



MOX-Report No. 06/2025

**Enhancing Bayesian model updating in structural health monitoring  
via learnable mappings**

Torzoni, M.; Manzoni, A.; Mariani, A.

MOX, Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

[mox-dmat@polimi.it](mailto:mox-dmat@polimi.it)

<https://mox.polimi.it>

# Enhancing Bayesian model updating in structural health monitoring via learnable mappings

Matteo Torzoni<sup>a</sup>, Andrea Manzoni<sup>b</sup>, Stefano Mariani<sup>a</sup>

<sup>a</sup>*Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133, Milano, Italy*

<sup>b</sup>*MOX – Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133, Milano, Italy*

---

## Abstract

In the context of structural health monitoring (SHM), the selection and extraction of damage-sensitive features from raw sensor recordings represent a critical step towards solving the inverse problem underlying the structural health identification. This work introduces a new way to enhance stochastic approaches to SHM through the use of deep neural networks. A learnable feature extractor and a feature-oriented surrogate model are synergistically exploited to evaluate a likelihood function within a Markov chain Monte Carlo sampling algorithm. The feature extractor undergoes a supervised pairwise training to map sensor recordings onto a low-dimensional metric space, which encapsulates the sensitivity to structural health parameters. The surrogate model maps the structural health parameters onto their feature description. The procedure enables the updating of beliefs about structural health parameters, effectively replacing the need for a computationally expensive numerical (finite element) model. A preliminary offline phase involves the generation of a labeled dataset to train both the feature extractor and the surrogate model. Within a simulation-based SHM framework, training vibration responses are cost-effectively generated by means of a multi-fidelity surrogate modeling strategy to approximate sensor recordings under varying damage and operational conditions. The multi-fidelity surrogate exploits model order reduction and artificial neural networks to speed up the data generation phase while ensuring the damage-sensitivity of the approximated signals. The proposed strategy is assessed through three synthetic case studies, demonstrating remarkable results in terms of accuracy of the estimated quantities and computational efficiency.

---

## 1. Introduction

The safety of civil structural systems is a key challenge of our society. This is daily threatened by material deterioration, cyclic and extraordinary loading conditions, and more and more by effects triggered by the climate change, such as anomalous heat waves and destructive storms [1]. Since the lifecycle (economic, social and safety) costs entailed by such structural systems may be extremely high, enabling a condition-based maintenance approach in place of time-based ones is nowadays critical [2, 3]. To this aim, non-destructive tests and in situ inspections are not suitable to implement a continuous and automated global monitoring; on the other hand, by assimilating vibration response data acquired with permanently installed data collecting systems [4, 5], vibration-based structural health monitoring (SHM) techniques allow for damage identification and evolution tracking.

Data-driven approaches to SHM [6, 7, 8] rely on a pattern recognition paradigm [9] involving the following steps: (i) operational evaluation; (ii) data acquisition; (iii) feature selection and extraction; (iv) statistical modeling to unveil the relationship between the selected features and sought damage patterns [10, 11]. In this process, the selection of synthetic and informative features is the most critical step, as it requires problem-specific knowledge subject to the available expertise. To this aim, deep learning (DL) represents a promising solution to automatize the selection and extraction of features optimized for the task at hand.

Within a different strategy, Bayesian model-based approaches to SHM [12, 13, 14, 15] assess damage from a parameter estimation perspective, through a model updating strategy. Such a probabilistic framework has the advantage of naturally dealing with the ill-posedness of the SHM problem, and allows to account for and quantify uncertainty due to, e.g., measurement noise, modeling assumptions, and environmental and operational variability.

In this paper, we propose a hybrid approach to SHM leveraging on the strengths of both data-driven and model-based approaches. Learnable features, optimized for the structure to be monitored, are automatically selected and extracted by a DL-based feature extractor. The feature extractor maps the input vibration recordings onto their feature representation in a low-dimensional space, and relies on an autoencoder architecture useful to perform a dimensionality reduction of the input data. During training, the autoencoder is equipped with a Siamese appendix [16] of the encoder, optimized through a pairwise contrastive learning strategy [17, 18]. Such a deep metric learning [19, 20] strategy enables learning a distance function that conforms to a task-specific definition of similarity, so that the neighbors of a data point are mapped closer than non-neighbors in the learned metric space [21]. The resulting mapping encodes the sensitivity to the sought parameters according to the chosen metric, thereby enabling a manifold to suitably describe the parametric space underlying the processed measurements. The extracted features are exploited within a Markov chain Monte Carlo (MCMC) algorithm [22, 23, 24], to address the estimation of parameters describing the variability of the structural system. The likelihood function underlying the MCMC sampler is evaluated by means of a feature-oriented surrogate model, to map the parameters that need to be updated onto the corresponding feature representation.

The proposed strategy takes advantage of a preliminary offline learning phase. The training of the feature extractor and the feature-oriented surrogate model is carried out in a supervised fashion. Labeled data pertaining to specific damage conditions are generated in an inexpensive way through a multi-fidelity (MF) surrogate modeling strategy. In this work, such a MF surrogate modeling is chosen as an effective strategy to reduce the computational cost, while ensuring the accuracy of the approximated signals in terms of damage-sensitivity. The vibration response data required to fit the MF surrogate are generated by physics-based numerical simulations, so that the effect of damage on the structural response can be systematically reproduced.

A graphical abstraction of the proposed framework is reported in Fig. 1. Vibration responses of different fidelity levels are simulated offline using physics-based full/reduced-order numerical models, similarly to [25, 26]. These data are then exploited to train a MF surrogate model, following the strategy proposed in [27]. Once trained, the MF surrogate model is employed to provide an arbitrarily large training dataset. This dataset is used to train the deep-metric-learning-based feature extractor, following a strategy similar to that proposed in [6], and the surrogate model, employed to approximate the functional link between the parameters to be updated and the low-dimensional feature space. During the online monitoring phase, the trained feature extractor and the surrogate model are eventually exploited by an MCMC sampling algorithm to update the prior belief about the structural state.

The elements of novelty that characterize this work are the following. First, the assimilation of data related to vibration responses is carried out by exploiting DL models, which allow the automatic selection and extraction of optimized features from raw vibration recordings. Second, the employed low-dimensional feature space benefits from a geometrical structure, which encodes the sensitivity to the parameters to be updated. The resulting MCMC framework enjoys: a competitive computational cost due to the low dimensionality of the involved features; fast convergence due to the geometrical structure characterizing the feature space; accurate estimates due to the informativeness of the extracted features.

The remainder of the paper is organized as follows. In Sec. 2, we review the MF surrogate modeling strategy that we employ for dataset population purposes. In Sec. 3, we describe the proposed parameter estimation framework. In Sec. 4, the computational procedure is assessed on three test cases, respectively related to a cantilever beam, a portal frame, and a railway bridge. Conclusions and future developments are finally drawn in Sec. 5.

## 2. Population of training datasets

In this section, we describe how the population of training datasets is performed with reference to the simulation-based paradigm of SHM. The composition of the handled vibration responses is specified in Sec. 2.1. The numerical models underlying the generation of labeled data pertaining to specific damage conditions are described in Sec. 2.2. The MF surrogate modeling strategy employed to populate large training datasets is reviewed in Sec. 2.3.

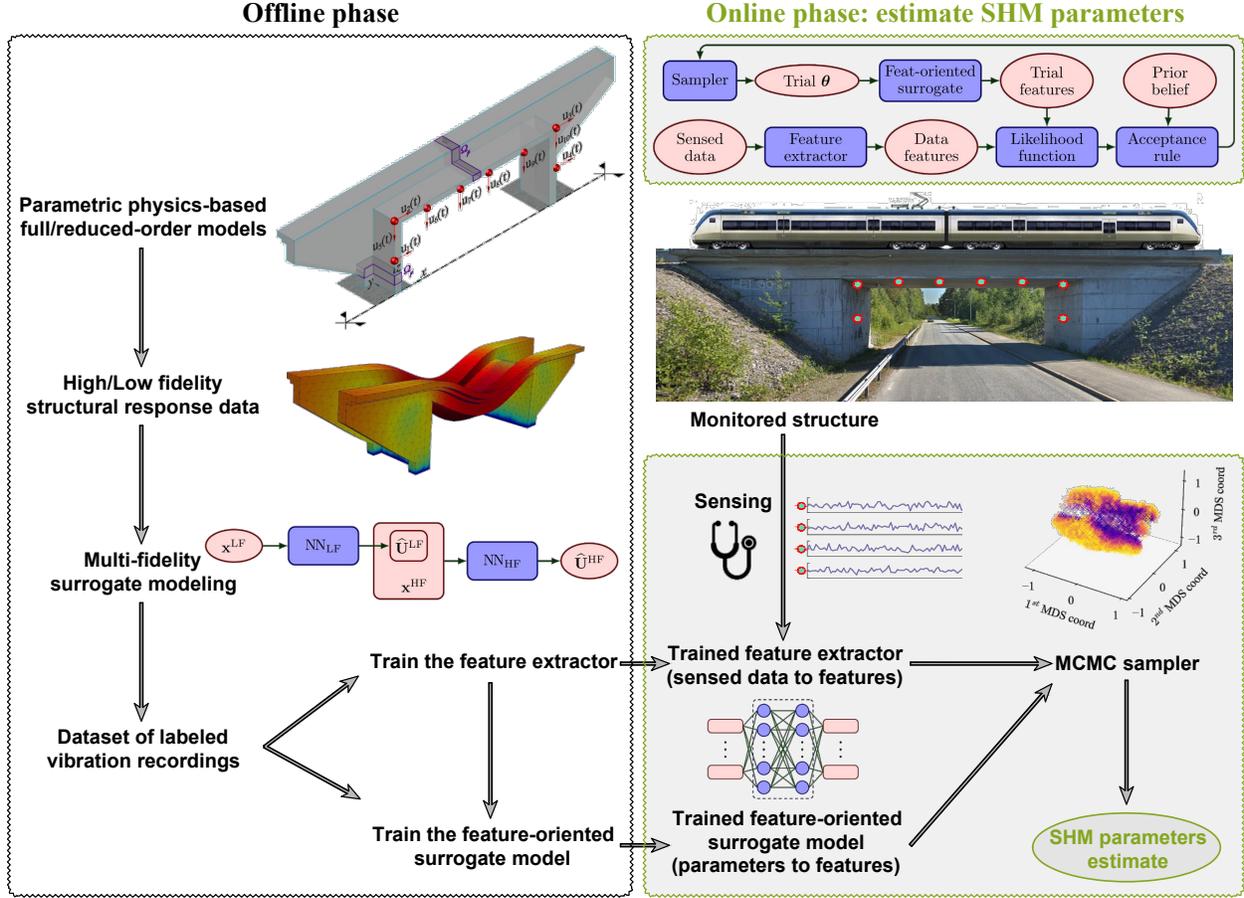


Figure 1. Graphical abstraction of the proposed methodology.

### 2.1. Data specification

The monitoring of structural systems relies on the assimilation of vibration recordings shaped as multivariate time series  $\mathbf{U}^{\text{EXP}}(\boldsymbol{\theta}) = [\mathbf{u}_1^{\text{EXP}}(\boldsymbol{\theta}), \dots, \mathbf{u}_{N_u}^{\text{EXP}}(\boldsymbol{\theta})] \in \mathbb{R}^{L \times N_u}$ , consisting of  $N_u$  series, each one consisting of  $L$  measurements equally spaced in time. For instance, measurements can be provided as accelerations or displacements at structural nodes. The vector  $\boldsymbol{\theta} \in \mathbb{R}^{N_{\text{par}}}$  comprises  $N_{\text{par}}$  parameters, representing the variability of the monitored system in terms of structural health and, possibly, operational conditions, for which we seek to update the relative belief. Each recording refers to a time interval  $(0, T)$ , within which measurements are recorded with a sampling rate  $f_s$ .

For the problem setting we consider herewith, the time interval  $(0, T)$  is assumed short enough for the operational, environmental, and damage conditions to be considered time-invariant, yet long enough to not compromise the identification of the structural behavior.

### 2.2. Low/high fidelity physics-based models

The labeled dataset required to train the feature extractor and the feature-oriented surrogate model is populated by exploiting the MF surrogate modeling strategy proposed in [27]. The resulting surrogate model relies on a composition of deep neural network (DNN) models and is therefore termed MF-DNN. The MF surrogate model is trained on synthetic data, generated by means of physics-based models. In this section, we describe the models employed to systematically reproduce the effect of damage on the structural response, while the MF-DNN surrogate model is reviewed in Sec. 2.3.

The chosen physics-based numerical models are: a low-fidelity (LF) reduced-order model (ROM), obtained by relying on a proper orthogonal decomposition (POD)-Galerkin reduced basis method for parametrized

finite element models [28, 29, 25, 26]; and a high-fidelity (HF) finite element model. The two models are employed to simulate the structural responses under varying operational conditions, respectively in the absence or in the presence of a structural damage. In particular, LF data are generated by always referring to a baseline condition, while HF data have to account for potential degradation processes. Thanks to this modeling choice, it is never necessary to update the LF component, and whenever a deterioration of the structural health is detected, the MF surrogate can be updated by adjusting only its HF component. Without loss of generality, in the following we will refer to the initial monitoring phase of an undamaged reference condition, see also [4].

The HF model describes the dynamic response of the monitored structure to the applied loadings, under the assumption of a linearized kinematics. By modeling the structure as a linear-elastic continuum, and by discretizing it in space through finite elements, the HF model consists of the following semi-discretized form:

$$\begin{cases} \mathbf{M}_{\text{HF}}\ddot{\mathbf{d}}^{\text{HF}}(t) + \mathbf{C}_{\text{HF}}(\mathbf{x}^{\text{HF}})\dot{\mathbf{d}}^{\text{HF}}(t) + \mathbf{K}_{\text{HF}}(\mathbf{x}^{\text{HF}})\mathbf{d}^{\text{HF}}(t) = \mathbf{f}_{\text{HF}}(t, \mathbf{x}^{\text{HF}}), & t \in (0, T) \\ \mathbf{d}^{\text{HF}}(0) = \mathbf{d}_0^{\text{HF}} \\ \dot{\mathbf{d}}^{\text{HF}}(0) = \dot{\mathbf{d}}_0^{\text{HF}}, \end{cases} \quad (1)$$

which is referred to as the HF full-order model (FOM). In problem (1):  $t \in (0, T)$  denotes time;  $\mathbf{d}^{\text{HF}}(t)$ ,  $\dot{\mathbf{d}}^{\text{HF}}(t)$ ,  $\ddot{\mathbf{d}}^{\text{HF}}(t) \in \mathbb{R}^{N_{\text{FE}}}$  are the vectors of nodal displacements, velocities and accelerations, respectively;  $N_{\text{FE}}$  is the number of degrees of freedom (dofs);  $\mathbf{M}_{\text{HF}} \in \mathbb{R}^{N_{\text{FE}} \times N_{\text{FE}}}$  is the mass matrix;  $\mathbf{C}_{\text{HF}}(\mathbf{x}^{\text{HF}}) \in \mathbb{R}^{N_{\text{FE}} \times N_{\text{FE}}}$  is the damping matrix, assembled according to the Rayleigh's model;  $\mathbf{K}_{\text{HF}}(\mathbf{x}^{\text{HF}}) \in \mathbb{R}^{N_{\text{FE}} \times N_{\text{FE}}}$  is the stiffness matrix;  $\mathbf{f}_{\text{HF}}(t, \mathbf{x}^{\text{HF}}) \in \mathbb{R}^{N_{\text{FE}}}$  is the vector of nodal forces induced by the external loadings;  $\mathbf{x}^{\text{HF}} \in \mathbb{R}^{N_{\text{par}}^{\text{HF}}}$  is a vector of  $N_{\text{par}}^{\text{HF}}$  input parameters ruling the operational, damage and (possibly) environmental conditions, such that  $\boldsymbol{\theta} \subseteq \mathbf{x}^{\text{HF}}$ ;  $\mathbf{d}_0^{\text{HF}}$  and  $\dot{\mathbf{d}}_0^{\text{HF}}$  are the initial conditions at  $t = 0$ , respectively in terms of nodal displacements and velocities. The solution of problem (1) is advanced in time using an implicit Newmark integration scheme (constant average acceleration method).

