Applications of Mathematics

Radim Blaheta; Michal Béreš; Simona Domesová; Pengzhi Pan

A comparison of deterministic and Bayesian inverse with application in micromechanics

Applications of Mathematics, Vol. 63 (2018), No. 6, 665–686

Persistent URL: http://dml.cz/dmlcz/147563

63 (2018) APPLICATIONS OF MATHEMATICS No. 6, 665–686

A COMPARISON OF DETERMINISTIC AND BAYESIAN INVERSE WITH APPLICATION IN MICROMECHANICS

Radim Blaheta, Michal Béreš, Simona Domesová, Ostrava,
Pengzhi Pan, Wuhan

Received July 16, 2018. Published online November 19, 2018.

Abstract. The paper deals with formulation and numerical solution of problems of identification of material parameters for continuum mechanics problems in domains with heterogeneous microstructure. Due to a restricted number of measurements of quantities related to physical processes, we assume additional information about the microstructure geometry provided by CT scan or similar analysis. The inverse problems use output least squares cost functionals with values obtained from averages of state problem quantities over parts of the boundary and Tikhonov regularization. To include uncertainties in observed values, Bayesian inversion is also considered in order to obtain a statistical description of unknown material parameters from sampling provided by the Metropolis-Hastings algorithm accelerated by using the stochastic Galerkin method. The connection between Bayesian inversion and Tikhonov regularization and advantages of each approach are also discussed.

Keywords: inverse problems; Bayesian approach; stochastic Galerkin method

MSC 2010: 86-08, 82-08, 65N21, 65C60, 60-08

1. Introduction

The paper begins with a description of a kind of inverse problems for identification of local material characteristics of heterogeneous (composite) materials, which

The authors acknowledge the support from the scholarship for visiting scholars program (Grant No. Z016001) of the State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences. The work was also supported by the projects CZ.1.05/1.1.00/02.0070 and National Programme of Sustainability (NPS II) project "IT4Innovations excellence in science – LQ1602" funded by the Ministry of Education, Youth and Sports of the Czech Republic and by the State Key Research Development Program of China (Grant No. 2017YFC0804203) and the Key Research Program of Frontier Sciences, CAS (QYZDB-SSW-DQC029).

appear in micromechanics. Then, it shows an application of both deterministic and stochastic tools for analysis of these inverse problems. It finishes with conclusions concerning a comparison of both approaches.

For a description of the inverse problems, we restrict ourselves to the analysis of stationary physical processes described by a differential equation $D(k,u) = f$ in the domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with boundary conditions $C(k,u) = g$ on the boundary $\partial\Omega$. Moreover, we also consider a sequence of boundary value problems (tests, state problems) for $i = 1, \ldots, m$,

$$D(k,u_i) = f_i \text{ in } \Omega \quad \text{and} \quad C_i(k,u_i) = g_i \text{ on } \partial\Omega,$$

with the same material coefficient k but possibly different sources f_i mainly with different localization and data for boundary conditions.

As an example, we consider a (Darcy) flow in saturated porous media or heat conduction described by the equation $D(k,u_i) = -\mathrm{div}(k\nabla u_i)$. We assume isotropic material, i.e. $k\colon \Omega \to \mathbb{R}^1$ is a scalar function on Ω. The boundary is divided into disjoint Dirichlet and Neumann parts, $\partial\Omega = \Gamma_i^D \cup \Gamma_i^N$, where

$$\Gamma_i^D = \bigcup_{j=1}^{\#D_i} \Gamma_{ij}^D \quad \text{and} \quad \Gamma_i^N = \bigcup_{j}^{\#N_i} \Gamma_{ij}^N.$$

The boundary conditions are of the form

$$\begin{aligned} C_i(k,u_i) &= u_i = g_{ij}^D \quad \text{on } \Gamma_{ij}^D \subset \partial\Omega, \\ C_i(k,u_i) &= -n \cdot k\nabla u_i = g_{ij}^N \quad \text{on } \Gamma_{ij}^N \subset \partial\Omega, \end{aligned}$$

where n denotes the unit outer normal to $\partial\Omega$.

If there are enough measurements of the state variable u_i or values derived from u_i, then we can think about using these measurements for determination or approximation of the function k, which represents material characteristics (permeability, conductivity). But in the case considered in this paper, we typically do not have enough measurements to recover k without additional information. Suitable additional information can be a knowledge of the decomposition of the domain $\Omega = \bigcup_{l=1}^{\#S} \Omega_l$ in such way that the function k is constant on Ω_l for all $l = 1, \ldots, \#S$. In the case of material samples, such decomposition can be naturally determined e.g. by X-ray CT scanning, see e.g. [4]. Consequences of such a priori given domain decomposition are discussed e.g. in [12].

To solve the inverse problems, we first use a deterministic approach using output least squares minimization, see e.g. [22]. Second, we use a stochastic approach

with Bayesian inverse and its implementation via delayed acceptance variant of the Metropolis-Hastings algorithm (DAMH) and surrogate model provided by the stochastic Galerkin method.

The paper is organized as follows. In Section 2, we specify the considered inverse problem in more details. Especially, we restrict ourselves to the data provided by measurements of integral flux over parts of the boundary. The inverse identification problem is formulated as an output least squares minimization. Section 3 then describes a model problem and experience with inverse analysis realized through a minimization of the cost functional by a classical method (quasi-Newton). It is shown that the minimization of the cost functional without regularization is not sufficient in the case of measurements corrupted by noise. It is also shown that Tikhonov type regularization with a suitable weighting parameter can help. In Section 4, we consider a Bayesian inversion instead of the regularization. The Bayesian inversion requires additional statistical information but provides a broader and more natural information. An implementation of the Bayesian inversion with the use of the Metropolis-Hastings algorithm and the stochastic Galerkin method is discussed in Section 5. Numerical experiments with the Bayesian inversion are described in Section 6 and a summary with a comparison of deterministic and stochastic approach and possible generalizations are outlined in Section 7.

2. Inverse problems in micromechanics

In micromechanics, one considers mainly tests and their evaluation on laboratory size samples. We assume samples of heterogeneous (composite) materials, similarly to [3]. In many cases, the geometry of the microstructure can be determined by CT scans. This geometry can be of different types with separated and mixed material distributions, see Fig. 2.1.

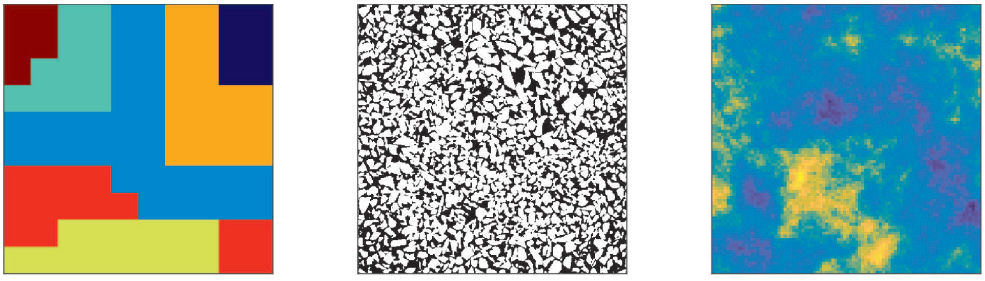


Figure 2.1. Different microstructure geometries, left: separated material subdomains, middle: a CT scan of a binary material (marble in PE resin), right: a randomly generated log-normal random field.

The testing of material samples is usually performed with zero volume sources but with different boundary conditions. The testing includes measurements of quantities derived from the state variables u_i. In our case, it will be measurements of the flux over parts Γ_{ij}^D of the Dirichlet boundary Γ_i^D. More precisely, let us consider one or more tests T_i, $i = 1, \dots, m$, of the process (flow) described by the boundary value problem

$$\begin{cases} -\operatorname{div}(k\nabla u_i) = 0 & \text{in } \Omega, \\ u_i = g_{ij}^D & \text{on } \Gamma_{ij}^D,\ j = 1, \dots, \#D_i, \\ n \cdot k\nabla u_i = k\dfrac{\partial u_i}{\partial n} = 0 & \text{on } \Gamma_i^N,\ j = 1, \dots, \#N_i. \end{cases} \tag{2.1}$$

We assume that we are able to measure some integral quantities on parts $\Gamma_{ij}^D \subset \Gamma_i^D \subset \partial\Omega$, $j = 1, \dots, \#D_i$, $\#D_i > 1$. In particular, we consider the integral flux

$$q_{ij} = \int_{\Gamma_{ij}^D} n \cdot k\nabla u_i \,\mathrm{d}s = \int_{\Gamma_{ij}^D} k\frac{\partial u_i}{\partial n} \,\mathrm{d}s \tag{2.2}$$

over $\Gamma_{ij}^D \subset \partial\Omega$ in the direction of (unit) outer normal n. It is recommended to get as many measurements as possible, which means using more parts Γ_{ij}^D as well as more tests ($m > 1$).

