Adaptive method for indirect identification of the statistical properties of random fields in a Bayesian framework

Guillaume Perrin, Christian Soize

Noname manuscript No.
(will be inserted by the editor)

Adaptive method for indirect identification of the statistical properties of random fields in a Bayesian framework

Guillaume Perrin · Christian Soize

Received: date / Accepted: date

Abstract This work considers the challenging problem of identifying the statistical properties of random fields from indirect observations. To this end, a Bayesian approach is introduced, whose key step is the nonparametric approximation of the likelihood function from limited information. When the likelihood function is based on the evaluation of an expensive computer code, this work also proposes a method to select iteratively new design points to reduce the uncertainties on the results that are due to the approximation of the likelihood. Two applications are finally presented to illustrate the efficiency of the proposed procedure: a first one based on analytic data, and a second one dealing with the identification of the random elasticity field of an heterogeneous microstructure.

Keywords Bayesian framework · uncertainty quantification · statistical inference · stochastic process · kernel density estimation

1 Introduction

Random field analysis has become a major tool in many scientific fields, such as uncertainty quantification, material science, biology, medicine, signal processing, quantitative finance, etc. However, in most of these applications, the knowledge of these random fields, which we write $\boldsymbol{X}$, is limited. Numerical methods are therefore needed to identify the probability distribution of $\boldsymbol{X}$ from the available information.

G. Perrin
CEA/DAM/DIF, F-91297, Arpajon, France
E-mail: guillaume.perrin2@cea.fr

C. Soize
Université Paris-Est, MSME UMR 8208 CNRS, Marne-la-Vallée, France
E-mail: christian.soize@u-pem.fr

When the information is constituted of direct measurements of the field to model, several techniques have been proposed to perform such an identification. For instance, the AutoRegressive-Moving-Average (ARMA) models [Whittle, 1951, Whittle, 1983, Box and Jenkins, 1970], allow the description of Gaussian stationary random fields as a parameterized integral of a Gaussian white noise. When considering a priori non-Gaussian and nonstationary random fields, the identification is generally based on a two-step procedure. The first step is the approximation of the random field by its projection on a reduced number of deterministic functions [Ghanem and Spanos, 2003, Le Maître and Knio, 2010], using for instance the proper orthogonal decomposition [Atwell and King, 2001], the proper generalized decomposition [Nouy, 2010], or the Karhunen-Loève expansion [Williams, 2011, Perrin et al., 2014, Perrin et al., 2013]. The second step is the identification of general stochastic representations of the projection coefficients in high stochastic dimension [Soize, 2010, Soize, 2011, Perrin et al., 2012, Nouy and Soize, 2014, Soize and Ghanem, 2016, Perrin et al., 2018].

The main specificity of this work comes from the fact that only indirect observations are available for the identification, in the sense that the experimental data is made of the transformations of a limited number of independent realizations of $\boldsymbol{X}$ through a black-box time-consuming nonlinear mapping, denoted by $\boldsymbol{g}$. To make this identification tractable, we assume that the random field to identify belongs to a known parametric class. Thus, identifying the distribution of $\boldsymbol{X}$ amounts to identifying the values of these parameters, which are gathered in the vector $\boldsymbol{z}$. A Bayesian framework is then considered [Marzouk and Najm, 2009, Stuart, 2010, Arnst et al., 2010, Matthies et al., 2016, Emery et al., 2016]: parameter $\boldsymbol{z}$ is supposed to be random, and we search its posterior distribution given the available data.

Markov Chain Monte Carlo (MCMC) [Rubinstein and Kroese, 2008, Tian et al., 2016] is generally considered as a powerful tool to explore the posterior distribution for these parameters. However it can be computationally prohibitive when each posterior evaluation requires evaluations of a computationally expensive code, as it the case here. To circumvent this problem, a standard approach is to replace the code by a surrogate model, and to directly sample from the approximated posterior distribution associated with the modified likelihood using classical MCMC procedures. The surrogate model can be based on polynomial representations [Marzouk and Najm, 2009, Marzouk and Xiu, 2009, Wan and Zabaras, 2011, Li and Marzouk, 2014, Tsilifis et al., 2017], Gaussian process regression [Kennedy and O'Hagan, 2001, Santner et al., 2003, Higdon et al., 2008, Bilionis and Zabaras, 2015, Sinsbeck and Nowak, 2017, Damblin et al., 2013], or runs of the code at different resolution levels [Higdon et al., 2003, Chen and Schwab, 2015]. Alternatively, the surrogate model can be used to adapt the proposal distribution. In that case, the number of expensive posterior evaluations per MCMC step can be strongly reduced, while sampling asymptotically from the exact posterior distribution (see [Rasmussen, 2003, Fielding et al., 2011, Conrad et al., 2016, Conrad et al., 2018] for further details about this approach).

This work can be seen as an extension of these methods to the case of stochastic codes. Indeed, for a given value of $\boldsymbol{z}$, as $\boldsymbol{X}(\boldsymbol{z})$ is random, $\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$

is also a random quantity. But if the distribution of $\boldsymbol{X}(\boldsymbol{z})$ is known once $\boldsymbol{z}$ is fixed, the distribution of $\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$ is unknown, and its identification is computationally demanding. Therefore, instead of constructing a surrogate model of the code, we focus on the approximation of the probability density function (PDF) of $\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$.

To run a MCMC procedure based on the associated approximated likelihood in a reasonable computational time, this approximation of the PDF of $\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$ in any $\boldsymbol{z}$ has to be constructed from a fixed number of already computed code evaluations. To this end, we first propose to directly work on the joint PDF of $(\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z})), \boldsymbol{z})$. Then, we focus on the Gaussian kernel density estimation (G-KDE) [Wand and Jones, 1995, Scott and Sain, 2004, Perrin et al., 2018] for the PDF approximation. Indeed, this method is particularly interesting for its ability to model non-Gaussian distributions with complex dependence structures, but also because it allows an explicit derivation of the PDF of $\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))|\boldsymbol{z}$ once the joint PDF is known. To construct relevant PDF approximations of this potentially high-dimensional random vector from a reduced number of code evaluations, we finally introduce two adaptations of the classical G-KDE formalism. First, an optimal partitioning of the components of $\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$ is introduced, which consists in decomposing the random vector to model in well-chosen groups of components that can reasonably be considered as independent. Secondly, a sequential strategy is proposed to choose the evaluations points on which the G-KDE relies. Starting from a space-filling design, the objective is to sequentially add new code evaluations in the regions where the posterior distribution of the parameters is high. We refer to [McKay et al., 1979, Fang and Lin, 2003, Fang et al., 2006, Draguljić et al., 2012, Joseph et al., 2015] for the construction of the initial space-filling designs when the input spaces is an hyperrectangle, and to [Stinstra et al., 2003, Stinstra et al., 2010, Auffray et al., 2012, Draguljić et al., 2012, Lekivetz and Jones, 2015, Mak and Joseph, 2016, Perrin and Cannamela, 2017] for the general case.

The outline of this work is as follows. Section 2 presents the theoretical framework of the proposed method. Section 3 first illustrates the efficiency of the method on an analytical example, and then shows its potential for the identification of the mechanical properties of an unknown heterogeneous medium.

2 Indirect identification of the statistical properties of random fields

The objective of this section is to describe the adaptive procedure we propose for the identification of the statistical properties of random fields when the available information is a set of indirect observations.

2.1 Definitions and notations

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. For $d_x, d_y, d_z \geq 1$,

- $\mathcal{P}(\mathbb{X}, \mathbb{R}^{d_x})$ denotes the space of all the second-order random fields defined on $(\Omega, \mathcal{A}, \mathbb{P})$, with values in $\mathbb{R}^{d_x}$, indexed by a compact and connected space $\mathbb{X}$;
- $L^2(\mathbb{X}, \mathbb{R}^{d_x})$ is the space of all the square-integrable functions defined on $\mathbb{X}$ with values in $\mathbb{R}^{d_x}$;
- $\boldsymbol{g}$ is a nonlinear measurable mapping whose computational cost can be high:

$$\boldsymbol{g} : \begin{cases} L^2(\mathbb{X}, \mathbb{R}^{d_x}) \ \to \ \mathbb{R}^{d_y} \\ \boldsymbol{h} \ \mapsto \ \boldsymbol{g}(\boldsymbol{h}) \end{cases} ; \tag{1}$$

- $\mathcal{X}(\mathbb{R}^{d_z}, \mathbb{R}^{d_x})$ refers to a particular class of random fields in $\mathcal{P}(\mathbb{X}, \mathbb{R}^{d_x})$, whose statistical properties are parameterized by a deterministic vector $\boldsymbol{z} \in \mathbb{R}^{d_z}$. For instance, $\mathcal{X}(\mathbb{R}^{d_z}, \mathbb{R}^{d_x})$ can correspond to the set of Gaussian random fields, whose mean and covariance functions are parameterized by the same d_z coefficients.
- For all $\boldsymbol{z}$ in $\mathbb{R}^{d_z}$, $\boldsymbol{X}(\boldsymbol{z})$ is an element of $\mathcal{X}(\mathbb{R}^{d_z}, \mathbb{R}^{d_x})$.

Let $\boldsymbol{X}^\star$ be a particular element of $\mathcal{P}(\mathbb{X}, \mathbb{R}^{d_x})$, which can belong or not to $\mathcal{X}(\mathbb{R}^{d_z}, \mathbb{R}^{d_x})$, and $\boldsymbol{Y}^\star$ be its transformation by $\boldsymbol{g}$. By construction, $\boldsymbol{Y}^\star$ is a d_y-dimensional random vector. For each realization of $\boldsymbol{X}^\star$, which we denote by $\boldsymbol{X}^\star(\theta)$ with $\theta \in \Omega$, $\boldsymbol{Y}^\star(\theta) := \boldsymbol{g}(\boldsymbol{X}^\star(\theta))$ defines a particular realization of $\boldsymbol{Y}^\star$.

Given N independent realizations of $\boldsymbol{Y}^\star$, gathered in the set

$$\mathcal{S}(N) := \{\boldsymbol{Y}^\star(\theta_n)\}_{1 \leq n \leq N}, \quad \theta_n \in \Omega,$$

the purpose of this work is to propose a Bayesian formalism for the identification of $\boldsymbol{z}^\star$, such that the probability distribution of $\boldsymbol{X}(\boldsymbol{z}^\star)$ is the closest to the one of $\boldsymbol{X}^\star$.

Remarks

- As mentioned in Introduction, it is important to notice that for each $\boldsymbol{z} \in \mathbb{R}^{d_z}$, $\boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$ is random. This strongly limits the possibility of replacing mapping $\boldsymbol{z} \mapsto \boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$ by a surrogate model, as it is classically done when solving inverse problems that invoke computationally expensive models.
- In the following, for the sake of simplicity, we assume that $\boldsymbol{X}^\star \in \mathcal{X}(\mathbb{R}^{d_z}, \mathbb{R}^{d_x})$. If it was not the case, it could be necessary to introduce an error term to model the difference between $\boldsymbol{X}^\star$ and $\boldsymbol{X}(\boldsymbol{z})$ [Kennedy and O'Hagan, 2001].

