



# Integrating automatic order determination with response prediction error minimization for nonlinear subspace identification in structural dynamics

Dong Jiang · Ang Li · Yusheng Wang · Shitao Xie · Zhifu Cao · Rui Zhu

Received: 15 November 2023 / Accepted: 19 August 2024  
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

**Abstract** The singular value decomposition order determination method used in nonlinear subspace identification may encounter challenges due to the noise in signals, resulting in the omission of modes or the occurrence of spurious mode. Additionally, the lack of an iterative process in the nonlinear subspace identification, which primarily relies on matrix operations, will lead to suboptimal solutions. To address these challenges, an improved framework for nonlinear subspaces identification is proposed in this paper. False modes are eliminated through data preprocessing and modal stability criteria, followed by the clustering of stable modes using the Density-Based Spatial Clustering of Applications with Noise algorithm for automatic order determination; Simultaneously, an iterative optimization approach based on

response prediction error minimization is introduced to enhance the accuracy of the state-space model estimation results. The effectiveness of proposed method is validated through two simulation cases and one experimental verification. The results show that the clustering algorithm effectively distinguishes real modes from false ones and achieves automatic system order determination across various SNR conditions. The iterative optimization process notably enhances state-space model estimation accuracy. Compared to original nonlinear subspace identification, proposed method significantly improves identification accuracy.

**Keywords** Nonlinear subspace identification · DBSCAN algorithm · Automatic order determination · State-space model optimization

---

D. Jiang (✉) · A. Li · Y. Wang · S. Xie  
School of Mechanical and Electronic Engineering,  
Nanjing Forestry University, Nanjing 210037, China  
e-mail: jiangdong@njfu.edu.cn

Z. Cao  
College of Aerospace Engineering, Chongqing  
University, Chongqing 400044, China

R. Zhu  
School of Engineering and Design, Technical University  
of Munich, 80539 Munich, Germany

R. Zhu  
Institute of Aerospace Machinery and Dynamics,  
Southeast University, Nanjing 211189, China

## 1 Introduction

Nonlinear system identification refers to determining the mathematical model of a nonlinear dynamical system from input and output data [1–3]. Accurate dynamic models play a crucial role in various fields such as vibration isolation, structural optimization, and damage identification [4–6]. Several papers have reviewed the development history and major methods of nonlinear system identification [7–9].

In recent years, many nonlinear system identification methods have been proposed and developed [10–12]. One notable method is the Nonlinear Subspace Identification (NSI) introduced by Marchesiello and Garibaldi [13, 14]. This method combines nonlinear feedback interpretation with the classical subspace identification framework, showcasing strong effectiveness and robustness, making it one of the classic approaches in nonlinear system identification [15–17]. Anastasio et al. [18] integrated NSI with the harmonic balance method, utilizing this enhanced approach to identify and analyze periodically excited dynamic systems to investigate the stability of system periodic solutions, showcasing promising prospects for applications in rotating machinery such as rotors. Zhu et al. [19] applied NSI to the aerodynamics domain, enabling systematic identification and analysis of the nonlinear airfoil-store system, demonstrating its robustness and flexibility across various wind speeds. Liu et al. [20] combined NSI with the Krotov method to establish a nonlinear Hammerstein model for multivariate molten iron quality, addressing the control issues encountered in the blast furnace iron-making process.

However, NSI still faces significant challenges, notably in accurately determining the system order, a crucial aspect of the method. In NSI method, the system order is determined using Singular Value Decomposition (SVD), where the number of singular values before the largest difference between adjacent singular values corresponds to the system order [13]. But the higher-order singular values of the Hankel matrix are often susceptible to noise, hindering the clear detection of singular value jumps and accurate system order determination, leading to mode omission or pseudo-mode phenomena [21, 22]. To address this challenge, numerous scholars have conducted research [23–25]. Among these efforts, the introduction of stability diagram techniques has provided a novel approach to determining system order [26, 27]. Zhou et al. [28] proposed a stability diagram method that combines Monte Carlo techniques, effectively distinguishing between system modes and spurious modes caused by noise, thereby accurately determining the system's order. The application of clustering techniques also supports the automatic determination of system order [29, 30]. Bakir [31] introduced the modal phase collinearity index to remove false modal poles, followed by hierarchical clustering to group real

modes into large clusters. Zhang et al. [32] proposed an algorithm utilizing fast-density peak clustering, enabling the automatic determination of system modal parameters without requiring user intervention.

Moreover, the NSI method suffers from the limitation that its computed identification results might not always represent the optimal solution, thereby compromising the accuracy of the identification process [13]. To address this issue, Wei et al. [33] introduced an enhanced NSI method that integrates the prediction error method (PEM) to re-estimate the coefficient matrix within the state-space representation. This refinement step occurs after obtaining the initial system model through the NSI method. The enhanced approach has demonstrated significantly improved recognition accuracy, especially in challenging and noisy environments, surpassing the performance of the conventional NSI method.

This paper addresses the issues above by proposing an improved framework for nonlinear subspace identification. Through the amalgamation of the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering algorithm and the stabilization graph method, automatic determination of system order is achieved. Before ordering, eliminate false modes and erroneous estimates caused by noise through data preprocessing and modal stability criteria, thereby enhancing the algorithm's robustness to noise and improving computational efficiency. At the same time, an iterative optimization algorithm based on response prediction error minimization is introduced in the process of estimating the state-space model, ensuring that the obtained state-space matrices are all optimal solutions, further enhancing the identification accuracy. Finally, the feasibility of the proposed method is validated through two multi-degree-of-freedom simulation cases, complemented by experimental verification on a multilayer building with nonlinear characteristics.

## 2 Theoretical basis

### 2.1 Automatic order determination

NSI is a nonlinear system identification method that establishes a relationship between the underlying frequency response function and nonlinear parameters. Appendix A offers a succinct introduction to the

NSI method. Estimating the state-space model constitutes the central aspect of the NSI method. For a nonlinear system with  $N$  degrees of freedom, the state-space model can be represented as follows:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}_c \mathbf{x}(t) + \mathbf{B}_c \mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C} \mathbf{x}(t) + \mathbf{D} \mathbf{u}(t) \end{aligned} \quad (1)$$

$$\left\{ \begin{array}{l} \mathbf{A}_c = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1} \mathbf{K} & -\mathbf{M}^{-1} \mathbf{C}_v \end{bmatrix}, \mathbf{B}_c = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{M}^{-1} & \theta_1 \mathbf{M}^{-1} \mathbf{L}_1 & \cdots & \theta_p \mathbf{M}^{-1} \mathbf{L}_p \end{bmatrix} \\ \mathbf{C} = [\mathbf{I}_{N \times N} & \mathbf{0}_{N \times N}], \mathbf{D} = [\mathbf{0}_{N \times N} & \mathbf{0}_{N \times 1} & \cdots & \mathbf{0}_{N \times 1}] \end{array} \right.$$

The continuous time state-space model can be discretized using a sampling interval  $\Delta t$ , with the relationship between the continuous time and discrete time models expressed as:

$$\mathbf{A} = e^{\mathbf{A}_c \Delta t}, \mathbf{B} = (\mathbf{e}^{\mathbf{A}_c \Delta t} - \mathbf{I}) \mathbf{A}_c^{-1} \mathbf{B}_c \quad (2)$$

In Eq. (1),  $\mathbf{A}_c$  represents the dynamical system matrix,  $\mathbf{B}_c$  represents the input matrix,  $\mathbf{C}$  represents the output matrix, and  $\mathbf{D}$  represents the direct feedthrough matrix. The mass matrix  $\mathbf{M}$ , stiffness matrix  $\mathbf{K}$ , and damping matrix  $\mathbf{C}_v$  constitute the underlying linear components, while local nonlinear factors are governed by position vector  $\mathbf{L}_i$  and nonlinear parameter  $\theta_i$ . The output  $\mathbf{y}(t)$  is a  $q$ -dimensional column vector, where  $t$  denotes time. The input  $\mathbf{u}(t)$  is an  $m$ -dimensional column vector, and the system order, denoted by the dimension of the state vector  $\mathbf{x}(t)$ , is  $n$ .

As mentioned in the introduction, NSI methods determine the system order  $n$  through singular value decomposition, which is susceptible to noise interference. The following section introduces an improved method to address automatic order determination in noisy environments.

### 2.1.1 Data preprocessing

After each NSI run, matrices  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{D}$  are identified. The system matrix  $\mathbf{A}$  contains modal information and the system order  $n$  can be determined by the stable modal parameter count.

The number of identifiable modal parameters varies with different orders of  $\mathbf{A}$ . Initially, the system order is

changed from a small value  $n_{min}$ , to a sufficiently large value  $n_{max}$  and NSI is performed separately for each. Assuming  $R$  identifications are conducted during this process, performing eigenvalue decomposition on the  $R$  identified  $\mathbf{A}$  matrices yields  $R_n$  complex eigenval-

ues. The eigenvalues are represented by the vector  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_{Rn}]^T$  and the eigenvectors by the matrix  $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \dots, \boldsymbol{\phi}_{Rn}]$ . To match the modal shape measured from the sensor, each complex eigenvector is multiplied by the output matrix  $\mathbf{C}$  associated with the identification, and the resulting matrix is still denoted by  $\boldsymbol{\Phi}$ . The modal frequency and damping ratios of the system can be obtained from  $\boldsymbol{\lambda}$ , and the modal shapes can be obtained from  $\boldsymbol{\Phi}$ .

During the process, noise can introduce errors in the decomposition of matrix  $\mathbf{A}$ . Hence, some evidently incorrect decomposition outcomes should be eliminated. Specifically, the correct eigenvalues should possess the following properties:

1. The real part of the eigenvalues is directly linked to the damping ratio. It is crucial for this real part to be non-negative, as negative values would not make physical sense.
2. The eigenvalues need to belong to the complex number field, indicating that their imaginary parts should not be zero.
3. The eigenvalues always appear as complex conjugate pairs. This property reflects the inherent symmetry of the system and should be maintained throughout the decomposition process.

These properties can be expressed mathematically:

$$\begin{cases} \text{Re}(\lambda_i) > 0 \\ \text{Im}(\lambda_i) \neq 0 \\ \lambda_j = \lambda_i^* \end{cases} \quad (3)$$

According to Eq. (3), any eigenvalue  $\lambda_i$  and its corresponding eigenvector  $\phi_i$  that do not satisfy any of the three conditions are eliminated. For each pair of complex conjugate modes, only one eigenvalue and one eigenvector are retained. Consequently, the correct eigenvalue decomposition is obtained, consisting of the remaining  $R_r$  eigenvalues and eigenvectors denoted by  $\lambda_r = [\lambda_1, \lambda_2, \dots, \lambda_{R_r}]^T$  and  $\Phi = [\phi_1, \phi_2, \dots, \phi_{R_r}]$  respectively.