Let μ_{ij} denote the measurements of q_{ij}. Then the inverse problem for the identification of k can be written as a minimization problem:

$$\begin{gathered} \text{find } k^* \in \mathcal{U}_{\mathrm{ad}} \text{ such that } \mathcal{J}(k^*) \leqslant \mathcal{J}(k) \quad \forall\, k \in \mathcal{U}_{\mathrm{ad}}, \\ \mathcal{J}(k) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{\#D_i} \mathcal{J}_{ij}(k), \quad \mathcal{J}_{ij}(k) = (q_{ij} - \mu_{ij})^2, \\ \mathcal{U}_{\mathrm{ad}} = \{k \in L^\infty(\Omega);\ 0 < k_{\min} \leqslant k \leqslant k_{\max} \text{ in } \Omega,\ k|_{\Omega_l} \in P_0(\Omega_i),\ l = 1, \dots, \#S\}, \end{gathered} \tag{2.3}$$

where P_0 denotes a set of constant functions. Note that for $k \in \mathcal{U}_{\mathrm{ad}}$ the values $k|_{\Omega_l}$ are positive and can be represented as values of the exponential function. Therefore, these values can be represented by a vector $\boldsymbol{k}_{\exp}$ from a compact set $\boldsymbol{\mathcal{U}}_{\mathrm{ad}} \subset \mathbb{R}_+^{\#S} \subset \mathbb{R}^{\#S}$, where $\mathbb{R}_+^{\#S}$ is the set of vectors with all components positive. For the numerical optimization, we shall use the transformation to $\boldsymbol{k} = \ln(\boldsymbol{k}_{\exp})$ componentwise.

To enable the computation of integral fluxes q_{ij} directly, we use a mixed variational formulation of problem (2.1) with the quantities u (pressure) and v (velocity) as two state variables. More precisely, let

$$U = L_2(\Omega) \quad \text{and} \quad V_i = \{v \in H(\operatorname{div}, \Omega);\ v \cdot n = 0 \text{ on } \Gamma_i^N\}.$$

Then we seek for $(u_i, v_i) \in U \times V_i$ such that

$$\int_\Omega k^{-1} v_i \cdot w \,\mathrm{d}x - \int_\Omega \operatorname{div}(w) \cdot u_i \,\mathrm{d}x = -\sum_{j=1}^{\#D_i} \int_{\Gamma_{ij}^D} g_{ij}^D (w \cdot n) \,\mathrm{d}s \quad \forall\, w \in V_i,$$

$$-\int_\Omega \operatorname{div}(v_i) z \,\mathrm{d}x = 0 \quad \forall\, z \in U.$$

It can be shown that there exists a unique solution $(u_i, v_i) \in U \times V_i$ and

$$q_{ij} = -\int_{\Gamma_{ij}^D} n \cdot v_i \,\mathrm{d}s.$$

See e.g. [5] or [11].

For the numerical implementation, we replace the function spaces U, V_i by their finite element approximations U_h, V_{ih}. Especially, we use a triangulation of the domain Ω aligned with the decomposition of Ω into Ω_l, $l = 1, \ldots, \#S$, and spaces U_h of piece-wise constant functions and $V_{ih} \subset V_h$, where V_h is the lowest order Thomas-Raviart space. Then there exists a unique solution $(u_{ih}, v_{ih}) \in U_h \times V_h$ and

$$q_{ij} \sim q_{ijh} = -\int_{\Gamma_{ij}^D} n \cdot v_{ih} \,\mathrm{d}s. \tag{2.4}$$

See again [5] or [11]. Using the numerically computed fluxes, we can solve a discrete minimization problem (2.3) with $\mathcal{J}$ replaced by $\mathcal{J}_h$,

$$\mathcal{J}_h(k) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{\#D_i} \mathcal{J}_{ijh}(k), \quad \mathcal{J}_{ijh}(k) = (q_{ijh} - \mu_{ij})^2. \tag{2.5}$$

Choosing bases in the finite element spaces, one can formulate the discretized state problem algebraically: find $\boldsymbol{v}_i$, $\boldsymbol{p}_i$ which solve the algebraic system

$$A_i(\boldsymbol{k}_{\mathrm{exp}}) \begin{bmatrix} \boldsymbol{v}_i \\ \boldsymbol{u}_i \end{bmatrix} = F = \begin{bmatrix} F_{v,i} \\ F_P \end{bmatrix}, \quad A_i(\boldsymbol{k}_{\mathrm{exp}}) = \begin{bmatrix} M_i(\boldsymbol{k}_{\mathrm{exp}}) & B_i^{\mathrm{T}} \\ B_i & 0 \end{bmatrix},$$

see [5] or [11]. Note that here $\boldsymbol{k}_{\mathrm{exp}} \in \mathbb{R}^{\#S}$ denotes an algebraic vector, $\boldsymbol{k}_{\mathrm{exp},l} = k|_{\Omega_l}$. The average fluxes q_{ijh} can be grouped into column vectors $q_{ih} = [q_{ijh}\,;\ j = 1, \ldots, \#D_i]$, $q_h = [q_{ih}\,;\ i = 1, \ldots, m]$ and computed from $\boldsymbol{v}_i$, $q_{ih} = L_i \boldsymbol{v}_i \in \mathbb{R}^{\#D_i}$, where L_i are linear operators.

The cost functions can be expressed as $\mathcal{J}(k) = \frac{1}{2}\|G(k) - \boldsymbol{\mu}\|^2$ and $\mathcal{J}_h(\boldsymbol{k}_{\mathrm{exp}}) = \frac{1}{2}\|G_h(\boldsymbol{k}_{\mathrm{exp}}) - \boldsymbol{\mu}\|^2$, where $\boldsymbol{\mu} = [\mu_{ij}] \in \mathbb{R}^{\#M}$, $\#M = \sum_{i=1}^{m} \#D_i$, and the mappings

$G\colon\ k \to q(k)$ and $G_h\colon\ \boldsymbol{k}_{\rm exp} \to q_h(\boldsymbol{k}_{\rm exp})$ are continuous and discretized observation operators, respectively. The following theorem concerns the smoothness properties of G_h and $\mathcal{J}_h$, which are important for the numerical solution of inverse problems.

Theorem 2.1. (i) *The discretized observation operator G_h and the cost function $\mathcal{J}_h$ are continuous in $\mathbb{R}_+^{\#S}$.*

(ii) *The function $\boldsymbol{k}_{\rm exp} \to A_i(\boldsymbol{k}_{\rm exp})$ and the discretized observation operator G_h are continuously differentiable in $\mathbb{R}_+^{\#S}$.*

P r o o f. In the discretized case

$$G_h(\boldsymbol{k}_{\rm exp}) = [q_{1h}, \ldots, q_{mh}]^{\rm T}, \quad q_i = [L_i, 0]A_i^{-1}(\boldsymbol{k}_{\rm exp})F, \quad i = 1, \ldots, m.$$

The elements of $A_i(\boldsymbol{k}_{\rm exp}) = [a_{i,rs}(\boldsymbol{k}_{\rm exp})]$ depend linearly on $\boldsymbol{k}_{\rm exp}$, because they are assembled from local contributions depending linearly on $\boldsymbol{k}_{\rm exp}$. Denote $A_i^{-1}(\boldsymbol{k}_{\rm exp}) = [\bar{a}_{i,rs}(\boldsymbol{k}_{\rm exp})]$. Then

$$\bar{a}_{i,rs}(\boldsymbol{k}_{\rm exp}) = (-1)^{r+s} \det(A_i^{(sr)}(\boldsymbol{k}_{\rm exp}))/\det(A_i(\boldsymbol{k}_{\rm exp})),$$

where det denotes determinant, $A_i^{(rs)}(\boldsymbol{k}_{\rm exp})$ is a submatrix (minor) of A arising after deleting sth row and rth column, $\det(A_i(\boldsymbol{k}_{\rm exp})) \neq 0$ because the finite element system is non-singular. As the elements $a_{i,rs}(\boldsymbol{k}_{\rm exp})$ and consequently $\det(A_i^{(rs)}(\boldsymbol{k}_{\rm exp}))$ and $\det(A_i(\boldsymbol{k}_{\rm exp}))$ depend continuously on $\boldsymbol{k}_{\rm exp}$, it holds that $\bar{a}_{i,rs} \in C(\mathbb{R}_+^{\#S})$. Moreover, $\bar{a}_{i,rs} \in C^l(\mathbb{R}_+^{\#S})$ for any integer $l \geqslant 0$, where C and C^l stand for the spaces of continuous functions and functions with continuous derivations up to order l, respectively. □

Corollary 2.2. (i) *Due to continuity of $\mathcal{J}_h$ and compactness of $\mathcal{U}_{\rm ad}$ from* (2.3), *there exists a solution of the discrete minimization problem.*

(ii) *Minimization is possible with the use of gradient methods.*

Note that for the numerical solution, the gradient should be computed numerically, using numerical differentiation or adjoint equation techniques. It is also possible to show continuity of the observation operator G with consequences similar as discussed in [12].