2.2 Bayesian formulation of the problem

In this work, $\boldsymbol{z}^\star$ is modeled by the random vector $\boldsymbol{Z}$, to take into account the fact that its value is unknown. Let $f_{\boldsymbol{Z}}$ be the probability density function

(PDF) of $\boldsymbol{Z}$, which is supposed to be known as a prior model. Hence, identifying $\boldsymbol{z}^\star$ amounts to searching the posterior PDF of $\boldsymbol{Z} \mid \mathcal{S}(N)$, which we denote by $f_{\boldsymbol{Z}|\mathcal{S}(N)}$. Using the Bayes theorem, it comes:

$$f_{\boldsymbol{Z}|\mathcal{S}(N)}(\boldsymbol{z}) = \frac{\mathcal{L}_{\mathcal{S}(N)}(\boldsymbol{z}) f_{\boldsymbol{Z}}(\boldsymbol{z})}{\mathbb{E}\left[\mathcal{L}_{\mathcal{S}(N)}(\boldsymbol{Z})\right]},\ \boldsymbol{z} \in \mathbb{R}^{d_z}. \tag{2}$$

There, $\mathbb{E}[\cdot]$ is the mathematical expectation and $\mathcal{L}_{\mathcal{S}(N)}$ is the likelihood function. The elements of $\mathcal{S}(N)$ being statistically independent, it follows:

$$\mathcal{L}_{\mathcal{S}(N)}(\boldsymbol{z}) = \prod_{n=1}^{N} f_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{Y}^\star(\theta_n)),\ \boldsymbol{z} \in \mathbb{R}^{d_z}, \tag{3}$$

in which $f_{\boldsymbol{Y}(\boldsymbol{z})}$ is the PDF of $\boldsymbol{Y}(\boldsymbol{z}) := \boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$ for given $\boldsymbol{z}$ in $\mathbb{R}^{d_z}$, and is unknown. To approximate $f_{\boldsymbol{Y}(\boldsymbol{z})}$, a first possibility is to generate M independent realizations of $\boldsymbol{Y}(\boldsymbol{z})$. Thus, based on this set, the value $f_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{y})$ of $f_{\boldsymbol{Y}(\boldsymbol{z})}$ in any point $\boldsymbol{y}$ in $\mathbb{R}^{d_y}$ can be approximated using any parametric or nonparametric statistical learning technique. However, this means that function $\boldsymbol{g}$ has to be evaluated $M \times Q$ times to evaluate function $\mathcal{L}_{\mathcal{S}(N)}$ in Q points for $\boldsymbol{z}$. This quickly becomes burdensome when the computational cost for each evaluation of $\boldsymbol{g}$ is relatively high (between several minutes to several hours CPU for the considered applications). One possible approach to circumvent this problem is to directly approximate the joint PDF of the $(d_y + d_z)$-dimensional random vector $(\boldsymbol{Y}(\boldsymbol{Z}), \boldsymbol{Z})$ [Soize and Ghanem, 2017]. Indeed, M independent realizations of $(\boldsymbol{Y}(\boldsymbol{Z}), \boldsymbol{Z})$ can be obtained from the following two-step procedure:

- we first draw at random M independent realizations of $\boldsymbol{Z}$ according to the distribution $f_{\boldsymbol{Z}}$, which we denote by $\boldsymbol{Z}(\omega_1)$, ..., $\boldsymbol{Z}(\omega_M)$, where $\omega_1, \ldots, \omega_M$ are in Ω;
- for each value of $\boldsymbol{z}$ in $\{\boldsymbol{Z}(\omega_1), \ldots, \boldsymbol{Z}(\omega_M)\}$, we draw, at random and independently the ones from the others, a particular realization of $\boldsymbol{X}(\boldsymbol{z})$, and we deduce a realization of $\boldsymbol{Y}(\boldsymbol{z})$ by evaluating $\boldsymbol{g}$ in this realization of $\boldsymbol{X}(\boldsymbol{z})$.

For the sake of simplicity, we denote these realizations by $\boldsymbol{Y}(\omega_m)$, $1 \leq m \leq M$. Based on these realizations, the kernel estimator of $f_{\boldsymbol{Y},\boldsymbol{Z}}$ is:

$$\widehat{f}_{\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{y}, \boldsymbol{z}; \boldsymbol{H}) := \frac{\det(\boldsymbol{H})^{-1/2}}{M} \sum_{m=1}^{M} K\left(\boldsymbol{H}^{-1/2}\left((\boldsymbol{y}, \boldsymbol{z}) - (\boldsymbol{Y}(\omega_m), \boldsymbol{Z}(\omega_m))\right)\right). \tag{4}$$

Here, $\det(\cdot)$ is the determinant operator, K is any positive function whose integral over $\mathbb{R}^{d_y+d_z}$ is one, and $\boldsymbol{H}$ is a $((d_y + d_z) \times (d_y + d_z))$-dimensional positive-definite symmetric matrix, which is generally referred as the "bandwidth matrix". In the following, we focus on the case where K is the Gaussian multidimensional density, and where $\boldsymbol{H}$ is proportional to the empirical estimation of the covariance matrix of $(\boldsymbol{Y}(\boldsymbol{Z}), \boldsymbol{Z})$, denoted by $\widehat{\boldsymbol{C}}$:

$$\boldsymbol{H} = h^2 \widehat{\boldsymbol{C}},\quad h \in \mathbb{R}. \tag{5}$$

The main interest of this hypothesis comes from the fact that it strongly reduces the number of parameters that need to be identified for the construction of $\boldsymbol{H}$, while generally leading to very interesting results for the modeling of multivariate PDFs (see [Perrin et al., 2018] for more details). Other parsimonious parameterizations could be proposed for $\boldsymbol{H}$, such as diagonal representations, but for sufficiently high values of M, the influence of this choice on the identification results is expected to be small.

Hence, the PDF of $(\boldsymbol{Y}(\boldsymbol{Z}), \boldsymbol{Z})$ is approximated by a mixture of M Gaussian PDFs, for which the means are the available realizations of $(\boldsymbol{Y}(\boldsymbol{Z}), \boldsymbol{Z})$ and the covariance matrices are all parameterized by a unique scalar h:

$$\widehat{f}_{\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{y},\boldsymbol{z};h) = \frac{1}{M}\sum_{m=1}^{M} \phi\left((\boldsymbol{y},\boldsymbol{z}); (\boldsymbol{Y}(\omega_m), \boldsymbol{Z}(\omega_m)), h^2\widehat{\boldsymbol{C}}\right). \tag{6}$$

There, for any $\mathbb{R}^d$-dimensional vector $\boldsymbol{\mu}$ and for any $(\mathbb{R}^d \times \mathbb{R}^d)$-dimensional symmetric positive-definite matrix $\boldsymbol{C}$, $\phi(\cdot;\boldsymbol{\mu},\boldsymbol{C})$ is the PDF of any $\mathbb{R}^d$-dimensional Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}$:

$$\phi\left(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{C}\right) := \frac{\exp\left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^T \boldsymbol{C}^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)\right)}{(2\pi)^{d/2}\sqrt{\det(\boldsymbol{C})}}, \quad \boldsymbol{x}\in\mathbb{R}^d. \tag{7}$$

In addition, the block decomposition of $\widehat{\boldsymbol{C}}$ is written as:

$$\widehat{\boldsymbol{C}} = \begin{bmatrix} \widehat{\boldsymbol{C}}_{YY} & \widehat{\boldsymbol{C}}_{YZ} \\ \widehat{\boldsymbol{C}}_{YZ}^T & \widehat{\boldsymbol{C}}_{ZZ} \end{bmatrix}. \tag{8}$$

For all $(\boldsymbol{y},\boldsymbol{z}) \in \mathbb{R}^{d_y}\times\mathbb{R}^{d_z}$, the kernel approximation of $f_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{y})$, which we denote by $\widehat{f}_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{y};h)$, can therefore be written as follows (see Appendix for more details about this expression):

$$\begin{aligned} \widehat{f}_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{y};h) &= \frac{\widehat{f}_{\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{y},\boldsymbol{z};h)}{\int_{\mathbb{R}^{d_y}} \widehat{f}_{\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{v},\boldsymbol{z};h)d\boldsymbol{v}} \\ &= \sum_{m=1}^{M} \frac{\gamma_m(\boldsymbol{z};h)}{\sum_{m'=1}^{M}\gamma_{m'}(\boldsymbol{z};h)} \phi\left(\boldsymbol{y};\boldsymbol{\mu}_m(\boldsymbol{z}),\boldsymbol{C}_m(h)\right), \end{aligned} \tag{9}$$

$$\gamma_m(\boldsymbol{z};h) := \exp\left(-\frac{1}{2h^2}\left(\boldsymbol{z}-\boldsymbol{Z}(\omega_m)\right)^T \widehat{\boldsymbol{C}}_{\boldsymbol{ZZ}}^{-1}\left(\boldsymbol{z}-\boldsymbol{Z}(\omega_m)\right)\right), \tag{10}$$

$$\boldsymbol{\mu}_m(\boldsymbol{z}) := \boldsymbol{Y}(\omega_m) + \widehat{\boldsymbol{C}}_{\boldsymbol{YZ}}\widehat{\boldsymbol{C}}_{\boldsymbol{ZZ}}^{-1}(\boldsymbol{z}-\boldsymbol{Z}(\omega_m)), \tag{11}$$

$$\boldsymbol{C}_m(h) := h^2\left(\widehat{\boldsymbol{C}}_{\boldsymbol{YY}} - \widehat{\boldsymbol{C}}_{\boldsymbol{YZ}}\widehat{\boldsymbol{C}}_{\boldsymbol{ZZ}}^{-1}\widehat{\boldsymbol{C}}_{\boldsymbol{YZ}}^T\right). \tag{12}$$

It follows that the posterior PDF of $\boldsymbol{Z}$ is estimated for each $\boldsymbol{z}$ in $\mathbb{R}^{d_z}$ by:

$$f_{\boldsymbol{Z}|\mathcal{S}(N)}(\boldsymbol{z}) \approx \frac{\widehat{\mathcal{L}}_{\mathcal{S}(N)}(\boldsymbol{z};h) f_{\boldsymbol{Z}}(\boldsymbol{z})}{\mathbb{E}\left[\widehat{\mathcal{L}}_{\mathcal{S}(N)}(\boldsymbol{Z})\right]}, \quad \widehat{\mathcal{L}}_{\mathcal{S}(N)}(\boldsymbol{z};h) := \prod_{n=1}^{N} \widehat{f}_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{Y}^{\star}(\theta_n);h). \quad (13)$$

