Training artificial neural networks using substructuring techniques: Application to joint identification

Jure Korbar[a], Domen Ocepek[a], Gregor Čepon[a,*], Miha Boltežar[a]

[a] *University of Ljubljana, Faculty of Mechanical Engineering, Aškerčeva 6, 1000 Ljubljana, Slovenia*

Cite as:
Jure Korbar, Domen Ocepek, Gregor Čepon, Miha Boltežar,
Training artificial neural networks using substructuring techniques: Application to joint identification,
Mechanical Systems and Signal Processing, 2023, https://doi.org/10.1016/j.ymssp.2023.110426

Abstract

The dynamic properties of assembled structures are governed by the substructure dynamics as well as the dynamics of the joints that are part of the assembly. It can be challenging to describe the physical interactions within the joints analytically, as slight modifications, such as static preload, temperature, etc. can lead to significant changes in the assembly's dynamic properties. Therefore, characterizing the dynamic properties of joints typically involves experimental testing and subsequent model updating. In this paper, a machine-learning-based approach to joint identification is proposed that utilizes a physics-based computational model of the joint. The idea is to combine the computational model of the joint with dynamic substructuring techniques to train the machine-learning model. The flexibility of dynamic substructuring permits the enforcement of compatibility and equilibrium conditions between the component models from the experimental and numerical domains, facilitating the development of machine-learning models that can predict the dynamic properties of joints. The proposed approach provides an accurate data-driven method for joint identification in real structures, while reducing the number of measurements needed for the identification. The approach permits the identification of a full 12-DoF joint, enabling the coupling of 3D dynamic models of substructures. Compared to the standard decoupling approach, no spurious peaks are present in the reconstructed assembly response. The proposed approach is validated numerically and experimentally by reconstructing the assembly response and comparing the results with known assembly dynamics.

Keywords: joint identification, dynamic substructuring, frequency-based substructuring, artificial neural networks, physics-based computational model

1. Introduction

The dynamics of typical elements within assembled structures, such as beams, bars, plates, etc., can be accurately described using updated finite-element (FE) models. On the other hand, accurately modeling an assembly's dynamics is more difficult due to the inherent dynamics of the joints connecting the substructures. However, by properly identifying the dynamic properties of the joints, the assembly's dynamics can be accurately predicted. Due to the complex interactions at the interface of the connected substructures, the joints are difficult to model using the FEM formulation. Therefore, the dynamic properties of joints are typically obtained with experimental testing.

Over the years, various approaches to joint identification have been proposed. Tsai and Chou [1] proposed an experimental approach to identifying the properties of bolted joints based on the measured frequency

*Corresponding author
Email address: gregor.cepon@fs.uni-lj.si (Gregor Čepon)

response functions. This approach was validated on a simple assembly of two beams and a bolted joint, resulting in a two-DoF joint model. The authors showed how neglecting the joint dynamics leads to an erroneous prediction of the assembly dynamics. Using the proposed approach, a good prediction of the assembly dynamics can be achieved, provided that the mass of the joint can be neglected. Ren and Beards [2] presented an FRF-based joint identification method in which the joint can be represented in terms of stiffness, mass, and damping matrices. A weighted least squares approach was presented to identify the joint properties, with the emphasis on obtaining a set of equations that is the least sensitive to measurement error. However, care must be taken to choose a suitable weighting approach to ensure that the joint-identification equations are not underdetermined or that the noise is not amplified. Another method for joint identification was presented by Čelič and Boltežar in [3], extending Ren and Beards' least squares solution. The authors emphasized the importance of the rotational DoFs in the identification procedure and demonstrated that neglecting the RDoFs leads to an erroneous prediction of the assembly's response. Due to the difficulty in measuring the RDoFs, the RDoF FRFs were estimated using accurately calibrated FE models. The authors published a second paper in which they discussed the influence of coordinate reduction on the identification of the dynamic properties of joints [4]. Two approaches to joint identification were presented: the direct and the average solutions. The direct solution is frequency dependent, whereas the average solution finds the best joint parameters in a least-squares sense for the frequency range of interest. It was shown that all the joint DoFs must be considered to obtain a valid average solution. However, the latter is significantly more robust in the presence of noise.

Wang et al. [5] presented four equations for joint identification. By proposing a method for estimating the unmeasured FRFs, both the measured and unmeasured FRFs were used for the identification. Mehrpouya et al. used a known set of equations to identify the dynamics of multiple joints [6]. By neglecting the rotational and in-plane joint parameters, the number of measurements required to identify the joint parameters was reduced. A methodology for the identification of the joint dynamics in 3D structures using a similar inverse receptance coupling method was presented in [7], where the joint's inertial properties were also considered. A two-node twelve-DoF joint was proposed, with three translational and three rotational DoFs per node, allowing the assembly of 3D structures.

In [8] equations for identifying the dynamic stiffness of a joint were deduced, resulting in the same set of equations as in the work of Ren and Beards. The dynamics of the joint were identified using a least squares approach to find the equivalent mass, stiffness, and damping matrices. In all of the above research, the dynamics at the interface were estimated by using accurately calibrated numerical models or by using measurements near the interface, which introduces errors to the identified joint dynamics, which might not be negligible, especially when there is a need for a highly accurate dynamic model of the joint. In [9] the dynamic response of the coupling elements was obtained using an in situ decoupling method without measurements of the individual substructures. In [10] the dynamic properties of rubber isolators were characterized using two approaches: a frequency-based substructuring (FBS) approach and the in-situ decoupling approach mentioned above and the results were compared with measurements of the isolator stiffness using a servo hydraulic testing machine. Since direct measurements of the rubber-isolator dynamics are typically not feasible, two fixtures, which were assumed to behave rigidly in the frequency range of interest, were attached to either side of the isolator and subsequently decoupled using FBS. The virtual point transformation (VPT) was applied to reconstruct the rotational DoFs from conventional translational measurements, yielding a full twelve-DoF dynamic model of the joint. The in-situ method is attractive since no measurements of the separate substructure dynamics are needed; however, the joint mass is neglected in the process, which is not always feasible. To account for the joint mass, correction terms were considered; however, the connecting substructures must be assumed to behave rigidly, and therefore an in-situ approach can only account for the joint mass if the connecting substructures are rigid, which is typically not the case in practice. On the other hand, the decoupling approach does not assume a negligible mass; however, a large number of measurements has to be performed and spurious peaks can occur in the vicinity of the natural frequencies of the rigidly-coupled substructures. Joint dynamics were also obtained using system equivalent model mixing (SEMM) and an updating scheme without measuring the dynamics at the interface [11]. Although this method can identify joints without interface measurements, the identified joint is prone to errors near the natural frequencies of a rigidly coupled assembly.

Recently, a characterization of rolling tires was presented by Minervini et al. [12]. The tire dynamics were characterized using four distinct FBS decoupling approaches by altering the compatibility and equilibrium conditions. Since the physical interactions inside the joints are typically complex, they can represent one of the main sources of nonlinearity in assemblies. For example, Jamia et al. [13] used a detailed, three-dimensional model of a bolted flange joint to study the frictional behavior at the contact interface due to microslip phenomena. An equivalent model of the joint consisting of Jenkins elements was proposed, and the parameters of the equivalent model were identified for the detailed model.

Purely physics-based modeling requires a detailed understanding of all mechanisms, which often proves infeasible or ineffective in practice. In contrast, machine learning (ML) offers the possibility to solve complex engineering and scientific problems with a data-driven approach [14]. The ML domain provides a wide range of algorithms with the common goal of automating the learning from acquired data [15]. In the subdomain of ML called supervised learning, an ML model is trained to predict target values (outputs) based on the acquired data (inputs) [16]. The advantage of such an approach is that it does not require an understanding of the physical relationships between the inputs and outputs. Typical ML applications include face, speech, and handwriting recognition [17], spam detection [18], autonomous vehicle control [19], object detection [20], etc. ML algorithms are also applied in structural dynamics and vibro-acoustics, e.g., in multi-bolt looseness detection [21], active noise cancellation [22], structural damage detection [23], prognostics and health monitoring [24], and brake-noise detection, characterization and prediction [25].

In the field of structural dynamics, the use of ML is challenging due to its data-driven nature, typically requiring a large training data set to compensate for the lack of a physical background. It is costly and time-consuming to acquire the measurements for all the system configurations. Therefore, the amount of available training data is often insufficient to obtain reliable models. Considering the limitations of the physics-based and ML modeling approaches, the idea of their integration has gained momentum in recent years. The objective of such an integration can be, among others, to employ physics-based computational models (e.g., digital twins [26]) to generate samples for training ML models [27]. Since it is particularly difficult to obtain large amounts of training data in structural dynamics, using a physics-based computational model to generate the training data provides a viable approach to implementing ML algorithms. Ritto et al. [28] proposed a digital-twin-based approach to structural damage identification using a physics-based computational model to generate the dynamic response data needed to train various ML algorithms. The latter were subsequently used to identify the severity and location of the damage in the observed structure.

The challenge when using ML for joint identification is to obtain a sufficiently large training data set of assembly responses. Obtaining assembly admittances experimentally is impractical, since obtaining just a single experimental assembly admittance is time consuming. Even if the experimental approach was practical, the joint would have to be replaced after each test, making the acquisition of an experimental training data set infeasible. The problem of obtaining training data can be alleviated by numerically generating a set of joint admittances using a computational model. Since the objective is to identify the dynamic properties of joints based on the assembly responses, the generated joint admittances must be coupled with substructures to form an assembly response. Therefore, the use of dynamic substructuring is essential as it makes it possible to evaluate the dynamic behavior of large and complex systems by considering the individual substructures separately through compatibility and equilibrium constraints [29]. Dynamic substructuring makes is possible to combine the experimental and numerical component models, providing a hybrid modeling approach that can efficiently generate assembly responses for a given set of joint response models.

In general, the available joint-identification approaches can obtain satisfactory results, although they typically assume a negligible joint mass, perform measurements near the interface due to the inaccessibility of the interface and can suffer from spurious peaks, which decreases the accuracy of the identified joint dynamics. Furthermore, they typically require a large set of measurements. This paper presents a conceptually new framework for identifying a joint's dynamic properties using a physics-based computational model to facilitate the training of an artificial neural network (ANN). Since dynamic substructuring can combine component models from the experimental and numerical domains, it proves to be well suited to obtain large sets of coupled assembly responses from the known substructure admittances and the generated joint admittances. The objective is to combine elements of dynamic substructuring with the data-driven

aspect of machine learning by exploiting their complementary advantages, i.e., eliminating the need for a physical description of the interactions at the interface and efficiently generating a large set of assembly admittances. This concept is similar to the approach presented in [28], as the procedure therein relies on the use of a physics-based computational model to generate training data, although in this case for the purpose of damage detection. However, the approach described therein was not experimentally validated on a real structure. In our approach the use of an ANN facilitates the prediction of the dynamic properties of joints based on the FRFs of assemblies. The data-recognition capability of ANNs is utilized to identify the parameters of the joints connecting substructures based on the FRFs of a real assembled structure. Furthermore, the proposed approach can identify the joint parameters without the need to perform measurements at the interface between the substructures in the assembled state. The interface dynamics of the individual substructures in the decoupled state are obtained by transforming the measurements near the interface using the VPT, thereby eliminating the need for a numerical model of the individual substructures, as is the case for some of the earlier research and the SEMM-based approaches. The contribution of the proposed approach is threefold: first, the approach demonstrates how significantly fewer measurements are needed to identify the joint dynamics. Second, the approach can be used in-situ, i.e. the joint can be identified from the dynamics of an existing assembly. The final contribution is successfully training an ANN to identify the parameters of the joint based on the FRF data, which is the easily measurable property of real structures. The proposed approach is benchmarked against an established joint identification approach based on the FBS decoupling process. By accurately identifying the joint dynamics, the accuracy of the predicted assembly dynamics can be greatly improved, which is important for characterizing the full-vehicle NVH profile [12], minimizing NVH-related problems in compressors [30] or decreasing the vibrations of the camera on unmanned aerial vehicles by actively controlling the gimbal system [31].

