

Research Article

Comparison of Different Independent Component Analysis Algorithms for Output-Only Modal Analysis

Jiaying Wang,¹ Cheng Wang,^{1,2} Tianshu Zhang,¹ and Bineng Zhong¹

¹College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China

²State Key Laboratory for Strength and Vibration of Mechanical Structures, Xi'an Jiaotong University, Xi'an 710049, China

Correspondence should be addressed to Cheng Wang; wangcheng@hqu.edu.cn

Received 30 May 2016; Revised 17 July 2016; Accepted 26 July 2016

Academic Editor: Nuno M. Maia

Copyright © 2016 Jiaying Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

From the principle of independent component analysis (ICA) and the uncertainty of amplitude, order, and number of source signals, this paper expounds the root reasons for modal energy uncertainty, identified order uncertainty, and modal missing in output-only modal analysis based on ICA methods. Aiming at the problem of lack of comparison and evaluation of different ICA algorithms for output-only modal analysis, this paper studies the different objective functions and optimization methods of ICA for output-only modal parameter identification. Simulation results on simply supported beam verify the effectiveness, robustness, and convergence rate of five different ICA algorithms for output-only modal parameters identification and show that maximization negentropy with quasi-Newton iterative of ICA method is more suitable for modal parameter identification.

1. Introduction

Identification of modal parameters including modal frequencies, modal shapes, and damping ratio is an important research for structure dynamic characteristics. Modal analysis is a kind of “inverse problem” analysis method in structural dynamics field, which is different from traditional methods including finite element analysis (FEA) method and theoretical modeling method. On the basis of actual dynamic test on engineering structure, modal analysis deals with the dynamic problems through combining the experiment with the theory.

In output-only modal analysis, modal parameters are identified only according to the output response data when the input excitation and system are unknown. Comparing with experimental modal analysis, output-only modal analysis has numerous positive features. It is effective to solve the modal identification problem when the input excitation of engineering structure is difficult to measure. What is more, there is no need to consume plenty of materials in the laboratory to simulate the engineering structure. In addition, because the response data are tested in the working condition, the identified modal parameters can well reflect the actual situation of engineering structure and boundary conditions.

Blind source separation (BSS) is a technique that recovers unknown independent source signals only from observed signals. At present, there are some BSS methods developed for output-only modal analysis [1]. Kerschen et al. [2] explained the one-to-one relationship between vibration modes of mechanical systems and modes computed through a BSS technique called independent component analysis (ICA) and then applied BSS technique to the free and random responses of mechanical systems [3, 4]. Zhou and Chelidze [5] proposed a method for linear normal mode identification based on BSS and verified its effectiveness in a vibration system. Antoni [6] advocated the use of Fourier-domain separation techniques based on the short-time Fourier transformation. In recent years, experts and scholars continued to study BSS of operation modal analysis [7, 8]. Abazarsa et al. [9] adopted a second-order blind identification (SOBI) technique to nonclassically damped systems. Antoni et al. [10] showed that several separation criteria purposely dedicated to operational modal analysis (OMA) can be deduced from general physical considerations. Sadhu et al. [11–13] took use of multirank parallel factor decomposition, time-series analysis-based BSS, and a decentralized BSS algorithm for ambient modal identification. Hazra et al. [14] applied hybrid time-frequency BSS to full-scale structure data for a variety

of conditions including the presence of supplemental damping devices. At the same time, it also can be applied in earthquake-excited structures [15], building vibration [16], and civil engineering structures [17]. There are some methods to deal with decentralized modal identification with limited sensors by taking use of sparse component analysis [18, 19]. When the mechanical system is complex, such as nonlinear system, a convolutive BSS towards modal identification was employed [20].

In recent years, output-only modal analysis based on BSS has been further developed [21, 22] and widely applied in practice. Modal identification issue was incorporated into the BSS formulation and transformed into a time-frequency framework [23]. Furthermore, independent component analysis (ICA) combined with wavelet transform has been used to identify damage for time-varying systems [24]. Therefore, the study of output-only modal analysis based on ICA is of great significance [7–13].

