CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Civil Engineering

132P03C Project 3C

Inverse Analysis of a Stochastic PDE Using Bayes Method

Author:
Bc. Liya Gaynutdinova

Project supervisors:
Ing. Jan Havelka, Ph.D.,
Ing. Ondřej Rokoš, Ph.D.,
doc. RNDr. Ivana Pultarová, Ph.D.,
prof. Ing. Jan Zeman, Ph.D.

Prague, WS 2019/2020

1 Introduction

The steady increase in computing power over the years allowed for substituting physical experiments that are often demanding in terms of time, manpower and funding with numerical computer simulations. One of the objectives of numerical analysis is to predict outputs of complex engineering systems, especially those burdened with a high amount of input data uncertainty, such as material parameters and boundary conditions. Using numerical models has an advantage of controlled environments with at least a priory knowledge of their inevitable numerical errors.
In this project, an inverse problem for parameter identification of an elliptic partial differential equation in the form of the steady-state heat equation is considered. The parameters are involved in the data of the problem, more precisely in the random fields that represent the thermal conductivity λ of the material and the Dirichlet boundary conditions. The parameters are to be identified based on temperature distributions resulting from the finite element method solution of the problem. The Bayesian approach is used by applying the Metropolis-Hastings algorithm to approximate the parameter's distribution.

2 The Physical Model

We assume the steady state heat equation

$$-\nabla \cdot (a(\boldsymbol{x})\nabla u(\boldsymbol{x})) = f(\boldsymbol{x}), \tag{1}$$ {eq1}

with the Dirichlet boundary conditions

$$u(\boldsymbol{x}) = b(\boldsymbol{x}) \quad \text{for } \boldsymbol{x} \in \Gamma. \tag{2}$$ {eq2}

The function $a(\boldsymbol{x})$ represents the material's conductivity and is a random field over the domain Ω, where $\Omega \subset \mathbb{R}^2$ is a rectangular domain. Conversely, $b(\boldsymbol{x})$ is the random field on the boundary Γ of Ω. The heat source function $f(\boldsymbol{x}) \in \Omega$ is constant, and the function $a(\boldsymbol{x})$ is uniformly positive definite in Ω.

Equation (1) is transformed into the weak form and discretized using the finite element method with N continuous piece-wise linear basis functions [4]. We denote

the resulting discretized system by

$$\mathsf{A}_\Omega \mathsf{u}_\Omega = \mathsf{f}_s - \mathsf{A}_\Gamma u_\Gamma, \tag{3}$$ {Auf}

where, in the case of a square domain Ω with the nodes on a regular $N \times N$ mesh, $\mathsf{A}_\Omega(\boldsymbol{\xi})$ is a $(\sqrt{N}-2)^2 \times (\sqrt{N}-2)^2$ symmetric, positive definite matrix of internal nodes, u_Ω is the approximate solution. The right-hand side vector consists of the heat source vector f_s and a boundary condition vector that is a product of A_Γ, a $(\sqrt{N}-2)^2 \times 4(\sqrt{N}-1)$ matrix, and boundary values u_Γ given by the Dirichlet condition function $b(\boldsymbol{x})$.

3 The Parametrization of the Model

We assume two random fields $a(\boldsymbol{x})$ and $b(\boldsymbol{x})$ which are stochastic processes in space. We state the parametrization problem as follows [5]: let $(Y_t,\ t \in T)$ be a stochastic process of the random inputs, where the index t belongs to an index set T, and find a suitable transformation function R such that $Y_t = R(Z)$, where $Z = (Z_1, \ldots, Z_d),\ d \geq 1$, are mutually independent. As the index set T is infinite-dimensional domain and d is a finite integer, the transformation is an approximate. To consider the finite-dimensional version of Y_t we first discretize the index domain T into a set of finite indices and then study the process

$$(Y_{t_1}, \ldots, Y_{t_n}),\ t_1, \ldots, t_n \in T,$$

which is now a finite-dimensional random vector. To reduce the dimensionality of the random processes we introduce two parameters into the model by applying the truncated Karhunen-Loeve (KL) expansion. Let $\boldsymbol{\mu}_Y(t)$ be the mean of the input process $Y_t(\boldsymbol{\omega}) = (a(\boldsymbol{x}), b(\boldsymbol{x})) \in \Omega$ and let $C(t,s) = \mathrm{cov}(Y_t, Y_s)$ be its covariance function. We assume that the functions $a(\boldsymbol{x})andb(\boldsymbol{x})$ are mutually independent. The Karhunen-Loeve expansion of Y_t is given by:

$$Y_t(\omega) \approx \mu_Y(t) + \sum_{i=1}^{d} \sqrt{\lambda_i}\psi_i(t)Y_i(\omega), \quad d \geq 1, \tag{4}$$ {eq3}

where ψ_i are the orthogonal eigenfunctions and λ_i are the corresponding eigenvalues of the eigenvalue problem

$$\int_T C(t,s)\psi_i(s)ds = \lambda_i\psi_i(t),\ t \in T,$$

and $Y_i(\omega)$ are mutually uncorrelated random variables satisfying

$$\mathbb{E}[Y_i] = 0, \quad \mathbb{E}[Y_iY_j] = \delta_{ij}$$

where δ_{ij} is Kronecker's delta and defined by

$$Y_i(\omega) = \frac{1}{\sqrt{\lambda_i}}\int_T \left(Y_t(\omega) - \mu_Y(t)\right)\psi_i(t)dt, \quad i = 1,\ldots,d.$$

