Mechanical dissimilarity of defects in welded joints via Grassmann manifold and machine learning

David Ryckelynck, Thibault Goessel, Franck Nguyen

To cite this version:

David Ryckelynck, Thibault Goessel, Franck Nguyen. Mechanical dissimilarity of defects in welded joints via Grassmann manifold and machine learning. 2020. hal-02559710v2

HAL Id: hal-02559710

https://hal.archives-ouvertes.fr/hal-02559710v2

Preprint submitted on 6 May 2020 (v2), last revised 23 Jul 2020 (v3)
Mechanical dissimilarity of defects in welded joints via Grassmann manifold and machine learning

David Ryckelynck\textsuperscript{a}, Thibault Goessel\textsuperscript{b}, Franck Nguyen\textsuperscript{a}

\textsuperscript{a}. Mines ParisTech, Centre des Matériaux, Evry | david.ryckelynck@mines-paristech.fr \textsuperscript{b}. Mines ParisTech | thibault.goessel@mines-paristech.fr

Keywords : Data encoding, Hyper-reduction, Reduced Order Model, ROM-net

Abstract

Assessing the harmfulness of defects based on images is becoming more and more common in industry. Today these defects can be insert in digital twins that aim to replicated in a mechanical model what is observed on a component. We propose a methodology for defect classification and defect labeling in view of fast prediction of their harmfulness by using hyper-reduced order models. Mechanical models studied here are non parametric. This problematic issue is circumvent by introducing a classification of defect. Then, a local hyper-reduced order model is defined in each class of defect. The proposed methodology for the classification, consists in seeing each defect by its mechanical effects more than by its morphology. Each defect is encoded as a point on a Grassmann manifold. This data encoding uses a projection-based hyper-reduction. We restrict our attention to a mechanical modeling of elastic stresses around defects. We show that the hyper-reduced equations are similar to an oblique projection of the finite element prediction. Hence we obtain an upper bound of the approximation error for displacement that depends on a Chordal distance between points on a Grassmann manifold. From the available data set of defects, we have observed that the Grassmann distance is magnifying the Chordal distance between defects. A k-medoids algorithm is used in order to create clusters of defects from a matrix of Grassmann distance between defects. A simulation-based labeling of defect is proposed. This labeling procedure relies on hyper-reduction and a related error indicator. This error indicator measures the discrepancy between the stresses predicted via hyper-reduction and equilibrated stresses. In spite of the large variety of defects in the data set, we show accurate predictions of stresses for most of defects in the test set, while training the clustering and the hyper-reduced order model on a different training set. These accurate results are obtained by combining two contributions for a reduced basis for displacements : macroscopic modes related to a defect-free mechanical problem and elastic fluctuation modes specific to each class of defect.

In this paper, we pay particular attention to the compromise to be made between the accuracy of the numerical approximations and the memory space required to save the reduced bases.

Version française abrégée

L’évaluation de la nocivité des défauts à partir d’images est de plus en plus courante dans l’industrie. Aujourd’hui, ces défauts peuvent être insérés dans des jumeaux numériques qui visent à reproduire dans un modèle mécanique ce qui est observé sur un objet. Nous proposons une méthodologie pour la classification et l’étiquetage des défauts en vue de prédire rapidement leur nocivité en utilisant des modèles d’ordre hyperréduits. Les modèles mécaniques étudiés sont non paramétriques. Cette difficulté est contournée par l’introduction d’une classification des défauts. Ensuite, un modèle d’ordre hyperréduit local est défini dans chaque classe de défaut. La méthodologie proposée pour la classification, consiste à voir chaque défaut par ses effets mécaniques plus que par sa morphologie. Chaque défaut est encodé comme un point dans une variété de Grassmann. Ce codage des données utilise un modèle d’ordre réduit par projection. Nous
limitons notre attention à une modélisation mécanique des contraintes élastiques autour des défauts. Nous montrons qu’une limite supérieure de l’erreur d’approximation pour le déplacement, dépend de la distance Chordal entre les points d’une variété de Grassmann, si une projection de Galerkin est effectuée pour la réduction du modèle. A partir de l’ensemble des données disponibles sur les défauts, nous avons observé que la distance de Grassmann amplifie la distance Chordal entre les défauts. Un algorithme de k-médoides est utilisé pour créer des groupes de défauts à partir d’une matrice de la distance de Grassmann entre les défauts. Un étiquetage des défauts basé sur la simulation est proposé. Cette procédure d’étiquetage repose sur l’hyper-réduction et un indicateur d’erreur associé. Cet indicateur d’erreur mesure l’écart entre les contraintes prédites par hyper-réduction et des contraintes équilibrées. Malgré la grande variété de défauts dans l’ensemble de données, nous montrons des prédictions précises des contraintes pour la plupart des défauts dans l’ensemble de test, tout en entraînant la classification et l’hyper-réduction sur un ensemble d’entraînement différent. Ces résultats précis sont obtenus en combinant deux contributions dans une base réduite pour les déplacements : des modes macroscopiques liés à un problème mécanique sans défaut et des modes de fluctuation élastique spécifiques à chaque classe de défaut. Dans cet article, nous accordons une attention particulière au compromis à faire entre la précision des approximations numériques et l’espace mémoire nécessaire pour enregistrer les bases réduites.

1 Introduction

Mechanical modeling based on images is becoming increasingly important in material science and in industrial applications for the assessment of defects harmfulness. The early detection of defects in industrial processes has been studied for more than a decade now. As a matter of fact, in 2011, scientists tried to improve the efficiency of early fault detection on gear, which are critical in many machinery operations [1]. Moreover, non destructive inspection techniques are able to detect and locate voids for a wide range of materials and welding processes : resistance seam welding of aluminium, zinc and galvanised steel [2], resistance spot welding of ferritic/martensitic steels [3], electron beam welding of steel to Fe-Al alloy [4] or laser welding of stainless steels [5] and aluminium alloys [6]. In the current work, data are 2D slices of experimental 3D images of voids obtained via X-ray computed tomography by L. Lacourt [7]. In this framework, image-based meshing methods [8] enable to generate complex finite element meshes of digital images obtained by such techniques. Unfortunately, it cannot be used as a tool to assess the quality of a component in a serial production framework. The required fine meshes of defective components generally lead to prohibitive computational time as explained in [9]. It’s a pity, because finite element models are feed by a huge knowledge on the mechanical behavior for various materials. The use of this knowledge would ensure a data continuity between the component design phase and the diagnostic phase on the quality of the production of these components.

In the framework of image-based diagnosis, machine learning aims to consider the following assertion : if two defects are similar, they have equivalent harmfulness. And, accounting for this similarity should facilitate the prediction of the harmfulness of new defects, by using a data set of defects. The last sentence precise the gain we expect by using machine learning. The higher this gain the more complex the machine learning can be. In this paper, we propose a similarity criterion between defects that is not based on their morphology, except their center and principal axes. In mechanics, defects produce displacement fluctuation around them. We follow the following assertion : If two defects have similar reduced bases for displacement fluctuations they should share a common surrogate model for the prediction of harmfulness. The proposed dissimilarity criterion accounts for mechanical effects of the defects, via model reduction. The gain here is to speedup mechanical predictions by using machine learning instead of finite element predictions. In [9], we have shown that these reduced basis of fluctuations can predict all mechanical simulation outputs in a zone of interest, similarly to a finite element prediction.

Pure data driven approaches have been proposed in literature. Provided that a wide range of data is available, machine learning methods can eventually detect automatically the defects and classify them in different classes, given a prescribed criteria. For instance, machine learning is used for defect classification issues related to a freezing process in [10]. Convolutional Neural Networks (CNNs) [11], which are very
helpful in computer vision [12], are trained to detect and diagnose defects. In [13], scientists used in particular CNN to diagnose automotive damper defects. Another encouraging study published this year uses CNN to detect defects on freight trains, with a precision of around 80% [14]. An amazing property of CNN is their ability to learn features, or kernels [12], solely by using labeled data and deep learning. In our opinion, using deep learning do not necessarily impose to forget the available knowledge in mechanics of materials.

