

Identification of Elastic Properties in the Belief Function Framework

Liqi Sui^a, Pierre Feissel^a, Thierry Denœux^{b,*}

^a *Université de Technologie de Compiègne, CNRS,
UMR 7337 Roberval, 60203 Compiègne, France*

^b *Université de Technologie de Compiègne, CNRS,
UMR 7253 Heudiasyc, 60203 Compiègne, France*

Abstract

Handling the uncertainty of information sources is a key issue in parameter identification. In this work, we address this issue using the theory of belief functions. First, measurement information is described through likelihood-based belief functions, and prior information is represented by an arbitrary belief function. Second, both belief functions are combined by Dempster's rule using point-cloud representation of focal sets and Monte Carlo simulation. Lastly, to summarize the combined belief function, we propose to find the minimal-area region in the parameter space, whose belief and plausibility values exceed given thresholds. As compared to Bayesian inference, this approach is more flexible, as it allows us to specify weak prior information. Experimental results show that it is also more robust than Bayesian inference to unreliable prior information.

Keywords: Mechanics, measurement field, prior information, aleatory uncertainty, epistemic uncertainty, likelihood, Dempster-Shafer theory, evidence theory, random set.

1. Introduction

In computational mechanics, more and more complex material models are used to meet the need for more predictive and accurate mechanical simulations. The parameters of these models have to be identified from experiments; consequently, identification of material properties remains a top-priority objective. As models get more complex, more information is needed from more elaborate experiments. In recent years, the use of full-field displacement measurements (based, in particular, on Digital Image Correlation (DIC) [1]) has rapidly spread

*Corresponding author. Université de Technologie de Compiègne, Rue Roger Couttolenc, CS 60319, 60203 Compiègne cedex, France.

Email addresses: suiliqi@126.com (Liqi Sui), pierre.feissel@utc.fr (Pierre Feissel), thierry.denoeux@utc.fr (Thierry Denœux)

within the experimental mechanics community. Such data (a full field of displacement over an area of the loaded structure) are very rich in comparison to previous measurements (e.g., strain gauges); they offer more possibilities in terms of material characterization. However, straightforward identification is no longer possible and dedicated approaches need to be developed.

Identification is typically an inverse problem [2]. Inverse problems are often ill-posed [3], in the sense that their solution may suffer from a lack of existence, unicity or continuity with respect to the data. To ensure the existence of a solution, inverse problems are often reformulated as optimization problems (e.g., least-squares). To tackle the non-unicity issue, we need to introduce some form of regularization [4, 5]. An effective way to regularize an inverse problem is to take into account prior information. This can be done through Bayesian inference [6], which makes it possible to deal with uncertainties in both the measurements and the prior information. However, under the Bayesian approach, uncertain information has to be described by probabilities, which is quite restrictive. Over the years, different alternative formalisms for modeling epistemic uncertainties have been developed, including interval analysis [7], possibility theory [8], and p-boxes [9]. In this work, we aim to develop new tools for parameter identification allowing us to incorporate weak prior knowledge in the belief function framework, which encompasses all the formalisms mentioned above.

This paper focuses on developing a method to identify material elastic parameters from displacement fields. There are two main challenges: (1) faithfully representing prior information, and (2) quantifying uncertainties from various sources and propagating them. Recently, a lot of methods have been proposed to identify material properties from kinematic fields, including Finite Element Model Updating (FEMU) [10, 11, 12], the Constitutive Relation Error Method (CREM) [13, 14], the Modified Constitutive Relation Error Method (M-CREM) [15], the Equilibrium Gap Method (EGM) [16, 17, 18], and the Virtual Fields Method (VFM) [19, 20]. These methods do not specifically address the issues of exploiting prior information and quantifying uncertainty. The Bayesian framework [21, 22] accounts for prior information and quantifies uncertainty using probabilities. However, the probabilistic formalism is limited when it comes to describing near-ignorance situations. This limitation, discussed by many authors (see, e.g., [23, 24, 25]), has motivated the development of alternative theories of uncertainty in the last quarter of the 20th century. In particular, the theory of *belief functions*, also referred to as *Dempster-Shafer theory* or *Evidence Theory* [26, 23, 27, 28], offers a suitable framework for encoding and quantifying both epistemic and aleatory (random) uncertainty. Moreover, it includes comprehensive mechanisms for combining information from multiple sources.

Most previous studies using the theory of belief functions in mechanics have focused on the *propagation* of uncertainty through mechanical models, i.e., on the *direct* problem. For instance, in [29, 30], the authors consider the problem of computing the uncertain response of a numerical model with uncertain parameters, seen as a black box. Other authors have used belief functions to predict the remaining lifetime of mechanical systems and quantify the corresponding

55 uncertainty [31, 32]. However, to our knowledge, none of these studies has been devoted to handling uncertainty in *inverse* problems, i.e., *identifying* the parameters of mechanical models from measurements based on the theory of belief functions, which is the topic of this paper¹. In Section 2, we introduce the basic concepts needed in this paper and we outline our identification strategy. Section 60 3 describes the numerical implementation of this strategy, including the description and discretization of multidimensional random sets and the summarization of the combined belief function. In Section 4, we present two applications to static tensile tests of homogeneous and heterogeneous plates. Section 5 concludes the paper.

65 2. Identification Strategy

This section is composed of two parts. In Section 2.1, we introduce the general identification framework, including the mechanical models and the basic concepts of the theory of belief functions. The encoding and merging of information are then addressed in Section 2.2.

70 2.1. Framework and Proposed Strategy

We consider the identification of elastic parameters of a solid body under static loading based on displacement field data obtained using DIC [35, 36, 37]. This framework corresponds to the standard case of a mechanical test on a specimen. The first step is to model the experiment, through the continuum 75 mechanics partial differential equations and boundary conditions that describe the mechanical behavior of the specimen. For the sake of simplicity, we will limit ourselves to the case where the available data make it possible to define a well-posed direct problem. Otherwise, some methods have been proposed to deal with lack of data [38], but they will not be addressed here. This direct 80 problem is parametrized by unknown elastic parameters collected in a vector θ , taking values in a set Θ .

Mechanical model and information for identification. We sum up the modeling of the experiment as follows. The specimen is considered as a 2D domain Ω . The displacement data \tilde{u} are measured over a domain Ω_m , which can be Ω itself or a subset of Ω . They have to be compared to the displacement computed from the model for a given θ , denoted by $u(\theta)$. The displacement $u(\theta)$ is the solution of the continuum mechanics problem on domain Ω , driven by the standard equations: equilibrium, Neumann and Dirichlet boundary equations and constitutive equation. The latter is assumed here to be linear elastic and isotropic so that it can be written as

$$C\varepsilon = \lambda \text{tr}(\varepsilon)I + 2\mu\varepsilon \quad \text{on all the domain,} \quad (1)$$

¹This paper is based on the first author's PhD thesis [33], and is an extended version of Ref. [34].

where \mathbf{I} denotes the identity matrix, tr is the trace operator and $\boldsymbol{\varepsilon}$ is the strain field associated to the displacement ($\boldsymbol{\varepsilon} = \frac{1}{2}(\nabla \mathbf{u} + \nabla^T \mathbf{u})$). \mathbf{C} is the Hook tensor corresponding to the elastic behaviour and (λ, μ) are the two Lamé parameters that can be constant over Ω or not. Vector $\boldsymbol{\theta}$ collects the parameters that allow us to describe the λ and μ fields.

The continuum mechanics direct problem does not usually have a closed-form solution. Therefore, an approximate solution has to be computed based on a discretization of the problem. The standard method used here is the finite element (FE) method [39, 40], which can provide an accurate estimation of $\mathbf{u}(\boldsymbol{\theta})$. The FE method is based on an approximation of the displacement field thanks to a mesh of Ω , where \mathbf{u} is described from its interpolation between its values at the nodes of the mesh (see Figure 4(b) in Section 4 for an illustration, where the nodes correspond to the corners of the triangles).

The purpose of identification through the inverse approach is to find the elastic parameters $\boldsymbol{\theta}$ such that $\mathbf{u}(\boldsymbol{\theta})$ is as close as possible to the measured displacement $\tilde{\mathbf{u}}$, taking uncertainty into account. The available information can be divided into three categories:

1. *Theoretical information*, considered as reliable and deterministic. For any given $\boldsymbol{\theta}$, this information is described by $\mathbf{u}(\boldsymbol{\theta})$ and corresponds to the displacement predicted by the model.
2. *Experimental information*, tainted with aleatory uncertainty (random error). It is assumed that the measured displacement can be written as

$$\tilde{\mathbf{u}} = \mathbf{u}(\boldsymbol{\theta}) + \mathbf{e}, \quad (2)$$

where \mathbf{e} is the measurement error, which is a random vector with a known probability distribution $\mathbf{e} \sim P_e(\mathbf{e})$.

3. *Background information* corresponding to prior knowledge on $\boldsymbol{\theta}$; it is subject to epistemic uncertainty.

Standard identification methods. Based on parametric model (2), the likelihood function after observing measurement vector $\tilde{\mathbf{u}}$ is defined from the error probability distribution $P_e(\mathbf{e})$ as

$$L(\boldsymbol{\theta}; \tilde{\mathbf{u}}) = P_e(\mathbf{u}(\boldsymbol{\theta}) - \tilde{\mathbf{u}}). \quad (3)$$

Deterministic identification of $\boldsymbol{\theta}$ then consists in finding the most likely parameter vector

$$\hat{\boldsymbol{\theta}}_L = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta}; \tilde{\mathbf{u}}). \quad (4)$$

If \mathbf{e} has a Gaussian distribution, the maximization problem in Eq. (4) is equivalent to the minimization of a least-square criterion; this is the case, e.g., in the FEMU method [41, 42, 43]. In order to quantify uncertainty and handle prior information, Bayesian inference applies Bayes' theorem to derive a posterior distribution on parameters $\boldsymbol{\theta}$ based on the measurement field $\tilde{\mathbf{u}}$. To do so, prior knowledge has to be expressed as a probability distribution, which amounts to

considering θ as a random vector with distribution $p(\theta)$. The posterior distribution is then

$$p(\theta|\tilde{\mathbf{u}}) = c \cdot p(\theta) \cdot L(\theta; \tilde{\mathbf{u}}), \quad (5)$$

where c is a normalizing constant. We can see that the introduction of prior probability distribution $p(\theta)$ is unavoidable in Bayesian inference. However, the use of probabilities to represent weak knowledge has been questioned (see, e.g., [44, 45, 46, 47]). This ongoing debate motivates us to explore alternative strategies for inference.