For our particular problem, we set the covariance functions as follows:

$$\begin{aligned}
&\mathrm{cov}_{\mathsf{A}}(\boldsymbol{x}_i,\boldsymbol{x}_j) = e^{\frac{-||(\boldsymbol{x}_i-\boldsymbol{x}_j)||_2^2}{l_{\mathsf{A}}}},\ \forall \boldsymbol{x} \in \Omega;\\
&\mathrm{cov}_{\mathsf{u}_\Gamma}(\boldsymbol{s}_i,\boldsymbol{s}_j) = e^{\frac{-|\sin\frac{\boldsymbol{s}_i-\boldsymbol{s}_j}{2}|}{l_{\mathsf{u}_\Gamma}}},\ S:\boldsymbol{x}\to\boldsymbol{s}\in[0,2\pi),\ \forall \boldsymbol{x}\in\Gamma
\end{aligned}$$

where $l_{\mathsf{A}}, l_{\mathsf{u}_\Gamma} \in \mathbb{R}$ are respective correlation lengths, S is a mapping from the rectangular boundary to a circle. We can write the input process A as:

$$\mathsf{A}(\boldsymbol{\xi}) = \boldsymbol{\xi}_1\mu_{\mathsf{A}} + \sum_{k=2}^{K}\boldsymbol{\xi}_k\sqrt{\lambda_k^{\mathsf{A}}}\Psi_k^{\mathsf{A}}, \tag{5}$$ {A}

where Ψ_k^{A} is the eigenvector $\psi_k^{\mathsf{A}} \in \mathbb{R}^{N^2}$ reshaped into a $N \times N$ matrix. Then from the resulting matrix A we can extract the matrices A_Ω and A_Γ:

$$\mathsf{Au} \equiv \left[\underbrace{\begin{pmatrix} a_1^1 & & \\ & \ddots & \\ & & a_{(\sqrt{N}-2)^2}^{(\sqrt{N}-2)^2} \end{pmatrix}}_{\mathsf{A}_\Omega} \underbrace{\begin{pmatrix} a_1^{(\sqrt{N}-2)^2+1} & & \\ & \ddots & \\ & & a_{(\sqrt{N}-2)^2}^{N} \end{pmatrix}}_{\mathsf{A}_\Gamma} \right] \left[\begin{array}{l} \mathsf{u}_\Omega \left\{ \begin{array}{l} \mathsf{u}_1 \\ \vdots \\ \mathsf{u}_{(\sqrt{N}-2)^2} \end{array} \right. \\ \mathsf{u}_\Gamma \left\{ \begin{array}{l} \mathsf{u}_{(\sqrt{N}-2)^2+1} \\ \vdots \\ \mathsf{u}_N \end{array} \right. \end{array} \right]$$

$$\begin{array}{ccc} \ddots & & \\ & \ddots & \\ & & a_N^N \end{array}$$

Similarly to (5), right-hand vector f can be written as:

$$\mathsf{f} = \mathsf{f_s} - \mathsf{A}_\Gamma \sum_{l=1}^{L} \theta_l \sqrt{\lambda_l^{\mathsf{u}_\Gamma}} \psi_l^{\mathsf{u}_\Gamma}, \tag{6}$$

Note that we normalized Ψ_1^{A} to a zero mean to eliminate linear dependency with μ_{A}, and that $\Psi_1^{\mathsf{u}_\Gamma}$ is constant, so we can treat it as the mean of the process.

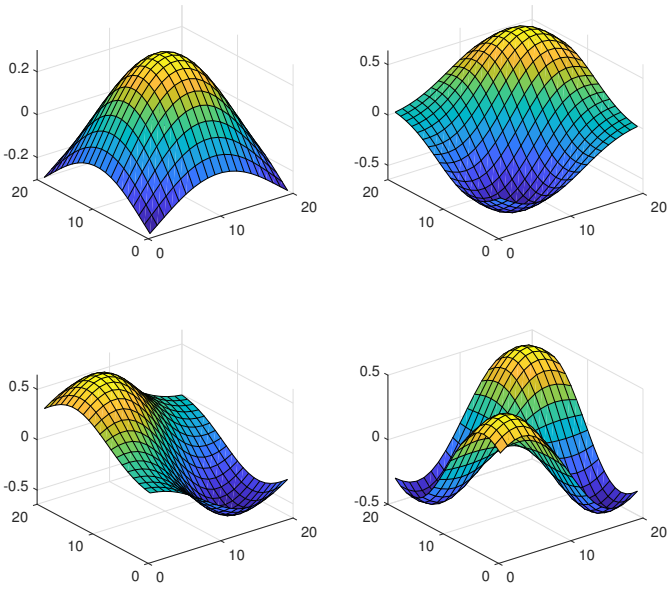

Figure 1: The first four eigenfunctions of cov_A.

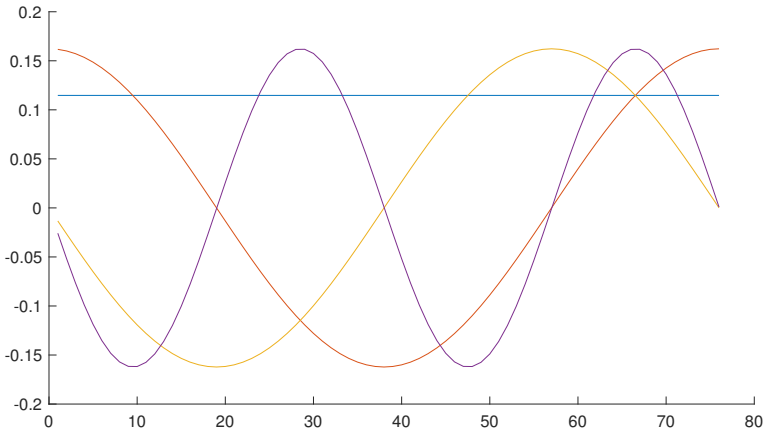

Figure 2: The first four eigenfunctions of cov_{u_Γ}.