Transfer Learning, is "motivated by the fact that people can intelligently apply knowledge learned previously to solve new problems faster or with better solutions" [15]. We refer the reader to [15] for a survey on transfer learning. It has practical application in aeronautic industry [16] and for the training of CNN [17]. In mechanics of materials, projection-based model order reduction methods, enable to reuse our knowledge in materials science while using machine learning. In this paper, the proposed method aims to learn a reduced approximation space from simulation data in the framework of weak formulation for partial differential equations. Hence, physics principles and material constitutive equations are preserved in the weak formulations, and the machine learning is restricted to the construction of an approximation space. Usually, the manifold containing the solution of a partial differential equation can be accurately depicted as living in a small vector space. Therefore, linear machine learning methods, like the proper orthogonal decomposition [18] or the singular value decomposition, aim at learning this reduced vector space. This sub-space then used as the approximation space for projection-based model order reduction. In other situations, the solution of the governing PDEs lies in a manifold which cannot be covered by a single vector space without increasing its dimension and thus degrading the efficiency of projection-based model reduction. Then deep learning algorithms are useful to reduce the complexity of approximation space by using simulation data. For instance, in [19], physics-informed neural networks are proposed to solve supervised learning tasks while respecting any given laws of physics described by general nonlinear partial differential equations. These networks are no more projection-based model order reduction schemes. But such reduced schemes are found in [20] and [21], where a deep classifier recommends a local reduced order model depending on input variables having a tensor format. In [20], a CNN recommends a reduced order model related to a loading environment seen as an image of an experimental setup. In [21], a deep classifier using CNN is trained to recommend hyper-reduced order models for lifetime prediction depending on a 3D stochastic temperature field. These to contributions follow the same neural network architecture, termed ROM-net [21]. Such ROM-net are trained by using simulation data encoded as points on a Grassmann manifold. A Grassmann manifold is a set of vector spaces of given common dimension in the same ambient space. This data encoding is quite general when considering simulation data.

In this paper, we show the relevance of data encoding as points on a Grassmann manifolds for 2D images of voids in materials, for harmfulness diagnosis via hyper-reduction. This approach is consistent with the hyper-reduction scheme proposed in [9] for defective components. In this last paper, the reduced approximation space is obtained via the concatenation of a defect-free reduced basis with fluctuation modes that are specific to each defect observed in the component. In that scheme, fluctuation modes are computed separately for each defect, by using scale separation assumption and dilution condition (the volume fraction of the defect is negligible). This paper aims to show that similar defects, according to the Grassmann distance, should share the same local reduced basis that enables relevant hyper-reduced predictions (i.e. fast and accurate enough).

We restrict our attention to a mechanical modeling of elastic stresses around 2D defects. Displacement fluctuation modes are the solution of three macroscopic loadings of each defect under dilution condition. For the sake of simplicity the target problem used for defect diagnosis is similar to the computation of fluctuation modes. The reader can find in [9] more complex target mechanical problems that are solved by the hyper-reduction method used in this paper. Then, the reduced basis of fluctuation modes for displacements is the point that represent the defect in a Grassmann manifold. We show that hyper-reduced equations are equivalent to an oblique projection of the finite element prediction. Hence, we obtain an upper bound of the approximation error for displacement, that depends on a Chordal distance between points in this Grassmann manifold. From the available data set of defects, we have observed that the Grassmann distance is magnifying the Chordal distance between defects. Hence, a k-medoids algorithm is used to creating clusters of defects from a matrix of Grassmann distance between defects. The center of each cluster generated by the k-medoids algorithm is a point in the training set. They are
termed medoids. Here, the simulation data related to a medoid are specific to an image in the training set. The simulation data of each medoid are saved in a dictionary for hyper-reduction. The attribution of a medoid to a new defect requires to set a label on this defect. A simulation-based labeling of defect is proposed. This labeling procedure relies on hyper-reduction and a related error indicator. This error indicator measures the discrepancy between the stresses predicted via hyper-reduction and equilibrated stresses. The best medoid in the dictionary is the one that have the lowest error indicator, when using its data for the hyper-reduction of the target problem. In spite of the large variety of defects in the data set, we show accurate predictions of stresses for most of defect in the test set, while training the clustering and hyper-reduced order model on a different training set. The finite-element approximation-space being specific to each defect, we need to design a common ambient space for the computation of Grassmann distances. We pay particular attention to the compromise to be made between the accuracy of the numerical approximations and the memory space required to save the simulation data.

The present paper is structured as follows. Section 2 presents the data set, including notions of image-based meshing for the solution of a target problem. Section 3 presents the projection-based model reduction method and the upper bound for errors on displacement prediction. Section 4 shows the results we have about the data encoding according to the Grassmann distance. Section 5 highlights the accuracy of the stress predictions by using a clustering of defects in a training set and the defect labelling for data in a testing set. Section 6 is the conclusion of this paper.

2 Data set and target mechanical problem

2.1 2D Images and related meshes

We restrict our attention to voids inside isotropic elastic materials, and 2D elastic problems. Each defect is a void placed in a surrounding box, the volumetric center of the void is taken as the center for the box.

3D images of voids have been obtained by Laurent Lacourt [9] via X-ray computed tomography of welded joints. These 3D images have been cut in slices so that we have more 2D samples. The data set contains $n_l = 2,745$ samples of 2D images, with only one defect per image. But one defect, may involves several voids that have a high mechanical interaction. In Figure 1 we show two samples of these images.

![Figure 1](image_url) - Examples of defects (defect in blue, steel in red).

Thanks to algorithms of Image Segmentation, automatic image-based meshing is possible, even when we wish to have a refined mesh around the defect. We refer the reader to [20] for more details about the meshing procedure. Then, each image has its own finite element mesh. Figure 2 shows two examples of image-based meshing. One can notice a difference in the mesh between the two, which is due to the geometry of the defect.
2.2 Target mechanical problem

For each defect, we define a specific target problem. This target problem is a macroscopic mechanical problem including the defect. The proposed numerical method aims to predict the stress field around the defect by using a hyper-reduced order model.

The hyper-reduced order model is created by using defect-free macroscopic-modes for displacement and defect-dependent modes for stresses and displacements. We follow the protocol proposed in [9]: the macroscopic modes are computed once for all possible defects and defect-dependent modes for displacement are fluctuation modes. For 3D welding-joints, such hyper-reduced predictions are 1000 times faster than usual finite element predictions [9].

In this paper, we propose to select the defect-dependent modes in a dictionary according to an error indicator \( \eta \in \mathbb{R}^D \), where \( D \) is the number of items in the dictionary. The dictionary contains simulation data required for the construction of \( D \) hyper-reduced order models that are specific to the input defect. The target hyper-reduced prediction will be performed for each item of the dictionary, via parallel computation. The smallest component of \( \eta \) will indicate the most relevant hyper-reduced model for a given defect.

Compared to a direct simulation of the macroscopic problem including the void, the speedup is the speedup of hyper-reduction divided by \( D \), by using the same computational resources. In the sequel, the number of items in the dictionary is arbitrary setup to \( D = 5 \). For larger values of \( D \), one has to consider the deep-classifier ROM-net proposed in [21] for a faster selection of a hyper-reduced order model.

In this paper, for the sake of simplicity, the macroscopic problem is the same as the mechanical problem used to generate the simulation data related to fluctuation modes. Equations of this problem are presented in the next section. This choice is not mandatory for future applications. More complex target problems can be found in [9].

3 Fluctuation modes

The domain occupied by the elastic material, in the surrounding box around each defect, is denoted by \( \Omega^* \). We give an example in Figure 3. Formally, \( \Omega^* \) has no parameter. Its morphology totally depends on the observation of a defect through a digital image. In the following, the superscript * refers to mathematical objects that are specific to a digital image of a defect. This notation emphasizes the variability due to the input image. All mathematical object with the superscript * can be seen as the output of an implicit function depending on \( \Omega^* \). Then, the proposed fluctuations modes modeling for voids is a non-parametric modeling. In this paper, we avoid giving a morphological definition of an input space that contains all possible \( \Omega^* \).