The paper is organized as follows. The next section summarizes the theory of the Lagrange-multiplier frequency-based substructuring (LM-FBS) method and the VPT. Then, the joint's computational model is presented along with the framework for generating training samples and a neural network scheme. Section 4 demonstrates the proposed approach in a numerical case study. Section 5 provides an application of the presented method on an experimental case, and finally, the conclusions are drawn in Section 6.

2. Frequency-based substructuring

In order to couple substructures based on their frequency-response models, the LM-FBS method [29] is adopted here.

2.1. LM-FBS method

The LM-FBS method makes it possible to determine the assembled system admittance $\mathbf{Y}^{\mathrm{AB}}$, in which the FRFs of the individual subsystems are considered. For the sake of brevity, the derivation presented here is limited to the coupling of two substructures. However, an arbitrary number of substructures can be used. The equation of motion in the frequency domain for the uncoupled substructures, shown in Fig. 1a, is:

$$\mathbf{u}(\omega) = \mathbf{Y}^{\mathrm{A|B}}(\omega)\big(\mathbf{f}(\omega) + \mathbf{g}(\omega)\big)\,. \tag{1}$$

(a) (b)

Figure 1: Schematic overview of the substructuring problem: (a) Uncoupled substructures A and B, (b) Coupled assembly AB.

The vector $\mathbf{u}(\omega)$ represents the responses to the external force vector $\mathbf{f}(\omega)$ acting on the assembly, and $\mathbf{g}(\omega)$ is the vector of interface forces holding the substructures together. To account for all the subsystems,

their individual admittance matrices must be stacked into a block diagonal matrix $\mathbf{Y}^{\mathrm{A|B}1}$:

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1^{\mathrm{A}} \\ \mathbf{u}_2^{\mathrm{A}} \\ \mathbf{u}_2^{\mathrm{B}} \\ \mathbf{u}_3^{\mathrm{B}} \end{bmatrix}, \quad \mathbf{Y}^{\mathrm{A|B}} = \begin{bmatrix} \mathbf{Y}_{11}^{\mathrm{A}} & \mathbf{Y}_{12}^{\mathrm{A}} & \mathbf{0} & \mathbf{0} \\ \mathbf{Y}_{21}^{\mathrm{A}} & \mathbf{Y}_{22}^{\mathrm{A}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_{22}^{\mathrm{B}} & \mathbf{Y}_{23}^{\mathrm{B}} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_{32}^{\mathrm{B}} & \mathbf{Y}_{33}^{\mathrm{B}} \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} \mathbf{f}_1^{\mathrm{A}} \\ \mathbf{f}_2^{\mathrm{A}} \\ \mathbf{f}_2^{\mathrm{B}} \\ \mathbf{f}_3^{\mathrm{B}} \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} \mathbf{0} \\ \mathbf{g}_2^{\mathrm{A}} \\ \mathbf{g}_2^{\mathrm{B}} \\ \mathbf{0} \end{bmatrix}. \tag{2}$$

The relations between the substructures are expressed in terms of the compatibility and equilibrium conditions. The inclusion of both is achieved using the Boolean matrix $\mathbf{B}$. The compatibility conditions are expressed with Eq. (3), which ensures the same displacement of the substructures at the interface in the coupled state (Fig. 1b). The equilibrium conditions (Eq. (4)) are introduced by replacing the interface forces with a set of unknown Lagrange multipliers $\boldsymbol{\lambda}$:

$$\mathbf{B}\,\mathbf{u} = \mathbf{0}\,, \tag{3}$$

$$\mathbf{g} = -\mathbf{B}^{\mathrm{T}}\boldsymbol{\lambda}\,. \tag{4}$$

Solving the set of Eqs. (2) – (4) by eliminating the Lagrange multipliers yields the response of the coupled structure:

$$\mathbf{u} = \mathbf{Y}^{\mathrm{AB}}\,\mathbf{f} = \left[\mathbf{Y}^{\mathrm{A|B}} - \mathbf{Y}^{\mathrm{A|B}}\,\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\,\mathbf{Y}^{\mathrm{A|B}}\,\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\,\mathbf{Y}^{\mathrm{A|B}}\right]\mathbf{f}\,. \tag{5}$$

The dynamic properties of the assembled system are governed by the admittance matrix $\mathbf{Y}^{\mathrm{AB}}$. The LM-FBS method requires the full-DoF response models of the individual substructures with DoFs at the interface collocated for all substructures. However, the DoFs measured on both sides of the interface do not usually coincide in the case of experimental testing of substructures with complex geometries. Even if collocation could be achieved, the lack of RDoFs, which are not measurable with conventional translational transducers, would pose a problem. Both problems mentioned above can be resolved using the VPT.

2.2. Virtual point transformation

A virtual point is chosen near the physical interface of the substructures where the admittance matrix is obtained using a geometric transformation. Several responses and excitations are measured near this point ($\mathbf{Y}_{\mathrm{uf}} \in \mathbb{C}^{n_{\mathrm{u}} \times n_{\mathrm{f}}}$) and then projected onto the interface deformation modes (IDMs). If only the rigid body IDMs are included in the transformation (rigid interface behavior), we have $m = 6$ DoFs for each virtual point. If the interface exhibits a more complex dynamic behavior, flexible IDMs can also be added [32]. The transformation is achieved using the following equation:

$$\mathbf{Y}_{\mathrm{qm}} = \mathbf{T}_{\mathrm{u}}\mathbf{Y}_{\mathrm{uf}}\mathbf{T}_{\mathrm{f}}^{\mathrm{T}}\,, \tag{6}$$

where $\mathbf{T}_{\mathrm{u}}$ is the displacement transformation matrix, and $\mathbf{T}_{\mathrm{f}}$ is the force transformation matrix. $\mathbf{Y}_{\mathrm{qm}} \in \mathbb{C}^{m \times m}$ is the VP admittance matrix with perfectly collocated force/moment and translation/rotation DoFs.

Assuming the interface exhibits only rigid behavior, the displacement vector $\mathbf{q}$ consists of three translations $\mathbf{q}_{\mathrm{t}} = [q_X,\, q_Y,\, q_Z]^{\mathrm{T}}$ and three rotations $\mathbf{q}_\theta = [q_{\theta_X},\, q_{\theta_Y},\, q_{\theta_Z}]^{\mathrm{T}}$. The following relation can be written between the m VP responses $\mathbf{q}$ and the n_{u} sensor displacements $\mathbf{u}$:

$$\mathbf{u} = \mathbf{R}_{\mathrm{u}}\mathbf{q} + \boldsymbol{\mu}_{\mathrm{u}}\,. \tag{7}$$

The columns of $\mathbf{R}_{\mathrm{u}} \in \mathbb{R}^{n_{\mathrm{u}} \times m}$ consist of IDMs constructed from the relative sensor locations and orientations with respect to the VP (Fig. 2). It is a common practice to overdetermine Eq. (7) so that $n_{\mathrm{u}} > m$. For more information about the construction of $\mathbf{R}_{\mathrm{u}}$, the reader is referred to [33]. Vector $\boldsymbol{\mu}_{\mathrm{u}}$ captures any residual motion not included in the subspace of the IDMs. If the rigid assumption of the interface is valid in the considered frequency range, the residual motion will most likely be negligible. By considering only

[1] An explicit dependence on frequency is omitted to improve the readability of the notation, as will be the case for the remainder of the paper.

rigid IDMs, flexible motion is filtered out of $\mathbf{q}$ and the interface problem is weakened due to the reduction. Solving Eq. (7) for $\mathbf{q}$ in a least-mean-square sense yields the displacements of the VP:[2]

$$\mathbf{q} = \left(\mathbf{R}_\mathrm{u}^\mathrm{T}\mathbf{R}_\mathrm{u}\right)^{-1}\mathbf{R}_\mathrm{u}^\mathrm{T}\mathbf{u} = \mathbf{T}_\mathrm{u}\mathbf{u} \quad \Rightarrow \quad \mathbf{T}_\mathrm{u} = \left(\mathbf{R}_\mathrm{u}^\mathrm{T}\mathbf{R}_\mathrm{u}\right)^{-1}\mathbf{R}_\mathrm{u}^\mathrm{T}\,. \tag{8}$$

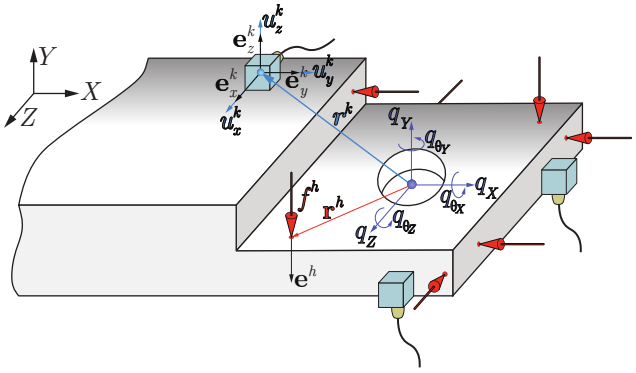

Figure 2: Interface connection example using a virtual point[3].

Similarly, the forces and moments at the virtual point $\mathbf{m}$ are obtained for a measured vector of forces $\mathbf{f}$. Again, assuming rigid IDMs, the force vector $\mathbf{m}$ consists of three forces and three moments ($\mathbf{m} = [m_X, m_Y, m_Z, m_{\theta_X}, m_{\theta_Y}, m_{\theta_Z}]^\mathrm{T}$). The contributions of all the input forces can be combined and expressed as follows:

$$\mathbf{m} = \mathbf{R}_\mathrm{f}^\mathrm{T}\mathbf{f}, \tag{9}$$

where $\mathbf{R}_\mathrm{f}^\mathrm{T} \in \mathbb{R}^{m \times n_\mathrm{f}}$ is the matrix containing the positions and orientations for all the impact locations with respect to the VP. A more detailed description of $\mathbf{R}_\mathrm{f}$ can be found in [33]. By solving Eq. (9) for $\mathbf{f}$ we obtain the following expression:

$$\mathbf{f} = \mathbf{R}_\mathrm{f}\left(\mathbf{R}_\mathrm{f}^\mathrm{T}\mathbf{R}_\mathrm{f}\right)^{-1}\mathbf{m} = \mathbf{T}_\mathrm{f}^\mathrm{T}\mathbf{m} \quad \Rightarrow \quad \mathbf{T}_\mathrm{f}^\mathrm{T} = \mathbf{R}_\mathrm{f}\left(\mathbf{R}_\mathrm{f}^\mathrm{T}\mathbf{R}_\mathrm{f}\right)^{-1}. \tag{10}$$

To alleviate the inconvenience of constructing the IDM matrices, the application of the VPT can be facilitated by using the open-source Python package pyFBS [35].

3. Joint's computational model and neural networks

In this paper an ANN is used to identify the joint parameters based on the response model of an assembly. The proposed identification approach relies on the assembly of an analytical joint model and two known substructures. The substructuring approach can be used to generate a large training data set of assembly responses, facilitating the training of an ANN. The objective is to identify the joint parameters based on the FRFs remote from the assembly interface. The advantage of such an approach is twofold: fewer measurements are required since no measurements need to be performed at the assembly interface (as opposed to identifying the joint parameters with the conventional FBS approach), and multiple joints can be characterized using the same trained ANN as long as the joint is the only varying component in the assembly.