Now the parameter identification problem can be written as a minimization problem:

$$\textit{find such } \boldsymbol{\rho}^* = \underset{\boldsymbol{\rho}\in\mathbb{R}^M}{\arg\min} \, ||\mathsf{u}_e + \eta - \mathsf{u}(\boldsymbol{\rho}^*)||_{L^2} \tag{7}$$ {P}

where $\mathsf{u}_e = \mathsf{u}_{e,\Omega} \cup \mathsf{u}_{e,\Gamma}$ is a given approximate solution based upon which the probability distributions of the parameter $\boldsymbol{\rho}$, $m = 1, \ldots, M$ must be identified. To represent discretization and measurement errors, we suppose η to be a random field of N_b independent elements, each normally distributed with zero mean and variance σ_η.

4 Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm [3] (MHA) can be used to identify the probability distribution of $\boldsymbol{\rho}$ based on measurements (7). Let us denote the prior distribution (our original guess about the distributions) of the parameters ρ_m as π_m, $m = 1, \ldots, K + L$. In our numerical experiments, we will choose the priors as standard normal distributions.

In MHA, we start with an initial sample $\boldsymbol{\rho}^i$ and set it as the current state $\widehat{\boldsymbol{\rho}}^i = \boldsymbol{\rho}^i$. Then each new proposal $\boldsymbol{\rho}^{i+1}$ is generated based on the proposal distribution q which is symmetric in the sense that $q(\boldsymbol{\rho}^i|\boldsymbol{\rho}^{i+1}) = q(\boldsymbol{\rho}^{i+1}|\boldsymbol{\rho}^i)$ for all $(\boldsymbol{\rho}^1, \boldsymbol{\rho}^2) \in \mathbb{R}^{M=K+L}$.

Usually q is Gaussian. In our numerical experiments, we will assume that its mean is zero and its variance is known and equal to σ_q. After evaluating $\mathsf{Gu}(\widehat{\boldsymbol{\rho}}^1)$ and $\mathsf{Gu}(\boldsymbol{\rho}^2)$, we compute the likelihoods (probabilities of observing) of g for given parameter vectors $\widehat{\boldsymbol{\rho}}^i$ and $\boldsymbol{\rho}^{i+1}$ [1]:

$$\begin{aligned} L(\mathsf{g}|\widehat{\boldsymbol{\rho}}^1) &\propto (\sigma_\eta^2)^{-N_b/2} \exp\left(-\frac{1}{2\sigma_\eta^2}\|\mathsf{Gu}(\widehat{\boldsymbol{\rho}}^{\mathrm{i}}) - \mathsf{g}\|_2^2\right) \\ L(\mathsf{g}|\boldsymbol{\rho}^2) &\propto (\sigma_\eta^2)^{-N_b/2} \exp\left(-\frac{1}{2\sigma_\eta^2}\|\mathsf{Gu}(\boldsymbol{\rho}^{\mathrm{i+1}}) - \mathsf{g}\|_2^2\right). \end{aligned}$$

Now we can compute the approximate probabilities

$$\begin{aligned} \pi(\widehat{\boldsymbol{\rho}}^i|\mathsf{g}) &= C\, L(\mathsf{g}|\widehat{\boldsymbol{\rho}}^i)\pi_1(\widehat{\rho}_1^i)\cdots\pi_M(\widehat{\rho}_M^i) \\ \pi(\boldsymbol{\rho}^{i+1}|\mathsf{g}) &= C\, L(\mathsf{g}|\boldsymbol{\rho}^{i+1})\pi_1(\rho_1^{i+1})\cdots\pi_M(\rho_M^{i+1}), \end{aligned}$$

where C represents an unknown normalization constant. In MHA, the newly proposed state $\boldsymbol{\rho}^{i+1}$ is then accepted with the probability of $\min(1, p)$, where

$$p = \frac{\pi(\boldsymbol{\rho}^{i+1}|\mathsf{g})}{\pi(\widehat{\boldsymbol{\rho}}^i|\mathsf{g})}.$$

This is usually implemented by generating a uniformly randomly distributed $\alpha \in [0, 1]$. If $\alpha < p$ then accept $\boldsymbol{\rho}^{i+1}$ and set $\widehat{\boldsymbol{\rho}}^{i+1} = \boldsymbol{\rho}^{i+1}$, otherwise, reject $\boldsymbol{\rho}^{i+1}$ and set $\widehat{\boldsymbol{\rho}}^{i+1} = \widehat{\boldsymbol{\rho}}^i$. Continuing in this way, we obtain a sequence $\widehat{\boldsymbol{\rho}}^j$, $j = 1, 2, \ldots, N_\rho$. After discarding the first I_0 elements (burn in), we can finally obtain approximation of marginal or joint posterior distributions of parameters ρ_m as distributions of $\widehat{\rho}_m^j$, $j = N_0 + 1, 2, \ldots, N_\rho$, for $m = 1, \ldots, M = K + L$.

The Algorithm:

1. Draw initial state $\boldsymbol{\rho}^1 \in \mathbb{R}^M$.
   Set $\widehat{\boldsymbol{\rho}}^1 = \boldsymbol{\rho}^1$.

2. For $i = 2, 3, \ldots, I$ do:

(a) draw proposal $\boldsymbol{\rho}^{i+1} \sim \mathcal{N}(\widehat{\boldsymbol{\rho}}^i, \sigma_q^2)$

(b) set $p = \pi(\boldsymbol{\rho}^{i+1}|\mathsf{g})/\pi(\widehat{\boldsymbol{\rho}}^i|\mathsf{g})$

(c) draw $a \sim \mathcal{U}(0,1)$

(d) if $a < p$ set $\widehat{\boldsymbol{\rho}}^{i+1} = \boldsymbol{\rho}^{i+1}$ (accept $\boldsymbol{\rho}^{i+1}$) else set $\widehat{\boldsymbol{\rho}}^{i+1} = \widehat{\boldsymbol{\rho}}^i$ (reject $\boldsymbol{\rho}^{i+1}$)

3. Discard $\widehat{\boldsymbol{\rho}}^1, \widehat{\boldsymbol{\rho}}^2, \ldots, \widehat{\boldsymbol{\rho}}^{I_0}$ (burn in).