This data preprocessing operation effectively eliminates numerous erroneous estimates, resulting in an improved quality of analysis. Moreover, the computational workload of the algorithm is significantly reduced as a result.

### 2.1.2 Mode stability criteria

After removing erroneous estimates, stable modes are determined by specifying modal stability criteria.

The vectors  $f \in \mathbb{R}^{R_r}$  and  $\zeta \in \mathbb{R}^{R_r}$  are introduced to capture the frequency and damping ratio from the eigenvalue vector  $\lambda_r$ , respectively. Consider a pair of identified modes characterized by their respective frequency, damping ratio, and mode shape vector, denoted as  $f_i, \zeta_i$  and  $\phi_i$  for one mode, and  $f_j, \zeta_j$  and  $\phi_j$  for the other mode. For the quantitative assessment of the correlation between these modes, the following calculation criteria are introduced:

$$\begin{aligned} \Delta f_{ij} &= \frac{|f_i - f_j|}{f_j} \\ \Delta \zeta_{ij} &= \frac{|\zeta_i - \zeta_j|}{\zeta_j} \end{aligned} \tag{4}$$

$$MAC_{ij} = MAC(\phi_i, \phi_j)$$

where **MAC** is defined as the mode confidence criterion, it quantifies the correlation between the identified mode shape  $\phi_i$  and  $\phi_j$ . **MAC** is calculated using the formula  $|\phi_i^T \phi_j|^2 / (\phi_i^T \phi_i) \cdot (\phi_j^T \phi_j)$ , with values ranging from 0 to 1. **MAC** value closer to 1 indicates a stronger correlation, suggesting that the identified two modes are of the same order. Conversely, **MAC** value of 0 indicates no relationship between the modes.

Equation (4) indicate the degree of similarity between the two modes. Define a matrix  $\Delta$  where the elements in the matrix  $\Delta$  are represented as:

$$\Delta_{ij} = \begin{cases} 1, & \Delta f_{ij} \leq \varepsilon_f, \Delta \zeta_j \leq \varepsilon_\zeta, 1 - MAC_{ij} \leq \varepsilon_{MAC} \\ 0, & \text{else} \end{cases} \tag{5}$$

where  $\varepsilon_f, \varepsilon_\zeta$ , and  $\varepsilon_{MAC}$  represent custom tolerances for frequencies, damping ratios, and mode shapes, respectively, all pertaining to modes of the same order. In this paper, the tolerance values were determined as follows:  $\varepsilon_f = 0.01, \varepsilon_\zeta = 0.04, \varepsilon_{MAC} = 0.05$ .

When  $\Delta_{ij} = 1$ , it signifies that the  $i_{th}$  and  $j_{th}$  modes belong to the same mode order. Consequently, by performing a statistical analysis on the  $j_{th}$  column of the matrix  $\Delta$ , the count of modes belonging to the same mode order as the  $j_{th}$  mode can be determined. Define a new vector  $z$  to represent the statistical result of all modes, whose elements are integers, and the element values can be expressed as:

$$z_j = \begin{cases} -1 + \sum_{i=1}^{N_r} \Delta_{ij}, & \zeta_j \in [0, \zeta_{max}] \\ 0, & \zeta_j \notin [0, \zeta_{max}] \end{cases} \tag{6}$$

where  $\zeta_{max}$  is the maximum value of the estimated value of the mode damping ratio of the structure, which can be assumed to be 0.1–0.2 depending on the structural characteristics.

By utilizing the vector  $z$ , stable and spurious modes can be distinguished. A significant coefficient  $z_j$  in  $z$  indicates that  $R_r$  modes possess comparable frequencies, damping ratios, and mode shapes to the  $j_{th}$  mode, thus suggesting its stability. Conversely, a smaller  $z_j$  value indicates sporadic recognition of the mode, rendering it unsuitable for classification as stable. To identify a set of stable modes rather than noisy modes, it is essential to establish a minimum number of similar modes,  $z_{min}$ . The selection of  $z_{min}$  should not be excessively large, as it may inadvertently eliminate low-frequency stable modes, albeit enhancing the reliability of false mode detection. Conversely, opting for a smaller  $z_{min}$  value effectively removes most noise modes without compromising low-frequency modes. In this study, the value of  $z_{min}$  is established as the smallest integer greater than or equal to 0.3 times  $R$ .

To determine the desired number of stable modes, a new binary vector  $s = [s_1, s_2, \dots, s_{R_r}]$  is derived from the vector  $z$ . Here,  $s_j$  takes a value of either 0 or 1, for all  $j = 1, 2, \dots, R_r$ . If  $z_j$  is greater than or equal to  $z_{min}$ ,  $s_j$  is set to 1, indicating the  $j_{th}$  mode as stable and thus

retained. Conversely, if  $s_j$  is assigned a value of 0, it signifies that the  $j_{th}$  mode is considered noise and subsequently discarded.

Calculating the non-zero terms in  $s$  yields  $P$  stable modes. At this point, vector  $f$ ,  $\zeta$ , and matrix  $\Phi$  become:

$$f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_P \end{bmatrix}, \zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_P \end{bmatrix}, \Phi = [\phi_1 \quad \phi_2 \quad \cdots \quad \phi_P] \quad (7)$$

Clearly, many components of the obtained vectors  $f$  and  $\zeta$ , as well as the corresponding columns in matrix  $\Phi$ , are highly like each other, representing information about the same-order modes. Therefore, the number of clusters for determining stable modes can be identified through clustering, thereby achieving automatic determination of the system order.

### 2.1.3 Determining system order using DBSCAN

The DBSCAN clustering algorithm [34, 35] is employed to determine the system order. This algorithm operates on the premise that each point within a cluster must have a minimum number of neighboring points (MinPts) within a specified radius  $e$ , and the density within this neighborhood should surpass a certain threshold. The process involves iteratively examining whether the mode quantity associated with any of the  $P$  modes corresponds to the representative of a specific cluster. The algorithm terminates when all  $P$  modes are assigned to clusters or rejected as outliers.

In mode clustering, a set of integers  $\{a_1, a_2, \dots, a_g\}$  represents the positions of cluster elements within the vectors  $f$ ,  $\zeta$ , and matrix  $\Phi$ , where the dimension of the cluster is denoted by  $g$ . Thus, for the  $i_{th}$  cluster, denoted as  $C_i$  with a dimension of  $g_i$ , a set of integers  $\{a_1, a_2, \dots, a_{g_i}\}$  is employed for representation. The modal information within cluster  $C_i$  are expressed as  $C_{if} = \{f_{a_1} f_{a_2} \dots f_{a_{g_i}}\}$ ,  $C_{i\zeta} = \{\zeta_{a_1} \zeta_{a_2} \dots \zeta_{a_{g_i}}\}$ ,  $C_{i\phi} = \{\phi_{a_1}, \phi_{a_2}, \dots, \phi_{a_{g_i}}\}$ , respectively. The mean values of these sets are  $f_{C_i}$ ,  $\zeta_{C_i}$ , and  $\phi_{C_i}$ , representing the centroid coordinates of cluster  $C_i$ .

To do clustering, first define a vector  $p = [1, 2, \dots, P]^T$  and  $v = [v_1, v_2, \dots, v_p]^T$ , the elements of  $v$  are a random permutation of the elements in  $p$ . At the

beginning, each element of the vector  $v$  is marked as “not accessed”.

In the  $q_{th}$  step, assuming that  $l$  clusters have already been identified, each cluster contains a different number of modal parameters. The  $q_{th}$  element  $v_q$  in  $v$  is used to compute the distance  $d(v_q)$  between the  $q_{th}$  modal parameter and the rest of the modal parameters. The distance calculation is performed using the Euclidean distance measure. Subsequently, a new modal set  $S_q$  is created, containing  $K_q$  modes that all within a radius distance  $e$  of the modal corresponding to  $v_q$ . Based on the minimum number of points MinPts, a new cluster  $C_{l+1}$  can be initialized.

$$K_q \geq \text{MinPts} \Rightarrow C_{l+1} = \{v_q\} \quad (8)$$

Alternatively, this point can be classified as an outlier or noise mode and subsequently assigned to cluster  $C_0$ , which is specifically designated as the outlier cluster:

$$K_q < \text{MinPts} \Rightarrow C_0 = C_0 \cup \{v_q\} \quad (9)$$

If the mode corresponding to  $v_q$  is identified as an outlier or noise mode, the next “not accessed” element in vector  $v$  is considered. In case the new cluster  $C_{l+1}$  has been initialized, the set  $S_q$  consisting of all points within the vicinity of  $e$  is examined. Each point is then reassigned either to the new  $C_{l+1}$  cluster or to a predefined cluster  $C_i$ .

The iterative process continues until all elements of vector  $v$  are marked as “accessed,” resulting in the completion of the clustering. Consequently, every point in vector  $p$  will be assigned to a mode class or an outlier class.

The DBSCAN algorithm needs to select two parameters: the neighborhood distance  $e$ , and the minimum number of points MinPts to form a dense area. In general, for datasets with high noise levels and large datasets, larger MinPts values are usually better to form more significant clusters; For the threshold distance  $e$ , if the value selected is too small, a large part of the data cannot be correctly clustered. In the follow-up study of this paper, MinPts was set as  $\text{MinPts} = \ln(P)$  and  $e$  was set as 0.01.

## 2.2 Iterative optimization of state-space matrices

Automatic determination of system order relies on estimating the system matrix  $A$  in the state-space

model. Likewise, NSI for nonlinear parameter identification is grounded in state-space model estimation. Nevertheless, NSI lacks nonlinear optimization methods, resulting in its numerical procedures (e.g., projection, QR decomposition, SVD decomposition) being non-iterative and thus yielding suboptimal solutions for the state-space model. To further enhance the identification performance of the proposed method, an iterative optimization method based on response prediction error minimization is integrated into the framework.

The predicted value of the output  $y$  of the system can be determined iteratively by Eq. (1).

$$\hat{y}_k = \mathbf{CA}^k \mathbf{x}_0 + \mathbf{CA}^{k-1} \mathbf{Bu}_0 + \dots + \mathbf{CA} \mathbf{Bu}_{k-2} + \mathbf{CBu}_{k-1} + \mathbf{Du}_k \quad (10)$$

Then the error between the measured output and the predicted output can be expressed as:

$$\mathbf{e}_k = \mathbf{y}_k - \hat{\mathbf{y}}_k \quad (11)$$

The error matrix can be further determined:

$$\Delta \mathbf{e} = [\mathbf{e}_k \quad \mathbf{e}_{k+1} \quad \dots \quad \mathbf{e}_{k+sh}] \quad (12)$$

where  $sh$  is the number of selected output samples.

Different models have different error criterion functions. Here the function is set as:

$$\mathbf{J} = h \left( \frac{1}{sh} \det(\Delta \mathbf{e}^T \Delta \mathbf{e}) \right) \quad (13)$$

where  $\det(\cdot)$  represents the determinant of the computed matrix  $(\cdot)$  and  $h(\cdot)$  represents a scalar monotonically increasing function.

