. $\partial_{\boldsymbol{\zeta}}$ w.r.t. $\boldsymbol{\zeta}$ in (3)—this is a concise form of writing a variational inequality. Specific forms of W and F will be shown and used later.

In [21,43] it is shown how, by collecting the state variables $\boldsymbol{u}$ and $\boldsymbol{w}$ as fields $z(\boldsymbol{x}) = (\boldsymbol{u}(\boldsymbol{x}), \boldsymbol{w}(\boldsymbol{x})) \in \mathcal{Z}_M$ for the whole body as elements of a space $\mathcal{Z}_M = \mathcal{U}_M \times \mathcal{W}_M$, the relations Eq. (1)—with the help of Eq. (2)—and Eq. (3) may be used to define a global energy like function W_M—from W in Eq. (2) and $\boldsymbol{f}, \boldsymbol{g}$ in Eq. (1)—and a global dissipation like function F_M from F in Eq. (3), which describe the equilibrium Eq. (1) and evolution Eq. (3) of z globally. Let $(\mathcal{Z}_M, W_M, F_M)$ be such an abstract structure of a general deterministic rate-independent small-strain homogeneous macro-model, in which $\mathcal{Z}_M$ denotes the state space, $W_M : [0, T] \times \mathcal{Z}_M \to \mathbb{R}$ is a time-dependent energy-like functional which encompasses the loading, and $F_M : \mathcal{Z}_M \times \mathcal{Z}_M \to \mathbb{R}_+$ is an in the second variable $s = \dot{z}$ convex and lower-semicontinuous dissipation pseudo-potential—i.e. $\forall z \in \mathcal{Z}_M : s \mapsto F_M(z, s)$— satisfying the homogeneity property $\forall z, s \in \mathcal{Z}_M : F_M(z, \lambda s) = \lambda F_M(z, s)$ for all $\lambda \geq 0$ for a rate-independent system [20,21,40,42,43]—something inherited from the local function F in Eq. (3). Then the evolution of the state $z \in \mathcal{Z}_M$ of the macro-mechanical system can be described mathematically in an abstract variational manner by the subdifferential inclusion

$$z : [0, T] \to \mathcal{Z}_M : \qquad \partial_s F_M(z, \dot{z}) + \delta_z W_M(t, z) \ni 0 \tag{6}$$

in which δ_z stands for the Gâteaux derivative w.r.t. the state variable $z \in \mathcal{Z}_M$, and the derivative of F_M is given in terms of the set-valued subdifferential $\partial_s F_M$ w.r.t. the second variable s in the sense of the convex analysis, e.g. see [40,43]—see also Eqs. (4) and (5) for reference. Furthermore, we assume that the rate-independent system in Eq. (6) is parameterised by a field Q representing the global form of the variables $\boldsymbol{Q}$ in Eqs. (2) and (3), i.e. the specific material characteristics. This means that the full list of arguments looks like

$$W_M = W_M(t, z; Q); \quad \text{and } F_M = F_M(z, s; Q).$$

Given the mathematical description of the macro-scale model in Eq. (6), the goal is to find the unknown variables Q such that the structural response of the macro-scale model matches the response of the meso-scale model as well as possible. As remarked before, the meso-scale could be any kind of model with the ability to produce values for stored and dissipated energy. Here, for testing purposes and to show how the procedure works, we take as meso-scale structure a more detailed and spatially resolved description of the macro-scale counterpart—one that accounts for material and geometrical heterogenety at a lower-scale level, and hence represents the system that we can evaluate at possibly high computational cost. The macro-scale model is hence one that has to represent equivalent physical behaviour at a much coarser spatial resolution. Thus the meso-scale mathematical

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.* (2020)7:50

Page 6 of 35

model is formally equal to Eq. (6), and reads

$$\partial_{s_m} F_m(z_m, \dot{z}_m) + \delta_{z_m} W_m(t, z_m) \ni 0, \tag{7}$$

and is subjected to equivalent outside actions—incorporated into W_m—as the one in Eq. (6). As before we assume that there are meso-scale parameters Q_m which describe possibly different meso structures. Hence the full list of arguments for the meso-scale energy and dissipation is

$$W_m = W_m(t, z_m; Q_m); \quad \text{and } F_m = F_m(z_m, s_m; Q_m).$$

As $\mathcal{Z}_M \neq \mathcal{Z}_m$, the states z_M and z_m cannot be directly compared, and the two models are to be compared by some observables or measurements $y \in \mathcal{Y}$, where $\mathcal{Y}$ is typically some vector space like $\mathbb{R}^m$. In other words, let

$$y_m = Y_m(z_m(Q_m, f_m)) + \hat{\epsilon}, \tag{8}$$

be the meso-scale observable (e.g. energy, stress or strain etc.) in which Y_m describes the measurement operator, f_m is the external excitation and $\hat{\epsilon}$ is a random variable which represents the measurement noise. On the other side, let

$$y_M = Y_M(Q, z_M(Q, f_M)) \tag{9}$$

be the prediction of the same observation on the macro-scale level, this time described by the measurement operator Y_M and the external excitation f_M of the same type as f_m. To incorporate the possible model error and other discrepancies between the meso- and macro-scale, we model y_M as a noisy variant of Y_M. For this purpose we introduce a probability space $(\Omega_\epsilon, \mathfrak{B}_\epsilon, \mathbb{P}_\epsilon)$—$\Omega_\epsilon$ is the space of elementary events or realisations, $\mathfrak{B}_\epsilon$ is the σ-algebra of measurable events, and $\mathbb{P}_\epsilon$ is the probability measure—and add to $Y_M(Q, z_M(Q, f_M))$ the random variable $\epsilon(\omega_\epsilon) \in L_2(\Omega_\epsilon, \mathfrak{B}_\epsilon, \mathbb{P}_\epsilon)$ that best describes our knowledge about those discrepancies. Hence, Eq. (9) becomes stochastic and reads

$$y_M(\omega_\epsilon) = Y_M(Q, z_M(Q, f_M)) + \epsilon(\omega_\epsilon). \tag{10}$$

Typically, $\epsilon(\omega_\epsilon)$ is modelled as a zero-mean Gaussian random variable $\epsilon \sim \mathcal{N}(0, C_\epsilon)$ with covariance C_ϵ. However, other models for $\epsilon(\omega_\epsilon)$ can also be introduced without modifying the general setting presented in this paper.

This is now the point to discuss different possibilities and choices for identification. Recall that Bayesian identification (e.g. [38,39]) proceeds such that the unknown object is considered as uncertain in an *epistemic* sense—an uncertainty of our knowledge— and modelled as a random variable. New information about that object is then obtained through some connected observation or measurement via *conditioning*, which hopefully will reduce the epistemic uncertainty. Now, if different observations or measurements come from a fixed specimen, or fixed computational model—fixed values of Q_m in Eq. (7) in our case—for different external actions, then it is not an unreasonable interpretation to say what one "really" wants to know is a fixed value of Q, and the uncertainty introduced in the Bayesian modelling is purely epistemic.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.* (2020) 7:50

Page 7 of 35

Trying to achieve a theoretical description in a Bayesian framework, as now Q is to be modelled as a random variable, one needs a stochastic version of the theory sketched so far. This means one needs a stochastic version of continuum thermodynamics [45], with a new interpretation of equilibrium Eq. (1), and a *stochastic generalised standard material* with re-interpreted Eqs. (2) and (3), see [51]. An early version of this for plasticity is [46]. This is achieved by modelling Q as a random quantity due to epistemic uncertainty, and assume that it is a random variable on the probability space $(\Omega_u, \mathfrak{B}_u, \mathbb{P}_u)$ for epistemic uncertainty, for the sake of simplicity we assume that it has finite variance, $Q \in L_2(\Omega_u, \mathfrak{B}_u, \mathbb{P}_u)$. As Q is now a random quantity, the macro-scale Eq. (6) has to be interpreted stochastically, the state becomes a stochastic quantity $z : \Omega_u \to \mathcal{Z}_M$, and Eq. (6) is translated into [46]

$$\partial_s \boldsymbol{F}_M(z, \dot{z}; Q) + \delta_z \boldsymbol{W}_M(t, z; Q) \ni 0, \tag{11}$$

where

$$\boldsymbol{W}_M(t, z; Q) = \mathbb{E}\left(W_M(t, z(\omega_u); Q(\omega_u))\right)_{\Omega_u},$$
$$\boldsymbol{F}_M(z, y; Q) = \mathbb{E}\left(F_M(z(\omega_u), y(\omega_u); Q(\omega_u))\right)_{\Omega_u},$$

the expectation being defined as $\mathbb{E}(W)_{\Omega_u} := \int_{\Omega_u} W(\omega_u)\,\mathbb{P}(d\omega_u)$. The random solution $z(\omega_u)$ from Eq. (11) is the input to Eq. (10) to provide the prediction $y_M(\omega_u, \omega_\epsilon)$. This leads to:

Problem 1 *Find an epistemically uncertain $Q(\omega_u)$ in Eq.* (11)—*a random variable on the space Ω_u —such that the predictions of Eq.* (10) *match those of Eq.* (8) *in a measurement sense.*

The upscaling process that is related to Problem 1 was already considered in [47,49–51], and hence will not be repeated here. However, as we show later, this problem is a special case of the upcoming Problem 2, which is the main topic of this paper.

This Problem 2 deals with the situation that different observations or measurements come from a population of different RVE realisations, further called *aleatoric* uncertainty. In our case observations come from different realisations of Q_m in Eq. (7). In such a case the natural approach would be to capture that randomness also in the macro-scale model, i.e. we want to identify in the macro-scale model a random variable Q.

Thus we model Q_m at each point as a random variable in $L_2(\Omega_Q, \mathfrak{B}_Q, \mathbb{P}_Q; \mathcal{Q}_m)$, i.e. a random field defined by the mapping

$$Q_m(\omega_Q) := Q_m(x, \omega_Q) : \mathcal{G} \times \Omega_Q \mapsto \mathcal{Q}_m, \tag{12}$$

where $\mathcal{Q}_m$ is the parameter space, which depends on the application. As a consequence, the deterministic evolution problem in Eq. (7) also becomes uncertain, and Eq. (7) rewrites to

$$\partial_{s_m} \boldsymbol{F}_m(z_m, \dot{z}_m; Q_m) + \delta_{z_m} \boldsymbol{W}_m(t, z_m; Q_m) \ni 0 \text{ a.s.,} \tag{13}$$

in which

$$\boldsymbol{W}_m(t, z_m; Q_m) = \mathbb{E}\left(W_m(t, z_m(\omega_Q); Q_m(\omega_Q))\right)_{\Omega_Q},$$
$$\boldsymbol{F}_m(z_m, s_m; Q_m) = \mathbb{E}\left(F_m(z_m(\omega_Q), s_m(\omega_Q); Q_m(\omega_Q))\right)_{\Omega_Q},$$

respectively.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.* (2020)7:50

Page 8 of 35

The formal changes are now that the observation in Eq. (8) not only contains the noise described by $\hat{\epsilon}$, but is a *random quantity* of an aleatoric nature, because so is $Q_m(\omega_Q)$ and hence $z_m(\omega_Q)$. Hence, one has:

$$y_m = Y_m(z_m(\omega_Q), Q_m(\omega_Q), f_m)) + \hat{\epsilon}. \tag{14}$$

As we want to see the randomness from Ω_Q reflected in Q on the macro-scale, these quantities now have to become a random variable on the probability space $\Omega_u \times \Omega_Q$, and hence so does the Eq. (11), with the change where

$$\boldsymbol{W}_M(t, z; Q) = \mathbb{E}\left(W_M(t, z(\omega_u, \omega_Q); Q(\omega_u, \omega_Q))\right)_{\Omega_u \times \Omega_Q},$$
$$\boldsymbol{F}_M(z, s; Q) = \mathbb{E}\left(F_M(z(\omega_u, \omega_Q), s(\omega_u, \omega_Q); Q(\omega_u, \omega_Q))\right)_{\Omega_u \times \Omega_Q},$$

i.e. the expected values are now on the space $\Omega_u \times \Omega_Q$.

The prediction of the observation Eq. (10) has now to be interpreted in this way too, it contains also the aleatoric randomness from Ω_Q. With this re-interpretation of Eq. (10), this in turn modifies Problem 1 into

Problem 2 *Find an epistemically uncertain random (aleatoric) variable $Q_M(\omega_Q)$ in Eq. (11)—a random variable defined on the space $\Omega_u \times \Omega_Q$—such that the re-interpreted predictions of Eq. (10) match those of Eq. (14) in a measurement sense.*

The designation of epistemic and aleatoric uncertainty is a kind of interpretation, mathematically Ω_u and Ω_Q are just probability spaces. Once such an object like Q is identified in a Bayesian framework, it does usually not really matter what caused the uncertainty described in a probabilistic sense. Thus, for the sake of simplicity, in the following we shall take on the macro scale just *one* probability space $(\Omega, \mathfrak{B}, \mathbb{P})$ for the description of the uncertainty, where one might think of $\omega \in \Omega$ as a realisation $\omega = (\omega_u, \omega_Q) \in \Omega_u \times \Omega_Q = \Omega$, i.e. the product of those two probability spaces.

Bayesian upscaling of random meso-structures

The goal is to identify the quantity Q defined on Ω conditioned on the observation y_m using Bayes's rule. As some or all of the components of Q may be required to be positive definite—as is often the case for material quantities—, this constraint has to be taken into consideration. In our case all components of Q have to fulfil that requirement. In most updating methods it is advantageous if the quantities to be identified have *no* constraints. We shall explain how to achieve this by considering a scalar component Q of Q. The first is to scale Q by a reference Q_0 to obtain a dimensionless quantity, and consider now $q := \log(Q/Q_0)$. As numerically $\log(Q/Q_0) = \log Q - \log Q_0$, it is convenient to choose Q_0 such that $Q_0 = 1$ in the units used, and hence numerically $\log Q_0 = 0$ and $q := \log Q$. Henceforth we assume that this has been done. The variable q now has no constraints, it is a *free* variable on all of $\mathbb{R}$. This procedure may be extended to all even order positive tensors, but will only be needed for scalars here. So instead of identifying the collection Q directly, we identify the logarithms of its components giving in this way a collection q where we write symbolically $q = \log Q$, as in this way whatever approximations or linear operations are performed computationally on the numerical representation of $q(x, \omega)$, in the end $\exp(q(x, \omega))$ is always going to be positive. This also gives the right kind of mean—the geometric mean—for positive quantities. The underlying reason is that

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 9 of 35

the multiplicative group of positive real numbers—a (commutative) one-dimensional Lie group—is thereby put into correspondence with the additive group of reals, which also represents the (one-dimensional) tangent vector space at the group unit, the number one. This is the corresponding Lie algebra. A positive quadratic form on the Lie algebra—in one dimension necessarily proportional to Euclidean distance squared—can thereby be carried to a Riemannian metric on the Lie group. A similar argument holds for positive tensors of any even order.

Therefore, instead of Problem 2, we consider its modified version:

Problem 3 *Find a random variable $q := \log Q : \Omega \to \mathcal{Q}$ for $Q = \exp q$ in Eq.* (11)*, such that the re-interpreted predictions of Eq.* (10) *match those of Eq.* (14) *in a measurement sense.*

Bayesian updating is in essence a probabilistic conditioning, the foundation of which is the conditional expectation operator [4]. Here we are interested in the case where the conditioning occurs w.r.t. another random variable, namely $y(q(\omega))$, which depends on the quantity q to be updated. For any function $\varphi : \mathcal{Q} \to \mathcal{F}$ of finite variance of q, the conditional expectation of it is defined [4] by the projection onto a closed subspace $\mathcal{C}_\varphi \subset L_2$, which is in simple terms the L_2-closure of all multivariate polynomials in the components of y with coefficients from the vector space $\mathcal{F}$, i.e.

$$\mathcal{C}_\varphi := \mathrm{cl}\ \{p(y(q)) \mid p(y(q)) = \sum_\alpha f_\alpha V_\alpha(y(q))\}. \tag{15}$$

where $f_\alpha \in \mathcal{F}$ and the V_α are real-valued multivariate polynomials in $y = (y_1, \ldots, y_j, \ldots)$, which means that for a multi-index $\alpha = (\alpha_1, \alpha_2, \ldots)$ the polynomial V_α is of degree α_j in the variable y_j. It turns out [4] that $\mathcal{C}_\varphi$ contains all measurable functions $g : \mathcal{Y} \to \mathcal{F}$ so that $g(y(q(\omega)))$ is of finite variance.