4. Compute estimates of parameter characteristics from $\widehat{\boldsymbol{\rho}}^{I_0+1}, \ldots, \widehat{\boldsymbol{\rho}}^I$.

5 Experiments

There are several important input parameters of the considered problem that are influencing the behaviour of the MHA and therefore are subjected to fine-tuning, namely the variance σ_η that represents the measurement accuracy, the variance σ_q of the proposal distribution q and the intensity of the heat source $f(\boldsymbol{x})$. We try to set these parameters for a given constant heat source f, so that the overall acceptance rate of the samples is roughly between 20% and 30%, close to the recommended value of 23,4%. [6]. In a real-life version of the experiment, the accuracy of the measuring instrument (e.g. thermographic cameras) would be given relative to the measured value, so to reflect that we set σ_η in percentage based on the solution u^g. The variance σ_q determines the step size of the MHA walker, and it should be scaled reasonably according to the finite number of samples and the estimated distance between the starting point of the algorithm and the area of high probability.

Before applying the MHA to identify the full set of parameters $\boldsymbol{\rho}$, we first identify the material parameters $\boldsymbol{\xi}$ with fixed boundary parameters $\boldsymbol{\theta}$ and vice versa. For the sake of simplicity, we constrained the respective number of dimensions of $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ to 2.

We noticed that the heat source f greatly influences the shape of the joint posteriori distribution of the parameters. For the heat source $f(\boldsymbol{x}) = 0$, there seems to be a linear dependency for the components of $\boldsymbol{\xi}$ (Fig. 3) while $\boldsymbol{\theta}$ remains constant, and we are not able to accurately identify the high-probability region.

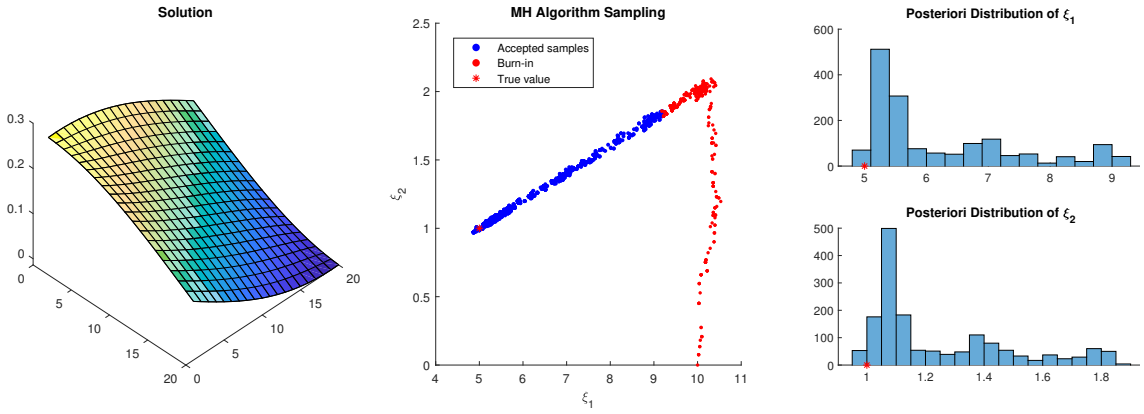

Figure 3: Identification of $\boldsymbol{\xi}$ for $f(\boldsymbol{x}) = 0$, measurement accuracy 0.1%.

For increasing values of f we are able to decrease the measurement accuracy σ_η while it seems we are better at establishing the mean of the conductivity field than the fluctuations parameter Ψ_2^A.

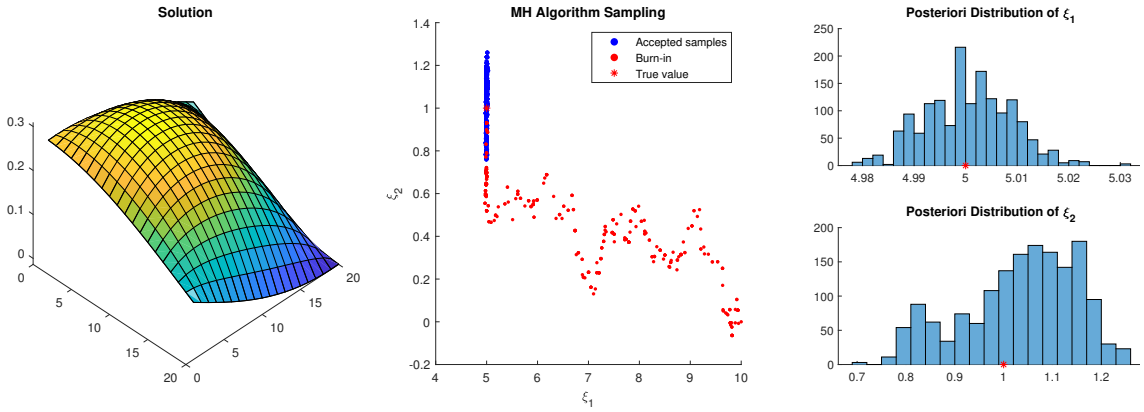

Figure 4: Identification of $\boldsymbol{\xi}$ for $f(\boldsymbol{x}) = 10$, measurement accuracy 1.33%.