The outer boundary of the surrounding box is denoted by \( \partial E \Omega^* \). The boundary of the void is denoted by \( \partial V \Omega^* \). The local frame follows the principal axes of the second-order moments of the volume distribu-
tion in the defect. These principal axes are denoted by $e_1$ and $e_2$. The displacement field and the Cauchy stress are denoted by $u^*$ and $\sigma^*$ respectively. The fluctuation of the displacement is denoted by $\Delta u^*$. It is defined as the counterpart of the homogeneous displacement field such that:

$$u^*(x) = E \cdot x + \Delta u^*(x) \quad \forall x \in \Omega^*, \quad \Delta u^* \in \mathcal{V}^*$$

(1)

where $E$ is a macroscopic strain tensor imposed on $\partial_E \Omega^*$. It is a given symmetric second order tensor. $E \cdot x$ is a defect-free macroscopic mode for displacements. $\mathcal{V}^*$ is an usual finite element approximation space that accounts for the defect geometry and Dirichlet boundary condition (2) on $\partial_E \Omega^*$.

$\Omega^*$ is large enough to fulfill a dilution assumption and the following boundary condition on $\Delta u^*$:

$$\Delta u^*(x) = 0, \quad \forall x \in \partial_E \Omega^*$$

(2)

The Hooke tensor for elasticity is denoted by $C$. Hence, the constitutive equation reads:

$$\sigma^* = C : \varepsilon(u^*)$$

(3)

where $\varepsilon(u^*)$ is the deformation tensor, i.e. the symmetric part of the displacement gradient. The weak form of the elastic equilibrium equation reads: find $\Delta u^* \in \mathcal{V}^*$ such that,

$$\Delta u^*(x) = \sum_{i=1}^{N^*} \varphi_i^*(x) q_i^*$$

(4)

$$\int_{\Omega^*} \varepsilon(\varphi_i^*): C : (E + \varepsilon(\Delta u^*)) \, d\Omega = 0 \quad \forall i \in \{1, \ldots, N^*\}$$

(5)

where $(\varphi_i^*)_{i=1}^{N^*}$ are the finite element shape functions that span $\mathcal{V}^*$, such that $\varphi_i^* = 0$ on $\partial_E \Omega^*$. $E$ being a symmetric second order tensor, it has only three components. Therefore, the elastic equations being linear, it exists an ideal reduced basis containing three fluctuation modes denoted by $(\psi_k^*)_{k=1}^{3}$ such that:

$$\Delta u^* = \sum_{k=1}^{N} \psi_k^* \tau_k^*, \quad \psi_k^* \in \mathcal{V}^*, \quad k = 1, \ldots, N$$

(6)

where $\tau^* \in \mathbb{R}^3$ is a vector of exact reduced coordinates. The fluctuations modes being vectors of the approximation space, one can introduce the reduction matrix $V^*$, that contains the finite element coordinates of the fluctuation modes:

$$\psi_k^* = \sum_{i=1}^{N^*} \varphi_i^*(x) V_{ik}^*, \quad k = 1, 2, 3 \forall x \in \Omega^*$$

(7)

In the sequel, approximate reduced bases for fluctuation modes are denoted by $V$. $V$ may be not totally specific to $\Omega^*$. The approximate continuous modes are:

$$\psi_k = \sum_{i=1}^{N^*} \varphi_i^*(x) V_{ik}, \quad k = 1, 2, 3 \forall x \in \Omega^*$$

(8)

The two reduced bases span two vector spaces of same dimension $N$ in the same ambient space $\mathbb{R}^{N^*}$, $N < N^*$. Each of these vector space is a point in Grassmann manifold, denoted by $Gr(N, N^*)$.

The hyper-reduction method [22] aims at computing reduced coordinates $\gamma$, by projecting the equilibrium equation on $V$, via a restriction of the domain $\Omega^*$ to a reduced integration domain (RID) $\Omega_R^*$. By following the empirical interpolation method [23], interpolation points are computed for column vectors in $V$. We choose the RID $\Omega_R^*$ such that it contains the interpolation points related to both the reduced basis for displacement and a reduced basis for stresses. We give more details about the construction of $\Omega_R^*$ in Appendix 6. We define, a set of test reduced functions denoted by $\psi_{R_j}$:

$$\mathcal{F} = \left\{ i \in \{1, \ldots, N^*\}, \int_{\Omega^* \setminus \Omega_{R^*}} (\varphi_i^*)^2 \, d\Omega = 0 \right\}$$

(9)

$$\psi_{R_j}(x) = \sum_{i \in \mathcal{F}} \varphi_i^*(x) V_{ij}, \quad \forall x \in \Omega^*, \quad j = 1, 2, 3$$

(10)
As explained in [22], these test functions are null on the interface between $\Omega^*_R$ and the counterpart of the domain, as if Dirichlet boundary conditions were imposed. On this interface, the displacement follows the shape of the modes $\psi_k$. The hyper-reduction method gives access to reduced coordinates $\gamma$ that fulfill the following balance equations:

$$\nabla \mathbf{u} = \sum_{k=1}^{N} \psi_k \gamma_k$$  \hspace{1cm} (11)

$$\int_{\Omega^*_R} \varepsilon(\psi_{Rj}) : \mathbf{C} : (\mathbf{E} + \varepsilon(\Delta \mathbf{u})) \, d\Omega = 0 \quad \forall j = 1, \ldots, N$$  \hspace{1cm} (12)

The matrix form of the hyper-reduced balance equations reads: find $\gamma \in \mathbb{R}^N$ such that,

$$\Delta \mathbf{u} = \sum_{i=1}^{N^*} \varphi^*_i(\mathbf{x}) q^*_i$$  \hspace{1cm} (13)

$$\mathbf{q}^* = \mathbf{V} \gamma$$  \hspace{1cm} (14)

$$\mathbf{K}^* \gamma = -\mathbf{V}[\mathbf{F},:]^T \mathbf{F}^*[\mathbf{F}]$$  \hspace{1cm} (15)

$$K^*_{ij} = \int_{\Omega^*_R} \varepsilon(\varphi^*_i) : \mathbf{C} : \varepsilon(\varphi^*_j) \, d\Omega, \; i, j = 1, \ldots, N^*$$  \hspace{1cm} (16)

$$F^*_i = -\int_{\Omega^*_R} \varepsilon(\varphi^*_i) : \mathbf{C} : \mathbf{E} \, d\Omega; \; i = 1, \ldots, N^*$$  \hspace{1cm} (17)

where $\mathbf{V}[\mathbf{F},:]$ denotes a row restriction of matrix $\mathbf{V}$ to indices in $\mathbf{F}$. We assume that the matrix $\mathbf{K}^*$ is full rank. This assumption is always checked during numerical solution of hyper-reduced equations. Rank deficiency may appear when the RID construction do not account for the contribution of a reduced basis dedicated to stresses.

In this paper, this hyper-reduced prediction is supplemented by the following equilibrium step over the RID: find $\delta \mathbf{u}^*$ such that,

$$\delta \mathbf{u}^*(\mathbf{x}) = \sum_{i \in \tilde{\mathbf{F}}} \varphi^*_i(\mathbf{x}) \delta q^*_i$$  \hspace{1cm} (19)

$$\int_{\Omega^*_R} \varepsilon(\varphi^*_i) : \mathbf{C} : (\mathbf{E} + \varepsilon(\Delta \mathbf{u} + \delta \mathbf{u}^*)) \, d\Omega = 0 \quad \forall i \in \tilde{\mathbf{F}}$$  \hspace{1cm} (20)

where $\tilde{\mathbf{F}}$ is the set of all degrees of freedom in $\Omega^*_R$ excepted those belonging to elements connected to the interface between $\Omega^*_R$ and its counter part. This correction step has been proposed in [24] for the evaluation of contact forces. This is a local correction step. The solution $\mathbf{u}^{eq} = \mathbf{E} x + \Delta \mathbf{u} + \delta \mathbf{u}^*$ is an hybrid solution that weakly couples hyper-reduction and a finite element approximation over the RID. We refer the reader to [24, 25, 26] for more details about hybrid hyper-reduction schemes. The equilibrated stress is quite accurate compared to the stress computed via hyper-reduction solely. The hyper-reduced stress prediction is denoted by $\sigma^{HR}$. The equilibrated stress is denoted by $\sigma^{eq}$, $\delta \sigma^*$ is the correction term for stress predictions:

$$\sigma^{HR} = \mathbf{C} : (\mathbf{E} + \varepsilon(\Delta \mathbf{u}))$$  \hspace{1cm} (21)

$$\delta \sigma^* = \mathbf{C} : \varepsilon(\delta \mathbf{u}^*)$$  \hspace{1cm} (22)

$$\sigma^{eq} = \sigma^{HR} + \delta \sigma^*$$  \hspace{1cm} (23)

