



Article

A Frequency-Based Approach for the Detection and Classification of Structural Changes Using t -SNE[†]

David Agis  and Francesc Pozo * 

Control, Modeling, Identification and Applications (CoDALab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya (UPC), Campus Diagonal-Besòs (CDB), Eduard Maristany, 16, 08019 Barcelona, Spain; david.agis@upc.edu

* Correspondence: francesc.pozo@upc.edu; Tel.: +34-934-137-316

† This paper is an extended version of our paper published in International Conference on Structural Engineering Dynamics (ICEDyn 2019).

Received: 24 October 2019; Accepted: 20 November 2019; Published: 21 November 2019



Abstract: This work presents a structural health monitoring (SHM) approach for the detection and classification of structural changes. The proposed strategy is based on t -distributed stochastic neighbor embedding (t -SNE), a nonlinear procedure that is able to represent the local structure of high-dimensional data in a low-dimensional space. The steps of the detection and classification procedure are: (i) the data collected are scaled using mean-centered group scaling (MCGS); (ii) then principal component analysis (PCA) is applied to reduce the dimensionality of the data set; (iii) t -SNE is applied to represent the scaled and reduced data as points in a plane defining as many clusters as different structural states; and (iv) the current structure to be diagnosed will be associated with a cluster or structural state based on three strategies: (a) the smallest point-centroid distance; (b) majority voting; and (c) the sum of the inverse distances. The combination of PCA and t -SNE improves the quality of the clusters related to the structural states. The method is evaluated using experimental data from an aluminum plate with four piezoelectric transducers (PZTs). Results are illustrated in frequency domain, and they manifest the high classification accuracy and the strong performance of this method.

Keywords: classification detection; principal component analysis (PCA); structural changes; structural health monitoring (SHM); t -distributed stochastic neighbor embedding (t -SNE)

1. Introduction

Structural health monitoring (SHM) is a crucial process for engineering structures because it checks the correct behavior of the structure and determines whether it needs some type of maintenance. The healthy state of the structure has to remain between the specified limits or threshold, but these limits may change due to the aging of the structure and its use, or due to the environmental and operational conditions (EOC). Hence, in SHM systems, detection and classification of structural changes are essential in order to know the current state of the structure for security and to reduce costs of inspection and maintenance. If damage is detected and classified precisely at the time it occurs, some action may be taken before a human and/or economic disaster occurs, thus reducing the probability of accidents and the maintenance costs. SHM has been applied in many structures such as wind turbines [1–3], buildings [4,5], and aircraft [6,7], among others, and a review of the state-of-the-art manifests that SHM is a very active research field.

With the goal of obtaining information about the state of the structure, data are collected by a sensor network, which is placed along the structure. The information obtained from multi-sensor signals

creates a high-dimensional data set with a large volume of data due to continuous measurements of the monitoring system. Various methods have been proposed for the handling of high-dimensional, big, and complex data. Among these methods, plane or spatial representation techniques stick out as they offer a way to handle this type of data by means of an interface that allows an easy detection of natural clusters, identifying hidden patterns, et cetera [8]. Plane or spatial representation techniques are also somehow related to dimensionality reduction. Dimensionality reduction is the mechanism of reducing the dimension of the original data, while keeping mostly the same intrinsic information [9]. One of the proposed dimensionality reduction methods in the literature is *t*-distributed stochastic neighbor embedding (*t*-SNE), a technique developed by L. van der Maaten and G. Hinton [10], which is able to represent the local structure of original high-dimensional data in a low-dimensional space (for example, a simple 2-D plot). This technique detects patterns by identifying clusters based on similarity of data points. *t*-SNE is widely used in the literature as a dimensionality reduction technique, as a classification or pattern recognition method, or as a visualization and compression method of big data sets, but although *t*-SNE has been applied in several applications, this is one of the first approaches of *t*-SNE in the field of SHM [11].

In the present approach, *t*-SNE is applied in the frequency domain. In the field of SHM and condition monitoring (CM) this is sometimes common, and the combination of time-frequency domain is also used. Some examples are Tsogka et al. [12], who propose a novel vibration-based SHM method for damage detection in the frequency domain, which illustrates its practical application in the case of a historic bell tower. Xu et al. [13] propose a clustering method based on ensemble empirical mode decomposition and affinity propagation for bearing performance degradation assessment. To prove the superiority of the approach, the proposed methodology is compared to various popular clustering methods and commonly used time-domain indicators. The results show that the proposed method outperforms these popular clustering methods and time-domain indicators. Cheng et al. [14] propose a multisensory data-driven health degradation monitoring system by using a generalized multiclass support vector machine. In this method, multidimensional feature extraction is implemented in the time domain, frequency domain, and time-frequency domain.

In this work, a SHM strategy for detection and classification of structural changes based on a two-step data integration (type *E* unfolding [15] and the so-called mean-centered group scaling (MCGS)), data transformation using PCA, and a two-step data reduction combining PCA and *t*-SNE has been proposed. PCA is an extensively used technique that is mainly used for dimensionality reduction or feature extraction in the framework of pattern recognition [16], and it can be applied differently to detect and classify structural changes or faults [17]. In some cases, however, it can be observed that the projection into the first principal components does not allow a visual grouping, clustering, or separation. For this reason, we propose the damage or fault detection based on the combination of PCA and *t*-SNE. As a consequence, the basic steps of the detection and classification procedure that we apply are: (i) the data collected, in the frequency domain, are first scaled using MCGS due to the different scales and magnitudes in the measurements; (ii) then PCA is applied to obtain a better representation of the original data, by reducing the dimensionality of the scaled data and projecting the scaled data into the vectorial space spanned by the principal components; and (iii) *t*-SNE is finally applied to the projected data to represent these points as points in a plane. It will be shown that, with respect to the time domain, the quality of the clusters related to the different structural states is significantly improved. More precisely, the current structure to be diagnosed will then be associated with a structural state based on three different strategies: (i) the smallest point-centroid distance, i.e., when a single actuation phase is considered; (ii) majority voting; and (iii) sum of the inverse distances, i.e., when several actuation phases are combined. Therefore, in this work, *t*-SNE is used, in combination with a particular data integration, data transformation, and data reduction, for the first time in the field of SHM in a frequency-based approach. In comparison to previous

strategies found in the literature, this novel method is able to yield a best detection and classification of structural changes, thus leading to a best performance.

The proposed method for the detection and classification of structural changes is assessed using experimental data from a plate with four piezoelectric transducers (PZTs). Since guided wave propagation-based SHM strategies have proven their ability to adequately identify defects in structures [18–21], in the present work, we have also considered the paradigm of guided waves. In this paradigm, the structure is excited by a signal and the response is measured to create a baseline pattern. When a new structure has to be diagnosed, it has to be excited by the exact same signal and the response is measured and compared with the baseline pattern. Results reveal the high classification accuracy and the strong performance of this methodology, with a percentage of correct decisions of about 100% in various scenarios. In the present work, the environmental conditions were not considered, as it will be the topic for further developments.

The structure of the paper is as follows: Section 2 describes the objective of t -SNE and how the plane or spatial representation is obtained. Section 3 includes how the baseline data are collected and pre-processed, how the global dimension of the data is reduced, and how the clusters are created using t -SNE. The damage detection and classification procedure of a structure that has to be diagnosed is presented in Section 4. The experimental case study is described in Section 5. In Section 6, the results are shown. Finally, in Section 7, some conclusions are drawn.

2. t -Distributed Stochastic Neighbor Embedding (t -SNE)

2.1. The Objective of t -SNE

t -SNE is an improved variation of the technique so-called stochastic neighbor embedding (SNE) [22]. With respect to SNE, t -SNE is much easier to optimize and yields better plane or spatial representations of the high-dimensional data, since it reduces the tendency to crowd points in the center of the distribution (the so-called *crowding problem*). Part of the enhancements of t -SNE with respect to SNE are due to the fact that the cost function used by t -SNE differs from the one used by SNE in two features: (i) t -SNE uses a symmetrized version of the SNE cost function with simpler gradients; and (ii) t -SNE uses a Student's t -distribution, instead of a Gaussian, to calculate the similarity between two points in the low-dimensional space.

Given a collection of high-dimensional data points:

$$\mathcal{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^v\} \subset \mathbb{R}^D, \quad v, D \in \mathbb{N}, \quad (1)$$

the aim is to find a collection of low-dimensional map points

$$\mathcal{Y} = \{\mathbf{y}^1, \dots, \mathbf{y}^v\} \subset \mathbb{R}^d, \quad d \in \mathbb{N}$$

that form a faithful representation of the original points $\{\mathbf{x}^1, \dots, \mathbf{x}^v\}$ in a lower-dimensional space. Typical values for d are 2 (*plane* representation) or 3 (*spatial* representation), where $d \ll D$. By saying *faithful* representation, we mean that the points $\{\mathbf{y}^1, \dots, \mathbf{y}^v\}$ in the lower-dimensional space preserve, as much as possible, the local structure of the original data \mathcal{X} .

2.2. Pairwise Similarities

To preserve local similarities of the original data \mathcal{X} by this embedding, t -SNE first converts the high-dimensional Euclidean distances between data points \mathbf{x}^i and \mathbf{x}^j into conditional probabilities by

centering a Gaussian distribution at \mathbf{x}^i , computing the density of \mathbf{x}^j under this Gaussian distribution, and renormalizing:

$$p_{j|i} = \frac{\exp\left(\frac{-\|\mathbf{x}^i - \mathbf{x}^j\|_2^2}{2\sigma_i^2}\right)}{\sum_{\substack{l=1 \\ l \neq i}}^v \exp\left(\frac{-\|\mathbf{x}^i - \mathbf{x}^l\|_2^2}{2\sigma_i^2}\right)}, \quad i, j = 1, \dots, v, i \neq j, \quad (2)$$

where $\frac{\|\mathbf{x}^i - \mathbf{x}^j\|_2^2}{2\sigma_i^2}$ (affinity or scaled squared Euclidean distance) is the dissimilarity between data points \mathbf{x}^i and \mathbf{x}^j . The variance of the Gaussian distribution, σ_i^2 , is computed automatically. Since only pairwise similarities between data points are of interest, t -SNE imposes $p_{i|i} = 0$. This conditional probability measures the similarity of \mathbf{x}^j to \mathbf{x}^i . If two data points are close, $p_{j|i}$ will be large. However, if two data points are far, $p_{j|i}$ will be small.

Then, by symmetrizing the conditional probability in Equation (2), the joint probability is defined as follows:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2v}, \quad i, j = 1, \dots, v, i \neq j, \\ p_{ii} = 0.$$

The joint probability also measures the pairwise similarity between data points \mathbf{x}^i and \mathbf{x}^j . As a result, let us define the similarity matrix $\mathcal{P} \in \mathcal{M}_{v \times v}(\mathbb{R})$ for the high-dimensional data points as $\mathcal{P} = (p_{ij})_{i,j=1,\dots,v}$.

