

SPARSE BAYESIAN POLYNOMIAL CHAOS APPROXIMATIONS OF ELASTO-PLASTIC MATERIAL MODELS

Bojana Rosić and Hermann G. Matthies

Institute of Scientific Computing
Technische Universität Braunschweig
Mühlenpfordstrasse 23, 38106 Braunschweig, Germany
e-mail: bojana.rosic@tu-bs.de

Key words: sparse Bayesian inference, sparse polynomial chaos expansion, uncertainty quantification, spectral Kalman filtering, stochastic elastoplasticity, iterative spectral filter

Abstract. In this paper we studied the uncertainty quantification in a functional approximation form of elastoplastic models parameterised by material uncertainties. The problem of estimating the polynomial chaos coefficients is recast in a linear regression form by taking into consideration the possible sparsity of the solution. Departing from the classical optimisation point of view, we take a slightly different path by solving the problem in a Bayesian manner with the help of new spectral based sparse Kalman filter algorithms.

1 INTRODUCTION

Uncertainty quantification currently becomes the focus of many scientific areas, especially engineering ones, due to the presence of aleatoric and epistemic uncertainties in the models describing for example heterogeneous media, loadings, geometry, etc. Due to the rapid development of experimental devices and massive production of sensors, uncertainty quantification is also extensively used in the process of probabilistic solving of inverse problems, i.e. in the prediction step of the Bayesian inference, for example. However, most of the practically used methods are still based on some kind of sampling either in the process of solving the forward problem or estimation itself. The forward propagation of uncertainty usually employs a large number of deterministic software calls. However, real time applications cannot afford this, as the estimation has to be performed under severe time constraints.

The focus of this paper is to show that both stages of Bayesian inference can be resolved in essentially the same manner—by resolving the corresponding inverse problem. Namely, the forward loop of Bayesian inference can be seen as an inverse problem in which the measurement data are given samples. Since the number of samples is usually much smaller than the number of parameters which describe the model response (i.e. random variables or their functional representative — polynomial chaos coefficients), the corresponding problem is ill-posed and thus has to be regularised. This can be done in a Bayesian setting similar to the process of estimation of model parameters given real measurement data. Such an approach turns out to be the generalisation of ℓ_1 and ℓ_2 -norm optimisation problems by taking the appropriate priors on the approximation coefficients. However, the sparsity of the solution appearing in the ℓ_1 minimisation in the Bayesian point of view turns out to be computationally difficult, because the distributions in Bayes rule are not conjugate. To resolve this, the sparsity priors are usually assumed in a hierarchical setting such as the normal prior with Gamma distributed hyperparameters used in relevance vector machine approaches [5]. On the other hand, the stochastic search can be also done in a numerical Markov chain Monte Carlo (MCMC) setting [6], however, this can lead to high computational costs. Instead, in this paper the conditional expectation setting of Kolmogorov is used in order to estimate the unknown polynomial chaos coefficients of the solution. The methods were recently developed by the authors and formulated in a purely algebraic setting [1, 3]. Here, the approaches are extended to include the sparsity of the solution.

The paper is organised as follows: in Section 2 the model problem is introduced, Section 3 is discusses its functional approximation. Section 4 presents Bayesian regression, and the new approach is presented in Section 5. Finally, numerical results are depicted in Section 6.

2 MODEL PROBLEM

Let $(\Omega_\theta, \mathcal{B}_\theta, \mathbb{P}_\theta)$ be a probability space, in which Ω_θ denotes the space of all events, \mathcal{B}_θ is a σ -algebra of subsets of Ω_θ , and \mathbb{P}_θ is a probability measure. In the presence of material uncertainties, here assumed to have finite variance and belonging to $L_2(\Omega_\theta)$, the elastoplastic material model is described by an uncertain infinitesimal elastoplastic state $w(\omega) := (u(\omega), \epsilon_p(\omega), \eta(\omega))$ in which u denotes the displacements, ϵ_p is the plastic strain, and η is an internal variable. The state satisfies the equilibrium equation \mathbb{P}_θ -almost surely, i.e.:

$$\begin{aligned} -\operatorname{div} \sigma(x, \omega) &= f(x, \omega) \quad \forall x \in \mathcal{G}_t, \\ \sigma(x, \omega) \cdot n(x, \omega) &= \sigma_N(x, \omega) \quad \forall x \in \Gamma_N, \\ u(x, \omega) &= u_0(x, \omega) \quad \forall x \in \Gamma_D, \end{aligned} \tag{1}$$

with appropriate Dirichlet and Neumann boundary conditions implied on parts Γ_D and Γ_N of the piecewise smooth Lipschitz continuous boundary $\Gamma = \partial\mathcal{G}$ such that $\Gamma_D \subseteq \partial\mathcal{G}$ and $\Gamma_N \subset \partial\mathcal{G}$, respectively. For reasons of simplicity, the last ones are assumed to be deterministic.