**Property 1**: If $\mathbf{K}^{HR}$ is full rank, then the hyper-reduced balance equations are equivalent to an oblique projection of the finite element prediction:

$$\mathbf{q}^{HR} = \mathbf{V}(\Pi^T \mathbf{V})^{-1} \Pi^T \mathbf{q}^*$$  \hspace{1cm} (24)
with \( \mathbf{\Pi} = \mathbf{K}^* \mathbf{F} \mathbf{V} \mathbf{F} \mathbf{V}^* \). Hence the hyper-reduced prediction is a minimizer for \( f(\beta) : \)
\[
\beta \in \mathbb{R}^N, \quad f(\beta) = \| \mathbf{\Pi}^T (\mathbf{V} \beta - q^*) \|_2^2
\]  
(25)

Here \( \mathbf{\Pi} \) is a projector for elastic stresses in \( \Omega^*_R \) according to the reduced test-functions :
\[
\sum_{i=1}^{N^*} \Pi_{ik} (\mathbf{V} \gamma - q^*)_i = \int_{\Omega^*_R} \mathbf{\varepsilon}(\psi_R) : (\sigma^{HR} - \sigma^*) \ d\Omega
\]  
(26)

where \( \sigma^* \) is the finite-element stress prediction.

The proof is straightforward. \( \mathbf{K}^{HR} = \mathbf{\Pi}^T \mathbf{V} \). The Jacobian matrix reads \( \mathbf{J} = \mathbf{V}^T \mathbf{\Pi} \mathbf{\Pi}^T \mathbf{V} = (\mathbf{K}^{HR})^T \mathbf{K}^{HR} \). If \( \mathbf{K}^{HR} \) is full rank, then \( \mathbf{J}^{-1} = (\mathbf{K}^{HR})^{-1} (\mathbf{K}^{HR})^{-T} \). Then, both the minimization problem and the hyper-reduced equation have a unique solution. The solution of the minimization problem is :

\[
q^f = \mathbf{V} (\mathbf{J})^{-1} \mathbf{\Pi}^T \mathbf{\Pi}^T q^*
\]  
(27)

\[
= \mathbf{V} (\mathbf{K}^{HR})^{-1} \mathbf{\Pi}^T q^*
\]  
(28)

Since \( \mathbf{\Pi}^T q^* = \mathbf{V} \mathbf{F} \mathbf{F}^* \mathbf{V}^* \mathbf{F} \mathbf{V} \mathbf{F}^* \mathbf{V}^* \mathbf{F}^* \mathbf{V}^* \mathbf{F}^* \mathbf{V}^* / \mathbf{F}^* \mathbf{V}^* \mathbf{F}^* \mathbf{V}^* \mathbf{F}^* \mathbf{V}^* = \mathbf{q}^{HR} \).

Let’s introduce the three canonical macroscopic strains, \( \mathbf{E}^{(1)}, \mathbf{E}^{(2)}, \mathbf{E}^{(3)} : \)
\[
\mathbf{E}^{(1)} = \mathbf{\varepsilon}(x_1 e_1)
\]  
(29)

\[
\mathbf{E}^{(2)} = \mathbf{\varepsilon}(x_2 e_2)
\]  
(30)

\[
\mathbf{E}^{(3)} = \mathbf{\varepsilon}(\frac{x_1 e_2 + x_2 e_1}{2})
\]  
(31)

In the sequel, the three finite element solutions obtained for each canonic macroscopic strains \( \mathbf{E}^{(1)}, \mathbf{E}^{(2)}, \mathbf{E}^{(3)} \), are saved in a matrix \( \mathbf{Q}^* \) such that the three displacement fields read :
\[
\mathbf{u}^{(j)}(\mathbf{x}) = \mathbf{E}^{(j)} . \mathbf{x} + \Delta \mathbf{u}^{(j)}(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega^*, \ j = 1, 2, 3
\]  
(32)

\[
\Delta \mathbf{u}^{(j)}(\mathbf{x}) = \sum_{i=1}^{N^*} \varphi_i^*(\mathbf{x}) Q_{ij}^*, \quad \forall \mathbf{x} \in \Omega^*, \ j = 1, 2, 3
\]  
(33)

All mechanical simulations are run with the Z-Set software suite. One can find more information on this website : http://www.zset-software.com/. Then, the reduced basis \( \mathbf{V}^* \) is obtained by using a truncated singular value decomposition (SVD) of \( \mathbf{Q}^* : \)
\[
\mathbf{Q}^* = \mathbf{V}^* \mathbf{S}^* \mathbf{W}^* \mathbf{T}, \quad \mathbf{V}^* \in \mathbb{R}^{N^* \times 3}, \ \mathbf{V}^{* T} \mathbf{V}^* = \mathbf{I}
\]  
(34)

where \( \mathbf{I} \) is the \( 3 \times 3 \) identity matrix. Because of linearity of the elastic problem above, \( \mathbf{C} \) is proportional to the Young modulus, hence \( \mathbf{V}^* \) does not depend on the value of the Young modulus, thanks to the normalization of the modes. Regarding the Poisson coefficient, for the sake of simplicity, we restrict our attention to isotropic materials having a Poisson coefficient equal to 0.3. An example of displacement fluctuations is shown in Figure 3.
Similarly to Céa’s lemma, but in finite dimension, it exists an upper bound for the approximation error observed through the projector $\Pi$. The best projection of the exact solution in the approximation space via the 2-norm is denoted by $\gamma_P$:

$$\gamma_P = \text{argmin}_{q = \mathcal{V}, g} ||q^* - q||_2$$

with $q^* = \mathcal{V}^* \gamma^*$ and $K^* q^* = F^*$.

**Property 2**: It exists a stability coefficient $c^*$, that does not depend on $F^*$ (the loading condition), such that the partial approximation error has the following upper bound:

$$||\Pi^T (q^* - V \gamma)||_2 \leq c^* ||q^* - V \gamma_P||_2$$

where $||\Pi^T (q^* - V \gamma)||_2$ is a partial approximation error. Hence, the smaller the Euclidian distance between the subspace span by $V$ and the finite element prediction, the better the prediction of the stress in $\Omega^*$ according to the projector $\Pi$. So, it is relevant to train $V$ by using a training set of images.

The proof of the previous property is straightforward in finite dimension. Let’s denote by $\alpha^*$ an upper bound of the highest singular value of $\Pi$.

Moreover, $\gamma$ being a minimizer for $f(\cdot)$, we obtain:

$$||\Pi^T (q^* - q)||_2 \leq (\alpha^*)^2 ||q^* - q||_2 \quad \forall q$$

Property 3: If, $V^* \in \mathbb{R}^{N^* \times N}$ and $V \in \mathbb{R}^{N^* \times N}$ are two orthonormal matrices of the same ambient space $\mathbb{R}^{N^*}$, then they span subspaces that belong to the same Grassmann manifold. The partial approximation error has a bound depending on the Chordal distance, denoted by $d^{Ch}(V^*, V)$, between the subspaces span by $V^*$ and $V$ respectively:

$$||\Pi^T (q^* - V \gamma)||_2 \leq \alpha^* d^{Ch}(V^*, V) \sqrt{N} ||\gamma^*||_2$$

where the Chordal distance uses the principal angles $\theta \in \mathbb{R}^N$ computed via the following singular value decomposition:

$$V^T V^* = U \cos(\theta) U^*, \quad U^T U = U^* U = I$$

$$d^{Ch}(V^*, V) = ||\sin(\theta)||_F$$

**Figure 3** – On the left, a experimental defect and a zoom on it. On the right, the horizontal component of $\Delta u^{(j)}$, for 2 traction modes related to $E^{(1)}$ and $E^{(2)}$, and one shearing mode related to $E^{(3)}$. 
Thus:

\[ \gamma = U^* \gamma^* - U \gamma \]

Hence:

\[ \|q^* - V \gamma_p\|_2^2 = \gamma^T (V^* - V V^T) \gamma^* = \gamma^T (I - V V^T) \gamma^* \]

For all matrices \( A \in \mathbb{R}^{n \times m} \) and \( B \in \mathbb{R}^{m \times n} \) the following property holds:

\[ \|AB\|_F \leq \|A\|_F \|B\|_F \]

and for \( a \in \mathbb{R}^n \) : \( \|a\|_F = \|a\|_2 \)

Thus:

\[ \|q^* - V \gamma_p\|_2^2 \leq \|\sin(\theta)\|_F \|\gamma^*\|_2^2 \leq \|\sin(\theta)\|_F \|\gamma^*\|_2^2 \]

The Property 3 is a convergence property for hyper-reduced predictions. If \( \theta = 0 \), then \( V U = V^* U^* \)

Hence the hyper-reduced prediction is the same as the finite element prediction. But it uses less floating point operations.

