



Approximate solutions of Lagrange multipliers for information-theoretic random field models

B Staber, Johann Guilleminot

► To cite this version:

B Staber, Johann Guilleminot. Approximate solutions of Lagrange multipliers for information-theoretic random field models. SIAM/ASA Journal on Uncertainty Quantification, 2015, pp.1-23. hal-01166830

HAL Id: hal-01166830

<https://hal.science/hal-01166830>

Submitted on 23 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approximate solutions of Lagrange multipliers for information-theoretic random field models*

B.STABER

J.GUILLEMINOT [†]

Abstract

This work is concerned with the construction of approximate solutions for the Lagrange multipliers involved in information-theoretic non-Gaussian random field models. Specifically, representations of physical fields with invariance properties under some orthogonal transformations are considered. A methodology for solving the optimization problems raised by entropy maximization (for the family of first-order marginal probability distributions) is first presented and exemplified in the case of elasticity fields exhibiting fluctuations in a given symmetry class. Results for all classes ranging from isotropy to orthotropy are provided and discussed. The derivations are subsequently used for proving a few properties that are required in order to sample the above models by solving a family of stochastic differential equations – along the lines of the algorithm constructed in [9]. The results thus allow for forward simulations of the probabilistic models in stochastic boundary value problems, as well as for a reduction of the computational cost associated with model calibration through statistical inverse problems.

Keywords: Elasticity; information-theoretic stochastic model; maximum entropy principle; random field; stochastic elliptic boundary value problem; uncertainty quantification.

1 Preliminaries

1.1 Introduction

Uncertainty quantification involving stochastic boundary value problems (SBVP) has received much attention during the past two decades in both computational mechanics and applied mathematics. Since the seminal work from Ghanem and Spanos [6], most of the related contributions have been devoted to the construction of efficient stochastic (Galerkin-type intrusive or non-intrusive) solvers for uncertainty propagation and rely on functional representations (such as polynomial chaos expansions [4]) of random variables and random fields (see *e.g.* [16]). As far as physical modeling and inverse identification are concerned, computational simulations based on such spectral approaches deeply depend on the modeling of the uncertainties, and the *a priori* selection of probability laws critically impacts the robustness and fidelity of the analysis. In this prospect, information-theoretic stochastic models [13, 14] have been proposed by invoking Shannon’s maximum entropy principle [21] and promoted as optimal prior choices for

*Preprint version, accepted for publication in SIAM/ASA Journal on Uncertainty Quantification on May 26, 2015. This work was supported by the French National Research Agency (ANR) (MOSAIC project, ANR-12-JS09-0001-01).

[†]Université Paris-Est Marne-la-Vallée, Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, 5 Bd Descartes, 77545 Marne-la-Vallée, France.

the modeling of random data and uncertainties. Fruitful applications of this framework can be found in [1, 2] in nuclear physics, in the pioneering work by Soize [22] for the modeling of uncertainties in structural dynamics (see [26] and the references therein for a thorough synthesis) or in [24, 9] for the representation of tensor-valued coefficients for stochastic elliptic operators, to name a few.

In particular, the construction of tensor-valued random field representations modeling the coefficients of stochastic elliptic operators is of great importance and finds applications in a broad range of engineering or scientific fields (such as biomedical engineering, geophysics or wave propagation through random media for instance). In addition to some usual mathematical properties (*e.g.* positive-definiteness) preserving the well-posedness of the associated SBVP, such random fields typically exhibit some invariance properties under the action of a given group of isometries. The algebraic structure of the state space induced by these constraints then makes the above stochastic models depend on a set of Lagrange multipliers whose expressions cannot be obtained explicitly. Consequently, forward simulations of these physically-sound random fields are not readily at hand and require the solving of a set of constrained optimization problems as a preprocessing step. In this work, and without loss of methodological generality, we consider the class of non-Gaussian elasticity random fields constructed in [9] and exhibiting some invariance properties under the action of a group of rotations. The aim is twofold. First, we address the construction of an approximate solution for the aforementioned Lagrange multipliers, for all material symmetry classes of practical interest (that is, for the six symmetry classes from isotropy to orthotropy). Second, we make use of the general form of these solutions to demonstrate some properties allowing the non-Gaussian random fields to be generated by solving a family of Itô stochastic differential equations (ISDEs) - following the sampling scheme proposed in [9].

This paper is organized as follows. The stochastic model for tensor-valued non-Gaussian elasticity random fields is first recalled in § 2. The methodology for solving the optimization problem raised by the maximum entropy formulation is then presented in § 3. The results for all symmetry classes, together with a specific application to the case of transversely isotropic materials, are provided. Some mathematical properties satisfied by the approximate solution are further investigated in § 4, where the ISDEs-based sampling scheme is also recalled for the sake of completeness. An application to a random field exhibiting the orthotropic symmetry (and related to composite modeling) is finally provided.

1.2 Notation

Throughout this paper, use will be made of the following matrix sets:

- i. $\mathbb{M}_n(\mathbb{R})$ the set of real $(n \times n)$ matrices,
- ii. $\mathbb{M}_n^S(\mathbb{R})$ the set of real $(n \times n)$ symmetric matrices,
- iii. $\mathbb{M}_n^+(\mathbb{R})$ the set of real $(n \times n)$ symmetric and positive definite matrices,
- iv. $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ the set of real $(n \times n)$ symmetric and positive definite matrices defined by the symmetry group $\mathcal{O}^{\text{sym}}$ of all isometries leaving elements of $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ invariant by composition (see [29]).

Let $[I_n]$ and $[0_n]$ be the identity and zero n -by- n matrices respectively, and let $\mathbf{0}_n$ denote the null vector of length n . Fourth-order and second-order tensors, vectors and scalars are denoted by $\llbracket A \rrbracket$, $[A]$, $\mathbf{a}$ and a (resp. by $\llbracket \mathbf{A} \rrbracket$, $[\mathbf{A}]$, $\mathbf{A}$ and A) in a deterministic (resp. stochastic) framework. For any $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, let $\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^n x_i y_i$ and $\|\mathbf{x}\|_{\mathbb{R}^n} := \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$ be the the Euclidean scalar product and associated norm, respectively. The matrix scalar product is denoted by $\ll [A], [B] \gg := \text{Tr}([B]^T [A])$ for all $([A], [B]) \in$

$\mathbb{M}_n(\mathbb{R}) \times \mathbb{M}_n(\mathbb{R})$ and defines the Frobenius norm $\|[A]\|_F := \ll [A], [A] \gg^{1/2}$. The level of statistical fluctuations of any random matrix $[A]$ is characterized by the scalar parameter $\delta_{[A]}$ defined as:

$$\delta_{[A]} := \sqrt{\mathbb{E}\{\|[A] - [A]\|_F^2\} / \|[A]\|_F^2},$$

where $\mathbb{E}$ stands for the mathematical expectation and $[A] = \mathbb{E}\{[A]\}$. Similarly, we denote by $\underline{a}$ and $\underline{a}$ the mean values of random variables A and A , respectively; notation σ_A represents the standard deviation of random variable A . Finally, c denotes a normalization constant whose value may change from line to line.

2 Stochastic modeling of $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ -valued random fields

Let $\{[C(\mathbf{x})], \mathbf{x} \in \Omega\}$ be a $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ -valued second-order random field, defined on a probability space $(\Theta, \mathcal{T}, P)$ and indexed by an open bounded domain $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, $n \in \{3, 6\}$. Following [9], let $\{[N(\mathbf{x})], \mathbf{x} \in \Omega\}$ be the auxiliary $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ -valued random field such that

$$\forall \mathbf{x} \in \Omega, \quad [C(\mathbf{x})] := [\underline{C}(\mathbf{x})]^{1/2} [N(\mathbf{x})] [\underline{C}(\mathbf{x})]^{1/2}, \quad (2.1)$$

with $[\underline{C}(\mathbf{x})] = \mathbb{E}\{[C(\mathbf{x})]\}$, and thus

$$\forall \mathbf{x} \in \Omega, \quad \mathbb{E}\{[N(\mathbf{x})]\} = [I_n]. \quad (2.2)$$

Let $\{[E_i]\}_{i=1}^N$ be the basis of matrix set $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$, with $2 \leq N \leq 21$ in three-dimensional linear elasticity ($n = 6$), defined in [29]. In the context of elasticity, the value of N depends on the symmetry class under consideration. In accordance with usual terminology in micromechanics, this basis will be referred to as a Walpole basis in the sequel. It can then be shown that for any $\mathbf{x}$ fixed in Ω , there exists a unique symmetric random matrix $[G(\mathbf{x})]$ such that

$$[N(\mathbf{x})] := \text{expm}([G(\mathbf{x})]), \quad (2.3)$$

with expm the matrix exponential, and

$$[G(\mathbf{x})] = \sum_{i=1}^N G_i(\mathbf{x}) [E_i], \quad (2.4)$$

hence defining a $\mathbb{R}^N$ -valued random field $\{G(\mathbf{x}) := (G_1(\mathbf{x}), \dots, G_N(\mathbf{x})), \mathbf{x} \in \Omega\}$ of coordinates on $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$. It should be emphasized that the relaxation of the support constraint (for the coordinates on the matrix basis) induced by the exponential mapping is a key ingredient for the selected sampling scheme (see section § 4.2), and that this property does not hold for square-type (*e.g.* Cholesky) representations (see [10] for a sampling scheme corresponding to this case).

2.1 Information-theoretic probabilistic model for random field $\{[G(\mathbf{x})], \mathbf{x} \in \Omega\}$

The system of marginal probability distributions of random field $\{G(\mathbf{x}), \mathbf{x} \in \Omega\}$ completely defines the one of random field $\{[G(\mathbf{x})], \mathbf{x} \in \Omega\}$ (see Eq. (2.4)). Let $\{g \mapsto p_{G(\mathbf{x})}(g)\}_{\mathbf{x} \in \Omega}$ be the family of first-order marginal probability density functions of $\{G(\mathbf{x}), \mathbf{x} \in \Omega\}$. For simplicity, we assume below that the above family is independent of $\mathbf{x}$ (this assumption can be readily relaxed at the expense of notational complexity). The methodology for deriving the random field model essentially consists

- in constructing the density $p_{\mathbf{G}(\mathbf{x})}$ through the maximization of Shannon's differential entropy, given some available information on random matrix $[\mathbf{G}(\mathbf{x})]$ – following Jaynes' framework of information theory;
- in defining $\mathbf{G}(\mathbf{x})$ as

$$\mathbf{G}(\mathbf{x}) := \mathcal{H}(\{\mathbf{W}(r, \mathbf{x}), r \geq 0\}) , \quad (2.5)$$

for any $\mathbf{x}$ fixed in Ω , where $\mathcal{H}$ is a non-linear operator (which depends on $\mathbf{x}$ whenever $p_{\mathbf{G}(\mathbf{x})}$ does) and $\{\mathbf{W}(r, \mathbf{x}), r \geq 0\}$ is a standard $\mathbb{R}^N$ -valued Wiener process such that $\mathbf{G}(\mathbf{x})$ exhibits the probability density function defined above.