110

Theory of belief functions. In the following, we will use the belief function framework to model different types of uncertainty. Considering a probability space (S, σ_S, P) , where σ_S is a σ -algebra of subsets of S ; a non-empty set Θ and its power set 2^Θ , a strongly measurable multi-valued mapping $\Gamma: S \rightarrow 2^\Theta$ is called a *random set* [48]. For any measurable subset $\mathbf{A} \subseteq \Theta$, the uncertainty of the proposition $\theta \in \mathbf{A}$ can be quantified by *belief* and *plausibility* functions [26, 23, 49],

$$Bel(\mathbf{A}) = \frac{P(\{\omega \in S \mid \Gamma(\omega) \subseteq \mathbf{A}, \Gamma(\omega) \neq \emptyset\})}{P(\{\omega \in S \mid \Gamma(\omega) \neq \emptyset\})}, \quad (6a)$$

$$Pl(\mathbf{A}) = \frac{P(\{\omega \in S \mid \Gamma(\omega) \cap \mathbf{A} \neq \emptyset\})}{P(\{\omega \in S \mid \Gamma(\omega) \neq \emptyset\})} = 1 - Bel(\overline{\mathbf{A}}), \quad (6b)$$

where $\overline{\mathbf{A}}$ denotes the complement of \mathbf{A} . The quantity $Bel(\mathbf{A})$ is interpreted as the degree of support given to the proposition $\theta \in \mathbf{A}$, while the quantity $Pl(\mathbf{A})$ measures the lack of support given to the proposition $\theta \notin \mathbf{A}$. Each set $\Gamma(\omega)$ for $\omega \in S$ is called a *focal set* of the belief function. The random set Γ and the corresponding belief function Bel are said to be *consonant* if, for any ω and ω' in S , we have either $\Gamma(\omega) \subseteq \Gamma(\omega')$ or $\Gamma(\omega') \subseteq \Gamma(\omega)$. The restriction of the plausibility function to singletons is called the *contour function*,

$$pl(\theta) = Pl(\{\theta\}), \quad \forall \theta \in \Theta. \quad (7)$$

In the case of a consonant belief function, we have

$$Pl(\mathbf{A}) = \sup_{\theta \in \mathbf{A}} pl(\theta), \quad (8)$$

for all $\mathbf{A} \subseteq \Theta$. A consonant belief function can, thus, be represented by its contour function.

The formal equivalence between random sets and belief functions has been shown by Nguyen in [48]: for any random set Γ , there exists a belief function defined by (6a) and, conversely, for any belief function, there exists a random set that generates it. In this paper, we will adopt the “random set” or “belief function” points of view, depending on the context.

115

Proposed inference strategy. Our strategy can be decomposed into three steps. In the first step, we aim at expressing both measurement and prior information using belief functions. In a second step, we merge these two pieces of information in order to construct a posterior belief function combining prior and measurement information. Finally, we summarize this combined belief function set to provide usable and interpretable information about parameter θ .

2.2. Measurement, Prior Information and Merging

In this section, we show how to express both measurement information and prior knowledge in the belief function framework, and how to combine these two pieces of information.

Measurement information. The measurement $\tilde{\mathbf{u}}$ is assumed to be a noisy version of the theoretical displacement $\mathbf{u}(\theta)$, corrupted by additive error (see Eq. (2)). The error is assumed to be Gaussian, with known covariance \mathbf{D} : $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$. The value of \mathbf{D} can actually be estimated by the DIC algorithm used for the measurement of the displacement field [50, 51]. The evidential likelihood-based approach, introduced in [23], represent the uncertainty on θ by a consonant belief function, whose contour function equals the normalized likelihood function

$$pl_L(\theta; \tilde{\mathbf{u}}) = \frac{L(\theta; \tilde{\mathbf{u}})}{\sup_{\theta} L(\theta; \tilde{\mathbf{u}})}. \quad (9)$$

Here, $L(\theta; \tilde{\mathbf{u}})$ is equal to

$$L(\theta; \tilde{\mathbf{u}}) = (2\pi \det \mathbf{D})^{-1/2} \exp \left[-\frac{1}{2} (\mathbf{u}(\theta) - \tilde{\mathbf{u}})^T \mathbf{D}^{-1} (\mathbf{u}(\theta) - \tilde{\mathbf{u}}) \right]. \quad (10)$$

In [28, 52], the author showed that Eq. (9) can be derived from three basic principles: the likelihood principle, compatibility with Bayes' rule when a prior probability distribution is available and the minimal commitment principle. This method of inference is more flexible than Bayesian inference. In particular, it does not assume the existence of a prior probability distribution. However, it boils down to Bayesian inference if a probabilistic prior is provided.

Prior information. Measurement information expressed by the likelihood-based belief function (9) can be combined with prior information also expressed in the belief function setting [30, 53]. In theory, the prior belief function can be arbitrary but, for practical reasons, it is convenient to use a simple form such as a consonant belief function [54, 55]. Thanks to Eq. (8), a consonant belief function can be represented by its contour function $pl : \Theta \rightarrow [0, 1]$. In the case of a multidimensional parameter $(\theta_1, \dots, \theta_d) \in \mathbb{R}^d$, it may be difficult to elicit directly the joint contour function pl . Rather, we may elicit its marginals

$$pl_i(\theta_i) = \sup_{\{\theta_j, j \neq i\}} pl(\theta_1, \dots, \theta_d). \quad (11)$$

From (11), we can deduce that $pl(\theta_1, \dots, \theta_d) \leq pl_i(\theta_i)$ and, thus,

$$pl(\theta_1, \dots, \theta_d) \leq \min\{pl_1(\theta_1), \dots, pl_d(\theta_d)\}.$$

Without any other information on pl , a conservative attitude leads to selecting the *least committed* one [56], defined by

$$pl(\theta_1, \dots, \theta_d) = \min\{pl_1(\theta_1), \dots, pl_d(\theta_d)\}, \quad (12)$$

for all $\theta_1, \dots, \theta_d$.

135 It is sometimes desirable to weaken, or *discount* [23] a belief function, to express partial lack of confidence in a source of information. Let us assume that, for instance, an expert gives us a belief function Bel about a parameter θ and we have a degree of confidence $1 - \alpha \in [0, 1]$ in the expert's opinion. The discounted belief function is defined as ${}^\alpha Bel = (1 - \alpha)Bel + \alpha Bel_?$, where
 140 $Bel_?$ is the *vacuous* belief function defined by $Bel_?(A) = 0$ for all $A \subset \Theta$. If Bel is consonant with contour function pl , the discounted contour function is ${}^\alpha pl(\theta) = \max[\alpha, pl(\theta)]$ for all $\theta \in \Theta$.

Information merging. The likelihood-based and prior belief functions defined, respectively, by (9) and (12) can be merged by *Dempster's rule* [26] [23, Chapter 3] [57], which is the standard operation for combining pieces of information from independent sources in the Dempster-Shafer framework. Let us consider two random sets $(S_k, \sigma_{S_k}, P_k, \Gamma_k)$, $k = 1, 2$. Let $S = S_1 \times S_2$ be the product space, $P = P_1 \times P_2$ the product measure on $\sigma_S = \sigma_{S_1} \otimes \sigma_{S_2}$, and Γ_p the multi-valued mapping defined by $\Gamma_p(\omega_1, \omega_2) = \Gamma_1(\omega_1) \cap \Gamma_2(\omega_2)$ for all $(\omega_1, \omega_2) \in S$. The combined random set is defined as $(S, \sigma_S, P, \Gamma_p)$. It induces the following belief and plausibility functions: for any $A \subseteq \Theta$,

$$Bel(A) = \frac{P(\{(\omega_1, \omega_2) \in S \mid \Gamma_p(\omega_1, \omega_2) \subseteq A, \Gamma_p(\omega_1, \omega_2) \neq \emptyset\})}{P(\{(\omega_1, \omega_2) \in S \mid \Gamma_p(\omega_1, \omega_2) \neq \emptyset\})}, \quad (13a)$$

$$Pl(A) = \frac{P(\{(\omega_1, \omega_2) \in S \mid \Gamma_p(\omega_1, \omega_2) \cap A \neq \emptyset\})}{P(\{(\omega_1, \omega_2) \in S \mid \Gamma_p(\omega_1, \omega_2) \neq \emptyset\})}. \quad (13b)$$

The *degree of conflict* is the quantity

$$\kappa = P(\{(\omega_1, \omega_2) \in S \mid \Gamma_p(\omega_1, \omega_2) = \emptyset\}). \quad (14)$$

It is a measure of the inconsistency between the two sources of information.

Remark 1. As shown in [58] and [28], combining the likelihood-based consonant belief function defined by contour function (9) with a probabilistic prior $p(\theta)$
 145 using Dempster's rule (13) yields the posterior probability distribution $p(\theta|\tilde{u})$. Dempster's rule (13) thus generalizes Bayes' rule (5). Consequently, Bayesian reasoning and Dempster-Shafer reasoning yield the same result if provided with the same prior information. However, thanks to its greater expressivity, the
 150 belief function formalism makes it possible to represent weak forms prior knowledge that cannot be expressed in the probabilistic framework.

Remark 2. *Dempster’s rule could be used not only to combine measurement and prior information, but also to combine prior information about different parameters. In the case of a multidimensional parameter, we could specify the prior joint belief function by fixing its marginals and assuming that they are independent. The joint belief function would then be obtained by combining the marginals using Dempster’s rule. If the marginals are consonant, the joint is then no longer consonant. While this approach is defensible, we prefer to assume the consonance of the joint and use the least commitment principle, which dispenses us from assuming independence and results in a simpler prior joint belief function.*

We have presented an identification strategy based on (i) the representation of measurement information by a consonant belief function whose contour function is identical to the normalized likelihood function, (ii) the representation of prior information by an arbitrary belief function (that can also be chosen to be consonant for simplicity), and (iii) the combination of these two belief functions by Dempster’s rule. The next section will be devoted to the numerical implementation of this strategy.

3. Numerical Implementation

We propose to implement the identification strategy outlined in Section 2 using Monte Carlo simulation, by randomly generating focal sets of the likelihood-based and prior belief functions, and computing their intersections. For that purpose, we need a method to represent and intersect regions of the parameter space of arbitrary shape. Such a method will first be presented in Section 3.1. The implementation of Dempster’s rule will then be described in Section 3.2. Finally, the summarization of the combined belief function will be addressed in Section 3.3.

3.1. Description of Multidimensional Belief Functions

To combine belief functions defined in multidimensional spaces, we need a method to describe domains with arbitrary shapes without losing too much information. There are two main categories of methods to describe a geometric domain. The first one makes use of functions to parametrize a domain, e.g., regular shape functions (triangle, square, ellipsoid, etc.); the level-set method [59] also belongs to this category. The second class of methods is based on discretizing a domain of interest into small regions; a geometric domain is then described by assigning a value (1, -1 or 0) to each region. In particular, *point clouds* are sets of data points in a coordinate system, which allow us to make detailed scans of complex objects [60].

Methods based on parametrized functions lead to difficult implementation of Boolean operations (intersection, union, and complementation), which are needed for Dempster-Shafer reasoning (see Section 3.2). There are three reasons for choosing point clouds to describe the domains. Firstly, a point cloud can be used to describe any domain in arbitrary dimensions using indicator functions.

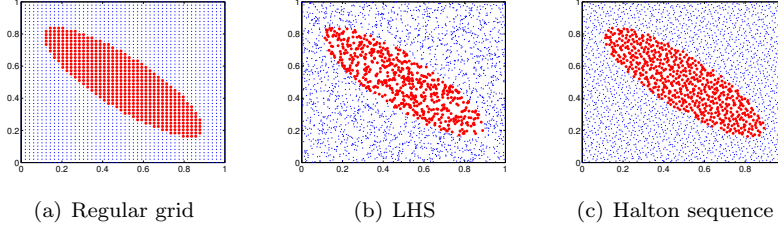


Figure 1: Ellipse described by point clouds.

