Glossary

Conditional probability distribution	Conditional probability distribution $\pi(y\|x)$ provides the plausibility of proposition y, given proposition x.
Correlation	A general term for the dependence between pairs of random variables.
Correlation coefficient	A measure for the strength of the dependence between pairs of random variables.
Covariance	A measure that shows how two random variables depend on each other.
Covariance matrix	A symmetric matrix in which the off-diagonal elements are covariances of pairs of random variables and the diagonal elements are variances of random variables.
Credible interval (region)	An interval (or a region in the multivariate case) of a distribution in which it is believed that one or more random variables (parameters in this study) lie with a certain probability.
Dependence and independence	Two events are statistically independent if the occurrence of one has no influence on the probability of the occurrence of the other one (i.e. $\pi(x) = \pi(x\|y)$). They are dependent if the occurrence of one has an influence on the probability of the occurrence of the other one (i.e. $\pi(x) \neq \pi(x\|y)$).
Event	A set of outcomes of an experiment.
Joint distribution	A multivariate distribution.
Laplace approximation	An approximation of a distribution with a Gaussian distribution centred at the MAP.
Likelihood function	If the conditional probability distribution $\pi(y\|x)$ is regarded as a function of x for given fixed y, the function is called a likelihood function. The likelihood describes the plausibility of a parameter, given observations.
Marginal distribution	A probability distribution as a function of a single variable or a combination of subsets of variables associated with a multivariate distribution (e.g. $\pi(x)$, $\pi(y)$, $\pi(x, y)$, $\pi(x, z)$ and $\pi(y, z)$, for joint distribution $\pi(x, y, z)$). A marginal distribution is obtained by integrating a multivariate distribution over one or more (but not all) other variables.

Markov chain	A stochastic model to describe a sequence of events in which the probability of each event only depends on the previous event.
Markov chain Monte Carlo (MCMC) methods	A set of techniques to draw samples (i.e. simulate observations) from probability distributions by the construction of a Markov chain.
Maximum a posteriori probability (MAP) point	A point at which the posterior distribution is (globally) maximum.
Mean (expected value)	A measure for the central value of the underlying distribution.
Multivariate distribution	A probability distribution of two or more random variables.
Point estimate	A scalar that measures a feature of a population, e.g. the mean value, the MAP point.
Population	The total set of all possible observations that can be made.
Posterior distribution (posterior)	The probability distribution that describes one's knowledge about a random variable (parameter in this study) after obtaining new measurements.
Posterior predictive distribution (PPD)	The distribution of unobserved measurements (observations), given the measured (observed) data.
Prior distribution (prior)	The probability distribution that describes one's a-priori knowledge about a random variable (parameter in this study).
Probability	The likelihood (or plausibility) that a certain event occurs.
Probability density function (PDF)	The equation that describes a continuous probability distribution.
Probability distribution	A function that provides the probabilities of the occurrence of the possible outcomes of an experiment.
Random sample	A randomly chosen sample.
Random variable	A variable of which the value depends on the outcome of a random experiment.
Realisation	The value that a random variable takes or the outcome of an experiment after its occurrence.
Sample	A set of observations from a population with the purpose of investigating particular properties of the population.
Standard deviation	A measure for the possible deviation of a random variable from its mean. Large standard deviations indicate large possible differences; and vice versa.

Validation point	A measurement (observation) used to assess the quality of a prediction based on the identified parameters, that is not used for the identification itself.
Variance	The standard deviation squared.

More definitions on statistical concepts can be found in [1] and [2].

A tutorial on Bayesian inference to identify material parameters in solid mechanics

H. Rappel[a,b], L.A.A. Beex[a], J.S. Hale[a], L. Noels[b], S.P.A. Bordas[a,c,*]

[a] *Institute of Computational Engineering, Faculty of Science, Technology and Communication, University of Luxembourg, Maison du Nombre, 6, Avenue de la Fonte, 4364, Esch-sur-Alzette, Luxembourg.*
[b] *Computational & Multiscale Mechanics of Materials (CM3), Department of Aerospace and Mechanical Engineering, University of Liège, Quartier Polytech 1, Allée de la Découverte 9, B-4000 Liège, Belgium.*
[c] *School of Engineering, Cardiff University, Queen's Buildings, The Parade, Cardiff CF243AA, Wales, UK.*

Abstract

The aim of this contribution is to explain in a straightforward manner how Bayesian inference can be used to identify material parameters of material models for solids. Bayesian approaches have already been used for this purpose, but most of the literature is not necessarily easy to understand for those new to the field. The reason for this is that most literature focuses either on complex statistical and machine learning concepts and/or on relatively complex mechanical models. In order to introduce the approach as gently as possible, we only focus on stress-strain measurements coming from uniaxial tensile tests and we only treat elastic and elastoplastic material models. Furthermore, the stress-strain measurements are created artificially in order to allow a one-to-one comparison between the true parameter values and the identified parameter distributions.

Keywords: Bayesian inference, Bayes' theorem, stochastic identification, statistical identification, parameter identification, elastoplasticity, plasticity

1. Introduction

The most commonly used approach to identify parameters of mechanical descriptions for solid materials is to formulate an error function that measures the difference between the model response and the experimental data [3]. This error function is then minimised with respect to the material parameters in order to determine their most suitable values. Such an approach provides a deterministic estimate of parameter values, unable to account for the unavoidable uncertainties of each parameter associated with experimental observations.

An alternative, and rather different approach is to use Bayesian inference (BI). Using Bayes' theorem, a probability density function (PDF), the so-called posterior distribution (or the posterior for short) can be formulated as a function of the material parameters of interest. Subsequently, the PDF is analysed to determine relevant summaries, such as the mean of the material parameters, the material properties at which the PDF is maximum (called the 'maximum-a-posteriori-probability' point, or MAP for short) and the covariance matrix (i.e. a matrix that measures the correlation between the parameters). The PDF can only be explored analytically for a limited number of straightforward cases. Hence, numerical methods are commonly employed, e.g. Markov chain Monte Carlo (MCMC) techniques [4–7]. An alternative is to first approximate the PDF (e.g. by a Laplace approximation) and then determine the statistical summaries of the approximated distribution (e.g. mean, MAP point and covariance matrix) [8, 9].

In contrast to deterministic identification approaches, approaches using BI can quite straightforwardly incorporate several uncertainty sources, such as the noises coming from different experimental devices, as well as the uncertainty caused by the fact that the model cannot perfectly capture reality (i.e. the model uncertainty) [10]. In addition, all parameter values come with their individual uncertainty, that is also

*Corresponding author
Email address: stephane.bordas@alum.northwestern.edu (S.P.A. Bordas)

measured in terms of the parameter values. This is in contrast to conventional deterministic identification approach which come with one residual as a measure for the full set of parameters, which is obviously not expressed in terms of the parameters themselves. Additionally, BI provides an intrinsic statistical regularisation which makes inverse problems with limited observations solvable [11]. On the other hand, applying Bayes' theorem for material parameter identification does require the measurement noise to be known, i.e. the noise distributions and their parameters must be established. One must note that the noise distribution and its parameter values can be treated as unknown parameters that need to be identified and hence, they appear in the posterior. The numerical techniques to analyse the posterior PDFs may furthermore need careful attention.

The developments of BI in the field of parameter identification for mechanical models started with the identification of elastic constants. Isenberg [12] proposed a Bayesian approach for the identification of elastic parameters in 1979. Various researchers subsequently used the framework to identify elastic material parameters based on dynamic responses [8, 13, 14]. Lai and Ip [15] used BI to identify the elastic properties of a thin composite plate. Daghia et al. [16] used the Bayesian framework for the identification of the elastic constants of thick laminated composite plates in a dynamical setting. Koutsourelakis [17] used Bayesian inference to identify spatially varying elastic material parameters. In 2010, Gogu et al. [3] presented an introduction in the Bayesian approach for the identification of elastic constants, and compared the results with those of a deterministic identification approach. The influence of the prior distributions was however not systematically studied. In another study, Gogu et al. [18] used a Bayesian framework to identify elastic constants in multi-directional laminates.

BI is also used for the parameter identification of nonlinear constitutive models. Muto and Beck [19] and Liu and Au [20] applied the approach to hysteretic models, whereas Fitzenz et al. [21] used BI to identify parameters of a creep model of quartz. Most [22] used a Bayesian updating procedure for the parameter identification of an elastoplastic model without hardening (perfect plasticity). Rosić et al. [23] used linear Bayesian updating via polynomial chaos expansion for an elastoplastic system. BI is also employed for the identification of viscoelastic material parameters in [24, 25].

Another study that uses Bayes' theorem to identify material parameters is the work of Nichols et al. [26]. They employed the theorem to identify the nonlinear stiffness of a dynamic system. Furthermore, Nichols et al. [26] used the method to find the location, size and depth of delamination in a composite beam. Abhinav and Manohar [27] used BI to characterise the dynamic parameters of a structural system with geometrical nonlinearities. The approach is also employed to assess the quality of different models with respect to measured data (i.e. model selection): e.g. hyperelastic constitutive models for soft tissue [28], phenomenological models for tumour growth [29], models for damage progression in composites due to fatigue [30] and fatigue models for metals [31]. Sarkar et al. [32] used the Bayesian method to identify thermodynamical parameters of cementitious materials. BI is also used in the fields of heat transfer and fluid mechanics for inverse problems [5, 33].

Bayesian inference relies on concepts that may be complex to grasp for those who are only familiar with deterministic identification methods. The primary objective of this contribution is to show how Bayesian inference can be applied for the stochastic identification of material parameters. We focus on elastoplastic material models in this contribution for two reasons: (1) they are widely used in the field of mechanics, and (2) the family of elastoplastic material models contains both simple (linear elasticity) and more complex, nonlinear, C_0-continuous descriptions (elastoplasticity with nonlinear hardening). Our contribution focuses on results of uniaxial tensile tests in order to be as straightforward as possible. In addition to introducing the general idea of identification approaches based on BI as gently as possible (to our abilities), we also discuss some more complex extensions, such as not only incorporating the error in the stress, but also incorporating the error in the strain and the uncertainty of the model. Those extensions are not discussed in much detail, but references to other works are included.

The structure of this article is as follows. Section 2 briefly discusses the employed material models in this contribution. Section 3 discusses the theoretical fundaments behind Bayes' theorem. Section 3 also describes a Bayesian approach for the stochastic identification of elastoplastic material parameters, if only the stress measurements include stochastic errors. In Section 4, MCMC methods as the numerical techniques

to analyse the posterior distribution are explained. In this section, we also explain the posterior predictive distribution (PPD) as an approach for predicting unobserved measurements. In Section 5, a considerable number of results are presented. In Section 6, some advanced concepts, such as incorporating the error in the strain and the model uncertainty, are discussed. We also briefly discuss how the approach differs if a viscoelastic material model is used instead. In Section 7, conclusions are presented.

Remark 1. *As mentioned before, we only consider stress-strain data coming from uniaxial tensile tests in this contribution. As the force is measured in tensile tests, the measured stress and its error is assumed to be proportional to the measured force and its error (the parameters of its distribution are to be identified using separate calibration experiments). Strain measurements are commonly based either on the clamp displacement or determined using digital image correlation (DIC). In both cases, the parameters of the error distribution of the strain can be determined using calibration experiments.*

Remark 2. *Throughout this paper bold letters and symbols denote vectors and matrices. Capitals furthermore denote random variables.*

2. Material models

In this contribution, BI is developed to identify the parameters of four one-dimensional material models: linear elasticity, linear elasticity with perfect plasticity, linear elasticity with linear hardening and linear elasticity with nonlinear hardening. Hardening is considered to be isotropic and associative. For each model, the identification is based on the results of monotonic uniaxial tensile tests. Below, material responses are given for monotonic tensile loading.

2.1. Linear elasticity

The linear elastic model assumes a linear relationship between the stress and the strains. In the case of uniaxial tension/compression, this writes:

$$\sigma(\epsilon, \mathbf{x}) = E\epsilon, \tag{1}$$

where σ denotes the stress, ϵ the strain, $\mathbf{x}$ the material parameter vector (here $\mathbf{x} = E$) and E the Young's modulus.

2.2. Linear elasticity-perfect plasticity

The linear elastic-perfectly plastic model neglects the effect of work hardening, assuming that purely plastic deformation occurs when the stress reaches its yield value. The total strain (ϵ) in this contribution is additively split into an elastic part, ϵ_e, and a plastic part, ϵ_p:

$$\epsilon = \epsilon_e + \epsilon_p, \tag{2}$$

and the stress is defined as a function of the elastic strain, ϵ_e:

$$\sigma(\epsilon, \mathbf{x}) = E\epsilon_e = E(\epsilon - \epsilon_p). \tag{3}$$

The yield condition at which plastic yielding occurs, is written as:

$$f(\sigma) = |\sigma| - \sigma_{y0} \leq 0, \tag{4}$$

where σ_{y0} denotes the initial yield stress and f the yield function. Consequently, $\mathbf{x} = \begin{bmatrix} E & \sigma_{y0} \end{bmatrix}^T$.

The flow rule for the plastic strain is furthermore expressed as:

$$\dot{\epsilon}_p = \dot{\alpha}\frac{\partial f}{\partial \sigma}, \tag{5}$$

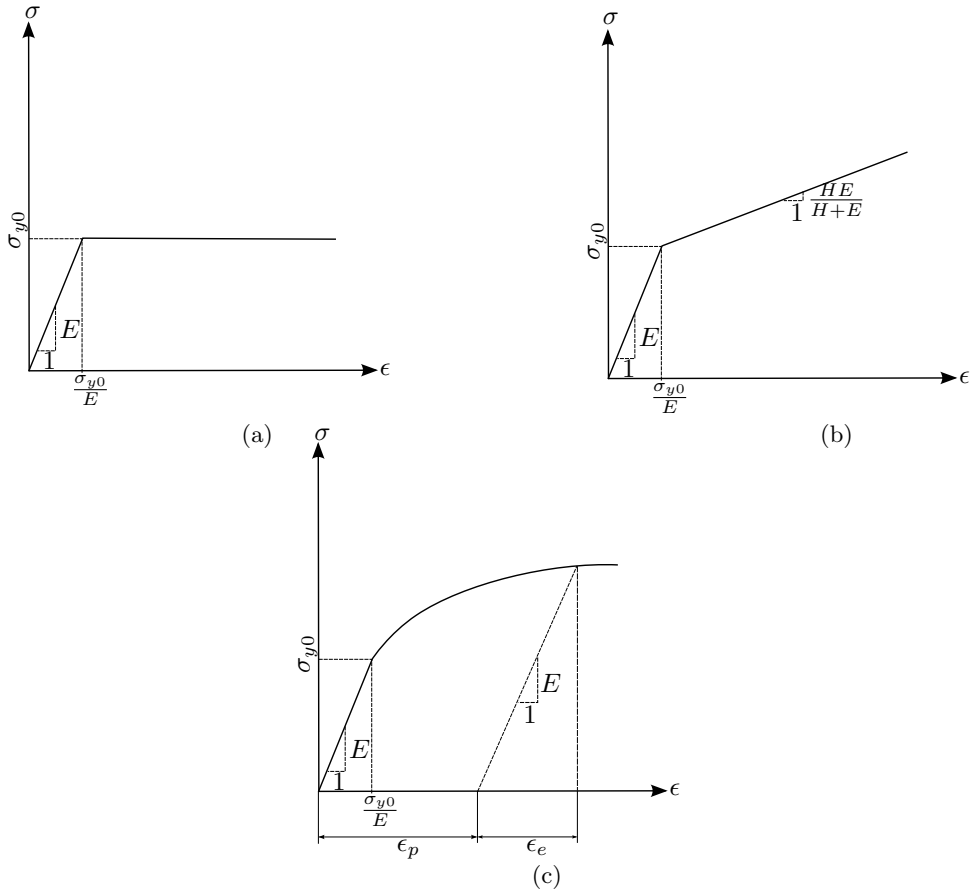


Figure 1: The stress-strain response of (a) linear elastic-perfectly plastic model, (b) linear elastic-linear hardening model and (c) linear elastic-nonlinear hardening model during monotonic tension.

where α denotes the cumulative plastic strain. Finally, the Kuhn-Tucker conditions [34] ensure that dissipation is irreversible:

$$\dot{\alpha} \geq 0, \quad f(\sigma) \leq 0, \quad \dot{\alpha} f(\sigma) = 0. \tag{6}$$

The stress-strain response of the linear elastic-perfectly plastic model during monotonic tension can be written as:

$$\sigma(\epsilon, \mathbf{x}) = \begin{cases} E\epsilon & \text{if } \epsilon \leq \frac{\sigma_{y0}}{E} \\ \sigma_{y0} & \text{if } \epsilon > \frac{\sigma_{y0}}{E} \end{cases}. \tag{7}$$

Using the Heaviside step function ($h(\cdot)$), Eq. (7) can alternatively be expressed as:

$$\sigma(\epsilon, \mathbf{x}) = E\epsilon\Big(1 - h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big)\Big) + \sigma_{y0} h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big), \tag{8}$$

Fig. 1(a) presents this response graphically.