Optimal estimates for the state-space matrices  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{D}$  are obtained by solving a minimization problem

$$\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\} = \arg \min_{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}} (\mathbf{J}), \quad (14)$$

several established methods can be used for this purpose, such as the steepest descent method, the Gauss–Newton method, or the Levenberg–Marquardt algorithm. In this study, the Levenberg–Marquardt algorithm was chosen for its ability to quickly converge to an optimal solution.

After obtaining the optimal matrix estimation, the parameter identification is completed according to Eq. (15).

$$\begin{aligned} \mathbf{H}_E(\omega) &= \mathbf{D} + \mathbf{C}(i\omega \mathbf{I} - \mathbf{A}_c)^{-1} \mathbf{B}_c \\ &= [\mathbf{H}(\omega) \quad \theta_1 \mathbf{L}_1 \mathbf{H}(\omega) \quad \dots \quad \theta_p \mathbf{L}_p \mathbf{H}(\omega)] \end{aligned} \quad (15)$$

### 2.3 Algorithm implementation process

Figure 1 summarizes the proposed method in this paper. Specifically, the main steps include:

1. Perform eigenvalue decomposition of the system matrix  $\mathbf{A}$  for different orders and store the complex eigenvalues and eigenvectors in vectors  $\lambda$  and  $\Phi$ , respectively.
2. Preprocess the data and eliminate incorrect eigenvalues and eigenvectors using Eq. (3);.
3. Calculate the modal parameters using the eigenvalues and eigenvectors, and compute the matrix  $\Delta$  based on the stability criterion in Eq. (4).
4. Determine the minimum number of similar modes  $z_{min}$ , create a binary vector  $s$ , and select which modes to retain to eliminate noise modes and keep only the stable modes.
5. Apply the DBSCAN algorithm to cluster the stable modes, with the number of clusters representing the mode system order.
6. Identify the state matrix  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{D}$  of the system after determining the system order.
7. Predict the measured response using the state matrix based on Eq. (10) and determine the error criterion function.
8. Solve the optimized  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{D}$  in Eq. (14) using a nonlinear optimization algorithm.
9. Identify the nonlinear parameters using the optimized matrix.

## 3 Simulation cases

### 3.1 Four degrees of freedom system

Consider a four-degree-of-freedom system with cubic stiffness nonlinearity. The schematic of this system is depicted in Fig. 2, while its pertinent physical parameters are delineated in Table 1. The nonlinear restoring force operative within the system can be expressed as:

$$f_{nl} = k_n x_2^3 \quad (16)$$

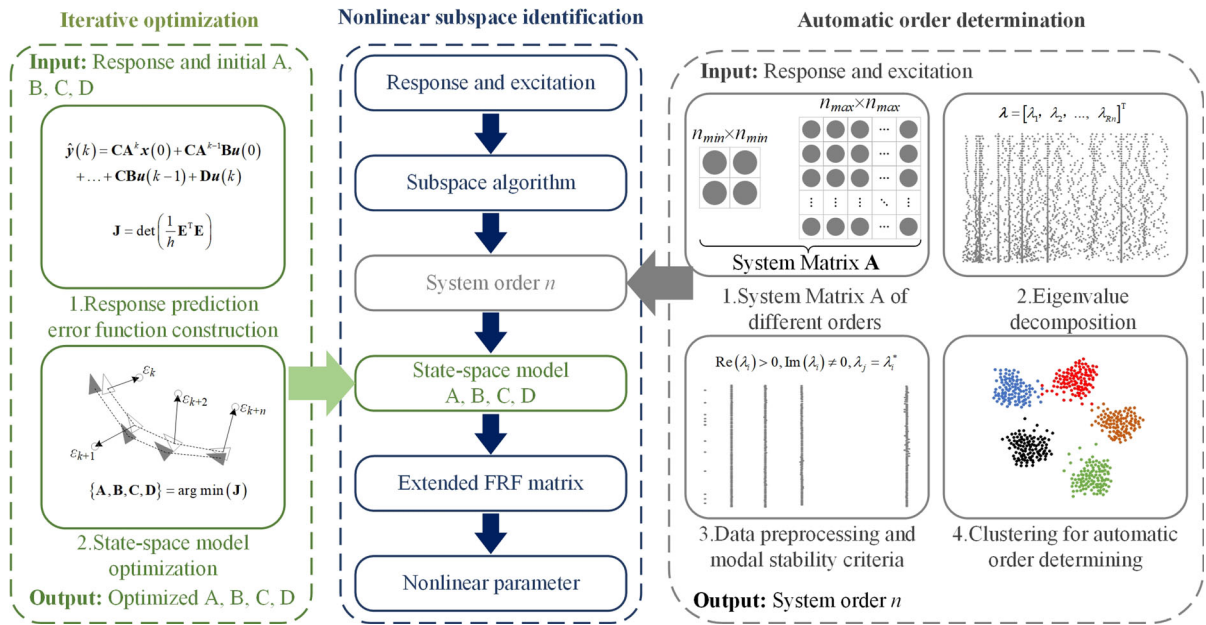


Fig. 1 Improved framework for nonlinear subspace identification

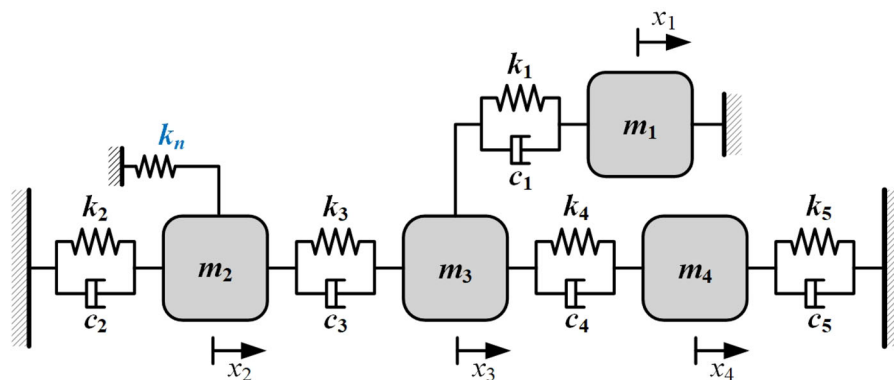


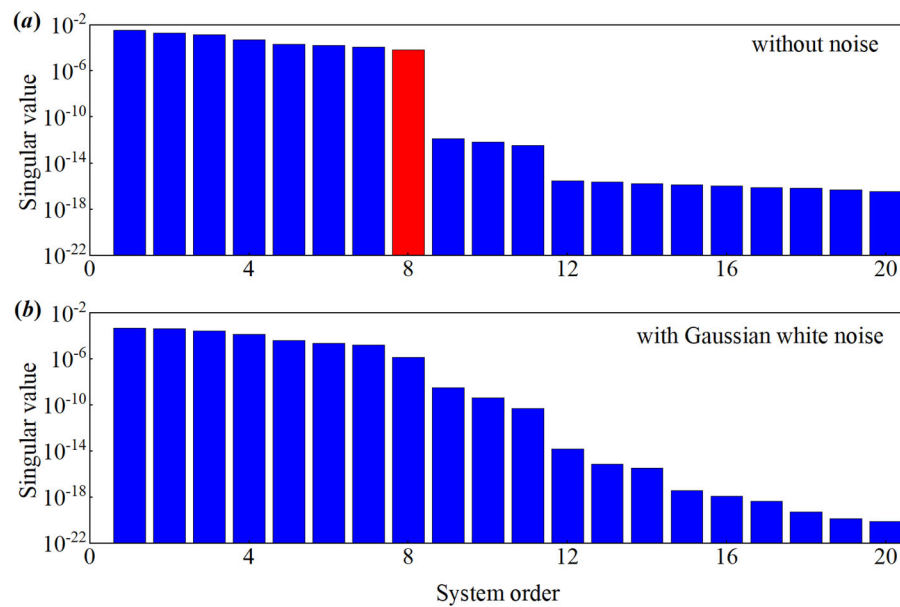
Fig. 2 Four-degree-of-freedom nonlinear systems

Table 1 Four-degree-of-freedom nonlinear system parameters

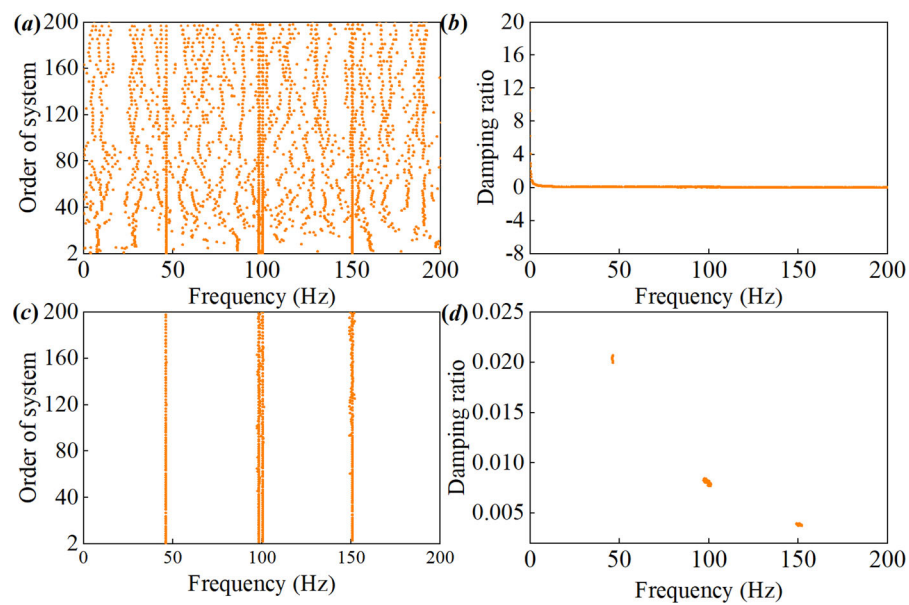
Mass (kg)	$m_1 = 0.55, m_2 = m_3 = m_4 = 1$
Stiffness (N/m)	$k_1 = k_2 = k_3 = k_4 = k_5 = 2 \times 10^5$
Viscous damping coefficient (Ns/m)	$c_1 = c_2 = c_3 = c_4 = c_5 = 5$
Nonlinear parameters	$k_n = 8 \times 10^{12} \text{ N/m}^3$

At  $t = 0$  s, a zero-mean Gaussian random force (r.m.s = 500 N) is applied to DOF 2 to simulate the forced response of the system at a sampling frequency  $f_s = 1000$  Hz and a total simulation time of 10 s. The

fourth-order Runge-Kutta time integration scheme was used to obtain the time response history of the system, the response data length was  $10^4$ , and a zero-mean Gaussian white noise with SNR = 20 dB was added to each analog output.