Given a domain $\mathbf{A} \subseteq \Theta$, we can associate to any point θ_i in the cloud a binary value $v_{\mathbf{A}}(i)$ such that

$$v_{\mathbf{A}}(i) = \begin{cases} 1 & \text{if } \theta_i \in \mathbf{A}, \\ 0 & \text{if } \theta_i \notin \mathbf{A}. \end{cases} \quad (15)$$

Domain \mathbf{A} can then be described by the vector $\mathbf{v}_{\mathbf{A}} = (v_{\mathbf{A}}(1), \dots, v_{\mathbf{A}}(N))$, where N is the number of points. Secondly, there are many methods and techniques to generate point clouds, including grid sequences, random sequences and quasi-random sequences [61, 62, 63]. Thirdly, the point-cloud representation lends itself to easy and fast Boolean operations.

Let us mention the various techniques to generate point clouds. *Regular grids* have been applied historically on discretized general domains [61]. The finer the grid, the more accurate the discretization, but the computational and storage cost quickly become prohibitive in high dimension due to the curse of dimensionality. Random sampling is another way to generate point clouds, but it also becomes inefficient in high-dimensional spaces. *Latin hypercube sampling (LHS)* [62, 64] is a form of stratified sampling, in which we divide each unidimensional domain into segments, and we insure that there is only one sample value in each segment of each dimension. *Quasi-random sequence* generation methods [63, 65] make use of deterministic low-discrepancy sequences to generate sample points. The low-discrepancy sequences are more “uniformly distributed” than random ones: consequently, they usually perform better than random sequences in Monte Carlo simulations. *Halton sequences* [66] are particular low-discrepancy sequences that will be used in this work.

After generating a point cloud, the domains are described by their discretized indicator functions defined by Eq. (15). Figure 1 shows an ellipse described by point clouds generated by a regular grid, LHS, and a Halton sequence. We use the estimator of relative volume (surface) error to assess the point cloud performance:

$$\epsilon = \frac{|N_{in}/N_{total} - \mathcal{V}_{rel}|}{\mathcal{V}_{rel}}, \quad (16)$$

where N_{in} is the number of points inside the domain and N_{total} is the total number of points in the cloud. The value \mathcal{V}_{rel} is the relative volume (surface) of the domain to be described with respect to the total volume described by

the whole point cloud. One hundred ellipsoids (hyper-ellipsoids) with different sizes and orientations in three, four and ten dimensions were described by point clouds generated by regular grids, LHS, and standard Halton sequences. The relations between the number of points and the error estimators are presented in
 215 Figures 2(a), 2(b), and 2(c) for numbers of dimensions equal, respectively, to 3, 4 and 10. (With 10 dimensions, the regular grid approach cannot be implemented anymore). We can see that the Halton sequences have the best performance in all three cases. These results coincide with the findings reported in [67, 68], where it was already noticed that random sequences outperform the regular grids in
 220 high-dimensional cases and low-discrepancy sequences have smaller error bounds than the average error bounds of random sequences. Therefore, in this paper, we will use a point cloud generated using Halton sequences to describe focal sets.

A point cloud can be stored in an $N \times D$ matrix \mathbf{X} of point coordinates, where N is the number of points, and D is the number of dimensions. A domain
 225 \mathbf{A} can then be represented by the values of the indicator function at each point of the cloud, stored in an N -vector $\mathbf{v}_{\mathbf{A}}$. Operations on sets can be implemented using Boolean operations on the components of their vectors. For instance, the intersection of two domains \mathbf{A} and \mathbf{B} is represented by $\mathbf{v}_{\mathbf{A}} \wedge \mathbf{v}_{\mathbf{B}}$ with components
 230 $\min\{v_{\mathbf{A}}(i), v_{\mathbf{B}}(i)\}$ for $i = 1, \dots, N$. The conditions $\mathbf{A} \neq \emptyset$ and $\mathbf{A} \subseteq \mathbf{B}$ translate, respectively, to $\max_i v_{\mathbf{A}}(i) = 1$ and $\min_i I[v_{\mathbf{A}}(i) \leq v_{\mathbf{B}}(i)] = 1$, where $I[\cdot]$ is the indicator function.

3.2. Monte Carlo Approximation of Dempster's Rule

Dempster's rule can be approximated by Monte Carlo simulation [69]. Consider two independent random sets $\Gamma_1(\omega_1)$ and $\Gamma_2(\omega_2)$, where ω_1 and ω_2 are independent random variables with distributions P_1 and P_2 . For each ω_j ($j = 1, 2$), we assume that we can determine the set of points approximating $\Gamma_j(\omega_j)$, and the corresponding vector representation $\mathbf{v}_{\Gamma_j(\omega_j)}$. For instance, if Γ_j is consonant with contour function π_j , then $\mathbf{v}_{\Gamma_j(\omega_j)}(i) = I[\pi_j(\theta_i) \geq \omega_j]$. The combined random set $\Gamma_p(\omega_1, \omega_2) = \Gamma_1(\omega_1) \cap \Gamma_2(\omega_2)$ can be approximated by iteratively drawing pairs (ω_1, ω_2) from the joint distribution $P_1 \times P_2$. For each pair, the point-cloud representations $\mathbf{v}_{\Gamma_1(\omega_1)}$ and $\mathbf{v}_{\Gamma_2(\omega_2)}$ of focal sets $\Gamma_1(\omega_1)$ and $\Gamma_2(\omega_2)$ are determined. The intersection $\Gamma_p(\omega_1, \omega_2)$ is represented by vector $\mathbf{v}_{\Gamma_1(\omega_1)} \wedge \mathbf{v}_{\Gamma_2(\omega_2)}$. This vector is kept in memory if $\Gamma_p(\omega_1, \omega_2)$ is nonempty,
 240 which can be checked by the condition $\max_i (\mathbf{v}_{\Gamma_1(\omega_1)} \wedge \mathbf{v}_{\Gamma_2(\omega_2)})(i) = 1$. The process is iterated until M nonempty focal sets have been obtained. This procedure is described formally in Algorithm 1.

The combined random set is finally represented by M vectors

$$\mathbf{v}_j = (v_j(1), \dots, v_j(N)), \quad j = 1, \dots, M,$$

corresponding to M nonempty focal sets. Given a subset $\mathbf{A} \subset \Theta$, focal set j is included in \mathbf{A} iff $\min_i I[v_j(i) \leq v_{\mathbf{A}}(i)] = 1$, and it has a non empty intersection with \mathbf{A} iff $\max_i \min\{v_j(i), v_{\mathbf{A}}(i)\} = 1$. The degrees of belief and plausibility

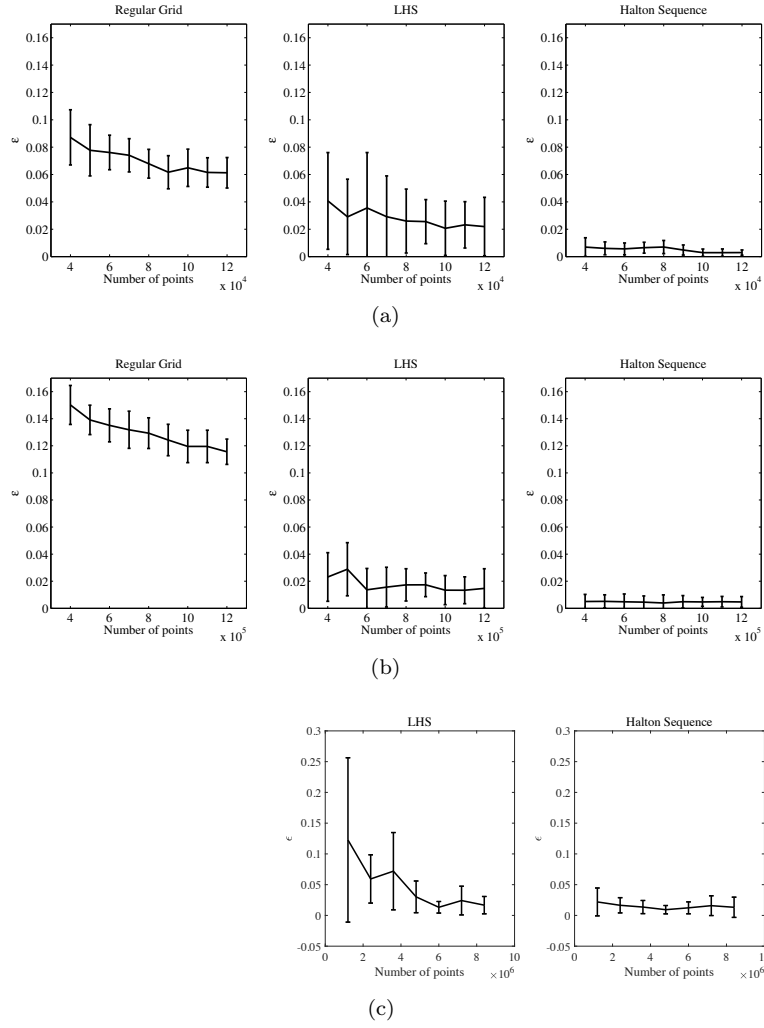


Figure 2: Relative volume error vs number of points in 3D (a), 4D (b) and 10D (c).

Algorithm 1 Monte Carlo algorithm for approximating Dempster's rule.

Require: M

$j \leftarrow 0, k \leftarrow 0$

while $j < M$ **do**

$k \leftarrow k + 1$

Draw (ω_1, ω_2) from $P_1 \times P_2$

Get $\Gamma_1(\omega_1), \Gamma_2(\omega_2)$ and their point-cloud representations $\mathbf{v}_{\Gamma_1(\omega_1)}$ and $\mathbf{v}_{\Gamma_2(\omega_2)}$

Compute the point-cloud representation $\mathbf{v}_{\Gamma_1(\omega_1)} \wedge \mathbf{v}_{\Gamma_2(\omega_2)}$ of $\Gamma_p(\omega_1, \omega_2)$

if $\max_i (\mathbf{v}_{\Gamma_1(\omega_1)} \wedge \mathbf{v}_{\Gamma_2(\omega_2)})(i) = 1$ **then**

$j \leftarrow j + 1$

$\mathbf{v}_j \leftarrow \mathbf{v}_{\Gamma_1(\omega_1)} \wedge \mathbf{v}_{\Gamma_2(\omega_2)}$

end if

end while

$N_s \leftarrow k$

return $\{\mathbf{v}_j \mid j = 1, \dots, M\}, N_s$

can be approximated by the proportion of focal sets, respectively, included in \mathbf{A} , and intersecting \mathbf{A} [69], and can thus be computed as follows,

$$\widehat{Bel}(\mathbf{A}) = \frac{1}{M} \# \left\{ 1 \leq j \leq M \mid \min_i I[v_j(i) \leq v_{\mathbf{A}}(i)] = 1 \right\} \quad (17a)$$

$$\widehat{Pl}(\mathbf{A}) = \frac{1}{M} \# \left\{ 1 \leq j \leq M \mid \max_i \min \{v_j(i), v_{\mathbf{A}}(i)\} = 1 \right\}. \quad (17b)$$

The only parameter to be determined in Algorithm 1 is the number M of focal sets. Here, we will discuss the relationship between sample size and numerical error using an example of reconstruction of a combined contour function. Let pl_1 and pl_2 be the contour functions of the two belief functions to be combined. Let pl denote the contour function after combination by Dempster's rule. The following relation holds,

$$pl(\boldsymbol{\theta}) = \frac{pl_1(\boldsymbol{\theta})pl_2(\boldsymbol{\theta})}{1 - \kappa}, \quad (18)$$

where κ is the degree of conflict (14), which can be estimated by $\hat{\kappa} = 1 - M/N_s$. The contour function pl can be estimated at $\boldsymbol{\theta}_i$ by

$$\widehat{pl}(\boldsymbol{\theta}_i) = \frac{1}{M} \sum_{j=1}^M v_j(i). \quad (19)$$

The relative numerical error can be defined as

$$\epsilon_{MC} = \frac{\frac{1}{N} \sum_{i=1}^N \left| (1 - \hat{\kappa}) \widehat{pl}(\boldsymbol{\theta}_i) - pl_1(\boldsymbol{\theta}_i)pl_2(\boldsymbol{\theta}_i) \right|}{\max_i pl_1(\boldsymbol{\theta}_i)pl_2(\boldsymbol{\theta}_i)}. \quad (20)$$

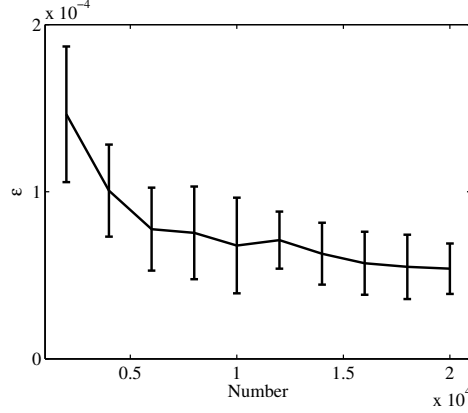


Figure 3: Relative numerical error of Monte Carlo sampling as a function of number M of draws, for the combination of two unidimensional trapezoidal contour functions.