As explained in [27, 28], Grassmann manifolds are the adequate concept when considering interpolation of reduced order models. Here, we show that it is also relevant for data encoding and the classification of voids via their mechanical effects.

Several geodesic distances are available between points in Grassmann manifolds. The above Chordal distance is one of them. The Grassmann distance is another one. Let’s consider two subspaces span by the column vector of orthonormal matrices \( V^* \in \mathbb{R}^{N^* \times N} \) and \( V \in \mathbb{R}^{N \times N} \). The Grassmann distance between these two subspaces is denoted by \( d_{GR}(V^*, V) \). It reads:

\[ d_{GR}(V^*, V) = \|\theta\|_2 \]

where \( \theta \in \mathbb{R}^N \) is the vector of principal angles between \( V^* \) and \( V \), computed via a truncated SVD in Equation (39). In the following numerical experiment, the Grassmann distance magnifies the distance between subspaces compared to the Chordal distance, similarly to the one dimensional equation \( \theta \geq \sin(\theta) \) for \( \theta \in [0, \pi/2] \).

In the sequel, we assume that there is a correlation between the approximation error \( \sigma^{HR} - \sigma^* \), in \( \Omega_{GR} \), and the upper bound in Equation (38). This correlation is evaluated through a numerical experiment in Section 5.
Depending on the variety of defect morphology in the data set, the manifold occupied by all the possible fluctuations may often not be approximated by a small vector subspace. In practice, a hyper-reduced order model gives faster predictions if the dimension of the reduced space is smaller than \(N\), because \(K^*\) is sparse and \(K^{HR}\) is full. Therefore, we propose to develop a dictionary for local reduced bases related to a classification of the defects in the data set. This classification uses the Grassmann distance as a dissimilarity measure between defects. Unfortunately, the Grassmann manifold \(Gr(N,N^*)\) is image-specific. Therefore we introduce an encoding mesh in the sequel and a common Grassmann manifold. Moreover, the evaluation of the Grassmann distance and the Chordal distance require the finite element solution of the target problem. Such distance computations are not relevant for fast defect diagnosis. Then, an error indicator is proposed to select the local reduced basis that forecasts the best mechanical predictions via hyper-reduction.

4 Encoding data as points on a common Grassmann manifold

We propose to introduce an encoding mesh in order to define a common Grassmann manifold for all the defects in the data set. It turns out that the evaluation of dissimilarities between defects is regularized by using a continuous manifold although there is no continuous parameter setting for defect morphology. The encoding mesh has \(N\) degrees of freedom. The image-specific predictions \(\Delta u^*(j), j = 1, 2, 3\), are transferred to this encoding mesh prior the computation of a relevant reduced basis \(\tilde{V}^* \in \mathbb{R}^{N^* \times 3}\) for data clustering. In practice, the transfer procedure gives a matrix \(\tilde{Q}^*\) such that:

\[
\Delta u^*(j) = \sum_{i=1}^{N} \varphi_i^{*} \tilde{Q}_{ij}^{*} + \tilde{r}^*(j), \quad j = 1, 2, 3
\]

where \(\tilde{r}^*(j)\) is an error due to the transfer from the original mesh to the encoding mesh, \((\varphi_i^{*})_{i=1}^{N^*}\) are the shape functions of the encoding mesh. They are no more component-specific. Hence, a more convenient reduced basis for the computation of Grassmann distances is \(\tilde{V}^* \in Gr(3,N)\):

\[
\tilde{Q}^* = \tilde{V}^* \tilde{S}^* \tilde{W}^* T, \quad \tilde{V}^* \in \mathbb{R}^{N^* \times 3}, \tilde{V}^* T \tilde{V}^* = I
\]

In order to facilitate this transfer, the prediction \(\Delta u^*(j)\) is extended inside the defect, by introducing a very small Young modulus in the void \((10^{-3} E)\). All reduced bases \(\tilde{V}^*\) span the subspace of dimension 3 in the same ambient space \(\mathbb{R}^N\). So, each image is encoded as a point in the Grassmann manifold \(Gr(3,N)\).

The potential similarity between defects is evaluated via the Grassmann distance \(d_{Gr}\), between \(A \in \mathbb{R}^{N^* \times 3}\) and \(B \in \mathbb{R}^{N^* \times 3}\):

\[
\tilde{d}_{Gr}(A,B) = \frac{||\theta||}{\pi/2}, \quad \text{with} \quad A^T A = B^T B = I \quad (58)
\]

\[
A^T B = U^A \cos(\theta) U^{B T} \quad (59)
\]

This data encoding via the reduced mechanical response of input data presents the huge advantage to be universal and may be applied in other fields. Thomas Daniel in his paper released in 2020 [21] uses that kind of data encoding for stochastic temperature fields. Whether we consider the displacement field or the temperature field, the methodology remains the same. Each data in the data set is supplemented by a reduced basis \(\tilde{V}^*(i)\), here \(i\) is the defect index.

The encoding mesh must have a number of degrees of freedom \(N\) as small as possible to lower the memory usage in a storage system, but it must be fine enough:

- to lower the norm of transfer errors \(\tilde{r}^*(j), j = 1, 2, 3\) from the original mesh to the encoding mesh,
- to detect dissimilarities between defects, up to a given accuracy.

The transfer error is computed via the transfer of \(\Delta u^*(j) - \tilde{r}^*(j)\) from the encoding mesh to the original mesh. Then, we obtain:

\[
\Delta u^*(j) - \tilde{r}^*(j) = \sum_{i=1}^{N^*} \varphi_i^{*} \tilde{Q}_{ij}^{*} + r^*(j) \quad (60)
\]
where $r^*(j)$ is the additional error due to this new transfer. It turns out that the Frobenius norm $\|Q^* - \tilde{Q}^*\|_F$ evaluates the error $\tilde{r}^*(j) + r^*(j)$.

In order to reduce as much as possible the loss of information due to the mesh transfer, the shape of the encoding mesh must look like the majority of the defects. In our pool of data, many defects have a characteristic size of 0.1 mm. We thus decided to design a encoding mesh with a 0.15 mm refined square at the middle. Figure 4 shows the chosen encoding mesh, with here 4 elements on each half-side of the middle square.

![Figure 4 – Encoding mesh with 4 elements on each half-side of the middle square. From left to right, the surrounding box, a zoom, the same zoom on a defect.](image)

To highlight the influence of the mesh’s refining on the transfer error, we varied the number of elements on each half-side of the middle square. The projection error is calculated on two different defects: one arbitrary anisotropic defect and one isotropic defect, both with the same characteristic size of 0.1 mm. Figure 1 shows these defects. The image-based meshes have 837 nodes and 242 nodes respectively for the isotropic and the anisotropic defect.

Figure 5 presents the projection error in function of the refinement of the encoding mesh. For each defect, one projection has been made from the initial mesh to the encoding one, and then reciprocally from the encoding mesh to the initial one. The Grassmann distance is then calculated on the reduced basis from the original mesh and from the re-projected one.
As expected, the more refined the encoding mesh is, the less the projection error is. Its influence is consequent as we can lose 20% of information even for a defect whose mesh is close to the encoding one (isotropic defect in green), and more than 35% for an anisotropic defect. Moreover, the loss of information is logically more important for the anisotropic defect as the original mesh is very different from the encoding mesh. Nevertheless, for both defects, when the data encoding is enough refined (more than 4 elements on each half-side of the middle square), the loss of information is negligible and the Grassmann distance is less than $0.1$.