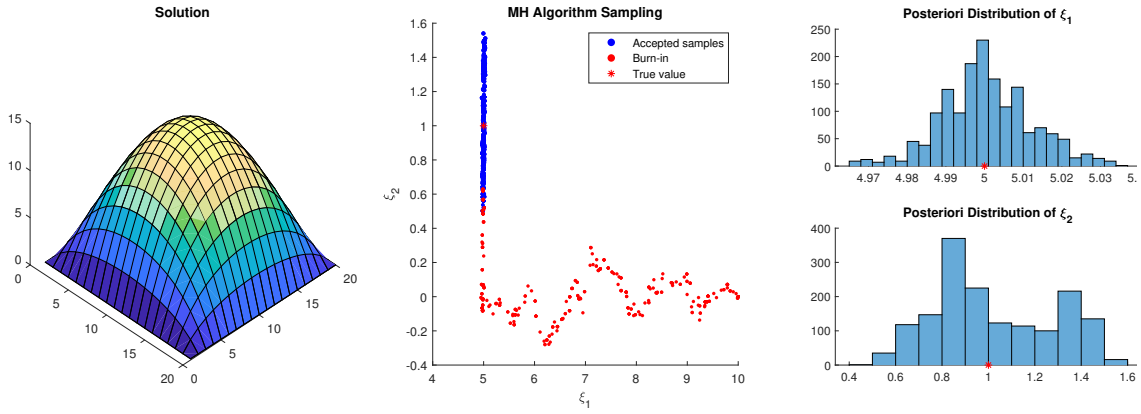

Figure 5: Identification of $\boldsymbol{\xi}$ for $f(\boldsymbol{x}) = 1000$, measurement accuracy 5%.

An opposite situation holds for the case of identifying boundary conditions $\boldsymbol{\theta}$ with fixed material parameters $\boldsymbol{\xi}$. Even with the heat source $f(\boldsymbol{x}) = 0$, we are able to determine the parameters of both modes (Fig. 6). Increasing the magnitude of the heat source improves the accuracy of the measurement if we went to keep the high-probability region from expanding (Fig. 7). Realistically, as modern thermovision cameras have an accuracy of about 2% [7], the end result from real-world experiment would look more like Fig. 8.

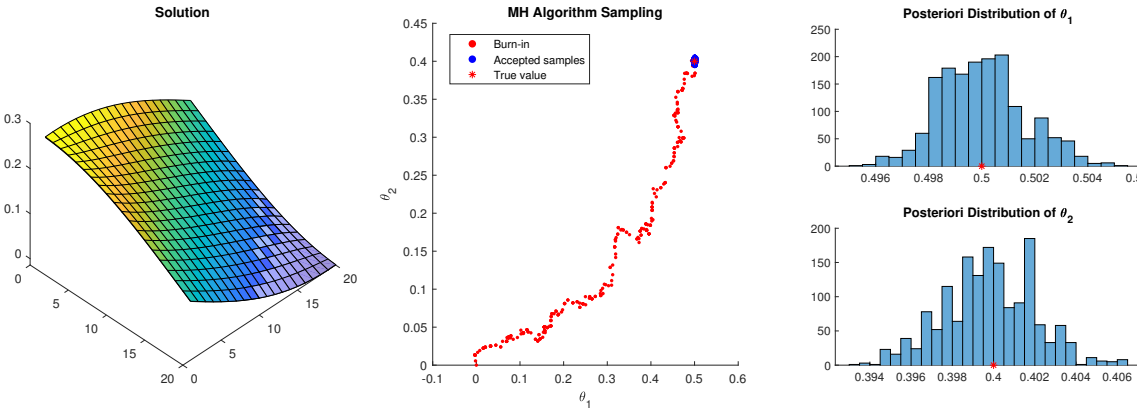

Figure 6: Identification of $\boldsymbol{\theta}$ for $f(\boldsymbol{x}) = 0$, measurement accuracy 6.67%.

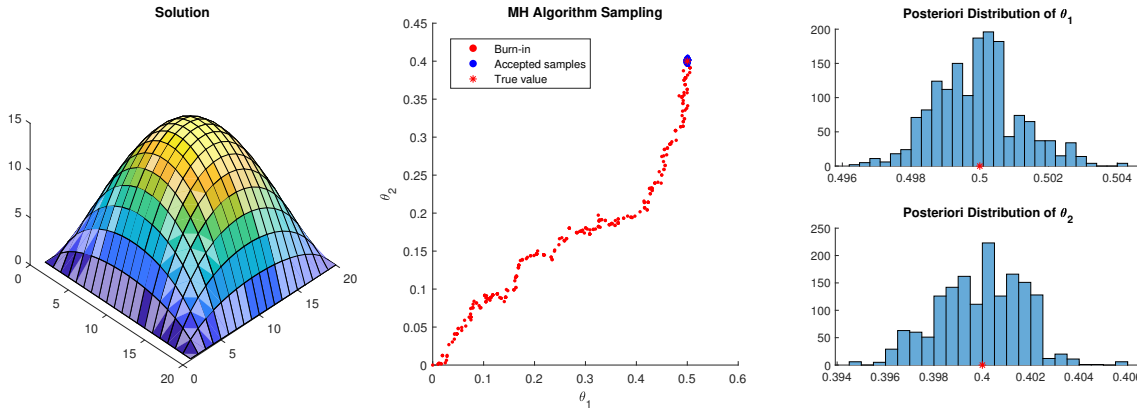

Figure 7: Identification of $\boldsymbol{\theta}$ for $f(\boldsymbol{x}) = 1000$, measurement accuracy 0.1%.