It is clear that $\epsilon_{MC} = 0$ in the case of perfect approximation, i.e., when $\hat{pl}(\theta_i) = pl(\theta_i)$ for all i and $\hat{\kappa} = \kappa$. the numerator in the right-hand side of (20) ensures that the error does not become artificially small in case of high conflict, i.e., when $pl_1(\theta_i)pl_2(\theta_i)$ is close to zero for all i . As an illustration, we computed the relative numerical error ϵ_{MC} as a function of sample size M in the special case of the combination of two unidimensional trapezoidal contour functions. Samples of each size were drawn 10 times. The error as a function of sample size is shown in Figure 3. The curve corresponds to the mean values of the errors; the vertical intervals represent the standard deviations. We can see that the error curve has an “elbow point” around a sample size of 5000, and the convergence rate becomes very small after 15,000 iterations. These values can, of course vary for different problems. Formula (20) to compute the relative numerical error can be used as long as we can compute the contour functions at any θ_i . The numerical study performed here can be reproduced or used as a guideline for further studies.

3.3. Summarization of the combined belief function

After combining the likelihood-based belief function with the prior belief function using the Monte Carlo approximation of Dempster’s rule as explained in Section 3.2, we get an approximation of the combined belief function in the form of M vectors $\mathbf{v}_j = (v_j(1), \dots, v_j(N))$, $j = 1, \dots, M$. To exploit this information, we need to synthesize it so that it can be interpreted by the user. The most straightforward approach is to compute the maximum plausibility estimate of θ ,

$$\hat{\theta}_P = \arg \max_{\theta \in \Theta} pl_P(\theta), \quad (21)$$

where pl_P is the posterior contour function. This estimate can be approximated by θ_i maximizing $\hat{pl}_P(\theta_i)$ given by (19). However, a point estimate is a poor summary of the combined belief function, as it does not reflect the estimation

uncertainty. A more interesting summary may consist in the minimal-size subset of Θ whose belief and plausibility exceed some predefined values. Such sets will be called *plausible subsets*. Here, we present a tractable method to compute plausible subsets.

Plausible Subsets. As a way to extract more useful information from the posterior random set samples, we can find a minimal-size subset $\mathbf{R} \subseteq \Theta$ such that $Pl(\mathbf{R})$ and $Bel(\mathbf{R})$ are larger than given threshold values δ_{Pl} and δ_{Bel} . Let \mathcal{V} be the volume of \mathbf{R} . This problem can be formalized as follows:

$$\hat{\mathbf{R}} = \arg \min_{\mathbf{R} \subseteq \Theta} \mathcal{V}(\mathbf{R}), \quad (22)$$

subject to:

$$\begin{cases} Pl(\mathbf{R}) & \geq \delta_{Pl}, \\ Bel(\mathbf{R}) & \geq \delta_{Bel}, \end{cases}$$

Because this is a constrained optimization problem, we can use a penalty method to transform it into an unconstrained one. The reformulation of the optimization problem can be written as

$$\hat{\mathbf{R}} = \arg \min_{\mathbf{R} \subseteq \Theta} \mathcal{V}'(\mathbf{R}), \quad (23)$$

where

$$\mathcal{V}'(\mathbf{R}) = \mathcal{V}(\mathbf{R}) + \varrho_{Pl} \cdot \varphi(\delta_{Pl} - Pl(\mathbf{R})) + \varrho_{Bel} \cdot \varphi(\delta_{Bel} - Bel(\mathbf{R})).$$

The quantities ϱ_{Pl} and ϱ_{Bel} are the penalty coefficients. They should be given large values (we chose $\varrho_{Pl} = \varrho_{Bel} = 10^3$ in the simulations reported in Section 4). Function $\varphi(t)$ is a penalty function defined as $\varphi(t) = \max(0, t)$.

In order to solve Eq. (23), we need again to parametrize \mathbf{R} . The point-cloud representation cannot be used here, because it uses too many parameters, which would make the optimization problem intractable. To reduce the number of parameters, we propose to use the Proper Orthogonal Decomposition² (POD) method [70]. This method extracts a smaller number $n \ll N$ of basis vectors; a region \mathbf{R} can then be described by its coordinates in this basis. POD proceeds as follows. We recall that the posterior belief function is represented by M N -vectors \mathbf{v}_j , $j = 1, \dots, M$. First, these vectors are centered,

$$\mathbf{w}_j = \mathbf{v}_j - \bar{\mathbf{v}}, \quad (24)$$

where

$$\bar{\mathbf{v}} = \frac{1}{M} \sum_{j=1}^M \mathbf{v}_j. \quad (25)$$

²In statistics, the same method is known as Principal Component Analysis (PCA).

The $N \times M$ matrix of centered vectors is written as

$$\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_M]. \quad (26)$$

We use the eigendecomposition of matrix $\mathbf{W}^T \mathbf{W}$ to extract basis vectors of POD:

$$\mathbf{W}^T \mathbf{W} = \mathbf{\Phi} \mathbf{\Upsilon} \mathbf{\Phi}^T, \quad (27)$$

where $\mathbf{\Phi}$ is the $M \times M$ matrix whose i -th column ϕ_i is the eigenvector of $\mathbf{W}^T \mathbf{W}$ corresponding to the i -th largest eigenvalue γ_i , and $\mathbf{\Upsilon}$ is the diagonal matrix with diagonal elements $(\gamma_1, \dots, \gamma_M)$. Then

$$\varphi_i = \mathbf{W} \phi_i, \quad (28)$$

is the i -th POD basis vector. The target subsets can be parametrized through their discrete indicator functions decomposed on the first n POD basis vectors,

$$\mathbf{v}^{POD} = \mathbf{1}_\delta \left(\bar{\mathbf{v}} + \sum_{k=1}^n \alpha_k \varphi_k \right), \quad (29)$$

where $\mathbf{1}_\delta$ is the threshold function such that the i -th component of $\mathbf{1}_\delta(\mathbf{v})$ is $I(v(i) \geq \delta)$. In order to choose the truncation number n , the energy proportion (or proportion of explained variance) of a basis vector φ_i is defined as

$$E_i = \gamma_i / \sum_j \gamma_j. \quad (30)$$

We can choose the first n basis vectors whose cumulated energy proportion exceeds some threshold. Finally, Eq. (23) is transformed into minimization with regard to $(\alpha_1, \dots, \alpha_n)$:

$$(\hat{\alpha}_1, \dots, \hat{\alpha}_n) = \arg \min_{\alpha_1, \dots, \alpha_n} \mathcal{V}'(\mathbf{v}^{POD}(\alpha_1, \dots, \alpha_n)). \quad (31)$$

4. Numerical Application

275 We now present two numerical examples to illustrate the general methodology outlined in Section 3. The first example deals with the identification of the elastic properties of a homogeneous 2D plate under loading (Section 4.1). Classical Bayesian inference is used as a comparison. In the second example, presented in Section 4.2, we consider a heterogeneous 2D plate under loading.

280 4.1. Identification of Homogeneous Material

As a first example, we consider the identification of the elastic properties of a 2D homogeneous plate as shown in Figure 4(a). The general mechanical model was described in Section 2.1. Here, we consider the case where the elastic properties are homogeneous over Ω . Thus, the parameters $\boldsymbol{\theta} = (\lambda, \mu)$ are the two Lamé parameters determining the plate's elastic properties, assumed to be constant over the whole domain. As for the boundary condition, the plate is fixed on the left-hand side and loaded on the right-hand side by a traction \mathbf{f} with horizontal projection $\mathbf{f}_x = f_0 y(1 - y)$ and with null vertical projection.

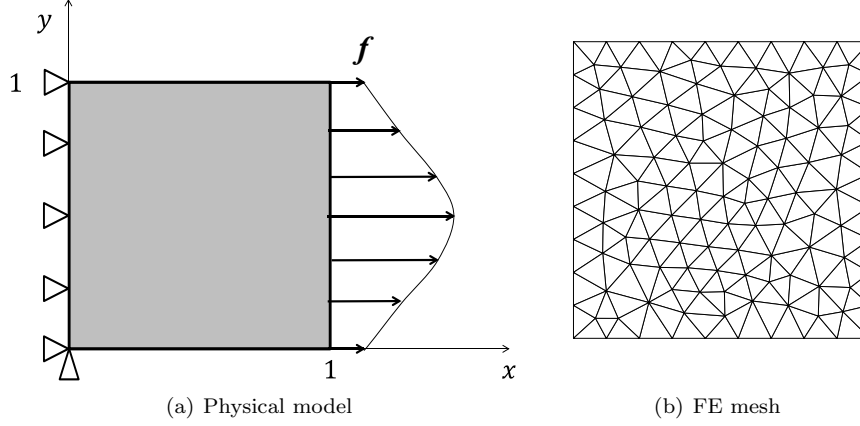


Figure 4: Homogeneous plate.

Measurement and prior information. Since we deal with a numerical example, the first step is to create synthetic measurements, thanks to a reference FE calculation playing the role of the experiment. This reference FE analysis was performed with reference values of the elastic parameters $(\lambda_0, \mu_0) = (1, 1)$ (see Figure 4(b)). The FE reference displacement was transferred to a 10×10 measurement grid covering the whole plate, and a Gaussian white noise sample was added to simulate measurement error. The standard deviation of the noise was 5% of the maximum displacement value.

These synthetic measurements along with the boundary conditions are the only data kept from the reference FE calculation as an input to the inverse identification. The same FE model (parameterized by the unknown parameter vector θ) was then used in the identification process. The measurement information about θ was represented by the likelihood-based belief function described by Eq. (9). Prior information from expert opinions was expressed as trapezoidal contour functions. In order to assess the robustness of this approach to wrong prior information, two scenarios were considered. In the first scenario, the expert opinions were correct, i.e., the reference values had plausibility values equal to 1, as shown by the solid lines in Figure 5. In the second scenario, the expert opinions were completely wrong, i.e., the reference values had zero plausibility, as shown by the broken lines in Figure 5. We assumed 80% confidence in the expert opinions in both scenarios. Using the discounting operation introduced in Section 2.2, the contour functions in both scenarios were transformed into the distributions shown in Figure 6. We note that, after discounting, the reference values belong to the support of both distributions. Using Eq. (12), we constructed a joint contour function for λ and μ , shown as rectangular contour lines in Figures 7(a) and 7(b). The quasi elliptic contour lines in Figure 7(a) and 7(b) represent the normalized likelihood pl_L expressing the measurement information.