With reference to civil structures, we focus on the early detection of damage patterns characterized by a small evolution rate, whose prompt identification can reduce lifecycle costs and increase the safety and availability of the structure. In this context, structural damage is often modeled as a localized reduction of the material stiffness [30, 31, 32], that is here obtained by means of a suitable parametrization of the stiffness matrix. In practical terms, we parametrize a damage condition through its position  $\boldsymbol{y} \in \mathbb{R}^3$  and magnitude  $\delta \in \mathbb{R}$ , both included in the parameter vector  $\mathbf{x}^{\text{HF}}$ .

The POD-based LF model approximates the solution to problem (1) through  $\mathbf{d}^{\text{LF}}(t, \mathbf{x}^{\text{LF}}) \approx \mathbf{W}\mathbf{r}(t, \mathbf{x}^{\text{LF}})$ , where  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_{N_{\text{RB}}}] \in \mathbb{R}^{N_{\text{FE}} \times N_{\text{RB}}}$  is a basis matrix featuring  $N_{\text{RB}} \ll N_{\text{FE}}$  POD basis functions as columns, and  $\mathbf{r}(t, \mathbf{x}^{\text{LF}}) \in \mathbb{R}^{N_{\text{RB}}}$  is the vector of unknown POD coefficients. The approximation is provided for a given vector of LF parameters  $\mathbf{x}^{\text{LF}} \in \mathbb{R}^{N_{\text{par}}^{\text{LF}}}$ , collecting  $N_{\text{par}}^{\text{LF}}$  parameters that rule the operational conditions undergone by the structure, with  $N_{\text{par}}^{\text{LF}} < N_{\text{par}}^{\text{HF}}$ . By enforcing the orthogonality between the residual and the subspace spanned by the first  $N_{\text{RB}}$  POD modes through a Galerkin projection, the following  $N_{\text{RB}}$ -dimensional semi-discretized form is obtained:

$$\begin{cases} \mathbf{M}_r \ddot{\mathbf{r}}(t) + \mathbf{C}_r \dot{\mathbf{r}}(t) + \mathbf{K}_r \mathbf{r}(t) = \mathbf{f}_r(t, \mathbf{x}^{\text{LF}}), & t \in (0, T) \\ \mathbf{r}(0) = \mathbf{W}^\top \mathbf{d}_0^{\text{LF}} \\ \dot{\mathbf{r}}(0) = \mathbf{W}^\top \dot{\mathbf{d}}_0^{\text{LF}}. \end{cases} \quad (2)$$

The solution of this low-dimensional dynamical system is advanced in time using the same strategy employed for the HF model, and then projected onto the original LF-FOM space as  $\mathbf{d}^{\text{LF}}(t, \mathbf{x}^{\text{LF}}) \approx \mathbf{W}\mathbf{r}(t, \mathbf{x}^{\text{LF}})$ . Here, the reduced-order matrices  $\mathbf{M}_r$ ,  $\mathbf{C}_r$ , and  $\mathbf{K}_r$ , and vector  $\mathbf{f}_r$  play the same role of their HF counterparts, yet with dimension  $N_{\text{RB}} \times N_{\text{RB}}$  instead of  $N_{\text{FE}} \times N_{\text{FE}}$ , and read:

$$\begin{aligned} \mathbf{M}_r &\equiv \mathbf{W}^\top \mathbf{M}_{\text{HF}} \mathbf{W}, & \mathbf{C}_r &\equiv \mathbf{W}^\top \mathbf{C}_{\text{HF}} \mathbf{W}, \\ \mathbf{K}_r &\equiv \mathbf{W}^\top \mathbf{K}_{\text{LF}} \mathbf{W}, & \mathbf{f}_r(t, \mathbf{x}^{\text{LF}}) &\equiv \mathbf{W}^\top \mathbf{f}_{\text{HF}}(t, \mathbf{x}^{\text{LF}}). \end{aligned} \quad (3)$$

The matrix  $\mathbf{W}$  is obtained by exploiting the so-called method of snapshots as follows [33, 34, 35]. First, a LF-FOM, resembling that defined by problem (1) but not accounting for the presence of damage, is employed to assemble a snapshot matrix  $\mathbf{S} = [\mathbf{d}_1^{\text{LF}}, \dots, \mathbf{d}_{N_{\text{S}}}^{\text{LF}}] \in \mathbb{R}^{N_{\text{FE}} \times N_{\text{S}}}$  from  $N_{\text{S}}$  solution snapshots, computed by integrating in time the LF-FOM for different values of parameters  $\mathbf{x}^{\text{LF}}$ . The computation of an optimal

reduced basis is then carried out by factorizing  $\mathbf{S}$  through a singular value decomposition. We use a standard energy-based criterion to set the order  $N_{\text{RB}}$  of the approximation; for further details, see [25, 26, 27, 6].

To populate the LF and HF datasets, respectively denoted as  $\mathbf{D}_{\text{LF}}$  and  $\mathbf{D}_{\text{HF}}$ , the parametric spaces of vectors  $\mathbf{x}^{\text{LF}}$  and  $\mathbf{x}^{\text{HF}}$  are taken as uniformly distributed, and then sampled via the latin hypercube rule. Although this is not a restrictive choice, the number of samples is equal to the number  $I_{\text{LF}}$  and  $I_{\text{HF}}$ , with  $I_{\text{LF}} > I_{\text{HF}}$ , of instances collected in  $\mathbf{D}_{\text{LF}}$  and  $\mathbf{D}_{\text{HF}}$ , respectively, as:

$$\mathbf{D}_{\text{LF}} = \{(\mathbf{x}_i^{\text{LF}}, \mathbf{U}_i^{\text{LF}})\}_{i=1}^{I_{\text{LF}}}, \quad \mathbf{D}_{\text{HF}} = \{(\mathbf{x}_j^{\text{HF}}, \mathbf{U}_j^{\text{HF}})\}_{j=1}^{I_{\text{HF}}}, \quad (4)$$

where the LF and HF vibration recordings  $\mathbf{U}_i^{\text{LF}}(\mathbf{x}_i^{\text{LF}}) = [\mathbf{u}_1^{\text{LF}}(\mathbf{x}_i^{\text{LF}}), \dots, \mathbf{u}_{N_u}^{\text{LF}}(\mathbf{x}_i^{\text{LF}})]_i \in \mathbb{R}^{L \times N_u}$  and  $\mathbf{U}_j^{\text{HF}}(\mathbf{x}_j^{\text{HF}}) = [\mathbf{u}_1^{\text{HF}}(\mathbf{x}_j^{\text{HF}}), \dots, \mathbf{u}_{N_u}^{\text{HF}}(\mathbf{x}_j^{\text{HF}})]_j \in \mathbb{R}^{L \times N_u}$ , are labeled by the corresponding  $i$ -th sampling of  $\mathbf{x}^{\text{LF}}$  and  $j$ -th sampling of  $\mathbf{x}^{\text{HF}}$ , respectively, and are obtained as detailed in the following. By dropping indices  $i$  and  $j$  for ease of notation and with reference to displacement recordings, nodal values in  $(0, T)$  are first collected as  $\mathbf{V}_{\text{LF}} = [\mathbf{W}\mathbf{r}_1, \dots, \mathbf{W}\mathbf{r}_L] \in \mathbb{R}^{N_{\text{FE}} \times L}$  and  $\mathbf{V}_{\text{HF}} = [\mathbf{d}_1^{\text{HF}}, \dots, \mathbf{d}_L^{\text{HF}}] \in \mathbb{R}^{N_{\text{FE}} \times L}$ , by solving problem (2) and problem (1), respectively. The relevant vibration recordings  $\mathbf{U}^{\text{LF}}$  and  $\mathbf{U}^{\text{HF}}$  are then obtained as:

$$\mathbf{U}^{\text{LF}} = (\mathbf{T}\mathbf{V}_{\text{LF}})^\top, \quad \mathbf{U}^{\text{HF}} = (\mathbf{T}\mathbf{V}_{\text{HF}})^\top, \quad (5)$$

where  $\mathbf{T} \in \mathbb{B}^{N_u \times N_{\text{FE}}}$  is a Boolean matrix whose  $(n, m)$ -th entry is equal to 1 only if the  $n$ -th sensor output coincides with the  $m$ -th dof. For the problem setting we consider, the sampling frequency  $f_s$ , and the number  $N_u$  and location of the monitored dofs are supposed to be the same for both fidelity levels. However, there are no restrictions in this regard, and LF and HF data with different dimensions can be equally considered. Moreover, we note that the matrix product  $\mathbf{T}\mathbf{W} \in \mathbb{R}^{N_u \times N_{\text{RB}}}$  can be computed, once and for all, to extract  $\mathbf{U}^{\text{LF}}$  for any given set of LF input parameters  $\mathbf{x}^{\text{LF}}$ .

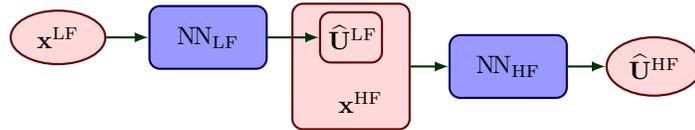
### 2.3. Multi-fidelity surrogate modeling for structural health monitoring

We review now the MF-DNN surrogate modeling strategy proposed in [27], which is here employed to generate data pertaining to specific damage conditions in an inexpensive way. The generated data will serve to carry out the foreseen training of the feature extractor and of the feature-oriented surrogate. The employed surrogate modeling strategy falls into the wider framework of MF methods, see for instance [36, 37, 38]. These methods are characterized by the use of multiple models with varying accuracy and computational cost. By blending LF and HF models, MF methods allow for improved approximation accuracy compared to the LF solution, while carrying a lower computational burden than the HF solver. Indeed, LF samples often supply useful information on the major trends of the problem, allowing the MF setting to outperform single-fidelity methods in terms of prediction accuracy and computational efficiency. In addition, MF surrogate models based on DNNs enjoy several appealing features: they are suitable for high-dimensional problems and benefit from large LF training datasets, provide real-time predictions, can deal with linear and nonlinear correlations in an adaptive fashion without requiring prior information, and can handle the approximation of strongly discontinuous trajectories.

Our MF-DNN surrogate model is devised to map damage and operational parameters onto sensor recordings. It leverages on an LF part and an HF part, sequentially trained, and respectively denoted by  $\text{NN}_{\text{LF}}$  and  $\text{NN}_{\text{HF}}$ . The resulting surrogate model reads as:

$$\text{NN}_{\text{MF}}(\mathbf{x}^{\text{HF}}, \mathbf{x}^{\text{LF}}) = \text{NN}_{\text{HF}}(\mathbf{x}^{\text{HF}}) \circ \text{NN}_{\text{LF}}(\mathbf{x}^{\text{LF}}), \quad (6)$$

where  $\circ$  stands for function composition, see Fig. 2.

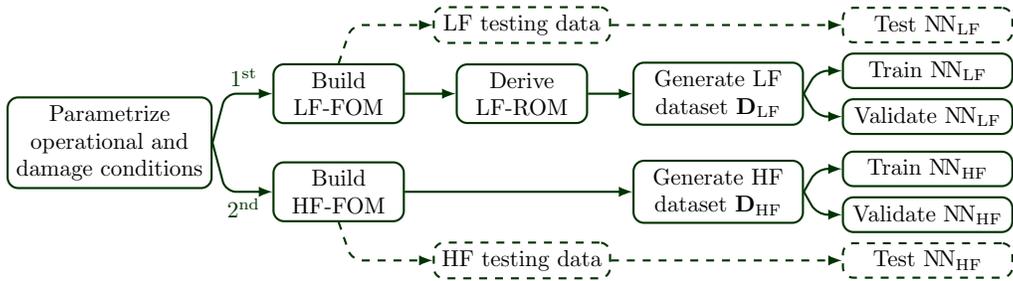


**Figure 2.** Scheme of the MF-DNN surrogate model: red nodes denote the input/output quantities, while blue nodes refer to the learnable components of the surrogate model; hat variables denote quantities obtained from neural network approximations. Figure adapted from [27].

$\text{NN}_{\text{LF}}$  is set as a fully-connected DL model, exploited to approximate the LF vibration recordings for any given set of LF input parameters  $\mathbf{x}^{\text{LF}}$ . In particular,  $\text{NN}_{\text{LF}}$  provides an approximation to a set of POD coefficients encoding  $\mathbf{U}^{\text{LF}}$ , allowing the number of trainable parameters of  $\text{NN}_{\text{LF}}$  to be largely reduced.

$\text{NN}_{\text{HF}}$  is a DNN built upon the long short-term memory (LSTM) model, useful to exploit the time correlation between the two fidelity levels. Indeed, an LSTM model for  $\text{NN}_{\text{HF}}$  can exploit the temporal structure of the LF signals  $\widehat{\mathbf{U}}^{\text{LF}}$  provided through  $\text{NN}_{\text{LF}}$ . At each time step,  $\text{NN}_{\text{HF}}$  takes the HF input parameters  $\mathbf{x}^{\text{HF}}$ , the current time instant  $t$ , and the corresponding LF approximation  $\widehat{\mathbf{U}}_t^{\text{LF}}$ , to enrich the latter with the effects of damage and provide the HF approximation  $\widehat{\mathbf{U}}^{\text{HF}}(t)$ .

The main steps involved in our MF-DNN surrogate modeling strategy are outlined in Fig. 3 and consist of: the definition of a parametric LF-FOM; the construction of a parametric LF-ROM by means of POD; the population of  $\mathbf{D}_{\text{LF}}$  with LF vibration recordings at sensor locations via LF-ROM simulations; the training and validation of the LF component  $\text{NN}_{\text{LF}}$ , employed to approximate  $\mathbf{U}^{\text{LF}}$  for any given  $\mathbf{x}^{\text{LF}}$ ; the testing of the generalization capabilities of  $\text{NN}_{\text{LF}}$  on LF-FOM data; the definition of a parametric HF structural model accounting for the effects of damage; the population of  $\mathbf{D}_{\text{HF}}$  via HF-FOM simulations; the training and validation of the HF component  $\text{NN}_{\text{HF}}$ , employed to enrich the  $\widehat{\mathbf{U}}^{\text{LF}}$  approximation with the effects of damage for any given  $\mathbf{x}^{\text{HF}}$ ; the testing of the generalization capabilities of  $\text{NN}_{\text{MF}}$ . For the interested reader, the detailed steps of our MF-DNN surrogate modeling strategy are reported in [27].



**Figure 3.** Flowchart of the MF-DNN surrogate modeling strategy. Figure adapted from [27].

The key feature of  $\text{NN}_{\text{MF}}$  is that the effect of damage on the structural response is reproduced with the HF model only, which is considered to be the most accurate description enabling to account for unexperienced damage scenarios. The  $\text{NN}_{\text{MF}}$  training is carried out offline once and for all, and is characterized by a limited number of evaluations of the HF finite element solver. At the same time, the computational time required to evaluate  $\text{NN}_{\text{MF}}$  for new input parameters is negligible. This latter aspect enables to greatly speed up the generation of a large number of training instances, compared to what would be required by relying solely on the HF finite element solver. Finally, it is worth noting that the MF-DNN surrogate modeling paradigm can be easily adapted to application domains other than SHM, even in the case of a different number of fidelity levels, and potentially extended to handle full-field approximation or feature-based data.

The trained MF-DNN surrogate model is eventually exploited to populate a large labeled dataset  $\mathbf{D}_{\text{train}}$ , according to:

$$\mathbf{D}_{\text{train}} = \{(\mathbf{x}_k^{\text{HF}}, \widehat{\mathbf{U}}_k^{\text{HF}} = \text{NN}_{\text{MF}}(\mathbf{x}_k^{\text{HF}}, \mathbf{x}_k^{\text{LF}}))\}_{k=1}^{I_{\text{train}}}, \quad (7)$$

where  $I_{\text{train}}$  is the number of instances collected in  $\mathbf{D}_{\text{train}}$ . These instances are provided through  $\text{NN}_{\text{MF}}$  for varying input parameters  $\mathbf{x}^{\text{HF}}$  (with  $\mathbf{x}^{\text{LF}}$  being a subset of them) sampled via the latin hypercube rule. In order to mimic measurement noise, each vibration recording in  $\mathbf{D}_{\text{train}}$  is then corrupted by adding an independent, identically distributed Gaussian noise, whose statistical properties depend on the target accuracy of the sensors.