Here we will be only interested in the function $\varphi(q) = q$, i.e. the conditional mean of q. To compute it, one may use the variational characterisation and compute the minimal distance from the subspace $\mathcal{C} := \mathcal{C}_q$ to the point q:

$$\phi(y(q)) := \mathbb{E}\,(q \mid y) := P_{\mathcal{C}}q = \underset{\hat{\phi} \in \mathcal{C}}{\arg\min}\ \mathbb{E}\left(\left\|q - \hat{\phi}(y(q))\right\|^2\right). \tag{16}$$

In this section, the expectation operators are to be understood as acting only on the variables which describe the uncertainty in the estimation, i.e. in the notation of "Abstract model problem" section only on the variables from Ω_u. One may observe from Eq. (16) that ϕ is the best "inverse" of $y(q)$ in a least square sense, the orthogonal projection $P_{\mathcal{C}}q$ of q onto $\mathcal{C}$. Following Eq. (16), one may decompose the random variable q into the projected component $q_p \in \mathcal{C}$ and the orthogonal residual $q_r \in \mathcal{C}^\perp$, such that

$$q = q_p + q_r = P_{\mathcal{C}}q + (I - P_{\mathcal{C}})q = \phi(y(q)) + (q - \phi(y(q))) \tag{17}$$

holds. Here, $q_p = P_{\mathcal{C}}q = \mathbb{E}\,(q \mid y) = \phi(y(q))$ is the orthogonal projection onto the subspace $\mathcal{C}$ of all random variables consistent with the data, whereas $q_r := (I - P_{\sigma(y)})q$ is its orthogonal residual.

This can be used to build a filter—filtering the observation y_m together with the prior forecast q_f—which is optimal in this least square sense [38,39] and returns an assimilated

random variable q_a which has the correct conditional expectation. The first term in the sum in Eq. (17) is taken to be the conditional expectation given the observation y_m, i.e. $\phi(y_m) = \mathbb{E}\left(q_f \mid y_m\right)$, whereas $(q_f - \phi(y(q_f)))$ is the residual component. Following this, Eq. (17) can be recast to obtain the update q_a for the prior random variable q_f as

$$q_a = q_f + \phi(y_m) - \phi(y_M(q_f)), \tag{18}$$

in which $y_M(q_f)$ (see Eq. (10)) is the random variable representing our prior prediction/forecast of the measurement data, and q_a is the assimilated random variable. By recalling $\phi(y(q_f)) = P_C q_f$, one sees immediately that $\mathbb{E}(q_a \mid y_m) = \mathbb{E}\left(q_f \mid y_m\right)$, and the assimilated random variable q_a has the correct conditional expectation. As in engineering practice one is often not interested in estimating the full posterior measure, and the conditional expectation is the most important characterisation, we will use this computationally simpler procedure.

Therefore, to estimate q_a one requires only information on the map ϕ in Eq. (18). To make the determination of this map computationally feasible, and for the sake of simplicity, the map ϕ in Eq. (18) can be approximated by a n-th order polynomial—i.e. the minimisation in Eq. (16) is not over *all* measurable maps, but only over n-th order polynomials—such that the map ϕ in Eq. (18) becomes

$$\phi_n(y; \beta) = \sum_\alpha K^{(\alpha)} V_\alpha(y) \tag{19}$$

with characterising coefficients $\beta = \{K^{(\alpha)}\}_\alpha, K^{(\alpha)} \in \mathcal{Q}$, multi-indices $\alpha := (\alpha_1, \dots)$ with $\forall j : 0 \leq \alpha_j \leq n$, and multivariate polynomials V_α as before in Eq. (15). In the affine case, when $n = 1$ and $\phi_1(y; \beta) = Ky + b$ in the previous formula Eq. (19), the Eq. (18) reduces to the Gauss-Markov-Kalman filter [38,39]:

$$q_a = q_f + K(y_m - y_M(q_f)), \tag{20}$$

a generalisation of the well known Kalman filter.

In order to estimate the macro-scale properties using Eq. (18), one requires both y_m and y_M, preferably in the functional approximation form. Note that y_M is the prediction of the measurement data on the macro-scale level, and is obtained by propagating the prior knowledge q_f (here a spatially homogeneous quantity) through the macro-scale model. In this paper we use Bayesian regression —not related to the Bayesian updating—to estimate the functional approximation of $y_M(q_f)$, as will be presented in "Approximating the macro-scale response by Bayesian regression" section. On the other hand, y_m represents the response of the high-dimensional meso-scale model in which meso-scale properties q_m are heterogeneous and uncertain. By modelling q_m one may estimate y_m in a similar manner as y_M, see the first case scenario in Fig. 1. However, due to the high-dimensionality of the input q_m, and the nonlinearity of the meso-scale model, a straightforward uncertainty quantification of y_m is often not computationally affordable. The estimate would require too many data points as thousands of input parameters can easily be involved in the description of the meso-scale properties. Therefore, we use an unsupervised learning algorithm to reduce the stochastic dimension of the meso-scale measurement, as further described in "Approximation of the meso-scale observation by unsupervised learning" section.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 11 of 35

Approximating the macro-scale response by Bayesian regression

The measurement prediction y_M is approximated by a surrogate model

$$\hat{y}_M = \phi_M(q_f; \beta) \tag{21}$$

in which ϕ_M is usually taken to be nonlinear map of $q_f \in L_2(\Omega_u, \mathfrak{B}_u, \mathbb{P}_u)$ with the coefficients β. Furthermore, we assume that y_M is in general only known as a set of samples, and our goal is to match $\hat{y}_M$ with y_M. Let $x := (q_f(\omega_i), y_M(\omega_i))_{i=1}^N$ be the full set of N data samples describing the forward propagation of q_f to y_M via macro-scale model. In order to specify $\hat{y}_M$ the only thing we need to find are the coefficients of the map ϕ_f. Therefore, we infer β given data x using Bayes's rule

$$p(\beta|x) = \frac{p(x, \beta)}{\int p(x, \beta)\, \mathrm{d}\beta} \tag{22}$$

In general case the marginalisation in Eq. (22) can be expensive, and therefore in this paper we use the variational Bayesian inference instead [31]. The idea is to introduce a family $\mathcal{D} := \{g(\beta) := g(\beta|\lambda, w)\}$ over β indexed by a set of free parameters (w, λ) such that $\hat{y}_M \sim y_M$. Thus, the idea is to optimise the parameter values by minimising the Kullback-Leibler divergence

$$g^*(\beta) = \underset{g(\beta)\in\mathcal{D}}{\arg\min}\, D_{KL}(g(\beta)||p(\beta|x)) = \underset{g(\beta)\in\mathcal{D}}{\arg\min} \int g(\beta)\, \log \frac{g(\beta)}{p(\beta|x)}\, \mathrm{d}\beta. \tag{23}$$

After few derivation steps as depicted in [31], the previous minimisation problem reduces to

$$\beta^* = \arg\max \mathcal{L}(g(\beta)) := \mathbb{E}_g(\log p(x, \beta)) - \mathbb{E}_g(\log g(\beta)) \tag{24}$$

in which $\mathcal{L}(g)$ is the evidence lower bound (ELBO), or variational free energy. To obtain a closed-form solution for β^*, the usual practice is to assume that both the posterior $p(\beta|x)$ as well as its approximation $g(\beta)$ can be factorised in a mean sense, i.e.

$$p(\beta|x) = \prod p_i(\beta_i|x), \quad g(\beta) = \prod g_i(\beta_i) \tag{25}$$

in which each factor $p_i(\beta_i|x), g_i(\beta_i)$ is independent and belongs to an exponential family. Similarly, their complete conditionals given all other variables and observations are also assumed to belong to exponential families, and are assumed to be independent. Obviously these assumptions lead to conjugacy relationships, and closed form solution of Eq. (24) as further discussed in more detail in [31].

To approximate $y_M(\omega)$, we take Eq. (21) to be described in a form of a polynomial chaos expansion (PCE) or generalised PCE (gPCE) [67]. In other words, $y_M(\omega)$ and $q_f(\omega)$ are taken to be functions of known RVs $\{\theta_1(\omega), \dots, \theta_n(\omega), \dots\}$. Often, when for example stochastic processes or random fields are involved, one has to deal here with infinitely many RVs, which for an actual computation have to be truncated to a finite vector $\boldsymbol{\theta}(\omega) = [\theta_1(\omega), \dots, \theta_L(\omega)] \in \Theta \cong \mathbb{R}^L$ of significant RVs. We shall assume that these have been chosen such as to be independent, and often even normalised Gaussian and independent. The reason to not use q_f directly is that in the process of identification of q they

may turn out to be correlated, whereas $\boldsymbol{\theta}$ can stay independent as they are. Thus a RV $\boldsymbol{y}_M(\boldsymbol{\theta})$ is replaced by a functional approximation

$$\boldsymbol{y}_M(\boldsymbol{\theta}) = \sum_{\alpha \in \mathcal{J}_Z} \boldsymbol{y}_M^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}), \tag{26}$$

and analogously $\boldsymbol{q}_f$ by

$$\boldsymbol{q}_f(\boldsymbol{\theta}) = \sum_{\alpha \in \mathcal{J}_Z} \boldsymbol{q}_f^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) \tag{27}$$

in which the multi-index $\alpha = (\ldots, \alpha_k, \ldots)$, and the set $\mathcal{J}_Z$ of multi-indices is a finite set with cardinality (size) Z.

The coefficients $\boldsymbol{q}_f^{(\alpha)}, \boldsymbol{y}_M^{(\alpha)}$, e.g. $\boldsymbol{v} := \{\boldsymbol{y}_M^{(\alpha)}\}_{\alpha \in \mathcal{J}_Z}$, are estimated by minimising the ELBO analogous to the one in Eq. (24) by using the variational relevance vector machine method [3]. Namely, the measurement forecast $\boldsymbol{y}_s := \{\boldsymbol{y}_M(\boldsymbol{\theta}_j)\}_{j=1}^N$ can be rewritten in a vector form as

$$\boldsymbol{y}_s = \boldsymbol{v}\boldsymbol{\Psi} \tag{28}$$

in which $\boldsymbol{\Psi}$ is the matrix of collection of basis functions $\Psi_\alpha(\boldsymbol{\theta})$ evaluated at the set of sample points $\{\boldsymbol{\theta}_j\}_{i=1}^N$. However, the expression in the previous equation is not complete, as the PCE in Eq. (26) is truncated. This implies the presence of the modelling errors. Under a Gaussian assumption, the data then can be modelled as

$$p(\boldsymbol{y}_s) \sim \mathcal{N}(\boldsymbol{v}\boldsymbol{\Psi}, \varsigma^{-1}\boldsymbol{I}) \tag{29}$$

in which $\varsigma \sim \Gamma(a_\varsigma, b_\varsigma)$ denotes the imprecision parameter, here assumed to follow Gamma distribution. The coefficients $\boldsymbol{v}$ are given a normal distribution under the independency assumption:

$$p(\boldsymbol{v}|\boldsymbol{a}) \sim \prod_{i=0}^Z \mathcal{N}(0, \zeta_i^{-1}) \tag{30}$$

in which Z denotes the cardinality of the PCE, and $\boldsymbol{\zeta} := \{\zeta_i\}$ is a vector of hyper-parameters. To promote for sparsity, the vector of hyper-parameters is further assumed to follow Gamma distribution

$$p(\zeta_i) \sim \Gamma(a_i, b_i) \tag{31}$$

under the independency assumption. In this manner the posterior for $\boldsymbol{\beta} := \{\boldsymbol{v}, \boldsymbol{\zeta}, \varsigma\}$, i.e. $p(\boldsymbol{\beta}|\boldsymbol{y}_s)$, can be approximated by a variational mean field form

$$g(\boldsymbol{\beta}) = g_v(\boldsymbol{v})g_\zeta(\boldsymbol{\zeta})g_\varsigma(\varsigma), \tag{32}$$

the factors of which are chosen to take same distribution type as the corresponding prior due to the conjugacy reasons. Once this assumption is made, one may maximise the corresponding ELBO in order to estimate the parameter set.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 13 of 35

Finally, we have everything to describe the macro-scale response y_M, and therefore we may fill this term in the filtering equation as presented in Eq. (18) to obtain:

$$q_a = \sum_{\alpha \in \mathcal{J}} q_f^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) + \phi_n(y_m) - \phi_n \left(\sum_{\alpha \in \mathcal{J}} y_M^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) \right) \tag{33}$$

Note that Eq. (33) is not yet fully computationally operational, the RV y_m has to be put into a computationally accessible form. This is considered in the following "Approximation of the meso-scale observation by unsupervised learning" section.

Approximation of the meso-scale observation by unsupervised learning

Let the measurement y_m be approximated by

$$y_m = \phi_m(w, \eta) \tag{34}$$

in which ϕ_m is an analytical function (e.g. a Gaussian mixture model, a neural network, etc.) parameterised by global variables/parameters w describing the whole data set, and the latent local/hidden variables η that describe each data point. An example is the generalised mixture model in which parameters w include statistics of individual components, and the mixture weights, whereas the hidden variable η stands for the indicator variable that describes the membership of data points to the mixture components. The goal is to estimate the pair $\beta := (w, \eta)$ given data $y_d := \{y_m(\hat{\omega}_i)\}, i = 1, .., M$ with the help of Bayes's rule. Note that we do not take full set of the input–output data $(q_m(\hat{\omega}_i), y_m(\hat{\omega}_i))$ with q_m defined on $(\Omega_Q, \mathfrak{B}_Q, \mathbb{P}_Q)$, but only its incomplete version generated only by the output $y_d := \{y_m(\hat{\omega}_i)\}, i = 1, .., M$. Following this, the coefficients of y_m can be estimated as

$$p(\beta|y_d) = \frac{p(y_d, \beta)}{\int p(y_d, \beta) \, d\beta}. \tag{35}$$

The previous equation is more general than Eq. (25), and hence includes the problem described in "Approximating the macro-scale response by Bayesian regression" section as a special case. The main reason is that next to the coefficients w we also need to estimate the argument η such that the functional approximation in Eq. (34) is minimally parametrised.

Following theory in "Approximating the macro-scale response by Bayesian regression" section, Eq. (35) is reformulated to the computationally simpler variational inference problem. In other words, we introduce a family of density functions $\mathcal{D} := \{g(\beta) := g(\beta|\lambda, \varpi)\}$ over β indexed by a set of free parameters (ϖ, λ) that approximate the posterior density $p(\beta|y_d)$, and further optimise the variational parameter values by minimising the Kullback–Leibler divergence between the approximation $g(\beta)$ and the exact posterior $p(\beta|y_d)$. Hence, following Eq. (24), we maximise the ELBO

$$\mathcal{L}(g) = \mathbb{E}_{g(\beta)}(\log p(y_d, \beta)) - \mathbb{E}_{g(\beta)}(\log g(\beta)))) \tag{36}$$

by using the mean-field factorisation assumption, and conjugacy relationships. The optimisation problem attains a closed form solution in which the lower bound is iteratively optimised with respect to the global parameters keeping the local parameters fixed, and in the second step the local parameters are updated and the global parameters are held

fixed. The algorithm can be improved by considering the stochastic optimisation in which a noisy estimate of the gradient is used instead of the natural one.