When the similarity matrix for the data points \mathcal{X} in Equation (1) is obtained, t -SNE also defines the similarity matrix $\mathcal{Q} \in \mathcal{M}_{v \times v}(\mathbb{R})$ for the map points \mathcal{Y} . Essentially, we build matrix \mathcal{Q} following the same idea as for the similarity matrix \mathcal{P} with respect to original data points. The one and only difference is that we use for matrix \mathcal{Q} a renormalized Student's t -distribution with one degree of freedom and $\sigma_i^2 = \frac{1}{2}$ for all i , instead of a Gaussian distribution:

$$q_{ij} = \frac{1}{\frac{1 + \|\mathbf{y}^i - \mathbf{y}^j\|_2^2}{\sum_{k=1}^v \sum_{\substack{l=1 \\ l \neq k}}^v \frac{1}{1 + \|\mathbf{y}^k - \mathbf{y}^l\|_2^2}}}, \quad i, j = 1, \dots, v, i \neq j, \quad (3)$$

$$q_{ii} = 0, \quad (4)$$

where q_{ij} represents the local structure of the data points in the low-dimensional space.

2.3. Comparing Similarity Matrices: Cost Function

The goal is to select the map points so that the two similarity matrices, \mathcal{P} and \mathcal{Q} , are as similar as possible. The similarity between these two matrices will be defined in terms of the Kullback–Leibler (KL) divergence. The KL divergence between the joint probability distributions \mathcal{P} and \mathcal{Q} is a measure of the *distance* between the two similarity matrices, and it can be defined as [10,22,23]:

$$\mathcal{C} = \mathcal{D}_{KL}(\mathcal{P} \parallel \mathcal{Q}) = \sum_{i=1}^v \sum_{\substack{j=1 \\ j \neq i}}^v p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right). \quad (5)$$

Therefore, minimizing the KL divergence reduces the *distance* between these two matrices. And to minimize the cost function \mathcal{C} , the gradient descent method is used: $\frac{\partial \mathcal{C}}{\partial \mathbf{y}^i}$. It is worth noting that the gradient descent is an iterative optimization algorithm and therefore it updates the map point \mathbf{y}^i at each step.

For more details, see the original *t*-SNE paper [10].

3. Data Collection, Pre-Processing, and Clustering: Baseline Data

3.1. Data Collection and Pre-Processing

The data collected are made up of different response signals measured, in the time domain, by sensors on a vibrating structure. Multiple realizations of these responses are measured, under different structural states. Coming up next, these responses signals are transformed into the frequency domain using the fast Fourier transform (FFT) algorithm, and features are extracted from the spectrum to reduce the data dimension, dividing by two and adding one to the number of components in each signal. A matrix which collects all the realizations under different structural states in the frequency domain is defined as:

$$\mathbf{X} = \left(x_{i,l}^{k,j} \right) = \begin{bmatrix} x_{1,1}^{1,1} & \cdots & x_{1,1}^{1,L} & \cdots & x_{1,1}^{N,1} & \cdots & x_{1,1}^{N,L} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n_1,1}^{1,1} & \cdots & x_{n_1,1}^{1,L} & \cdots & x_{n_1,1}^{N,1} & \cdots & x_{n_1,1}^{N,L} \\ \hline x_{1,2}^{1,1} & \cdots & x_{1,2}^{1,L} & \cdots & x_{1,2}^{N,1} & \cdots & x_{1,2}^{N,L} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n_2,2}^{1,1} & \cdots & x_{n_2,2}^{1,L} & \cdots & x_{n_2,2}^{N,1} & \cdots & x_{n_2,2}^{N,L} \\ \hline \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{1,E}^{1,1} & \cdots & x_{1,E}^{1,L} & \cdots & x_{1,E}^{N,1} & \cdots & x_{1,E}^{N,L} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n_E,E}^{1,1} & \cdots & x_{n_E,E}^{1,L} & \cdots & x_{n_E,E}^{N,1} & \cdots & x_{n_E,E}^{N,L} \end{bmatrix} \quad (6)$$

$$= \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_E \end{bmatrix} \in \mathcal{M}_{(n_1+\dots+n_E) \times (N \cdot L)}(\mathbb{R}), \quad (7)$$

$$= \left[\mathbf{X}^1 \mid \mathbf{X}^2 \mid \dots \mid \mathbf{X}^N \right], \quad (8)$$

where $N \in \mathbb{N}$ is the number of sensors and $k = 1, \dots, N$ identifies the sensor that is measuring; $L \in \mathbb{N}$ is the number of components in each signal and $j = 1, \dots, L$ indicates the j -th measurement in the frequency domain; $E \in \mathbb{N}$ is the number of different structural states that are considered and $l = 1, \dots, E$ represents the structural state that is been measured; and finally, n_l , $l = 1, \dots, E$, is the number of realizations per structural state and $i = 1, \dots, n_l$ is the i -th realization related to the l -th structural state. Note that matrix \mathbf{X} in Equation (7) is formed by E horizontal blocks, \mathbf{X}_l , $l = 1, \dots, E$, each one of them related to the different structural states. At the same time, this matrix \mathbf{X} can also be viewed as formed by N vertical blocks, \mathbf{X}^k , $k = 1, \dots, N$, each one of them related to the different sensors. Due to the different scales and magnitudes in the measurements, the matrix \mathbf{X} in Equation (7) is rescaled using MCGS, which is suggested by Pozo et al. [24].

3.2. Dimensionality Reduction

One of the main reasons of using the MCGS is that the covariance matrix of matrix $\check{\mathbf{X}}$, i.e., the scaled data set, can be computed in a very simple way as:

$$\mathbf{C}_{\check{\mathbf{X}}} = \frac{1}{n-1} \check{\mathbf{X}}^T \check{\mathbf{X}} \in \mathcal{M}_{(N \cdot L) \times (N \cdot L)}(\mathbb{R}). \quad (9)$$

The eigenvectors ρ_k , $k = 1, \dots, N \cdot L$, and eigenvalues λ_k , $k = 1, \dots, N \cdot L$, of the covariance matrix $\mathbf{C}_{\check{\mathbf{X}}}$ define the subspaces in the PCA model. The eigenvalues λ_k are then ordered in decreasing order as

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N \cdot L},$$

and the matrix $\mathbf{P} \in \mathcal{M}_{(N \cdot L) \times (N \cdot L)}(\mathbb{R})$ contains, written as columns, their corresponding eigenvectors ρ_k . These eigenvectors are known as the principal components. The eigenvalues define the partial variance of each eigenvector. When the *column scaling* is applied to matrix \mathbf{X} in Equation (7), although not the case of this paper, we have that the trace of the covariance matrix $\mathbf{C}_{\check{\mathbf{X}}}$, which is the sum of the eigenvalues, is equal to the number of columns of \mathbf{X} , that is, $N \cdot L$. This means that the proportion of the variance directed along the first $\ell \in \mathbb{N}$ principal components is given by $\frac{\lambda_1 + \dots + \lambda_\ell}{N \cdot L}$. However, when the MCGS is applied to scale the raw data in matrix \mathbf{X} in Equation (7), the trace of the covariance matrix $\mathbf{C}_{\check{\mathbf{X}}}$ is no longer necessarily equal to $N \cdot L$. As a consequence, the proportion of the variance explained by the first ℓ principal components is given by:

$$\frac{\lambda_1 + \dots + \lambda_\ell}{\lambda_1 + \dots + \lambda_{N \cdot L}}.$$

In this work, we use PCA to reduce the dimensionality of the scaled data set $\check{\mathbf{X}}$ by selecting a reduced, but still significant, number $\ell = D < N \cdot L$ of principal components. This dimensionality reduction is performed through the reduced matrix

$$\mathbf{P}_\ell = \mathbf{P}_D = \begin{bmatrix} \rho_1 & \rho_2 & \dots & \rho_\ell \end{bmatrix} \in \mathcal{M}_{(N \cdot L) \times \ell}(\mathbb{R}), \quad (10)$$

which is composed by the concatenation of the eigenvectors ρ_k related to the highest eigenvalues. Matrix $\mathbf{P}_\ell = \mathbf{P}_D$ in Equation (10) is the *model* or *PCA model*. The scaled data set $\check{\mathbf{X}}$ is then projected into the vectorial space spanned by the $\ell = D$ first principal components through the premultiplication of \mathbf{P}_ℓ by $\check{\mathbf{X}}$. More precisely,

$$\mathbf{T}_\ell = \mathbf{T}_D = \check{\mathbf{X}} \mathbf{P}_\ell \in \mathcal{M}_{n \times \ell}(\mathbb{R}). \quad (11)$$

\mathbf{P}_ℓ in Equation (10) has been defined as the PCA model that includes multiple realizations under different structural states. At the same time, \mathbf{T}_ℓ in Equation (11) is the projection of the scaled data set $\check{\mathbf{X}}$ into the subspace spanned by the PCA model. The number of principal components $\ell = D \in \mathbb{N}$ is chosen so that the proportion of the variance explained is greater than or equal to 95%.

3.3. Clustering Effect

In Section 3.2, the dimensionality reduction has been performed. More precisely, n realizations under different structural states (the rows of matrix \mathbf{X} in Equation (7)), that may be seen as $N \cdot L$ -dimensional vectors, are projected and transformed into $\ell = D$ -dimensional vectors. This reduction of the dimension

of the original data is performed with a small loss of information, less than 5%, and it is also expected that ℓ is much smaller than $N \cdot L$.

A second transformation is performed to the projected data in matrix \mathbf{T}_ℓ in Equation (11) using the t -SNE presented in Section 2. Let us define

$$\mathbf{x}^i = \mathbf{e}_i^\top \mathbf{T}_\ell = \mathbf{e}_i^\top \check{\mathbf{X}} \mathbf{P}_\ell \in \mathbb{R}^\ell, \quad i = 1, \dots, n$$

as the i -th row of matrix \mathbf{T}_ℓ in Equation (11). The vector $\mathbf{e}_i \in \mathbb{R}^n$ is the i -th element of the canonical basis. Let us also define

$$\mathcal{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^n\} \subset \mathbb{R}^\ell \quad (12)$$

as a collection of high-dimensional data points. The objective is to find a collection of 2-dimensional map points

$$\mathcal{Y} = \{\mathbf{y}^1, \dots, \mathbf{y}^n\} \subset \mathbb{R}^2$$

that represent the original set \mathcal{X} with no explicit loss of information and preserving the local structure of this set. After the application of t -SNE, we expect E clusters to be observed, related to the E different structural states. These clusters are formed by the map points:

$$\begin{aligned} \{\mathbf{y}^1, \dots, \mathbf{y}^{n_1}\} &\subset \mathcal{Y}, \text{ related to the } 1\text{st structural state;} \\ \{\mathbf{y}^{n_1+1}, \dots, \mathbf{y}^{n_1+n_2}\} &\subset \mathcal{Y}, \text{ related to the } 2\text{nd structural state;} \\ \{\mathbf{y}^{n_1+n_2+1}, \dots, \mathbf{y}^{n_1+n_2+n_3}\} &\subset \mathcal{Y}, \text{ related to the } 3\text{rd structural state;} \\ &\vdots \\ \{\mathbf{y}^{n-n_{E-1}-n_E+1}, \dots, \mathbf{y}^{n-n_E}\} &\subset \mathcal{Y}, \text{ related to the } penultimate \text{ structural state;} \text{ and} \\ \{\mathbf{y}^{n-n_E+1}, \dots, \mathbf{y}^n\} &\subset \mathcal{Y}, \text{ related to the } last \text{ structural state.} \end{aligned} \quad (13)$$

4. Damage Detection and Classification Procedure: Structure to Diagnose

In Section 3.3, we have seen how the original realizations under different structural states are finally projected on a plane to define a set of clusters. In this section, we will present the damage detection and classification procedure of a structure that has to be diagnosed.

