Use of physics-augmented neural networks for unsupervised learning of material constitutive relations - Comparison of the NN-Euclid and NN-mCRE methods

Edgar Zembra, Antoine Benady, Emmanuel Baranger, Ludovic Chamoin

Use of physics-augmented neural networks for unsupervised learning of material constitutive relations

–

Comparison of the NN-EUCLID and NN-mCRE methods

Zembra Edgar
Master: *Modeling and Simulation in Structural Mechanics and Coupled Systems*

Supervisors:
Benady Antoine
Baranger Emmanuel
Chamoin Ludovic

Report submitted to
Barbarulo Andréa
Néron David

09/2023

Contents

1 Introduction

Model identification and updating is a core area of research in structural mechanics and computational engineering. Over the past decades, identification methods have evolved benefiting from improved measuring technologies, such as Digital Image Correlation (DIC)[36, 41], and from new concepts portraying material behaviors, like the modified Constitutive Relation Error (mCRE) [32]. More recently, identification procedures have benefited from numerical breakthroughs regarding artificial intelligence, such as the Efficient Unsupervised Constitutive Law Identification and Discovery (EUCLID)[17].

Traditionally, experimental observations of simple mechanical tests (uni-axial tension, torsion and bending) were used to update specific constitutive parameters of predefined state laws [8]. Thanks to advances in measuring techniques and computing power, experimental mechanics has adapted to data-driven methods. Many of them have benefited from DIC, as it enables reliable full-field measurements for a small number of tests, and from optic fibres [15, 14, 10], a highly flexible tool that can be embedded in multiple structures for partial-field measurements. An overview of identification strategies taking advantage of these measurement tools is presented in [4]. These include the Finite Element Model Updating (FEMU) framework [12, 35] and the Constitutive Equation Gap Method (CEGM)[9].

In such studies, as it is classically done, the analytical form of the constitutive behavior is pre-established (with unknown constitutive parameters). Although such procedures benefit from interpretable models, many issues have forced modern data-driven methods to develop identification strategies that overcome the need to postulate a behavior law. Indeed, developing a constitutive relation is difficult and time-consuming in terms of research and experimentation. Also, model bias arises as behavior laws become insufficient to describe the complexity of the experimental observations.

The idea to develop a neural network-based constitutive model was first introduced in 1991 in [20]. The neural network used is trained in a supervised manner, which means that the learning process consists in mapping known strain-stress pairs. Such strategies benefit from neural networks as a way to relax the model form and as universal approximators [22] in the sense that they can, in theory, describe any generic constitutive behavior without introducing any model bias. However, machine learning techniques, that are purely data-driven and for which no physical knowledge is embedded, suffer from difficulties to be trained, lack physical consistency and generalize poorly when operating on new data [44].

Also, supervised training procedures are not practically applicable within the data-driven constitutive identification framework. Indeed, stress labels are measured with great complexity and only for simple mechanical experiments. Otherwise, for a majority of tests, it is not possible to experimentally measure strain-stress pairs. Overall, the available data is mainly composed of measured kinematic variables, thus restraining training processes to be carried out in an unsupervised manner. However, it should be noted that supervised training processes are actively studied for optimizing computational time using parallel structures [3, 42, 19].

Recently, these issues have been tackled thanks to physics-augmented machine learning [25], where deep learning approaches are coupled to physical concepts. These are integrated into neural networks at three main levels:

- Initialization: known as transfer learning [38], the aim is to reduce the sensitivity to random initialization by configuring the initial weights and bias of the network with respect to *a priori* knowledge.

- Architecture: the idea is to verify physical and thermodynamics constraints (namely regarding convexity of free energy) throughout a neural network's architecture [26, 3]. The latter is referred to as physics-augmented architecture and is based on Input Convex Neural Networks (ICNNs) [2].
- Loss function: known as physics-informed loss functions [37], their goal is to evaluate the satisfaction of physical requirements and/or experimental observations.

Among other procedures that embody all of the above recent developments [31, 6], the NN-EUCLID [40] and NN-mCRE [5] methods are representative of the two main approaches there are for constitutive modeling using ICNNs.

The NN-EUCLID method lies within the EUCLID scope [17, 18, 24], which is an unsupervised data-driven method that uses sparse regression over a library of candidate functions to form interpretable constitutive relations. In [40], NN-EUCLID was tested to learn in an unsupervised way isotropic and anisotropic hyperelastic constitutive laws. Namely, ICNNs were trained for describing Mooney-Rivlin, Arruda-Boyce, Ogden and Holzapfel models with a loss function penalizing the non-conformity of equilibrium.

The NN-mCRE method lies within the Dynamic Data Driven Application Systems (DDDAS) framework, which aims to pave the way for Structural Health Monitoring (SHM) in a continuous and dynamic manner. The goal is to dynamically train with real-time measurements a computational model that continuously predicts the evolution of a quantity of interest and reacts accordingly. The NN-mCRE approach benefits from the mCRE-based loss function as a model error [32]. During the training phase, reliable information is enforced through an admissibility space, while uncertain information is relaxed through the functional. The mCRE strategy has namely been applied for identification problems regarding forced vibrations dynamics [29, 21], transient dynamics [1, 7], acoustics [13, 43] and nonlinear material behaviors [11, 33].

This study aims to review and compare the NN-EUCLID and NN-mCRE strategies. A main goal is to highlight the conceptual differences of these methods and to understand the choices made for each methodology within the neural network-based identification framework. Also, these approaches will be evaluated regarding the configurable parameters adjustment time (hyper-parameters, scaling factors...), the quantity of experimental data needed, the computational time, the quality and consistency of their predictions and their robustness to noise. Note that the implementation procedures for NN-EUCLID and NN-mCRE will not be taken into account, as the codes used are not necessarily optimized.

This study will focus on nonlinear elasticity (hyperelasticity with finite deformation context). Also, noisy synthetic data, simulating the measurements obtained through DIC, will be used.

The report is structured as follows. Section 2 sets the framework within the study. Section 3 and 4 focus on developing the methodology of NN-EUCLID and NN-mCRE. Section 5 layouts the experimental background for which these methods are tested and compared. Finally, in Section 6, conclusion and prospects are reached.

2 Problem definition

This section aims to set the identification problem to be solved, to define the experimental data on which the NN-EUCLID and NN-mCRE methods will operate and finally to detail the benefits of using neural networks within this framework.

2.1 Problem setting

Let Ω, composed of a material of unknown constitutive relation, be a subset of $\mathbb{R}^d$, with $d \in \{1, 2, 3\}$, and $\partial\Omega$ its boundary. This domain is subject to known Dirichlet and Neumann conditions respectively on $\partial\Omega_u$ and $\partial\Omega_f$ such that $\partial\Omega_u \cap \partial\Omega_f = \emptyset$. The corresponding imposed displacements and imposed forces are denoted $\mathbf{u}_\mathrm{d}$ and $\mathbf{f}_\mathrm{s}$. A known body force $\mathbf{f}_\mathrm{v}$ is also applied on Ω. This study case is represented Figure 1.

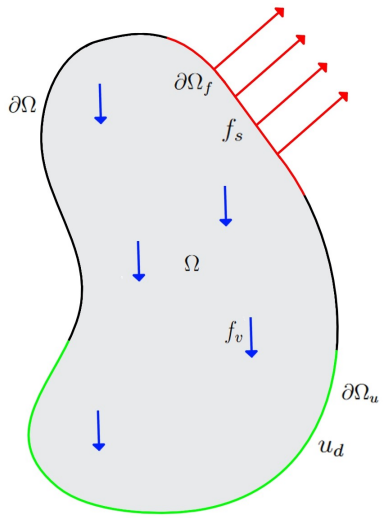


Figure 1: Study case

The associated mechanical problem is described by three sets of equations (the kinematic admissibility, the static admissibility and the constitutive relation), by a displacement field $\mathbf{u}$ and by a stress field, both fields being defined on Ω. Depending on the method (NN-EUCLID or NN-mCRE), either the first Piola-Kirchhoff stress field $\mathbf{P}$ or the second Piola-Kirchhoff stress field $\mathbf{S}$ is used. Hereunder the equations mentioned above are developed for each type of solution:

- **Kinematic admissibility**:

$$\mathbf{u} \in \mathcal{U}^{ad}_{\mathbf{u}_\mathrm{d}} = \{\mathbf{w}, \mathbf{w}\big|_{\partial\Omega_u} = \mathbf{u}_\mathrm{d}\} \tag{1}$$

- **Static admissibility**:

$$\mathbf{P} \in \mathcal{P}^{ad} = \left\{\mathbf{P},\ \forall \mathbf{v} \in \mathcal{U}^{ad}_{\mathbf{0}},\ \int_\Omega \mathbf{P} : \nabla\mathbf{v} = \int_\Omega \mathbf{f}_\mathrm{v}.\mathbf{v} + \int_{\partial\Omega_f} \mathbf{f}_\mathrm{s}.\mathbf{v}\right\} \tag{2}$$

or

$$\mathbf{S} \in \mathcal{S}^{ad} = \left\{\mathbf{S},\ \forall \mathbf{v} \in \mathcal{U}^{ad}_{\mathbf{0}},\ \int_\Omega \mathbf{S} : (\mathbf{F}^T.\nabla\mathbf{v}) = \int_\Omega \mathbf{f}_\mathrm{v}.\mathbf{v} + \int_{\partial\Omega_f} \mathbf{f}_\mathrm{s}.\mathbf{v}\right\} \tag{3}$$

- **Constitutive relation**:

$$\mathbf{P} = \frac{\partial\psi(\mathbf{F}, \mathbf{p})}{\partial\mathbf{F}} \quad \text{or} \quad \mathbf{S} = \frac{\partial\psi(\mathbf{E}, \mathbf{p})}{\partial\mathbf{E}} \tag{4}$$

  Note that $\psi(\mathbf{F}, \mathbf{p})$ and $\psi(\mathbf{E}, \mathbf{p})$ are two different functions. For the sake of brevity and because they satisfy the same properties (see below), they will be both mentioned as ψ.

Let $\mathbf{F} = \nabla\mathbf{u} + \mathbf{I_d}$ denote the deformation gradient, $\mathbf{E} = \frac{1}{2}(\nabla\mathbf{u} + \nabla^T\mathbf{u} + \nabla^T\mathbf{u}\nabla\mathbf{u})$ the strain tensor, $\mathbf{p}$ the state law's parameters and ψ a potential such that the constitutive relation satisfies physics and thermodynamics. This translates into having:

- ψ such that, at the reference state (no deformation), the stress field is null:

$$\mathbf{P}(\mathbf{F} = \mathbf{I_d}) = \mathbf{0} \quad \text{or} \quad \mathbf{S}(\mathbf{E} = \mathbf{0}) = \mathbf{0} \tag{5}$$

- a convex potential ψ with respect to $\mathbf{F}$ (for $\mathbf{P}$) or $\mathbf{E}$ (for $\mathbf{S}$)

A way to establish a constitutive model is to identify an analytical form for ψ that satisfies the above constraints. For example, for a quadratic potential $\psi = \frac{1}{2}\mathbf{E} : \mathrm{K} : \mathbf{E}$, the corresponding constitutive law is the one describing a linear elasticity problem: $\mathbf{S} = \mathrm{K}\mathbf{E}$.

2.2 Available experimental data

This study focuses on the use of data obtained by Digital Image Correlation (DIC) [36, 41]. It is a full-field measurement method adapted to kinematic variables such as displacement and strain. It allows great flexibility in the sense that it provides rich and reliable information for a small number of tests that can be conducted under non-homogeneous conditions. Note that the choice was made to work with DIC measurements as both NN-EUCLID and NN-mCRE can operate on such experimental data, making it possible to compare these methods.

It is important to underline that DIC (like most measurement techniques) does not allow to measure stress fields. Thus the pair $(\mathbf{F}, \mathbf{P})$ or $(\mathbf{E}, \mathbf{S})$, characteristic of the constitutive relation, is unknown.

Also, while using a full-field measurement method is practical for test studies, it is not adapted for a large scope of structures. To illustrate, this issue could be problematic while carrying out *System Health Monitoring* (SHM) on wind turbines, aircraft or spacecrafts. Hence in Section 5, an investigation regarding if the NN-EUCLID and NN-mCRE methods are adapted or can be adapted to partial-field measurements, will be led. This situation typically occurs when acquiring data using optic fibers: while it is a more flexible tool for concrete studies, measurements can only be done along the fiber axis.

2.3 Thermodynamics-consistent neural networks for representing constitutive laws

To sum up, the study considers a structure of known geometry, of unknown constitutive relation, submitted to known boundary conditions and for which displacement measurements are available on the whole domain. With these pieces of information, the objective is to recover the constitutive relation (4).

2.3.1 Use of neural networks

The solution proposed by the NN-EUCLID and NN-mCRE methods is to train in an unsupervised way Input Convex Neural Networks (ICNNs) [2] to predict the stress field associated to the measured displacement field and hence describe the constitutive relation thanks to the pair $(\mathbf{F}, \mathbf{P})$ or $(\mathbf{E}, \mathbf{S})$.

The use of neural networks is justified inherently. First of all, elaborating an analytical form of an unknown behavior law can be complex and time-consuming in terms of research and experimentation. Second, while the latter strategy benefits from interpretable models, these

may be insufficient to describe all the observations' complexity. This phenomena reflects the issue of model bias. In the end, neural networks overcome the need to postulate a constitutive relation and, benefiting from their quality as universal approximators [22], they can theoretically describe any behavior without introducing model bias.

In the case of this study, to satisfy the above mentioned physical and thermodynamic constraints, a ICNN-type neural network is used as it enforces physical knowledge throughout a physics-augmented architecture (represented Figure 2 for the constitutive relation associated to $(\mathbf{E}, \mathbf{S})$). Note that the ICNN's outputs are potentials ψ as predicting a stress field is equivalent to predicting ψ, according to (4).

2.3.2 ICNN's architecture

The neural network is composed of an input layer, where the deformation gradients $\mathbf{F}$ or the strain measurements $\mathbf{E}$ are specified, of m hidden layers, each formed of $(n_k)_{k=1,\dots,m}$ neurons, and of an output layer, where the prediction ψ_{NN} is returned.