3. Solution of inverse problem with exact and noisy measurements: An example

Let us consider a model problem of isotropic Darcy flow in a unit square with 7 materials in separated subdomains as shown in Fig. 3.1. The material is represented by $\boldsymbol{k} = \ln(\boldsymbol{k}_{\text{exp}}) \in \mathbb{R}^7$. The localization of the parts with Dirichlet boundary condition can be seen in Fig. 3.2. Note that we use $u_i = g_{ij}^D = 1$ on red (input) ports and $u_i = g_{ij}^D = 0$ on green (output) ports. Different localization of these segments is taken for three considered tests.

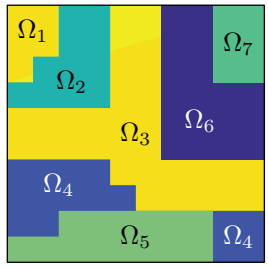


Figure 3.1. Domain Ω divided into 7 artificially selected subdomains.

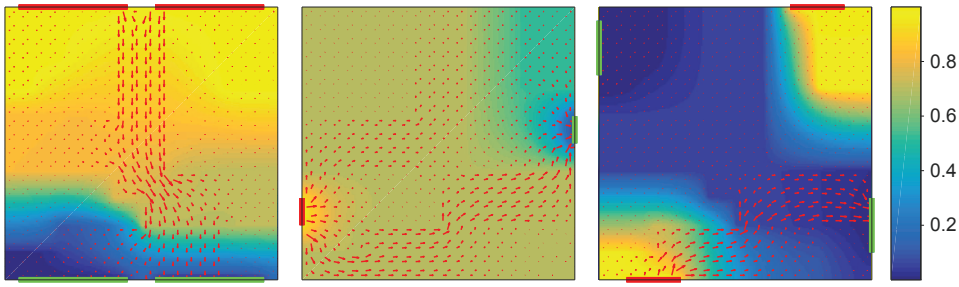


Figure 3.2. Three tests with input (red) boundary segments ($g_i^D = 1$ on Γ_{ij}^D) and output (green) boundary segments ($g_i^D = 0$ on Γ_{ij}^D). The images show velocity (arrows) and pressures (colour scale).

One measurement of the inflow/outflow is performed at each boundary segment providing 4, 2, and 4 values for tests 1, 2, and 3, respectively. We shall work with artificial measurements $\boldsymbol{\mu}_0 = [\mu_{11}, \ldots, \mu_{14}, \mu_{21}, \mu_{22}, \mu_{31}, \ldots, \mu_{34}]$ for all three tests computed for given permeabilities $\boldsymbol{k}_{\text{ref}} = \ln([9000, 100, 5000, 4, 300, 2, 200])$. As an initial guess for the optimization methods, the vector $\boldsymbol{k}_0 = [9, 5, 9, 1, 5, 1, 5]$ was used. As an alternative to $\boldsymbol{\mu}_0$, we consider noisy measurements $\boldsymbol{\mu}_p = \boldsymbol{\mu}_0 + \boldsymbol{\eta}$, where $\boldsymbol{\eta} \sim N(0, p^2\boldsymbol{\Sigma})$ with p being a scaling parameter. The matrix $\boldsymbol{\Sigma}$ is diagonal with values computed as squares of the measurements $\boldsymbol{\mu}_0$.

The finite element discretization of the described type uses a 500×500 grid. The optimization problem (2.3) with approximate $q_{ij} \sim q_{ijh}$ can be solved by various

numerical methods, see e.g. [22]. We assumed that there are box (element-wise) constraints $k_{\min} \leqslant k \leqslant k_{\max}$, but the minimization can be also done under a weaker condition $k > 0$. After the transformation we get an unconstrained minimization problem for $\boldsymbol{k} = \ln(\boldsymbol{k}_{\exp})$. This transformation also brings computational benefits for $\boldsymbol{k}_{\exp}$ with components of different orders of magnitude. We used the quasi-Newton method from MATLAB Optimization Toolbox, which uses a mixed quadratic/cubic line search procedure and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula for updating the approximation of the Hessian matrix. The derivatives are approximated as differences. For more details see [14].

In the case of measurements without noise, the parameters are identified quite precisely, see Fig. 3.3. But for noisy measurements, the identification without regularization fails and we need a regularization, e.g. the Tikhonov regularization with the weight ϱ, $\mathcal{J}_{TR}(\boldsymbol{k}) = \mathcal{J}_h(\boldsymbol{k}) + \frac{1}{2}\varrho\|\boldsymbol{k} - \boldsymbol{k}_0\|^2$. Note that there is also a possibility of adaptive choice of the regularization parameter, see e.g. [22]. Another natural possibility is to use the Bayesian inversion, which will be discussed in Section 4.

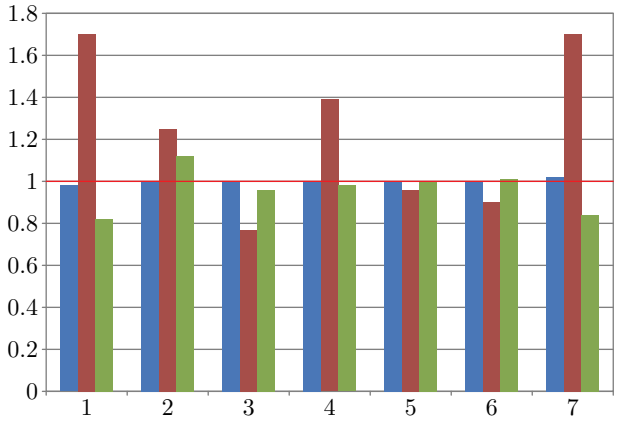


Figure 3.3. Ratio of identified parameters to the exact values from $\boldsymbol{k}_{\text{ref}}$. Blue (left) columns—no noise; red (middle) columns—noise with $p = 0.02$ and no regularization (after 78 iterations, 704 function evaluation); green (right) columns—noise with $p = 0.02$ and regularization with the parameter $\varrho = 20$ (after 42 iterations, 352 function evaluation).

4. Bayesian inversion

The case with noisy measurements can be investigated more thoroughly by the Bayesian inverse approach [20]. It assumes given statistical properties of the measurement error and does not attempt to get deterministic material characteristics. Instead, we assume that the material is in some level uncertain and attempt to obtain its statistical description, i.e. to estimate the joint probability distribution of

a vector representing the piece-wise constant material property. A thorough presentation on the topic of stochastic computations is e.g. [13]. The results presented in this article are direct continuation of the work presented in [8], [9].

The Bayesian inverse approach uses again the observation operator $G\colon \boldsymbol{\mathcal{U}}_{\rm ad} \to \mathbb{R}^{\#M}$, now expressed as a function of transformed parameters $\boldsymbol{k} = \ln(\boldsymbol{k}_{\rm exp})$, $G\colon \mathbb{R}^{\#S} \to \mathbb{R}^{\#M}$, and vector of the measurements $\boldsymbol{\mu} \in \mathbb{R}^{\#M}$ corrupted by noise $\boldsymbol{\eta} \in \mathbb{R}^{\#M}$. The aim is to characterize the vector of material parameters $\boldsymbol{k}$ for which

$$\boldsymbol{\mu} = G(\boldsymbol{k}) + \boldsymbol{\eta}.$$

The Bayesian inversion uses a prior statistical information given in the form of the joint probability density function (pdf) π_0 to preliminarily characterize material parameters $\boldsymbol{k}$, and statistical information about the noise $\boldsymbol{\eta}$ given by its joint pdf π_η. The aim is then to find the posterior pdf $\pi(\boldsymbol{k}|\boldsymbol{\mu})$ of the vector $\boldsymbol{k}$ of uncertain (random) material parameters.

