

# Parameter Identification in a Probabilistic Setting

Bojana V. Rosić<sup>a,b</sup>, Anna Kučerová<sup>c</sup>, Jan Sýkora<sup>c</sup>, Oliver Pajonk<sup>d,a</sup>,  
Alexander Litvinenko<sup>a</sup>, and Hermann G. Matthies<sup>a</sup>

<sup>a</sup>Technische Universität Braunschweig

[wire@tu-bs.de](mailto:wire@tu-bs.de)

<sup>b</sup>University of Kragujevac

<sup>c</sup>Czech Technical University in Prague

<sup>d</sup>SPT Group Hamburg

December 2011

## Abstract

Parameter identification problems are formulated in a probabilistic language, where the randomness reflects the uncertainty about the knowledge of the true values. This setting allows conceptually easily to incorporate new information, e.g. through a measurement, by connecting it to Bayes's theorem. The unknown quantity is modelled as a (may be high-dimensional) random variable. Such a description has two constituents, the measurable function and the measure. One group of methods is identified as updating the measure, the other group changes the measurable function. We connect both groups with the relatively recent methods of functional approximation of stochastic problems, and introduce especially in combination with the second group of methods a new procedure which does not need any sampling, hence works completely deterministically. It also seems to be the fastest and more reliable when compared with other methods. We show by example that it also works for highly nonlinear non-smooth problems with non-Gaussian measures.

**Keywords:** parameter identification, Bayesian update, linear Bayes, Kalman filter, polynomial chaos

## Classification:

(MSC2010) 62F15, 65N21, 62P30, 60H15, 60H25, 74G75, 80A23, 74C05

(PACS2010) 46.65.+g, 46.35.+z, 44.10.+i

(ACM1998) G.1.8, G.3, J.2

## 1 Introduction

In trying to predict the behaviour of physical systems, one is often confronted with the fact that although one has a mathematical model of the system which carries some confidence as to its fidelity, some quantities which characterise the system may only be incompletely known, or in other words they are uncertain. See [27] for a synopsis on our approach to such parametric problems.

Here we want to identify these parameters through observations or measurement of the response of the system. Such an identification can be approached in different ways. One way is to measure the difference between observed and predicted system output and try to find parameters such that this difference is minimised, this optimisation approach leads to regularisation procedures [7].

Here we take the view that our lack of knowledge or uncertainty of the actual value of the parameters can be described in a *Bayesian* way through a probabilistic model [14, 38]. The unknown parameter is then modelled as a random variable—also called

the *prior* model—and additional information on the system through measurement or observation changes the probabilistic description to the so-called *posterior* model. The second approach is thus a method to update the probabilistic description in such a way as to take account of the additional information.

To be more specific, let us consider the following situation: we are investigating some physical system which is modelled by an evolution equation for its state:

$$\frac{\partial}{\partial t}u(t) + A(p; u(t)) = f(p; t), \quad (1)$$

where  $u(t) \in \mathcal{U}$  describes the state of the system at time  $t \in [0, T]$  lying in a Hilbert space  $\mathcal{U}$  (for the sake of simplicity),  $A$  is an operator modelling the physics of the system, and  $f \in \mathcal{U}^*$  is some external influence (action / excitation / loading). The model depends on some parameters  $p \in \mathcal{P}$ , and by  $q$  we denote that component of the parameters  $p$  which we are uncertain about and would thus like to identify the actual value.

Now assume that we observe a function of the state  $Y(u(q), q)$ , and from this observation we would like to identify the corresponding  $q$ . This is called the *inverse* problem, and as the mapping  $q \mapsto Y(q)$  is usually not invertible, the inverse problem is *ill-posed*. Embedding this problem of finding the best  $q$  in a larger class by modelling our knowledge about it with the help of probability theory, then in a Bayesian manner the task becomes to estimate conditional expectations, e.g. see [14, 38] and the references therein. The problem now is *well-posed*, but at the price of ‘only’ obtaining probability distributions on the possible values of  $q$ , which now is modelled as a  $\mathcal{Q}$ -valued random variable (RV). Predicting what the measurement  $Y(q)$  should be from some assumed  $q$  is computing the *forward* problem. The inverse problem is then approached by comparing the forecast from the forward problem with the actual information.

Since the parameters of the model to be estimated are uncertain, all relevant information may be obtained via their stochastic description. In order to extract information from the posterior most estimates take the form of expectations w.r.t. the posterior. These expectations—mathematically integrals, numerically to be evaluated by some quadrature rule—may be computed via asymptotic, deterministic or sampling methods. In our review of current work we follow our recent reports [35, 31].

One often used technique is a Markov chain Monte Carlo (MCMC) method [21, 10], constructed such that the asymptotic distribution of the Markov chain is the Bayesian posterior distribution. This can be then sampled by letting the Markov chain run for a sufficiently long time, although the samples are not independent in this case. With the intention of accelerating the MCMC method some authors [22, 16, 32, 17] have introduced stochastic spectral methods into the computation. Expanding the prior random process into a polynomial chaos (PCE) or a Karhunen-Loève expansion (KLE) (e.g. [24]), the inverse problem becomes an inference on the weights of the Karhunen-Loève modes. Pence et al. [32] combine polynomial chaos theory with maximum likelihood estimation, where the parameter estimates are calculated in a recursive or iterative manner. Christen and Fox [6] have applied a local linearisation of the forward model to improve the acceptance probability of proposed moves, while in [2, 20, 18, 39] collocation methods are employed as a more efficient sampling technique.

The approaches mentioned above require a large number of samples in order to obtain satisfactory results. While showing some results in this direction, the main idea here is to do the Bayesian update directly on the polynomial chaos expansion (PCE) without any sampling [31, 35, 27]. This idea has appeared independently in [3] in a simpler context, whereas in [36] it appears as a variant of the Kalman filter (e.g. [15]). A PCE for a push-forward of the posterior measure is constructed in [28].

From this short overview it becomes apparent that the update may be seen abstractly in two different ways. Regarding the uncertain parameters

$$q : \Omega \rightarrow \mathcal{Q} \text{ as a RV on a probability space } (\Omega, \mathfrak{A}, \mathbb{P}) \quad (2)$$

where the set of elementary events is  $\Omega$ ,  $\mathfrak{A}$  a  $\sigma$ -algebra of events, and  $\mathbb{P}$  a probability measure, one set of methods performs the update by changing the probability measure  $\mathbb{P}$  and leaving the mapping  $q(\omega)$  as it is, whereas the other set of methods leaves the probability measure unchanged and updates the function  $q(\omega)$ . In any case, the push forward measure  $q_*\mathbb{P}$  on  $\mathcal{Q}$  is changed from prior to posterior.

The organisation of the paper is as follows: in Section 2 we review the Bayesian update and recall the link between the conditional measure and conditional expectation. This allows to recover a Kalman filter like update for the RV via the conditional expectation. In the following Section 3 different ways of numerically computing the Bayesian update are indicated. The resolution of the forward problem in the context of stochastic discretisation procedures is outlined in Section 4. The numerical examples for such parameter identification procedures, contained in Section 5, are representative of some linear and nonlinear problems in structural and continuum mechanics.

## 2 Bayesian Updating

In the setting of Eq. (1) let us pose the following problem: Some components—let us denote these by  $q$ —of the parameters  $p \in \mathcal{P}$  are uncertain. To be more specific, assume that  $q \in \mathcal{Q}$  are elements of some vector space. By making observations  $z_k$  at times  $0 < t_1 < \dots < t_k \dots \in [0, T]$  one would like to infer what they are. But we can not observe the entity  $q$  directly—like in Plato’s cave allegory we can only see a ‘shadow’ of it, formally given by a ‘measurement operator’

$$Y : \mathcal{Q} \times \mathcal{U} \ni (q, u(t_k)) \mapsto y_k = Y(q; u(t_k)) \in \mathcal{Y}; \quad (3)$$

at least this is our model of what we are measuring. Frequently the space  $\mathcal{Y}$  may be regarded as finite dimensional, as one can only observe a finite number of quantities. Usually the observation will deviate from what we expect to observe even if we knew the right  $q$  as Eq. (1) is only a *model*—so there is some model error  $\epsilon$ , and the measurement will be polluted by some measurement error  $\varepsilon$ . Hence we observe  $z_k = y_k + \epsilon + \varepsilon$ . From this one would like to know what  $q$  and  $u(t_k)$  are. For the sake of simplicity we will only consider one error term  $z_k = y_k + \varepsilon$  which subsumes all the errors.

The mapping in Eq. (3) is usually not invertible and hence the problem is called ill-posed. One way to address this is via regularisation (see e.g. [7]), but here we follow a different track. Modelling our lack-of-knowledge about  $q$  and  $u(t_k)$  in a Bayesian way [38] by replacing them with a  $\mathcal{Q}$ - resp.  $\mathcal{U}$ -valued random variable (RV), the problem becomes well-posed [37]. But of course one is looking now at the problem of finding a probability distribution that best fits the data; and one also obtains a probability distribution, not just *one* pair  $q$  and  $u(t_k)$ . Here we focus on the use of a linear Bayesian approach [11] in the framework of ‘white noise’ analysis.

The mathematical setup then is as follows: we assume that  $\Omega$  is a measure space with  $\sigma$ -algebra  $\mathfrak{A}$  and with a probability measure  $\mathbb{P}$ , and that  $q : \Omega \rightarrow \mathcal{Q}$  and  $u : \Omega \rightarrow \mathcal{U}$  are random variables. For simplicity, we shall also require  $\mathcal{Q}$  to be a Hilbert space where each vector is a possible realisation. This is in order to allow to measure the distance between different  $q$ ’s as the norm of their difference, and to allow the operations of linear algebra to be performed.

### 2.1 Bayesian updating of the measure

Bayes’s theorem is commonly accepted as a consistent way to incorporate new knowledge into a probabilistic description [14, 38]. The elementary textbook statement of the theorem is about conditional probabilities

$$\mathbb{P}(I_q | M_y) = \frac{\mathbb{P}(M_y | I_q)}{\mathbb{P}(M_y)} \mathbb{P}(I_q), \quad (4)$$

where  $I_q$  is some subset of possible  $q$ 's, and  $M_y$  is the information provided by the measurement. This becomes problematic when the set  $M_y$  has vanishing probability measure, but if all measures involved have probability density functions (pdf), it may be formulated as ([38] Ch. 1.5)

$$\pi_q(q|y) = \frac{p(y|q)}{\varpi} p_q(q), \quad (5)$$

where  $p_q$  is the pdf of  $q$ ,  $p(y|q)$  is the likelihood of  $y = Y(q)$  given  $q$ , as a function of  $q$  sometimes denoted by  $L(q)$ , and  $\varpi$  is a normalising factor such that the conditional density  $\pi_q(\cdot|y)$  integrates to unity. These terms are in direct correspondence with those in Eq. (4). Most computational approaches determine the pdfs [23, 37, 16].

However, Kolmogorov already defined conditional probabilities via conditional expectation, e.g. see [4]. Given the conditional expectation  $\mathbb{E}(\cdot|M_y)$ , the conditional probability is easily recovered as  $\mathbb{P}(I_q|M_y) = \mathbb{E}(\chi_{I_q}|M_y)$ , where  $\chi_{I_q}$  is the characteristic function of the subset  $I_q$ .

## 2.2 Conditional expectation

The easiest point of departure for conditional expectation in our setting is to define it not just for one piece of measurement  $M_y$ , but for sub- $\sigma$ -algebras  $\mathfrak{S} \subset \mathfrak{A}$ . The connection with an event  $M_y$  is to take  $\mathfrak{S} = \sigma(Y)$ , the  $\sigma$ -algebra generated by  $Y$ .

For RVs with finite variance—elements of  $\mathcal{S} := L_2(\Omega, \mathfrak{A}, \mathbb{P})$ —the conditional expectation  $\mathbb{E}(\cdot|\mathfrak{S})$  is defined as the orthogonal projection onto  $L_2(\Omega, \mathfrak{S}, \mathbb{P})$ . It can then be extended as a contraction onto all  $L_p(\Omega, \mathfrak{A}, \mathbb{P})$  with  $p \geq 1$ , e.g. see [4].

Let us define the space  $\mathcal{Q} := \mathcal{Q} \otimes \mathcal{S}$  of  $\mathcal{Q}$ -valued RVs of finite variance, and set  $\mathcal{Q}_n := \mathcal{Q} \otimes \mathcal{S}_n$  with  $\mathcal{S}_n := L_2(\Omega, \mathfrak{S}, \mathbb{P})$  for the  $\mathcal{Q}$ -valued RVs with finite variance on the sub- $\sigma$ -algebra  $\mathfrak{S}$ , representing the new information.

The Bayesian update as conditional expectation is now simply formulated:

$$\mathbb{E}(q|\sigma(Y)) = P_{\mathcal{Q}_n}(q) = \arg \min_{\tilde{q} \in \mathcal{Q}_n} \|q - \tilde{q}\|_{\mathcal{Q}}^2, \quad (6)$$

where  $P_{\mathcal{Q}_n}$  is the orthogonal projector onto  $\mathcal{Q}_n$ . Already in [15] it was noted that the conditional expectation is the best estimate not only for the *loss function* ‘distance squared’, as in Eq. (6), but for a much larger class of loss functions under certain distributional constraints. However for the above loss function this is valid without any restrictions.

Requiring the derivative of the loss function in Eq. (6) to vanish—equivalently remembering from elementary geometry that the line to the closest point is perpendicular to the approximating subspace—one arrives at the Galerkin orthogonality conditions

$$\forall \tilde{q} \in \mathcal{Q}_n : \quad \langle q - \mathbb{E}(q|\sigma(Y)), \tilde{q} \rangle_{\mathcal{Q}} = 0. \quad (7)$$

To continue, note that the *Doob-Dynkin* lemma [4] assures us that if a RV like  $\mathbb{E}(q|\sigma(Y))$  is measurable w.r.t.  $\sigma(Y)$ , then  $\mathbb{E}(q|\sigma(Y)) = \psi(Y)$  for some measurable  $\psi \in L_0(\mathcal{Y}; \mathcal{Q})$ . More precisely one should write  $\mathbb{E}(q|\sigma(Y)) = \psi(Y(q)) = \psi \circ Y \circ q$ .