Now that the encoding mesh and the projection error have been presented, we can study the efficiency of that data encoding to classify defects. One pilot circular defect has been created, with the same method as the one for the encoding mesh. Only a circle with a radius of 0.1 mm was introduced inside the middle square. Figure 6 shows the node distribution for mesh of the circular defect.

Figure 6 – Node distribution for the mesh of the pilot circular defect.

Figure 7 presents the convergence of the Grassmann distance for the isotropic and anisotropic defects with respect to the pilot circular defect. The Grassmann distances are calculated for various encoding mesh.
The Grassmann distance and the encoding mesh successfully distinguish the two defects and we observe a convergence of the distance when the mesh exceeds 4 elements on each half-side of the middle square. A difference of 20% in the normalized Grassmann distance between the two defects is consistent as the pilot circular defect and the anisotropic one are mechanically very different. The gap may be enough to distinguish them during the clustering procedure.

If the encoding mesh is too coarse, no dissimilarity can be found between the defects and the dictionary would be inaccurate. The encoding mesh must be accurate enough to classify efficiently the defects and mitigate the projection error. Nevertheless, if the encoding mesh is too refined, the dictionary construction is consuming too much memory and computational time. Therefore, we suggest to keep the encoding mesh with 609 nodes (4 elements on each half-side).

5 Data clustering

5.1 Clustering of simulation data in the training set

All available images in the dataset have been converted into finite element predictions of fluctuation modes $\hat{V}^{s(i)} (i = 1, \ldots, n_I)$, transferred on the encoding mesh. Two dissimilarity matrices have been computed, one related to the Grassmann distances and the other one related to the Chordal distances. They are denoted by $D^{Gr}$ and $D^{Ch}$ respectively:

$$D^{Gr}_{ij} = d^{Gr}(\hat{V}^{s(i)}, \hat{V}^{s(j)}), \quad i, j \in \{1, \ldots, n_I\} \quad (61)$$

$$D^{Ch}_{ij} = d^{Ch}(\hat{V}^{s(i)}, \hat{V}^{s(j)}), \quad i, j \in \{1, \ldots, n_I\} \quad (62)$$

These matrices are obviously symmetric. For all couples of simulation data, we observe the following relationship: $D^{Ch}_{ij} \approx \sin(D^{Gr}_{ij})$. In practice the three principal angles are very close to each other, for each singular value decomposition. Therefore $D^{Gr}_{ij} \geq D^{Ch}_{ij}$ (as an empirical extension of $\theta \geq \sin(\theta)$ for $\theta \in [0, \pi/2]$). The Grassmann distance magnifies the dissimilarities between the defect in the data set. For this reason, the clustering procedure is performed with $D^{Gr}$.

From these data, we have extracted two subsets: a training set of simulation data and a testing set of simulation data. The test set contains the simulation data for 20% of images.

In this section, we restrict our attention to simulation data in the training set. A k-medoids clustering [29] has been used to classify the simulation data according to their Grassmann distance. The k-medoids algorithm proposed in [29] can be summarized as follows:

- Initialization step: select $K$ rows in $D^{Gr}$ as indices of initial medoids ($m_1, \ldots, m_K$).
- Repeat the two following steps until convergence:
  - Data assignment step: assign each point of the data set to the cluster corresponding to its closest medoid:
    $$L^G_{i} = \underset{m \in \{m_1, \ldots, m_K\}}{\arg\min}(D^{Gr}[m, i]) \quad i \in \text{train set} \quad (63)$$
— Medoid update step: for each cluster, update the medoid by finding the point which minimizes the sum of distances to all the points in the cluster.

\[ C_k = \{ i \in \text{train set} \mid L_{Gr}^i = k \}, \quad k = 1, \ldots, K \]  \hspace{2cm} (64)

\[ m_k = \arg \min_j \left( \sum_{i \in C_k} D_{Gr}^{[j,i]} \right) \quad i \in \text{train set} \]  \hspace{1cm} (65)

We arbitrary set the number of cluster to \( K = 5 \). The set \( C_k \) contains the indices of defects in the cluster number \( k \). The size of the clusters is: \( \text{card}(C_1) = 554 \), \( \text{card}(C_2) = 200 \), \( \text{card}(C_3) = 202 \), \( \text{card}(C_4) = 901 \), \( \text{card}(C_5) = 339 \).

The defects located in the medoids are shown in Figure 8. These defects have different anisotropy. \( m_4 \) is the most isotropic defect. \( m_2 \) and \( m_3 \) are the most anisotropic defects. The defects are rather symmetric. In future work this symmetry should be removed from the data set. In the target application the orientation of the defect is taken into account during the transfer of the fluctuation modes on the target mesh.

![Figure 8 — Defects located at the medoids.](image)

Defects \( m_2 \) and \( m_3 \) contain two voids are so close to each other, that fluctuation modes account for the local interactions between these voids.

5.2 HROM predictions for the training set

Hyper-reduction of the fluctuation mode computations set are used to choosing the meta-parameters related to the construction of the reduced integration domain \( \Omega_R \), with the data in the training set.

For each medoid \( m^{(k)} \) and for each target image number \( i \), we define an image-specific hyper-reduced order model denoted by \( HR(k,i) \), without using the encoding mesh here. The HROM corresponds to a couple of reduced basis for displacement fluctuations and a reduced basis for stress around defects, \( (V^{(k,i)}, V^{\sigma,(k,i)})) \). These reduced bases are computed after the transfer of simulation data from the mesh of the medoid onto the mesh of the target problem which is defect-dependent. The macroscopic modes related to a defect-free mechanical problem is denoted \( V^{\text{macro}} \in \mathbb{R}^{N \times 3} \). The column number \( j \) in \( V^{\text{macro}} \) is related to the macroscopic displacement \( E^{(j)} x \). The complete reduced basis for displacement reads:

\[ V^{(k,i)} = [V^{\text{macro}}, V^{(k,i)}] \in \mathbb{R}^{N \times 6} \]

The hyper-reduced equilibrium equation is Equation (12) with empirical modes obtained by substituting \( V^{(k,i)} \) for \( V \) in Equation (8). A zone of interest is designed automatically around each defect. It contains 4
layers of elements from the border of the defect \( \partial \Omega^* \). The construction of the RID follows the procedure explained in Appendix 1. It involves the zone of interest.

For each image in the training set, we build a HROM \( HR(k,i) \) for each medoids \( (k = 1, \ldots, 5) \) and we compute the HROM prediction for the three macroscopic strains \( E^{(j)} \) \( (j = 1, \ldots, 3) \). The predicted stresses on the RID are denoted by \( \sigma^{HR(k,i,j)} \). The finite element prediction of this stress is denoted by \( \sigma^{\star(i,j)} \). The exact error on the stress prediction in \( \Omega^*_R \), \( i \in \{1, \ldots, n_I \} \), reads:

\[
e^{HR(k,i)} = 100 \sqrt{\frac{\sum_{j=1}^{3} \| \sigma^{\star(i,j)} - \sigma^{HR(k,i,j)} \|_2^2}{\sum_{j=1}^{3} \| \sigma^{\star(i,j)} \|_2^2}}
\]

where \( \| \cdot \|_\Omega_R = \int_{\Omega_R} \cdot : \cdot d\Omega \) for stress tensors. The Figure 9 reports the correlation between the exact error and square root of the Chordal distance to the medoid. This plot contains \( 5 \times n_I \) points. The coordinates of these points are \(( \sqrt{D_{Ch}}^{Ch_m,k}, e^{HR(k,i)} )\). The average square root of Chordal distance is 0.7. The average error is 29%.

\[\text{Figure 9 – Correlation plot between } e^{HR(k,i)} \text{ and } \sqrt{d_{Ch}} \text{ for all points in the training set and for all HROM. The red dot is the average point } (x = 0.7, y = 29\%)\]

The correlation between \( e^{HR} \) and \( \sqrt{d_{Ch}} \) is not perfect, but it is sufficient here for the clustering of simulation data. In some situations, \( e^{HR} \) and \( \sqrt{d_{Ch}} \) may not be correlated. A good correlation requires that the target mechanical problem activates all the modes in \( V^{\star(i)} \) and \( V^{(k,i)} \), in the sense that the related reduced coordinates have no null component. It is the case here. If one of the reduced component is null, the upper-bound can be obviously simplified without using \( d_{Ch}(V^{\star(i)}, V^{(k,i)}) \).