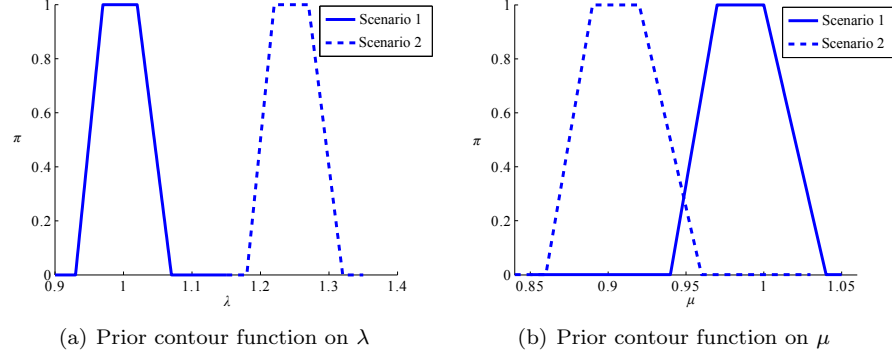


Figure 5: Prior information expressed as trapezoidal contour functions.

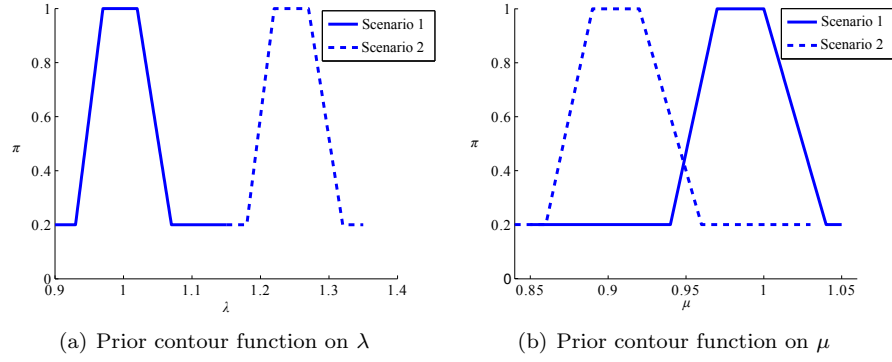


Figure 6: Discounted contour functions, based on a 80% confidence in expert opinions. All parameter values have a plausibility at least equal to 0.2.

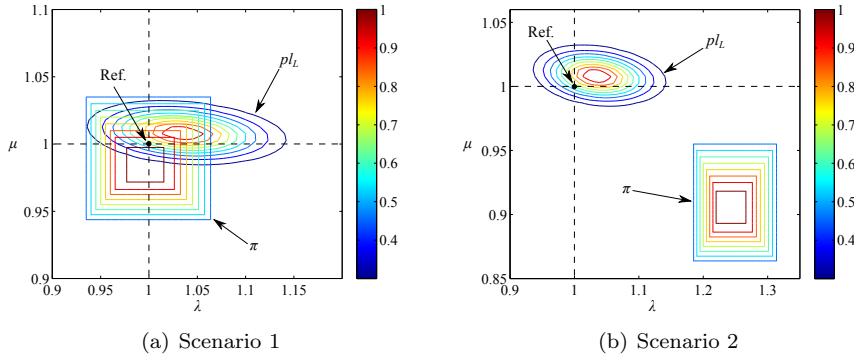


Figure 7: Prior contour function π and normalized likelihood pL_L for Scenarios 1 and 2.

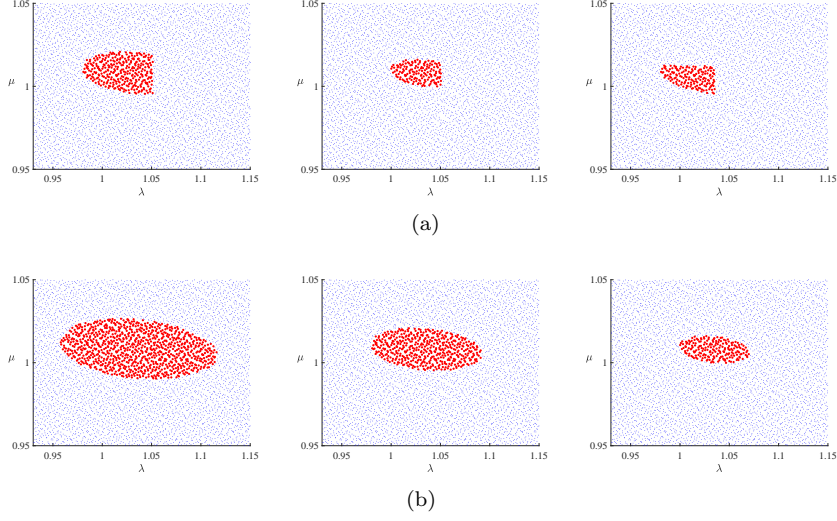


Figure 8: Focal sets of combined belief functions in Scenarios 1 (a) and 2 (b).

Information merging. The likelihood based and prior belief functions were combined by Dempster's rule (13), using the Monte Carlo algorithm described in Section 3.2, with $M = 10^4$. Figure 8(a) and 8(b) show examples of focal sets of the combined belief function in Scenarios 1 and 2, respectively. We can see that the focal sets in Scenario 1 are truncated versions of the ω -level cuts of pl_L , as they have been intersected with level cuts of prior contour function π . In contrast, in Scenario 2, the focal sets result from the intersection of the ω -level cuts of pl_L with the whole frame \mathbb{R}^2 . The estimated degrees of conflict in Scenarios 1 and 2 were, respectively, $\hat{\kappa}_1 = 0.020$ and $\hat{\kappa}_2 = 0.80$.

The combined contour functions in the two scenarios are shown in Figure 9. The maximum plausibility estimate was $\hat{\theta}_P = (1.0275, 1.0063)$ in Scenario 1 and $\hat{\theta}_P = (1.0344, 1.0082)$ in Scenario 2. These estimates are to be compared with reference values $\theta_0 = (1, 1)$ and with the maximum likelihood estimate $\hat{\theta}_L = (1.0344, 1.0082)$. We can see that the correct prior information provided in Scenario 1 has resulted in a slightly smaller estimation error.

Comparison with Bayesian inference. To compare belief function inference with Bayesian inference, we need to express prior information as a probability distribution. To do so, we transformed the prior contour function into Gaussian probability density functions (PDFs) $p(\theta)$. The mean of $p(\theta)$ was taken as the central value of the core of the contour function (defined as the set of parameter values with plausibility one), and the standard deviation was defined in such a way that $P(\text{Supp}) = 1 - m(\Theta)$, where Supp is the support of the contour function (defined as the set of parameter values with strictly positive plausibility).

The marginal distributions of prior PDFs (dot-dash line), likelihood functions (solid line) and posterior PDFs (dashed line) are shown in Figure 10 for the

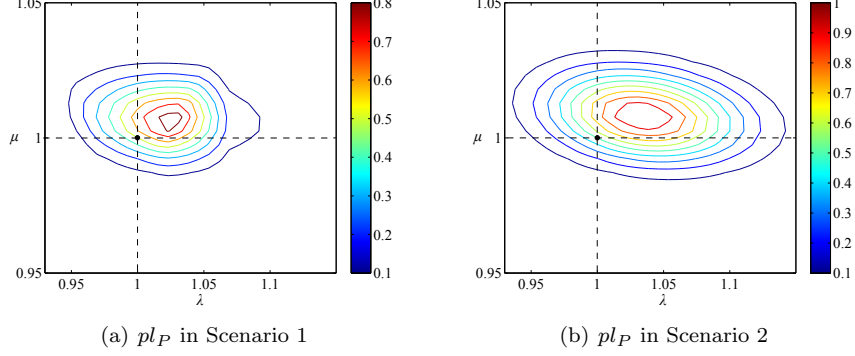


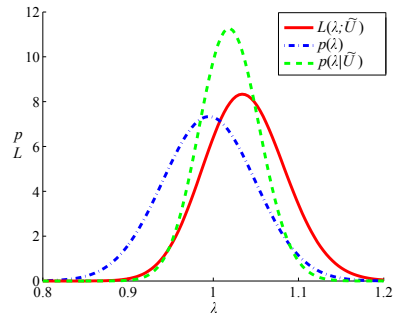
Figure 9: Combined contour functions in Scenarios 1 (a) and 2 (b).

two scenarios. These results can be compared with the marginal contour functions π (dot-dash line), pl_L (solid line) and pl_P (dashed line) shown in Figure 11. We can see that, in Scenario 1, both the posterior PDFs and the combined contour functions “move” towards the reference values. In contrast, comparing Figures 10(c)-10(d) with Figures 11(c)-11(d), we can see that the two methods yield very different results in Scenario 2. As the posterior PDFs of Bayesian inference averages the two pieces of information, the incorrect prior information (dashed-dotted curve in Figures 10(c) and 10(d)) corrupts the measurement information. As a consequence, the posterior PDF $p(\theta|\tilde{u})$ is less accurate than the likelihood function $L(\theta; \tilde{u})$. In contrast, the belief function formalism makes it possible to specify a weak prior in the form of a non-additive belief function, such as the discounted consonant belief function used in this example. In particular, the discounting operation allows the expert to express doubt in his assessment, thus avoiding the undesirable effect of a high conflict between the prior and the data.

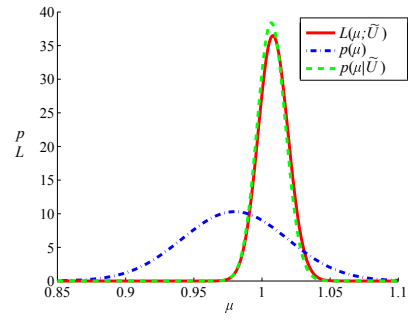
Summarization of the combined belief function. For further exploitation of the combined belief function, we constructed plausible subsets using the POD-based method described in Section 3.3. This method relies on two parameters: the truncation number n and the threshold δ in Eq. (29). Figure 12 shows the energy proportion of the first ten basis vectors. We can see that the first 10 basis vectors account for 82.55% of the total energy. It thus seems sufficient to select the first 10 basis vectors to parametrize the target subsets.

It is interesting to study the sensitivity of the basis vectors to the sample size. Figure 13 shows, as a function of sample size, the error on the first 10 basis vectors defined as

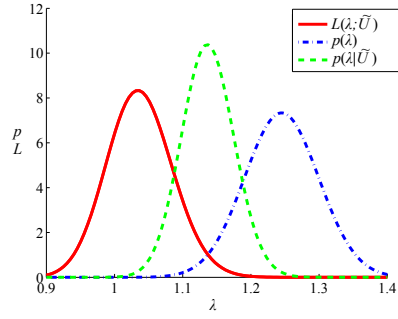
$$\epsilon^{basis} = \frac{1}{\sum_{i=1}^{10} \gamma_i} \sum_{i=1}^{10} \gamma_i \|\varphi_i - \varphi_i^{ref}\|^2, \quad (32)$$



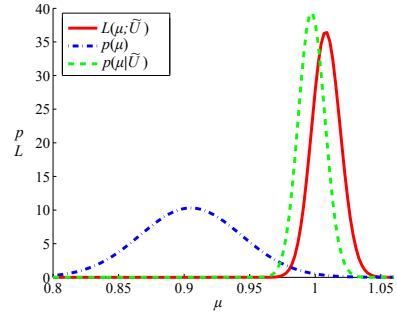
(a) Scenario 1: λ



(b) Scenario 1: μ



(c) Scenario 2: λ



(d) Scenario 2: μ

Figure 10: Marginal probability density distributions in Scenarios 1 (a)-(b) and 2 (c)-(d).