### 3. Deep learning-enhanced Bayesian model updating

In this section, we describe the proposed methodology to enhance an MCMC algorithm for model updating purposes through learnable mappings. The key components are a learnable feature extractor, which extracts informative features from the sensed structural response, and a feature-oriented surrogate model, which maps the  $\boldsymbol{\theta}$  parameters to be updated onto the low-dimensional feature space. Both the feature extractor

and the feature-oriented surrogate model rely on DL models. These models are trained by exploiting the  $\mathbf{D}_{\text{train}}$  dataset, populated through the MF-DNN surrogate model described above. The architectures of the two models and the technical aspects related to their training and evaluation are discussed in Sec. 3.1. In Sec. 3.2, we explain how the feature extractor and the feature-oriented surrogate model are employed to sample the posterior distribution of  $\boldsymbol{\theta}$  conditioned on observational data.

### 3.1. Feature extractor and feature-oriented surrogate: models specification and training

In what follows, we describe the models and the relevant training process underlying the feature extractor and the feature-oriented surrogate. Before training, the synthetic data generated through the MF-DNN surrogate model and collected in  $\mathbf{D}_{\text{train}}$  are preprocessed to be transformed into images as described below. We remark that this is not a restrictive choice; indeed, the proposed methodology is general and can be easily adapted to deal with data of different nature.

The recent developments in computer vision suggest the possibility of transforming time series onto images for SHM purposes, see for instance [39, 40, 41]. Imaging time series is reported to help highlighting local patterns that might otherwise be spread over or laying outside the time domain. In particular, the Markov transition field (MTF) technique [42] is here employed to preprocess the multivariate time histories collected in  $\mathbf{D}_{\text{train}}$ . The MTF technique is chosen over other conversion methods, such as the Gramian angular fields [42], the recurrence plots [43] and the grey-scale encoding [44], as it has been reported to offer better performance for SHM purposes [40, 41]. However, the MTF is a signal processing algorithm not employed in the practice as frequently as those based on spectral analysis, such as the spectrogram or scalogram representations. The MTF technique is reviewed in Appendix A.

Each instance  $\widehat{\mathbf{U}}_k^{\text{HF}}$ , with  $k = 1, \dots, I_{\text{train}}$ , is transformed into a grey-scale mosaic  $\mathcal{J}_k \in \mathbb{R}^{h_{\mathcal{J}} \times w_{\mathcal{J}}}$ , with  $h_{\mathcal{J}}$  and  $w_{\mathcal{J}}$  respectively being the height and the width of the mosaic. Each mosaic is composed of the juxtaposition of  $N_u$  MTF representations, or tesserae, obtained via MTF encoding of the  $N_u$  time series collected in  $\widehat{\mathbf{U}}_k^{\text{HF}}$ . Accordingly,  $\mathbf{D}_{\text{train}}$  is reassembled as:

$$\mathbf{D}_{\text{train}}^{\mathcal{J}} = \{(\mathbf{x}_k^{\text{HF}}, \mathcal{J}_k)\}_{k=1}^{I_{\text{train}}} . \quad (8)$$

The feature extractor and the feature-oriented surrogate model are learned through a sequential training process involving two learning steps. (see Fig. 4). A first learning step involves training the feature extractor to map structural response data onto their feature representation in a low-dimensional space. A second learning step involves training the surrogate model to map the parametric space that needs be updated onto the low-dimensional feature space. Once trained, the two components are exploited within an MCMC algorithm to sample the posterior distribution of  $\boldsymbol{\theta}$  conditioned on observational data, as detailed next.

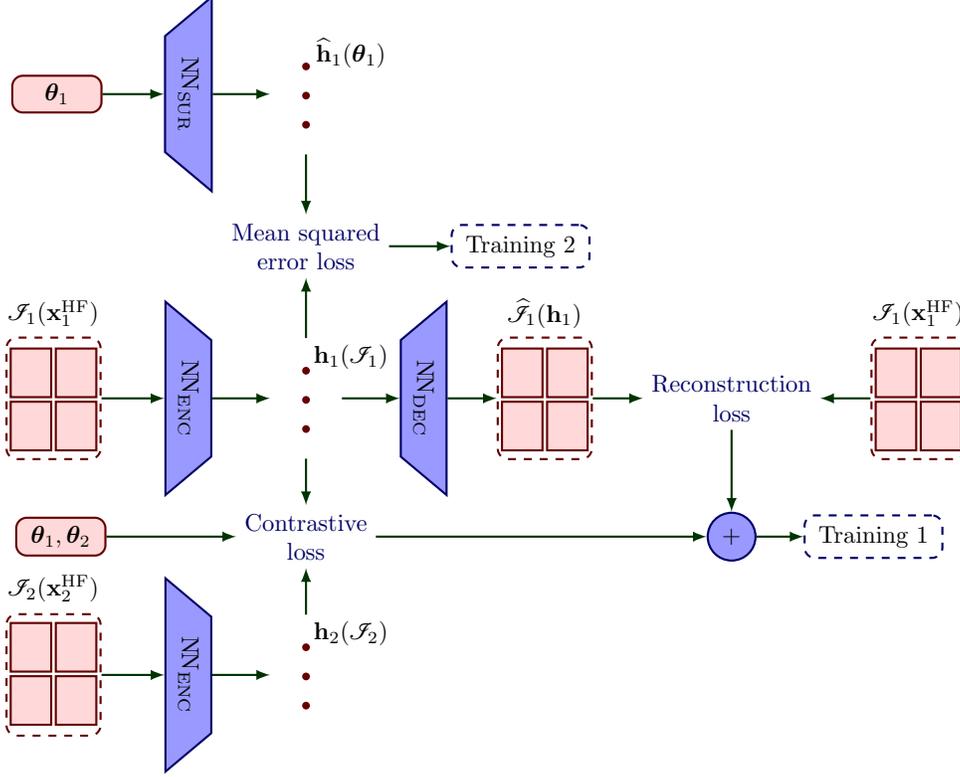
The feature extractor is built upon an autoencoder equipped with a Siamese appendix [16] of the encoder branch (refer to ‘‘Training 1’’ in Fig. 4). This model enhances the dimensionality reduction capabilities provided by the unsupervised training of an autoencoder, by enabling a distance function for the relative latent space through pairwise contrastive learning [18]. Within the resulting latent space, features extracted from similar data points are pushed to be as close as possible, while those provided for dissimilar data points are kept away. The concept of similarity refers to a task-specific distance measure, in terms of the  $\boldsymbol{\theta}$  parameters describing the variability of the monitored system.

The learnable components involved in the training of the feature extractor are the encoder  $\text{NN}_{\text{ENC}}$  and decoder  $\text{NN}_{\text{DEC}}$  branches of an autoencoder.  $\text{NN}_{\text{ENC}}$  provides the feature representation  $\mathbf{h} \in \mathbb{R}^{D_h}$  of the input mosaic  $\mathcal{J}$  in a low-dimensional space of size  $D_h$ , while  $\text{NN}_{\text{DEC}}$  takes  $\mathbf{h}$  and provides the reconstructed mosaic  $\widehat{\mathcal{J}}$ , as follows:

$$\mathbf{h}(\mathcal{J}) = \text{NN}_{\text{ENC}}(\mathcal{J}(\mathbf{x}^{\text{HF}})) , \quad (9)$$

$$\widehat{\mathcal{J}}(\mathbf{h}) = \text{NN}_{\text{DEC}}(\mathbf{h}(\mathcal{J})) . \quad (10)$$

The key component that links  $\text{NN}_{\text{ENC}}$  and  $\text{NN}_{\text{DEC}}$  is the bottleneck layer characterized by the low-dimensional feature size  $D_h$ .  $D_h$  is much smaller than the dimension of the input and output layers of the autoencoder, thus forcing the data through a compressed representation while attempting to recreate the input as closely as possible at the output. The unsupervised training of an autoencoder is a well-known procedure in the literature, see for instance [45]. On the other hand, the Siamese appendix of the encoder branch affects the



**Figure 4.** Learnable feature extractor and feature-oriented surrogate: flowchart of the sequential training process. Red nodes refer to the input/output quantities, while blue nodes denote the relevant computational blocks.  $\text{NN}_{\text{ENC}}$  is the feature extractor,  $\text{NN}_{\text{DEC}}$  is the decoder branch, and  $\text{NN}_{\text{SUR}}$  is the feature-oriented surrogate model.  $\mathcal{S}(\mathbf{x}^{\text{HF}})$  denotes an input mosaic,  $\hat{\mathcal{S}}(\mathbf{h})$  denotes a reconstructed mosaic, and  $\boldsymbol{\theta}$  is the vector of parameters for which we seek to update the relative belief.  $\mathbf{h}(\mathcal{S})$  is the low-dimensional feature representation of  $\mathcal{S}(\mathbf{x}^{\text{HF}})$  provided by  $\text{NN}_{\text{ENC}}$ , and  $\hat{\mathbf{h}}(\boldsymbol{\theta})$  is the corresponding approximation provided by  $\text{NN}_{\text{SUR}}$ .

training process through a contrastive loss function linking two twins  $\text{NN}_{\text{ENC}}$ . Data points are thus processed in pairs, yielding two outputs  $\mathbf{h}_1 = \text{NN}_{\text{ENC}}(\mathcal{S}_1(\mathbf{x}_1^{\text{HF}}))$  and  $\mathbf{h}_2 = \text{NN}_{\text{ENC}}(\mathcal{S}_2(\mathbf{x}_2^{\text{HF}}))$ . The required data pairing process is carried out as follows. First, a threshold distance  $\bar{\mathcal{E}}_\theta$  is fixed to characterize the similarity for the parametric space of  $\boldsymbol{\theta}$ . The mosaics dataset  $\mathbf{D}_{\text{train}}^{\mathcal{S}}$  is then augmented by assembling  $\zeta_+$  positive pairs for each instance, characterized by  $\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2 \leq \bar{\mathcal{E}}_\theta$ , and  $\zeta_-$  negative pairs, characterized by  $\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2 > \bar{\mathcal{E}}_\theta$ , according to:

$$\mathbf{D}_{\text{P}}^{\mathcal{S}} = \{(\mathbf{x}_1^{\text{HF}}, \mathcal{S}_1, \mathbf{x}_2^{\text{HF}}, \mathcal{S}_2)\}_{\ell=1}^{I_{\text{train}}^{\text{P}}}, \quad (11)$$

with  $I_{\text{train}}^{\text{P}} = I_{\text{train}}(\zeta_+ + \zeta_-)$  being the total number of pairs.

The set of weights and biases parametrizing the autoencoder is denoted as  $\boldsymbol{\Omega}_{\text{AE}}$ . During “Training 1”, this is optimized by minimizing the following loss function over  $\mathbf{D}_{\text{P}}^{\mathcal{S}}$ :

$$\mathcal{L}_{\text{AE}}(\boldsymbol{\Omega}_{\text{AE}}, \mathbf{D}_{\text{P}}^{\mathcal{S}}) = \frac{1}{I_{\text{train}}^{\text{P}}} \sum_{\ell=1}^{I_{\text{train}}^{\text{P}}} \left\{ \|\mathcal{S}_1(\mathbf{x}_1^{\text{HF}}) - \text{NN}_{\text{DEC}}(\text{NN}_{\text{ENC}}(\mathcal{S}_1(\mathbf{x}_1^{\text{HF}})))\|_2^2 + \left[ \frac{1-\gamma}{2} (\mathcal{E}_h)^2 + \frac{\gamma}{2} [\max(0, \psi - \mathcal{E}_h)]^2 \right] \right\} + \lambda_{\text{AE}} \|\boldsymbol{\Omega}_{\text{AE}}\|_2^2, \quad (12)$$

where: the first term is the reconstruction loss function, typically employed to train autoencoders; the second term is the pairwise contrastive loss function, useful to induce a geometrical structure in the feature space; and the last term is an  $L^2$  regularization of rate  $\lambda_{\text{AE}}$  applied over the model parameters  $\boldsymbol{\Omega}_{\text{AE}}$ . In Eq. (12);  $\gamma = \{0, 1\}$ , if  $\boldsymbol{\theta}_1$  and  $\boldsymbol{\theta}_2$  identify either a positive or a negative pair, respectively;  $\psi > 0$  is a margin beyond which negative pairs do not contribute to  $\mathcal{L}_{\text{AE}}$ ;  $\mathcal{E}_h = \|\mathbf{h}_1 - \mathbf{h}_2\|_2$  is the Euclidean distance between any pair

of mappings  $\mathbf{h}_1 = \text{NN}_{\text{ENC}}(\mathcal{J}_1(\mathbf{x}_1^{\text{HF}}))$  and  $\mathbf{h}_2 = \text{NN}_{\text{ENC}}(\mathcal{J}_2(\mathbf{x}_2^{\text{HF}}))$ . Minimizing the contrastive loss function is equivalent to learning a distance function  $\mathcal{E}_h$  that approximates, at least semantically, the Euclidean distance  $\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2$  between the target labels  $\boldsymbol{\theta}_1$  and  $\boldsymbol{\theta}_2$  of the processed pair of data points. The labels information is thus exploited to guide the dimensionality reduction, so that the sensitivity to damage and (possibly) operational conditions described via  $\boldsymbol{\theta}$  is encoded in the low-dimensional feature space.

After the first learning step,  $\text{NN}_{\text{DEC}}$ , the Siamese appendix, and  $\mathbf{D}_{\text{train}}^{\mathcal{J}}$  are discarded, and only  $\text{NN}_{\text{ENC}}$  and  $\mathbf{D}_{\text{train}}^{\mathcal{J}}$  are retained to train the feature-oriented surrogate  $\text{NN}_{\text{SUR}}$  (refer to ‘‘Training 2’’ in Fig. 4).  $\text{NN}_{\text{SUR}}$  is set as a fully-connected DL model, and it approximates the functional link between the parametric space of  $\boldsymbol{\theta}$  and the low-dimensional feature space described by  $\text{NN}_{\text{ENC}}$  as follows:

$$\hat{\mathbf{h}} = \text{NN}_{\text{SUR}}(\boldsymbol{\theta}) , \quad (13)$$

where  $\hat{\mathbf{h}}$  denotes the  $\text{NN}_{\text{SUR}}$  approximation to the low-dimensional features provided through  $\text{NN}_{\text{ENC}}$ .

The dataset dedicated to the training of  $\text{NN}_{\text{SUR}}$  is derived from the mosaics dataset  $\mathbf{D}_{\text{train}}^{\mathcal{J}}$  in Eq. (8) by mapping the mosaics in  $\mathbf{D}_{\text{train}}^{\mathcal{J}}$  onto the feature space, once and for all, to provide:

$$\mathbf{D}_{\text{train}}^h = \{(\boldsymbol{\theta}_k, \mathbf{h}_k)\}_{k=1}^{I_{\text{train}}} , \quad (14)$$

collecting the feature representations  $\mathbf{h}$  of the training data and the relative labels, in terms of the sought parameters  $\boldsymbol{\theta}$ . The set of weights and biases  $\boldsymbol{\Omega}_{\text{SUR}}$  parametrizing  $\text{NN}_{\text{SUR}}$  is then learned through the minimization of the following loss function:

$$\mathcal{L}_{\text{SUR}}(\boldsymbol{\Omega}_{\text{SUR}}, \mathbf{D}_{\text{train}}^h) = \frac{1}{I_{\text{train}}} \sum_{k=1}^{I_{\text{train}}} \|\mathbf{h}_k(\mathcal{J}_k) - \text{NN}_{\text{SUR}}(\boldsymbol{\theta}_k)\|_2^2 + \lambda_{\text{SUR}} \|\boldsymbol{\Omega}_{\text{SUR}}\|_2^2 . \quad (15)$$

Eq. (15) provides a measure of the distance between the target low-dimensional features vector  $\mathbf{h}(\mathcal{J})$ , obtained through the feature extractor  $\text{NN}_{\text{ENC}}$ , and its approximated counterpart  $\hat{\mathbf{h}} = \text{NN}_{\text{SUR}}(\boldsymbol{\theta})$ , provided through the feature-oriented surrogate model.

