C. R. Mecanique 347 (2019) 831–844

Contents lists available at ScienceDirect

Comptes Rendus Mecanique

www.sciencedirect.com

Data-Based Engineering Science and Technology / *Sciences et technologies de l'ingénierie basées sur les données*

Data-driven computation for history-dependent materials

Pierre Ladevèze *, David Néron, Paul-William Gerbaud

LMT (ENS Paris-Saclay, CNRS, Université Paris-Saclay), 61, av. du Président-Wilson, 94235 Cachan, France

ARTICLE INFO

Article history:
Received 26 September 2019
Accepted 1 November 2019
Available online 18 November 2019

Keywords:
Data-driven
History-dependent materials
Computational mechanics
Big data
Experimental constitutive manifold
Material mechanics

ABSTRACT

This paper introduces a new vision of data-driven structure computation taking advantage of Material Science, especially for highly nonlinear and time-dependent material behaviours. Technical solutions are also derived, in order to build internal hidden variables defining the so-called "Experimental Constitutive Manifold".

1. Introduction

Data-driven structure calculation has recently become one of the issues impacted by the area of research linked to big data. The pioneering works were due to Ortiz and Chinesta and their teams [1–6], in which Material Science was either annihilated nor minimised. Several developments have followed, based on the same concepts, and one can cite for example [7–11]. However, as in [7] a pertinent material knowledge is added. These works provide original and very effective responses in some situations but they are limited to really specific material behaviours. It follows that the treatment of complex material behaviours, such as viscoplasticity with several internal variables, remains a real challenge.

The present work is a tentative answer for history-dependent material behaviours, such as viscoplasticity under small perturbation hypothesis, and especially a paradigm for the most complex behaviours [12,13]. The approach proposed herein is based, first of all, on the separation of the governing equations of the mechanical problem, defined over the space-time domain: equilibrium and compatibility equations in one hand and constitutive equations in the other hand. This separation is the basis of the LATIN solver, which will be the well-suited calculation method [14–17]. Moreover, this separation is also the basis of the Constitutive Relation Error method, developed for both validation and verification [14,18,17,19]. The first group of equations can be qualified as "exact", in contrast to the constitutive equations that strongly depends on experimental data.

Only the material constitutive relations are impacted by the experimental data. In the future, these data will be capitalised and therefore will densely cover the area of interest. These experimental points define what we call the "Experimental Constitutive Manifold" (ECM) whose construction is not a triviality when one considers nonlinear and time-dependent behaviours. The data-driven computational approach that is proposed is based on the classical internal variable approach that is used today to describe the material state. However, here the internal hidden variables are not a priori known and

* Corresponding author.
E-mail address: ladeveze@ens-paris-saclay.fr (P. Ladevèze).

https://doi.org/10.1016/j.crme.2019.11.008

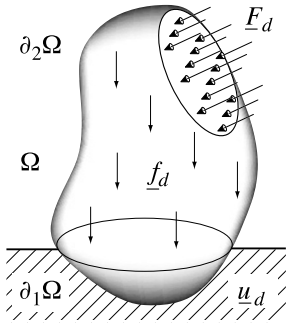

Fig. 1. The computational problem.

computation tools are introduced to determine them from raw experimental data. They are used to obtain the necessary structuring of the Experimental Constitutive Manifold.

This approach is a real challenge because its application requires an amount of experimental information that even if it is minimised, remains gigantic and therefore prohibitive. It is no longer a question of placing oneself beside the knowledge of Material Science accumulated for centuries and which continues to develop, but on the contrary, to exploit it, to incorporate it to try to do better, of course, with the methods developed in the field of big data. It is a change of philosophy when compared to the pioneering works.

The first part of the paper introduces the basic ideas of data-driven computation. In order to simplify the presentation, it will be applied to relatively simple material behaviours, even if it allows us to highlight the difficulties. Then, and this is the essence of this work, the data-driven computation will be extended to complex material behaviours. In particular, it will be shown that approximation computation does not lead to serious difficulties using the LATIN solver. Difficulties are elsewhere, in the construction of the Experimental Constitutive Manifold, for which several computational tools inspired from big data methods are given.

2. Basic ideas and difficulties

2.1. Today's computational approach

Let us consider the quasi-static evolution of a structure Ω whose state is defined by its displacement field $\underline{u}$, strain field $\boldsymbol{\varepsilon}$ and stress field $\boldsymbol{\sigma}$. This structure is subjected to prescribed body forces $\underline{f}_d$, traction forces $\underline{F}_d$ over a part $\partial_2\Omega$ of the boundary, and displacements $\underline{u}_d$ over the complementary part $\partial_1\Omega$ (see Fig. 1). For the sake of simplicity, let us consider that the material behaviour can be described as nonlinear elastic.

In order to get closer to the formulation used in the data-driven computational approach, the computational problem that must be solved consists in *finding the strain-stress pair* $\boldsymbol{s} = (\boldsymbol{\varepsilon}, \boldsymbol{\sigma}) \in \mathbf{L}^2(\Omega)^{12}$, *which satisfies:*

$$(\mathbf{A_d}): \begin{cases} \text{compatibility equation: } \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\underline{u}), \text{ with } \underline{u} \in \boldsymbol{U}_{\text{ad}} \\ \text{equilibrium equation: } \boldsymbol{\sigma} \in \boldsymbol{S}_{\text{ad}} \end{cases}$$

$$(\boldsymbol{\Gamma}): \text{ constitutive relations: } \boldsymbol{\sigma} = \mathbb{K} : \boldsymbol{\varepsilon}(\underline{u}) \text{ over } \Omega,$$

where $\mathbb{K}$ is the elasticity tensor.

The first important point is to see the first group of equations as "exact", in contrast to the second group, which strongly depends on experimental data. Such splitting is not new, it is the starting point of the "Constitutive Relation Error" (CRE) [19,18,14] and also of the LATIN-PGD [17,16]. To go further, one introduces an abstract geometrical scheme, as illustrated in Fig. 2. In this scheme, the "exact" solution lies at the intersection of the manifold ($\boldsymbol{\Gamma}$) and the "exact" admissible space ($\mathbf{A_d}$). It is also the admissible point $\boldsymbol{s}$ that minimises the distance to the manifold ($\boldsymbol{\Gamma}$). Today, Material Science is used to construct this manifold ($\boldsymbol{\Gamma}$). From experimental observations, an analytical material model is proposed, whose parameters are identified through several experimental tests.

2.2. Data-driven computation

Ideal situation — The only change consists in replacing manifold ($\boldsymbol{\Gamma}$) by experimental points. This discrete manifold is then named "Experimental Constitutive Manifold" (ECM). The geometrical interpretation of the problem that must be solved is then described in Fig. 3.