The mean field factorisation as presented previously is computationally simple, but not accurate. For example, one cannot assume independence between the stored energy and dissipation coming from the same experiment. In other words, the correlation among the latent variables is not explored. As a result, the covariance of the measurement will be underestimated. To allow dependence in the factorisation, one may extend the mean-field approach via copula factorisations [63,64]:

$$g(\beta) = c(F_1(\beta_1), ..., F_m(\beta_m), \chi) \prod_{i=1}^{m} g_i(\beta_i) \tag{37}$$

in which $c(F_1(\beta_1), ..., F_m(\beta_m), \chi)$ is the representative of a copula family, $F_i(\beta_i)$ is the marginal cumulative distribution function of the random variable β_i, and χ is the set of parameters describing the copula family. Similarly, $g_i(\beta_i)$ represent the independent marginal densities. In this manner any distribution type can be represented by a formulation as given in Eq. (37) according to Sklar's theorem [56].

Following Eq. (37), the goal is to find $g(\beta)$ such that the Kullback-Leibler divergence to the exact posterior distribution is minimised. Note that if the true posterior is described by

$$p(\beta|y_d) = c_t(F_1(\beta_1), ..., F_m(\beta_m), \chi_t) \prod_{i=1}^{m} f_i(\beta_i), \tag{38}$$

then the Kullback-Leibler divergence reads:

$$D_{KL}(g(\beta)||p(\beta|y_m) = D_{KL}(c||c_t) + \sum_{i=1}^{m} D_{KL}(g_i(\beta_i)||f_i(\beta_i)), \tag{39}$$

and contains one additional term compared to the mean field approximation. When the copula is uniform, the previous equation reduces to the mean field one, and hence only the second term is minimised. On the other hand, if the mean field factorisation is not a good assumption and the dependence relations are neglected, then the total approximation error will be dominated by the first term. To avoid this, the ELBO derived in Eq. (36) modifies to

$$\mathcal{L}(g) = \mathbb{E}_{g(\beta)}(\log p(y_m, \beta)) - \log g(\beta, \chi) \tag{40}$$

and is a function of parameters of the latent variables β, as well as of the copula parameters χ. Therefore, the algorithm applied here consists of iteratively finding the parameters of the mean field approximation, as well as those of the copula. The algorithm is adopted from [63], and is a black-box algorithm as it only depends on the likelihood $p(y_m, \beta)$ and copula description in a vine form. Note that when the copula is equal to identity, i.e. uniform, the previous factorisation collapses to the mean field one.

Once the copula dependence structure is found, the measurement data y_m are represented in a functional form—here taken as generalised mixture model—as in Eq. (34), which is different than the polynomial chaos representation. In other words, the measurement is given in terms of dependent random variables, and not independent ones.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 15 of 35

Therefore, the dependence structure has to be mapped to an independent one. In a Gaussian copula case, the Nataf transformation can be used, and otherwise the Rosenblatt transformation is applied. For high-dimensional copulas, such as a regular vine copula, [1] provides algorithms to compute the Rosenblatt transform and its inverse. The result of the transformation are mutually independent and marginally uniformly distributed random variables, which further can be mapped to Gaussian ones or other types of standard random variables via marginals [62]. Let the functional approximation of the measurement be given as

$$
\boldsymbol{y}_m(\boldsymbol{\xi}) \approx \sum_{\alpha \in \mathcal{J}_m} \boldsymbol{y}_m^{(\alpha)} G_\alpha(\boldsymbol{\xi}) \tag{41}
$$

in which $\mathcal{J}_m$ is a multi-index set, and $G_\alpha(\boldsymbol{\xi})$ is a set of functions (e.g. orthogonal polynomials) with random variables $\boldsymbol{\xi}$ as arguments. With this, we have obtained the measurement $\boldsymbol{y}_m$ in a minimised functional approximation form, which further can be plugged into Eq. (33) to obtain the final filter discretisation. By combining the random variables $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$, one may re-write Eq. (33) in the following form

$$
\sum_{\alpha \in \mathcal{J}_a} \boldsymbol{q}_a^{(\alpha)} H_\alpha(\boldsymbol{\theta}, \boldsymbol{\xi}) = \sum_{\alpha \in \mathcal{J}} \boldsymbol{q}_f^{(\alpha)} H_\alpha(\boldsymbol{\theta}, \boldsymbol{\xi}) + \phi_n \left(\sum_{\alpha \in \mathcal{J}_m} \boldsymbol{y}_m^{(\alpha)} H_\alpha(\boldsymbol{\theta}, \boldsymbol{\xi}) \right) - \phi_n \left(\sum_{\alpha \in \mathcal{J}} \boldsymbol{y}_M^{(\alpha)} H_\alpha(\boldsymbol{\theta}, \boldsymbol{\xi}) \right) \tag{42}
$$

in which H_α is a generalised polynomial chaos expansion with random variables $(\boldsymbol{\theta}, \boldsymbol{\xi})$ as arguments. Note that the coefficients $\boldsymbol{q}_f^{(\alpha)}$, as well as $\boldsymbol{y}_M^{(\alpha)}$ and $\boldsymbol{y}_m^{(\alpha)}$, are sparse as they only depend on $\boldsymbol{\theta}$ or $\boldsymbol{\xi}$, respectively. As $\boldsymbol{\theta}$ describes the a priori (epistemic) uncertainty, one may take the mathematical expectation of the previous equation w.r.t. $\boldsymbol{\theta}$ to obtain the natural (aleatoric) variability of the macro-scale parameters:

$$
\sum_{\alpha \in \mathcal{J}_m} \boldsymbol{q}_a^{(\alpha)} G_\alpha(\boldsymbol{\xi}) = \mathbb{E}_{\boldsymbol{\theta}} \left(\sum_{\alpha \in \mathcal{J}_a} \boldsymbol{q}_a^{(\alpha)} H_\alpha(\boldsymbol{\theta}, \boldsymbol{\xi}) \right)
$$
$$
= \phi_n \left(\sum_{\alpha \in \mathcal{J}_m} \boldsymbol{y}_m^{(\alpha)} \Gamma_\alpha(\boldsymbol{\xi}) \right) + \mathbb{E}_{\boldsymbol{\theta}} \left(\sum_{\alpha \in \mathcal{J}} \boldsymbol{q}_f^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) - \phi_n \left(\sum_{\alpha \in \mathcal{J}} \boldsymbol{y}_M^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) \right) \right). \tag{43}
$$

In general, the approximation of the meso-scale information as previously described can be cubersome due to high nonlinearity and time-dependence of $\boldsymbol{y}_m$. Therefore, instead of approximating $\boldsymbol{y}_m$ in a form as in Eq. (34), one may discretise $\boldsymbol{y}_m$ in a Monte Carlo sampling manner such that Eq. (33) rewrites to $\forall \hat{\omega}_i : i = 1, \ldots, M$

$$
\boldsymbol{q}_a^{(i)}(\boldsymbol{\theta}) := \sum_{\alpha \in \mathcal{J}} \boldsymbol{q}_i^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) = \sum_{\alpha \in \mathcal{J}} \boldsymbol{q}_f^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) + \phi_n(\boldsymbol{y}_m(\hat{\omega}_i)) - \phi_n \left(\sum_{\alpha \in \mathcal{J}} \boldsymbol{y}_M^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}) \right). \tag{44}
$$

In other words we repeat the update formula M times for each instance of the measurement $\boldsymbol{y}_m$, and thus obtain M posteriors $\boldsymbol{q}_a^{(i)}$, $i = 1, \ldots, M$ that depend only on the epistemic uncertainty embodied in $\boldsymbol{\theta}$. By averaging over $\boldsymbol{\theta}$ one obtains a set of samples:

$$
\forall \omega_i : \bar{\boldsymbol{q}}_i = \mathbb{E}_{\boldsymbol{\theta}}(\boldsymbol{q}_a^{(i)}(\boldsymbol{\theta})), \quad i = 1, \ldots, M, \tag{45}
$$

i.e. the data which are to be used for the estimation of the functional approximation form of the macro-scale parameter q_M similar to Eq. (34). To achieve this, we search for an approximation
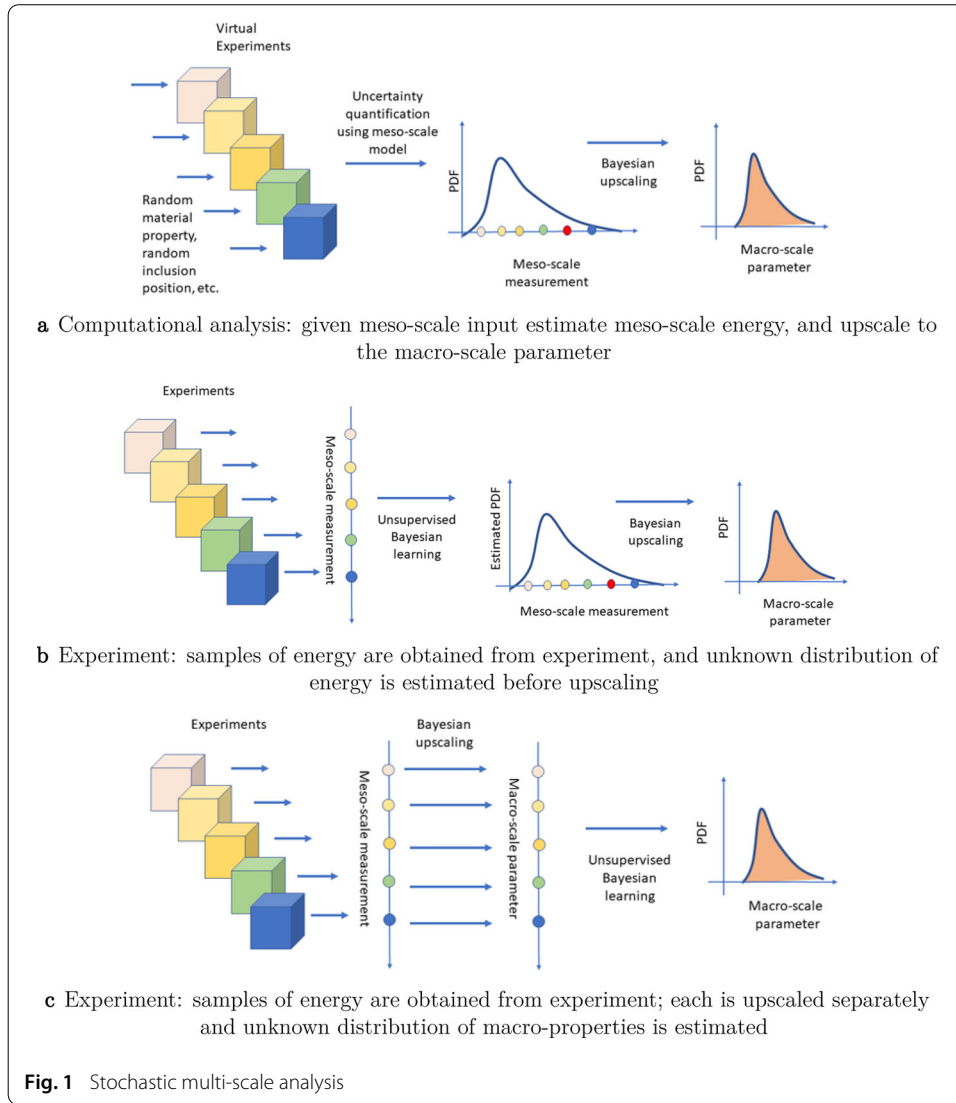
$$q_M = \varphi_q(\boldsymbol{w}_q, \eta_q) \tag{46}$$

given the incomplete data set $q_d := (\bar{\boldsymbol{q}}_i)_{i=1}^n$. Here, w_q and η_q have same meaning as in Eq. (34), and therefore can be estimated by using same unsupervised algorithm as previously described. This approach is computationally more convenient, as the correlation structure between the material parameters is easier to learn than the one between measurement data on the meso-scale.

For better clarity, we re-capitulate the upscaling procedure in Fig. 1 for comparison reasons. In Fig. 1a) is shown the direct computational approach in which Eq. (42) is used with y_m being approximated in same manner as y_M by supervised Bayesian regression described in "Approximating the macro-scale response by Bayesian regression" section. Due to a high computational footprint, this approach is not considered in this paper, for more information please see [47]. The upscaling approach presented in Eq. (42), in which y_m is approximated by Eq. (34) via an unsupervised learning algorithm, is further depicted in Fig. 1b. Here one first uses the Bayesian unsupervised learning algorithm to learn the distribution of the meso-scale measurement, and later a Bayesian upscaling procedure to estimate the macro-scale parameters. Finally, the upscaling approach given by Eqs. (44) and (46) is shown in Fig. 1c. In this approach one first uses a Bayesian upscaling procedure and estimates the macro-scale parameter sample-wise, after which the Bayesian unsupervised learning algorithm is used to approximate the distribution of the macro-scale parameters. The choice of algorithm depends on the application and dimensionality, as well as on the nonlinearity of the meso-scale model.

Bayesian upscaling via energy considerations

As an example of the abstract model in "Abstract model problem" section in Eqs. (2) and (3), we choose a prototypical version of the compressive behaviour of a cementitious-like material, one which displays in its irreversible behaviour both a "softening" component, as well as an "hardening" component, and therefore interaction with the Helmholtz free energy, to test the merits of the identification algorithm on such behaviour. It is a simple version of a coupled elasto-damage model introduced in [24] (Sect. 3.5.2), in which the state variable $z = (u, w)$ from Eq. (6) is locally at some point $\boldsymbol{x} \in \mathcal{G}$ in the body $z(\boldsymbol{x}) = (\boldsymbol{u}(\boldsymbol{x}), \boldsymbol{w}(\boldsymbol{x})) \in \mathcal{Z}_M = \mathcal{U}_M \times \mathcal{W}_M$, giving the local resulting strain $\boldsymbol{\varepsilon}(\boldsymbol{x}) = \nabla^s \boldsymbol{u}(\boldsymbol{x})$, as well as the internal variables $\boldsymbol{w} = (D, \varsigma)$. Here $D \in [1, \infty[$ is a damage variable which will modify the stiffness through the elasticity tensor, and ς is a scalar hardening variable. Observe that often $\tilde{D} = (1 - 1/D)$ is regarded as the "real damage", as $\tilde{D} = 0$ corresponds to virgin or undamaged material, and $\tilde{D} = 1$ corresponds to totally damaged material. For the elastic part, we choose to identify an isotropic material, as the meso-model description is stochastically isotropic. Splitting the strain in its volumetric $\mathrm{vol}\,\boldsymbol{\varepsilon} = (\mathrm{tr}\,\boldsymbol{\varepsilon}/3)\boldsymbol{I}$ and deviatoric part $\mathrm{dev}\,\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon} - \mathrm{vol}\,\boldsymbol{\varepsilon}$, the elastic relations may be written in the isotropic case (e.g. [24]) as

$$\boldsymbol{\sigma} = \frac{3\kappa}{D}\,\mathrm{vol}\,\boldsymbol{\varepsilon} + 2\mu\,\mathrm{dev}\,\boldsymbol{\varepsilon} = \boldsymbol{C}(D) : \boldsymbol{\varepsilon},$$

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 17 of 35

Fig. 1 Stochastic multi-scale analysis

where κ is the bulk modulus and μ the shear modulus, in total describing a damage depended elasticity tensor $\mathbf{C}(D)$, the damage only acting on the bulk response.

For undamaged material the damage variable is $D = 1$, and as it grows $D \to \infty$, the bulk response weakens.