Remarks

- One key step of these methods is the exploration of the whole space of the input variables. To maximize this covering, it is generally worth choosing $\{\boldsymbol{Z}(\omega_1), \ldots, \boldsymbol{Z}(\omega_M)\}$ as a space filling design of experiments that preserves good projection properties for each scalar input (see [Fang and Lin, 2003, Fang et al., 2006, Perrin and Cannamela, 2017] for the construction of such designs when prior density $f_{\boldsymbol{Z}}$ is uniform or not).
- Another crucial aspect of these Bayesian approaches is the choice of prior distribution $f_{\boldsymbol{Z}}$. Indeed, the more informative it is, the less measurements we need to get a useful posterior distribution for $\boldsymbol{Z}$. But if it is overconfident around values that are potentially biased, the uncertainty carried by the posterior distribution may not be large enough to adequately capture the true value of $\boldsymbol{Z}$ (see [Marin and Robert, 2007] for more details on the construction of this prior distribution).
- In the standard case, the M code evaluations are generally used to construct a surrogate model of a computationally expensive but deterministic code. Hence, depending on the dimension of the input space and the regularity of the code output with respect to the inputs, interesting approximations can be obtained using relatively small values of M [Perrin et al., 2017]. On the contrary, in our case, as $\boldsymbol{z} \mapsto \boldsymbol{g}(\boldsymbol{X}(\boldsymbol{z}))$ is a stochastic simulator, the value of M is likely to be higher, as we want the code evaluations to allow a precise approximation of the dependence structure between $\boldsymbol{Y}(\boldsymbol{Z})$ and $\boldsymbol{Z}$ in the construction of their joint PDF. And the higher $d_y + d_z$ is, the higher value of M we may need. However, when confronted to expensive simulators, the maximal number of code evaluations is generally limited (M must be less than 1000 for instance). In that case, it is particularly important to work on methods that allow the most precise identification of the parameters at the minimal cost. This is the objective of the following sections. In Section 2.3, we first propose to decompose $\boldsymbol{Y}(\boldsymbol{Z})$ in several groups to improve the relevance of the nonparametric representation of PDF $f_{\boldsymbol{Y},\boldsymbol{Z}}$ for a fixed value of M. Then, selection criteria are proposed in Section 2.5 to sequentially concentrate the code evaluations in the most likely regions for $\boldsymbol{Z}$, and therefore reduce the uncertainties on its posterior PDF $f_{\boldsymbol{Z}|\mathcal{S}(N)}$.

2.3 Optimal partitioning

As it is explained in [Perrin et al., 2018], when d_y becomes high, separating in different groups the components of $\boldsymbol{Y}(\boldsymbol{Z})|\boldsymbol{Z}$ that could reasonably be considered as independent can strongly improve the relevance of $\widehat{f}_{\boldsymbol{Y}(\boldsymbol{z})}$ for a fixed

number of code evaluations. Let $\boldsymbol{b} = (b_1, \ldots, b_{d_y})$ be a particular group decomposition of $\boldsymbol{Y}(\boldsymbol{Z})|\boldsymbol{Z}$ in the sense that:

- if $b_i = b_j$, $Y_i(\boldsymbol{Z})|\boldsymbol{Z}$ and $Y_j(\boldsymbol{Z})|\boldsymbol{Z}$ are supposed to be dependent and therefore belong to the same block,
- if $b_i \neq b_j$, $Y_i(\boldsymbol{Z})|\boldsymbol{Z}$ and $Y_j(\boldsymbol{Z})|\boldsymbol{Z}$ are supposed to be independent and they can belong to two different blocks.

To avoid redundancies in this block by block representation, vector $\boldsymbol{b}$ can be chosen in the set:

$$\mathbb{B}(d_y) := \left\{ \boldsymbol{b} \in \{1, \ldots, d_y\}^{d_y} \ \mid b_1 = 1,\ 1 \leq b_j \leq 1 + \max_{1 \leq i \leq j-1} b_i,\ 2 \leq j \leq d_y \right\}. \tag{14}$$

Hence, for any $\boldsymbol{b}$ in $\mathbb{B}(d_y)$, we can define

- $\mathrm{Max}(\boldsymbol{b})$ as the maximal value of $\boldsymbol{b}$,
- $\boldsymbol{Y}^{(\ell)}(\boldsymbol{z}, \boldsymbol{b})$ as the random vector that gathers all the components of $\boldsymbol{Y}(\boldsymbol{Z})|\boldsymbol{Z} = \boldsymbol{z}$ with a block index equal to ℓ,
- $\boldsymbol{y}^{(\ell)}(\boldsymbol{y}, \boldsymbol{b})$ as the vector that gathers all the components of $\boldsymbol{y}$ with a block index equal to ℓ.

For all $\boldsymbol{b}$ in $\mathbb{B}(d_y)$, $\boldsymbol{z}$ in $\mathbb{R}^{d_z}$ and $\boldsymbol{h} := (h_1, \ldots, h_{\mathrm{Max}(\boldsymbol{b})})$ in $\mathbb{R}^{\mathrm{Max}(\boldsymbol{b})}$, if $\widehat{f}_{\boldsymbol{Y}^{(\ell)}(\boldsymbol{z},\boldsymbol{b})}(\boldsymbol{y}^{(\ell)}(\boldsymbol{y}, \boldsymbol{b}); h_\ell)$ is the kernel estimator of the PDF of $\boldsymbol{Y}^{(\ell)}(\boldsymbol{z}, \boldsymbol{b})$, it comes:

$$f_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{y}) \approx \widetilde{f}_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{y}; \boldsymbol{h}, \boldsymbol{b}) := \prod_{\ell=1}^{\mathrm{Max}(\boldsymbol{b})} \widehat{f}_{\boldsymbol{Y}^{(\ell)}(\boldsymbol{z},\boldsymbol{b})}(\boldsymbol{y}^{(\ell)}(\boldsymbol{y}, \boldsymbol{b}); h_\ell),\ \boldsymbol{y} \in \mathbb{R}^{d_y}, \tag{15}$$

leading to another approximation of $f_{\boldsymbol{Z}|\mathcal{S}(N)}(\boldsymbol{z})$ for each $\boldsymbol{z}$ in $\mathbb{R}^{d_z}$:

$$f_{\boldsymbol{Z}|\mathcal{S}(N)}(\boldsymbol{z}) \approx \widetilde{f}_{\boldsymbol{Z}|\mathcal{S}(N)}(\boldsymbol{z}) := \frac{\widetilde{\mathcal{L}}_{\mathcal{S}(N)}(\boldsymbol{z}; \boldsymbol{h}, \boldsymbol{b}) f_{\boldsymbol{Z}}(\boldsymbol{z})}{\mathbb{E}\left[\widetilde{\mathcal{L}}_{\mathcal{S}(N)}(\boldsymbol{Z})\right]}, \tag{16}$$

$$\widetilde{\mathcal{L}}_{\mathcal{S}(N)}(\boldsymbol{z}; \boldsymbol{h}, \boldsymbol{b}) := \prod_{n=1}^{N} \widetilde{f}_{\boldsymbol{Y}(\boldsymbol{z})}(\boldsymbol{Y}^\star(\theta_n); \boldsymbol{h}, \boldsymbol{b}). \tag{17}$$

2.4 Estimation of the kernel parameters

To evaluate $\widetilde{\mathcal{L}}_{\mathcal{S}(N)}$, the values of $\boldsymbol{h}$ and $\boldsymbol{b}$ have to be identified. This can be done by solving the following optimization problem:

$$(\boldsymbol{h}^{\mathrm{AIC}}, \boldsymbol{b}^{\mathrm{AIC}}) \approx \arg \min_{\boldsymbol{h} \in]0, +\infty[^{\mathrm{Max}(\boldsymbol{b})},\ \boldsymbol{b} \in \mathbb{B}(d_y)} \mathrm{AIC}^{\mathrm{LOO}}(\boldsymbol{h}, \boldsymbol{b}), \tag{18}$$

$$\mathrm{AIC}^{\mathrm{LOO}}(\boldsymbol{h},\boldsymbol{b}) := 2\mathrm{Max}(\boldsymbol{b}) - 2\log\left(\prod_{m=1}^{M}\prod_{\ell=1}^{\mathrm{Max}(\boldsymbol{b})} \widehat{f}^{(-m)}_{\boldsymbol{Y}^{(\ell)}(\boldsymbol{Z}(\omega_m),\boldsymbol{b})}(\boldsymbol{Y}^{(\ell)}(\boldsymbol{Y}(\omega_m),\boldsymbol{b});h_\ell)\right), \tag{19}$$

where $\widehat{f}^{(-m)}_{\boldsymbol{Y}^{(\ell)}(\boldsymbol{Z}(\omega_m),\boldsymbol{b})}$ is the kernel estimator of the PDF of $\boldsymbol{Y}^{(\ell)}(\boldsymbol{Z}(\omega_m),\boldsymbol{b})$ that is based on all the evaluations of $\boldsymbol{g}$ but the m^{th} one. Indeed, given Eq. (9), this amounts to minimizing a "Leave-One-Out" version of the Akaike information criterion (AIC) [Akaike, 1974] associated with the PDF of $\boldsymbol{Y}(\boldsymbol{Z})|\boldsymbol{Z}$ (very close results would be obtained by considering another information criterion such as the Bayesian information criterion (BIC)). We refer to [Perrin et al., 2018] for more details about the solving of this optimization problem.

2.5 Adaptive strategy

By construction, the precision of the estimation of $\boldsymbol{z}^\star$ depends on the number of experimental measurements, N, and the number of code evaluations, M. Classically, the value of N is fixed, whereas it should be possible to improve the accuracy of $\widetilde{f}_{\boldsymbol{Y}(\boldsymbol{z})}$, which is defined by Eq. (15), by adding new code evaluations in the learning set. For instance, M^{new} new points could be added to the learning set by evaluating the code in M^{new} independent realizations of $\boldsymbol{Z}|\mathcal{S}(N)$ (we remind that no code evaluations are required to choose these new points). However, as the kernel density estimator is based on the post-processing of independent and identically distributed realizations of the random vector to model, non consistent results could be obtained by mixing realizations of $\boldsymbol{Z}|\mathcal{S}(N)$ with realizations of $\boldsymbol{Z}$. If such a selection criterion was chosen, this would mean that the M code evaluations at the initial step should not be used for the refining.

