A new reliability-based data-driven approach for noisy experimental data with physical constraints

Jacobo Ayensa-Jiménez[a], Mohamed H. Doweidar[a], Jose A. Sanz-Herrera[b], Manuel Doblaré[1,*]

[a] *Mechanical Engineering Department, University of Zaragoza, Spain; Aragón Institute of Engineering Research (I3A), University of Zaragoza, Spain; Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain. Campus Río Ebro, Edificio I+D, Mariano Esquillor S. N.*
[b] *School of Engineering (ETSI), University of Seville, Spain*

Abstract

Data Science has burst into simulation-based engineering sciences with an impressive impulse. However, data are never uncertainty-free and a suitable approach is needed to face data measurement errors and their intrinsic randomness in problems with well-established physical constraints. As in previous works, this problem is here faced by hybridizing a standard mathematical modeling approach with a new data-driven solver accounting for the phenomenological part of the problem, with the aim of finding a solution point, satisfying some constraints, that minimizes a distance to a given data-set. However, unlike such works that are established in a deterministic framework, we use the Mahalanobis distance in order to incorporate statistical second order uncertainty of data in computations, i.e. spread and correlations. We develop the underlying stochastic theoretical framework and establish the fundamental mathematical and statistical properties. The performance of the resulting reliability-based data-driven procedure performance is evaluated in a simple but illustrative unidimensional problem as well as in a more realistic solution of a 3D structural problem with a material with intrinsically random constitutive behavior as concrete. The results show, in comparison with other data-driven solvers, better

[*] Corresponding author
Email address: mdoblare@unizar.es (Manuel Doblaré)

convergence, higher accuracy, clearer interpretation, and major flexibility besides the relevance of allowing uncertainty management, with low computational demand.

Keywords: Data-Driven, Reliability, Mahalanobis distance

1. Introduction

Nowadays, Data Science and disciplines such as Big Data or Data Analytics [1] are essential in our everyday life. Photos and videos handling, control of patients data, consumer preferences data, census information and police incident reports are just some examples of the daily huge data treatment.

These methodologies permit the extraction of patterns and/or relevant information from available unstructured data [2]. Since the main ideas and concepts were introduced at the beginning of the century [3, 4], an extensive literature may be found on this broad area [5, 6, 7, 8, 9, 10, 11, 12, 13, 14].

For example, in Machine Learning approaches [3, 15], the idea is improving continuously the accuracy of predictions by means of new available data, adding new explicit knowledge from the actual response to previous predictions. A particular subdiscipline of Machine Learning is Manifold Learning [16, 17] in which the particular aim is getting newer and richer hidden knowledge related to the underlying structure or, in mathematical terms, the dimensionality and local bases of the relevant working space. There are many methods for building the underlying manifold from data, ranging from pure interpolation to pure regression, including all Manifold Learning techniques (kernel Principal Component Analysis (kPCA) [18], Self Organizing Map (SOM) [19], Locally Linear Embedding (LLE) [20], Isomap [21], Laplacian Eigenmap [22], t-distributed Stochastic Neighbor Embedding (t-SNE) [23] among others [8]).

In the same direction, since Rosenblatt developed the *perceptron* [24], artificial neural networks is another field where new concepts as Deep Learning and dynamic networks are in continuous development and are able to identify more abstract features and solve more complex problems [25, 26, 27].

Despite the wide application of Data Science in areas such as marketing and e-commerce [28], social sciences [29], or healthcare [30], there are other fields where very little has been done. An example are the disciplines where physical models and the corresponding mathematical and numerical simulation tools are well established like Computational Physics, Computational Chemistry or Computational Engineering (Simulation-Based Engineering and Sciences -SBES-). A straightforward application of these techniques is Dynamic Data-Driven Applications Systems (DDDAS) [31], in which the idea is providing both predictive and learning capabilities to the control system of data acquired from a sufficient set of sensors. This paradigm was settled down by Kalman [32] in the sixties with his groundbreaking filter and is still nowadays a hot topic of research opening up a huge range of possibilities [33]. Some work has been done for dynamical systems [34] and for parameterized PDEs systems [35].

SBES may incorporate, in addition to data, some *a priori* characteristic physical knowledge of the analyzed system. At this point, it is crucial to distinguish between two kinds of knowledge. On one hand, physical general principles, such as conservation and thermodynamic laws that are universally accepted as able to describe the underlying universe structure. On the other hand, we find phenomenological models, such as macroscopic material constitutive relations. The latter is an intelligent simplification of the real interactions at molecular level extracted from available experimental data.

From above, it is clear that *Data Analytics* techniques would be very useful in SBES to extrapolate the phenomenological model, but now constrained with the mathematical expression of first principles. This approach, of increasing importance, is known as Data-Driven Simulation-Based Engineering and Sciences (DDSBES). In this mixed approach, the absence of physical constraints implies recovering the Data Science and Machine Learning framework, while total *a priori* parameterization of experimental data recovers classical SBES. Actually, all linear and non-linear phenomenological models are formulated in terms of parametric mathematical equations, where the variables of interest are forced to remain within a given pre-established manifold. DDSBES may be considered then as an *a posteriori* manifold constructor that may be context-dependent. In other words: *let the data tell us which physical variables persist without forcing them a priori, except for universal physical laws.*

Recently, F. Chinesta and coworkers defined a strategy for Data-Driven (DD) Computational Mechanics [36], combining Manifold Learning techniques and a (possibly optimized) directional search strategy inspired in the LaTin method [37]. In that work, it is highlighted that manifold construction step is not compulsory, but could give some insight about underlying physical structure and could result in less computationally demanding solutions. M. Ortiz and his group presented a material model-free method based on the minimization of the distance between the searched solution and a set of experimental data, using a proper energy norm, while remaining in the equilibrium manifold, or equivalently, a well-posed penalty approach [38]. Other DD hybrid approaches use Gaussian processes (GP) in a given data-set for dynamical feedback of the parametrical model [39], or fusion prognosis [40, 41], to combine DD and physical models, but these methodologies remain encapsulated in the underlying specific physical phenomenological model.

None of these works take into consideration the inherent inaccuracy of the data. Conversely, empirical data are considered as error-free for both directional search and penalty approaches [36, 38]. Only for the latter, some mathematical convergence results are derived for zero uncertainty approaches, which in practice is never the case.

In this paper, a new family of methods, called reliability-based data-driven solvers (RBDD), based on a metric accounting for uncertainty are defined and some new mathematical results are derived. It is highlighted how DD solver methodology naturally allows incorporating reliability along the statement of the modeling. The penalty approach suggested in [38] is chosen as starting point, but now, a metric taking into account the uncertainty, the Mahalanobis distance, is employed, in order to deal with spread and correlations of the data. With this, data-driven simulations become sensitive to measurements precision and incorporate uncertainty considerations.

Through this paper, we will discuss the main problems of the classical and the new simulation-based techniques when dealing with noisy data, highlighting the limitations of each methodology. An easy but illustrative one-dimensional problem is used to compare results and to show improvements using this methodology. We also present another more realistic example with real concrete test data, emphasizing the implications of nonexistence of an explicit set of well-defined hypotheses and the corresponding material model.

2. Data-driven solvers

Following [38] and [36], DD solvers may be seen as iterative solvers searching for the intersection of a (data based) empirical manifold and a physical manifold. The first one is in many practical applications experimentally based and has, therefore, a discrete nature. The second is usually established in terms of sound laws particular to the problem in hands, but otherwise derived from first principles universally accepted as the basis of Physics. For the sake of simplicity, we may consider the elastic three-dimensional problem. In that case, the physical manifold is the set of states that verify global and local equilibrium (i.e. conservation of linear and angular momenta), that in the static case (negligible inertial effects) is written in differential form as:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = \mathbf{0} \tag{1}$$

with $\boldsymbol{\sigma}$ the stress tensor.

Equation (1) is usually approximated and solved in a discrete form using numerical methods like Finite Elements (FEM). In that case, after a convenient discretization we can state:

$$\mathbf{B}\mathbf{y} = \mathbf{0} \tag{2}$$

where $\mathbf{y}$ is a finite dimensional vector containing the full stress tensor field information related to a given discretization (for FEM, this vector contains the components of the stress tensor for all the integration points) and $\mathbf{B}$ is a matrix encoding the geometry and connectivity of the domain.

The empirical manifold is defined via a set $\mathcal{E} = \{(\mathbf{x}^j; \mathbf{y}^j)\}_{j=1,\cdots,m}$ of data points, resulting from experimental measurements (and therefore not uncertainty free) as it will be illustrated subsequently. The set $\mathcal{E}$ may be seen as a representation of the underlying material behavior in the following asymptotic sense: (i) if $\mathcal{E}$ approximates a mathematical manifold and (ii) uncertainty of each point approximates to zero. Some basic mathematical results related to these considerations may be found in [38].

When solving the problem, there are two main approaches:

1. The first one is based on identifying, at least locally, manifolds from data. Here, regression techniques (based on least squares or other optimization approach) [36] or interpolation techniques [42, 43] are generally used. For high dimensional spaces, regression algorithms are expensive and therefore a previous step including dimensionality reduction, i.e. Manifold Learning, is generally compulsory [36]. Once the underlying manifold is built, locally tangent spaces may be computed and tangent-based iterative solvers such as Newton-Raphson (NR), quasi-Newton or arc-length strategies may be used. We call this the *linearization* approach.
2. The second one is searching for the solution point directly from data, following [38], or equivalently the third approach presented in [36]. In that case, it is necessary to define a distance, i.e. a metric, in order to select the nearest data point to the physical manifold. In [36], the euclidean norm is selected, despite it is dimensionally inconsistent, while in [38] a more physically-meaningful norm (energy norm) is selected. Both norms, however, do not consider uncertainty in data. The problem is then formulated as a constrained minimization problem, and solved iteratively. We call this the *pure DD* approach.

2.1. Problem formulation

We present here the general framework for DDSBES problems. With this aim, we postulate that a model-free engineering problem may be defined in terms of state variables (X, Y) that are related through a latent and unknown relationship $F(X, Y) = 0$. For most computational frameworks, state variables are presented in a discrete manner such as $(X, Y) = (\mathbf{x}_i, \mathbf{y}_i)_{i=1,\cdots,N}$ where $\mathbf{x}_i$, and $\mathbf{y}_i$ are vectors whose dimension n is the size of the state vector and $\mathcal{N} = N \times n$ is the number of scalar state variables of the problem. Returning to the elastic problem, $\mathbf{x}_i$ is the vector containing all strain components (ε_{kl}) at the point i and $\mathbf{y}$ the vector containing all stress components (σ_{kl}) at the point i. It is now necessary to define a distance (a metric) in the state space for $\mathbf{x}_i$ and $\mathbf{y}_i$. That is, for example for $\mathbf{x}_i$, to define a symmetric and positive-definite matrix $\mathbf{M}_\mathbf{x}$ and:

$$||\mathbf{x}_i||^2_{x,i} = \frac{1}{2}\mathbf{x}_i^T \mathbf{M}_{\mathbf{x},i}\mathbf{x}_i \tag{3}$$

Therefore:

$$d^2_{x,i}(\mathbf{x}_i, \mathbf{x}'_i) = ||\mathbf{x}_i - \mathbf{x}'_i||^2_{x,i} = \frac{1}{2}(\mathbf{x}_i - \mathbf{x}'_i)^T \mathbf{M}_{\mathbf{x},i}(\mathbf{x}_i - \mathbf{x}'_i) \tag{4}$$

As we are considering engineering problems, we have physical constraints. For the sake of simplicity, but without any conceptual limitation, we shall consider linear constraints only, so they can be written as:

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{a} \\ \mathbf{C}\mathbf{y} &= \mathbf{c} \end{aligned} \tag{5}$$

At each point i, we have a trial set $\mathcal{E}_i$ that may be thought as the result of experimental tests. We then define a local penalty function for each point i as:

$$F_i(\mathbf{x}_i, \mathbf{y}_i) = \min_{(\mathbf{x}', \mathbf{y}') \in \mathcal{E}_i} \{d_{x,i}(\mathbf{x}_i, \mathbf{x}') + d_{y,i}(\mathbf{y}_i, \mathbf{y}')\} \tag{6}$$

It is obvious that the penalty function vanishes for each point in $\mathcal{E}_i$, $F_i|_{\mathcal{E}_i} = 0$.