Hence  $\mathcal{Q}_n = \overline{\text{span}}\{\phi \circ Y \circ q \in \mathcal{Q} \mid \phi \in L_0(\mathcal{Y}; \mathcal{Q})\}$ , where  $L_0(\mathcal{Y}; \mathcal{Q})$  is the vector space of measurable maps from  $\mathcal{Y}$  to  $\mathcal{Q}$ . In particular one sees that  $\mathbb{E}(q|\sigma(Y))$  is of this form. In this light the task of finding the conditional expectation may be seen as rephrasing Eq. (6) as: find  $\psi \in L_0(\mathcal{Y}; \mathcal{Q})$  such that

$$\psi = \arg \min_{\phi \in L_0(\mathcal{Y}; \mathcal{Q})} \|q - \phi \circ Y \circ q\|_{\mathcal{Q}}^2. \quad (8)$$

Then  $q_a := \psi(y) = \mathbb{E}(q|y)$  is called the *analysis, assimilated, or posterior* value, incorporating the new information.

We would like to emphasise that it is the vector space setting of  $\mathcal{Q}$  and  $\mathcal{Y}$  which has made this well-known formulation possible [15], and it will also allow for easy numerical computation. To work with measures as in Eq. (4) is cumbersome, as probability

measures are on the intersection of the unit sphere and the positive cone in the space of signed finite measures. Similarly, in Eq. (5) the pdfs are in the positive cone of  $L_1(\Omega)$  and on the unit sphere in  $L_1(\Omega)$ , as they have to integrate to unity. A bit easier would be to work with RVs which are in a metric space, the so-called Fréchet-type conditional expectation then minimises the metric distance squared; but the Hilbert space setting is certainly the simplest instance of this, and we will adhere to it here.

In case the  $q$ 's are not without constraints, or not in a vector space, then they should be mapped to such quantities. For example, if  $q$  is a diffusion tensor field, then it has to be symmetric and positive definite. The symmetric tensors are of course a subspace, but the sub-manifold of positive definite ones is not a subspace, nor is this subset closed, making a minimisation as in Eq. (6) ill-defined. However the symmetric positive definite tensors can be given the structure of a Lie group and a Riemannian manifold [1], and then distance is measured as the length of a path along a geodesic. Furthermore, the associated Lie algebra—the tangent space at the neutral element of the Lie group—is in one to one correspondence with the geodesics; hence one can play everything back to a vector space. A simple case of this are positive scalars; through the logarithm they are transformed into a vector space without constraints.

### 2.3 The linear Bayesian update

As we work in a vector space, we make an approximation to simplify the computations by replacing  $L_0(\mathcal{Y}; \mathcal{Q})$  above by  $\mathcal{L}(\mathcal{Y}, \mathcal{Q}) \subset L_0(\mathcal{Y}; \mathcal{Q})$ , the subspace of linear continuous maps. The minimisation Eq. (8) is then translated to: find  $K \in \mathcal{L}(\mathcal{Y}, \mathcal{Q})$  such that [15, 19]

$$K = \arg \min_{H \in \mathcal{L}(\mathcal{Y}, \mathcal{Q})} \|q - H \circ Y \circ q\|_{\mathcal{Q}}^2, \quad (9)$$

and we set  $\mathbb{E}(q|\sigma(Y))_{\ell} := K \circ Y \circ q$ , a linear in  $Y$  approximation to  $\mathbb{E}(q|\sigma(Y))$ . As the projection is now onto the smaller space  $\mathcal{Q}_{\ell} := \overline{\text{span}}\{H \circ Y \circ q \in \mathcal{Q} \mid H \in \mathcal{L}(\mathcal{Y}, \mathcal{Q})\} \subset \mathcal{Q}_n \subset \mathcal{Q}$ , we are not using all the information available. Hence the error—the value of the functional being minimised—will remain larger, but the computation is simpler. The optimal  $K$  is not hard to find by taking the derivative in Eq. (9) w.r.t. the linear map  $H$  (see e.g. [15] and [19] Ch. 3.2 Thm. 4.711), equivalently one may rewrite the variational Galerkin orthogonality condition Eq. (7), for the final result see Eq. (10).

In the case of prior information represented by the *forecast* RV  $q_f$ , which results in the measurement forecast  $y_f = Y(q_f)$ , the projection is adjusted by an affine shift to [15, 19]

$$q_a = q_f + K(z - y_f), \text{ with } K = C_{q,y}(C_y + C_{\varepsilon})^{-1}, \quad (10)$$

here the operator  $K$  is also known as the *Kalman gain*. This includes the errors  $\varepsilon$  assumed independent of  $q$  with zero mean and covariance operator  $C_{\varepsilon}$ , where  $C_y := \mathbb{E}(\tilde{Y}(q, u) \otimes \tilde{Y}(q, u))$  and  $C_{q,y} = \mathbb{E}(\tilde{q} \otimes \tilde{Y}(q, u))$ , where for any RV like  $q$  for the sake of brevity we set  $\bar{q} := \mathbb{E}(q)$  such that  $\tilde{q} := q - \bar{q}$  is the zero-mean part. In case  $C_y + C_{\varepsilon}$  is not invertible or close to singularity, its inverse in Eq. (10) should be replaced by the Moore-Penrose pseudo-inverse. This update is in some ways very similar to the ‘Bayes linear’ approach [11]. If the mean is taken in Eq. (10), one obtains the familiar Kalman filter formula [15] for the update of the mean, and one may show [31] that Eq. (10) also contains the Kalman update for the covariance.

Stated differently, in the situation of Eq. (10)  $q_a$  is the orthogonal projection of  $q$  onto the subspace  $\mathcal{Q}_a = \mathcal{Q}_f + \mathcal{Q}_{\ell}$ , which is generated jointly by the prior information and the measurement. This may be written as  $\mathcal{Q}_a = \mathcal{Q}_f + \mathcal{Q}_{\ell} = \mathcal{Q}_f \oplus \mathcal{Q}_i$ , where the information gain or innovation  $\mathcal{Q}_i$  is the part of  $\mathcal{Q}_{\ell}$  orthogonal to  $\mathcal{Q}_f$ , the last orthogonal sum reflecting the terms in Eq. (10).

Each new measurement enlarges the space we project onto, and thus constrains the error  $q - q_a$ —which is in the orthogonal complement—further. In case the space generated by the measurements is not dense in  $\mathcal{Q}$  a residual error will thus remain, as

the measurements do not contain enough information to resolve our lack of knowledge about  $q$ . Anyway, finding  $q$  is limited by the presence of the error  $\varepsilon$ , as obviously the error influences the update in Eq. (10). If the measurement operator is approximated in some way—as it will be in the computational examples to follow—this will introduce a new error, further limiting the resolution.

### 3 Numerical realisation

In the instances where we want to employ the theory detailed in the previous Section 2, the spaces  $\mathcal{U}$  and  $\mathcal{Q}$  are usually infinite dimensional, as is the space  $\mathcal{S} = L_2(\Omega)$ . For an actual computation they have to be discretised or approximated by finite dimensional spaces. In our examples we will chose finite element discretisations and corresponding subspaces. Hence let  $\mathcal{Q}_M := \text{span} \{\varrho_m : m = 1, \dots, M\} \subset \mathcal{Q}$  be an  $M$ -dimensional subspace with basis  $\{\varrho_m\}_{m=1}^M$ . An element of  $\mathcal{Q}_M$  will be represented by the vector  $\mathbf{q} = [q^1, \dots, q^M]^T \in \mathbb{R}^M$  such that  $\sum_{m=1}^M q^m \varrho_m \in \mathcal{Q}_M$ . The space of possible measurements can usually be taken to be finite dimensional, whose elements similarly are represented by a vector of coefficients  $\mathbf{z} \in \mathbb{R}^R$ .

On  $\mathbb{R}^M$ , representing  $\mathcal{Q}_M$ , the minimisation in Eq. (9) is translated to

$$\mathbf{K} = \arg \min_{\mathbf{H} \in \mathbb{R}^{M \times R}} \|\mathbf{q} - \mathbf{H}(\mathbf{Y}(\mathbf{q}))\|_{\mathcal{Q}}^2, \quad (11)$$

where the mapping induced by  $Y$  has been denoted by boldface  $\mathbf{Y}$ , and the norm  $\|\mathbf{q}\|_{\mathcal{Q}}$  results from the inner product  $\langle \mathbf{q}_1 | \mathbf{q}_2 \rangle_{\mathcal{Q}} := \mathbb{E}(\mathbf{q}_1^T \mathbf{Q} \mathbf{q}_2)$  with  $\mathbf{Q}_{mn} = \langle q_m | q_n \rangle_{\mathcal{Q}}$ , the Gram matrix of the basis. We will later choose an orthonormal basis, so that  $\mathbf{Q} = \mathbf{I}$  is the identity matrix. Then the update corresponding to Eq. (10) is

$$\mathbf{q}_a = \mathbf{q}_f + \mathbf{K}(\mathbf{z} - \mathbf{y}_f), \text{ with } \mathbf{K} = \mathbf{C}_{q,y}(\mathbf{C}_y + \mathbf{C}_\varepsilon)^{-1}. \quad (12)$$

Here the covariances are naturally  $\mathbf{C}_{q,y} := \mathbb{E}(\tilde{\mathbf{q}} \tilde{\mathbf{y}}^T) = \mathbb{E}(\tilde{\mathbf{q}} \otimes \tilde{\mathbf{y}})$ , and similarly for  $\mathbf{C}_y$  and  $\mathbf{C}_\varepsilon$ .

#### 3.1 Markov chain Monte Carlo

We shall shortly sketch the Markov chain Monte Carlo (MCMC) method for the sake of completeness [21, 10], as it will be used on some examples. It is a method which changes the underlying probability measure according to Eq. (4) and Eq. (5) in Subsection 2.1.

The idea is to construct a Markov chain with an equilibrium distribution which corresponds to some desired distribution. In our case this will be the posterior or conditional distribution. The method has the distinct advantage that it does not need target probabilities but only ratios of target probabilities to work. This means that the normalisation constant  $\varpi$  in Eq. (5) does not have to be computed at all—this would involve a cumbersome high-dimensional integration—as it cancels out. Having such a Markov Chain, all that is required to obtain samples from the posterior is to run the Markov chain for a sufficiently long time. As the posterior is the equilibrium distribution, one is indeed sampling the posterior.

The Metropolis algorithm to achieve the construction is described very quickly for the simplest case [21]: assume that the datum  $y$  has been observed, and that the range of possible  $q$ 's has been quantised into  $X$  equal bins with representatives  $\{q_\xi\}_{\xi=1}^X$ , to each  $q_\xi$  assign the number  $\rho_\xi := L(q_\xi)p_q(q_\xi)$ —the product of likelihood and prior pdf. The representatives of the bins  $q_\xi$  will be the states of the Markov chain, which will be denoted by  $\{\varsigma_k\}_{k=1, \dots}$  in the order visited at step  $k = 1, \dots$ .

For each step  $k = 1, 2, \dots$  do the following:

if currently in state  $\varsigma_k = q_\xi$ , then

1. pick any state  $q_\zeta (\zeta \neq \xi)$  randomly with probability  $1/(X - 1)$ , this is the *proposal*;

2. let  $\alpha = \min\{1, \rho_\zeta/\rho_\xi\}$ , this is the *acceptance probability*;
3. accept  $q_\zeta$  with probability  $\alpha$ ,  
i.e. pick a sample of a uniformly distributed RV  $U \in [0, 1]$ , and  
if  $U \leq \alpha$  then set  $\varsigma_{k+1} = q_\zeta$ ,  
else set  $\varsigma_{k+1} = q_\xi$ .

As the acceptance probability in step 2 only involves ratios of the  $\rho_\xi$ , the—hard to compute—normalising constant  $\varpi$  in Eq. (5) is indeed not needed.

The simplicity of the formulation and possibility to correctly estimate the posterior without any approximation are certainly the main advantages of the method. We should warn that although the formulation is so simple, to use the method efficiently and correctly may not be so simple.

First, MCMC is a Monte Carlo method, which means that estimates converge only slowly with increasing number of samples. Second, the samples—coming from a Markov chain—are obviously not independent. This makes estimating any statistic other than the mean (e.g. the variance) difficult. Also usual statistical formulas for the spread or accuracy of the estimates assume *independent* samples and are not applicable. Third, the sequence of visited states is not even *stationary* as the equilibrium distribution is only the asymptotic distribution. This means that one has to wait until  $k$  is *large enough* before actually starting to sample, this is called the *burn-in* period. Usually estimating that the burn-in period is over has to be tested by tests on the stationarity of the sequence. And finally, if the likelihood function and prior pdf differ very much, the acceptance probabilities  $\alpha$  in step 2 may be very low—meaning that practically the chain has got stuck in some state and does not move on.

Here we have used one of the simplest variants, where the underlying transition probabilities of the Markov chain are all equal. For extensions which partly try to alleviate the problems alluded to above and thorough explanations the reader should refer to the literature cited and the references therein.

Now, the two methods to be considered in the following are based on the update via conditional expectation as outlined in Subsection 2.2, and they both employ the linear approximation from Subsection 2.3. This means that in contrast to MCMC they make an additional approximation error, or stated differently, they do not use all the information available for the update. After the discretisation of the space  $\mathcal{Q}$  this boils down to using Eq. (12) in some way on  $\mathcal{Q}_M$ . How this is done differs in the two methods.

### 3.2 Ensemble Kalman Filter

The ensemble Kalman filter (EnKF) is a Monte Carlo method, a sampling interpretation of Eq. (12), for details see e.g. [8, 9]. The basic form is actually simple to describe and is a good preparation for the PCE-based form in the next Subsection 3.3.