Such a construction requires a suitable definition for the (Gaussian) random field $\{\mathbf{W}(r, \mathbf{x}), r \geq 0, \mathbf{x} \in \mathbb{R}^d\}$ (see [9] for details, as well as section 4.2.1 for a summary), and the transformation defined by operator $\mathcal{H}$ then defines the complete system of marginal laws for $\{\mathbf{G}(\mathbf{x}), \mathbf{x} \in \Omega\}$. As regards the MaxEnt formulation, the following constraints are considered:

$$\mathbb{E} \left\{ \expm \left(\sum_{i=1}^N G_i(\mathbf{x})[E_i] \right) \right\} = [I_n] , \quad (2.6a)$$

$$\mathbb{E} \left\{ \log \left(\det \left(\expm \left(\sum_{i=1}^N G_i(\mathbf{x})[E_i] \right) \right) \right) \right\} = \nu , \quad |\nu| < +\infty . \quad (2.6b)$$

The first constraint is a direct consequence of Eqs. (2.2–2.3), whereas the second one implies that $[\mathbf{N}(\mathbf{x})]$ and $[\mathbf{N}(\mathbf{x})]^{-1}$ are both second-order matrix-valued random variables (as shown in [22]). Note that the combination of the above equations, together with Jensen's inequality, basically shows that $\nu \leq 0$. The MaxEnt probability density function $p_{\mathbf{G}(\mathbf{x})}$ then reads as [9]:

$$p_{\mathbf{G}(\mathbf{x})}(\mathbf{g}) = c \exp \left(- \ll [\Lambda^{\text{sol}}], \expm \left(\sum_{i=1}^N g_i[E_i] \right) \gg - \lambda^{\text{sol}} \sum_{i=1}^N g_i \text{Tr}([E_i]) \right) , \quad (2.7)$$

where $[\Lambda^{\text{sol}}]$ and λ^{sol} are the unknown Lagrange multipliers (raised by the MaxEnt formulation) such that the two constraints (2.6a) and (2.6b) are satisfied. Clearly, $[\Lambda^{\text{sol}}]$ belongs to a N -dimensional matrix set. In addition, it is assumed that it can be decomposed as $[\Lambda^{\text{sol}}] = \sum_{i=1}^N \lambda_i^{\text{sol}}[E_i]$, so that an independent scalar-valued Lagrange multiplier λ_i^{sol} , $1 \leq i \leq N$, can be assigned to the i -th coordinate involved in the decomposition on the matrix basis. In the sequel, we will consider the vector-valued representation $\boldsymbol{\lambda}^{\text{sol}} := (\lambda_1^{\text{sol}}, \dots, \lambda_N^{\text{sol}}, \lambda^{\text{sol}})$ of the solution Lagrange multipliers.

3 Calculation of the Lagrange multipliers

3.1 Overview of the methodology

Apart from the isotropic and cubic symmetry classes, for which explicit solutions can be easily constructed thanks to the particular properties of the matrix basis, no explicit solutions can be obtained for the Lagrange multiplier $\boldsymbol{\lambda}^{\text{sol}}$ involved in the above MaxEnt formulation (see Eq. (2.7)). In order to tackle this issue, a methodology relying on sequential optimization problems is proposed in order to obtain admissible values of the parameter ν (which does not belong to $\mathbb{R}$) involved in the second constraint equation (see

Eq. (2.6b)), as well as to derive suitable cost functions for the computation of the Lagrange multipliers. The numerical strategy basically proceeds along three main steps that are described below (for a given material symmetry class).

- *Step 1: Determination of a sequence $\{\nu_i\}_{i=1}^{m_\nu}$ of admissible values for parameter ν .* This first step is concerned with the definition of a sequence of well-posed optimization problems, the well-posedness being understood here in the sense of sampling parameter ν in its admissible set $\mathbb{R}_-$ (see Eq. (2.6b)), in accordance with the underlying structure of $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ (see section § 3.2.1 for an illustration). This sampling is achieved by having recourse to a projection procedure that is detailed in § 3.1.1.
- *Step 2: Construction of optimization problems for the Lagrange multiplier λ^{sol} .* In this step, we first introduce an optimization problem related to the constraints defined by Eqs. (2.6), for each given value of ν determined in Step 1. The algebraic form of a solution guess is subsequently obtained and allows one to solve an alternative optimization problem depending on a single parameter. These optimization problems are detailed in § 3.1.2 and § 3.1.3, respectively.
- *Step 3: Construction of an interpolation $\delta_{[N]} \mapsto \lambda^{\text{sol}}$.* In this final step, a simple interpolation relating $\delta_{[N]}$ (which is the parameter characterizing the level of statistical fluctuations of random matrix $[N]$; recall that $\delta_{[N]}$ is assumed to be independent of $\mathbf{x}$ for simplicity) to λ^{sol} is constructed, hence facilitating the practical use of the constructed random field model (regardless of the material symmetry class). This construction relies on the one-to-one mapping between parameters ν and $\delta_{[N]}$ (see *e.g.* [22]).

A detailed application of the methodology for the transversely isotropic symmetry is provided in section § 3.2.1. The results for all other classes are then given in § 3.2.3.

3.1.1 Definition of admissible values of ν through an Euclidean projection

Let us consider a given material symmetry class $\mathbb{M}_n^{\text{sym}}$, as well as a particular related crystallographic system (if need be). When a triclinic system (which corresponds to the case $\mathbb{M}_n^{\text{sym}} = \mathbb{M}_n^+(\mathbb{R})$) is considered, the stochastic model for random matrix $[N]$ turns out to correspond to the information-theoretic random matrix class for fully anisotropic random matrices, named SG^+ , constructed in [22] [23]. In this particular case, an exact algebraic solution for the Lagrange multipliers and a random generator exist (see the above references). Consequently, samples of a random matrix with mean value $[I_n]$ and exhibiting triclinic fluctuations (say, $[N^{\text{triclinic}}(\delta^{\text{triclinic}})]$, where $\delta^{\text{triclinic}}$ refers to the parameter measuring the level of fluctuations) can be easily obtained and projected on $\mathbb{M}_n^{\text{sym}}$, hence providing realizations of a random matrix (say, $[N^{\text{sym}}(\delta^{\text{sym}})]$) with fluctuations in the symmetry class under consideration. Clearly, the parameter δ^{sym} depends on $\delta^{\text{triclinic}}$. Note further that this mapping is not one-to-one, so that it can not be considered as a valuable alternative for defining, through a measure transformation, $\mathbb{M}_n^{\text{sym}}$ -valued random matrices. However, a set of M_S realizations of $[N^{\text{sym}}(\delta^{\text{sym}})]$ can be obtained as

$$[N^{\text{sym}}(\delta^{\text{sym}}, \omega_k)] := \underset{[A] \in \mathbb{M}_n^{\text{sym}}}{\operatorname{argmin}} \mathcal{D}([A], [N^{\text{triclinic}}(\delta^{\text{triclinic}}, \omega_k)]) , \quad 1 \leq k \leq M_S , \quad (3.1)$$

where $\mathcal{D}$ is the Euclidean distance defined by

$$\mathcal{D}([A_1], [A_2]) := \|[A_1] - [A_2]\|_F , \quad \forall ([A_1], [A_2]) \in \mathbb{M}_n^+(\mathbb{R}) \times \mathbb{M}_n^+(\mathbb{R}) , \quad (3.2)$$

the argument ω_k denoting the k -th realization on both sides of Eq. (3.1). Explicit solutions for the above projection problem can be found in *e.g.* [19]. The strategy can then be summarized as follows:

1. Select m_ν equally spaced values of $\delta^{\text{triclinic}}$ in the admissible range $[0, \sqrt{(n+1)/(n+5)}]$.
2. For each selected value $\delta_i^{\text{triclinic}}$, $1 \leq i \leq m_\nu$,
 - a. perform Monte-Carlo simulations to obtain the set $\{[N^{\text{triclinic}}(\delta_i^{\text{triclinic}}, \omega_k)]\}_{k=1}^{M_S}$ of realizations of the triclinic random matrix $[N^{\text{triclinic}}(\delta_i^{\text{triclinic}})]$;
 - b. compute the corresponding set of realizations of $[N^{\text{sym}}(\delta_i^{\text{sym}})]$ using Eq. (3.1);
 - c. compute the value of ν_i using a classical statistical estimator:

$$\nu_i \simeq \frac{1}{M_S} \sum_{k=1}^{M_S} \log(\det([N^{\text{sym}}(\delta_i^{\text{sym}}, \omega_k)])) .$$

Such a methodology allows for the definition of well-posed optimization problems related to the constraints given by Eqs. (2.6), as discussed in the next section.

3.1.2 Construction of an optimization problem for the computation of the Lagrange multipliers

Let Φ be the potential function defined from $\mathbb{R}^N$ into $\mathbb{R}$ by:

$$\Phi_{\boldsymbol{\lambda}}(\mathbf{u}) := \ll \sum_{i=1}^N \lambda_i [E_i], \expm \left(\sum_{j=1}^N u_j [E_j] \right) \gg + \lambda \sum_{k=1}^N u_k \text{Tr}([E_k]) , \quad (3.3)$$

with $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_N, \lambda) \in \mathbb{D}_{\boldsymbol{\lambda}}$. The admissible set $\mathbb{D}_{\boldsymbol{\lambda}} \subset \mathbb{R}^{N+1}$ is such that the function $\mathbf{u} \mapsto \exp\{-\Phi_{\boldsymbol{\lambda}}(\mathbf{u})\}$ is integrable over $\mathbb{R}^N$ for any $\boldsymbol{\lambda} \in \mathbb{D}_{\boldsymbol{\lambda}}$. For any $\boldsymbol{\lambda}$ fixed in $\mathbb{D}_{\boldsymbol{\lambda}}$, the above potential function allows for the introduction of a $\mathbb{R}^N$ -valued random variable $\mathbf{Z}_{\boldsymbol{\lambda}}$ defined by the following probability density function:

$$p_{\boldsymbol{\lambda}}(\mathbf{z}) := c \exp\{-\Phi_{\boldsymbol{\lambda}}(\mathbf{z})\} . \quad (3.4)$$

It follows that

$$\mathbf{G} = \mathbf{Z}_{\boldsymbol{\lambda}^{\text{sol}}} \quad (3.5)$$

in probability distribution, in view of Eq. (2.7).