A single realization of the current structure to diagnose is needed. The data collected are made up, in this case, of different response signals measured by the same number of sensors N and the same number of components in each signal L , as in Equation (7). When these measures are obtained and they are transformed into the frequency domain, a new data vector \mathbf{z} is constructed:

$$\mathbf{z}^\top = \left[z^{1,1} \quad \dots \quad z^{1,L} \mid \dots \mid z^{N,1} \quad \dots \quad z^{N,L} \right] \in \mathbb{R}^{N \cdot L}.$$

4.1. Scaling (MCGS)

Before the collected data coming from the structure to diagnose is projected into the space spanned by the principal components, the row vector \mathbf{z}^\top has to be scaled to define a scaled row vector $\check{\mathbf{z}}^\top$:

$$\check{z}^{k,j} = \frac{z^{k,j} - \mu^{k,j}}{\sigma^k}, \quad k = 1, \dots, N, \quad j = 1, \dots, L, \quad (14)$$

where μ^{kj} is the arithmetic mean of all the elements in the $[(k-1)L+j]$ -th column of matrix \mathbf{X} in Equation (6) i.e., the j -th column of the vertical block \mathbf{X}^k in Equation (8); and σ^k is the standard deviation of all the elements in the vertical block \mathbf{X}^k in Equation (8) with respect to the mean value μ^k (the arithmetic mean of all the elements in the vertical block \mathbf{X}^k in Equation (8)).

4.2. Projection (PCA)

The projection of the scaled row vector $\mathbf{z}^\top \in \mathbb{R}^{N \cdot L}$ into the space spanned by the ℓ first principal components in \mathbf{P}_ℓ is performed through the following vector to matrix multiplication:

$$\mathbf{x}^{n+1} = \mathbf{z}^\top \cdot \mathbf{P}_\ell \in \mathbb{R}^\ell.$$

It is worth noting that the $N \cdot L$ -dimensional vector that contains the collected data coming from the structure to be diagnosed is now transformed into an ℓ -dimensional vector. This new point will be added to the data set \mathcal{X} in Equation (12) to define a new set:

$$\mathcal{X}' = \mathcal{X} \cup \{\mathbf{x}^{n+1}\} = \{\mathbf{x}^1, \dots, \mathbf{x}^n, \mathbf{x}^{n+1}\} \subset \mathbb{R}^\ell. \quad (15)$$

4.3. t-SNE and Final Classification

t-SNE is applied to the ℓ -dimensional data set \mathcal{X}' in Equation (15) to find a collection of 2-dimensional map points:

$$\mathcal{Y}' = \{\mathbf{y}^1, \dots, \mathbf{y}^n, \mathbf{y}^{n+1}\} \subset \mathbb{R}^2$$

that represent the original set \mathcal{X} with no explicit loss of information and preserving the local structure of this set, as well as including the map point \mathbf{y}^{n+1} associated to the data point \mathbf{x}^{n+1} . The same E clusters have to be observed, related to the E different structural states. As in Section 3.3, these clusters are formed by the map points in Equation (13).

For each cluster, we compute its centroid, that is, the mean of the values of the points of data in the cluster. For instance, the centroid associated with the first structural state is:

$$\mathcal{Y}_1 := \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{y}^i = \frac{\mathbf{y}^1 + \dots + \mathbf{y}^{n_1}}{n_1} \in \mathbb{R}^2,$$

whereas the centroid associated with the second structural state is:

$$\mathcal{Y}_2 := \frac{1}{n_2} \sum_{i=1}^{n_2} \mathbf{y}^{n_1+i} = \frac{\mathbf{y}^{n_1+1} + \dots + \mathbf{y}^{n_1+n_2}}{n_2} \in \mathbb{R}^2.$$

In general, the centroid associated with the l -th structural state, $l = 1, \dots, E$, is the point of the plane defined as

$$\mathcal{Y}_l := \frac{1}{n_l} \sum_{i=1}^{n_l} \mathbf{y}^{\left(\sum_{j=0}^{l-1} n_j\right)+i} \in \mathbb{R}^2, \quad l = 1, \dots, E, \quad (16)$$

where $n_0 = 0$. Therefore, the current structure to diagnose is associated to the l -th structural state if

$$l = \arg \min_{l=1, \dots, E} \|\mathcal{Y}_l - \mathbf{y}^{n+1}\|_2,$$

that is, if the minimum distance between y^{n+1} and each one of the centroids corresponds to the Euclidean distance between y^{n+1} and \mathcal{Y}_j . We call this approach the smallest point-centroid distance (see Figure 1).

A flowchart of the proposed approach and how it is applied is given in Figure 2.

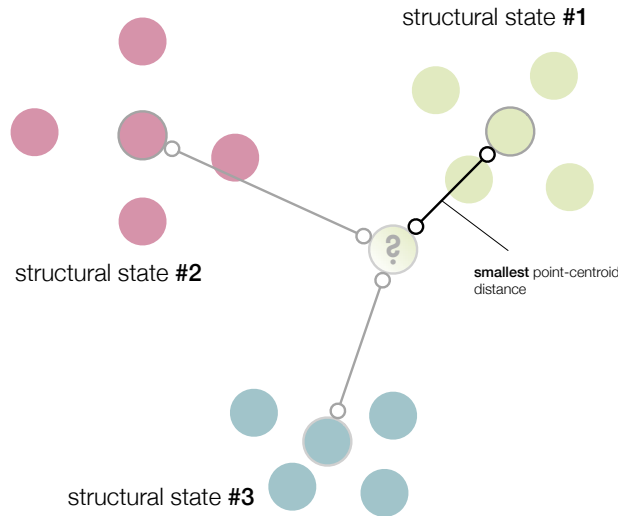


Figure 1. The current structure to diagnose is associated to the structural state with the smallest point-centroid distance.

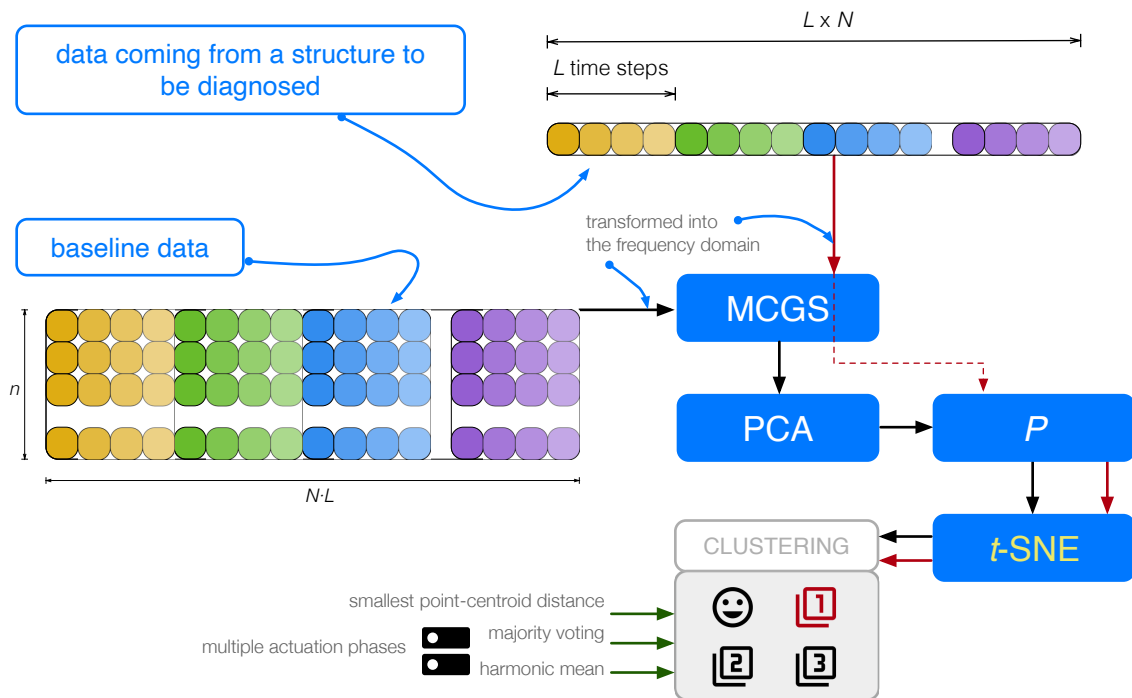


Figure 2. Flowchart of the proposed approach. Data coming from a structure are first transformed into the frequency domain and scaled, and then projected into the principal component analysis (PCA) model. Finally, *t*-distributed stochastic neighbor embedding (*t*-SNE) is used to create the clusters that will be used in the detection and classification of structural changes. MCGS = mean-centered group scaling.

5. Case Study: Aluminum Plate with Four PZTs

5.1. Structure

In this section, a square aluminum plate with an area of 1600 cm^2 ($40 \text{ cm} \times 40 \text{ cm}$, and a thickness of 0.2 cm) and instrumented with four PZTs is considered to demonstrate the reliability of the damage detection and classification methodology introduced in Sections 3 and 4. The piezoelectric transducer discs are attached to the surface and their location is shown in Figure 3. Assuming that the lower left corner of the plate in Figure 3 represents the origin of coordinates, the PZTs are installed at these positions (units in centimeters):

- PZT1 at (20, 35)
- PZT2 at (35, 20)
- PZT3 at (20, 5)
- PZT4 at (5, 20)

These PZTs are able to work both in actuator mode and in sensor mode. In actuator mode, the burst signal in Figure 4 is applied to the PZTs, and they produce a mechanical vibration; and in sensor mode, they detect time varying mechanical response. It is worth keeping in mind that the distance between the four sensors is not the same. More precisely, for example, the distance between sensor 1 and sensor 2 and the distance between sensor 1 and sensor 4 is equal. However, the distance between sensor 1 and sensor 3 is relatively larger.