Finally, a global penalty function is defined, $F(\mathbf{x}, \mathbf{y}|\mathcal{E}) = \sum_{i=1}^{N} F_i(\mathbf{x}_i, \mathbf{y}_i)$ where $\mathcal{E} = \prod_{i=1}^{N} \mathcal{E}_i$, $\mathbf{x} = (\mathbf{x}_i)_{i=1,\cdots,N}$ and $\mathbf{y} = (\mathbf{y}_i)_{i=1,\cdots,N}$. Here we have, for each $(\mathbf{x}, \mathbf{y}) \in \mathcal{E}$, $F(\mathbf{x}, \mathbf{y}) = 0$, $F|_{\mathcal{E}} = 0$ and we have a global norm $||(\mathbf{x}, \mathbf{y})||^2 = \sum_{i=1}^{N} ||(\mathbf{x}_i, \mathbf{y}_i)||^2$.

Therefore, a DDSBES problem is defined by the constrained optimization problem:

$$\begin{aligned} \min_{(\mathbf{x},\mathbf{y})} \quad & F(\mathbf{x}, \mathbf{y}|\mathcal{E}) \\ \text{subject to} \quad & \\ & \mathbf{A}\mathbf{x} = \mathbf{a} \\ & \mathbf{C}\mathbf{y} = \mathbf{c} \end{aligned} \tag{7}$$

For the elastic problem, problem (7) takes the form:

$$\begin{aligned} \min_{(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_N,\boldsymbol{\sigma}_1,\cdots,\boldsymbol{\sigma}_N)} \quad & F(\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_N, \boldsymbol{\sigma}_1, \cdots, \boldsymbol{\sigma}_N|\mathcal{E}) \\ \text{subject to} \quad & \\ & \mathbf{C}[\boldsymbol{\sigma}_1, \cdots, \boldsymbol{\sigma}_N]^T = \mathbf{c} \end{aligned} \tag{8}$$

where $\mathbf{C}$ is a matrix encoding connectivity and geometry of the problem, depending on the particular discretization.

This formulation is similar to the one proposed in [38], except for the fact that it is formulated in a slightly more general context, including a generalized distance (and therefore a more flexible way of measuring how far is a point from the data-set).

It is important to note that state variables $\mathbf{x}_i$, $\mathbf{y}_i$ may be in a lower dimensional space obtained after dimensionality reduction. For example, for the elastic problem, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ live in a space of 6 dimensions, so $(\boldsymbol{\varepsilon}, \boldsymbol{\sigma})$ has 12 dimensions. These dimensions may be reduced if additional simplifications are imposed *a priori* onto the material behavior. For example, if a homogeneous isotropic linear material is considered, this dimension is actually two, because of the Hooke law states $\boldsymbol{\sigma} = \lambda \mathrm{Tr}(\boldsymbol{\varepsilon}) + 2\mu\boldsymbol{\varepsilon}$ where λ and μ are Lame's constants, related to phenomenological parameters E and ν.

Proposition 1. *The problem defined by (7) has a unique solution if* $\mathrm{rang}(\mathbf{A}) = \mathrm{rang}(\mathbf{C}) = r$ *where* r *is the number of restrictions.*

Proof:. Let $\mathbf{M_x} = \bigoplus_{i=1}^{n} \mathbf{M}_{\mathbf{x},i}$ y $\mathbf{M_y} = \bigoplus_{i=1}^{n} \mathbf{M}_{\mathbf{x},i}$. Let $(\mathbf{x}^*, \mathbf{y}^*)$ a pair of state variables verifying $F_i(\mathbf{x}_i, \mathbf{y}_i) = d^2_{x,i}(\mathbf{x}_i, \mathbf{x}^*_i) + d^2_{y,i}(\mathbf{x}_i, \mathbf{y}^*_i)$ which exists because of $\mathcal{E}_i$ finiteness. With these definitions we can write

$$F(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \mathbf{M_x}(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{y} - \mathbf{y}^*)^T \mathbf{M_y}(\mathbf{y} - \mathbf{y}^*) \tag{9}$$

We define the lagrangian function $\mathcal{L}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = F(\mathbf{x}, \mathbf{y}) - \boldsymbol{\lambda}^T(\mathbf{A}\mathbf{x} - \mathbf{a}) - \boldsymbol{\mu}^T(\mathbf{B}\mathbf{y} - \mathbf{b})$. Then:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \mathbf{x}} &= \mathbf{M_x}(\mathbf{x} - \mathbf{x}^*) - \mathbf{A}^T\boldsymbol{\lambda} \\
\frac{\partial \mathcal{L}}{\partial \mathbf{y}} &= \mathbf{M_y}(\mathbf{y} - \mathbf{y}^*) - \mathbf{B}^T\boldsymbol{\mu} \\
\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}} &= \mathbf{A}\mathbf{x} - \mathbf{a} \\
\frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}} &= \mathbf{B}\mathbf{y} - \mathbf{b}
\end{aligned} \tag{10}$$

Using Lagrange multipliers theorem: $\frac{\partial \mathcal{L}}{\partial \mathbf{x}} = \mathbf{0}$, $\frac{\partial \mathcal{L}}{\partial \mathbf{y}} = \mathbf{0}$, $\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}} = \mathbf{0}$ y $\frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}} = \mathbf{0}$, therefore:

$$\begin{aligned}
\mathbf{M_x}\mathbf{x} - \mathbf{A}^T\boldsymbol{\lambda} &= \mathbf{M_x}\mathbf{x}^* \\
\mathbf{M_y}\mathbf{y} - \mathbf{B}^T\boldsymbol{\mu} &= \mathbf{M_y}\mathbf{y}^* \\
\mathbf{A}\mathbf{x} &= \mathbf{a} \\
\mathbf{B}\mathbf{y} &= \mathbf{b}
\end{aligned} \tag{11}$$

We define $\mathbb{K}$, $\mathbb{X}$ and $\mathbb{F}$ as:

$$\mathbb{K} = \begin{pmatrix} \mathbf{M_x} & -\mathbf{A}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M_y} & -\mathbf{B}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{0} \end{pmatrix} \tag{12}$$

$$\mathbb{X} = \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \\ \mathbf{y} \\ \boldsymbol{\mu} \end{pmatrix} \tag{13}$$

$$\mathbb{F} = \begin{pmatrix} \mathbf{M_x x}^* \\ \mathbf{M_y y}^* \\ \mathbf{a} \\ \mathbf{b} \end{pmatrix} \tag{14}$$

Then (11) writes as $\mathbb{K}\mathbb{X} = \mathbb{F}$ and, would have a single solution if and only if $\det(\mathbb{K}) \neq 0$. Using block decomposition of determinant:

$$\det(\mathbb{K}) = \det(\mathbf{M_x})\det(\mathbf{M_y})\det(\mathbf{A M_x A}^T)\det(\mathbf{B M_y B}^T) \tag{15}$$

As $\det(\mathbf{M_x}) = \prod_{i=1}^{n} \det(\mathbf{M}_{\mathbf{x},i})$ and $\det(\mathbf{M_y}) = \prod_{i=1}^{n} \det(\mathbf{M}_{\mathbf{y},i})$ and $\mathbf{M}_{\mathbf{x},i}$ and $\mathbf{M}_{\mathbf{y},i}$ are positive definite matrices, $\det(\mathbb{K}) \neq 0 \Leftrightarrow \det(\mathbf{A M_x A}^T)\det(\mathbf{B M_y B}^T) \neq 0 \Leftrightarrow \mathbf{A M_x A}^T$ and $\mathbf{B M_y B}^T$ are regular. Finally, if $\mathbf{D}$ is a positive definite matrix and $\text{rang}(\mathbf{B D B}^T) = \text{rang}(\mathbf{B})$ then regularity condition is equivalent to $\text{rang}(\mathbf{A}) = \text{rang}(\mathbf{B}) = r$.

□

The reason of why the solution may be not unique relies on the (possible) existence of many points $(\mathbf{x}^*, \mathbf{y}^*)$ on the set to minimize the penalty function.

When solving the nonlinear problem (7), two steps are required:

- Local search of a minimum of the penalty function F_i for each element i using the nearest neighbor algorithm. This search looks for the most representative datum in the empirical discrete manifold.
- Global resolution of the linear system $\mathbb{K}\mathbb{X} = \mathbb{F}$. This equation states that the searched points should remain on the physical manifold.

An easy algorithm for data-driven problem solving is:

```
1. Initialization $\mathbf{x}^{*(0)}$ $\mathbf{y}^{*(0)}$ and $k = 0$.
2. While $(\mathbf{x}^{*(k-1)}, \mathbf{y}^{*(k-1)}) \neq (\mathbf{x}^{*(k)}, \mathbf{y}^{*(k)})$.
   (a) Compute $\mathbb{F}^k$.
   (b) Solve $\mathbb{K}\mathbb{X}^k = \mathbb{F}^k$.
   (c) Extraction of components $\mathbf{x}^k$ and $\mathbf{y}^k$ of $\mathbb{X}$.
   (d) Compute $(\mathbf{x}^{*(k+1)}, \mathbf{y}^{*(k+1)})$, nearest sample point to $(\mathbf{x}^k, \mathbf{y}^k)$.
   (e) Update: $k := k + 1$.
3. Solution is $(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^k, \mathbf{y}^k)$.
```

2.2. Reliability-based data-driven solver

Let's suppose we have a method for creating data couples, i.e, state pairs, (X^j, Y^j), $j = 1, \cdots, m$. Now, each of the pairs $U^j = (X^j, Y^j)$ is considered to have random nature. Returning to the discrete case, $U = (X, Y) = (\mathbf{X}_i, \mathbf{Y}_i)_{i=1,\cdots,N}$ where $\mathbf{X}_i$, and $\mathbf{Y}_i$ are now random vectors whose dimension n is the size of the state vector and, as before, $\mathcal{N} = N \times n$ is the number of scalar state variables. Now, we may define the stochastic analogous problem to the deterministic one (7):

$$\begin{aligned} \min_{(\mathbf{x},\mathbf{y})} \quad & \mathbb{E}[F(\mathbf{x}, \mathbf{y}|\mathcal{E})] \\ \text{subject to} \quad & \\ & \mathbf{A}\mathbf{x} = \mathbf{a} \\ & \mathbf{B}\mathbf{y} = \mathbf{b} \end{aligned} \tag{16}$$

Note that with this formulation, the solution candidate $\mathbf{u} = (\mathbf{x}, \mathbf{y})^T$ is not random, while $F(\mathbf{x}, \mathbf{y}|\mathcal{E}) = F(\mathbf{u}, \mathcal{E})$ is a random variable due to the random nature of $\mathcal{E}$.

Proposition 2 (Second-order properties of minimal distance). *Let $D^2_{\mathcal{E}} = F(\mathbf{x}, \mathbf{y}|\mathcal{E})$ the random variable representing the squared distance between $\mathbf{u}$ and the set $\mathcal{E}$ and $d^2_{\mathcal{E}}$ the certain equivalent squared distance, obtained by substituting in the penalty function (6), $\mathbf{x}'$ and $\mathbf{y}'$ by $\mathbb{E}[\mathbf{X}']$ and $\mathbb{E}[\mathbf{Y}']$ respectively. Besides, we denote as $\mathbf{U}^*$ the random vector associated to the minimization of $d_{\mathcal{E}}$, i.e, verifying $F(\mathbf{u}|\mathcal{E}) = ||\mathbf{u} - \mathbb{E}[\mathbf{U}^*]||^2$. If $\mathbf{\Sigma}$ is the variance-covariance matrix of $\mathbf{U}^*$ and $\mathbf{\Omega}$ is the fourth order moment tensor of $\mathbf{u} - \mathbf{U}^*$, that is, tensor defined by $\Omega_{ijkl}(\mathbf{u} - \mathbf{U}^*) = \mathbb{E}[(u_i - U^*_i)(u_j - U^*_j)(u_k - U^*_k)(u_l - U^*_l)]$ then:*

$$\mu(D^2_{\mathcal{E}}) = \mathbb{E}[D^2_{\mathcal{E}}] = \frac{1}{2}\mathrm{Tr}(\mathbf{M}\mathbf{\Sigma}) + d^2_{\mathcal{E}} \tag{17}$$

$$\sigma^2(D^2_{\mathcal{E}}) = \mathrm{Var}(D^2_{\mathcal{E}}) = \mathbf{M} : \mathbf{\Omega} : \mathbf{M} - (\mathrm{Tr}(\mathbf{M}\mathbf{\Sigma}) + d^2_{\mathcal{E}})^2 \tag{18}$$

with $\mathbf{M} = \mathbf{M_x} \oplus \mathbf{M_y}$.