Let ν_i be a given value of parameter ν obtained through the strategy detailed in section § 3.1.1, $1 \leq i \leq m_\nu$. Let $\mathcal{J}_\alpha : \mathbb{D}_{\boldsymbol{\lambda}} \rightarrow \mathbb{R}^+$ be the cost function given by

$$\mathcal{J}_\alpha(\boldsymbol{\lambda}) := \left(\frac{1-\alpha}{n} \right) \|\mathbb{E}\{[\mathbf{Q}_{\boldsymbol{\lambda}}]\} - [I_n]\|_F^2 + \left(\frac{\alpha}{\nu_i^2} \right) (\mathbb{E}\{\log(\det([\mathbf{Q}_{\boldsymbol{\lambda}}]))\} - \nu_i)^2 , \quad (3.6)$$

in which

$$[\mathbf{Q}_{\boldsymbol{\lambda}}] := \expm \left(\sum_{i=1}^N Z_{\boldsymbol{\lambda}i} [E_i] \right) \quad (3.7)$$

and $\mathbf{Z}_{\boldsymbol{\lambda}} := (Z_{\boldsymbol{\lambda}1}, \dots, Z_{\boldsymbol{\lambda}N})$ is the random variable defined by Eqs. (3.3-3.4). In Eq. (3.6), $\alpha \in [0, 1]$ is a free parameter that allows one to numerically balance the respective contributions of the two constraints (note that the latter are not conflicting, since the two terms are respectively related to the mean value and coefficient of variation). In practice, the sensitivity of the results with respect to α can be investigated,

although selecting $\alpha = 0.5$ generally provides robust predictions (because of the normalization of each term in the cost function). The solution Lagrange multiplier is then defined as:

$$\boldsymbol{\lambda}^{\text{sol}} := \underset{\boldsymbol{\lambda} \in \mathbb{D}_{\boldsymbol{\lambda}}}{\operatorname{argmin}} \mathcal{J}_{\alpha}(\boldsymbol{\lambda}) . \quad (3.8)$$

The above constrained nonconvex nonlinear optimization problem can be solved by using an interior-point method (such as the one available in the `fmincon` function for Matlab users, for instance) or a pattern search algorithm running in the neighborhood of a starting point. The definition of this optimization guess, denoted by $\boldsymbol{\lambda}^{\text{ini}}$ hereafter, is addressed below.

3.1.3 Definition of the initial guess through an auxiliary optimization problem

By using a first order approximation of the constraint in mean given by Eq. (2.6a), namely

$$\mathbb{E} \left\{ \expm \left(\sum_{i=1}^N G_i(\mathbf{x})[E_i] \right) \right\} \approx [I_n] + \mathbb{E} \left\{ \sum_{i=1}^N G_i(\mathbf{x})[E_i] \right\} , \quad (3.9)$$

it can be deduced that

$$\forall 1 \leq i \leq N , \quad \mathbb{E} \{ G_i(\mathbf{x}) \} \approx 0 . \quad (3.10)$$

Upon matching the mode coordinates with the mean values thus approximated, it is then assumed that

$$\nabla_{\mathbf{u}} \Phi_{\boldsymbol{\lambda}}(\mathbf{u})|_{\mathbf{u}=\mathbf{0}_N} = \mathbf{0}_N . \quad (3.11)$$

Owing to the normal convergence of the matrix exponential series, it can be shown that

$$(\nabla \Phi_{\boldsymbol{\lambda}}(\mathbf{u}))_{\ell} = \sum_{k=1}^{+\infty} \frac{1}{k!} \left(\sum_{p=1}^k \operatorname{Tr} ([\Lambda][U]^{p-1}[E_{\ell}][U]^{k-p}) \right) + \lambda \operatorname{Tr}([E_{\ell}]) . \quad (3.12)$$

Hence

$$\operatorname{Tr}([\Lambda][E_{\ell}]) + \lambda \operatorname{Tr}([E_{\ell}]) = 0 \quad (3.13)$$

in virtue of Eq. (3.11), and thus

$$\lambda_{\ell} = -\lambda \operatorname{Tr}([E_{\ell}]) / \| [E_{\ell}] \|_F^2 , \quad 1 \leq \ell \leq N . \quad (3.14)$$

For all symmetry classes, the above relation yields a very simple form for the matrix-valued Lagrange multiplier, namely

$$\sum_{i=1}^N \lambda_i [E_i] = -\lambda [I_n] , \quad (3.15)$$

so that the probability density function defined by Eq. (3.4) is integrable at infinity if and only if $\lambda < 0$. We assume at this stage that this probability density function is integrable over $\mathbb{R}^N$ for $\lambda < 0$. For any fixed value of λ , let $\mathcal{J}_{\alpha}^0$ be the auxiliary cost function defined as

$$\mathcal{J}_{\alpha}^0(\lambda) := \mathcal{J}_{\alpha}(\varphi_1(\lambda), \dots, \varphi_N(\lambda), \lambda) , \quad (3.16)$$

where

$$\varphi_i(\lambda) := -\lambda \operatorname{Tr}([E_i]) / \| [E_i] \|_F^2 , \quad 1 \leq i \leq N . \quad (3.17)$$

The initial guess $\boldsymbol{\lambda}^{\text{ini}}$ is then defined by

$$\boldsymbol{\lambda}^{\text{ini}} := (\varphi_1(\lambda^0), \dots, \varphi_N(\lambda^0), \lambda^0) , \quad (3.18)$$

with

$$\lambda^0 := \underset{\lambda \in \mathbb{R}_+^*}{\operatorname{argmin}} \mathcal{J}_\alpha^0(\lambda) . \quad (3.19)$$

In this work, the one-dimensional optimization problem given by Eq. (3.19) is solved by evaluating the cost function on a regular grid. The mathematical expectations involved in the cost function $\mathcal{J}_\alpha$ are computed by invoking the ergodic theorem. For any given value of λ , the probability distribution defined by Eqs. (3.3-3.4) is sampled by using a Metropolis-Hastings algorithm (MHA) [12] [17] (with an acceptance rate set to 1/4, following [20]). For later use, this algorithm is denoted as $Z_{\boldsymbol{\lambda}}^{(k+1)} = \mathcal{F}_{\text{MHA}}(Z_{\boldsymbol{\lambda}}^{(k)})$, $k \geq 1$ (with initial value $Z_{\boldsymbol{\lambda}}^{(1)}$), where $Z_{\boldsymbol{\lambda}}^{(k)}$ is the k -th iterate along the Markov chain. By using a second-order approximation of the exponential term in the potential function, it is seen that a Gaussian transition kernel may be advantageously selected. In this case, the mean and covariance matrix of the (Gaussian) kernel are given by the null vector $\mathbf{0}_N$ and by the diagonal matrix $[\text{Cov}_{\boldsymbol{\lambda}}]$ with entries:

$$[\text{Cov}_{\boldsymbol{\lambda}}]_{ii} := -\lambda^{-1} \| [E_i] \|_F^{-2} , \quad 1 \leq i \leq N . \quad (3.20)$$

Here, it is worthwhile to notice that the MHA does not require specific properties on the potential function (apart from asymptotic ones related to integrability) and can thus be selected for constructing an approximate solution for the Lagrange multipliers, for a given value of the parameters involved in the constraint equations. Once this solution has been determined, additional properties of Φ [25] will be investigated in order to address the random field generation through the alternative algorithm (involving stochastic differential equations) proposed in [9] (see section § 4.1).

3.2 Application to all symmetry classes (ranging from isotropy to orthotropy) in three-dimensional elasticity

3.2.1 Detailed analysis for the transversely isotropic case

In this section, the above methodology is applied to the transversely isotropic symmetry (for which $N = 5$). Let $\mathbf{n} = \mathbf{e}_3$ be the unit vector normal to the plane of isotropy. A matrix basis can then be obtained through the Kelvin-Voigt matrix representation of the fourth-rank tensor basis given by [29]:

$$\begin{aligned} \llbracket E_1 \rrbracket &:= [p] \otimes [p] , \quad \llbracket E_2 \rrbracket := \frac{1}{2} [q] \otimes [q] , \quad \llbracket E_3 \rrbracket := \frac{1}{\sqrt{2}} ([p] \otimes [q] + [q] \otimes [p]) , \\ \llbracket E_4 \rrbracket &:= [q] \otimes [q] - \llbracket E_2 \rrbracket , \quad \llbracket E_5 \rrbracket := \llbracket I \rrbracket - \llbracket E_1 \rrbracket - \llbracket E_2 \rrbracket - \llbracket E_4 \rrbracket , \end{aligned} \quad (3.21)$$

where $[p] := \mathbf{n} \otimes \mathbf{n}$, $[q] := [I_3] - [p]$ and $\llbracket I \rrbracket_{ijkl} := (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})/2$. It should be noticed that the algebraic properties of the matrix basis, together with the constraints used in the maximum entropy principle, yield some specific statistical dependences between the components of random vector $\mathbf{G}(\mathbf{x})$ ($\mathbf{x}$ being fixed; see [8]). Following the strategy introduced in section § 3.1, a set of admissible values for parameter ν is first determined by means of the projection-based methodology (see section § 3.1.1). The graph of mapping $\delta_{[N]} \mapsto \nu$ thus obtained is shown in Fig. 3.1, where it is seen that the aforementioned function is monotonically decreasing. Note that this figure shows that most of the values of practical

interest for ν are concentrated near the origin (from a numerical standpoint, it is seen that the range of concentration depends on the symmetry class under consideration). This important fact illustrates that completing the first step of the methodology (see section § 3.1) is absolutely required in order to set well-posed optimization problems associated with Eqs. (2.6). For a given value of λ , the burn-in period

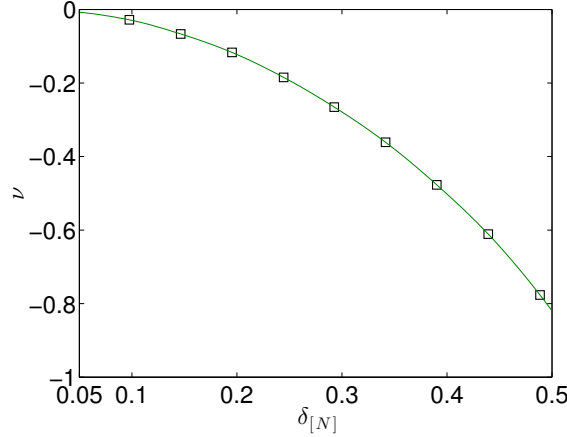


Figure 3.1: Graph of $\delta_{[N]} \mapsto \nu$ (cubic spline interpolation) for transverse isotropy.

for the Metropolis-Hastings algorithm is estimated by characterizing the convergence of the mapping $N_{iter} \mapsto \text{ConvMes}(N_{iter})$ defined as:

$$\text{ConvMes}(N_{iter}) := \frac{1}{N_{iter}} \sum_{k=1}^{N_{iter}} \left\| Z_{(\varphi_1(\lambda), \dots, \varphi_N(\lambda), \lambda)}^{(k)} \right\|^2. \quad (3.22)$$

Note that for transverse isotropy, one has $(\varphi_1(\lambda), \dots, \varphi_N(\lambda), \lambda) = \lambda(-1, -1, 0, -1, -1, 1)$. The graph of ConvMes is shown in Fig. 3.2 for $\lambda = -5$, showing that the convergence is achieved for $k \geq M_0$, with $M_0 = 3 \times 10^4$. The ergodic estimators are then computed between iterates M_0 and M_e , where the value of M_e is again determined through a convergence analysis on the cost function defined by Eq. (3.16). The graph of the ergodic estimator for $\mathcal{J}_{1/2}^0$ is shown in Fig. 3.3 for $\lambda = -5$. It is seen that for this case, the convergence is clearly reached for $M_e = 5 \times 10^7$. The graphs of mapping $\lambda \mapsto \mathcal{J}_{1/2}^0(\lambda)$ obtained for the $m_\nu = 7$ retained values of parameter ν are shown in Fig. 3.4. For each selected value, it is seen that an optimal initialization value λ^0 can be clearly identified. By using this guess (for a particular value of ν) in the solving of the global optimization problem defined by Eq. (3.6) (in which the underlying structure of the search space for each multiplier is relaxed), it is found that numerically,

$$\boldsymbol{\lambda}^{\text{sol}} \simeq (\varphi_1(\lambda^0), \dots, \varphi_5(\lambda^0), \lambda^0), \quad (3.23)$$

with a relative error that is typically less than 1% – hence indicating the relevance of the strategy for defining the guesses. In addition, the relation $\delta_{[N]} \mapsto |\lambda^0|$ turns out to be linear in log-log scale, as shown in Fig. 3.5 (see also Tab. 3.1 below for specific values), thus suggesting the use of a power-law interpolation:

$$\lambda^0 = -\alpha_1(\delta_{[N]})^{\alpha_2}, \quad (3.24)$$

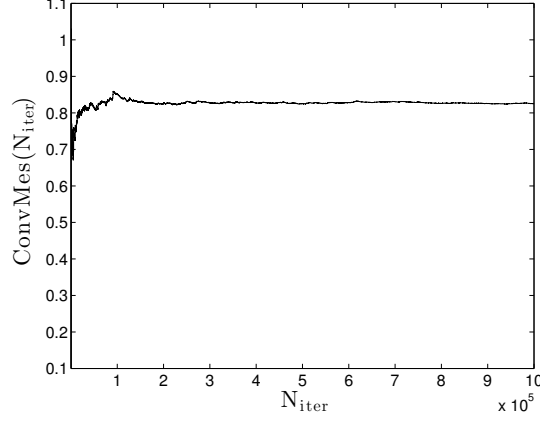


Figure 3.2: Convergence towards the invariant measure: graph of $N_{iter} \mapsto \text{ConvMes}(N_{iter})$ for $\lambda = -5$.