The local stored energy is given by

$$W(\boldsymbol{\varepsilon}, \boldsymbol{w}, \kappa) = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbf{C}(D) : \boldsymbol{\varepsilon} + \frac{1}{2}\varsigma K_d \varsigma = \frac{3\kappa}{2D}(\text{vol}\,\boldsymbol{\varepsilon} : \text{vol}\,\boldsymbol{\varepsilon}) + \mu(\text{dev}\,\boldsymbol{\varepsilon} : \text{dev}\,\boldsymbol{\varepsilon}) + \frac{K_d}{2}\varsigma^2,$$

(47)

with the components of the characterising parameter vector $\boldsymbol{Q} = (\kappa, \mu, \sigma_f, K_d)$, where K_d is a hardening modulus for the scalar internal variable ς for hardening, and σ_f is a failure stress, used in the failure criterion, where we choose one of *crushing damage*—as it may occur e.g. for cementitious material [24] (Sect. 3.5.2)—with failure function

$$f_d(\text{tr}\,\boldsymbol{\sigma}, \chi_d) = \langle -\text{tr}\,\boldsymbol{\sigma}\rangle - (\sigma_f - \chi_d),$$

(48)

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 18 of 35

where $\langle x \rangle = 1/2(x + |x|)$ is the Macauley bracket, and $\chi_d = -K_d\varsigma$, the thermodynamic force corresponding to hardening. The elastic domain is $f_d(\mathrm{tr}\,\boldsymbol{\sigma}, \chi_d) \leq 0$, i.e. damage occurs when the pressure satisfies $p = -\mathrm{tr}\,\boldsymbol{\sigma} \geq \sigma_f + K_d\varsigma$. The internal variables are $\boldsymbol{w} = (D, \varsigma)$. The Legendre-Fenchel dual F^* of the dissipation pseudo-potential $F(\boldsymbol{\varepsilon}, \boldsymbol{w}, \dot{\boldsymbol{w}}, \boldsymbol{Q})$ — the indicator function of the elastic domain—is then

$$F^*(\boldsymbol{\varepsilon}, \boldsymbol{w}, \boldsymbol{\zeta}, \boldsymbol{Q}) = \begin{cases} 0 & \text{if } f_d(\mathrm{tr}\,\boldsymbol{\sigma}, \chi_d) \leq 0, \\ +\infty & \text{otherwise}. \end{cases}$$

The thermodynamic forces are

$$\boldsymbol{\zeta} = \mathrm{D}_{\boldsymbol{w}} W = \left(\frac{\kappa}{2D^2}(\mathrm{tr}\,\boldsymbol{\varepsilon})^2, K_d\varsigma \right) = \left(\frac{1}{18\kappa}(\mathrm{tr}\,\boldsymbol{\sigma})^2, -\chi_d \right),$$

so that the instantaneous dissipation density becomes $\eta = \dot{D}(\mathrm{tr}\,\boldsymbol{\sigma})^2/(18\kappa) + \dot{\varsigma}K_d\varsigma$.

The measurement or observation prediction on the macro-scale y_M will be only registered at certain intervals Δt of the pseudo-time variable t, $t_0 = 0, t_1 = \Delta t, \ldots, t_\ell = \ell\Delta t, \ldots$. The observation prediction is specified by the spatial average of the energy-type prediction of the macro model $y_M = (\mathcal{E}_e, \mathcal{E}_h, \mathcal{E}_d)$ at observation time t_ℓ, given by

$$\mathcal{E}_e = \frac{1}{2} \int_{\mathcal{G}} \boldsymbol{\varepsilon}(\boldsymbol{x}, t_\ell) : \mathbf{C}(D)(\boldsymbol{x}) : \boldsymbol{\varepsilon}(\boldsymbol{x}, t_\ell) \, \mathrm{d}\boldsymbol{x},$$
$$\mathcal{E}_h = \frac{1}{2} \int_{\mathcal{G}} K_d(\boldsymbol{x})\varsigma(\boldsymbol{x}, t_\ell)^2 \, \mathrm{d}\boldsymbol{x},$$
$$\mathcal{E}_d = \int_{t_{\ell-1}}^{t_\ell} \int_{\mathcal{G}} \eta(\boldsymbol{x}, t) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}t, \tag{49}$$

i.e. the integrated or averaged stored elastic and hardening energy, and the energy dissipated between the last observation at $t = t_{\ell-1}$ and now at $t = t_\ell$.

In the numerical experiments the previous model is used on both the meso- and macro-scale. Finally, the upscaling is considered for the energy-type of measurement from the meso-scale $y_m = (\mathcal{E}_{me}, \mathcal{E}_{mh}, \mathcal{E}_{md})$, and in our case defined exactly in an analogous way as in Eq. (49), using the meso-scale model respectively. On the macro-scale, the integrand quantities $\mathbf{C}(D)(\boldsymbol{x})$ and $K_d(\boldsymbol{x})$ in Eq. (49)—in fact all components of $\boldsymbol{Q}$—are assumed spatially homogeneous or constant, whereas on the heterogeneous meso-scale model they do vary spatially and are modelled as random fields. In this manner, the stored as well as the dissipated portion of the total energy is faithfully mapped from the meso- to the macro-scale model.

To represent the measurement data y_m for all meso-structure realisations, one may use generalised mixture models. In our particular application the measurement is positive. Therefore, we use samples $x := (\log y_m(\omega_i))_{i=1}^N$ to approximate $\log y_m$ as a Gaussian mixture model [2]

$$p(x) = \sum_{k=1}^{K} \pi_k \, \mathcal{N}(x|\nu_k), \quad \sum_{k=1}^{K} \pi_k = 1, \quad 0 \leq \pi_k \leq 1 \tag{50}$$

described by parameter set $\boldsymbol{\nu} = (\mu_k, \Sigma_k)_{k=1}^K$ with $\nu_k := (\mu_k, \Sigma_k)$ being the statistics parameters of Gaussian components, and π_k are the mixing coefficients. These constitute the parameter vector w. The hidden variable η is the indicator vector z_k of dimension N that

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.* (2020) 7:50

Page 19 of 35

describes the membership of each data point to the Gaussian cluster. Following this, the joint distribution is given as

$$p(x, Z, \mu, \Sigma, \pi) = p(x|z, \mu, \Sigma, \pi)\, p(z|\pi)\, p(\pi)\, p(\mu|\Sigma)\, p(\Sigma) \tag{51}$$

in which

$$p(z|\pi) = \prod_{n=1}^{N} \prod_{k=1}^{K} \pi_k^{z_{nk}}, \quad p(z_k = 1) = \pi_k \tag{52}$$

and

$$p(x|z, \mu, \Sigma, \pi) = \prod_{n=1}^{N} \prod_{k=1}^{K} \mathcal{N}(x_n|\mu_K, \Sigma_K^{-1})^{z_{nk}}. \tag{53}$$

The priors are chosen such that $p(\pi)$ is a Dirichlet prior, whereas $p(\mu, \Sigma)$ follows an independent Gaussian-Wishart prior governing the mean and precision of each Gaussian component. Hence, our parameter set β is described by a set of global parameters $w := (\mu, \Sigma, \pi)$ and the hidden variable z. To incorporate correlations, the copula dependence structure of Gaussian mixture as in Eq. (37) is found, and the measurement data are represented in a functional form, as discussed in "Bayesian upscaling of random meso-structures" section.

Numerical results

Bayesian upscaling of a heterogeneous linear elastic material model

In this section the proposed upscaling scheme is first applied on a linear elastic random heterogeneous material. Here it is known from homogenisation theory that for a large enough representative volume element (RVE) one may indeed find a *spatially homogeneous* constant value for the elastic parameters. This is also quickly seen from a general consideration: for a homogeneous strain, on the macro-scale the stored energy is a quadratic function of the strain. The meso-scale local strain on the other hand is a linear function of the homogeneous boundary strain on the meso-scale model, and locally the stored energy is a quadratic function of the local strain, hence it is locally a quadratic function of the homogeneous boundary strain. The observable, the spatial average of the local meso-scale stored energy, is as a spatial integral of those quadratic functions again a quadratic function of the homogeneous boundary strain. Hence there exists a unique set of coefficients on the macro-scale for a perfect match of the energies. In computational practice, the RVE will often not be large enough, and the above consideration only applies if the macro-scale allows a large enough elastic symmetry class to capture possible anisotropies. Here we want an isotropic macro-scale model, as the meso-scale description is stochastically also isotropic and homogeneous. But as the RVE may not be large enough, residual model errors can be expected in numerical practice.

All our experiments will be performed only for 2D situations, as this is sufficient to demonstrate the approach. The example meso-scale specimen consists of a 2D block described by 64 circular inclusions of equal size randomly distributed in the domain. In the first case scenario only one meso-scale realisation is observed for verification purposes. The computational FE-model uses regular 50×50 mesh of standard bi-linear Q4 quad

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

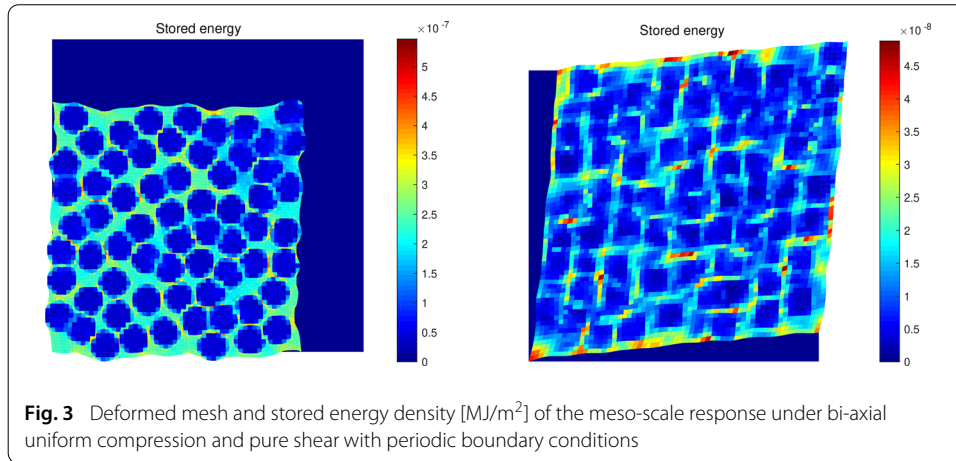Page 20 of 35

Fig. 2 Experimental setup
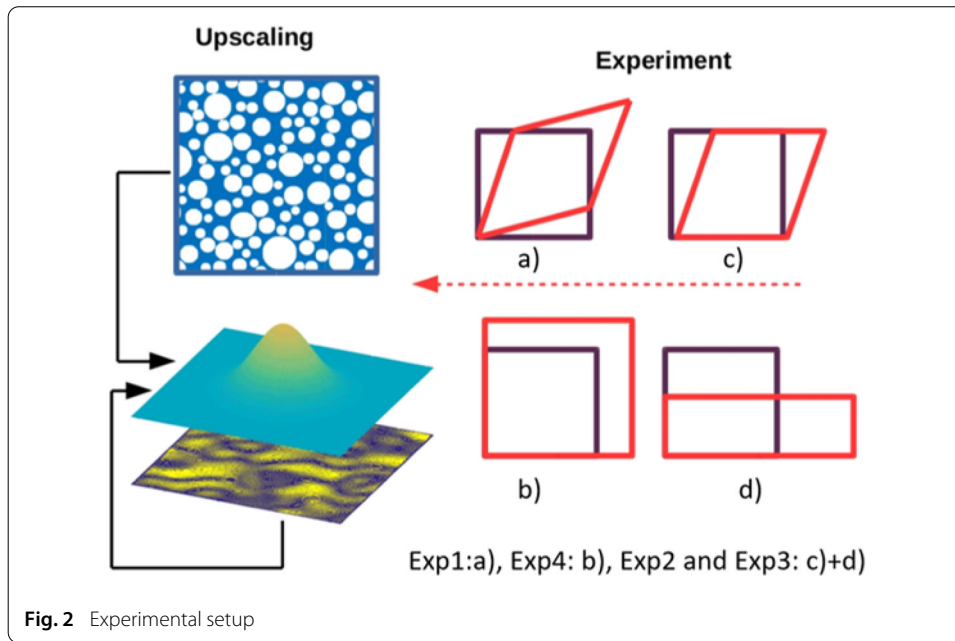
elements for the meso-scale, and *one* such element for the macro-scale. The material properties are taken as follows: the bulk modulus is $K_m = 4$ MPa, and the shear modulus is $G_m = 1$ MPa for the matrix phase, whereas the inclusions characteristics are prescribed to be ten times higher. The volume fraction of the inclusions is taken as 40%.

The meso-scale characteristics are upscaled in a Bayesian manner to the coarse scale homogeneous isotropic finite element described by material properties taking the form of a posteriori random variable as schematically shown in Fig. 2. To gather as much as information as possible in observation data, we consider different types of loading conditions including pure shear or compression, or their combination as shown in Fig. 2.

Verification

To verify our method, we compare Bayesian upscaling procedure to the deterministic homogenisation approach (as presented in [61], Exercise 4). Therefore, we initially observe only one realisation of the random meso-structure and apply periodic boundary conditions. For the sake of clarity, we describe here the deterministic approach briefly: the FE simulation is performed on the meso-structure, the resulting response is used to compute the macro-scale deformation or stress tensor (depending on the deformation, traction or mixed boundary conditions on the RVE), finally the computed average macro-scale quantities are used to compute the homogeneous/effective material parameters i.e. bulk and shear moduli in the current setting. For further details, the interested reader can consult [61]. An example of the fine-scale response in terms of the element level energy density is shown in Fig. 3 for Experiment 1 and 4. Here we recall that one of the distinguishing features of Bayesian upscaling approach is that it uses energy to estimate macro-scale material parameters.
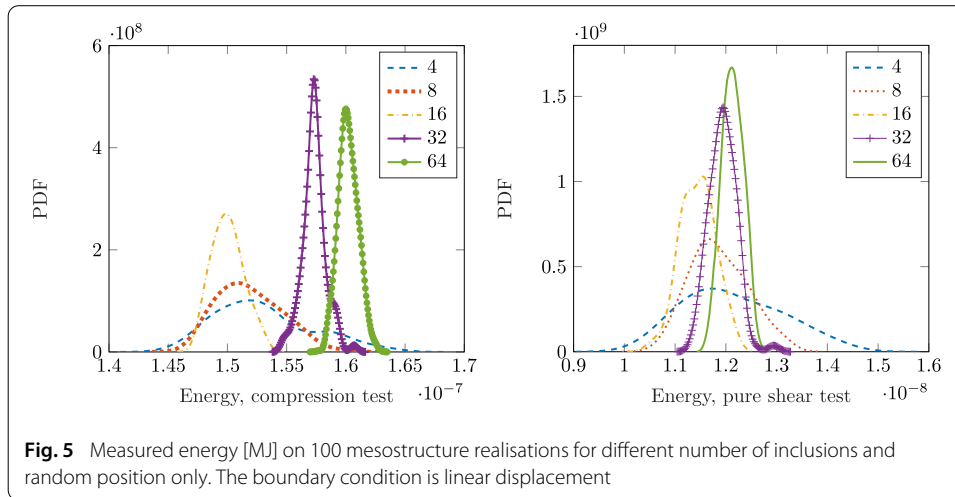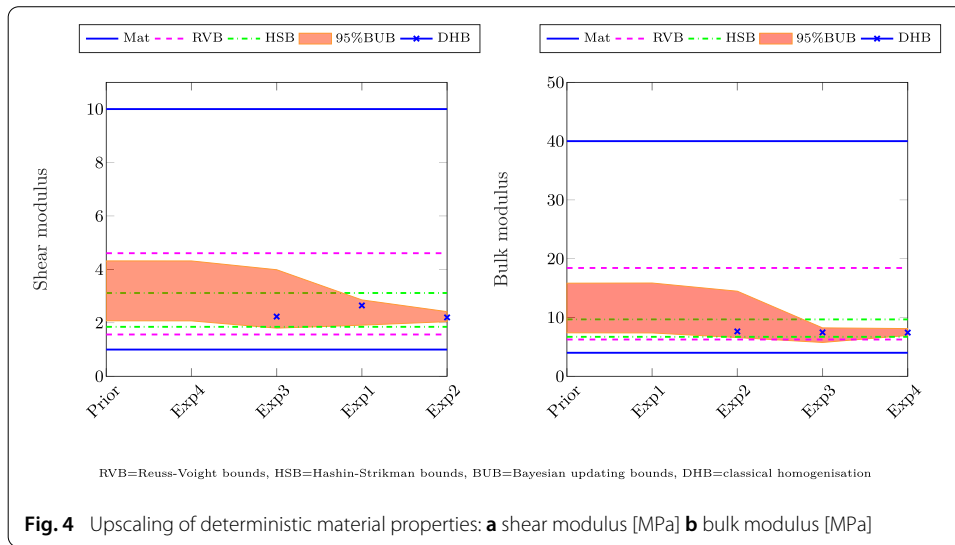
The comparison of the deterministic homogenization and Bayesian upscaling results (abbreviated as DHB and BUB respectively) are shown in Fig. 4, along with different analytical bounds (computed using the given material properties and volume fraction).