Proof:. We define $\mathbf{u} = (\mathbf{x}, \mathbf{y})^T$, $\mathbf{U}^* = (\mathbf{X}^*, \mathbf{Y}^*)^T$ and $\boldsymbol{\mu} = \mathbb{E}[\mathbf{U}^*]$, so we have $\mathbb{E}[F(\mathbf{u}|\mathcal{E})] = \frac{1}{2}\mathbb{E}[(\mathbf{u} - \mathbf{U}^*)^T\mathbf{M}(\mathbf{u} - \mathbf{U}^*)]$. It is possible to define a random quadratic form:

$$Q_{\frac{1}{2}\mathbf{M}}(\mathbf{u} - \mathbf{U}^*) = F(\mathbf{u}|\mathcal{E}) \tag{19}$$

Then, we have $D^2_{\mathcal{E}} = F(\mathbf{u}|\mathcal{E}) = Q_{\frac{1}{2}\mathbf{M}}(\mathbf{u} - \mathbf{U}^*)$, $d^2_{\mathcal{E}} = Q_{\frac{1}{2}\mathbf{M}}(\mathbf{u} - \boldsymbol{\mu}^*)$. In the Appendix we show that for a stochastic quadratic form $Q_A(\mathbf{Z})$ with expected value $\boldsymbol{\mu}$, variance-covariance matrix $\mathbf{\Sigma}$ and fourth order moment tensor $\mathbf{\Upsilon}$, it is possible to write:

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})] = \mathrm{Tr}(\mathbf{M}\mathbf{\Sigma}) + \boldsymbol{\mu}^T\mathbf{M}\boldsymbol{\mu} \tag{20}$$

$$\mathrm{Var}(Q_{\mathbf{A}}(\mathbf{Z})) = \mathbf{A} : \mathbf{\Upsilon} : \mathbf{A} - (\mathrm{Tr}(\mathbf{A}\mathbf{\Sigma}) + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu})^2 \tag{21}$$

Then, using $\mathbf{A} = \frac{1}{2}\mathbf{M}$ and $\mathbf{Z} = \mathbf{u} - \mathbf{U}^*$, the final statement is easily obtained.

□

Proposition 3 (Second-order properties of minimal distance under normality). *Using the same conditions and notations of the later and assuming that* $\mathbf{U}^*$ *is a multivariate normally distributed random vector,* $\mathbf{U}^* \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*, then*

$$\mu(D^2_{\mathcal{E}}) = \mathbb{E}[D^2_{\mathcal{E}}] = \frac{1}{2}\mathrm{Tr}(\mathbf{M}\boldsymbol{\Sigma}) + d^2_{\mathcal{E}} \tag{22}$$

$$\sigma^2(D^2_{\mathcal{E}}) = \mathrm{Var}(D^2_{\mathcal{E}}) = \frac{1}{2}\mathrm{Tr}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{M}\boldsymbol{\Sigma}) + (\mathbf{u} - \boldsymbol{\mu})^T\mathbf{M}\boldsymbol{\Sigma}\mathbf{M}(\mathbf{u} - \boldsymbol{\mu}) \tag{23}$$

Proof:. It is again a consequence of the definition of $D^2_{\mathcal{E}} = Q_{\frac{1}{2}\mathbf{M}}(\mathbf{u} - \mathbf{U}^*) = F(\mathbf{u}|\mathcal{E})$ and the result for quadratic forms shown in the Appendix, $Q_{\mathbf{A}}(\mathbf{Z})$, when $\mathbf{Z}$ is multivariate normally distributed random vector $\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$:

$$\mathrm{Var}(Q_{\mathbf{A}}(\mathbf{Z})) = 2\mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}) + 4\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu} \tag{24}$$

□

Now, the crucial point is to select a suitable norm for this stochastic approach of the problem. A very recommended one is Mahalanobis distance [44]:

$$d(\mathbf{U}, \mathbf{U}') = \sqrt{(\mathbb{E}[\mathbf{U}] - \mathbb{E}[\mathbf{U}'])^T(\boldsymbol{\Sigma}_{\mathbf{U}})^{-1}((\mathbb{E}[\mathbf{U}] - \mathbb{E}[\mathbf{U}']))} \tag{25}$$

This is equivalent to choose as metric matrix $\mathbf{M} = 2(\boldsymbol{\Sigma})^{-1}$. the expected value of the optimal penalty function is thus:

$$\mu(D^2_{\mathcal{E}}) = \mathbb{E}[D^2_{\mathcal{E}}] = 2\mathcal{N} + d^2_{\mathcal{E}} \tag{26}$$

and, the variance:

$$\sigma^2(D^2_{\mathcal{E}}) = \mathrm{Var}(D^2_{\mathcal{E}}) = 2\boldsymbol{\Sigma}^{-1} : \boldsymbol{\Omega} : \boldsymbol{\Sigma}^{-1} - (4\mathcal{N} + d^2_{\mathcal{E}})^2 \tag{27}$$

Under normality conditions, the variance writes

$$\sigma^2(D^2_{\mathcal{E}}) = \mathrm{Var}(D^2_{\mathcal{E}}) = 2(2\mathcal{N} + 2d^2_{\mathcal{E}})$$

Again, under normality conditions, we can state the following:

Proposition 4 (Squared distance distribution under normality conditions). *Let $D^2_{\mathcal{E}} = F(\mathbf{u}|\mathcal{E})$ the (random) squared optimal distance to $\mathcal{E}$ using Mahalanobis distance and $d^2_{\mathcal{E}}$ the optimal distance of the certain equivalent problem. Assume that $\mathbf{U}^*$ is a multivariate normally distributed random vector, $\mathbf{U}^* \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $D^2_{\mathcal{E}}$ follows a non-central chi-squared distribution with $n = 2\mathcal{N}$ degrees of freedom and non-centrality parameter $\lambda = (\mathbf{u} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{u} - \boldsymbol{\mu})$.*

$$D^2_{\mathcal{E}} \sim \chi^2\left(2\mathcal{N}, (\mathbf{u} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{u} - \boldsymbol{\mu})\right)$$

Proof:. Given $\mathbf{u}$, we have $\mathbf{du} = \mathbf{u} - \mathbf{U}^* \sim \mathcal{N}(\mathbf{u} - \boldsymbol{\mu}, \boldsymbol{\Sigma})$, therefore, $\boldsymbol{\Sigma}^{-1/2}\mathbf{du} \sim \mathcal{N}(\boldsymbol{\Sigma}^{-1/2}(\mathbf{u}-\boldsymbol{\mu}), \mathbb{I})$. By using the non-central chi-squared distribution definition we get:

$$D^2_{\mathcal{E}} = \mathbf{du}^T \boldsymbol{\Sigma}^{-1} \mathbf{du} = (\boldsymbol{\Sigma}^{-1/2}\mathbf{du})^T (\boldsymbol{\Sigma}^{-1/2}\mathbf{du}) \sim \chi^2(n, \lambda)$$

where $n = 2\mathcal{N}$ and $\lambda = (\boldsymbol{\Sigma}^{-1/2}(\mathbf{u}-\boldsymbol{\mu}))^T (\boldsymbol{\Sigma}^{-1/2}(\mathbf{u}-\boldsymbol{\mu})) = (\mathbf{u}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{u}-\boldsymbol{\mu})$.

□

Having some knowledge onto the expected value, variance and distributional properties of the optimal distance to the initial data set, gives us tools for some uncertainty considerations. In this sense, low mean values are related to good convergence while low variance implies neighborhood to certain convergence.

It is important to highlight that in practical common applications, the expected value and the variance-covariance matrix are not known and must be estimated. This can be easily done using parametric estimation from a data sample of size K. Thus, the expected value and variance-covariance matrix may be estimated using the standard formulas:

$$\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}^*] \simeq \overline{\mathbf{X}^*} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{X}^{*k}$$

$$\mathbf{\Sigma} = \mathrm{COV}(\mathbf{X}^*) \simeq \mathbf{Q} = \frac{1}{K-1}\sum_{k=1}^{K}(\mathbf{X}^{*k} - \overline{\mathbf{X}^*})(\mathbf{X}^{*k} - \overline{\mathbf{X}^*})^T$$

Distributional properties of $D^2_{\mathcal{E}}$ when substituting population parameters by sample estimators could be derived but are out of the scope of this work.

3. Numerical experiments

3.1. Unidimensional problem

Now we evaluate the performance of different data-driven solvers, including the reliability-based one proposed herein. As it could be predicted, the main problem of the linearization approach appears when dealing with irregular (non-smooth) empirical manifolds. This is typical in Physics when working with models that have discontinuities, like in many mechanical problems such as plasticity, damage, fracture and contact problems. A very basic unidimensional trivial problem exemplifies well their main pathologies.

Let us consider a simple uniaxial loaded rod, as schematized in Figure 1, with $F = 100$ kN, $A = a^2 = 200$ cm^2 and $L = 10$ m. This problem may be easily solved through traditional model-based techniques. The solution is based on the combination of three equations. Equilibrium equation, stays $\sigma A = F$, compatibility equation, stays $\varepsilon = \frac{u}{L}$. For this problem to be mathematically closed, we need a mathematical relation, i.e. a model, relating the internal (state) variable stress, σ, and the measurable variable strain, ε, what is known as constitutive relation of the material $\sigma = f(\epsilon)$. For linear elasticity, $\sigma = E\varepsilon$.

Here the approach is different. Let us consider that the constitutive relation is not known and the material behavior could be linear, smoothly nonlinear or non-smoothly nonlinear. In any case, what we have to describe the material behavior is a considerable amount of experimental pair values (ε, σ), $\mathcal{E} = \{(\varepsilon^j; \sigma^j)\}_{j=1,\cdots,m}$. For testing DD solvers based on linearization, let us compare the computed results when considering a non-smoothly nonlinear behavior and using the well-known iterative tangent Newton-Raphson method, with the analytic results obtained through the exact linear model.

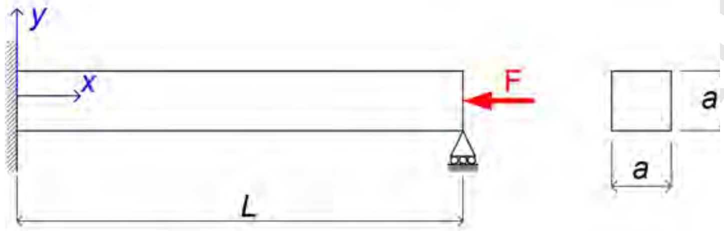


Figure 1: Rod under uniaxial load.

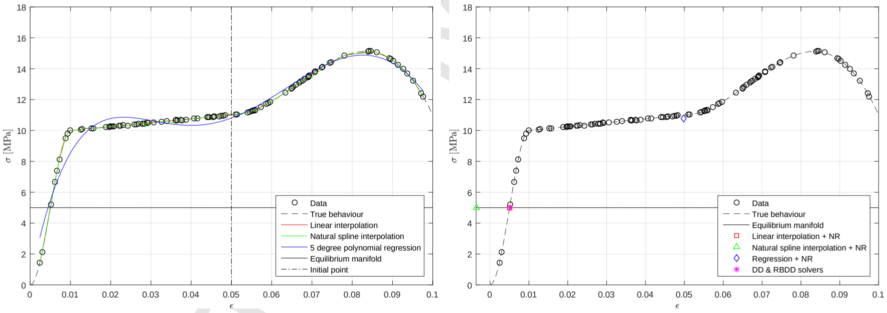
It is important to note in the next figures, corresponding to the numerical experiments, that the grey dashed line, drawn as a function graph $\sigma = f(\epsilon)$, represents the actual material behavior obtained in the experiments, whose expression is not known *a priori*. When using linearization approaches, we need to define, at least locally, a smooth manifold in order to work with tangent spaces. This can be done by using the multiple Manifold Learning techniques presented at the Introduction. As a rule of thumb, the more accurate and structured empirical data set, the better interpolation-based techniques perform. On the other hand, regression techniques are preferred when dealing with noisy and unstructured data but low dimensional and regular underlying manifolds are desirable.

Besides, $\mathcal{E}$ can be generated either allowing control in one of the variables (laboratory controlled tests) or control is impossible (for example sensors in dynamic DD systems). The later case is the most general and challenging. We are going to test the convergence for these two cases using the typical Newton-Raphson solver. We fixed a maximum number of iterations to 10^4 which is huge taking into consideration that, usually, this kind of solvers achieve convergence in a few iterations.

Four analyses are considered, varying the number of data sample points, m, and the error measure related to uncertainty, s. Data generation is as follows: for each $\varepsilon^j \sim \mathcal{U}(0; \varepsilon_{max})$, $j = 1, \cdots, m$ and, as before, $\sigma^j \sim \mathcal{N}(\mu^j, s)$, with

$\mu^j = f(\varepsilon^j)$ and $s = \alpha\sigma_{max}$, where σ_{max} is the maximum stress.