A generic approach to ICA consists of preprocessing the data, defining measures, and optimizing an objective function [25]. The classical measure method of maximization non-Gaussianity includes maximization, kurtosis, and negentropy. And the three common approaches in ICA algorithms are minimum mutual information, Infomax, and maximum likelihood estimation [25, 26]. Based on the theoretically described method as mentioned, a comparison of different ICA algorithms for output-only modal analysis is made. This paper mainly focuses on comparison and evaluation of different ICA algorithms for output-only modal analysis and then finds out which method is of the best accuracy and robustness.

The remainder of this paper is organized as follows. In Section 2, from the principle of ICA and the uncertainty of amplitude, order, and number of source signals, we expound the root causes for modal energy uncertainty, order uncertainty, and modal missing in output-only modal analysis based on ICA methods. Different objective functions and optimization methods of ICA algorithms for output-only modal analysis are compared and evaluated in Section 3. The simulation verifies effectiveness, robustness, and convergence rate of different ICA algorithms for output-only modal analysis and judges the most suitable method for output-only modal analysis in Section 4. At the end of this paper, the conclusions and further research directions of the study are summarized in Section 5.

2. ICA Based Output-Only Modal Analysis

2.1. Modal Decomposition of Dynamic Response and Output-Only Modal Analysis of Multi-Degree-of-Freedom System. Dynamic characteristic of the system is an intrinsic property, which reflects the linear relationship between the input and output of the system. Mechanical system is usually thought of as a proportional viscous damping system. Based on structural dynamic theory, for n -degree-of-freedom vibration system, the equation of motion is

$$\mathbf{M}\ddot{\mathbf{X}}(t) + \mathbf{C}\dot{\mathbf{X}}(t) + \mathbf{K}\mathbf{X}(t) = \mathbf{F}(t), \quad (1)$$

where $\mathbf{F}(t) \in \mathbb{R}^{m \times T}$ is excitation vector of additional load and $\ddot{\mathbf{X}}(t)$, $\dot{\mathbf{X}}(t)$, and $\mathbf{X}(t) \in \mathbb{R}^{n \times T}$ are on behalf of acceleration vector, velocity vector, and displacement vector, respectively. \mathbf{M} , \mathbf{C} , and $\mathbf{K} \in \mathbb{R}^{n \times n}$ are mass matrix, damping matrix, and stiffness matrix of the vibration system, respectively.

For general engineering of small damping, the displacement response can be expressed in modal coordinates as follows:

$$\mathbf{X}(t) = \mathbf{\Phi}\mathbf{Q}(t) = \sum_{i=1}^n \vec{\phi}_i \vec{q}_i(t), \quad (2)$$

where $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$ is modal matrix made up of order modal shape $\vec{\phi}_i \in \mathbb{R}^n$ and $\vec{\phi}_i$ and $\vec{\phi}_j$ are normalized and orthogonal mutually, while $\mathbf{Q}(t) \in \mathbb{R}^{n \times T}$ is the vector matrix composed of order modal response $\vec{q}_i(t) \in \mathbb{R}^T$, while $\vec{q}_i(t)$ and $\vec{q}_j(t)$ are mutually independent.

Classical modal parameter identification is to extract a series of mode shape ϕ_i , mode frequency ω_i , and modal damping ratio ξ_i from external excitation $\mathbf{F}(t)$ and response signal $\mathbf{X}(t)$ of the system. While in output-only modal analysis, modal parameters are identified only according to the output response signal $\mathbf{X}(t)$ when the input excitation $\mathbf{F}(t)$ and system are unknown.

2.2. ICA for BSS. After Jutten and Herault [26] published the paper on the signal processing in 1991, ICA became an effective solution to BSS. ICA is a method for finding underlying source signals from observed mixing signals. In order to estimate the source signals, ICA assumes that the source signals are statistically independent and the distributions of the components are non-Gaussian. Finally, to keep ICA model simpler, the linear system defined by the mixing matrix is invertible.