**Fig. 3** Singular values of the skew projection matrix



**Fig. 4** Modal parameter estimation results

The NSI method determines the system order through singular value decomposition. If the signal is noise-free, as shown in Fig. 3a, the singular value decomposition yields clear jumps in singular values at

the 8th order, allowing for the determination of the system order. However, in the presence of noise, depicted in Fig. 3b, no significant singular value



jumps are apparent, rendering the NSI method inadequate for system order determination.

Using the automated order determination method introduced in Sect. 2.1, the system order was initially varied from 2 to 200, and the eigenvalues and eigenvectors of the system matrix **A** were extracted respectively. Throughout this process, a total of 10,100 complex eigenvalues and eigenvectors were obtained. The system mode parameters estimated from all eigenvalues and eigenvectors are illustrated in Fig. 4a and b. These figures exhibit numerous false modes generated by noise, such as damping ratio values exceeding 100% at low frequencies in Fig. 4b.

Subsequently, the 10,100 modes were preprocessed and assessed for modal stability using the principles outlined in Eq. (3)–(6). Throughout this process, numerous false modes and erroneous estimations were eliminated, reducing the total number of stable modes to 525. The stable modal results are illustrated in Fig. 4c and d, where Fig. 4c displays four stable frequency axes, and Fig. 4d shows clusters of four modal damping ratios.

The DBSCAN clustering algorithm was employed to cluster the 525 frequency and damping data points obtained in the preceding step. The modal clustering results are illustrated in Fig. 5. During the grouping process facilitated by DBSCAN, a total of four clusters were identified within the stable modes, as shown in Fig. 5a. The distribution of modes in each cluster is depicted in Fig. 5b, and more fourth-order modes are identified. Particularly noteworthy is the algorithm's efficacy in distinguishing closely aligned second and third-order modes. Regarding the clustering results, the system order is determined to be  $4 \times 2 = 8$ .

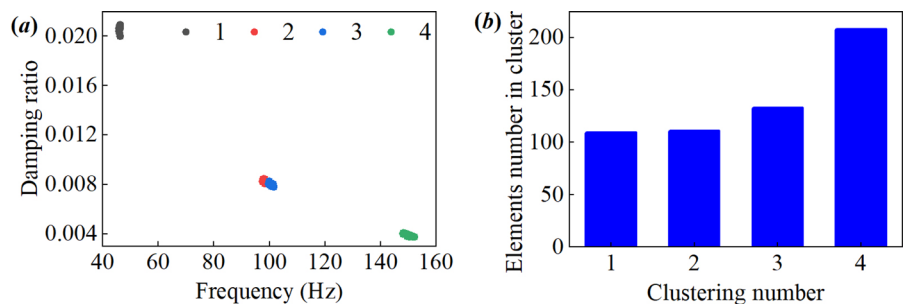
Once the system order is determined, the **A**, **B**, **C**, and **D** matrices in the state-space model can be readily established. The frequency response curve  $H_{22}$  is estimated directly using Eq. (15), as depicted by the

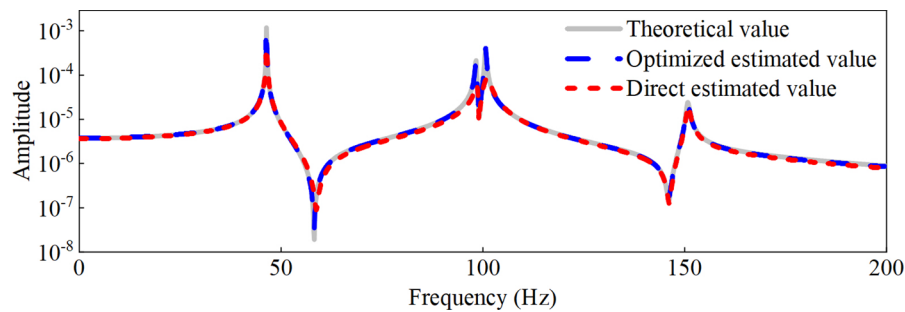
red dashed line in Fig. 6. The directly estimated results exhibit considerable deviations from the theoretical values, particularly at the peaks. Subsequently, by optimizing the **A**, **B**, **C**, and **D** matrices based on the principles outlined in Sect. 2.3, the optimized estimated frequency response curve is represented by the blue dashed line in Fig. 6. The optimized results demonstrate a higher level of consistency with the theoretical values.

Figure 7 presents the identification results of non-linear parameters in the form of histograms. Figure 7a displays the identification results directly obtained through NSI method, while Fig. 7b illustrates the identification results of proposed method. It can be observed that the parameter results obtained directly from identification exhibit greater dispersion, whereas the results from proposed method are more concentrated and closer to the theoretical values. The identification results, determined by taking the average, are shown in Table 2. Evidently, the proposed method significantly enhances the accuracy of parameter identification in noisy environments. In this case study, the error of the optimized estimation results is reduced by approximately 20%.

The simulation case was then replicated under different noise levels, and the outcomes of parameter identification errors are depicted in Fig. 8. From the figure, it can be observed that under low noise intensity conditions, the proposed method exhibits a slight improvement in identification accuracy. However, as the noise level increases, a significant discrepancy in accuracy between the NSI method and the proposed method emerges, with the gap widening as the noise level rises.

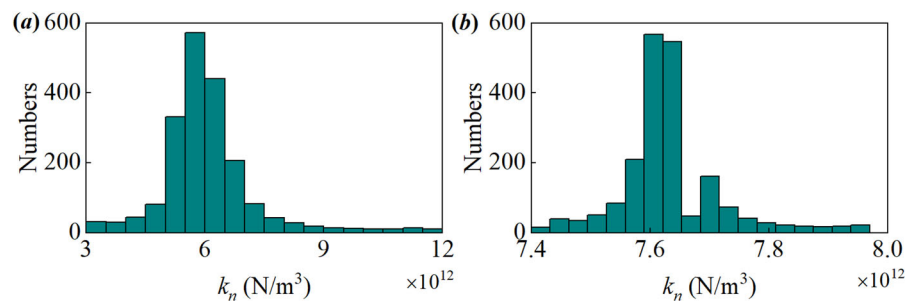
**Fig. 5** Modal clustering results





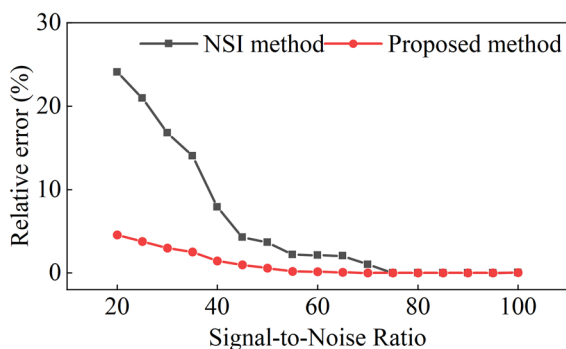
**Fig. 6** Frequency response curve  $H_{22}$

**Fig. 7** Histograms of nonlinear parameters identified



**Table 2** Nonlinear parameter identification values and errors

Nonlinear parameters	Theoretical value	NSI method	Error (%)	Proposed method	Error (%)
$k_n$	$8 \times 10^{12}$	$6.07 \times 10^{12}$	24.09	$7.63 \times 10^{12}$	4.56



**Fig. 8** Recognition results under different intensity noise

### 3.2 Complex multi-degree-of-freedom example

To further verify the recognition performance of the algorithm in complex structures, consider a complex multi-degree-of-freedom system as shown in Fig. 9 where nonlinearity consists of cubic stiffness between

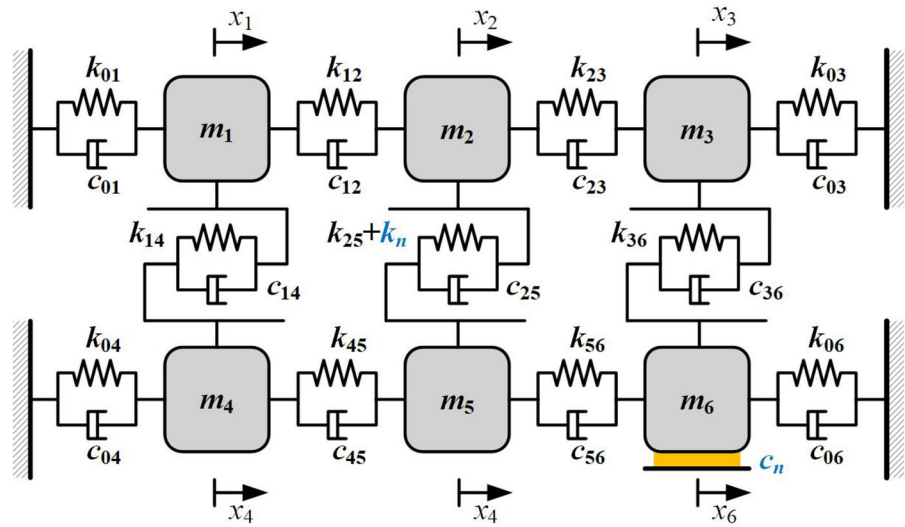
degrees of freedom 2 and 5, Coulomb friction at degrees of freedom 6, in the form of:

$$f_{nl1} = k_n(x_2 - x_5)^3, f_{nl2} = c_n \text{sgn}(\dot{x}_6) \quad (17)$$

Detailed values for the physical and nonlinear parameters of the system are shown in Table 3. At  $t = 0$  s, a zero-mean Gaussian random force (r.m.s = 30 N) is applied to DOF 1 to simulate the forced response of the system at a sampling frequency  $f_s = 1000$  Hz and a total simulation time of 10 s. The response is calculated using the fourth-order Runge–Kutta algorithm, adding zero-mean white Gaussian noise with SNR = 20 dB to each analog output.