Figures 2a, 3a, 4a, 4c, 5a and 5c show the considered empirical set, the equilibrium manifold and the constitutive manifold built for some fitting techniques (linear interpolation, natural spline interpolation and 5 degree polynomial regression). The vertical dashed line shows initial points considered for the Newton-Raphson solver. Figures 2b, 3b, 4b, 4d, 5b and 5d show the empirical set, the equilibrium manifold and final point for each solver. Both reliability-based data-driven (RBDD) and DD solvers converge to the same point. Convergence is not achieved by the Newton-Raphson solver based on regression fit because of the untrue local convexity of the built manifold, which is inherent to parametric regression. In the case of natural spline regression, almost linear behavior in the hardening part of the curve causes bad convergence. In Figure 2, solvers are based on an empirical set of $m = 100$ pairs $(\varepsilon; \sigma)$ and a low homogeneous uncertainty is considered ($\alpha = 0.001$).

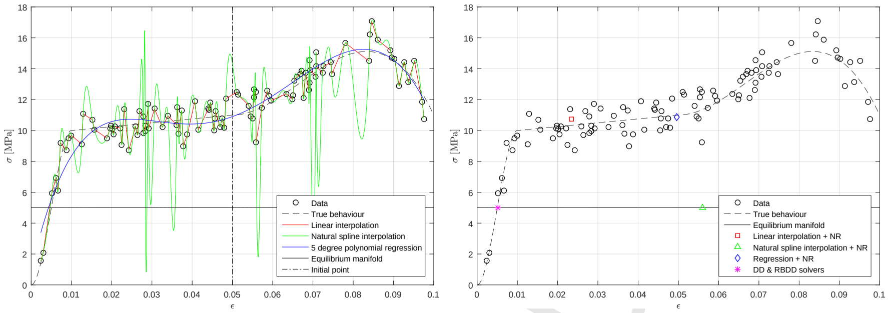


(a) Built manifold for linearization techniques.

(b) Solution point for different solvers.

Figure 2: Performance of different solvers for $m = 100$ and $\alpha = 0.001$.

In Figure 3, we use an empirical set of $m = 100$ pairs $(\varepsilon; \sigma)$ but higher homogeneous uncertainty is considered ($\alpha = 0.05$). Even if polynomial regression is not sensitive to noise, convergence is again not achieved because of the untrue local convexity of the built manifold. Besides, due to noise, natural splines suffer spurious oscillations provoking bad convergence. This can be avoided using linear interpolation, but in this case, non-smoothness of the broken line is incompatible with a tangent-based solver, which in turns results in non-convergence.



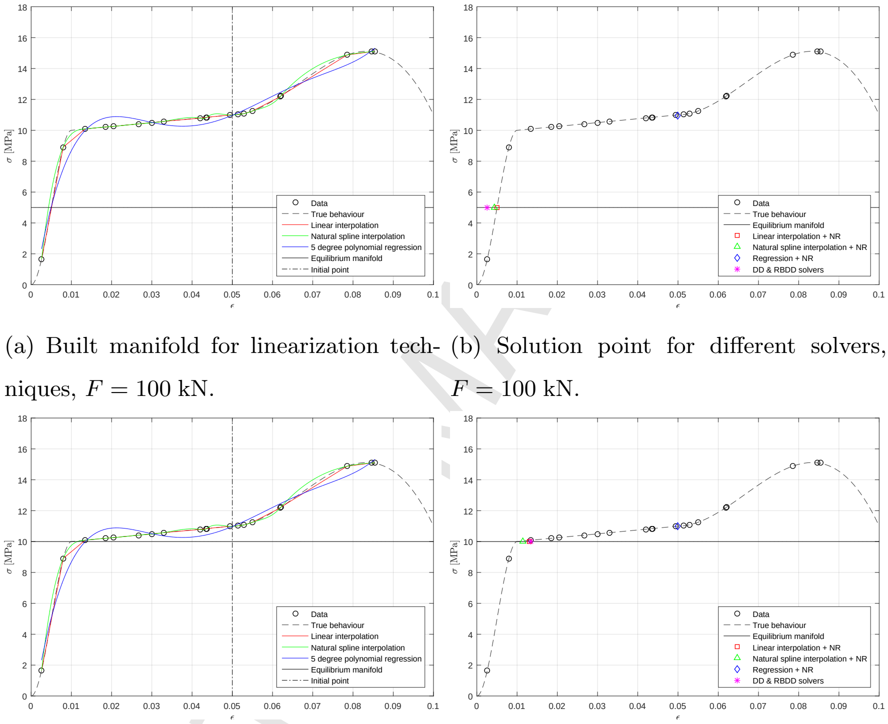
(a) Built manifold for linearization techniques.

(b) Solution point for different solvers.

Figure 3: Performance of different solvers for $m = 100$ and $\alpha = 0.05$.

We next analyze the solver behavior for reduced sample sizes. Using an empirical set of $m = 20$ pairs (ε, σ). First, we consider accurate data ($\alpha = 0.001$). For soundness considerations, we analyze the case with $F = 100$ kN and $F = 200$ kN. Empirical sets are different but are associated to the same m and α. Results are shown in Figure 4. The regression based solver is not convergent for any method. Obviously, due to the lack of data, the DD solver has, in that case, less accuracy than linearization approaches based on interpolation techniques. However, spline interpolation may also have bad convergence, depending on the empirical set mesh and the equilibrium manifold. Linear interpolation would give accurate results only for quasi-linear behavior and/or fine constitutive manifold meshes. Convergence problems increase dramatically when considering greater noise, as seen in Figure 5, where only DD solvers converge to an accurate enough solution.

For homogeneous uncertainty, the DD solver and the RBDD solver give the same result, as pointed out before. Table 1 shows the squared distance results for DD solvers. RBDD is more informative in the following sense: for $m = 100$, the distance to the empirical set increases when passing from $\alpha = 0.001$ to



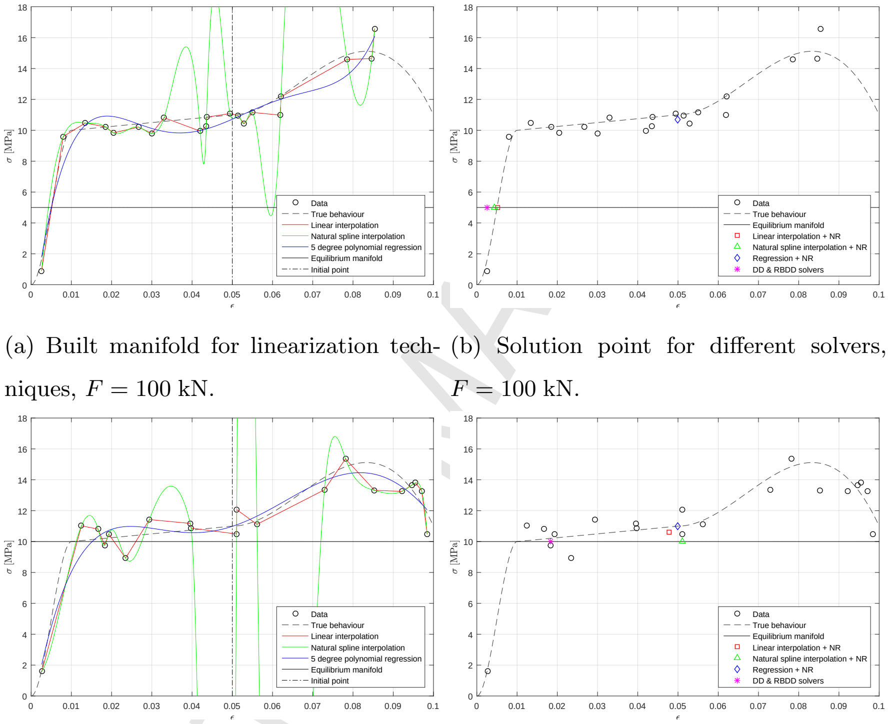
(a) Built manifold for linearization techniques, $F = 100$ kN.

(b) Solution point for different solvers, $F = 100$ kN.

(c) Built manifold for linearization techniques, $F = 200$ kN.

(d) Solution point for different solvers, $F = 200$ kN.

Figure 4: Performance of different solvers for $m = 20$ and $\alpha = 0.001$.



(a) Built manifold for linearization techniques, $F = 100$ kN.

(b) Solution point for different solvers, $F = 100$ kN.

(c) Built manifold for linearization techniques, $F = 200$ kN.

(d) Solution point for different solvers, $F = 200$ kN.

Figure 5: Performance of different solvers for $m = 20$ and $\alpha = 0.05$.

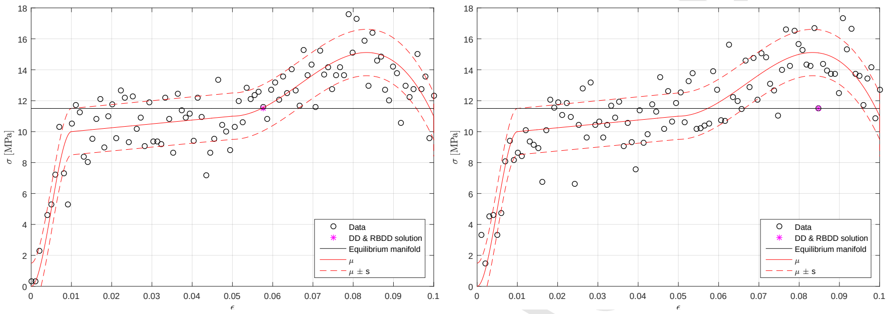
$\alpha = 0.05$. This is due to pure hazard; each realization will give us a different distance depending only on the empirical set sample. RBDD solver does not have this problem because it is uncertainty dependent and can detect when uncertainty is of the order of the optimal distance. Only when $\alpha \to 0$ and the empirical set is almost a subset of the constitutive manifold, this distance can be used as a good uncertainty-free indicator. Otherwise, the locus of the underlying manifold is unknown and there is no way to interpret DD optimal distance in a coherent manner.

m	100		20		20	
F [kN]	100		100		200	
$\boldsymbol{\alpha}$	0.001	0.05	0.001	0.05	0.001	0.05
DD solver	275	6040	75251	112431	6692	1602
RBDD solver	183.43	1.61	50167.62	29.98	29.59	0.096

Table 1: Squared distance results for DD solvers.

To analyze the statistical properties of the squared distance, we consider $\alpha = 0.1$, $m = 100$ and $F = 230$ kN to generate two possible empirical sets with the same statistical properties. Let us suppose that the expected value of the empirical set is known. This should approximate the true underlying manifold but actually it may be estimated from experimental samples. Figure 6 shows the considered data points, expected values μ and an error band, defined by $\mu \pm s$, where s is the standard deviation. Table 2 shows statistical properties of the RBDD solver for both cases, assuming normality. It is important to note that here we consider ε as an uncertainty free variable, and therefore chi-squared distribution of $D^2_{\mathcal{E}} = \left(\frac{\sigma-\mu}{s}\right)^2$ has $n = 1$ degree of freedom. The squared distance computed from data is almost zero in both cases, but a deeper knowledge about empirical set statistics ($\mu = \mu(\varepsilon)$, $s = s(\varepsilon)$) highlights the distance to the true manifold. Anyway, for the case analyzed, 96 simulations have had to be carried out to obtain such result (Case 2).

RBDD solver is not only a more suitable and more informative solver. It



(a) Solution near to the underlying manifold (Case 1).

(b) Solution far from the underlying manifold (Case 2).

Figure 6: Performance of DD and RBDD solvers for $m = 100$ and $\alpha = 0.1$ considering two different samples.