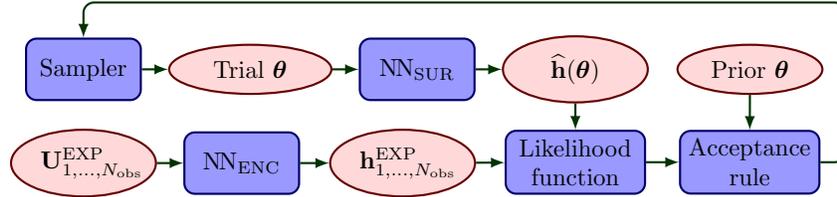
The implementation details of the DL models are reported in [Appendix B](#). It is worth noting that the modeling choices for the feature extractor and the feature-oriented surrogate are suited to the specific characteristics of the observational data considered in this paper. However, the overall framework presented is rather general, admitting different modeling choices, tailored to the data and the characteristics of the problem at hand. In this specific case, the vibration data of interest are encoded into images via MTF preprocessing to highlight structures and patterns in the data. While we thus show how to extract informative features in a low-dimensional metric space in the case of image data, data of different nature can be addressed in a similar way through an appropriate choice of the architectures of the DL models. For instance, one-dimensional convolutional layers could be exploited in place of two-dimensional ones to deal with time series data. Moreover, there may be cases where the decoding branch  $\text{NN}_{\text{DEC}}$  should be discarded. The reason why  $\text{NN}_{\text{DEC}}$  should be kept is that the reconstruction term in Eq. (12) regularizes the overall learning process. As a by-product, the trained  $\text{NN}_{\text{DEC}}$  and  $\text{NN}_{\text{SUR}}$  models can also serve as a surrogate model of the type  $\text{NN}_{\text{DEC}}(\text{NN}_{\text{SUR}}(\boldsymbol{\theta}))$ , following an approach similar to [46], to approximate the observational data for any given parameters vector  $\boldsymbol{\theta}$ . In our case, the contrastive term in Eq. (12) is minimized by exploiting the label information that completely describe the parametrization underlying the physics-based modeling of the problem. However, when the number  $N_{\text{par}}^{\text{HF}}$  of parameters in  $\mathbf{x}^{\text{HF}}$  becomes large, the paring process underlying the minimization of the contrastive loss function becomes computationally demanding. Although this issue does not show up in this work, it would be possible to address it by including in  $\boldsymbol{\theta}$  only a subset of  $\mathbf{x}^{\text{HF}}$ , limited to the parameters for which we seek to update the relative belief. In this eventuality, the decoding branch  $\text{NN}_{\text{DEC}}$  should be discarded to avoid a latent space showing dependence on parameters not included in  $\boldsymbol{\theta}$ , which could not be captured by  $\text{NN}_{\text{SUR}}$ . The same consideration also applies to systems subject to unknown stochastic inputs. In this eventuality, the dependency of the features vector  $\mathbf{h}$  on parameters describing stochastic inputs can not be uniquely defined and could not be modeled correctly by  $\text{NN}_{\text{SUR}}$ . This is, for instance, the case of seismic or wind loads acting on civil structures.

### 3.2. Feature-based MCMC sampling algorithm

The feature extractor  $\text{NN}_{\text{ENC}}$  and the feature-oriented surrogate model  $\text{NN}_{\text{SUR}}$ , trained as described in the previous section, are exploited in the online monitoring phase to enhance an MCMC sampler for model

updating purposes. The MCMC algorithm is here employed to update the prior probability density function (pdf)  $p(\boldsymbol{\theta})$  of the parameters vector  $\boldsymbol{\theta}$ , to provide a posterior pdf  $p(\boldsymbol{\theta}|\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}})$ , conditioned on a batch of gathered sensor recordings  $\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}$ . Here,  $N_{\text{obs}}$  represents the batch size of processed observations, each consisting of  $N_u$  series of  $L$  measurements over time.

By exploiting the Metropolis-Hastings sampler [47], the updating procedure is carried out by iteratively generating a chain of samples  $\{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{L_{\text{chain}}}\}$  from a proposal distribution, and then deciding whether to accept or reject each sample, on the basis of the likelihood of the current  $\boldsymbol{\theta}$  sample to represent  $\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}$ . To this aim,  $\text{NN}_{\text{ENC}}$  and  $\text{NN}_{\text{SUR}}$  are synergistically exploited to provide informative features  $\mathbf{h}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}$  via assimilation of the observational data, and to surrogate the functional link between the parametric space to be updated and the feature space, respectively, as sketched in Fig. 5. The resulting parameter estimation framework enjoys a greatly reduced computational cost due to the low dimensionality of the involved features, an improved convergence rate due to the geometrical structure characterizing the feature space, and more accurate estimates due to the informativeness of the extracted features.



**Figure 5.** Scheme of the MCMC procedure to update the probability distribution of the structural state. Red nodes refer to the input/output quantities, while blue nodes denote the relevant computational blocks.

According to the Bayes' rule, the posterior pdf  $p(\boldsymbol{\theta}|\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}})$  is given as:

$$p(\boldsymbol{\theta}|\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}) = \frac{p(\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}}, \quad (16)$$

where:  $p(\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}|\boldsymbol{\theta})$  is the likelihood function that provides the mechanism informing the posterior about the observations; the denominator is a normalizing factor, that is typically analytically intractable. To address this challenge,  $p(\boldsymbol{\theta}|\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}})$  is approximated through an MCMC sampling algorithm. By assuming an additive Gaussian noise to represent the uncertainty due to modeling inaccuracies and measurement noise, the likelihood function is assumed to be Gaussian too and to read:

$$p(\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}}|\boldsymbol{\theta}) = \prod_{n=1}^{N_{\text{obs}}} c^{-1} \exp\left(-\frac{(\mathbf{h}_n^{\text{EXP}} - \hat{\mathbf{h}}(\boldsymbol{\theta}))^\top (\mathbf{h}_n^{\text{EXP}} - \hat{\mathbf{h}}(\boldsymbol{\theta}))}{2\sigma^2}\right). \quad (17)$$

In Eq. (17), the term  $c = \sqrt{2\pi\sigma^2}$  is a normalization constant, with  $\sigma \in \mathbb{R}$  being the root mean square of the prediction error at each MCMC iteration, which serves as the standard deviation of the uncertainty under the zero-mean assumption. Due to its dependence on  $\boldsymbol{\theta}$ ,  $\sigma$  must be computed at each MCMC iteration; however, this does not affect the computational cost of the methodology due to the low dimensionality of the feature vectors.

The proposal pdf is taken as Gaussian. The covariance matrix is initialized as diagonal, with entries small enough so that the sampler gets moving, and then tuned as the sampling evolves by exploiting the adaptive Metropolis algorithm [48]. It is worth noting that the proposed procedure is a general one, admitting different choices for the sampling algorithm. The procedure can be similarly exploited with more advanced samplers, such as the transitional MCMC or Hybrid Monte Carlo algorithms and their recently proposed extensions, see for instance [49, 50]. Moreover, the entire methodology can be easily adapted to solve inverse problems in application domains other than SHM, even when dealing with data other than vibration recordings.

In order to check the quality of the estimates and stop the MCMC simulation, the estimated potential scale reduction (EPSR) metric [51] is employed to monitor the converge to a steady distribution. Since it is not possible to monitor the convergence of an MCMC simulation from a single chain, the EPSR test exploits multiple chains from parallel runs. Only when all the chains converge to the (same) stationary

distribution, the convergence criterion is considered satisfied. The EPSR metric  $\widehat{\mathcal{E}}$  tests the convergence of a multivariate chain by measuring the ratio between the estimate of the between-chain variance of samples and the average within-chain variance of samples. In this work, each MCMC simulation is carried out by randomly initializing five Markov chains, that are simultaneously evolved to meet the EPSR convergence criterion set to  $\widehat{\mathcal{E}} \leq 1.01$ ,  $\widehat{\mathcal{E}} = 1.1$  being a safe tolerance value [51]. The first half of each chain is then removed to get rid of the initialization effect, and 3 out of 4 samples are discarded to reduce the within chain autocorrelation of samples.

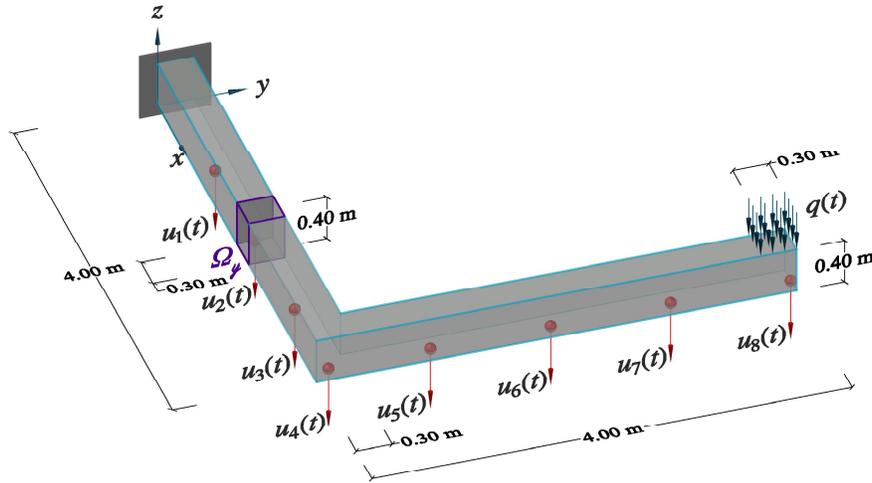
#### 4. Numerical results

This section aims at demonstrating the capability and performance of the proposed strategy in cases of simulated monitoring of three structural systems of increasing structural complexity: an L-shaped cantilever beam, a portal frame and a railway bridge.

The FOM and ROM have been solved in the `Matlab` environment, using the `redbKIT` library [52]. All computations have been carried out on a PC featuring an AMD Ryzen™ 9 5950X CPU @ 3.4 GHz and 128 GB RAM. The DL models have been implemented through the `Tensorflow`-based `Keras` API [53], and trained on a single `Nvidia GeForce RTX™ 3080` GPU card.

##### 4.1. L-shaped cantilever beam

The first test case involves the L-shaped cantilever beam depicted in Fig. 6. The structure is made of two arms, each one having a length of 4 m, a width of 0.3 m and a height of 0.4 m. The assumed mechanical properties are those of concrete: Young’s modulus  $E = 30$  GPa, Poisson’s ratio  $\nu = 0.2$ , density  $\rho = 2500$  kg/m<sup>3</sup>. The structure is excited by a distributed vertical load  $q(t)$ , acting on an area of  $(0.3 \times 0.3)$  m<sup>2</sup> close to its tip. The load varies in time according to  $q(t) = Q \sin(2\pi ft)$ , with  $Q \in [1, 5]$  kPa and  $f \in [10, 60]$  Hz respectively being the load amplitude and frequency. Following the setup described in Sec. 2,  $Q$  and  $f$  have a uniform distribution within their reported ranges.



**Figure 6.** L-shaped cantilever beam: details of synthetic recordings related to displacements  $u_1(t), \dots, u_8(t)$ , loading condition and damageable region  $\Omega_y$ .

Synthetic displacement time histories are gathered in relation to  $N_u = 8$  dofs along the bottom surface of the structure, to mimic a monitoring system arranged as depicted in Fig. 6. Each recording is provided for a time interval  $(0, T = 1)$  s with an acquisition frequency  $f_s = 200$  Hz. Recordings are corrupted with an additive Gaussian noise yielding a signal-to-noise ratio of 100.

The HF numerical model is obtained with a finite element discretization using linear tetrahedral elements and resulting in  $N_{FE} = 4659$  dofs. The structural dissipation is modeled by means of a Rayleigh’s damping matrix, assembled to account for a 5% damping ratio on the first four structural modes. Damage is simulated by reducing the material stiffness within a subdomain  $\Omega_y$  of size  $0.3 \times 0.3 \times 0.4$  m<sup>3</sup>. The position of  $\Omega_y$  is

parametrized by the coordinates of its center of mass  $\mathbf{y} = (x_\Omega, y_\Omega)^\top$ , with either  $x_\Omega$  or  $y_\Omega$  varying in the range  $[0.15, 3.85]$  m. The magnitude of the stiffness reduction is set to  $\delta = 25\%$  and held constant within the considered time interval. Accordingly, the vector of HF input parameters is  $\mathbf{x}^{\text{HF}} = (Q, f, x_\Omega, y_\Omega)^\top$ .

The basis matrix  $\mathbf{W}$  ruling the LF-ROM is obtained from a snapshot matrix  $\mathbf{S}$ , assembled through 200 evaluations of the LF-FOM, at varying values of the LF parameters  $\mathbf{x}^{\text{LF}} = (Q, f)^\top$  sampled via the latin hypercube rule. By prescribing a tolerance  $\epsilon = 10^{-3}$  on the fraction of energy content to be disregarded in the approximation, the order of the LF-ROM approximation turns out to be  $N_{\text{RB}} = 14$ .

The dataset  $\mathbf{D}_{\text{LF}}$  is built with  $I_{\text{LF}} = 10,000$  LF data instances collected using the LF-ROM. The dataset  $\mathbf{D}_{\text{HF}}$  is instead built with only  $I_{\text{HF}} = 1000$  additional HF data instances. The two datasets are exploited to train the MF-DNN surrogate model, with  $\mathbf{D}_{\text{LF}}$  employed to learn  $\text{NN}_{\text{LF}}$  and  $\mathbf{D}_{\text{HF}}$  employed to learn  $\text{NN}_{\text{HF}}$ . The trained MF-DNN surrogate model is then exploited to populate  $\mathbf{D}_{\text{train}}$  with  $I_{\text{train}} = 20,000$  instances, generated for varying values of the HF input parameters  $\mathbf{x}^{\text{HF}}$ .

To train the feature extractor and the feature-oriented surrogate, the vibration recordings in  $\mathbf{D}_{\text{train}}$  are transformed into images via MTF encoding. Each Mosaic  $\mathcal{J}_k$  in  $\mathbf{D}_{\text{train}}^{\mathcal{J}}$ , with  $k = 1, \dots, I_{\text{train}}$ , is obtained by disposing the  $N_u = 8$  MTF tesserae into a  $2 \times 4$  grid, with each MTF tessera being a  $40 \times 40$  pixel image (see Fig. 7). The size of the MTF tesserae depends on the length of the time series and on the width of the blurring kernel. For the detailed steps of the mosaics generation via MTF encoding, see Appendix A. In this case, the length  $L$  of the vibration recordings in  $\mathbf{D}_{\text{train}}$  is reduced by removing the initial 20% of each time history, to get rid of potential inaccuracies induced by the  $\text{NN}_{\text{HF}}$  LSTM model, and the chosen width of the blurring kernel is equal to 4. Moreover, each vibration recording in  $\mathbf{D}_{\text{train}}$  is normalized to follow a standard Gaussian distribution, thus allowing the dependence on the load amplitude  $Q$  to be neglected thanks to the linear-elastic modeling behind  $\mathbf{D}_{\text{HF}}$ . The mosaics dataset  $\mathbf{D}_{\text{train}}^{\mathcal{J}}$  is eventually exploited to minimize the loss functions in Eq. (12) and in Eq. (15), as described in Sec. 3.1 and according to the implementation details reported in Appendix B.

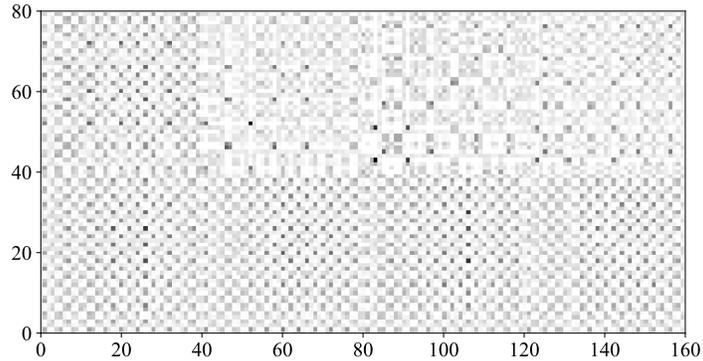
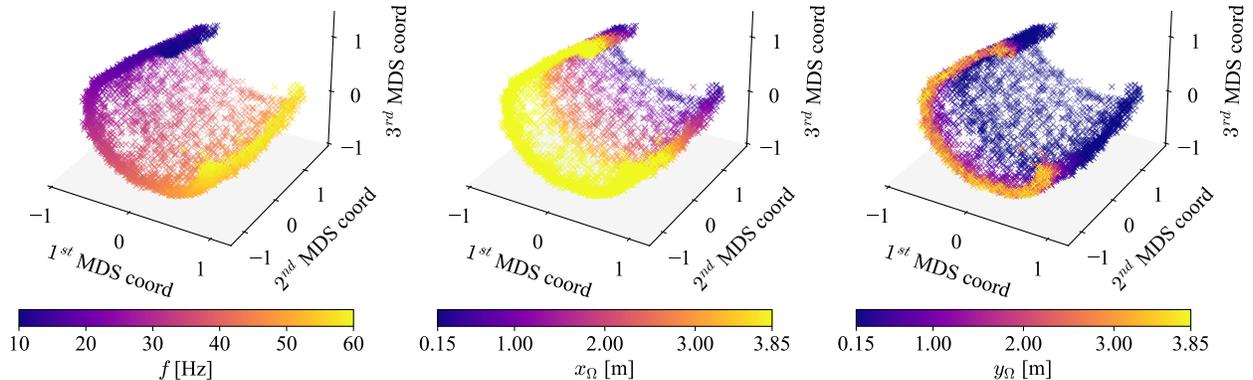


Figure 7. L-shaped cantilever beam - Exemplary MTF mosaic.