The constitutive law for the elastic part is assumed to be of Hooke's type and is described by an uncertain isotropic homogeneous $C(\omega)$ constitutive tensor modelled via bulk $\kappa(\omega)$ and shear $G(\omega)$ moduli taken as independent positive definite lognormally distributed random variables. Finally, the rate of the plastic strain is assumed to follow the associative plastic flow rule $\dot{E}_p(x, \omega) \in N_{\mathcal{K}}(\Sigma(x, \omega))$ in which $N_{\mathcal{K}}(\Sigma(x, \omega))$ is the normal cone on the convex set of admissible stresses $\mathcal{K}(\omega) = \{\Sigma(\omega) \mid \phi_{\mathcal{K}}(\Sigma(\omega)) \leq 0 \text{ } \mathbb{P}_{\theta} \text{ a.s.}\}$ described by the von Mises yield function $\phi_{\mathcal{K}}$ with uncertain yield stress $\sigma_y(\omega)$ and the hardening variables $h(\omega)$ as arguments. Here, $E_p := (\varepsilon_p, \eta)$ denotes the generalised plastic strain, and $\Sigma := (\sigma, \chi)$ stands for the generalised stress.

For computational purposes the problem given in Eq. (1) is rewritten on weak form following discussions in [2]. The goal is to estimate the state $w \in H^1(\mathcal{T}, \mathcal{Z})$ with $w(0) = 0$, its dual $w^* \in H^1(\mathcal{T}, \mathcal{Z}^*)$, $w^*(0) = 0$ and $\dot{w} \in \mathcal{K}^\infty$ such that

$$a(w(t), z) + \langle\langle \dot{w}(t), z \rangle\rangle = \langle\langle f, z \rangle\rangle \quad (2)$$

for all $z = (v, (\mu, \nu)) \in \mathcal{Z}$ and

$$\langle\langle \dot{w}, z^* - w^* \rangle\rangle \leq 0, \quad \forall z^* \in \mathcal{K} \subset \mathcal{Z}^*. \quad (3)$$

hold a.s. in Ω_θ and a.e. in \mathcal{T} . Here, and $\langle\langle \cdot, \cdot \rangle\rangle$ is the duality pairing given in terms of the mathematical expectation $\mathbb{E}(\langle\langle \cdot, \cdot \rangle\rangle) = \int_{\Omega_\theta} \langle\langle \cdot, \cdot \rangle\rangle \mathbb{P}_\theta(d\omega)$, and $\langle s_1, s_2 \rangle = \int_{\mathcal{G}} s_1 s_2 dx$. The existence and uniqueness are already studied by authors in [2], and will be not be discussed here. After time (by the implicit Euler) and spatial finite element discretisation, the formulation in Eq. (2)-Eq. (3) reduces to a nonlinear stochastic residual equation to be solved globally for the increment of the displacement $\Delta \mathbf{u}_n(\omega)$, and the first order variational inequality which corresponds to the constrained stochastic quadratic convex optimisation problem (the so-called closest point projection scheme) to be solved locally in each integration point of the finite element scheme and the stochastic space for the increments of the strain-like $\Delta \mathbf{E}_{pn}$ and the stress like $\Delta \Sigma_n$ variables. These algorithms are very well known in the classical deterministic setting [9], albeit, their extension to the stochastic counterpart is not an easy task, see [2].

In this paper we would like to keep the deterministic algorithms per se, and to use them to estimate the statistics of random variables (fields) $\Delta \mathbf{w}_n(\omega)$ and $\Delta \mathbf{w}_n^*(\omega)$, i.e. moments

$$\mathbb{E}(\Delta \mathbf{w}_n^m) = \int_{\Omega_\theta} (\Delta \mathbf{w}_n(\omega))^m \mathbb{P}_\theta(d\omega), \quad \mathbb{E}((\Delta \mathbf{w}_n^*)^m) = \int_{\Omega_\theta} (\Delta \mathbf{w}_n^*(\omega))^m \mathbb{P}_\theta(d\omega) \quad (4)$$

non-intrusively. This matches with the high-dimensional numerical integration and the Monte Carlo type of algorithms. Due to their slow convergence rates, a large number of evaluation points (i.e. deterministic executions of the finite element code) are needed to achieve the desired accuracy. Therefore, to reduce the computational cost, here we consider the functional approximation algorithms as further described in the text.