Exact solution and approximations — The exact solution is still the admissible point that minimises the distance to the manifold (ECM). Concerning the practical computation, the natural way is to use the solver LATIN, which separates the exact equations from the experimental data-dependent equations. More details will be given later in this article, but at the point,

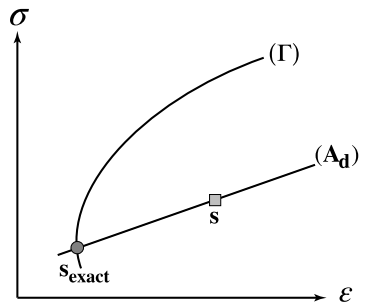

Fig. 2. A geometrical scheme of the computation method.

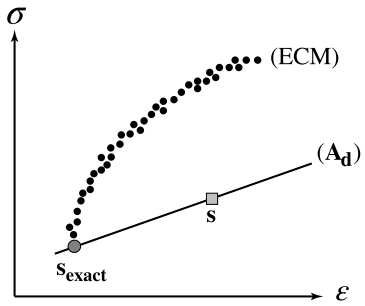

Fig. 3. The data-driven computational approach.

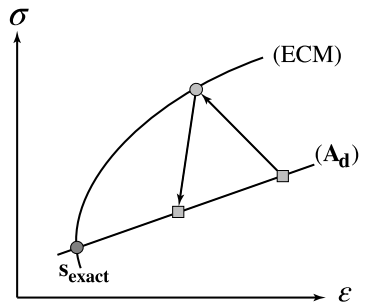

Fig. 4. Approximation scheme.

let us note that it is an iterative algorithm that consists in solving alternatively the two groups of equations. An illustration of the iterative solver is represented in Fig. 4.

An illustration — Here, we consider the 2D example given in [5], where the experimental data are simulated for a nonlinear elastic material model. The loading and boundary conditions are illustrated in Fig. 5. 10 loading directions and 10 increments of loading amplitude are considered, leading to 100 different loading scenarios. The experimental data taken into account are the displacement of 11 points equally distributed on the upper side of the specimen and that are supposed to be experimentally accessible. The ECM can then be computed under the assumption of a homogeneous specimen. In the last step, a dimensionality reduction is done using the kPCA method. The final results are represented in Fig. 6, the ECM being represented in a 2D space. Two remarks shall be addressed, which both emphasise the necessity to work in synergy with Material Science.

Remarks — Firstly, 2D ECM obtained in [5] is only one part of the complete one, as Material Science points out that the ECM dimension is greater or equal to the number of the strain components, i.e. three here. Therefore, additional "tests" have been performed to fully compute the ECM, which is represented in Fig. 7. Secondly, it must be noted that the number of coordinates obtained through kernel-PCA method can be larger than three. If it is the case, this means that the behaviour cannot be seen as nonlinear elastic. One should, therefore, turn toward Material Science to know exactly what to perform in terms of observations and additional tests.

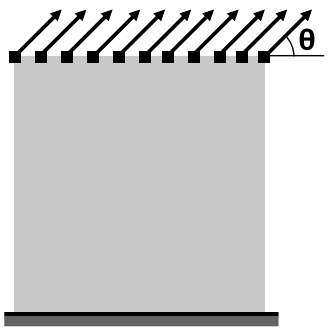

Fig. 5. Illustration from [5]: loading cases.

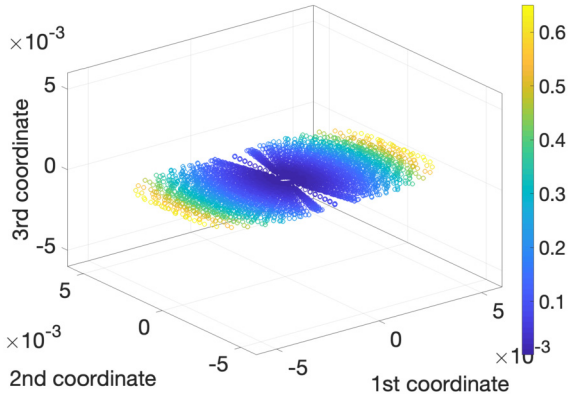

Fig. 6. Final ECM (from [5]).

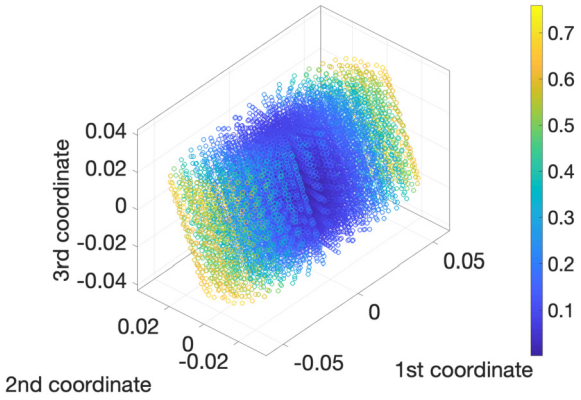

Fig. 7. The complete ECM.

2.3. Data-driven computation: a general approach

Main lines — First, we consider that it is not possible today to get enough experimental points in order to properly describe the material behaviour, due to the "holes" between experimental points. Consequently, we propose to first construct from available experimental points our reference, the discrete ECM. Its dimension can be reduced through big data methods. In addition, extrapolation techniques based on Material Science knowledge and observations should be used.

Interest — The main interest of this data-driven computational approach is the easy comparison between calculated points and experimental ones. In particular, the coincidence of those points is the best criterion one can have in terms of validation. In practice, the comparison is done using "experimental proximity indicators" that point out when additional tests are required. Based on this synergy between data and Material Science, one can think that a new generation of ECM descriptors can be proposed and solve several open questions in today's Material Science (e.g. anisotropy in viscoplasticity, ageing, fatigue, etc).

3. Extension to complex behaviours

3.1. Today's computational approach

Let us consider again the quasi-static evolution over time interval $[0, T]$ of a structure Ω, subjected to prescribed body forces $\underline{f}_d$, traction forces $\underline{F}_d$ over a part $\partial_2\Omega$ of the boundary, and displacements $\underline{u}_d$ over the complementary part $\partial_1\Omega$ (see Fig. 1). The state is assumed to be defined by $\boldsymbol{s} = (\dot{\boldsymbol{\varepsilon}}_p, \boldsymbol{\sigma})$ where $\dot{\boldsymbol{\varepsilon}}_p$ designates the inelastic strain rate and $\boldsymbol{\sigma}$ the stress. Assuming the elasticity operator to be known, the problem is defined as: *Find* $\boldsymbol{s} = (\dot{\boldsymbol{\varepsilon}}_p, \boldsymbol{\sigma}) \in \boldsymbol{s}^{[0,T]}$ *such that:*

$$(\mathbf{A_d}): \begin{cases} \text{compatibility equation: } \boldsymbol{\varepsilon}_p = \boldsymbol{\varepsilon}(\underline{u}) \ - \ \mathbb{K}^{-1}\boldsymbol{\sigma}, \ \text{with } \underline{u} \in \boldsymbol{U}_{\text{ad}}^{[0,T]} \\ \text{equilibrium equation: } \boldsymbol{\sigma}, \ \text{with } \boldsymbol{\sigma} \in \boldsymbol{S}_{\text{ad}}^{[0,T]} \end{cases}$$

$$(\boldsymbol{\Gamma}): \ \text{constitutive relations over } \Omega \times [0, T]$$