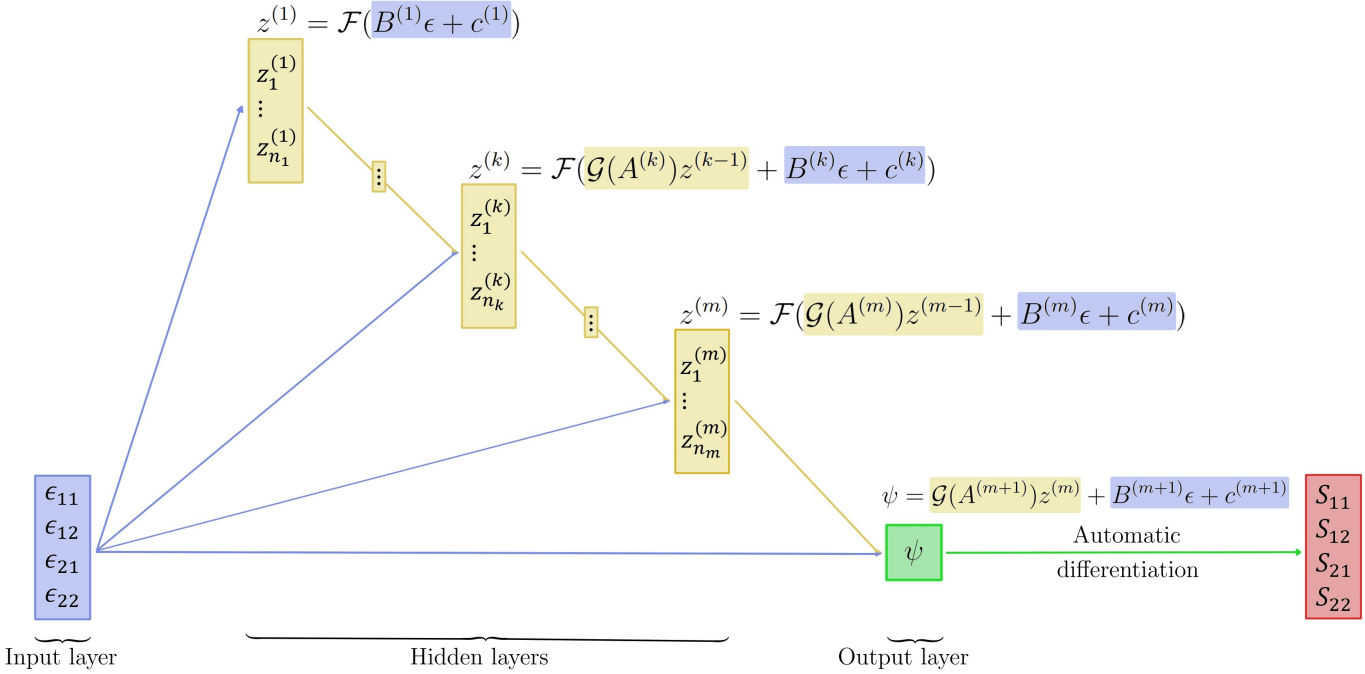


Figure 2: Input Convex Neural Network Architecture

To certify that the stress field and the potential vanish at the reference state, ψ is considered of the form [3, 26]:

$$\psi(\mathbf{F}, \mathbf{p}) = \psi_{NN}(\mathbf{F}, \mathbf{p}) - \psi_{NN}(\mathbf{I_d}, \mathbf{p}) - \frac{\partial \psi_{NN}}{\partial \mathbf{F}}\Big|_{\mathbf{F}=\mathbf{I_d}} : \mathbf{E} \tag{6}$$

or

$$\psi(\mathbf{E}, \mathbf{p}) = \psi_{NN}(\mathbf{E}, \mathbf{p}) - \psi_{NN}(\mathbf{0}, \mathbf{p}) - \frac{\partial \psi_{NN}}{\partial \mathbf{E}}\Big|_{\mathbf{E}=\mathbf{0}} : \mathbf{E} \tag{7}$$

To satisfy the convexity constraint of ψ with respect to $\mathbf{F}$ or $\mathbf{E}$, the activation function $\mathcal{F}$ is convex and non-decreasing and the weights of the intermediate hidden layers are positive. The latter is schematically represented by $\mathcal{G}$. In the case of this study:

$$\mathcal{F}(x) = \begin{cases} x^2 & if\ x > 0 \\ 0 & otherwise \end{cases} \qquad \text{and} \qquad \mathcal{G}(x) = \begin{cases} x & if\ x > 0 \\ 0 & otherwise \end{cases} \tag{8}$$

To avoid any loss of generality, skip connections link every hidden layer with the input layer throughout a linear transformation represented by $(B^{(k)}\mathbf{E} + c^{(k)})_{k=1,\dots,m}$ [2]. Note that the first hidden layer is considered fully as a skip connection layer. The output layer is calculated using the same transformation as for the hidden layers but without passing it through the activation function $\mathcal{F}$. Finally, $\mathbf{P}$ or $\mathbf{S}$ is obtained using automatic differentiation, where each mathematical operation is tracked in a computational graph enabling the calculation of the necessary gradients by following the chain rule for differentiation.

As a side note, this architecture can be adapted to other constraints. To illustrate, within the hyperelastic framework of [40], the objectivity of ψ with respect to $\mathbf{F}$ is verified by specifying the invariants based on the Cauchy-Green deformation tensor in the input layer instead of $\mathbf{F}$.

2.3.3 Initialization and hyper-parameters

Overall, the trainable parameters of the neural network (representing the constitutive parameters $\mathbf{p}$) are the weights $(A^{(k)}, B^{(k)})_{k=1,\dots,m}$ and the biases $(c^{(k)})_{k=1,\dots,m}$. The studied problem being highly ill-posed (absence of stress values), an important aspect of the training procedure is the initialization of the weights and biases. The outcome of the learning process is highly dependent on the initial values of these parameters. Multiple strategies exist, such as sorting the most accurate network out of an ensemble of neural networks trained with different random initializations, as it is done in [40]. The initialization method employed here, and implemented in [5], consists in operating on *a priori* knowledge and pre-training a network in a supervised manner with a database created from a known constitutive relation assumed to be close to that sought. The parameters obtained from this pre-training are used to initialize the neural network.

Another important aspect is the strategy employed to tune neural network's hyper-parameters. These parameters are well-known to be very sensitive and they require time-consuming research to be correctly set. This issue will be developed in the following sections as each method has a specific procedure for setting these parameters.

3 NN-EUCLID method

3.1 Overall presentation

The NN-EUCLID method is a continuation of the study denoted *Efficient Unsupervised Constitutive Law Identification and Discovery* (EUCLID) [17], which is an unsupervised data-driven method that uses sparse regression over a wide library of candidate functions to form interpretable constitutive relations. A main objective of the NN-EUCLID method is to overcome the limitation brought by a handcrafted feature library and expand the space of possible constitutive models by benefiting from neural networks as universal approximators. It should be noted that by doing so, interpretability of the model is lost.

The NN-EUCLID methodology is based on the *force balance method* (FEMU-F) [12, 35], also denoted as the *input residual method* or the *equilibrium gap method*. This identification process consists in minimizing a loss function evaluating the deviation from equilibrium. It requires full-field measurements and known boundary conditions (prescribed displacements and forces). To illustrate, for a given elasticity problem, the FEMU-F least-squared functional is written such that:

$$J_{\mathrm{F}}(\mathbf{p}) = \frac{1}{2}(\mathbf{F}_{obs} - \mathrm{K}(\mathbf{p})\mathbf{u}_{obs})^T \mathrm{W}_{\mathrm{F}}(\mathbf{F}_{obs} - \mathrm{K}(\mathbf{p})\mathbf{u}_{obs}) \tag{9}$$

where $\mathbf{u}_{obs}$ are the measured displacements, $\mathbf{F}_{obs}$ is a generalized load vector which incorporates the prescribed forces and $\mathrm{W_F}$ is a weighting matrix.

As is, the NN-EUCLID method cannot operate on partial-field measurements. However it could be adapted to the *displacement method* (FEMU-U) [12, 16, 35], also known as the *output residual method*, that was developed as an alternative approach from FEMU-F in the case where kinematic variables are partially known.

3.2 Methodology

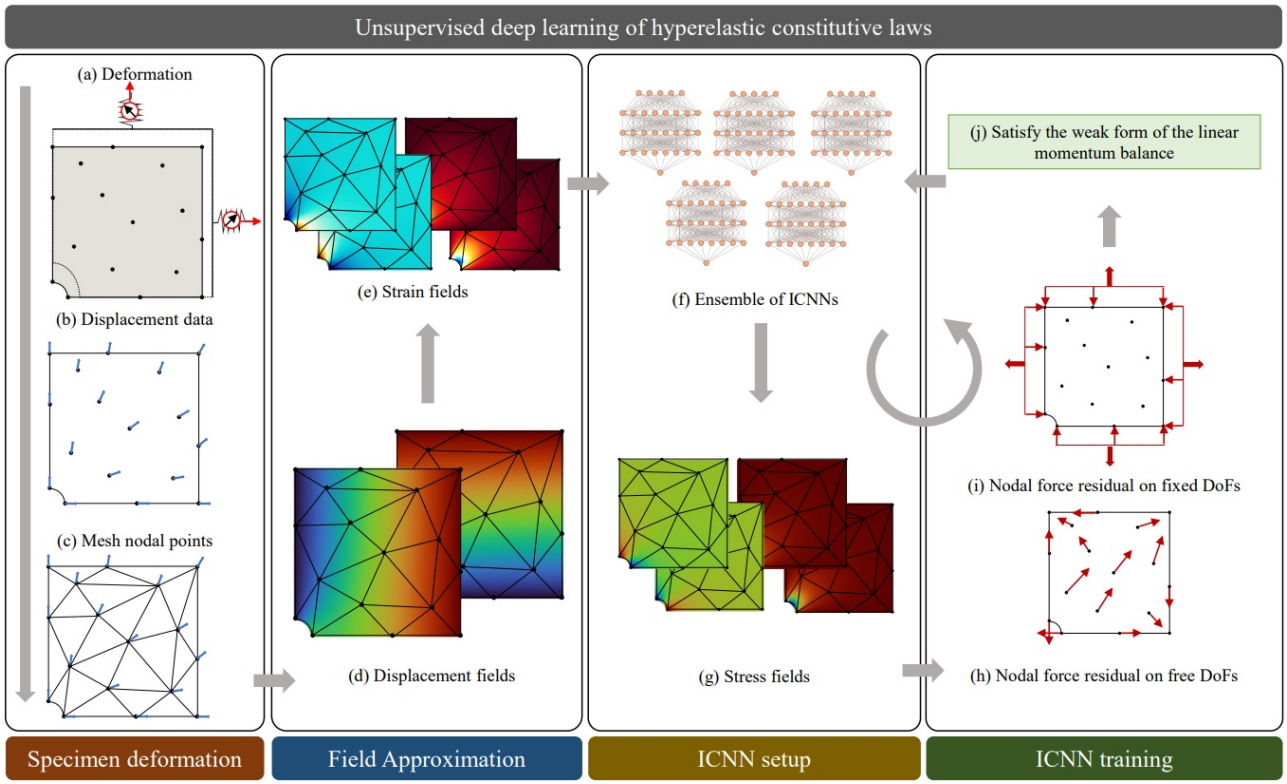


Figure 3: Schematic of the NN-EUCLID approach (Figure credited to [40])

3.2.1 Notations

In the case of this study, the experimental setup consists in taking n_t snapshots of the domain Ω composed of n_n measurement points where the displacements $\mathbf{u}_{obs}$ are collected. Both Dirichlet and Neumann conditions are considered and, to simplify calculations, no body force $\mathbf{f}_{\mathrm{v}}$ is applied. For each snapshot, n_β reaction forces, denoted $(R^{\beta,t})_{\beta=1,\dots,n_\beta,\ t=1,\dots,n_t}$, are measured.

Furthermore, the overall NN-EUCLID loss function is the sum of every snapshot's associated NN-EUCLID functional. So, for the sake of brevity in the following section, only one snapshot will be considered and the notation $t = 1, ..., n_t$ will be dropped.

Finally, within the NN-EUCLID framework, the choice was made to work with the following constitutive relation :

$$\mathbf{P} = \frac{\partial \psi(\mathbf{F}, \mathbf{p})}{\partial \mathbf{F}} \tag{10}$$

3.2.2 Minimization process

The training process of the ICNN (represented in Figure 3) consists in minimizing a loss function that tests the validity of the network's prediction in relation to the balance of forces:

$$\mathbf{p}_{opt} = \underset{\mathbf{p}}{\operatorname{argmin}}\, \mathcal{E}_{NN-EUCLID}(\mathbf{F}(\mathbf{u}_{obs}), \mathbf{p}) \tag{11}$$

To do so, the NN-EUCLID functional $\mathcal{E}_{NN-EUCLID}$ is computed from the static admissibility constraint:

$$\int_\Omega \mathbf{P} : \nabla \mathbf{v} = \int_\Omega \mathbf{f}_{\mathrm{v}}.\mathbf{v} + \int_{\partial\Omega_f} \mathbf{f}_{\mathrm{s}}.\mathbf{v} \quad \forall \mathbf{v} \in \mathcal{U}_{\mathbf{0}}^{ad} \tag{12}$$

Applying a Galerkin discretization using Finite Element shape functions (denoted $(\mathbf{N}^a)_{a=1,\dots,n_n}$), the above system can be written as:

$$\sum_{a=1}^{n_n} \mathrm{v}_i^a \mathrm{g}_i^a = 0 \qquad \text{where} \qquad \mathrm{g}_i^a = \int_\Omega \mathbf{P}_i \nabla \mathbf{N}^a - \int_\Omega (\mathbf{f}_{\mathrm{v}})_i \mathbf{N}^a - \int_{\partial\Omega_f} (\mathbf{f}_{\mathrm{s}})_i \mathbf{N}^a \qquad \forall i = 1, 2 \tag{13}$$

As (13) is verified for all fields $\mathbf{v}$ that vanish on the Dirichlet boundary, depending on the degree of freedom (DOF), two constraints can be identified:

- For DOFS that are not subjected to Dirichlet conditions:

  $$\mathrm{g}_i^a = 0 \qquad \forall (a, i) \in \mathcal{A}_{free} \tag{14}$$

  These DOFs are referred to as free DOFs and compose the space $\mathcal{A}_{free} = \{(a, i) \in \{1, ..., n_n\} \times \{1, 2\},\ \mathrm{v}_i^a \neq 0\}$
  This constraint is representative of the fact that, for free DOFs at equilibrium, the internal forces arising from the constitutive relation compensate with the external forces resulting from the imposed surface and body forces.