Case	1	2
Squared Optimal distance	0.0056	0.0000
Expected value	1.001	6.725
Variance	2.003	24.899
95%-Confident bound	3.84	16.30

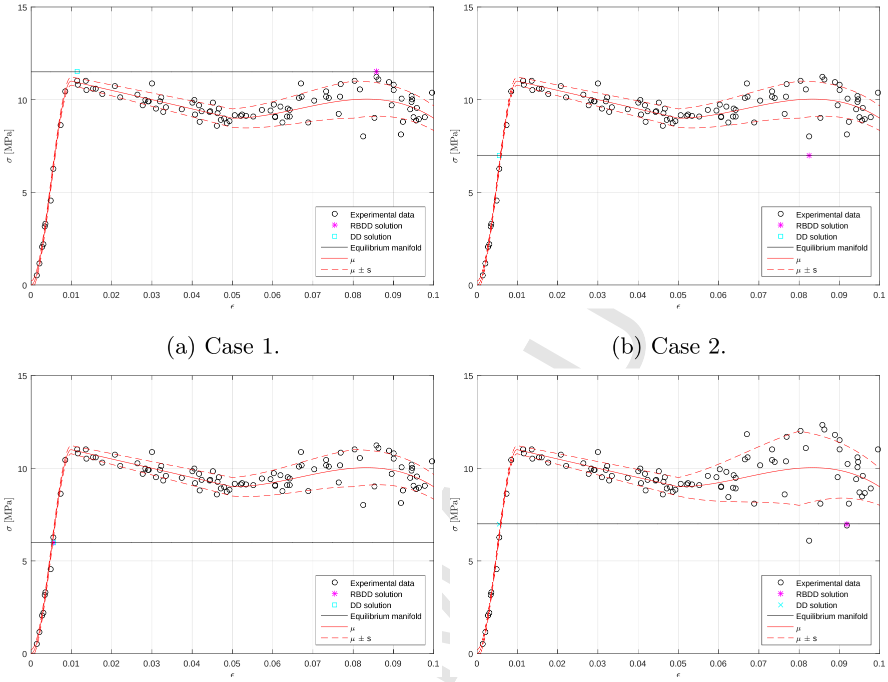
Table 2: Statistical characteristics of the two solution points.

can, for non-homogeneous uncertainty, result in a proper convergence in the following sense. Figure 7 shows solution points for the DD and RBDD solvers for a material with different (ε, σ) constitutive relationship. Note that the uncertainty associated with the actual material behavior is not homogeneous: in the elastic zone, where the material is very well characterized, uncertainty is low, but it increases when strains are higher. RBDD solver is sensitive to this variation, while DD solver is not. For complete information, Figure 7 should be complemented by the statistical properties summarized in Table 3. Thus, in Figure 7a, we can see that the DD solver converges to a very unlikely point while RBDD converges to a more likely one. This is due to the smaller ratio between the geometric distance from the solution point to the empirical data set point and the bandwidth, in case of the RBDD. However, even though the squared distance is small, the expected value indicates that the RBDD solver has converged to a point not very close to the mean manifold. In Figure 7b, a different convergence point is also observed. Now, the RBDD solver has an undesirable behavior because of the lack of data in the linear and certain region. This is detected by means of the expected value and variance, as well as the 95% upper bound. RBDD solvers may be therefore used for sampling strategy considerations. In any case, this is a very unreasonable case, because often, more sample points are associated with less uncertainty. In Figure 7c we can see that both solvers converge to the same values and the statistical properties are similar to those of the first case, indicating a reliable convergence. Finally, the fourth case is similar to the second one, but with higher uncertainty, which in turn reduces the expected value of the squared distance, although the variance remains relatively high.

Note that knowledge of the upper confident bound of the squared distance, $D^2_{\mathcal{E}}$, could be interesting for defining a quantitative criterion for convergence.

3.2. Scale data reduction

An interesting application of the here introduced RBDD solver in the domain of Computational Mechanics appears when dealing with several scales.



(a) Case 1.

(b) Case 2.

(c) Case 3.

(d) Case 4.

Figure 7: Performance of DD and RBDD solvers for heterogeneous uncertainty.

	Case	1	2	3	4
DD	Squared distance	$24.72 \cdot 10^2$	$54.14 \cdot 10^2$	$6.98 \cdot 10^2$	$54.14 \cdot 10^2$
RBDD	Squared distance	0.08	1.12	1.44	0.01
	Expected value	3.70	10.98	4.20	4.94
	Variance	12.82	41.91	14.18	17.76
	95%-Confident bound	10.82	23.07	11.79	13.18

Table 3: Statistical properties of both solvers for each of the presented cases.

One of the main strategies used when coupling two scales (multiscale approach) is selecting a representative volume element (RVE) and establishing a sound transition procedure between the microscale properties and the macroscale response [45, 46]. This strategy has allowed setting up implicitly material constitutive relationships that were not known explicitly at the macroscale [47, 48]. Recent works foreground the crucial point of scales decoupling in the averaging process and the need of uncertainty quantification when building the restriction operator [49]. As answer, many works have incorporated microscale randomness in the multiscale procedure, either using Montecarlo Method (MCM) sampling [50, 51, 52] or Perturbation Method [53]. However, these considerations are still used for model validation and uncertainty has not been incorporated routinely in macroscale computations, except through expensive MCM sampling. A different alternative has been proposed by using stochastic partial differential equations (SPDEs) [54].

We can apply the presented RBDD solver for uncertainty propagation from the microscale to the macroscale allowing to incorporate it in macroscale computations. Let us assume that we have at the microscopic scale a (discrete) coupled field $(\mathbf{x}_i, \mathbf{y}_i),\ i = 1, \cdots K$. Therefore, classical RVE techniques allow us to define a macroscopic reference value $(\mathbf{X}, \mathbf{Y})$ where $\mathbf{X} = \overline{\mathbf{x}_i}$ and $\mathbf{Y} = \overline{\mathbf{y}_i}$. For elastic problems, this could be strain-stress pairs $(\boldsymbol{\varepsilon}_i, \boldsymbol{\sigma}_i)$. It is possible to compute the variance-covariance matrix $\mathbf{\Sigma}$ of the sample $\mathbf{u}_i = (\mathbf{x}_i, \mathbf{y}_i)$. Geometrically, this means to define a $2n$-dimensional ellipsoid in the state space (X, Y) associated to each single macroscopic point, where n is the space dimension of state variable $\mathbf{x}$ or $\mathbf{y}$. Figure 8 shows an ellipsoid in a two-dimensional plane for two possible microstructural fields. Note that accounting only for average values, as done in classical RVE techniques, gives the same result in both of them. We have presented here an approach from the point of view of dimensionality reduction (we use first and second order statistics instead of the whole microscopic field). Besides, this could be exploited in multiscale computational mechanics if a resourceful method allows second order statistical characterization of the microscale without the whole microscale fields computation.

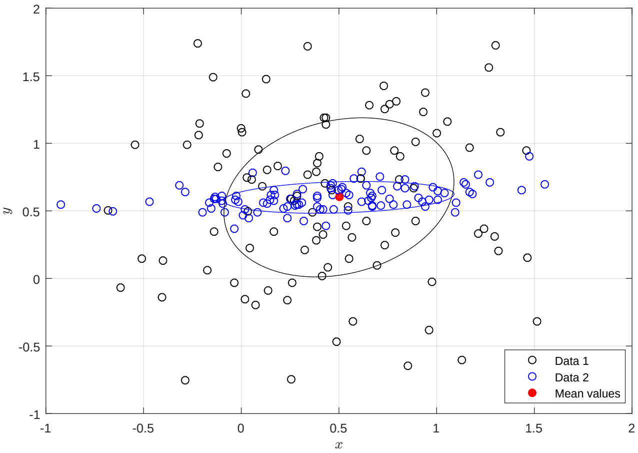


Figure 8: Two possible data sets with associated error ellipsoids.

450 An application of the methodology described above is shown in Figure 9 with the rod problem presented in the previous section. For each macroscopic point (ε, σ), we have the mean strain $\mu(\varepsilon)$, mean stress, $\mu(\sigma)$, strain variance $s^2(\varepsilon)$, stress variance $s^2(\sigma)$ and correlation coefficient $\rho(\varepsilon, \sigma)$ computed from data in the microscale. For the sake of simplicity, these data have been randomly generated using a parametric law but should be interpreted as the result of measurements in the lower scale (pure dimensionality reduction) or obtained through more complex multiscale procedures and techniques. In Figure 9, stresses and strains are normalized using $\varepsilon_0 = 0.1$ and $\sigma_0 = 10$ MPa.

As it can be seen, RBDD allows uncertainty propagation, through a second order moment characterization of state variables. In other words, geometry of the state space is distorted by means of uncertainty: the solution point is agreed to be the nearest point to a given uncertainty ellipsoid, built from input data or specific computations.

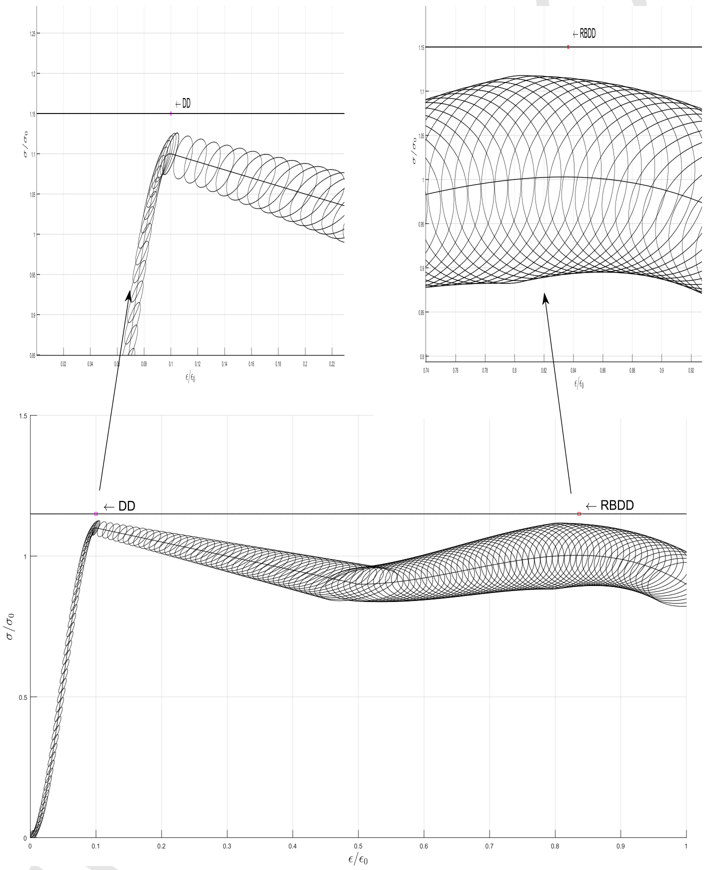


Figure 9: Illustration of a coupling scale strategy using DD and RBDD solvers, with the latter accounting for uncertainty.

3.3. 3D Example

In this section, the potential of the proposed methodology is highlighted in a real example of application. For that, the RBDD solver is implemented in a 3D model using actual data of concrete behavior, thus introducing a more complex level of numerical implementation (now based on finite element methodology). Besides, additional hypotheses are required for the practical use of the RBDD methodology. Both issues are introduced next.

Data regarding mechanical characterization of concrete are obtained from the experimental setup shown in Figure 10. A squared mortar concrete specimen of 100 mm size is subjected to a uniaxial stress state by means of two compression plates, as sketched in Figure 10. Concrete includes Portland cement and a calibrated dosage to get an ultimate strength of 40 MPa expected value. Four experimental tests were carried out at a compression rate of 0.015 mm/s with displacement control. Displacement values and loading were recorded up to rupture of the specimen as seen in Figure 10. These values are treated to build a 3D data-set as commented above.

On the other hand, a concrete specimen subjected to a compressive load - reminiscent to the bottom part of a structural column (see Figure 11) was selected as the 3D problem of interest for the application of the RBDD methodology. Three steps of loading were considered in order to check the performance of the solver at different regions of the mechanical behavior shown in Figure 12.

In this 3D example (but also extended to other scenarios of common practical use), stress-strain data are available along the direction of the load only. To extrapolate this 1D behavior to a multiaxial situation some hypothesis have to be assumed, that are explicitly stated when defining the 3D constitutive model, but are not so explicit (while still necessary) when applying directly the experimental data. The most important are the following:

1. The material is considered homogeneous, or at last, with the same level of homogeneity than the experimental sample used.
2. A given stress state is associated to a certain strain state disregard the

particular material orientation. This implies that the material is isotropic in average at the microstructural level.

3. Only the behavior in one direction is known, so the stress-strain relation in other directions has to be assumed as equal (again isotropy) while the relation between different directions (e.g. Poisson ratio) has to be assumed and estimated. This is a classical hypothesis made during characterization of the mechanical behavior of materials. A value of 0.2 was assumed for concrete in accordance with standard codes of practice.
4. The measured behavior (in principal components) is extrapolated to a multiaxial state using (a) and (b). This extrapolation is made using the same sampling interval than original data.

This rises again the problem of having enough data to extract all possible situations (point location, direction, level of strain, etc.) in order to have the possibility of accurately extrapolating every conditions possible in our particular application. This is rarely the case in reality, so, at least a profound reflection onto the applicability of the data to the particular context and the assumptions it implies is mandatory.

The 3D numerical RBDD solver implemented herein partially follows the work by Kirchdoerfer and Ortiz [38]. Briefly, the algorithm proceeds iteratively based on a finite element methodology to search at each Gauss point of every element the closest solution to the material experimental data-set, i.e.

$$(\sigma_I^{k+1}, \sigma_{II}^{k+1}, \sigma_{III}^{k+1}) - (\varepsilon_I^{k+1}, \varepsilon_{II}^{k+1}, \varepsilon_{III}^{k+1})$$

to

$$(\sigma_I^{D-k+1}, \sigma_{II}^{D-k+1}, \sigma_{III}^{D-k+1}) - (\varepsilon_I^{D-k+1}, \varepsilon_{II}^{D-k+1}, \varepsilon_{III}^{D-k+1})$$

The optimality criterion is based on minimizing the Mahalanobis metric in Equation (25) and therefore follows strictly the methodology explained in previous sections. For the sake of simplicity and computational cost, the searching algorithm proceeds in the space of principal directions. Convergence is considered to be achieved once $\|\mathcal{W}\| < TOL$, being TOL a certain (tolerance) value and $\mathcal{W}$ a certain criterion defined in this section as follows,

$$\mathcal{W} = \sqrt{\frac{1}{s}\|\sigma^{k+1} - \sigma^{k}\|^2 + \frac{1}{e}\|\varepsilon^{k+1} - \varepsilon^{k}\|^2} \tag{28}$$

being s and e representative values of the stress and strain ranges in the test data, respectively. The code was implemented in Matlab software.