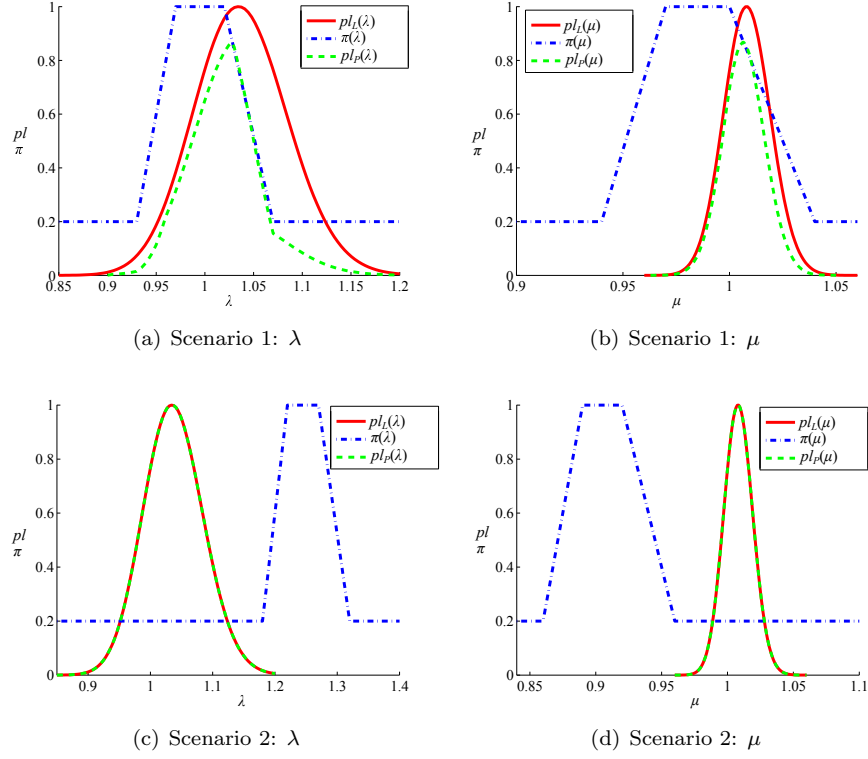


Figure 11: Marginal contour functions pl_L , π and pl_P in Scenarios 1 (a)-(b) and 2 (c)-(d).

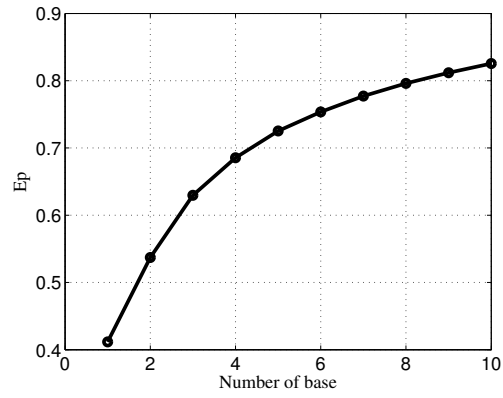


Figure 12: Cumulative energy proportion vs. number of base vectors.

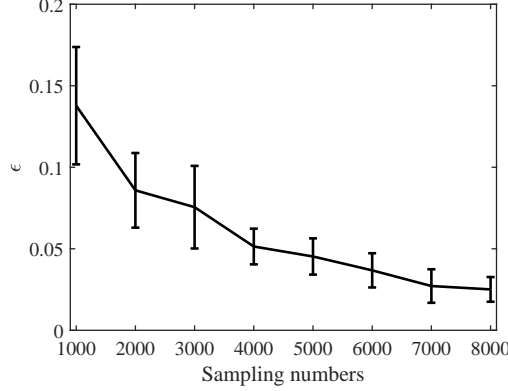


Figure 13: Error of first 10 basis vectors error as a function of sample size.

where φ_i is the i -th basis vector corresponding to eigenvalue γ_i , and φ_i^{ref} is the i -th reference basis vectors calculated with 9796 focal sets of the combined belief function. We can see that the error does not decrease significantly when the sample size is greater than 7000.

In order to determine δ and n , we define the approximation error of the POD as

$$\epsilon^{POD} = \frac{\mathcal{V}(\mathbf{v} - \mathbf{v}^{POD})}{\mathcal{V}(\mathbf{v})}, \quad (33)$$

where $\mathcal{V}(\mathbf{v} - \mathbf{v}^{POD})$ is the volume estimated using the point cloud of the difference between the original \mathbf{v} and the approximated \mathbf{v}^{POD} , and $\mathcal{V}(\mathbf{v})$ is the volume of the domains associated with \mathbf{v} . First, we provisionally fixed $\delta = 0.5$, and we reconstructed 100 subsets using Eq. (29) with different truncation numbers. The approximation errors are shown in Figure 14(a). In Figure 12 and 14(a), we can see that the first six basis vectors represent more than 75% energy and the approximation error does not decrease significantly when $n \geq 6$. Therefore, we chose $n = 6$. Next, we computed the approximation errors with $n = 6$ and different thresholds δ ; the resulting curve is shown in Figure 14(b). We can see that the value $\delta = 0.5$ minimizes the approximation error. Parameter δ was thus fixed to 0.5 in our simulations.

When choosing the truncation number $n = 6$, plausible subsets are described by six parameters α_k , $k = 1, \dots, 6$ in Eq. (31). We randomly sampled 10^4 parameter vectors (α_k), generated the corresponding subsets, and calculated their Bel , Pl and \mathcal{V} values. These quantities are shown in Figure 15 for the two scenarios. The frontier of the point cloud in Figure 15 can be seen as an approximation of the Pareto frontier [71]. The Pareto frontier is helpful to solve multi-objective optimization by revealing the potentially optimal solutions. This frontier reflects the relationship among the extreme values of Pl , Bel and \mathcal{V} . The points in Figure 15(b) are all in the same plane ($Pl = 1$), because the combined belief function in Scenario 2 is consonant. To enhance the precision of

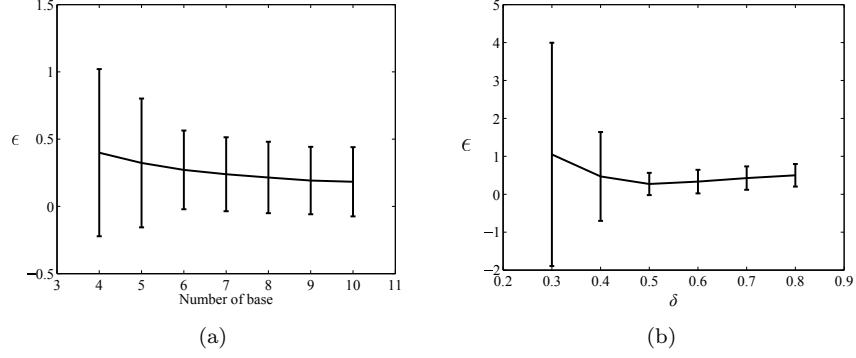


Figure 14: Approximation error vs threshold truncation number (a) and δ (b).

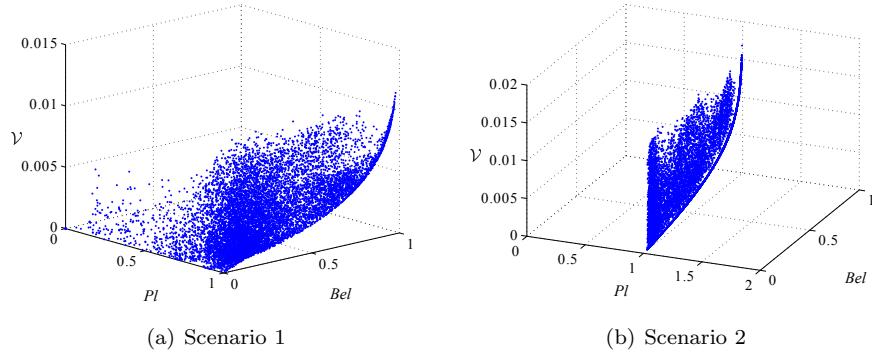


Figure 15: Samples of Bel , Pl and \mathcal{V} .

the Pareto frontier, some authors [72, 73, 74] proposed efficient tools for various applications. Here, we can make use of the Pareto frontier to get suitable initial values for finding the global minimum in Eq. (31). For example, we set (1) $\delta_{Bel} = 0.10$ and $\delta_{Pl} = 0.95$; (2) $\delta_{Bel} = 0.30$ and $\delta_{Pl} = 0.95$ (3) $\delta_{Bel} = 0.50$ and $\delta_{Pl} = 0.95$; the resulting subsets for the two scenarios are shown in Figure 16. It is clear that the size of the plausible subset increases as the threshold increases. The subsets in Scenario 1 move toward the bottom left because of the influence of prior information. In contrast, in Scenario 2, the prior information does not affect the measurement information, and the resulting subsets reflect the form of the likelihood function. For the third group of thresholds, the resulting subset covers the reference values in Scenario 1.

4.2. Identification of Heterogeneous Material

Measurement and prior information. In this second example, we consider the identification of the elastic properties of a heterogeneous plate with an inclusion

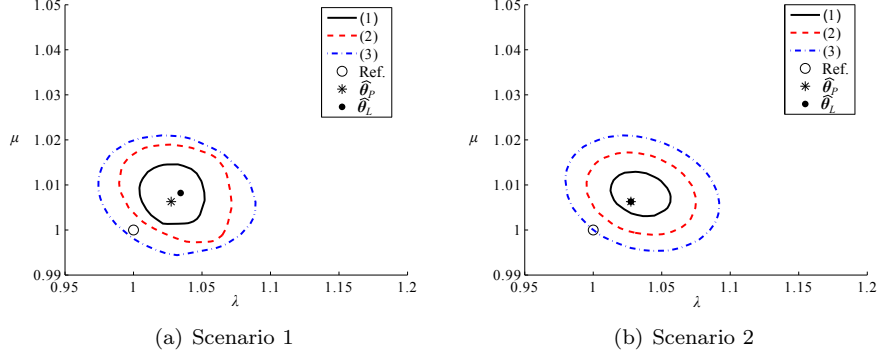


Figure 16: Plausible subsets and point estimates. (1) $\delta_{Bel} = 0.10$ and $\delta_{Pl} = 0.95$; (2) $\delta_{Bel} = 0.30$ and $\delta_{Pl} = 0.95$; (3) $\delta_{Bel} = 0.50$ and $\delta_{Pl} = 0.95$.