2.3. Linear elasticity-linear hardening

The linear elastic-linear hardening model is identical to the linear elastic-perfectly plastic model, except for the yield function, which writes:

$$f(\sigma) = |\sigma| - \sigma_{y0} - H\alpha \leq 0, \tag{9}$$

where H denotes the plastic modulus. Hence, $\mathbf{x} = \begin{bmatrix} E & \sigma_{y0} & H \end{bmatrix}^T$.

Consequently, the stress-strain response of the model during monotonic tension writes:

$$\sigma(\epsilon, \mathbf{x}) = \begin{cases} E\epsilon & \text{if } \epsilon \leq \frac{\sigma_{y0}}{E} \\ \sigma_{y0} + H\epsilon_p & \text{if } \epsilon > \frac{\sigma_{y0}}{E} \end{cases}, \tag{10}$$

which can again be expressed using the Heaviside step function:

$$\sigma(\epsilon, \mathbf{x}) = E\epsilon\Big(1 - h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big)\Big) + \Big(\sigma_{y0} + \frac{HE}{H+E}\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big)\Big)h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big). \tag{11}$$

Fig. 1(b) shows this response graphically.

2.4. Linear elasticity-nonlinear hardening

The linear elastic-nonlinear hardening model also only differs from the linear elastic-perfectly plastic model through the yield function, which writes:

$$f(\sigma) = |\sigma| - \sigma_{y0} - H\alpha^n \leq 0, \tag{12}$$

where n is an additional plastic material parameter and hence, $\mathbf{x} = \begin{bmatrix} E & \sigma_{y0} & H & n \end{bmatrix}^T$.

For monotonic uniaxial tension, the stress-strain response can be written as:

$$\sigma(\epsilon, \mathbf{x}) = \begin{cases} E\epsilon & \text{if } \epsilon \leq \frac{\sigma_{y0}}{E} \\ \sigma_{y0} + H\epsilon_p^n & \text{if } \epsilon > \frac{\sigma_{y0}}{E} \end{cases}, \tag{13}$$

or using the Heaviside step function:

$$\sigma(\epsilon, \mathbf{x}) = E\epsilon\Big(1 - h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big)\Big) + \Big(\sigma_{y0} + H\Big(\epsilon - \frac{\sigma(\epsilon, \mathbf{x})}{E}\Big)^n\Big)h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big). \tag{14}$$

Fig. 1(c) shows this stress-strain response.

It is worth noting that Eq. (14) is an implicit function of the stress ($\sigma(\epsilon, \mathbf{x})$ appears both on the left hand side and right hand side of Eq. (14) and cannot analytically be determined if ϵ is known). This is in contrast to the stress-strain expressions of the previous material models for monotonic tension (Eqs. (1), (8) and (11)), which are all explicit functions (i.e. $\sigma(\epsilon, \mathbf{x})$ can analytically be computed when one has ϵ).

3. Bayesian inference

3.1. Concepts

We start by considering random events A and B, and the discrete probabilities of each event: $\mathrm{P}(a)$ and $\mathrm{P}(b)$. The probability that events A and B both occur, is given by the joint probability, $\mathrm{P}(a, b)$, which can be expanded as:

$$\mathrm{P}(a, b) = \mathrm{P}(a|b)\mathrm{P}(b) = \mathrm{P}(b|a)\mathrm{P}(a), \tag{15}$$

where $\mathrm{P}(a|b)$ and $\mathrm{P}(b|a)$ are conditional probabilities. Conditional probability $\mathrm{P}(a|b)$ expresses the probability that event A occurs, if it is certain that event B occurs. Using Eq. (15), the simplest form of Bayes' theorem can be written as:

$$\mathrm{P}(a|b) = \frac{\mathrm{P}(a)\mathrm{P}(b|a)}{\mathrm{P}(b)}. \tag{16}$$

If one regards two continuous random variables $\mathbf{X} \in \mathbb{R}^{n_p}$ and $\mathbf{Y} \in \mathbb{R}^{n_m}$, instead of discrete variables, where $\mathbf{X}$ denotes a random vector with n_p unknown parameters and $\mathbf{Y}$ a random vector with n_m measurements, Eq. (16) can be rewritten in terms of the following probability distribution functions (where π denotes a PDF):

$$\pi(\mathbf{x}|\mathbf{y}) = \frac{\pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x})}{\pi(\mathbf{y})}, \tag{17}$$

where $\pi(\mathbf{x})$, $\pi(\mathbf{y}|\mathbf{x})$ and $\pi(\mathbf{x}|\mathbf{y})$ are referred to as the prior distribution (i.e. the PDF that includes one's prior knowledge), the likelihood function (i.e. the PDF of the observed data $\mathbf{y}$, given unknown parameters $\mathbf{x}$) and the posterior distribution (i.e. the PDF of the unknown parameters $\mathbf{x}$, given the observations $\mathbf{y}$), respectively.

Using the law of total probabilities [35] which relates the marginal probabilities ($\pi(\mathbf{x})$ and $\pi(\mathbf{y})$) to the conditional probabilities ($\pi(\mathbf{y}|\mathbf{x})$), the denominator in Eq. (17) can be written as:

$$\pi(\mathbf{y}) = \int_{\mathbb{R}^{n_p}} \pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x}) d\mathbf{x}. \tag{18}$$

Since the data ($\mathbf{y}$) is already measured, the denominator in Eq. (17) is a positive constant number, $C \in \mathbb{R}^+$. This constant number can be regarded as a normalisation factor that ensures that the integral of the posterior ($\pi(\mathbf{x}|\mathbf{y})$) over $\mathbf{x}$ equals 1:

$$\pi(\mathbf{x}|\mathbf{y}) = \frac{1}{C}\pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x}). \tag{19}$$

Hence, one can rewrite Eq. (19) as:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x}). \tag{20}$$

Note that the statistical summaries of the posterior distribution (i.e. the mean, the MAP point and the covariance matrix) do not depend on the absolute posterior, but only on its shape.

In order to obtain the posterior in Eq. (20) (i.e. the PDF of the vector of unknown parameters, given the observations $\pi(\mathbf{x}|\mathbf{y})$), the likelihood function ($\pi(\mathbf{y}|\mathbf{x})$) and the prior ($\pi(\mathbf{x})$) need to be formulated. First, the likelihood function is considered.

In order to construct the likelihood function, a noise model has to be formulated and a noise distribution (π_{noise}) has to be determined. For the moment, we assume that the noise distribution is known (including its parameters). The noise model used in this study is additive, which is frequently employed, amongst others in [11, 16]. The additive noise model can be written as follows:

$$\mathbf{Y} = \mathbf{f}(\mathbf{X}) + \mathbf{\Omega}, \tag{21}$$

where $\mathbf{X} \in \mathbb{R}^{n_p}$ denotes again the vector with the unknown material parameters, $\mathbf{Y} \in \mathbb{R}^{n_m}$ the vector with the measured data and $\mathbf{\Omega} \in \mathbb{R}^{n_m}$ the noise vector. $\mathbf{f} : \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_m}$ denotes the material description and is a function of the unknown material parameters ($\mathbf{X}$). Given realisations $\mathbf{X} = \mathbf{x}$ and $\mathbf{\Omega} = \boldsymbol{\omega}$, and assuming that the parameters ($\mathbf{X}$) and the error ($\mathbf{\Omega}$) are statistically independent, the likelihood function reads:

$$\pi(\mathbf{y}|\mathbf{x}) = \pi_{\text{noise}}(\mathbf{y} - \mathbf{f}(\mathbf{x})), \tag{22}$$

where $\pi_{\text{noise}}(\boldsymbol{\omega})$ is the PDF of the noise (which is assumed to be identified based on separate calibration experiments, see Subsection 5.1). Substitution of Eq. (22) in Eq. (20) yields:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \pi(\mathbf{x})\pi_{\text{noise}}(\mathbf{y} - \mathbf{f}(\mathbf{x})). \tag{23}$$

A critical aspect of the Bayesian framework is the selection of the prior distribution ($\pi(\mathbf{x})$) [11] in which *a-priori* knowledge about the parameters is translated in terms of a PDF. The influence of the prior distribution diminishes if the number of observation increases [36], which is considered in more detail in Section 5.

Once the posterior is formulated (Eq. (23)), the mean parameter values, MAP parameter values and the covariance matrix can be extracted from it. We will use Markov chain Monte Carlo methods for this, since they are the most commonly employed approaches to do so. These techniques are discussed in Subsection 4.1. Only for linear elasticity, we analytically analyse the posterior.

Remark 3. *The parameters of the noise distribution can also be treated as unknown parameters that need to be identified. In that case, they appear as variables in the posterior, together with the material parameters. For tensile testers in well-controlled environments however, it is fair to say that the noise distribution and its parameters can be identified using a separate calibration process (see subsection 5.1).*

3.2. Application to the material responses during monotonic uniaxial tension

In this subsection, we apply the aforementioned Bayesian framework to the four material descriptions for monotonic uniaxial tension. Effectively, this means that we replace model $\mathbf{f}(\mathbf{x})$ in Eqs. (21) and (23) by the four material responses $\sigma(\epsilon, \mathbf{x})$. We will see however that the resulting posterior distribution for linear elasticity can be analysed analytically, but that the posteriors for the other three material descriptions need to be analysed numerically. This is because all elastoplastic descriptions are only C_0-continuous. On top of that, the description of the elastoplastic model with nonlinear hardening is implicit.

3.2.1. Linear elasticity

The only unknown material parameter in the linear elastic model is the Young's modulus (E). Based on Section 3, the additive noise model for *a single stress measurement* can be written as follows:

$$Y = E\epsilon + \Omega, \tag{24}$$

where Y denotes the measured stress and Ω denotes the random variable representing the noise in the stress measurement. We consider the noise distribution to be normal (Gaussian) and hence, it can be written as:

$$\pi_{\text{noise}}(\omega) = \frac{1}{\sqrt{2\pi}s_{\text{noise}}}\exp\Big(-\frac{\omega^2}{2s_{\text{noise}}^2}\Big). \tag{25}$$

Using Eq. (22), the likelihood function for a single stress measurement can now be expressed as:

$$\pi(y|E) = \pi_{\text{noise}}(y - E\epsilon) = \frac{1}{\sqrt{2\pi}s_{\text{noise}}}\exp\Big(-\frac{(y-E\epsilon)^2}{2s_{\text{noise}}^2}\Big). \tag{26}$$

Substitution of Eq. (26) in Eq. (23) then yields the following expression for the posterior:

$$\pi(E|y) \propto \pi(E)\exp\Big(-\frac{(y-E\epsilon)^2}{2s_{\text{noise}}^2}\Big). \tag{27}$$

If we use a prior in the form of a modified normal distribution as follows:

$$\pi(E) \propto \begin{cases} \exp\Big(-\frac{(E-\overline{E})^2}{2s_E^2}\Big) & \text{if } E \geq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{28}$$

the posterior distribution for a single stress measurement reads:

$$\pi(E|y) \propto \begin{cases} \exp\Big(-\Big[\frac{(E-\overline{E})^2}{2s_E^2} + \frac{(y-E\epsilon)^2}{2s_{\text{noise}}^2}\Big]\Big) & \text{if } E \geq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{29}$$

where $\overline{E}$ and s_E denote the mean and standard deviation of the prior distribution, respectively. Note that the Young's modulus cannot be negative which is taken into account in the prior distribution (Eq. (28)).

If we now consider the posterior distribution of the previous measurement to be the prior distribution of the current measurement, the posterior for all n_m measurements can be expressed as:

$$\pi(E|\mathbf{y}) \propto \exp\Big(- \Big[\frac{(E-\overline{E})^2}{2s_E^2} + \frac{\sum_{i=1}^{n_m}(y_i - E\epsilon_i)^2}{2s_{\text{noise}}^2} \Big] \Big), \quad E \geq 0, \tag{30}$$

where $\pi(E|\mathbf{y}) = \pi(E|y_1, \cdots, y_{n_m})$. Eq. (30) can now be written in the following form:

$$\pi(E|\mathbf{y}) \propto \exp\Big(- \frac{(E-\overline{E}_{\text{post}})^2}{2s_{\text{post}}^2} \Big), \quad E \geq 0, \tag{31}$$

where $\overline{E}_{\text{post}}$ and s_{post} denote the mean and standard deviation of the posterior distribution, which is again a normal distribution (with the condition $E \geq 0$). Both can be expressed as:

$$\overline{E}_{\text{post}} = \frac{s_{\text{noise}}^2 \overline{E} + s_E^2 \sum_{i=1}^{nm} \epsilon_i y_i}{s_{\text{noise}}^2 + s_E^2 \sum_{i=1}^{nm} \epsilon_i^2}, \quad s_{\text{post}} = \sqrt{\frac{s_{\text{noise}}^2 s_E^2}{s_{\text{noise}}^2 + s_E^2 \sum_{i=1}^{nm} \epsilon_i^2}}. \tag{32}$$

Hence, it is possible to analytically examine the posterior distribution for linear elasticity if the noise model is additive and the noise distribution as well as the prior distribution are (modified) normal distributions [11]. For the other cases below, we use numerical techniques.