A compact representation of the low-dimensional features provided through  $\text{NN}_{\text{ENC}}$  for the validation set of  $\mathbf{D}_{\text{train}}^{\mathcal{J}}$  is reported in Fig. 8. The scatter plots report a downsized version of the extracted features, obtained by means of the metric multidimensional scaling (MDS) implemented in `scikit-learn` [54]. The three-dimensional (3D) MDS representations are reported with a color channel referring to the target values of the load frequency and of the damage position along the  $x$  and  $y$  directions. Note how the resulting manifold suitably encodes the sensitivity of the processed measurements on the parameters employed to describe the variability of the system. This visual check provides a first qualitative indication about the positive impact of adopting the feature space described by  $\text{NN}_{\text{ENC}}$ , to address the foreseen Bayesian model updating task.

In the absence of experimental data, the MCMC simulations are carried out considering batches of  $N_{\text{obs}} = 8$  HF noisy observations. Each observation batch is relative to the same  $\theta_\Omega$ , where  $\theta_\Omega \in [0.15, 7.55]$  m is an abscissa running along the axis of the structure and encoding the position of  $\Omega_y$  in place of  $x_\Omega$  and  $y_\Omega$ . Each data instance in the observation batch is generated by sampling parameters  $Q$  and  $f$  from a Gaussian pdf centered at the ground truth values of the parameters, and featuring a standard deviation equal to 0.25% of their respective ranges.



**Figure 8.** L-shaped cantilever beam - 3D multidimensional scaling representations of the low-dimensional features obtained for the validation data, against the target values of (left) load frequency, and damage position along (center) the  $x$ -direction and (right) the  $y$ -direction.

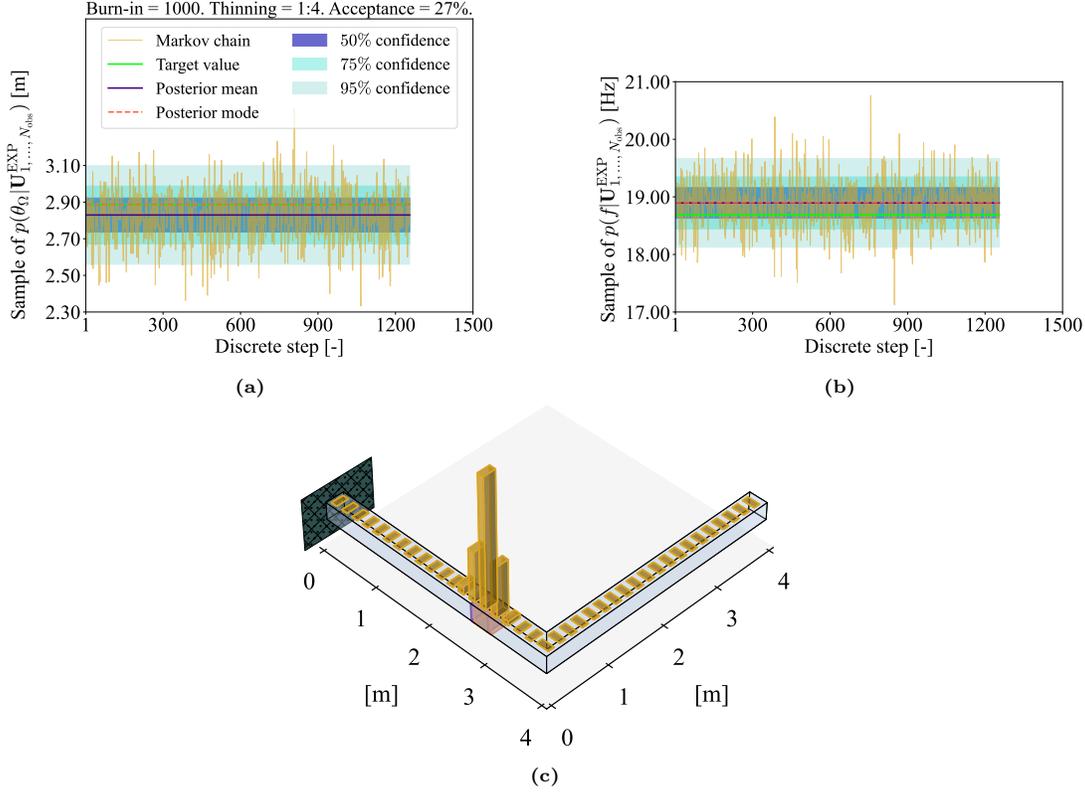
In the following, results are reported for six MCMC analyses, carried out under different operational conditions while moving the damage position from the clamp to the free-end. Tab. 1 reports the outcome of the identification of the damage position, in terms of target value, posterior mean, posterior mode, standard deviation and chain length. The quality of the estimates is highlighted by the small discrepancy between the target and the posterior mean values, which is only a few centimeters (less than 3% of the admissible support length). Also note the relatively low values of standard deviation, which however increase as the damage position gets far from the clamped side of the structure. This is a quite expected outcome, and is due to a smaller sensitivity of sensor recordings to damage when the damage is located near the free-end of the beam. The only case characterized by a large discrepancy between the target and the posterior mean values, as well as by a larger uncertainty, is in fact the last one, featuring a damage position close to the free-end of the beam. For instance, in case 4, the discrepancy between the target and the posterior mean values is only 0.044 m over an admissible support of 7.4 m, while it reaches 0.845 m in case 6. Despite the larger discrepancy between the target and the posterior mean, the target value falls within the 95% confidence interval, as in the other cases, demonstrating the reliability of the estimates provided.

**Table 1.** L-shaped cantilever beam - Damage localization results for different operational and damage conditions, in terms of: target value; posterior mean; posterior mode; standard deviation; chain length.

Case	Target( $\theta_\Omega$ )	Mean( $\theta_\Omega$ )	Mode( $\theta_\Omega$ )	Stdv( $\theta_\Omega$ )	$L_{\text{chain}}$
1	0.564 m	0.580 m	0.600 m	0.043 m	2200
2	2.200 m	2.195 m	2.225 m	0.110 m	2000
3	2.888 m	2.830 m	2.887 m	0.137 m	2000
4	4.435 m	4.391 m	4.362 m	0.077 m	2000
5	5.204 m	5.403 m	5.412 m	0.315 m	2150
6	7.380 m	6.535 m	6.200 m	0.511 m	2250

An exemplary MCMC simulation outcome is reported in Fig. 9 for case 3. The graphs show the sampled Markov chain alongside the estimated posterior mean and credibility intervals, for both  $\theta_\Omega$  and  $f$ . Note that the chains are plotted over a relatively small range of values for the sake of visualization. Thanks to the low-dimensionality of the involved features, the procedure also enjoys a considerable computational efficiency. The computing time for the parameter estimation is only about 5 s; this is a remarkable result, highlighting the real-time damage identification capabilities of the proposed strategy, all with quantified uncertainty.

To quantify the impact of using the learnable features, additional results relevant to the identification of the damage location are reported in Tab. 2, as obtained in [27]. In this latter,  $p(\theta|\mathbf{U}_{1,\dots,N_{\text{obs}}}^{\text{EXP}})$  was sampled without employing the feature extractor and the feature-oriented surrogate, but directly using the MF-DNN surrogate model. The comparison with Tab. 1 shows that  $\text{NN}_{\text{ENC}}$  and  $\text{NN}_{\text{SUR}}$  allow the parameter identification outcomes to be improved in all the considered performance indicators. In case 4, for instance, the discrepancy between the target and the posterior mean values, and the standard deviation value, are shown in Tab. 2 to increase by 7 and 10 times, respectively.



**Figure 9.** L-shaped cantilever beam - Exemplary MCMC result (case 3): Markov chain, target value, posterior mean and credibility intervals relative to the estimation of **a** damage position  $\theta_\Omega$  and **b** load frequency  $f$ ; **c** histogram of the approximated, unnormalized posterior pdf  $p(\theta_\Omega | \mathbf{U}_{1, \dots, N_{\text{obs}}}^{\text{EXP}})$  over the admissible support.

**Table 2.** L-shaped cantilever beam - Damage localization results for different operational and damage conditions, without leveraging the feature extractor and the feature-oriented surrogate model. Table adapted from [27].

Case	Target( $\theta_\Omega$ )	Mean( $\theta_\Omega$ )	Mode( $\theta_\Omega$ )	Stdv( $\theta_\Omega$ )	$L_{\text{chain}}$
1	0.564 m	0.631 m	0.587 m	0.170 m	2000
2	2.200 m	2.474 m	2.414 m	0.511 m	2000
3	2.888 m	3.088 m	2.844 m	0.710 m	3400
4	4.435 m	4.834 m	4.198 m	0.969 m	2000
5	5.204 m	5.759 m	5.397 m	0.962 m	3000
6	7.380 m	6.080 m	7.136 m	0.866 m	4000

#### 4.2. Portal frame

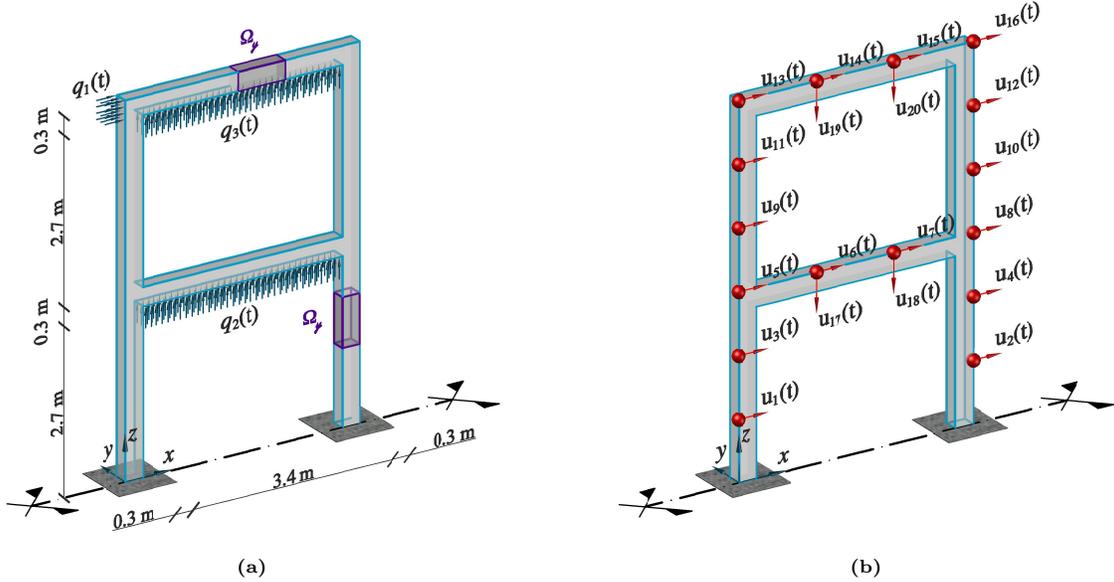
The second test case involves the two-story portal frame depicted in Fig. 10. The columns have a width of 0.3 m, the beams have a height of 0.3 m, the inter-story height is 2.7 m, the span of the beams is 3.4 m, and the out of plane thickness is 0.45 m. The assumed mechanical properties are: Young's modulus  $E = 34$  GPa, Poisson's ratio  $\nu = 0.2$ , density  $\rho = 2500$  kg/m<sup>3</sup>.

The structure is excited by three distributed loads  $q_1(t), q_2(t), q_3(t)$ , respectively applied on top of the left column and on the bottom surface of the two horizontal beams, as shown in Fig. 10a. The three loads vary in time according to:

$$q_{1,2,3}(t) = \begin{cases} Q \frac{t}{T_q}, & \text{if } t \leq T_q, \\ 0, & \text{if } t > T_q, \end{cases} \quad (18)$$

with  $Q = 10$  kPa and  $T_q = 0.08$  s. This fast-linear-ramp actuation may be connected to smart instrumented structures, equipped with an excitation system designed for forced vibration tests.

Displacement time histories are obtained in relation to  $N_u = 20$  dofs, mimicking a monitoring system deployed as depicted in Fig. 10b. The recordings are provided for a time interval  $(0, T = 1.12$  s) with an



**Figure 10.** Portal frame: details of (a) loading condition, damageable region  $\Omega_y$ , and (b) synthetic recordings related to displacements  $u_1(t), \dots, u_{20}(t)$ .

acquisition frequency  $f_s = 125$  Hz, and corrupted with an additive Gaussian noise yielding a signal to noise ratio of 150.

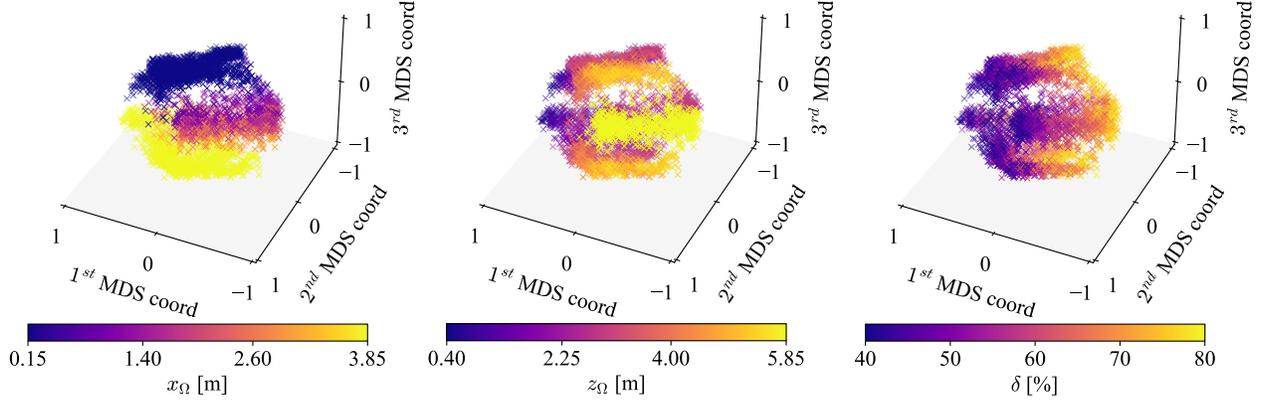
The HF numerical model features  $N_{FE} = 4827$  dofs. The Rayleigh’s damping matrix is assembled to account for a 2.5% damping ratio on the first two structural modes. In this case, damage is simulated by means of a localized stiffness reduction that can take place anywhere in the frame, within subdomains  $\Omega_y$  featuring a different layout for the columns and for the beams (see Fig. 10a). The position of  $\Omega_y$  is parameterized by the coordinates of its center of mass  $\mathbf{y} = (x_\Omega, z_\Omega)^\top$ , with  $x_\Omega$  and  $z_\Omega$  varying in the ranges  $[0.15, 3.85]$  m and  $[0.4, 5.85]$  m, respectively. The magnitude of the stiffness reduction can range in  $\delta \in [40\%, 80\%]$ , and remains constant during an excitation event.

In the present case, the LF structural response is not parametrized. The LF dataset  $\mathbf{D}_{LF}$  consists of a single instance underlying the structural response in the absence of damage. This is thus employed in place of  $\mathbf{NN}_{LF}$  in the MF-DNN surrogate. The HF component  $\mathbf{NN}_{HF}$  is trained on  $I_{HF} = 1000$  HF data instances, to enrich the LF instance with the effects of damage for any given  $\mathbf{x}^{HF} = (x_\Omega, z_\Omega, \delta)^\top$ . The trained MF-DNN surrogate is then employed to populate  $\mathbf{D}_{train}$  with  $I_{train} = 20,000$  instances, generated for varying values of the  $\mathbf{x}^{HF}$  input parameters.