A 17.2916 grams mass is added to simulate the damage, in a non-destructive way, in the aluminum plate. This mass is an attached magnet in both sides of the plate, since aluminum is non-magnetic metal. This kind of damage is used to change the properties of the structure and to produce changes in the propagated wave, therefore providing different scenarios for validating the proposed method. The location of the mass defines each damage. These locations are (units in centimeters):

- damage 1 at (12.5, 27.5)
- damage 2 at (27.5, 27.5)
- damage 3 at (12.5, 12.5)

$E = 4$ structural states are considered here:

- the first structural state corresponds to the healthy state of the structure, that is, the square aluminum plate with no damage;
- the second, third, and fourth structural states correspond to the plate with an added mass at the positions indicated in Figure 3 as damage 1, damage 2, and damage 3, respectively.

The aluminum plate is isolated from the vibration and noise that could affect the laboratory, as can be observed in Figure 5.

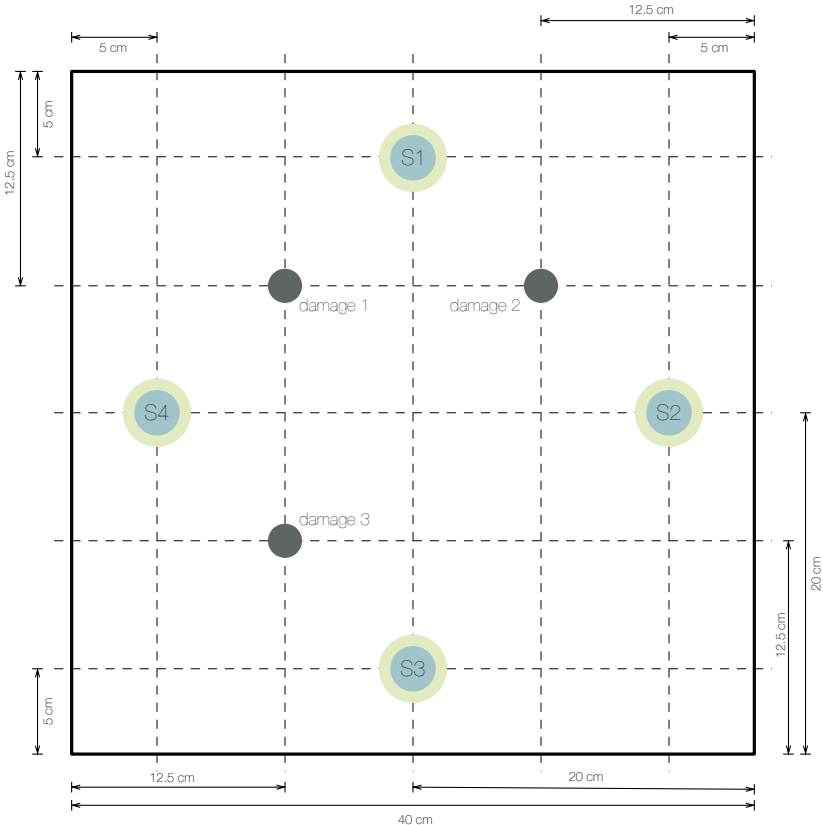


Figure 3. Aluminum plate instrumented with four piezoelectric sensors (S1, S2, S3, and S4).

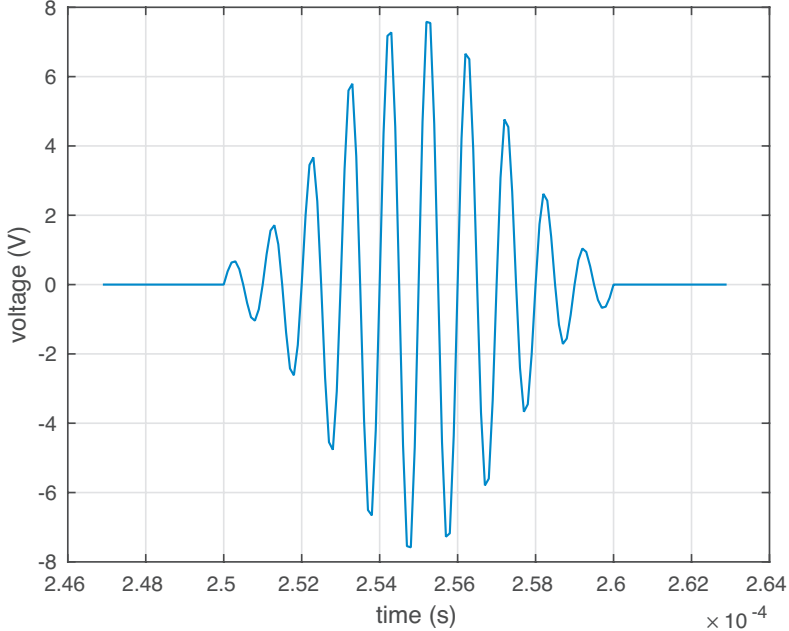


Figure 4. In actuator mode, this burst signal is applied to the piezoelectric transducers (PZTs) to produce a mechanical vibration.

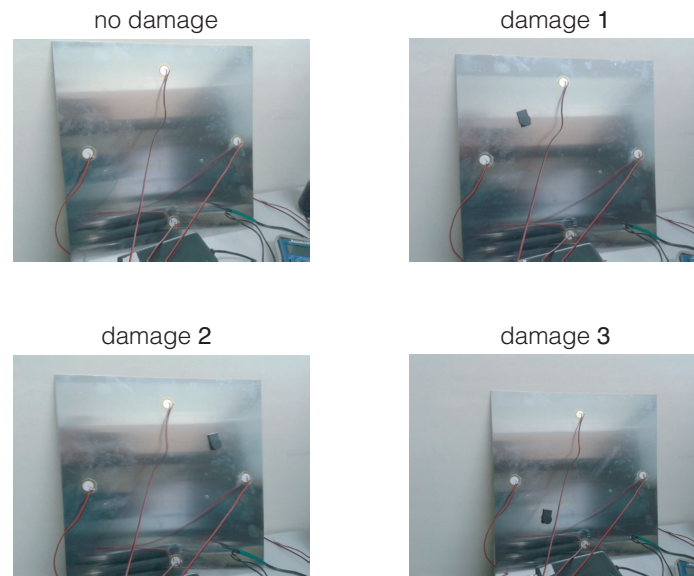


Figure 5. Aluminum plate with four PZTs and with four different structural states.

5.2. Scenarios and Actuation Phases

The experimental setup includes three different scenarios to determine the behavior of the methodology under the presence of white Gaussian noise, filters, and with respect to the length of the wire that is used from the digitizer to the sensors:

- **Scenario 1.** The signals are obtained using a short wire (0.5 m) from the digitizer to the PZTs, and these signals are filtered with a Savitzky–Golay (SG) [25] filter algorithm after adding white Gaussian noise. The filter is applied for the intention of smoothing the data.
- **Scenario 2.** The signals are obtained using a short wire (0.5 m) from the digitizer to the PZTs, but these signals are not filtered.
- **Scenario 3.** The signals are obtained using a long wire (2.5 m) from the digitizers to the PZTs. Signals are also filtered with the SG algorithm.

In this manner, we can observe the effect of the attenuation with short and long wires, the effect of adding white Gaussian noise to the measured signals, and the effect of the use of a SG filter in the detection and classification procedure.

As stated in Section 5.1, four PZTs (PZT1, PZT2, PZT3, and PZT4) are used to excite the aluminum plate and collect the measured response. This sensor network works in what we call *actuation phases*. In each actuation phase, a single PZT is used as an actuator (active sensor: the PZT excites the structure with a given excitation signal), and the rest of the PZTs are used as sensors (passive sensors: PZTs measure signals). Therefore, we have as many *actuation phases* as *sensors*:

- **Actuation phase 1.** PZT1 is used as the actuator, and PZT2, PZT3, and PZT4 are used as sensors.
- **Actuation phase 2.** PZT2 is used as the actuator, and PZT1, PZT3, and PZT4 are used as sensors.
- **Actuation phase 3.** PZT3 is used as the actuator, and PZT1, PZT2, and PZT4 are used as sensors.
- **Actuation phase 4.** PZT4 is used as the actuator, and PZT1, PZT2, and PZT3 are used as sensors.

It is very common in the literature, when using a sensor data fusion as in J. Vitola et al. [26,27], to merge the data that come from the different actuation phases in a single data matrix. In this paper, the approach with a single data matrix is also considered, but the case where each actuation phase is used as a classifier is additionally examined in Section 5.5.

5.3. Data Collection

Given a particular scenario, as the three defined in Section 5.2, four matrices $\mathbf{X}[\varphi]$, $\varphi = 1, \dots, 4$, one for each actuation phase, are obtained. Each matrix $\mathbf{X}[\varphi]$, $\varphi = 1, \dots, 4$, is organized as follows:

- $n_1 = n_2 = n_3 = n_4 = 25$ experiments or realizations are performed for each structural state. Consequently, each matrix $\mathbf{X}[\varphi]$, $\varphi = 1, \dots, 4$, consists of 100 rows, that is $n_1 + n_2 + n_3 + n_4 = 25 \cdot 4$. More precisely, the first 25 rows represent the structure with no damage, the next 25 are realization where damage 1 is present in the structure, and so on.
- For each actuation phase φ , $\varphi = 1, \dots, 4$, we measure $N = 3$ PZTs working as sensors during 60000 time instants. Then, these measurements are transformed into the frequency domain. Therefore, the number of columns of matrix $\mathbf{X}[\varphi]$, $\varphi = 1, \dots, 4$, is equal to $N \cdot L = 3 \cdot ((60000/2) + 1) = 90003$.

Therefore, the matrix that collects all the realizations under the four different structural states in the frequency domain is (see Equation (6): here $L = 30001$ and $E = 4$):

$$\mathbf{X}[\varphi] = \left(x[\varphi]_{i,l}^{k,j} \right) \in \mathcal{M}_{100 \times 90003}(\mathbb{R}). \quad (17)$$

The damage detection and classification procedure introduced in Sections 3 and 4 can be applied to each one of the matrices $\mathbf{X}[\varphi]$, $\varphi = 1, \dots, 4$, in Equation (17), thus leading to one classification per actuation phase. But we can also use the horizontal concatenation of the four matrices $\mathbf{X}[\varphi]$, $\varphi = 1, \dots, 4$, to obtain the matrix:

$$\mathbf{X}[1,2,3,4] = \left[\mathbf{X}[1] \mathbf{X}[2] \mathbf{X}[3] \mathbf{X}[4] \right] \in \mathcal{M}_{100 \times (4 \cdot 90003)}(\mathbb{R}) = \mathcal{M}_{100 \times 360012}(\mathbb{R}). \quad (18)$$

If matrix $\mathbf{X}[1,2,3,4]$ in Equation (18) is used for the damage detection and classification procedure introduced in Sections 3 and 4, which this allows analyzation of the information of all the actuation phases at one time, a single classification is obtained that combines these four phases. Finally, we can also use the separate classification obtained for each actuation phase so that each actuation phase casts a vote thus leading to a final decision based on the four actuation phases. These strategies will be explained in detail in Section 5.5.