- For DOFs submitted to Dirichlet conditions:

  $$\sum_{(a,i) \in \mathcal{A}_{fixed}^\beta} \mathrm{g}_i^a = R^\beta \qquad \forall \beta = 1, ..., n_\beta \tag{15}$$

  These DOFs are referred to as fixed DOFs and compose the space $\mathcal{A}_{fixed}^\beta = \{(a, i) \in \{1, ..., n_n\} \times \{1, 2\},\ \mathrm{v}_i^a = 0\}$ This constraint is representative of the fact that, for fixed DOFs at equilibrium, the internal forces arising from the constitutive relation are balanced by the reaction forces deriving from the imposed displacements.

Overall, the NN-EUCLID loss function evaluates how relevant the predicted stress field is, with respect to the above mentioned constraints:

$$\mathcal{E}_{NN-EUCLID}(\mathbf{F}(\mathbf{u}_{obs}), \mathbf{p}) = \mathcal{E}_{free}(\mathbf{F}(\mathbf{u}_{obs}), \mathbf{p}) + \mathcal{E}_{fixed}(\mathbf{F}(\mathbf{u}_{obs}), \mathbf{p}) \tag{16}$$

where:

$$\mathcal{E}_{free}(\mathbf{F}(\mathbf{u}_{obs}), \mathbf{p}) = \sum_{(a,i) \in \mathcal{A}_{free}} (\mathrm{g}_i^a)^2 \qquad \text{and} \qquad \mathcal{E}_{fixed}(\mathbf{F}(\mathbf{u}_{obs}), \mathbf{p}) = \sum_{\beta=1}^{n_\beta} \left[R^\beta - \sum_{(a,i) \in \mathcal{A}_{fixed}^\beta} \mathrm{g}_i^a \right]^2 \tag{17}$$

The minimizing procedure is carried out via gradient descent:

$$\mathbf{p}_{updated} = \mathbf{p} - l_r \nabla_p \mathcal{E}_{NN-EUCLID}(\mathbf{F}(\mathbf{u}_{obs}), \mathbf{p}) \tag{18}$$

The learning rate l_r and the number of epochs are both manually tuned through trial and error.

In the Appendix A, the NN-EUCLID method is implemented in a simple case involving a one dimensional beam built in one of its ends and subjected to traction/compression on the other.

4 NN-mCRE method

4.1 Overall presentation

The NN-mCRE's training process is based on the minimization of the *modified Constitutive Relation Error* (mCRE) [32], also known as the *constitutive equation gap method* (CEGM) [9].

The *Constitutive Relation Error* (CRE) [28, 30] was introduced as *a posteriori* error estimation method for Finite Element computations. This approach consists in comparing a statically admissible stress field with one resulting from a given state law evaluated for a kinematically admissible strain field. The behavior model being the most uncertain data compared to the admissibility constraints, the CRE functional therefore evaluates an error on the constitutive relation. To illustrate, for a given elasticity problem and for $(\hat{\mathbf{u}}, \hat{\mathbf{S}}) \in \mathcal{U}^{ad}_{\mathbf{u}_\mathrm{d}} \times \mathcal{S}^{ad}$, the CRE is written such that:

$$\begin{aligned} \mathcal{E}^2_{CRE}(\hat{\mathbf{u}}, \hat{\mathbf{S}}) &= \int_\Omega (\hat{\mathbf{S}} - \mathrm{K}\mathbf{E}(\hat{\mathbf{u}})) : \mathrm{K}^{-1} : (\hat{\mathbf{S}} - \mathrm{K}\mathbf{E}(\hat{\mathbf{u}})) \\ &= 2 \times \int_\Omega \left(\frac{1}{2}\mathbf{E}(\hat{\mathbf{u}})\mathrm{K}\mathbf{E}(\hat{\mathbf{u}}) + \frac{1}{2}\hat{\mathbf{S}}\mathrm{K}^{-1}\hat{\mathbf{S}} - \hat{\mathbf{S}}\mathbf{E}(\hat{\mathbf{u}}) \right) \end{aligned} \tag{19}$$

This concept was then adapted to identification problems and parameter updating [27]. In this case, experimental data is considered as reliable information and embedded directly in the kinematic variables of the CRE-based loss function. Since measurement noise is a major obstacle to solving identification problems, this methodology may be to rigid, as it enforces noisy data.
An improvement to this aspect was proposed with the mCRE [32]. The key concept of this approach is to enforce reliable information, such as kinematic and static admissibility, and to weaken uncertain information, such as the state law, experimental data and unreliable boundary conditions, in the minimizing procedure of the mCRE-based loss function.
A main objective of the NN-mCRE method is to benefit from neural networks as universal approximators to overcome the need of a behavior model, as is it classically required for mCRE applications, and hence correct model bias.

In the following paragraphs, it will be shown that a main aspect of the minimization process is to entirely rebuild the kinematic and static fields to satisfy the admissibility constraints. So, inherently, the NN-mCRE method can operate on partial-field measurements.

4.2 Methodology

4.2.1 Notations

In the case of this study, the experimental setup consists in taking n_t snapshots of the domain Ω composed of n_n measurement points where the displacements are collected and from which the strain field $\mathbf{E}_{obs}$ is deduced. Both Dirichlet and Neumann conditions are considered and to simplify calculations, no body force $\mathbf{f}_\mathrm{v}$ is applied.

Furthermore, as for the NN-EUCLID methodology, the overall NN-mCRE loss function is a combination of every snapshot's associated NN-mCRE functional. So, for the sake of brevity in the following section, only one snapshot will be considered and the notation $t = 1, ..., n_t$ will be dropped.

Finally, within the NN-mCRE framework, the choice was made to work with the following

constitutive relation :

$$\mathbf{S} = \frac{\partial \psi(\mathbf{E}, p)}{\partial \mathbf{E}} \tag{20}$$

4.2.2 mCRE-based loss function

The training process of the ICNN consists in a two step minimization of the mCRE functional that evaluates the constitutive relation error and the distance to experimental data for kinematically and statically admissible fields ($\mathcal{A}_d = \mathcal{U}^{ad}_{\mathbf{u}_d} \times \mathcal{S}^{ad}$):

$$\mathbf{p}_{opt} = \underset{\mathbf{p}}{\operatorname{argmin}} \left[\min_{(\hat{\mathbf{u}}, \hat{\mathbf{S}}) \in \mathcal{A}_d} \mathcal{E}^2_{mCRE}(\hat{\mathbf{u}}, \hat{\mathbf{S}}, \mathbf{p}) \right] \tag{21}$$

where the mCRE functional is written such that:

$$\mathcal{E}^2_{mCRE}(\hat{\mathbf{u}}, \hat{\mathbf{S}}, \mathbf{p}) = \mathcal{E}^2_{CRE}(\hat{\mathbf{u}}, \hat{\mathbf{S}}, \mathbf{p}) + \alpha \parallel \Pi \mathbf{E}(\hat{\mathbf{u}}) - \mathbf{E}_{obs} \parallel^2 \tag{22}$$

with α a scaling factor and Π a projector.

Overall, the mCRE-based loss function is a weighted sum balanced by α that combines the two following characteristic terms:

- The CRE functional $\mathcal{E}^2_{CRE}$ evaluates model error for admissible fields such that:

$$\mathcal{E}^2_{CRE}(\hat{\mathbf{u}}, \hat{\mathbf{S}}, \mathbf{p}) = \int_\Omega \left[\psi(\mathbf{E}(\hat{\mathbf{u}}), \mathbf{p}) + \psi^*(\hat{\mathbf{S}}, \mathbf{p}) - \hat{\mathbf{S}} : \mathbf{E}(\hat{\mathbf{u}}) \right] \tag{23}$$

  The potential ψ^* defines the material law $\mathbf{E} = \frac{\partial \psi^*(\mathbf{S}, \mathbf{p})}{\partial \mathbf{S}}$ and is dual to ψ in the Legendre-Fenchel sense:

$$\psi^*(\hat{\mathbf{S}}, \mathbf{p}) = \sup_{\mathbf{E}} \left[\hat{\mathbf{S}} : \mathbf{E} - \psi(\mathbf{E}, \mathbf{p}) \right] \tag{24}$$

  In short, CRE tests the discrepancy from this duality given admissible fields. A graphic interpretation of the CRE is given Figure 4: for a given couple $(\mathbf{E}(\hat{\mathbf{u}}), \hat{\mathbf{S}})$, the hatched area (representing CRE) is what remains when the red rectangle (representing $\hat{\mathbf{S}} : \mathbf{E}(\hat{\mathbf{u}})$) is subtracted to the orange and blue area (respectfully representing $\psi(\mathbf{E}(\hat{\mathbf{u}}), \mathbf{p})$ and $\psi^*(\hat{\mathbf{S}}, \mathbf{p})$).

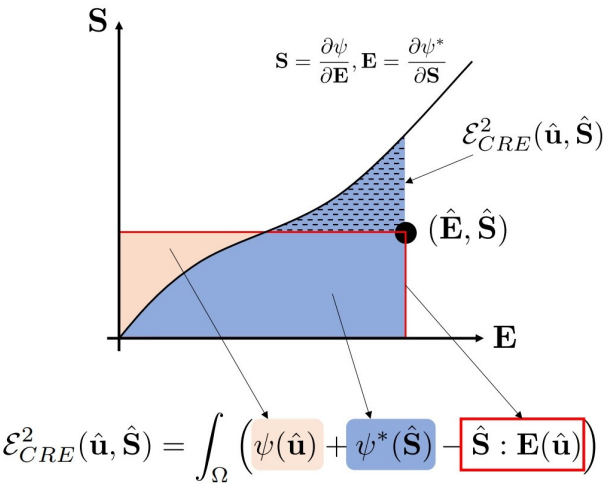


Figure 4: CRE interpretation with the stress-strain curve (Figure credited to [5])

- $\parallel \Pi \mathbf{E}(\hat{\mathbf{u}}) - \mathbf{E}_{obs} \parallel^2$ quantifies the distance to observations by projecting $\hat{\mathbf{u}}$ through Π onto the measured variables.

To express $\psi^*(\hat{\mathbf{S}}, \mathbf{p})$ in a more convenient way, let $\hat{\mathbf{v}}$ be the displacement field such that the maximum regarding (24) is reached for $\mathbf{E}(\hat{\mathbf{v}})$. Noticing that $\hat{\mathbf{S}} = \frac{\partial \psi}{\partial \mathbf{E}}\Big|_{\mathbf{E}(\hat{\mathbf{v}})}$, the potential $\psi^*(\hat{\mathbf{S}}, \mathbf{p})$ can be rewritten such that:

$$\psi^*(\hat{\mathbf{S}}, \mathbf{p}) = \frac{\partial \psi}{\partial \mathbf{E}}\Big|_{\mathbf{E}(\hat{\mathbf{v}})} : \mathbf{E}(\hat{\mathbf{v}}) - \psi(\mathbf{E}(\hat{\mathbf{v}}, \mathbf{p})) \tag{25}$$

Finally, the mCRE functional to minimize is given by:

$$\mathcal{E}^2_{mCRE}(\hat{\mathbf{u}}, \hat{\mathbf{v}}, \mathbf{p}) = \int_\Omega \left[\psi(\mathbf{E}(\hat{\mathbf{u}}), \mathbf{p}) - \psi(\mathbf{E}(\hat{\mathbf{v}}), \mathbf{p}) + \frac{\partial \psi}{\partial \mathbf{E}}\Big|_{\mathbf{E}(\hat{\mathbf{v}})} : (\mathbf{E}(\hat{\mathbf{v}}) - \mathbf{E}(\hat{\mathbf{u}})) \right] + \alpha \parallel \Pi \mathbf{E}(\hat{\mathbf{u}}) - \mathbf{E}_{obs} \parallel^2 \tag{26}$$

4.2.3 Minimization of the loss function

The minimization procedure is carried out iteratively in two steps.

Step 1: The first step consists in finding admissible fields $(\hat{\mathbf{u}}, \hat{\mathbf{S}}) \in \mathcal{A}_d$ for given parameters $\mathbf{p}$ by means of a constrained minimization. This procedure is carried out through the Newton scheme associated to the following Lagrangian functional $\mathcal{L}$:

$$\mathcal{L} = \mathcal{E}^2_{mCRE}(\hat{\mathbf{u}}, \hat{\mathbf{v}}) - h(\hat{\mathbf{v}}, \boldsymbol{\lambda}) \quad \forall \boldsymbol{\lambda} \in \mathcal{U}^{ad}_{\mathbf{0}} \tag{27}$$

where h is the static admissibility constraint:

$$\underbrace{\int_\Omega \hat{\mathbf{S}}(\hat{\mathbf{v}}) : (\mathbf{F}(\hat{\mathbf{v}})^T . \nabla \boldsymbol{\lambda}) - \int_\Omega \mathbf{f}_{\mathrm{v}} . \boldsymbol{\lambda} - \int_{\partial \Omega_f} \mathbf{f}_{\mathrm{s}} . \lambda}_{h(\hat{\mathbf{v}}, \boldsymbol{\lambda})} = 0 \quad \forall \boldsymbol{\lambda} \in \mathcal{U}^{ad}_{\mathbf{0}} \tag{28}$$

It is important to note that the Dirichlet conditions are directly embedded in $\hat{\mathbf{u}}$ such that the kinematic admissibility is satisfied.