As an alternative, we propose to evaluate the function

$$\boldsymbol{z} \mapsto \widetilde{f}(\boldsymbol{z}) := \widetilde{\mathcal{L}}_{\mathcal{S}(N)}(\boldsymbol{z};\boldsymbol{h}^{\mathrm{AIC}},\boldsymbol{b}^{\mathrm{AIC}}) f_{\boldsymbol{Z}}(\boldsymbol{z})$$

in each value of $\{\boldsymbol{Z}(\omega_1),\ldots,\boldsymbol{Z}(\omega_M)\}$. For each $1 \leq m \leq M$, let π_m be the following weights:

$$0 \leq \pi_m := \frac{\widetilde{f}(\boldsymbol{Z}(\omega_m))}{\sum_{m'=1}^{M}\widetilde{f}(\boldsymbol{Z}(\omega_{m'}))} \leq 1. \tag{20}$$

Without loss of generality, these weights are assumed to be sorted in decreasing order, $\pi_1 \geq \pi_2 \geq \ldots \geq \pi_M$. Hence, for $0 < \alpha < 1$, if we denote by Q_α the smallest integer such that:

$$\sum_{m=1}^{Q_\alpha - 1} \pi_m \geq \alpha, \tag{21}$$

the domain $\mathcal{Z}_\alpha := \{\boldsymbol{z} \in \mathbb{R}^{d_z} \mid \widetilde{f}(\boldsymbol{z}) \geq \widetilde{f}(\boldsymbol{Z}(\omega_{Q_\alpha})\}$ can be seen as a conservative α-credible set for $\boldsymbol{Z}|\mathcal{S}(N)$, in the sense that the probability for $\boldsymbol{Z}|\mathcal{S}(N)$ to be in $\mathcal{Z}_\alpha$ is likely to be higher than α. Therefore, adding new realizations of $\boldsymbol{Z}|\boldsymbol{Z} \in \mathcal{Z}_\alpha$ seems a good mean to enrich the set of points on which the kernel density estimator is based. Indeed, the most likely values of $\boldsymbol{z}$ at the former step are kept in the adaptive procedure, while a good exploration of the input domain is expected if the value of α is chosen sufficiently high.

Finally, choosing $\boldsymbol{f}_{\boldsymbol{Z}|\boldsymbol{Z}\in\mathcal{Z}_\alpha}$ instead of $f_{\boldsymbol{Z}}$ for the prior distribution of $\boldsymbol{Z}$, and repeating several times this procedure, it is possible to iteratively reduce the uncertainties about $\boldsymbol{z}^\star$.

Remarks

- By adding new code evaluations, the objective is to make $\widetilde{f}_{\boldsymbol{Y}(\boldsymbol{z})}$ be as close to $f_{\boldsymbol{Y}(\boldsymbol{z})}$ as possible, such that the approximate posterior $\widetilde{f}_{\boldsymbol{Z}|\mathcal{S}(N)}$ is as close to the true (but unknown) posterior $f_{\boldsymbol{Z}|\mathcal{S}(N)}$ as possible. Choosing a value of α that is strictly inferior to one only aims at limiting the number of new code evaluations that will be in the region where true posterior $f_{\boldsymbol{Z}|\mathcal{S}(N)}$ is almost zero. However, this value has not to be chosen too small, as it would artificially reduce the uncertainty associated with the estimation of $\boldsymbol{z}^\star$ by cutting too much the tails of the true posterior. Hence, in the applications that will be presented in Section 3, α is chosen equal to 0.99.
- According to Eq. (21), we deliberately add one to the value of Q_α to be conservative for the estimation of the α-credible set. This is particularly important for cases when after the first iteration, $\pi_1 \approx 1$. Indeed, even if one value of $\boldsymbol{z}$ appears to be much more relevant than the others, we do not want to focus too much around a single mode.

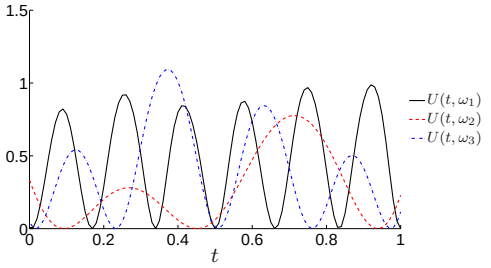
3 Applications

The purpose of this section is to illustrate the method proposed in Section 2 on two applications.

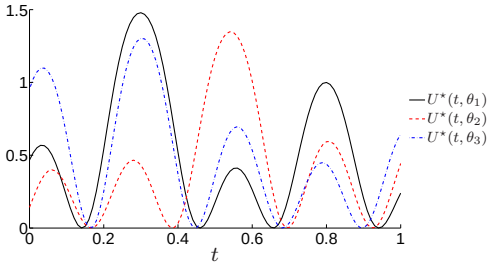
3.1 Analytical application

In this first application, $X(\boldsymbol{z})$ refers to the Gaussian random fields whose mean is equal to $t \mapsto \sin(2\pi z_3 t + z_4)$, and whose covariance function is equal to $(t, t') \mapsto z_1^2 \exp\left(-\frac{(t-t')^2}{2z_2^2}\right)$.

This class of random fields is therefore parameterized by four quantities: two parameters for the mean value, denoted by z_3 and z_4, and two parameters for the covariance function, denoted by z_1 and z_2. We then introduce $U(X(\boldsymbol{z}))$ as the image of $X(\boldsymbol{z})$ by the following nonlinear mapping:



(a) Realizations of $U(X(\boldsymbol{Z}))$



(b) Realizations of $U(X(\boldsymbol{Z}))|\boldsymbol{Z}=\boldsymbol{z}^\star$

Fig. 1 Comparison of independent realizations $U(X(\boldsymbol{Z}))$ and $U(X(\boldsymbol{Z}))|\boldsymbol{Z}=\boldsymbol{z}^\star$.

$$U(X(\boldsymbol{z})) := \{X(t;\boldsymbol{z})\sin(X(t;\boldsymbol{z})),\ t\in[0,1]\}\,. \tag{22}$$

The value of $\boldsymbol{z}^\star$ is chosen equal to $(0.3, 0.2, 2, 1)$, and it is *a priori* modeled by a uniformly distributed over $[0.1,1]\times[0.05,1]\times[1,3]\times[0,2]$ random vector, denoted by $\boldsymbol{Z}$. To identify $\boldsymbol{z}^\star$, the available information is made of $N=10$ independent realizations of $U(X(\boldsymbol{Z}))|\boldsymbol{Z}=\boldsymbol{z}^\star$, denoted by $U^\star(\theta_1),\ldots,U^\star(\theta_N)$. To solve the inference problem, $M=500$ independent realizations of $\boldsymbol{Z}$ have been drawn, which we write $\{\boldsymbol{Z}(\omega_1),\ldots,\boldsymbol{Z}(\omega_M)\}$. For each $1\le m\le M$, we then draw at random one realization of $U(X(\boldsymbol{Z}(\omega_m)))$, and we denote it by $U(\omega_m)$ for the sake of simplicity. As an illustration, several realizations of $U(X(\boldsymbol{Z}))$ and $U(X(\boldsymbol{Z}))|\boldsymbol{Z}=\boldsymbol{z}^\star$ are compared in Figure 1.

In principle, the Bayesian formulation can be applied to any multi-variate output code. But in practice, it is generally very convenient to condense (if it is possible) the statistical content of the code output in a low-dimensional vector [Perrin, ress]. In our context, it is even more important, as a key step of the proposed method is the identification of the joint distribution between the parameters to be identified and the associated code output, whose complexity strongly increases with the dimension of the code output. In that prospect, we introduce ψ_p, $p\ge 1$ as the solutions of the following eigenvalue problem:

$$\int_0^1 \sum_{m=1}^{M} U(t,\omega_m)U(t',\omega_m)\psi_p(t')dt' = \lambda_p\psi_p(t), \tag{23}$$

$$\lambda_1 \geq \lambda_2 \geq \cdots \to 0, \quad \int_0^1 \psi_p(t')\psi_q(t')dt' = \delta_{pq}, \tag{24}$$

where δ_{pq} is the Kronecker symbol that is equal to 1 if p = q and 0 otherwise. To solve the inference problem, we finally introduce $\boldsymbol{Y}(\boldsymbol{z})$ as the vector gathering the projection coefficients of $U(X(\boldsymbol{Z}))$ on the former eigenfunctions associated with the d_y highest eigenvalues:

$$\boldsymbol{Y}(\boldsymbol{Z}) := \left(\int_0^1 U(t; X(\boldsymbol{Z}))\psi_1(t)dt, \ldots, \int_0^1 U(t; X(\boldsymbol{Z}))\psi_{d_y}(t)dt \right). \tag{25}$$

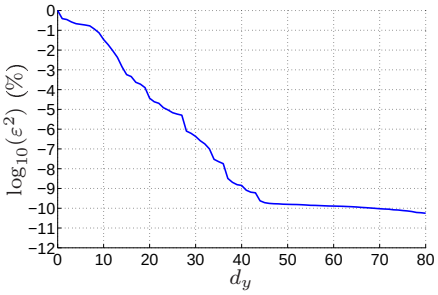
The value of d_y can then be chosen to guarantee a relevant representation of the observations. To this end, we introduce ε^2 as the following quantity:

$$\varepsilon^2(d_y) := \frac{\sum_{n=1}^N \int_0^1 \left(U^\star(t, \theta_n) - \widehat{U}^\star(t, \theta_n; d_y) \right)^2 dt}{\sum_{n=1}^N \int_0^1 \left(U^\star(t, \theta_n) \right)^2 dt}, \tag{26}$$

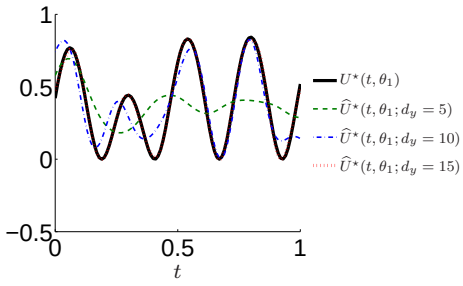
$$\widehat{U}^\star(t, \theta_n; d_y) := \sum_{p=1}^{d_y} \psi_p(t) \left(\int_0^1 U^\star(t', \theta_n)\psi_p(t')dt' \right). \tag{27}$$

As an illustration, Figure 2 shows the evolution of $\varepsilon^2(d_y)$ with respect to d_y, as well as the difference between $U^\star(t, \theta_1)$ and $\widehat{U}^\star(t, \theta_1; d_y)$ for three values of d_y. For this application, d_y was chosen equal to 12, which corresponds to a value of ε^2 that is less than 1%.

Based on these M realizations of $(\boldsymbol{Y}(\boldsymbol{Z}), \boldsymbol{Z})$, and on these N realizations of $\boldsymbol{Y}(\boldsymbol{z}^\star) := \boldsymbol{Y}(\boldsymbol{Z})|\boldsymbol{Z} = \boldsymbol{z}^\star$, the adaptive Bayesian formalism presented in Section 2 is now applied. For this application, the parameter α, which was introduced in Section 2.5, is chosen equal to 0.99. At each iteration, new samples are therefore added in the region where $f_{\boldsymbol{Z}|\mathcal{S}(N)}$ is not too small using a rejection approach until we get a total of M points (including the points computed at the former iterations) in the α-credible set $\mathcal{Z}_\alpha$, whose definition is also given in Section 2.5. After 5 iterations, the total number of calls to the code is equal to 2300, which means that around 450 new points have been added at each iteration. The results are summarized in Table 3.1 and Figures 3 and 4. As a first comment, we verify that the identification of $\boldsymbol{z}^\star$ after only one iteration is not very precise, in the sense that the prediction uncertainties are very high. This is not surprising, as we are trying to approximate the PDF of a 16-dimensional random vector ($d_y = 12$, $d_z = 4$) on its whole definition domain from only 500 realizations. Moving from $M = 500$ to $M = 2300$, that is to say spending the total budget at the first iteration, does not really help. Indeed, the results we get in terms of mean and variance of $\boldsymbol{Z}|\mathcal{S}(N)$ are approximatively the same. This is explained by the fact that even if the number of points is almost multiplied by five, the coverage of the definition



(a) Error convergence



(b) Reduced basis representation

Fig. 2 Evolution of the projection error with respect to d_y.