ν_i	-0.0295	-0.0767	-0.1225	-0.2	-0.2767	-0.4	-0.5019
λ^0	-85.9962	-32.9983	-20.9264	-12.7144	-9.1724	-6.4327	-5.1673
$\delta_{[N]}$	0.0984	0.1587	0.1991	0.2554	0.3005	0.3576	0.3985

Table 3.1: Table of selected values for parameter ν_i , together with the corresponding values of initial guesses and associated parameters $\delta_{[N]}$.

where $\alpha_1 = 0.8156$ and $\alpha_2 = -2.01$. The final solution for the transverse isotropy is then defined as

$$\lambda^{\text{sol}} = -\alpha_1(\delta_{[N]})^{\alpha_2}(-1, -1, 0, -1, -1, 1), \quad (3.25)$$

or in matrix form,

$$[\Lambda^{\text{sol}}] = -\lambda^{\text{sol}}[I_n], \quad \lambda^{\text{sol}} = -\alpha_1(\delta_{[N]})^{\alpha_2}, \quad (3.26)$$

with $\alpha_1 = 0.8156$ and $\alpha_2 = -2.01$.

3.2.2 Numerical application for the transversely isotropic case

In this section, we provide a comparison between a database obtained by projecting a set of realizations of a random matrix with triclinic fluctuations ($\delta^{\text{triclinic}} \simeq 0.68$) on the set of transversely isotropic matrices (for a given unit vector normal to the plane of isotropy), and data simulated by using the results derived in the previous section. To this aim, the parameter $\delta_{[N]}$ is estimated from the projected data, and the probability density function involving the approximate potential is sampled through the MHA. Note that the above level of fluctuations does not correspond to a value that was considered for constructing the approximate solution. The graphs of some first- and second-order marginal probability density functions (obtained by using a classical kernel density estimator with 60,000 realizations) are shown in Figs. 3.6, 3.7, 3.8, 3.9 and 3.10. A very good match is observed (no matter the components involved), thus showing the relevance of the approximate solution. It should be noted that the capability of the model to reproduce projected data has been numerically checked for all symmetry classes of interest. This property can be

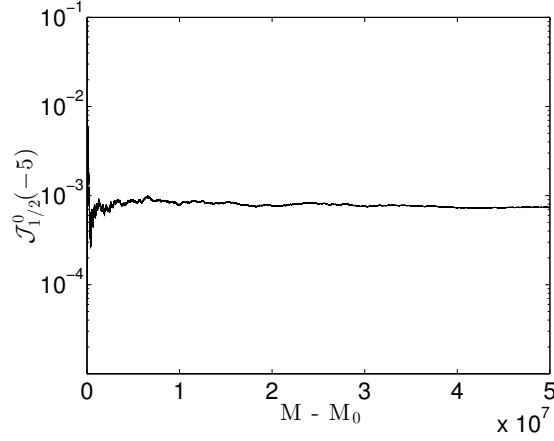


Figure 3.3: Convergence of the ergodic estimator for $\mathcal{J}_{1/2}^0$ and $\lambda = -5$.

explained in part by the linearity of the relation between the triclinic tensor and its projection, so that the database and the model essentially carries information of the same probabilistic nature.

3.2.3 Results summary for all symmetry classes

For all symmetry classes, a general form for the potential function can be then derived and expressed in terms of parameters $\delta_{[N]}$, α_1 and α_2 . Following the previous sections, it can be deduced that the potential function reduces to

$$\Phi(\mathbf{u}) = \alpha_1 \delta_{[N]}^{\alpha_2} \times \text{Tr} \left(\expm \left(\sum_{i=1}^N u_i [E_i] \right) - \sum_{j=1}^N u_j [E_j] \right), \quad \forall \mathbf{u} \in \mathbb{R}^N, \quad (3.27)$$

where the parameters α_1 and α_2 are summarized in Tab. 3.2 for the six symmetry classes (it is recalled that the triclinic case has been addressed in [24], and that the monoclinic class is not considered here because of its limited range of application – apart from the modeling of crystals) and $\delta_{[N]} > 0$. It is seen that the

Symmetry	N	α_1	α_2	Crystallographic preferred directions
Isotropic	2	0.3246	-2.0300	-
Cubic	3	0.4914	-2.0293	Canonical basis of $\mathbb{R}^3$
Trans. Iso.	5	0.8156	-2.01	Normal to the plane of isotropy $\mathbf{n} = \mathbf{e}_3$
Tetragonal	6	0.9906	-2.0040	Canonical basis of $\mathbb{R}^3$
Trigonal	6	0.9858	-2.0056	$\mathbf{a} = \mathbf{e}_1$, $\mathbf{b} = (-1/2, \sqrt{3}/2, 0)$, $\mathbf{c} = \mathbf{e}_3$
Orthotropic	9	1.4323	-2.0293	Canonical basis of $\mathbb{R}^3$

Table 3.2: Regression coefficients for the six symmetry classes and description of the crystallographic systems used in the calculations (the canonical basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ of $\mathbb{R}^3$ is defined as $(\mathbf{e}_i)_j = \delta_{ij}$, $1 \leq i, j \leq 3$).

real solution multiplier λ turns out to be proportional to $\delta_{[N]}^{-2}$ and depends almost linearly on N . These

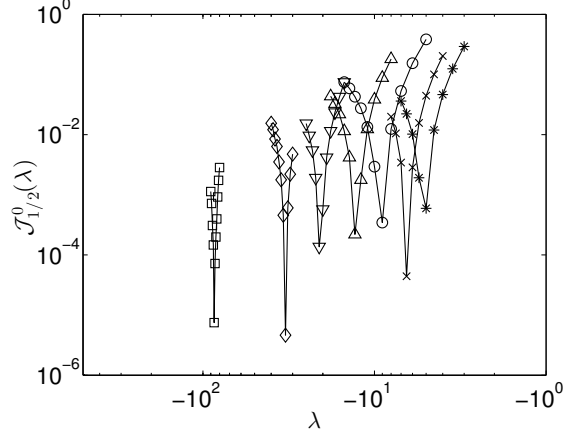


Figure 3.4: Analysis of the cost function in terms of a single parameter. Identification of the initializations for the last Lagrange multiplier. The symbols $\square$, $\diamond$, ∇ , $\triangle$, $\circ$, $\times$ and $*$ correspond in this order to the decreasing values of the parameters ν considered.

results are consistent with the ones obtained for the triclinic class (see [24]) and allow for a significant reduction of the computational cost associated with the statistical inverse identification of the stochastic model (since the first-order marginal distribution then depends on a single statistically-sound parameter, namely $\delta_{[\mathbf{N}]}$ – rather than on $N+1$ non-physical parameters). It is worth mentioning that for all symmetry classes, the traces of all basis matrices turn out to be independent from the crystallographic parameters [29], hence making the above form for the potential function valid regardless of these parameters. Finally, the gradient associated with the final form of the potential function (defined by Eq. (3.27)) takes the following simple expression,

$$1 \leq \ell \leq N, \quad (\nabla \Phi_{\boldsymbol{\lambda}}(\mathbf{u}))_{\ell} = -\lambda \ll \expm[U], [E_{\ell}] \gg + \lambda \text{Tr}([E_{\ell}]), \quad (3.28)$$

which satisfies Eq. (3.11) by construction.

3.2.4 Analytical solutions for the isotropic and cubic symmetries

For the isotropic and cubic symmetries, analytical solutions for the Lagrange multipliers can be obtained thanks to some specific properties of the tensorial basis. Let us first consider the isotropic case, and recall that $\mathbf{x}$ is fixed. From Eq. (2.7), it is deduced that $p_{\mathbf{G}(\mathbf{x})}(\mathbf{g}) = p_{G_1(\mathbf{x})}(g_1) \times p_{G_2(\mathbf{x})}(g_2)$, with:

$$\begin{cases} \forall g_1 \in \mathbb{R}, & p_{G_1(\mathbf{x})}(g_1) = c \exp(-\lambda_1^{\text{sol}} \exp(g_1) - \lambda^{\text{sol}} g_1), \\ \forall g_2 \in \mathbb{R}, & p_{G_2(\mathbf{x})}(g_2) = c \exp(-5\lambda_2^{\text{sol}} \exp(g_2) - 5\lambda^{\text{sol}} g_2). \end{cases} \quad (3.29)$$

The integrability of the marginal probability density functions at infinity implies that $\lambda_1^{\text{sol}} > 0$, $\lambda_2^{\text{sol}} > 0$ and $\lambda^{\text{sol}} < 0$. Using the normalization constraint and the measure transformation, it can be shown that the marginal probability density functions associated with the random variables $N_1(\mathbf{x})$ and $N_2(\mathbf{x})$ take

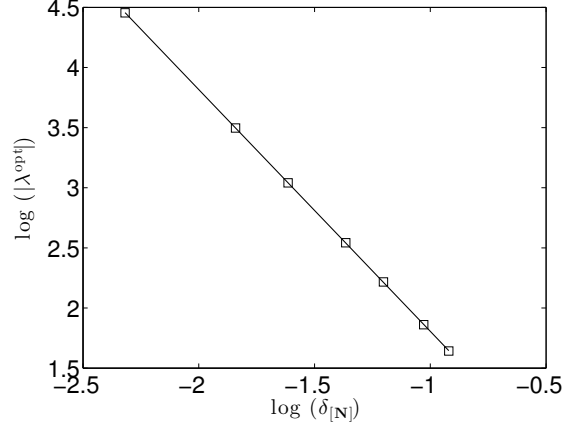


Figure 3.5: Linear interpolation of the last optimized Lagrange multiplier in logarithmic scale.