For each cluster of data we have reported on Figure 10 the histogram of the density distribution of \( e^{HR(k,i)} \), for \( i \in C_k \), inside each cluster separately.
The clustering results are interesting because, most of errors are lower to 29% (the average error on all the data set). But many points have an error $e^{HR}$ larger than 20%. The clustering may not perform well for cluster where there is a lack of similar data for this cluster. Here cluster $C_2$ and $C_3$ have a low number of points with an error lower to 20%.

Some cluster could be rejected for hyper-reduction, in a sense that a full finite element prediction may be preferable for the defects in these clusters. In the sequel we don’t reject any cluster, and we include simulation data related to all the medoids in the dictionary for hyper-reduction.

5.3 Results on testing set

In this section, we restrict our attention to simulation data in the testing set. This testing set aims to evaluate the full modeling procedure for fast prediction of stresses around defects. The average speed up for the solution of linear system is $0.03 \times 4 = 0.12$, for the hyper-reduced equilibrium equation. The average speed up for the computation of the stresses, in this equation, is $0.06 \times 2 = 0.12$. Speedups around 1000 are obtained for similar 3D problems [9].

An error indicator has been developed for the selection of the medoid in the dictionary, that is
expected to give the best stress prediction via hyper-reduction. This error indicator is similar to the error \( e^{HR(k,i)} \) where the exact stress \( \sigma^* \) is replaced by the equilibrated stress \( \sigma^{eq} \). The error indicator reads:

\[
\eta^{(k,i)} = 100 \sqrt{\frac{\sum_{j=1}^{3} ||\delta \sigma^{(k,i,j)}||^2_{\Omega_R}}{\sum_{j=1}^{3} ||\sigma^{eq(i,j)}||^2_{\Omega_R}}} \quad (67)
\]

where \( \delta \sigma^{(k,i,j)} \) is the stress correction (i.e. Equation (22)) computed by the equilibrium step related to \( HR(k,i) \).

For each defect in the test set, we select a medoid \( k^{*(i)} \) for stress prediction such that:

\[
k^{*(i)} = \text{argmin}_k \eta^{(k,i)} \quad (68)
\]

The perfect medoid selection aims to lower the error prediction on stresses:

\[
k^{*(i)} = \text{argmin}_k e^{HR(k,i)} \quad (69)
\]

Then we get an automatic labeling of the defects in the test set. The estimated labels are denoted by \( L^\eta_i : L^\eta_i = \text{argmin}_k \eta^{(k,i)} \). The perfect labels are denoted by \( L_i : L_i = \text{argmin}_k e^{HR(k,i)} \). The Figure 11, reports the correlation between \( e^{HR(k,i)} \) and the error indicator \( \eta^{(k,i)} \).

**Figure 11** – Correlation plot between \( e^{HR(k,i)} \) and \( \eta^{(k,i)} \) for points in the testing set. The dot color is related to the cluster index \( k \).

There is a strong correlation between \( e^{HR(k,i)} \) and \( \eta^{(k,i)} \) for low error: \( e^{HR(k,i)} \approx \eta^{(k,i)} \) for \( e^{HR(k,i)} < 40\% \). This correlation does not depend on defect cluster. Then, this indicator helps to find the medoid as if we have the exact finite element prediction.

The better \( \sigma^{eq} \) approximates \( \sigma^* \) the more accurate is the error indicator. The error on equilibrated stresses reads:

\[
e^{eq(k,i)} = 100 \sqrt{\frac{\sum_{j=1}^{3} ||\sigma^{\star(i,j)} - \sigma^{eq(k,i,j)}||^2_{\Omega_R}}{\sum_{j=1}^{3} ||\sigma^{\star(i,j)}||^2_{\Omega_R}}} \quad (70)
\]

The Figure 12 shows the correlation between \( e^{eq(k,i)} \) and \( \eta^{(k,i)} \). Such a correlation is sufficient to select the same hyper-reduced order model for the predictions with or without an equilibrium step. The range of variation of error on equilibrated stress is much smaller than the range of error on \( \sigma^{HR} \). Then, the equilibrated stresses are the simulation output of interest.
**Figure 12** – Correlation plot between $e^{eq(k,i)}$ and $\eta^{(k,i)}$ for points in the testing set. The spot color is related to the cluster index $k$.

For each cluster of data we have reported on Figure 13 the histogram of the density distribution of $e^{eq(k,i)}$, for $L^\eta_i = k$, inside each cluster separately.
Figure 13 – Density distribution for $e^{eq(k,i)}$ in each cluster, $k = 1, \ldots, 5$, for $L^\eta_i = k$.

Most of errors on equilibrated stresses are lower than 5%. The modeling procedure is very accurate. Errors related to medoids $m_1$, $m_4$ and $m_5$ are more concentrated below 5% than for medoids $m_2$ and $m_3$. Hyper-reduced predictions attached to medoids $m_2$ and $m_3$ are less accurate. Such results have been anticipated during the clustering procedure.

The following confusion matrix, denoted by $C$, is proposed to evaluate the error due to selection of medoid via $\eta^{(k,i)}$.

\[
C_{kk} = \sum_{i \text{ such that } L_i = k} e^{eq(k,i)}
\]

\[
C_{kp} = \sum_{i \text{ such that } L_i^\eta = p \text{ and } L_i = k} e^{eq(p,i)} - e^{eq(k,i)}
\]

(71)

(72)
We obtain the following results:

$$
\begin{bmatrix}
22 & 59 & 36 & 0 & 164 \\
10 & 25 & 87 & 4 & 125 \\
30 & 135 & 13 & 14 & 94 \\
162 & 36 & 26 & 43 & 85 \\
0 & 83 & 6 & 8 & 44
\end{bmatrix}
$$

(73)

Large confusion values are obtained for $C_{15}$ and $C_{51}$. It is probably because the related medoids are very similar. It is quite difficult to comment the other values.

5.4 Detailed numerical results on hyper-reduced prediction

In this section we report local numerical results for images in the test set. The index of the first image is $i = 1664$. Its label via the error indicator is $L^0_i = 3$. This label is the best: $L^0_i = L_i$. It is intentionally related to a medoid of the small cluster $C_3$, for which there is certainly a lack of observed defects. The prediction is globally accurate $e^{eq}(3, 1664) = 5\%$.

![Figure 14](image1.png)

**Figure 14** – Test defect #1664, from left to right: finite element prediction of the shear stress related to $E^{(3)}$, the related equilibrated stress $\sigma^{eq}$ on the RID around the defect, the error map for this stress component.

Better results are obtained for test data attached to $m_4$, the medoid of the larger cluster. The index of the second image is $i = 1987$. The global error is $e^{eq}(4, 1987) = 0.2\%$. The local predictions of the shear stress for $E^{(3)}$ are reported in Figure 15.

![Figure 15](image2.png)

**Figure 15** – Test defect #1987, from left to right: finite element prediction of the shear stress related to $E^{(3)}$, the related equilibrated stress $\sigma^{eq}$ on the RID around the defect, the error map for this stress component.

Local predictions are very accurate for this defect, certainly because cluster $C_4$ involves a large number of observed defects.
6 Conclusion

Many strong hypotheses have been made here. Only elastic deformation has been considered, with low stresses applied in consequence. Plasticity is widely present when fractures and harmfulness of defects are studied. The mechanical response of the defect, and therefore the classification, would significantly change in plasticity. But elastic classification of defects can be understood as a first classification step, because usually plastic transformation start with elastic strains. A study on 3D images would also be more relevant for speedup analysis. The higher the speedup obtained by hyper-reduced predictions the larger could be the number of medoids, because the computational complexity of medoid selection for test data is proportional to the hyper-reduction speedup times the number of medoids. Rom-nets are available to accelerate this selection by using a deep classifier. Such machine learning approaches require data augmentation schemes for future work. The simulation-based labeling of defects proposed in this paper should be very interesting for the development of deep classifiers that require a large set of labeled images.

Hyper-reduction of linear elastic problems is an opportunity to propose new theoretical results on the convergence of partial errors about stress predictions in reduced integration domains.