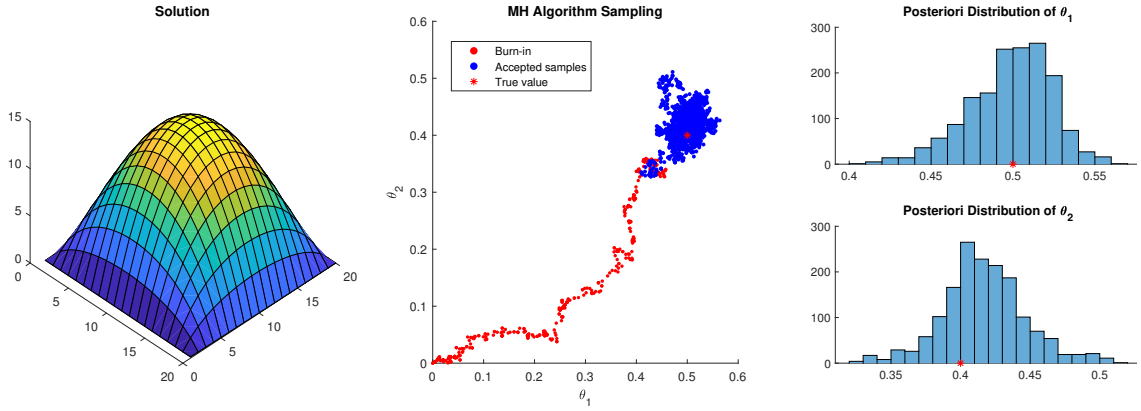

Figure 8: Identification of $\boldsymbol{\theta}$ for $f(\boldsymbol{x}) = 0$, measurement accuracy 2%.

Finally we applied the MHA to identify both parameters of the material $\boldsymbol{\xi}$ and boundary conditions $\boldsymbol{\theta}$, with the consideration for the effect the heat source has.

Naturally, with a heat source turned off , the material parameters cannot be identified with no negative effect on the high-probability region of the boundary condition parameters (Fig. 9), while the opposite is true of the other extreme example with $f(\boldsymbol{x}) = 1000$ (Fig. 11).

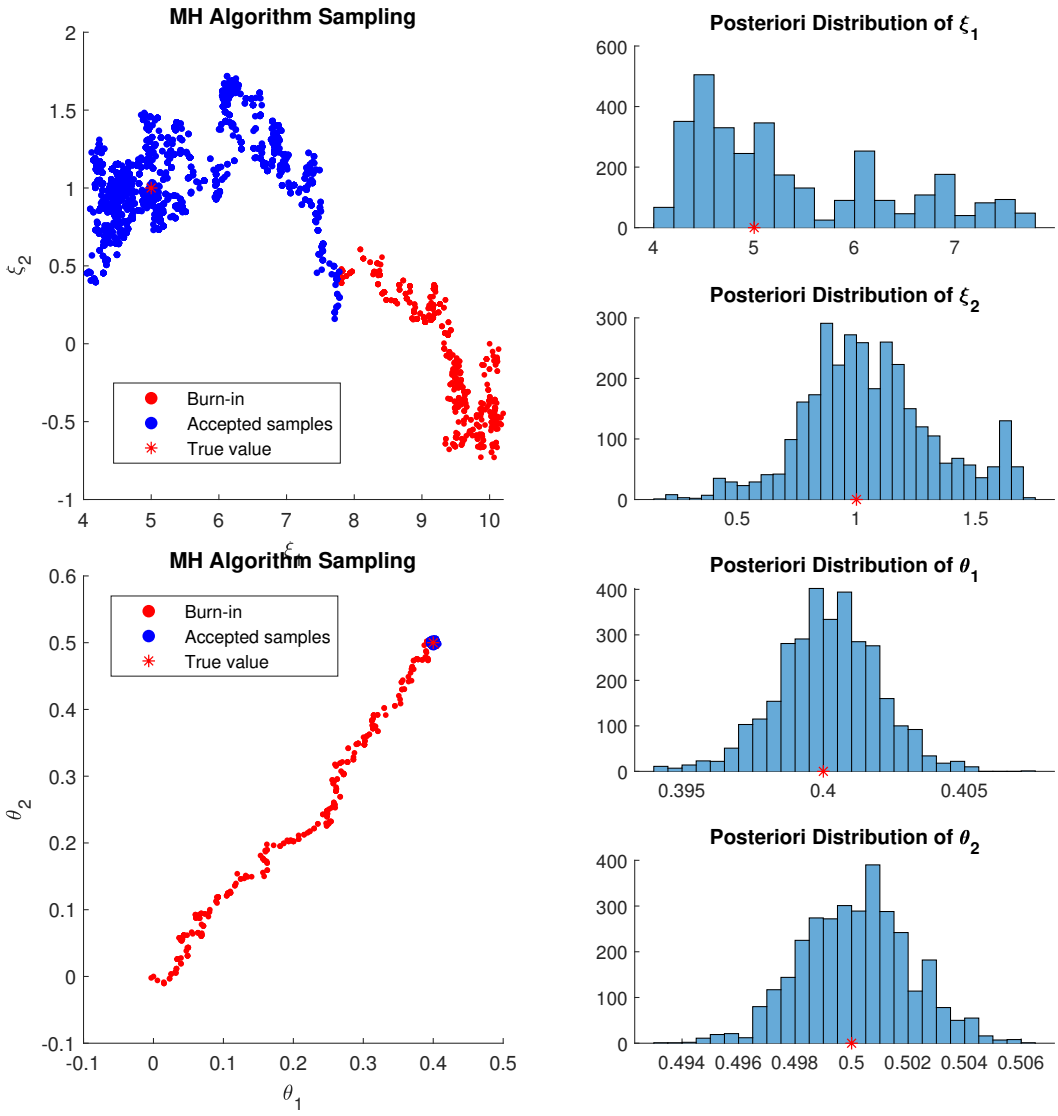

Figure 9: Identification of $\boldsymbol{\xi}$ *and* $\boldsymbol{\theta}$ *for* $f(\boldsymbol{x}) = 0$*, measurement accuracy 4%.*

There is a considerably wide range of values for the heat source for which we can establish the areas of interest for both of the parameters $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$, although it is worth noting that, again, increasing the heat source at a certain point immensely expands the area of interest for the parameter $\boldsymbol{\theta}$, while having little to no effect on the accuracy of the mean's guess and slightly decreasing the area of interest for the

second mode of $\boldsymbol{\xi}$ (Fig. 12).