The linear model of ICA without measurement noise is

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t), \quad t = 1, 2, \dots, T, \quad (3)$$

where $\mathbf{X}(t) = [\vec{x}_1(t), \vec{x}_2(t), \dots, \vec{x}_m(t)]^T \in \mathbb{R}^{m \times T}$ is observation signal, $\mathbf{A} \in \mathbb{R}^{m \times n}$ is mixing matrix, and $\mathbf{S}(t) = [\vec{s}_1(t), \vec{s}_2(t), \dots, \vec{s}_n(t)]^T \in \mathbb{R}^{n \times T}$ represents source signal.

The goal of ICA algorithm is to isolate the source signals and the information about mixing matrix only from observation signals. In order to separate the statistically independent source signals from observation signals, it is necessary to find a separate matrix, shown in Figure 1. Its equation can be expressed as follows:

$$\mathbf{Y}(t) = \mathbf{W}\mathbf{X}(t), \quad t = 1, 2, \dots, T, \quad (4)$$

where $\mathbf{Y}(t) = [\vec{y}_1(t), \vec{y}_2(t), \dots, \vec{y}_n(t)]^T \in \mathbb{R}^{n \times T}$ is optimal estimation of source signals $\mathbf{S}(t)$.

It is impossible to determine specific values of mixing matrix \mathbf{A} and source signals $\mathbf{S}(t)$ simultaneously without any prior knowledge. But it is easy to find that there are some fuzzy factors and uncertainties from the model of ICA. It can be summarized as three uncertainties.

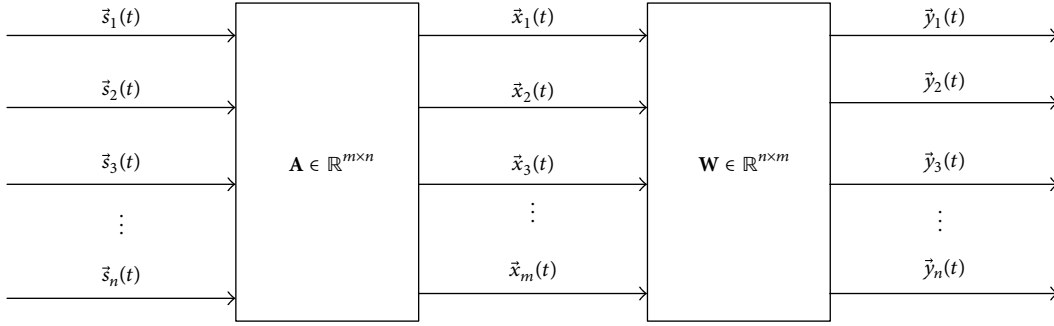


FIGURE 1: The essence of ICA method for BSS.

2.2.1. The Energy of ICA Is Uncertain. Suppose \mathbf{D} is a diagonal matrix; then, the linear instantaneous mixture modal can be expressed as

$$\begin{aligned} \mathbf{X}(t) &= [\mathbf{AD}^{-1}] \mathbf{DS}(t) \iff \\ \mathbf{X}(t) &= \mathbf{AS}(t). \end{aligned} \quad (5)$$

So, the amplitude of separated signals is inconsistent with the amplitude of source signals.

2.2.2. The Order of ICA Is Uncertain. Suppose \mathbf{P} is a permutation matrix; the linear instantaneous mixture modal can be expressed as

$$\begin{aligned} \mathbf{X}(t) &= [\mathbf{AP}^{-1}] \mathbf{PS}(t) \iff \\ \mathbf{X}(t) &= \mathbf{AS}(t), \end{aligned} \quad (6)$$

where $\mathbf{PS}(t)$ is a new source signals matrix after reordering and \mathbf{AP}^{-1} is a new mixing matrix.

2.2.3. The Number of Independent Components Is Uncertain. If the contribution of independent components is not enough, it is hard to identify and separate them. So, without any prior knowledge, determination of the number of independent components or sources by ICA algorithm is difficult.

2.3. ICA Based Output-Only Modal Analysis. Comparing (2) with (3), the dynamic response modal coordinates decomposition of multi-degree-of-freedom system under the effect of free vibration or random excitation can be regarded as a special case of ICA [2–4]. Therefore, as shown in Figure 2, there are a one-to-one relationship between vibration modal shapes of dynamic mechanical systems and separating matrix and a one-to-one relationship between modal responses and independent components.