Pick  $Z$  independent samples  $\{\mathbf{q}_f(\omega_z)\}_{z=1}^Z$  according to the probability measure  $\mathbb{P}$ . Now Eq. (12) holds for the the RV  $\mathbf{q} : \Omega \rightarrow \mathcal{Q}_M$ , i.e. for all  $\omega \in \Omega$  (a.s.—almost  $\mathbb{P}$ -surely). One takes this to mean that it holds for each  $\omega_z \in \Omega$ . Arranging the the samples in a matrix  $\mathbf{Q}_f := [\mathbf{q}_f(\omega_1), \dots, \mathbf{q}_f(\omega_Z)]$ , and similarly for the forecasts  $\mathbf{Y}_f := [\mathbf{y}_f(\omega_1), \dots, \mathbf{y}_f(\omega_Z)]$  and measurements  $\mathbf{Z}$ , Eq. (12) is now formulated in matrix notation as

$$\mathbf{Q}_a = \mathbf{Q}_f + \mathbf{K}(\mathbf{Z} - \mathbf{Y}_f), \quad (13)$$

where  $\mathbf{K}$  is the same as in Eq. (12).

But the variances needed to compute  $\mathbf{K}$  have to be estimated from the sample. This simply takes the form

$$\mathbf{C}_{q,y} \approx \frac{1}{Z-1} \tilde{\mathbf{Q}}_f \tilde{\mathbf{Y}}_f^T \quad (14)$$

$$\mathbf{C}_y \approx \frac{1}{Z-1} \tilde{\mathbf{Y}}_f \tilde{\mathbf{Y}}_f^T. \quad (15)$$

The normalisation terms  $(Z - 1)^{-1}$  do not really have to be used as they cancel in the computation of  $\mathbf{K}$ . With the estimated means

$$\bar{\mathbf{q}}_f = \frac{1}{Z} \sum_{z=1}^Z \mathbf{q}_f(\omega_z) \quad \text{and} \quad \bar{\mathbf{y}}_f = \frac{1}{Z} \sum_{z=1}^Z \mathbf{y}_f(\omega_z),$$

the terms in Eq. (14) and Eq. (15) are

$$\tilde{\mathbf{Q}}_f = \mathbf{Q}_f - \bar{\mathbf{q}}_f \mathbf{1}_Z^T \quad (16)$$

$$\tilde{\mathbf{Y}}_f = \mathbf{Y}_f - \bar{\mathbf{y}}_f \mathbf{1}_Z^T, \quad (17)$$

where  $\mathbf{1}_Z$  is a vector of ones of size  $Z$ .

This method is a Monte Carlo method, hence it also suffers from the slow convergence with increasing  $Z$ . On the other hand it is fairly simple to implement, all it needs are samples. In practice the number of samples is often low, and then special care is needed when computing the covariances—whose rank can not be higher than the number of samples—and the Kalman gain  $\mathbf{K}$ , see [8, 9]. Another closely connected problem is the possible underestimation of the posterior sample variance with low sample numbers. Still, this method is currently a favourite for problems where the computation of the predicted measurement  $\mathbf{y}_f(\omega_z)$  is difficult or expensive. It needs far fewer samples for meaningful results than MCMC, but on the other hand it uses the linear approximation inherent in Eq. (12).

### 3.3 Polynomial chaos projection

In Subsection 3.2 the Eq. (12) was discretised in the variables  $\omega \in \Omega$  through sampling. Here another track is taken, relying on the tensor product representation of the RVs  $\mathcal{Q} = \mathcal{Q} \otimes \mathcal{S}$  from Subsection 2.2. The first factor already was discretised to  $\mathcal{Q}_N \subset \mathcal{Q}$ , and here an explicit discretisation of the second factor is used. In a numerical sense the sampling in Subsection 3.2 is of course also a discretisation. As for  $\mathcal{Q}_M$  we pick a finite set of independent vectors in  $\mathcal{S}$ . As  $\mathcal{S} = L_2(\Omega)$ , these abstract vectors are in fact RVs with finite variance. Here we will use *Wiener's polynomial chaos* expansion (PCE) as basis [13, 24], this allows to use Eq. (12) without sampling, see [31, 35, 27], and also [36, 3].

The PCE is an expansion in multivariate *Hermite polynomials* [13, 24]; denote by  $H_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \prod_{k \in \mathbb{N}} h_{\alpha_k}(\theta_k) \in \mathcal{S}$  the multivariate polynomial in standard independent Gaussian RVs  $\boldsymbol{\theta}(\omega) = (\theta_1(\omega), \dots, \theta_k(\omega), \dots)_{k \in \mathbb{N}}$ , where  $h_j$  is the usual univariate Hermite polynomial, and  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k, \dots)_{k \in \mathbb{N}} \in \mathcal{N} := \mathbb{N}_0^{(\mathbb{N})}$  is a multi-index of generally infinite length but with only finitely many entries non-zero. As  $h_0 \equiv 1$ , the infinite product is effectively finite and always well-defined.

The *Cameron-Martin* theorem assures us [13] that the set of these polynomials is dense in  $\mathcal{S} = L_2(\Omega)$ , and in fact  $\{H_{\boldsymbol{\alpha}}/\sqrt{(\boldsymbol{\alpha}!)}\}_{\boldsymbol{\alpha} \in \mathcal{N}}$  is a complete orthonormal system (CONS), where  $\boldsymbol{\alpha}! := \prod_{k \in \mathbb{N}} (\alpha_k!)$  is the product of the individual factorials, also well-defined as except for finitely many  $k$  one has  $\alpha_k! = 0! = 1$ . So we may assume that  $\mathbf{q}(\omega) = \sum_{\boldsymbol{\alpha} \in \mathcal{N}} \mathbf{q}^{\boldsymbol{\alpha}} H_{\boldsymbol{\alpha}}(\boldsymbol{\theta}(\omega))$ , and similarly for  $\mathbf{z}$  and  $\mathbf{y}$ . In this way the RVs are expressed as functions of other, known RVs, and not through samples.

The space  $\mathcal{S}$  may now be discretised by taking a finite subset  $\mathcal{J} \subset \mathcal{N}$  of size  $J = |\mathcal{J}|$ , and setting  $\mathcal{S}_J = \text{span} \{H_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{J}\} \subset \mathcal{S}$ . The orthogonal projection  $P_J$  onto  $\mathcal{S}_J$  is then simply

$$P_J : \mathcal{Q}_M \otimes \mathcal{S} \ni \sum_{\boldsymbol{\alpha} \in \mathcal{N}} \mathbf{q}^{\boldsymbol{\alpha}} H_{\boldsymbol{\alpha}} \mapsto \sum_{\boldsymbol{\alpha} \in \mathcal{J}} \mathbf{q}^{\boldsymbol{\alpha}} H_{\boldsymbol{\alpha}} \in \mathcal{Q}_M \otimes \mathcal{S}_J. \quad (18)$$

We then take Eq. (12) and rewrite it as

$$\mathbf{q}_a = \mathbf{q}_f + \mathbf{K}(\mathbf{z} - \mathbf{y}_f) = \quad (19)$$

$$\sum_{\boldsymbol{\alpha} \in \mathcal{N}} \mathbf{q}_a^{\boldsymbol{\alpha}} H_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha} \in \mathcal{N}} (\mathbf{q}_f^{\boldsymbol{\alpha}} + \mathbf{K}(\mathbf{z}^{\boldsymbol{\alpha}} - \mathbf{y}_f^{\boldsymbol{\alpha}})) H_{\boldsymbol{\alpha}}. \quad (20)$$



Projecting both sides of Eq. (20) is then very simple and results in

$$\sum_{\alpha \in \mathcal{J}} \mathbf{q}_a^\alpha H_\alpha = \sum_{\alpha \in \mathcal{J}} (\mathbf{q}_f^\alpha + \mathbf{K}(\mathbf{z}^\alpha - \mathbf{y}_f^\alpha)) H_\alpha. \quad (21)$$

Obviously the projection  $P_J$  commutes with the Kalman operator  $K$  and hence with its finite dimensional analogue  $\mathbf{K}$ . One may actually concisely write Eq. (21) as

$$P_J \mathbf{q}_a = P_J \mathbf{q}_f + P_J \mathbf{K}(\mathbf{z} - \mathbf{y}_f) = P_J \mathbf{q}_f + \mathbf{K}(P_J \mathbf{z} - P_J \mathbf{y}_f). \quad (22)$$

Elements of the discretised space  $\mathcal{Q}_{M,J} = \mathcal{Q}_M \otimes \mathcal{S}_J \subset \mathcal{Q}$  thus may be written as  $\sum_{m=1}^M \sum_{\alpha \in \mathcal{J}} q^{\alpha,m} \varrho_m H_\alpha$ . The tensor representation  $\mathbf{q} := (q^{\alpha,m}) = \sum_{\alpha \in \mathcal{J}} \mathbf{q}^\alpha \otimes \mathbf{e}^\alpha$ , where the  $\mathbf{e}^\alpha$  are the unit vectors in  $\mathbb{R}^J$ , may be used to express Eq. (21) or Eq. (22) succinctly as

$$\mathbf{q}_a = \mathbf{q}_f + \mathbf{K}(\mathbf{z} - \mathbf{y}_f), \quad (23)$$

where  $\mathbf{K} = \mathbf{K} \otimes \mathbf{I}$  with  $\mathbf{K}$  from Eq. (12). Hence the update equation is naturally in a tensorised form. This is how the update can finally be computed in the PCE representation without any sampling [31, 35, 27].

It remains to say how to compute the Kalman gain—or rather the covariance matrices—in this approach. Given the PCEs of the RVs, this is actually quite simple as any moment can be computed directly from the PCE [24, 31, 35]; remembering that  $\mathbf{q}^0 = \mathbb{E}(\mathbf{q})$ , due to the orthogonality of the  $H_\alpha$  one has for the variance

$$\mathbf{C}_y \approx \mathbf{C}_{P_J y} = \mathbb{E}((P_J \mathbf{y}) \otimes (P_J \mathbf{y})) = \sum_{\alpha \in \mathcal{J}, \alpha \neq 0} (\alpha!) \mathbf{y}^\alpha \otimes \mathbf{y}^\alpha, \quad (24)$$

$$\mathbf{C}_{q,y} = \sum_{\alpha \in \mathcal{N}, \alpha \neq 0} (\alpha!) \mathbf{q}^\alpha \otimes \mathbf{y}^\alpha \approx \sum_{\alpha \in \mathcal{J}, \alpha \neq 0} (\alpha!) \mathbf{q}^\alpha \otimes \mathbf{y}^\alpha. \quad (25)$$

As was already remarked Subsection 3.2, the covariance matrix can not have higher rank than the number of tensor products in the sum. And certainly any such truncation as is implicit in the projection operator  $P_J$  Eq. (18) or explicit as in Eq. (24) and Eq. (25) will reduce the variance. The difference to the procedure in Subsection 3.2 is that here the  $\mathbf{q}^\alpha$  and  $\mathbf{y}^\alpha$  are components of the orthogonal basis  $H_\alpha$ . The main ‘difficulty’ to apply the the update as described in this section is the need to have a PCE of  $\mathbf{q}_f(\omega)$  and especially of the forecast measurement  $\mathbf{y}_f(\omega) = Y(\mathbf{q}_f(\omega))$ . How this might be done will be sketched in the next Section 4.

One should not forget to draw attention to the fact that once such an expansion is available, evaluating an expression such as  $\sum_{\alpha \in \mathcal{J}} \mathbf{y}^\alpha H_\alpha(\omega)$  for some specific  $\omega \in \Omega$  is computationally relatively cheap, and this is an approximation to a sample of  $\mathbf{y}_f(\omega)$ —it is precisely  $P_J \mathbf{y}_f(\omega)$ . Hence this can be used wherever samples are needed, and this device has actually been employed, i.e. in the MCMC method [22, 16, 32, 17] as described in Subsection 3.1, or in the EnKF method [18] described in Subsection 3.2.

## 4 The forward problem

As was noted already in Section 3, the spaces  $\mathcal{U}$  and  $\mathcal{Q}$  are usually infinite dimensional, as is the space  $\mathcal{S} = L_2(\Omega)$ . Similarly to the discretisation described there, the space  $\mathcal{U}$  has to be discretised as well. In an analogous fashion, choosing an  $N$ -dimensional subspace  $\mathcal{U}_N = \text{span} \{v_n : n = 1, \dots, N\} \subset \mathcal{U}$  with basis  $\{v_n\}_{n=1}^N$ . An element of  $\mathcal{U}_N$  will similarly be represented by the vector  $\mathbf{u} = [u^1, \dots, u^N]^T \in \mathbb{R}^N$  such that  $\sum_{n=1}^N u^n v_n \in \mathcal{U}_N$ .

The state equation may then be discretised by inserting the ‘ansatz’ that the solution is from  $\mathcal{U}_N$  by assuming in Eq. (1) that  $u = \sum_n u^n v_n$ . We obtain equations for the

coefficients  $u^n$  for example by projecting in general Galerkin manner Eq. (1) onto  $\mathcal{U}_N$ ; most straightforward is to use the basis  $\{v_n\}_{n=1}^N$  also for the projection:

$$\forall k: \quad \langle v_k | \frac{\partial}{\partial t} \sum_n u^n(t) v_n \rangle_{\mathcal{U}} + \langle v_k | A(q; \sum_n u^n(t) v_n) \rangle_{\mathcal{U}} = \langle v_k | f(q; t) \rangle_{\mathcal{U}}, \quad (26)$$

where we have now only included the dependence on the parameter  $q \in \mathcal{Q}$  which has to be identified. Eq. (26) may in standard manner succinctly be written as an equation on  $\mathbb{R}^N$ :

$$\frac{\partial}{\partial t} \mathbf{u}(t) + \mathbf{A}(\mathbf{q}; \mathbf{u}) = \mathbf{f}(\mathbf{q}; t). \quad (27)$$

The measurement operator will, in combination with the developments in Section 3, be denoted as (see Eq. (3))

$$\mathbf{y}_f = \mathbf{Y}(\mathbf{u}(t_k), \mathbf{q}_{k-1}) = \mathbf{Y}(\mathbf{q}_{k-1}). \quad (28)$$

The unknown parameter is modelled as a RV—see Section 2, which in the discrete setting of Section 3 means that Eq. (27) reads

$$\frac{\partial}{\partial t} \mathbf{u}(\omega) + \mathbf{A}(\mathbf{q}(\omega); \mathbf{u}(\omega)) = \mathbf{f}(\mathbf{q}(\omega)) \quad \mathbb{P}\text{- a.s. in } \omega \in \Omega. \quad (29)$$

For MCMC (Subsection 3.1) and EnKF (Subsection 3.2) or any other Monte Carlo or collocation method, this equation has to be solved for each sample or collocation point  $\omega_z$  and  $\mathbf{q}(\omega_z)$ , to obtain  $\mathbf{u}(\omega_z)$ , and with that predict the measurement  $\mathbf{y}_f(\omega_z) = \mathbf{Y}(\mathbf{u}(\omega_z), \mathbf{q}(\omega_z))$ . Obviously, this may be computationally quite costly. As remarked at the end of Subsection 3.3, if one can compute a PCE of  $\mathbf{y}_f(\omega) = \sum_{\alpha \in \mathcal{J}} \mathbf{y}_f^\alpha H_\alpha(\boldsymbol{\theta}(\omega))$ , this may be used instead to great advantage, as it will be usually much less work to evaluate instead of going via Eq. (29).