According to the Bayes theorem, see e.g. [20], the posterior pdf can be characterized as

$$\pi(\boldsymbol{k}|\boldsymbol{\mu}) = \Theta^{-1}[\pi_\eta(\boldsymbol{\mu} - G(\boldsymbol{k}))\pi_0(\boldsymbol{k})], \tag{4.1}$$

where Θ is a normalizing factor

$$\Theta = \int_{\mathbb{R}^{\#S}} \pi_\eta(\boldsymbol{\mu} - G(\boldsymbol{k}))\pi_0(\boldsymbol{k})\,{\rm d}\boldsymbol{k}. \tag{4.2}$$

Therefore, the expression

$$\pi(\boldsymbol{k}|\boldsymbol{\mu}) \propto \pi_\eta(\boldsymbol{\mu} - G(\boldsymbol{k}))\pi_0(\boldsymbol{k}), \tag{4.3}$$

where $\propto$ denotes a proportionality, is used for specification of the posterior pdf. Note that $\pi_\eta(\boldsymbol{\mu} - G(\boldsymbol{k})) = f_\eta(\boldsymbol{\mu}|\boldsymbol{k})$ is called the data likelihood.

As we consider the transformed parameters $\boldsymbol{k} = \ln(\boldsymbol{k}_{\rm exp})$, it is natural to consider the normal distribution as a prior guess, $\boldsymbol{k} \sim \mathcal{N}(\boldsymbol{k}_0, \sigma_k^2 I_{\#S})$,

$$\pi_0(\boldsymbol{k}) = ((2\pi)^{\#S}\sigma_k^{2\#S})^{-1/2} \exp\Big(-\frac{(\boldsymbol{k} - \boldsymbol{k}_0)^{\rm T}(\boldsymbol{k} - \boldsymbol{k}_0)}{2\sigma_k^2}\Big), \tag{4.4}$$

and normal distribution for noise, $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{0},\, \sigma_\eta^2 I_{\#M})$, the data likelihood gets the form of

$$\pi_\eta(\boldsymbol{\mu} - G(\boldsymbol{k})) = ((2\pi)^{\#M}\sigma_\eta^{2\#M})^{-1/2} \exp\Big(-\frac{(\boldsymbol{\mu} - G(\boldsymbol{k}))^{\rm T}(\boldsymbol{\mu} - G(\boldsymbol{k}))}{2\sigma_\eta^2}\Big), \tag{4.5}$$

and consequently

$$\pi(\boldsymbol{k}|\boldsymbol{\mu}) \propto \exp\Big(-\frac{\|\boldsymbol{\mu} - G(\boldsymbol{k})\|^2}{2\sigma_\eta^2} - \frac{\|\boldsymbol{k} - \boldsymbol{k}_0\|^2}{2\sigma_k^2}\Big). \tag{4.6}$$

Therefore, the minimization of observation least squares with regularization used in Section 3 corresponds to seeking the parameter $\boldsymbol{k}$, where the posterior pdf $\pi(\boldsymbol{k}|\boldsymbol{\mu})$ is maximal (mode). Formula (4.6) provides some idea about the weights for regularization. Note that if G is linear and both the measurements and prior distributions are Gaussian, then the posterior distribution is also Gaussian. But this is not the case of our application.

For the numerical realization, the operator G is fully replaced by an inexact model G_h, i.e. a model which primarily involves the discretization of the state problem by the finite element method. (Besides the discretization error possibly also some other numerical errors from numerical implementation can be present.) We could also consider these errors as a numerical noise. It means that our analysis aims at the posterior distribution

$$\pi_h(\boldsymbol{k}|\boldsymbol{\mu}) \propto \pi_\eta(\boldsymbol{\mu} - G_h(\boldsymbol{k}))\pi_0(\boldsymbol{k}).$$

5. Implementation of the Bayesian inverse analysis

In standard cases, we cannot evaluate the posterior distribution $\pi_h(\boldsymbol{k}|\boldsymbol{\mu})$ analytically, but we can generate samples from it using the Metropolis-Hastings (MH) algorithm [17]. The MH algorithm is based on proposing samples from a pre-defined instrumental density q and accepting them with some computed probability.

Algorithm 5.1 (MH algorithm).

▷ Choose an initial guess $\boldsymbol{k}^{(0)}$.

▷ For $t = 1, 2, \ldots, T$:

(1) Generate $\boldsymbol{\kappa}$ from $q(\boldsymbol{\kappa}|\boldsymbol{k}^{(t-1)})$.

(2) Compute the acceptance probability

$$\alpha_t = \min\Big\{1, \frac{\pi_h(\boldsymbol{\kappa}|\boldsymbol{\mu})q(\boldsymbol{k}^{(t-1)}|\boldsymbol{\kappa})}{\pi_h(\boldsymbol{k}^{(t-1)}|\boldsymbol{\mu})q(\boldsymbol{\kappa}|\boldsymbol{k}^{(t-1)})}\Big\}. \tag{5.1}$$

(3) (a) Set $\boldsymbol{k}^{(t)} = \boldsymbol{\kappa}$ with probability α_t, otherwise
(b) set $\boldsymbol{k}^{(t)} = \boldsymbol{k}^{(t-1)}$.

The *basic MH algorithm* can be seen in Algorithm 5.1. We choose a symmetric random-walk proposal distribution with a conditional pdf

$$q(\boldsymbol{\kappa}|\boldsymbol{k}) = ((2\pi)^{\#S}\sigma_{MH}^{2\#S})^{-1/2}\exp\Big(-\frac{(\boldsymbol{k}-\boldsymbol{\kappa})^{\mathrm{T}}(\boldsymbol{k}-\boldsymbol{\kappa})}{2\sigma_{MH}^2}\Big). \tag{5.2}$$

In this case (symmetry of q), the fraction in (5.1) can be reduced.

Theorem 5.2. *The basic MH algorithm described in Algorithm* 5.1 *with* (5.2) *generates a Markov chain with the limiting distribution* $\pi_h(\boldsymbol{k}|\boldsymbol{\mu})$ *despite of the choice of* $\sigma_{MH} > 0$.

For the proof of a more general form of this theorem see [18]. Note that the implementation of this algorithm requires the number of steps T to be estimated, which is usually done adaptively by watching some values as e.g. the average of generated $\boldsymbol{k}^{(t)}$ and stopping the iterations when this value stabilizes enough. When the Markov chain is created, its autocorrelation is analyzed. The autocorrelation time can also be estimated to determine the distance between almost uncorrelated samples.

Computational labour of the MH algorithm is concentrated in evaluation of the observation operator G, in practice of its discretized version G_h. Having an approximation $\widetilde{G}_h$ to G_h, we can use the modified MH algorithm with the following extended Step (2):

(2a) Compute the pre-acceptance probability

$$\tilde{\alpha}_t = \min\left\{1, \frac{\tilde{\pi}_h(\boldsymbol{\kappa}|\boldsymbol{\mu})q(\boldsymbol{k}^{(t-1)}|\boldsymbol{\kappa})}{\tilde{\pi}_h(\boldsymbol{k}^{(t-1)}|\boldsymbol{\mu})q(\boldsymbol{\kappa}|\boldsymbol{k}^{(t-1)})}\right\}.$$

(2b) Pre-accept $\boldsymbol{\kappa}$ with probability $\tilde{\alpha}_t$, otherwise go to Step (3b), i.e. $\boldsymbol{k}^{(t)} = \boldsymbol{k}^{(t-1)}$.

(2c) If the sample $\boldsymbol{\kappa}$ is pre-accepted, compute the modified acceptance probability

$$\alpha_t = \min\left\{1, \frac{\pi_h(\boldsymbol{\kappa}|\boldsymbol{\mu})\tilde{\pi}_h(\boldsymbol{k}^{(t-1)}|\boldsymbol{\mu})}{\pi_h(\boldsymbol{k}^{(t-1)}|\boldsymbol{\mu})\tilde{\pi}_h(\boldsymbol{\kappa}|\boldsymbol{\mu})}\right\}$$

and continue with Step (3).