3.2.2. Linear elasticity-perfect plasticity

The parameters to be identified for the linear elastic-perfectly plastic model are the Young's modulus and the initial yield stress, which are stored in the parameter vector $\mathbf{x} = \begin{bmatrix} E & \sigma_{y0} \end{bmatrix}^T$. Since we consider the same experimental equipment and conditions as in the case of linear elasticity (i.e. the measured stresses are still polluted by noise stemming from the same normal distribution and the measured strains are still exact), the same additive noise model applies:

$$Y = \sigma(\epsilon, \mathbf{x}) + \Omega, \tag{33}$$

where $\sigma(\epsilon, \mathbf{x})$ denotes the model response and is expressed in Eq. (8). Using Eq. (25) for the noise distribution, the likelihood function for a single stress measurement reads:

$$\pi(y|\mathbf{x}) = \pi_{\text{noise}}(y - \sigma(\epsilon, \mathbf{x})) = \frac{1}{\sqrt{2\pi} s_{\text{noise}}} \exp\Big(- \frac{(y - \sigma(\epsilon, \mathbf{x}))^2}{2s_{\text{noise}}^2} \Big), \tag{34}$$

or:

$$\pi(y|\mathbf{x}) = \frac{1}{\sqrt{2\pi} s_{\text{noise}}} \exp\left(- \frac{\left(y - E\epsilon \left(1 - h\left(\epsilon - \frac{\sigma_{y0}}{E} \right) \right) - \sigma_{y0} h\left(\epsilon - \frac{\sigma_{y0}}{E} \right) \right)^2}{2s_{\text{noise}}^2} \right). \tag{35}$$

Taking the physical constraints into account that the Young's modulus and the initial yield stress must be nonnegative, the following prior distribution is selected:

$$\pi(\mathbf{x}) \propto \begin{cases} \exp\Big(- \frac{(\mathbf{x}-\overline{\mathbf{x}})^T \mathbf{\Gamma}_{\mathbf{x}}^{-1} (\mathbf{x}-\overline{\mathbf{x}})}{2} \Big) & \text{if } E \geq 0 \text{ and } \sigma_{y0} \geq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{36}$$

where $\overline{\mathbf{x}}$ denotes the mean value vector of the prior distribution and $\mathbf{\Gamma}_{\mathbf{x}}$ the covariance matrix of the prior. Substitution of Eq. (33) and Eq. (34) in the reduced variant of Bayes' formula of Eq. (23), yields the following posterior distribution for n_m measurements:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \exp\left(- \left[\frac{(\mathbf{x} - \overline{\mathbf{x}})^T \mathbf{\Gamma}_{\mathbf{x}}^{-1} (\mathbf{x} - \overline{\mathbf{x}})}{2} + \frac{\sum_{i=1}^{n_m} \left(y_i - E\epsilon_i \left(1 - h\left(\epsilon_i - \frac{\sigma_{y0}}{E} \right) \right) - \sigma_{y0} h\left(\epsilon_i - \frac{\sigma_{y0}}{E} \right) \right)^2}{2s_{\text{noise}}^2} \right] \right). \tag{37}$$

We again note that the probability of obtaining a negative Young's modulus and yield stress is zero thanks to the selected prior distribution.

It is worth noting that the presence of the Heaviside function in the posterior makes the posterior difficult to analyse analytically. We therefore resort to the adaptive MCMC approach in Section 5.

3.2.3. Linear elasticity-linear hardening

The parameter vector for the linear elastic-linear hardening model reads $\mathbf{x} = \begin{bmatrix} E & \sigma_{y0} & H \end{bmatrix}^T$. Assuming again the same experimental equipment and conditions (and hence, the same noise model and noise distribution), the likelihood function for a single measurement reads:

$$\pi(y|\mathbf{x}) \propto \exp\Bigg(- \frac{\Big(y - E\epsilon\Big(1 - h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big)\Big) - \Big(\sigma_{y0} + \frac{HE}{H+E}\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big)\Big)h\Big(\epsilon - \frac{\sigma_{y0}}{E}\Big)\Big)^2}{2s^2_{\text{noise}}}\Bigg). \tag{38}$$

In addition to the physical constraints for the Young's modulus and the initial yield stress, we also use that plastic modulus H must be nonnegative. The following prior distribution is therefore selected:

$$\pi(\mathbf{x}) \propto \begin{cases} \exp\Big(- \frac{(\mathbf{x}-\overline{\mathbf{x}})^T \mathbf{\Gamma}_{\mathbf{x}}^{-1} (\mathbf{x}-\overline{\mathbf{x}})}{2}\Big) & \text{if } E \geq 0 \text{ and } \sigma_{y0} \geq 0 \text{ and } H \geq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{39}$$

Using Bayes' formula, the posterior distribution for n_m measurements reads:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \exp\Bigg(- \Bigg[\frac{(\mathbf{x} - \overline{\mathbf{x}})^T \mathbf{\Gamma}_{\mathbf{x}}^{-1} (\mathbf{x} - \overline{\mathbf{x}})}{2} + \sum_{i=1}^{n_m} \frac{\Big(y_i - E\epsilon_i\Big(1 - h\Big(\epsilon_i - \frac{\sigma_{y0}}{E}\Big)\Big) - \Big(\sigma_{y0} + \frac{HE}{H+E}\Big(\epsilon_i - \frac{\sigma_{y0}}{E}\Big)\Big)h\Big(\epsilon_i - \frac{\sigma_{y0}}{E}\Big)\Big)^2}{2s^2_{\text{noise}}} \Bigg]\Bigg). \tag{40}$$

3.2.4. Linear elasticity-nonlinear hardening

The parameter vector for the linear elastic-nonlinear hardening material description is denoted by $\mathbf{x} = \begin{bmatrix} E & \sigma_{y0} & H & n \end{bmatrix}^T$. Considering no change of experimental equipment (and hence, the same noise model and noise distribution), the expression for the measured stress again reads as Eq. (33), where $\sigma(\epsilon, \mathbf{x})$ is given by Eq. (14). It is important to note that in contrast to the previous cases, model response $\sigma(\epsilon, \mathbf{x})$ is not a closed form expression (see Eq. (14)). The likelihood function for a single measurement is:

$$\pi(y|\mathbf{x}) \propto \exp\Big(- \frac{(y - \sigma(\epsilon, \mathbf{x}))^2}{2s^2_{\text{noise}}}\Big), \tag{41}$$

where $\sigma(\epsilon, \mathbf{x})$ is numerically determined by solving Eq. (14). Choosing the prior distribution in the form of a modified normal distribution as:

$$\pi(\mathbf{x}) \propto \begin{cases} \exp\Big(- \frac{(\mathbf{x}-\overline{\mathbf{x}})^T \mathbf{\Gamma}_{\mathbf{x}}^{-1} (\mathbf{x}-\overline{\mathbf{x}})}{2}\Big) & \text{if } E \geq 0 \text{ and } \sigma_{y0} \geq 0 \text{ and } H \geq 0 \text{ and } n \geq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{42}$$

and employing the Bayes' theorem, the posterior distribution for n_m measurements reads:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \exp\Bigg(- \Bigg[\frac{(\mathbf{x} - \overline{\mathbf{x}})^T \mathbf{\Gamma}_{\mathbf{x}}^{-1} (\mathbf{x} - \overline{\mathbf{x}})}{2} + \frac{\sum_{i=1}^{n_m} (y_i - \sigma(\epsilon_i, \mathbf{x}))^2}{2s^2_{\text{noise}}} \Bigg]\Bigg). \tag{43}$$

4. Numerical procedures

4.1. Markov chain Monte Carlo method (MCMC)

Once the posterior is constructed, it needs to be analysed to determine the statistical summaries. For the aforementioned case of linear elasticity, the statistical summaries were established analytically, but for the other cases we need to determine them numerically because they are only C_0-continuous. We will use a Markov chain Monte Carlo (MCMC) technique for this.

MCMC techniques are frequently employed, derivative-free numerical approaches to investigate posteriors [37–39]. They draw samples from the posterior to do so. Below, the fundamental concepts of the Monte Carlo method are discussed, as well as the adaptive Metropolis algorithm to perform the sampling.

4.1.1. Monte Carlo method

The main purpose of the Monte Carlo method is to approximate integrals of the following form:

$$\mathbf{I} = \int_{\mathbb{R}^{n_p}} \mathbf{g}(\mathbf{x})\pi(\mathbf{x})d\mathbf{x}, \tag{44}$$

where π denotes the PDF of interest (in our case the posterior) and $\mathbf{g} : \mathbb{R}^{n_p} \to \mathbb{R}^{n_g}$ denotes an integrable function over $\mathbb{R}^{n_p}$. This integral can be approximated using the following quadrature:

$$\hat{\mathbf{I}} = \frac{1}{n_s}\sum_{i=1}^{n_s} \mathbf{g}(\mathbf{x}_i), \tag{45}$$

where $\{\mathbf{x}_i\}_i^{n_s}$ denotes a set of samples drawn from the PDF of interest (π) and the hat on $\hat{\mathbf{I}}$ represents the numerically approximated equivalent of $\mathbf{I}$. Drawing samples from π implies that most of the samples are in the domain in which numerical evaluations of π are nonzero. Note that $\hat{\mathbf{I}}$ converges according to [40]:

$$\lim_{n_s\to+\infty} \frac{1}{n_s}\sum_{i=1}^{n_s} \mathbf{g}(\mathbf{x}_i) = \mathbf{I}. \tag{46}$$

The numerical approximation of the components of the covariance matrix for $\mathbf{g}(\mathbf{x})$ ($\widehat{\mathbf{\Gamma}}_{\mathbf{g}}$) is [41]:

$$(\widehat{\Gamma}_{\mathbf{g}})_{jk} = \frac{1}{n_s-1}\sum_{i=1}^{n_s} \Big(g_j(\mathbf{x}_i) - I_j\Big)\Big((g_k(\mathbf{x}_i) - I_k\Big),\ j = 1, 2, \cdots, n_g,\ k = 1, 2, \cdots, n_g. \tag{47}$$

The mean of the posterior ($\overline{\mathbf{x}}_{\text{post}}$) can be computed by substituting $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ and $\pi = \pi_{\text{post}}$ in Eq. (44), which yields:

$$\overline{\mathbf{x}}_{\text{post}} = \int_{\mathbb{R}^{n_p}} \mathbf{x}\,\pi_{\text{post}}(\mathbf{x})d\mathbf{x} = \lim_{n_s\to+\infty} \frac{1}{n_s}\sum_{i=1}^{n_s} \mathbf{x}_i. \tag{48}$$

Furthermore, the components of the posterior's covariance matrix are approximated as follows:

$$(\widehat{\Gamma}_{\text{post}})_{jk} = \frac{1}{n_s-1}\sum_{i=1}^{n_s} \Big((x_i)_j - (\overline{x}_{\text{post}})_j\Big)\Big((x_i)_k - (\overline{x}_{\text{post}})_k\Big),\ j = 1, 2, \cdots, n_p,\ k = 1, 2, \cdots, n_p. \tag{49}$$

If we assume that a sufficiently large number of samples is taken (i.e. n_s is large), the MAP point can furthermore be approximated as [40]:

$$\widehat{\mathbf{MAP}} = \underset{\mathbf{x}_i; i=1,\ldots,n_s}{\text{argmax}}\ \pi(\mathbf{x}_i). \tag{50}$$

The essential part of a Monte Carlo procedure is the drawing of admissible samples ($\mathbf{x}_i$). Below, the standard and adaptive Metropolis algorithms are discussed as means to draw samples. The adaptive one is the algorithm used in Section 5.

4.1.2. The standard Metropolis-Hastings and the adaptive Metropolis algorithms

The standard Metropolis-Hastings approach is a frequently employed MCMC algorithm [40]. The basic idea of the Metropolis-Hastings algorithm is to explore the PDF of interest by making a random walk through parameter space $\mathbf{x}$. Considering sample $\mathbf{x}_i$ and its evaluation of the PDF, $\pi(\mathbf{x}_i)$, new sample $\mathbf{x}_p$ is proposed by drawing from a proposal distribution (q in Algorithm 1). If the PDF evaluated at the proposed sample ($\pi(\mathbf{x}_p)$) multiplied by the evaluation of the proposal distribution evaluated at $\mathbf{x}_i$, given the proposed sample ($q(\mathbf{x}_i|\mathbf{x}_p)$), is larger than the PDF at the current sample ($\pi(\mathbf{x}_i)$) multiplied by the evaluation of the proposal distribution at the proposed sample given the current sample ($q(\mathbf{x}_p|\mathbf{x}_i)$), the proposed sample is *always* accepted as the new sample. If $\pi(\mathbf{x}_p)q(\mathbf{x}_i|\mathbf{x}_p) < \pi(\mathbf{x}_i)q(\mathbf{x}_p|\mathbf{x}_i)$ however, the proposed sample *may* be accepted. The fact whether or not it is accepted depends on the ratio of scalar r in Algorithm 1. The ratio is compared to a random number generated from a uniform distribution. If the ratio is greater than the random number, the proposed sample is accepted. If the ratio is smaller than the random number, the proposed sample is rejected, and the current sample becomes the new sample. Otherwise, the proposed sample becomes the new sample. The algorithm is repeated for n_s samples.

Algorithm 1 The standard Metropolis-Hastings algorithm

1: select the initial sample $\mathbf{x}_0 \in \mathbb{R}^{n_p}$
2: **for** $i = 0, 1, 2, ..., n_s - 1$ **do**
3: draw $\mathbf{x}_p \in \mathbb{R}^{n_p}$ from the proposal distribution $q(\mathbf{x}_p|\mathbf{x}_i)$ in Eq. (53)
4: calculate the ratio $r(\mathbf{x}_i, \mathbf{x}_p) = \min\Big(1, \frac{\pi(\mathbf{x}_p)q(\mathbf{x}_i|\mathbf{x}_p)}{\pi(\mathbf{x}_i)q(\mathbf{x}_p|\mathbf{x}_i)}\Big)$
▷ $\pi(\cdot)$ denotes the target distribution (i.e. posterior).
5: draw $u \in [0, 1]$ from uniform probability density
6: **if** $r(\mathbf{x}_i, \mathbf{x}_p) \geq u$ **then**
7: $\mathbf{x}_{i+1} = \mathbf{x}_p$
8: **else**
9: $\mathbf{x}_{i+1} = \mathbf{x}_i$
10: **end if**
11: **end for**

In case of a symmetric proposal distribution (as in this contribution), the following relation holds:

$$q(\mathbf{x}_i|\mathbf{x}_p) = q(\mathbf{x}_p|\mathbf{x}_i). \tag{51}$$

Consequently, step 4 in Algorithm 1 simplifies to:

$$r(\mathbf{x}_i, \mathbf{x}_p) = \min\Big(1, \frac{\pi(\mathbf{x}_p)}{\pi(\mathbf{x}_i)}\Big). \tag{52}$$

Note that the Metropolis-Hastings algorithm with a symmetric proposal distribution is commonly called the Metropolis algorithm.

A commonly employed approach to check the stability and convergence of the algorithm is to trace the generated samples and analyse their characteristics, after the algorithm is finished. The evolution of the mean value and the standard deviation can for instance be checked for convergence [32]. We refer the readers for more information to the review on assessing the convergence of the MCMC by Sinharay [42].

The efficiency of the algorithm is influenced by the initial sample ($\mathbf{x}_0$) and the proposal distribution (q) [11]. The most common proposal distribution for the Metropolis-Hastings algorithm (as employed here) is of the following Gaussian form:

$$q(\mathbf{x}_i|\mathbf{x}_p) = q(\mathbf{x}_p|\mathbf{x}_i) \propto \exp\Big(-\frac{1}{2\gamma^2}\left\|\mathbf{x}_i - \mathbf{x}_p\right\|^2\Big), \tag{53}$$

where γ denotes the parameter that determines the width of the proposal distribution and must be tuned to obtain an efficient and converging algorithm. An efficient starting value is $\gamma = \frac{2.38}{\sqrt{n_p}}$ [43], where n_p denotes

the number of unknown parameters and hence, the dimension of the posterior for the cases presented above.