On account of three assumptions of ICA method, the identified modal parameters are of the following characteristics.

- (1) Every order of modal shape is with different amplitude. In case of ICA method, the energy of separation matrix is not unique, and independent component also loses amplitude information. Unlike principal

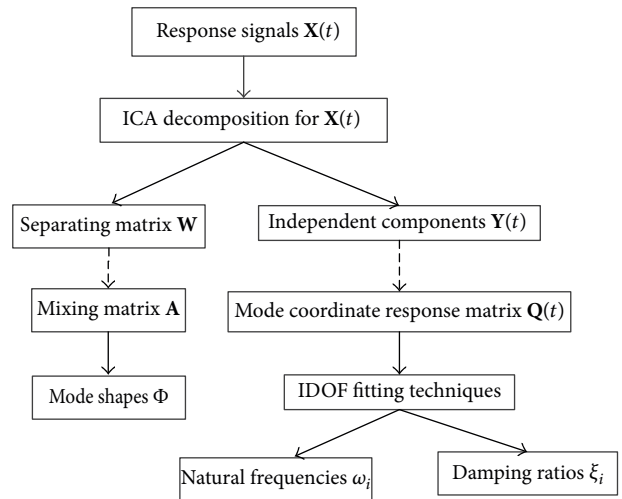


FIGURE 2: One-to-one corresponding relations between dynamic response modal coordinates decomposition and ICA.

component analysis method [27], ICA method cannot get the contribution ratio information of each modal. Modal shape is a relative quantity rather than an absolute value. So, in order to compare the modal shape with the real modal shape, the separation matrix and modal shape identified by ICA method should be normalized.

- (2) The order of modal parameter identified by ICA is uncertain. Modal identified by ICA method is not in accordance with the order from small to large. In fact, the first separated output source and vector are the ones whose independence is the strongest rather than the first-order modal parameter. Therefore, in order to compare the natural frequencies with the real natural frequencies, the modal parameters identified by ICA need to be reordered by the modal frequencies.
- (3) If the contribution of independent components is not enough, it is hard to identify and separate them and to determine the number of modal parameters. So, in ICA based output-only modal analysis, if the modal parameter is with small contribution to

independence, ICA may barely identify it, which would cause modal parameter missing.

3. Different Objective Functions and Optimization Methods of ICA for Output-Only Modal Analysis

3.1. Different Objective Functions of ICA Algorithm for Output-Only Modal Analysis. The objective functions of ICA algorithm mainly include the measurement of maximization non-Gaussianity (kurtosis and negentropy), minimization of mutual information, Infomax method, and maximum likelihood estimation.

3.1.1. Maximization Kurtosis. A non-Gaussian measure is kurtosis, and kurtosis represents fourth-order statistics of a random variable. A random variable $\tilde{y}(t)$ with zero-mean and kurtosis is defined as

$$\text{kurt}(\tilde{y}(t)) = E\{(\tilde{y}(t))^4\} - 3(E\{(\tilde{y}(t))^2\})^2. \quad (7)$$

If $\tilde{y}(t)$ is Gaussian random variables, $\text{kurt}(\tilde{y}(t)) = 0$. However, for a non-Gaussian variable, its kurtosis usually is not equal to zero.

When $\text{kurt}(\tilde{y}(t)) > 0$, $\tilde{y}(t)$ represents a super-Gaussian variable, and when $\text{kurt}(\tilde{y}(t)) < 0$, $\tilde{y}(t)$ represents a sub-Gaussian variable, and the greater the non-Gaussianity of $\tilde{y}(t)$, the greater the absolute value of kurtosis. Therefore, ICA problem can be converted into calculating the maximum of the absolute value or square of kurtosis of source signal component.

3.1.2. Maximization Negentropy. Another non-Gaussian measure is negentropy. According to information theory, Gaussian variables have the maximum entropy of all the random variables with the same variance.