5.4. κ -Fold Non-Exhaustive Leave- p -Out Cross Validation

The analysis of the proposed approach is done by comparing *test* data, i.e., the new experiments in *unknown* state under the same conditions, with *baseline* data, which is data from the structure under different structural states. To this end, we use the κ -fold non-exhaustive leave- p -out cross validation described in the subsequent paragraphs.

For the sake of clarity, let us write $\mathbf{X}[\Phi]$ to refer to both matrix $\mathbf{X}[\varphi]$ in Equation (17) and matrix $\mathbf{X}[1,2,3,4]$ in Equation (18). Some of the rows in $\mathbf{X}[\Phi]$ will be used as the baseline data to build the model and the clusters, i.e., $v = 5$ rows per structural state, and the rest of the rows are used for the validation. More precisely, we will perform five iterations ($\kappa = 5$) of a non-exhaustive leave- p -out cross validation, where $p = \sum_{i=1}^E (n_i - v) = n_1 + n_2 + n_3 + n_4 - v \cdot E = 80$, to estimate the overall accuracy and avoid overfitting. Let us define, for each structural state $l = 1, \dots, E$, the permutation σ_l :

$$\begin{aligned} \sigma_l : \{1, 2, \dots, n_l\} &\rightarrow \{1, 2, \dots, n_l\} \\ i &\rightarrow \sigma_l(i) \end{aligned}$$

In this particular case, $n_1 = n_2 = n_3 = n_4 = 25$. Therefore, in the first iteration, the baseline data to build the model are the matrix:

$$\mathbf{X} = \mathbf{S}^\top \cdot \mathbf{X}[\Phi], \quad (19)$$

$$\mathbf{S} = [\mathbf{e}_{\sigma_1(1)} \cdots \mathbf{e}_{\sigma_1(5)} | \mathbf{e}_{n_1+\sigma_2(1)} \cdots \mathbf{e}_{n_1+\sigma_2(5)} | \cdots | \mathbf{e}_{n_1+n_2+n_3+\sigma_4(1)} \cdots \mathbf{e}_{n_1+n_2+n_3+\sigma_4(5)}], \quad (20)$$

where $\mathbf{e}_j \in \mathbb{R}^{n_1+n_2+n_3+n_4} = \mathbb{R}^{100}$ is the j -th element of the canonical basis of the real vector space $\mathbb{R}^{n_1+n_2+n_3+n_4} = \mathbb{R}^{100}$, and $\mathbf{S} \in \mathcal{M}_{(n_1+n_2+n_3+n_4) \times (v \cdot E)}(\mathbb{R})$ is the *selector* matrix. Basically, matrix \mathbf{X} in Equation (19) has been built by randomly selecting $v = 5$ rows per structural state. The $\sum_{i=1}^E (n_i - v) = 80$ rows of matrix $\mathbf{X}[\Phi]$ that are not used to build the model are used for the validation.

In the i -th iteration, $i = 1, \dots, \kappa$, the baseline data to build the model are the matrix:

$$\mathbf{X} = \mathbf{S}^\top \cdot \mathbf{X}[\Phi],$$

$$\mathbf{S} = \left[\mathbf{e}_{\sigma_1(5(i-1)+1)} \cdots \mathbf{e}_{\sigma_1(5(i-1)+5)} | \mathbf{e}_{n_1+\sigma_2(5(i-1)+1)} \cdots \mathbf{e}_{n_1+\sigma_2(5(i-1)+5)} | \cdots \right. \\ \left. \cdots | \mathbf{e}_{n_1+n_2+n_3+\sigma_4(5(i-1)+1)} \cdots \mathbf{e}_{n_1+n_2+n_3+\sigma_4(5(i-1)+5)} \right]$$

where, as in Equation (19), $\mathbf{e}_j \in \mathbb{R}^{n_1+n_2+n_3+n_4} = \mathbb{R}^{100}$ is the j -th element of the canonical basis of the real vector space $\mathbb{R}^{n_1+n_2+n_3+n_4} = \mathbb{R}^{100}$ and, as in Equation (20), $\mathbf{S} \in \mathcal{M}_{(n_1+n_2+n_3+n_4) \times (v \cdot E)}(\mathbb{R})$ is the *selector* matrix. Since $\sum_{i=1}^E (n_i - v) = 80$ rows of matrix $\mathbf{X}[\Phi]$ will be used for the validation step and with respect to $\kappa = 5$ iterations, the sum of all the elements in the *confusion matrices* that we will present in Section 6 is equal to $(\sum_{i=1}^E (n_i - v)) \cdot \kappa = 400$.

5.5. Application of the Damage Detection and Classification Procedure

In this section, two strategies are presented to apply the damage detection and classification procedure. These two strategies are:

- (1) the classification is based on a single matrix: $\mathbf{X}[1], \mathbf{X}[2], \mathbf{X}[3], \mathbf{X}[4]$ or $\mathbf{X}[1, 2, 3, 4]$, as defined in Equations (17) and (18), respectively, with κ -fold non-exhaustive leave- p -out cross validation;
- (2) the classification is based on the four matrices $\mathbf{X}[1], \mathbf{X}[2], \mathbf{X}[3]$ and $\mathbf{X}[4]$, defined in Equation (17), with respect to the four actuation phases, with κ -fold non-exhaustive leave- p -out cross validation. Each actuation phase will cast a vote and a final decision is taken.

In the first case, in a succinct way, the following seven steps are performed:

- **Step 1.** The data in matrix \mathbf{X} are scaled using MCGS to define a new matrix $\check{\mathbf{X}}$.
- **Step 2.** PCA is applied to $\check{\mathbf{X}}$ to obtain the PCA model \mathbf{P} .
- **Step 3.** The number $\ell = D \in \mathbb{N}$ of principal components is chosen so that the proportion of variance explained is greater than or equal to 95%. Therefore, the reduced PCA model is \mathbf{P}_ℓ .
- **Step 4.** A realization $\mathbf{z}^\top \in \mathbb{R}^{3 \cdot 30001} = \mathbb{R}^{90003}$ —for $\mathbf{X}[1], \mathbf{X}[2], \mathbf{X}[3]$ and $\mathbf{X}[4]$ — or $\mathbf{z}^\top \in \mathbb{R}^{4 \cdot 90003} = \mathbb{R}^{360012}$ —for $\mathbf{X}[1, 2, 3, 4]$ — of the current structure to diagnose is needed. Then, vector \mathbf{z}^\top is scaled as in Equation (14) to define $\check{\mathbf{z}}^\top$.
- **Step 5.** The data points set \mathcal{X} is defined as:

$$\mathcal{X}' = \{\mathbf{x}^1, \dots, \mathbf{x}^{20}, \mathbf{x}^{21}\} \subset \mathbb{R}^\ell,$$

where

$$\begin{aligned}\mathbf{x}^i &= \mathbf{e}_i^\top \check{\mathbf{X}} \mathbf{P}_\ell, \quad i = 1, \dots, 20, \\ \mathbf{x}^{21} &= \check{\mathbf{z}}^\top \mathbf{P}_\ell.\end{aligned}$$

Subsequently, t -SNE is applied to this ℓ -dimensional data set \mathcal{X}' to find a collection of 2-dimensional map points:

$$\mathcal{Y}' = \{\mathbf{y}^1, \dots, \mathbf{y}^{20}, \mathbf{y}^{21}\} \subset \mathbb{R}^2.$$

- **Step 6.** $E = 4$ clusters are obtained, that are related to the $E = 4$ different structural states. These clusters are formed by the map points:

$$\begin{aligned}\{\mathbf{y}^1, \dots, \mathbf{y}^5\} &\subset \mathcal{Y}, \text{ related to the 1st structural state;} \\ \{\mathbf{y}^6, \dots, \mathbf{y}^{10}\} &\subset \mathcal{Y}, \text{ related to the 2nd structural state;} \\ \{\mathbf{y}^{11}, \dots, \mathbf{y}^{15}\} &\subset \mathcal{Y}, \text{ related to the 3rd structural state;} \\ \{\mathbf{y}^{16}, \dots, \mathbf{y}^{20}\} &\subset \mathcal{Y}, \text{ related to the 4th structural state.}\end{aligned}$$

The centroid \mathcal{Y}_l , $l = 1, \dots, E$, associated with the l -th structural state is computed as in Equation (16).

- **Step 7.** Finally, the current structure to diagnose is associated to the l -th structural state if

$$l = \arg \min_{l=1, \dots, E} \|\mathcal{Y}_l - \mathbf{y}^{21}\|_2.$$

In the second case, we follow **Step 1** to **Step 6** above for the four matrices $\mathbf{X}[\varphi]$, $\varphi = 1, \dots, 4$, related to the four actuation phases. With the information provided by the four actuation phases, several approaches can be considered to finally classify the structure that has to be diagnosed. One of these approaches, majority voting, is widely used in standard fusion schemes [28], as well as weighted majority vote or soft voting. For our case of the small aluminum plate, the majority voting will be used, as well as an approach based on the sum of the inverse distances between the centroids and the map point, which is somehow related to a weighted majority vote. Here are the details of both approaches:

- **Majority voting.** In this case, the strategy of the smallest point-centroid distance is performed four times, one per actuation phase. Therefore, four classifications are obtained for a single structure to diagnose. More precisely, each actuation phase acts as a *classifier*. Figure 6 illustrates this idea with respect to three actuation phases.

The current structure to diagnose, in the φ -th actuation phase, $\varphi = 1, \dots, 4$, is associated to the l_φ -th structural state if

$$l_\varphi = \arg \min_{l=1, \dots, E} \|\mathcal{Y}_l^\varphi - \mathbf{y}_\varphi^{21}\|_2.$$

It is worth remembering that $\mathbf{y}_\varphi^{21} \in \mathbb{R}^2$ is the map point associated to the realization of the current structure to diagnose. The structure is finally classified according to the most repeated classification. That is, the current structure to diagnose is associated to the l -th structural state if

$$l = \text{mode}\{l_1, l_2, l_3, l_4\},$$

in the case of a unimodal set. In the case of a bi-modal set, if the two modal values are l_α and l_β , the current structure to diagnose is associated to the l -th structural state if

$$l = \arg \min_{l \in \{l_\alpha, l_\beta\}} \sum_{\varphi=1}^4 \|\mathcal{Y}_l^\varphi - \mathbf{y}_\varphi^{21}\|_2.$$

Finally, if the set $\{l_1, l_2, l_3, l_4\}$ is a set with no mode, the structure is associated to the l -th structural state if

$$l = \arg \min_{l=1, \dots, E} \sum_{\varphi=1}^4 \|\mathcal{Y}_l^\varphi - \mathbf{y}_\varphi^{21}\|_2.$$

- **Sum of the inverse distances.** In this case, for a given structural state, we sum the inverse of the distances between the centroids \mathcal{Y}_l^φ and the map point \mathbf{y}_φ^{21} , for all the actuation phases $\varphi = 1, \dots, 4$. The assigned structural state is the one that obtains the highest sum. More precisely, the current structure to diagnose is associated to the l -th structural state if

$$l = \arg \max_{l=1, \dots, E} \sum_{\varphi=1}^4 \frac{1}{\|\mathcal{Y}_l^\varphi - \mathbf{y}_\varphi^{21}\|_2}.$$

It is worth remarking that the arguments of the maxima of the sum of the inverse distances is equivalent to the arguments of the minima of the harmonic mean of these distances. More precisely, for a given structural state, the harmonic mean of the distances between the centroids \mathcal{Y}_l^φ and the map point \mathbf{y}_φ^{21} , for all the actuation phases $\varphi = 1, \dots, 4$ is

$$\frac{1}{\frac{1}{4} \sum_{\varphi=1}^4 \frac{1}{\|\mathcal{Y}_l^\varphi - \mathbf{y}_\varphi^{21}\|_2}}.$$

Therefore,

$$l = \arg \max_{l=1, \dots, E} \sum_{\varphi=1}^4 \frac{1}{\|\mathcal{Y}_l^\varphi - \mathbf{y}_\varphi^{21}\|_2} = \arg \min_{l=1, \dots, E} \frac{1}{\frac{1}{4} \sum_{\varphi=1}^4 \frac{1}{\|\mathcal{Y}_l^\varphi - \mathbf{y}_\varphi^{21}\|_2}}.$$

S. Mehta et al. [29] also uses the harmonic distance to define a pattern classification technique similar to k -nearest neighbors classifier.