The Newton scheme is given by:

$$\begin{bmatrix} \mathbf{u}_{k+1} - \mathbf{u}_k \\ \mathbf{v}_{k+1} - \mathbf{v}_k \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 \mathcal{E}^2_{mCRE}}{\partial \mathbf{u}^2} & \frac{\partial^2 \mathcal{E}^2_{mCRE}}{\partial \mathbf{u} \partial \mathbf{v}} & \mathbf{0} \\ \frac{\partial^2 \mathcal{E}^2_{mCRE}}{\partial \mathbf{v} \partial \mathbf{u}} & \frac{\partial^2 \mathcal{E}^2_{mCRE}}{\partial \mathbf{v}^2} & \frac{\partial h(\mathbf{v}, \lambda)}{\partial \mathbf{v}} \\ \mathbf{0} & \frac{\partial h(\mathbf{v}, \lambda)}{\partial \mathbf{v}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial \mathcal{E}^2_{mCRE}}{\partial \mathbf{u}} \\ \frac{\partial \mathcal{E}^2_{mCRE}}{\partial \mathbf{v}} \\ h(\mathbf{v}, \lambda) \end{bmatrix} \tag{29}$$

Step 2: The second step involves updating the parameters $\mathbf{p}$ throughout a gradient based method using the above computed admissible fields $(\hat{\mathbf{u}}, \hat{\mathbf{S}})$:

$$\mathbf{p}_{updated} = \mathbf{p} - l_r \nabla_p \mathcal{E}^2_{mCRE}(\hat{\mathbf{u}}, \hat{\mathbf{v}}, \mathbf{p}) \tag{30}$$

The gradient of the loss function with respect to the parameters is computed using the adjoint-state method, from which the following result is obtained:

$$\begin{aligned} \frac{\mathrm{d}\mathcal{E}^2_{mCRE}(\hat{\mathbf{u}}, \hat{\mathbf{v}}, \mathbf{p})}{\mathrm{d}\mathbf{p}} &= \frac{\partial \mathcal{L}}{\partial \mathbf{p}} \\ &= \int_\Omega \frac{\partial \psi}{\partial \mathbf{p}}\Big|_{\mathbf{E}(\hat{\mathbf{u}})} - \frac{\partial \psi}{\partial \mathbf{p}}\Big|_{\mathbf{E}(\hat{\mathbf{v}})} \end{aligned} \tag{31}$$

In the Appendix B, the NN-mCRE method is implemented in a simple case involving a one dimensional beam built in one of its ends and subjected to traction/compression on the other.

4.2.4 Automation of user configurable parameters

In the DDDAS context where the NN-mCRE training process is carried out online, special attention was brought to automate the tuning of hyper-parameters and the tuning of the scaling factor α [5]. As it is known, these parameters are very sensitive and require time-consuming research to be correctly determined. So, within this framework, a manual configuration is not adapted.

Automation of the scaling factor $\boldsymbol{\alpha}$:
Minimizing the mCRE functional is finding a compromise between satisfying the model and satisfying the experimental data. The weight α adjusts the influence of each of these terms in the loss function. The more reliable (respectively unreliable) the measurements are, the higher (respectively the lower) the value of α should be set.

In the case of this study, the noise level (known *a priori*) is a good indicator to evaluate to what extent the experimental data should be relied on. The presented strategy, following the Morozov criterion [34], aims to update $(\hat{\mathbf{u}}, \hat{\mathbf{S}})$ at the noise level. Let σ be the standard deviation of the measurement noise. The scaling factor is written such that $\alpha = \alpha'/n_n\sigma^2$. This way, updating the admissible fields at the noise level translates into tuning α' such that:

$$\frac{\| \Pi\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{E}_{obs} \|^2}{n_n\sigma^2} = 1 \tag{32}$$

The process for setting α' is as follows. For each training cycle, the updated admissible fields (computed throughout step 1) are tested with respect to the above constraint. If (32) is not satisfied, then α' is tuned using dichotomy and new admissible fields are computed. This procedure is repeated until the constraint is satisfied.

Automation of the learning rate $\mathbf{l_r}$:
The proposed approach is empirical and can be adjusted depending on the test case. The key idea is to adapt the learning rate depending on the number of iterations needed for the Newton scheme to compute the updated admissible fields during step 1 of the training cycle (composed of step 1 and step 2).
Let us consider the $(n+1)^{th}$ epoch of the training process, where $(\hat{\mathbf{u}}^n, \hat{\mathbf{S}}^n)$,$\mathbf{p}^n$ and l_r^n have already been computed. As a reminder, $(\hat{\mathbf{u}}^n, \hat{\mathbf{S}}^n)$ (respectfully $(\hat{\mathbf{u}}^{n+1}, \hat{\mathbf{S}}^{n+1})$) are obtained by running step 1 for fixed parameters $\mathbf{p}^{n-1}$ (respectfully $\mathbf{p}^n$). The number of iterations needed for the Newton scheme to go from $(\hat{\mathbf{u}}^n, \hat{\mathbf{S}}^n)$ to $(\hat{\mathbf{u}}^{n+1}, \hat{\mathbf{S}}^{n+1})$ is a good indicator to quantify how much the parameters $\mathbf{p}$ have changed from $\mathbf{p}^{n-1}$ to $\mathbf{p}^n$. If the number of iteration is small, then the parameters have changed little and the learning rate can be increased ($l_r^{n+1} > l_r^n$). If the number of iteration is high, then they have changed a lot, which can cause instability. Therefore, the learning rate should be reduced ($l_r^{n+1} < l_r^n$).
Defining an adequate number of iterations is empirically achieved throughout grid search. In the case of this study, the learning rate is suitable when four iterations are needed.

Automation of the end of training:
Rather than setting a fixed number of epochs, the methodology consists in defining a physics-based stop criterion using the CRE. This functional benefits from having a strong physical sense as it can be interpreted as a model error. Also, since it is homogeneous to an energy, the CRE functional can be compared to the structure's energy $\int_\Omega \psi(\mathbf{E}, p)$ such that:

$$\eta = \frac{\mathcal{E}_{CRE}^2}{\int_\Omega \psi(\mathbf{E}, p)} < tolerance \tag{33}$$

where the tolerance is set according to the user's requirements.

5 Comparative results

5.1 Comparison methodology

The aim of this section is to highlight the advantages and disadvantages of NN-EUCLID and NN-mCRE and to understand the choices made for each methodology with respect to the issues raised by a neural networked-based identification process.

5.1.1 Conceptual differences between the NN-EUCLID and NN-mCRE methods

The NN-EUCLID and NN-mCRE method share the same goal which is to propose a neural network-based identification process that only requires kinematic field measurements to operate. However, as these methods are developed within different frameworks, the strategies employed to solve this problem differ.

As a reminder, in [40], the attention is focused on learning isotropic and anisotropic hyperelasticity. In this context, NN-EUCLID must be able to operate on highly heterogeneous displacement fields. To fully grasp the entire spectrum of such material behavior, it is in their favor to use DIC. Indeed, this technique enables full-field measurements, thus maximizing data collection, and is particularly adapted to provide rich and reliable information for tests conducted under non-homogeneous conditions. The aim for the NN-EUCLID method is to make the most of the abundant data to predict at best complex behaviors. The proposed training process is hence entirely based on experimental measurements in the sense that the neural network learns to satisfy equilibrium regarding experimental displacement fields. In the following sections, it will be shown that this strategy benefits from a straightforward and fast training phase. However, the resulting downfall is a loss function that does not focus on what should be learned, namely the constitutive relation.

On the other hand, in [5], the attention is focused on Dynamic Data Driven Application Systems (DDDAS) which aims to conduct Structural Health Monitoring (SHM) in a continuous and dynamic manner. In this context, the challenge lies within detecting early damage and predicting its evolution in sensor-equipped structures to control their integrity throughout their life cycle. By driving the systems accordingly, the goal is to increase durability and safety. Overall, NN-mCRE has been thought to be applied over whole structures like wind turbines, aircraft and spacecrafts, making it impossible to use measurement techniques such as DIC.
Highlighting a main difference between NN-EUCLID and NN-mCRE, these two methods are not meant to operate on the same kind of data. Here, a measurement tool adapted to large structures is necessary. With this in mind, optic fibers have shown to provide accurate experimental information on the damage state for a large scope of structures. The main downfall is that measurements are only enabled along the fiber's axis, thus compelling NN-mCRE to operate on partial-field measurements. The proposed training strategy is hence based on the mCRE functional which benefits from working with admissible fields, instead of partially known experimental fields. The measurements are not as valued compared to the NN-EUCLID method as they are considered uncertain and relaxed in the loss function. A main aspect of the NN-mCRE training procedure is to reconstruct admissible fields on the whole structure regarding reliable information. Consequently, this strategy suffers from a more complex and computationally expensive training phase. However, the mCRE functional is specially adapted for neural network-based identification problem, as it evaluates model error directly on what needs to be learned, namely the constitutive relation.
Emphasizing on another focal difference between the NN-EUCLID and NN-mCRE methods, while the first has favored rapid training over a learning process that evaluates the predicted state law, the other values constitutive model error at the expense of costly computations.

Finally, another key difference between the two studied methods is the approach adopted for tuning hyper-parameters. While the NN-EUCLID context does not limit the possible strategies for setting these parameters, the DDDAS framework prohibits manual configuration as the neural network could be trained online. In the following sections, it will be shown that the NN-mCRE pre-training phase is much shorter than the NN-EUCLID one.

5.1.2 Training time

A central comparing criterion is the time required for implementing a functional neural network. This process includes the pre-training phase, for which research is dedicated to correctly tuning the hyper-parameters, and the training phase, for which the hyper-parameters are set and the weights and biases of the neural network are iteratively updated. Once again, being developed within two different frameworks, the NN-EUCLID and NN-mCRE methods have tackled this issue in two different manners.

Only considering the computational time for fixed hyper-parameters, the NN-EUCLID benefits from a straightforward training process. The computational cost is mainly due to the calculation of the stress fields, of the loss function (using numerical quadrature for the integrals) and of the trainable parameters' updating phase.
The NN-mCRE training process is much denser. The main difference with the NN-EUCLID methodology, is that, for each training cycle, in addition to the above, new admissible fields are computed by means of a constrained minimization carried out through a Newton scheme. This considerably weighs up the computational cost as it is an iterative process.
So, for correctly tuned networks, the NN-EUCLID training process is expected to be faster than the NN-mCRE one.

However, this statement is to be qualified with regard to the time required for correctly tuning the hyper-parameters. For NN-EUCLID, these parameters are manually set through trial and error. In the case of this study, the NN-EUCLID code has been adapted such that only the number of epochs and the learning rate (which includes the scheduler, the range and the steps) are configurable. But it should be noted that in [40], the configurable hyper-parameters also include the number of hidden layers, the number of neurons per hidden layer, the dropout rate in the hidden layers, the scaling parameters in the activation functions, and the number of neural networks in the ensemble used for reducing the sensitivity regarding initialization. As it is known, these parameters are very sensible and demand time-consuming research to be correctly set.
To tackle this issue and to be in line within the DDDAS framework, the main goal of NN-mCRE is to automate the setting of these variables. The strategies proposed by the NN-mCRE method benefit from the strong physical sense incorporated in the mCRE functional (Section 4). The stop criterion takes advantage of the fact that the CRE term of the loss function is homogeneous to an energy and can be interpreted as a model error. The scaling factor is set based on the loss evaluating the distance to observations in a way that the latter is equal to the noise level. Finally, the learning rate is empirically adjusted depending on how much the trainable parameters have been updated. Even though the strategies proposed are not optimal (α is updated through dichotomy, which is a computationally costly process, and the space of possible learning rates is empirically obtained throughout testing), they can be improved and adapted depending on the study case.
Overall, the pre-training phase is expected to last much longer for NN-EUCLID than for NN-mCRE.

5.1.3 Flexibility to incomplete experimental data

To have a full-field measurement setup available is ideal yet unpractical for a broad field of applications. Indeed, while it is adapted for studies carried out in laboratories or in a short zone of the considered subject, such measurements are unfeasible over whole structures like wind turbines, aircraft or spacecraft. So, testing whether or not NN-EUCLID and NN-mCRE are adapted or can be adapted to partial-field measurements is an important comparison criterion.

As is, NN-EUCLID can not operate on incomplete data. To work properly, this method requires full-field displacement measurements, known Neumann conditions and the reaction forces associated to the Dirichlet conditions. As presented in Section 3, the NN-EUCLID loss function is based on the static admissibility constraint, which is integrated on Ω. So, for the latter to be meaningful, the stress field must be defined on the whole domain and the imposed (surface and body) forces must be known. The first Piola-Kirchhoff stress field being fully computed from the experimental data, these necessarily have to be measured in a full-field manner. Furthermore, one of the characteristic terms of the loss function compares the internal forces arising from the constitutive relation to the reaction forces. If one of the latter were to be unknown, the comparison with the internal forces would be impossible. This means that the displacement measurements needed to compute these internal forces would be unused. Hence, all the reaction forces deriving from the Dirichlet conditions must be measured or else the neural network could not train on all the available experimental data. However, NN-EUCLID could be adapted to incomplete data by establishing a loss function based on the displacement method.

Considering the DDDAS framework, the NN-mCRE method was established to inherently operate on partial-field measurements. Unlike NN-EUCLID, the strategy is not to work directly with experimental data but with new admissible fields computed on the whole domain. These fields are only partly representative of the measurements as they are hybrid solutions resulting from a compromise between verifying the constitutive model and verifying the available experimental data. Also, the comparison is partial as only the admissible displacements projected on the measurement points can be compared to the experimental data. Overall, the measured fields are not used to define but merely to guide the construction and minimization of the loss function. Hence, only partial-field measurements are required for the proper functioning of the method. In [5], NN-mCRE is tested in the case where strain measurements are obtained using optic fibers. While the latter is a very flexible tool and can be embedded in a large scope of structures, the measurements can only be carried out along the the axis of the fiber.