The mosaics dataset  $\mathbf{D}_{train}^{\mathcal{F}}$  is obtained by encoding each training instance in  $\mathbf{D}_{train}$  into a  $4 \times 5$  MTF mosaic, with each MTF tessera being a  $32 \times 32$  pixel image. Before undergoing the MTF encoding, the vibration recordings in  $\mathbf{D}_{train}$  are normalized to follow a Gaussian distribution with zero mean and unit standard deviation, and the initial 8% of each time history is removed to get rid of potential inaccuracies induced by  $\mathbf{NN}_{HF}$ . In the present case, the width of the blurring kernel is set equal to 4.

The MDS representation of the features provided through  $\mathbf{NN}_{ENC}$  for the validation set of  $\mathbf{D}_{train}^{\mathcal{F}}$  is reported in Fig. 11. In this case, the color channels correspond to  $x_\Omega$ ,  $z_\Omega$ , and  $\delta$ . The three plots qualitatively demonstrate also in this case the presence of an underlying manifold, which encodes the sensitivity of the structural response to the health parameters.

The learned feature space is employed to update the prior belief of  $\boldsymbol{\theta} = (x_\Omega, z_\Omega, \delta)^\top$  under varying damage conditions via MCMC simulations. The MCMC algorithm is fed with batches of  $N_{obs} = 8$  noisy observations, all related to the same damage location and magnitude. In the following, we provide an analysis of the results obtained from the six MCMC simulations summarized in Tab. 3. In general, both the damage location and the damage magnitude are identified with very high accuracy and relatively low uncertainty. There are no cases characterized by a significant discrepancy between the target and the posterior mean values. As expected, the standard deviation of either  $x_\Omega$  or  $z_\Omega$  is larger along the axis in which  $\Omega_y$  can move.



**Figure 11.** Portal frame - 3D multidimensional scaling representations of the low-dimensional features obtained for the validation data, against the target values of damage position along (left) the  $x$ -direction and (center) the  $z$ -direction, and (right) damage magnitude.

Additionally, the uncertainty in  $\delta$  increases as  $\Omega_u$  gets far from the clamped sides, due to a smaller sensitivity of sensor recordings to damage in such cases. For visualization purposes, an exemplary MCMC-recovered posterior is reported in Fig. 12 for case 3.

**Table 3.** Portal frame - Damage localization and quantification results for different operational and damage conditions, in terms of: target value; posterior mean; posterior mode; standard deviation; chain length.

Case	Target( $x_\Omega; z_\Omega; \delta$ )	Mean( $x_\Omega; z_\Omega; \delta$ )	Mode( $x_\Omega; z_\Omega; \delta$ )	Stdv( $x_\Omega; z_\Omega; \delta$ )	$L_{chain}$
1	0.15 m; 0.58 m; 74.32%	0.21 m; 0.66 m; 74.54%	0.15 m; 0.40 m; 76.00%	0.06 m; 0.22 m; 2.21%	4550
2	0.15 m; 3.68 m; 77.69%	0.21 m; 3.47 m; 74.92%	0.15 m; 3.67 m; 76.00%	0.06 m; 0.16 m; 3.29%	3350
3	3.85 m; 2.65 m; 67.58%	3.76 m; 2.61 m; 68.63%	3.85 m; 2.58 m; 68.00%	0.07 m; 0.19 m; 2.65%	4050
4	3.85 m; 4.94 m; 53.30%	3.76 m; 5.04 m; 53.13%	3.85 m; 5.30 m; 52.00%	0.07 m; 0.19 m; 4.02%	5850
5	1.94 m; 2.85 m; 56.64%	1.85 m; 2.84 m; 56.37%	1.63 m; 3.58 m; 56.00%	0.26 m; 0.23 m; 2.99%	3850
6	1.70 m; 5.85 m; 63.70%	1.90 m; 5.70 m; 69.06%	2.00 m; 5.85 m; 68.00%	0.29 m; 0.16 m; 3.35%	2950

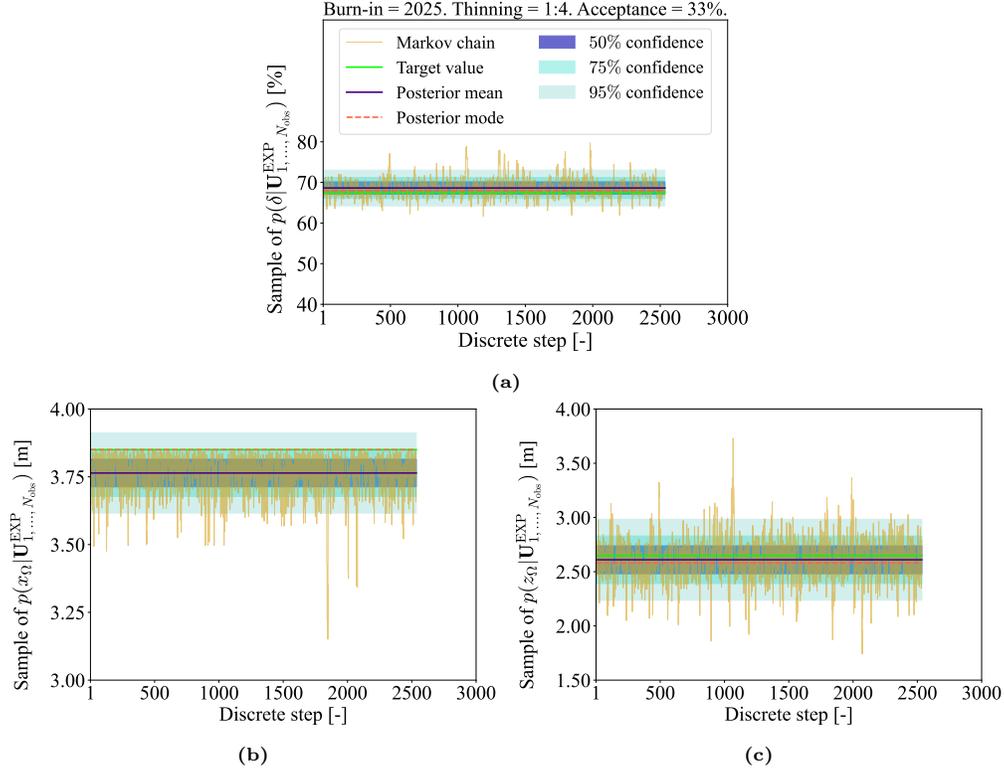
#### 4.3. Hörnefors railway bridge

This third test case aims to assess the performance of the proposed strategy in a more complex situation, involving the railway bridge depicted in Fig. 13. It is an integral concrete portal frame bridge located along the Bothnia line in Hörnefors, Sweden. It features a span of 15.7 m, a free height of 4.7 m and a width of 5.9 m (edge beams excluded). The thickness of the structural elements is 0.5 m for the deck, 0.7 m for the frame walls, and 0.8 m for the wing walls. The bridge is founded on two plates connected by stay beams and supported by pile groups. The concrete is of class C35/45, whose mechanical properties are:  $E = 34$  GPa,  $\nu = 0.2$ ,  $\rho = 2500$  kg/m<sup>3</sup>. The superstructure consists of a single track with sleepers spaced 0.65 m apart, resting on a ballast layer 0.6 m deep, 4.3 m wide and featuring a density of  $\rho_B = 1800$  kg/m<sup>3</sup>. The geometrical and mechanical modeling data have been adapted from former research activities [55, 56].

The bridge is subjected to the transit of trains of type Gröna Tåget, at a speed  $v \in [160, 215]$  km/h. Only trains composed of two wagons are considered, thus characterized by 8 axles, each one carrying a mass  $\phi \in [16, 22]$  ton. The corresponding load model is described in [25], and consists of 25 equivalent distributed forces transmitted by the sleepers to the deck through the ballast layer with a slope 4 : 1, according to Eurocode 1 [57].

The monitoring system features  $N_u = 10$  sensors and is deployed as depicted in Fig. 14. Displacement time histories are provided for a time interval  $(0, T = 1.5$  s), with an acquisition frequency  $f_s = 400$  Hz.

The HF numerical model features  $N_{FE} = 17,292$  dofs, resulting from a finite element discretization with an element size of 0.15 m for the deck, to enable a smooth propagation of the traveling load, and 0.80 m elsewhere. The presence of the ballast layer is accounted for through an increased density for the deck and for the edge beams. The embankments are accounted for through distributed springs over the surfaces facing the ground, modeled as a Robin mixed boundary condition (with elastic coefficient  $k_{Robin} = 10^8$  N/m<sup>3</sup>).



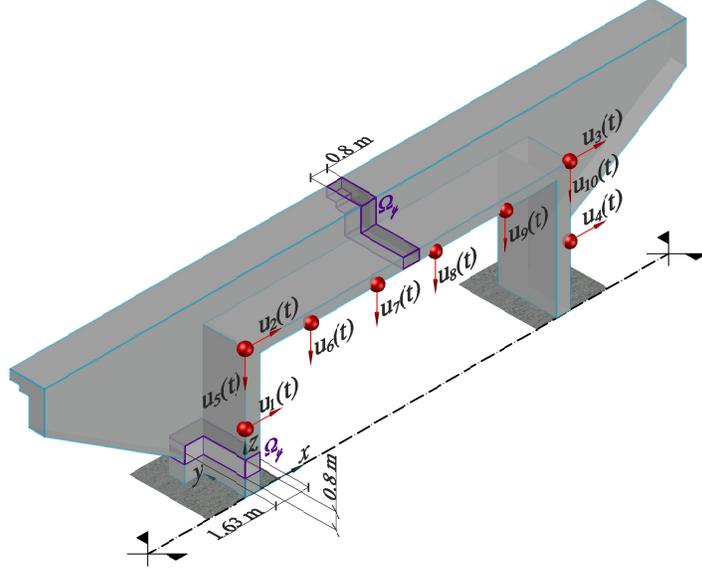
**Figure 12.** Portal frame - Exemplary MCMC result (case 3): Markov chain, target value, posterior mean, posterior mode and credibility intervals relative to the estimation of (a) damage magnitude  $\delta$ , (b) damage position along the  $x$ -direction, and (c) damage position along the  $z$ -direction.



**Figure 13.** Hörnefors railway bridge.

The Rayleigh's damping matrix accounts for a 5% damping ratio on the first two structural modes. In this case, damage is simulated by means of a localized stiffness reduction that can take place anywhere over the two lateral frame walls and the deck, within subdomains  $\Omega_y$  featuring a different layout in the two cases (see Fig. 14). The position of  $\Omega_y$  is parametrized through  $\mathbf{y} = (x_\Omega, z_\Omega)^\top$ , with  $x_\Omega$  and  $z_\Omega$  varying in the ranges  $[-0.115, 16.515]$  m and  $[0.4, 6.25]$  m, respectively. The stiffness reduction can occur with a magnitude  $\delta \in [40\%, 80\%]$ , which is kept fixed while a train travels across the bridge. To summarize, the vector of HF input parameters is  $\mathbf{x}^{\text{HF}} = (v, \phi, x_\Omega, z_\Omega, \delta)^\top$ .

The basis matrix  $\mathbf{W}$  is obtained from a snapshot matrix  $\mathbf{S}$ , assembled through 200 evaluations of the LF-FOM for different values of parameters  $\mathbf{x}^{\text{LF}} = (v, \phi)^\top$ . By setting the error tolerance to  $\epsilon = 10^{-3}$ ,  $N_{\text{RB}} = 312$  POD modes are retained in  $\mathbf{W}$ .



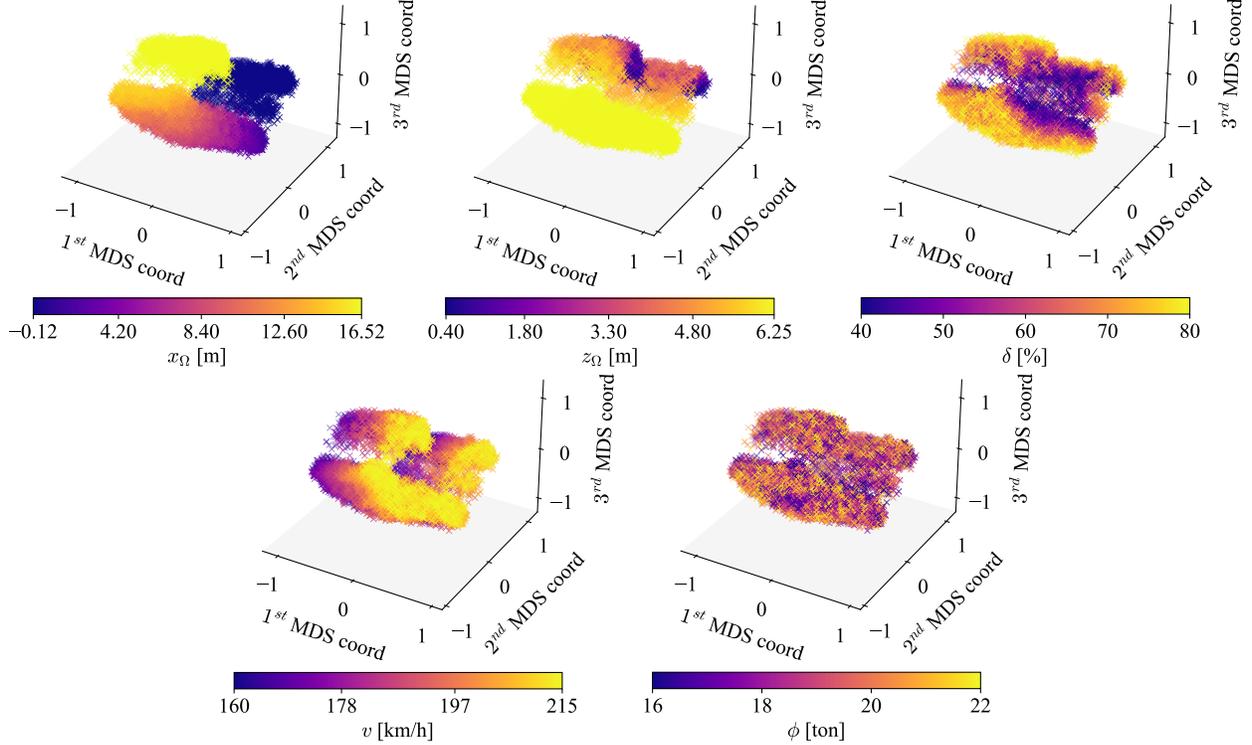
**Figure 14.** Railway bridge: details of synthetic recordings related to displacements  $u_1(t), \dots, u_{10}(t)$ , and damageable region  $\Omega_y$ .

The MF-DNN surrogate model is trained using  $I_{LF} = 5000$  LF data instances for  $\text{NN}_{LF}$ , and only  $I_{HF} = 500$  HF data instances for  $\text{NN}_{HF}$ . The MF-DNN surrogate is then employed to populate  $\mathbf{D}_{\text{train}}$  with  $I_{\text{train}} = 30,000$  instances, generated for varying values of the  $\mathbf{x}^{\text{HF}}$  input parameters.

The mosaics dataset  $\mathbf{D}_{\text{train}}^{\mathcal{F}}$  is obtained by encoding each training instance in  $\mathbf{D}_{\text{train}}$  into a  $2 \times 5$  MTF mosaic, with each MTF tessera being a  $64 \times 64$  pixel image. Before undergoing the MTF encoding, the initial 4% of each time history in  $\mathbf{D}_{\text{train}}$  is removed to get rid of potential inaccuracies induced by  $\text{NN}_{HF}$ . Moreover, since in this case the vibration recordings are characterized by data distributions mainly spread over the tails, each time history in  $\mathbf{D}_{\text{train}}$  is normalized to take values between 0 and 1, and quantized through a uniform bin assignment instead of a Gaussian one. In this case, the chosen width of the blurring kernel is equal to 9.

The visual check on the MDS representation of the features extracted from the validation data is reported in Fig. 15. In this case, the color channels refer to each entry of  $\mathbf{x}^{\text{HF}}$ . It is interesting to note how the overall shape defined by the scatter plot resembles the structural layout of the bridge (rotated and extruded), which is automatically retrieved from  $x_{\Omega}$  and  $z_{\Omega}$ . These plots qualitatively show a clear sensitivity of the low-dimensional feature space to the damage location, the damage magnitude, and the train velocity. The axle mass is instead characterized by a fuzzier representation, which does not present a manifold topology capable of adequately capturing its influence on the processed measurements.