Above, the computation of $\tilde{\alpha}_t$ with $\tilde{\pi}_h(\boldsymbol{\kappa}|\boldsymbol{\mu}) = \pi_\eta(\boldsymbol{\mu} - \widetilde{G}_h(\boldsymbol{k}))\pi_0(\boldsymbol{k})$ saves the computational labour-expensive operator G_h should be evaluated only for samples, which are more likely to be accepted. This technique is called the Delayed acceptance MH (DAMH) algorithm (see [7]). In comparison with the direct replacement of the operator G_h with its approximation $\widetilde{G}_h$, the DAMH algorithm preserves the limiting

distribution. Intuitively, this comes from the fact that the pre-acceptance using $\widetilde{G}_h$ can be understood as a different proposal density

$$q^*(\boldsymbol{\kappa}|\boldsymbol{k}) = g(\boldsymbol{k},\boldsymbol{\kappa})q(\boldsymbol{k}|\boldsymbol{\kappa}) + \left(1 - \int_{\mathbb{R}^{\#S}} g(\boldsymbol{k},\boldsymbol{\kappa})q(\boldsymbol{k}|\boldsymbol{\kappa})\,\mathrm{d}\boldsymbol{\kappa}\right)\delta_{\boldsymbol{k}}(\boldsymbol{\kappa}),$$

where

$$g(\boldsymbol{k},\boldsymbol{\kappa}) = \min\left\{1, \frac{\tilde{\pi}_h(\boldsymbol{\kappa}|\boldsymbol{\mu})q(\boldsymbol{k}|\boldsymbol{\kappa})}{\tilde{\pi}_h(\boldsymbol{k}|\boldsymbol{\mu})q(\boldsymbol{\kappa}|\boldsymbol{k})}\right\}$$

and $\delta_{\boldsymbol{k}}(\cdot)$ denotes the Dirac mass at $\boldsymbol{k}$. It can be shown that the MH algorithm with proposal density $q^*(\boldsymbol{\kappa}|\boldsymbol{k})$ has the same limiting distribution as with $q(\boldsymbol{\kappa}|\boldsymbol{k})$.

Theorem 5.3. *If $q(\boldsymbol{\kappa}|\boldsymbol{k}) > 0$ implies $\tilde{\pi}_h(\boldsymbol{\kappa}|\boldsymbol{\mu}) > 0$, the DAMH algorithm described above with* (5.2) *generates a Markov chain with the limiting distribution $\pi_h(\boldsymbol{k}|\boldsymbol{\mu})$ despite of the choice of $\sigma_{MH} > 0$.*

For the proof, see [7], [18]. Note that for (4.4) and (4.5), the assumption is fulfilled despite of the choice of $\widetilde{G}_h$.

There are many ways of constructing $\widetilde{G}_h$ including the FEM computation on a coarser grid, the stochastic collocation method and various methods for surrogate models construction. In our paper, the approximation of the FEM solution using the stochastic Galerkin method (SGM) is considered, see [2]. The SGM fitted to our problem is described in the next subsection.

5.1. Stochastic Galerkin method. The SGM denotes the Galerkin method applied to the PDE with uncertain parameters in the form of functions of a random vector, for a more thorough introduction see [13], [23], [2]. The SGM assumes a discretization of both the physical space (functions on the domain) and the stochastic/parametric space (functions of random variables). The SGM solution is then a function of both space variable and vector of random parameters, see (5.4).

Here, we aim to create a suitable approximation $\widetilde{G}_h$ to the discretized observation operator G_h for numerical realization of the DAMH algorithm. This means to create an approximate operator assigning integral fluxes over each Dirichlet part of the boundary in each of m tests to a parameterized material field. For this, we first solve each of m boundary value problems (tests) separately and then we obtain the fluxes for each Dirichlet window. Note that here we will use standard variational and FEM formulation instead of the mixed formulation.

Pointing out the stochastic nature, each of the boundary value problems (tests) can be described by the following equation:

$$\begin{cases} -\nabla_x \cdot (k(x;\boldsymbol{Z})\nabla_x u(x;\boldsymbol{Z})) = 0 & \forall\, x \in \Omega,\ \boldsymbol{Z} \in \mathbb{R}^{\#S}, \\ u(x;\boldsymbol{Z}) = g(x) & \forall\, x \in \Gamma^D,\ \boldsymbol{Z} \in \mathbb{R}^{\#S}, \\ n(x) \cdot k(x;\boldsymbol{Z})\nabla_x u(x;\boldsymbol{Z}) = 0 & \forall\, x \in \Gamma^N,\ \boldsymbol{Z} \in \mathbb{R}^{\#S}. \end{cases} \tag{5.3}$$

To complete the formulation, we need to specify some probability distribution of vector $\boldsymbol{Z} = (Z_1, \ldots, Z_{\#S})$. This should reflect the random nature of the material field from the previous sections-constant on given subdomains. Within the DAMH algorithm, $k(x;\boldsymbol{Z})$ is defined through the vector $\boldsymbol{k}$. It means that we work with a random material field, which can be expressed as

$$k(x;\boldsymbol{Z}) = \sum_{i=1}^{\#S} \chi_{\Omega_i}(x) \exp(a_i Z_i + b_i),$$

where $\chi_{\Omega_i}(x)$ is a characteristic function of the subdomain Ω_i, $k_i = a_i Z_i + b_i$ (for the convenience of the computation, parameters of the problem are normalized), b_i is the mean value of the material on subdomain Ω_i (see values of $\boldsymbol{k}_0$ in Section 6) and a_i is the standard deviation of the material on subdomain Ω_i (see σ_k in Section 6). Then the components of $\boldsymbol{Z}$ are independent standard normal random variables. Note that for all $\boldsymbol{Z} \in R^{\#S}$ the function $k(x;\boldsymbol{Z})$ is bounded and positive on Ω, therefore the deterministic counterpart of problem (5.3) has a unique solution for all $\boldsymbol{Z} \in \mathbb{R}^{\#S}$.

Problem (5.3) can be formulated variationally with bilinear form a in a tensor product of standard Sobolev space $H^1(\Omega)$ and $L^2_{\mathrm{d}F\boldsymbol{Z}}(\mathbb{R}^{\#S})$, which is a space of square integrable functions on $\mathbb{R}^{\#S}$ with respect to the measure/distribution of $\boldsymbol{Z}$, see e.g. [1], [16]. For the Galerkin discretization we shall use finite dimensional space with the basis created as a tensor product of basis $\langle \varphi_1(x), \ldots, \varphi_{N_d+N_{Dd}}(x) \rangle$ of standard linear elements on Ω and basis $\langle \psi_1(\boldsymbol{Z}), \ldots, \psi_{N_s}(\boldsymbol{Z}) \rangle$ of orthonormal polynomials (with respect to the distribution of $\boldsymbol{Z}$) on $\mathbb{R}^{\#S}$ (here it means the products of one dimensional Hermite polynomials). Note that we assume that the last basis functions $\varphi_{N_d+1}(x), \ldots, \varphi_{N_d+N_{Dd}}(x)$ correspond to the degrees of freedom on the Dirichlet part of the boundary $\partial\Omega$.

The Galerkin approximation of $u(x, Z)$ then takes the form

$$\tilde{u}(x, Z) = \sum_{i=1}^{N_d+N_{Dd}} \sum_{j=1}^{N_s} \overline{u}_{ij} \varphi_i(x) \psi_j(\boldsymbol{Z}). \tag{5.4}$$

Note that once we have computed $\tilde{u}$, the evaluation of the approximation of u for some given sample of $\boldsymbol{Z}$ consists only of a matrix-vector multiplication and an evaluation of polynomials $\langle \psi_1(\boldsymbol{Z}), \ldots, \psi_{N_s}(\boldsymbol{Z}) \rangle$, which is a fairly cheap operation. Due

to separable nature of the material field, the values of bilinear form on elements of tensor product basis can be written as

$$a(\varphi_i\psi_j, \varphi_k\psi_l) = \sum_{t=1}^{\#S} \int_{\mathbb{R}^{\#S}} \psi_j\psi_l \exp(a_t Z_t + b_t)\,\mathrm{d}F\boldsymbol{Z} \int_{\Omega_t} \nabla\varphi_i \nabla\varphi_k \,\mathrm{d}x.$$