3 FUNCTIONAL APPROXIMATION

Instead of integrating the functions $\mathbf{y} := \{\Delta \mathbf{w}_n, \Delta \mathbf{w}_n^*\}$ directly, one may approximate the integrand in Eq. (4) by some known elementary functions, the integration of which is algebraically computable. A typical example is the generalised polynomial chaos expansion

$$y(x, \omega) = \sum_{\alpha \in \mathcal{J}} y^{(\alpha)}(x) \Psi_{\alpha}(\boldsymbol{\theta}(\omega)) \quad (5)$$

in which Ψ_{α} are the multi-variate polynomials with the standard random variables $\boldsymbol{\theta}(\omega)$ as arguments, and x denotes the spatial position (i.e. the finite element node for displacements or the Gauss integration point for stress- and strain-like variables). Other kinds of approximation functions can be also used, however this will not be further discussed here. The random variables $\boldsymbol{\theta}(\omega)$ represent the parameterisation of existing uncertainties in model parameters. They are usually taken as independent, uncorrelated random variables of a simpler kind such as normal or uniform random variables corresponding to the Askey scheme as discussed in [10].

Given N samples of $y(x, \omega)$ one may rewrite Eq. (5) as a linear system of equations

$$y(x, \omega_i) = \sum_{\alpha \in \mathcal{J}} y^{(\alpha)}(x) \Psi_{\alpha}(\boldsymbol{\theta}(\omega_i)), \quad i = 1, \dots, N \quad (6)$$

with $y^{(\alpha)}$ being unknown coefficients. Denoting $\mathbf{s} := [y(x, \omega_i)] \in \mathbb{R}^N$, $\boldsymbol{\Psi} := [\Psi_{\alpha}(\boldsymbol{\theta}(\omega_i))] \in \mathbb{R}^{N \times P}$ and $\mathbf{v} := [y^{(\alpha)}(x)] \in \mathbb{R}^P$, one may rewrite the previous equation in a matrix-vector form

$$\mathbf{s} = \boldsymbol{\Psi} \mathbf{v} \quad (7)$$

which is equivalent to the more robust projected version

$$\mathbf{d} := \boldsymbol{\Psi}^T \mathbf{u} = \boldsymbol{\Psi}^T \boldsymbol{\Psi} \mathbf{v} =: \mathbf{W} \mathbf{v}. \quad (8)$$

The system in the previous equation or in Eq. (7) is quite often depicted as underdetermined, especially when one deals with very expensive solvers, i.e. finely discretised problems, for which $N < P$ with P being the cardinality of the polynomial expansion in Eq. (5). To tackle this problem, different kinds of regularisation

procedures are used in the literature, the most popular among which are the regularised least square (i.e. the Tikhonov regularisation)

$$\mathbf{v} = \arg \min \left(\frac{1}{2} \|\mathbf{W}\mathbf{v} - \mathbf{d}\|_2^2 + \frac{\lambda}{2} \|\mathbf{v}\|_2^2 \right) \quad (9)$$

and the basis pursuit denoising

$$\mathbf{v} = \arg \min \left(\frac{1}{2} \|\mathbf{W}\mathbf{v} - \mathbf{d}\|_2^2 + \frac{\lambda}{2} \|\mathbf{v}\|_1 \right) \quad (10)$$

methods, also known as ℓ_2 and ℓ_1 minimisation procedures. These consist of the squared error part used to enforce closeness of \mathbf{v} to the data, and the regularisation term enforcing the smoothness of \mathbf{v} . To balance these two terms, the regularisation parameter λ is used. However, in general the regularisation parameter λ representing the noise variance is known to be an uneducated guess, and it is difficult to find the most optimal value. If $\lambda = 0$ both of problems are equivalent and correspond to the classical least squares procedure. If $\lambda > 0$, then the ℓ_1 minimisation is preferable here compared to the ℓ_2 minimisation as it promotes the sparsity of the solution. On the other hand, in a computational setting the ℓ_2 problem is easier to solve as the solution \mathbf{v} is linear in the data \mathbf{b} in contrast to the ℓ_1 minimisation. The objective function in Eq. (10) is convex but not differentiable and thus requires special methods such as subgradient methods [8]. In computational practice the problem in Eq. (10) is transformed to the quadratic convex optimisation with linear inequality constraints, which can be solved by interior point methods. On the other hand, in order to be able to recover the sparse solution, the sensing matrix \mathbf{W} in Eq. (10) has to satisfy the so-called restricted isometry property [11], which is usually not the case. To ensure this, the principle of random projections [12] is used, such that the \mathbf{d} in Eq. (8) is projected onto a basis that consists of random linear combination of basis functions in Ψ , i.e. the problem given in Eq. (8) is rewritten to