Fig. 3 Deformed mesh and stored energy density [MJ/m^2] of the meso-scale response under bi-axial uniform compression and pure shear with periodic boundary conditions

These bounds are defined as follows (with increasing degree of refinement): the material properties bound (Mat) with inclusion and matrix properties defined as higher and lower limits respectively, the *Reuss-Viogt* bounds (RBH) and *Hashin-Strickman* bounds (HSB). As expected and depicted in Fig. 4, the DHB fails to predict the shear moduli in a pure compression state (Experiment 4). An analogous result is expected and also obtained for the bulk modulus in case when only shear loading conditions are applied (Experiment 1). On the other hand, the BUB in the form of Eq. (33) regularises the problem by introducing prior information. When the data are not informative about the parameter set, this is recognised in that the posterior estimate is unchanged from the prior. Otherwise, the posterior mode and the deterministic homogenised value are identical after all experiments. Note that the blue crosses representing the DHB results, taken from [61], are fluctuating before the final estimate in contrast to the Bayesian estimate. We also observe from Fig. 4 that both the DHB and BUB results remain within the confines of analytical bounds. In particular for BUB results, the bounds for shear modulus reside inside all of the considered analytical limits for the last two experiments: 1 and 2, whereas for bulk modulus, a similar behaviour is observed for experiments: 3 and 4. To conclude, the Bayesian upscaling procedure is more robust than the classical one, and it additionally reflects possible model errors due an insufficient size of the RVE in the residual uncertainty. In addition, in the Bayesian upscaling procedure one may sequentially introduce the measurement data into the upscaling process. For example, one may first use the measurement information coming from the fourth experiment to obtain the upscaled material properties. These further can be used as a new prior for the third experiment, and so on, see Fig. 4.
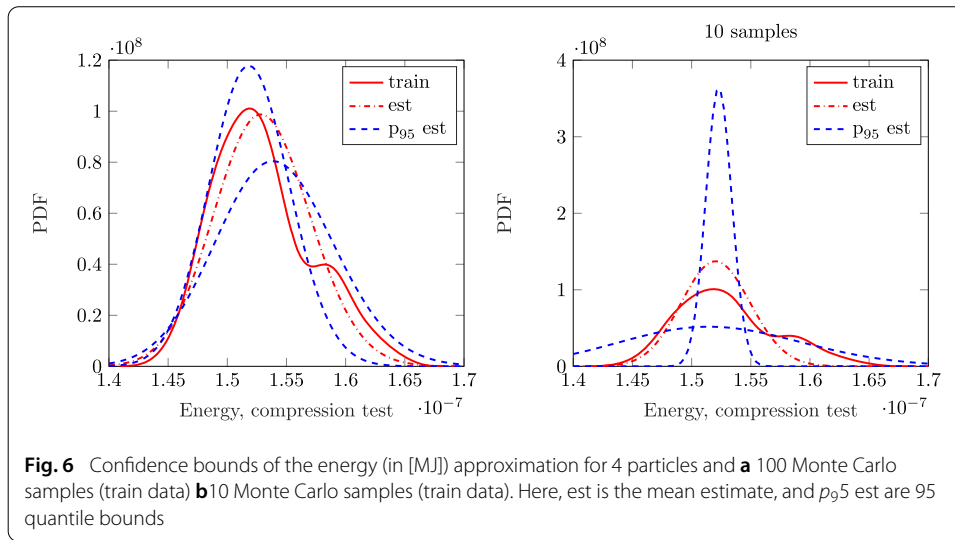
Upscaling of random elastic material

To quantify randomness on the meso-scale level, the previously described experiment is repeated several times, and the averaged stored elastic energies per experiment are collected. In particular, we observe realisations of the meso-scale elastic material described by randomly placed inclusions with a volume fraction of 40%. Initially, the stored energy is identified given the observed data by using the variational Bayesian inference method as described in "Bayesian upscaling of random meso-structures" section, resp. "Approximation of the meso-scale observation by unsupervised learning" section. The logarithm of the energy is modelled by a copula Gaussian mixture model, and the individual compo-

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 22 of 35

Fig. 4 Upscaling of deterministic material properties: **a** shear modulus [MPa] **b** bulk modulus [MPa]

Fig. 5 Measured energy [MJ] on 100 mesostructure realisations for different number of inclusions and random position only. The boundary condition is linear displacement

nents are identified. The optimal number of mixture components is further decoupled by an inverse transform. The resulting uncorrelated random variables are then further used to obtain the polynomial chaos surrogate of the measurement data.

The simulation is performed on the 2D meso-structure with an increasing number of particles and linear displacement based boundary condition.The material properties for the matrix and inclusion phases are kept the same as considered previously in the verification procedure. For a given number of inclusions embedded in the matrix phase, an ensemble of 100 realisations of stored energy is considered to gather corresponding measurement set. In Fig. 5 the PDFs for the identified elastic energies are shown for the pure shear and the bi-axial compression test, respectively. As expected, the variation of stored energy reduces with the increase of the number of particles in the matrix phase. One would expect that these would converge to the same PDF after taking large number of inclusions. However, here the largest taken number was not large enough. In order to be able to compare results to [61], we therefore did not increase this number any more. It is interesting to note that in the compression case the mean responses of stored energies vary more than in a pure shear test. This is closely related to the way how boundary

Fig. 6 Confidence bounds of the energy (in [MJ]) approximation for 4 particles and **a** 100 Monte Carlo samples (train data) **b** 10 Monte Carlo samples (train data). Here, est is the mean estimate, and $p_9$5 est are 95 quantile bounds
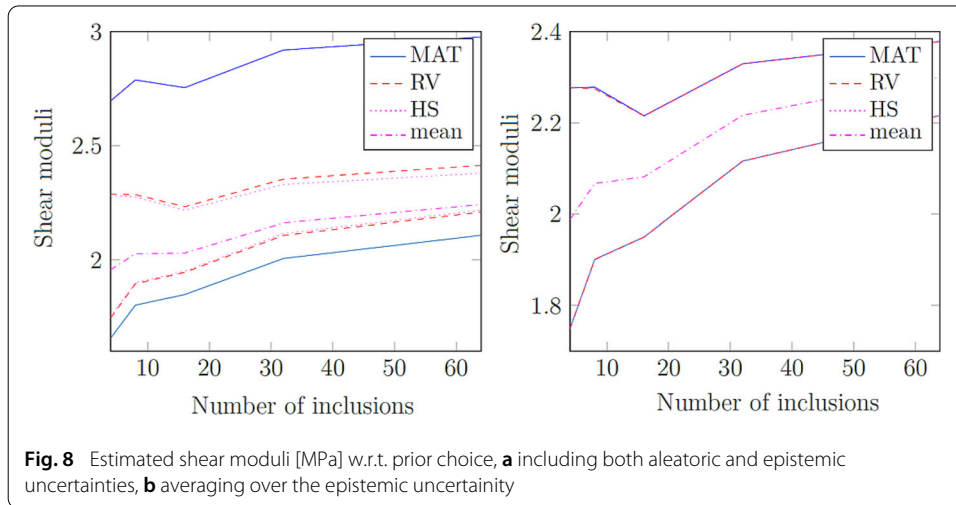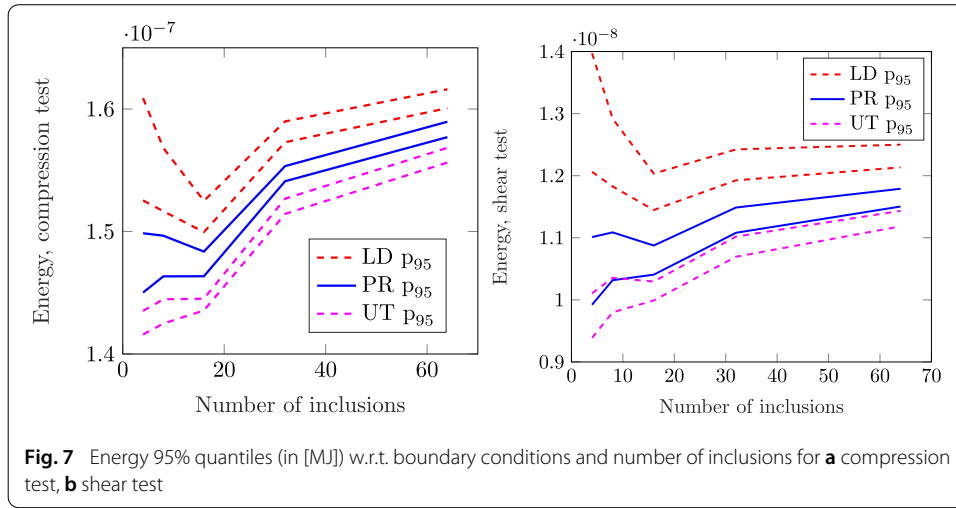
conditions are imposed. Namely, we take into consideration directly the element on which the loadings are imposed, and thus one may recognise the strong influence of boundary conditions on the obtained results. In would be better if one would take into consideration only internal elements, which are away from the boundary. This means that the averaging would be performed only over an RVE embedded in a large domain on which the boundary conditions are applied.

The previously discussed results are estimates of the stored energy given its samples following Eq. (41), obtained by the unsupervised learning algorithm described in "Approximation of the meso-scale observation by unsupervised learning" section. However, the residual uncertainty consists of two kinds of uncertainties: aleatoric (meso-scale randomness) and epistemic (prior information in the unsupervised learning algorithm, see Eq. (35)). Estimating the confidence intervals w.r.t. to the epistemic uncertainties one obtains the corresponding PDFs of energy: the mean PDF which represents purely aleatory uncertainty and p_{95} upper and lower PDF's that describe 95% epistemic quantiles on the mean PDF, see Fig. 6a. Naturally the epistemic quantile intervals strongly depend on the size of the measurement set. From Fig. 6b one may conclude that with the smaller measurement set by using only 10 samples our confidence about the estimated PDF is lower than in case of higher number of measurements, as expected.

Besides the previous analysis, the impact of boundary conditions on the upscaled quantities is another important factor to study. In Fig. 7 are depicted the 95% quantiles of energy for linear displacement (LD), periodic (PR) and uniform tension (UT) boundary conditions. According to these results, linear displacement defines the upper bound on the estimated energy, whereas uniform tension gives its lower limit. On the other hand, variations of the energies are similar for all three types of boundary conditions, and are inverse proportion to the number of inclusions.

Once the measurement energy is identified, in the second step we use the proxy of y_m to identify the elastic macro-scale material characteristics by using the filter of polynomial order 2 as given in Eq. (42). When using this type of upscaling ,one is biased to the prior knowledge of the material characteristics on the macro-scale. In a multi-scale analysis, however, it is not an easy task to define the prior knowledge, or better to say the limits of the

Fig. 7 Energy 95% quantiles (in [MJ]) w.r.t. boundary conditions and number of inclusions for **a** compression test, **b** shear test

Fig. 8 Estimated shear moduli [MPa] w.r.t. prior choice, **a** including both aleatoric and epistemic uncertainties, **b** averaging over the epistemic uncertainty

prior distribution. Therefore, in Fig. 8 is investigated the posterior change of shear moduli w.r.t. prior knowledge. The prior distributions are chosen such that their 95% limits match the interval described by the material properties of the matrix phase and inclusions (in the figure denoted by MAT), Reuss-Voigt (RV) or Hashin-Shtrikman (HS) bounds. Their corresponding 95% posterior limits w.r.t. number of inclusions are depicted in Fig. 8a. It is interesting to note that even though the posterior (including both aleatoric and epistemic uncertainties) of the upscaled shear moduli changes w.r.t. the prior assumption, its posterior averaged over the epistemic uncertainty as in Eq. (43) remains the same, see Fig. 8b, and does not depend on the prior knowledge.

To verify our result further, in Fig. 9 we compare the aleatoric part of the posterior distribution with the posterior distribution obtained by repeating the deterministic homogenisation, see [61], on each of the meso-scale samples. As one may notice, the distribution coming from the deterministic homogenisation (denoted by det) and the aleatoric one obtained by our approach (denoted by partial) are matching. They are further compared with the full posterior distribution (denoted by total), i.e. the total uncertainty that includes both aleatory and epistemic knowledge.
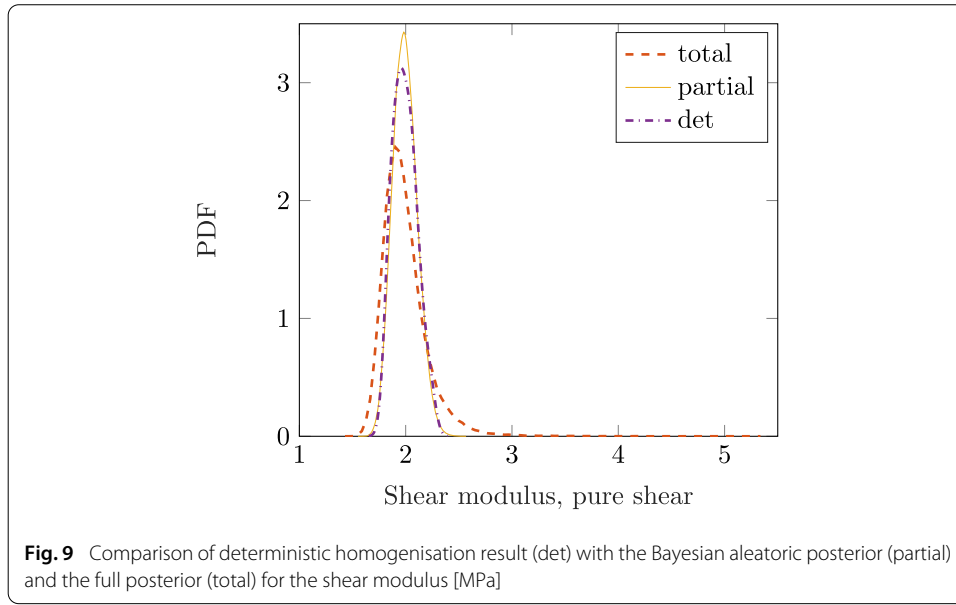
Fig. 9 Comparison of deterministic homogenisation result (det) with the Bayesian aleatoric posterior (partial) and the full posterior (total) for the shear modulus [MPa]

Table 1 Mean values of random meso-scale material parameters [MPa]