The negentropy of random variable $\tilde{x}(t)$ is defined as

$$J(\tilde{x}(t)) = H(\tilde{x}_{\text{Gauss}}(t)) - H(\tilde{x}(t)), \quad (8)$$

where $\tilde{x}_{\text{Gauss}}(t)$ and $\tilde{x}(t)$ are Gaussian random variables with the same covariance matrix and $H(\tilde{x}(t)) = -\int p_{\tilde{x}(t)}(\eta) \log p_{\tilde{x}(t)}(\eta) d\eta$ is differential entropy. The negentropy value is always nonnegative and $J(\tilde{x}(t)) = 0$ if and only if $\tilde{x}(t)$ is the Gaussian distribution.

To overcome the disadvantage of dyscalculia of negentropy, an approximate representation with high-order cumulant can be expressed as

$$J(\tilde{x}(t)) \propto [E\{G(\tilde{x}(t))\} - E\{G(\tilde{y}(t))\}], \quad (9)$$

where $\tilde{y}(t)$ is a Gaussian variable with zero-mean and unit variance and $\tilde{x}(t)$ is a variable with zero-mean and unit variance, while $G(\cdot)$ and $E\{\cdot\}$ are nonquadratic function and expectation, respectively.

3.1.3. Minimization of Mutual Information. For m random variables $\tilde{y}_i(t)$, $i = 1, 2, \dots, m$, their mutual information is defined as

$$\begin{aligned} I(\tilde{y}_1(t), \tilde{y}_2(t), \dots, \tilde{y}_m(t)) \\ = \sum_{i=1}^m H(\tilde{y}_i(t)) - H(\mathbf{Y}(t)) \end{aligned} \quad (10)$$

for invertible linear transformation $\mathbf{Y}(t) = \mathbf{W}\mathbf{X}(t)$, so

$$\begin{aligned} I(\tilde{y}_1(t), \tilde{y}_2(t), \dots, \tilde{y}_m(t)) \\ = \sum_{i=1}^m H(\tilde{y}_i(t)) - H(\mathbf{X}(t)) - \log |\det \mathbf{W}|. \end{aligned} \quad (11)$$

When $\tilde{y}_i(t)$ and $\tilde{y}_j(t)$ are not related to each other and the variance of $\tilde{y}_i(t)$ and $\tilde{y}_j(t)$ is a unit matrix, then $E\{\mathbf{Y}(t)\mathbf{Y}^T(t)\} = \mathbf{W}E\{\mathbf{X}(t)\mathbf{X}^T(t)\}\mathbf{W}^T = \mathbf{I}$.

So,

$$\begin{aligned} \det \mathbf{I} = 1 &= \det(\mathbf{W}E\{\mathbf{X}(t)\mathbf{X}^T(t)\}\mathbf{W}^T) \\ &= (\det \mathbf{W}) \det(E\{\mathbf{X}(t)\mathbf{X}^T(t)\}) (\det \mathbf{W}^T). \end{aligned} \quad (12)$$

From (12), $\det \mathbf{W}$ must be a constant, because $E\{\mathbf{X}(t)\mathbf{X}^T(t)\}$ is independent of \mathbf{W} , so

$$I(\tilde{y}_1(t), \tilde{y}_2(t), \dots, \tilde{y}_m(t)) = C - \sum_i J(\tilde{y}_i(t)). \quad (13)$$

According to (13), it suggests the basic relationship of maximization negentropy and mutual information. So the ICA problem of maximizing non-Gaussian transforms into the problem of minimizing mutual information.

3.1.4. Maximum Likelihood Estimation. Maximum likelihood estimation can be interpreted as follows: when likelihood function reaches the maximum, its parameter corresponds to the separated matrix. The goal is to find the inverse matrix of mixing matrix; namely, $\mathbf{W} = \mathbf{A}^{-1}$. The likelihood function is defined as density function of variable density, with mixed matrix \mathbf{A} as parameter.