For the linear Bayesian update in Subsection 3.3 via PCE this comes natural, but of course this means that Eq. (29) has to be solved so that this is possible. Just as the parameter was represented in Subsection 3.3 via its truncated PCE Eq. (18)  $\mathbf{q}(\omega) = \sum_{\alpha \in \mathcal{J}} \mathbf{q}^\alpha H_\alpha(\boldsymbol{\theta}(\omega))$ , or through the tensor of coefficients  $\mathbf{q} = \sum_{\alpha \in \mathcal{J}} \mathbf{q}^\alpha \otimes \mathbf{e}^\alpha$ , the same is assumed by an ‘ansatz’ for the solution to Eq. (29)  $\mathbf{u}(\omega) = \sum_{\alpha \in \mathcal{J}} \mathbf{u}^\alpha H_\alpha(\boldsymbol{\theta}(\omega))$ , represented through  $\mathbf{u} = \sum_{\alpha \in \mathcal{J}} \mathbf{u}^\alpha \otimes \mathbf{e}^\alpha$ .

Inserting this into Eq. (29) and projecting with  $P_J$  Eq. (18) onto  $\mathcal{S}_J$ , one obtains the equations for  $\mathbf{u}(t) \in \mathbb{R}^N \otimes \mathbb{R}^J$  through Galerkin conditions with the basis  $\{H_\alpha\}_{\alpha \in \mathcal{J}}$  analogous to Eq. (26):

$$\frac{\partial}{\partial t} \mathbf{u} + \mathbf{A}(\mathbf{q}; \mathbf{u}) = \mathbf{f}(\mathbf{q}), \quad (30)$$

with the obvious interpretation of the terms. The same procedure applied to Eq. (28) results in  $\mathbf{y}_f = \mathbf{Y}(\mathbf{u}(t_k); \mathbf{q}_{k-1})$ . Identifying  $\mathbf{q}_f$  with  $\mathbf{q}_{k-1}$  at step  $k$  of the updating, all the terms needed to use the update Eq. (23) are present.

Let us remark that the tensor product representation like the one employed for the state  $\mathbf{u}$  in Eq. (30) may be extended [27] to all entities in Eq. (30). Low-rank approximations to those entities in tensor representation then become possible in different guises, as model reduction [5], as so called ‘generalised spectral decomposition’ [30], during the solution process progressively as ‘proper generalised decomposition’ [29], or in the iteration as approximate, perturbed, or compressed iteration [12, 26], and this may lead to considerable numerical savings. As Eq. (23) is in such a format, these savings carry over to the Bayesian update.

## 5 Numerical examples

A comparison of the different updating methods is outside the scope of this paper and will be presented elsewhere. Here we rather want to present some examples of identification

procedures which illustrate how the methods work. Note that these examples are not intended to describe real world applications but are rather set up to illustrate the class of full and linear Bayesian identification methods. Specifically, the measurement operation will not be performed on a real system, but will also be simulated through computation. This has the advantage that we know what the ‘truth’ is, i.e. the value of the parameter field with which we simulate the measurement, so that one can compare this with the estimate from the identification.

These examples will be both linear and nonlinear problems described through partial differential equations, namely linear and nonlinear diffusion and elasto-plastic behaviour. Hence the Eq. (1) is here a partial differential equation, and the ‘parameter’ to be identified will be some coefficient fields in the equation, either the diffusion coefficient, or the shear modulus. Let us denote such a generic random field as  $q(x, \omega)$ , where now  $x \in \mathcal{G}$  is a point in the spatial domain  $\mathcal{G} \subset \mathbb{R}^d$ . More precisely the spatial function  $q(\cdot, \omega)$  is the parameter to be identified. We assume that this field is transformed in such a way that it is without constraints in some vector space  $\mathcal{Q}$ .

To start the numerical computations, one more step is needed to be able to effectively describe the random fields. What one needs for the developments described in Subsection 3.3 is the PCE of the random field  $q(x, \omega) = \sum_{\alpha \in \mathcal{N}} q^\alpha(x) H_\alpha(\boldsymbol{\theta}(\omega))$ . The spatial coefficient functions are given by simple projection  $q^\alpha(x) = \mathbb{E}(q(x, \cdot) H_\alpha(\boldsymbol{\theta}(\cdot)))$ .

The computational problem with this approach is that the PCE is completely general and is defined without any reference to the random field  $q(x, \omega)$ . This means that an excessive number of RVs  $\theta_1(\omega), \dots, \theta_k(\omega), \dots$  may be needed to give an accurate enough approximation when the above PCE is truncated to some  $\alpha \in \mathcal{J} \subset \mathcal{N}$ .

One way to have an accurate description with a fairly small number of RVs is to use the Karhunen-Loève expansion (KLE), e.g. see [24]. With the covariance function of the random field  $c_q(x_1, x_2) := \mathbb{E}(q(x_1, \cdot)q(x_2, \cdot))$ —for the sake of simplicity we will only consider scalar random fields—one computes the eigenvalues and -functions of the Fredholm eigenvalue problem with the covariance function as kernel:

$$\int_{\mathcal{G}} c_q(x_1, x_2) q_j(x_2) dx_2 = \lambda_j q_j(x_1), \quad (31)$$

which results in the spectral decomposition  $c_q(x_1, x_2) = \sum_j \lambda_j q_j(x_1) q_j(x_2)$ .

The Karhunen-Loève expansion then is

$$q(x, \omega) = \bar{q}(x) + \sum_j \sqrt{\lambda_j} \xi_j(\omega) q_j(x), \quad (32)$$

where the non-negative eigenvalues  $\lambda_j$  and orthonormal eigenfunctions  $q_j(x)$  are usually ordered in a decreasing sequence, and the zero-mean unit-variance uncorrelated (i.e. orthonormal) RVs  $\xi_j(\omega)$  are given through a simple projection of the random field

$$\xi_j(\omega) = \frac{1}{\sqrt{\lambda_j}} \int_{\mathcal{G}} (q(x, \omega) - \bar{q}(x)) q_j(x) dx. \quad (33)$$

As the variance of each term in the KLE Eq. (32) is  $\lambda_j$ —the total variance is  $\sigma_t^2 := \sum_j \lambda_j = \int_{\mathcal{G}} c_q(x, x) dx$ —it is usually truncated after say  $M$  terms so that the residual variance  $\sigma_r^2 := \sum_{j>M} \lambda_j$  is small.

The field  $q(x, \omega)$  is then approximated or represented by

$$q(x, \omega) = \bar{q}(x) + \sigma_c \sum_{j \leq M} \sqrt{\lambda_j} \xi_j(\omega) q_j(x), \quad (34)$$

which has the same mean  $\bar{q}(x)$  as the original field Eq. (32) and covariance  $c_q(x_1, x_2) = \sigma_c^2 \sum_{j \leq M} \lambda_j q_j(x_1) q_j(x_2)$ . If the correction factor  $\sigma_c$  is chosen as  $\sigma_c = \sqrt{\sigma_t^2 / (\sigma_t^2 - \sigma_r^2)}$ , the approximate field Eq. (34) will also have the correct total variance—only distributed on the modes  $q_j(x)$  with  $j \leq M$ .

Through the KLE the random field is best approximated in variance with the least number of RVs. For further computation the RVs  $\xi_j(\omega)$  are sometimes unwieldy, so they in turn may now be expanded in a PCE

$$\xi_j(\omega) = \sum_{\alpha \in \mathbb{N}_0^M} \xi_j^\alpha H_\alpha(\boldsymbol{\theta}_M(\omega)), \quad (35)$$

where we only need  $M$  isonormal RVs  $\boldsymbol{\theta}_M(\omega) = (\theta_1(\omega), \dots, \theta_M(\omega))$ . To become computationally viable, the series Eq. (35) will have to be reduced to a finite sum, usually by choosing a finite subset  $\mathcal{J}_j \subset \mathbb{N}_0^M$ . One criterion could of course be again to choose those  $\alpha$  which give the highest contribution. By doing this one will obviously decrease the variance, and this may be compensated by using an additional scaling factor  $\sigma_j$  computed similarly to  $\sigma_c$  above. For the sake of simplicity we shall avoid doing this here. Inserting the thus truncated series into Eq. (34) gives

$$q(x, \omega) = \bar{q}(x) + \sigma_c \sum_{j \leq M} \sum_{\alpha \in \mathcal{J}_j} \sqrt{\lambda_j} \xi_j^\alpha H_\alpha(\boldsymbol{\theta}_M(\omega)) q_j(x). \quad (36)$$

Let  $\mathcal{J} := \bigcup_j \mathcal{J}_j$ , set  $\xi_j^\alpha = 0$  for  $\alpha \in \mathcal{J} \setminus \bigcup_{k \neq j} \mathcal{J}_k$ , and let  $\xi^\alpha(x) := \sigma_c \sum_{j \leq M} \sqrt{\lambda_j} \xi_j^\alpha q_j(x)$  and  $\xi^0(x) = \bar{q}(x)$ ; then the truncated PCE of  $q(x, \omega)$  in  $M$  Gaussian RVs is

$$q(x, \omega) = \sum_{\alpha \in \mathcal{J}} \xi^\alpha(x) H_\alpha(\boldsymbol{\theta}_M(\omega)). \quad (37)$$

Depending on the specific situation, either the expression Eq. (34) or Eq. (37) can be used in an actual computation.

## 5.1 Examples with MCMC updating

Our study will start with a relatively simple system and then progress to more difficult cases, illustrating the Markov chain Monte Carlo (MCMC) methods of Subsection 3.1. Let us recall that in that case the posterior is described by changing the distribution of the underlying RVs.

We start with a simple model of steady state heat transfer [16], where the energy balance equation leads to

$$-\operatorname{div}(\kappa(x, \omega) \nabla u(x, \omega)) = f(x, \omega), \quad \text{a.e. } x \in \mathcal{G} \subset \mathbb{R}^2 \quad (38)$$

where both the heat conductivity  $\kappa$  and the right hand side  $f$  may be considered as random fields over a probability space  $\Omega$ . Thus Eq. (38) is required to hold almost surely in  $\omega$ , i.e.  $\mathbb{P}$ -almost everywhere. For the sake of simplicity the conductivity  $\kappa(x, \omega)$  is taken to be a scalar valued random field, although a conductivity tensor would be a more appropriate choice. First we will not let  $\kappa$  depend on  $u$  and Eq. (38) represents linear model. To the balance equation one has to add appropriate Dirichlet and Neumann boundary conditions, given as a prescribed temperature  $u_0 = 20^\circ\text{C}$  and a heat flux  $\nu_x = 100 \text{ Wm}^{-2}$  on the left side and zero flux on top and bottom respectively, as shown in Fig. 1. This completes the concrete description of the abstract Eq. (1), where here there is no time dependence and thus the time derivative terms vanish.

After spatial discretisation by a standard finite element method into 120 elements (see Fig. 1), the problem is further discretised with the help of random variables as described in Section 4. Following this the problem reduces to the form given by Eq. (30) without the first term as mentioned already. The system in Eq. (30) is then solved via the methods given in [5, 29, 26].

The conductivity parameter  $\kappa(x, \omega)$  is first considered in two different scenarios: first as a simple random variable—i.e. no dependence on  $x \in \mathcal{G}$ —and then as a random field.

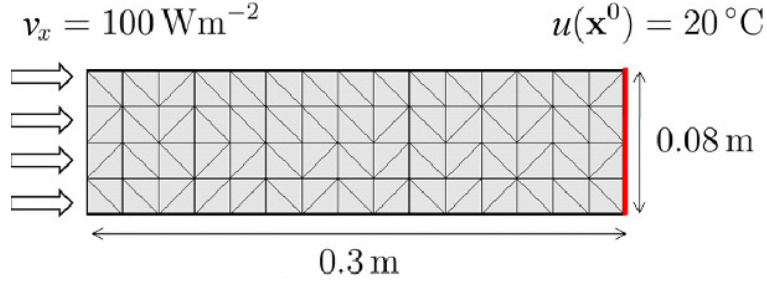


Figure 1: Experimental setup.

### 5.1.1 Conductivity as a simple random variable

We take the thermal conductivity  $\kappa$  as independent of the coordinates  $x \in \mathcal{G}$ , and for simplicity it is we assume it as an a priori normally distributed variable  $\kappa_f = q_f$  with mean value  $\mathbb{E}(q_f) = 2 \text{ Wm}^{-1}\text{K}^{-1}$  and standard deviation  $0.3 \text{ Wm}^{-1}\text{K}^{-1}$ .

The measurements will be simulated with a ‘true’ value of  $\kappa_t = 2.5 \text{ Wm}^{-1}\text{K}^{-1}$ . They are performed in each node of the finite element mesh; and further polluted at each node by independent centred Gaussian noise with a standard deviation of  $\sigma_\varepsilon = 10^\circ\text{C}$ .

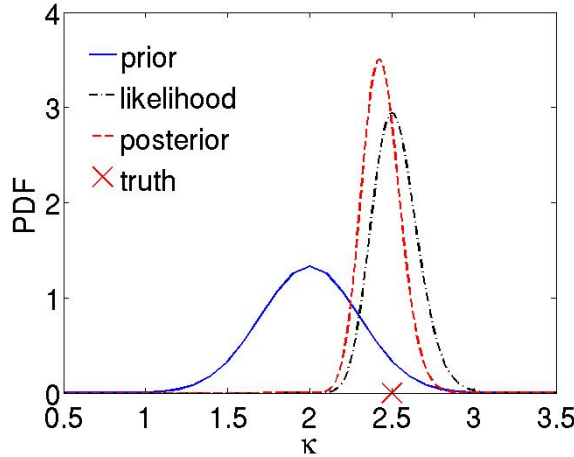


Figure 2: Comparison of prior and posterior pdf and the likelihood function.