[2]To obtain more flexibility in the transformation, a frequency-dependent symmetric weighting matrix $\mathbf{W}$ can be considered in the derivation of $\mathbf{q}$, see [34].

[3]The position vector from the VP to the center of the triaxial accelerometer is denoted by $\mathbf{r}^k$. The unit vector for each sensor axis is $\mathbf{e}_i^k$ and the response in each axis is denoted by u_i^k ($i \in (x, y, z)$). The position vector from VP to the force impact is $\mathbf{r}^h$, the impact direction is $\mathbf{e}^h$ and the impact magnitude is f^h.

3.1. Analytical joint model

To generate the training set of joint response models, the analytical model of the joint must first be established. A two-node joint model with six DoFs per node is adopted, resulting in a twelve-DoF joint model. Each node contains three translational and three rotational DoFs. The nodes interact with each another via linear springs. The joint is schematically depicted in Fig. 3 as a superposition of objects with mass (or mass moment of inertia) connected by springs.

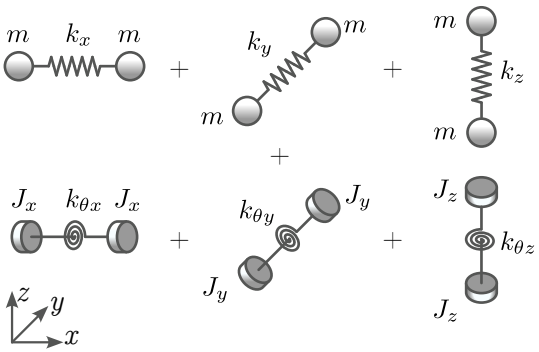

Figure 3: Schematic of the DoF-to-DoF joint model.

The joint mass and stiffness matrices are defined by the submatrices $\mathbf{K}_\mathrm{s}^\mathrm{J}$ and $\mathbf{M}_\mathrm{s}^\mathrm{J}$ of each node:

$$\mathbf{K}_\mathrm{s}^\mathrm{J} = \begin{bmatrix} k_x & 0 & 0 & 0 & 0 & 0 \\ 0 & k_y & 0 & 0 & 0 & 0 \\ 0 & 0 & k_z & 0 & 0 & 0 \\ 0 & 0 & 0 & k_{\theta x} & 0 & 0 \\ 0 & 0 & 0 & 0 & k_{\theta y} & 0 \\ 0 & 0 & 0 & 0 & 0 & k_{\theta z} \end{bmatrix}, \quad \mathbf{M}_\mathrm{s}^\mathrm{J} = \begin{bmatrix} m & 0 & 0 & 0 & 0 & 0 \\ 0 & m & 0 & 0 & 0 & 0 \\ 0 & 0 & m & 0 & 0 & 0 \\ 0 & 0 & 0 & J_x & 0 & 0 \\ 0 & 0 & 0 & 0 & J_y & 0 \\ 0 & 0 & 0 & 0 & 0 & J_z \end{bmatrix}. \tag{11}$$

The connection of nodes leads to the following joint mass and stiffness matrices:

$$\mathbf{K}^\mathrm{J} = \begin{bmatrix} \mathbf{K}_\mathrm{s}^\mathrm{J} & -\mathbf{K}_\mathrm{s}^\mathrm{J} \\ -\mathbf{K}_\mathrm{s}^\mathrm{J} & \mathbf{K}_\mathrm{s}^\mathrm{J} \end{bmatrix}, \quad \mathbf{M}^\mathrm{J} = \begin{bmatrix} \mathbf{M}_\mathrm{s}^\mathrm{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_\mathrm{s}^\mathrm{J} \end{bmatrix}. \tag{12}$$

The joint admittance matrix $\mathbf{Y}^\mathrm{J}$ can be obtained by inverting the joint dynamic stiffness matrix $\mathbf{Z}^\mathrm{J}$ [36]:

$$\mathbf{Y}^\mathrm{J} = \left(\mathbf{Z}^\mathrm{J}\right)^{-1}, \quad \text{where} \quad \mathbf{Z}^\mathrm{J} \triangleq -\omega^2 \mathbf{M}^\mathrm{J} + \mathbf{K}^\mathrm{J}. \tag{13}$$

The dynamic response of an undamped joint can be expressed using Eq. (13). Alternatively, the joint admittance matrix can be obtained by solving the eigenvalue problem $-\omega^2\mathbf{M}^\mathrm{J}+\mathbf{K}^\mathrm{J} = \mathbf{0}$, yielding n eigenfrequencies ω_r and the corresponding eigenvectors $\boldsymbol{\varphi}_r$. A damping ratio ζ_r can be defined for each non-rigid body mode $\boldsymbol{\varphi}_r$ and the mode-superposition method can be used to synthesize the FRFs of the joint admittance matrix [36]:

$$\mathbf{Y}^\mathrm{J} \approx -\frac{1}{\omega^2} \sum_{r=1}^{n_\mathrm{RB}} \boldsymbol{\phi}_r \boldsymbol{\phi}_r^\mathrm{T} + \sum_{r=n_\mathrm{RB}+1}^{n} \frac{\boldsymbol{\phi}_r \boldsymbol{\phi}_r^\mathrm{T}}{-\omega^2 + 2\mathrm{j}\omega\zeta_r\omega_r + \omega_r^2}. \tag{14}$$

The modes $\boldsymbol{\phi}_r$ used in Eq. (14) are mass normalized. The number of rigid-body modes, corresponding to the eigenfrequencies $\omega_r = 0$, is denoted by n_RB, while the total number of eigenfrequencies and mode shapes of the joint is denoted by n. The joint admittance matrix $\mathbf{Y}^\mathrm{J}$ can be obtained using Eq. (13) for an undamped joint or using Eq. (14) for a damped joint.

3.2. Generating training data using FBS

The objective of this paper is to identify the dynamic properties of a joint based on the frequency response of an assembly. If the proposed analytical joint model can describe the true joint's dynamics sufficiently well, the joint can be identified by predicting the parameters of the joint model. The prediction is facilitated by using an ANN, which can be viewed as a function for predicting joint parameters based on the FRF matrix of a given assembly. The ANN can correctly predict the joint parameters only if the ANN parameters are set properly, which is achieved during the training process. To effectively train the ANN, a significant amount of data is required. Each training sample contains a set of randomly generated joint parameters and the FRF matrix of the corresponding assembly.

Each set of joint parameters should be randomly generated according to a chosen distribution. For example, if all the values of a joint parameter within a certain range should have equal probability, a uniform distribution should be chosen. If the value of a parameter x_k spans multiple orders of magnitude, the log-uniform distribution can be considered, resulting in the order of magnitude (i.e. $\log(x_k)$) being uniformly distributed.

After the joint parameter samples are generated, the coupled dynamic response for each joint sample must be calculated. First, one of the Eqs. (13) or (14) should be used to obtain the FRF matrix $\mathbf{Y}^{\mathrm{J}^i}$ of the joint training sample, where J^i denotes the ith joint sample for training. Once the joint FRFs are generated, LM-FBS coupling (Eq. (2)) is used to obtain the FRF matrix $\mathbf{Y}^{\mathrm{AJB}^i}$ of the assembly for each generated joint. The details of the coupling procedure are presented in Section 4.

3.3. Artificial neural networks

Once the assembled FRFs are calculated, the configuration of the ANN must be selected. In this work a deep feedforward ANN was chosen. The building blocks of ANNs are called neurons. In the case of fully connected ANNs, multiple neurons are stacked in layers, with each neuron connected to all the neurons in the adjacent layers. Each neuron can be viewed as a set of mathematical operations, i.e. a weighted sum and an activation function. Let us denote the output of the jth neuron in layer $(l-1)$ as ${}^{l-1}a_j$. The weighted sum ${}^{l}z_i$ computed by the ith neuron in the lth layer is expressed as:

$$ {}^{l}z_i = \sum_j {}^{l-1}w_{ij}\,{}^{l-1}a_j + {}^{l}b_i\,, \quad l \geq 1\,, \tag{15} $$

with weights and biases denoted by ${}^{l-1}w_{ij}$ and ${}^{l}b_i$, respectively. Eq. (15) can also be written in matrix form:

$$ {}^{l}\mathbf{z} = {}^{l-1}\mathbf{W}\,{}^{l-1}\mathbf{a} + {}^{l}\mathbf{b}\,. \tag{16} $$

To calculate the output of the ith neuron, an activation function[4] ${}^{l}f$ is used:

$$ {}^{l}a_i = {}^{l}f({}^{l}z_i)\,. \tag{17} $$

The Eqs. (15) and (17) are used to calculate the output of each layer until the final layer is reached, where the outputs should approach the values of the joint parameters. A visual representation of a simple three-layer ANN can be found in Fig. 4.

The layer $l = 0$ represents the input of the ANN, where the input values are denoted by $\mathbf{x}$. These are used as the input for the second layer, i.e., $l = 1$. Thus, in Eq. (16), when $l = 1$, the term ${}^{l-1}\mathbf{a}$ is replaced by $\mathbf{x}$. The weights are typically set randomly during the ANN initialization, with the biases set to zero. For the ANN to predict the values of the joint parameters, the ANN must be trained by optimizing the weights and biases. First, a loss function is defined that evaluates the error between the predicted and target values. Typically, a quadratic loss function is implemented to evaluate the difference between the

[4] A typical activation function is the sigmoid function $f(z) = 1/(1+\mathrm{e}^{-z})$.

[5] For the sake of simplicity, only the output values ${}^{l}\mathbf{a}$ are displayed to represent the neurons, with a detailed illustration of the mathematical operations for the first neuron of layer $l = 2$.

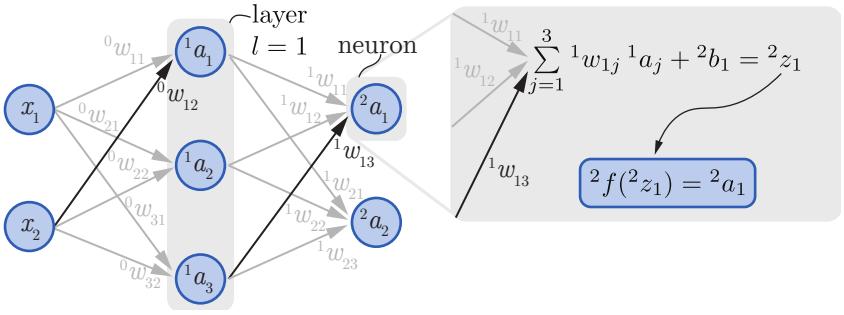

Figure 4: A simple artificial neural network with three layers[5].

predicted and target (i.e., generated) values of the joint parameters. A gradient-descent[6] algorithm can be used to iteratively minimize the loss function, resulting in optimization of the ANN parameters (i.e., weights and biases). Since the loss function depends on the values of the ANN parameters, the gradient of the loss function with respect to the ANN parameters can be calculated. In each iteration, the gradient (which points in the direction of the steepest ascent of the loss function) is calculated. The gradient is scaled and subtracted from the current ANN parameters, updating the values of the parameters. This process is repeated until the minimum of the loss function is found.