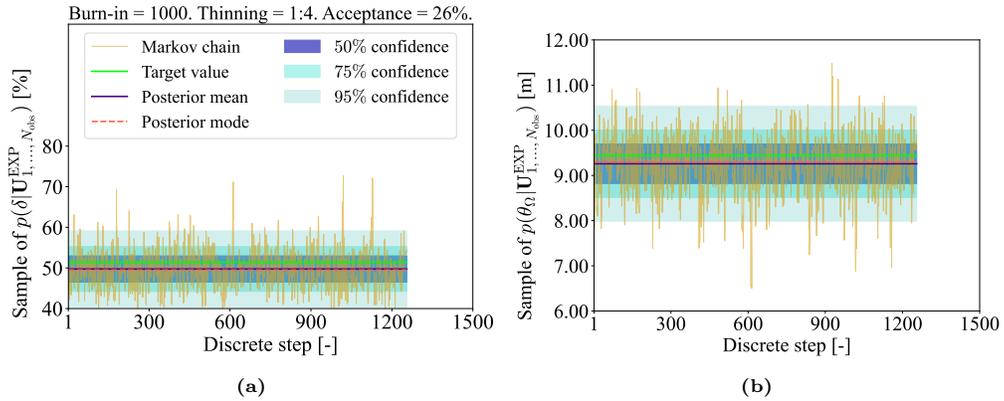
Results of six MCMC simulations, carried out for different operational and damage conditions, are next considered. The MCMC algorithm is fed with batches of  $N_{\text{obs}} = 8$  HF observations. Each observation batch is relative to the same damage location  $\theta_{\Omega} \in [0.4, 26]$  m and damage magnitude  $\delta$ , but each data instance in the batch is obtained for a random value of train velocity  $v$  and axle mass  $\phi$ . The train speed and axle mass are provided by the train on-board system; since these measurements are able to be taken accurately, the relative posterior is deterministically set to the measured values. The results relevant to the sampling of the posterior pdf of the damage location and magnitude are reported in Tab. 4. The damage location is always identified with relatively low uncertainty, except in case 2. Nevertheless, the relative discrepancy between the target and the posterior mean values is only 1.58 m over an admissible support of 25.6 m. On the other hand, the damage magnitude always falls within the estimated 95% confidence interval. Again, the worst outcome is obtained in case 2, which is characterized by a discrepancy between the target and the posterior mean values of about 7.5%. An exemplary MCMC outcome is reported in Fig. 16 for case 4: note how the recovered posterior present good post-inference diagnostic statistics, with no divergences and high homogeneity between and within chains.



**Figure 15.** Railway bridge - 3D MDS representations of the low-dimensional features obtained for the validation data, against the target values of damage position along (top-left) the  $x$ -direction and (top-center) the  $z$ -direction, (top-right) damage magnitude, (bottom-left) train velocity, and (bottom-right) axle mass.

**Table 4.** Railway bridge - Damage localization and quantification results for different operational and damage conditions, in terms of: target value; posterior mean; posterior mode; standard deviation; chain length.

Case	Target( $\theta_\Omega; \delta$ )	Mean( $\theta_\Omega; \delta$ )	Mode( $\theta_\Omega; \delta$ )	Stdv( $\theta_\Omega; \delta$ )	$L_{\text{chain}}$
1	2.31 m; 73.42%	2.15 m; 70.96%	2.00 m; 75.00%	0.81 m; 7.03%	2650
2	3.96 m; 63.75%	2.38 m; 56.28%	2.30 m; 56.50%	0.38 m; 8.43%	2000
3	6.07 m; 47.53%	5.72 m; 50.91%	5.75 m; 41.5%	0.18 m; 8.37%	2000
4	9.44 m; 51.41%	9.19 m; 49.33%	9.15 m; 50.5%	0.72 m; 5.35%	2000
5	13.37 m; 41.25%	13.06 m; 43.97%	13.40 m; 41.50%	0.84 m; 3.95%	2000
6	17.13 m; 52.07%	16.27 m; 48.02%	16.20 m; 41.50%	1.45 m; 8.31%	2150



**Figure 16.** Railway bridge - Exemplary MCMC result (case 4): Markov chain, target value, posterior mean, posterior mode and credibility intervals relative to the estimation of (a) damage magnitude  $\delta$  and (b) damage position  $\theta_\Omega$ .

## 5. Conclusions

In this work, we have proposed a deep learning-based strategy to enhance stochastic approaches to structural health monitoring. The presented strategy relies upon a learnable feature-extractor and a feature-oriented surrogate model. The two data-driven models are synergically exploited to improve the accuracy and efficiency of the parameter estimation workflow. The feature extractor makes the selection and the extraction of informative features from raw sensor recordings almost automated. The extracted features allow the sensitivity of the observational data to the sought parameters to be encoded in a low-dimensional metric space. The surrogate model approximates the functional link between the parametric input space, for which we seek to update the relative belief, and the low-dimensional feature space. The methodology can be easily adapted to solve inverse problems in application domains other than structural health monitoring, such as, e.g., scattering problems, medical diagnoses, and inverse kinematics.

The computational procedure takes advantage of a preliminary offline phase that: (i) employs physics-based numerical models and reduced-order modeling, to overcome the lack of experimental data for civil applications under varying damage and operational conditions; (ii) exploits a multi-fidelity surrogate modeling strategy to generate a large labeled dataset; (iii) trains the feature extractor and the feature-oriented surrogate model.

The proposed strategy has been assessed on the simulated monitoring of an L-shaped cantilever beam, a portal frame, and a railway bridge. In the absence of experimental data under the effect of varying operational and damage conditions, the tests have been carried out by exploiting high-fidelity simulation data corrupted with an additive Gaussian noise. The obtained results have shown that learnable features used instead of raw vibration recordings, enables to largely improve the parameter identification outcomes. The presented strategy also enjoys a high computational efficiency due to the low-dimensionality of the involved features.

The upcoming activities will be devoted to the integration of the proposed strategy within a digital twin concept, see for instance [58, 31]. Along this path, the assimilation of observational data to provide real-time structural health estimates would be useful to inform an optimal planning of maintenance and management actions, within a dynamic decision-making framework.

**Acknowledgments:** This work is supported in part by the interdisciplinary Ph.D. Grant “Physics-Informed Deep Learning for Structural Health Monitoring” at Politecnico di Milano. Andrea Manzoni acknowledges the project “Dipartimento di Eccellenza” 2023-2027, funded by MUR, and the project FAIR (Future Artificial Intelligence Research), funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, Investment 1.3, Line on Artificial Intelligence).

## References

- [1] A. S. of Civil Engineers, *Adaptive Design and Risk Management*, ch. 7, pp. 173—226. Reston, Virginia: American Society of Civil Engineers, 2018.
- [2] S. D. Glaser and A. Tolman, “Sense of Sensing: From Data to Informed Decisions for the Built Environment,” *J Infrastruct Syst*, vol. 14, no. 1, pp. 4–14, 2008.
- [3] J. D. Achenbach, “Structural health monitoring – What is the prescription?,” *Mech Res Commun*, vol. 36, no. 2, pp. 137–142, 2009.
- [4] L. Rosafalco, M. Torzoni, A. Manzoni, S. Mariani, and A. Corigliano, “A Self-adaptive Hybrid Model/data-Driven Approach to SHM Based on Model Order Reduction and Deep Learning,” in *Structural Health Monitoring Based on Data Science Techniques*, pp. 165–184, Springer International Publishing, 2022.
- [5] E. García-Macías and F. Ubertini, “Integrated SHM Systems: Damage Detection Through Unsupervised Learning and Data Fusion,” in *Structural Health Monitoring Based on Data Science Techniques*, pp. 247–268, Springer International Publishing, 2022.
- [6] M. Torzoni, A. Manzoni, and S. Mariani, “Structural health monitoring of civil structures: A diagnostic framework powered by deep metric learning,” *Comput Struct*, vol. 271, p. 106858, 2022.

- [7] K. Worden, “Structural fault detection using a novelty measure,” *J Sound Vib*, vol. 201, no. 1, pp. 85–101, 1997.
- [8] L. Rosafalco, A. Manzoni, S. Mariani, and A. Corigliano, “Fully convolutional networks for structural health monitoring through multivariate time series classification,” *Adv Model and Simul in Eng Sci*, vol. 7, p. 38, 2020.
- [9] C. R. Farrar, S. W. Doebling, and D. A. Nix, “Vibration-Based Structural Damage Identification,” *Phil Trans R Soc A*, vol. 359, no. 1778, pp. 131–149, 2001.
- [10] C. M. Bishop, *Pattern Recognition and Machine Learning*. Information Science and Statistics, New York, NY: Springer-Verlag, 2006.
- [11] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. Inman, “A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications,” *Mech Syst Signal Process*, vol. 147, p. 107077, 2021.
- [12] L. Ierimonti, N. Cavalagli, E. García-Macías, I. Venanzi, and F. Ubertini, “Bayesian-Based Damage Assessment of Historical Structures Using Vibration Monitoring Data,” in *International Workshop on Civil Structural Health Monitoring*, pp. 415–429, Springer, 2021.
- [13] D. Cristiani, C. Sbarufatti, and M. Giglio, “Damage diagnosis and prognosis in composite double cantilever beam coupons by particle filtering and surrogate modelling,” *Struct Health Monit*, vol. 20, no. 3, pp. 1030–1050, 2021.
- [14] A. Kamariotis, E. Chatzi, and D. Straub, “Value of information from vibration-based structural health monitoring extracted via Bayesian model updating,” *Mech Syst Signal Process*, vol. 166, p. 108465, 2022.
- [15] S. Eftekhar Azam and S. Mariani, “Online damage detection in structural systems via dynamic inverse analysis: A recursive Bayesian approach,” *Eng Struct*, vol. 159, pp. 28–45, 2018.
- [16] J. Bromley, I. Guyon, Y. Lecun, E. Säckinger, and R. Shah, “Signature Verification using a ”Siamese” Time Delay Neural Network,” *Int J Pattern Recognit Artif Intell*, vol. 7, p. 25, 1993.
- [17] S. Chopra, R. Hadsell, and Y. LeCun, “Learning a similarity metric discriminatively, with application to face verification,” in *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, pp. 539–546, 2005.
- [18] R. Hadsell, S. Chopra, and Y. Lecun, “Dimensionality Reduction by Learning an Invariant Mapping,” in *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, pp. 1735–1742, 2006.
- [19] M. Kaya and H. Bilge, “Deep Metric Learning: A Survey,” *Symmetry*, vol. 11, p. 1066, 2019.
- [20] A. Bellet, A. Habrard, and M. Sebban, “A Survey on Metric Learning for Feature Vectors and Structured Data.” arXiv preprint arXiv:1306.6709, 2013.
- [21] F. Cakir, K. He, X. Xia, B. Kulis, and S. Sclaroff, “Deep Metric Learning to Rank,” in *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, pp. 1861–1870, 2019.
- [22] R. Hou, X. Wang, and Y. Xia, “Vibration-Based Structural Damage Detection Using Sparse Bayesian Learning Techniques,” in *Structural Health Monitoring Based on Data Science Techniques*, pp. 1–25, Springer International Publishing, 2022.
- [23] P. L. Green and K. Worden, “Bayesian and Markov chain Monte Carlo methods for identifying nonlinear systems in the presence of uncertainty,” *Phil Trans R Soc A*, vol. 373, 2015.
- [24] H. F. Lam, J. H. Yang, and S. K. Au, “Markov chain Monte Carlo-based Bayesian method for structural model updating and damage detection,” *Struct Contr Health Monit*, vol. 25, no. 4, pp. 1–22, 2018.

- [25] L. Rosafalco, M. Torzoni, A. Manzoni, S. Mariani, and A. Corigliano, “Online structural health monitoring by model order reduction and deep learning algorithms,” *Comput Struct*, vol. 255, p. 106604, 2021.
- [26] M. Torzoni, L. Rosafalco, A. Manzoni, S. Mariani, and A. Corigliano, “SHM under varying environmental conditions: An approach based on model order reduction and deep learning,” *Comput Struct*, vol. 266, p. 106790, 2022.
- [27] M. Torzoni, A. Manzoni, and S. Mariani, “A multi-fidelity surrogate model for structural health monitoring exploiting model order reduction and artificial neural networks,” *Mech Syst Signal Process*, vol. 197, p. 110376, 2023.
- [28] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced basis methods for partial differential equations: an introduction*. Springer, 2015.
- [29] L. Rosafalco, A. Manzoni, S. Mariani, and A. Corigliano, “Combined Model Order Reduction Techniques and Artificial Neural Network for Data Assimilation and Damage Detection in Structures,” in *Computational Sciences and Artificial Intelligence in Industry: New Digital Technologies for Solving Future Societal and Economical Challenges*, pp. 247–259, Springer International Publishing, 2022.
- [30] C. Farrar and K. Worden, *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley & Sons, 2013.
- [31] M. G. Kapteyn, J. V. R. Pretorius, and K. E. Willcox, “A probabilistic graphical model foundation for enabling predictive digital twins at scale,” *Nat Comput Sci*, vol. 1, no. 5, pp. 337–347, 2021.
- [32] A. Teughels, J. Maeck, and G. De Roeck, “Damage assessment by FE model updating using damage functions,” *Comput Struct*, vol. 80, no. 25, pp. 1869–1879, 2002.
- [33] L. Sirovich, “Turbulence and the dynamics of coherent structures. I. Coherent structures,” *Q Appl Math*, vol. 45, no. 3, pp. 561–571, 1987.
- [34] G. Kerschen and J.-C. Golinval, “Physical interpretation of the proper orthogonal modes using the singular value decomposition,” *J Sound Vib*, vol. 249, no. 5, pp. 849–865, 2002.
- [35] G. Kerschen, J.-c. Golinval, A. F. Vakakis, and L. A. Bergman, “The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview,” *Nonlinear Dyn*, vol. 41, no. 1, pp. 147–169, 2005.
- [36] B. Peherstorfer, K. Willcox, and M. Gunzburger, “Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization,” *SIAM Review*, vol. 60, no. 3, pp. 550–591, 2018.
- [37] X. Meng and G. E. Karniadakis, “A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse PDE problems,” *Journal of Computational Physics*, vol. 401, p. 109020, 2020.
- [38] P. Conti, M. Guo, A. Manzoni, and J. Hesthaven, “Multi-fidelity surrogate modeling using long short-term memory networks,” *Comput Methods Appl Mech Eng*, vol. 404, p. 115811, 2023.
- [39] V. Giglioni, I. Venanzi, A. E. Baia, V. Poggioni, A. Milani, and F. Ubertini, “Deep Autoencoders for Unsupervised Damage Detection with Application to the Z24 Benchmark Bridge,” in *European Workshop on Structural Health Monitoring*, pp. 1048–1057, Springer International Publishing, 2023.
- [40] G. R. Garcia, G. Michau, M. Ducoffe, J. S. Gupta, and O. Fink, “Temporal signals to images: Monitoring the condition of industrial assets with deep learning image processing algorithms,” *Proc Inst Mech Eng O J Risk Reliab*, vol. 236, no. 4, pp. 617–627, 2022.
- [41] I. M. Mantawy and M. O. Mantawy, “Convolutional neural network based structural health monitoring for rocking bridge system by encoding time-series into images,” *Struct Control Health Monit*, vol. 29, no. 3, p. e2897, 2022.