To overcome the tuning of γ, Haario et al. [44] introduced the adaptive proposal (AP). The AP method updates the width of the proposal distribution, using the existing knowledge of the posterior. The existing knowledge is based on the previous samples. For sample $n_{\mathbf{K}}+1$, the update employs the following formulation:

$$q(\mathbf{x}_p|\mathbf{x}_i) \sim N(\mathbf{x}_i, \gamma^2 \mathbf{R}_{n_{\mathbf{K}}}), \tag{54}$$

where $N(\mathbf{x}_i, \gamma^2 \mathbf{R}_{n_{\mathbf{K}}})$ denotes a normal distribution with mean $\mathbf{x}_i$ and covariance matrix $\gamma^2 \mathbf{R}_{n_{\mathbf{K}}}$, of size $n_p \times n_p$. To establish $\mathbf{R}_{n_{\mathbf{K}}}$, all $n_{\mathbf{K}}$ previous samples are first stored in matrix $\mathbf{K}$ of size $n_{\mathbf{K}} \times n_p$. $\mathbf{R}_{n_{\mathbf{K}}}$, is then computed as:

$$\mathbf{R}_{n_{\mathbf{K}}} = \frac{1}{n_{\mathbf{K}} - 1} \widetilde{\mathbf{K}}^T \widetilde{\mathbf{K}}, \tag{55}$$

where $\widetilde{\mathbf{K}} = \mathbf{K} - \mathbf{K}_{\text{mean}}$ and $\mathbf{K}_{\text{mean}}$ reads:

$$\mathbf{K}_{\text{mean}} = \begin{bmatrix} \mathbf{k}_{\text{mean}} \\ \mathbf{k}_{\text{mean}} \\ \vdots \\ \mathbf{k}_{\text{mean}} \end{bmatrix}_{n_{\mathbf{K}} \times n_p}, \tag{56}$$

and $\mathbf{k}_{\text{mean}}$ denotes a row matrix of length n_p which is determined as follows:

$$\mathbf{k}_{\text{mean}} = \frac{1}{i} \left[\sum_{i=1}^{n_{\mathbf{K}}} (K)_{i1} \quad \sum_{i=1}^{n_{\mathbf{K}}} (K)_{i2} \quad \cdots \quad \sum_{i=1}^{n_{\mathbf{K}}} (K)_{in_p} \right]. \tag{57}$$

The following relation is used for $N(\mathbf{x}_i, \gamma^2 \mathbf{R}_{n_{\mathbf{K}}})$ in this contribution:

$$N(\mathbf{x}_i, \gamma^2 \mathbf{R}_{n_{\mathbf{K}}}) \sim \mathbf{x}_i + \frac{\gamma}{\sqrt{n_{\mathbf{K}} - 1}} \widetilde{\mathbf{K}}^T N(0, \mathbf{I}_{n_{\mathbf{K}}}), \tag{58}$$

where $\mathbf{I}_{n_{\mathbf{K}}}$ denotes the identity matrix of size $n_{\mathbf{K}} \times n_{\mathbf{K}}$ and $N(0, \mathbf{I}_{n_{\mathbf{K}}})$ denotes the $n_{\mathbf{K}}$-dimensional normal distribution.

Note that it is computationally inefficient to update the proposal distribution after each new sample is generated. In the numerical examples in this study therefore, updating takes place once per 1000 sample generations. Algorithm 2 shows the Metropolis-Hastings algorithm with the symmetric AP proposal (Eq. (58)) that is employed here.

Algorithm 2 The Metropolis algorithm with AP proposal

1: select the initial sample $\mathbf{x}_0 \in \mathbb{R}^{n_p}$ and set $\gamma = \frac{2.38}{\sqrt{n_p}}$
2: **for** $i = 0, 1, 2, ..., n_s - 1$ **do**
3: draw $\mathbf{x}_p \in \mathbb{R}^{n_p}$ from the proposal distribution $q(\mathbf{x}_p|\mathbf{x}_i)$ in Eq. (58)
4: calculate the ratio $r(\mathbf{x}_i, \mathbf{x}_p) = \min\left(1, \frac{\pi(\mathbf{x}_p)}{\pi(\mathbf{x}_i)}\right)$
▷ $\pi(\cdot)$ denotes the target distribution (i.e. posterior).
5: draw $u \in [0, 1]$ from uniform probability density
6: **if** $r(\mathbf{x}_i, \mathbf{x}_p) \geq u$ **then**
7: $\mathbf{x}_{i+1} = \mathbf{x}_p$
8: **else**
9: $\mathbf{x}_{i+1} = \mathbf{x}_i$
10: **end if**
11: **per 1000 samples**
12: update matrix $\widetilde{\mathbf{K}}$
13: **end for**

4.2. Posterior predictive distribution (PPD)

Once the posterior is established, the posterior predictive distribution can be used to predict new measurements, given the current measurements. A comparison between the newly generated measurements and the current (yet observed) measurements may then indicate if a different model and/or prior need to be used. For instance, if the envelope of the new measurements differs substantially from the envelope of the current measurements, the current measurements are unlikely to be generated using the current model and/or prior and the user may want to select a new model and/or prior.

Considering measurements $\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_{n_m} \end{bmatrix}^T$, the posterior predictive distribution of new measurement y^{new} for new strain ϵ^{new}, given the current measurements, reads [45]:

$$\pi(y^{\text{new}}|\mathbf{y}, \boldsymbol{\epsilon}, \epsilon^{\text{new}}) = \int_{\mathbb{R}^{n_p}} \pi(y^{\text{new}}|\mathbf{x}, \epsilon^{\text{new}}) \pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\epsilon}) d\mathbf{x}, \tag{59}$$

where $\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 & \cdots & \epsilon_{n_m} \end{bmatrix}^T$ denotes the vector of the strains at which stresses $\mathbf{y}$ were measured. Note that we have so far neglected $\boldsymbol{\epsilon}$ in the notation.

Computing the integral in Eq. (59) is usually challenging for high dimensional problems. However, the Monte Carlo Markov chain can be employed to draw samples from the PPD for a new measurement (y^{new}), given measurement vector $\mathbf{y}$. This can be achieved by employing a sampling procedure twice. First, samples are drawn from the posterior distribution for the parameters, given the measurements ($\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\epsilon})$). Note that this is already performed during the numerical analysis of the posterior and hence, if those samples are saved, this procedure does not have to be applied again. Second, the i^{th} sample is replaced in $\pi(y^{\text{new}}|\mathbf{x}_i, \epsilon^{\text{new}})$, which is subsequently used to generate a sample for new measurement y_i^{new}.

5. Examples

All formulations derived in the previous section are investigated below. The effect of the prior distribution on the posterior distribution is studied, as well as the ability of the current formulations to recover a material parameter distribution when they are taken from a specific distribution. Also, BI's ability to recover correlations between different material parameters is exposed. First however, we will identify the noise distribution and its parameters.

5.1. Noise distribution

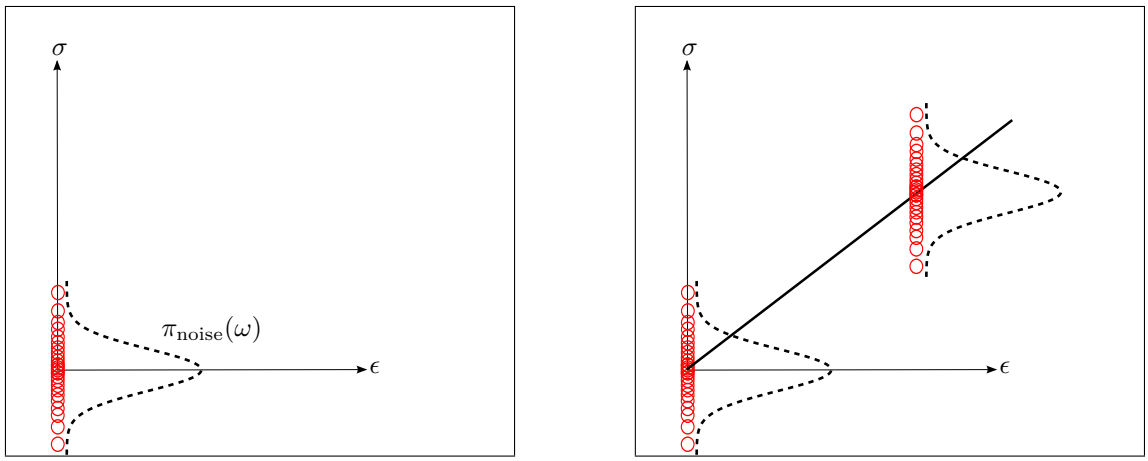
To determine the noise distribution and its parameters, two sets of 'calibration experiments' can be performed. First, a test is performed without any specimen. The stress-strain measurements of this test are shown in Fig. 2(a). It shows that the PDF of the noise in the 'stress measurements' is a normal distribution with a zero mean and a standard deviation of s_{noise}.

Second, the evolution of the noise distribution (including its parameters) is determined. To this purpose, a tensile test is performed on a calibration specimen (of which the Young's modulus is known). The artificially generated results are presented in Fig. 2(b). The mean stress value varies linearly with the strain. Standard deviation s_{noise} remains the same however.

Thus, the 'noise calibration measurements' indicate that an additive noise model can be used and the stresses are polluted by realisations coming from a normal noise distribution with standard deviation s_{noise}. Now, we will employ BI to identify the Young's modulus of the linear elastic model.

5.2. Linear elasticity

Identification of the Young's modulus. In the first example, a specimen with a Young's modulus of 210 GPa is considered, which is to be identified. 'Noise calibration experiments' were performed and the noise in the stress follows the normal distribution of Eq. (25) with $s_{\text{noise}} = 0.01$ GPa. For only one stress measurement of $y = 0.1576$ GPa with corresponding strain $\epsilon = 7.25 \times 10^{-4}$, the posterior distribution is calculated using



(a) Fitted noise distribution (b) The shifted noise distribution

Figure 2: Schematic of the stress-strain measurements (red circles) of the 'noise calibration experiments', including the noise distributions (dashed). The theoretical stress-strain relation (which is exact for the calibration experiments) is presented as a solid line in the diagram on the right.

Eq. (31). Selecting the prior distribution as in Eq. (28) with mean $\overline{E} = 150$ GPa and a relatively large standard deviation of $s_E = 50$ GPa, the posterior reads:

$$\pi(E|y) \propto \exp\Big(- \tfrac{(E-\overline{E}_{\text{post}})^2}{2s^2_{\text{post}}} \Big), \quad E \geq 0, \tag{60}$$

where $\overline{E}_{\text{post}} = 212.6486$ GPa and $s_{\text{post}} = 13.2964$ GPa.

Fig. 3 shows this posterior distribution, as well as the prior distribution and the value predicted by the least squares method, for one and five measurements. Fig. 4 presents the linear elastic responses for one and for ten measurements. The figure also shows the stress-strain responses using Young's moduli drawn from the 95% credible region (i.e. the region that contains 95% of the posterior) of the posterior as well as the posterior predictions. One can see that envelope associated with the 95% credible region is narrower than the 95% prediction interval. Note that the 95% prediction interval is obtained using the posterior predictive distribution and its upper and lower bounds read:

$$\text{prediction bounds} = \overline{\text{PPD}} \pm 2s_{\text{PPD}}, \tag{61}$$

where $\overline{\text{PPD}}$ denotes the mean of posterior predictive distribution for the new measurement (i.e. y^{new} in (59)) and s_{PPD} denotes its standard deviation.

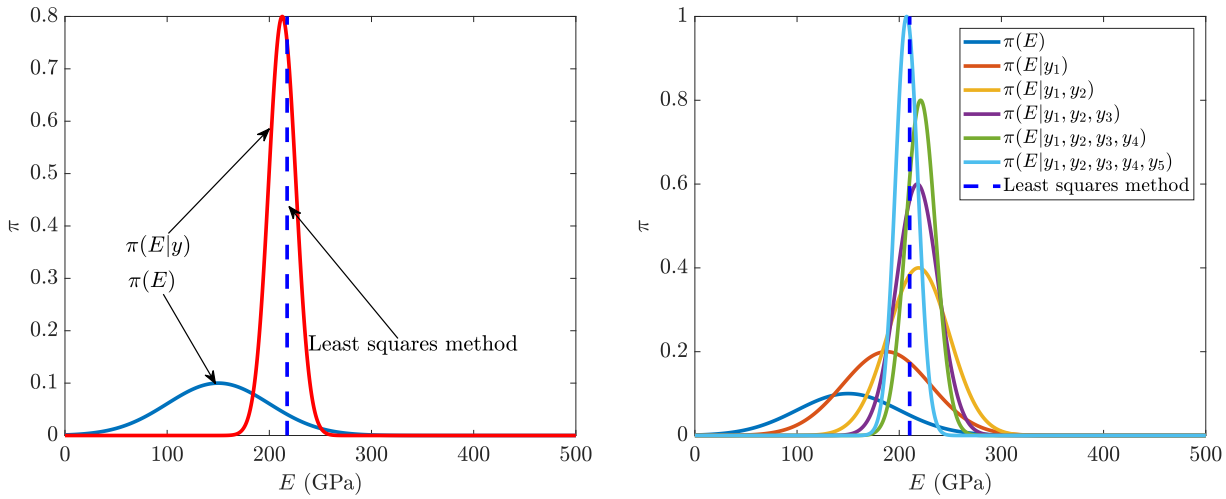
Two points can be observed in Fig. 3. First, the strain at which a measurement is made has a strong influence on the posterior. This can be observed by comparing the posterior of Fig. 3(a) with that in Fig. 3(b) for only one measurement (the distribution in red, denoted by $\pi(E|y_1)$). The latter distribution is significantly wider and its MAP point is relatively distant from the specimen's Young's modulus. Hence, a measurement made at a comparatively large strain reduces the width of the posterior distribution (i.e. reduces the uncertainty).

The second observation is that for an increasing number of measurements, the posterior becomes narrower and the MAP point moves closer to the specimen's Young's modulus.

By comparing the MAP point for a single measurement in Fig. 3(b) ($\overline{E}_{\text{post}} = 207.2821$ GPa) with the result of the least squares method for the same measurement ($E_{\text{ls}} = 210.2216$ GPa), one can notice the effect of the selected prior distribution. One interpretation of this is that the least squares method gives a more accurate result than BI (although this depends the selected prior), as the result of the least squares method is closer to the specimen's Young's modulus than the MAP point determined using BI. On the other hand, the result determined using the least squares method is not the actual Young's modulus of the specimen

(210 GPa), whereas the posterior distribution of BI does include this value. Furthermore, the MAP point and mean value of BI, come with an uncertainty in terms of the parameter value itself. This can be considered as an advantage if one wants to include this uncertainty, instead of including one deterministic value.