5.1.4 Flexibility to noise

Throughout the studied methods, two different strategies are presented to account for noisy measurements.

In [40], the proposed approach is to denoise the experimental data beforehand using the Kernel Ridge Regression (KRR) algorithm [39]. This allows to interpolate the noisy data by performing a non-linear regression based on the construction of a linear regression function in a high dimensional feature space. The interpolated displacements are then used to train the neural network. This approach is consistent with the NN-EUCLID methodology, as its training process does not account for measurement noise. If the data was not treated beforehand, the minimization of the loss function, which is fully computed from these measurements, would be flawed. It should be noted that, even though the Dirichlet and Neumann conditions are supposed to be known in a reliable way, no strategy is proposed to adapt the NN-EUCLID method to uncertain and noisy boundary conditions.

Considering the NN-mCRE strategy, measurement noise is taken into account in the training

process at multiple levels. First of all, benefiting from the mCRE functional, the experimental data is considered as uncertain information and relaxed in the loss function. So, the measurements will at best be satisfied when minimizing the mCRE functional. In an ideal situation with no measurement noise, the experimental data could be enforced through an admissibility constraint. Second, the scaling factor α quantifies the trust placed into the measurements and into the constitutive model. The more reliable the measured displacements are (respectively the more uncertain the model is) the higher α should be set and vice versa. In [5], this weight is tuned such that the experimental data is verified up to the noise level (which is supposed to be known). It should be noted that, Dirichlet and Neumann conditions can also be considered as unreliable information and incorporated in the mCRE functional. Many applications can be found in [23].

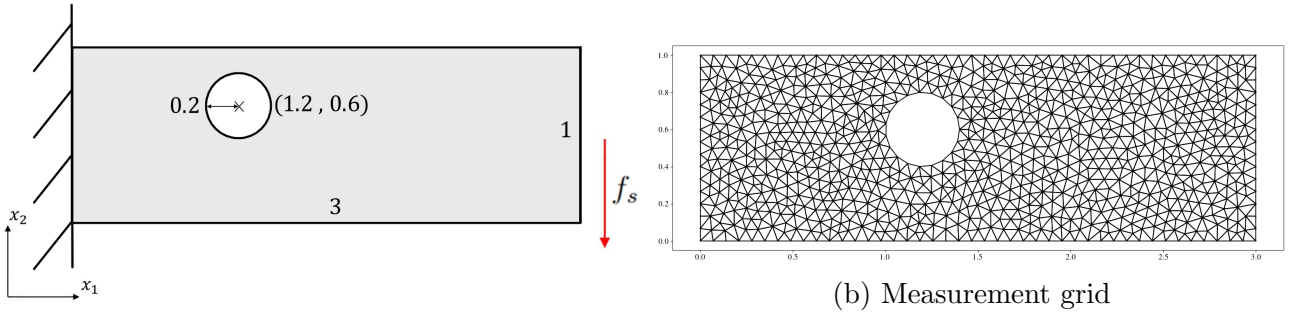
In the case of this study, one of the main goals is to test the robustness of each method with respect to measurement noise. So, the choice was made to use noisy data for the training of the neural networks without any upstream denoising technique. The expected results are that the NN-EUCLID predictions are much more influenced by noisy data than the NN-mCRE predictions.

5.2 Numerical comparison

This section aims to present the numerical framework in which the NN-EUCLID and NN-mCRE methods are compared. The neural networks are trained on synthetic data disturbed by artificial noise and obtained from FEM simulations carried out on a structure presented below. The methods are then compared based on the research-time required for tuning the ICNNs, the computational time of the training processes and the quality of the predictions made by the trained models.

5.2.1 Numerical framework

Let Ω be a 2D rectangle with a hole, clamped on one of its ends and subject to a pure quasi-static bending force on the other (Figure 5a). For this study, the choice was made to focus on small perturbations.



(a) Structure

(b) Measurement grid

Figure 5: Numerical framework

The associated constitutive relation to learn is described by a non-quadratic potential representative of non-linear elasticity. Here, the material behavior differs depending on whether the subject is submitted to traction or compression along the horizontal axis:

$$\psi = \frac{1}{2}\mathrm{E}^{+}(\mathbf{E}_{11}^{+})^2 + \frac{1}{2}\mathrm{E}^{-}(\mathbf{E}_{11}^{-})^2 + \frac{1}{2}\mathrm{E}^{-}(\mathbf{E}_{22})^2 + \frac{1}{2}\mathrm{G}(\mathbf{E}_{12}+\mathbf{E}_{21})^2 \tag{34}$$

$\mathrm{E}^+ = 8GPa$ is the traction Young modulus associated to the positive strain $\mathbf{E}_{11}^+$, $\mathrm{E}^- = 20GPa$ is the compression Young modulus associated to the negative strain $\mathbf{E}_{11}^-$ and $\mathrm{G} = 8GPa$ is the Coulomb modulus.

Note that the geometry and the applied load (pure bending along the vertical axis) are specifically chosen to highlight a key issue of neural network-based identification procedures. For a neural network to fully learn a constitutive model, it requires to train on data that is representative of the entire spectrum of a material's behavior. Neural networks can hardly generalize when operating on data that is not relevant to the training database. In this context, pure bending along the vertical axis is interesting to study as it involves both traction and compression along the horizontal axis, allowing the neural network to learn the nonlinearity aspect of the behavior. However, for this geometry and this load case, the strain components $\mathbf{E}_{12}$, $\mathbf{E}_{21}$ and $\mathbf{E}_{22}$ are much less solicited. So, it is expected from the neural networks to learn less in these directions.

5.2.2 Available data

In the case of this study, the NN-EUCLID and NN-mCRE methods operate on synthetic data generated from FEM simulations. The numerical database aims to simulate full-field measurements carried out on the structure using DIC. Hence, the numerical meshing of Ω is assumed to overlap the experimental measurement grid such that each node corresponds to a measurement point. The grid is such that 874 measurement points are spread across the entire structure (Figure 5b).

The numerical information is disturbed by artificial noise μ. The latter is added to the displacement measurements such that: $\mu \sim \mathcal{N}(0, \sigma^2)$. This means that the noise follows a zero-centered normal distribution of standard deviation σ. Two noise levels are considered: a low noise for $\sigma = 10^{-4}$ and a high noise for $\sigma = 10^{-3}$.

Finally, the boundary conditions are supposed to be known and reliable. For the numerical simulations, five snapshots will be considered where the following loads are applied:

$$f_s = \{-5 \times 10^{-4},\ -2 \times 10^{-4},\ 1 \times 10^{-4},\ 3 \times 10^{-4},\ 7 \times 10^{-4}\} \tag{35}$$

It is important to note that, while the displacement and stress fields are computed through FEM simulations, only the displacements will be used to train the neural network.

5.2.3 ICNN configuration

In the context of this study, the considered ICNNs are composed of 2 hidden layers formed by 50 neurons each. Regarding initialization (Figure 6), the networks are pre-trained in a supervised manner to represent linear elasticity. The initial model is supposed to be close to the one sought and is described by the following quadratic potential:

$$\psi \ = \ \frac{1}{2}\mathrm{E}(\mathbf{E}_{11})^2 \ + \ \frac{1}{2}\mathrm{E}(\mathbf{E}_{22})^2 \ + \ \frac{1}{2}\mathrm{G}(\mathbf{E}_{12} + \mathbf{E}_{21})^2 \tag{36}$$

where $\mathrm{E} = 24GPa$ is the Young modulus and $\mathrm{G} = 11GPa$ is the Coulomb modulus.

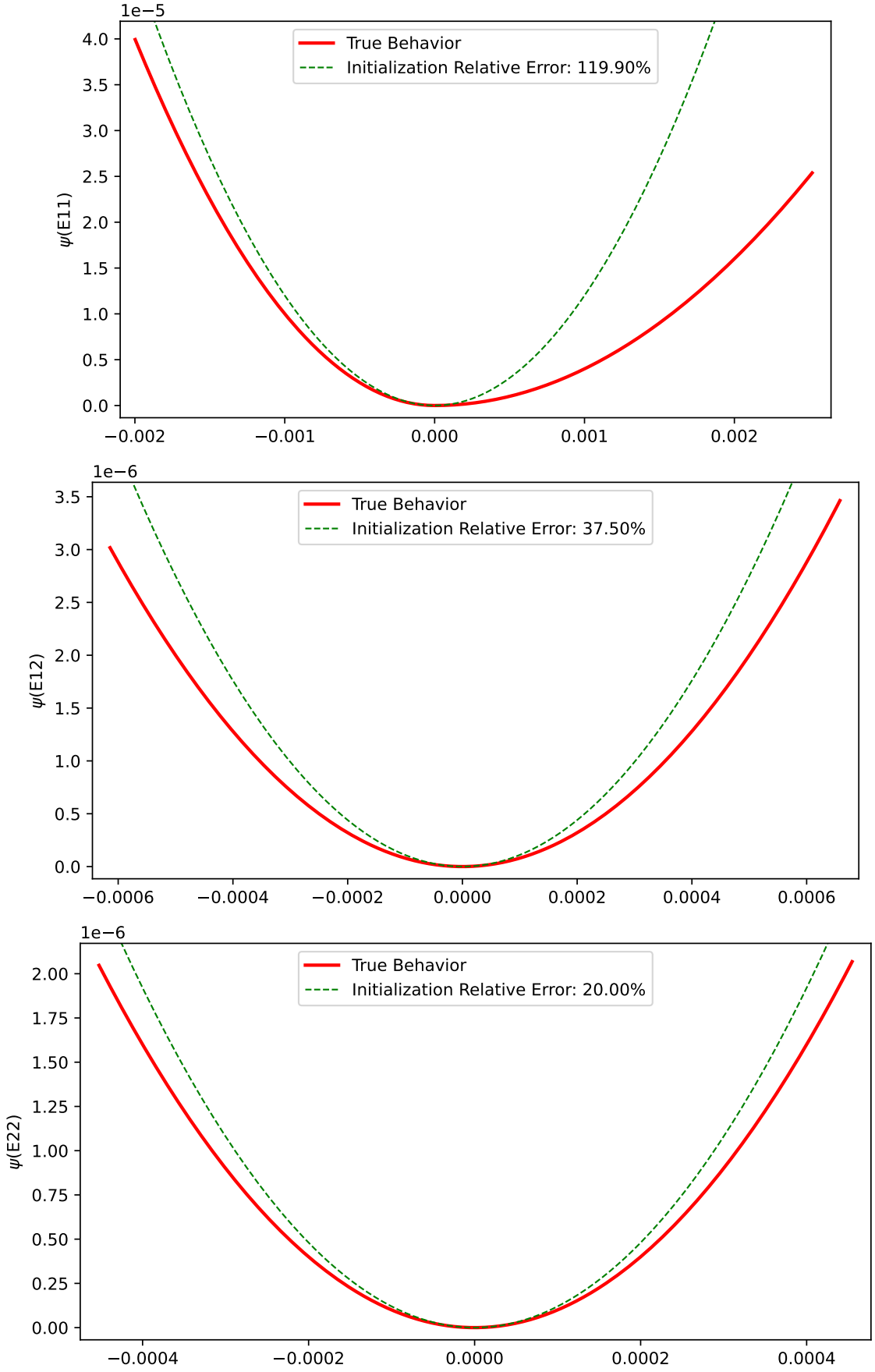


Figure 6: Initialization such that the NN-EUCLID and NN-mCRE ICNN represent linear elasticity

The configuration of hyper-parameters specific to each method are detailed in the following sections, as the settings change with respect to the noise level.

5.2.4 Results

Research phase for correct tuning of the ICNNs:
For the following comparison, the author has worked on the NN-EUCLID and NN-mCRE methods, respectively one after the other. For each method, the number of days required for correctly setting the neural networks was counted. Note that for each method, there are three networks to tune as three noise levels are studied.

Overall, for the NN-EUCLID method, the pre-training phase of the ICNNs required 15 days. The following strategy was adopted to establish satisfactory configurations. First, the search was purely random, the goal being to understand the trends of the training process with respect to different hyper-parameter setups. This search was not conclusive as there are too many parameters to define, each one having no reference point. Also, no clear trends were identified as the training process is very sensitive to any small changes in the settings. A second strategy consisted in simplifying as much as possible the numerical framework in which the study is carried out. The idea was to first consider a 2D square (instead of a pierced rectangle) with only 5 measurement points (instead of 874) and to tune an ICNN of small architecture accordingly. As this was achieved, the numerical framework was then slowly developed by adding measurement points and changing the geometry of the structure. For each change, the neural network was updated starting from the configuration of the previous stage. This iterative process was carried out until the numerical framework presented initially was reached. Note that this strategy was applied once in the case of zero noise. From this configuration, the settings for the case of low and high noise were deduced.

For the NN-mCRE method, the tuning of the main hyper-parameters being automated, the pre-training phase of the ICNNs required 3 days. Depending on the noise level, the search range for the scale factor α and the learning rate updating rule had to be adapted. The latter were successfully configured through trial and error.

The final settings for the NN-EUCLID and NN-mCRE methods are presented Table 1.

NN-EUCLID	Optimizer	Epochs	l_r scheduler	l_r base	l_r max	l_r step (up and down)
No Noise	Adam	700	cyclic	0.001	0.11	50
Low Noise	Adam	700	cyclic	0.001	0.05	50
High Noise	Adam	6000	cyclic	10^{-5}	0.0015	50

NN-mCRE	α search range	tolerance
No Noise	$[10^6, 10^9]$	2×10^{-4}
Low Noise	$[10^{-1}, 5 \times 10^2]$	2×10^{-4}
High Noise	$[10^{-2}, 5 \times 10^1]$	2×10^{-4}

Table 1: Tuned hyper-parameters for the NN-EUCLID and NN-mCRE methods

Note that, for the NN-EUCLID method, the main hyper-parameters to change depending on the test case are the number of epochs and the learning rate sweep range. For the NN-mCRE method, the learning rate updating rules are not presented as it is not relevant within the scope of this comparison.

Training phase of the ICNNs:
Table 2 presents the training time of the NN-EUCLID and NN-mCRE neural networks.