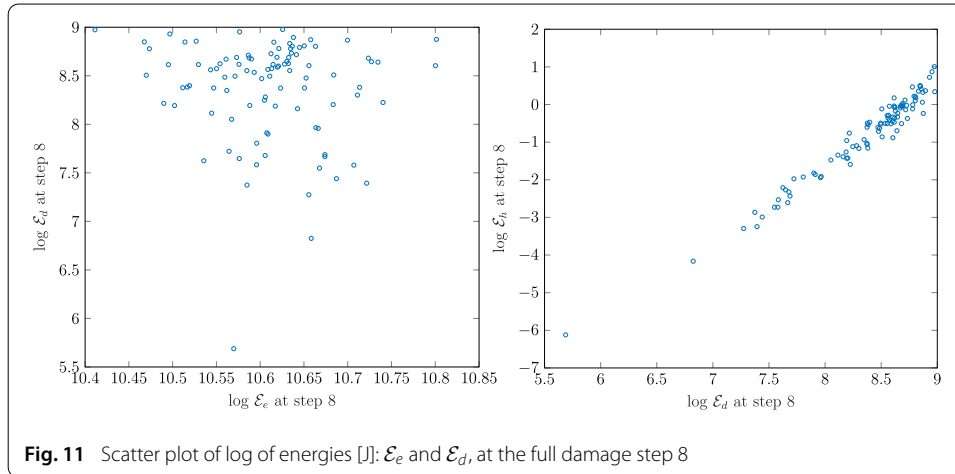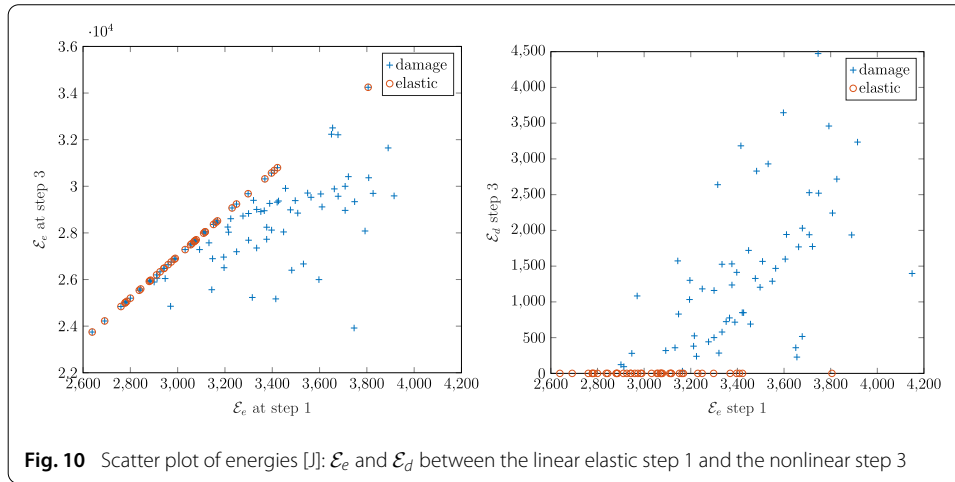| κ	μ	σ_f	K_d
204,440	92,000	310	470

Upscaling of damage phenomena

In this subsection, the proposed approach is applied to another interesting problem. For this purpose a phenomenological elasto-damage model is considered as described in the beginning of this section. The goal is to compute a homogenised description of random material parameters on the macro-scale given meso-scale measurements. The meso-scale is assumed to follow same constitutive model as the macro-scale. For verification purposes we assume that the meso-scale has homogeneous material properties, which are modelled by a random variable. Hence, the meso- and macro-scale are identical. Furthermore, in a second experiment we model the meso-scale material properties as spatially varying, and apply the upscaling procedure in order to estimate the homogeneous macro-scale material properties. In both experiments we only simulate the displacement controlled uniform bi-axial compression of a 2D block with unit length (similar to the experiment in the previous example). The volumetric strain $\varepsilon_v = \text{tr}\,\boldsymbol{\varepsilon}/2$ in 2D is calculated for a given time step through piece-wise linear interpolation from the set of values given as $(t, \varepsilon_v) : \{(0, 0), (3, -0.00025), (10, -0.00035)\}$, where t denotes the pseudo-time. For the experiments under consideration, the displacement is applied in 8 equidistant steps in t.

Verification

For verification purposes the material parameters $\boldsymbol{Q}_m$ on the meso-scale are modelled as lognormal random variables with the mean values shown in Table 1 and the coefficient of variation 5%. After propagating the variables through the elasto-damage model, the corresponding measurements as in Eq. (49) are estimated. We assume that the polynomial chaos approximation of the measurement is not given, but only a set of 100 samples. Therefore, the log of measurements are modelled as copula Gaussian mixtures with the

Fig. 10 Scatter plot of energies [J]: $\mathcal{E}_e$ and $\mathcal{E}_d$ between the linear elastic step 1 and the nonlinear step 3

Fig. 11 Scatter plot of log of energies [J]: $\mathcal{E}_e$ and $\mathcal{E}_d$, at the full damage step 8

unknown number of components. The simulation is run in 8 equidistant time steps, the first two being elastic. In the third to sixth steps the behaviour is a combination of elasticity and damage, whereas in the last step is dominated by the damage component.

In Fig. 10 are shown scatter plots of energies in the first and the third step, both depicting two states in the response: elastic and damage. The red circles denote samples that are in the elastic state in the third step, whereas blue crosses denote samples that experience damage behaviour in the third step. The third step is the first time step in which damage behaviour initiates. Hence, the correlation between elastic energies in the third and the first step for samples that are undergoing elastic behaviour is linear, see red circles in left plot in Fig. 10, as expected. However, the elastic part of energy in the third step has nonlinear correlation to the elastic part of energy in the first step for the samples that are switching from elastic to damage state, see blue crosses in the left plot in Fig. 10. Similar holds for the right plot in Fig. 10. Here one may see that the samples that remain in elastic state from the first to the third step do not have $\mathcal{E}_d$ in the third step (therefore the straight line made of red circles), whereas samples that change their state from elastic to damage have non-zero $\mathcal{E}_d$ nonlinearly correlated to the elastic part of energy in the first step.

To estimate the macro-scale properties, we observe measurements at the last time step as depicted in Fig. 11. Clearly, the $\mathcal{E}_h$ and $\mathcal{E}_d$ are almost linearly related in the log-space,
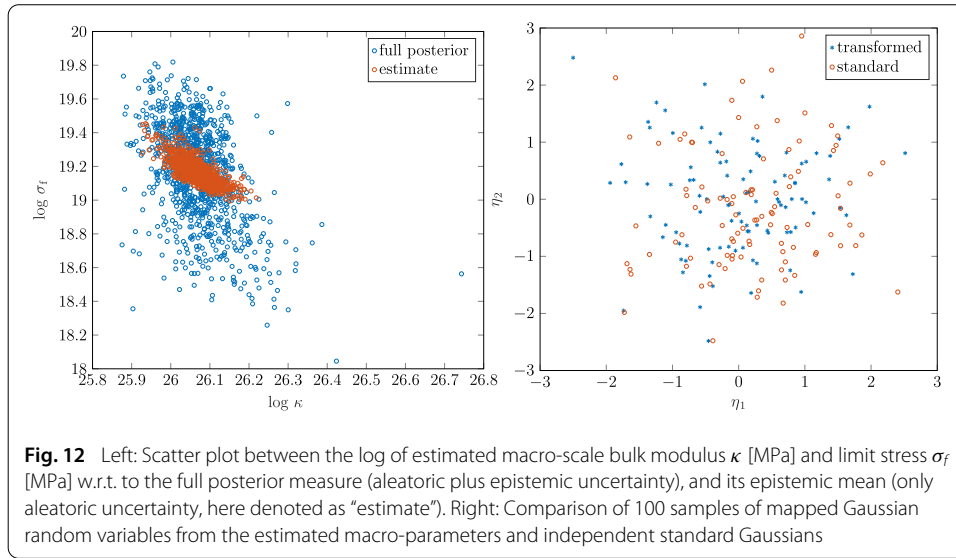
Fig. 12 Left: Scatter plot between the log of estimated macro-scale bulk modulus κ [MPa] and limit stress σ_f [MPa] w.r.t. to the full posterior measure (aleatoric plus epistemic uncertainty), and its epistemic mean (only aleatoric uncertainty, here denoted as "estimate"). Right: Comparison of 100 samples of mapped Gaussian random variables from the estimated macro-parameters and independent standard Gaussians
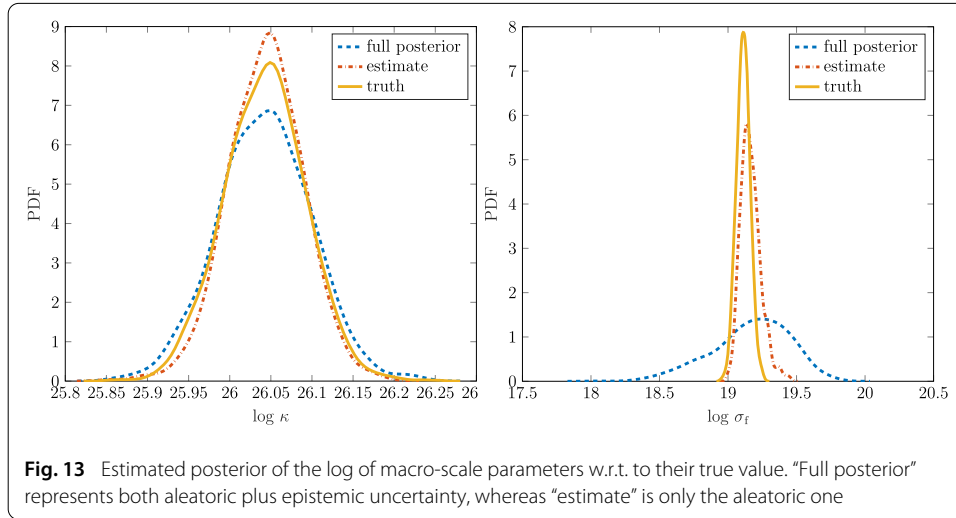

Fig. 13 Estimated posterior of the log of macro-scale parameters w.r.t. to their true value. "Full posterior" represents both aleatoric plus epistemic uncertainty, whereas "estimate" is only the aleatoric one

whereas this does not hold for the $\mathcal{E}_d$ and $\mathcal{E}_e$. Furthermore, we employ a copula approach, see "Approximation of the meso-scale observation by unsupervised learning" section, to uncouple these measurement data, and estimate their functional approximations. Once they are mapped to Gaussian random variables, we may easily generate the approximation of measurements at other time steps. For this we utilise the approach described in "Approximating the macro-scale response by Bayesian regression" section.

Given the approximation of the measurement y_m, we may estimate homogeneous macro-scale properties by the approach described in "Bayesian upscaling of random meso-structures" section. The a priori description of the macro-scale properties is taken to be also modelled as lognormal random variables with the mean 20% larger than in the meso-scale case, and a coefficient of variation of 20%.

The resulting updated macro-scale properties are shown in Fig. 13. The bulk modulus κ and the limit stress σ_f that initiates the damage are both updated and match the true distribution, whereas their correlation and the mapping to the normal space is shown in Fig. 12. The left plot in Fig. 12 depicts the correlation between upscaled bulk modulus κ
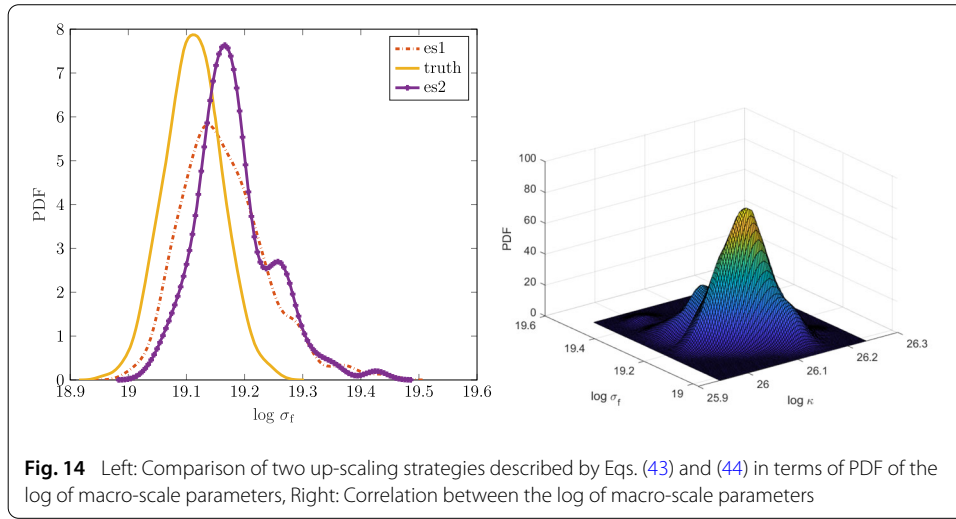
Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 28 of 35

Fig. 14 Left: Comparison of two up-scaling strategies described by Eqs. (43) and (44) in terms of PDF of the log of macro-scale parameters, Right: Correlation between the log of macro-scale parameters

and the limit stress σ_f. Here, one may distinguish the correlation calculated by taking into account the full posterior (both aleatory uncertainty due to uncertain RVEs, as well as epistemic uncertainty or only its aleatoric (estimate) part (obtained by averaging the posterior over the prior uncertainty). Note that the remaining epistemic uncertainty is bigger for the limit stress σ_f than the bulk modulus κ, as the relationship between the parameter and the measurement is more nonlinear than in case of the bulk modulus. The right plot in Fig. 12 describes the correlation between variables (η_1, η_2) obtained after mapping the measurement data $(\mathcal{E}_e, \mathcal{E}_d)$ to the Gaussian space by the use of algorithms in "Approximation of the meso-scale observation by unsupervised learning" section. These are referred as transformed variables, and are further compared to the sample set of standard uncorrelated Gaussians in order to verify the mapping algorithm. As can be seen, the mapped Gaussians are indeed uncorrelated, and hence can be used for further approximations. On the other hand, due to the chosen experiment, both shear and hardening moduli stay unidentified as they are not observable. Hence, their analysis is not considered.

In the previously described experiment the relationships between the measurement data and their approximations are too complex in order to be properly modelled. Therefore, the experiment is repeated in same setting, only this time the measurement is not functionally approximated. Instead, the inverse problem is solved for each individual sample of measurement (each RVE), and then the updated parameters are collected into the set of parameter samples as described in Eqs. (44) and (46). This calculation is expected to be simpler than the previous one as the relationships between the material parameters are easier to model. In Fig. 14 is depicted the difference between this approach (est1) and the previous one (est2), as well as the joint distribution between the bulk modulus κ and σ_f. Hence, by upscaling we obtain a simpler representation of our meso-scale data, however, at the expense of a correlation of the material parameters.
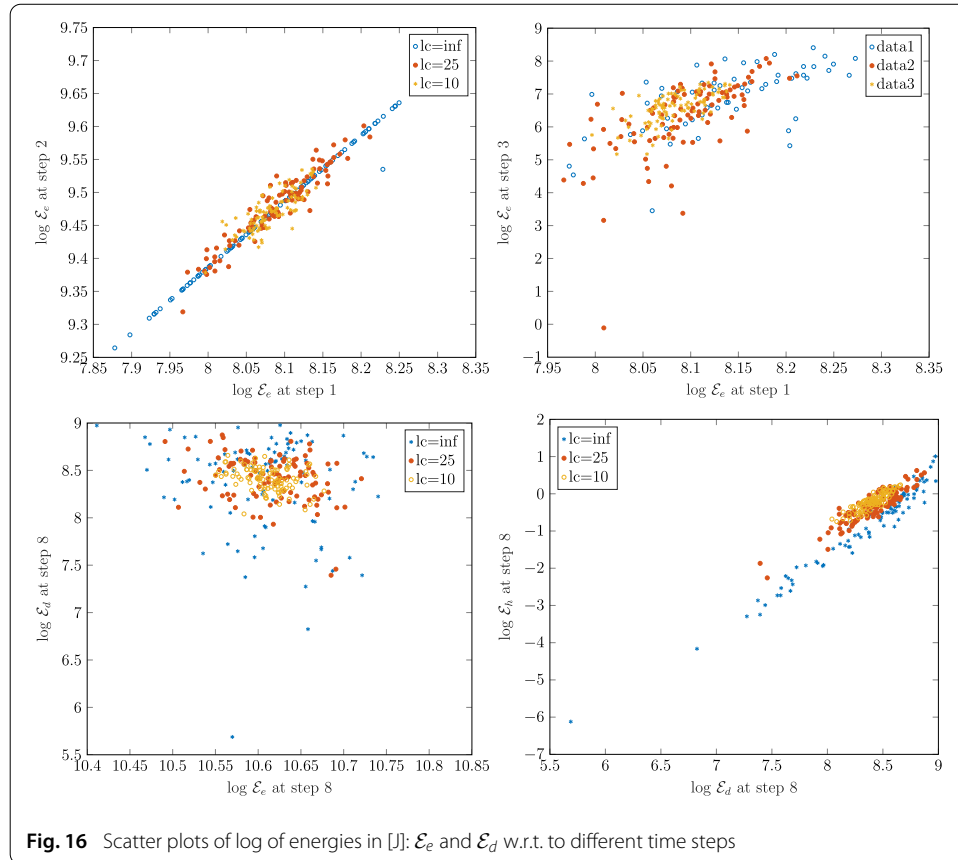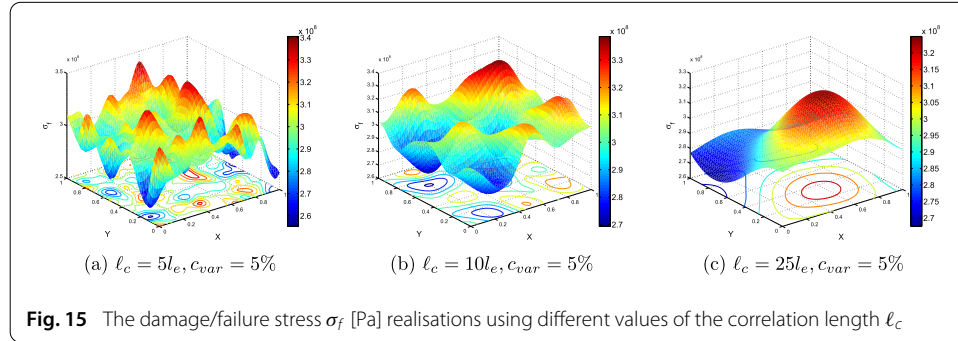
Upscaling of a heterogeneous medium

As before, the block is deformed by displacement controlled uniform bi-axial compression.