The system order was increased from 2 to 220, and the modal parameters of the system were estimated by extracting the eigenvalues and eigenvectors of the system matrix  $\mathbf{A}$ , as shown in Fig. 10. Like the four-degree-of-freedom case, without data preprocessing and modal stability criterion, Fig. 10a and b exhibits

**Fig. 9** Complex multi-degree-of-freedom systems



**Table 3** Complex multi-degree-of-freedom systems parameters

Mass (kg)	$m_1 = m_3 = m_5 = 1$ $m_2 = m_4 = m_6 = 0.5$
Stiffness (N/m)	$k_{01} = k_{04} = 6 \times 10^5$ $k_{12} = k_{23} = 2 \times 10^5$ $k_{45} = 4000, k_{56} = 5 \times 10^5$ $k_{14} = k_{25} = k_{36} = 1 \times 10^5$ $k_{03} = k_{06} = 1 \times 10^5$
Viscous damping coefficient (Ns/m)	$c_{01} = c_{12} = c_{23} = c_{04} = c_{45} = c_{56} = 5$ $c_{14} = c_{25} = c_{36} = 1$ $c_{03} = c_{06} = 2$
Nonlinear parameters	$k_{n1} = 1 \times 10^{12} \text{ N/m}^3, c_n = 1 \text{ N/m}$

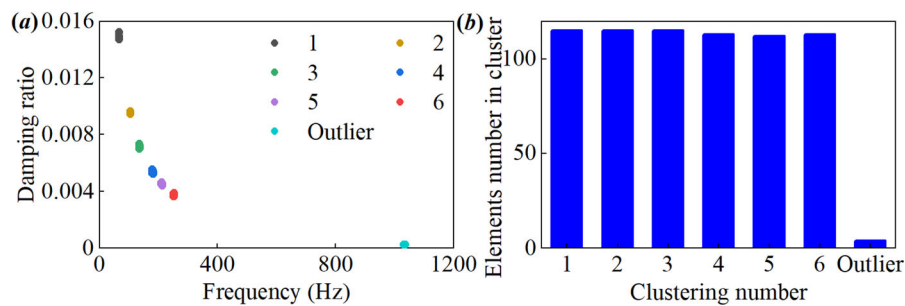
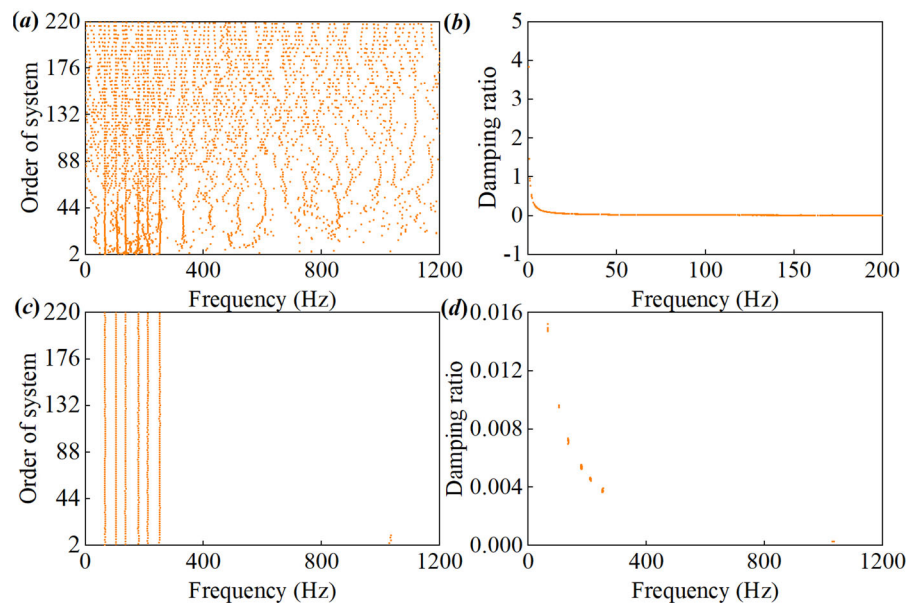
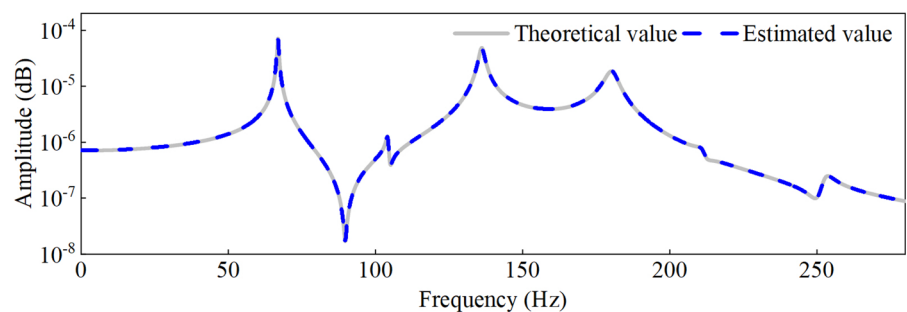
numerous spurious frequency axes and inaccurate damping ratios. The identification results in Fig. 10c and d show significant improvement, revealing six stable frequency axes and a more concentrated relationship between frequency and damping ratios. However, unlike the four-degree-of-freedom case, there are still some erroneous estimation points at high frequencies.

Figure 11 depicts the clustering outcomes of the DNSCAN algorithm for stable mode frequency and damping. Six normal clusters and one outlier cluster are discernible. The elements within the normal clusters exhibit a balanced distribution, whereas the outlier cluster comprises only a few elements, signifying their outlier status. By disregarding the outlier cluster, the system order can be established as  $6 \times 2 = 12$ .

Similar to Sect. 3.1, once the system order is determined, the **A**, **B**, **C**, and **D** matrices can be estimated within the state-space framework. By incorporating the estimated **A**, **B**, **C**, and **D** matrices into the optimization algorithm detailed in Sect. 2.3, the frequency response curve  $H_{21}$  is estimated, as depicted in Fig. 12. The estimated frequency response curve demonstrates a strong alignment with the theoretical values.

Figure 13 presents parameter identification results in the form of a statistical histogram. From the graph, it is evident that both the nonlinear stiffness parameters and damping parameters are concentrated near their theoretical values.

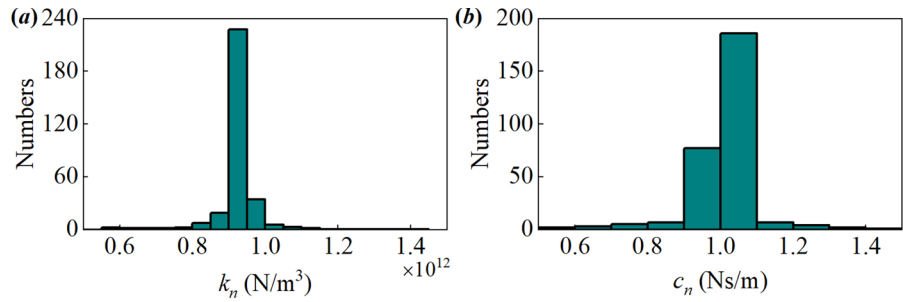
Table 4 summarizes the parameter identification results by averaging. The results indicate that despite the complexity of the nonlinear system, this method

**Fig. 10** Modal parameter estimation results**Fig. 11** Modal clustering results**Fig. 12** Frequency response curve  $H_{21}$ 

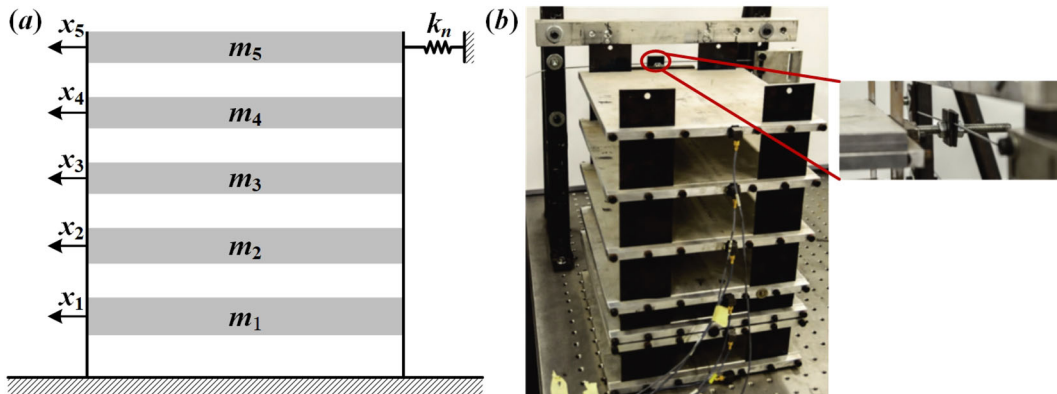
maintains a commendable level of identification accuracy.

#### 4 Experimental verification

An experimental investigation was undertaken to explore the practical application of the methodology proposed in this paper to real structures. The experimental dataset was generously provided by Prof. Stefano Marchesiello [36, 37] from the Department of

**Fig. 13** Histograms of nonlinear parameters identified

**Table 4** Nonlinear parameter identification results and errors

Nonlinear parameters	Theoretical value	identification results	Error (%)
$k_n$	$1 \times 10^{12}$	$9.29 \times 10^{11}$	7.06
$c_n$	1	1.06	6.11


**Fig. 14** multilayer nonlinear structure

**Table 5** Physical parameters of the test device

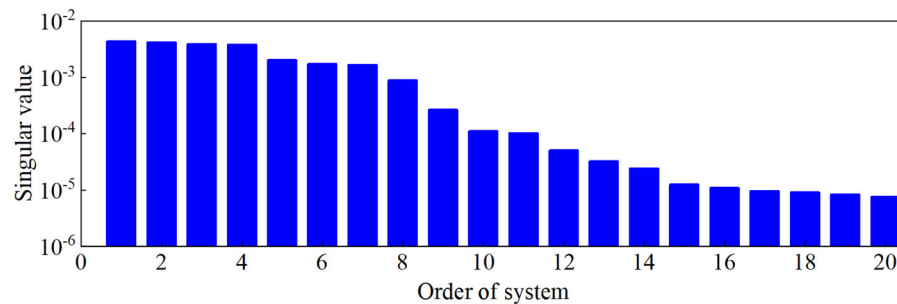
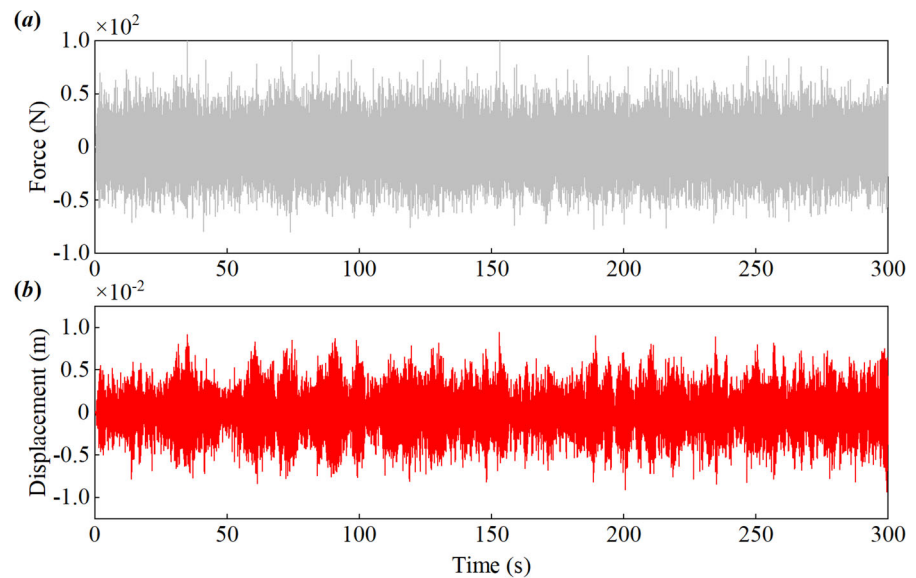
No	Laminate				Thin steel beams	
	Mass (kg)	Width (mm)	Length (mm)	Thickness (mm)	Length (mm)	Sectional area (mm <sup>2</sup> )
1	4.30	270	250	24	50	60 × 0.3
2	2.15	270	250	12	30	60 × 0.3
3	1.97	270	250	10	60	60 × 0.3
4	1.79	270	250	10	60	60 × 0.3
5	1.99	270	250	10	60	60 × 0.3

Mechanics and Aerospace at Politecnico di Torino, Italy.