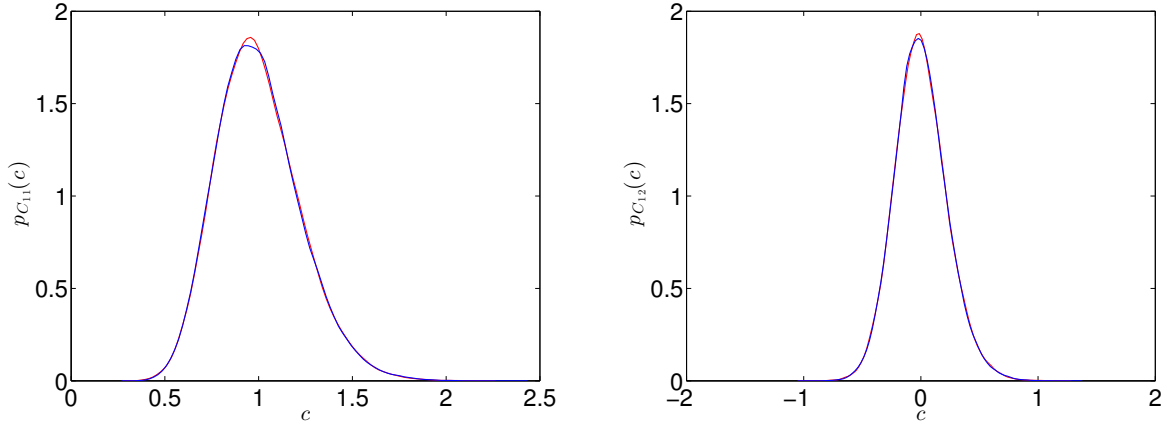


Figure 3.6: Graph of the probability density function of random variables C_{11} (left panel) and C_{12} (right panel). Blue line: data. Red line: model-based simulations.

the form:

$$\begin{cases} \forall n_1 \in \mathbb{R}_+^* , & p_{N_1(\mathbf{x})}(n_1) = \mathcal{I}_{\mathbb{R}_+}(n_1) \frac{n_1^{k_1-1}}{\theta_1^{k_1} \Gamma(k_1)} \exp\left(-\frac{n_1}{\theta_1}\right) , \\ \forall n_2 \in \mathbb{R}_+^* , & p_{N_2(\mathbf{x})}(n_2) = \mathcal{I}_{\mathbb{R}_+}(n_2) \frac{n_2^{k_2-1}}{\theta_2^{k_2} \Gamma(k_2)} \exp\left(-\frac{n_2}{\theta_2}\right) , \end{cases} \quad (3.30)$$

in which $\mathcal{I}_{\mathbb{R}_+}$ is the indicator function of $\mathbb{R}_+$ and Γ is the Gamma function defined as:

$$\Gamma(z) := \int_0^{+\infty} t^{z-1} \exp(-t) dt , \quad \forall z > 0 . \quad (3.31)$$

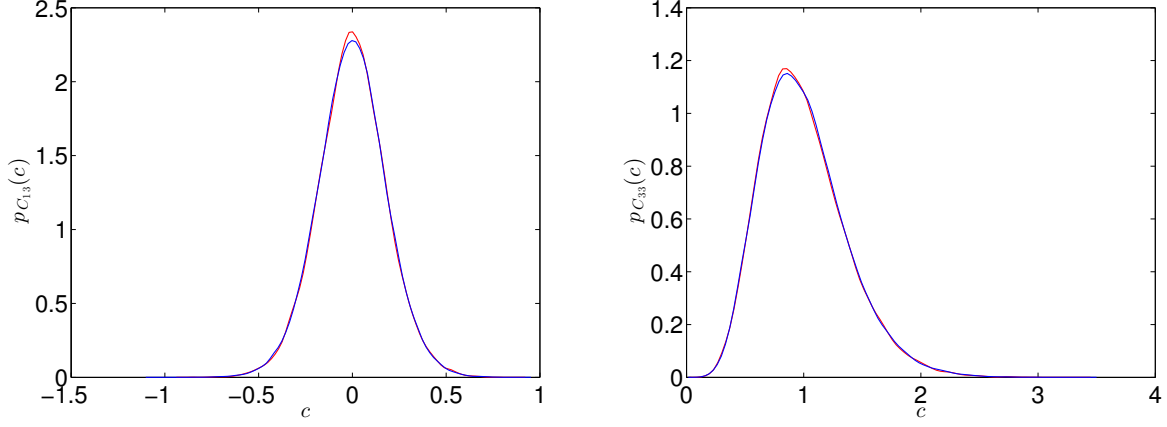


Figure 3.7: Graph of the probability density function of random variables C_{13} (left panel) and C_{33} (right panel). Blue line: data. Red line: model-based simulations.

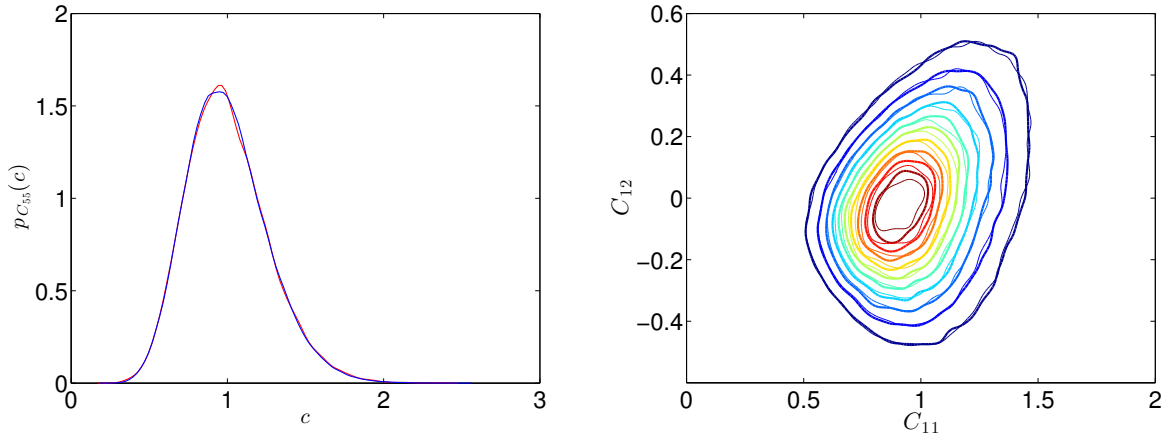


Figure 3.8: Graph of the probability density function of random variable C_{55} (left panel) and contour plot of the joint probability density function of random variables C_{11} and C_{12} (right panel). Blue (or thick) line(s): data. Red (or thin) line(s): model-based simulations.

It follows that $N_1(\mathbf{x})$ and $N_2(\mathbf{x})$ are statistically independent Gamma-distributed random variables, with parameters $(k_1, \theta_1) = (-\lambda^{\text{sol}}, 1/\lambda_1^{\text{sol}})$ and $(k_2, \theta_2) = (-5\lambda^{\text{sol}}, 1/(5\lambda_2^{\text{sol}}))$. Since $\mathbb{E}\{N_i\} = k_i\theta_i$, $\text{Var}\{N_i\} = k_i\theta_i^2$ and $\mathbb{E}\{N_i\} = 1$ for $1 \leq i, j \leq 2$, one has $\delta_{[N]}^2 = (\theta_1 + 5\theta_2)/6$. Therefore, it is seen that for the isotropic case,

$$\lambda^{\text{sol}} = -1/(3\delta_{[N]}^2), \quad \lambda_1^{\text{sol}} = \lambda_2^{\text{sol}} = -\lambda^{\text{sol}}, \quad (3.32)$$

which is in complete accordance with the numerical result provided in Tab. 3.2.

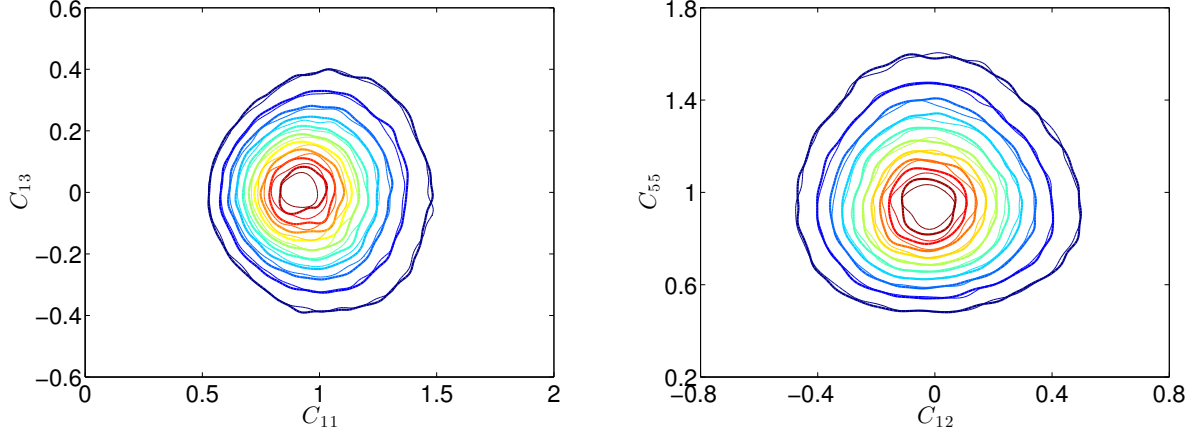


Figure 3.9: Contour plots of the joint probability density function of random variables C_{11} and C_{13} (left panel); C_{12} and C_{55} (right panel). Thick lines: data. Thin lines: model-based simulations.

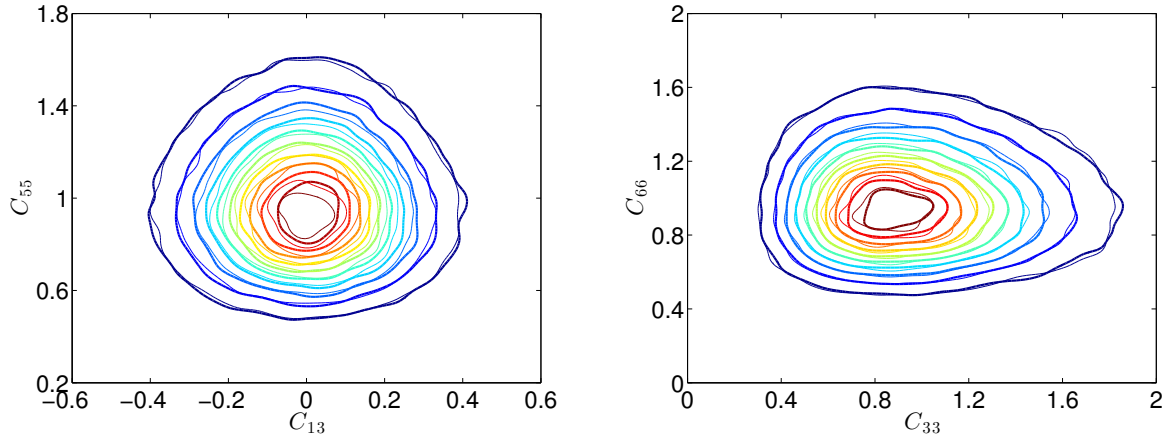


Figure 3.10: Contour plots of the joint probability density function of random variables C_{13} and C_{55} (left panel); C_{33} and C_{66} (right panel). Thick lines: data. Thin lines: model-based simulations.