As far as the material description is concerned, the material properties on the heterogeneous meso-scale are a priori assumed to be realisations of log-normal random fields with the statistics depicted in Table 2, and Gaussian covariance functions. These are simulated
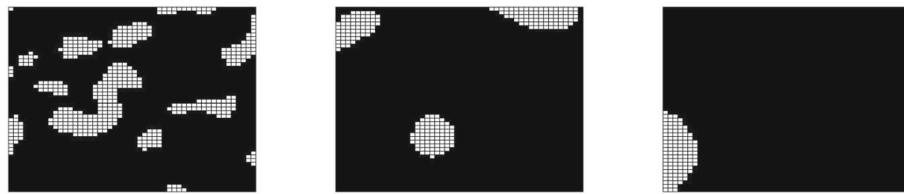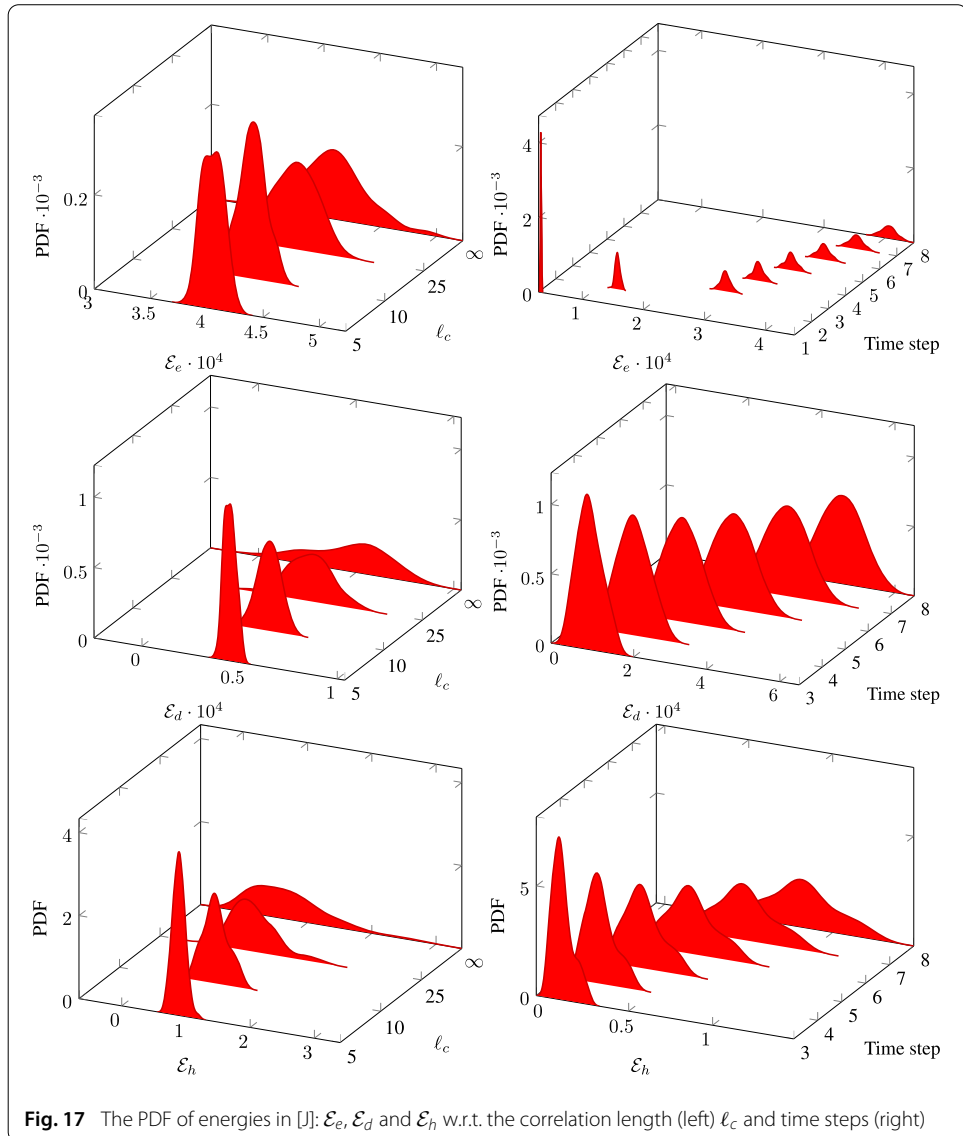
Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 29 of 35

Table 2 Macro-scale statistics for elasto-damage constitutive model in [MPa]

Property	κ	σ_f	K_d
Mean	204, 440	300	450
Std.dev.	10, 222	15	22.5

(a) $\ell_c = 5l_e, c_{var} = 5\%$ (b) $\ell_c = 10l_e, c_{var} = 5\%$ (c) $\ell_c = 25l_e, c_{var} = 5\%$

Fig. 15 The damage/failure stress σ_f [Pa] realisations using different values of the correlation length ℓ_c

Fig. 16 Scatter plots of log of energies in [J]: $\mathcal{E}_e$ and $\mathcal{E}_d$ w.r.t. to different time steps

using different values of the correlation length $\ell_c \in \{5l_e, 10l_e, 25l_e\}$ (l_e is the element length on the meso-scale) and coefficients of variation $c_{var} \in \{5\%, 10\%\}$.

In Fig. 15 is shown an example of the meso-scale random field realisations given different correlation lengths. The realisation is becoming smoother when the correlation length increases. This means that the material becomes more homogeneous in the limit $\ell_c = \infty$. On the other hand, the macro-scale material properties are taken a priori as a log-

Fig. 17 The PDF of energies in [J]: $\mathcal{E}_e$, $\mathcal{E}_d$ and $\mathcal{E}_h$ w.r.t. the correlation length (left) ℓ_c and time steps (right)

Fig. 18 The presence of damage (elements marked in black) on the meso-scale for $c_{var} = 0.1$ and $\ell_c = 5\ell_e$, $\ell_c = 10\ell_e$, $\ell_c = 25\ell_e$, respectively

normal random variables, with the same mean and standard deviation as their meso-scale counterpart.

The measurement data are made of three type of averaged data: $\mathcal{E}_e$, $\mathcal{E}_d$, and $\mathcal{E}_h$. Their logarithms are simulated using mixture models and vine copulas, and further identified using a variational Bayesian rule. Similarly to the experiment in the verification section, in the first two simulation steps one may observe only the elastic energy, as the dissipation effects do not appear yet. Therefore, we start the simulation with the last step, and

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.* (2020)7:50

Page 31 of 35
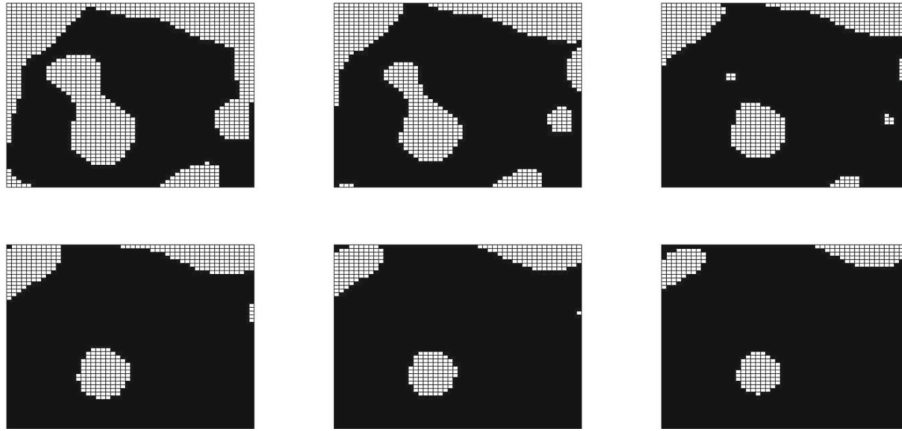
Fig. 19 The presence of damage (elements marked in black) on the meso-scale for one random field realisation with $c_{var} = 0.1$ and $\ell_c = 10\ell_e$ w.r.t. time step

approximate the corresponding measurements by mixture models and vine copulas, see Fig. 16. The complete simulation steps from the previous section are repeated. However, it is interesting to note the change of the scatter plots with the correlation length. The scatter plots of $\mathcal{E}_e$ in two consecutive linear steps are wider with the reduction of the correlation length. The opposite holds true when it comes to relationship between $\mathcal{E}_e$ and $\mathcal{E}_d$ in the last simulation step.

The distribution of damage is graphically illustrated in Figs. 18 and 19. The elements without any color are undamaged, whereas the ones marked in black are damaged. As shown in Fig. 17 the variation of measurements increases with the correlation length size for the case when the c_{var} of the meso-scale random field is taken to be 10%. The reason for this is that measurement realisations are less fluctuating with increasing correlation length, but their average value is more pronounced as prospective fluctuations do not cancel out, as similar can be concluded when observing Fig. 15. The previous conclusion holds for all measurements, and can be explained by Fig. 18 in which the presence of damage for one random field realisation and different correlation lengths is shown. With increase of the correlation length the damage is more pronounced, and hence one expects higher variations. In addition, one may also conclude that the corresponding PDFs are becoming more skewed when the material model approaches the homogeneous case, see Fig. 17. The skewness in terms of long tails is not completely caused by variations of the random meso-scale, but also by the inaccuracy of the variational method used for the PCE estimation of the measurement due to possible overestimation.

On the right side of Fig. 17, one may observe the energy evolution w.r.t. to time. Here, c_{var} of the corresponding meso-scale random field is chosen to be 10% and the correlation length is $\ell_c = 10\ell_e$. The top figure depicts the elastic energy. As expected, the energy variation grows in time. On the contrary, $\mathcal{E}_d$ seen in the middle does not alter much the PDF form. The damage initialises in the third step, and mostly shifts towards higher average value due to increased presence of damage as shown in Fig. 19. Finally, $\mathcal{E}_h$ increases, but also changes the PDF form significantly in time.

The upscaled parameter estimates behave similarly to the measurement estimates as shown in Fig. 20. The hardening parameter does not get updated, and stays constant over time.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 32 of 35

Fig. 20 The PDF of parameters κ (in Pa), Failure stress σ_d (in Pa) and Hardening modulus K_d (in Pa) w.r.t. the correlation length (left) ℓ_c and time steps (right)
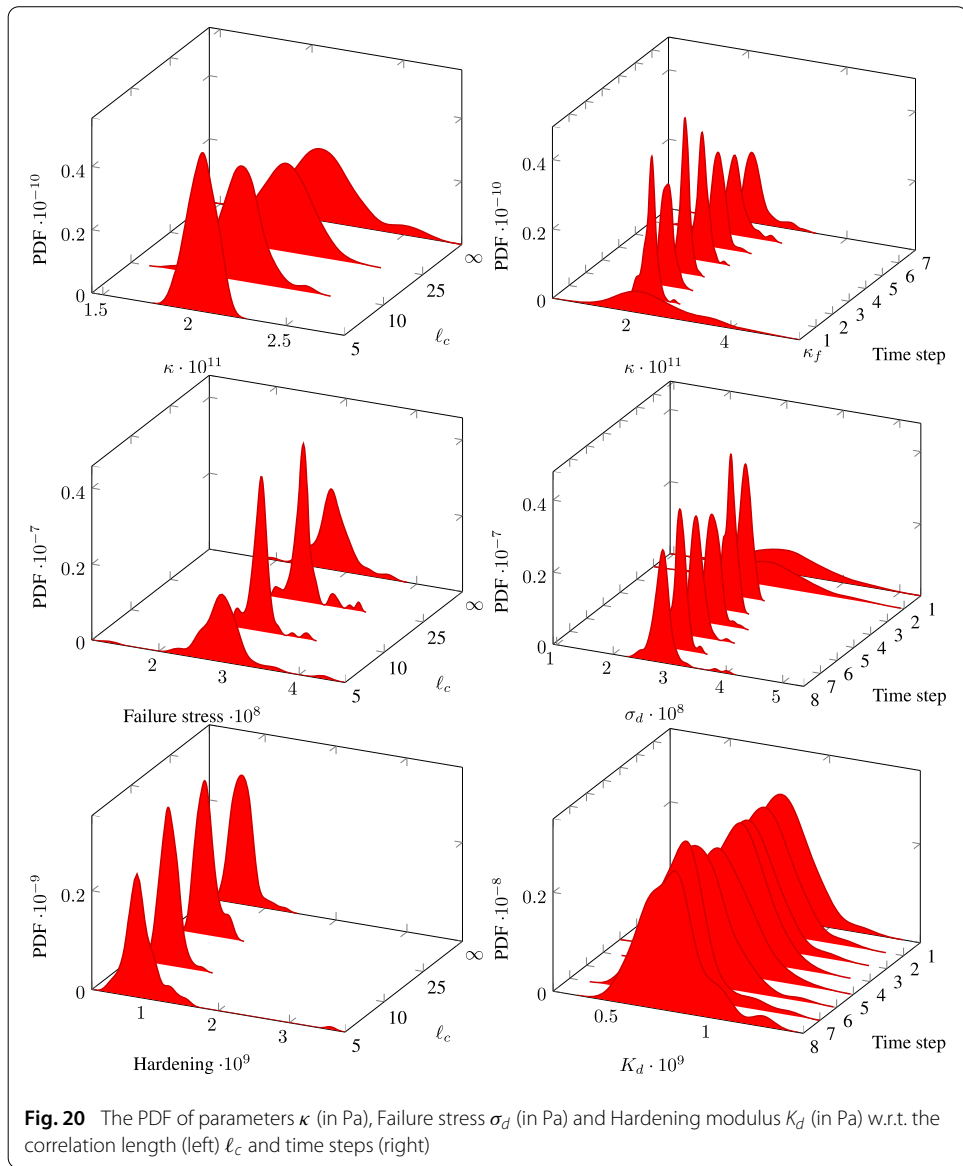
Conclusion

The stochastic multi-scale analysis as previously presented is one particular kind of inverse problem in which the macro-scale parameters are to be estimated given the meso-scale information. In this paper we employed an extended Hill-Mandel principle in order to estimate the macro-scale parameters given the meso-scale energy observations. Such an approach allows the fitting of appropriate constitutive laws on the macro-scale counterpart, the ones that are optimally matching the energy information. Furthermore, we show that in a case when the meso-scale energy information is of deterministic kind, i.e. describes one particular RVE, the process of estimation can be easily done by employing a nonlinear conditional expectation filter. The filter represents the map between the observation and the quantity of interest, i.e. the macro-scale model parameters, or the model itself. In addition, we have shown that this kind of mapping can be also used in a more general situation, in which one wants to upscale an ensemble of meso-structures, and the meso-scale information is described by aleatoric uncertainty. The only requirement to

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 33 of 35

achieve this is to fully specify the random variables representing the data, i.e. to describe its probability distribution. For this purpose we employ a Bayesian variational inference in combination with copula theory. Computationally, the measurement probability distribution is then represented by a functional approximation in terms of the polynomial chaos expansion obtained by mapping the measurement data to the Gaussian space, applying an inverse transform, and using an additional sparse Bayes variational regression for the purpose of estimation of the expansion coefficients. As the inverse map from the energy space to the Gaussian one is not easy to approximate, we recommend to first discretise the energy space (i.e. to sample), and then to map each sample to the macro-scale model parameters. As shown on both the linear elastic and the elasto-damage examples, the latter ones can be more accurately approximated. Note that in this paper we have only observed the elasto-damage models on two scales under one specific loading condition. This was possible due to the simple nature of the damage model. However, in practice this will not suffice to achieve a good macro-scale representation. Therefore, the next step to be considered is to add different loading conditions into estimation.