	No Noise	Low Noise	High Noise
NN-EUCLID	233 sec	236 sec	2012 sec
NN-mCRE	3257 sec	2662 sec	6425 sec

Table 2: Computational time for the training of the NN-EUCLID and NN-mCRE methods

As expected, the computational times are much faster for NN-EUCLID compared to NN-mCRE: regarding the no noise test case, the training time of NN-EUCLID is ten times faster than NN-mCRE.

Inference phase:
The trained models are tested with respect to the strain paths ($\mathbf{E}_{11}$, $\mathbf{E}_{12} = \mathbf{0}$, $\mathbf{E}_{22} = \mathbf{0}$), ($\mathbf{E}_{11} = \mathbf{0}$, $\mathbf{E}_{12}$, $\mathbf{E}_{22} = \mathbf{0}$) and ($\mathbf{E}_{11} = \mathbf{0}$, $\mathbf{E}_{12} = \mathbf{0}$, $\mathbf{E}_{22}$). Each path sweeps a range of values that is twice as big as the associated training database, the goal being to evaluate how each model generalizes when operating on new data. The predictions made by the NN-EUCLID and NN-mCRE models are presented Figure 7 for zero noise, Figure 8 for low noise and Figure 9 for high noise.

Considering NN-EUCLID for the zero noise test case, the model has perfectly learned the full spectrum of the material behavior. This is unexpected, as the training databases for $\mathbf{E}_{12}$ and $\mathbf{E}_{22}$ are five times smaller than the one for $\mathbf{E}_{11}$ (Figure 7). Hence, it was anticipated from the model to learn less regarding these paths.
For the low noise test case, the predictions are more aligned with what was expected. Indeed, the model perfectly learns the non-linear aspect of the material behavior but struggles to generalize within the $\mathbf{E}_{22}$ strain path.
Finally, for the high noise test case, the model is flawed because of the unreliable data it has trained on. Once again, it is important to note that the NN-EUCLID method does not take measurement noise into account during the training process. The strategy used in [40] to overcome this issue is to denoise the experimental data beforehand with the Kernel Ridge Regression (KRR) algorithm.

Considering NN-mCRE, the trained models are consistent with what was expected regardless of the noise level. Each model has perfectly learned the non-linearity aspect of the behavior. Also, the predictions along the $\mathbf{E}_{12}$ strain path are close to the truth. Finally, the models have not learned from the $\mathbf{E}_{22}$ strain path as the associated training database is negligeable compared to the others.

Overall, the results of all the above comparisons are gathered in the following table:

	Research-time for pre-training	Average learning time	Average relative error no noise	Average relative error low noise	Average relative error high noise
NN-EUCLID	15 days	827 sec	0.52%	12.83%	80.32%
NN-mCRE	3 days	4115 sec	7.27%	7.60%	9.89%

Table 3: Summary of the comparisons made between the NN-EUCLID and NN-mCRE methods

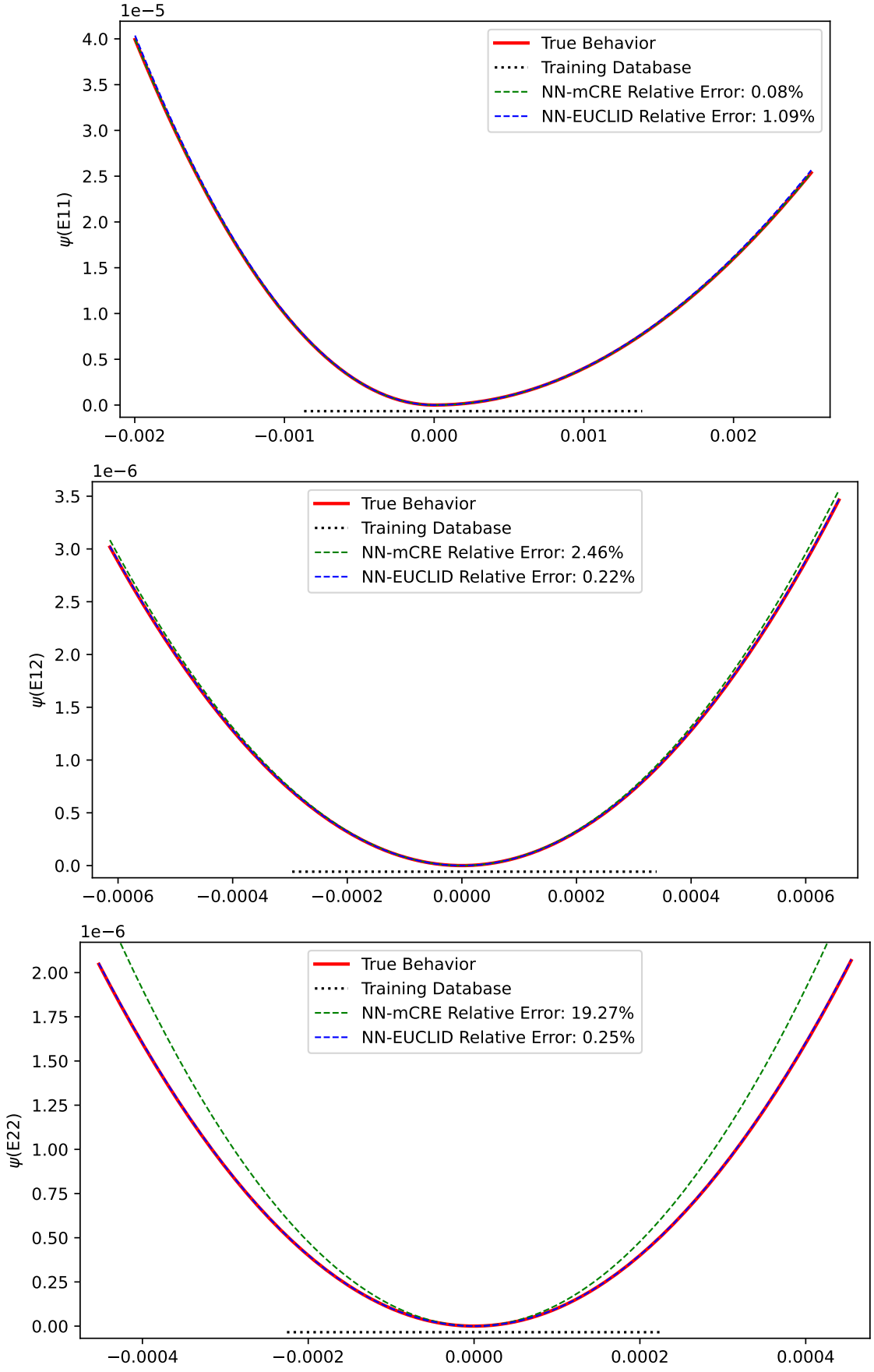


Figure 7: NN-EUCLID and NN-mCRE predictions for zero noise along the strain paths $(\mathbf{E}_{11}, \mathbf{E}_{12} = \mathbf{0}, \mathbf{E}_{22} = \mathbf{0})$, $(\mathbf{E}_{11} = \mathbf{0}, \mathbf{E}_{12}, \mathbf{E}_{22} = \mathbf{0})$ and $(\mathbf{E}_{11} = \mathbf{0}, \mathbf{E}_{12} = \mathbf{0}, \mathbf{E}_{22})$

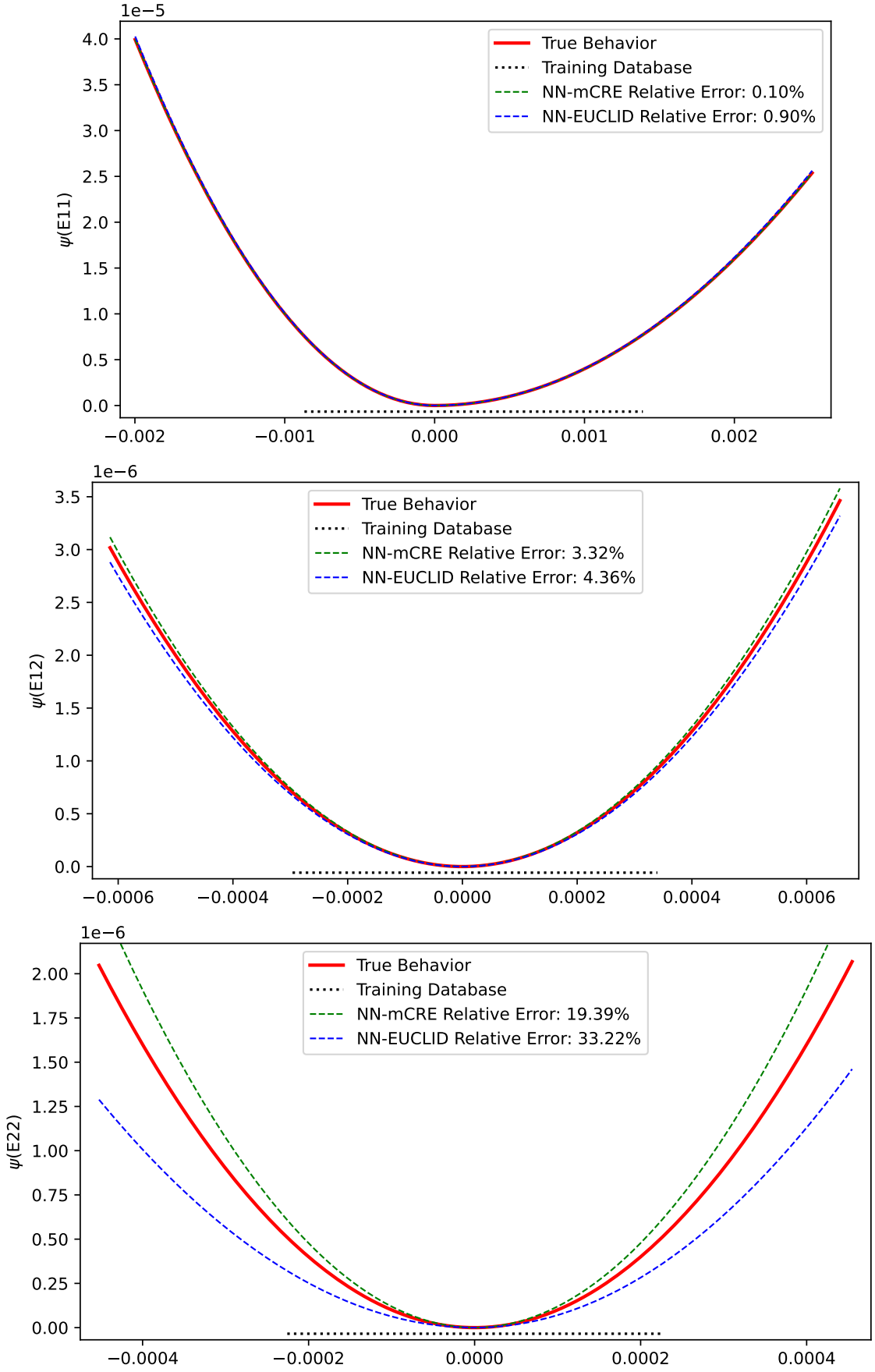


Figure 8: NN-EUCLID and NN-mCRE predictions for low noise along the strain paths $(\mathbf{E}_{11}, \mathbf{E}_{12} = \mathbf{0}, \mathbf{E}_{22} = \mathbf{0})$, $(\mathbf{E}_{11} = \mathbf{0}, \mathbf{E}_{12}, \mathbf{E}_{22} = \mathbf{0})$ and $(\mathbf{E}_{11} = \mathbf{0}, \mathbf{E}_{12} = \mathbf{0}, \mathbf{E}_{22})$

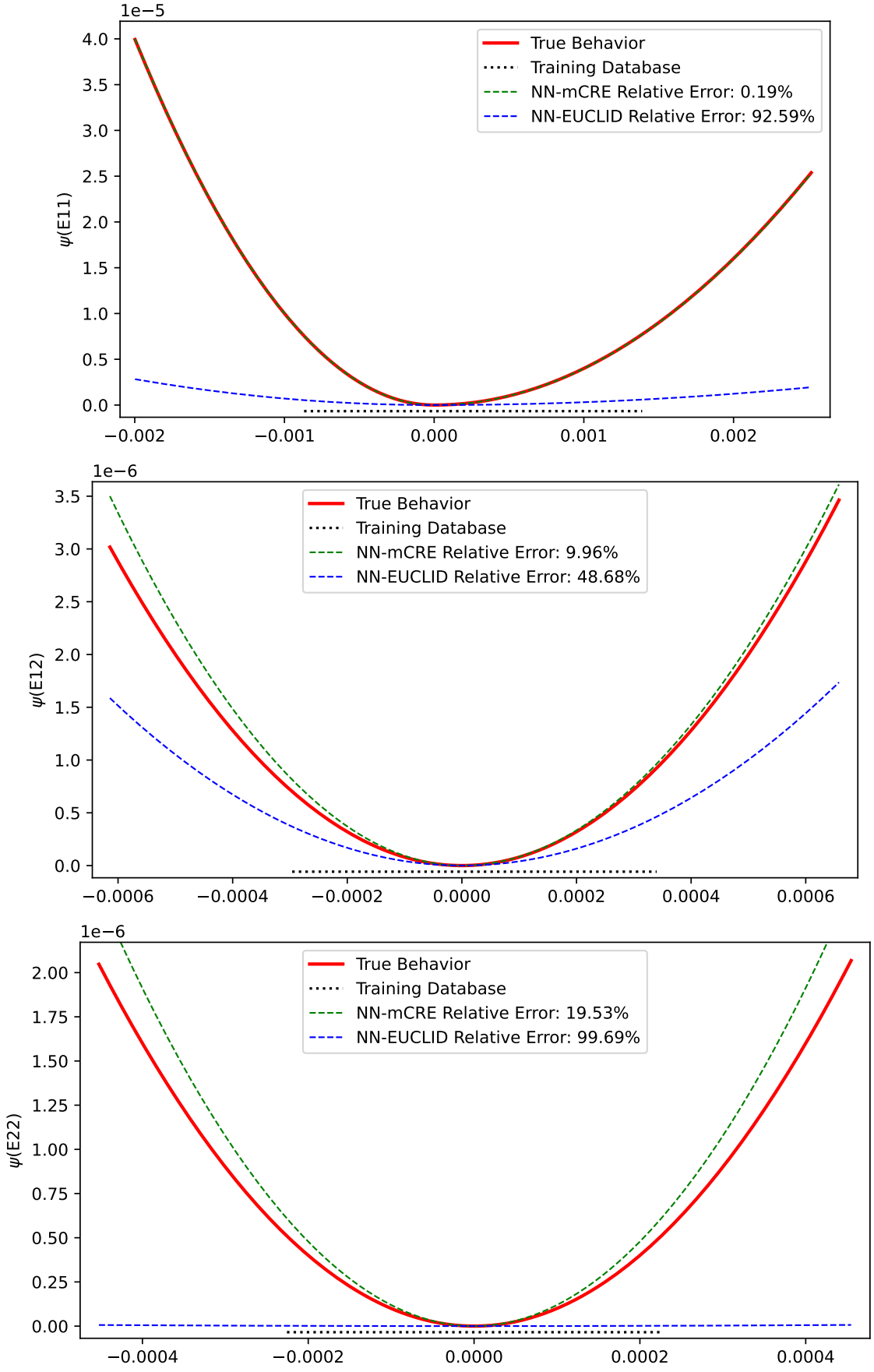


Figure 9: NN-EUCLID and NN-mCRE predictions for high noise along the strain paths $(\mathbf{E}_{11}, \mathbf{E}_{12} = \mathbf{0}, \mathbf{E}_{22} = \mathbf{0})$, $(\mathbf{E}_{11} = \mathbf{0}, \mathbf{E}_{12}, \mathbf{E}_{22} = \mathbf{0})$ and $(\mathbf{E}_{11} = \mathbf{0}, \mathbf{E}_{12} = \mathbf{0}, \mathbf{E}_{22})$