	Z_1	Z_2	Z_3	Z_4
Reference	0.3	0.2	2	1
$\mathbb{E}[\boldsymbol{Z}\|\mathcal{S}(N)]$, $M=2300$, $i=1$	0.25	0.57	2.00	1.02
$\mathbb{E}[\boldsymbol{Z}\|\mathcal{S}(N)]$, $M=500$, $i=1$	0.24	0.59	2.00	1.02
$\mathbb{E}[\boldsymbol{Z}\|\mathcal{S}(N)]$, $M=500$, $i=2$	0.35	0.31	2.00	1.04
$\mathbb{E}[\boldsymbol{Z}\|\mathcal{S}(N)]$, $M=500$, $i=3$	0.34	0.23	2.00	1.04
$\mathbb{E}[\boldsymbol{Z}\|\mathcal{S}(N)]$, $M=500$, $i=4$	0.34	0.24	2.00	1.04
$\mathbb{E}[\boldsymbol{Z}\|\mathcal{S}(N)]$, $M=500$, $i=5$	0.29	0.19	2.00	1.04

Table 1 Evolution of the posterior mean with respect to the iteration number.

domain stays very sparse. On the contrary, adding iteratively around 450 new code evaluations in the most likely region, whose volume is much smaller than the initial volume, allows $\mathbb{E}[\boldsymbol{Z}|\mathcal{S}(N)]$ to tend to $\boldsymbol{z}^\star$, and strongly reduces the credible intervals. This convergence is quicker for the mean parameters than for the covariance parameters, which was also expected, as the mean function is generally easier to identify than the covariance. Focusing on Figure 4, it is also interesting to notice that the reference value does not need to be in the 99%-credible ellipse associated with $\boldsymbol{Z}|\mathcal{S}(N)$ at the first iteration to be in the 99%-credible ellipses at the next iterations.

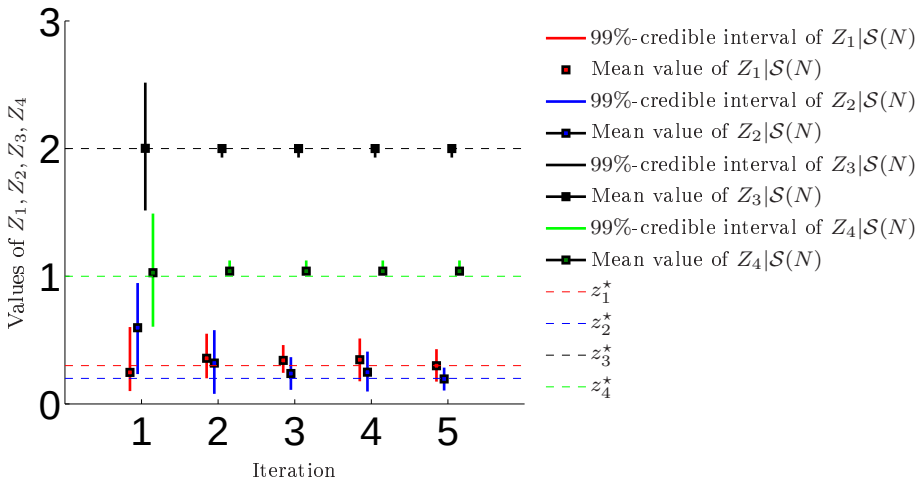


Fig. 3 Evolution of the 99%-credible intervals and of the mean values of the components of $\boldsymbol{Z}|\mathcal{S}(N)$ with respect to the iteration number.

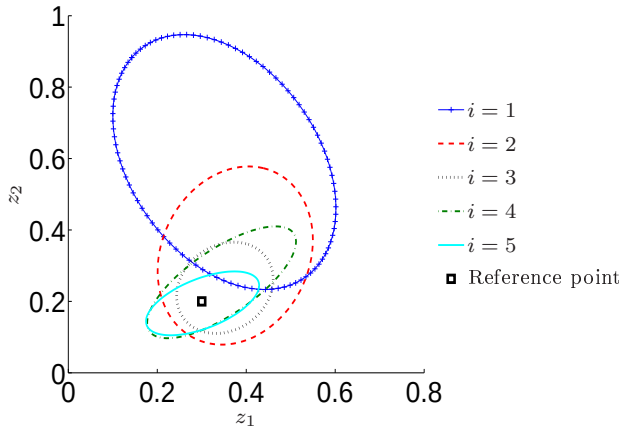
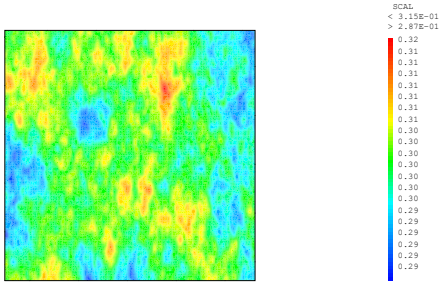


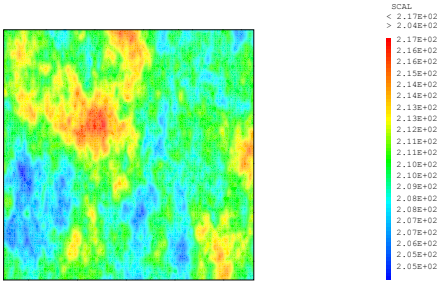
Fig. 4 Evolution of the 99%-credible ellipses of the two first elements of $\boldsymbol{Z}|\mathcal{S}(N)$ with respect to the iteration number.

3.2 Application to the identification of the mechanical properties of an unknown anisotropic material

The second application deals with the identification of the mechanical properties of an heterogeneous micro-structure, which is modeled by a random elastic medium. To this end, several experimental tests are performed on a series of specimens made of the same material. To be coherent with the notations introduced in Section 2, we denote by $\boldsymbol{X}$ the elasticity field characterizing the



(a) Evolution of the Poisson ratio



(b) Evolution of the Young modulus

Fig. 5 One potential spatial variation of the Young modulus and the Poisson ratio associated with the stochastic model $\boldsymbol{X}(\boldsymbol{z}^\star)$.

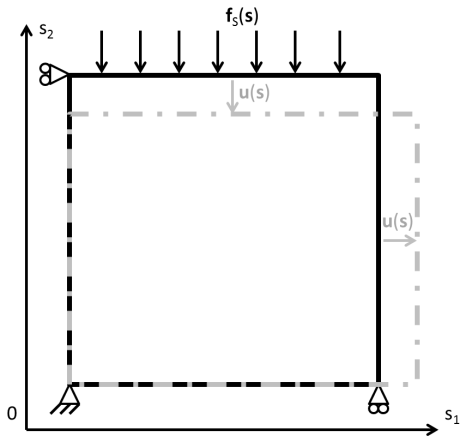


Fig. 6 Representation of the studied mechanical phenomenon.

mechanical properties of the material that constitutes the specimens. Several stochastic models have been proposed in the framework of the heterogeneous anisotropic linear elasticity [Soize, 2006, Soize, 2008, Clouteau et al., 2013, Guilleminot and Soize, 2013]. It should be noted that the elasticity field is not a real-valued random field, but a tensor-valued random field, and that the different components of this random field cannot be identified separately due to algebraic constraints. For this application, the stochastic model for the elasticity field is based on the model proposed in [Soize, 2006] and [Guilleminot and Soize, 2013] in 2D plan stresses for the sake of simplicity. Hence, the distribution of $\boldsymbol{X}$ is non-Gaussian, and it is parameterized by a 5-dimensional deterministic vector $\boldsymbol{z} = (z_1, \dots, z_5)$, where:

- z_1 is a positive dispersion coefficient that controls the level of fluctuations,
- z_2, z_3 are two spatial correlation lengths,
- z_4 is the mean value of the Young Modulus ($\times 10^9$ Pa);
- z_5 is the mean value of the Poisson ratio.

We then assume that $N = 100$ cubic specimens are available, whose respective mechanical properties are characterized by one particular realization of $\boldsymbol{X}(\boldsymbol{z}^\star)$, with $\boldsymbol{z}^\star = (2000,\ 0.1, 0.15, 210, 0.3)$. As an illustration, Figure 5 shows, for one particular specimen, the evolution of the Young modulus and the Poisson ratio in each point of $[0,1]^2$. The same pressure field $\boldsymbol{f}_S = -f_S \boldsymbol{e}_2$ is then imposed on the top of each specimen, and we only have access to the induced displacement field on the boundaries of these specimens (see Figure 6 for an illustration of the experimental protocol). Let $\boldsymbol{U}^\star(\theta_1), \dots, \boldsymbol{U}^\star(\theta_N)$ be these measured displacements.

Based on this set of measurements, the method described in Section 2 could directly be applied to the identification of $\boldsymbol{z}^\star$. To speed up this identification, following the works achieved in [Nguyen et al., 2015], we propose an alternative method, which is based on a two-step procedure. First, $z_4^\star$ and $z_5^\star$ will be identified by confronting the measured displacements to the homogeneous case. Once $z_4^\star$ and $z_5^\star$ have been found, a Bayesian formalism will be proposed for the identification of the three remaining components of $\boldsymbol{z}^\star$.

Indeed, if the specimens were made of a homogeneous material, characterized by its young modulus E and its Poisson ratio ν, it is well known [Lai et al., 2010] that the induced displacement in each point $\boldsymbol{s} \in [0,1]^2$ would be equal to $\boldsymbol{u}^{\text{homo}}(\boldsymbol{s}) = (as_1, bs_2)$, with

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{bmatrix} \lambda & \lambda + 2\mu \\ \lambda + 2\mu & \lambda \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ -f_S \end{pmatrix}, \quad \mu = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{2\mu\nu}{(1-2\nu)}. \tag{28}$$

Hence, as we are considering a class of stationary random processes, the values of $z_4^\star$ and $z_5^\star$ can be identified as the arguments that minimize the L2 distance between the N measured displacements and the associated homogeneous displacements. In this two-step approach, the Bayesian identification is

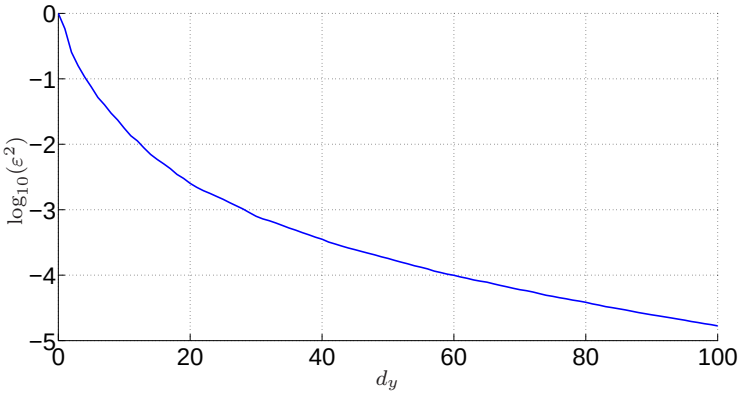
no longer carried out in dimension 5, but in dimension 3. This strongly reduces the number of code evaluations that will be needed for a correct identification of $(z_1^\star, z_2^\star, z_3^\star)$.