Assuming $\tilde{u} = \tilde{u}_h + \tilde{u}_0$, where $\tilde{u}_h = \sum_{i=1}^{N_d}\sum_{j=1}^{N_s} (\overline{u}_h)_{ij}\varphi_i\psi_j$ is a homogeneous part of the solution and $\tilde{u}_0 = \sum_{i=N_d+1}^{N_{Dd}} (\overline{u}_0)_i\varphi_i$ is a particular part of the solution fulfilling Dirichlet boundary condition, we obtain a large system $(N_s \times N_d)$ of linear equations in the form of

$$\begin{aligned}
\mathbf{A}\cdot\overline{u}_h = \overline{b},\ \mathbf{A} &= \sum_{t=1}^{\#S} \mathbf{G}_t \otimes \mathbf{K}_t,\ \overline{b} = \sum_{t=1}^{\#S} \overline{g}_t \otimes \overline{f}_t,\\
\mathbf{A}_{ij,kl} &= a(\varphi_i\psi_j, \varphi_k\psi_l),\ \overline{b}_{ij} = a(\varphi_i\psi_j, \tilde{u}_0),\\
(\mathbf{G}_t)_{j,l} &= \int_{\mathbb{R}^{\#S}} \psi_j\psi_l \exp(a_t Z_t + b_t)\,\mathrm{d}F\boldsymbol{Z},\ (\mathbf{K}_t)_{i,k} = \int_{\Omega_t} \nabla\varphi_i\nabla\varphi_k\,\mathrm{d}x,\\
(\overline{g}_t)_j &= \int_{\mathbb{R}^{\#S}} \psi_j \exp(a_t Z_t + b_t)\,\mathrm{d}F\boldsymbol{Z},\ (\overline{f}_t)_i = \int_{\Omega_t} \nabla\varphi_i\nabla\tilde{u}_0\,\mathrm{d}x,\\
&\qquad\qquad i,k = 1,\ldots,N_d;\ j,l = 1,\ldots,N_s,
\end{aligned}$$

where $\mathbf{G}_t/\overline{g}_t$ are matrices/vectors corresponding to the discretization of parameter space and $\mathbf{K}_t/\overline{f}_t$ are matrices/vectors corresponding to the discretization using linear finite elements. The form of the matrix and the right-hand side allows us to greatly reduce the memory requirements and computation costs. The multiplication by the matrix $\mathbf{A}$ can be performed in the following way:

$$\mathbf{A}\cdot\overline{u}_h = \mathrm{vec}\left(\sum_{t=1}^{\#S} \mathbf{K}_t \boldsymbol{u}_h \mathbf{G}_t^{\mathrm{T}}\right),$$

where $\boldsymbol{u}_h$ denotes the vector $\overline{u}_h$ reshaped into an $N_d \times N_s$ matrix and $\mathrm{vec}(\cdot)$ reshapes the same sized matrix back into a vector.

This form of the matrix allows us to view the system as a matrix equation. We solve the system using the reduced basis method, which involves a creation of the reduced rational Krylov approximation of matrices $\mathbf{K}_t$, for more details see [15].

Next part is to obtain the values of fluxes over the parts Γ_j^D of the Dirichlet boundary Γ^D. This can be done using the Green theorem and the fact, that we do not assume any volume sources and the Neumann boundary is homogeneous. The

flux over a Dirichlet window Γ_j^D can be calculated as

$$q_j(\boldsymbol{Z}) = \int_{\Gamma_j^D} k(x;\boldsymbol{Z}) \frac{\partial u(x;\boldsymbol{Z})}{\partial n(x)} \,\mathrm{d}s = \int_\Omega k(x;\boldsymbol{Z}) \nabla_x u(x;\boldsymbol{Z}) \nabla_x v_{D_j}(x) \,\mathrm{d}x,$$

where $v_{D_j}(x) \in H^1(\Omega)$ is an arbitrary function such as $v_{D_j}(x) = 1$ for all $x \in \Gamma_j^D$ and $v_{D_j}(x) = 0$ for all $x \in \Gamma^D \setminus \Gamma_j^D$. It is a generalization of the idea described e.g. in [6]. The approximation of $q_j(\boldsymbol{Z})$ can be obtained from the SGM solution $\tilde{u}$ in the following way:

$$\tilde{q}_j(\boldsymbol{Z}) = \overline{v}_{D_j}^{\mathrm{T}} \left[\sum_{t=1}^{\#S} \exp(a_t Z_t + b_t) \widehat{\mathbf{K}}_t \boldsymbol{u} \right] \overline{\psi}(\boldsymbol{Z}),$$

where $\overline{v}_{D_j}$ stands for a vector of length $N_d + N_{Dd}$ containing 0/1 denoting if a basis function $\varphi_i(x)$ is nonzero on Γ_j^D, $\widehat{\mathbf{K}}_t$ (a square matrix of size $N_d + N_{Dd}$) denotes an augmented form of the matrix $\mathbf{K}_t$, $\boldsymbol{u}$ denotes the solution $\overline{u}$ reshaped into a matrix of size $(N_d + N_{Dd}) \times N_s$ and $\overline{\psi}(\boldsymbol{Z})$ vector of basis functions $\langle \psi_1(\boldsymbol{Z}), \ldots, \psi_{N_s}(\boldsymbol{Z}) \rangle$. Note that the evaluation of fluxes over multiple Dirichlet parts of the boundary can be done cheaply at once (e.g. the value $\overline{\psi}(\boldsymbol{Z})$ is the same for each part and for each test/boundary value problem). Moreover, $\overline{v}_{D_j}^{\mathrm{T}} \widehat{\mathbf{K}}_t \boldsymbol{u}$ can be calculated only once for each Dirichlet part and each material parameter, therefore we store and work with reasonably small vectors ($\#S \times \#M$ vectors of size N_s).

6. Numerical experiments

Let us consider the model problem described in Section 3. First we illustrate the principle of the Bayesian inversion by studying the impact of the number of measurements to the posterior distribution. Further aim of the numerical experiments is to analyze sampling using the DAMH method and compare it with the basic MH algorithm.

6.1. Impact of the number of measurements. The following numerical experiments were performed three times: first by using only the measurements from test 1, then from tests 1 and 2 and finally using all 10 measurements from tests 1, 2, and 3. The measurements correspond to the noise distribution from Section 3 with $p = 0.02$. The prior pdf is defined by (4.4), where $\boldsymbol{k}_0 = [9, 5, 9, 1, 5, 1, 5]$ and $\sigma_k = 0.5$. This prior information expresses that the material is highly permeable in subdomains Ω_1 and Ω_3, medium permeable in subdomains Ω_2, Ω_5, and Ω_7 and hardly permeable in remaining subdomains Ω_4 and Ω_6. Additionally, this prior information corresponds to the initial guess of optimization methods used in Section 3.

subdomain	1	2	3	4	5	6	7
original values $\boldsymbol{k}_{\text{ref}}$	9.105	4.605	8.517	1.386	5.703	0.693	5.298
posterior mean	8.983	4.657	8.555	1.350	5.694	0.710	4.974
posterior std	0.509	0.106	0.085	0.178	0.025	0.067	0.493
$G_h(\boldsymbol{k}^*)$ sensitivity	1.050	32.30	121.8	9.805	334.9	7.327	0.172

Table 1. An overview of the parameters and the results obtained using all measurements (7 parameters).

Table 1 provides the results of the Bayesian inverse approach (approximated means and standard deviations). Here, the sensitivity of the material i is calculated as the approximation of

$$\sum_{j=1}^{\#M} \left| \left(\frac{\partial G_h(\boldsymbol{k}^*)}{\partial k_i} \right)_j \right|$$

(approximated using forward differences with a step equal to $0.01 \cdot \boldsymbol{k}^*$), where $\boldsymbol{k}^*$ denotes the least squares solution from Section 3. The marginal posterior distributions for the seven parameters of the original problem are approximately Gaussian, see Fig. 6.1. With the increasing number of measurements, the std decreases and the mean of the posterior distribution moves away from the mean of the prior distribution (in the case with the most parameters).

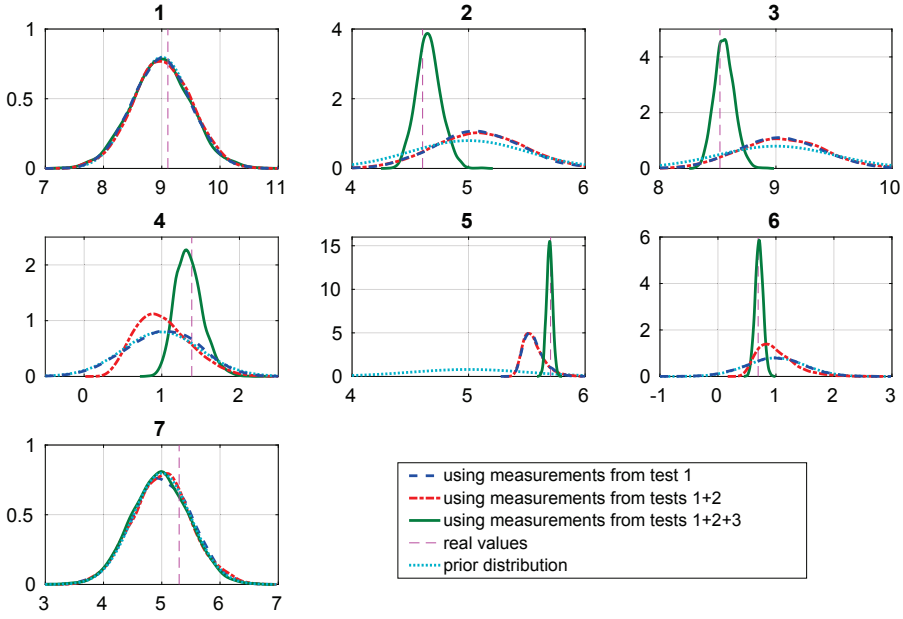


Figure 6.1. Marginal posterior distributions of the 7 parameters obtained using different numbers of measurements compared to the prior distribution.