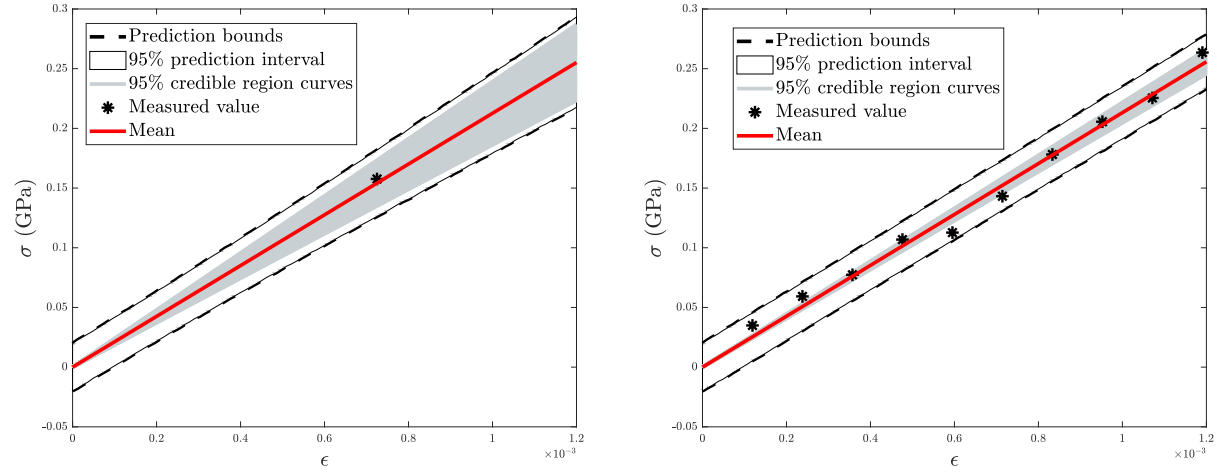
The main point is that BI cannot be directly compared to the least squares method, because in contrast to the latter, BI results in a posterior probability distribution that represents the probability of each possible value to occur.



(a) One observation

(b) Five successive observations

Figure 3: Linear elasticity: The prior, the posterior and the value predicted by least squares method for one measurement (a) and five measurements (b). The distributions are not normalised. The strain at which a measurement is made has a considerable influence on the posterior. This can be observed by comparing the posterior of (a) ($\pi(E|y)$, red line) with the posterior of (b), if only the first measurement is considered ($\pi(E|y_1)$, red line). An increase of the number of measurements results in narrower posteriors, with their MAP estimates closer to the true Young's modulus.



(a) One observation

(b) Ten observations

Figure 4: Linear elasticity: The measurements, the posterior prediction and the stress-strain curves created using the posterior and mean for (a) one measurement and (b) ten measurements. The envelope associated with the 95% credible region is narrower than the 95% prediction interval.

The influence of the prior. Now, we will study the effect of the prior distribution on the MAP point (which is the same as the mean value for the normal posteriors in this subsection). In Fig. 5 the MAP points are shown as a function of the mean and the standard deviation of the prior. The MAP points are presented for different numbers of measurements. As can be seen, an increase of the number of measurements results in a flatter surface, which means that the influence of the prior distribution decreases.

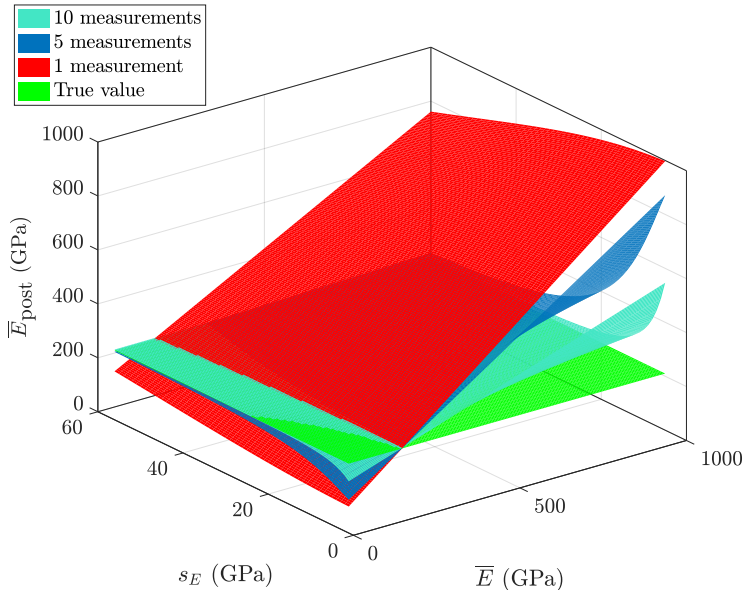


Figure 5: Linear elasticity: The influence of the prior (i.e. the mean value and the standard deviation) on the resulting MAP point for different numbers of measurements. Increasing the number of the measurements results in a flatter surface which indicates a decreasing influence of the prior distribution.

Recovering material heterogeneity. A last important point to show using the linear elastic model is BI's ability (or inability for the current formulation) to capture the intrinsic heterogeneity of the material parameters. The question here is thus if BI is able to recover the distribution of the Young's modulus if several specimens are tested and their Young's moduli are taken from a specific underlying distribution. To this end, 25 specimens are considered of which the Young's moduli are taken from a normal distribution with a mean value of 210 GPa and a standard deviation of 10 GPa (blue curve in Fig. 6). For each specimen ten measurements are made. The aforementioned noise model and noise distribution are applied.

The resulting posterior is presented by the red curve in Fig. 6, which is a (modified) normal distribution with $\overline{E}_{\text{post}} = 215.3971$ GPa and $s_{\text{post}} = 0.8561$ GPa. The posterior is substantially narrower than the distribution of the specimens' Young's moduli and hence, using the BI formulations of this contribution, the intrinsic heterogeneity of the material itself cannot be captured. This entails that the width of the posterior distributions (represented by s_{post} in this subsection) is only a measure of the uncertainty of the MAP points and the mean value and not of the material heterogeneity.

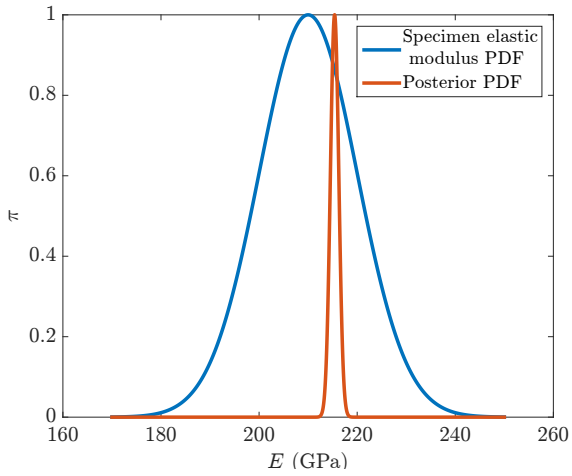


Figure 6: Linear elasticity: The distribution of the specimens' Young's moduli and the resulting posterior. The PDFs are not normalised. The current formulation is clearly not able to recover the material heterogeneity. To be able to recover the material heterogeneity, one needs to consider both the intrinsic uncertainty of the material parameters as well as that of the measurements.

5.3. Linear elasticity-perfect plasticity

Identification of the material parameters. In the first example of this subsection, a linear elastic-perfectly plastically behaving specimen is considered with Young's modulus $E = 210$ GPa and yield stress $\sigma_{y0} = 0.25$ GPa. Twelve measurements are generated by employing the same noise distribution as in the previous subsection. The prior distribution of Eq. (36) is furthermore selected with the following mean vector and covariance matrix:

$$\overline{\mathbf{x}} = \begin{bmatrix} 200 \\ 0.29 \end{bmatrix} \text{ GPa}, \ \boldsymbol{\Gamma}_{\mathbf{x}} = \begin{bmatrix} 2500 & 0 \\ 0 & 2.7778 \times 10^{-4} \end{bmatrix} \text{GPa}^2. \tag{62}$$

Consequently, the posterior of Subsection 3.2.2 is of the form of Eq. (37), which is investigated by the MCMC approach given in Subsection 4.1.2. Running the chain for 10^4 samples whilst burning the first 3000 samples (i.e. the first 3000 samples are not used to determine the mean, the covariance matrix and the MAP estimate) yields:

$$\widehat{\overline{\mathbf{x}}}_{\text{post}} = \begin{bmatrix} 208.9859 \\ 0.2578 \end{bmatrix} \text{ GPa}, \ \widehat{\boldsymbol{\Gamma}}_{\text{post}} = \begin{bmatrix} 29.807 & 4.1064 \times 10^{-4} \\ 4.1064 \times 10^{-4} & 1.5067 \times 10^{-5} \end{bmatrix} \text{ GPa}^2, \tag{63}$$

and

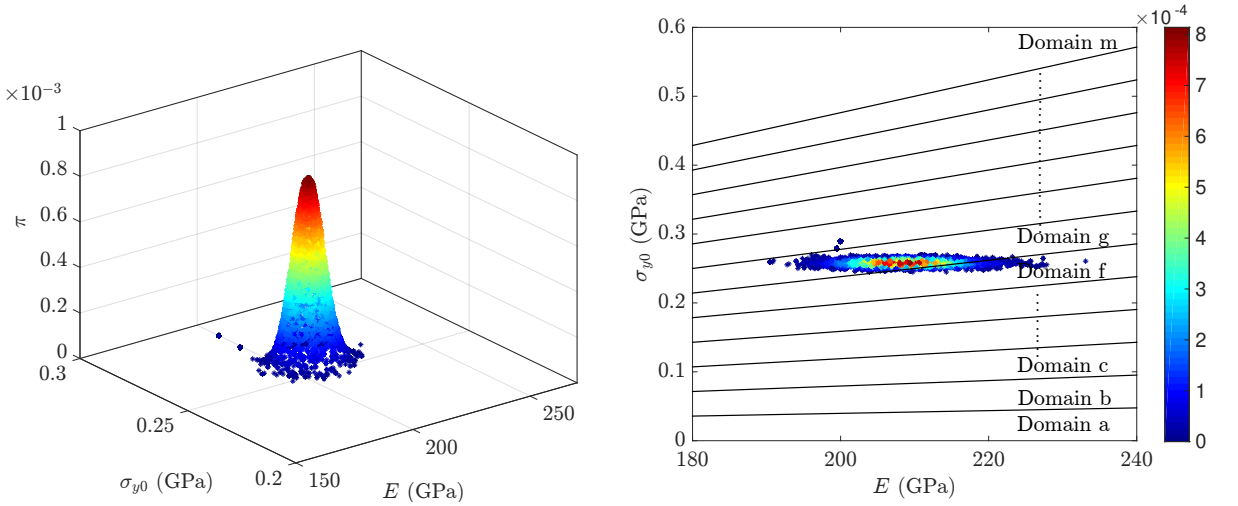
$$\widehat{\mathbf{MAP}} = \begin{bmatrix} 208.4475 \\ 0.2578 \end{bmatrix} \text{ GPa}, \tag{64}$$

where the hat sign $\widehat{(\cdot)}$ denotes the numerical approximation.

Fig. 7(a) shows the samples generated by the adaptive MCMC approach which are used to approximate the posterior distribution. The domains presented in Fig. 7(b) show which of the measurements are included in the purely elastic part and which fall within the elastoplastic part. These discrete domains are a result of the C_0-continuity of Eq. (8). In domain 'a' (in which no samples are generated by the adaptive MCMC approach), all the measurements are considered to be in the elastoplastic part. In domain 'b' on the other hand, the first measurement (the one with the smallest strain) is considered to be in the purely elastic part and the others remain in the elastoplastic part. Continuing like this, in domain 'c' the second measurement is also considered to be in the purely elastic part. Finally, in domain 'm' all measurements are considered

to fall within the elastoplastic domain. Based on Fig. 7(b) the MAP point is clearly located in the domain in which the first six measurements are considered to be in the purely elastic part and the remaining in the elastoplastic part.

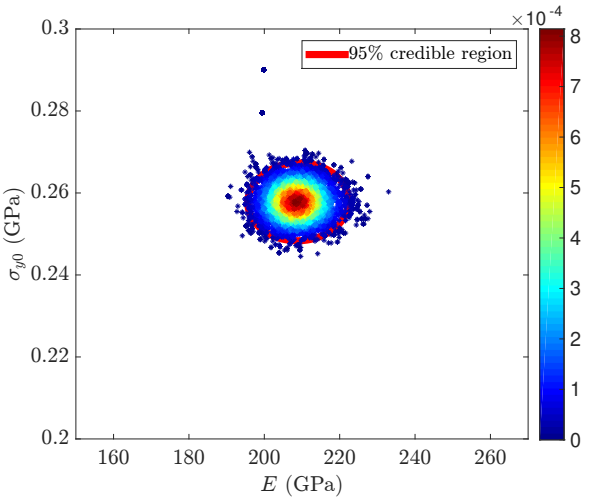
The 95% credible region is shown together with the posterior distribution in Fig. 8(a). The possible stress-strain responses inside the credible region as well as the posterior prediction are presented in Fig. 8(b). The posterior distribution seems to be roughly of an elliptical shape with its primary axes almost along the E-axis and σ_{y0}-axis. This entails that the correlation between the two material parameters is not significant. One has to notice though, that the assumed prior is uncorrelated. In other words, the prior covariance matrix ($\mathbf{\Gamma_x}$) is diagonal. It is therefore interesting to investigate the influence of the off-diagonal term of the prior covariance matrix on the posterior covariance matrix. In Fig. 9, this influence is graphically presented for the three terms of the posterior covariance matrix (note that both the prior covariance matrix and the posterior covariance matrix are symmetric). It seems that an increase of $(\Gamma_{\mathbf{x}})_{12}$ leads to some decreasing trend for $(\widehat{\Gamma}_{\text{post}})_{11}$ and some increasing trend for $(\widehat{\Gamma}_{\text{post}})_{12}$. However, it is difficult to assess whether or not these trends can be considered as meaningful.



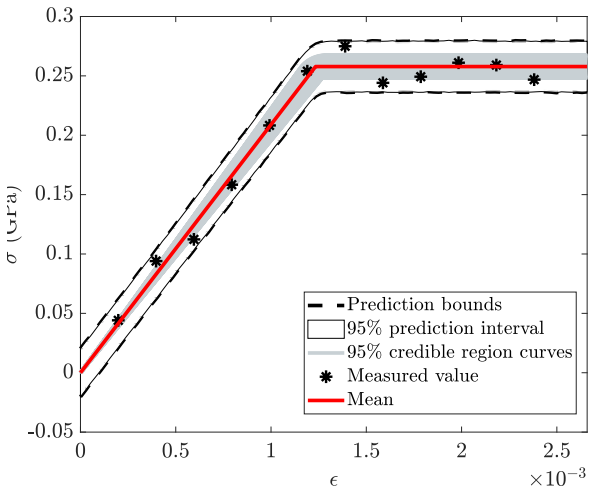
(a) Samples generated by the adaptive MCMC approach

(b) Samples generated by the adaptive MCMC approach (top view) including different domains

Figure 7: Linear elasticity-perfect plasticity: Two different views of the samples generated by the adaptive MCMC approach to approximate the posterior. The colours represent the value of the posterior, which in the left image is also shown along the z-axis. In Fig. 7(b) several domains are shown. Each of these domains corresponds to a region for which the number of measurements considered to be in the purely elastic part is constant (e.g zero in domain a and one in domain b).

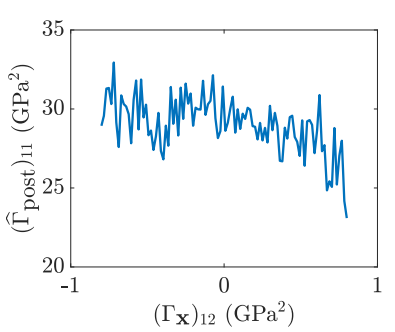


(a) The 95% credible region

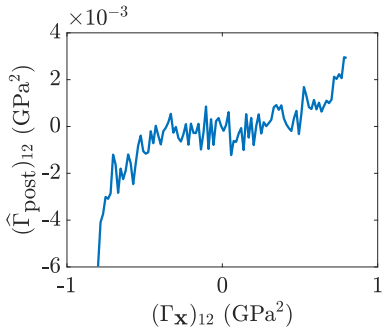


(b) The measurements, the posterior prediction and the stress-strain curves

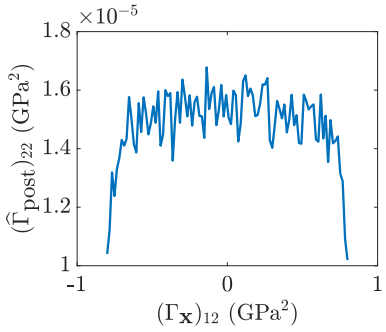
Figure 8: Linear elasticity-perfect plasticity: The 95% credible region and the posterior distribution (a), the measurements, the posterior prediction and the stress-strain curves created using the 95% credible region of the posterior (b).