$$\mathbf{b} := \mathbf{W}\mathbf{d} = \mathbf{W}\Psi^T\Psi\mathbf{v} = \mathbf{A}\mathbf{v} \quad (11)$$

in which \mathbf{W} denotes the carefully chosen random sensing matrix. This problem will be considered further instead of the one in Eq. (8).

4 BAYESIAN REGRESSION

In this paper the generalisation of the regularisation approach will be considered. The method relies on the probabilistic view of Eq. (8), in which the coefficients are assumed to be unknown, and hence priorly modelled as independent random variables \mathbf{v}_f in $(\Omega_\xi, \mathcal{B}_\xi, \mathbb{P}_\xi)$, with the joint probability density function (pdf) given as

$$p_f(\mathbf{v}) = \prod_{\alpha \in \mathcal{J}} p(v^{(\alpha)}) \quad (12)$$

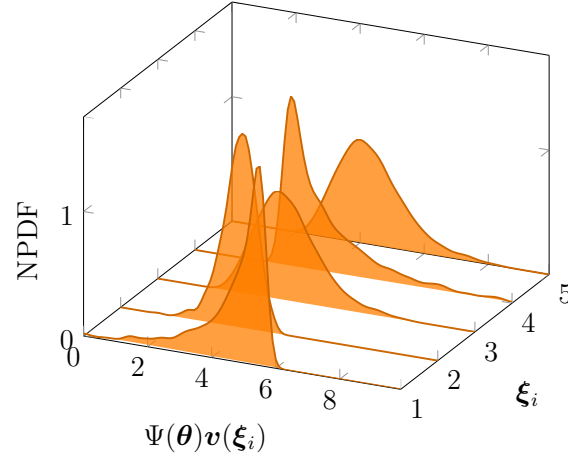


Figure 1: The stochastic version of the polynomial chaos expansion. Each realisation represents the pdf of the expansion given one realisation of the coefficient

with $p(v^{(\alpha)})$ being the pdf of individual parameters. Following this, the linear system in Eq. (8) becomes uncertain and is described by prediction:

$$\mathbf{b}_f(\xi) = \mathbf{A}\mathbf{v}_f(\xi). \quad (13)$$

This can be seen in Fig. 1, in which the schematic representation of the stochastic version of the polynomial chaos expansion from Eq. (8) is depicted.

The pdf of the coefficients can be further updated given data via Bayes's rule

$$\pi(\mathbf{v}|\mathbf{b}) \sim p(\mathbf{b}|\mathbf{v})p_f(\mathbf{v}) \quad (14)$$

in which $\pi(\mathbf{v}|\mathbf{b})$ denotes the posterior density, $p(\mathbf{b}|\mathbf{v})$ corresponds to the likelihood, and $p_f(\mathbf{v})$ is the prior. Assuming that the prior is normally distributed

$$p_f(\mathbf{v}) \sim \exp\left(-\frac{1}{2}\|\mathbf{v}\|_2^2\right) \quad (15)$$

as well as the likelihood, the posterior pdf obtains the form

$$\pi(\mathbf{v}|\mathbf{b}) \sim \exp\left(-\frac{1}{2}\|\mathbf{A}\mathbf{v} - \mathbf{b}\|_2^2\right) \exp\left(-\frac{1}{2}\|\mathbf{v}\|_2^2\right) \quad (16)$$

Its maximum a posteriori (MAP) estimate is the minimiser of the objective function given in Eq. (9). Under the same assumptions, only taking the prior to follow the Laplace distribution

$$p_f(\mathbf{v}) \sim \exp\left(-\frac{1}{2}\|\mathbf{v}\|_1\right), \quad (17)$$