To illustrate the Bayesian approach, a simplified problem with two parameters is also considered. It is assumed that the subdomains Ω_4 and Ω_6 form one subdomain with permeability exp(4.605) and the remaining subdomains form the second subdomain with permeability exp(5.704). The impact of the amount of measurements is shown in Fig. 6.2 (the red circle has a radius of one std of the prior distribution).

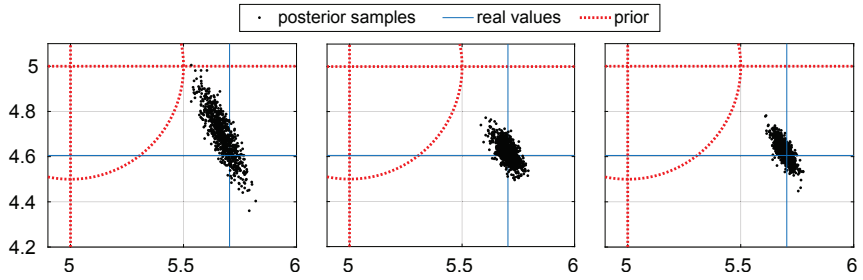


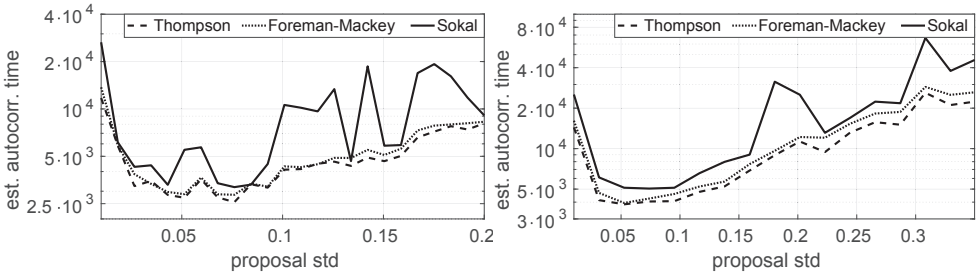
Figure 6.2. A simplified problem with 2 material parameters (logarithms of permeability). The points represent 400 samples from the posterior distribution obtained by the MH algorithm with the use of measurements from test 1 (left), 1+2 (middle), 1+2+3 (right).

6.2. Analysis of DAMH and MH. In the rest of the experiments, all 10 measurements were considered. Samples from the posterior distribution were provided using the DAMH algorithm (see Section 5) and for comparison also using the MH algorithm. The quality of the sampling process can be influenced by the choice of the proposal density. Here we choose (5.2) with various values of the proposal std σ_{MH}. The value of σ_{MH} controls the length of steps of the sampling algorithm and therefore, it influences the autocorrelation of the resulting Markov chain. Figure 6.3 shows the dependence of the autocorrelation time on the proposal std, both for the MH algorithm (left) and the DAMH algorithm (right). Generally, for a stable estimation of the autocorrelation time, long chains are required. We compared three different methods and in the rest of this section we use the estimation obtained by the method labeled as "Foreman-Mackey" in this figure. This method approximates the integrated autocorrelation time defined as

$$\tau = \sum_{i=-\infty}^{\infty} \varrho(i),$$

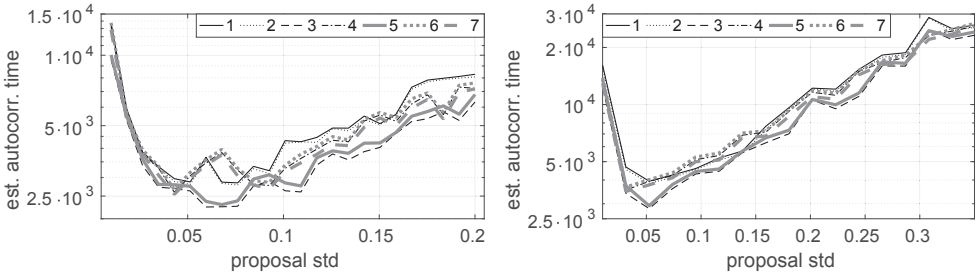
where $\varrho(i)$ is the normalized autocorrelation function of a studied stochastic process, by the value

$$1 + 2\sum_{i=1}^{M} \hat{\varrho}(i),$$



(a) MH algorithm (b) DAMH algorithm

Figure 6.3. Estimation of the autocorrelation time using several methods (see [21], [10], [19]) depending on the proposal std (for the 1st component of the Markov chain).



(a) MH algorithm (b) DAMH algorithm

Figure 6.4. Estimation of the autocorrelation time according to [10] depending on the proposal std (for all 7 components of the Markov chain).

where $\hat{\varrho}$ is the estimation of ϱ obtained from a finite chain $\{X^{(n)}\}_{n=1}^{N}$ and $M \ll N$. The comparison of these methods is performed on the first component of the Markov chain; however, the behavior is similar for all components (as illustrated in Fig. 6.4).

For the comparison of the efficiency of several DAMH based sampling processes, we use the quantity CpUS (Cost per Uncorrelated Sample) calculated as

$$\text{CpUS} = \left(\frac{\#[G_h \text{ evaluations}]}{[\text{chain length}]} + [\widetilde{G}_h \text{ evaluation cost}] \right) \cdot [\text{autocorr. time}], \tag{6.1}$$

where $[\widetilde{G}_h \text{ evaluation cost}]$ is the ratio between evaluation cost of $\widetilde{G}_h$ and of G_h. The quantity CpUS is meant to approximate the efficiency of the sampling procedure, it calculates the cost per one sample calculation relative to the cost of G_h as

$$\frac{\#[G_h \text{ evaluations}] + [\text{chain length}] \cdot [\widetilde{G}_h \text{ evaluation cost}]}{[\text{chain length}]}$$

and multiplies it by the auto-correlation time. Higher auto-correlation time means lower amount of information in the chain, we can view the auto-correlation time

as the number of chain steps to an almost "uncorrelated" sample. Therefore, this value expresses approximate computational cost per one almost uncorrelated sample in the units corresponding to the cost of one G_h evaluation. Figure 6.5 shows the dependence of the computational cost on the proposal std for several costs of $\widetilde{G}_h$ evaluations. For comparison, the efficiency of the MH algorithm is also shown in the same figure. Since the basic MH algorithm does not work with $\widetilde{G}_h$, (6.1) goes to the form CpUS = [autocorr. time].

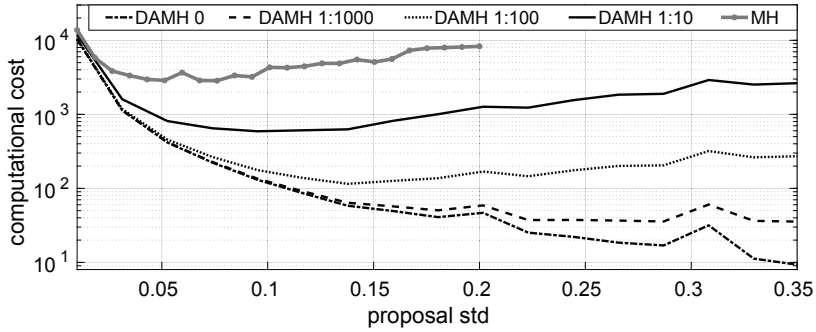


Figure 6.5. Efficiency of sampling processes depending on the proposal std (lower value is better)

First line of Table 2 shows additional properties of the MH sampling process for $\sigma_{\mathrm{MH}} = 0.0761$ that was chosen as optimal proposal std (according to previous experiments). Similarly, second line contains properties of the DAMH sampling process for $\sigma_{\mathrm{MH}} = 0.3509$. Average acceptance rates are calculated as the ratio between the number of accepted samples and the length of the chain. Generally, the efficiency of the DAMH algorithm is given by the quality of the approximation $\widetilde{G}_h$. In the DAMH implementation used in our experiments, the ratio between the $\widetilde{G}_h$ and G_h evaluation cost was approximately 1:1000 (evaluation cost of our surrogate model was approximately 1000 times lower than the evaluation cost of G_h). The efficiency was approximately 80 times higher compared to the basic MH algorithm.

method	σ_{MH}	chain length	average acceptance rate	number of G_h evaluations	estimated autocorr. time	CpUS
MH	0.0761	10^4	$4.80 \cdot 10^{-2}$	10^4	2848	2848
DAMH	0.3509	10^7	$2.55 \cdot 10^{-4}$	3516	26297	35.54

Table 2. Comparison of MH and DAMH sampling process.