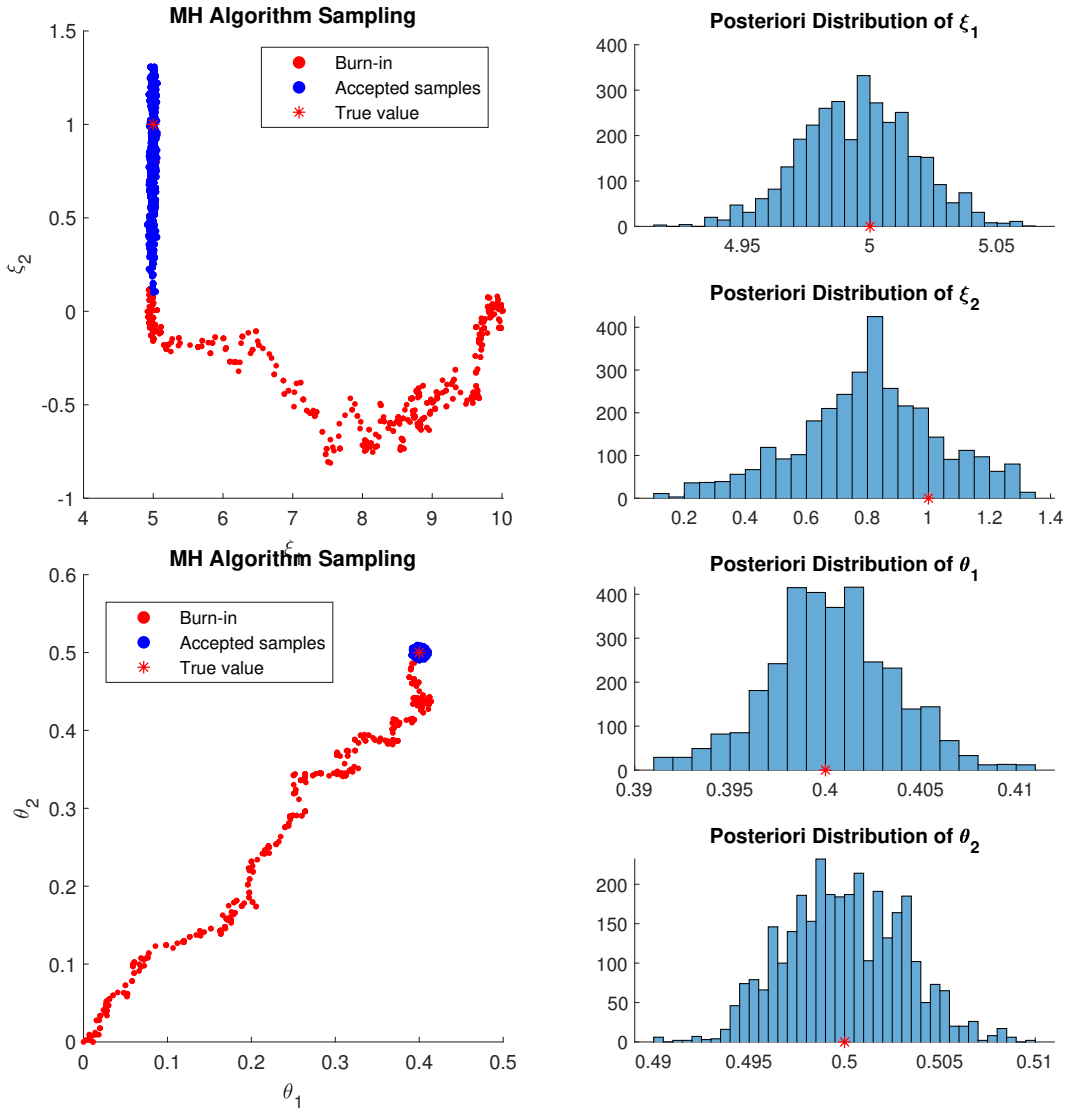

Figure 10: Identification of $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ for $f(\boldsymbol{x}) = 10$, measurement accuracy 4%.