3.4. Training data preparation

FRFs are evaluated at discrete frequency points, resulting in large arrays of complex values (e.g., in the range from 1 to 2000 Hz at a resolution of 1 Hz, a single FRF is represented by a set of 2000 complex values). Therefore, using FRFs as the training data could lead to an unnecessarily large input layer, which is unfavorable for three main reasons. First, the number of ANN parameters increases with the size of the input layer, extending the time required to train an ANN. Second, increasing the number of ANN parameters means that additional training samples should be obtained to prevent overfitting. Third, the assembly's FRFs might be less sensitive to the change in joint parameters at these frequencies, suggesting that the values of the FRFs at certain frequencies are not as relevant to the training process. In addition, some FRFs might not be sensitive to changes in the joint parameters overall, e.g., due to low transmissibility. Therefore, it does not seem reasonable to use the raw FRF data of the coupled AJB structure as the input data for the ANN. To overcome the problem of the input data size, a dimensionality-reduction technique must be used. Principal component analysis (PCA), which has been applied to FRF data in the past [37, 38], is widely used for this purpose. Due to the low transmissibility, some FRFs are ignored, as the joint's dynamic properties have a negligible effect on these FRFs. In the following, the process of converting the assembly admittance matrices $\mathbf{Y}^{\mathrm{AJB}^i}$ to the training data matrix $\mathbf{A}_p$ is presented.

First, the FRFs are selected from the assembly admittance matrix $\mathbf{Y}^{\mathrm{AJB}^i}$ of each generated joint sample. From the selected FRFs, a vector of FRF magnitudes $\mathbf{h}^i$ is constructed for each joint[7]. Each vector $\mathbf{h}^i$ contains $n = n_{\mathrm{f}} \cdot n_{\mathrm{FRFs}}$ components, where n_{f} and n_{FRFs} represent the number of frequency points and the number of selected FRFs, respectively. An initial data matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$, where m is equal to the number of generated joint samples, is constructed by stacking the vectors $\mathbf{h}^i$ in the rows of $\mathbf{H}$. A typical element of the original data matrix is denoted by h_{ij}, where the subscripts i and j denote the row and column of $\mathbf{H}$, respectively. Dimensionality reduction (i.e., reducing the number of columns of $\mathbf{H}$) can be achieved using PCA. Prior to the PCA the components should be standardized to obtain scale-independent principal components [39]. First, the mean response $\bar{H}_j$ and the standard deviation S_j of the jth component (i.e.,

[6]To be precise, stochastic gradient descent is used in conjunction with the backpropagation algorithm to calculate the gradients. For more information, the interested reader is referred to [16, 17].

[7]The selection of FRFs used in the ongoing joint identification is presented in Section 4.

the column of $\mathbf{H}$) must be evaluated:

$$\bar{H}_j = \frac{1}{m}\sum_{i=1}^{m} h_{ij}\,, \tag{18}$$

$$S_j = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(h_{ij} - \bar{H}_j\right)^2}\,. \tag{19}$$

The standardized counterpart $\tilde{h}_{ij}$ of a typical element is expressed as:

$$\tilde{h}_{ij} = \frac{h_{ij} - \bar{H}_j}{S_j\sqrt{m}}\,. \tag{20}$$

Replacing all the typical elements h_{ij} with the standardized elements $\tilde{h}_{ij}$ yields the response-variation matrix $\tilde{\mathbf{H}}$. The eigenvectors of the correlation matrix $\mathbf{C} = \tilde{\mathbf{H}}^{\mathrm{T}}\tilde{\mathbf{H}}$ are known as the principal axes of $\tilde{\mathbf{H}}$. The principal axes represent the orthogonal directions in an n-dimensional space that maximize the variance of the data projected onto each principal axis. Projecting the data onto the principal axes yields the principal components (PCs). The data dimensionality can be significantly reduced by considering only the first $p \ll n$ PCs with the highest variance. The principal axes can be determined by the eigendecomposition of the correlation matrix:

$$\mathbf{C} = \mathbf{V}\mathbf{L}\mathbf{V}^{\mathrm{T}}\,, \tag{21}$$

where the columns of $\mathbf{V}$ are the eigenvectors of the correlation matrix (i.e., the principal axes of $\tilde{\mathbf{H}}$), where $\mathbf{L}$ is a diagonal matrix of the eigenvalues of $\mathbf{C}$ that can be used as a measure of the significance of the principal axes. The principal axes can also be determined using the singular value decomposition (SVD) of the response variation matrix $\tilde{\mathbf{H}}$:

$$\tilde{\mathbf{H}} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}\,, \tag{22}$$

where $\mathbf{U}$ and $\mathbf{V}$ are, respectively, the $m \times m$ and $n \times n$ orthonormal matrices, and $\mathbf{S}$ is an $m \times n$ diagonal matrix [40]. The columns of $\mathbf{U}$ and $\mathbf{V}$ contain the left-singular and right-singular vectors of $\tilde{\mathbf{H}}$, which form two sets of orthonormal bases. The columns of $\mathbf{V}$ are again the principal axes. The diagonal elements of $\mathbf{S}$ are the singular values of $\tilde{\mathbf{H}}$, which are identical to the square roots of the eigenvalues of the correlation matrix $\mathbf{C}$ and are typically arranged in descending order. Analogous to the eigenvalues of $\mathbf{C}$, the singular values of $\tilde{\mathbf{H}}$ can also be used as a measure of the importance of the principal axes, i.e., the amount of variation in the direction of each principal axis. The PCs are obtained by multiplying $\tilde{\mathbf{H}}$ by the right-singular vectors, i.e., by projecting the data matrix onto the principal axes. Since $\mathbf{V}$ is orthonormal, the PCs, denoted by $\mathbf{A}$, can be expressed as:

$$\mathbf{A} = \tilde{\mathbf{H}}\mathbf{V} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{U}\mathbf{S}\,, \tag{23}$$

where the product $\mathbf{V}^{\mathrm{T}}\mathbf{V}$ is the identity matrix, since $\mathbf{V}$ is orthonormal. The data dimensionality can be significantly reduced by considering only $p \ll n$ PCs associated with the p highest singular values. This can be done by manually selecting the first p columns of $\mathbf{A}$ or by projecting $\tilde{\mathbf{H}}$ onto the first p principal axes $\mathbf{V}_p$, where $\mathbf{V}_p$ is obtained by retaining only the first p columns of $\mathbf{V}$ associated with the highest singular values. The reduced data matrix, denoted by $\mathbf{A}_p$, is calculated as follows:

$$\mathbf{A}_p = \tilde{\mathbf{H}}\mathbf{V}_p = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}\mathbf{V}_p\,. \tag{24}$$

The matrix $\mathbf{A}_p \in \mathbb{R}^{m\times p}$ contains m samples and p principal components used as the input to the ANN. Multiplying the reduced PCs $\mathbf{A}_p$ by $\mathbf{V}_p^{\mathrm{T}}$ yields a reconstruction $\tilde{\mathbf{H}}_p$ of the response-variation matrix $\tilde{\mathbf{H}}$ from the first p PCs. By rearranging Eq. (20), a reconstruction of the original data matrix $\mathbf{H}$ can be obtained. This can be used to compare the original and reconstructed data [37, 38].

3.5. Joint identification framework

This section briefly outlines the framework for the identification of the joint's dynamic properties. First, the ANN must be trained. An analytical joint model is defined, and m sets of joint parameters are randomly generated. The joint admittance matrices are computed from the generated sets of joint parameters, and LM-FBS coupling is used to obtain the assembly admittance matrix for each set of generated joint parameters. The coupling is followed by data preparation, where the assembly FRF magnitudes are selected, and dimensionality reduction by PCA, yielding the reduced data matrix $\mathbf{A}_p$. The latter is used to train the ANN. The aforementioned steps are depicted in Fig. 5.

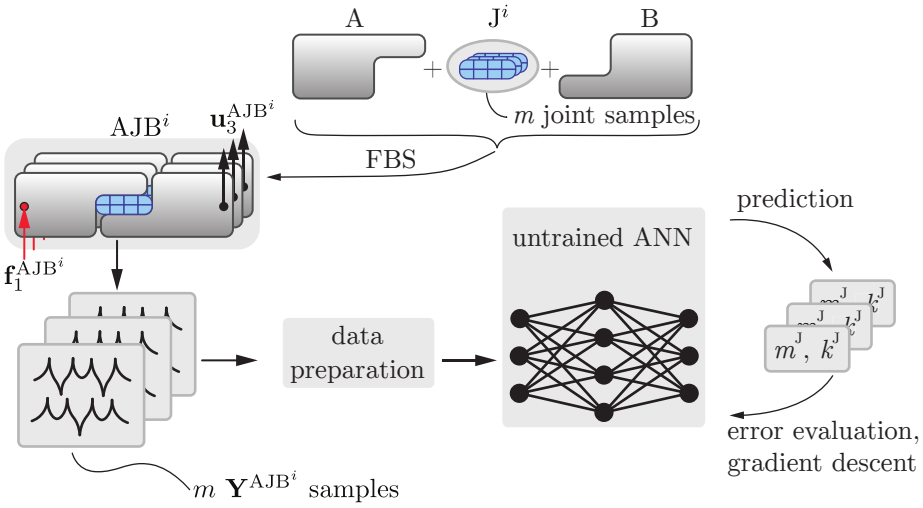

Figure 5: Schematic of ANN training process.

Both the numerical and experimental ANNs have a single hidden layer with 200 neurons. The number of neurons in each layer is also presented in Table 1.

Table 1: Number of neurons in each layer of the neural networks used in the numerical and experimental studies.

	Input layer	Hidden layer	Output layer
Numerical	200	200	7
Experimental	200	200	9

Once the training is complete, the joint parameters can be identified using the numerically or experimentally determined FRFs of the assembly. The identified joint parameters are used to calculate the admittance matrix using Eq. (13) or (14). The identified joint admittance can be coupled with A and B to obtain the reconstructed admittance of the assembly, which can then be compared with the original assembly. The comparison facilitates the validation of the proposed joint model, as well as the ANN's ability to correctly identify the joint parameters. The joint identification and ANN validation are depicted in Fig. 6.

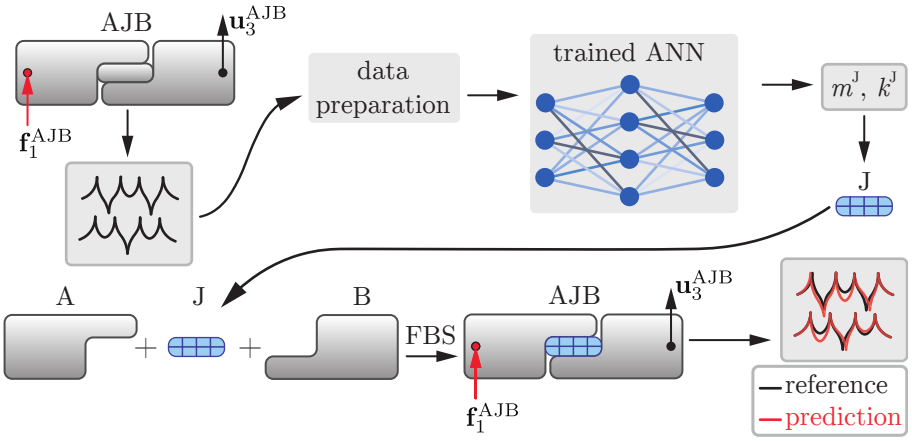

Figure 6: Validation of the ANN and joint identification.

4. Numerical study

4.1. Numerical study setup

A numerical study was conducted to test the proposed joint identification approach. An undamped joint model was chosen for the numerical study to minimize the number of parameters for the identification. It is assumed that the joint's mass moments of inertia have a negligible effect on the coupled FRFs. Therefore, the parameters J_x, J_y, and J_z were considered as small non-zero values to ensure the invertibility of the dynamic stiffness matrix $\mathbf{Z}^{\mathrm{J}}$. In the numerical study, a total of seven joint parameters were identified: the mass of the joint and six stiffness coefficients. The samples of the coupled dynamic responses were obtained by randomly generating the joint parameters and inverting the corresponding dynamic stiffness matrices, which were subsequently coupled with A and B, resulting in a set of coupled dynamic responses $\mathbf{Y}^{\mathrm{AJB}^i}$. The CAD models of the substructures A, J, B, and the assembled AJB are shown in Fig. 7.