6 Conclusion

This study aimed to compare the NN-EUCLID and NN-mCRE methods which propose two different procedures for the unsupervised learning of unknown state laws with physics-augmented neural networks. Both approaches only require kinematic field measurements and known boundary conditions to train neural networks to recover the constitutive behavior. These strategies benefit from neural networks as universal approximators to relax the model form and suppress model bias.

To compensate the lack of stress labels and ensure consistent and relevant predictions, physical knowledge is embedded in the main stages of the training process. First, to reduce sensitivity regarding initialization, the strategy presented in [5] was applied and consisted in pre-training the neural network in a supervised manner with respect to *a priori* knowledge. In the context of this study, the networks used were pre-trained to represent a state law that was considered close from the one sought. Second, to satisfy thermodynamics, the architecture used is the one of an Input Convex Neural Network (ICNN), which ensures convexity of the output with respect to the input. Finally, physical knowledge is embedded in the loss function. The NN-EUCLID and NN-mCRE methods consider two distinct loss functions and hence propose very different training procedures.

Within the framework of this study, the attention was focused on understanding the conceptual differences between these methods.
First, being developed within different scopes, the two strategies do not rely on the same form of data. In the hyperelastic context, NN-EUCLID aims to make to most of the experimental information and hence benefits from full-field measurements enabled through DIC. In the DDDAS paradigm, NN-mCRE focuses on carrying out SHM on large structures, making it impossible to use DIC. Measurements are obtained through optic fibers, thus compelling NN-mCRE to operate on partial-field measurements.
Second, NN-EUCLID and NN-mCRE value different aspects of the training. The first method favors a fast training process over being able to evaluate the learned constitutive model, which, in the end, is what matters. The second method values the constitutive relation error at the expense of a computationally costly training phase.
This study also aimed to compare the NN-EUCLID and NN-mCRE methods on more practical criteria, such as the research-time for the tuning of hyper-parameters, the computational time of the training phase and the quality and robustness of predictions with respect to measurement noise. Overall, NN-EUCLID suffers from a long hyper-parameter tuning phase, as they are to be set through trial and error. The training phase is very fast but does not adapt to noisy measurements, which have to be denoised beforehand. Hence, for high noise levels, the predicted models are flawed. Considering NN-mCRE, the hyper-parameter tuning phase is rapid, as this process is automated to be in line with the DDDAS framework. This method does however suffer from a long training phase, especially as it depends on the tuning strategies employed and as admissible fields are computed for each epoch. Finally, NN-mCRE shows perfect adaptation to measurement noise, as the experimental data is relaxed in the mCRE functional.

Future developments include comparing the NN-EUCLID and NN-mCRE methods for other constitutive relations, like the one regarding isotropic and anisotropic hyperelasticity.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 101002857). The author was assisted by the NN-mCRE team.

A NN-EUCLID application in the 1D case

Problem setting:
The NN-EUCLID method is applied on a simple case (represented Figure 10), where a 1D beam is clamped at one end, while the other end is subjected to a quasi-static tension/compression force. The beam is of length L, of cross-section S and of Young modulus E. The test is controlled in terms of displacement: a displacement δ is imposed on the free end.

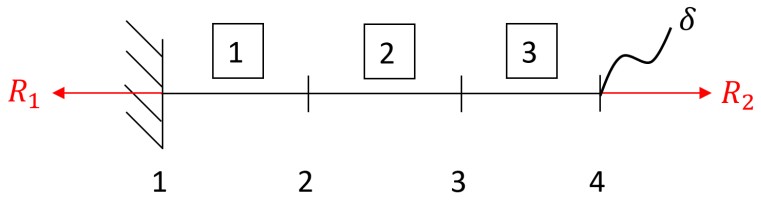


Figure 10: Study Case

The displacements u_a are measured at the points $a \in [\![1, n_n]\!]$ (here $n_n = 4$) and the reaction forces $(R_i)_{i=1,2}$ are measured at the ends. The study is carried out in the linear framework under the hypothesis of small perturbations. Thus, the material's behavior is governed by Hooke's law: $\sigma = E\epsilon$. The aim is therefore to determine the Young modulus using the NN-EUCLID method.

The first step is to determine the NN-EUCLID loss function based on the balance of forces. The latter is expressed as follows:

$$S \int_0^L \sigma(x) \times \frac{\partial v}{\partial x}(x) dx = 0 \quad \forall v \in \mathcal{U}_0^{ad} \tag{37}$$

Applying a Galerkin discretization using FEM shape functions, this system can be written as:

$$\sum_{a=1}^{n_n} v_a f_a \qquad \text{where} \qquad f_a = \begin{cases} ES \times \epsilon_1 & \text{for} \quad a = 1 \\ ES \times [\epsilon_{a-1} - \epsilon_a] & \forall a \in [\![2, n_n - 1]\!] \\ ES \times \epsilon_{n_n - 1} & \text{for} \quad a = n_n \end{cases} \tag{38}$$

Note that $(\epsilon_a = \frac{u_{a+1} - u_a}{\Delta x})_{a \in [\![1, n_n - 1]\!]}$ are the linearized strains defined at the Gauss point of each element.
Finally, the NN-EUCLID loss function is given by:

$$\mathcal{E}_{NN-EUCLID} = \sum_{a=2}^{n_n - 1} (f_a)^2 + (R_1 - f_1)^2 + (R_2 - f_{n_n})^2 \tag{39}$$

Overall, the identification process consists in minimizing this functional with respect to the Young modulus:

$$E_{opt} \leftarrow \underset{E}{argmin}[\mathcal{E}_{NN-EUCLID}] \tag{40}$$

The latter is performed via gradient descent: $E_{updated} = E - l_r \frac{\mathrm{d}\mathcal{E}_{NN-EUCLID}}{\mathrm{d}E}$

Implementation:
The NN-EUCLID method is implemented for the following parameters:

Surface	Length	Young modulus	δ	Number of elements	Learning rate	Number of epochs
25 mm^2	0.1 m	210 GPa	1 mm	50	10^{11}	1000

Table 4: Configurable parameters

The network is trained using synthetic data disturbed by artificial noise. The added noise is randomly generated according to a zero-centered normal distribution of standard deviation σ_u. Three noise levels are considered: a high ($\sigma_u = 10^{-3}$), a low ($\sigma_u = 10^{-4}$) and a null one ($\sigma_u = 0$). Note that these levels are normalized with respect to the displacement values.

Table 5 presents the learned Young moduli for the above experimental cases:

Target	No noise	Low noise	High noise
210 GPa	209.999 GPa	210.259 GPa	225.263 GPa

Table 5: Young modulus learned by the neural network for different noise levels

The trained models are then tested with respect to a strain path corresponding to traction. Figure 11 shows the predicted behaviors:

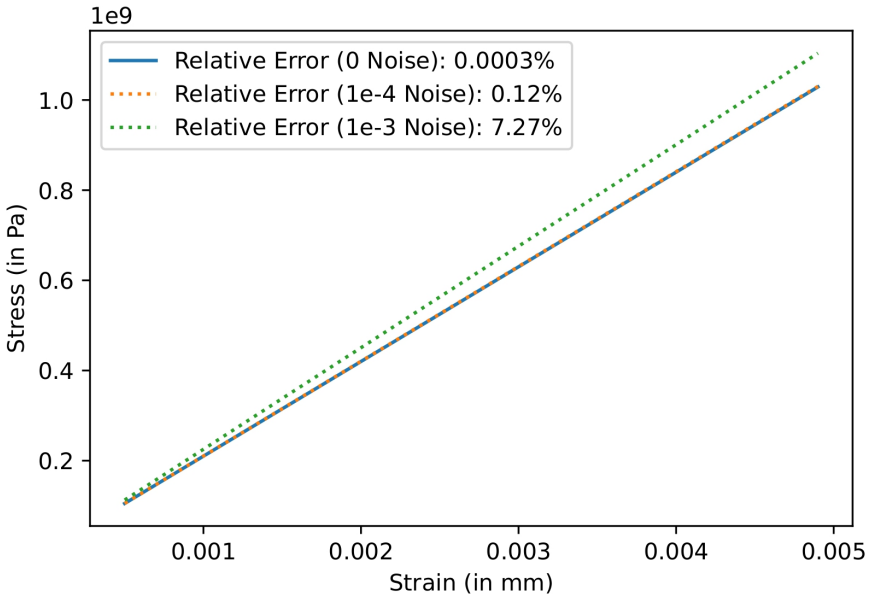


Figure 11: Stress prediction along the traction deformation path for no, low and high noise

B NN-mCRE application in the 1D case

Problem setting:
The NN-mCRE method is applied in a simple case (represented Figure 12), where a 1D beam is clamped at one end, while the other end is subjected to a quasi-static tension/compression force F. The beam is of length L, of cross-section S and of Young modulus E.

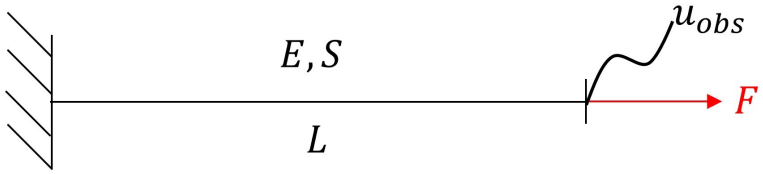


Figure 12: Study Case

The displacement u_{obs} is measured at the free end. The study is carried out in the linear framework under the hypothesis of small perturbations. Thus, the material's behavior is governed by Hooke's law: $\sigma = E\epsilon$. The aim is therefore to determine the Young modulus using the NN-mCRE method.

In a more general case, for an elasticity problem of stiffness matrix $\mathrm{K}(\mathbf{p})$, the mCRE functional is written such that:

$$\mathcal{E}^2_{mCRE}(\hat{\mathbf{u}}, \hat{\boldsymbol{\sigma}}, \mathbf{p}) = \int_\Omega (\hat{\boldsymbol{\sigma}} - \mathrm{K}(\mathbf{p})\mathbf{E}(\hat{\mathbf{u}})) : \mathrm{K}(\mathbf{p})^{-1} : (\hat{\boldsymbol{\sigma}} - \mathrm{K}(\mathbf{p})\mathbf{E}(\hat{\mathbf{u}})) \; + \; \alpha \parallel \Pi\hat{\mathbf{u}} - \mathbf{u}_{obs} \parallel^2 \quad (41)$$

The first step of the training process consists in finding the admissible fields $(\hat{\mathbf{u}}, \hat{\boldsymbol{\sigma}}) \in \mathcal{U}^{ad}_{\mathbf{u}_d} \times \mathcal{S}^{ad}$ for given parameters $\mathbf{p}$ by means of a constrained minimization using the following discretized Lagrangian functional $\mathcal{L}$:

$$\mathcal{L}(\hat{\mathbf{U}}, \hat{\mathbf{V}}, \boldsymbol{\Lambda}) = \frac{1}{2}(\hat{\mathbf{U}}-\hat{\mathbf{V}})^T\mathbb{K}(\mathbf{p})(\hat{\mathbf{U}}-\hat{\mathbf{V}})+\frac{\alpha}{2}(\Pi\hat{\mathbf{U}}-\mathbf{U}_{obs})^T\mathbb{G}_{obs}(\Pi\hat{\mathbf{U}}-\mathbf{U}_{obs})-\boldsymbol{\Lambda}^T(\mathbb{K}(\mathbf{p})\hat{\mathbf{V}}-\mathbf{F}) \quad (42)$$

where $\mathbb{K}(\mathbf{p})$ is the global stiffness matrix, $\mathbf{F}$ the global load vector, Π a projector, $\mathbb{G}_{obs}$ a scaling matrix, $\boldsymbol{\Lambda}$ the discretized form of the Lagrangian multiplier and $\hat{\mathbf{V}}$ the discretized form of $\hat{\mathbf{v}}$ such that $\hat{\boldsymbol{\sigma}} = \mathrm{K}(\mathbf{p})\mathbf{E}(\hat{\mathbf{v}})$.