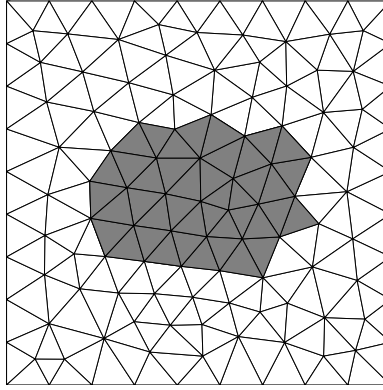


Figure 17: Heterogeneous material.

at its center, as shown in Figure 17. The matrix material (white area) and the inclusion material (grey area) are referred to, respectively, as materials 1 and 2. The traction is the same as in the previous example (Figure 4(a)). The Lamé parameters $\theta = \{\lambda_1, \mu_1, \lambda_2, \mu_2\}$ determine its elastic properties. Parameters λ_1 and μ_1 correspond to the matrix material, while λ_2 and μ_2 correspond to the inclusion material. The relative reference values were $(\lambda_{10}, \mu_{10}, \lambda_{20}, \mu_{20}) = (1, 1, 1.5, 1.5)$. The measurement error was assumed to be a Gaussian white noise with standard deviation equal to 5% of the maximum displacement value.

The prior information was represented by the marginal trapezoidal contour functions π shown in Figure 18. As before, an 80% confidence degree was postulated. Figure 18 also displays the marginal likelihood-based contour functions. Comparing Figures 18(b) and 18(c), we can see that the measurement information is more precise for parameter μ_1 , while the prior information is more precise for parameter λ_2 .

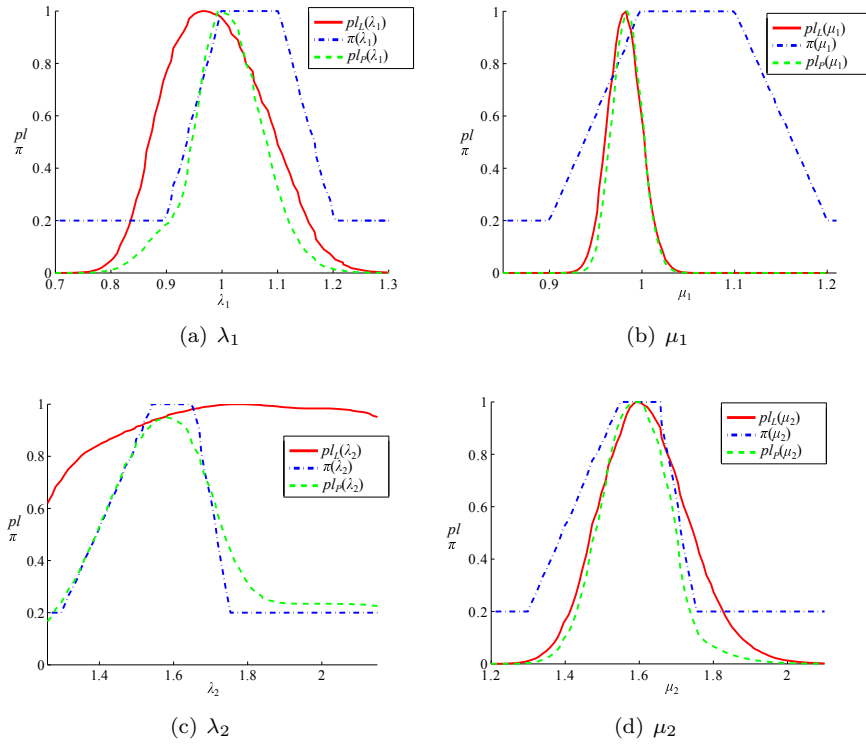


Figure 18: Marginal contour functions pL_L , π and pL_P .

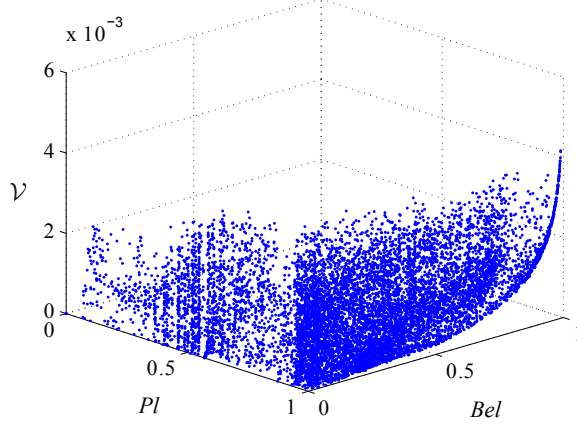


Figure 19: Samples of Bel , Pl and \mathcal{V} .

Information merging. As before, we combined the measurement and prior information by Dempster's rule. The marginal combined contour functions pl_P are shown in Figure 18. Function pl_P is always more specific than pl_L and π . Therefore, when measurement information is more precise, the influence of prior information is smaller. In contrast, when measurement information does not allow precise inference, prior information has a regularization effect on measurement information.

Summarization of the combined belief function. After applying the POD to the samples of the posterior random set, as explained in Section 4.1, we chose the first six parameters to parametrize the plausible subsets and we fixed δ to 0.5. Then, we sampled 10^4 parameter vectors (α_k) , generated corresponding subsets, and calculated their Bel , Pl and \mathcal{V} values. These quantities are shown in Figure 19. The plausible regions corresponding to $\delta_{Bel} = 0.30$ and $\delta_{Pl} = 0.95$ are shown in Figure 20. The location of the maximum of pl_P is closer to the reference values than that of pl_L . The reason for this improvement is the regularization by prior information. The shapes of the plausible subsets reflect the dependence between the parameters, while their sizes reflect the precision of input information. More precise information results in smaller subsets for the same threshold values. It should be noticed that the shapes of the plausible subsets are no longer regular, which indicates that using regular shape functions (such as ellipsoids or boxes) to parametrize plausible subsets would be inefficient in this case.

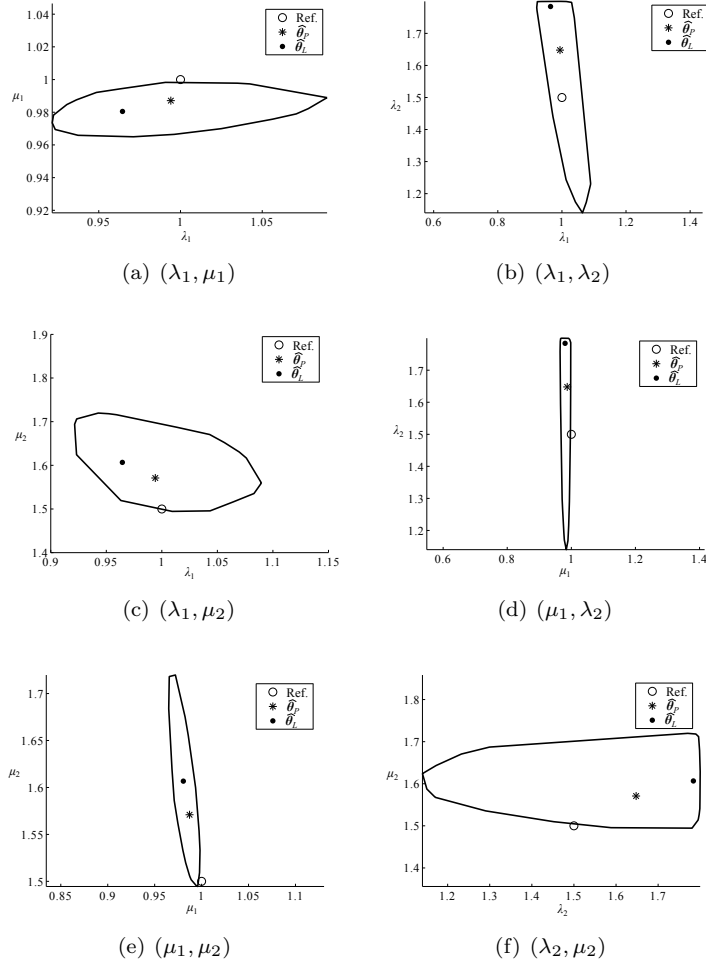


Figure 20: Plausible subsets and point estimates.

5. Conclusion

In this paper, we have presented an identification strategy based on the representation of measurement and prior information as belief functions, and on their combination by Dempster’s rule. In order to implement Dempster’s rule in a continuous multidimensional setting, we have proposed to represent focal sets using a point cloud. To summarize the combined belief function, we have introduced a new method based on the construction of plausible subsets, defined as minimal-size regions whose belief and plausibility exceed some thresholds. This methodology has been applied to two simulated examples concerning homogeneous and heterogeneous materials. Overall, our approach has been shown to be more robust than Bayesian inference to incorrect prior information, thanks to the possibility to encode weak prior information as non-additive belief functions. In particular, the discounting operation allows us to express prior near-ignorance, thus avoiding a high conflict with the data that may lead to wrong identification.

Many of the ideas presented in this paper can be applied not only to identification problems, but also to general problems involving the manipulation and combination of multidimensional belief functions in continuous spaces. However, computational complexity when the dimension becomes very large is an important issue that remains to be investigated. In mechanics, the presented identification method needs to be further explored to deal with more complex material models, such as non-linear ones. Accounting for model uncertainty is another important problem that will be tackled in future work. Finally, the uncertainty arising from unreliable boundary conditions also remains to be investigated.

Acknowledgments

This work was carried out and funded in the framework of the Labex MS2T. It was supported by the French Government, through the program “Investments for the future” managed by the National Agency for Research (Reference ANR-11-IDEX-0004-02) .

References

- [1] W. Peters, W. Ranson, Digital imaging techniques in experimental stress analysis, *Optical engineering* 21 (3) (1982) 427–431.
- [2] H. D. Bui, M. Tanaka, Inverse Problems in Engineering Mechanics: Proceedings of the 2nd international symposium, Paris, 2-4 November 1994, Balkema, 1994.
- [3] J. Hadamard, Sur les problèmes aux dérivées partielles et leur signification physique, *Princeton University Bulletin* 13 (49-52) (1902) 28.

- [4] H. W. Engl, M. Hanke, A. Neubauer, Regularization of inverse problems, Vol. 375, Springer Science & Business Media, 1996.
- [5] A. N. Tikhonov, Numerical methods for the solution of ill-posed problems, Vol. 328, Springer, 1995.
- 480 [6] A. Tarantola, Inverse problem theory and methods for model parameter estimation, Society for Industrial and Applied Mathematics, 2005.
- [7] R. E. Moore, Interval analysis, Vol. 4, Prentice-Hall Englewood Cliffs, 1966.
- [8] D. Dubois, E. Kerre, R. Mesiar, H. Prade, Fuzzy interval analysis, in: Fundamentals of fuzzy sets, Springer, 2000, pp. 483–581.
- 485 [9] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, K. Sentz, Constructing probability boxes and Dempster-Shafer structures, Vol. 835, Sandia National Laboratories, 2002.
- [10] J. Kajberg, G. Lindkvist, Characterisation of materials subjected to large strains by inverse modelling based on in-plane displacement fields, International Journal of Solids and Structures 41 (13) (2004) 3439 – 3459.
- 490 [11] M. B. Azzouna, P. Feissel, P. Villon, Identification of elastic properties from full-field measurements: a numerical study of the effect of filtering on the identification results, Measurement Science and Technology 24 (5) (2013) 055603.
- 495 [12] T. He, L. Liu, A. Makeev, B. Shonkwiler, Characterization of stress-strain behavior of composites using digital image correlation and finite element analysis, Composite Structures 140 (2016) 84 – 93.
- [13] F. Latourte, A. Chrysochoos, S. Pagano, B. Wattrisse, Elastoplastic behavior identification for heterogeneous loadings and materials, Experimental Mechanics 48 (4) (2008) 435–449.
- 500 [14] E. Florentin, G. Lubineau, Identification of the parameters of an elastic material model using the constitutive equation gap method, Computational Mechanics 46 (4) (2010) 521–531.
- [15] M. B. Azzouna, P. Feissel, P. Villon, Robust identification of elastic properties using the modified constitutive relation error, Computer Methods in Applied Mechanics and Engineering 295 (2015) 196–218.
- 505 [16] D. Claire, F. Hild, S. Roux, A finite element formulation to identify damage fields: the equilibrium gap method, International Journal for Numerical Methods in Engineering 61 (2) (2004) 189–208.
- 510 [17] J. Réthoré, S. Roux, F. Hild, An extended and integrated digital image correlation technique applied to the analysis of fractured samples: The equilibrium gap method as a mechanical filter, European Journal of Computational Mechanics/Revue Européenne de Mécanique Numérique 18 (3-4) (2009) 285–306.