Stress component along the compression direction is analyzed in Figure 13 for different regimes (steps) of the strain-strain curve, at two representative points (top and bottom) located at the surface of the specimen (see Figure 12). Figure 13 also shows the stress-strain level of points 1 and 2 along the data-set as well as mean and mean ± standard deviation curves. It is observed that stress keeps in the linear range at steps 1 and 2. Conversely, point 1 at step 3 falls into the so-called damaged region of the concrete behavior. It is convenient to note that DD numerical methodology naturally deals with nonlinear material behavior without the need of elaborated model-dependent formulations and associated nonlinear solvers, and RBDD solver turns out to be uncertainty robust as well. For completeness, Table 4 shows the optimal Mahalanobis distance of the obtained solution at points 1 and 2 for the different analyzed steps.

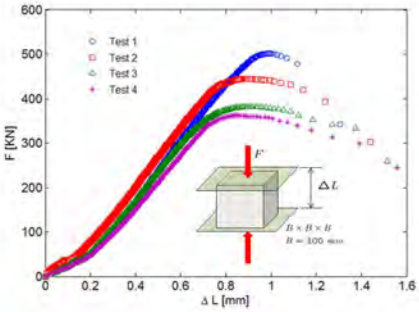


Figure 10: Experimental setup and obtained experimental data from four different tests.

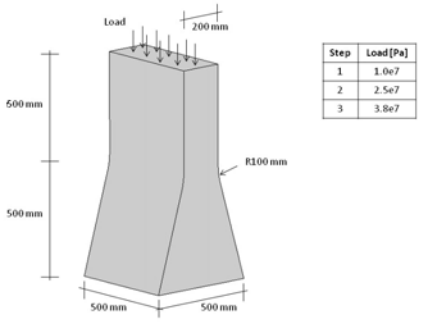
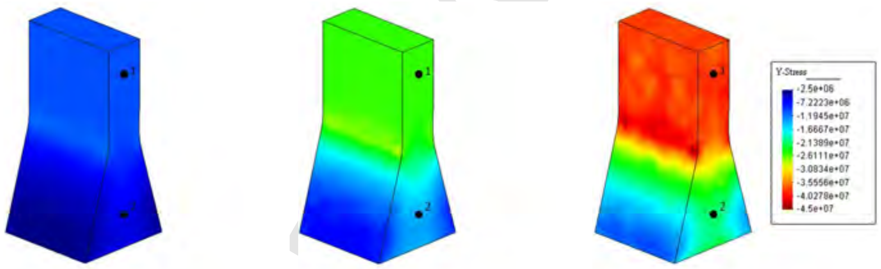


Figure 11: Geometry and dimensions of the concrete test piece used in numerical simulation.



(a) σ_{yy} field obtained at step 1.

(b) σ_{yy} field obtained at step 2.

(c) σ_{yy} field obtained at step 3.

Figure 12: Stress field obtained using RBDD methodology.

Point	1	2
Step 1	$1,44 \cdot 10^{-3}$	$7.93 \cdot 10^{-4}$
Step 2	$2.41 \cdot 10^{-3}$	$3.34 \cdot 10^{-3}$
Step 2	$6.89 \cdot 10^{-3}$	$3.60 \cdot 10^{-3}$

Table 4: Mahalanobis optimal distance for RBDD methodology applied to the 3D example of application

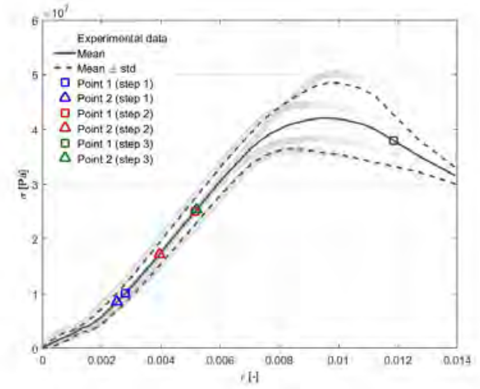


Figure 13: Experimental data, confidence band at level $\mu \pm s$ and numerical solution ad different points and steps.

4. Discussion and conclusions

In this work, a new RBDD solver has been formulated for DDSBES problems, allowing uncertainty considerations in the input data that are, therefore, not considered as uncertainty-free, but of random nature. The DDSBES problem is here defined as a constrained stochastic optimization problem. Constraints encode all relevant physical information of the system, such as fundamental conservation or physical laws. Calculations are carried out directly from data, avoiding any modeling error via state or constitutive equation assumptions. Optimality is sought in terms of a penalty function showing the distance between a candidate solution point and input data set.

It has been shown that employing a proper uncertainty dependent distance, the Mahalanobis distance, results in good statistical properties as well as an easily interpretable optimal distances. Indeed, this optimal distance is computed in terms of the data sample and data uncertainty, which allows assessing when the solution point is accurate enough up to data precision. Moreover, this distance offers the possibility of considering heterogeneous uncertainty, leading to most likely solution points, instead of getting deterministic solution points which may be very sampling-dependent.

Excluding very simple problems with uncertainty-free linear behavior and small sample sizes (Figure 4 and Figure 7), where conventional solvers could be used, the method here proposed has shown better convergence, higher precision, clearer interpretation, major flexibility and more soundness. Besides, the statistical interpretation, depending on sampling statistics, allows decision-oriented statistical inference.

RBDD solvers appear to be a very suitable tool for facing at least three important problems:

1. *Dynamic DD systems*, where predicting features and learning capabilities are combined. The presented solver could start predictions from scratch, where sample sets are small and the underlying true manifold is unknown. Further knowledge of the analyzed system due to the increase of the sample size will feed the RBDD solver, thus allowing faster updates of solution points and statistical inference. This will, additionally, enhance the possibility to define different sampling strategies, data coverage and improve solver performance. Moreover, it is sensitive to measurement errors, that depends on equipment and human precision. In this sense, RBDD solvers conform a robust framework that provides coherent results within the experimental context.
2. *Scale dimensionality reduction problems.* From purely theoretical (sound physical reasons) and/or practical (speed-up calculations) point of views, it could be interesting to define the transition from a small scale to a higher one, defining a hierarchical procedure. The presented RBDD is an ideal tool for uncertainty propagation from one scale to another. Moreover, this may be helpful in case of a dimensionality reduction strategy anchored to Big Data frameworks. For instance, if we work with an n-dimensional field and two scales with two mesh sizes N and M, the whole problem will have $N \cdot M \cdot n$ degrees of freedom. Averaging techniques could reduce the problem to a $M \cdot n$ degrees of freedom problem, but all the variability of the lower scale is lost. With the RBDD approach, lower scales variability

is conserved, at low computational cost, resulting in $M \cdot \frac{n(n+3)}{2} \sim \theta(Mn^2)$ degrees of freedom. If $n \ll M$, savings are evident.

3. *Model-free engineering based on empirical measurements.* The new presented solver offers the possibility of carrying out simulation directly from data, without explicit model assumptions. However, this rises an important limitation not only of this methodology, but of *Data Analytics* in general. This corresponds to the need of contextualize data in order to be sure that they can be extrapolated to a possibly different context corresponding to the particular application. The possibility (or not) of this extrapolation is analyzed explicitly in the standard model-driven approach when making explicit the assumptions that drive to such particular model.

The need of matching the DD methodology with existing (simplified) experimental setups today available to capture the mechanical behavior of materials implies the need of making some explicit assumptions or, at least, to think about the context in which the data have been obtained and the one of the application in hand to decide if they can be extrapolated, and if there is additional data required to fulfill the problems demands. Moreover, this method relies on the hypothesis that we have complete information for each point of the data-set, that is, for each point, all the coordinates are known (for example, in the problem arising from computational mechanics, all the components of both tensors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ are known for each data point). When this is not the case, an appropriate filling data strategy should be considered.

Regardless of this, RBDD solvers present a meeting point between theoretical sciences, through epistemologic constraints, and experimental sciences, through uncertain real world data. The elegance of the mathematical formulation enables many analysis and theoretical considerations for the whole spectrum of Continuum Physics. The ease of combining the presented concepts with all trendy Data Science and Deep Learning tools opens up huge possibilities for facing the most challenging problems because it offers a huge range of possibilities in dynamic DD applications, dimensionality reduction, decision-support systems and any kind of problem in which uncertainty plays a major role.

Acknowledgments

The authors gratefully acknowledge the financial support from the Spanish Ministry of Economy and Competitiveness (MINECO MAT2016-76039-C4-4-R, AEI/FEDER, UE), the Government of Aragon (DGA) and the Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN). CIBER-BBN is financed by the Instituto de Salud Carlos III with assistance from the European Regional Development Fund.

Appendix A. Mathematical proofs

Definition:. Let $\mathbf{X} = (X_1, \cdots, X_n)^T$ *a random vector and* $\mathbf{M}$ *a symmetric positive-definite matrix.*

The stochastic quadratic form (SQF) $Q_{\mathbf{M}}(\mathbf{X})$ *is the random variable defined as*:

$$Q_{\mathbf{M}}(\mathbf{X}) = \mathbf{X}^T \mathbf{M} \mathbf{X} \qquad \text{(Appendix A.1)}$$

Lemma Appendix A.1. *Let* $Q_{\mathbf{A}}(\mathbf{X})$ *and* $Q_{\mathbf{B}}(\mathbf{X})$ *two SQF, and let* $\{\mathbf{e}_k\}_{k=1,\cdots,n}$ *the standard basis in* $\mathbb{R}^n$*, then:*

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})] = A_{ij}\Omega_{ij}(\mathbf{X}) \qquad \text{(Appendix A.2)}$$

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})\mathbf{X}] = A_{ij}\Lambda_{ijk}(\mathbf{X})\mathbf{e}_k \qquad \text{(Appendix A.3)}$$

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})Q_{\mathbf{B}}(\mathbf{X})] = A_{ij}B_{kl}\Upsilon_{ijkl}(\mathbf{X}) \qquad \text{(Appendix A.4)}$$

$$\mathrm{Cov}(Q_{\mathbf{A}}(\mathbf{X}), Q_{\mathbf{B}}(\mathbf{X})) = A_{ij}B_{kl}\left(\Upsilon_{ijkl}(\mathbf{X}) - \Omega_{ij}(\mathbf{X})\Omega_{kl}(\mathbf{X})\right) \qquad \text{(Appendix A.5)}$$

where:

$$\Omega_{ij}(\mathbf{X}) = \mathbb{E}[X_i X_j] \qquad \text{(Appendix A.6)}$$

$$\Lambda_{ijk}(\mathbf{X}) = \mathbb{E}[X_i X_j X_k] \qquad \text{(Appendix A.7)}$$

$$\Upsilon_{ijkl}(\mathbf{X}) = \mathbb{E}[X_i X_j X_k X_l] \qquad \text{(Appendix A.8)}$$

Proof:. Using index notation $Q_{\mathbf{A}}(\mathbf{X}) = A_{ij}X_iX_j$ therefore:

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})] = \mathbb{E}[A_{ij}X_iX_j] = A_{ij}\mathbb{E}[X_iX_j] = A_{ij}\Omega_{ij}(\mathbf{X}) \qquad \text{(Appendix A.9)}$$

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})\mathbf{X}] = \mathbb{E}[A_{ij}X_iX_jX_k\mathbf{e}_k] = A_{ij}\mathbf{e}_k\mathbb{E}[X_iX_jX_k] = A_{ij}\Lambda_{ijk}(\mathbf{X})\mathbf{e}_k \qquad \text{(Appendix A.10)}$$

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})Q_{\mathbf{B}}(\mathbf{X})] = \mathbb{E}[A_{ij}X_iX_jB_{kl}X_kX_l] = A_{ij}B_{kl}\mathbb{E}[X_iX_jX_kX_l] = A_{ij}B_{kl}\Upsilon_{ijkl}(\mathbf{X}) \qquad \text{(Appendix A.11)}$$