- [42] Z. Wang and T. Oates, “Imaging Time-Series to Improve Classification and Imputation,” in *Proc of the International Conference on Artificial Intelligence*, vol. 24, pp. 3939–3945, 2015.
- [43] N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths, “Recurrence plots for the analysis of complex systems,” *Phys Rep*, vol. 438, no. 5, pp. 237–329, 2007.
- [44] G. Xu, M. Liu, Z. Jiang, W. Shen, and C. Huang, “Online Fault Diagnosis Method Based on Transfer Convolutional Neural Networks,” *IEEE Trans Instrum Meas*, vol. 69, no. 2, pp. 509–520, 2020.
- [45] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016. <http://www.deeplearningbook.org>.
- [46] S. Fresca, L. Dede, and A. Manzoni, “A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs,” *J Sci Comput*, vol. 87, pp. 1–36, 2021.
- [47] W. K. Hastings, “Monte Carlo Sampling Methods Using Markov Chains and Their Applications,” *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.
- [48] H. Haario, E. Saksman, and J. Tamminen, “An adaptive Metropolis algorithm,” *Bernoulli*, vol. 7, no. 2, pp. 223–242, 2001.
- [49] W. Betz, I. Papaioannou, and D. Straub, “Transitional Markov chain Monte Carlo: observations and improvements,” *J Eng Mech*, vol. 142, no. 5, p. 04016016, 2016.
- [50] M. D. Hoffman and A. Gelman, “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo,” *J Mach Learn Res*, vol. 15, no. 1, pp. 1593–1623, 2014.
- [51] A. Gelman and D. B. Rubin, “Inference from Iterative Simulation Using Multiple Sequences,” *Stat Sci*, vol. 7, no. 4, pp. 457–472, 1992.
- [52] F. Negri, “redbKIT, version 2.2,” 2016. <http://redbkit.github.io/redbKIT>.
- [53] F. Chollet *et al.*, “Keras,” 2015. <https://keras.io>.
- [54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine Learning in Python,” *J Mach Learn Res*, vol. 12, pp. 2825–2830, 2011.
- [55] M. Ülker-Kaustell, *Some aspects of the dynamic soil-structure interaction of a portal frame railway bridge*. PhD thesis, KTH Royal Institute of Technology, 2009.
- [56] T. Arvidsson and J. Li, *Dynamic analysis of a portal frame railway bridge using frequency dependent soil structure interaction*. Master thesis, KTH Royal Institute of Technology, 2011.
- [57] European Committee for Standardization, “Part 2: Traffic loads on bridges,” in *EN 1991-2 Eurocode 1: Actions on structures*, pp. 66–74, 2003.
- [58] M. Torzoni, M. Tezzele, S. Mariani, A. Manzoni, and K. E. Willcox, “A digital twin framework for civil engineering structures,” *Comput Methods Appl Mech Eng*, vol. 418, p. 116584, 2024.
- [59] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths, “Recurrence networks—a novel paradigm for nonlinear time series analysis,” *New J Phys*, vol. 12, no. 3, p. 033025, 2010.
- [60] A. S. Campanharo, M. I. Sireer, R. D. Malmgren, F. M. Ramos, and L. A. N. Amaral, “Duality between Time Series and Networks,” *PLoS One*, vol. 6, no. 8, pp. 1–13, 2011.
- [61] J. Lin, E. Keogh, L. Wei, and S. Lonardi, “Experiencing SAX: a novel symbolic representation of time series,” *Data Min Knowl Discov*, vol. 15, pp. 107–144, 2007.
- [62] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” *J Mach Learn Res*, vol. 9, pp. 249–256, 2010.

- [63] D. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” in *Int Conf Learn Represent*, vol. 3, pp. 1–13, 2015.

## Appendix A. Imaging time series via Markov transition field

In this Appendix, we review the MTF encoding [42] employed in this work to transform multivariate time series  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{N_u}] \in \mathbb{R}^{L \times N_u}$  into images. The technique is detailed with reference to a univariate time series, and it is applied identically to all the  $N_u$  input channels.

The MTF encoding can be traced back to the use of recurrence networks to analyze the structural properties of time series. As proposed in [59], the recurrence matrix of a time series can be interpreted as the adjacency matrix of an associated complex network. In [60], the concept of building adjacency matrices has been extended as follows, by extracting transition dynamics from first order Markov matrices. Given a time series  $\mathbf{u} = (u_1, \dots, u_L)^\top$ , this is first discretized into  $N_\omega$  quantile bins. Each entry  $u_l$ ,  $l = 1, \dots, L$ , is assigned to the corresponding bin  $\omega_j$ ,  $j = 1, \dots, N_\omega$ . A weighted adjacency matrix  $\mathbf{Z} \in \mathbb{R}^{N_\omega \times N_\omega}$  is then built with entries  $z_{j,\mathcal{K}} = \bar{z}_{j,\mathcal{K}} / \sum_{\mathcal{K}} \bar{z}_{j,\mathcal{K}}$ , where  $\mathcal{K} = 1, \dots, N_\omega$  and  $\bar{z}_{j,\mathcal{K}}$  is the number of transitions  $\omega_j \rightarrow \omega_{\mathcal{K}}$  between consecutive time steps.  $\mathbf{Z}$  is a Markov transition matrix. From a network perspective, each bin represents a node and each pair of nodes is connected with a weight proportional to the probability that a data point in bin  $\omega_j$  is followed by a data point in bin  $\omega_{\mathcal{K}}$ .

The MTF encoding [42]  $\bar{\mathbf{Z}} \in \mathbb{R}^{L \times L}$  extends  $\mathbf{Z}$  by measuring the probabilities of observing a change of value between any pair of points in the time series. Similarly to  $\mathbf{Z}$ , also  $\bar{\mathbf{Z}}$  encodes the Markovian dynamics, but the transition probabilities in  $\bar{\mathbf{Z}}$  are represented sequentially to avoid losing the time dependence of the conditional relationship. The MTF matrix  $\bar{\mathbf{Z}}$  reads:

$$\bar{\mathbf{Z}} = \begin{bmatrix} z_{j,\mathcal{K}} | u_1 \in \omega_j, u_1 \in \omega_{\mathcal{K}} & \dots & z_{j,\mathcal{K}} | u_1 \in \omega_j, u_L \in \omega_{\mathcal{K}} \\ z_{j,\mathcal{K}} | u_2 \in \omega_j, u_1 \in \omega_{\mathcal{K}} & \dots & z_{j,\mathcal{K}} | u_2 \in \omega_j, u_L \in \omega_{\mathcal{K}} \\ \vdots & \ddots & \vdots \\ z_{j,\mathcal{K}} | u_L \in \omega_j, u_1 \in \omega_{\mathcal{K}} & \dots & z_{j,\mathcal{K}} | u_L \in \omega_j, u_L \in \omega_{\mathcal{K}} \end{bmatrix}, \quad (\text{A.1})$$

and measures the probability of a transition  $\omega_j \rightarrow \omega_{\mathcal{K}}$  for each pair of time steps, not necessarily consecutive. This is equivalent to spread out matrix  $\mathbf{Z}$  on the time axis by considering the temporal positions of data points in  $\mathbf{u}$ . By measuring the quantiles transition probability at two arbitrary time steps, matrix  $\bar{\mathbf{Z}}$  encodes the multi-span transition probabilities of the time series.

The MTF requires the time series to be discretized on the amplitude axis into  $N_\omega$  quantile bins. Since the time series discretization is a surjective transformation, this is not reversible and involves the loss of a certain amount of information. The information content retained in the transformation is mainly controlled by the refinement level of the discretization. With an equally spaced discretization, a large  $N_\omega$  might lead to a sparse image (not suitable for highlighting structures and patterns in the data), while a small  $N_\omega$  might lead to a substantial loss of information. To achieve a good trade-off between sparsity in the image and information loss, the symbolic aggregate approximation algorithm [61] is exploited to perform a non-uniform bin assignment. As proposed in [40], the time series is discretized in bins that roughly follow a Gaussian distribution. This non-uniform bin assignment is suitable for handling the discretization of time histories that follow long-tailed distributions, and makes the choice of the number of bins a less critical task. In the present work, the number of bins has been set to  $N_\omega = 20$ , which provides satisfactory results without yielding a significant computational burden. Finally, to make the image size manageable and improve the computational efficiency of the downstream image processing, the MTF matrix  $\bar{\mathbf{Z}}$  is downsized by averaging the pixels in each non-overlapping square patch through a blurring kernel.

## Appendix B. Implementation details

In this Appendix, we discuss the implementation details of the DL models described in Sec. 3.1. The architectures, as well as the relevant hyperparameters and training options, have been chosen through a preliminary study, aimed at minimizing  $\mathcal{L}_{\text{AE}}$  and  $\mathcal{L}_{\text{SUR}}$ , while retaining the generalization capabilities of  $\text{NN}_{\text{ENC}}$ ,  $\text{NN}_{\text{DEC}}$ , and  $\text{NN}_{\text{SUR}}$ .

$\text{NN}_{\text{ENC}}$  and  $\text{NN}_{\text{DEC}}$  are set as the encoder and decoder of a convolutional autoencoder, whose architecture is described in Tab. B.5. The encoding branch consists of a stack of four two-dimensional (2D) convolutional and max pooling layers. The output is then flattened and run through a fully-connected layer featuring  $D_h = 20$  neurons, which provides the low-dimensional feature space. This bottleneck layer is linked to the decoding branch by means of a fully-connected layer, whose output is reshaped before undergoing through a stack of four transposed 2D convolutional layers useful to reconstruct the input mosaic. All convolutional layers feature  $3 \times 3$  kernels and Softsign activation function, except the last one that is Sigmoid-activated, while the two fully-connected layers are Softsign-activated.

Using Xavier’s weight initialization [62], the loss function  $\mathcal{L}_{\text{AE}}$  is minimized using Adam [63] for a maximum of 100 allowed epochs. The learning rate  $\eta_{\text{AE}}$  is initially set to 0.001, and decreased for 4/5 of the allowed training steps using a cosine decay schedule with weight decay equal to 0.05. The optimization is carried out considering an 80 : 20 splitting ratio of the dataset for training and validation purposes. We use an early stopping strategy to interrupt learning, whenever the loss function value attained on the validation set does not decrease for a prescribed number of patience epochs in a row. The relevant hyperparameters and training options are reported in Tab. B.5.

**Table B.5.**  $\text{NN}_{\text{ENC}}, \text{NN}_{\text{DEC}}$  - employed architecture (left) and selected hyperparameters and training options (right).

Layer	Type	Output shape	Activ.	Input layer	
0	Input	$(B_{\text{AE}}, h_{\mathcal{I}}, w_{\mathcal{I}}, 1)$	None	None	
1	Conv2D	$(B_{\text{AE}}, h_{\mathcal{I}}, w_{\mathcal{I}}, 4)$	Softsign	0	Convolution kernel size: $3 \times 3$
2	MaxPool2D	$(B_{\text{AE}}, h_{\mathcal{I}}/2, w_{\mathcal{I}}/2, 4)$	None	1	$L^2$ regularization rate: $\lambda_{\text{AE}} = 10^{-4}$
3	Conv2D	$(B_{\text{AE}}, h_{\mathcal{I}}/2, w_{\mathcal{I}}/2, 8)$	Softsign	2	Weight initializer: Xavier
4	MaxPool2D	$(B_{\text{AE}}, h_{\mathcal{I}}/4, w_{\mathcal{I}}/4, 8)$	None	3	Optimizer: Adam
5	Conv2D	$(B_{\text{AE}}, h_{\mathcal{I}}/4, w_{\mathcal{I}}/4, 16)$	Softsign	4	Batch size: $B_{\text{AE}} = 128$
6	MaxPool2D	$(B_{\text{AE}}, h_{\mathcal{I}}/8, w_{\mathcal{I}}/8, 16)$	None	5	Initial learning rate: $\eta_{\text{AE}} = 0.001$
7	Conv2D	$(B_{\text{AE}}, h_{\mathcal{I}}/8, w_{\mathcal{I}}/8, 32)$	Softsign	6	Allowed epochs: 100
8	MaxPool2D	$(B_{\text{AE}}, h_{\mathcal{I}}/16, w_{\mathcal{I}}/16, 32)$	None	7	Learning schedule: $\frac{4}{5}$ cosine decay
9	Flatten	$(B_{\text{AE}}, h_{\mathcal{I}}w_{\mathcal{I}}/8)$	None	8	Weight decay: 0.05
10	Dense	$(B_{\text{AE}}, D_h = 20)$	Softsign	9	Early stop patience: 15 epochs
11	Dense	$(B_{\text{AE}}, h_{\mathcal{I}}w_{\mathcal{I}}/8)$	Softsign	10	Positive pairings: $\zeta_+ = 2$
12	Reshape	$(B_{\text{AE}}, h_{\mathcal{I}}/16, w_{\mathcal{I}}/16, 32)$	None	11	Negative pairings: $\zeta_- = 2$
13	Conv2D <sup>T</sup>	$(B_{\text{AE}}, h_{\mathcal{I}}/8, w_{\mathcal{I}}/8, 16)$	Softsign	12	Similarity margin: $\psi = 1$
14	Conv2D <sup>T</sup>	$(B_{\text{AE}}, h_{\mathcal{I}}/4, w_{\mathcal{I}}/4, 8)$	Softsign	13	Train-val split: 80 : 20
15	Conv2D <sup>T</sup>	$(B_{\text{AE}}, h_{\mathcal{I}}/2, w_{\mathcal{I}}/2, 4)$	Softsign	14	
16	Conv2D <sup>T</sup>	$(B_{\text{AE}}, h_{\mathcal{I}}, w_{\mathcal{I}}, 1)$	Sigmoid	15	

$\text{NN}_{\text{SUR}}$  consists of four fully-connected layers featuring 10, 10, 40 and  $D_h = 20$  neurons, respectively. The three hidden layers are Softsign-activated, while no activation is applied to the output layer. The architecture of  $\text{NN}_{\text{SUR}}$  is outlined in Tab. B.6. Also in this case, the optimization is carried out using Adam together with the Xavier’s weight initialization. The learning rate  $\eta_{\text{SUR}}$  is decreased as the training advances using a cosine decay schedule. An early stop strategy is employed to prevent overfitting, by considering an 80 : 20 splitting ratio for training and validation purposes. The relevant hyperparameters and the training options are summarized in Tab. B.6.

**Table B.6.**  $\text{NN}_{\text{SUR}}$  - employed architecture (left) and selected hyperparameters and training options (right).

Layer	Type	Output shape	Activ.	Input layer	
0	Input	$(B_{\text{SUR}}, N_{\text{par}})$	None	None	
1	Dense	$(B_{\text{SUR}}, 10)$	Softsign	0	$L^2$ regularization rate: $\lambda_{\text{SUR}} = 10^{-4}$
2	Dense	$(B_{\text{SUR}}, 10)$	Softsign	1	Weight initializer: Xavier
3	Dense	$(B_{\text{SUR}}, 40)$	Softsign	2	Optimizer: Adam
4	Dense	$(B_{\text{SUR}}, D_h = 20)$	None	3	Batch size: $B_{\text{SUR}} = 128$

Initial learning rate:	$\eta_{\text{SUR}} = 0.001$
Allowed epochs:	5000
Learning schedule:	$\frac{4}{5}$ cosine decay
Weight decay:	0.01
Early stop patience:	100 epochs
Train-val split:	80 : 20

## MOX Technical Reports, last issues

Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 05/2025** Buchwald, S.; Ciaramella, G.; Verani, M.  
*Greedy reconstruction algorithms for function approximation*
- 04/2025** Andrini, D.; Magri, M.; Ciarletta, P.  
*Nonlinear morphoelastic theory of biological shallow shells with initial stress*
- 02/2025** Corda, A.; Pagani, S.; Vergara, C.  
*Influence of patient-specific acute myocardial ischemia maps on arrhythmogenesis: a computational study*
- 01/2025** Dede', L.; Parolini, N.; Quarteroni, A.; Villani, G.; Ziarelli, G.  
*SEIHRDV: a multi-age multi-group epidemiological model and its validation on the COVID-19 epidemics in Italy*
- 109/2024** Liverotti, L.; Ferro, N.; Matteucci, M.; Perotto, S.  
*A PCA and mesh adaptation-based format for high compression of Earth Observation optical data with applications in agriculture*
- 110/2024** Pederzoli, V.; Corti, M.; Riccobelli, D.; Antonietti, P.F.  
*A coupled mathematical and numerical model for protein spreading and tissue atrophy, applied to Alzheimer's disease*
- 108/2024** Arostica, R.; Nolte, D.; Brown, A.; Gebauer, A.; Karabelas, E.; Jilberto, J.; Salvador, M.; Bucelli, M.; Piersanti, R.; Osouli, K.; Augustin, C.; Finsberg, H.; Shi, L.; Hirschvogel, M.; Pfaller, M.; Africa, P.C.; Gsell, M.; Marsden, A.; Nordsletten, D.; Regazzoni, F.; Plank, G.; Sundnes, J.; Dede', L.; Peirlinck, M.; Vedula, V.; Wall, W.; Bertoglio, C.  
*A software benchmark for cardiac elastodynamics*
- 107/2024** Chen, J.; Ballini, E.; Micheletti, S.  
*Active Flow Control for Bluff Body under High Reynolds Number Turbulent Flow Conditions Using Deep Reinforcement Learning*
- 106/2024** Brunati, S.; Bucelli, M.; Piersanti, R.; Dede', L.; Vergara, C.  
*Coupled Eikonal problems to model cardiac reentries in Purkinje network and myocardium*
- 105/2024** Bartsch, J.; Barakat, A.A.; Buchwald, S.; Ciaramella, G.; Volkwein, S.; Weig, E.M.  
*Reconstructing the system coefficients for coupled harmonic oscillators*