Similarly, it is found for the cubic symmetry that

$$\lambda^{\text{sol}} = -1/(2\delta_{[\mathbf{N}]}) , \quad \lambda_1^{\text{sol}} = \lambda_2^{\text{sol}} = \lambda_3^{\text{sol}} = -\lambda^{\text{sol}} , \quad (3.33)$$

in agreement with Tab. 3.2.

4 Mathematical properties of the approximate potential function and sampling algorithm

4.1 Mathematical properties of the potential function Φ

Let $\mathbb{M}_n^{\text{sym}}$ be given, and let Φ be the potential function defined by Eq. (3.27), with $\lambda = -\alpha_1 \delta_{[N]}^{\alpha_2}$ and $\delta_{[N]} > 0$.

Property 4.1 *The potential function $\mathbb{R}^N \ni \mathbf{u} \mapsto \Phi(\mathbf{u})$ defined by Eq. (3.27) is continuous on $\mathbb{R}^N$ and strictly convex for all symmetry classes.*

Proof. The continuity first follows from the composition of continuous functions. Next, note that

$$\text{Tr} \left(\expm \left(\sum_{i=1}^N u_i [E_i] \right) - \sum_{j=1}^N u_j [E_j] \right) = \text{Tr} \left(\sum_{i=1}^n f(\rho_i) \mathbf{v}^i \otimes \mathbf{v}^i \right), \quad (4.1)$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ is the function defined by $f(x) = \exp(x) - x$ for any x in $\mathbb{R}$, $\rho_1, \dots, \rho_n$ are the (possibly equal) eigenvalues of $[U] = \sum_{i=1}^N u_i [E_i] \in \mathbb{M}_n^{\text{S}}(\mathbb{R})$ and $\{\mathbf{v}^i\}_{i=1}^n$ is the set of associated eigenvectors. Since the function f is continuous and strictly convex on $\mathbb{R}$, it follows that the function $[U] \mapsto \text{Tr}(\expm([U]) - [U])$ is strictly convex on $\mathbb{M}_n^{\text{S}}(\mathbb{R})$. Based on this property, it can easily be shown that the function $h : \mathbf{u} \mapsto h(\mathbf{u}) = \text{Tr}(\expm(\sum_{i=1}^N u_i [E_i]) - \sum_{j=1}^N u_j [E_j])$ is then strictly convex on $\mathbb{R}^N$, and so is the potential function. $\square$

Property 4.2 *The function $\mathbf{u} \mapsto \|\nabla_{\mathbf{u}} \Phi(\mathbf{u})\|_{\mathbb{R}^N}$ is locally bounded on $\mathbb{R}^N$, regardless of the symmetry class.*

Proof. For $\mathbf{u} \in \mathbb{R}^N$, let $[U] = \sum_{k=1}^N u_k [E_k]$ and $[L] = \sum_{j=1}^N \ell_j [E_j] = \expm([U])$. Eq. (3.12) can then be rewritten as

$$(\nabla_{\mathbf{u}} \Phi(\mathbf{u}))_{\ell} = -\lambda \ll [L] - [I_n], [E_{\ell}] \gg, \quad (4.2)$$

for $1 \leq \ell \leq N$. Being the composition of continuous functions, the function $\mathbf{u} \mapsto \nabla_{\mathbf{u}} \Phi(\mathbf{u})$ is clearly continuous over $\mathbb{R}^N$, and so is the function $\mathbf{u} \mapsto \|\nabla_{\mathbf{u}} \Phi(\mathbf{u})\|_{\mathbb{R}^N}$. Hence, it follows that $\mathbf{u} \mapsto \|\nabla_{\mathbf{u}} \Phi(\mathbf{u})\|_{\mathbb{R}^N}$ is locally bounded on $\mathbb{R}^N$. $\square$

Property 4.3 *The function $\mathbf{u} \mapsto \|\nabla_{\mathbf{u}} \Phi(\mathbf{u})\|_{\mathbb{R}^N}$ is such that*

$$\mathbb{E}\{\|\nabla_{\mathbf{u}} \Phi(\mathbf{u})\|_{\mathbb{R}^N}\} = \int_{\mathbb{R}^N} \|\nabla_{\mathbf{u}} \Phi(\mathbf{u})\|_{\mathbb{R}^N} c \exp\{-\Phi(\mathbf{u})\} d\mathbf{u} < +\infty, \quad (4.3)$$

with $c^{-1} = \int_{\mathbb{R}^N} \exp\{-\Phi(\mathbf{u})\} d\mathbf{u}$, regardless of the symmetry class.

Proof. Note first that $\text{Tr}([E_i]) \geq 0$, $1 \leq i \leq N$, for any symmetry class. Upon using Eq. (4.2) and by invoking the Cauchy-Schwartz inequality, it can then be shown that

$$\mathbb{E}\left\{\|\nabla_{\mathbf{u}} \Phi(\mathbf{u})\|_{\mathbb{R}^N}^2\right\} \leq \lambda^2 \sum_{i=1}^N \left(\mathbb{E}\{\|[L]\|_F^2\} \|[E_i]\|_F^2 + (\text{Tr}([E_i]))^2 \right), \quad (4.4)$$

where $[\mathbf{L}] = \expm([\mathbf{U}])$ and $[\mathbf{U}] = \sum_{j=1}^N u_j [E_j]$. Due to the constraint defined by Eq.(2.6b), the random matrix $[\mathbf{L}]$ is a second-order variable (that is, $\mathbb{E}\{\|[\mathbf{L}]\|_F^2\} < +\infty$), so that

$$\mathbb{E}\{\|\nabla_{\mathbf{u}}\Phi(\mathbf{u})\|_{\mathbb{R}^N}^2\} < +\infty \quad (4.5)$$

and hence, $\mathbb{E}\{\|\nabla_{\mathbf{u}}\Phi(\mathbf{u})\|_{\mathbb{R}^N}\} < +\infty$. $\square$

Property 4.4 *The potential function Φ is such that*

$$\inf_{\mathbf{u} \in \mathbb{R}^N} \Phi(\mathbf{u}) = \Phi^{\min}, \quad \Phi^{\min} > 0, \quad (4.6)$$

regardless of the symmetry class.

Proof. By construction, one has $\Phi^{\min} = \Phi(\mathbf{0}_N)$ and thus, $\Phi^{\min} = n \alpha_1 \delta_{[N]}^{\alpha_2} > 0$ in view of Eq. (3.27). $\square$

Property 4.5 *The potential function Φ is such that*

$$\inf_{\|\mathbf{u}\|_{\mathbb{R}^N} > R} \Phi(\mathbf{u}) \rightarrow +\infty \text{ as } R \rightarrow +\infty, \quad (4.7)$$

for all symmetry classes.

Proof. Note first that the potential function can be written as

$$\Phi(\mathbf{u}) = \alpha_1 \delta_{[N]}^{\alpha_2} \sum_{i=1}^q m_i (\exp(\rho_i) - \rho_i), \quad (4.8)$$

where $\rho_1, \dots, \rho_q$ are the q distinct eigenvalues of $[\mathbf{U}] = \sum_{i=1}^N u_i [E_i] \in \mathbb{M}_n^S(\mathbb{R})$ and m_i denotes the algebraic multiplicity of ρ_i , $1 \leq i \leq q$; hence, $\Phi(\mathbf{u}) > 0$ for any $\mathbf{u} \in \mathbb{R}^N$. Since $\|[E_i]\|_F^2 \geq 1$ for $i = 1, \dots, N$ (for all symmetry classes), it is easily seen that

$$\|\mathbf{u}\|_{\mathbb{R}^N} \leq \|[\mathbf{U}]\|_F \leq \sqrt{n} \|[\mathbf{U}]\|_2, \quad (4.9)$$

where $\|[\mathbf{U}]\|_2 = \max_{i \in \{1, \dots, q\}} |\rho_i| = |\rho_M|$. From Eq. (4.8), it can be deduced that

$$\alpha_1 \delta_{[N]}^{\alpha_2} \times m_{\rho_M} (\exp(\rho_M) - \rho_M) \leq \Phi(\mathbf{u}), \quad \forall \mathbf{u} \in \mathbb{R}^N, \quad (4.10)$$

in which $\rho_M \in \mathbb{R}$. If $\rho_M > 0$, a combination of Eqs. (4.9) and (4.10) yields

$$\inf_{\|\mathbf{u}\|_{\mathbb{R}^N} > R} \Phi(\mathbf{u}) \rightarrow +\infty \text{ as } R \rightarrow +\infty. \quad (4.11)$$

If $\rho_M < 0$, the left hand side of Eq. (4.10) tends to $+\infty$ as $\|\mathbf{u}\|_{\mathbb{R}^N} \rightarrow +\infty$ (which corresponds to $\rho_M \rightarrow -\infty$), so that Eq. (4.11) similarly holds. $\square$

The properties 4.2-4.3 are essentially related to the sampling scheme for the non-Gaussian random field, whereas the properties 4.4-4.5 show that $P(d\mathbf{u}) := c \exp\{-\Phi(\mathbf{u})\} d\mathbf{u}$ (with $c^{-1} = \int_{\mathbb{R}^N} \exp\{-\Phi(\mathbf{u})\} d\mathbf{u}$) can be interpreted as a probability measure whenever $\delta_{[N]} > 0$.

4.2 Sampling scheme for random fields with values in $\mathbb{M}_n^{\text{sym}}$

Based on the previous properties, it can be shown that the $\mathbb{M}_n^{\text{sym}}(\mathbb{R})$ -valued non-Gaussian random field defined in Section 2 can be sampled by solving a family of Itô stochastic differential equations indexed by Ω [9]. This sampling scheme is recalled for the sake of completeness below.

4.2.1 Definition of an auxiliary $\mathbb{R}^N$ -valued Gaussian random field

Let $\{\Xi(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ be a $\mathbb{R}^N$ -valued Gaussian homogeneous centered random field defined by the continuous correlation function $[\mathbf{R}_\Xi]$ (with values in the set of diagonal real matrices) given by:

$$\forall \mathbf{y} \in \mathbb{R}^d, \quad [\mathbf{R}_\Xi(\mathbf{y})]_{ii} = \mathbb{E}\{\Xi_i(\mathbf{x} + \mathbf{y})\Xi_i(\mathbf{x})\}, \quad [\mathbf{R}_\Xi(\mathbf{0})]_{ii} = 1, \quad 1 \leq i \leq N. \quad (4.12)$$

Let $\mathbb{R}^d \ni \mathbf{x} \mapsto \Xi(\mathbf{x}, \theta_k)$ be the k -th independent realization of the aforementioned random field. Let $\mathbf{W} = \{\mathbf{W}(r, \mathbf{x}), r \geq 0, \mathbf{x} \in \Omega\}$ be the $\mathbb{R}^N$ -valued second-order centered Gaussian random field defined as follows:

- $\forall \mathbf{x} \in \Omega, \mathbf{W}(0, \mathbf{x}) = \mathbf{0}$ almost surely;
- the generalized derivative $D_r \mathbf{W}$ of $\mathbf{W}$ with respect to r is the normalized Gaussian cylindrical white noise denoted by $\mathbf{B}$.