Following the Karush-Kuhn-Tucker (KKT) conditions, the stationarity of the Lagrangian functional is obtained for:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \hat{\mathbf{U}}} = \mathbb{K}(\mathbf{p})(\hat{\mathbf{U}} - \hat{\mathbf{V}}) \; + \; \alpha\Pi^T(\Pi\hat{\mathbf{U}} - \mathbf{U}_{obs}) = \mathbf{0} \\ \frac{\partial \mathcal{L}}{\partial \hat{\mathbf{V}}} = \mathbb{K}(\mathbf{p})(\hat{\mathbf{V}} - \hat{\mathbf{U}}) \; - \; \mathbb{K}(\mathbf{p})\boldsymbol{\Lambda} = \mathbf{0} \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\Lambda}} = \mathbb{K}(\mathbf{p})\hat{\mathbf{V}} \; - \; \mathbf{F} = \mathbf{0} \end{cases} \quad (43)$$

Note that the above system is verified only for free degrees of freedom (DOFs). The prescribed ones are directly embedded in the search space.

In the end, the admissible fields are given by:

$$\begin{cases} \boldsymbol{\Lambda} = \hat{\mathbf{V}} - \hat{\mathbf{U}} \\ \mathbb{K}(\mathbf{p})\hat{\mathbf{V}} = \mathbf{F} \\ (\mathbb{K}(\mathbf{p}) + \alpha\Pi^T\Pi)\hat{\mathbf{U}} = \mathbf{F} + \alpha\Pi^T(\mathbf{U}_{obs} - \Pi\hat{\mathbf{U}}) \end{cases} \quad (44)$$

The second step involves updating the parameters $\mathbf{p}$ throughout a gradient based method using the above computed admissible fields. The gradient of the loss function with respect to the parameters is computed using the adjoint-state method:

$$\begin{aligned} \frac{\mathrm{d}\mathcal{E}^2_{mCRE}(\hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{p})}{\mathrm{d}\mathbf{p}} &= \frac{\partial \mathcal{L}}{\partial \mathbf{p}} \\ &= \frac{1}{2}(\hat{\mathbf{U}} - \hat{\mathbf{V}})^T \frac{\partial \mathbb{K}(\mathbf{P})}{\partial \mathbf{p}}(\hat{\mathbf{U}} + \hat{\mathbf{V}}) \end{aligned} \quad (45)$$

Implementation:
The NN-mCRE method is implemented for the following parameters:

Surface	Length	Young modulus	F	Number of elements	Learning rate	Number of epochs
25 mm^2	0.1 m	210 GPa	50 kN	50	10^{18}	8000

Table 6: Configurable parameters

The network is trained using synthetic data disturbed by artificial noise. This noise is implemented the same way as for Appendix A.

Table 7 presents the learned Young moduli for the above experimental cases. The scaling factors α used for the training process are also specified. However, it should be noted that these parameters were not optimally tuned as it is not essential in such a simple case. The intention here is merely to illustrate the trend α should follow regarding the noise level.

	Target	No noise	Low noise	High noise
α	-	10^5	100	1
E	210 GPa	209.987 GPa	209.979 GPa	209.777 GPa

Table 7: Young modulus learned by the neural network for different noise levels

The trained models are then tested with respect to a strain path corresponding to traction. Figure 13 shows the predicted behaviors:

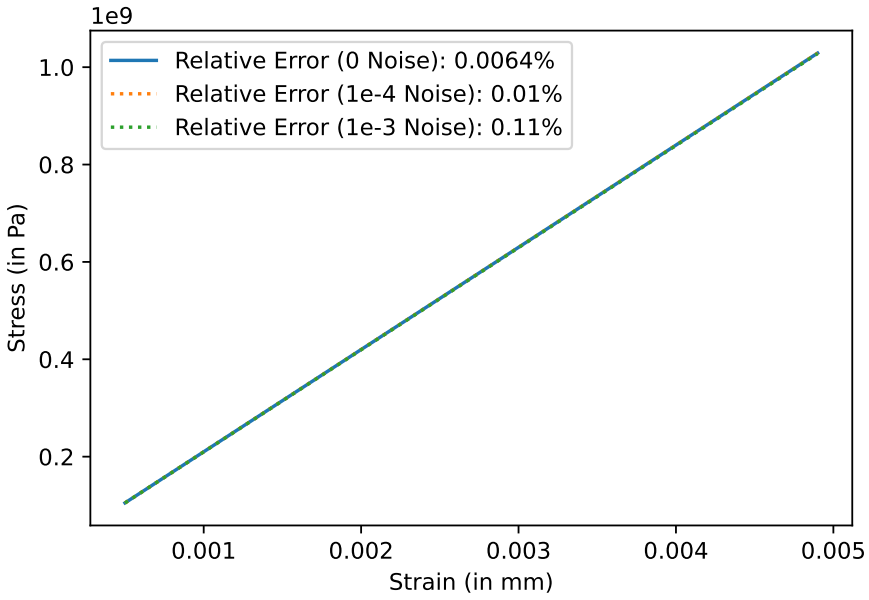


Figure 13: Stress prediction along the traction deformation path for no, low and high noise

References

[1] Olivier Allix, P Feissel, and P Thévenet. A delay damage mesomodel of laminates under dynamic loading: basic aspects and identification issues. *Computers & structures*, 81(12):1177–1191, 2003.

[2] Brandon Amos, Lei Xu, and J Zico Kolter. Input convex neural networks. In *International Conference on Machine Learning*, pages 146–155. PMLR, 2017.

[3] Faisal As' ad, Philip Avery, and Charbel Farhat. A mechanics-informed artificial neural network approach in data-driven constitutive modeling. *International Journal for Numerical Methods in Engineering*, 123(12):2738–2759, 2022.

[4] Stéphane Avril, Marc Bonnet, Anne-Sophie Bretelle, Michel Grédiac, François Hild, Patrick Ienny, Félix Latourte, Didier Lemosse, Stéphane Pagano, Emmanuel Pagnacco, et al. Overview of identification methods of mechanical parameters based on full-field measurements. *Experimental Mechanics*, 48:381–402, 2008.

[5] Antoine Benady, Emmanuel Baranger, and Ludovic Chamoin. Nn-mcre: a modified constitutive relation error framework for unsupervised learning of nonlinear state laws with physics-augmented neural networks. 2023.

[6] Pauline Böhringer, Daniel Sommer, Thomas Haase, Martin Barteczko, Joachim Sprave, Markus Stoll, Celalettin Karadogan, David Koch, Peter Middendorf, and Mathias Liewald. A strategy to train machine learning material models for finite element simulations on data acquirable from physical experiments. *Computer Methods in Applied Mechanics and Engineering*, 406:115894, 2023.

[7] Marc Bonnet and Wilkins Aquino. Three-dimensional transient elastodynamic inversion using an error in constitutive relation functional. *Inverse Problems*, 31(3):035010, 2015.

[8] Marc Bonnet and Andrei Constantinescu. Inverse problems in elasticity. *Inverse problems*, 21(2):R1, 2005.

[9] Sylvain Calloch, David Dureisseix, and François Hild. Identification de modèles de comportement de matériaux solides: utilisation d'essais et de calculs. *Technologies et Formations*, 100:36–41, 2002.

[10] Ludovic Chamoin, Antoine Benady, Sahar Farahbakhsh, Emmanuel Baranger, and Martin Poncelet. Data-based model updating, selection, and enrichment using the modified constitutive relation error concept. In *15th World Congress on Computational Mechanics*, 2022.

[11] Amar T Chouaki, Pierre Ladeveze, and Laurent Proslier. Updating structural dynamic models with emphasis on the damping properties. *AIAA journal*, 36(6):1094–1099, 1998.

[12] N Cottin, H-P Felgenhauer, and HG Natke. On the parameter identification of elastomechanical systems using input and output residuals. *Archive of Applied Mechanics*, 54(5):378–387, 1984.

[13] Vincent Decouvreur, Ph Bouillard, Arnaud Deraemaeker, and Pierre Ladevèze. Updating 2d acoustic models with the constitutive relation error method. *Journal of sound and vibration*, 278(4-5):773–787, 2004.

[14] Sahar Farahbakhsh, Ludovic Chamoin, and Martin Poncelet. Model updating with a modified dual kalman filter acting on distributed strain measurements. In *XI International*

Conference on Adaptive Modeling and Simulation (ADMOS 2023), 2023.

[15] Sahar Farahbakhsh, Martin Poncelet, and Ludovic Chamoin. Structural health monitoring and model updating with distributed optic fiber measurements. In *17th International Conference on Advances in Experimental Mechanics*, 2023.

[16] Charbel Farhat and Francois M Hemez. Updating finite element dynamic models using an element-by-element sensitivity methodology. *AIAA journal*, 31(9):1702–1711, 1993.

[17] Moritz Flaschel, Siddhant Kumar, and Laura De Lorenzis. Unsupervised discovery of interpretable hyperelastic constitutive laws. *Computer Methods in Applied Mechanics and Engineering*, 381:113852, 2021.

[18] Moritz Flaschel, Siddhant Kumar, and Laura De Lorenzis. Discovering plasticity models without stress data. *npj Computational Materials*, 8(1):91, 2022.

[19] Jan Niklas Fuhg, Craig M Hamel, Kyle Johnson, Reese Jones, and Nikolaos Bouklas. Modular machine learning-based elastoplasticity: generalization in the context of limited data. *Computer Methods in Applied Mechanics and Engineering*, 407:115930, 2023.

[20] Jamshid Ghaboussi, JH Garrett Jr, and Xiping Wu. Knowledge-based modeling of material behavior with neural networks. *Journal of engineering mechanics*, 117(1):132–153, 1991.

[21] Shyamal Guchhait and Biswanath Banerjee. Constitutive error based parameter estimation technique for plate structures using free vibration signatures. *Journal of Sound and Vibration*, 419:302–317, 2018.

[22] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.

[23] Shaojuan Huang, Pierre Feissel, and Pierre Villon. Modified constitutive relation error: An identification framework dealing with the reliability of information. *Computer Methods in Applied Mechanics and Engineering*, 311:1–17, 2016.

[24] Akshay Joshi, Prakash Thakolkaran, Yiwen Zheng, Maxime Escande, Moritz Flaschel, Laura De Lorenzis, and Siddhant Kumar. Bayesian-euclid: Discovering hyperelastic material laws with uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 398:115225, 2022.

[25] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[26] Dominik K Klein, Mauricio Fernández, Robert J Martin, Patrizio Neff, and Oliver Weeger. Polyconvex anisotropic hyperelasticity with neural networks. *Journal of the Mechanics and Physics of Solids*, 159:104703, 2022.

[27] Robert V Kohn and Bruce D Lowe. A variational method for parameter identification. *ESAIM: Mathematical Modelling and Numerical Analysis*, 22(1):119–158, 1988.

[28] Pierre Ladeveze and Dominique Leguillon. Error estimate procedure in the finite element method and applications. *SIAM Journal on Numerical Analysis*, 20(3):485–509, 1983.

[29] Pierre Ladevèze, Djamel Nedjar, and Marie Reynier. Updating of finite element models using vibration tests. *AIAA journal*, 32(7):1485–1491, 1994.

[30] Pierre Ladevèze and Jean-Pierre Pelle. *Mastering calculations in linear and nonlinear mechanics*, volume 171. Springer, 2005.

[31] LF Li and CQ Chen. Equilibrium-based convolution neural networks for constitutive modeling of hyperelastic materials. *Journal of the Mechanics and Physics of Solids*, 164:104931, 2022.

[32] Nuno MM Maia, Marie Reynier, Pierre Ladeveze, and ENS Cachan. Error localization for updating finite element models using frequency-response-functions. In *PROCEEDINGS-SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING*, pages 1299–1299. SPIE INTERNATIONAL SOCIETY FOR OPTICAL, 1994.

[33] Basile Marchand, Ludovic Chamoin, and Christian Rey. Parameter identification and model updating in the context of nonlinear mechanical behaviors using a unified formulation of the modified constitutive relation error concept. *Computer Methods in Applied Mechanics and Engineering*, 345:1094–1113, 2019.

[34] Vladimir Alekseevich Morozov. The error principle in the solution of operational equations by the regularization method. *USSR Computational Mathematics and Mathematical Physics*, 8(2):63–87, 1968.

[35] E Pagnacco, D Lemosse, François Hild, and Fabien Amiot. Inverse strategy from displacement field measurement and distributed forces using fea. In *2005 SEM annual conference and exposition on experimental and applied mechanics, Portland*, pages 7–9, 2005.

[36] B Pan, K Li, and W Tong. Fast, robust and accurate digital image correlation calculation without redundant computations. *Experimental Mechanics*, 53:1277–1289, 2013.

[37] Maziar Raissi, Paris Perdikaris, and George 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.

[38] Ricardo Ribani and Mauricio Marengoni. A survey of transfer learning for convolutional neural networks. In *2019 32nd SIBGRAPI conference on graphics, patterns and images tutorials (SIBGRAPI-T)*, pages 47–57. IEEE, 2019.

[39] Craig Saunders, Alexander Gammerman, and Volodya Vovk. Ridge regression learning algorithm in dual variables. 1998.

[40] Prakash Thakolkaran, Akshay Joshi, Yiwen Zheng, Moritz Flaschel, Laura De Lorenzis, and Siddhant Kumar. Nn-euclid: Deep-learning hyperelasticity without stress data. *Journal of the Mechanics and Physics of Solids*, 169:105076, 2022.

[41] Zvonimir Tomičevć, Francois Hild, and Stephane Roux. Mechanics-aided digital image correlation. *The Journal of Strain Analysis for Engineering Design*, 48(5):330–343, 2013.

[42] Nikolaos N Vlassis and WaiChing Sun. Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. *Computer Methods in Applied Mechanics and Engineering*, 377:113695, 2021.

[43] James E Warner, Manuel I Diaz, Wilkins Aquino, and Marc Bonnet. Inverse material identification in coupled acoustic-structure interaction using a modified error in constitutive equation functional. *Computational mechanics*, 54:645–659, 2014.

[44] Jared Willard, Xiaowei Jia, Shaoming Xu, Michael Steinbach, and Vipin Kumar. Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Computing Surveys*, 55(4):1–37, 2022.