In both the numerical and experimental studies, a total of 12 channels were used to measure the dynamic response of A, and nine channels were used to measure the dynamic response of B. The first three channels on A measured the response away from the interface ($\mathbf{u}_1^{\mathrm{A}}$), the other nine channels measured the response at the interface of A ($\mathbf{u}_2^{\mathrm{A}}$), and the nine channels on B measured the response at the interface of B ($\mathbf{u}_3^{\mathrm{B}}$). Similarly, nine impacts were located at the interface of A ($\mathbf{f}_2^{\mathrm{A}}$), nine impacts at the interface of B ($\mathbf{f}_3^{\mathrm{B}}$), and three impacts away from the interface of B ($\mathbf{f}_4^{\mathrm{B}}$). The same configuration of channels and impacts was used for the coupled AJB structures, which were generated to evaluate the performance of the proposed joint identification approach. The assemblies, denoted AJB^i, represent the training samples obtained by coupling A, B, and the generated joint training samples J^i. In contrast, the assemblies AJB denote the test structures obtained by modeling the assemblies as a whole and used for validation.

The VPT was applied to both A and B to facilitate coupling with the joint models. The VPT projects translational DoFs at the interface onto a set of three translational and three rotational DoFs for both substructures A and B, while ensuring that the DoFs of the substructures are collocated with the joint DoFs. No measurements at the interface of AJB are required to identify the joints with the ANN. However, as shown in Fig. 7a, measurements were performed at the interface of AJB to enable a comparison of the proposed approach with the standard decoupling method[8]. To obtain the joint dynamics by decoupling, the VPT was applied to the AJB assembly's response model. The interface DoFs associated with A were projected onto one VP and the interface DoFs associated with B were projected onto the other VP using Eq. (6). By collocating the interface DoFs, both A and B can be decoupled from the AJB assembly, retaining only the joint dynamics. In contrast to the proposed approach, an obvious shortcoming of decoupling is the need to measure the interface dynamics of the assembly.

[8]Decoupling can be achieved by coupling the negative admittance matrices of A and B with AJB due to the altered equilibrium conditions (i.e., coupling $\mathbf{Y}^{\mathrm{AJB}}$, $-\mathbf{Y}^{\mathrm{A}}$ and $-\mathbf{Y}^{\mathrm{B}}$ using Eq. (5)).

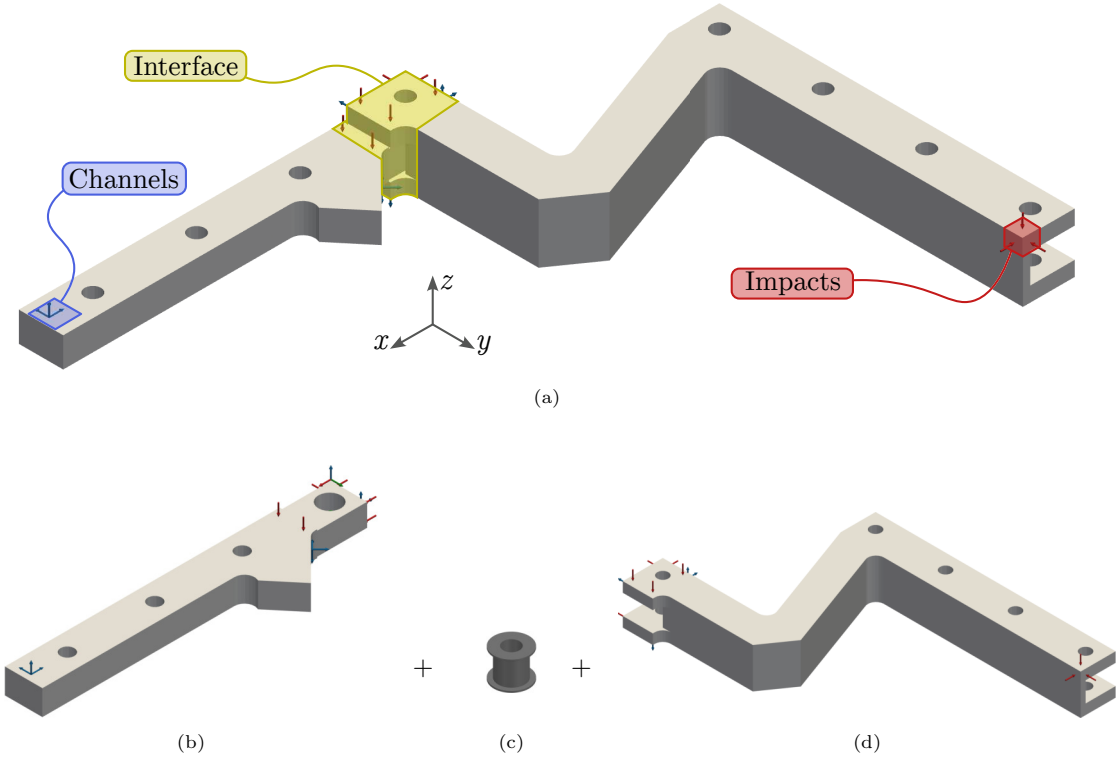

Figure 7: CAD models of the structures used for joint identification: (a) AJB, (b) A, (c) J, (d) B.

4.2. Numerical joint identification

The following presents a detailed explanation of how we obtained the initial data matrix $\mathbf{H}$. Applying the VPT to both A and B transforms the substructure response models to allow for the coupling with the joint response model samples J^i. The equation of motion for the uncoupled set of substructures A, J^i, and B is:

$$\mathbf{u} = \mathbf{Y}^{\mathrm{A}|\mathrm{J}^i|\mathrm{B}}(\mathbf{f} + \mathbf{g})\,. \tag{25}$$

The displacements, forces, and uncoupled admittance matrix are constructed using Eq. (26):

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}^{\mathrm{A}} \\ \mathbf{u}^{\mathrm{J}^i} \\ \mathbf{u}^{\mathrm{B}} \end{bmatrix}, \; \mathbf{f} = \begin{bmatrix} \mathbf{f}^{\mathrm{A}} \\ \mathbf{f}^{\mathrm{J}^i} \\ \mathbf{f}^{\mathrm{B}} \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} \mathbf{g}^{\mathrm{A}} \\ \mathbf{g}^{\mathrm{J}^i} \\ \mathbf{g}^{\mathrm{B}} \end{bmatrix}, \; \mathbf{Y}^{\mathrm{A}|\mathrm{J}^i|\mathrm{B}} = \begin{bmatrix} \mathbf{Y}^{\mathrm{A}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}^{\mathrm{J}^i} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y}^{\mathrm{B}} \end{bmatrix}, \tag{26}$$

with the elements of Eq. (26) defined by Eq. (27):

$$\mathbf{u}^{\mathrm{A}} = \begin{bmatrix} \mathbf{u}_1^{\mathrm{A}} \\ \mathbf{u}_2^{\mathrm{A}} \end{bmatrix}, \quad \mathbf{Y}^{\mathrm{A}} = \begin{bmatrix} \mathbf{Y}_{12}^{\mathrm{A}} \\ \mathbf{Y}_{22}^{\mathrm{A}} \end{bmatrix}, \quad \mathbf{f}^{\mathrm{A}} = \mathbf{f}_2^{\mathrm{A}}, \quad \mathbf{g}^{\mathrm{A}} = \mathbf{g}_2^{\mathrm{A}}, \tag{27a}$$

$$\mathbf{u}^{\mathrm{J}^i} = \begin{bmatrix} \mathbf{u}_2^{\mathrm{J}^i} \\ \mathbf{u}_3^{\mathrm{J}^i} \end{bmatrix}, \quad \mathbf{Y}^{\mathrm{J}^i} = \begin{bmatrix} \mathbf{Y}_{22}^{\mathrm{J}^i} & \mathbf{Y}_{23}^{\mathrm{J}^i} \\ \mathbf{Y}_{32}^{\mathrm{J}^i} & \mathbf{Y}_{33}^{\mathrm{J}^i} \end{bmatrix}, \quad \mathbf{f}^{\mathrm{J}^i} = \begin{bmatrix} \mathbf{f}_2^{\mathrm{J}^i} \\ \mathbf{f}_3^{\mathrm{J}^i} \end{bmatrix}, \quad \mathbf{g}^{\mathrm{J}^i} = \begin{bmatrix} \mathbf{g}_2^{\mathrm{J}^i} \\ \mathbf{g}_3^{\mathrm{J}^i} \end{bmatrix}, \tag{27b}$$

$$\mathbf{u}^{\mathrm{B}} = \mathbf{u}_3^{\mathrm{B}}, \quad \mathbf{Y}^{\mathrm{B}} = \begin{bmatrix} \mathbf{Y}_{33}^{\mathrm{B}} & \mathbf{Y}_{34}^{\mathrm{B}} \end{bmatrix}, \quad \mathbf{f}^{\mathrm{B}} = \begin{bmatrix} \mathbf{f}_3^{\mathrm{B}} \\ \mathbf{f}_4^{\mathrm{B}} \end{bmatrix}, \quad \mathbf{g}^{\mathrm{B}} = \begin{bmatrix} \mathbf{g}_3^{\mathrm{B}} \\ \mathbf{0} \end{bmatrix}. \tag{27c}$$

Using Eq. (5) yields the coupled admittance $\mathbf{Y}^{\mathrm{AJB}^i}$. The coupling of A and B with the analytical joint model is shown schematically in Fig. 8. The coupling is performed for each joint sample to obtain samples

of assembly admittances for the ANN training. A total of $m = 10\,000$ joint samples were generated to train the ANN. Fig. 9 presents the selection of FRFs from a coupled admittance matrix $\mathbf{Y}^{\mathrm{AJB}^i}$, leading to a row $\mathbf{h}^i$ of the initial data matrix $\mathbf{H}$. Only four FRFs were used for the joint identification using the ANN, three of which described the dynamic response to an impact in the same direction, while one FRF described the dynamic response in the y-direction to an impact in the x-direction. These FRFs were selected because of their relatively high dynamic response amplitudes compared to other FRFs. Alternatively, the FRFs can be selected by trial and error for different combinations of FRFs. By coupling the identified joints with A and B, a comparison can be made with the assembly FRFs and an optimal selection of FRFs can be achieved.

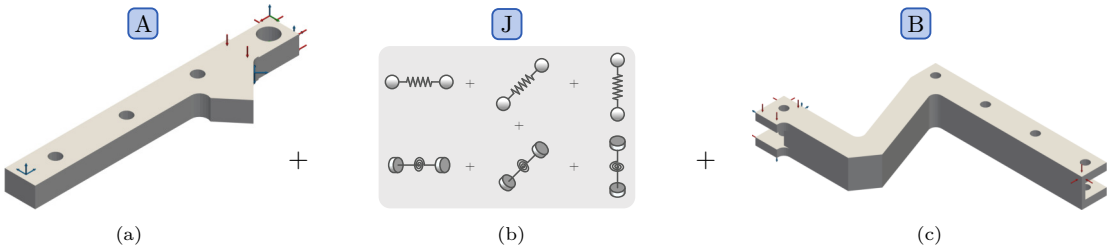

Figure 8: Substructures in the numerical study: (a) A, (b) J, (c) B

$$\mathbf{Y}^{\mathrm{AJB}^i} = \begin{bmatrix} \mathbf{Y}_{12}^{\mathrm{AJB}^i} & \mathbf{Y}_{13}^{\mathrm{AJB}^i} & \mathbf{Y}_{14}^{\mathrm{AJB}^i} \\ \mathbf{Y}_{22}^{\mathrm{AJB}^i} & \mathbf{Y}_{23}^{\mathrm{AJB}^i} & \mathbf{Y}_{24}^{\mathrm{AJB}^i} \\ \mathbf{Y}_{32}^{\mathrm{AJB}^i} & \mathbf{Y}_{33}^{\mathrm{AJB}^i} & \mathbf{Y}_{34}^{\mathrm{AJB}^i} \end{bmatrix} \quad \mathbf{Y}_{14}^{\mathrm{AJB}^i} = \begin{bmatrix} Y_{xx}^i & Y_{xy}^i & Y_{xz}^i \\ Y_{yx}^i & Y_{yy}^i & Y_{yz}^i \\ Y_{zx}^i & Y_{zy}^i & Y_{zz}^i \end{bmatrix}$$

$$\mathbf{h}^i = \mathrm{abs}\left(\begin{bmatrix} Y_{xx}^i & Y_{yx}^i & Y_{yy}^i & Y_{zz}^i \end{bmatrix}\right)$$

Figure 9: Selection of FRFs used in the joint identification using an ANN.