(a) Effect of $(\Gamma_{\mathbf{x}})_{12}$ on $(\widehat{\Gamma}_{\text{post}})_{11}$

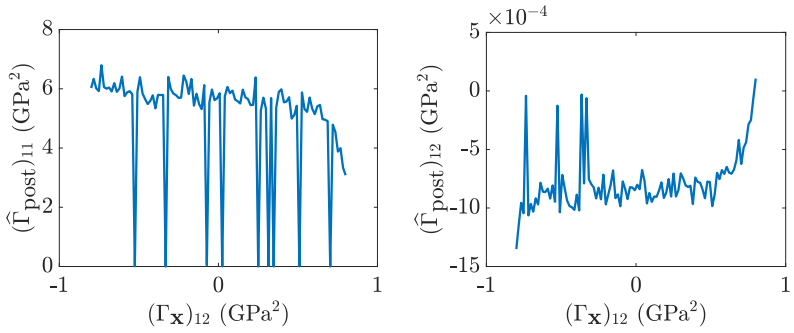


(b) Effect of $(\Gamma_{\mathbf{x}})_{12}$ on $(\widehat{\Gamma}_{\text{post}})_{12}$

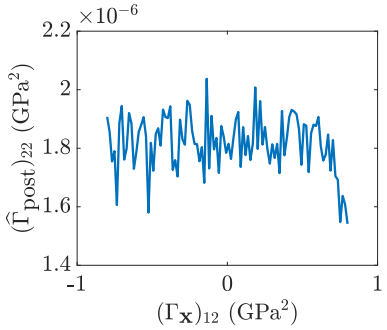


(c) Effect of $(\Gamma_{\mathbf{x}})_{12}$ on $(\widehat{\Gamma}_{\text{post}})_{22}$

Figure 9: Linear elasticity-perfect plasticity: Effect of the off-diagonal component of the prior's covariance matrix on the posterior's covariance matrix. It seems that an increase of $(\Gamma_{\mathbf{x}})_{12}$ leads to a decreasing trend of $(\widehat{\Gamma}_{\text{post}})_{11}$ and a increasing trend of $(\widehat{\Gamma}_{\text{post}})_{12}$. However, it is difficult to assess whether or not a true trend is present in these results.



(a) Effect of $(\Gamma_{\mathbf{x}})_{12}$ on $(\widehat{\Gamma}_{\text{post}})_{11}$ (b) Effect of $(\Gamma_{\mathbf{x}})_{12}$ on $(\widehat{\Gamma}_{\text{post}})_{12}$



(c) Effect of $(\Gamma_{\mathbf{x}})_{12}$ on $(\widehat{\Gamma}_{\text{post}})_{22}$

Figure 10: Linear elasticity-perfect plasticity: Effect of the off-diagonal components of the prior's covariance matrix on the posterior's covariance matrix if the measurements are generated from ten specimens with their material parameters drawn from a normal distribution given in Eq. (65). No real trends can be observed.

The influence of the prior on the correlation between the material parameters. The next example focuses on the ability of the current formulation to capture a correlation between the Young's modulus and the initial yield stress if they are correlated. To this end, ten specimens are considered of which the material parameters are governed by a normal distribution with the following mean vector and covariance matrix:

$$\overline{\mathbf{x}}_{\text{spc}} = \begin{bmatrix} 210 \\ 0.25 \end{bmatrix} \text{ GPa}, \ \boldsymbol{\Gamma}_{\text{spc}} = \begin{bmatrix} 100 & 10^{-4} \\ 10^{-4} & 1.1111 \times 10^{-4} \end{bmatrix} \text{GPa}^2. \tag{65}$$

For each specimen, twelve measurements are made. Using the same prior as in the previous example (see Eq. (62)) and the adaptive MCMC approach for 10^4 samples whilst burning the first 3000 samples, yields:

$$\widehat{\overline{\mathbf{x}}}_{\text{post}} = \begin{bmatrix} 211.1077 \\ 0.2519 \end{bmatrix} \text{ GPa}, \ \widehat{\boldsymbol{\Gamma}}_{\text{post}} = \begin{bmatrix} 5.5373 & -8.396 \times 10^{-4} \\ -8.396 \times 10^{-4} & 1.8174 \times 10^{-6} \end{bmatrix} \text{GPa}^2. \tag{66}$$

The MAP point is given by:

$$\widehat{\mathbf{MAP}} = \begin{bmatrix} 210.5923 \\ 0.2521 \end{bmatrix} \text{ GPa}. \tag{67}$$

These results show that the correlation of the posterior is not the same as that of the distribution of the actual material. This corresponds closely with the observation that the formulations in this contribution are unable to capture any of the intrinsic uncertainty of the material parameters. Fig. 10 shows the effect of the off-diagonal component of the prior's covariance matrix ($\boldsymbol{\Gamma}_{\mathbf{x}}$) on the components of the posterior's covariance matrix ($\boldsymbol{\Gamma}_{\text{post}}$). Again, no real trends can be observed.

5.4. Linear elasticity-linear hardening

Identification of the material parameters. This subsection deals with the Bayesian formulation for the linear elastic-linear hardening material description. A specimen with Young's modulus $E = 210$ GPa, initial yield stress $\sigma_{y0} = 0.25$ GPa and plastic modulus $H = 50$ GPa is regarded. Twelve measurements are created by employing the same noise distribution as in the previous subsection. The prior distribution is given by Eq. (39) with the following properties:

$$\overline{\mathbf{x}} = \begin{bmatrix} 200 \\ 0.29 \\ 60 \end{bmatrix} \text{ GPa}, \ \mathbf{\Gamma_x} = \begin{bmatrix} 2500 & 0 & 0 \\ 0 & 2.7778 \times 10^{-4} & 0 \\ 0 & 0 & 100 \end{bmatrix} \text{GPa}^2. \tag{68}$$

The adaptive MCMC algorithm for 10^4 samples whilst burning the first 3000 samples yields:

$$\widehat{\overline{\mathbf{x}}}_{\text{post}} = \begin{bmatrix} 207.4586 \\ 0.2533 \\ 55.9187 \end{bmatrix} \text{GPa}, \widehat{\mathbf{\Gamma}}_{\text{post}} = \begin{bmatrix} 36.5642 & -1.2746 \times 10^{-2} & -3.7886 \\ -1.2746 \times 10^{-2} & 4.0359 \times 10^{-5} & -2.6218 \times 10^{-2} \\ -3.7886 & -2.6218 \times 10^{-2} & 66.8214 \end{bmatrix} \text{GPa}^2, \tag{69}$$

and

$$\widehat{\mathbf{MAP}} = \begin{bmatrix} 206.9528 \\ 0.2548 \\ 55.2838 \end{bmatrix} \text{GPa}. \tag{70}$$

Fig. 11 shows the generated samples by the adaptive MCMC approach in the $E - \sigma_{y0} - H$ space, including the projections on the $E - \sigma_{y0}$, $E - H$ and $\sigma_{y0} - H$ planes.

The 95% credible region is presented in Fig. 12(a) and the stress-strain responses associated with it, as well as the posterior prediction are shown in Fig. 12(b).

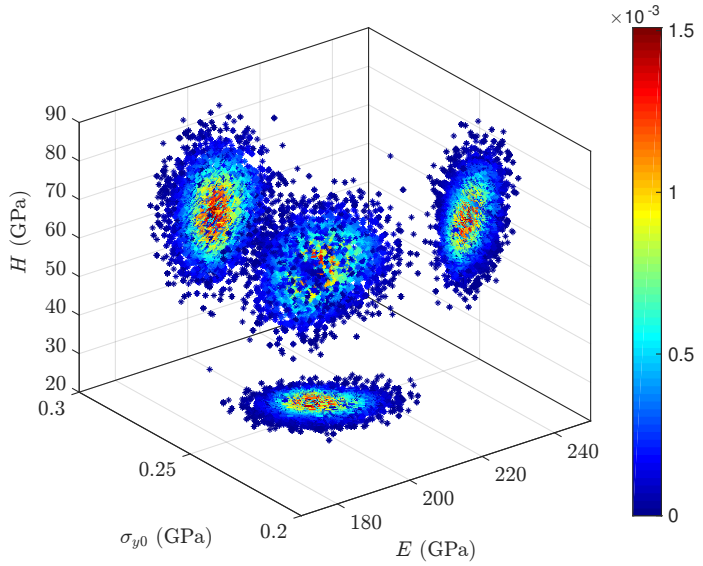
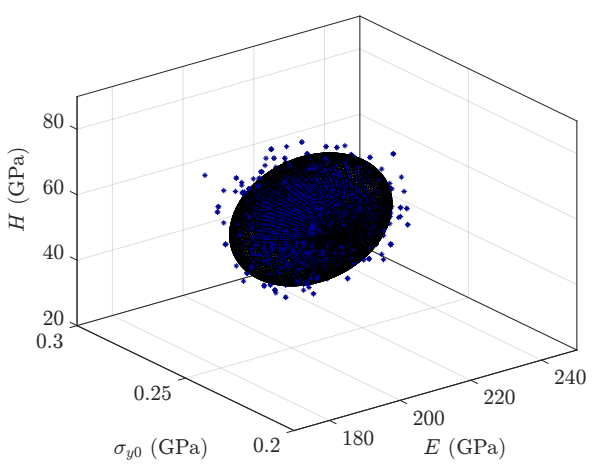
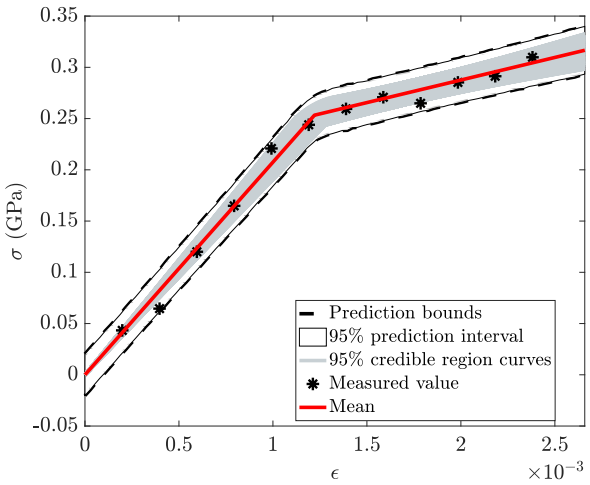


Figure 11: Linear elasticity-linear hardening: Samples generated by the adaptive MCMC approach to approximate the posterior distribution and its projection on the three planes.



(a) The 95% credible region



(b) The measurements, the posterior prediction and the stress-strain curves

Figure 12: Linear elasticity-linear hardening: The 95% credible region and the posterior distribution (a) and the measurements, the posterior prediction and the stress-strain curves associated with the 95% credible region (b).

5.5. Linear elasticity-nonlinear hardening

Identification of the material parameters. In this subsection, twelve measurements are generated using $E = 210$ GPa, $\sigma_{y0} = 0.25$ GPa, $H = 2$ GPa, $n = 0.57$ (which are to be identified) and the same noise distribution as in the previous subsections. The prior distribution is selected in the form of Eq. (42) with the following mean vector and covariance matrix:

$$\overline{\mathbf{x}} = \begin{bmatrix} 200 \\ 0.29 \\ 2.5 \\ 0.57\ \text{GPa}^{-1} \end{bmatrix} \text{GPa}, \quad \mathbf{\Gamma}_{\mathbf{x}} = \begin{bmatrix} 2500 & 0 & 0 & 0 \\ 0 & 2.7778 \times 10^{-4} & 0 & 0 \\ 0 & 0 & 0.1111 & 0 \\ 0 & 0 & 0 & 0.0025\ \text{GPa}^{-2} \end{bmatrix} \text{GPa}^2. \tag{71}$$

Running the adaptive MCMC approach for 10^4 samples and burning the first 3000 samples yields:

$$\widehat{\overline{\mathbf{x}}}_{\text{post}} = \begin{bmatrix} 210.444 \\ 0.254 \\ 2.1937 \\ 0.5988\ \text{GPa}^{-1} \end{bmatrix} \text{GPa},$$

$$\widehat{\mathbf{\Gamma}}_{\text{post}} = \begin{bmatrix} 24.3496 & -8.1743 \times 10^{-3} & 0.1501 & -2.2095 \times 10^{-3}\ \text{GPa}^{-1} \\ -8.1743 \times 10^{-3} & 9.5238 \times 10^{-5} & -6.8472 \times 10^{-4} & 1.8694 \times 10^{-4}\ \text{GPa}^{-1} \\ 0.1501 & -6.8472 \times 10^{-4} & 9.5319 \times 10^{-2} & 5.4179 \times 10^{-3}\ \text{GPa}^{-1} \\ -2.2095 \times 10^{-3}\text{GPa}^{-1} & 1.8694 \times 10^{-4}\ \text{GPa}^{-1} & 5.4179 \times 10^{-3}\ \text{GPa}^{-1} & 1.0629 \times 10^{-3}\ \text{GPa}^{-2} \end{bmatrix} \text{GPa}^2, \tag{72}$$

and

$$\widehat{\mathbf{MAP}} = \begin{bmatrix} 210.0794 \\ 0.2536 \\ 2.198 \\ 0.5978\ \text{GPa}^{-1} \end{bmatrix} \text{GPa}. \tag{73}$$

The stress-strain responses associated with the 95% credible region of the posterior and the posterior prediction are presented in Fig. 13.

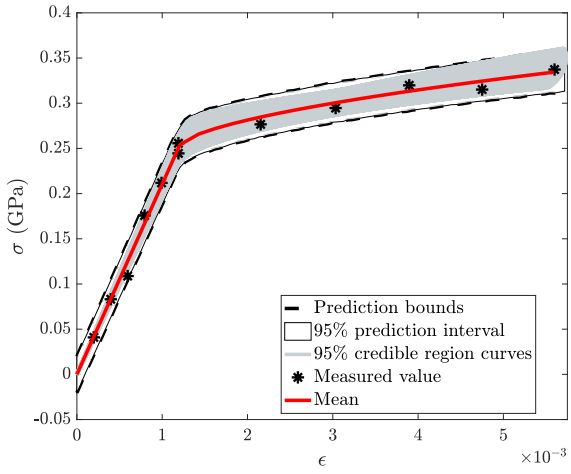


Figure 13: Liner elasticity-nonlinear hardening: The measurements, the stress-strain curves associated with the 95% credible region and the posterior prediction.

6. Additional concepts for parameter identification from uniaxial tensile results

In the previous sections, we have introduced BI for the identification of elastic and elastoplastic parameters. The formulation of the framework was relatively basic. In this section, we will discuss, without too many details, some changes to the framework if (1) a viscoelastic model is considered instead, (2) not only the stress measurements are uncertain, but also the strain measurements, and (3) the model itself is uncertain as well.

6.1. Viscoelasticity

Viscoelasticity differs from the aforementioned elastoplasticity by its rate-dependency (whilst no plastic deformation occurs). Consequently, the rate-dependency must be identified, besides one or more stiffness parameters. The rate at which uniaxial tensile tests are performed thus comes into play. On the other hand, the mechanical responses are C_1-continuous, which results in C_1-continuous posteriors. This may be considered as more straightforward to treat with MCMC approaches than the posteriors of the aforementioned elastoplastic descriptions.