In this particular example the identification of  $\kappa$  is done in a fully Bayesian manner (see Eq. (4)) with the help of a MCMC procedure with 100000 samples as described in Subsection 3.1. In Fig. 2 we display the shape of the prior, the likelihood function, and the posterior probability density function (pdf), and compare it to the truth. We see that the mean and mode of the posterior —both are often taken as single point estimates—when compared to the prior have moved in the direction of the truth; and the variance of the posterior (signifying the uncertainty about the true value) is less than that of the prior. Another frequent single point estimate is the maximum of the likelihood function—the well known maximum likelihood estimate  $L_{\max}(\kappa)$ —which may be seen as also being close to the truth. One may also observe that these (and other) point estimates give only an incomplete picture, as they will not contain information on the residual uncertainty. The Bayesian identification procedure on the other hand yields a complete probability distribution which informs also about the residual uncertainty.

The MCMC method performs very well as shown, it is computationally very expensive. It requires the calculation of the model response for a large number of samples, each time solving a FE-system with a different material parameter. In order to improve the performance [23, 16] we compute a polynomial chaos approximation of the model

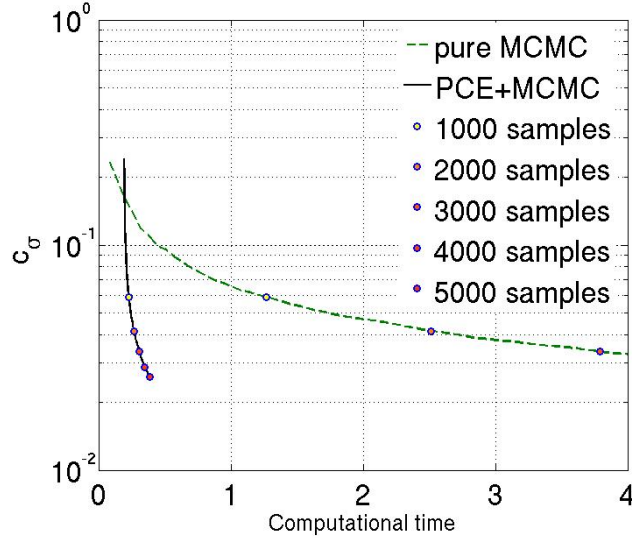


Figure 3: Comparison of the computational times of pure MCMC and PCE based MCMC methods versus the estimate  $c_\sigma$ .

response by a stochastic Galerkin procedure. Then the integration of the posterior distribution may be realised by sampling the PCE as alluded to at the end of Section 4. This significantly accelerates the update procedure as shown in Fig. 3, where the estimate  $c_\sigma := \sigma(L_{\max}(\kappa))/\mathbb{E}(L_{\max}(\kappa))$  is plotted versus computational time for both the pure and the PCE based MCMC approach. Clearly the PCE approximation for the same accuracy of estimate runs much quicker than the pure MCMC procedure.

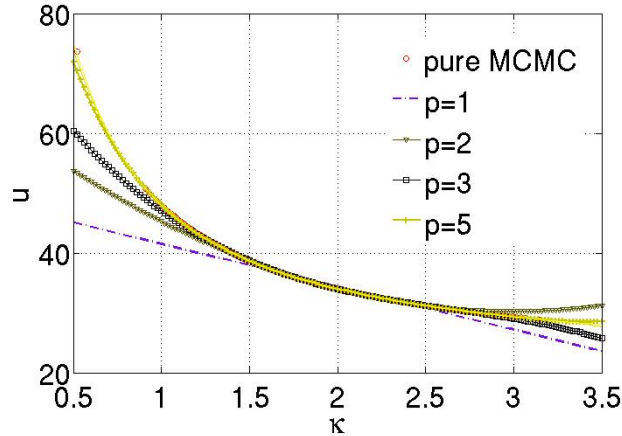


Figure 4: Polynomial chaos approximation of the model response in one node of computational domain

The dependence of the PCE on the polynomial order is investigated in Fig. 4 where the accuracy of the temperature approximation in one node of the domain is plotted for different polynomial orders. Comparing the approximated solution with the reference one obtained by MCMC with 100000 samples one concludes that only the fifth order approximation can be safely used. In this case both responses match. Otherwise, for lower values of the polynomial order the mean value is estimated correctly (the crossing point) but not the higher order moments (the lines do not match). We try to quantify

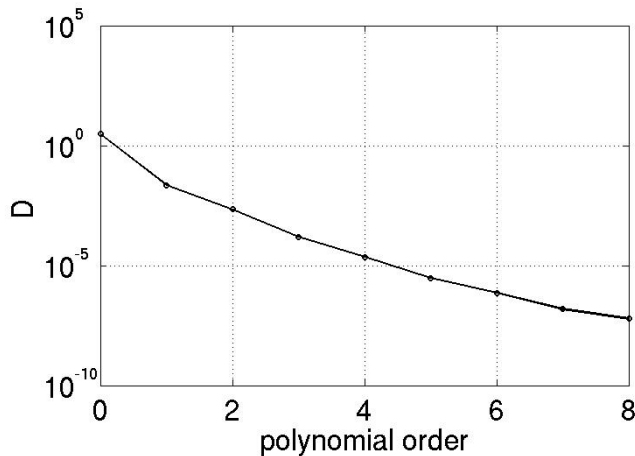


Figure 5: Kullback-Leibler distance  $D(\tilde{\pi}||\pi)$  from the posterior obtained by pure MCMC to the PCE reformulated posterior, versus polynomial order

this discrepancy in Fig. 5 in a quantitative assessment of the error in the posterior density of the PCE-based MCMC by computing the Kullback-Leibler divergence (KLD) of  $\tilde{\pi}_\kappa$  (posterior obtained by PCE/MCMC) from  $\pi_\kappa$  (posterior obtained by pure MCMC),

$$D(\tilde{\pi}||\pi) := \int \tilde{\pi}(\kappa) \log \frac{\tilde{\pi}(\kappa)}{\pi(\kappa)} d\kappa. \quad (39)$$

This is shown in Fig. 5, where one may observe a rapid rate of the convergence of the surrogate posterior  $\tilde{\pi}_\kappa$  to the true posterior.

### 5.1.2 Conductivity as random field

The next example is a bit more realistic, by not assuming a priori that the truth is spatially constant. We repeat the previous analysis for the prior  $\kappa_f$  being a random field—again for the sake of simplicity—assumed as Gaussian. The field  $\kappa_f = q_f$  is described by a covariance kernel  $\exp(-r/l_c)$  with  $r$  being the distance and  $l_c$  a correlation length. Furthermore, the mean and standard deviation are chosen in the same manner as before (see Section 5.1.1), while the truth is modelled in a more realistic way as one realisation of the prior field, as shown in Fig. 6. All other quantities such as measurement and measurement error stay the same as in the previous example.

The numerical simulation of the conductivity field is performed by truncating the prior KLE (see beginning of this Section 5) to six modes, i.e. the random variables  $\xi_1 \dots \xi_6$  in Eq. (32) have to be updated. Similarly as before we use the pure MCMC update procedure as well as the PCE based MCMC method. The pure MCMC Bayesian update uses 100000 samples in order to assimilate the posterior variables, while the PCE approximation is taken to be of second order. In both cases the updated variables obtain the similar non-Gaussian form as shown in Fig. 7 b) and d). In addition, by the comparison of the bar-graphs of prior and posterior distributions in Fig. 7 b) and d) one may see the significant reduction of the prior variance. This is clearly visible in Fig. 7 c) where the probability density function of  $\xi_5$  is plotted. Through the identification or assimilation process one obtains from a very wide prior distribution a much narrower posterior as expected. Also, both methods give very similar results leading to the conclusion that the error introduced by PCE truncation to second order seems to be negligible in this example.

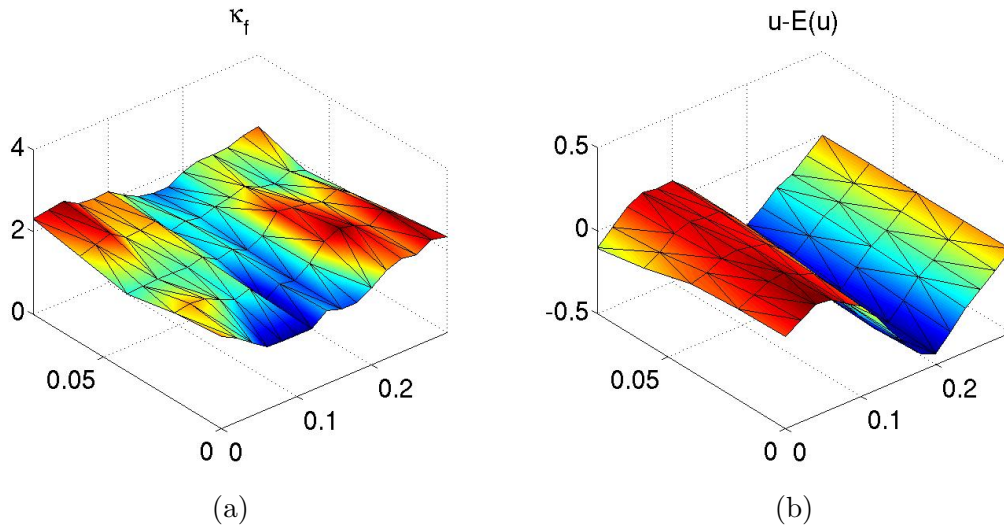


Figure 6: (a) A realisation of the prior random field, (b) Fluctuations of a corresponding temperature field.

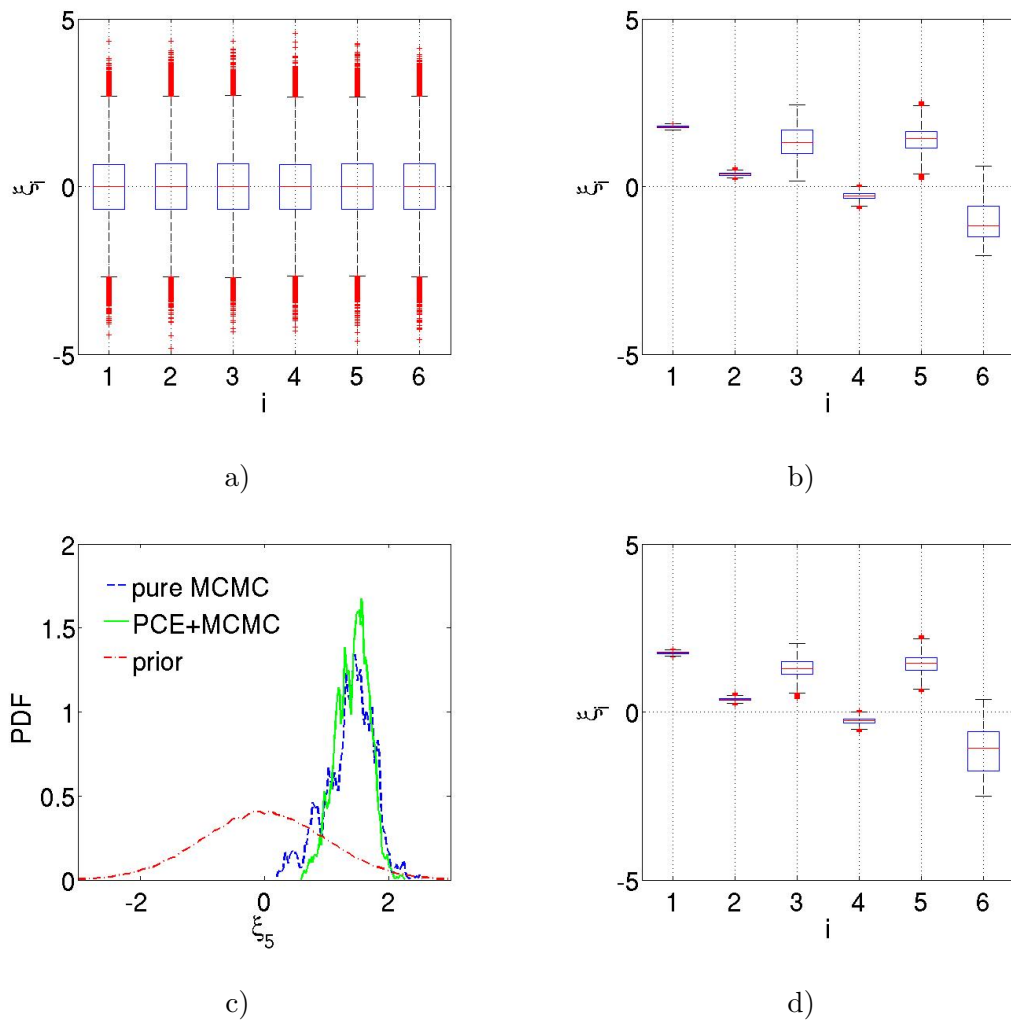


Figure 7: Distributions of particular random variables: (a) prior, (b) posterior obtained by pure MCMC, (c) pdfs of  $\xi_5$ , (d) posterior obtained using PCE approximation of model response. In a), b), and c), the box graphs show for each RV ( $\xi_1, \dots, \xi_6$ ) on the ordinate the median, the central 25th percentile as a box, the central 75th percentile, and outliers.



### 5.1.3 Non-linear diffusion

In the isotropic quasi-static diffusion equation given by Eq. (38), we now assume that the conductivity  $\kappa$  depends on the temperature field  $u(x, \omega)$  through

$$\kappa(u, x, \omega) = \kappa_0(x, \omega) + \kappa_1 u(x, \omega), \quad (40)$$

and the previously linear SPDE becomes nonlinear. Again, just like the prior assumption of a Gaussian conductivity before—chosen purely for reasons of simplicity and useful solely because the coefficient of variation was very small—is not consistent with the requirement by the second law of thermodynamics that the conductivity be positive, also this linear dependence of the conductivity can only be seen as a crude approximation, useful only for a small temperature range. Its only purpose is to show how the method works for a simple nonlinear state equation.