Finally,

$$\begin{aligned}
&\mathrm{Cov}(Q_{\mathbf{A}}(\mathbf{X}), Q_{\mathbf{B}}(\mathbf{X})) \\
&= \mathbb{E}\left[(A_{ij}X_iX_j - A_{ij}\mathbb{E}[X_iX_j])(B_{kl}X_kX_l - B_{kl}\mathbb{E}[X_kX_l])\right] \\
&= A_{ij}B_{kl}\mathbb{E}\left[X_iX_jX_kX_l - \mathbb{E}[X_iX_j]X_kX_l - X_iX_j\mathbb{E}[X_kX_l] + \mathbb{E}[X_iX_j]\mathbb{E}[X_kX_l]\right] \\
&= A_{ij}B_{kl}\left(\mathbb{E}[X_iX_jX_kX_l] - \mathbb{E}[X_iX_j]\mathbb{E}[X_kX_l] - \mathbb{E}[X_iX_j]\mathbb{E}[X_kX_l] + \mathbb{E}[X_iX_j]\mathbb{E}[X_kX_l]\right) \\
&= A_{ij}B_{kl}\left(\Upsilon_{ijkl}(\mathbf{X}) - \Omega_{ij}(\mathbf{X})\Omega_{kl}(\mathbf{X})\right) \qquad \text{(Appendix A.12)}
\end{aligned}$$

□

Proposition Appendix A.1 (Expectation of a SQF). *Let $Q_{\mathbf{M}}(\mathbf{X})$ a SQF and let $\boldsymbol{\mu}(\mathbf{X})$ the expected value of $\mathbf{X}$ and $\boldsymbol{\Sigma}(\mathbf{X})$ the variance - covariance matrix of $\mathbf{X}$. Therefore:* 630

$$\mathbb{E}[Q_{\mathbf{M}}(\mathbf{X})] = \mathrm{Tr}(\mathbf{M}\boldsymbol{\Sigma}(\mathbf{X})) + \boldsymbol{\mu}(\mathbf{X})^T\mathbf{M}\boldsymbol{\mu}(\mathbf{X}) \qquad \text{(Appendix A.13)}$$

Proof:. Following Lemma Appendix A.1, $\mathbb{E}[Q_{\mathbf{M}}(\mathbf{X})] = M_{ij}\Omega_{ij}$. Furthermore, $\Omega_{ij}(\mathbf{X}) = \Sigma_{ij}(\mathbf{X}) + \mu_i(\mathbf{X})\mu_j(\mathbf{X})$ then:

$$\begin{aligned}\mathbb{E}[Q_{\mathbf{M}}(\mathbf{X})] &= M_{ij}(\Sigma_{ij}(\mathbf{X}) + \mu_i(\mathbf{X})\mu_j(\mathbf{X}))\\ &= M_{ij}\Sigma_{ij}(\mathbf{X}) + M_{ij}\mu_i(\mathbf{X})\mu_j(\mathbf{X})\\ &= \mathrm{Tr}(\mathbf{M}\boldsymbol{\Sigma}(\mathbf{X})) + \boldsymbol{\mu}(\mathbf{X})^T\mathbf{M}\boldsymbol{\mu}(\mathbf{X}) \qquad \text{(Appendix A.14)}\end{aligned}$$

In the last equality we have used $\mathrm{Tr}(\mathbf{AB}) = \mathbf{A} : \mathbf{B}$.

635 □

Proposition Appendix A.2 (Variance and covarianze of SQF). *Let $Q_{\mathbf{A}}(\mathbf{X})$ and $Q_{\mathbf{B}}(\mathbf{X})$ two SQF and let $\boldsymbol{\mu}(\mathbf{X})$ the expected value of $\mathbf{X}$, $\boldsymbol{\Sigma}(\mathbf{X})$ the variance - covariance matrix of $\mathbf{X}$ and $\boldsymbol{\Upsilon}(\mathbf{X})$ the fourth order moment tensor of $\mathbf{X}$. Then:*

$$\begin{aligned}&\mathrm{Cov}(Q_{\mathbf{A}}(\mathbf{X}), Q_{\mathbf{B}}(\mathbf{X}))\\ &= \mathbf{A} : \boldsymbol{\Upsilon}(\mathbf{X}) : \mathbf{B} - \left(\mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma}(\mathbf{X})) + \boldsymbol{\mu}(\mathbf{X})^T\mathbf{A}\boldsymbol{\mu}(\mathbf{X})\right)\left(\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}(\mathbf{X})) + \boldsymbol{\mu}(\mathbf{X})^T\mathbf{B}\boldsymbol{\mu}(\mathbf{X})\right)\end{aligned}$$
(Appendix A.15)

In particular, if $\mathbf{A} = \mathbf{B}$:

$$\mathrm{Var}(Q_{\mathbf{A}}(\mathbf{X})) = \mathbf{A} : \boldsymbol{\Upsilon}(\mathbf{X}) : \mathbf{A} - (\mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma}(\mathbf{X})) + \boldsymbol{\mu}(\mathbf{X})^T\mathbf{A}\boldsymbol{\mu}(\mathbf{X}))^2$$
(Appendix A.16)

640 *Proof:.* Following Lemma Appendix A.1, $\mathrm{Cov}(Q_{\mathbf{A}}(\mathbf{X}), Q_{\mathbf{A}}(\mathbf{X})) = A_{ij}B_{kl}\Upsilon_{ijkl}(\mathbf{X}) + A_{ij}B_{kl}\Omega_{ij}(\mathbf{X})\Omega_{kl}(\mathbf{X})$. However, $\Omega_{ij}(\mathbf{X}) = \Sigma_{ij}(\mathbf{X}) + \mu_i(\mathbf{X})\mu_j(\mathbf{X})$ therefore:

$$\begin{aligned}&\mathrm{Cov}(Q_{\mathbf{A}}(\mathbf{X}), Q_{\mathbf{A}}(\mathbf{X}))\\ &= A_{ij}B_{kl}\Upsilon_{ijkl}(\mathbf{X}) + A_{ij}B_{kl}(\Sigma_{ij}(\mathbf{X}) + \mu_i(\mathbf{X})\mu_j(\mathbf{X}))(\Sigma_{kl}(\mathbf{X}) + \mu_k(\mathbf{X})\mu_l(\mathbf{X}))\\ &= A_{ij}B_{kl}\Upsilon_{ijkl}(\mathbf{X}) + (A_{ij}\Sigma_{ij}(\mathbf{X}) + A_{ij}\mu_i(\mathbf{X})\mu_j(\mathbf{X}))(B_{kl}\Sigma_{kl}(\mathbf{X}) + B_{kl}\mu_k(\mathbf{X})\mu_l(\mathbf{X}))\\ &= \mathbf{A} : \boldsymbol{\Upsilon}(\mathbf{X}) : \mathbf{B} - (\mathbf{A} : \boldsymbol{\Sigma}(\mathbf{X}) + \boldsymbol{\mu}(\mathbf{X})^T\mathbf{A}\boldsymbol{\mu}(\mathbf{X}))(\mathbf{B} : \boldsymbol{\Sigma}(\mathbf{X}) + \boldsymbol{\mu}(\mathbf{X})^T\mathbf{B}\boldsymbol{\mu}(\mathbf{X}))\end{aligned}$$
(Appendix A.17)

The final result is obtained noting that $\mathrm{Tr}(\mathbf{AB}) = \mathbf{A} : \mathbf{B}$.

□

Let's now assume normality. The following result may be found in [55]:

Remark Appendix A.1 (Fourth order moments of centered multivariate normal distribution). *Let $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ n-dimensional multivariate normally distributed random vector whose expected value is zero and variance - covariance matrix is $\mathbf{\Sigma}$. Then:*

$$\mu_i(\mathbf{Z}) = 0 \tag{Appendix A.18}$$

$$\Omega_{ij}(\mathbf{Z}) = \Sigma_{ij} \tag{Appendix A.19}$$

$$\Lambda_{ijk}(\mathbf{Z}) = 0 \tag{Appendix A.20}$$

$$\Upsilon_{ijkl}(\mathbf{Z}) = \Sigma_{ij}\Sigma_{kl} + \Sigma_{ik}\Sigma_{jl} + \Sigma_{jk}\Sigma_{il} \tag{Appendix A.21}$$

Lemma Appendix A.2. *Let $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ an n-dimensional multivariate normally distributed random vector with expected value $\boldsymbol{\mu} = \mathbf{0}$ and variance - covariance matrix $\mathbf{\Sigma}$. Therefore, for symmetric matrices $\mathbf{A}$ and $\mathbf{B}$:*

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})] = \mathrm{Tr}(\mathbf{A\Sigma}) \tag{Appendix A.22}$$

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})Q_{\mathbf{B}}(\mathbf{Z})] = \mathrm{Tr}(\mathbf{A\Sigma})\mathrm{Tr}(\mathbf{B\Sigma}) + 2\mathrm{Tr}(\mathbf{A\Sigma B\Sigma}) \tag{Appendix A.23}$$

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})\mathbf{Z}] = \mathbf{0} \tag{Appendix A.24}$$

Proof:. The first equation is obtained directly from linearity of the expected value operator and the fact that $\boldsymbol{\mu} = \mathbf{0}$. For the second, note that, following Lemma Appendix A.1 $\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})Q_{\mathbf{B}}(\mathbf{Z})] = A_{ij}B_{kl}\Upsilon_{ijkl}(\mathbf{Z})$, but, by virtue of Observation Appendix A.1, $\Upsilon_{ijkl}(\mathbf{Z}) = \Sigma_{ij}\Sigma_{kl} + \Sigma_{ik}\Sigma_{jl} + \Sigma_{jk}\Sigma_{il}$ and then, using $\mathbf{\Sigma}$ symmetry:

$$\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})Q_{\mathbf{B}}(\mathbf{Z})] = A_{ij}\Sigma_{ij}B_{kl}\Sigma_{kl} + A_{ij}\Sigma_{ki}B_{kl}\Sigma_{lj} + A_{ij}\Sigma_{jk}B_{kl}\Sigma_{li}$$
$$= \mathrm{Tr}(\mathbf{A\Sigma})\mathrm{Tr}(\mathbf{B\Sigma}) + 2\mathrm{Tr}(\mathbf{A\Sigma B\Sigma}) \qquad \text{(Appendix A.25)}$$

Finally, for the third one, following again Lemma Appendix A.1, $\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})\mathbf{Z}] = A_{ij}\Lambda_{ijk}\mathbf{e}_k$, and then, using Observation Appendix A.1 we obtain that $\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})\mathbf{Z}] = \mathbf{0}$.

□

660 We can then prove the following result:

Proposition Appendix A.3 (Variance and covariance of two SQF under normality). *Let $\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ an n-dimensional multivariate normally distributed random vector with expected value $\boldsymbol{\mu} = \mathbf{0}$ and variance - covariance matrix $\mathbf{\Sigma}$.*

Then, if $\mathbf{A}$ and $\mathbf{B}$ are symmetric:

$$\mathrm{Cov}(Q_{\mathbf{A}}(\mathbf{X}), Q_{\mathbf{B}}(\mathbf{X})) = 2\mathrm{Tr}(\mathbf{A\Sigma B\Sigma}) + 4\boldsymbol{\mu}^T\mathbf{A\Sigma B}\boldsymbol{\mu} \qquad \text{(Appendix A.26)}$$

665 *In particular:*

$$\mathrm{Var}(Q_{\mathbf{A}}(\mathbf{X})) = 2\mathrm{Tr}(\mathbf{A\Sigma A\Sigma}) + 4\boldsymbol{\mu}^T\mathbf{A\Sigma A}\boldsymbol{\mu} \qquad \text{(Appendix A.27)}$$

Proof:. We use the expression

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \qquad \text{(Appendix A.28)}$$

With $X = Q_{\mathbf{A}}(\mathbf{X})$ and $Y = Q_{\mathbf{B}}(\mathbf{X})$. The first term of the right hand side may be developed in terms of $\mathbf{Z} = \mathbf{X} - \boldsymbol{\mu}$ using symmetry of matrices $\mathbf{A}$ and $\mathbf{B}$ as:

$$
\begin{aligned}
&\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})Q_{\mathbf{B}}(\mathbf{X})] \\
&= \mathbb{E}[(\mathbf{Z}^T\mathbf{A}\mathbf{Z} + \mathbf{Z}^T\mathbf{A}\boldsymbol{\mu} + \boldsymbol{\mu}^T\mathbf{A}\mathbf{Z} + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu})(\mathbf{Z}^T\mathbf{B}\mathbf{Z} + \mathbf{Z}^T\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\mu}^T\mathbf{B}\mathbf{Z} + \boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu})] \\
&= \mathbb{E}[\mathbf{Z}^T\mathbf{A}\mathbf{Z}\mathbf{Z}^T\mathbf{B}\mathbf{Z}] + 2\boldsymbol{\mu}^T\mathbb{E}[\mathbf{Z}^T\mathbf{A}\mathbf{Z}\mathbf{Z}] + 2\boldsymbol{\mu}^T\mathbb{E}[\mathbf{Z}^T\mathbf{B}\mathbf{Z}\mathbf{Z}] + 4\boldsymbol{\mu}^T\mathbf{A}\mathbb{E}[\mathbf{Z}\mathbf{Z}^T]\mathbf{B}\boldsymbol{\mu} \\
&+ \mathbb{E}[\mathbf{Z}^T\mathbf{A}\mathbf{Z}]\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu} + +\mathbb{E}[\mathbf{Z}^T\mathbf{B}\mathbf{Z}]\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu} + 2\boldsymbol{\mu}^T\mathbf{A}\mathbb{E}[\mathbf{Z}] + 2\boldsymbol{\mu}^T\mathbf{B}\mathbb{E}[\mathbf{Z}] + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu} \\
&= \mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})Q_{\mathbf{B}}(\mathbf{Z})] + 2\boldsymbol{\mu}^T\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})\mathbf{Z}] + 2\boldsymbol{\mu}^T\mathbb{E}[Q_{\mathbf{B}}(\mathbf{Z})\mathbf{Z}] + 4\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\mu} \\
&+ \mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z})]\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu} + \mathbb{E}[Q_{\mathbf{B}}(\mathbf{Z})]\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu} \\
&= \mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma})\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}) + 2\mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\Sigma}) + \mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma})\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu} + \mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma})\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu} \\
&+ 4\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu}
\end{aligned}
\tag{Appendix A.29}
$$

In last equality we have used Lemma Appendix A.2.