Each FRF was evaluated at frequency points ranging from 1 to 2000 Hz with a resolution of 1 Hz, resulting in a set of $n_{\mathrm{f}} = 2000$ values. With four FRFs used to identify the joint model parameters, i.e., $n_{\mathrm{FRFs}} = 4$, each row $\mathbf{h}^i$ of the initial data matrix $\mathbf{H}$ consisted of $n = n_{\mathrm{f}} \cdot n_{\mathrm{FRFs}} = 8000$ components. Using PCA, the number of components in the numerical study was reduced to 200 PCs, i.e., $p = 200$, where $\mathbf{A}_p \in \mathbb{R}^{m \times p}$ is the reduced data matrix used for the ANN training. The trained ANN was used to identify the analytical joint model parameters, and the joint admittance matrix was obtained using Eq. (13). The resulting ANN contained approximately 40 000 parameters, which had to be optimized. Due to the large number of parameters (relative to the number of training samples), the ANN was prone to overfitting. As Zhang et al. have shown, a sufficiently large neural network is capable of learning any dataset perfectly, even if the labels are assigned randomly [41]. Such an ANN cannot predict the labels of a test dataset due to the randomly assigned labels, as the ANN simply adapts too well to the training data. To circumvent the problem of overfitting, regularization methods are usually integrated to serve as additional constraints for the ANN parameter optimization, leading to a better generalization of the ANN [42].

The term regularization describes any procedure with the aim of improving the generalization of the ANN, i.e., increasing the prediction performance on the test set. In our example, batch normalization and L_2 regularization were implemented. While both methods have the effect of regularization on the ANN, the use of batch normalization negates the regularization effect of L_2 regularization. Nonetheless, both methods were implemented, as the use of L_2 regularization significantly increases the effective learning rate, leading to a faster optimization of the ANN, which is discussed in greater detail in [43].

A twelve-DoF joint model was also obtained using LM-FBS decoupling to compare the ANN's perfor-

mance with an existing joint identification approach. To decouple the substructures A and B from the AJB assembly, the equation of motion of the system, prior to decoupling, is written as:

$$\mathbf{u} = \mathbf{Y}^{\mathrm{AJB}|-\mathrm{A}|-\mathrm{B}}(\mathbf{f} + \mathbf{g}) \,, \tag{28}$$

where

$$\mathbf{Y}^{\mathrm{AJB}|-\mathrm{A}|-\mathrm{B}} = \begin{bmatrix} \mathbf{Y}^{\mathrm{AJB}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Y}^{\mathrm{A}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{Y}^{\mathrm{B}} \end{bmatrix} . \tag{29}$$

Using Eq. (5), the dynamics of A and B are removed from the assembly, retaining only the dynamics of the joint itself. To compare the two joint identification approaches, both the ANN-based joint and the decoupled joint were coupled to A and B, resulting in two separate reconstructions of the assembly's response model. Both reconstructed admittances were compared to the assembly FRFs to evaluate the advantages and shortcomings of the proposed approach. Since decoupling is a well-established approach for joint identification, it served as a benchmark for the ANN-based method. In addition, A and B were rigidly coupled to observe the influence of the joint dynamics on the assembly.

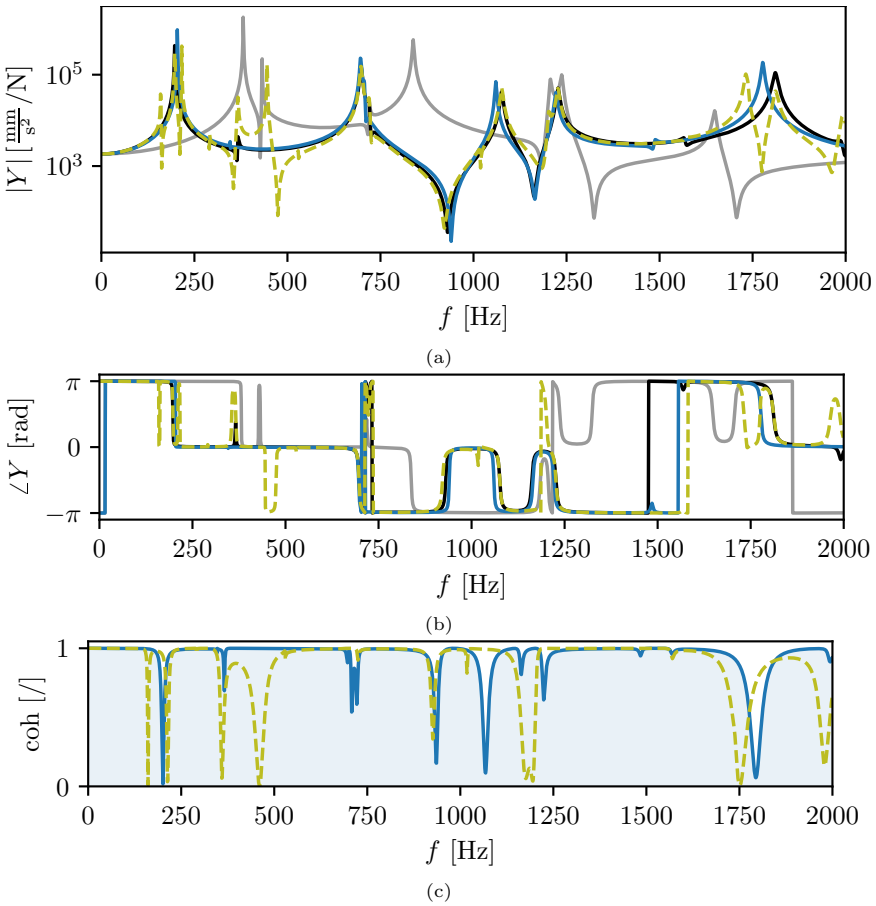

Figure 10: Numerical study FRF comparison of joint 1: (a) Magnitude, (b) Phase, (c) Coherence. Rigidly coupled AB (——), true response AJB (——), ANN-based AJB (——), decoupling-based AJB (- - -).

The magnitude plot (Fig. 10a) illustrates that neglecting the joint dynamics by rigidly coupling A and B cannot be justified, since this leads to an erroneous prediction. The reconstructed coupled FRFs based on decoupling are in reasonable agreement with the FRFs of the real assembly, except for spurious peaks, present across the observed frequency range, especially below 500 Hz. Performing additional simulations revealed

that the spurious peaks are present at the eigenfrequencies of the assembly with fixed displacements at the joint DoFs, as described in [44]. The ANN-based reconstructed coupled FRFs match the true assembly's FRFs better; therefore, the joint model can describe the true joint dynamics in the observed frequency range. Moreover, since no inversion takes place in the ANN-based approach, there are no spurious peaks present in the reconstructed FRFs.

A comparison of the reconstructed assembly FRFs, obtained using the ANN and the decoupling approach, is presented on the coherence plot (Fig. 10c), using the true dynamic response as a reference. The coherence is calculated by:

$$\mathrm{coh}(Y, Y^{\mathrm{ref}}) = \frac{(Y + Y^{\mathrm{ref}})(Y^* + Y^{\mathrm{ref}*})}{2(Y\,Y^* + Y^{\mathrm{ref}}\,Y^{\mathrm{ref}*})}\,. \tag{30}$$

The average coherence of the ANN and the decoupling approach are 0.94 and 0.89, respectively. It was observed that the coherence of the ANN-based assembly response drops slightly near the resonance peaks, as the ANN is not able to identify the joint parameters accurately enough so that the resonance frequencies would match exactly. The values of the identified joint parameters are listed in Table 2. It should be noted that the original joint model is a separate substructure and the proposed analytical joint model differs form the original joint model. Since the parameters of the analytical model are obtained using a trained ANN, the uncertainty of the parameters cannot be evaluated in comparison to the true joint model.

Table 2: Identified numerical model's joint parameters.

parameter	value	units
m	12.6	g
k_x	60.519	kN/mm
k_y	48.384	kN/mm
k_z	93.266	kN/mm
k_{tx}	62.697	kN m/rad
k_{ty}	18.732	kN m/rad
k_{tz}	4.297	kN m/rad

4.3. Discussion

Using the ANN-based approach, significantly fewer measurements are required to identify the joint dynamics, which is a major advantage over the decoupling approach. Both approaches require the dynamic response of A and B separately, accounting for 108 FRFs of A (12 channels and 9 impacts) and 108 FRFs of B (9 channels and 12 impacts). However, for the ANN-based approach, only nine assembly FRFs were measured for the ANN (using a single triaxial accelerometer and 3 distinct impacts). In contrast, 441 assembly FRFs were used for the decoupling approach (21 channels and 21 impacts). Moreover, the trained ANN can identify a variety of joints by measuring only nine FRFs for each new joint. This makes the approach experimentally appealing, as the number of measurements needed to identify the joint is significantly reduced compared to the decoupling approach.

The presented results can only be achieved if the dynamics of the joint measurably affect the assembly dynamics at the measured DoFs (observability and controllability). Although only four FRFs are used in the identification process, the joint dynamics can be observed through different vibration modes of the assembly in the chosen frequency range. It should be noted that the joint model must be complex enough to sufficiently describe the joint dynamics. On the other hand, an overly-complex joint model would eventually require more measurements to ensure that all of the parameters are observable in the assembly dynamics. If the reconstructed assembly dynamics significantly deviate from the true assembly dynamics, the joint model might not be complex enough or more measurements should be included in the identification procedure. To increase the model's complexity, additional mass and spring elements or cross-coupling terms between the joint DoFs can be considered. Since a good prediction of the assembly dynamics is achieved, the analytical model of the joint seems to be adequate for the presented application.

Due to efficient experimental testing, substructures A and B can be used as a test rig for identifying multiple joints. To demonstrate the ability to identify different joints, two additional joints were simulated by changing the Young's modulus of the joint material, and the corresponding coupled responses were obtained from the finite-element model. For the sake of brevity, only the magnitude plots are included in Fig. 11.