Again, we see the first group of equations as "exact", in contrast to the second group of equations, which strongly depends on experimental data. Let us recall that such a splitting is not new, it is also the basic step of the Constitutive Relation Error [14,18,19], and also of the LATIN-PGD method [15–17] for complex behaviour. To describe the material, Material Science proposes two approaches. The first one is the so-called "Functional Approach" where the stress at a given time t is a functional of the strain history until this time t. Its identification needs all possible experimental stress-strain histories, which is far from being easy. Today, one prefers the second approach name "Internal Variable Approach," for which identification needs much less data. With this approach, the state of the material at time t, which depends on the material history, is completely defined by the values at time t of the inelastic strain rate / stress pair $(\dot{\boldsymbol{\varepsilon}}_p, \boldsymbol{\sigma})$, and of several internal (hidden) variables $(\dot{\boldsymbol{X}}, \boldsymbol{Y})$. One can say that the Functional Approach is an Internal Variable Approach with an infinite number of hidden variables. Precisely, the Internal Variables Approach used here leads to a state defined by:

$$(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{X}}, \boldsymbol{\sigma}, \boldsymbol{Y}) \tag{1}$$

where $(\boldsymbol{X}, \boldsymbol{Y})$ are additional internal variables. $\boldsymbol{X}$ is the kinematic one and $\boldsymbol{Y}$ its associated force. $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same dimension. The constitutive relations can be split into:

- state equation:

$$\boldsymbol{Y} = \boldsymbol{\Lambda}(\boldsymbol{X}) \tag{2}$$

- state evolution laws:

$$\begin{bmatrix} \dot{\boldsymbol{\varepsilon}}_p \\ -\dot{\boldsymbol{X}} \end{bmatrix} = \boldsymbol{B}\left(\begin{bmatrix} \boldsymbol{\sigma} \\ \boldsymbol{Y} \end{bmatrix}\right), \ \text{with } \boldsymbol{\varepsilon}_p, \ \boldsymbol{X} = \boldsymbol{0} \text{ at } t = 0$$

A fundamental simplification comes from [16,20]. Additional internal (hidden) variables are not intrinsic and thus, an internal variables transformation can be done. A remarkable choice is the so-called "normal formulation," for which

$$\boldsymbol{X} = \boldsymbol{Y} \tag{3}$$

It follows that the only part of the constitutive relation depending on experimental data are the state evolution laws. Finally, the problem to solve can be defined by the following abstract geometrical sketch (Fig. 8). Again, the exact solution is the admissible point that lies at the minimum distance with the manifold $(\boldsymbol{\Gamma})$. Material Science is also used to identify the manifold $(\boldsymbol{\Gamma})$ from experimental data.

Remark — Of course, the two principles of Thermodynamics are satisfied. Moreover, the particular family of the so-called "standard materials" is defined by the energy and an additional potential.

3.2. Data-driven computation

The ideal situation is described in Fig. 9, in which experimental points cover $(\boldsymbol{\Gamma})$, which is now named Experimental Constitutive Manifold (ECM).

3.2.1. Exact solution and approximation

The exact solution is also here the admissible point that minimises its distance with the ECM. Approximations can be easily computed thanks to the LATIN method (see Fig. 10). The interested reader can refer to [15] to find more details on the method and, in particular, on the PGD usage. The linear stage $\hat{\boldsymbol{s}}_{n+1/2} \rightarrow \boldsymbol{s}_{n+1}$ does not lead to any difficulty.

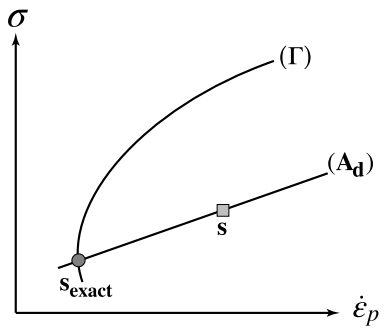

Fig. 8. Today's computational approach: a geometrical sketch.

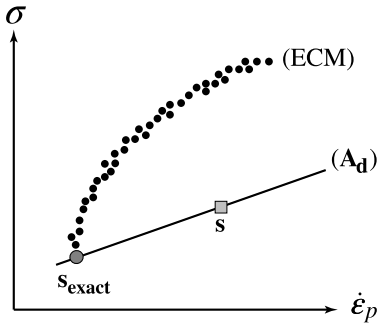

Fig. 9. The data-driven computational approach.

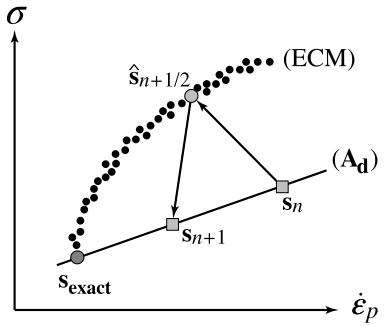

Fig. 10. The LATIN solver.

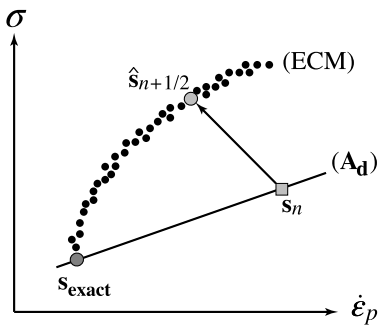

Fig. 11. The LATIN local stage.

Let us consider only the local stage $\boldsymbol{s}_n \to \hat{\boldsymbol{s}}_{n+1/2}$ illustrated in Fig. 11. Here, solution $\hat{\boldsymbol{s}}_{n+1/2}$ is find by minimizing the residual $\boldsymbol{R}^i$ associated with $\hat{\boldsymbol{s}}^i \in$ (ECM):

$$\boldsymbol{R}^i = (\hat{\boldsymbol{\varepsilon}}^i_{p,n+1/2} - \dot{\boldsymbol{\varepsilon}}^i_{p,n}) + \boldsymbol{H}(\hat{\boldsymbol{\sigma}}^i_{n+1/2} - \boldsymbol{\sigma}^i_n) \tag{4}$$

In the case where the dispersion is large, one can consider the barycentric manifold associated with (ECM) and thus:

$$\hat{\boldsymbol{s}}_{n+1/2} = \sum_{i\in N} p_i\, \hat{\boldsymbol{s}}^i \tag{5}$$

with $\sum_{i\in N} p_i = 1$, $p_i = \gamma_i/(\sum_{j\in N}\gamma_j)$ and $\gamma_i = \exp(-\beta|\mathbf{R}^i|^2)$, where β is a positive parameter. To summarise, one can say that approximation computation does not lead to serious difficulties. Difficulties are in the ECM construction.