The second term of the right hand side, is obtained analogously:

$$
\begin{aligned}
&\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})]\mathbb{E}[Q_{\mathbf{B}}(\mathbf{X})] \\
&= \mathbb{E}[\mathbf{Z}^T\mathbf{A}\mathbf{Z} + \mathbf{Z}^T\mathbf{A}\boldsymbol{\mu} + \boldsymbol{\mu}^T\mathbf{A}\mathbf{Z} + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}]\mathbb{E}[\mathbf{Z}^T\mathbf{B}\mathbf{Z} + \mathbf{Z}^T\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\mu}^T\mathbf{B}\mathbf{Z} + \boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu}] \\
&= (\mathbb{E}[\mathbf{Z}^T\mathbf{A}\mathbf{Z}] + 2\boldsymbol{\mu}^T\mathbf{A}\mathbb{E}[\mathbf{Z}] + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu})(\mathbb{E}[\mathbf{Z}^T\mathbf{B}\mathbf{Z}] + 2\boldsymbol{\mu}^T\mathbf{B}\mathbb{E}[\mathbf{Z}] + \boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu}) \\
&= (\mathbb{E}[Q_{\mathbf{A}}(\mathbf{Z}] + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu})(\mathbb{E}[Q_{\mathbf{B}}(\mathbf{Z}] + \boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu}) \\
&= (\mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu})(\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu})
\end{aligned}
\tag{Appendix A.30}
$$

Again, in the last equality, Lemma Appendix A.2 was used.

Subtracting Equation (Appendix A.30) to Equation (Appendix A.29), we obtain:

$$
\mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})Q_{\mathbf{B}}(\mathbf{X})] - \mathbb{E}[Q_{\mathbf{A}}(\mathbf{X})]\mathbb{E}[Q_{\mathbf{B}}(\mathbf{X})] = 2\mathrm{Tr}(\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\Sigma}) + 4\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\mu}
\tag{Appendix A.31}
$$

□

References

[1] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A. H. Byers, Big data: The next frontier for innovation, competition, and productivity.

[2] D. T. Larose, Discovering knowledge in data: an introduction to data mining, John Wiley & Sons, 2014.

[3] T. M. Mitchell, Machine learning. 1997, Burr Ridge, IL: McGraw Hill 45 (37) (1997) 870–877.

[4] V. Vapnik, The nature of statistical learning theory, Springer science & business media, 2013.

[5] H. Park, D. Cho, The use of the karhunen-loeve decomposition for the modeling of distributed parameter systems, Chemical Engineering Science 51 (1) (1996) 81–98.

[6] M. D. Graham, I. G. Kevrekidis, Alternative approaches to the karhunen-loeve decomposition for model reduction and data analysis, Computers & chemical engineering 20 (5) (1996) 495–506.

[7] I. Jolliffe, Principal component analysis, Wiley Online Library, 2002.

[8] P. E. LJP, H. H. Van Den, Dimensionality reduction: A comparative review, Tech. Rrep.

[9] G. Rozza, D. B. P. Huynh, A. T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations, Archives of Computational Methods in Engineering 15 (3) (2008) 229–275.

[10] H. Abdi, L. J. Williams, Principal component analysis, Wiley interdisciplinary reviews: computational statistics 2 (4) (2010) 433–459.

[11] F. Chinesta, P. Ladeveze, E. Cueto, A short review on model order reduction based on proper generalized decomposition, Archives of Computational Methods in Engineering 18 (4) (2011) 395.

[12] C. Ghnatios, F. Masson, A. Huerta, A. Leygue, E. Cueto, F. Chinesta, Proper generalized decomposition based dynamic data-driven control of thermal processes, Computer Methods in Applied Mechanics and Engineering 213 (2012) 29–41.

[13] A. Manzoni, A. Quarteroni, G. Rozza, Computational reduction for parametrized pdes: strategies and applications, Milan Journal of Mathematics 80 (2) (2012) 283–309.

[14] P. Benner, S. Gugercin, K. Willcox, A survey of projection-based model reduction methods for parametric dynamical systems, SIAM review 57 (4) (2015) 483–531.

[15] S. Khan, Introduction to machine learning (adaptive computation and machine learning series), Natural Language Engineering 14 (01) (2008) 133–137.

[16] J. A. Lee, M. Verleysen, Nonlinear dimensionality reduction, Springer Science & Business Media, 2007.

[17] M. Yunquan, F. Yun, Manifold learning theory and applications (2011).

[18] B. Schölkopf, A. Smola, K.-R. Müller, Nonlinear component analysis as a kernel eigenvalue problem, Neural computation 10 (5) (1998) 1299–1319.

[19] T. Kohonen, The self-organizing map, Proceedings of the IEEE 78 (9) (1990) 1464–1480.

[20] S. T. Roweis, L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (5500) (2000) 2323–2326.

[21] J. B. Tenenbaum, V. De Silva, J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, science 290 (5500) (2000) 2319–2323.

[22] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, Neural computation 15 (6) (2003) 1373–1396.

[23] L. v. d. Maaten, G. Hinton, Visualizing data using t-sne, Journal of Machine Learning Research 9 (Nov) (2008) 2579–2605.

[24] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain., Psychological review 65 (6) (1958) 386.

[25] A. Krenker, A. Kos, J. Bešter, Introduction to the artificial neural networks, INTECH Open Access Publisher, 2011.

[26] K. Suzuki, Artificial neural networks: methodological advances and biomedical applications, InTech, 2011.

[27] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al., Mastering the game of go with deep neural networks and tree search, Nature 529 (7587) (2016) 484–489.

[28] S. Hill, F. Provost, C. Volinsky, Network-based marketing: Identifying likely adopters via consumer networks, Statistical Science (2006) 256–276.

[29] C. S. Aneshensel, Theory-based data analysis for the social sciences, Sage, 2013.

[30] W. Raghupathi, V. Raghupathi, Big data analytics in healthcare: promise and potential, Health Information Science and Systems 2 (1) (2014) 1.

[31] F. Darema, Dynamic data driven applications systems: A new paradigm for application simulations and measurements, in: International Conference on Computational Science, Springer, 2004, pp. 662–669.

[32] R. E. Kalman, A new approach to linear filtering and prediction problems, Journal of basic Engineering 82 (1) (1960) 35–45.

[33] B. Peherstorfer, K. Willcox, Dynamic data-driven reduced-order models, Computer Methods in Applied Mechanics and Engineering 291 (2015) 21–41.

[34] M. Schmidt, H. Lipson, Distilling free-form natural laws from experimental data, science 324 (5923) (2009) 81–85.

[35] B. Peherstorfer, K. Willcox, Data-driven operator inference for nonintrusive projection-based model reduction, Computer Methods in Applied Mechanics and Engineering 306 (2016) 196–215.

[36] R. Ibanez, E. Abisset-Chavanne, J. V. Aguado, D. Gonzalez, E. Cueto, F. Chinesta, A manifold learning approach to data-driven computational elasticity and inelasticity, Archives of Computational Methods in Engineering (2016) 1–11.

[37] P. Ladevèze, The large time increment method for the analysis of structures with non-linear behavior described by internal variables, COMPTES RENDUS DE L ACADEMIE DES SCIENCES SERIE II 309 (11) (1989) 1095–1099.

[38] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, Computer Methods in Applied Mechanics and Engineering 304 (2016) 81–101.

[39] S. Mohanty, R. Teale, A. Chattopadhyay, P. Peralta, C. Willhauck, Mixed gaussian process and state-space approach for fatigue crack growth prediction, in: International workshop on structural heath monitoring, Vol. 2, 2007, pp. 1108–1115.

[40] S. Cheng, M. Pecht, A fusion prognostics method for remaining useful life prediction of electronic products, in: Automation Science and Engineering, 2009. CASE 2009. IEEE International Conference on, IEEE, 2009, pp. 102–107.

[41] Y. Xing, Q. Miao, K.-L. Tsui, M. Pecht, Prognostics and health monitoring for lithium-ion battery, in: Intelligence and Security Informatics (ISI), 2011 IEEE International Conference on, IEEE, 2011, pp. 242–247.

[42] C. Pellegrino, U. Galvanetto, B. Schrefler, Numerical homogenization of periodic composite materials with non-linear material components, International Journal for Numerical Methods in Engineering 46 (10) (1999) 1609–1637.

[43] T. Sussman, K.-J. Bathe, A model of incompressible isotropic hyperelastic material behavior using spline interpolations of tension–compression test data, Communications in numerical methods in engineering 25 (1) (2009) 53–63.

[44] R. De Maesschalck, D. Jouan-Rimbaud, D. L. Massart, The mahalanobis distance, Chemometrics and intelligent laboratory systems 50 (1) (2000) 1–18.

[45] R. Hill, On constitutive macro-variables for heterogeneous solids at finite strain, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, Vol. 326, The Royal Society, 1972, pp. 131–147.

[46] T. I. Zohdi, P. Wriggers, An introduction to computational micromechanics, Springer Science & Business Media, 2008.

[47] M. Brieu, F. Devries, Micro-mechanical approach and algorithm for the study of damage appearance in elastomer composites, Composite structures 46 (4) (1999) 309–319.

[48] C. Miehe, J. Schotte, J. Schröder, Computational micro–macro transitions and overall moduli in the analysis of polycrystals at large strains, Computational Materials Science 16 (1) (1999) 372–382.

[49] X. F. Xu, X. Chen, Stochastic homogenization of random elastic multiphase composites and size quantification of representative volume element, Mechanics of Materials 41 (2) (2009) 174–186.

[50] V. Kouznetsova, W. Brekelmans, F. Baaijens, An approach to micro-macro modeling of heterogeneous materials, Computational Mechanics 27 (1) (2001) 37–48.

[51] J. Ma, J. Zhang, L. Li, P. Wriggers, S. Sahraee, Random homogenization analysis for heterogeneous materials with full randomness and correlation in microstructure based on finite element method and monte-carlo method, Computational Mechanics 54 (6) (2014) 1395–1414.

[52] J. Ma, S. Sahraee, P. Wriggers, L. De Lorenzis, Stochastic multiscale homogenization analysis of heterogeneous materials under finite deformations with full uncertainty in the microstructure, Computational Mechanics 55 (5) (2015) 819–835.

[53] S. Sakata, F. Ashida, T. Kojima, Stochastic homogenization analysis for thermal expansion coefficients of fiber reinforced composites using the equivalent inclusion method with perturbation-based approach, Computers & structures 88 (7) (2010) 458–466.

[54] B. Ganapathysubramanian, N. Zabaras, Modeling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multiscale method, Journal of Computational Physics 226 (1) (2007) 326–353.

[55] L. Isserlis, On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, Biometrika 12 (1/2) (1918) 134–139.

Highlights

- A hybrid mathematical modeling and data-driven approach is proposed
- Data uncertainty managed by a proper metric, the Mahalanobis distance
- Good convergence, accuracy and flexibility with low computational demand were found

A new reliability-based data-driven approach for noisy experimental data with physical constraints

Jacobo Ayensa-Jiménez, Mohamed H. Doweidar, Jose A. Sanz-Herrera, Manuel Doblaré

PII: S0045-7825(17)30425-5
DOI: http://dx.doi.org/10.1016/j.cma.2017.08.027
Reference: CMA 11568

To appear in: *Comput. Methods Appl. Mech. Engrg.*

Received date : 13 March 2017
Revised date : 21 July 2017
Accepted date : 20 August 2017