The generalized covariance function $[\mathbf{C}_B]$ of $\mathbf{B}$ writes

$$\forall (\mathbf{x}, \mathbf{y}) \in \Omega \times \Omega, \quad \forall \tau \in \mathbb{R}, \quad [\mathbf{C}_B(\mathbf{x}, \mathbf{y}, t + \tau, t)] = \delta_0(\tau) [\mathbf{R}_\Xi(\mathbf{x} - \mathbf{y})], \quad (4.13)$$

with δ_0 the generalized Delta function at the origin of $\mathbb{R}$. It follows that by construction, and for any $\mathbf{x}$ fixed in Ω , the stochastic process $\{\mathbf{W}(r, \mathbf{x}), r \geq 0\}$ is a standard Wiener process with values in $\mathbb{R}^N$.

4.2.2 Sampling scheme

For $\mathbf{x}$ fixed in Ω , let $\{\mathbf{U}(r, \mathbf{x}), \mathbf{V}(r, \mathbf{x}), r \geq 0\}$ be the Markov process with values in $\mathbb{R}^N \times \mathbb{R}^N$ satisfying the following Itô stochastic differential equation:

$$\forall r \in \mathbb{R}^+, \quad \begin{cases} d\mathbf{U}(r, \mathbf{x}) = \mathbf{V}(r, \mathbf{x}) dr \\ d\mathbf{V}(r, \mathbf{x}) = \left(-\nabla_{\mathbf{u}} \Phi(\mathbf{U}(r, \mathbf{x})) - \frac{\eta}{2} \mathbf{V}(r, \mathbf{x}) \right) dr + \sqrt{\eta} d\mathbf{W}(r, \mathbf{x}) \end{cases}, \quad (4.14)$$

where $\eta \in \mathbb{R}_+^*$ is a free tunable parameter and $\{\mathbf{W}(r, \mathbf{x}), r \geq 0\}$ is the Wiener process defined in the previous section. The above stochastic differential equation is supplemented with initial conditions $\mathbf{U}(0, \mathbf{x}) = \mathbf{U}^0(\mathbf{x})$ and $\mathbf{V}(0, \mathbf{x}) = \mathbf{V}^0(\mathbf{x})$ almost surely, where the probability distribution of the random variable $(\mathbf{U}^0(\mathbf{x}), \mathbf{V}^0(\mathbf{x}))$ is assumed to be given. It can then be shown that

$$\lim_{r \rightarrow +\infty} \mathbf{U}(r, \mathbf{x}) = \mathbf{G}(\mathbf{x}) \quad (4.15)$$

in probability distribution, hence providing a sampling algorithm for the non-Gaussian random field $\{\mathbf{G}(\mathbf{x}), \mathbf{x} \in \Omega\}$ introduced in Section 2.1. In practice, the stochastic differential equation can be discretized by using any suitable algorithm (see *e.g.* [15] [18] and [27] for further details). In particular, the Störmer-Verlet algorithm [28] [5] [11] [3] turns out to be very efficient and is then recalled hereinafter. Let Δr be

the discretization time step, and let $r_k = (k-1)\Delta r$, $k \geq 1$, correspond to a regular time discretization. For $k \geq 1$, the scheme is defined as

$$\begin{cases} \mathbf{U}^{(k+1/2)}(\mathbf{x}) &= \mathbf{U}^{(k)}(\mathbf{x}) + \frac{\Delta r}{2} \mathbf{V}^{(k)}(\mathbf{x}) \\ \mathbf{V}^{(k+1)}(\mathbf{x}) &= \frac{1-\zeta}{1+\zeta} \mathbf{V}^{(k)}(\mathbf{x}) - \frac{\Delta r}{1+\zeta} \nabla_{\mathbf{u}} \Phi(\mathbf{U}^{(k+1/2)}(\mathbf{x})) + \frac{\sqrt{\eta}}{1+\zeta} \Delta \mathbf{W}^{(k+1)}(\mathbf{x}) \quad , \\ \mathbf{U}^{(k+1)}(\mathbf{x}) &= \mathbf{U}^{(k+1/2)}(\mathbf{x}) + \frac{\Delta r}{2} \mathbf{V}^{(k+1)}(\mathbf{x}) \end{cases} \quad (4.16)$$

where $\mathbf{U}^{(k)}(\mathbf{x}) = \mathbf{U}(r_k, \mathbf{x})$, $\mathbf{V}^{(k)}(\mathbf{x}) = \mathbf{V}(r_k, \mathbf{x})$, $\zeta = \eta\Delta r/4$ and

$$\Delta \mathbf{W}^{(k+1)}(\mathbf{x}) = \mathbf{W}(r_{k+1}, \mathbf{x}) - \mathbf{W}(r_k, \mathbf{x}) := \sqrt{\Delta r} \boldsymbol{\Xi}(\mathbf{x}, \theta_{k+1}) \quad . \quad (4.17)$$

The initial conditions are given by $\mathbf{U}^{(1)}(\mathbf{x}) = \mathbf{u}^{(0)}$ and $\mathbf{V}^{(1)}(\mathbf{x}) = \mathbf{v}^{(0)}$, with $\mathbf{u}^{(0)}$ and $\mathbf{v}^{(0)}$ two arbitrary vectors in $\mathbb{R}^N$. The convergence stated by Eq. (4.15) thus reads as

$$\mathbf{G}(\mathbf{x}) = \lim_{\Delta r \downarrow 0} \left(\lim_{k \rightarrow +\infty} \mathbf{U}(r_k, \mathbf{x}) \right) \quad . \quad (4.18)$$

From a computational standpoint, the sampling scheme requires $N \times M_0$ (with M_0 the number of iterations required to reach the stationary solution) realizations of a $\mathbb{R}$ -valued centered homogeneous Gaussian random field to be generated, for every realization of $\{\mathbf{G}(\mathbf{x}), \mathbf{x} \in \Omega\}$. Therefore, the CPU time associated with the algorithm mainly depends on the strategy that is used for simulating such a Gaussian random field; see *e.g.* [7] and the references therein for a discussion about the computational complexity. Note that whenever the realizations of the Gaussian fields can be stored in a preprocessing step, the algorithm can be distributed in a straightforward manner. In this case, the total CPU time is scaled by the number of cores available, as well as by the CPU time that is necessary to reach the stationary solution at any point of the domain (the latter being equal to $M_0 \times t_{\text{iter}}$, with t_{iter} the CPU time to complete one iteration in the discretization scheme). It should be finally noted that the value of M_0 depends on the symmetry class under consideration (or equivalently, on N), as well as on the level of statistical fluctuations.

4.2.3 Numerical illustration for an orthotropic material

Here, we consider a two-dimensional open bounded domain $\Omega =]0, 32[^2$, discretized with 64 points along each direction. The selected mean model is given by

$$\forall \mathbf{x} \in \Omega, \quad [\underline{\mathcal{C}}(\mathbf{x})] = \begin{bmatrix} 32 & 4.2 & 3.8 & 0 & 0 & 0 \\ 4.2 & 20 & 4 & 0 & 0 & 0 \\ 3.8 & 4 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \quad (4.19)$$

and corresponds to a typical stiffness matrix for an orthotropic composite (in GPa). The correlation function $[\mathbf{R}_{\boldsymbol{\Xi}}]$ of the Gaussian homogeneous random field $\{\boldsymbol{\Xi}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^2\}$ is defined as

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^2, \quad [\mathbf{R}_{\boldsymbol{\Xi}}(\mathbf{x} - \mathbf{y})]_{ii} = \prod_{j=1}^2 \left(\frac{2\mathcal{L}_c^{ij}}{\pi|x_j - y_j|} \right)^2 \sin^2 \left(\frac{\pi|x_j - y_j|}{2\mathcal{L}_c^{ij}} \right) \quad (4.20)$$

for all $1 \leq i \leq N$, $\mathcal{L}_c^{ij}$ being interpreted as the spatial correlation length of the Gaussian homogeneous random field $\{\Xi_i(\mathbf{x}), \mathbf{x} \in \mathbb{R}^2\}$ along the direction defined by the unit vector $\mathbf{e}^j$ of the canonical basis of $\mathbb{R}^2$. For simplicity, the correlation lengths are assumed to be all equal to the same value, denoted by $\mathcal{L}_c$. The level of statistical fluctuations is set to $\delta_{[N]} = 0.1$. The parameters involved in the sampling algorithm are given by $\eta = 9$ and $\Delta r = 0.001$, and the stationary solution is reached after $M_0 = 50,000$ iterations. The CPU time associated with the convergence to the stationary regime at any point is 0.888 sec on average. Realizations of some components of the orthotropic stiffness random field are shown below for two choices of correlation lengths, namely $\mathcal{L}_c = 5.3$ and $\mathcal{L}_c = 16$. As expected, it is seen that the random field

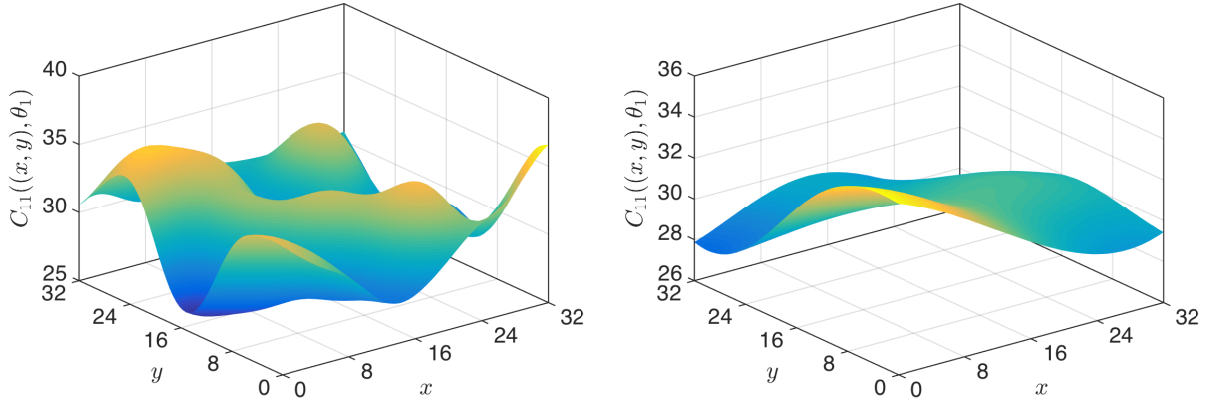


Figure 4.1: One realization of random field $\{C_{11}((x, y)), (x, y) \in \Omega\}$: $\mathcal{L}_c = 5.3$ (left panel), $\mathcal{L}_c = 16$ (right panel).