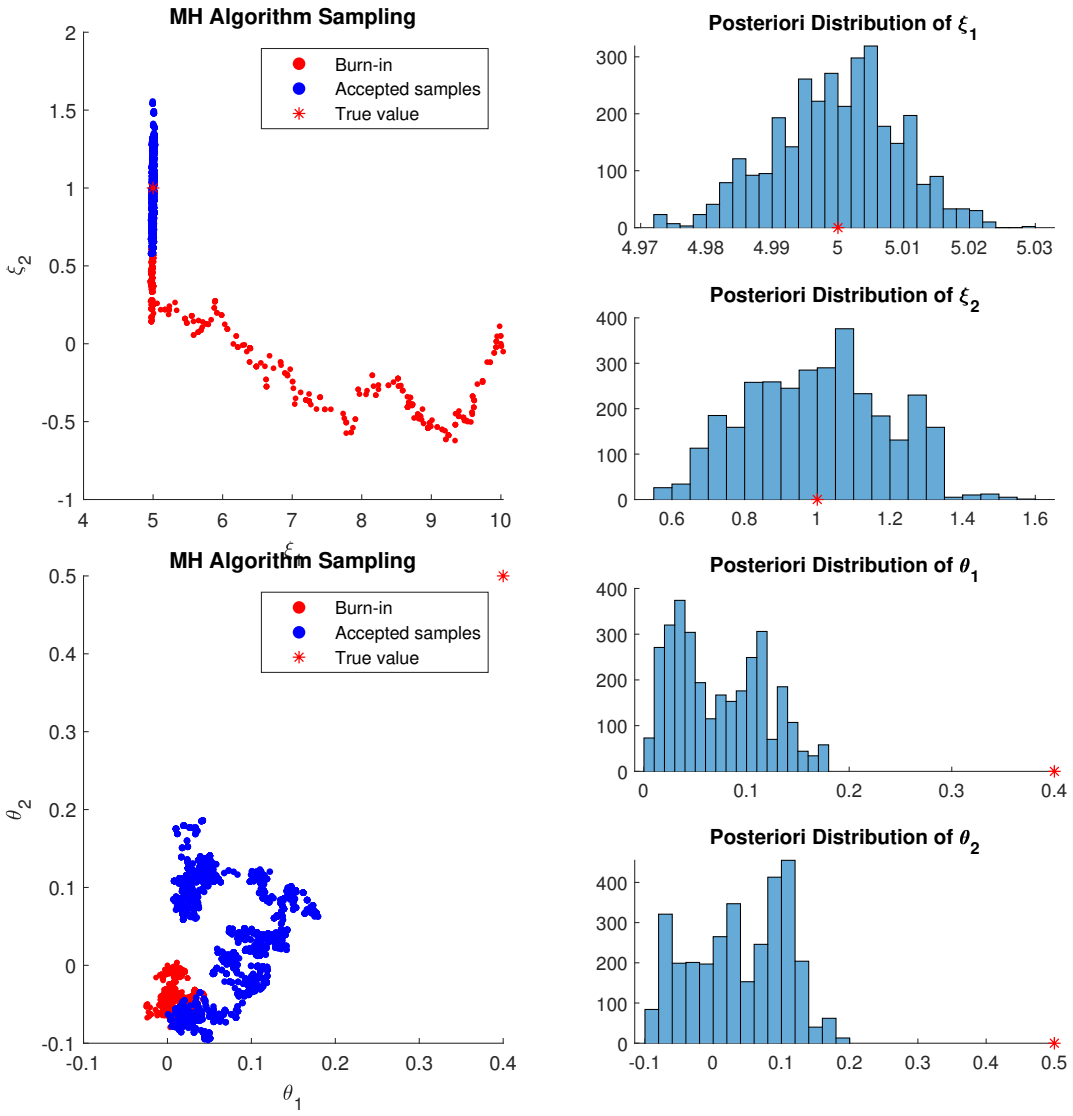

Figure 11: Identification of $\boldsymbol{\xi}$ *and* $\boldsymbol{\theta}$ *for* $f(\boldsymbol{x}) = 1000$*, measurement accuracy 4%.*

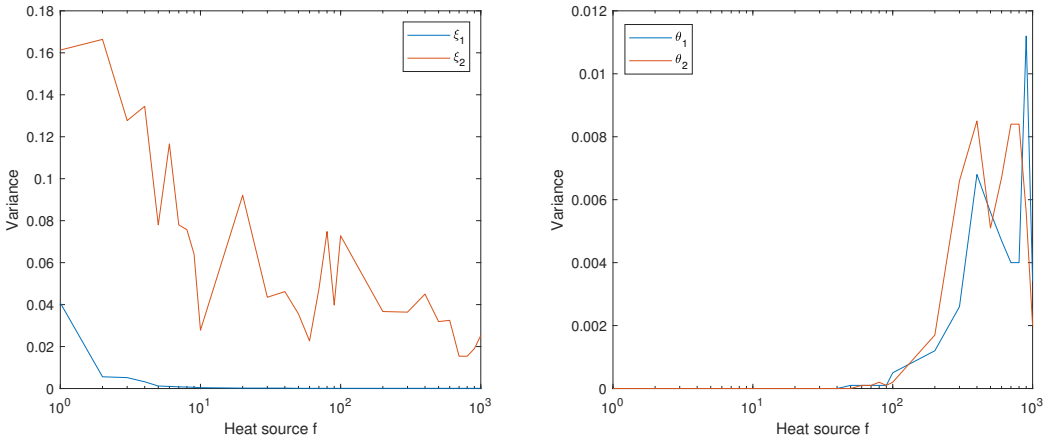

Figure 12: Parameters' variance depending on the intensity of the heat source

6 Conclusion

In this project we applied the Metropolis-Hastings algorithm to identify the distribution of parameters of the steady-state heat equation, such as the conductivity and boundary conditions. As the experiments have shown, there are other crucial input variables that we must consider to get satisfactory results, specifically there seems to be an optimal range for the magnitude of the heat source. In the future, we would like to increase the scope of this study to more complex problems with higher dimensionality with a potential application of surrogate models and identification of parameters from real-life experiments.

References

[1] R. Blaheta, M. Béreš, S. Domesová, P. Pan, A Comparison of Deterministic and Bayesian Inverse with Application in Micromechanics. Applications of Mathematics, 63 (6), 2018, pp. 665–686.

[2] T. Bui-Thanh, A Gentle Tutorial on Statistical Inversion using the Bayesian Paradigm, Institute for Computational Engineering and Sciences, The University of Texas at Austin. Published online.

[3] B. J. Brewer, STATS 331. Introduction to Bayesian Statistics, 2012. pp. 52–60. Published online.

[4] K. W. Morton, D. Mayers, Numerical Solution of Partial Differential Equations: An Introduction, Second Edition, Cambridge University Press, 2005.

[5] D. Xiu, Numerical Methods for Stochastic Computations: A Spectral Method Approach, Princeton University Press, 2010.

[6] G. O. Roberts, A. Gelman AND W. R. Gilks, Weak Convergence and Optimal Scaling Of Random Walk Metropolis Algorithms, 1997. The Annals of Applied Probability, Vol. 7, No. 1, pp. 110–120.

[7] Costello JT, McInerney CD, Bleakley CM, Selfe J, Donnelly AE (2012-02-01). The use of thermal imaging in assessing skin temperature following cryotherapy: a review" (PDF). Journal of Thermal Biology. 37 (2): 103–110.