The test object used in the study is a five-story structure consisting of five aluminum laminates

interconnected by thin steel beams. The schematic representation of the structure is shown in Fig. 14. By simplifying the frame structure as a 5-degree-of-freedom system, it can be assumed that the vertical

**Fig. 15** Excitation and displacement response for layer 2



**Fig. 16** Singular value distribution

thin steel beams contribute primarily to the bending stiffness. Table 5 provides the essential physical parameters of the structure.

The fifth layer of the structure is equipped with a thin tension metal wire that exhibits nonlinear stiffness when subjected to significant oscillations. The total restoring force exerted by the wire can be represented as the sum of a linear stiffness force  $k_l x_5$  and a cubic nonlinear stiffness force  $k_n x_5^3$ . This nonlinear term in the system can be expressed as:

$$f_{nl} = k_n x_5(t)^3 \quad (18)$$

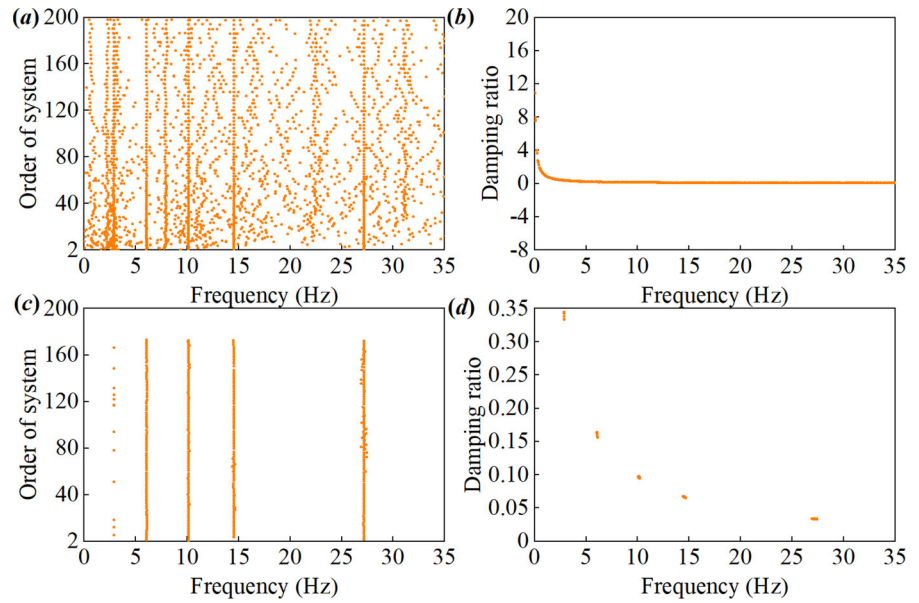
In the second layer of the structure, an electric shaker applies a random external force with an r.m.s. value of 20.89 N. Each layer is equipped with an acceleration sensor to obtain displacement signals by

integrating the acceleration response. Figure 15 illustrates the inputs and responses for the second layer.

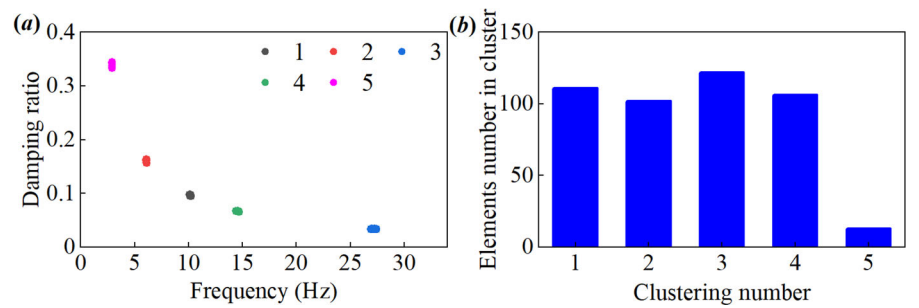
When applying the NSI method to all displacement and excitation signals, the singular value decomposition results in Fig. 16 are similar to the simulation cases. Noise components in the signals disrupt the decomposition, resulting in several singular value spikes. Consequently, determining the system order based on singular values becomes unfeasible.

System orders, ranging from 2 to 200 in even increments, were employed, and the NSI method was executed individually. Eigenvalue decomposition of the recognized system matrix  $\mathbf{A}$  generated a total of 10,100 eigenvalues and eigenvectors from 100 identifications. Subsequently, by employing data preprocessing and modal stability criteria to eliminate

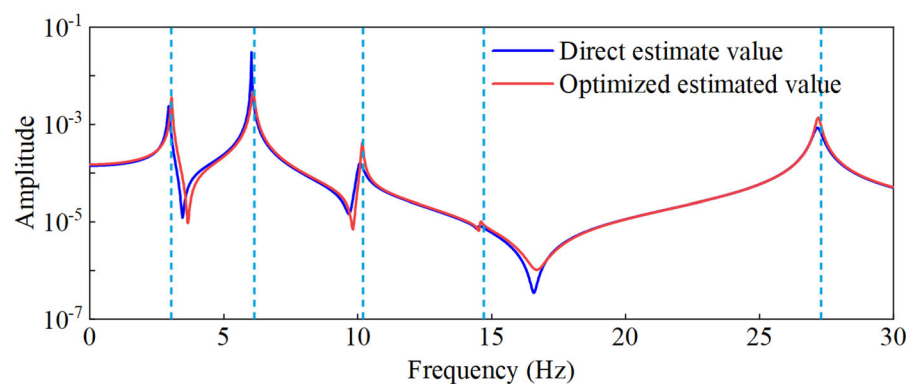
**Fig. 17** System mode identification results



**Fig. 18** Modal clustering results



**Fig. 19** Frequency response curves

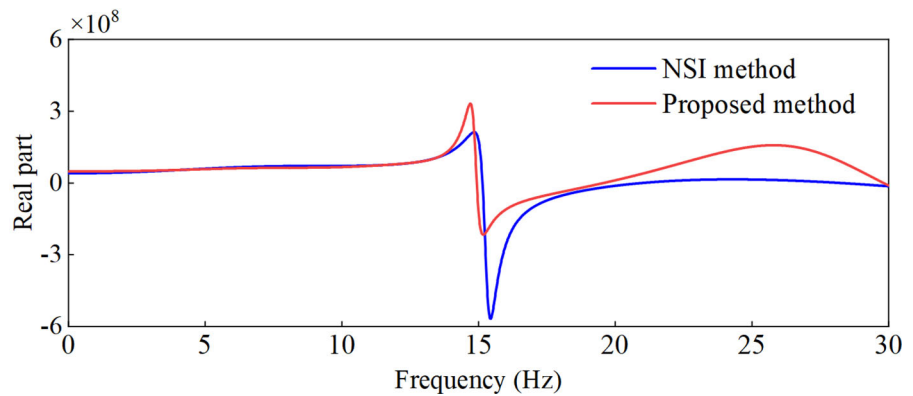


spurious modes and erroneous estimations, the modal identification results transition from Fig. 17a and b to Fig. 17c and d. In Fig. 17c and d, five frequency axes and five modal damping ratio points are evident. However, unlike the simulation cases, the first

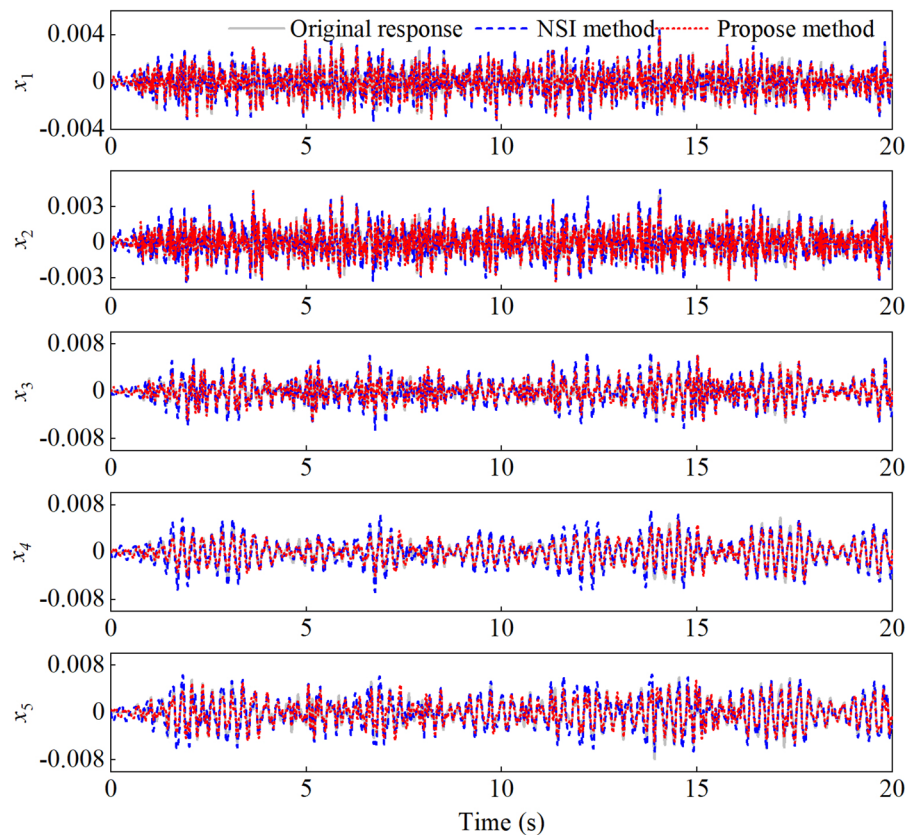
frequency axis appears sparse, indicating a potential occurrence of excessive denoising.

In Fig. 18, observe the frequency-damping clustering outcomes. These results reveal the presence of five distinct clusters within the stable modes, with the fifth

**Fig. 20** Real part of the identified nonlinear parameters



**Fig. 21** Reconstructed displacement response



**Table 6** R-squared coefficient

	$x_1$ (%)	$x_2$ (%)	$x_3$ (%)	$x_4$ (%)	$x_5$ (%)
NSI method	55.64	48.72	57.35	50.95	53.37
Proposed method	66.85	62.59	64.31	60.84	61.96

cluster corresponding to the recognition of the first-order mode. Notably, the elements in this cluster are notably lower, approximately 1/6 to 1/7 compared to

those in the other clusters, indicating relatively unstable recognition. Based on this clustering outcome, the automatic determination of the system order can be established as  $5 \times 2 = 10$ .