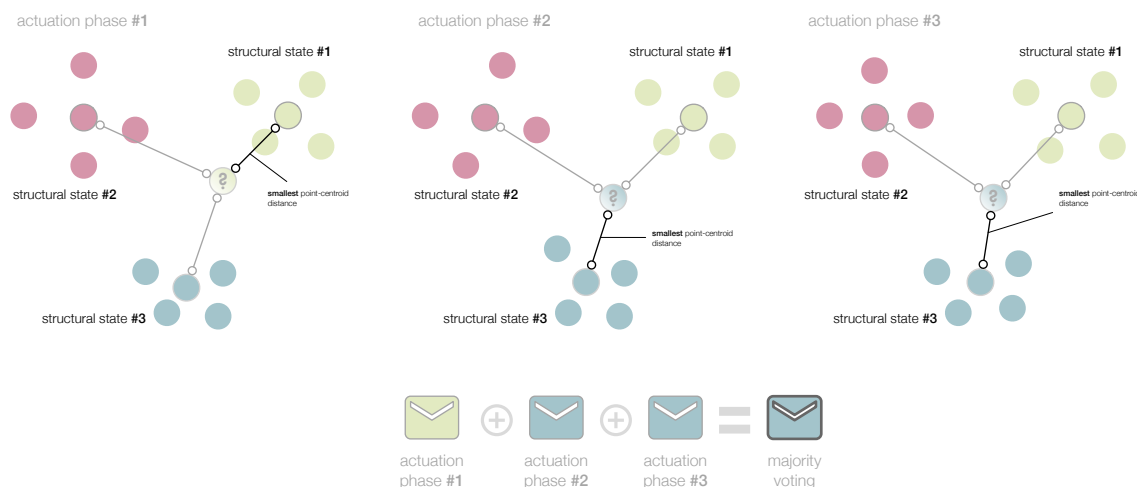


Figure 6. In the majority voting, the strategy of the smallest point-centroid distance is performed per actuation phase. The current structure to diagnose is associated to the most voted structural state.

6. Results

In this section, the results of the application of the damage detection and classification procedure, introduced in Sections 3 and 4 and detailed in Sections 5.3–5.5, to the aluminum plate are presented in terms of the confusion matrices and with respect to the scenarios defined in Section 5.2. The results for each scenario are presented in a different section. More precisely, in Section 6.1 the results with respect to **Scenario 1** are presented. Equivalently, Sections 6.2 and 6.3 present the results with respect to **Scenario 2** and **Scenario 3**, respectively. In the three scenarios, four different structural states have been considered:

- the first structural state corresponds to the healthy state of the structure, that is, the square aluminum plate with no damage, noted as D_0 ;
- the second, third, and fourth structural states correspond to the plate with an added mass at the positions indicated in Figures 3 and 5, noted as D_1 , D_2 , and D_3 , respectively.

To validate the damage detection and classification detailed in Sections 5.3–5.5, we will perform five iterations ($\kappa = 5$) of a non-exhaustive leave- p -out cross validation, where $p = 80$, as described in Section 5.4. At each iteration, a total of 80 realizations have been considered, according to the following distribution: 20 realization per structural state (D_0 , D_1 , D_2 , and D_3). Since 80 realizations have been used for the validation step and with respect to $\kappa = 5$ iterations, the sum of all the elements in the confusion matrices that we will present in Sections 6.1–6.3 is equal to $5 \cdot 80 = 400$.

Again, for the three scenarios, seven different confusion matrices are presented:

- **Actuation phase 1.** The damage detection and classification procedure is applied to a single matrix, $X[1]$, as in Equation (17), using the smallest point-centroid distance.
- **Actuation phase 2.** The damage detection and classification procedure is applied to a single matrix, $X[2]$, as in Equation (17), using the smallest point-centroid distance.
- **Actuation phase 3.** The damage detection and classification procedure is applied to a single matrix, $X[3]$, as in Equation (17), using the smallest point-centroid distance.
- **Actuation phase 4.** The damage detection and classification procedure is applied to a single matrix, $X[4]$, as in Equation (17), using the smallest point-centroid distance.

- **Actuation phases 1–4.** The damage detection and classification procedure is applied to a single matrix, i.e., the horizontal concatenation of the four matrices $X[\varphi]$, $\varphi = 1, \dots, 4$, $\mathbf{X}[1, 2, 3, 4]$, as in Equation (18), using the smallest point-centroid distance.
- **Majority voting.** The damage detection and classification procedure is applied to the four $X[\varphi]$, $\varphi = 1, \dots, 4$. Each actuation phase casts a vote, and a final decision is taken based on majority voting (Section 5.5).
- **Sum of the inverse distances.** The damage detection and classification procedure is applied to the four $X[\varphi]$, $\varphi = 1, \dots, 4$. Each actuation phase casts a vote and a final decision is taken based on the maximum sum of the inverse distances (Section 5.5).

Finally, in some cases, and with the purpose of comparing the performance of the current damage detection and classification approach, confusion matrices in the frequency and time domains have been included.

As an illustrating example, we have included in Figure 7 the clusters formed by the different structural states described in this section and in the case of Scenario 3. In this figure, the diamond represents the structure to diagnose. It can be clearly observed how the diamond is close to the cluster related to damage 3.

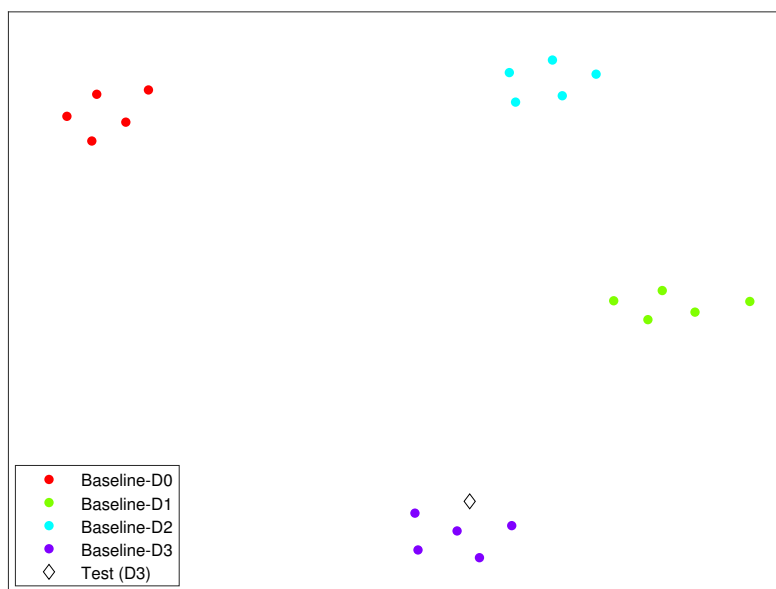


Figure 7. Clusters formed by the different structural states described in Section 6, for Scenario 3. The diamond represents the structure to diagnose.

6.1. Scenario 1

In this section, the results with respect to **Scenario 1** are presented. It is worth noting that, in this scenario, a short wire has been used, and the measured signals are filtered with a SG algorithm. The seven confusion matrices can be found in Tables 1 and 2. When the decision is based on a single actuation phase (Table 1), the overall accuracy is quite good. More precisely, with 397 in the actuation phase 1, 399 in the actuation phase 2, 395 in the actuation phase 3, and 397 in the actuation phase 4, realizations have been correctly classified out of 400 cases, which represents an overall accuracy of 99.25%, 99.75%, 98.75%, and 99.25%, respectively. When the four actuation phases are used at the same time (actuation phases 1–4, Equation (18), majority voting, and sum of the inverse distances), an overall accuracy of 99–100% is achieved, as it can be observed from Table 2.

Table 1. Confusion matrix of the application of the *t*-SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 1**, in the **frequency domain**. Rows represent true values, while columns represent predicted values.

	Actuation Phase 1				Actuation Phase 2				Actuation Phase 3				Actuation Phase 4			
	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3
D0	97	0	2	1	100	0	0	0	100	0	0	0	100	0	0	0
D1	0	100	0	0	0	100	0	0	0	100	0	0	0	100	0	0
D2	0	0	100	0	0	0	100	0	3	0	96	1	0	0	98	2
D3	0	0	0	100	0	0	1	99	1	0	0	99	0	0	1	99

D0 (healthy state of the structure); D1, D2, and D3 (added masses at the positions indicated in Figures 3 and 5).

Table 2. Confusion matrix of the application of the *t*-SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 1**, when the four actuation phases are used at the same time, in the **frequency domain**. Rows represent true values, while columns represent predicted values.

	Phases 1–4				Majority Voting				Inverse Distances			
	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3
D0	100	0	0	0	100	0	0	0	100	0	0	0
D1	0	99	1	0	0	100	0	0	0	100	0	0
D2	1	0	99	0	0	0	100	0	0	0	100	0
D3	0	0	0	100	0	0	0	100	0	0	0	100

D0 (healthy state of the structure); D1, D2, and D3 (added masses at the positions indicated in Figures 3 and 5).

6.2. Scenario 2

In this section, the results with respect to **Scenario 2** are presented. In this case, a short wire has been used but the measured signals are not filtered. The seven confusion matrices can be found in Tables 3 and 4. When the decision is based on a single actuation phase (Table 3), the overall accuracy is very remarkable. More precisely, with respect to actuation phase 1, 2, and 3, 400 realizations have been correctly classified out of 400 cases, which represents an overall accuracy of 100%. With respect to actuation phase 4, 399 realizations have been correctly classified out of 400 cases, that is to say, an overall accuracy of 99.75%. When the four actuation phases are used at the same time (actuation phases 1–4, Equation (18), majority voting, and sum of the inverse distances), an overall accuracy of 100% is achieved, as it can be observed from Table 4.