Different types of uniaxial tensile tests can be considered. In a constant rate experiment, a constant clamp velocity is prescribed, resulting in a stress-time response. In a relaxation test, a user-selected displacement (i.e. strain) is enforced as fast as possible in the beginning of the test and then kept constant. The result of this is a stress-time response. In a creep test, a user-selected force (i.e. stress) is prescribed as fast as possible in the beginning of the test and then kept constant. The result of this is a strain-time response. In relaxation and creep tests, the stress-strain response for the material descriptions in this contribution must be replaced by a stress-time response or a strain-time response. As these are both C_1-continuous, their posteriors are also C_1-continuous and hence, the MCMC algorithm to explore them is easier to implement than for elastoplasticity.

The study of Rappel et al. [25] shows that the effect of the prior on the mean and MAP point in viscoelasticity is larger than for elastoplasticity. The influence is especially larger for the damping parameter. Although an increase of the number of measurements decreases the influence of the prior, its influence on the damping parameter remains recognisable. An interesting result of using BI is that it was shown that the uncertainty level of the identified parameter values is substantially larger if uniaxial tensile tests at constant strain-rates are used than if relaxation or creep tests are used.

6.2. Noise in both stress and strain

Depending on the experimental setup, the strains may also be contaminated by noise. The noise of the strains may therefore be worth to incorporate. The additive noise model if both the stresses and strains are contaminated by their own stochastic noise, can be expressed as follows:

$$\begin{cases} Y = \sigma(\epsilon, \mathbf{x}) + \Omega_y \\ \epsilon^* = \epsilon + \Omega_{\epsilon^*} \end{cases}, \tag{74}$$

where ϵ^* denotes the measured strain, ϵ the true strain, Ω_y the error of the stress measurement and Ω_{ϵ^*} the error of the strain measurement. Because the information from both the measured stress and the measured strain is used here, Bayes' formula for multiple variables must be employed [46]:

$$\pi(\mathbf{x}|y, \epsilon^*) = \frac{\pi(\mathbf{x})\pi(\epsilon^*)\pi(y|\mathbf{x}, \epsilon^*)}{\pi(\epsilon^*)\pi(y|\epsilon^*)}. \tag{75}$$

Since the denominator in Eq. (75) is a constant number, the equation above can be written as:

$$\pi(\mathbf{x}|y, \epsilon^*) \propto \pi(\mathbf{x})\pi(y|\mathbf{x}, \epsilon^*). \tag{76}$$

The likelihood function, $\pi(y|\mathbf{x}, \epsilon^*)$, must be determined by integration (over ϵ [36]), because $\pi(y|\mathbf{x}, \epsilon)$ can be determined directly, but $\pi(y|\mathbf{x}, \epsilon^*)$ not. To this end, we write:

$$\pi(y|\mathbf{x}, \epsilon^*) = \int_0^a \pi(y|\mathbf{x}, \epsilon)\pi(\epsilon|\epsilon^*)d\epsilon, \tag{77}$$

where a denotes the physical upper bound of the tensile tester (i.e. the ratio of the original length of the specimen and the maximum distance that the clamps can move). Using Eq. (74), one can express conditional probabilities $\pi(y|\mathbf{x}, \epsilon)$ and $\pi(\epsilon|\epsilon^*)$ as follows:

$$\begin{cases} \pi(y|\mathbf{x}, \epsilon) = \pi_y(y - \sigma(\epsilon, \mathbf{x})) \\ \pi(\epsilon|\epsilon^*) = \pi_{\epsilon^*}(\epsilon^* - \epsilon) \end{cases}, \tag{78}$$

where $\pi_y(\omega_y)$ and $\pi_{\epsilon^*}(\omega_{\epsilon^*})$ denote the noise distributions of the errors in the stress measurements and the strain measurements, respectively. For n_m independent measurements, we write:

$$\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\epsilon}^*) = \prod_{i=1}^{n_m} \pi(y_i|\mathbf{x}, \epsilon_i^*), \tag{79}$$

where $\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_{n_m} \end{bmatrix}^T$ denotes the vector with the n_m stress measurements, $\boldsymbol{\epsilon}^* = \begin{bmatrix} \epsilon_1^* & \cdots & \epsilon_{n_m}^* \end{bmatrix}^T$ the vector with the n_m measured strains and $\pi(y_i|\mathbf{x}, \epsilon_i^*)$ is given in Eq. (77). Further details on the resulting likelihoods for different material models is presented in [47].

We now focus on a simple example for linear elasticity with one measurement point, given by $y = 0.1576$ GPa and $\epsilon^* = 7.25 \times 10^{-4}$. Both noise distributions are considered as normal distributions with a zero mean and $s_y = 0.01$ GPa and $s_{\epsilon^*} = 0.0001$ for the noise in stress and the noise in the strain, respectively.

In Fig. 14 the posteriors are shown if only the noise in the stress measurement is considered and if the noise in the stress measurement as well as in the strain measurement is considered. The prior is also presented, which is in form of Eq. (28) with mean $\overline{E} = 150$ GPa and standard deviation $s_E = 50$ GPa. Not only is the posterior for the double uncertainty case wider than the other one, it is also asymmetric.

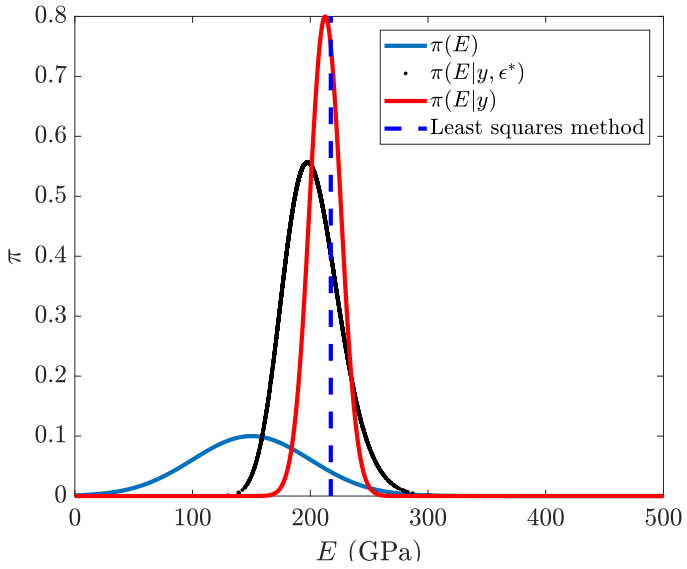


Figure 14: Noise in the stress and strain: The prior and posterior if both the stress and the strain are corrupted by noise (black dots), the posterior if only the stress is corrupted (red) and the value predicted by the least squares method (blue dashed). The posterior for the case with noise in the stress and strain is wider than the posterior for the case with noise in the stress only. Note furthermore that the posterior if only the noise in the stress is considered is a (modified) normal distribution (symmetric), but the posterior if both noises are incorporated neither is a (modified) normal distribution, nor is it symmetric.

Fig. 15 shows the posterior predictions for the same case, except that ten measurements are considered. One can see that incorporating the error results in a wider envelope that includes more validation points. Note that the validation points are only used to assess the quality of the predictions based on the identified material parameter (i.e. E) and not for the actual identification.

Incorporating not only the error in the stress, but also the error in the strain often results in a larger uncertainty (wider posterior) and consequently, the posterior prediction interval encompasses more measurement data. More information can be found in [47].

6.3. Model uncertainty

So far in this contribution, the modelling error (model uncertainty) has not been incorporated. However, no model is completely correct and model uncertainty as an error source may be incorporated. A framework able to do so was developed by Kennedy and O'Hagan [10] (the 'KOH' framework). In this framework, the difference between model response $\sigma(\epsilon, \mathbf{x})$ and true response σ_{true} is written as an additive uncertainty term [48]:

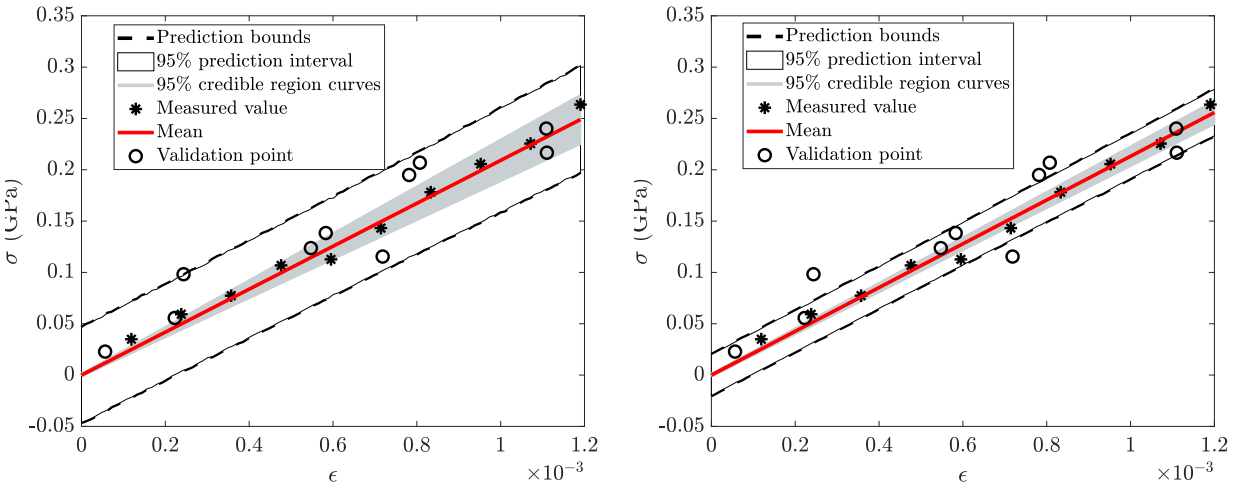
$$Y = \sigma_{\text{true}} + \Omega, \tag{80}$$

with

$$\sigma_{\text{true}} = \sigma(\epsilon, \mathbf{x}) + d(\epsilon, \mathbf{x}_d), \tag{81}$$

where $d(\epsilon, \mathbf{x}_d)$ denotes the model uncertainty, which may be assumed to depend on the input (i.e. strain ϵ here), and $\mathbf{x}_d$ denotes the parameter vector of the model uncertainty. Assuming that both the form and parameters of the noise distribution of the stress measurements are known, the likelihood function now reads:

$$\pi(y|\mathbf{x}, \mathbf{x}_d) = \pi_{\text{noise}}(y - \sigma(\epsilon, \mathbf{x}) - d(\epsilon, \mathbf{x}_d)). \tag{82}$$



(a) Uncertainty in both the stress and the strain measurements

(b) Uncertainty in the stress only

Figure 15: Noise in the stress and strain: The measurements, the validation points, the posterior prediction and the stress-strain curves associated with the 95% credible region of the posterior for (a) noise in the stress and the strain and (b) noise in the stress only. One can see that the uncertainty is larger for the case with noise in the stress and strain, than that for the case with noise in the stress only. Consequently, the posterior predictions if both the noise in the stress and strain is incorporated, includes more validation points.

Note that we have not incorporated the error in the strain measurements for simplicity. Readers are referred to [47] for cases in which the error in the strain is also considered.

Using Eqs. (23) and (82), the posterior distribution for a single measurement can be written as:

$$\pi(\mathbf{x}, \mathbf{x}_d|y) \propto \pi(y|\mathbf{x}, \mathbf{x}_d)\pi(\mathbf{x})\pi(\mathbf{x}_d). \tag{83}$$

For several independent measurements, the final likelihood function is the product of the likelihood function for each measurement:

$$\pi(\mathbf{y}|\mathbf{x}, \mathbf{x}_d) = \prod_{i=1}^{n_m} \pi_{\text{noise}}(y_i - \sigma(\epsilon_i, \mathbf{x}) - d(\epsilon_i, \mathbf{x}_d)). \tag{84}$$

After establishing the posterior, the posterior needs to be sampled numerically (see Subsection 4.1.1) in order to obtain the statistical summaries (e.g. mean value, MAP point or covariance matrix). Note that the posterior distribution of Eq. (83) is a joint distribution of $\mathbf{x}$ and $\mathbf{x}_d$. In order to sample the marginal distribution of each parameter (e.g. the Young's modulus), one only needs to consider the samples of that specific parameter in the joint posterior distribution (i.e. Eq.(83)) and ignore those of the other parameters [49].

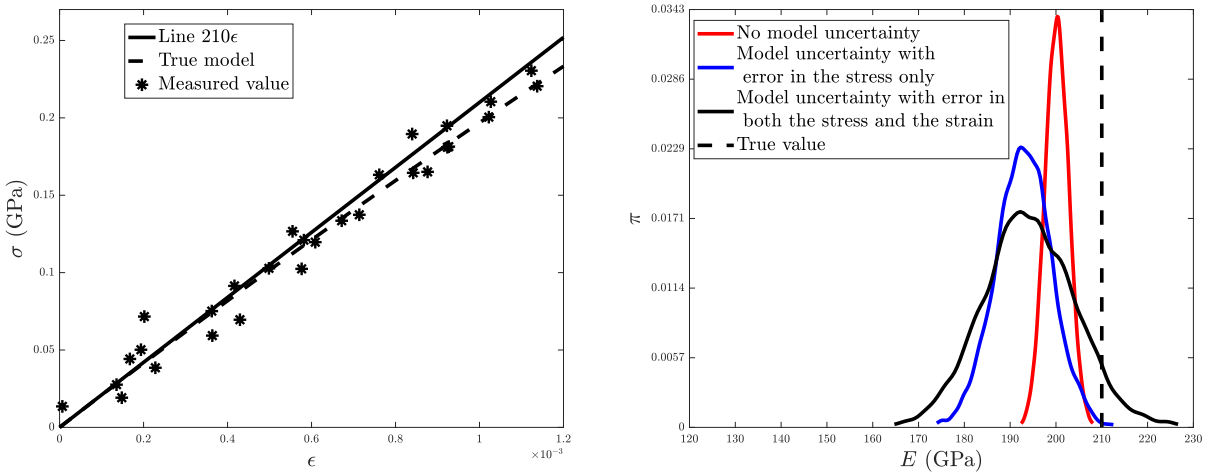
Various formulations have been employed in different studies to express the model uncertainty term in Eq. (81). Probably the simplest way is to represent the model uncertainty using a single deterministic variable [50]. It can also be described by a deterministic, input-dependent function [51]. Another way to express this uncertainty is to describe it by a random variable coming from a normal distribution and include the parameters of this distribution in the posterior distribution [47, 48, 50]. The parameters of the normal distribution from which the model uncertainty is originating can also be input-dependent functions (i.e. strain) [47, 48]. The model uncertainty can also be represented as a Gaussian process [10, 52, 53]. Some more formulations to describe model uncertainty can be found in [48].

As an example here, we consider a nonlinear curve (i.e. dashed line in Fig. 16(a)) as our true material response, whereas the model uses linear elasticity. Thirty measurements are generated artificially (see

Fig. 16(a)). Model uncertainty is described by a random variable coming from a normal distribution with constant parameters and hence, the mean and standard deviation of this normal distribution appear as parameters in the posterior.

The marginal posteriors of the Young's modulus are presented in Fig. 16(b). One can observe that the incorporation of model uncertainty in this example results in a wider posterior distribution that includes the true value, although its MAP point is located further away from the true value. If the error in the strain is incorporated as well, the posterior at the true value increases even more.