Thus, in the following, only $z_1^\star, z_2^\star$ and $z_3^\star$ are modeled by random quantities. They are gathered in the vector $\boldsymbol{Z}$, whose components are assumed independent and distributed according to the following distributions:

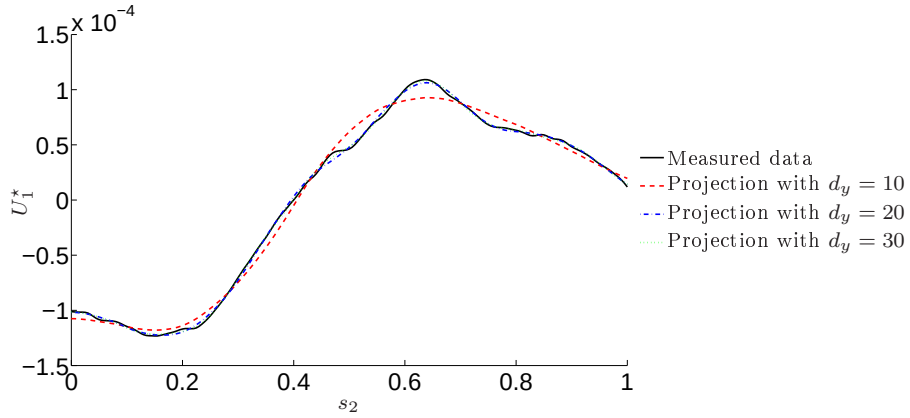
$$\log(Z_1) \sim \mathcal{U}(4.6, 11.5),\ Z_2 \sim \mathcal{U}(0.01, 0.3),\ Z_3 \sim \mathcal{U}(0.01, 0.3), \qquad (29)$$

where for all $a < b$, $\mathcal{U}(a, b)$ is the uniform distribution over $[a, b]$. For a given value of $\boldsymbol{Z}$, it is possible to simulate independent realizations of $\boldsymbol{X}(\boldsymbol{Z})$, and to approximate (using the Finite Element Method) the displacements induced by the experimental force field, which we write $\boldsymbol{U}(\boldsymbol{X}(\boldsymbol{Z}))$. Thus, for this second application, we first chose at random $M = 1000$ values of $\boldsymbol{Z}$ according to its prior distribution. For each of these values, a particular realization of the elasticity tensor was then generated over $[0, 1]^2$, and the mechanical problem that corresponds to the experimental protocol was solved (using the software Cast3M) to get the displacements at the boundary of the cube. In the same manner than in Section 3.1, we finally introduce $\boldsymbol{Y}(\boldsymbol{Z})$ as the projection of $\boldsymbol{U}(\boldsymbol{X}(\boldsymbol{Z}))$ on the d_y first eigenfunctions associated with the empirical estimation of the covariance of $\boldsymbol{U}(\boldsymbol{X}(\boldsymbol{Z}))$ based on the M code evaluations. In the same manner, we gather in $\mathcal{S}(N)$ the projection coefficients of each measured displacement $\boldsymbol{U}^\star(\theta_n)$ on this reduced basis. To choose the value of d_y, the normalized error defined by Eq. (26) is once again considered. For this application, d_y is chosen equal to 23 in order to correctly represent most of the local oscillations of the displacements. According to Figure 7, this corresponds to a projection error that is less than 0.1%.

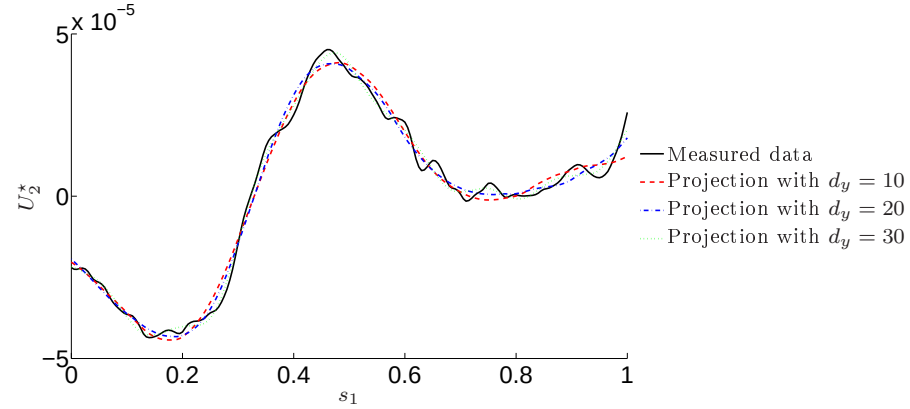
Following the framework proposed in Section 2, the PDF of $\boldsymbol{Z}|\mathcal{S}(N)$ is deduced from the kernel estimator of the PDF of $(\boldsymbol{Y}(\boldsymbol{Z}), \boldsymbol{Z})$. An adaptive procedure (with $\alpha = 0.99$) is moreover introduced to better concentrate the distribution of $\boldsymbol{Z}|\mathcal{S}(N)$ on the true value of $\boldsymbol{z}^\star$. To be more precise, 900 new code evaluations were added between the two first iterations, and 620 between the two last iterations, leading to a total budget of 2520 code evaluations. The relevance of this approach is shown in Figure 8, where the blue continuous lines correspond to the 95%-credible ellipses associated with the distribution of $\boldsymbol{Z}|\mathcal{S}(N)$. After three iterations, the values of $z_1^\star, z_2^\star$ and $z_3^\star$ are indeed identified with a high precision. To emphasize the interest of the partitioning presented in Section 3.2, these results are compared to the case where there is no optimization of the block structure (the ellipses in red dotted lines). Although these two approaches are based on the same information, there is no denying that searching groups of independent components of $\boldsymbol{Y}(\boldsymbol{Z})|\boldsymbol{Z}$ is really helpful. This is especially true for the first iteration, where 23 groups of independent components were chosen, and for the second iteration, where 8 groups of independent components were chosen. For the third iteration, as only 4 groups



(a) Error convergence

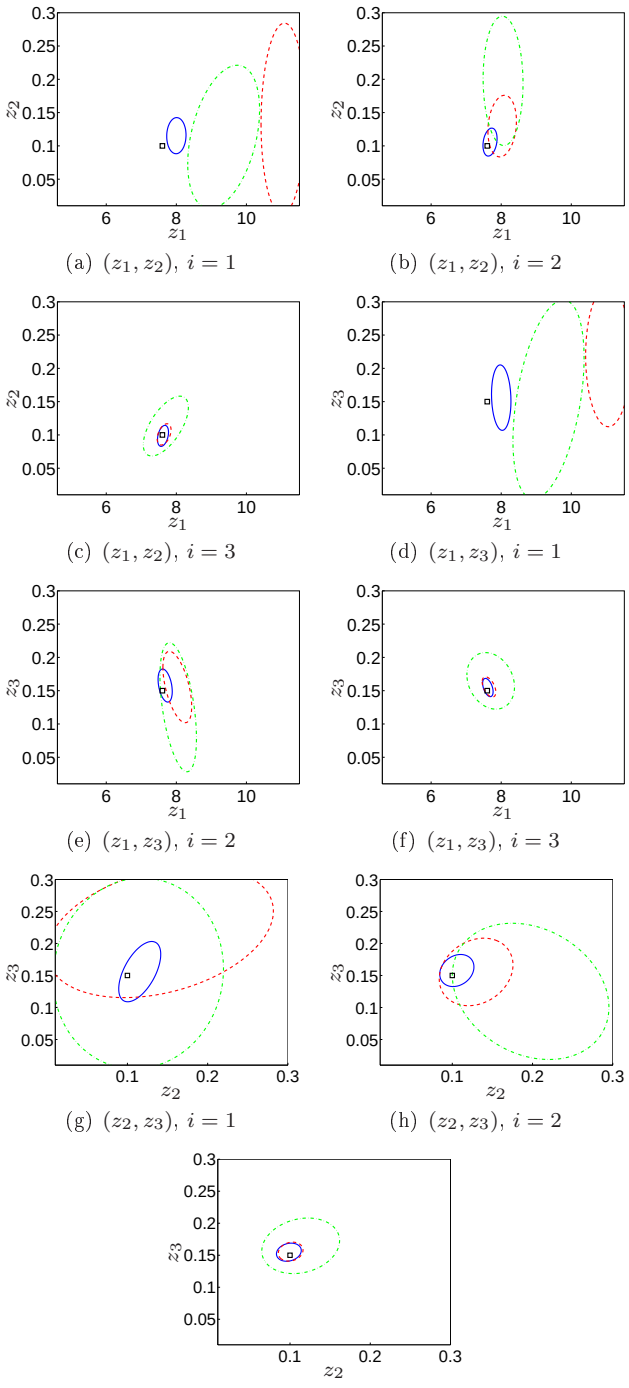


(b) Reduced basis representation of $U_1^\star(1, s_2; \theta_1) - \boldsymbol{u}^{\text{homo}}(1, s_2)$



(c) Reduced basis representation of $U_2^\star(s_1, 1; \theta_1) - \boldsymbol{u}^{\text{homo}}(0, 1)$

Fig. 7 Influence of truncation parameter d_y on the representation of the measured data.



(a) (z_1, z_2), $i = 1$ (b) (z_1, z_2), $i = 2$

(c) (z_1, z_2), $i = 3$ (d) (z_1, z_3), $i = 1$

(e) (z_1, z_3), $i = 2$ (f) (z_1, z_3), $i = 3$

(g) (z_2, z_3), $i = 1$ (h) (z_2, z_3), $i = 2$

(i) (z_2, z_3), $i = 3$

Fig. 8 Evolution of the 95%-credible ellipses with respect to the iteration number. Blue continuous line: $d_y = 23$ with optimization of the block structure. Red dotted line: $d_y = 23$ without optimization of the block structure. Green dashed line: $d_y = 5$ without optimization of the block structure.

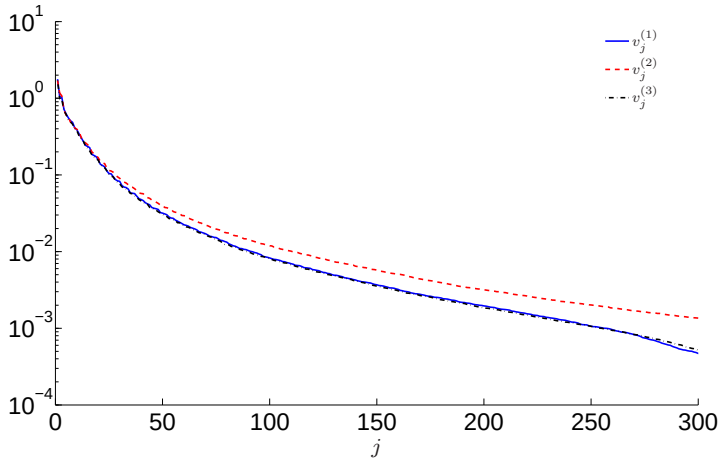
were chosen, introducing the partitioning does not make a big difference for the PDF identification, which explains the similarities between the blue and the red curves.