Table 3. Confusion matrix of the application of the *t*-SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 2**, in the **frequency domain**. Rows represent true values, while columns represent predicted values.

	Actuation Phase 1				Actuation Phase 2				Actuation Phase 3				Actuation Phase 4			
	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3
D0	100	0	0	0	100	0	0	0	100	0	0	0	100	0	0	0
D1	0	100	0	0	0	100	0	0	0	100	0	0	0	100	0	0
D2	0	0	100	0	0	0	100	0	0	0	100	0	0	0	100	0
D3	0	0	0	100	0	0	0	100	0	0	0	100	0	0	1	99

D0 (healthy state of the structure); D1, D2, and D3 (added masses at the positions indicated in Figures 3 and 5).

Table 4. Confusion matrix of the application of the *t*-SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 2**, when the four actuation phases are used at the same time, in the **frequency domain**. Rows represent true values, while columns represent predicted values.

	Phases 1–4				Majority Voting				Inverse Distances			
	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3
D0	100	0	0	0	100	0	0	0	100	0	0	0
D1	0	100	0	0	0	100	0	0	0	100	0	0
D2	0	0	100	0	0	0	100	0	0	0	100	0
D3	0	0	0	100	0	0	0	100	0	0	0	100

D0 (healthy state of the structure); D1, D2, and D3 (added masses at the positions indicated in Figures 3 and 5).

6.3. Scenario 3

The results with respect to **Scenario 3** are finally presented in this section. In the two previous scenarios, a short wire was used. However, in this case, the signals are acquired using a 2.5 m long wire. Tables 5 and 6 include the seven confusion matrices. When the decision is based on a single actuation phase (Table 5), the overall accuracy is quite good, too. More precisely, with 384 in the actuation phase 1, 395 in the actuation phase 2, 395 in the actuation phase 3, and 398 in the actuation phase 4, realizations have been correctly classified, which represents an overall accuracy of 96%, 98.75%, 98.75%, and 99.5%, respectively. When the four actuation phases are used at the same time (actuation phases 1–4, Equation (18), majority voting, and sum of the inverse distances), an overall accuracy of 99.5–100% is achieved, as it can be observed from Table 6.

Table 5. Confusion matrix of the application of the *t*-SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 3**, in the **frequency domain**. Rows represent true values, while columns represent predicted values.

	Actuation Phase 1				Actuation Phase 2				Actuation Phase 3				Actuation Phase 4			
	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3
D0	98	2	0	0	99	0	1	0	100	0	0	0	100	0	0	0
D1	6	90	2	2	0	99	1	0	0	95	5	0	1	99	0	0
D2	1	1	97	1	0	1	97	2	0	0	100	0	0	0	100	0
D3	0	1	0	99	0	0	0	100	0	0	0	100	0	0	1	99

D0 (healthy state of the structure); D1, D2, and D3 (added masses at the positions indicated in Figures 3 and 5).

Table 6. Confusion matrix of the application of the *t*-SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 3**, when the four actuation phases are used at the same time, in the **frequency domain**. Rows represent true values, while columns represent predicted values.

	Phases 1–4				Majority Voting				Inverse Distances			
	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3
D0	100	0	0	0	100	0	0	0	100	0	0	0
D1	0	99	0	1	0	100	0	0	0	100	0	0
D2	0	1	99	0	0	0	100	0	0	0	100	0
D3	0	0	0	100	0	0	0	100	0	0	0	100

D0 (healthy state of the structure); D1, D2, and D3 (added masses at the positions indicated in Figures 3 and 5).

The potential of the approaches where the four actuation phases are used can be observed in this last scenario, see Table 6:

- When the four actuation phases are merged in a single matrix as in Equation (18), 398 realizations have been correctly classified out of 400 cases, which represents an overall accuracy of 99.5%.
- When each actuation phase casts a vote and a final decision is taken based on majority voting, the overall accuracy is increased to 100%.
- Finally, when each actuation phase casts a vote and a final decision is taken based on the maximum sum of the inverse distances, the overall accuracy is increased to 100%, too.

In addition, in this scenario, the results in the frequency domain are compared to those in the time domain [11]. In the time domain, when the decision is based on a single actuation phase (Table 7), with 244 in the actuation phase 1, 398 in the actuation phase 2, 280 in the actuation phase 3, and 277 in the actuation phase 4, realizations have been correctly classified out of 400 cases. This represents an overall accuracy of 61%, 99.5%, 70%, and 69.25%, respectively. Clearly, the strategy in the frequency domain, i.e., the overall accuracy fluctuates between 96% and 99.5%, outperforms the approach in the time domain. In addition, the false positive rate (FPR), i.e., the number of false positives with respect to the total number of negatives, and the false negative rate (FNR), i.e., the number of false negatives with respect to the total number of positives, are clearly unsatisfactory in the time domain. However, FPR and FNR are significantly reduced to values close to 0% in the frequency domain. It is worth noting that in the computation of the FNR, the three different types of damage $D1$, $D2$, and $D3$ are considered as a single category, just the opposite of the healthy state of the structure. In the time domain, when the four actuation phases are used at the same time (Table 8), the overall accuracy is of 88.5% in actuation phases 1–4, of 91.25% in majority voting, and of 97% in sum of the inverse distances, whereas the overall accuracy is increased to 99.5%, 100%, and 100%, respectively, in the frequency domain. At the same time, FPR and FNR are reduced to 0% in the frequency domain, while they are slightly increased in the time domain. All this is an indication of the better quality of the clusters created in the frequency domain than the ones created in the time domain. Table 9 summarizes the values for the overall accuracy, the FPR, and the FNR in this scenario and in the time and frequency domains.

Table 7. Confusion matrix of the application of the t -SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in textbfScenario 3, in the **time domain**. Rows represent true values, while columns represent predicted values.

	Actuation Phase 1				Actuation Phase 2				Actuation Phase 3				Actuation Phase 4			
	$D0$	$D1$	$D2$	$D3$	$D0$	$D1$	$D2$	$D3$	$D0$	$D1$	$D2$	$D3$	$D0$	$D1$	$D2$	$D3$
$D0$	50	6	19	25	100	0	0	0	93	1	3	3	53	21	1	25
$D1$	19	66	11	4	0	100	0	0	16	50	23	11	17	61	2	20
$D2$	14	3	73	10	0	1	98	1	5	19	70	6	5	3	76	16
$D3$	15	9	21	55	0	0	0	100	0	23	10	67	10	3	0	87

$D0$ (healthy state of the structure); $D1, D2$, and $D3$ (added masses at the positions indicated in Figures 3 and 5).

Table 8. Confusion matrix of the application of the *t*-SNE based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 3**, when the four actuation phases are used at the same time, in the **time domain**. Rows represent true values, while columns represent predicted values.

	Phases 1–4				Majority voting				Inverse distances			
	D0	D1	D2	D3	D0	D1	D2	D3	D0	D1	D2	D3
D0	86	1	7	6	98	0	0	2	99	0	1	0
D1	8	88	4	0	12	85	2	1	1	99	0	0
D2	1	8	89	2	6	2	90	2	1	1	95	3
D3	3	4	2	91	0	3	5	92	1	2	2	95

D0 (healthy state of the structure); D1, D2, and D3 (added masses at the positions indicated in Figures 3 and 5).

Table 9. Overall accuracy, false positive rate (FPR) and false negative rate (FNR) of the application of the *t*-SNE-based damage detection and classification procedure presented in Sections 3 and 4 to the case of the aluminum plate in **Scenario 3**, when the four actuation phases are used separately and at the same time, in both domains, time and frequency.

	Accuracy		FPR		FNR	
	Time	Frequency	Time	Frequency	Time	Frequency
Actuation phase 1	61.0%	96.0%	50.0%	2.0%	16.0%	2.3%
Actuation phase 2	99.5%	98.8%	0.0%	1.0%	0.0%	0.0%
Actuation phase 3	70.0%	98.8%	7.0%	0.0%	7.0%	0.0%
Actuation phase 4	69.3%	99.5%	47.0%	0.0%	10.7%	0.3%
Phases 1-4	88.5%	99.5%	14.0%	0.0%	4.0%	0.0%
Majority voting	91.3%	100.0%	2.0%	0.0%	6.0%	0.0%
Inverse distances	97.0%	100.0%	1.0%	0.0%	1.0%	0.0%

6.4. General Comments

The results presented in Sections 6.1–6.3 reveal that it is better to make a decision considering all of the actuation phases (assembling these phases or using them to cast a vote) rather than working with the phases separately. On the other hand, the results also show that both strategies majority voting and sum of the inverse distances slightly outperform the horizontal concatenation of the four actuation phases in the frequency domain. However, in the time domain (**Scenario 3**, Table 8), the results reveal (i) the strong performance of the sum of the inverse distances strategy, which clearly classifies the practical totality of the kinds of damage that we have considered, compared to majority voting or the horizontal concatenation of the four actuation phases; and (ii) that the majority voting outperforms the horizontal concatenation of the four actuation phases, but it cannot completely classify damage D1.

It is worth noting that, in general, the healthy state of the structure is confused with the structure with damage in just a few cases. Similarly, the structure with damage is identified as a structure with no damage in a very limited number of realizations.

In general, the performance of the proposed methodology is very satisfactory when the signals are acquired using a short wire, with or without adding white Gaussian noise. In these two cases, using PCA as a pre-processing step, the noise is canceled. The third scenario presents the worst case because it used a long cable (2.5 m) from the digitizers to the sensors. In this scenario, the signals were badly digitized due to the impedance of the cable, the low voltage of the stimulus, and other experimental features. Therefore, it can be observed that the use of a long cable from the digitizer to the sensors affects in the detection and classification method. However, combining the four actuation phases with

- (i) the sum of the inverse distances strategy, in the time domain; or
- (ii) the majority voting strategy or the sum of the inverse distances strategy, in the frequency domain,

very accurate results can be obtained.

It should be noted that, in general, it is better to work in the frequency domain than in the time domain because the obtained results are significantly improved, as it can be clearly observed in the third scenario.

7. Conclusions

In this work, a SHM strategy for detection and classification of structural changes based on a two-step data integration (type *E* unfolding and MCGS), data transformation using PCA, and a two-step data reduction combining PCA and *t*-SNE has been proposed. The proposed approach is evaluated using experimental data. In general, the results obtained show that the performance of the proposed methodology is very satisfactory, given its high classification accuracy; and its behavior is very good and similar in all the data sets.