In this article, a method for data encoding relies on the mechanical response related to observational data. The Grassmann distance is therefore used to achieve such objective. It classifies the reduced models linked to each prediction of defect harmfulness, by assuming dilution condition. A common mesh, named encoding mesh, is required for the computation of Grassmann distances between the simulation data in the training set. This distance effectively detects a mechanical dissimilarity of defects, if the encoding mesh is refined enough.

A k-medoids clustering has been used to create medoids that compress the simulation data available in a training set. The dictionary used for hyper-reduced order models is the set of finite-element simulation-data of the computed medoids.

An error indicator is proposed to select a convenient medoid in the dictionary for the construction of the hyper-reduced order model that is expected to give the best stress prediction.

Appendix: details about the RID construction

Let’s introduce two mathematical operators. The first one, collects the degrees of freedom of a subdomain $\Omega_\alpha$:

$$\mathcal{L}(\Omega_\alpha) = \{ i \in \{1, \ldots, N\}, \int_{\Omega_\alpha} \phi_i^2 d\Omega > 0 \}$$

The second one aggregates the support of FE shape functions having their index in a set $\mathcal{G}$:

$$\mathcal{L}(\mathcal{G}) = \bigcup_{i \in \text{supp}\phi_i} \text{supp}(\phi_i), \mathcal{L}(\mathcal{G}) \subset \Omega$$

The extension of this subdomain by adding $n$ layers of connected elements reads:

$$(\mathcal{L} \circ \mathcal{C})^n \circ \mathcal{L}(\mathcal{G})$$

The operator $\mathcal{L}$ is fine for displacement fields, because they are approximated by the FE shape functions. A similar operator is also introduced for stresses. When collecting simulation data related to stresses, in the matrix $\bar{Q}^\sigma$, we store all the stress components at all Gauss points for all elements. Each row of $\bar{Q}^\sigma$ is related to one component of the stress tensor, at a Gauss point in an element. Then, the DEIM algorithm applied to $\bar{L}^\sigma$ gives a set of indices of components of the stress tensor, at some Gauss points in some elements. This set is denoted by $\mathcal{P}^\sigma$. We denote by $\mathcal{L}^\sigma(\mathcal{P}^\sigma)$ the support of the elements related to set $\mathcal{P}^\sigma$. $\mathcal{L}^\sigma(\mathcal{P}^\sigma)$ is a subdomain of $\Omega$.

In this paper, the RID construction was the following:

$$\Omega_R = (\mathcal{L} \circ \mathcal{C}) \circ (\mathcal{L}(\mathcal{P}) \cup \mathcal{L}^\sigma(\mathcal{P}^\sigma)) \cup \Omega_{ZOI}$$

where $\Omega_{ZOI}$ is the zone of interest.
Références

[1] Z. Li, X. Yan, C. Yuan, Z. Peng, and L. Li. Virtual prototype and experimental research on gear multi-fault diagnosis using wavelet-autoregressive model and principal component analysis method. *Mechanical Systems and Signal Processing*, 25 :2589–2607, 2011.

[2] T Mirá-Aguiar, C Leitão, and DM Rodrigues. Solid-state resistance seam welding of galvanized steel. *The International Journal of Advanced Manufacturing Technology*, 86(5) :1385–1391, Sep 2016.

[3] T Uwaba, Y Yano, and M Ito. Resistance spot weldability of 11cr-ferritic/martensitic steel sheets. *Journal of Nuclear Materials*, 421(1) :132 – 139, 2012.

[4] SK Dinda, JM Warnett, MA Williams, GG Roy, and P Srirangam. 3d imaging and quantification of porosity in electron beam welded dissimilar steel to fc-al alloy joints by x-ray tomography. *Materials and Design*, 96 :224 – 231, 2016.

[5] JD Madison and L K Aagesen. Quantitative characterization of porosity in laser welds of stainless steel. *Scripta Materialia*, 67(9) :783 – 786, 2012.

[6] A Haboudou, P Peyre, AB Vannes, and G Peix. Reduction of porosity content generated during nd :yag laser welding of a356 and aa5083 aluminium alloys. *Materials Science and Engineering : A*, 363(1) :40 – 52, 2003.

[7] Laurent Lacourt. *Lifetime assessment of welded structures containing defects*. PhD thesis, Mines ParisTech-PSL University, 2019.

[8] F N’Guyen. *Morphologie mathématique appliquée au développement d’outils de maillage EF automatis-iques dans le cas de microstructures hétérogènes bi et multiphasées*. PhD thesis, Lille 1 University, 2014.

[9] Laurent Lacourt, David Ryckelynck, Samuel Forest, Victor de Rancourt, and Sylvain Flouriot. Hyper-reduced direct numerical simulation of voids in welded joints via image-based modeling. *International Journal for Numerical Methods in Engineering*, n/a(n/a), 2020.

[10] S. Moomkesh, S.A. Mircei, M. Sadeghi, and M. Nazeri. Early detection of freezing damage in sweet lemons using vis/swnir spectroscopy. *Biosystems Engineering*, 164 :157–170, 2017.

[11] Y. Le Cun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and et al. L.D. Jackel. Hand-written digit recognition with a back-propagation network. *Advances in neural information processing systems*, 1990.

[12] A. Krizhevsky, I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2 :1097–1105, 2012. cited By 33336.

[13] T. Zehelein, T. Hemmert-Pottmann, and M. Lienkamp. Diagnosing automotive damper defects using convolutional neural networks and electronic stability control sensor signals. *Sensor and Actuator Networks*, pages 1–19, 2020.

[14] Ling Xiao, Bo Wu, Youmin Hu, and Jiwei Liu. A hierarchical features-based model for freight train defect inspection. *IEEE Sensors Journal*, 20 :2671–2678, 2020.

[15] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10) :1345–1359, 2010.

[16] A. T. W. Min, R. Sagarana, A. Gupta, Y. Ong, and C. K. Goh. Knowledge transfer through machine learning in aircraft design. *IEEE Computational Intelligence Magazine*, 12(4) :48–60, 2017.

[17] M. Oquab, L. Bottou, I. Laptev, and J. Sivic. Learning and transferring mid-level image representations using convolutional neural networks. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1717–1724, 2014.

[18] J Lumley. The structure of inhomogeneous turbulence. *Atmospheric turbulence and wave propagation*, Nauka,Moscow :166–178, 1967.

[19] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378 :686 – 707, 2019.
[20] Franck Nguyen, Selim M. Barhli, Daniel Pino Muñoz, and David Ryckelynck. Computer vision with error estimation for reduced order modeling of macroscopic mechanical tests. *Complexity*, 2018.

[21] Thomas Daniel, Fabien Casenave, Nissrine Akkari, and David Ryckelynck. Model order reduction assisted by deep neural networks (rom-net). *Advanced Modeling and Simulation in Engineering Sciences*, 7(1) :16, 2020.

[22] D Ryckelynck, K Lampoh, and S Quilici. Hyper-reduced predictions for lifetime assessment of elastoplastic structures. *Meccanica*, 51 :309–317, 2016.

[23] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathématiques*, 339(9) :667 – 672, 2004.

[24] J Fauque, I Ramiere, and D Ryckelynck. Hybrid hyper-reduced modeling for contact mechanics problems. *International Journal for Numerical Methods in Engineering*, 115(1) :117–139, JUL 6 2018.

[25] J. Baiges, R. Codina, and S. Idelson. A domain decomposition strategy for reduced order models. Application to the incompressible navier-stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 267 :23–42, 2013.

[26] P. Kerfriden, J. C. Passieux, and S. Bordas. Local/global model order reduction strategy for the simulation of quasi-brittle fracture. *International Journal for Numerical Methods in Engineering*, 89(2) :154–179, 2012.

[27] D. Amsallem and C. Farhat. Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA Journal*, 46 :1803–1813, 2008.

[28] R. Mosquera, A. Hamdouni, A. El Hamidi, and C. Allery. Pod basis interpolation via inverse distance weighting on grassmann manifolds. *Discrete and Continuous Dynamical Systems, Series S*, 12 :1743–1759, 2018.

[29] Park HS and Jun CH. A simple and fast algorithm for k-medoids clustering. *Expert Systems with Applications*, 36 :3336–3341, 2009.