3.3. The ECM construction

Raw experimental data — We suppose that the following quantities are known:

- the elasticity tensor $\mathbb{K}$,
- a set of time-histories of inelastic strain rate / stress pair,

$$\boldsymbol{s}^i = (\dot{\varepsilon}^i_p, \boldsymbol{\sigma}^i),\ i \in \bar{\mathbb{N}} \tag{6}$$

The $\boldsymbol{s}^i(t)$, $i \in \bar{\mathbb{N}}$ are rewritten in terms of "material time" thanks to the cumulated inelastic strain, that is:

$$p^i = \int_0^T |\dot{\boldsymbol{\varepsilon}}^i_p|\, \mathrm{d}t \tag{7}$$

Consequently, available data are:

$$\boldsymbol{s}^i(p) = (\boldsymbol{\varepsilon}^i_{p,p}, \boldsymbol{\sigma}^i, \dot{p}^i)(p), i \in \bar{\mathbb{N}}(p) \tag{8}$$

Finally, **ECM** is defined as:

$$\mathbf{ECM} = \{\boldsymbol{\Sigma}_p \mid p \in [0, \bar{p}]\} \tag{9}$$

with

$$\boldsymbol{\Sigma}_p = \{\boldsymbol{s}^i(p) \mid i \in \bar{\mathbb{N}}\} \tag{10}$$

Structured ECM — **ECM** without its additional hidden variable structuring cannot be used in the general case. This structuring for one additional hidden variable $\boldsymbol{X}$ is the following:

$$\boldsymbol{\varepsilon}_{p,p} = \boldsymbol{g}(\boldsymbol{\sigma}, \boldsymbol{X}, p) \quad \dot{p} > 0 \quad \text{sub-manifold } (\mathbf{G}(p)) \tag{11}$$

$$\boldsymbol{X}_{,p} = \boldsymbol{h}(\boldsymbol{\sigma}, \boldsymbol{X}, p) \quad \dot{p} > 0 \quad \text{sub-manifold } (\mathbf{H}(p)) \tag{12}$$

$$\dot{p} = f(\boldsymbol{\sigma}, \boldsymbol{X}, p) \quad \dot{p} \geqslant 0 \quad \text{sub-manifold } (\mathbf{F}(p)) \tag{13}$$

with $\boldsymbol{X} = 0$ and $\boldsymbol{\varepsilon}_p = 0$ at $t = 0$. Moreover, functions $\boldsymbol{g}, \boldsymbol{h}, f$ are single-valued. Building $\boldsymbol{X}$ and the associated sub-manifolds is named the "central problem". In the next paragraph, specific tools inspired by classical big data methods are derived.

Remarks — It is assumed, as usual, that $\dot{p} = 0$ implies $\dot{\boldsymbol{X}} = 0$. Moreover, the previous formulation is not restricted to "normal" material (see section 3.1).

3.4. The ECM central problem

3.4.1. Basic tools

The problem model — To introduce the different tools, let us consider the following central problem where the raw experimental data are:

$$\boldsymbol{\Sigma} = \{(\boldsymbol{A}, \boldsymbol{B})^i \mid \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^n,\ i \in \mathbb{N}\} \tag{14}$$

where $\boldsymbol{A}^i = \boldsymbol{\alpha}(\boldsymbol{B}^i, \boldsymbol{C}^i)$, $\boldsymbol{C}^i$ being a hidden variable. The central problem is then:

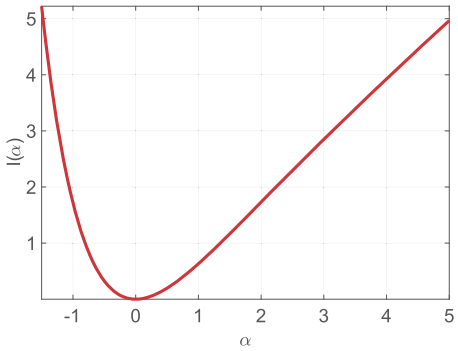

Fig. 12. The function **I**.

Find $\boldsymbol{C} = \{\boldsymbol{C}^i \mid i \in \mathbb{N}\}$ *such that function* f *is single valued:*

$$f \ : \ (\boldsymbol{B}^i, \boldsymbol{C}^i) \longrightarrow \boldsymbol{A}^i, \ \forall i \in \mathbb{N} \qquad (15)$$
$$\boldsymbol{\Sigma_B} \times \boldsymbol{\Sigma_C} \quad \boldsymbol{\Sigma_A}$$

Separator **P** — This operator is classical. It makes it possible to separate $\boldsymbol{B}^i$ and its close neighbours from the rest of the points:

$$P_{i,j} = \mathbf{P}(\boldsymbol{B}^i, \boldsymbol{B}^j) = \frac{\Gamma^{i,j}}{\sum\limits_{j \in N} \Gamma^{i,j}} \qquad (16)$$

with

$$\Gamma^{i,j} = \exp\Big(- \frac{||\boldsymbol{B}^i - \boldsymbol{B}^j||^2}{b^2 \epsilon^2} \Big) \qquad (17)$$

where b is a normalisation constant and ϵ a small parameter. One has:

$$\begin{cases} \dfrac{||\boldsymbol{B}^i - \boldsymbol{B}^j||}{b} = O(\epsilon) \Leftrightarrow P_{i,j} = O(1) \\ \text{otherwise} \qquad\qquad\qquad P_{i,j} \approx 0 \end{cases} \qquad (18)$$

The univocity indicator — Let us introduce the function **I**, which is a variant of the Kullback–Leibler divergence. **I** plotted in Fig. 12 has for expression:

$$\mathbf{I}(\alpha) = (\exp(-\alpha) - 1)(-\alpha), \ \alpha \ \in \ ]-\infty, +\infty[\qquad (19)$$

The univocity indicator is then defined by:

$$\boldsymbol{U}(\boldsymbol{B}; \boldsymbol{A}) = \sum_{i \in N} \sum_{j \in N} u^{ij}, \qquad (20)$$

with

$$u^{ij} = \mathbf{P}(\boldsymbol{B}^i, \boldsymbol{B}^j) \mathbf{I}(\alpha^{ij}), \qquad (21)$$

$$\alpha^{i,j} = \frac{1}{\epsilon^2} \left[\frac{||\boldsymbol{B}^i - \boldsymbol{B}^j||^2}{b^2} - \frac{||\boldsymbol{A}^i - \boldsymbol{A}^j||^2}{a^2} \right] \qquad (22)$$

and where ϵ is a small parameter, whereas a, b are normalisation constants. Its interpretation comes from the $\boldsymbol{U}^{ij}$ contribution:

$$\begin{cases} \text{if } ||\boldsymbol{B}^i - \boldsymbol{B}^j|| > O(\epsilon) & \blacktriangleright \ u^{ij} \approx 0 \\ \text{if } ||\boldsymbol{B}^i - \boldsymbol{B}^j|| = O(\epsilon) \begin{cases} \text{if } ||\boldsymbol{A}^i - \boldsymbol{A}^j|| = O(\epsilon) & \blacktriangleright \ u^{ij} = O(1) \\ \text{if } ||\boldsymbol{A}^i - \boldsymbol{A}^j|| \neq O(\epsilon) & \blacktriangleright \ u^{ij} \gg 1 \end{cases} \end{cases} \qquad (23)$$