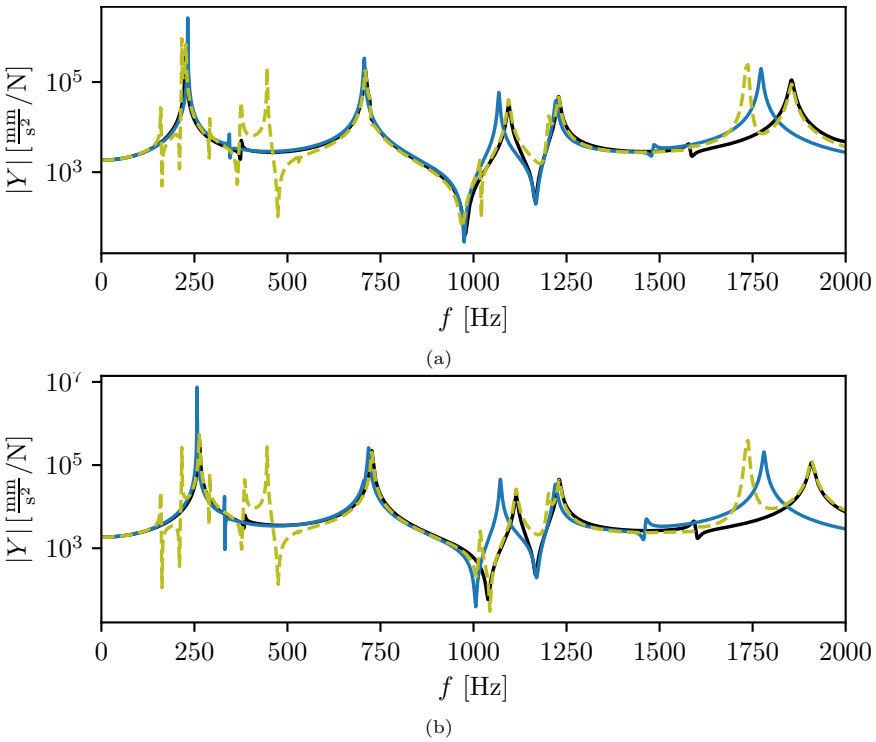

Figure 11: Numerical study FRF comparison of two additional joints: (a) Magnitude - joint 2, (b) Magnitude - joint 3. True response AJB (—), ANN-based AJB (—), decoupling-based AJB (- - -).

The average coherence of the decoupling-based AJB response was higher than the ANN-based AJB response, although spurious peaks were observed in the decoupling-based AJB response. The ANN-based AJB response matches the true assembly response for both joints sufficiently well. However, some eigenfrequencies were misidentified, leading to the discrepancy between the ANN-based response and the true response, as well as a decrease in the coherence near the eigenfrequencies.

5. Experimental case-study

Since the numerical study yielded encouraging results, an experimental case-study was conducted to evaluate the real-world applicability of the proposed joint identification approach. The geometry of the physical substructures used in the experimental example was consistent with the substructures from the numerical study. Due to the different material parameters, as well as the possibility for different interactions at the interface, the eigenfrequencies of the experimental assembly were not consistent with the eigenfrequencies of its numerical counterpart. In addition, the response of the experimental assembly was significantly affected by damping. Therefore, the damping parameters of the joint could not have been neglected, resulting in a more complex analytical joint model for which additional joint model parameters had to be included.

5.1. Experimental setup

The experimental setup consisted of substructures A, J, and B, seven triaxial accelerometers, and an impact hammer. The positioning of the sensors and impacts was consistent with the numerical study,

allowing the use of VPT. The main sources of error in the FRFs obtained by impact excitation include uncertainties in the location and orientation of the impacts. Inaccurate measurements can result in a significantly lower accuracy of the VPT. Therefore, performing the measurements precisely is necessary for accurate VP FRFs [45]. The experimental setup is shown in Fig. 12. The structures were suspended using foam pads to simulate free-free boundary conditions.

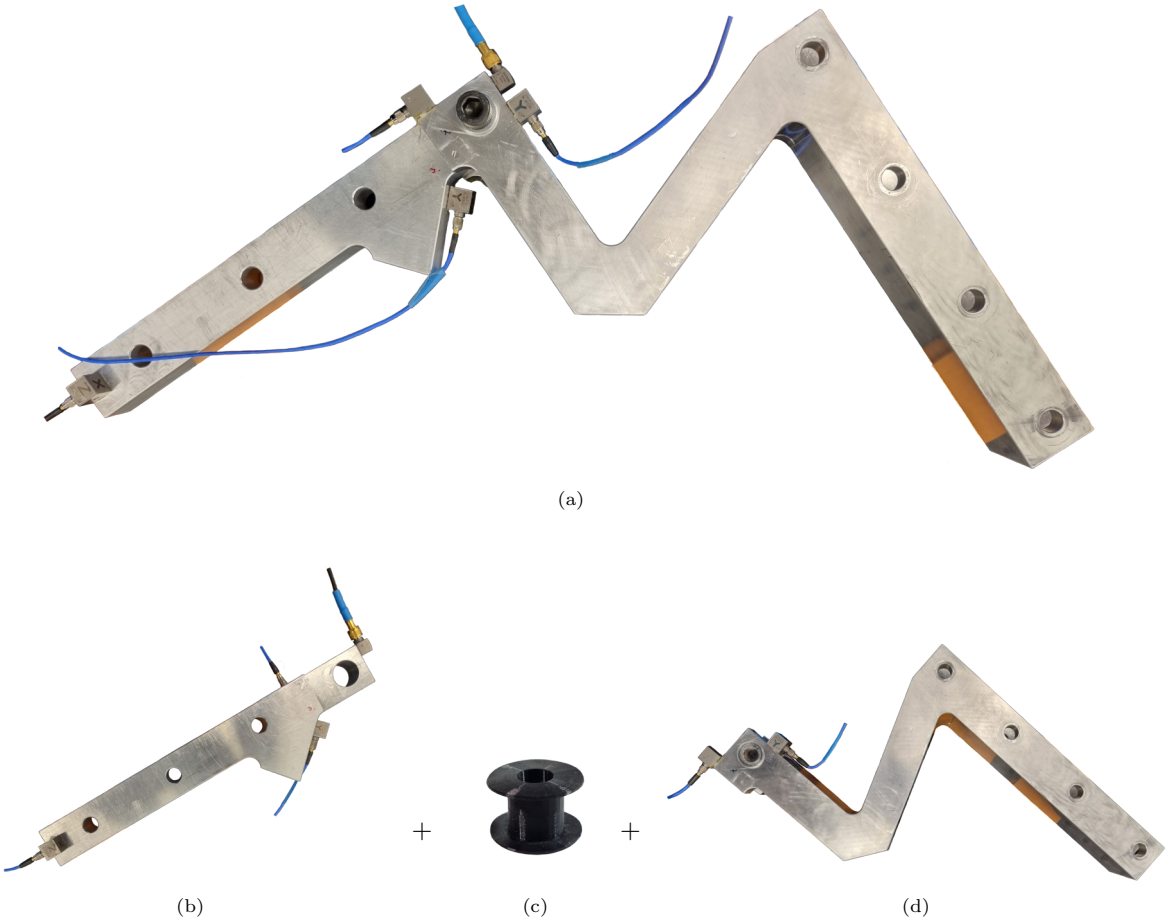

Figure 12: Experimentally equipped physical substructures: (a) AJB, (b) A, (c) J, (d) B.

5.2. Experimental joint identification

The joint model used in the experimental example contained 13 parameters: the joint mass, six stiffness coefficients, and six damping coefficients. Damping was accounted for using Eq. (14), which required the identification of six modal damping ratios. Successful identification becomes increasingly difficult as the number of joint parameters increases. The number of joint parameters was reduced by making all the modal damping ratios associated with translational modes equal, and the same was done for the rotational modal damping ratios. Therefore, the identification was facilitated by reducing the number of independent joint parameters. The number of joint model parameters for identification was reduced to nine.

Data preparation for the experimental joint identification followed the same procedure as described in Section 4, i.e., FRF selection and PCA. Again, 200 PCs were considered as the input data features for the joint identification. Due to the large number of neural network parameters, the ANN was again prone to overfitting. To overcome this problem, batch normalization was implemented in each layer of the ANN in

conjunction with L_2 regularization, analogous to the ANN in the numerical study. Batch normalization provided a good generalization, meaning the testing loss closely followed the training loss throughout the training iterations. As described in subsection 4.2, L_2 regularization has no regularization effect in the presence of batch normalization, although the effective learning rate is increased, justifying the use of both regularization methods.

The results of the experimental joint identification are shown in Fig. 13. Again, the decoupling approach leads to spurious peaks. In contrast, the response of the ANN-based approach resembles the true coupled response with some eigenfrequencies misaligned due to misidentified joint parameters or the lack of joint model complexity to adequately describe the true joint's dynamics. The average coherence of the decoupling-based response and the ANN-based response were 0.80 and 0.93, respectively. Again, the coherence was measured with respect to the true coupled response. The values of the identified parameters are listed in Table 3.

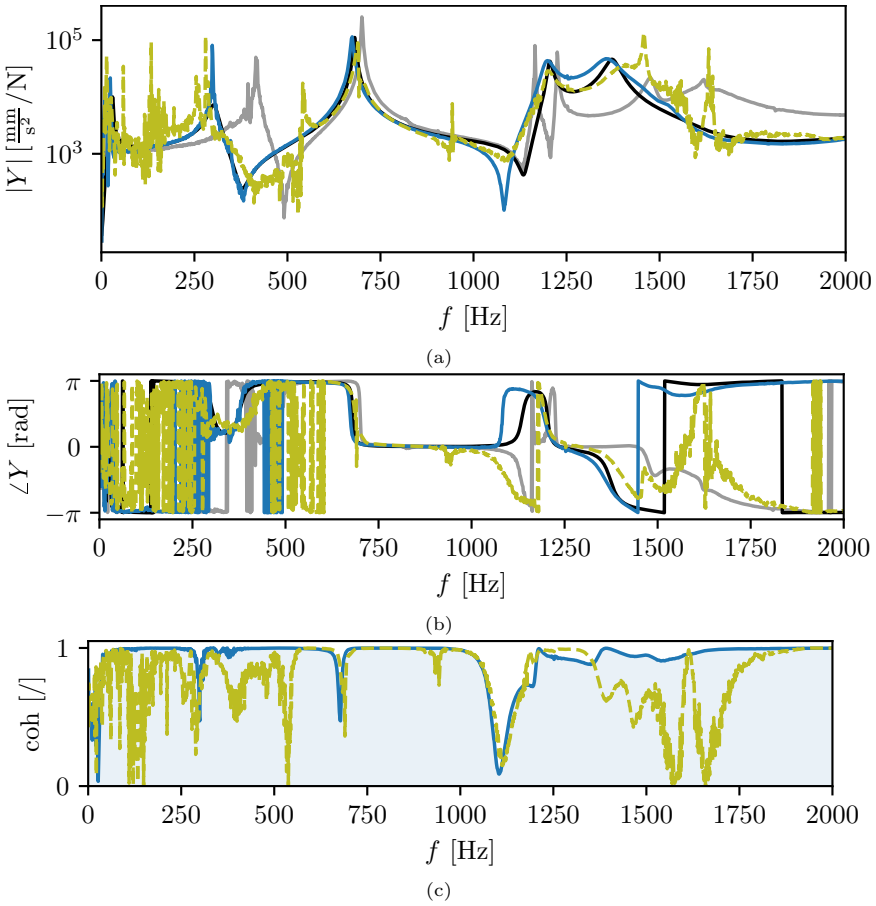

Figure 13: Experimental example FRF comparison: (a) Magnitude plot, (b) Phase plot, (c) Coherence. Rigidly coupled AB (—), true response AJB (—), ANN-based AJB (—), decoupling-based AJB (- - -).