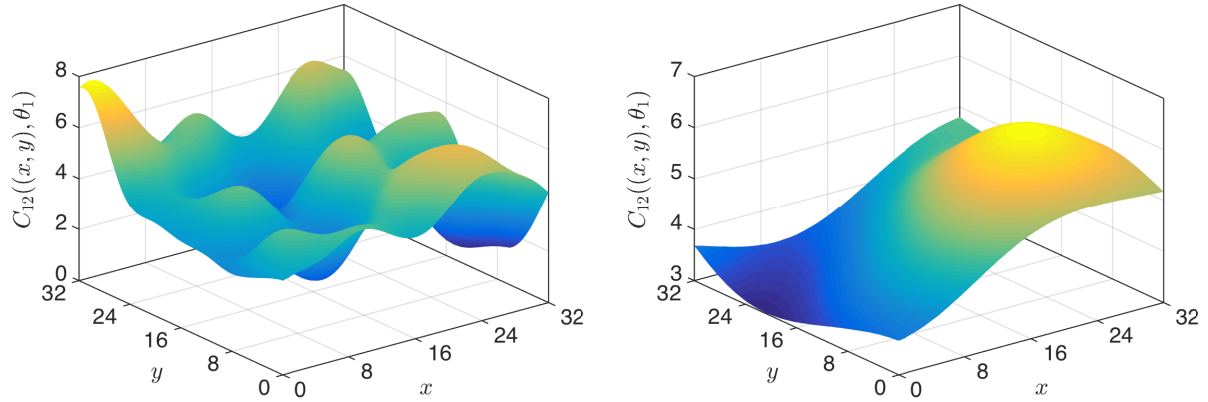


Figure 4.2: One realization of random field $\{C_{12}((x, y)), (x, y) \in \Omega\}$: $\mathcal{L}_c = 5.3$ (left panel), $\mathcal{L}_c = 16$ (right panel).

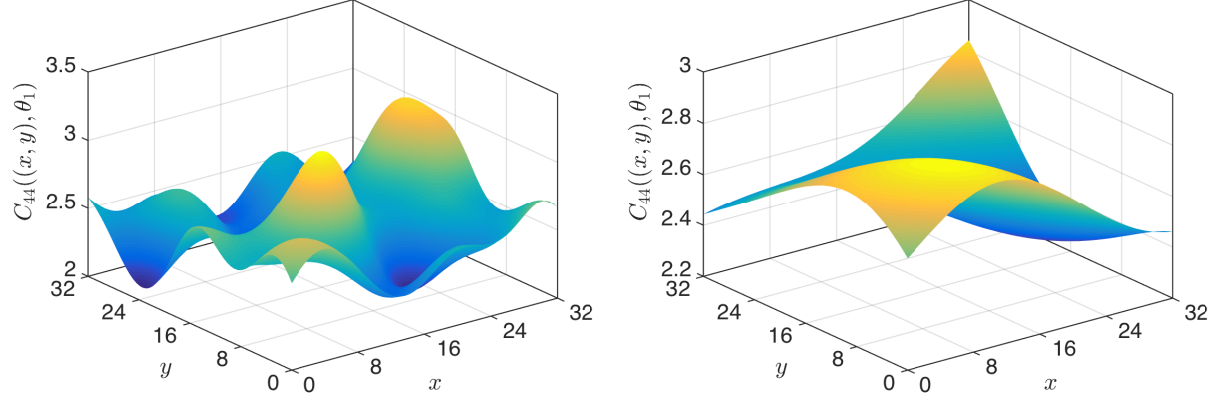


Figure 4.3: One realization of random field $\{C_{44}((x, y)), (x, y) \in \Omega\}$: $\mathcal{L}_c = 5.3$ (left panel), $\mathcal{L}_c = 16$ (right panel).

defined by the smallest correlation length exhibits fluctuations with higher spatial frequencies. Finally, the estimated correlation function of random field $\{C_{11}(\mathbf{x}), \mathbf{x} \in \Omega\}$ is shown in Fig. 4.4, where it is seen that the elasticity random field inherits a correlation structure that is similar to that of the input underlying Gaussian fields (as already observed in [9]).

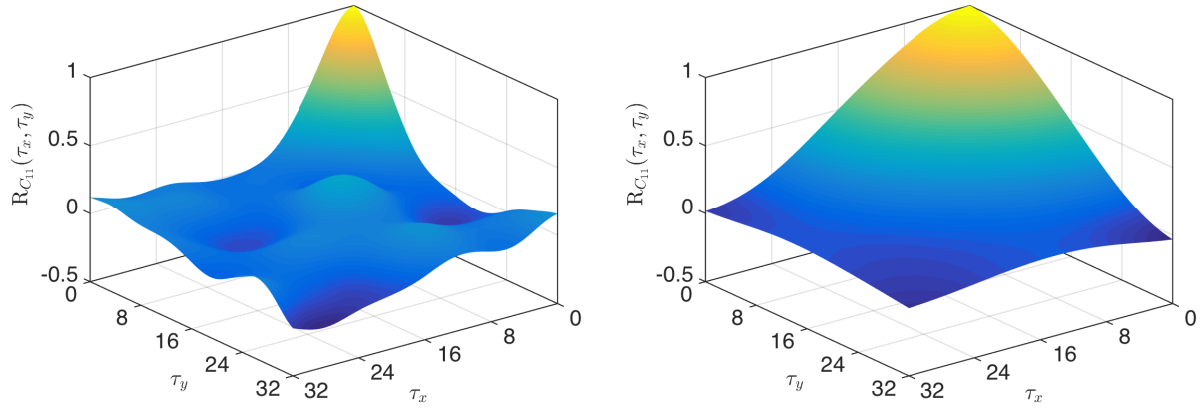


Figure 4.4: Plot of the estimated normalized correlation function for the random field $\{C_{11}(\mathbf{x}), \mathbf{x} \in \Omega\}$, for $\mathcal{L}_c = 5.3$ (left panel) and $\mathcal{L}_c = 16$ (right panel).

5 Conclusion

In this work, we have addressed the construction of approximate solutions for the Lagrange multipliers involved in some classes of information-theoretic non-Gaussian random field models. A general methodology for solving the optimization problems raised by entropy maximization (for the family of first-order marginal probability distributions) was first presented. Whereas the strategy is exemplified on tensor-valued elasticity fields exhibiting fluctuations in a given symmetry class, it can be readily extended for solving similar problems in formulations involving state spaces whose algebraic structure is induced by invariance properties. The results for all classes of practical interest (from isotropy to orthotropy) are then provided and are shown to be consistent with the ones algebraically obtained for the triclinic case. A numerical application to the case of transversely isotropic materials is further presented and illustrates the relevance of the approach. These derivations are subsequently used to establish a set of mathematical properties related to the ISDE-based sampling scheme proposed in [9]. This paper offers a self-contained treatment of information-theoretic probabilistic models and allows for forward simulations of such random field representations in uncertainty quantification for stochastic boundary value problems. It further allows, through a simple parametrization relying on a single statistically-sound parameter, for a substantial reduction of the computational cost whenever these models are plugged into optimization problems related to statistical inverse identification.

References

- [1] R. BALIAN, *Random Matrices and Information Theory*, Il Nuovo Cimento B, 57 (1968), pp. 183–193.
- [2] ———, *From Microphysics to Macrophysics (Vol. I)*, Springer, 1992.
- [3] K. BURRAGE, I. LENANE, AND G. LYTHE, *Numerical methods for second-order stochastic differential equations*, SIAM Journal on Scientific Computing, 29 (2007), pp. 245–264.
- [4] R.H. CAMERON AND W.T. MARTIN, *The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals*, The Annals of Mathematics, Second Series, 48 (1947), pp. 385–392.
- [5] R. DE VOGELAERE, *Methods of integration which preserve the contact transformation property of the Hamiltonian equations*, Tech. Report 4, University of Notre Dame, 1956.
- [6] R. GHANEM AND P.D. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, 1991.
- [7] T. GNEITING, H. SEVEIKOVA, D.B. PERCIVAL, M. SCHLATHER, AND Y. JIANG, *Fast and exact simulation of large gaussian lattice systems in $\mathbb{R}^2$: Exploring the limits*, Journal of Computational and Graphical Statistics, 15 (2006), pp. 483–501.
- [8] J. GUILLEMINOT AND C. SOIZE, *On the statistical dependence for the components of random elasticity tensors exhibiting material symmetry properties*, Journal of elasticity, 111(2) (2012), pp. 109–130.
- [9] ———, *Random fields with symmetry properties: Application to the mesoscopic modeling of elastic random media*, SIAM Multiscale Modeling and Simulation, 11(3) (2013), pp. 840–870.
- [10] ———, *Isde-based generator for a class of non-gaussian vector-valued random fields in uncertainty quantification*, SIAM Journal on Scientific Computing, 36 (2014), pp. 2763–2786.

- [11] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer, Berlin, 2002.
- [12] W. K. HASTINGS, *Monte carlo sampling methods using markov chains and their applications*, Biometrika, 109 (1970), pp. 57–97.
- [13] E.T. JAYNES, *Information theory and statistical mechanics*, Physical Review, 106/108 (1957), pp. 620–630/171–190.
- [14] ———, *The Probability Theory: the Logic of Science*, Cambridge University Press, Cambridge, 2003.
- [15] P.E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.
- [16] O.P. LE MAITRE AND O.M. KNIO, *Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics*, Springer, 2010.
- [17] N. METROPOLIS, W. ROSENBLUTH, M.N. ROSENBLUTH, A.H. TELLER, AND E. TELLER, *Equation of state calculations by fast computing machines*, J. Chem. Phys., 21 (1953), pp. 1087–1092.
- [18] G.N. MILSTEIN AND M.V. TRETYAKOV, *Stochastic Numerics for Mathematical Physics*, Springer, 2004.
- [19] M. MOAKHER AND A. N. NORRIS, *The closest elastic tensor of arbitrary symmetry to an elasticity tensor of lower symmetry*, Journal of Elasticity, 85 (2006), pp. 215–263.
- [20] G.O. ROBERTS, A. GELMAN, AND W.R. GILKS, *Weak convergence and optimal scaling of random walk metropolis algorithms*, The Annals of Applied Probability, 7 (1997), pp. 110–120.
- [21] C.E. SHANNON, *A mathematical theory of communication*, Bell System Technical Journal, 27 (1948), pp. 379–423/623–659.
- [22] C. SOIZE, *A nonparametric model of random uncertainties on reduced matrix model in structural dynamics.*, Probabilistic Engineering Mechanics, 15(3) (2000), pp. 277–294.
- [23] ———, *Maximum entropy approach for modeling random uncertainties in transient elastodynamics.*, Journal of the Acoustical Society of America, 109(5) (2001), pp. 1979–1996.
- [24] ———, *Non-gaussian positive-definite matrix-valued random fields for elliptic stochastic partial differential operators*, Computer Methods in Applied Mechanics and Engineering, 195 (2006), pp. 26–64.
- [25] ———, *Construction of probability distributions in high dimension using the maximum entropy principle: Applications to stochastic processes, random fields and random matrices*, International Journal for Numerical Methods in Engineering, 76 (2008), pp. 1583–1611.
- [26] ———, *Stochastic Models of Uncertainties in Computational Mechanics*, American Society of Civil Engineers, Reston, 2012.
- [27] D. TALAY, *Simulation and Numerical Analysis of Stochastic Differential Systems: a Review*, in Probabilistic Methods in Applied Physics, P. Kree and W. Wedig, eds., vol. 451 of Lecture Notes in Physics, Springer-Verlag, 1995, pp. 63–106.

- [28] L. VERLET, *Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules*, Physical Review, 159 (1967), pp. 98–103.
- [29] L. WALPOLE, *Fourth-rank tensors of the thirty-two crystal classes : Multiplication tables*, Proc. R. Soc. Lond. A, 391 (1984), pp. 149–179.