Assume that $\hat{p}_x(\mathbf{X}(t))$ is estimation of probability density function $p_x(\mathbf{X}(t))$ of observed signals $\mathbf{X}(t)$. The probability density function of source signals $p_s(\mathbf{S}(t))$ is known, according to the relationship of two probability density functions; therefore, $\hat{p}_x(\mathbf{X}(t))$ satisfies the following equation:

$$\hat{p}_x(\mathbf{X}(t)) = \frac{p_s(\mathbf{A}^{-1}\mathbf{X}(t))}{|\det \mathbf{A}|}. \quad (14)$$

For the given model, the likelihood function of observed signals $\mathbf{X}(t)$ is function of parameter \mathbf{A} and can be defined as

$$\begin{aligned} L(\mathbf{A}) &= E\{\log_2 \hat{p}_x(\mathbf{X}(t))\} \\ &= \int p_x(\mathbf{X}(t)) \log_2 p_s(\mathbf{A}^{-1}\mathbf{X}(t)) dx \\ &\quad - \log_2 |\det \mathbf{A}|. \end{aligned} \quad (15)$$

TABLE 1: Comparison of five objective functions.

Evaluation factor	Maximization kurtosis	Maximization negentropy	Minimization of mutual information	Maximum likelihood estimation	Infomax
Estimate the component at once	One	One	All	All	All
Estimated components must be irrelevant	Yes	Yes	No	Yes	Yes
Prior knowledge	No	No	No	Yes	No
Robustness to outliers	Low	High	Middle	Middle	Middle

When the parameter becomes separation matrix $\mathbf{W} = \mathbf{A}^{-1}$, the log-likelihood function is

$$L(\mathbf{W}) \approx \frac{1}{n} \sum_{i=1}^m \{\log_2 p_s(\mathbf{W}\mathbf{X}(t))\} + \log_2 |\det \mathbf{W}|, \quad (16)$$

where m represents the number of source signals and n represents the sample number of observed signals with independent and identical distribution.

3.1.5. Infomax. If we consider the flow of information in a neural network, in order to realize the effective information transmission, it is necessary to maximize the mutual information between the input and output.

Supposing that $\mathbf{X}(t)$ is the input of a neural network, then the output can be expressed as

$$\tilde{y}_i(t) = \phi_i(\tilde{w}_i^T \mathbf{X}(t)) + \tilde{n}(t), \quad (17)$$

where $\phi_i(\cdot)$ is a nonlinear scalar function, \tilde{w}_i represents weight vector of neural network, and $\tilde{n}(t)$ is additive measurement noise.

So, the output entropy is

$$H(\tilde{y}_i(t)) = H(\phi_1(\tilde{w}_1^T \mathbf{X}(t)), \dots, \phi_m(\tilde{w}_m^T \mathbf{X}(t))). \quad (18)$$

In the absence of measurement noise, maximization mutual information between input and output equals maximization output entropy.

3.1.6. Comparison of Different Objective Functions of ICA for Output-Only Modal Analysis. Five objective functions have been introduced as mentioned above, and all the objective functions are interrelated. In fact, all the estimation theories are the same, and the only difference is the concrete forms. So there are some differences between the objective functions [28]:

- Some objective functions such as maximization kurtosis and negentropy can estimate a component at once, but others estimate all the components at once.
- Some objective functions such as maximization negentropy adopt the nonpolynomial function, but others adopt polynomial function which is in connection with cumulant.
- Among the objective functions, the constraint of estimating independent component is irrelevant. However, when mutual information is the minimum, the

independent components may not be irrelevant from each other.

- In the objective function of maximum likelihood estimation, the density of independent component is certain with prior knowledge, and it does not meet the model of ICA. However, maximization negentropy adopts approximation algorithm, without need of any prior knowledge.
- An important evaluation index is the robustness to outliers. It means that single and wrong observed values will not influence estimating independent components.

Above all, the advantages and disadvantages are summed up in Table 1.

3.2. Different Optimization Methods of ICA for Output-Only Modal Analysis. Gradient descent [29], stochastic gradient descent [30], and quasi-Newton method [31] are three commonly used optimization methods.