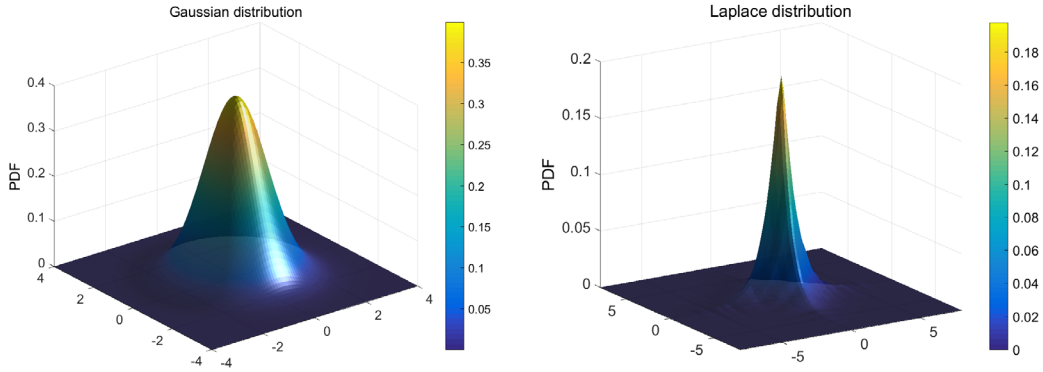


Figure 2: The Gaussian (left) and Laplace (right) probability distributions

one may show that the posterior MAP estimate is the minimiser of the objective function given in Eq. (10).

Similar to the deterministic setting, the process of computing the posterior distribution given the Gaussian prior is easier than in the Laplace case. The reason is that in the former case the prior and the likelihood are conjugate, whereas in the latter case they are not. Hence, the posterior corresponding to the ℓ_1 minimisation cannot be estimated algebraically, but only using some of the sampling based approaches such as the Markov chain Monte Carlo algorithm. To avoid this, one may consider a hierarchical type of prior which mimics the Laplace behaviour, but it is easier to evaluate. Following [5], one may model the polynomial chaos coefficients by normal distribution

$$p(\mathbf{v}|\mathbf{w}) = \prod_{\alpha} \mathcal{N}(0, w_{\alpha}^{-1}) \quad (18)$$

with zero mean and the precision (inverse variance) \mathbf{w} that follows the Gamma distribution. The posterior distribution is then represented as

$$p(\mathbf{v}, \mathbf{w}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{v}, \mathbf{w})p(\mathbf{v}|\mathbf{w})p(\mathbf{w}) \quad (19)$$

and cannot be computed analytically. Therefore, the posterior is re-written as

$$p(\mathbf{v}, \mathbf{w}|\mathbf{y}) = p(\mathbf{v}|\mathbf{y}, \mathbf{w})p(\mathbf{w}|\mathbf{y}) \quad (20)$$

in which the first term $p(\mathbf{v}|\mathbf{y}, \mathbf{w})$ follows the normal posterior distribution, whereas the second term is approximated by the delta function at its mode. The latter corresponds to the maximisation of the marginal likelihood as presented in [5]. However, this kind of approach relies on the hierarchical estimation in which the marginalisation over the hyperparameters have to be introduced, which in this paper will be avoided.

5 COMPRESSIVE SENSING SPECTRAL KALMAN FILTER

To allow for the algebraic evaluation of the posterior, this paper considers a more fundamental Kolmogorov's approach to estimation. The method is based on the definition of the conditional expectation as a projection onto the subspace generated by the σ -algebra of data:

$$\mathbb{E}(\mathbf{v}|\mathbf{b}) = P_{\sigma(\mathbf{b})}\mathbf{v} \quad (21)$$

Being an orthogonal projection, the conditional expectation matches the minimum mean square estimate [1, 3]

$$\min \mathbb{E}(\|\mathbf{v} - \mathbb{E}(\mathbf{v}|\sigma(\mathbf{b}))\|^2) \quad (22)$$

which according to [1] implies an orthogonal decomposition

$$\mathbf{v} = P_{\sigma(\mathbf{b})}\mathbf{v} + (I - P_{\sigma(\mathbf{b})})\mathbf{v}. \quad (23)$$

However, as $P_{\sigma(\mathbf{b})}\mathbf{v}$ is difficult to compute directly, one may further refer to the Doob-Dynkin lemma and search for an optimal map, i.e. a measurable function φ such that

$$\mathbb{E}(\mathbf{v}|\sigma(\mathbf{b})) = P_{\sigma(\mathbf{b})}\mathbf{v} = \varphi(\mathbf{b})$$

holds. Therefore, Eq. (23) becomes

$$\mathbf{v} = \varphi(\mathbf{b}) + (\mathbf{v} - \varphi(\mathbf{b})). \quad (24)$$