The posterior predictions for these three cases are shown in Fig. 17. Incorporating model uncertainty clearly results in a wider prediction interval. If both the error in the strain and model uncertainty are incorporated however, the prediction interval becomes even wider and all measurement points and validation points are present within its bounds.



(a) True curve

(b) Marginal posteriors of Young's modulus

Figure 16: Model uncertainty: (a) The true stress-strain curve from which the measurements are generated, as well as the linear stress-strain curve with the true Young's modulus. (b) The marginal posterior distribution of the Young's modulus with no model uncertainty (red curve), with model uncertainty and error in the stress (blue curve) and with model uncertainty and error in both the stress and the strain measurements (black curve). Incorporating model uncertainty results in a posterior distribution that includes the true Young's modulus. In case the error in the strain is considered as well, results furthermore in a higher possibility at the true value.

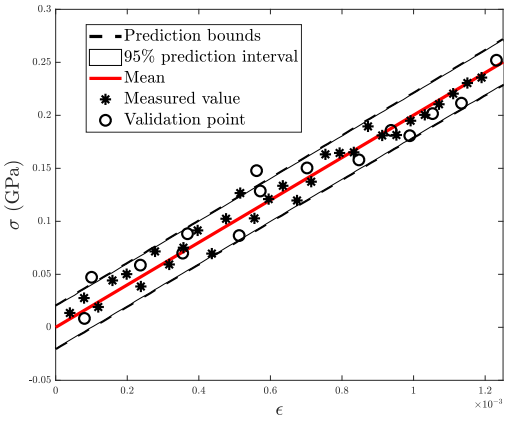
7. Conclusions

Although BI has been employed in various studies for parameter identification, most may not be straightforward to understand for those who are new to the subject. In this contribution, we have aimed to explain BI in a straightforward and didactic manner. For this purpose, a number of Bayesian inference formulations are presented to identify elastic and elastoplastic material parameters from uniaxial tensile results. Elastic and elastoplastic material models are chosen for two reasons: (1) they are widely used in solid mechanics and (2) they include the most simple material behaviour (linear elasticity), as well as increasingly complex descriptions such as linear elasticity-nonlinear hardening, which entails C_0-continuous, implicit responses.

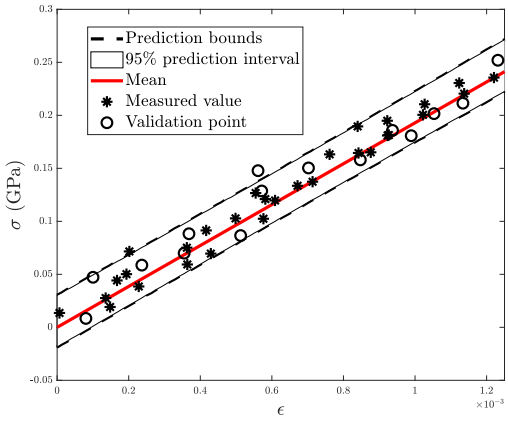
The following conclusions can be made based on the examples given in Section 5:

(1) The results of BI cannot directly be compared to those of the least squares method, since the result of BI is a distribution and that of the least squares method is a single value.

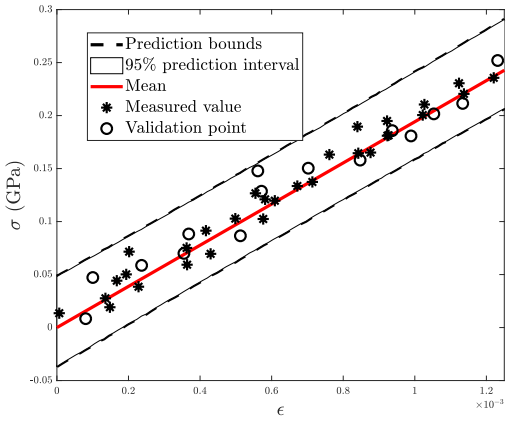
(2) If one wants to compare the two nevertheless, point estimators such as the mean and MAP point can be compared to the results of the least squares method. It is shown in Fig. 5 that the selected prior



(a) Error in the stress



(b) Error in the stress and model uncertainty



(c) Errors in the stress and strain and model uncertainty

Figure 17: Model uncertainty: The measurements, validation points and the posterior predictions. One can see that if model uncertainty is considered, the prediction interval becomes wider. Furthermore, only if model uncertainty as well as the errors in both the stress and the strain measurements are incorporated, all measured and validation points are inside the prediction interval.

distribution may significantly influence the results. Fig. 5 also shows that the influence of the prior decreases significantly if the number of measurements increases.

(3) The standard deviations and correlations of the material parameters established using the 'standard' BI formulations presented in this contribution, do *not* reflect the heterogeneity of the material parameters. In other words, they are *not* representative for the standard deviations and correlations of the intrinsic material parameter distributions, but only for the level of uncertainty. The reason is that the formulations in this contribution still assume that a unique set of parameter values is the solution of the identification problem.

Acknowledgements

Hussein Rappel, Lars A.A. Beex and Stéphane P.A. Bordas would like to acknowledge the financial support from the University of Luxembourg. Stéphane P.A. Bordas also thanks the European Research Council Starting Independent Research Grant (ERC Stg grant agreement No. 279578) entitled "Towards real time multiscale simulation of cutting in nonlinear materials with applications to surgical simulation and

computer guided surgery". Jack S. Hale is supported by the National Research Fund, Luxembourg, and cofunded under the Marie Curie Actions of the European Commission (FP7-COFUND Grant No. 6693582).

Compliance with Ethical Standards

The authors declare that they have no conflict of interest.

References

[1] B.S. Everitt, A. Skrondal, The Cambridge dictionary of statistics, Cambridge University Press, 2010.

[2] R.E. Walpole, R.H. Myers, S.L. Myers, K. Ye, Probability andstatistics for engineers and scientists, Pearson custom library, Pearson, 2013.

[3] C. Gogu, R. Haftka, R.L. Riche, J. Molimard, A. Vautrin, Introduction to the Bayesian approach applied to elastic constants identification, AIAA journal 48 (5) (2010) 893–903.

[4] D. Higdon, H. Lee, Z. Bi, A Bayesian approach to characterizing uncertainty in inverse problems using coarse and fine scale information, IEEE Transactions on Signal Processing 50 (2002) 388–399.

[5] J. Wang, N. Zabaras, A Bayesian inference approach to the inverse heat conduction problem, International Journal of Heat and Mass Transfer 47 (17-18) (2004) 3927–3941.

[6] P. Risholm, F. Janoos, I. Norton, A.J. Golby, W.M. Wells, Bayesian characterization of uncertainty in intra-subject non-rigid registration, Medical image analysis 17 (5) (2013) 538–555.

[7] S. Lan, T. Bui-Thanh, M. Christie, M. Girolami, Emulation of higher-order tensors in manifold Monte Carlo methods for Bayesian inverse problems, Journal of Computational Physics 308 (2016) 81–101.

[8] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, Journal of Engineering Mechanics 124 (4) (1998) 455–461.

[9] C.K. Oh, J.L. Beck, M. Yamada, Bayesian learning using automatic relevance determination prior with an application to earthquake early warning, Journal of Engineering Mechanics 134 (12) (2008) 1013–1020.

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

[11] J. Kaipio, E. Somersalo, Statistical and computational inverse problems, Vol. 160, Springer Science & Business Media, 2006.

[12] J. Isenberg, Progressing from least squares to Bayesian estimation, in: Proceedings of the 1979 ASME design engineering technical conference, New York, 1979, pp. 1–11.

[13] K.F. Alvin, Finite element model update via Bayesian estimation and minimization of dynamic residuals, AIAA journal 35 (5) (1997) 879–886.

[14] T. Marwala, S. Sibusiso, Finite element model updating using Bayesian framework and modal properties, Journal of Aircraft 42 (1) (2005) 275–278.

[15] T.C. Lai, K.H. Ip, Parameter estimation of orthotropic plates by Bayesian sensitivity analysis, Composite Structures 34 (1) (1996) 29–42.

[16] F. Daghia, S. de Miranda, F. Ubertini, E. Viola, Estimation of elastic constants of thick laminated plates within a Bayesian framework, Composite Structures 80 (3) (2007) 461–473.

[17] P.S. Koutsourelakis, A novel Bayesian strategy for the identification of spatially varying material properties and model validation: An application to static elastography, International Journal for Numerical Methods in Engineering 91 (3) (2012) 249–268.

[18] C. Gogu, W. Yin, R. Haftka, P. Ifju, J. Molimard, R. Le Riche, A. Vautrin, Bayesian identification of elastic constants in multi-directional laminate from moiré interferometry displacement fields, Experimental Mechanics 53 (4) (2013) 635–648.

[19] M. Muto, J.L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, Journal of Vibration and Control 14 (1-2) (2008) 7–34.

[20] P. Liu, S.K. Au, Bayesian parameter identification of hysteretic behavior of composite walls, Probabilistic Engineering Mechanics 34 (2013) 101–109.

[21] D.D. Fitzenz, A. Jalobeanu, S.H. Hickman, Integrating laboratory creep compaction data with numerical fault models: A Bayesian framework, Journal of Geophysical Research: Solid Earth 112 (B8), B08410.

[22] T. Most, Identification of the parameters of complex constitutive models: Least squares minimization vs. Bayesian updating, in: D. Straub (Ed.), Reliability and optimization of structural systems, CRC press, 2010, pp. 119–130.

[23] B.V. Rosić, A. Kčerová, J. Sýkora, O. Pajonk, A. Litvinenko, H.G. Matthies, Parameter identification in a probabilistic setting, Engineering Structures 50 (2013) 179–196, Engineering Structures: Modelling and Computations (special issue IASS-IACM 2012).

[24] W.P. Hernandez, F.C.L. Borges, D.A. Castello, N. Roitman, C. Magluta, Bayesian inference applied on model calibration of fractional derivative viscoelastic model, in: V. Steffen Jr, D.A. Rade, W.M. Bessa (Eds.), DINAME 2015-Proceedings of the XVII International symposium on dynamic problems of mechanics, Natal, 2015.

[25] H. Rappel, L.A.A. Beex, S.P.A. Bordas, Bayesian inference to identify parameters in viscoelasticity, Mechanics of Time-Dependent Materials 22 (2) (2018) 221–258.

[26] J.M. Nichols, W.A. Link, K.D. Murphy, C.C. Olson, A Bayesian approach to identifying structural nonlinearity using free-decay response: Application to damage detection in composites, Journal of Sound and Vibration 329 (15) (2010) 2995–3007.

[27] S. Abhinav, C.S. Manohar, Bayesian parameter identification in dynamic state space models using modified measurement equations, International Journal of Non-Linear Mechanics 71 (2015) 89–103.

[28] S. Madireddy, B. Sista, K. Vemaganti, A Bayesian approach to selecting hyperelastic constitutive models of soft tissue, Computer Methods in Applied Mechanics and Engineering 291 (2015) 102–122.

[29] J.T. Oden, E.E. Prudencio, A. Hawkins-Daarud, Selection and assesment of phenomenological models of tumor growth, Mathematical Models and Methods in Applied Sciences 23 (7) (2013) 1309–1338.

[30] J. Chiachío, M. Chiachío, A. Saxena, S. Sankararaman, G. Rus, K. Goebel, Bayesian model selection and parameter estimation for fatigue damage progression models in composites, International Journal of Fatigue 70 (2015) 361–373.

[31] I. Babuška, Z. Sawlan, M. Scavino, B. Szabó, R. Tempone, Bayesian inference and model comparison for metallic fatigue data, Computer Methods in Applied Mechanics and Engineering 304 (2016) 171–196.

[32] S. Sarkar, D.S. Kosson, S. Mahadevan, J.C.L. Meeussen, H. van der Sloot, J.R. Arnold, K.G. Brown, Bayesian calibration of thermodynamic parameters for geochemical speciation modeling of cementitious materials, Cement and Concrete Research 42 (7) (2012) 889–902.

[33] S.L. Cotter, M. Dashti, J.C. Robinson, A.M. Stuart, Bayesian inverse problems for functions and applications to fluid mechanics, Inverse Problems 25 (11) (2009) 115008.

[34] J.C. Simo, T.J.R. Hughes, Computational inelasticity, Springer Science & Business Media, New York, 2000.

[35] T.J. Ulrych, M.D. Sacchi, A. Woodbury, A Bayes tour of inversion: A tutorial, GEOPHYSICS 66 (1) (2001) 55–69.

[36] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, Bayesian data analysis, Chapman & Hall/CRC Texts in Statistical Science, Chapman & Hall/CRC, 2003.

[37] J.L. Beck, S.K. Au, Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation, Journal of Engineering Mechanics 128 (4) (2002) 380–391.

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

[39] J. Kristensen, N. Zabaras, Bayesian uncertainty quantification in the evaluation of alloy properties with the cluster expansion method, Computer Physics Communications 185 (11) (2014) 2885–2892.

[40] C. Andrieu, N. De Freitas, A. Doucet, M.I. Jordan, An introduction to MCMC for machine learning, Machine learning 50 (1-2) (2003) 5–43.

[41] S. Brooks, A. Gelman, G. Jones, X.L. Meng, Handbook of Markov chain Monte Carlo, CRC press, 2011.

[42] S. Sinharay, Assessing convergence of the Markov chain Monte Carlo algorithms: A review, ETS Research Report Series 2003 (1) (2003) i–52.

[43] A. Gelman, G.O. Roberts, W.R. Gilks, Efficient Metropolis jumping rules, in: J.M. Bernardo, J.O. Berger, A.P. Dawid, A.F.M. Smith (Eds.), Bayesian statistics, Vol. 5 of Oxford Sci. Publ., Oxford University Press, New York, 1996, pp. 599–607.

[44] H. Haario, E. Saksman, J. Tamminen, Adaptive proposal distribution for random walk Metropolis algorithm, Computational Statistics 14 (3) (1999) 375–396.

[45] J.L. Beck, Bayesian system identification based on probability logic, Structural Control and Health Monitoring 17 (7) (2010) 825–847.

[46] S.J.D. Prince, Computer vision: Models learning and inference, Cambridge University Press, 2012.

[47] H. Rappel, L.A.A. Beex, L. Noels, S.P.A. Bordas, Identifying elastoplastic parameters with Bayes' theorem considering double error sources and model uncertainty, Probabilistic Engineering Mechanics. URL `https://doi.org/10.1016/j.probengmech.2018.08.004`

[48] Y. Ling, J. Mullins, S. Mahadevan, Selection of model discrepancy priors in Bayesian calibration, Journal of Computational Physics 276 (Supplement C) (2014) 665–680.

[49] C. Bishop, Pattern recognition and machine learning, Information Science and Statistics, Springer, 2006.

[50] G.B. Arhonditsis, D. Papantou, W. Zhang, G. Perhar, E. Massos, M. Shi, Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management, Journal of Marine Systems 73 (1) (2008) 8–30.

[51] Y. Xiong, W. Chen, K.L. Tsui, D.W. Apley, A better understanding of model updating strategies in validating engineering models, Computer Methods in Applied Mechanics and Engineering 198 (15) (2009) 1327–1337.

[52] P.D. Arendt, D.W. Apley, W. Chen, Quantification of model uncertainty: Calibration, model discrepancy, and identifiability, Journal of Mechanical Design 134 (10) (2012) 100908.

[53] J. Brynjarsdóttir, A. O'Hagan, Learning about physical parameters: The importance of model discrepancy, Inverse Problems 30 (11) (2014) 114007.