In reference [36], modal parameters of the structure were estimated using low-level excitation, with the first five modal frequencies depicted by the blue dashed lines in Fig. 19. It can be observed from Fig. 19 that the peaks of the frequency response curves



estimated by the proposed method are closer to the natural frequencies, indicating higher credibility compared to the directly estimated results.

The real part of the identified nonlinear parameters at each frequency is shown in Fig. 20. Taking the mean value of the real part in the range of 0 ~ 30 Hz as the final identification result, the value of  $k_n$  obtained from NSI method is  $3.80 \times 10^7$  N/m<sup>3</sup>, and the value of  $k_n$  obtained from proposed method is  $5.78 \times 10^7$  N/m<sup>3</sup>.

Since an accurate nonlinear restoring force model cannot be determined, the identification results are evaluated by reconstructing responses. The response is reconstructed using the identified state-space model, and the reconstructed response is shown in Fig. 21.

The fitting between the response signal and the measured signal is evaluated using the R-squared coefficient. A higher R-squared coefficient, closer to 1, indicates a better accuracy of response reconstruction and consequently a better identification result of the system. The results of the R-squared coefficient are presented in Table 6, where it's clear that for each response, the R-squared coefficients obtained from NSI method are lower than the proposed method. The lowest R-squared coefficient obtained from NSI method is 48.72%, and the highest is 57.35%, whereas all optimized R-squared coefficients obtained from proposed method exceed 60%.

## 5 Conclusions

This paper presents an improved framework for nonlinear subspace identification. The proposed method utilizes the DBSCAN algorithm for cluster analysis of stable modes, enabling automatic determination of system order and enhancing the estimation accuracy of the state-space model through an iterative optimization algorithm based on response prediction error minimization. These measures effectively improve the accuracy of nonlinear parameter identification. The effectiveness of this method is demonstrated through simulation studies involving two multi-degree-of-freedom systems and experimental validation on a multilayer building with nonlinear characteristics.

The main conclusions of this paper are as follows:

1. Through data preprocessing and modal stability criteria, a significant number of false modes and erroneous estimates induced by noise can be eliminated. Results from two numerical examples demonstrate that over 90% of the erroneous estimates are eliminated through this process.
2. In noisy environments, the NSI method of determining system order via singular value decomposition in subspace methods may fail. The DBSCAN clustering algorithm introduced in this paper effectively automates the determination of system order.
3. The response prediction error minimization method optimizes the identification results of system matrices. Compared to NSI methods, the proposed approach exhibits strong noise resistance. Simulation results indicate that under SNR = 20 dB, the parameter identification accuracy of the method surpasses NSI methods by over 20%. Experimental results also show that the method achieves approximately 10% higher fidelity in reconstructing responses compared to NSI methods.

Despite these promising results, there are several limitations to the proposed method that warrant further investigation:

1. The current methodology cannot be applied to systems with inherent time delays. Time-delay systems introduce additional complexity into the model, which requires specialized techniques to accurately capture the delay dynamics.
2. Another limitation of the proposed method is its inability to perform real-time identification. This issue arises from the construction of the Hankel matrix in subspace identification methods. Future work might overcome this limitation by integrating the proposed approach with real-time identification techniques such as recursive subspace methods.

**Acknowledgements** This work is supported by the National Natural Science Foundation of China (11602112, 52202436), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (20KJB460003), Qinglan Project of Jiangsu Province of China (2020)

**Authors contribution** Dong Jiang: Conceptualization, Methodology, Software, Validation, Investigation, Visualization, Funding acquisition, Writing—original draft.

Ang Li: Conceptualization, Methodology, Supervision, Writing. Yusheng Wang: Conceptualization, Methodology, Supervision, Writing. Shitao Xie: Conceptualization, Methodology, Validation, Formal analysis. Zhifu Cao: Conceptualization, Methodology, Validation, Formal analysis. Rui Zhu: Conceptualization, Methodology, Validation, Formal analysis.

**Funding** This work was funded by National Natural Science Foundation of China (11602112, 52202436), Natural Science Research of Jiangsu Higher Education Institutions of China (20KJB460003), Qinglan Project of Jiangsu Province of China.

**Data availability** Data will be made available on request.

**Declarations**

**Conflict of 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.

**Appendix**

The dynamic control equations for an  $N$ -degree-of-freedom mechanical structure with localized nonlinearity can be formulated as follows:

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}_v\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) + \mathbf{f}_{nl}(t) = \mathbf{f}(t) \tag{A.1}$$

where  $\mathbf{M}\ddot{\mathbf{q}}(t)$  represents the inertial force with  $\mathbf{M}$  denoting the mass matrix and  $\ddot{\mathbf{q}}(t)$  indicating the acceleration at time  $t$ .  $\mathbf{C}_v\dot{\mathbf{q}}(t)$  denotes the damping dissipative force, with  $\mathbf{C}_v$  as the damping matrix, and

$$\left\{ \begin{aligned} \mathbf{A}_c &= \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C}_v \end{bmatrix}, \mathbf{B}_c = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{M}^{-1} & \theta_1\mathbf{M}^{-1}\mathbf{L}_1 & \dots & \theta_p\mathbf{M}^{-1}\mathbf{L}_p \end{bmatrix} \\ \mathbf{C} &= [\mathbf{I}_{N \times N} \quad \mathbf{0}_{N \times N}], \mathbf{D} = [\mathbf{0}_{N \times N} \quad \mathbf{0}_{N \times 1} \quad \dots \quad \mathbf{0}_{N \times 1}] \end{aligned} \right. \tag{A.5}$$

$\dot{\mathbf{q}}(t)$  representing velocity.  $\mathbf{K}\mathbf{q}(t)$  signifies the elastic force, with  $\mathbf{K}$  as the stiffness matrix, and  $\mathbf{q}(t)$  denoting displacement. These forces, damping dissipative and elastic, are linear restorative components within the system.  $\mathbf{f}_{nl}(t)$  characterizes the non-linear restorative force within the system, which depends on displacement and velocity and can be expressed as the sum of  $p$  forces:

$$\mathbf{f}_{nl}(t) = \sum_{i=1}^p \theta_i \mathbf{L}_i g_i(t) \tag{A.2}$$

where  $g_i(t)$  signifies the functional form of the  $i_{th}$  nonlinear elastic or damping force,  $\mathbf{L}_i$  represents the force location, and  $\theta_i$  denotes the stiffness and damping coefficients of the  $i_{th}$  non-linear force, which are the parameters to be identified in this study.

Moving the nonlinear terms to the right-hand side results in:

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}_v\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) &= \mathbf{f}(t) - \sum_{i=1}^p \theta_i \mathbf{L}_i g_i(t) \\ &= \mathbf{f}(t) - \mathbf{f}_{nl}(t) \end{aligned} \tag{A.3}$$

Equation (A.3) considers nonlinearity as internal feedback forces applied to the underlying linear system. Consequently, the measured output of a non-linear system can be seen as the result of the combined action of external force  $\mathbf{f}(t)$  and internal feedback force  $\mathbf{f}_{nl}(t)$  on the underlying linear system.

The state-space representation of Eq. (A.3) for the defined state vector  $\mathbf{x}(t) = [\mathbf{q}(t) \quad \dot{\mathbf{q}}(t)]^T$ , input vector  $\mathbf{u}(t) = [\mathbf{f}(t) \quad -g_1(t) \quad \dots \quad -g_p(t)]^T$ , and output vector  $\mathbf{y}(t) = \mathbf{q}(t)$  can be expressed as follows:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}_c \mathbf{x}(t) + \mathbf{B}_c \mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C} \mathbf{x}(t) + \mathbf{D} \mathbf{u}(t) \end{aligned} \tag{A.4}$$

where  $\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}, \mathbf{D}$ , respectively, denote the system matrices and can be expressed as follows: where  $\mathbf{I}$  denotes the unit matrix.

It is possible to transform the continuous-time state-space model into a discrete-time model. Given a sampling interval  $\Delta t$ , the relationship between the continuous-time and discrete-time state-space models are as follows:

$$\mathbf{A} = e^{\mathbf{A}_c \Delta t}, \mathbf{B} = (e^{\mathbf{A}_c \Delta t} - \mathbf{I}) \mathbf{A}_c^{-1} \mathbf{B}_c \tag{A.6}$$

The Fourier transform of Eq. (A.4) yields the extended frequency response function matrix:

$$\mathbf{H}_E = \mathbf{D} + \mathbf{C}(i\omega\mathbf{I} - \mathbf{A}_c)^{-1}\mathbf{B}_c \quad (\text{A.7})$$

where  $i$  denotes the unit imaginary number,  $i = \sqrt{-1}$ .

Defining the Matrix:

$$\begin{aligned} \mathbf{N} &= \begin{bmatrix} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \mathbf{N}_{21} & \mathbf{N}_{22} \end{bmatrix} = i\omega\mathbf{I} - \mathbf{A}_c, \mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} \\ &= (i\omega\mathbf{I} - \mathbf{A}_c)^{-1} \end{aligned} \quad (\text{A.8})$$

Substituting Eq. (A.8) into Eq. (A.7) yields:

$$\mathbf{H}_E(\omega) = \mathbf{P}_{12}[\mathbf{M}^{-1} \quad \theta_1\mathbf{M}^{-1}\mathbf{L}_1 \quad \cdots \quad \theta_p\mathbf{M}^{-1}\mathbf{L}_p] \quad (\text{A.9})$$

This is obtained according to the chunked matrix inversion rule:

$$\begin{aligned} \mathbf{P}_{12} &= -\mathbf{N}_{11}^{-1}\mathbf{N}_{12}(\mathbf{N}_{22} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12})^{-1} \\ &= (\mathbf{K} + i\omega\mathbf{C}_v - \omega^2\mathbf{M})^{-1}\mathbf{M} = \mathbf{H}^{-1}(\omega)\mathbf{M} \end{aligned} \quad (\text{A.10})$$

Thus, Eq. (A.7) can finally be expressed as:

$$\mathbf{H}_E(\omega) = [\mathbf{H}(\omega) \quad \theta_1\mathbf{L}_1\mathbf{H}(\omega) \quad \cdots \quad \theta_p\mathbf{L}_p\mathbf{H}(\omega)] \quad (\text{A.11})$$

where  $\mathbf{H}(\omega)$  is the underlying linear system frequency response function.

From Eq. (A.11), it is evident that nonlinear parameters can be obtained by calculating the extended frequency response function  $\mathbf{H}_E(\omega)$ . Considering the presence of the imaginary unit  $i$ , identified nonlinear parameters are complex quantities associated with frequency. At each frequency  $\omega$ , the real part represents the parameter estimate, ideally having a zero imaginary component. However, the presence of noise and nonlinear modeling errors can introduce non-zero imaginary components.