The first term is the projection and is altered by data, whereas the remaining orthogonal part stays unchanged, i.e. described by our prior knowledge. This finally leads to the filtering formula:

$$\mathbf{v}_a(\xi) = \mathbf{v}_f(\xi) + (\varphi(\mathbf{b}) - \varphi(\mathbf{b}_f(\xi))). \quad (25)$$

Assuming that the optimal map φ is linear, the formula reduces to the so-called Gauss-Markov-Kalman filter

$$\mathbf{v}_a(\xi) = \mathbf{v}_f(\xi) + \mathbf{K}(\mathbf{b} - \mathbf{b}_f(\xi)), \quad (26)$$

a generalisation of the classical Kalman filter form. For more details please see [1, 3]. Here, the factor \mathbf{K} is the Kalman gain

$$\mathbf{K} = \text{cov}_{\mathbf{v}_f, \mathbf{b}_f} (\text{cov}_{\mathbf{b}_f} + \text{cov}_{\varepsilon})^\dagger \quad (27)$$

with \dagger denoting the pseudo-inverse, and the covariance functions defined as

$$\text{cov}_{q,y} := \mathbb{E}((q - \mathbb{E}(q)) \otimes (y - \mathbb{E}(y))). \quad (28)$$

The modelling error ε is here introduced as an additional term representing the error of the truncated approximations in Eq. (11).

The linear filter as presented previously will not be optimal in the nonlinear case. To better account for nonlinearity, one may use higher order polynomial approximations as in [1], or turn to the iterative version of Eq. (26). The latter one approximates the nonlinear measurement operator $Y(\mathbf{v})$ in $\mathbf{b} := Y(\mathbf{v}) + \varepsilon$ by [4]

$$Y_\lambda(\mathbf{v}) = \mathbf{M}(\mathbf{v} - \tilde{\mathbf{v}}) + \mathbf{a} = M\tilde{\mathbf{v}} + \mathbf{a} \quad (29)$$

in which $\tilde{\mathbf{v}} := \mathbf{v} - \check{\mathbf{v}}$ (also known as the fluctuating part of the random variable \mathbf{v} when $\check{\mathbf{v}} := \mathbb{E}(\mathbf{v})$) and \mathbf{M} is the linear measurement matrix. Then, following [4], one may design the iterative formula:

$$\mathbf{v}_a^{(i+1)} = \mathbf{v}_f + \mathbf{K}_\lambda^{(i)}(\mathbf{b} - \mathbf{a}^{(i)} - \mathbf{M}^{(i)}(\mathbf{v}_f - \check{\mathbf{v}}^{(i)}) - \varepsilon). \quad (30)$$

Here, $\mathbf{M}^{(i)}$ and $\mathbf{a}^{(i)}$ denote either the exact Jacobian and $\mathbf{a} := Y(\mathbb{E}(\mathbf{v}_f))$, or the inexact Jacobian and $\mathbf{a} := \mathbb{E}(Y(\mathbf{v}_f))$ in case of an unbiased estimate [4].

The previously described filtering formulas become especially interesting when considered in the functional approximation setting. Instead of sampling, the random variables of interest in Eq. (26) or Eq. (30) can be represented by the polynomial chaos approximations similar to those given in Eq. (5). This leads to the purely deterministic (algebraic) filtering formula:

$$\sum_{\beta \in \mathcal{I}} \mathbf{v}_a^{(\beta)} \Gamma_\beta(\xi) = \sum_{\beta \in \mathcal{I}} \mathbf{v}_f^{(\beta)} \Gamma_\beta(\xi) + \mathbf{K} \left(\sum_{\beta \in \mathcal{I}} \mathbf{b}^{(\beta)} \Gamma_\beta(\xi) - \sum_{\beta \in \mathcal{I}} \mathbf{b}_f^{(\beta)} \Gamma_\beta(\xi) \right) \quad (31)$$

in which Γ denote the polynomials corresponding to the distribution on the polynomial chaos coefficients, and are not necessarily same as Ψ . Note that the first term in the innovation part corresponds to the deterministic measurements, and hence has only non-zero mean. By projecting the formula onto the polynomial basis Γ_β one obtains

$$\mathbf{v}_a = \mathbf{v}_b + \mathbf{K}(\mathbf{b} - \mathbf{b}_f) \quad (32)$$

in which $\mathbf{v}_f := [\mathbf{v}_f^{(\beta)}]_{\beta \in \mathcal{I}} = [v_f^{(\alpha, \beta)}]_{\alpha \in \mathcal{J}, \beta \in \mathcal{I}}$. Similarly, the Kalman gain \mathbf{K} can be computed using the algebraic expression for the covariance matrix