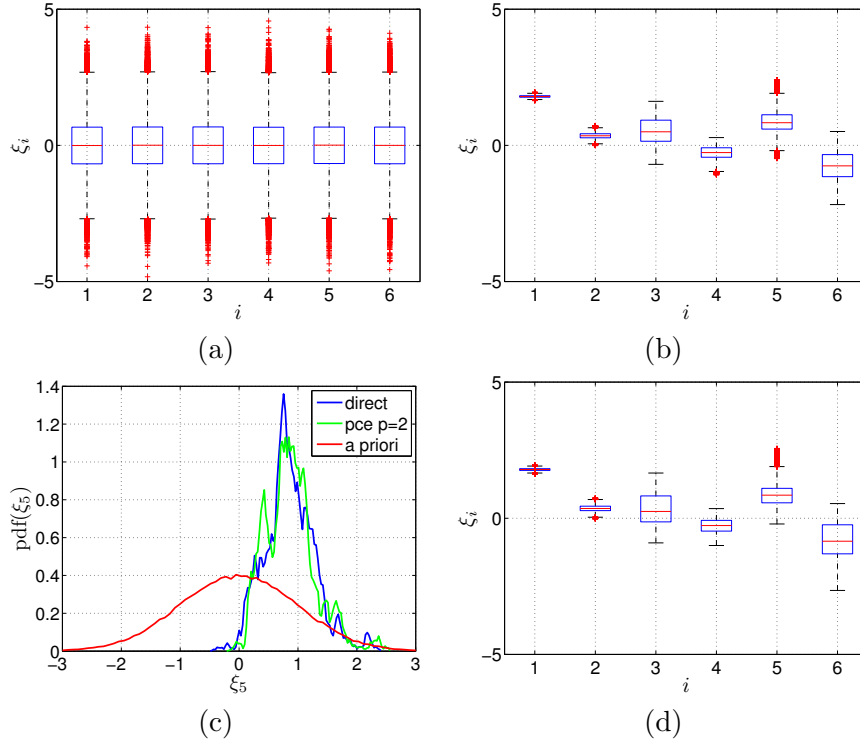


Figure 8: Distributions of particular random variables: (a) a priori, (b) a posteriori obtained by direct evaluation of the FE model, (c) probability density functions of  $\xi_5$ , (d) a posteriori obtained using PC approximation of model response. Box-graphs as in Fig. 7.

This very simple non-linear constitutive law is described by two artificial parameters  $\kappa_0$  and  $\kappa_1$ , where  $\kappa_1 = 0.05 \text{ Wm}^{-1}\text{K}^{-2}$  is assumed as known and hence is certain, and  $\kappa_0$  is a priori taken as a Gaussian random field as in the previous Section 5.1.2 with a six term KLE with mean  $\mathbb{E}(\kappa_0) = 2 \text{ Wm}^{-1}\text{K}^{-1}$  and standard deviation  $0.3 \text{ Wm}^{-1}\text{K}^{-1}$ . The truth is taken as one sample of this random field. The rest of the experimental set up is taken from the previous tests (see the previous Section 5.1.1 and 5.1.2).

Similarly as before, the MCMC method in its pure sampling and PCE based form is used to calculate the pdfs of the conductivity parameter  $\kappa_1$ . The update procedure consists of assimilation of the six random variables  $\xi_1, \dots, \xi_6$  of the truncated KLE of the prior field  $q_f = \kappa_f$  with the help of the temperature measurements in all points of the finite element mesh. Similarly as in the linear case, the box-graphs of the prior and posterior random variables in Fig. 8 a) and b) show significant reduction of the prior variance and a change of the mean, which causes the non-Gaussian form of the posterior random variables. Furthermore, the comparison of the pure and PCE based MCMC

procedure in Fig. 8 c) and d) shows that both methods give approximately the same results, even though the PCE approximation is only of second order.

We see that the MCMC based Bayesian update works also for random fields, even though the measurement (here the temperature) is not linear in the quantity to be identified (the conductivity), and also works in case the state equation is nonlinear. The inconsistent shape of the prior distribution (allowing negative conductivities with non-zero probability) is corrected in all the updates to a form where practically all the probability weight is in the positive range. The low order truncated PCE based MCMC performs almost just as well and is computationally much cheaper—this is practically a surrogate model.

## 5.2 Examples with linear Bayesian updating

The examples shown up to now have used the so called ‘full’ Bayesian update based on Eq. (4) or Eq. (5), i.e. by updating the measure. We now turn our attention to updates which change the random variable, based on Eq. (23). The first example will also be a diffusion equation, although on a different domain, and the second example will be an elastic-plastic system, hence non-smooth and strongly nonlinear.

In previous three examples we have considered the conductivity a priori as a normally distributed random field, which ignores the positivity requirement as already noted. In order to have an a priori model which takes care of this requirement in a way that our variables are free of constraints—the importance of this was already stressed in Subsection 2.2—we model the logarithm of the conductivity field; the a priori field is taken as a lognormal random field, according to maximum entropy considerations this is the right choice if one only presumes that the field has a finite variance. This makes the identification problem harder, as the logarithm resp. exponential is additionally involved.

As already pointed out in Subsection 2.2, this corresponds with [1] considering the positive cone in the space of all fields as a differentiable Riemann manifold, and more specifically also as a Lie group. The logarithm and exponential are then the equally named maps from Lie theory, carrying the tangent space at the identity into the corresponding Lie algebra and vice versa. As the manifold can be equipped with a Riemann metric which is carried from the Lie algebra via the exponential map to the tangent space of the identity, and from there via the group operations to any other place, distances on the manifold can be measured as path lengths along geodesics. This turns the manifold into a metric space and would allow to use the notions of Fréchet-type conditional expectations as alluded to in Subsection 2.2. But geodesics correspond uniquely to straight lines in the Lie algebra, hence to their direction vectors without any constraints [1]. It is in this space that we propose to do all the operations.

### 5.2.1 Diffusion with linear Bayes updates

We consider the same Eq. (38) as before, but now on an  $L$ -shaped domain  $\mathcal{G}$  with specified homogenous Dirichlet and Neumann boundary conditions. This makes the heat flow a bit more complicated than in the previous example, where inhomogeneities in the heat flow arise solely due conductivity variations, whereas here they are also caused through the geometry. The external loading is defined as a sinusoidal function  $f = f_0 \sin(\frac{2\pi}{\lambda} x^T v + \varphi)$ , with  $f_0$  being the amplitude,  $\lambda$  the wave-length,  $\varphi \in [0, 2\pi]$  the phase, and  $v = (\cos \alpha, \sin \alpha)$  the direction of the sinusoidal wave specified by an angle  $\alpha \in [-\pi/2, \pi/2]$ . For a more detailed description, please see [35].

The identification of the true conductivity field  $\kappa_t$  is done with the help of the linear Bayesian method in its direct PCE and MC sampling (EnKF) form. Additionally, the update process is realised in a sequential way as shown in Fig. 9 by repeating the measurements in the same experimental conditions for different values of the right hand side. In other words one takes for the prior in the next update the posterior from the

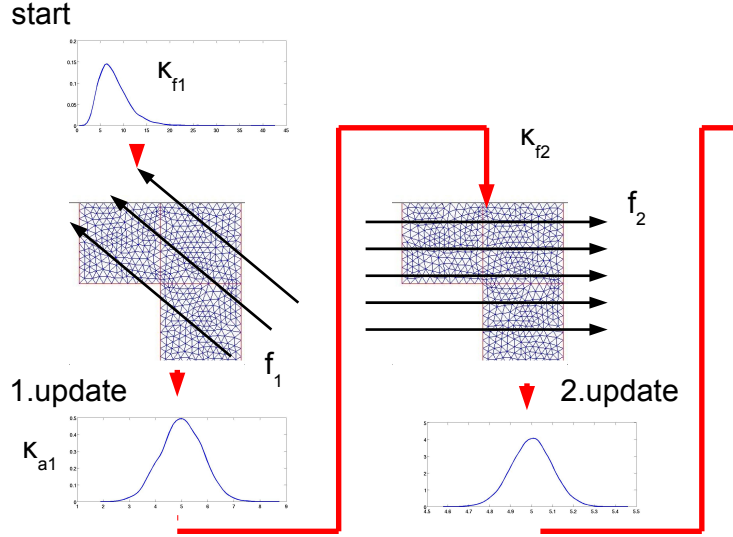


Figure 9: Schematic depiction of sequential update procedure

previous update and does the new measurement according to a new value of  $f$ . The right hand side values are altered by variation of the appropriate parameter values, such as the wave length  $\lambda$ , the phase  $\varphi$ , etc. In this way different regions of the domain are stimulated, aiding in the identification process.

The forward problem is solved within the finite element framework by discretising the spatial domain into 1032 triangular elements. The measurements  $y(u)$  of the state quantity  $u$ ,

$$y(u, \omega) := [\dots, y(x_j), \dots] \in \mathbb{R}^L, \quad y(x_j) = \int_{\mathcal{G}_j} u(x, \omega) dx, \quad (41)$$

are obtained on little patches  $\mathcal{G}_j \subset \mathcal{G}$  centred at the finite element node  $x_j \in \hat{\mathcal{G}} = \{x_1, \dots, x_B\}$ . The measurements are performed in only 10% of the total number of mesh nodes, equally distributed over the computational area. The measured values are disturbed by a centred Gaussian noise with the diagonal covariance  $\sigma_\varepsilon^2 \mathbf{I}$ , where  $\sigma_\varepsilon$  is approximately equal to 1% of the measured value.

For the ‘truth’ we take one realisation of a lognormal random field sampled from the modified lognormal distribution  $\kappa_t := 2 + \kappa_b(x, \omega)$ , where  $\kappa_b$  represents a homogenous lognormal random field with  $\mathbb{E}(\kappa_b) = 1$  and standard deviation  $\sigma_{\kappa_b} = 0.2236$  obtained as the exponential transformation of a Gaussian random field  $q_b$  described by an exponential covariance function  $\exp(-r/l_c)$ , with  $r$  being the spatial distance and  $l_c = 0.5$  the correlation length.

Similarly to this, the prior is chosen to be a homogeneous lognormal field  $\kappa_f = \exp(q_f)$  with  $\mathbb{E}(\kappa_f) = 2.4$  and standard deviation  $\sigma_{\kappa_f} = 0.8944$ . Its underlying Gaussian random field is described by a Gaussian covariance function  $\exp(-r^2/l_c^2)$  with a correlation length  $l_c = 2$ , which significantly distinguishes  $\kappa_f$  from  $\kappa_t$ , see Fig. 10 a) and b). For the PCE computations we used polynomials with total order up to  $p = 3$  and  $M = 10$  KLE modes, while for the EnKF we chose 290 ensemble members, as this has approximately the same workload as the PCE.

In Fig. 10 c) and d) the true field is compared with the first and the fourth sequential posterior update obtained by the PCE based method. Since the prior is very different from the truth by all its properties the ‘shapes’ of the first update and the true field are still relatively distinct. However, the mean and the variance of the prior are clearly moved in the direction of the truth. Doing further updates with different loadings, the posterior field approaches the truth although some residual uncertainty remains. This

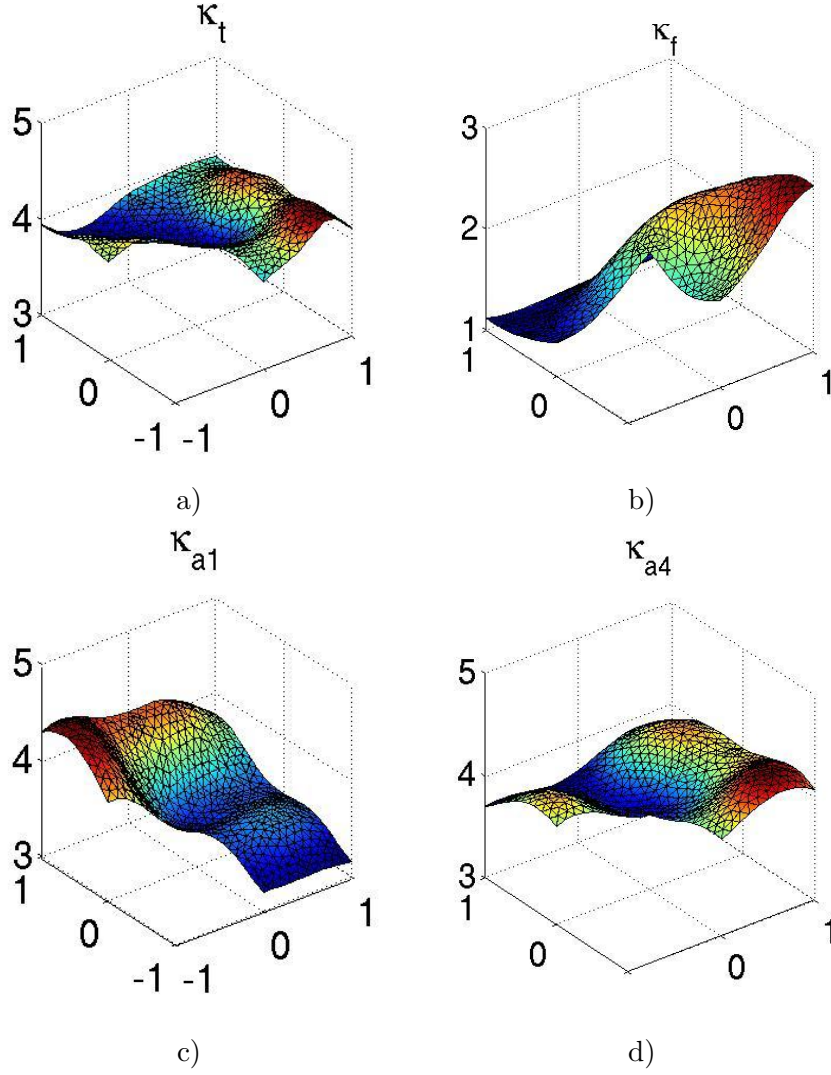


Figure 10: Direct PCE update ( $p = 3, M = 10$ ): a) Truth  $\kappa_t$ , b) prior, c) 1st update  $\kappa_{a1}$ , d) 4th update  $\kappa_{a4}$

Update	PCE			EnKF		
	$\varepsilon_m$	$\bar{\varepsilon}_a$	$\text{var}(\kappa_a)$	$\varepsilon_m$	$\bar{\varepsilon}_a$	$\text{var}(\kappa_a)$
A priori	0.53	0.42	0.8	0.53	0.42	0.8
1st update	0.18	0.14	0.13	0.19	0.15	0.08
2nd update	0.07	0.06	0.07	0.02	0.02	0.001
3rd update	0.06	0.04	0.07	0.02	0.02	5.38e-04
4th update	0.03	0.03	0.07	0.02	0.02	3.58e-04

Table 1: Comparison of relative error of the mode ( $\varepsilon_m$ ), relative error of the mean ( $\bar{\varepsilon}_a$ ), and the variance ( $\text{var}(\kappa_a)$ ) of the pdf in one point of the domain obtained by the direct (PCE) and the ensemble Kalman filter (EnKF) method.

behaviour is summarised in Table 1, where the PCE and EnKF results are shown for one point in the domain. Here one may notice that the EnKF method results in an unrealistically small variance (smaller than  $10^{-2}$  already in the second update), such that the truth is in a very low probability region. This undesirable behaviour, which gives such overconfident estimates of the residual uncertainty is one of the problems of the EnKF, and one has to take care not to be misled.