This set of figures also emphasizes the importance of considering a high value of d_y, even if it complicates the PDF identification. For instance, choosing $d_y = 5$ leads to the results in green dashed lines, which are clearly less relevant than the results in blue that correspond to $d_y = 23$. Intermediate results were obtained for values of d_y between 5 and 23, whereas still increasing d_y did not really change the results.

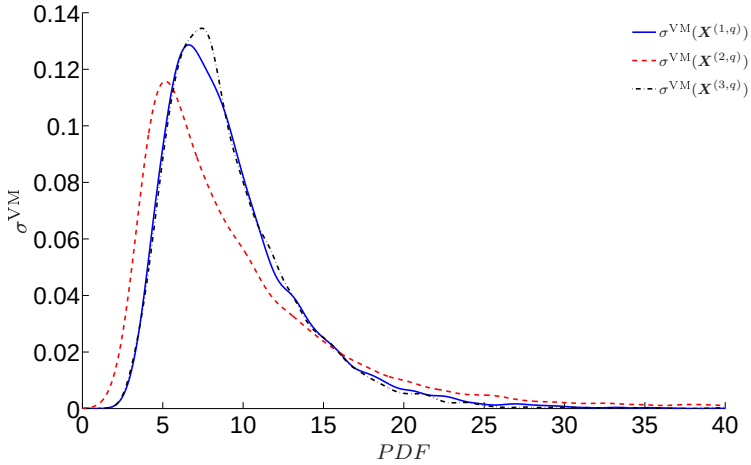
In order to emphasize the efficiency of the proposed method to recover the true underlying stochasticity, three additional batches of $Q = 10^4$ simulations are launched. These simulations are associated with the same cubic system than in Figure 6, but with different boundary conditions (by changing the boundary conditions, we want to verify that the identified values of $\boldsymbol{Z}$ are not dependent of a fixed configuration). While the boundary conditions on the inferior side of the cube do not change, the left and right sides are now free of constraints, and the displacements on the superior side are chosen equal to $0.002\boldsymbol{e}_1 - 0.01\boldsymbol{e}_2$. We then denote by $\left\{\boldsymbol{X}^{(1,q)}, 1 \leq q \leq Q\right\}$, $\left\{\boldsymbol{X}^{(2,q)}, 1 \leq q \leq Q\right\}$ and $\left\{\boldsymbol{X}^{(3,q)}, 1 \leq q \leq Q\right\}$ the elasticity fields characterizing the material properties of the different cubes of the three sets respectively. For all $1 \leq q \leq Q$,

- $\boldsymbol{X}^{(1,q)}$ is an independent realization of the true elasticity field, $\boldsymbol{X}(\boldsymbol{z}^\star)$;
- $\boldsymbol{X}^{(2,q)}$ is an independent realization of $\boldsymbol{X}((\boldsymbol{z}^{q,\text{prior}}, z_4^\star, z_5^\star))$, where $\boldsymbol{z}^{q,\text{prior}}$ is a realization of $\boldsymbol{Z}$, whose distribution is given by Eq. (29),
- $\boldsymbol{X}^{(3,q)}$ is an independent realization of $\boldsymbol{X}((\boldsymbol{z}^{q,\text{post}}, z_4^\star, z_5^\star))$, where $\boldsymbol{z}^{q,\text{post}}$ is a realization of $\boldsymbol{Z}|\mathcal{S}(N)$ after the three formerly presented iterations.

For each simulation, we denote by $\boldsymbol{U}(\boldsymbol{X}^{(i,q)})$, $1 \leq i \leq 3$, the concatenation of the vertical and horizontal displacements that are induced on the left and right sides of the cube. To compare the statistical information gathered in these displacements, we then compute, for each $1 \leq i \leq 3$, the eigenvalues $\left\{v_j^{(i)},\ j \geq 0\right\}$ associated with the empirical estimate of their covariance matrices. In addition, we denote by $\sigma^{\text{VM}}(\boldsymbol{X}^{(i,q)})$ the maximum value over the cubic domain of the Von Mises stress. This Von Mises criterion is commonly used to characterize the resistance of the system (see [Lai et al., 2010] for more details). The decrease of these eigenvalues and the PDF of these Von Mises criteria are finally compared in Figure 9. Looking at these figures, we see that the results associated with the posterior distribution of $\boldsymbol{Z}$ are very close to the ones associated with the true elasticity field, which is not true for the results associated with the prior distribution of $\boldsymbol{Z}$. This underlines the capacity of the proposed method to take into account indirect observations for the identifica-



(a) Eigenvalues



(b) Von Mises criterion

Fig. 9 Representation of the eigenvalues decreases (a) and of the PDFs of the Von Mises criteria (b) for the three compared configurations. In each figure, the blue continuous lines are associated with the reference case, the red dashed line correspond to the results associated with the prior distribution of $\boldsymbol{Z}$, when the black two-dashed lines correspond to the results associated with the posterior distribution of $\boldsymbol{Z}$.

tion of the parameters characterizing the distribution of an unknown random process of interest.

4 Conclusion

The increasing of the computational resources and the generalization of the monitoring of mechanical systems have encouraged many scientific fields to take into account random fields in their modeling. In that prospect, this work proposes an adaptive Bayesian framework to efficiently identify the statistical properties of these random fields when the available information is a reduced set of indirect observations. Two examples based on simulated data are finally presented to show the potential of this approach.

Extending this approach to the cases where the number of parameters to identify and the number of observations are very high would be interesting for future work.

Appendix

A.1. Proof of the equality of Eq. 9

Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}$ be the block decomposition matrices of $\widehat{\boldsymbol{C}}^{-1}$:

$$\widehat{\boldsymbol{C}}^{-1} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^T & \boldsymbol{D} \end{bmatrix}. \tag{30}$$

Using the Schur complement, if follows that:

$$\begin{cases} \widehat{\boldsymbol{C}}_{\boldsymbol{ZZ}}^{-1} = \boldsymbol{D} - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B}, \\ (\widehat{\boldsymbol{C}}_{\boldsymbol{YY}} - \widehat{\boldsymbol{C}}_{\boldsymbol{YZ}} \widehat{\boldsymbol{C}}_{\boldsymbol{ZZ}}^{-1} \widehat{\boldsymbol{C}}_{\boldsymbol{YZ}}^T)^{-1} = \boldsymbol{A}, \\ -\widehat{\boldsymbol{C}}_{\boldsymbol{YZ}} \widehat{\boldsymbol{C}}_{\boldsymbol{ZZ}}^{-1} = \boldsymbol{A}^{-1} \boldsymbol{B}. \end{cases} \tag{31}$$

It comes

$$\begin{aligned} &((\boldsymbol{y}, \boldsymbol{z}) - (\boldsymbol{Y}(\omega_m), \boldsymbol{Z}(\omega_m)))^T (h^2 \widehat{\boldsymbol{C}})^{-1} ((\boldsymbol{y}, \boldsymbol{z}) - (\boldsymbol{Y}(\omega_m), \boldsymbol{Z}(\omega_m))) \\ &= \frac{1}{h^2} \begin{pmatrix} (\boldsymbol{y} - \boldsymbol{Y}(\omega_m))^T \boldsymbol{A} (\boldsymbol{y} - \boldsymbol{Y}(\omega_m)) + 2(\boldsymbol{y} - \boldsymbol{Y}(\omega_m))^T \boldsymbol{A} \boldsymbol{A}^{-1} \boldsymbol{B} (\boldsymbol{z} - \boldsymbol{Z}(\omega_m)) \\ + (\boldsymbol{z} - \boldsymbol{Z}(\omega_m))^T \boldsymbol{D} (\boldsymbol{z} - \boldsymbol{Z}(\omega_m)) \end{pmatrix} \\ &= \frac{1}{h^2} \begin{pmatrix} (\boldsymbol{y} - \boldsymbol{Y}(\omega_m) + \boldsymbol{A}^{-1} \boldsymbol{B} (\boldsymbol{z} - \boldsymbol{Z}(\omega_m)))^T \boldsymbol{A} (\boldsymbol{y} - \boldsymbol{Y}(\omega_m) + \boldsymbol{A}^{-1} \boldsymbol{B} (\boldsymbol{z} - \boldsymbol{Z}(\omega_m))) \\ + (\boldsymbol{z} - \boldsymbol{Z}(\omega_m))^T \left(\boldsymbol{D} - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B} \right) (\boldsymbol{z} - \boldsymbol{Z}(\omega_m)) \end{pmatrix} \\ &= (\boldsymbol{y} - \boldsymbol{\mu}_n(\boldsymbol{z}))^T \boldsymbol{C}_n^{-1} (\boldsymbol{y} - \boldsymbol{\mu}_n(\boldsymbol{z})) + \frac{1}{h^2} (\boldsymbol{z} - \boldsymbol{Z}(\omega_m))^T \boldsymbol{C}_{\boldsymbol{ZZ}}^{-1} (\boldsymbol{z} - \boldsymbol{Z}(\omega_m)) \end{aligned} \tag{32}$$

This leads to the searched result.

References

Akaike, 1974. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6):716–723.

Arnst et al., 2010. Arnst, M., Ghanem, R., and Soize, C. (2010). Identification of bayesian posteriors for coefficients of chaos expansions. *Journal of Computational Physics*, 229 (9):3134–3154.

Atwell and King, 2001. Atwell, J. and King, B. (2001). Proper orthogonal decomposition for reduced basis feedback controllers for parabolic equations. *Math. Comput. Modell.*, 33 (1-3):1–19.

Auffray et al., 2012. Auffray, Y., Barbillon, P., and Marin, J. M. (2012). Maximin design on non hypercube domains and kernel interpolation. *Statistics and Computing*, 22(3):703–712.

Bilionis and Zabaras, 2015. Bilionis, I. and Zabaras, N. (2015). Bayesian uncertainty propagation using gaussian processes. *In: Ghanem R., Higdon D., Owhadi H. (eds) Handbook of Uncertainty Quantification. Springer.*

Box and Jenkins, 1970. Box, G. and Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

Chen and Schwab, 2015. Chen, P. and Schwab, C. (2015). Sparse-grid, reduced basis bayesian inversion. *Computer Methods in Applied Mechanics and Engineering*, 297:84–115.

Clouteau et al., 2013. Clouteau, D., Cottereau, R., and Lombaert, G. (2013). Dynamics of structures coupled with elastic media - a review of numerical models and methods. *J. Sound Vibr.*, 332:2415–2436.