Function f introduced previously is singular-valued if the manifold Σ satisfies to the following criterion:

$$\boldsymbol{U}(\boldsymbol{B};\boldsymbol{A}) < \sum_{i\in N}\sum_{j\in N}\mathbf{P}(\boldsymbol{B}^i,\boldsymbol{B}^j)\mathbf{I}\Big(k\frac{||\boldsymbol{B}^i-\boldsymbol{B}^j||^2}{b^2\epsilon^2}\Big),\quad k=O(1) \tag{24}$$

where k is a parameter. Otherwise, one internal hidden variable should be added to get a single-valued function f. Unknowns are $\boldsymbol{C}$ and Σ_C. The separator $\mathbf{P}$ and the univocity indicator $\boldsymbol{U}$ are extended to take into account the variable $\boldsymbol{C}$. Therefore, $\boldsymbol{C}$ is defined modulo a $\boldsymbol{B}$-function according to:

$$\begin{aligned}&\boldsymbol{C}=\arg\min[\boldsymbol{U}(\boldsymbol{B},\boldsymbol{C}';\boldsymbol{A})]\\&\quad\boldsymbol{C}'\in[\mathbb{R}^n]^N\end{aligned} \tag{25}$$

Therefore, to get a unique solution, additional constraints are added through the potential $\boldsymbol{s}(\boldsymbol{B},\boldsymbol{C})$. They are related to material symmetries, $\boldsymbol{B}/\boldsymbol{C}$ non correlation, etc. The potential associated with the $\boldsymbol{B}/\boldsymbol{C}$ non correlation is:

$$\boldsymbol{s}(\boldsymbol{B},\boldsymbol{C})=\delta\Big[\sum_{i\in N}\sum_{j\in N}\mathbf{P}(\boldsymbol{B}^i,\boldsymbol{B}^j)(\boldsymbol{B}^i-\boldsymbol{B}^j)\boldsymbol{C}^{i\intercal}\Big]^2 \tag{26}$$

where δ is a positive constant. Finally, one can give an explicit solution to the model of the problem model:

$$\boldsymbol{C}^i=\boldsymbol{A}^i-\sum_{j\in N}\mathbf{P}_{ij}(\boldsymbol{B}^i,\boldsymbol{B}^j)\boldsymbol{A}^j \tag{27}$$

*3.4.2. Application to the **ECM***

Let us consider the situation where the **ECM** to be structured needs one additional hidden variable $\boldsymbol{X}\in[\mathbb{R}^n]^{\bar{N}}$, $\bar{N}$ being the total number of experimental points and $\bar{p}$ being the final value of the cumulated inelastic strain. Introducing the following univocity indicators:

$$\begin{aligned}\bar{\boldsymbol{U}}(\boldsymbol{\sigma},\boldsymbol{X};\boldsymbol{\varepsilon}_{p,p}) &= \int_0^{\bar{p}}\mathrm{d}p\;\gamma(p)\,\boldsymbol{U}(\boldsymbol{\sigma},\boldsymbol{X};\boldsymbol{\varepsilon}_{p,p})\\ \bar{\boldsymbol{U}}(\boldsymbol{\sigma},\boldsymbol{X};\boldsymbol{X}_{,p}) &= \int_0^{\bar{p}}\mathrm{d}p\;\gamma(p)\,\boldsymbol{U}(\boldsymbol{\sigma},\boldsymbol{X};\boldsymbol{X}_{,p}),\qquad \gamma(p)=1-\exp\Big(\frac{-\dot{p}}{\epsilon}\Big)\\ \bar{\boldsymbol{U}}(\boldsymbol{\sigma},\boldsymbol{X};\dot{p}) &= \int_0^{\bar{p}}\mathrm{d}p\;\boldsymbol{U}(\boldsymbol{\sigma},\boldsymbol{X};\dot{p})\end{aligned} \tag{28}$$

with ϵ being a positive parameter and $\gamma(p)$ allowing a smooth separation of the domains $\dot{p}>0$ and $\dot{p}=0$, and additional constraints, such as the non correlation between $\boldsymbol{\sigma}$ and $\boldsymbol{X}$:

$$\bar{\boldsymbol{s}}(\boldsymbol{\sigma},\boldsymbol{X})=\int_0^{\bar{p}}\mathrm{d}p\;\boldsymbol{s}(\boldsymbol{\sigma},\boldsymbol{X}) \tag{29}$$

the $\boldsymbol{X}$-solution is defined as:

$$\begin{aligned}&\boldsymbol{X}=\arg\min[\bar{\boldsymbol{U}}(\boldsymbol{\sigma},\boldsymbol{X}';\boldsymbol{\varepsilon}_{p,p})+\bar{\boldsymbol{U}}(\boldsymbol{\sigma},\boldsymbol{X}';\boldsymbol{X}'_{,p})+\bar{\boldsymbol{U}}(\boldsymbol{\sigma},\boldsymbol{X}';\dot{p})+\bar{\boldsymbol{s}}(\boldsymbol{\sigma},\boldsymbol{X}')]\\&\quad\boldsymbol{X}'\in[\mathbb{R}^n]^N\end{aligned} \tag{30}$$

Additional properties, especially on $\boldsymbol{X}$ are welcome.

*3.5. The extended **ECM***

The data-driven computational approach needs to extend **ECM** to any $\dot{\boldsymbol{\varepsilon}}_p$ history (Fig. 13). That is done in two stages:
ECM construction — **ECM** is defined by the raw experimental data and its associated internal hidden variables obtained from the solving of the central problem. The last step is to reduce its dimensionality using a big data method.
*Extended **ECM*** — **ECM** is completed through extrapolation, for which observations and Material Science are absolutely essential. To manipulate **ECM**, the analytical model can be used (polynomial, FE, Material Science). Pointwise description can be also used.

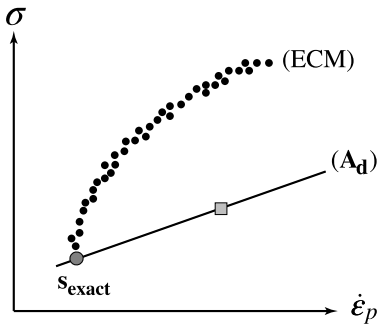

Fig. 13. The true data-driven computational approach.