For the construction of the SGM surrogate model we choose the finite element grid 500×500 and complete orthogonal polynomials up to the degree 4 (330 polynomials). The resulting SGM system of equations has approximately $82 \cdot 10^6$ unknowns.

Figure 6.6 shows the convergence of the reduced basis method (RB). We can see that for obtaining the relative residual (of the large SGM matrix) less or equal to 10^{-9} we need only 16 iterations of the RB method. The computational cost of one iteration of the RB method consists of $\#S$ solutions of finite element systems (same in each iteration, same finite element grid), orthogonalization of $\#S$ vectors and rank compression. Therefore, we can roughly compare the computational cost of the SGM solution to $16 \times \#S$ evaluations of G_h.

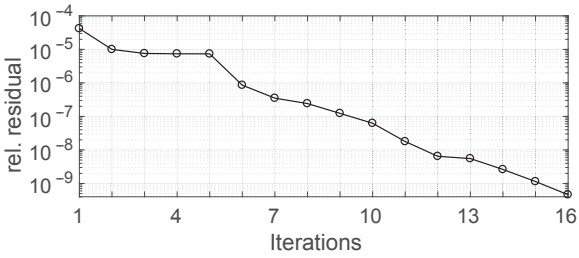


Figure 6.6. Convergence of reduced basis method

7. Summary and possible generalizations

In this paper, we described a specific type of inverse problems in micromechanics and outlined deterministic and stochastic approaches to their solution. We should stress the importance of enough data-multiple measurements within one test and multiple tests. The state equation can be efficiently solved by the standard or mixed finite element method with a special care for accurate computing of averaged boundary fluxes. Concerning comparison of deterministic and stochastic approaches, we could see that the deterministic approach is cheaper, but not robust without regularization. The proper setting of regularization is then a special problem, which should be overcome. Stochastic approach with the Bayesian inversion requires some additional statistical inputs-a guess of a prior distribution of the analyzed quantities and statistical characterization of measurement errors. It also requires to think how these inputs can be obtained in real situations. The benefits of the Bayesian inversion are robustness and more thorough analysis of the analyzed quantities (material parameters in our case). A disadvantage is in highly increased computational demands in the numerical realization of the Bayesian inversion. In this respect, a lot of effort has been recently devoted to the research on efficient methods for implementation of the Bayesian inversion with the construction of an approximate observation operator. In this paper, we used the delayed acceptance Metropolis-Hastings method with the stochastic Galerkin method for the construction of an approximate observation operator and showed the potential and benefits of this approach.

We skipped some points, like a suitable localization of the measurements on the boundary, which could enhance the sensitivity and robustness of the identification. We also skipped the generalization to other types of problems, e.g. anisotropy or elasticity considered in [4]. Applications of the considered type of inverse analysis to fractured porous media, multiphysics, as e.g. poroelasticity, and to nonlinear problems are planned in the future.

References

[1] *I. Babuška, R. Tempone, G. E. Zouraris*: Galerkin finite element approximations of Stochastic elliptic partial differential equations. SIAM J. Numer. Anal. *42* (2004), 800–825. zbl MR doi

[2] *M. Béreš, S. Domesová*: The stochastic Galerkin method for Darcy flow problem with log-normal random field coefficients. Adv. Electr. Electron. Eng. *15* (2017), 267–279. doi

[3] *R. Blaheta, M. Béreš, S. Domesová*: A study of stochastic FEM method for porous media flow problem. Proc. Int. Conf. Applied Mathematics in Engineering and Reliability. CRC Press, 2016, pp. 281–289. doi

[4] *R. Blaheta, R. Kohut, A. Kolcun, K. Souček, L. Staš, L. Vavro*: Digital image based numerical micromechanics of geocomposites with application to chemical grouting. Int. J. Rock Mechanics and Mining Sciences *77* (2015), 77–88. doi

[5] *D. Boffi, F. Brezzi, M. Fortin*: Mixed Finite Element Methods and Applications. Springer Series in Computational Mathematics 44, Springer, Berlin, 2013. zbl MR doi

[6] *G. F. Carey, S. S. Chow, M. K. Seager*: Approximate boundary-flux calculations. Comput. Methods Appl. Mech. Eng. *50* (1985), 107–120. zbl MR doi

[7] *J. A. Christen, C. Fox*: Markov chain Monte Carlo using an approximation. J. Comput. Graph. Statist. *14* (2005), 795–810. MR doi

[8] *S. Domesová, M. Béreš*: Inverse problem solution using Bayesian approach with application to Darcy flow material parameters estimation. Adv. Electr. Electron. Eng. *15* (2017), 258–266. doi

[9] *S. Domesová, M. Béreš*: A Bayesian approach to the identification problem with given material interfaces in the Darcy flow. Int. Conf. High Performance Computing in Science and Engineering, 2017 (T. Kozubek et al., eds.). Springer International Publishing, Cham, 2018, pp. 203–216. doi

[10] *D. Foreman-Mackey, D. W. Hogg, D. Lang, J. Goodman*: emcee: The MCMC hammer. Publ. Astron. Soc. Pacific *125* (2013), 306–312. doi

[11] *G. N. Gatica*: A Simple Introduction to the Mixed Finite Element Method. Theory and Applications. SpringerBriefs in Mathematics, Springer, Cham, 2014. zbl MR doi

[12] *J. Haslinger, R. Blaheta, R. Hrtus*: Identification problems with given material interfaces. J. Comput. Appl. Math. *310* (2017), 129–142. zbl MR doi

[13] *G. J. Lord, C. E. Powell, T. Shardlow*: An Introduction to Computational Stochastic PDEs. Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2014. zbl MR doi

[14] *Mathworks*: Matlab Optimization Toolbox User's Guide. Available at `https://uk.mathworks.com/products/optimization.html` (2017). sw

[15] *C. E. Powell, D. Silvester, V. Simoncini*: An efficient reduced basis solver for Stochastic Galerkin matrix equations. SIAM J. Sci. Comput. *39* (2017), A141–A163. zbl MR doi

[16] *I. Pultarová*: Hierarchical preconditioning for the stochastic Galerkin method: Upper bounds to the strengthened CBS constants. Comput. Math. Appl. *71* (2016), 949–964. MR doi

[17] *C. P. Robert*: The Bayesian Choice. From Decision-Theoretic Foundations to Computational Implementation. Springer Texts in Statistics, Springer, New York, 2007. zbl MR doi
[18] *C. P. Robert, G. Casella*: Monte Carlo Statistical Methods. Springer Texts in Statistics, Springer, New York, 2004. zbl MR doi
[19] *A. Sokal*: Monte Carlo methods in statistical mechanics: Foundations and new algorithms. Functional Integration: Basics and Applications, 1996 (C. DeWitt-Morette et al., eds.). NATO ASI Series. Series B. Physics. 361, Plenum Press, New York, 1997, pp. 131–192. zbl MR doi
[20] *A. M. Stuart*: Inverse problems: A Bayesian perspective. Acta Numerica *19* (2010), 451–559. zbl MR doi
[21] *M. B. Thompson*: A comparison of methods for computing autocorrelation time. Available at `https://arxiv.org/abs/1011.0175` (2010).
[22] *C. R. Vogel*: Computational Methods for Inverse Problems. Frontiers in Applied Mathematics 23, Society for Industrial and Applied Mathematics, Philadelphia, 2002. zbl MR doi
[23] *D. Xiu*: Numerical Methods for Stochastic Computations. A Spectral Method Approach. Princeton University Press, Princeton, 2010. zbl MR doi

Authors' addresses: *Radim Blaheta*, Institute of Geonics of the CAS, Studentská 1768, 708 00 Ostrava, Czech Republic, e-mail: `radim.blaheta@ugn.cas.cz`; *Michal Béreš*, *Simona Domesová*, Institute of Geonics of the CAS, Studentská 1768, 708 00 Ostrava, Czech Republic; Department of Applied Mathematics, Faculty of Electrical Engineering and Computer Science, VŠB – Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava, Czech Republic; IT4Innovations National Supercomputing Center, VŠB – Technical University of Ostrava, Studentská 6231/1B, 708 33 Ostrava, Czech Republic, e-mail: `michal.beres@vsb.cz`, `simona.domesova@vsb.cz`; *Pengzhi Pan*, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Xiaohongshan, Wuchang, Wuhan 430071, China, e-mail: `pzpan@whrsm.ac.cn`.