Table 3: Experimentally identified joint model parameters.

parameter	value	units
m	37.4	g
k_x	57.344	kN/mm
k_y	14.379	kN/mm
k_z	149.767	kN/mm
k_{tx}	9.005	kN m/rad
k_{ty}	15.401	kN m/rad
k_{tz}	3.266	kN m/rad
ζ_r^{tr}	1.49	/
ζ_r^{rot}	93.8	/

5.3. Discussion

The identified values of the joint parameters follow a similar trend in both the numerical and experimental studies. In both cases, the translational stiffness k_z is higher than k_x and k_y, while the rotational stiffness k_{tz} is lower than k_{tx} and k_{ty}. A high stiffness k_z is caused by the geometry of the joint. Only two thin rubber flanges ensure the substructures A and B are not in direct contact, therefore a relatively high stiffness in the z-direction is expected compared to the x and y directions. The same flanges resist the relative rotation of the substructures about the x and y axes, therefore the rotational stiffness about the z-axis is expected to be relatively low. In the experimental study, the identified rotational damping ratio ζ_r^{rot} is significantly higher than the translational damping ratio ζ_r^{tr}. Again, due to the geometry of the joint, less effort was needed to induce a relative angular displacement of the substructures A and B in the assembled state due to slipping. Therefore, slipping about the z-axis could explain the higher rotational damping.

The proposed ANN-based approach has several advantages. Compared to the decoupling joint identification approach, an overall increase in the average coherence was achieved. The ANN-based response is not only less noisy, but also more accurately matches the response of the real assembly since there are no spurious peaks. Further research could improve the proposed approach, e.g., by using a different joint model or ML model. The computational model of the joint is relatively simple and does not consider the effects of cross-coupling between the individual DoFs (the displacement of a node in a particular direction can only be caused by a load in the same direction). The results can be improved by considering cross-coupling, which is permitted in the FBS approach.

6. Conclusion

In this work an ANN-based approach was used to identify the dynamic properties of a joint. The approach relies on the VPT to obtain collocated DoF models of the substructures A and B, connected by the joint. A physics-based computational model (i.e., an analytical joint model) is proposed to describe the true joint's dynamics. Using the computational model, joint samples are obtained by randomly generating joint model parameters. The LM-FBS coupling is used to obtain the dynamic response of the assembly for each joint sample, and PCA is used for the dimensionality reduction. An ANN is constructed, and the PCs are used as the input with the corresponding values of the joint model parameters as the output. After the ANN training, the approach is validated by calculating the joint admittance matrix from the identified parameters, coupling the joint with substructures A and B, and comparing the coupled response with the original assembly response.

The ANN was first validated in a numerical study to evaluate its ability to identify the joint dynamics. Based on the promising results, an experimental example was conducted. The damping parameters of the joint model could not be neglected in the experimental example since the response was significantly damped. Although additional joint parameters had to be identified, the predicted response matched well with the original response of the assembly.

The proposed ANN-based approach is experimentally appealing. It allows an efficient identification of different joints, as significantly fewer measurements are required for the identification compared to the decoupling approach. Although there might be some loss of accuracy in the identified eigenfrequencies of the coupled response, the spurious peaks are not present, in contrast to the decoupling approach. Further research could improve the ANN-based approach, e.g., by proposing a different analytical joint model or by modifying the ML algorithm used to predict the values of the joint model parameters.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge financial support from the European Union's Horizon Europe research and innovation programme under the grant agreement No 101091536 (DiCiM project).

References

[1] J.-S. Tsai, Y.-F. Chou, The identification of dynamic characteristics of a single bolt joint, Journal of Sound and Vibration 125 (3) (1988) 487–502.
[2] Y. Ren, C. F. Beards, Identification of joint properties of a structure using FRF data, Journal of Sound and Vibration 186 (4) (1995) 567–587.
[3] D. Čelič, M. Boltežar, Identification of the dynamic properties of joints using frequency–response functions, Journal of Sound and Vibration 317 (1-2) (2008) 158–174.
[4] D. Čelič, M. Boltežar, The influence of the coordinate reduction on the identification of the joint dynamic properties, Mechanical Systems and Signal Processing 23 (4) (2009) 1260–1271.
[5] M. Wang, D. Wang, G. Zheng, Joint dynamic properties identification with partially measured frequency response function, Mechanical Systems and Signal Processing 27 (2012) 499–512.
[6] M. Mehrpouya, E. Graham, S. S. Park, Identification of multiple joint dynamics using the inverse receptance coupling method, Journal of Vibration and Control 21 (16) (2015) 3431–3449.
[7] M. Mehrpouya, M. Sanati, S. S. Park, Identification of joint dynamics in 3D structures through the inverse receptance coupling method, International Journal of Mechanical Sciences 105 (2016) 135–145.
[8] G. Dong, W. Wu, J. Zhou, Joints dynamic identification and modeling based on FRFs data, in: MATEC Web of Conferences, Vol. 189, EDP Sciences, 2018, p. 01010.
[9] J. W. R. Meggitt, A. T. Moorhouse, In-situ sub-structure decoupling of resiliently coupled assemblies, Mechanical Systems and Signal Processing 117 (2019) 723–737.
[10] M. Haeussler, S. W. B. Klaassen, D. J. Rixen, Experimental twelve degree of freedom rubber isolator models for use in substructuring assemblies, Journal of Sound and Vibration 474 (2020) 115253.
[11] S. W. B. Klaassen, D. J. Rixen, Using SEMM to identify the joint dynamics in multiple degrees of freedom without measuring interfaces, in: Dynamic Substructures, Volume 4, Springer, 2020, pp. 87–99.
[12] D. Minervini, A. Stella, Experimental frequency-based decoupling for tire characterization (09 2022).
[13] N. Jamia, H. Jalali, J. Taghipour, M. I. Friswell, H. H. Khodaparast, An equivalent model of a nonlinear bolted flange joint, Mechanical Systems and Signal Processing 153 (2021) 107507.
[14] S. L. Brunton, J. N. Kutz, Data-driven science and engineering: Machine learning, dynamical systems, and control, Cambridge University Press, 2022.
[15] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, D. J. Schwab, A high-bias, low-variance introduction to machine learning for physicists, Physics reports 810 (2019) 1–124.
[16] C. M. Bishop, N. M. Nasrabadi, Pattern recognition and machine learning, Vol. 4, Springer, 2006.
[17] S. Shalev-Shwartz, S. Ben-David, Understanding machine learning: From theory to algorithms, Cambridge university press, 2014.
[18] M. Crawford, T. M. Khoshgoftaar, J. D. Prusa, A. N. Richter, H. Al Najada, Survey of review spam detection using machine learning techniques, Journal of Big Data 2 (1) (2015) 1–24.
[19] S. Kuutti, R. Bowden, Y. Jin, P. Barber, S. Fallah, A survey of deep learning applications to autonomous vehicle control, IEEE Transactions on Intelligent Transportation Systems 22 (2) (2020) 712–733.
[20] Y. Xiao, Z. Tian, J. Yu, Y. Zhang, S. Liu, S. Du, X. Lan, A review of object detection based on deep learning, Multimedia Tools and Applications 79 (33) (2020) 23729–23791.
[21] F. Wang, G. Song, A novel percussion-based method for multi-bolt looseness detection using one-dimensional memory augmented convolutional long short-term memory networks, Mechanical Systems and Signal Processing 161 (2021) 107955.

[22] D. P. Yang, D. F. Song, X. H. Zeng, X. L. Wang, X. M. Zhang, Adaptive nonlinear ANC system based on time-domain signal reconstruction technology, Mechanical Systems and Signal Processing 162 (2022) 108056.
[23] A. Movsessian, D. G. Cava, D. Tcherniak, An artificial neural network methodology for damage detection: Demonstration on an operating wind turbine blade, Mechanical Systems and Signal Processing 159 (2021) 107766.
[24] W. Booyse, D. N. Wilke, S. Heyns, Deep digital twins for detection, diagnostics and prognostics, Mechanical Systems and Signal Processing 140 (2020) 106612.
[25] M. Stender, M. Tiedemann, D. Spieler, D. Schoepflin, N. Hoffmann, S. Oberst, Deep learning for brake squeal: Brake noise detection, characterization and prediction, Mechanical Systems and Signal Processing 149 (2021) 107181.
[26] D. Jones, C. Snider, A. Nassehi, J. Yon, B. Hicks, Characterising the digital twin: A systematic literature review, CIRP Journal of Manufacturing Science and Technology 29 (2020) 36–52.
[27] J. Willard, X. Jia, S. Xu, M. Steinbach, V. Kumar, Integrating physics-based modeling with machine learning: A survey, arXiv preprint arXiv:2003.04919 1 (1) (2020) 1–34.
[28] T. G. Ritto, F. A. Rochinha, Digital twin, physics-based model, and machine learning applied to damage detection in structures, Mechanical Systems and Signal Processing 155 (2021) 107614.
[29] D. De Klerk, D. J. Rixen, S. N. Voormeeren, General framework for dynamic substructuring: history, review and classification of techniques, AIAA journal 46 (5) (2008) 1169–1181.
[30] M. Haeussler, D. Kobus, D. Rixen, Parametric design optimization of e-compressor NVH using blocked forces and substructuring, Mechanical Systems and Signal Processing 150 (2021) 107217.
[31] A. Altan, R. Hacıoğlu, Model predictive control of three-axis gimbal system mounted on UAV for real-time target tracking under external disturbances, Mechanical Systems and Signal Processing 138 (2020) 106548.
[32] E. A. Pasma, M. van der Seijs, S. W. B. Klaassen, M. W. van der Kooij, Frequency based substructuring with the virtual point transformation, flexible interface modes and a transmission simulator, in: Dynamics of Coupled Structures, Volume 4, Springer, 2018, pp. 205–213.
[33] M. van der Seijs, D. van den Bosch, D. J. Rixen, D. de Klerk, An improved methodology for the virtual point transformation of measured frequency response functions in dynamic substructuring, COMPDYN (2013).
[34] M. van der Seijs, Experimental dynamic substructuring: Analysis and design strategies for vehicle development, Ph.D. thesis, Delft University of Technology (2016).
[35] T. Bregar, A. El Mahmoudi, M. Kodrič, D. Ocepek, F. Trainotti, M. Pogačar, M. Göldeli, G. Čepon, M. Boltežar, D. J. Rixen, pyFBS: A Python package for Frequency Based Substructuring, Journal of Open Source Software 7 (69) (2022) 3399.
[36] M. Häußler, Modular sound & vibration engineering by substructuring, Ph.D. thesis, Universität München (2021).
[37] C. Zang, M. Imregun, Structural damage detection using artificial neural networks and measured FRF data reduced via principal component projection, Journal of Sound and Vibration 242 (5) (2001) 813–827.
[38] C. Zang, M. Imregun, Combined neural network and reduced FRF techniques for slight damage detection using measured response data, Archive of Applied mechanics 71 (8) (2001) 525–536.
[39] I. T. Jolliffe, Principal Component Analysis, 2nd Edition, Springer Series in Statistics, Springer, 2002.
[40] S. Banerjee, A. Roy, Linear algebra and matrix analysis for statistics, Vol. 181, Crc Press Boca Raton, FL, USA, 2014.
[41] C. Zhang, S. Bengio, M. Hardt, B. Recht, O. Vinyals, Understanding deep learning (still) requires rethinking generalization, Communications of the ACM 64 (3) (2021) 107–115.
[42] J. Kukačka, V. Golkov, D. Cremers, Regularization for deep learning: A taxonomy, arXiv preprint arXiv:1710.10686 (2017).
[43] T. Van Laarhoven, L2 regularization versus batch and weight normalization, arXiv preprint arXiv:1706.05350 (2017).
[44] S. W. B. Klaassen, M. van der Seijs, D. de Klerk, System equivalent model mixing, Mechanical Systems and Signal Processing 105 (2018) 90–112.
[45] G. Čepon, D. Ocepek, J. Korbar, T. Bregar, M. Boltežar, Sensitivity-based characterization of the bias errors in frequency based substructuring, Mechanical Systems and Signal Processing 170 (2022) 108800.