Author details
[1]Institute of Scientific Computing, Technische Universität Braunschweig, Braunschweig 38106, Germany, [2]Chair for Applied Mechanics and Data Analysis, University of Twente, Drienerlolaan 5, Enschede, NB 7522, Netherlands, [3]Chair for Computational Mechanics, Université de Technologie de Compiègne, 60205 Compiègne, France.

References
1. Aas K, Czado C, Frigessi A, Bakken H. Pair-copula construction of multiple dependence. Insur Math Econ. 2009;44(2):182–98.
2. Bishop CM. Pattern recognition and machine learning. Information science and statistics. Berlin: Springer; 2006.
3. Bishop CM, Tipping ME. *Variational relevance vector machines*, 2013. arxiv:1301.3838.
4. Bobrowski A. Functional analysis for probability and stochastic processes: an introduction. Cambridge: Cambridge University Press; 2005.
5. Bruder L, Koutsourelakis P-S. *Beyond black-boxes in Bayesian inverse problems and model validation: applications in solid mechanics of elastography*, 2018. arXiv:1803.00930.
6. Clément A, Soize C, Yvonnet J. High-dimension polynomial chaos expansions of effective constitutive equations for hyperelastic heterogeneous random microstructures, Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS. TU Wien. 2012;2012:1–2.
7. Clément A, Soize C, Yvonnet J. Computational nonlinear stochastic homogenization using a nonconcurrent multiscale approach for hyperelastic heterogeneous microstructures analysis. Int J Numer Methods Eng. 2012;91(8):799–824.
8. Clément A, Soize C, Yvonnet J. Uncertainty quantification in computational stochastic multiscale analysis of nonlinear elastic materials. Computer Methods Appl Mech Eng. 2013;254:61–82.
9. Colliat J-B, Hautefeuille M, Ibrahimbegović A, Matthies HG. Stochastic approach to size effect in quasi-brittle materials. Comptes Rendus Mécanique. 2007;335:430–5.
10. de SouzaNeto EA, Peric D, Owen DRJ. Computational methods for plasticity: theory and applications. New York: Wiley; 2011.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 34 of 35

11. Felsberger L, Koutsourelakis P-S. *Physics-constrained, data-driven discovery of coarse-grained dynamics*, 2018. arXiv:1802.03824.

12. Feyel F. A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua. Computer Methods Appl Mech Eng. 2003;192:3233–44. https://doi.org/10.1016/S0045-7825(03)00348-7.

13. Feyel F, Chaboche J-L. FE2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials. Computer Methods Appl Mech Eng. 2000;183(3):309–30. https://doi.org/10.1016/S0045-7825(99)00224-8.

14. Franck IM, Koutsourelakis P-S. Sparse variational Bayesian approximations for nonlinear inverse problems: applications in nonlinear elastography. Computer Methods Appl Mech Eng. 2016;299:215–44.

15. Franck IM, Koutsourelakis P-S. Multimodal, high-dimensional, model-based, Bayesian inverse problems with applications in biomechanics. J Comput Phys. 2017;329:91–125.

16. Franck IM, Koutsourelakis P. Constitutive model error and uncertainty quantification. PAMM. 2017;17(1):865–8.

17. Geers MG, Kouznetsova VG, Matouš K, Yvonnet J. Homogenization methods and multiscale modeling: nonlinear problems. In: de Borst R, Hughes TJR, editors. Encyclopaedia of computational mechanics. 2nd ed. New York: Wiley; 2017. p. 1–34.

18. Germain P, Nguyen QS, Suquet P. Continuum thermodynamics. Trans ASME. 1983;50:1010–20.

19. Graham-Brady L. Statistical characterization of meso-scale uniaxial compressive strength in brittle materials with randomly occurring flaws. Int J Solids Struct. 2010;47(18):2398–413. https://doi.org/10.1016/j.ijsolstr.2010.04.034.

20. Halphen B, Nguyen QS. Sur les matériaux standards généralisés. J de Mécanique. 1975;14:39–63.

21. Han W, Reddy BD. Plasticity: mathematical theory and numerical analysis. Interdisciplinary applied mathematics, vol. 9. 2nd ed. New York: Springer; 2013. https://doi.org/10.1007/978-1-4614-5940-8.

22. Hautefeuille M. *Numerical modeling strategy for heterogeneous materials: a FE multi-scale and component-based approach*, Ph.D. thesis, Université Technologique de Compiègne, Technische Universität Braunschweig, and École Normale Supérieure de Cachan; 2009.

23. Hautefeuille M, Colliat J-B, Ibrahimbegović A, Matthies HG, Villon P. A multi-scale approach to model localized failure with softening. Computers Struct. 2012;94–95:83–95. https://doi.org/10.1016/j.compstruc.2011.11.007.

24. Ibrahimbegović A. Nonlinear solid mechanics: theoretical formulations and finite element solution methods. Berlin: Springer; 2009. https://doi.org/10.1007/978-90-481-2331-5.

25. Ibrahimbegović A, Matthies HG. Probabilistic multiscale analysis of inelastic localized failure in solid mechanics. Computer Assisted Methods Eng Sci. 2012;19:277–304.

26. Koutsourelakis P-S. Variational Bayesian strategies for high-dimensional, stochastic design problems. J Comput Phys. 2016;308:124–52.

27. Koutsourelakis P-S. Stochastic upscaling in solid mechanics: an excercise in machine learning. J Comput Phys. 2007;226(1):301–25.

28. Le BA, Yvonnet J, He Q-C. Computational homogenization of nonlinear elastic materials using neural networks: neural networks-based computational homogenization. Int J Numer Methods Eng. 2015;104:1061–84. https://doi.org/10.1002/nme.4953.

29. Liu J, Graham-Brady L. Perturbation-based surrogate models for dynamic failure of brittle materials in a multiscale and probabilistic context. Int J Multiscale Comput Eng. 2016;14(3):273–90.

30. Lu X, Giovanis DG, Yvonnet J, Papadopoulos V, Detrez F, Bai J. A data-driven computational homogenization method based on neural networks for the nonlinear anisotropic electrical response of graphene/polymer nanocomposites. Comput Mech. 2019;64(2):307–21. https://doi.org/10.1007/s00466-018-1643-0.

31. Hoffman CWMD, Blei DM, Pasley J. Stochastic variational inference. J Mach Learn Res. 2013;14:1303–47.

32. Ma J, Sahraee S, Wriggers P, De Lorenzis L. Stochastic multiscale homogenization analysis of heterogeneous materials under finite deformations with full uncertainty in the microstructure. Comput Mech. 2015;55:819–35. https://doi.org/10.1007/s00466-015-1136-3.

33. Ma J, Temizer I, Wriggers P. Random homogenization analysis in linear elasticity based on analytical bounds and estimates. Int J Solids Struct. 2011;48:280–91. https://doi.org/10.1016/j.ijsolstr.2010.10.004.

34. Ma J, Wriggers P, Li L. Homogenized thermal properties of 3d composites with full uncertainty in the microstructure. Struct Eng Mech. 2016;57:369–87. https://doi.org/10.12989/sem.2016.57.2.369.

35. Ma J, Zhang S, Wriggers P, Gao W, De Lorenzis L. Stochastic homogenized effective properties of three-dimensional composite material with full randomness and correlation in the microstructure. Computers Struct. 2014;144:62–74. https://doi.org/10.1016/j.compstruc.2014.06.003.

36. Markovič D, Ibrahimbegović A, Niekamp R, Matthies HG. A multi-scale finite element model for inelastic behaviour of heterogeneous structures and its parallel computing implementation. Engineering Structures under Extreme Conditions. Multi-physics and multi-scale computer models in non-linear analysis and optimal design of engineering structures under extreme conditions — NATO-ARW (A. Ibrahimbegović and B. Brank, eds). Narodna i univerzitetna knjižica, Ljubljana; 2004. p. 342–351. http://arw-bled2004.scix.net/Files/acceptedpapers/Accepted/Markovic.pdf.

37. Markovič D, Niekamp R, Ibrahimbegović A, Matthies HG, Taylor RL. Multi-scale modeling of heterogeneous structures with inelastic constitutive behavior: Part I—physical and mathematical aspects. Engrng Comput. 2005;22:664–83.

38. Matthies HG, Zander E, Rosić B, Litvinenko A. Parameter estimation via conditional expectation: a Bayesian inversion. Adv Model Simul Eng Sci. 2016;3(1):1–21. https://doi.org/10.1186/s40323-016-0075-7.

39. Matthies HG, Zander E, Rosić B, Litvinenko A, Pajonk O. *Inverse problems in a Bayesian setting*, Computational Methods for Solids and Fluids—Multiscale Analysis, Probability Aspects and Model Reduction (A Ibrahimbegović, ed) Computational Methods in Applied Sciences, vol. 41, New York, Springer, 2016; p. 245–286, https://doi.org/10.1007/978-3-319-27996-1_10.

40. Matthies HG. Computation of constitutive response, Nonlinear Computational Mechanics-State of the Art (P Wriggers and W Wagner, eds). New York: Springer; 1991.

Matthies *et al. Adv. Model. and Simul. in Eng. Sci.*(2020)7:50

Page 35 of 35

41. Matthies HG, Ibrahimbegović A. Stochastic multiscale coupling of inelastic processes in solid mechanic, Multiscale Modelling and Uncertainty Quantification of Materials and Structures (M Papadrakakis and G Stefanou eds), vol. 3. New York: Springer; 2014. p. 135–57.
42. Maugin GA. The thermomechanics of plasticity and fracture. Cambridge: Cambridge University Press; 1992.
43. Mielke A, Roubiček T. Rate independent systems: theory and application. New York: Springer; 2015.
44. Niekamp R, Markovič D, Ibrahimbegović A, Matthies HG, Taylor RL. Multi-scale modeling of heterogeneous structures with inelastic constitutive behavior: Part II—software coupling implementation aspects. Engrng Computations. 2009;26:6–26.
45. Ostoja-Starzewski M. Towards stochastic continuum thermodynamics. J Non-Equilib Thermodyn. 2002;27:335–48.
46. Rosić B, Matthies HG. Variational theory and computations in stochastic plasticity. Arch Comput Methods Eng. 2015;22:457–509. https://doi.org/10.1007/s11831-014-9116-x.
47. Rosić B, Sarfaraz MS, Matthies HG, Ibrahimbegović A. Stochastic upscaling of random microstructures. PAMM. 2017;17:869–70. https://doi.org/10.1002/pamm.201710401.
48. Sagiyama K, Garikipati K. Machine learning materials physics: deep neural networks trained on elastic free energy data from martensitic microstructures predict homogenized stress fields with high accuracy, 2019. arXiv:1901.00524.
49. Sarfaraz MS, Rosić B, Matthies HG. Stochastic upscaling of heterogeneous materials. PAMM. 2016;16:679–80. https://doi.org/10.1002/pamm.201610328.
50. Sarfaraz SM, Rosić BV, Matthies HG, Ibrahimbegović A. Stochastic upscaling via linear Bayesian updating. Coupled Syst Mech. 2018;7(2):211–32. https://doi.org/10.12989/csm.2018.7.2.211.
51. Sarfaraz SM, Rosić BV, Matthies HG, Ibrahimbegović A. *Stochastic Upscaling via Linear Bayesian Updating*, Multiscale Modeling of Heterogeneous Structures (J. Sorić, P. Wriggers, and O. Allix, eds.), Lecture Notes in Applied and Computational Mechanics, vol. 86, Springer, 2018; pp. 163–181, https://doi.org/10.1007/978-3-319-65463-8_9.
52. Savvas D, Stefanou G. Determination of random material properties of graphene sheets with different types of defects. Composites Part B: Eng. 2018;143:47–54. https://doi.org/10.1016/j.compositesb.2018.01.008.
53. Savvas D, Stefanou G, Papadrakakis M. Determination of RVE size for random composites with local volume fraction variation. Computer Methods Appl Mech Eng. 2016;305:340–58. https://doi.org/10.1016/j.cma.2016.03.002.
54. Schöberl M, Zabaras N, Koutsourelakis P-S. Predictive collective variable discovery with deep Bayesian models, 2018. arXiv:1809.06913.
55. Simo JC, Hughes TJR. Computational Inelasticity, Interdisciplinary Applied Mathematics, vol. 7. New York: Springer; 2006.
56. Sklar A. Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut Statistique de l'Université de Paris. 1959;8:229–31.
57. Staber B, Guilleminot J. Functional approximation and projection of stored energy functions in computational homogenization of hyperelastic materials: A probabilistic perspective. Computer Methods Appl Mech Eng. 2017;313:1–27. https://doi.org/10.1016/j.cma.2016.09.019.
58. Stefanou G, Savvas D, Papadrakakis M. Stochastic finite element analysis of composite structures based on mesoscale random fields of material properties. Computer Methods Appl Mech Eng. 2017;326:319–37. https://doi.org/10.1016/j.cma.2017.08.002.
59. Stefanou G. Simulation of heterogeneous two-phase media using random fields and level sets. Front Strut Civil Eng. 2015;9(2):114–20. https://doi.org/10.1007/s11709-014-0267-5.
60. Stefanou G, Savvas D, Papadrakakis M. Stochastic finite element analysis of composite structures based on material microstructure. Composite Struct. 2015;132:384–92. https://doi.org/10.1016/j.compstruct.2015.05.044.
61. Temizer I. *Lecture Notes in Micromechanics: Analysis of Heterogeneous Materials*, Department of Mechanial Engineering, Bilkent University 06800 Ankara, Turkey, July 2012.
62. Torre E, Marelli S, Emberchts P, Sudret B. A general framework for data-driven uncertainty quantification under complex input dependencies using vine copulas. Probabilistic Eng Mech. 2019;55:1–16.
63. Tran D, Blei DM, Airoldi EM. Variational inference with copula augmentation. In: 29th Conference on Neural Information Processing Systems (NIPS), 2015. https://nips.cc/Conferences/2015.
64. Tran VH. Copula variational Bayes inference via information geometry, 2018. arXiv:1803.10998.
65. Unger JF, Könke C. Coupling of scales in a multiscale simulation using neural networks. Comput Struct. 2008;86(21):1994–2003. https://doi.org/10.1016/j.compstruc.2008.05.004.
66. Unger J, Könke C. An inverse parameter identification procedure assessing the quality of the estimates using Bayesian neural networks. Appl Soft Comput. 2011;11:3357–67. https://doi.org/10.1016/j.asoc.2011.01.007.
67. Xiu D. Numerical methods for stochastic computations: A spectral method approach. Princeton: Princeton University Press; 2010.

Publisher's Note

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