## References

- Chen, G.Y., Gan, M., Chen, J., Chen, L.: Embedded point iteration based recursive algorithm for online identification of nonlinear regression models. *IEEE Trans. Autom. Control* **68**(7), 4257–4264 (2023). <https://doi.org/10.1109/tac.2022.3200950>
- Moghaddam, M.J.: Online system identification using fractional-order Hammerstein model with noise cancellation. *Nonlinear Dyn.* **111**(9), 7911–7940 (2023). <https://doi.org/10.1007/s11071-023-08249-5>
- Xavier, J., Patnaik, S.K., Panda, R.C.: Nonlinear system identification in coherence with nonlinearity measure for dynamic physical systems-case studies. *Nonlinear Dyn.* **112**(8), 6475–6501 (2024). <https://doi.org/10.1007/s11071-023-09258-0>
- Liu, Q.H., Hou, Z.H., Zhang, Y., Jing, X.J., Kerschen, G., Cao, J.Y.: Nonlinear restoring force identification of strongly nonlinear structures by displacement measurement. *J. Vibrot. Acoust. Trans. ASME* (2022). <https://doi.org/10.1115/1.4052334>
- Cui, N.Y., Liu, Y., Liang, H.Y., Bao, K.Y., Shan, Y., Gao, C.Y.: Improved frequency sweep modeling method based on model prediction output error for rub-impact rotor system. *Nonlinear Dyn.* (2024). <https://doi.org/10.1007/s11071-024-09463-5>
- Zuo, H., Guo, H.Y.: Structural nonlinear damage identification based on the information distance of GNPAX/GARCH model and its experimental study. *Struct. Health Monit.* **23**(2), 991–1012 (2024). <https://doi.org/10.1177/14759217231176958>
- Schoukens, J., Ljung, L.: Nonlinear system identification: a user-oriented road map. *IEEE Control. Syst. Mag.* **39**(6), 28–99 (2019). <https://doi.org/10.1109/mcs.2019.2938121>
- Xavier, J., Patnaik, S.K., Panda, R.C.: Process modeling, identification methods, and control schemes for nonlinear physical systems—a comprehensive review. *ChemBioEng Rev.* **8**(4), 392–412 (2021). <https://doi.org/10.1002/cben.202000017>
- Noël, J.P., Kerschen, G.: Nonlinear system identification in structural dynamics: 10 more years of progress. *Mech. Syst. Signal Proc.* **83**, 2–35 (2017). <https://doi.org/10.1016/j.ymssp.2016.07.020>
- Brunton, S.L., Proctor, J.L., Kutz, J.N.: Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U. S. A.* **113**(15), 3932–3937 (2016). <https://doi.org/10.1073/pnas.1517384113>
- Liu, Q.H., Zhang, Y., Hou, Z.H., Qiao, Y.T., Cao, J.Y., Lei, Y.G.: Optimal Hilbert transform parameter identification of bistable structures. *Nonlinear Dyn.* **111**(6), 5449–5468 (2023). <https://doi.org/10.1007/s11071-022-08120-z>
- Anastasio, D., Marchesiello, S., Gatti, G., Gonçalves, P.J.P., Shaw, A.D., Brennan, M.J.: An investigation into model extrapolation and stability in the system identification of a nonlinear structure. *Nonlinear Dyn.* **111**(19), 17653–17665 (2023). <https://doi.org/10.1007/s11071-023-08770-7>
- Marchesiello, S., Garibaldi, L.: A time domain approach for identifying nonlinear vibrating structures by subspace methods. *Mech. Syst. Signal Proc.* **22**(1), 81–101 (2008). <https://doi.org/10.1016/j.ymssp.2007.04.002>
- Noël, J.P., Kerschen, G.: Frequency-domain subspace identification for nonlinear mechanical systems. *Mech. Syst. Signal Proc.* **40**(2), 701–717 (2013). <https://doi.org/10.1016/j.ymssp.2013.06.034>
- Pang, Z.Y., Ma, Z.S., Ding, Q., Yang, T.Z.: An improved approach for frequency-domain nonlinear identification through feedback of the outputs by using separation

- strategy. *Nonlinear Dyn.* **105**(1), 457–474 (2021). <https://doi.org/10.1007/s11071-021-06595-w>
16. Zhu, R., Marchesiello, S., Anastasio, D., Jiang, D., Fei, Q.G.: Nonlinear system identification of a double-well Duffing oscillator with position-dependent friction. *Nonlinear Dyn.* **108**(4), 2993–3008 (2022). <https://doi.org/10.1007/s11071-022-07346-1>
  17. Zhu, R., Chen, S.F., Jiang, D., Xie, S.T., Ma, L., Marchesiello, S., Anastasio, D.: Enhancing nonlinear subspace identification using sparse Bayesian learning with spike and slab priors. *J. Vib. Eng. Technol.* (2023). <https://doi.org/10.1007/s42417-023-01030-3>
  18. Anastasio, D., Marchesiello, S.: Nonlinear frequency response curves estimation and stability analysis of randomly excited systems in the subspace framework. *Nonlinear Dyn.* **111**(9), 8115–8133 (2023). <https://doi.org/10.1007/s11071-023-08280-6>
  19. Zhu, R., Jiang, D., Hang, X.C., Zhang, D.H., Fei, Q.G.: Using novel nonlinear subspace identification to identify airfoil-store system with nonlinearity. *Aerosp. Sci. Technol.* **142**, 108647 (2023). <https://doi.org/10.1016/j.ast.2023.108647>
  20. Liu, Y., Zhou, P., Sun, X.Y., Chai, T.Y.: Optimal tracking control of blast furnace molten iron quality based on Krotov's method and nonlinear subspace identification. *IEEE Trans. Ind. Electron.* (2023). <https://doi.org/10.1109/tie.2023.3327555>
  21. Zhou, P.C., He, L., Yi, C., Zhou, Q.Y.: Impulses recovery technique based on high oscillation region detection and shifted rank-1 reconstruction-its application to bearing fault detection. *IEEE Sens. J.* **22**(8), 8084–8093 (2022). <https://doi.org/10.1109/jSEN.2022.3159116>
  22. Yu, M.Y., Zhang, Y., Yang, C.X.: Rolling bearing faults identification based on multiscale singular value. *Adv. Eng. Inf.* **57**, 102040 (2023). <https://doi.org/10.1016/j.aei.2023.102040>
  23. Hou, J., Chen, F., Li, P., Zhu, Z., Liu, F.: An improved consistent subspace identification method using parity space for state-space models. *Int. J. Control. Autom. Syst.* **17**(5), 1167–1176 (2019). <https://doi.org/10.1007/s12555-018-0499-6>
  24. Zhu, R., Jiang, D., Marchesiello, S., Anastasio, D., Zhang, D.H., Fei, Q.G.: Automatic nonlinear subspace identification using clustering judgment based on similarity filtering. *AIAA J.* **61**(6), 2666–2674 (2023). <https://doi.org/10.2514/1.J062816>
  25. Li, K., Luo, H., Yin, S., Kaynak, O.: A novel bias-eliminated subspace identification approach for closed-loop systems. *IEEE Trans. Ind. Electron.* **68**(6), 5197–5205 (2021). <https://doi.org/10.1109/tie.2020.2989717>
  26. He, Y.C., Li, Z., Fu, J.Y., Wu, J.R., Ng, C.T.: Enhancing the performance of stochastic subspace identification method via energy-oriented categorization of modal components. *Eng. Struct.* **233**, 111917 (2021). <https://doi.org/10.1016/j.engstruct.2021.111917>
  27. Feng, W.H., Wu, C.Y., Fu, J.Y., Ng, C.T., He, Y.C.: Automatic modal identification via eigensystem realization algorithm with improved stabilization diagram technique. *Eng. Struct.* **291**, 116449 (2023). <https://doi.org/10.1016/j.engstruct.2023.116449>
  28. Zho, K., Li, Q.S., Han, X.L.: Modal identification of civil structures via stochastic subspace algorithm with Monte Carlo-based stabilization diagram. *J. Struct. Eng.* (2022). [https://doi.org/10.1061/\(asce\)st.1943-541x.0003353](https://doi.org/10.1061/(asce)st.1943-541x.0003353)
  29. Jiang, D., Wang, Y., Hu, J., Qian, H., Zhu, R.: Automatic modal identification based on similarity filtering and fuzzy clustering. *J. Vib. Control* **30**(5–6), 1036–1047 (2023). <https://doi.org/10.1177/10775463231155714>
  30. Rainieri, C., Fabbrocino, G.: Development and validation of an automated operational modal analysis algorithm for vibration-based monitoring and tensile load estimation. *Mech. Syst. Signal Proc.* **60–61**, 512–534 (2015). <https://doi.org/10.1016/j.ymsp.2015.01.019>
  31. Bakir, P.G.: Automation of the stabilization diagrams for subspace based system identification. *Expert Syst. Appl.* **38**(12), 14390–14397 (2011). <https://doi.org/10.1016/j.eswa.2011.04.021>
  32. Zhang, X., Zhou, W., Huang, Y., Li, H.: Automatic identification of structural modal parameters based on density peaks clustering algorithm. *Struct. Control. Health Monit.* (2022). <https://doi.org/10.1002/stc.3138>
  33. Wei, S., Peng, Z.K., Dong, X.J., Zhang, W.M.: A nonlinear subspace-prediction error method for identification of nonlinear vibrating structures. *Nonlinear Dyn.* **91**(3), 1605–1617 (2018). <https://doi.org/10.1007/s11071-017-3967-2>
  34. Chen, Y.W., Zhou, L.D., Bouguila, N., Wang, C., Chen, Y., Du, J.X.: BLOCK-DBSCAN: Fast clustering for large scale data. *Pattern Recogn.* **109**, 107624 (2021). <https://doi.org/10.1016/j.patcog.2020.107624>
  35. Chen, Y.W., Zhou, L.D., Pei, S.W., Yu, Z.W., Chen, Y., Liu, X., Du, J.X., Xiong, N.X.: KNN-BLOCK DBSCAN: fast clustering for large-scale data. *IEEE Trans. Syst. Man Cybern. Syst.* **51**(6), 3939–3953 (2021). <https://doi.org/10.1109/tsmc.2019.2956527>
  36. Marchesiello, S., Fasana, A., Garibaldi, L.: Modal contributions and effects of spurious poles in nonlinear subspace identification. *Mech. Syst. Signal Proc.* **74**, 111–132 (2016). <https://doi.org/10.1016/j.ymsp.2015.05.008>
  37. Anastasio, D., Marchesiello, S.: Free-decay nonlinear system identification via mass-change scheme. *Shock Vib.* **2019**, 1–14 (2019). <https://doi.org/10.1155/2019/1759198>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.