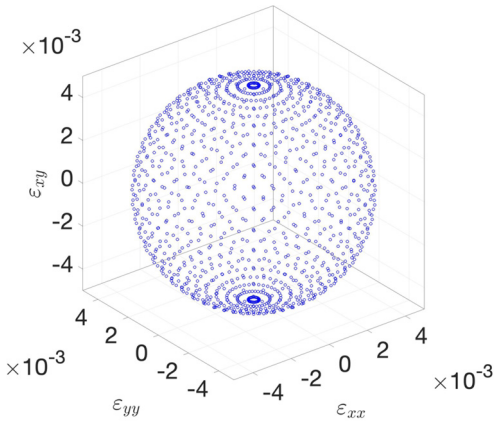

Fig. 14. Deformation trajectories.

3.6. ECM identification

A remark should be addressed, concerning the identification of ECM, which will be the subject of a companion paper. Large amount of data can be obtained today thanks to new experimental methods [21]. Data for the experimental test $\mathbf{T}(k)$, $k \in \mathbb{N}$ are defined by the admissible manifold $(\mathbf{A_d}^k)$, $k \in \mathbb{N}$. Consequently, **ECM** is determined by:

$$\mathbf{ECM} = \underset{\overline{\mathbf{ECM}}' \in \{\overline{\mathbf{ECM}}\}}{\arg\min} \left[\sum_{k \in Z} \min_{s_a \in A_d^k} \int_{\Omega^k} d^2\,(s_a, \overline{\mathbf{ECM}}')\ \mathrm{d}\Omega \right] \tag{31}$$

A large literature is devoted to such an identification problem, which is basically ill-posed. Iterative process can be proposed such that

$$\mathbf{ECM}^{n+1} = \mathbf{ECM}^{n} + \boldsymbol{\Delta}\mathbf{ECM}^{n+1} \tag{32}$$

The resulting **ECM** describes the material behaviour through the experimental tests, observations and Material Science. The number of its hidden internal variables increases with the degree of complexity of the considered material's histories.

Important remark — "Experimental" points can be obtained from microstructure simulations related to a defined class of time-history loadings. Here, the proposed data-driven computational approach gives a new answer to the construction of homogenized constitutive relation in nonlinear mechanics from a known microstructure.

3.7. A first illustration

Here, one studies a viscoplastic material under the plane strain assumption.

Data: experimental points — "Experimental" points are generated using the classical Chaboche viscoplastic model with linear kinematic hardening. A set of triaxial deformation cyclic loadings is applied, proportional to the one described on Fig. 14 (spherical envelope). It is parametrised with two angular parameters. The loading amplitude evolution is represented on Fig. 15. A total of 1,296 tests are done. The p-interval is discretised with 300 points for both the $\dot{p} > 0$ and $\dot{p} = 0$ domains. For a given p, the **ECM** is defined by the following amount of experimental points:

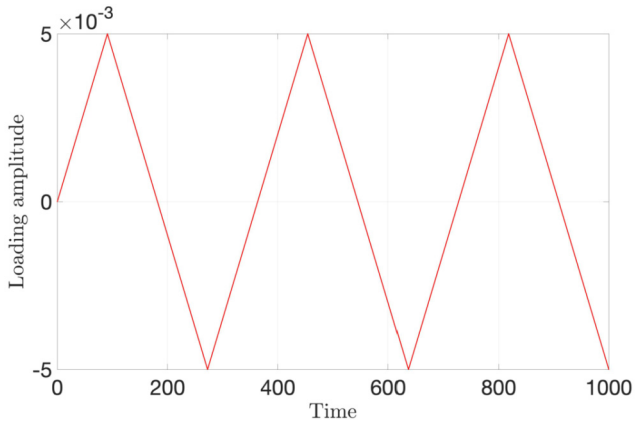

Fig. 15. Loading amplitude evolution.

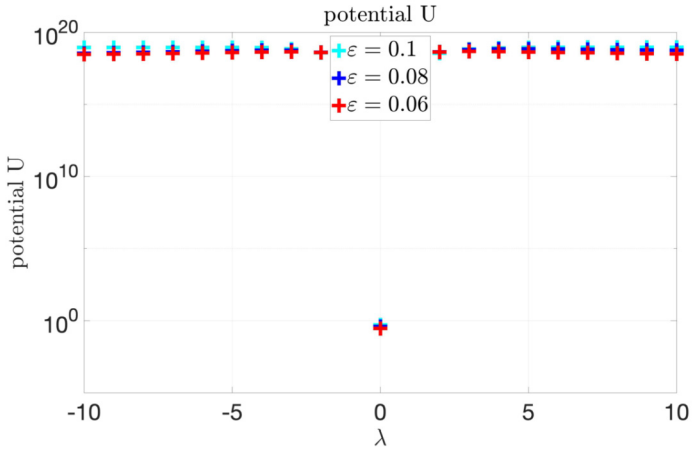

Fig. 16. Potential U as a function of λ, for 3 values of the parameter ϵ.

$$\text{for } \dot{p} > 0,\ N_p = 1,296 \tag{33}$$
$$\text{for } \dot{p} = 0,\ N_p \approx 100$$

For the domain $\dot{p} = 0$, the amount of points is variable for a given p.
The total amount of 'experimental" points resulting is 388,800 for the domain $\dot{p} > 0$ and $\approx$ 30, 000 for the domain $\dot{p} = 0$.

Hidden variable computation — The total potential associated with the raw ECM is defined as

$$\bar{\mathbf{U}}_{\text{total}} = \bar{\mathbf{U}}(\boldsymbol{\sigma}; \boldsymbol{\varepsilon}_{p,p}) + \bar{\mathbf{U}}(\boldsymbol{\sigma}; \dot{p}) \tag{34}$$

Its value is $\approx 3 \cdot 10^8$, which must be compared to the criterion value of 0.44 (see Equation (24) with the field $\boldsymbol{\sigma}$ here instead of $\mathbf{B}$). The parameters values are $k = 1$ and $\epsilon = 0.06$. Therefore, it is necessary to introduce one hidden variable, which can be a scalar, a vector, etc.
The total potential without the 4th potential $\bar{\mathbf{s}}(\boldsymbol{\sigma}, \mathbf{X})$ is defined as

$$\bar{\mathbf{U}}_{\text{total}} = \bar{\mathbf{U}}(\boldsymbol{\sigma}, \mathbf{X}; \boldsymbol{\varepsilon}_{p,p}) + \bar{\mathbf{U}}(\boldsymbol{\sigma}, \mathbf{X}; \mathbf{X}_{,p}) + \bar{\mathbf{U}}(\boldsymbol{\sigma}, \mathbf{X}; \dot{p}) \tag{35}$$

It should be minimised in order to get the new hidden internal variable $\mathbf{X}$. The algorithm necessary to obtain such solution will be developed in a companion paper. Here, the aim is to show that a clear minimum exists. For such a purpose, one introduces a perturbation to the "exact" $\mathbf{X}_{,p}$, which is random at any value of p:

$$\tilde{\mathbf{X}}_{,p} = \mathbf{X}_{,p} + \lambda \cdot \Delta \mathbf{X}_{,p} \tag{36}$$

where $\Delta \mathbf{X}_{,p} = \mathbf{X}_{,p} - \mathbf{r} \cdot \mathbf{X}_{,p}$, $\mathbf{r}$ is a field of random values in $[0.8, 1.2]$, $\lambda \in [-10, 10]$. Fig. 16 and its zoom for small values of λ Fig. 17 show the behaviour of the potential in terms of the perturbation. It exhibits a clear minimum for the exact solution.