Conrad et al., 2018. Conrad, P. R., Davis, A., Marzouk, Y. M., Pillai, N. S., and Smith, A. (2018). Parallel local approximation MCMC for expensive models. *SIAM/ASA J. Uncertainty Quantification*, 6(1):39–373.

Conrad et al., 2016. Conrad, P. R., Marzouk, Y. M., Pillai, N. S., and Smith, A. (2016). Accelerating asymptotically exact MCMC for computationally intensive models via local approximations. *Journal of the American Statistical Association*, 111:1591–1607.

Damblin et al., 2013. Damblin, G., Barbillonz, P., Keller, M., Pasanisi, A., and Parent, E. (2013). Adaptive Numerical Designs for the Calibration of Computer Codes. *SIAM/ASA J. Uncertainty Quantification*, 6(1):151–179.

Draguljić et al., 2012. Draguljić, D., Santner, T. J., and Dean, A. M. (2012). Noncollapsing Space-Filling Designs for Bounded Nonrectangular Regions. *Technometrics*, 54(2):169–178.

Emery et al., 2016. Emery, J., Grigoriu, M., and Jr., R. F. (2016). Bayesian methods for characterizing unknown parameters of material models. *Applied Mathematical Modelling*, 13-14:6395–6411.

Fang et al., 2006. Fang, K., Li, R., and Sudjianto, A. (2006). *Design and modeling for computer experiments.* Chapman & Hall, Computer Science and Data Analysis Series, London.

Fang and Lin, 2003. Fang, K. and Lin, D. (2003). Uniform experimental designs and their applications in industry. *Handbook of Statistics*, 22:131–178.

Fielding et al., 2011. Fielding, M., Nott, D. J., and Liong, S. Y. (2011). Efficient MCMC schemes for Computationally Expensive Posterior Distributions. *Technometrics*, 53(1):16–28.

Ghanem and Spanos, 2003. Ghanem, R. and Spanos, P. D. (2003). *Stochastic Finite Elements: A Spectral Approach, rev. ed.* Dover Publications, New York.

Guilleminot and Soize, 2013. Guilleminot, J. and Soize, C. (2013). On the statistical dependence for the components of random elasticity tensors exhibiting material symmetry properties. *Journal of Elasticity*, 111:109–130.

Higdon et al., 2008. Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008). Computer model calibration using high-dimensional output. *Journal of the American Statistical Association*, 103(482):570–583.

Higdon et al., 2003. Higdon, D., Lee, H., and Holloman, C. (2003). Markov chain monte carlo based approaches for inference in computationally intensive inverse problems. *Bayesian Statistics*, 7:181–197.

Joseph et al., 2015. Joseph, V. R., Gul, E., and Ba, S. (2015). Maximum projection designs for computer experiments. *Biometrika*, 102(2):371–380.

Kennedy and O'Hagan, 2001. Kennedy, M. and O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the royal statistical society*, 63:425–464.

Lai et al., 2010. Lai, W. M., Rubin, D., and Krempl, E. (2010). *Introduction to Continuum Mechanics*. Elsevier, Inc.

Le Maître and Knio, 2010. Le Maître, O. and Knio, O. (2010). *Spectral Methods for Uncertainty Quantification*. Springer.

Lekivetz and Jones, 2015. Lekivetz, R. and Jones, B. (2015). Fast Flexible Space-Filling Designs for Nonrectangular Regions. *Quality and Reliability Engineering International*, 31(5):829–837.

Li and Marzouk, 2014. Li, J. and Marzouk, Y. M. (2014). Adaptive construction of surrogates for the bayesian solution of inverse problems. *SIAM Journal on Scientific Computing*, 36:A1163–A1186.

Mak and Joseph, 2016. Mak, S. and Joseph, V. R. (2016). Minimax designs using clustering. pages 1–24.

Marin and Robert, 2007. Marin, J. M. . and Robert, C. P. (2007). *Bayesian core*. Springer-Verlag, New York.

Marzouk and Najm, 2009. Marzouk, Y. M. and Najm, H. N. (2009). Dimensionality reduction and polynomial chaos acceleration of bayesian inference in inverse problems. *Journal of Computational Physics*, 228 (6):1862–1902.

Marzouk and Xiu, 2009. Marzouk, Y. M. and Xiu, D. (2009). A stochastic collocation approach to bayesian inference in inverse problems. *Communications in Computational Physics*, 6:826–847.

Matthies et al., 2016. Matthies, H., Zander, E., Rosi, B., and Litvinenko, A. (2016). Parameter estimation via conditional expectation: a Bayesian inversion. *Adv. Model. and Simul. in Eng. Sci.*, 3(24).

McKay et al., 1979. McKay, M., Beckman, R., and Conover, W. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21:239–245.

Nguyen et al., 2015. Nguyen, M. T., Desceliers, C., Soize, C., Allain, J., and Gharbi, H. (2015). Multiscale identification of random elasticity field at mesoscale of a heterogeneous microstructure using multiscale experimental observations. *Journal for Multiscale Computational Engineering*, 13 (4):281–295.

Nouy, 2010. Nouy, A. (2010). Proper generalized decomposition and separated representations for the numerical solution of high dimensional stochastic problems. *Archives of computational methods in engineering*, 17:403–434.

Nouy and Soize, 2014. Nouy, A. and Soize, C. (2014). Random fields representations for stochastic elliptic boundary value problems and statistical inverse problems. *Eur. J. Appl. Math.*, 25:339–373.

Perrin, ress. Perrin, G. (in press). Adaptive calibration of a computer code with time-series output. *Reliability Engineering and System Safety*.

Perrin and Cannamela, 2017. Perrin, G. and Cannamela, C. (2017). A repulsion-based method for the definition and the enrichment of opotimized space filling designs in constrained input spaces. *Journal de la Société Française de Statistique*, 158(1):37–67.

Perrin et al., 2012. Perrin, G., Soize, C., Duhamel, D., and Funfschilling, C. (2012). Identification of polynomial chaos representations in high dimension from a set of realizations. *SIAM J. Sci. Comput.*, 34(6):2917–2945.

Perrin et al., 2013. Perrin, G., Soize, C., Duhamel, D., and Funfschilling, C. (2013). Karhunen–loève expansion revisited for vector-valued random fields: Scaling, errors and optimal basis. *Journal of Computational Physics*, 242:607–622.

Perrin et al., 2014. Perrin, G., Soize, C., Duhamel, D., and Funfschilling, C. (2014). A posteriori error and optimal reduced basis for stochastic processes defined by a finite set of realizations. *SIAM/ASA J. Uncertainty Quantification*, 2:745–762.

Perrin et al., 2017. Perrin, G., Soize, C., Marque-Pucheu, S., and Garnier, J. (2017). Nested polynomial trends for the improvement of gaussian process-based predictors. *Journal of Computational Physics*, 346:389–402.

Perrin et al., 2018. Perrin, G., Soize, C., and Ouhbi, N. (2018). Data-driven kernel representations for sampling with an unknown block dependence structure under correlation constraints. *Journal of Computational Statistics and Data Analysis*, 119:139–154.

Rasmussen, 2003. Rasmussen, C. E. (2003). Gaussian processes to speed up hybrid monte carlo for expensive bayesian integrals. *Bayesian Statistics*, 7:651–659.

Rubinstein and Kroese, 2008. Rubinstein, R. T. and Kroese, D. (2008). *Simulation and the Monte Carlo method*. John Wiley and Sons, Inc., Hoboken, New Jersey.

Santner et al., 2003. Santner, T. J., Williams, B., and Notz, W. (2003). *The design and analysis of computer experiments*. Springer, New York.

Scott and Sain, 2004. Scott, D. W. and Sain, S. R. (2004). Multidimensional Density Estimation.

Sinsbeck and Nowak, 2017. Sinsbeck, M. and Nowak, W. (2017). Sequential Design of Computer Experiments for the Solution of Bayesian Inverse. *SIAM/ASA J. Uncertainty Quantification*, 5:640–664.

Soize, 2006. Soize, C. (2006). Non-gaussian positive-definite matrix-valued random fields for elliptic stochastic partial differential operators. *Comput. Methods Appl. Mech. Eng.*, 195:26–64.

Soize, 2008. Soize, C. (2008). Tensor-valued random fields for meso-scale stochastic model of anisotropic elastic microstructure and probabilistic analysis of representative volume element size. *Probab. Eng. Mech.*, 23:307–323.

Soize, 2010. Soize, C. (2010). Identification of high-dimension polynomial chaos expansions with random coefficients for non-Gaussian tensor-valued random fields using partial and limited experimental data. *Computer Methods in Applied Mechanics and Engineering*, 199(33-36):2150–2164.

Soize, 2011. Soize, C. (2011). A computational inverse method for identification of non-Gaussian random fields using the Bayesian approach in very high dimension. *Computer Methods in Applied Mechanics and Engineering*, 200(45–46):3083–3099.

Soize and Ghanem, 2016. Soize, C. and Ghanem, R. (2016). Data-driven probability concentration and sampling on manifold. *Journal of Computational Physics*, 321(September 2015):242–258.

Soize and Ghanem, 2017. Soize, C. and Ghanem, R. (2017). Probabilistic learning on manifold for optimization under uncertainties. *Proceeding of Uncecomp 2017*, pages 1–15.

Stinstra et al., 2003. Stinstra, E., den Hertog, D., Stehouwer, P., and Vestjens, A. (2003). Constrained maximin designs for computer experiments. *Technometrics*, 45(4):340–346.

Stinstra et al., 2010. Stinstra, E., den Hertog, D., Stehouwer, P., and Vestjens, A. (2010). Uniform designs over general input domains with applications to target region estimation in computer experiments. *Computational Statisticsand Data Analysis*, 51(1):219–232.

Stuart, 2010. Stuart, A. M. (2010). Inverse problems: A bayesian perspective. *Acta Numerica*, 19:451–559.

Tian et al., 2016. Tian, M., Li, D., Cao, Z., Phoon, K., and Wang, Y. (2016). Bayesian identification of random field model using indirect test data. *Engineering Geology*, 210:197–211.

Tsilifis et al., 2017. Tsilifis, P., Ghanem, R. G., and Hajali, P. (2017). Efficient Bayesian Experimentation Using an Expected Information Gain Lower Bound. *SIAM/ASA J. Uncertainty Quantification*, 5:30–62.

Wan and Zabaras, 2011. Wan, J. and Zabaras, N. (2011). A Bayesian approach to multiscale inverse problems using the sequential Monte Carlo method. *Inverse Problems*, 27.

Wand and Jones, 1995. Wand, M. P. and Jones, M. C. (1995). Kernel Smoothing. *Encyclopedia of Statistics in Behavioral Science*, 60(60):212.

Whittle, 1951. Whittle, P. (1951). *Hypothesis testing in time series, PhD thesis*. University of Uppsala.

Whittle, 1983. Whittle, P. (1983). *Prediction and Regulation by Linear Least-Square Methods*. University of Minnesota Press.

Williams, 2011. Williams, M. (2011). The eigenfunctions of the karhunen-loeve integral equation for a spherical system. *Propabilistic Engineering Mechanics*, 26:202–207.