$$\mathbf{C}_{v_f} = \mathbb{E}_\xi((\hat{\mathbf{v}}_f - \bar{\mathbf{v}}_f) \otimes (\hat{\mathbf{v}}_f - \bar{\mathbf{v}}_f)) = \sum_{\alpha, \beta \in \mathcal{J}_p} \mathbb{E}_\xi(\Gamma_\alpha \Gamma_\beta) \mathbf{v}_f^{(\alpha)} \otimes \mathbf{v}_f^{(\beta)} - \bar{\mathbf{v}}_f \otimes \bar{\mathbf{v}}_f \quad (33)$$

in which $\bar{\mathbf{v}}_f := \mathbb{E}_\xi(\mathbf{v}_f)$. The last relation can be further rewritten in a matrix form as

$$\mathbf{C}_{v_f} = \tilde{\mathbf{V}}_f \Delta \tilde{\mathbf{V}}_f^T \quad (34)$$

in which $(\mathbf{\Delta})_{\alpha\beta} = \mathbb{E}_\xi(\Gamma_\alpha\Gamma_\beta) = \text{diag}(\alpha!)$ and $\tilde{\mathbf{V}}_f$ is equal to $\mathbf{v}_f := (\dots, \mathbf{v}_f^{(\alpha)}, \dots)^T$ without the mean part. In a similar manner one derives algebraic formula for the iterative filter.

The filter in Eq. (32) does not lead to the sparse solution, as only the ℓ_2 minimisation is performed. To allow for the sparsity of the solution, one solves

$$\begin{aligned} \min \quad & \mathbb{E}(\|\mathbf{v} - \mathbb{E}(\mathbf{v}|\sigma(\mathbf{b}))\|^2) \\ \text{such that} \quad & \|\mathbb{E}(\mathbf{v}|\sigma(\mathbf{b}))\|_1 \leq \epsilon \end{aligned} \quad (35)$$

The inequality in Eq. (35) is nonlinear, and its subgradient can be rewritten as the pseudo-measurement equation

$$\mathbf{Z}(\mathbf{v}) := H(\mathbf{v})\mathbf{v} - \epsilon = 0 \quad (36)$$

instead of Eq. (22). Here, $H(\mathbf{v}) := \text{sign}(\mathbf{v})$, and ϵ is the given tolerance with the covariance \mathbf{C}_ϵ chosen as the regularisation parameter. The computational algorithm consists of a sequential estimation in which the first update is obtained by using the real measurement and Eq. (26), and the second by using the pseudo-one and the iterative formula in Eq. (30).

6 NUMERICAL RESULTS AND CONCLUSIONS

The algorithm as described previously is tested on Cook's membrane benchmark problem in five loading steps (two of which are plastic). The three material parameters (bulk and shear moduli, as well as yield stress) are modelled as independent lognormally distributed random variables, and the response is approximated by Hermite polynomial chaos expansion of fourth order leading to 35 polynomial coefficients. The approximation is computed by using 15 random evaluation points, and its accuracy is tested against the result obtained given $1e4$ Monte Carlo simulations. As depicted in Fig. 3, for the last time increment the error in both stress and strain variables is low. Hence, the Bayesian method can catch the sparsity of the solution. However, the error in strain-like variables is slightly higher. The reason lies in the polynomial chaos basis which does not change in time. This can be seen in Fig. 4, where clearly the approximation error increase with time.

According to the previous results, the suggested methods seem to be promising, and they will be further analysed with respect to the adaptivity of the polynomial chaos scheme.

Acknowledgment: This project is partially funded by DFG.

REFERENCES

- [1] H. G. Matthies, E. Zander, B. Rosić and A. Litvinenko. Parameter Estimation via Conditional Expectation: A Bayesian Inversion, *Advanced Modeling and Simulation in Engineering Sciences*, available online, 2016