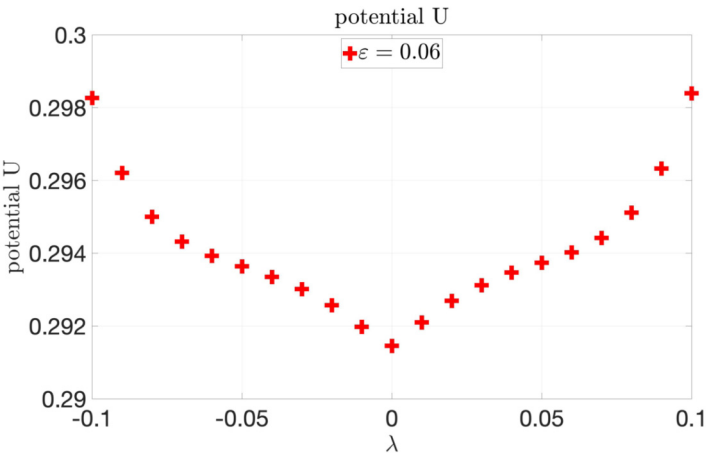

Fig. 17. Potential U for a restricted interval of λ.

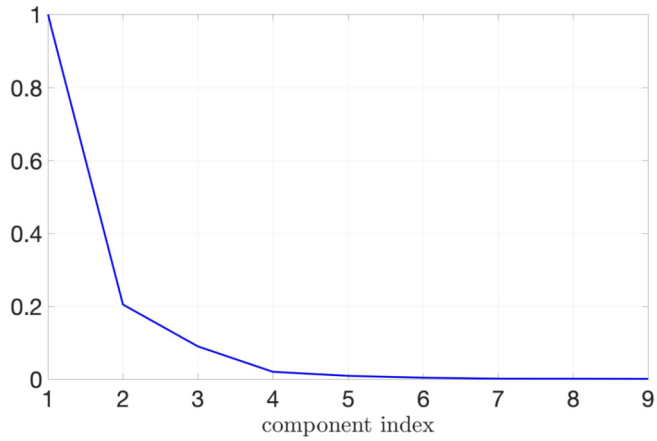

Fig. 18. Normalised kPCA eigenvalues for the sub-manifolds $\mathbf{G}(p)$.

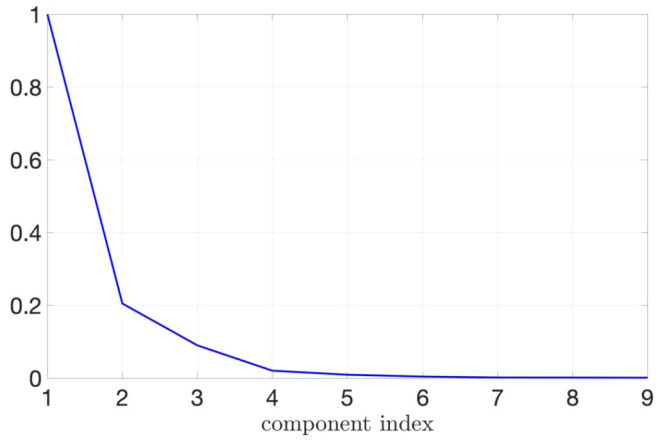

Fig. 19. Normalised kPCA eigenvalues for the sub-manifold $\mathbf{H}(p)$.

Geometric representation of the sub-manifolds $\mathbf{G}(p)$, $\mathbf{H}(p)$ *and* $\mathbf{F}(p)$ — The kPCA method is applied to reduce the dimensionality of the sub-manifolds $\mathbf{G}(p)$ and $\mathbf{H}(p)$. It happens that they are independent of p, and Figs. 18 and 19 show that their dimension is 2 with a very good accuracy. Figs. 20 and 21 show the sub-manifolds $\mathbf{G}(p)$ and $\mathbf{H}(p)$. The sub-manifold $\mathbf{F}(p)$ depends on p. Its extension $\{\mathbf{F}(p)|\ p \in [0, \bar{p}]\}$ is considered here. The kPCA applied to this extension is represented on

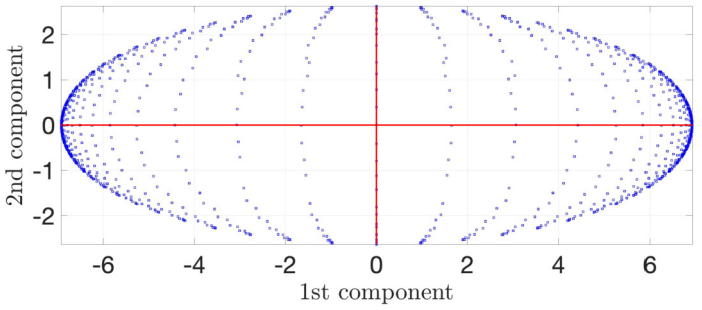

Fig. 20. Sub-manifold $\mathbf{G}(p)$.

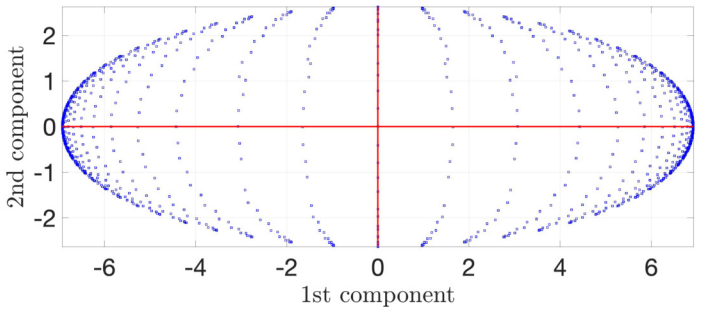

Fig. 21. Sub-manifold $\mathbf{H}(p)$.

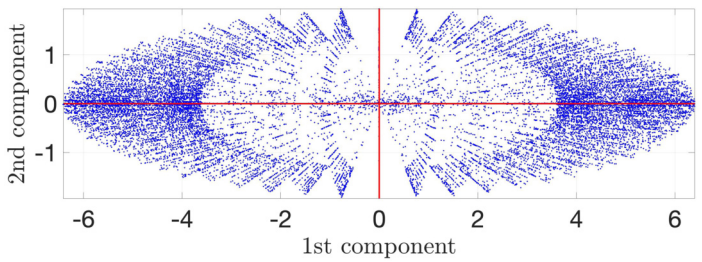

Fig. 22. Sub-manifold $\mathbf{F}(p)$.