In the case study, very accurate results are obtained with or without adding white Gaussian noise, since PCA cancels the noise. However, the use of a long wire (2.5 m) from the digitizers to the sensors negatively affects the detection and classification method. But combining the four actuation phases with the sum of the inverse distances strategy, in the time domain, and with the majority voting strategy or the sum of the inverse distances strategy, in the frequency domain, accurate results can be obtained. Results also show that the quality of the two-dimensional clusters created with *t*-SNE in the frequency domain is better than the quality of the two-dimensional clusters created with *t*-SNE in the time domain, thus leading to a better classification. Therefore, the strategy in the frequency domain significantly outperforms the approach in the time domain.

Some aspects to highlight in the proposed methodology are: the *t*-SNE technique has been extended and adapted to the field of SHM, in the detection and classification of structural changes; the method classifies the current state of the structure by means of a data-driven analysis, that is, using collected data from the structure under different structural states and without the use of complex mathematical models; it is better to make a decision considering all of the actuation phases (assembling these phases or using them to cast a vote) rather than working with the phases separately; both strategies, majority voting and sum of the inverse distances, slightly outperform the horizontal concatenation of the four actuation phases in the frequency domain; in the time domain, sum of the inverse distances strategy outperforms majority voting, and this last strategy outperforms the horizontal concatenation of the four actuation phases; it is better to work in the frequency domain than in the time domain because better results are obtained; and finally, in general, the healthy state of the structure is confused with the structure with damage in just a few cases, and similarly, the structure with damage is identified as a structure with no damage in a very limited number of realizations. With respect to the possible fields of application, similar aluminum plates have been used to represent parts of a plane (wings or fuselage). We think that we can also apply this approach for the damage and fault detection of wind turbines. In general, there is no a prescribed field of application: if a sensor network can be installed in a structure, and several actuation phases can be considered, the proposed approach can be implemented a priori.

As a future work, we plan to develop further the proposed method for different EOC to determine its effectiveness, as well as to handle imbalanced data. In addition, we aim to investigate the parametric version of *t*-SNE.

Author Contributions: D.A. and F.P. developed the idea and designed the exploration framework; D.A. developed the algorithms; F.P. supervised the results of the validation; D.A. and F.P. drafted the manuscript.

Funding: This work has been partially funded by the Spanish Agencia Estatal de Investigación (AEI) - Ministerio de Economía, Industria y Competitividad (MINECO), and the Fondo Europeo de Desarrollo Regional (FEDER) through

the research project DPI2017-82930-C2-1-R; and by the Generalitat de Catalunya through the research project 2017 SGR 388.

Acknowledgments: We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. We thank the Universitat Politècnica de Catalunya (UPC) for pre-doctoral fellowship (to David Agis). We also thank Jaime Vitola and Diego A. Tibaduiza for useful discussions with respect to the experimental set-up.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

\mathcal{X}	High-dimensional data set.
\mathcal{X}'	High-dimensional data set including the data point to diagnose.
\mathcal{Y}	Low-dimensional map points.
\mathcal{Y}'	Low-dimensional map points including the map point to diagnose.
\mathcal{Y}_l	Centroid associated with the l -th structural state.
\mathcal{P}	Similarity matrix for the high-dimensional data points.
\mathcal{Q}	Similarity matrix for the low-dimensional map points.
\mathcal{C}	KL divergence.
η	Learning rate.
$\alpha(t)$	Momentum term at iteration t .
p	Perplexity.
\mathbf{X}	Matrix that collects all the realizations under different structural states.
$\tilde{\mathbf{X}}$	Scaled original—in the frequency domain— data matrix.
N	Number of sensors.
L	Number of components in each signal.
E	Number of different structural states.
$\mathbf{C}_{\tilde{\mathbf{X}}}$	Covariance matrix of $\tilde{\mathbf{X}}$.
λ_k	Eigenvalues.
ρ_k	Eigenvectors.
Λ	Eigenvalues in a diagonal matrix.
\mathbf{P}	PCA model (full case).
\mathbf{P}_ℓ	PCA model (reduced).
\mathbf{T}	Projection of the scaled data set $\tilde{\mathbf{X}}$ into the subspace spanned by the PCA model.
\mathbf{e}_i	i -th element of the canonical basis.
\mathbf{z}	Current original—in the frequency domain— data vector to diagnose.
$\tilde{\mathbf{z}}$	Current scaled original—in the frequency domain— data vector to diagnose.
φ	φ -th actuation phase.
\mathbf{S}	Selector matrix

Abbreviations

CM	Condition monitoring
EOC	Environmental and operational conditions
FFT	Fast Fourier transform
FNR	False negative rate
FPR	False positive rate
KL	Kullback-Leibler
MCGS	Mean-centered group scaling
PCA	Principal component analysis

PZT	Piezoelectric transducer
SG	Savitzky–Golay
SHM	Structural health monitoring
SNE	Stochastic neighbor embedding
<i>t</i> -SNE	<i>t</i> -distributed stochastic neighbor embedding

References

1. Rolfes, R.; Zerbst, S.; Haake, G.; Reetz, J.; Lynch, J.P. Integral SHM-system for offshore wind turbines using smart wireless sensors. In Proceedings of the 6th International Workshop on Structural Health Monitoring, Stanford, CA, USA, 11–13 September 2007; pp. 11–13.
2. Ciang, C.C.; Lee, J.R.; Bang, H.J. Structural health monitoring for a wind turbine system: A review of damage detection methods. *Meas. Sci. Technol.* **2008**, *19*, 122001.
3. Loh, C.H.; Loh, K.J.; Yang, Y.S.; Hsiung, W.Y.; Huang, Y.T. Vibration-based system identification of wind turbine system. *Struct. Control Health Monit.* **2017**, *24*, e1876.
4. Raju, K.S.; Pratap, Y.; Sahni, Y.; Babu, M.N. Implementation of a WSN system towards SHM of civil building structures. In Proceedings of the IEEE 9th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, India, 9–10 January 2015; pp. 1–7.
5. Blanco, H.; Boffill, Y.; Lombillo, I.; Villegas, L. An integrated structural health monitoring system for determining local/global responses of historic masonry buildings. *Struct. Control Health Monit.* **2018**, *25*, e2196.
6. Nisha, M. Structural health monitoring of aircraft wing using wireless network. *Int. J. Technol. Explor. Learn. Wwv. Ijtel. Org Struct* **2014**, *3*, 341–343.
7. Ochôa, P.; Groves, R.M.; Benedictus, R. Systematic multiparameter design methodology for an ultrasonic health monitoring system for full-scale composite aircraft primary structures. *Struct. Control Health Monit.* **2019**, *26*, e2340.
8. Ward, M.O.; Grinstein, G.; Keim, D. *Interactive Data Visualization: Foundations, Techniques, and Applications*; AK Peters/CRC Press: Boca Raton, FL, USA, 2015.
9. Tenenbaum, J.B.; De Silva, V.; Langford, J.C. A global geometric framework for nonlinear dimensionality reduction. *Science* **2000**, *290*, 2319–2323.
10. van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
11. Agis, D.; Pozo, F. Detection and classification of structural changes using t-distributed stochastic neighbor embedding. In Proceedings of the International Conference on Structural Engineering Dynamics, Viana do Castelo, Portugal, 24–26 June 2019; pp. 1–10.
12. Tsogka, C.; Daskalakis, E.; Comanducci, G.; Ubertini, F. The stretching method for vibration-based structural health monitoring of civil structures. *Comput. Aided Civ. Infrastruct. Eng.* **2017**, *32*, 288–303.
13. Xu, F.; Song, X.; Tsui, K.L.; Yang, F.; Huang, Z. Bearing performance degradation assessment based on ensemble empirical mode decomposition and affinity propagation clustering. *IEEE Access* **2019**, *7*, 54623–54637.
14. Cheng, Y.; Zhu, H.; Hu, K.; Wu, J.; Shao, X.; Wang, Y. Multisensory data-driven health degradation monitoring of machining tools by generalized multiclass support vector machine. *IEEE Access* **2019**, *7*, 47102–47113.
15. Westerhuis, J.A.; Kourti, T.; MacGregor, J.F. Comparing alternative approaches for multivariate statistical analysis of batch process data. *J. Chemom. A J. Chemom. Soc.* **1999**, *13*, 397–413.
16. Mujica, L.; Rodellar, J.; Fernandez, A.; Güemes, A. Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. *Struct. Health Monit.* **2011**, *10*, 539–553.
17. Vidal, Y.; Pozo, F.; Tutivén, C. Wind turbine multi-fault detection and classification based on SCADA data. *Energies* **2018**, *11*, 3018.
18. Sikdar, S.; Kundu, A.; Jurek, M.; Ostachowicz, W. Nondestructive analysis of debonds in a composite structure under variable temperature conditions. *Sensors* **2019**, *19*, 3454.
19. Yan, J.; Jin, H.; Sun, H.; Qing, X. Active monitoring of fatigue crack in the weld zone of bogie frames using ultrasonic guided waves. *Sensors* **2019**, *19*, 3372.

20. Cho, H.; Hasanian, M.; Shan, S.; Lissenden, C.J. Nonlinear guided wave technique for localized damage detection in plates with surface-bonded sensors to receive Lamb waves generated by shear-horizontal wave mixing. *NDT E Int.* **2019**, *102*, 35–46.
21. Jiménez, A.A.; Muñoz, C.Q.G.; Márquez, F.P.G. Dirt and mud detection and diagnosis on a wind turbine blade employing guided waves and supervised learning classifiers. *Reliab. Eng. Syst. Saf.* **2019**, *184*, 2–12.
22. Hinton, G.; Roweis, S.T. Stochastic neighbor embedding. *Adv. Neural Inf. Process. Syst.* **2003**, *15*, 857–864.
23. Min, R. *A Non-Linear Dimensionality Reduction Method for Improving Nearest Neighbour Classification*; University of Toronto: Toronto, ON, Canada, 2005.
24. Pozo, F.; Vidal, Y.; Salgado, Ó. Wind turbine condition monitoring strategy through multiway PCA and multivariate inference. *Energies* **2018**, *11*, 749.
25. Orfanidis, S.J. *Introduction to Signal Processing*; Prentice-Hall, Inc.: Upper Saddle River, NJ, USA, 1995.
26. Vitola, J.; Pozo, F.; Tibaduiza, D.A.; Anaya, M. A sensor data fusion system based on k -nearest neighbor pattern classification for structural health monitoring applications. *Sensors* **2017**, *17*, 417.
27. Vitola, J.; Pozo, F.; Tibaduiza, D.A.; Anaya, M. Distributed piezoelectric sensor system for damage identification in structures subjected to temperature changes. *Sensors* **2017**, *17*, 1252.
28. Tardy, B.; Inglada, J.; Michel, J. Assessment of optimal transport for operational land-cover mapping using high-resolution satellite images time series without reference data of the mapping period. *Remote Sens.* **2019**, *11*, 1047.
29. Mehta, S.; Shen, X.; Gou, J.; Niu, D. A new nearest centroid neighbor classifier based on k local means using harmonic mean distance. *Information* **2018**, *9*, 234.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).