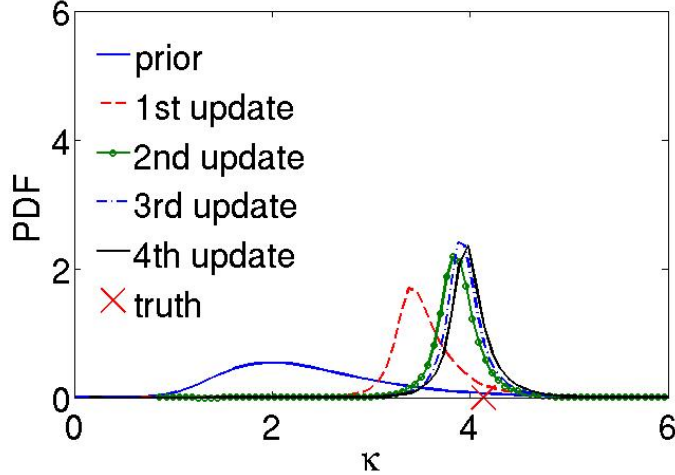


Figure 11: Probability density function (pdf) of the posterior for four updates by the direct (PCE) method.

In a similar vein, comparing the posterior pdfs of the EnKF and the PCE update in Fig. 11 and Fig. 12, one may observe that after first update the posteriors are fairly similar. However, in further updates (from second to fourth), the EnKF appears to be very certain about the posterior value, although the truth stays in a very low probability region. In contrast to this, the PCE update is able to keep a credible residual variance such that the truth stays inside a high probability area.

### 5.2.2 Elasto-plasticity with linear Bayesian update

To describe quasi-static elasto-plasticity with hardening, we are lead a bit beyond the format in Eq. (1) as one has to consider the non-smooth evolution of the internal parameters [25, 33, 34], which gives a variational inequality as a generalisation of the differential equation. The state variable is  $u = (v, \epsilon_p, \nu)$ , where  $v$  denotes the displacement field,  $\epsilon_p$  is the plastic deformation, and  $\nu$  the appropriate internal hardening variable. In a mixed formulation [33, 34] one has to consider at the same time a dual variable  $u^*$  of stress-like quantities. These quantities have to stay inside a non-empty closed convex set  $\mathcal{K}$ , the so-called elastic domain. The abstract mixed formulation then is to find functions  $u(t)$  and  $u^*(t)$  such that  $u^*(t) \in \mathcal{K}$  and

$$\begin{aligned} A(u(t)) + u^*(t) &= f(t) \\ \forall z^* \in \mathcal{K} : \quad \langle \dot{u}(t), z^* - u^*(t) \rangle &\leq 0. \end{aligned} \quad (42)$$

In this description, the first equation represents the equilibrium condition, while the second is the so-called flow rule, stating that the rate of change  $\dot{u}(t) = (\partial/\partial t)u(t)$  is in the normal cone of  $\mathcal{K}$  at  $u^*(t)$ ; for more details see [25, 33, 34].

Note that while  $u^*(t)$  is not on the boundary of  $\mathcal{K}$ , the stress-strain relationship stays linear and the model reduces to a linear equation mathematically similar to the diffusion equation, with  $\kappa$  replaced by a constitutive tensor described by the bulk  $K$  and shear modulus  $G$ . Here we investigate the identification of the shear modulus  $G$  via

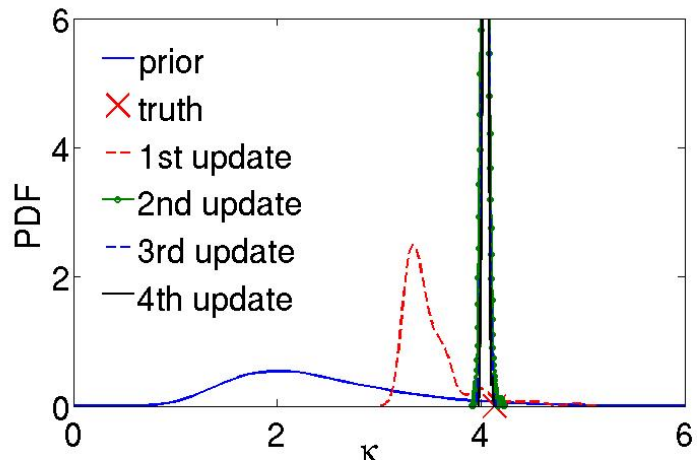


Figure 12: Probability density function (pdf) of the posterior for four updates by the EnKF method.

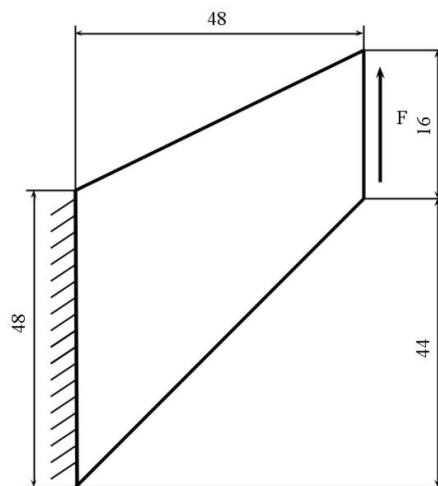


Figure 13: Geometry and loading of Cook's membrane test

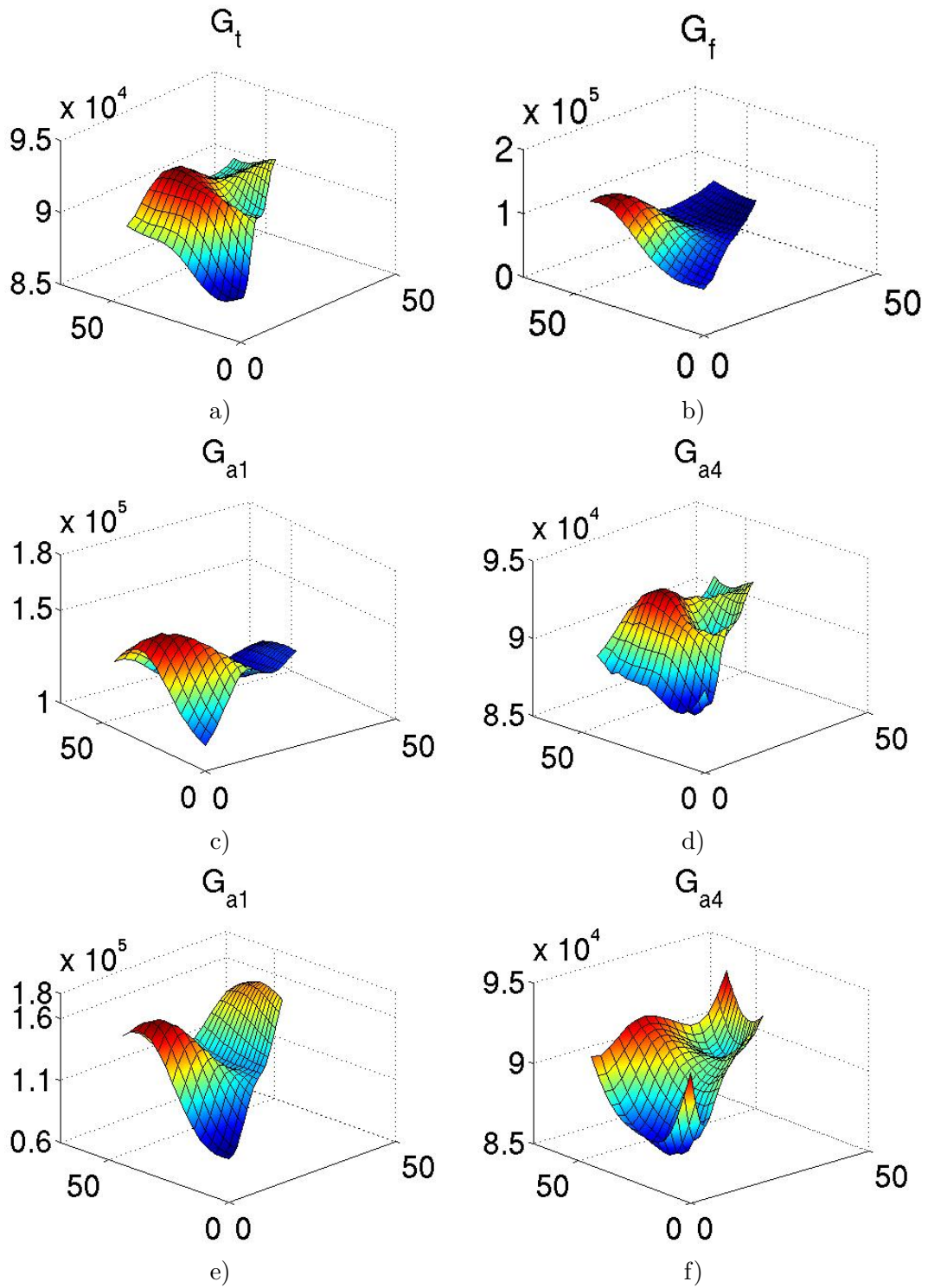


Figure 14: Shear modulus  $G$  via elastic response: a) truth, b) prior, c) first update, (linear elastic range), d) fourth update (linear elastic range), e) first update (nonlinear elasto-plastic range) f) fourth update (nonlinear elasto-plastic range).

Model	Update						
	A priori	1st	2nd	3rd	4th	5th	6th
Elastic	0.45	0.15	0.04	0.02	0.01	0.01	0.01
Elasto-Plastic	0.45	0.30	0.21	0.15	0.08	0.03	0.02

Table 2: The comparison of the relative root mean square error  $\varepsilon_a$  in each update for purely elastic and elasto-plastic response.

the linear sequential Bayesian procedure and shear stress measurements. The numerical tests are performed on a well-known and often used example.

Cook’s membrane is clamped on one end and loaded by a shear force in the vertical direction at the other end as shown in Fig. 13. The plate is discretised into 225 quadrilateral quadratic eight-noded serendipity finite elements (see [40]), of which 30% are chosen to place the measurements. For the sake of simplicity we do not investigate which nodes are the most appropriate. Instead, the measurement points are equally distributed over the domain, and the measurements are polluted by a measurement error modelled by independent Gaussian noise with zero mean with a diagonal covariance  $\sigma_\varepsilon^2 \mathbf{I}$ , and  $\sigma_\varepsilon$  approximately equal to 1% of the measured value.

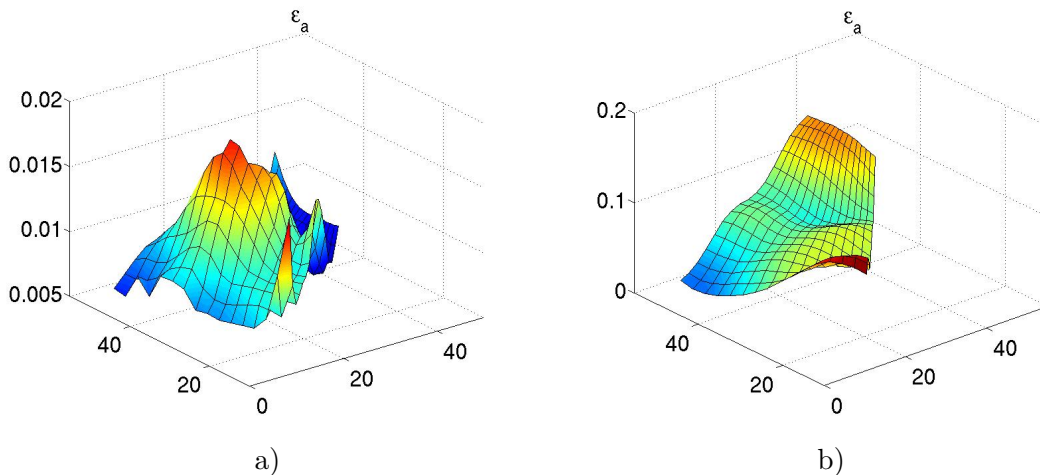


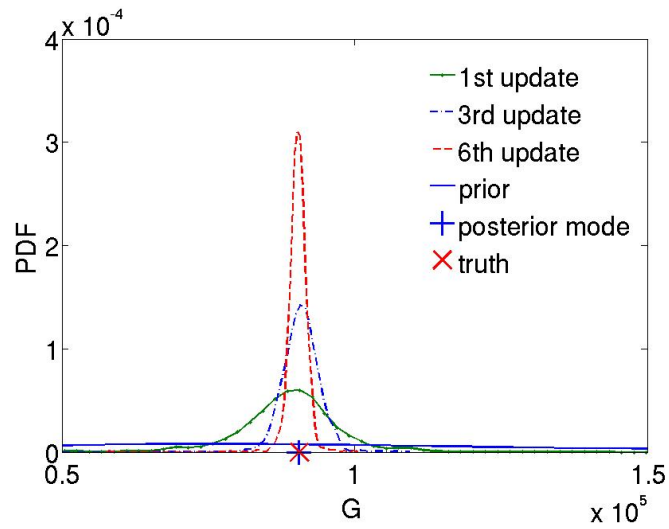
Figure 15: Comparison of relative root mean square error  $\varepsilon_a$  in the fourth update for a) purely elastic, b) elasto-plastic response.