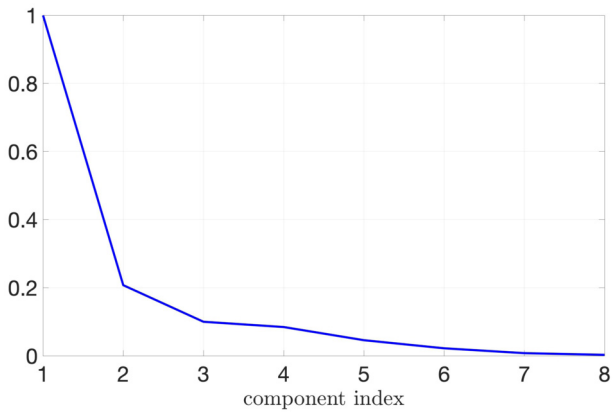

Fig. 23. Normalised kPCA eigenvalues for the sub-manifold $\mathbf{F}(p)$.

Fig. 22. Fig. 23 shows that its dimension is also 2 with a very good accuracy. Consequently, the material behaviour is completely defined by the 2D sub-manifolds $\mathbf{G}(p)$, $\mathbf{H}(p)$ and $\mathbf{F}(p)$, described in Figs. 20, 21, and 22.

4. Conclusion

Data-driven computation is a promising approach, and certainly the future of simulation in nonlinear solid mechanics. The vision developed in this paper for complex material behaviours is based on the synergy between data and Material Science. The key is validation tools that allow us to characterise the proximity of calculation points to experimental ones. Indeed, if the calculation points are too far from the experimental ones, additional tests must be carried out. It would also put into question the actual ECM, for which new hidden internal variables should also be added. This approach is accompanied by many scientific challenges, such as the extension to multiscale experimental data. Taking into account material variability and the different errors is also an important research issue.

References

[1] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, Comput. Methods Appl. Mech. Eng. 304 (2016) 81–101, https://doi.org/10.1016/j.cma.2016.02.001, arXiv:1510.04232.
[2] T. Kirchdoerfer, M. Ortiz, Data Driven Computing with noisy material data sets, Comput. Methods Appl. Mech. Eng. 326 (2017) 622–641, https://doi.org/10.1016/j.cma.2017.07.039, arXiv:1702.01574.
[3] T. Kirchdoerfer, M. Ortiz, Data-driven computing in dynamics, Int. J. Numer. Methods Eng. 113 (2018) 1697–1710, https://doi.org/10.1002/nme.5716, arXiv:1706.04061.
[4] F. Chinesta, P. Ladeveze, R. Ibanez, J.V. Aguado, E. Abisset-Chavanne, E. Cueto, Data-Driven Computational Plasticity, Procedia Engineering, vol. 207, Elsevier B.V., 2017, pp. 209–214.
[5] R. Ibañez, D. Borzacchiello, J.V. Aguado, E. Cueto, P. Ladevèze, F. Chinesta, R. Ibañez, D. Borzacchiello, J.V. Aguado, E. Abisset-chavanne, R. Ibañez, D. Borzacchiello, J. Vicente, E.A-c. Elias, C. Pierre, F. Chinesta, Data-driven non-linear elasticity: constitutive manifold construction and problem discretization, Comput. Mech. (2017).
[6] E. Lopez, D. Gonzalez, J.V. Aguado, E. Abisset-Chavanne, E. Cueto, C. Binetruy, F. Chinesta, A manifold learning approach for integrated computational materials engineering, Arch. Comput. Methods Eng. 25 (2018) 59–68, https://doi.org/10.1007/s11831-016-9172-5.
[7] D. González, F. Chinesta, E. Cueto, Thermodynamically consistent data-driven computational mechanics, Contin. Mech. Thermodyn. 31 (2019) 239–253, https://doi.org/10.1007/s00161-018-0677-z.
[8] M. Guo, J.S. Hesthaven, Data-driven reduced order modeling for time-dependent problems, Comput. Methods Appl. Mech. Eng. 345 (2019) 75–99, https://doi.org/10.1016/j.cma.2018.10.029.
[9] A. Leygue, M. Coret, J. Réthoré, L. Stainier, E. Verron, Data-based derivation of material response, Comput. Methods Appl. Mech. Eng. 331 (2018) 184–196, https://doi.org/10.1016/j.cma.2017.11.013.
[10] Z. Liu, M.A. Bessa, W.K. Liu, Self-consistent clustering analysis: an efficient multi-scale scheme for inelastic heterogeneous materials, Comput. Methods Appl. Mech. Eng. 306 (2016) 319–341, https://doi.org/10.1016/j.cma.2016.04.004.
[11] D. Versino, A. Tonda, C.A. Bronkhorst, Data driven modeling of plastic deformation, Comput. Methods Appl. Mech. Eng. 318 (2017) 981–1004, https://doi.org/10.1016/j.cma.2017.02.016.
[12] P. Ladevèze, Data Driven Structure Computation, Technical Report, LMT, 2018.
[13] P. Ladevèze, F. Chinesta, On the Concept of Constitutive Experimental Manifold in Nonlinear Mechanics, Technical Report, LMT, 2017.
[14] P. Ladevèze, A. Chouaki, Application of a posteriori error estimation for structural model updating, Inverse Probl. 15 (1999) 49–58, https://doi.org/10.1088/0266-5611/15/1/009.
[15] P. Ladevèze, On reduced models in nonlinear solid mechanics, Eur. J. Mech. A, Solids 60 (2016) 227–237, https://doi.org/10.1016/j.euromechsol.2016.08.005.
[16] P. Ladevèze, The large time increment method for the analysis of structures with non linear constitutive relation described by internal variables, C. R. Acad. Sci. Paris, Ser. IIb 309 (1989) 1095–1099.
[17] P. Ladevèze, Sur une famille d'algorithmes en mécanique des structures, C. R. Acad. Sci. Paris, Ser. IIb 300 (1985) 40–44.
[18] P. Ladevèze, J.-P. Pelle, Mastering Calculations in Linear and Nonlinear Mechanics, Springer, New York, 2005, https://doi-org.ezproxy.universite-paris-saclay.fr/10.1007/b138705.
[19] P. Ladeveze, D. Leguillon, Error estimate procedure in the finite element method and applications, SIAM J. Numer. Anal. 20 (1983) 485–509.
[20] P. Ladevèze, Nonlinear Computational Structural Mechanics—New Approaches and Non-Incremental Methods of Calculation, Springer Verlag, 1999.
[21] F. Hild, A. Bouterf, L. Chamoin, H. Leclerc, F. Mathieu, J. Neggers, F. Pled, Z. Tomičević, S. Roux, Toward 4D mechanical correlation, Adv. Model. Simul. Eng. Sci. 3 (2016), https://doi.org/10.1186/s40323-016-0070-z.