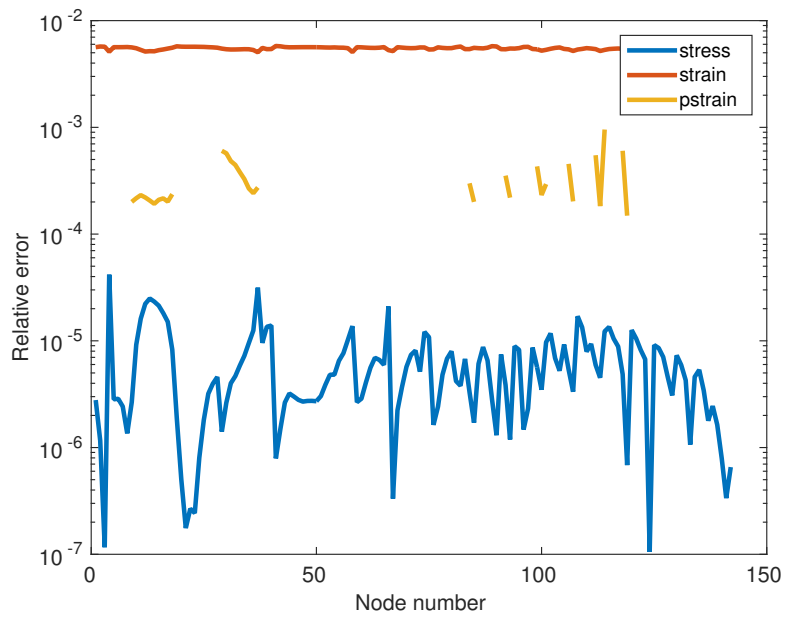


Figure 3: Relative error of approximated solutions across the finite element nodes

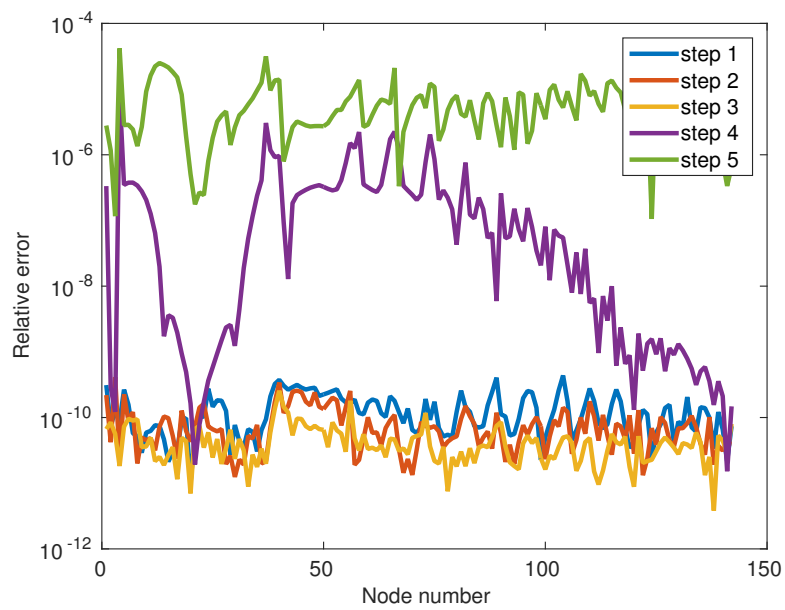


Figure 4: The change of relative error of approximated stress across the finite element nodes and in time

- [2] B. Rosić and H. G. Matthies. Variational Theory and Computations in Stochastic Plasticity. *Archives of Computational Methods in Engineering*, 22(3): 457-509, 2015
- [3] B. Rosić, A. Kučerová, J. Sýkora, O. Pajonk, A. Litvinenko and H. G. Matthies. Parameter Identification in a Probabilistic Setting. *Engineering Structures*, 50: 179–196, 2013
- [4] B. Rosić. A spectral iterative filter based on conditional expectation for non-linear parameter estimation, submitted, 2017
- [5] M. E. Tipping. Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*. 1: 211–244, 2001
- [6] X. Chen, Z. Jane Wang, M. J. McKeown. A Bayesian Lasso via reversible-jump MCMC. *Signal Processing*, 91:1920–1932, 2011
- [7] K. Koh, S.-J. Kim, and S. Boyd. An Interior-Point Method for Large-Scale l_1 -Regularized Logistic Regression. *Journal of Machine Learning Research*, 8:1519-1555, 2007.
- [8] D. A. Lorenz, M. E. Pfetsch and A. M. Tillmann. Solving Basis Pursuit: Subgradient Algorithm, Heuristic Optimality Check, and Solver Comparison. *ACM Transactions on Mathematical Software*, 41(2), 2015
- [9] J. C. Simo and T. R. Hughes. *Computational Inelasticity*. Springer, 1998
- [10] D. Xiu and G. Em Karniadakis. The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM J. Sci. Comput.*, 24(2), 619–644, 2002
- [11] A. M. Tillmann and M. E. Pfetsch, The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing,” *IEEE Trans. Inf. Th.*, 60(2): 1248–1259 (2014)
- [12] D. L. Donoho, Compressed sensing, *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.