In this particular example the measurement represents the shear stress  $\sigma_{xy}$  numerically computed in a virtual experiment with the adopted ‘true’ value of  $G_t$ . The truth  $G_t$  is taken to be one realisation of the modified lognormal field  $G_t = G_0 + G_1 \kappa_G$ , where  $\kappa_G := \exp(q_t)$ ,  $G_0 = 82760$  MPa,  $G_1 = 0.1$  MPa,  $\mathbb{E}(\kappa_G) = G_0$  and standard deviation  $\sigma_{\kappa_G} = 0.1G_0$ . The underlying Gaussian random field  $q_t$  is described by a Gaussian covariance function  $\exp(-r^2/l_c^2)$ , with  $r$  being the spatial distance and  $l_c = 30$  the correlation length.

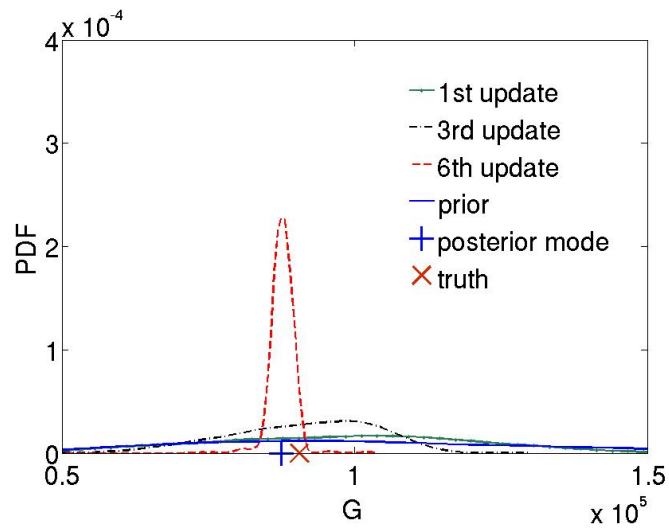
Similarly as for the previous diffusion example in Section 5.2.1, the shear modulus has to be positive and the prior distribution is hence assumed to be lognormal, though with different characteristics than for the truth. Namely, the prior is assumed to be the lognormal random field  $G_f = \exp(q_f)$  with  $\mathbb{E}(G_f) = 107590$  MPa, a standard deviation of  $0.4G_0$ , and an underlying covariance function of  $q_f$  given by  $\exp(-r/l_c)$  with  $l_c = 20$ . In this way the realisation of the prior is very different from the realisation of the true field as shown in Fig. 14.

We compute the direct linear Bayesian update via polynomial chaos expansion of order  $p = 3$  and  $M = 10$  random variables. The updated field, i.e. the posterior, is then adopted for the new prior in the next sequential update, see Fig. 9. Thereby one introduces new information into update process by changing the value of the right hand





a)



b)

Figure 16: Comparison of the posterior and prior pdfs over the updates: a) elastic (linear) response, b) elasto-plastic (nonlinear) response.

side. With each update the force decreases or increases in every second update for 20% starting from 1 MN/m.

The sequential update is done in six stages using both purely elastic and elasto-plastic response. As expected, the nonlinearity is detrimental to the updating and the posterior obtained via elastic response is better than the one from an elasto-plastic model as can be seen in Table 2, where the relative root mean square error reduces faster with each update for purely elastic response. This is confirmed from the plots of the relative error over the domain. In Fig. 15 one may see that the elasto-plastic response has 10 times bigger error than the elastic response. In addition, the error is similar in almost all parts of domain besides the region where the plastification starts to occur.

Finally in Fig. 16 we compare the pdfs for different updates for both the elastic/linear and elasto-plastic (nonlinear) response. The overall picture is similar to what was observed before, namely that the severe nonlinearity in the elasto-plastic model makes the identification process much harder.

## 6 Conclusion

The problem of identifying parameters or quantities in a computational model by comparison with either real world measurements or other computational models (e.g. more refined models) is an old and frequent one. Practically all approaches start from the idea that the choice of parameters should be such as to minimise a certain error functional. Classical methods to do this lead to regularisation. Another point of view—taken here—is to embed the unknown quantity in a probability distribution, where the spread of the probability distribution should reflect the uncertainty about that quantity.

We have, starting from the elementary textbook formulation of Bayes’s theorem, shown how it connects to conditional expectation. Conditional expectation has the minimisation of variance at its background. This in turn gives rise to a computational characterisation of updates. By approximating the space of all measurable functions through the subspace of linear functions, we disregard a certain amount of information, but we are rewarded with a simple computation. Our result contains the well-known Kalman filter as a special case.

We then discuss how these theoretical constructs can be implemented in a real computation. The first group is based on the classical formula for measures and probability densities. We sketch the Markov chain Monte Carlo (MCMC) methods, which may be used in this case. We also discuss how the Monte Carlo sampling may be accelerated through the use of functional approximations for the stochastic part. We here use standard Hermite polynomials in Gaussian RVs, i.e. Wiener’s polynomial chaos.

The second group is based on the conditional expectation idea, leaving the underlying measure unchanged and updating the relevant RV. We show how this idea may be implemented in a sampling Monte Carlo manner—i.e. the ensemble Kalman filter (EnKF)—or alternatively in the PCE setting. We then demonstrate the workings of all these methods on some simple examples of slowly increasing difficulty. We find that MCMC needs a huge number of samples, EnKF needs considerably less. On the other hand the variance estimate of EnKF seems to be often erring on the optimistic side, thus severely underestimating the residual uncertainty. The PCE based methods on the other hand do not seem to suffer from this illness and give much more reliable results at a comparable workload. Our final identification problem is a very difficult one, an elasto-plastic system, or mathematically speaking a variational inequality. The non-smoothness inherent in such problems slows also the PCE based methods down, but they still succeed. In our view these PCE based methods show great promise for these taxing parameter identification problems.

## References

- [1] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2006. ISSN 0895-4798.
- [2] Suhrid Balakrishnan, Amit Roy, Marianthi G. Ierapetritou, Gregory P. Flach, and Panos G. Georgopoulos. Uncertainty reduction and characterization for complex environmental fate and transport models: An empirical Bayesian framework incorporating the stochastic response surface method. *Water Resources Research*, 39(12):1350, December 2003. ISSN 0043-1397. doi: 10.1029/2002WR001810.
- [3] Emmanuel D. Blanchard, Adrian Sandu, and Corina Sandu. A polynomial chaos-based Kalman filter approach for parameter estimation of mechanical systems. *Journal of Dynamic Systems, Measurement, and Control*, 132(6):061404, 2010. doi: 10.1115/1.4002481.
- [4] Adam Bobrowski. *Functional Analysis for Probability and Stochastic Processes*. Cambridge University Press, Cambridge, 2005. ISBN 0-521-53937-4.
- [5] S. Boyaval, C. Le Bris, T. Lelièvre, Y. Maday, N. C. Nguyen, and A. T. Patera. Reduced basis techniques for stochastic problems. *Archives of Computational Methods in Engineering*, 17:435–454, 2010. doi: 10.1007/s11831-010-9056-z.
- [6] J. A. Christen and C. Fox. MCMC using an approximation. *Journal of Computational and Graphical Statistics*, 14(4):795–810, 2005. doi: 10.1198/106186005X76983.
- [7] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*. Kluwer, Dordrecht, 2000.
- [8] Geir Evensen. *Data Assimilation — The Ensemble Kalman Filter*. Springer-Verlag, Berlin, 2<sup>nd</sup> edition, 2009.
- [9] Geir Evensen. The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Systems Magazine*, 29:82–104, 2009. doi: 10.1109/MCS.2009.932223.
- [10] D. Gamerman and H. F. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall/CRC, 2006. ISBN 1584885874.
- [11] Michael Goldstein and David Wooff. *Bayes Linear Statistics—Theory and Methods*. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester, 2007.
- [12] W. Hackbusch, B. N. Khoromskij, and E. E. Tyrtysnikov. Approximate iterations for structured matrices. *Numerische Mathematik*, 109:365–383, 2008. doi: 10.1007/s00211-008-0143-0.
- [13] Svante Janson. *Gaussian Hilbert spaces*. Cambridge Tracts in Mathematics, 129. Cambridge University Press, Cambridge, 1997.
- [14] E. T. Jaynes. *Probability Theory, The Logic of Science*. Cambridge University Press, Cambridge, 2003.
- [15] R. E. Kálmán. A new approach to linear filtering and prediction problems. *Transactions of the ASME—J. of Basic Engineering (Series D)*, 82:35–45, 1960.
- [16] A. Kučerová and H. G. Matthies. Uncertainty updating in the description of heterogeneous materials. *Technische Mechanik*, 30(1-3):211–226, 2010.

- [17] A. Kučerová, J. Sýkora, B. Rosić, and H. G. Matthies. Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials. *arXiv:1110.2049v1 [cs.CE]*, 2011. URL <http://arxiv.org/abs/1110.2049>.
- [18] Jia Li and Dongbin Xiu. A generalized polynomial chaos based ensemble Kalman filter with high accuracy. *Journal of Computational Physics*, 228(15):5454–5469, August 2009. ISSN 0021-9991. doi: 10.1016/j.jcp.2009.04.029.
- [19] David G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, Chichester, 1969.
- [20] Xiang Ma and Nicholas Zabaras. An efficient Bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method. *Inverse Problems*, 25(3):035013, 2009. ISSN 0266-5611. doi: 10.1088/0266-5611/25/3/035013.
- [21] N. Madras. *Lectures on Monte Carlo Methods*. American Mathematical Society, Providence, RI, 2002.
- [22] Youssef Marzouk and Dongbin Xiu. A stochastic collocation approach to Bayesian inference in inverse problems. *Communications in Computational Physics*, 6(4): 826–847, 2009.
- [23] Youssef M. Marzouk, Habib N. Najm, and Larry A. Rahn. Stochastic spectral methods for efficient Bayesian solution of inverse problems. *Journal of Computational Physics*, 224(2):560–586, 2007. ISSN 0021-9991. doi: 10.1016/j.jcp.2006.10.010.
- [24] H. G. Matthies. Uncertainty quantification with stochastic finite elements. In Erwin Stein, René de Borst, and Thomas J. R. Hughes, editors, *Encyclopaedia of Computational Mechanics*. John Wiley & Sons, Chichester, 2007. doi: 10.1002/0470091355.ecm071.
- [25] H. G. Matthies and B. Rosić. Inelastic media under uncertainty: stochastic models and computational approaches. In B. D. Reddy, editor, *IUTAM Symposium on Theoretical, Computational and Modelling Aspects of Inelastic Media*, volume 11 of *IUTAM Bookseries*, pages 185–194. Springer Verlag, Berlin, 2008.
- [26] H. G. Matthies and E. Zander. Solving stochastic systems with low-rank tensor compression. *Linear Algebra and its Applications*, page accepted, 2010.
- [27] Hermann G. Matthies, Alexander Litvinenko, Oliver Pajonk, Bojana V. Rosić, and Elmar Zander. Parametric and uncertainty computations with tensor product representations. In Andrew Dienstfrey and Ronald Boisvert, editors, *Uncertainty Quantification in Scientific Computing*, IFIP Advances in Information and Communication Technology, Berlin, 2011. Springer.
- [28] T. A. El Moselhy and Y. M. Marzouk. Bayesian inference with optimal maps. *submitted to Journal of Computational Physics*, 2011.
- [29] A. Nouy. Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems. *Archives of Computational Methods in Engineering*, 17:403–434, 2010. doi: 10.1007/s11831-010-9054-1.
- [30] Anthony Nouy and Olivier P. Le Maître. Generalized spectral decomposition for stochastic nonlinear problems. *Journal of Computational Physics*, 228(1):202–235, January 2009. ISSN 0021-9991. doi: 10.1016/j.jcp.2008.09.010.
- [31] Oliver Pajonk, Bojana V. Rosić, Alexander Litvinenko, and Hermann G. Matthies. A deterministic filter for non-Gaussian Bayesian estimation. Informatikbericht 2011-04, TU Braunschweig, Brunswick, 2011. URL <http://www.digibib.tu-bs.de/?docid=00038994>.

- [32] B.L. Pence, H.K. Fathy, and J.L. Stein. A maximum likelihood approach to recursive polynomial chaos parameter estimation. In *American Control Conference (ACC)*, pages 2144–2151, 2010.
- [33] B. Rosić and H. G. Matthies. Computational approaches to inelastic media with uncertain parameters. *J. Serbian Soc. Comp. Mech.*, 2:28–43, 2008.
- [34] B. V. Rosić, H. G. Matthies, and M. Živković. Mathematical formulation and numerical simulation of stochastic elasto-plastic behaviour. In *IV European Conference on Computational Mechanics*, Paris, 2010. URL [http://www.eccm2010.org/complet/fullpaper\\_489.pdf](http://www.eccm2010.org/complet/fullpaper_489.pdf).
- [35] Bojana V. Rosić, Alexander Litvinenko, Oliver Pajonk, and Hermann G. Matthies. Direct Bayesian update of polynomial chaos representations. Informatikbericht 2011-02, TU Braunschweig, Brunswick, 2011. URL <http://www.digibib.tu-bs.de/?docid=00039000>.
- [36] G. Saad and R. Ghanem. Characterization of reservoir simulation models using a polynomial chaos-based ensemble Kalman filter. *Water Resources Research*, 45: W04417, 2009. doi: 10.1029/2008WR007148.
- [37] A. M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19: 451–559, 2010. doi: 10.1017/S0962492910000061.
- [38] Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, 2004.
- [39] L. Zeng and D. Zhang. A stochastic collocation based Kalman filter for data assimilation. *Comput Geosci*, 14:721–744, 2010. doi: 10.1007/s10596-010-9183-5.
- [40] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method*. Butterworth-Heinemann, Oxford, 5th ed., 2000.