- 515 [18] J. N. Périé, H. Leclerc, S. Roux, F. Hild, Digital image correlation and
biaxial test on composite material for anisotropic damage law identification,
International Journal of Solids and Structures 46 (11) (2009) 2388–2396.
- [19] M. Grédiac, Principe des travaux virtuels et identification, Comptes rendus
de l’Académie des sciences. Série 2, Mécanique, Physique, Chimie, Sciences
520 de l’univers, Sciences de la Terre 309 (1) (1989) 1–5.
- [20] H. Chalal, S. Avril, F. Pierron, F. Meraghni, Experimental identification of
a nonlinear model for composites using the grid technique coupled to the
virtual fields method, Composites Part A: Applied Science and Manufac-
turing 37 (2) (2006) 315–325.
- 525 [21] E. Zhang, J. Chazot, J. Antoni, Parametric identification of elastic modu-
lus of polymeric material in laminated glasses, IFAC Proceedings Volumes
45 (16) (2012) 422 – 427.
- [22] C. Gogu, W. Yin, R. Haftka, P. Ifju, J. Molimard, R. Le Riche, A. Vautrin,
Bayesian identification of elastic constants in multi-directional laminate
530 from Moiré interferometry displacement fields, Experimental Mechanics
53 (4) (2013) 635–648.
- [23] G. Shafer, A mathematical theory of evidence, Vol. 1, Princeton university
press Princeton, 1976.
- [24] P. Walley, Statistical reasoning with imprecise probabilities, Chapman and
535 Hall London, 1991.
- [25] P. Smets, R. Kennes, The transferable belief model, Artificial Intelligence
66 (2) (1994) 191–234.
- [26] A. P. Dempster, Upper and lower probabilities induced by a multivalued
mapping, Annals of Mathematical Statistics 38 (1967) 325–339.
- 540 [27] P. Smets, Belief functions on real numbers, International Journal of Ap-
proximate Reasoning 40 (3) (2005) 181–223.
- [28] T. Denœux, Likelihood-based belief function: Justification and some exten-
sions to low-quality data, International Journal of Approximate Reasoning
55 (7) (2014) 1535–1547.
- 545 [29] F. Tonon, Using random set theory to propagate epistemic uncertainty
through a mechanical system, Reliability Engineering & System Safety
85 (1) (2004) 169–181.
- [30] F. Tonon, H.-R. Bae, R. V. Grandhi, C. L. Pettit, Using random set theory
to calculate reliability bounds for a wing structure, Structures and Infras-
550 tructure Engineering 2 (3-4) (2006) 191–200.

- [31] E. Ramasso, R. Gouriveau, Remaining useful life estimation by classification of predictions based on a neuro-fuzzy system and theory of belief functions, *IEEE Transactions on Reliability* 63 (2) (2014) 555–566.
- 555 [32] P. Baraldi, F. Mangili, E. Zio, A belief function theory based approach to combining different representation of uncertainty in prognostics, *Information Sciences* 303 (2015) 134–149.
- [33] L. Sui, Uncertainty management in parameter identification, Ph.D. thesis, Université de Technologie de Compiègne, Compiègne, France (2017).
- 560 [34] L. Sui, P. Feissel, T. Denœux, Identification of elastic properties based on belief function inference, in: J. Vejnarová, V. Kratochvíl (Eds.), *Belief Functions: Theory and Applications*, Springer International Publishing, Cham, 2016, pp. 182–189.
- 565 [35] F. Ackermann, Digital image correlation: performance and potential application in photogrammetry, *The Photogrammetric Record* 11 (64) (1984) 429–439.
- [36] M.-T. Nguyen, J.-M. Allain, H. Gharbi, C. Desceliers, C. Soize, Experimental multiscale measurements for the mechanical identification of a cortical bone by digital image correlation, *Journal of the Mechanical Behavior of Biomedical Materials* 63 (2016) 125–133.
- 570 [37] E. Verhulp, B. van Rietbergen, R. Huiskes, A three-dimensional digital image correlation technique for strain measurements in microstructures, *Journal of biomechanics* 37 (9) (2004) 1313–1320.
- 575 [38] S. Huang, P. Feissel, P. Villon, Modified constitutive relation error: An identification framework dealing with the reliability of information, *Computer Methods in Applied Mechanics and Engineering* 311 (2016) 1 – 17.
- [39] K. Bathe, *Finite Element Method*, John Wiley & Sons, Inc., 2000.
- [40] O. C. Zienkiewicz, R. L. Taylor, *The finite element method for solid and structural mechanics*, Butterworth-heinemann, 2005.
- 580 [41] M. Friswell, J. E. Mottershead, *Finite element model updating in structural dynamics*, Vol. 38 of *Solid Mechanics and Its Applications*, Springer Netherlands, 1995.
- [42] R. Mahnken, E. Stein, A unified approach for parameter identification of inelastic material models in the frame of the finite element method, *Computer Methods in Applied Mechanics and Engineering* 136 (3) (1996) 225–258.
- 585 [43] M. I. Friswell, J. E. Mottershead, H. Ahmadian, Finite-element model updating using experimental test data: parametrization and regularization, *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 359 (1778) (2001) 169–186.

- [44] S. Ferson, L. R. Ginzburg, Different methods are needed to propagate ignorance and variability, *Reliability Engineering & System Safety* 54 (2) (1996) 133–144.
- [45] J. C. Helton, W. Oberkampf, Alternative representations of epistemic uncertainty, *Reliability Engineering & System Safety* 85 (1) (2004) 1–10.
- [46] G. J. Klir, Is there more to uncertainty than some probability theorists might have us believe?, *International Journal of General System* 15 (4) (1989) 347–378.
- [47] W. L. Oberkampf, J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, S. Ferson, Challenge problems: uncertainty in system response given uncertain parameters, *Reliability Engineering & System Safety* 85 (1) (2004) 11–19.
- [48] H. T. Nguyen, On random sets and belief functions, *Journal of Mathematical Analysis and Applications* 65 (3) (1978) 531–542.
- [49] G. Shafer, Allocations of probability, *The Annals of Probability* 7 (5) (1979) 827–839.
- [50] B. Michel, H. François, O. Jean-José, R. Stéphane, Digital image correlation, in: M. Grédiac, F. Hild (Eds.), *Full-field measurements and identification in solid mechanics*, John Wiley & Sons, 2012, pp. 157–190.
- [51] G. Besnard, F. Hild, S. Roux, Finite-element displacement fields analysis from digital images: Application to Portevin-Le Châtelier bands, *Experimental Techniques* 46 (2006) 789–803.
- [52] T. Denœux, Rejoinder on “Likelihood-based belief function: Justification and some extensions to low-quality data”, *International Journal of Approximate Reasoning* 55 (7) (2014) 1614–1617.
- [53] A. Nasekhian, H. F. Schweiger, Random set finite element method application to tunnelling, *International Journal of Reliability and Safety* 5 (3-4) (2011) 299–319.
- [54] G. Bellenfant, D. Guyonnet, D. Dubois, O. Bouc, Uncertainty theories applied to the analysis of CO₂ plume extension during geological storage, *Energy Procedia* 1 (1) (2009) 2447–2454.
- [55] N. Ben Abdallah, N. Mouhous-Voyneau, T. Denœux, Combining statistical and expert evidence using belief functions: Application to centennial sea level estimation taking into account climate change, *International Journal of Approximate Reasoning* 55 (1) (2014) 341–354.
- [56] P. Smets, Belief functions: the disjunctive rule of combination and the generalized bayesian theorem, *International Journal of Approximate Reasoning* 9 (1) (1993) 1–35.

- [57] G. Shafer, Dempster's rule of combination, *International Journal of Approximate Reasoning* 79 (2016) 26–40.
- [58] L. A. Wasserman, Belief functions and statistical evidence, *The Canadian Journal of Statistics* 18 (3) (1990) 183–196.
- 630 [59] M. Y. Wang, X. Wang, D. Guo, A level set method for structural topology optimization, *Computer Methods in Applied Mechanics and Engineering* 192 (1) (2003) 227–246.
- [60] N. J. Mitra, A. Nguyen, L. Guibas, Estimating surface normals in noisy point cloud data, *International Journal of Computational Geometry & Applications* 14 (04n05) (2004) 261–276.
- 635 [61] J. F. Thompson, B. K. Soni, N. P. Weatherill, *Handbook of grid generation*, CRC press, 1998.
- [62] M. D. McKay, R. J. Beckman, W. J. Conover, Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics* 21 (2) (1979) 239–245.
- 640 [63] R. E. Caflisch, Monte Carlo and quasi-Monte Carlo methods, *Acta numerica* 7 (1998) 1–49.
- [64] E. Saliby, F. Pacheco, An empirical evaluation of sampling methods in risk analysis simulation: quasi-monte carlo, descriptive sampling, and latin hypercube sampling, in: *Simulation Conference, 2002. Proceedings of the Winter*, Vol. 2, IEEE, 2002, pp. 1606–1610.
- 645 [65] H. Niederreiter, *Quasi-Monte Carlo Methods*, Wiley Online Library, 2010.
- [66] J. H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numerische Mathematik* 2 (1) (1960) 84–90.
- 650 [67] C. Lemieux, *Monte Carlo and Quasi-Monte Carlo Sampling*, Springer Series in Statistics, Springer New York, 2009.
- [68] H. Niederreiter, Random number generation and quasi-Monte Carlo methods, *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992.
- 655 [69] N. Wilson, A Monte Carlo algorithm for Dempster-Shafer belief, in: B. D'Ambrosio, P. Smets, P. Bonissone (Eds.), *Proc. of the 7th Conference on Uncertainty in AI*, Morgan Kaufmann, 1991, pp. 414–417.
- 660 [70] Y. Liang, H. Lee, S. Lim, W. Lin, K. Lee, C. Wu, Proper orthogonal decomposition and its applications. Part I: Theory, *Journal of Sound and vibration* 252 (3) (2002) 527–544.

- 665 [71] B. Wilson, D. Cappelleri, T. W. Simpson, M. Frecker, Efficient Pareto frontier exploration using surrogate approximations, *Optimization and Engineering* 2 (1) (2001) 31–50.
- [72] K. Deb, A. Anand, D. Joshi, A computationally efficient evolutionary algorithm for real-parameter optimization, *Evolutionary computation* 10 (4) (2002) 371–395.
- 670 [73] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE Transactions on Evolutionary Computation* 3 (4) (1999) 257–271.
- [74] J. Knowles, D. Corne, The Pareto archived evolution strategy: A new baseline algorithm for Pareto multiobjective optimisation, in: *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, Vol. 1, IEEE, 1999, pp. 98–105.
- 675