3.2.1. Gradient Descent. Gradient descent is an optimization algorithm, the essence of which is to make the change of steps and finally attain the minimal error. We minimize or maximize the objective function $J(\tilde{w}) = E\{g(\tilde{w}, \tilde{x}(t))\}$ iteratively by starting from an initial point $\tilde{w}(0)$, then compute the gradient of $J(\tilde{w})$ at this point, and move in the direction of the negative gradient or with the steepest descent from a suitable distance. After that, we repeat the same procedure at the new point and so on. Finally, we get the update rule as follows:

$$\tilde{w}(k) = \tilde{w} - \alpha(k) \frac{\partial E\{g(\tilde{w}, \tilde{x}(t))\}}{\partial \tilde{w}} \Big|_{\tilde{w}=\tilde{w}(k-1)}, \quad (19)$$

where $\partial E\{g(\tilde{w}, \tilde{x}(t))\}/\partial \tilde{w}$ represents the gradient of $J(\tilde{w})$ and $\alpha(k)$ is step of moving in the direction of the negative gradient.

3.2.2. Stochastic Gradient Descent. Stochastic gradient descent is a method combining random procedure with optimization method. Its basic idea is not to calculate the exact value of gradient directly but unbiasedly estimate the gradient as the direction of the negative gradient.

If the objective function of ICA is $J(\tilde{w}) = E\{g(\tilde{w}, \tilde{x}(t))\}$, we would like to minimize or maximize it. Firstly, we also need to initialize point $\tilde{w}(0)$ and compute the gradient. Unlike gradient descent, what stochastic gradient descent method

TABLE 2: Comparison of three optimization methods.

Evaluation factor	Gradient descent	Stochastic gradient descent	Quasi-Newton
Computational complexity	Middle	Simple	Complex
Need step parameter	Yes	Yes	No
Iterative times	Middle	High	Low
Convergence rate	Middle	Slow	Fast
Iterative efficiency	High	Middle	Low

computes is $g(\bar{w}, \bar{x}(t))$ rather than $E\{g(\bar{w}, \bar{x}(t))\}$. And then we move in the direction of the negative gradient or the steepest descent by a suitable distance and repeat the same procedure until we find the optimum point. Therefore, the update rule is as follows:

$$\bar{w}(k) = \bar{w} - \alpha(k) \left. \frac{\partial g(\bar{w}, \bar{x}(t))}{\partial \bar{w}} \right|_{\bar{w}=\bar{w}(k-1)}. \quad (20)$$

3.2.3. Quasi-Newton. In the optimal aspect, Newton's method is with fast convergence speed and high precision while it has the shortcoming of expensive computing, because it needs to calculate the second derivative and inverse of Hessian matrix. In order to keep its merits and to overcome the shortcoming, quasi-Newton method appeared. Quasi-Newton method calculates the second derivative and inverse of a quadratic function instead of Hessian matrix.

For the objective function of ICA is $J(\bar{w})$, if we want to minimize or maximize it, its update rule is

$$\bar{w}(k) = \bar{w} - H^{-1} \left. \frac{\partial J(\bar{w})}{\partial \bar{w}} \right|_{\bar{w}=\bar{w}(k-1)}, \quad (21)$$

where $H_{ij} = \partial^2 J(\bar{w}) / \partial \bar{w}_i \partial \bar{w}_j$ is a Hessian matrix.

However, it is quite difficult to calculate the inverse of Hessian matrix. And quasi-Newton method takes the measure by replacing the Hessian matrix with a quadratic function, so the update rule becomes

$$\bar{w}(k) = \bar{w} - G(J(\bar{w})) \left. \frac{\partial J(\bar{w})}{\partial \bar{w}} \right|_{\bar{w}=\bar{w}(k-1)}, \quad (22)$$

where $G(\cdot)$ represents a quadratic function.

3.2.4. Comparison of Different Optimization Methods of ICA for Output-Only Modal Analysis. Gradient descent, stochastic gradient descent, and quasi-Newton are optimization algorithms, but each method is with respective advantages and disadvantages. Gradient descent is relatively easy but it is with low convergence rate and easy to be trapped in local minima. Stochastic gradient descent is simpler than gradient descent, but with the shortcoming of reaching the convergence precision more hardly. Quasi-Newton method is with fast convergence speed and high precision but with the more complex computation than the methods mentioned above. Table 2 shows comparison of three optimization methods with different evaluation factors.

4. Simulation Verification and Results Analysis

After weighing the advantages and disadvantages of different objective functions and optimization methods, we think the following five methods are more suitable to output-only modal parameters identification. The methods include maximization of kurtosis with gradient descent method, maximization negentropy with quasi-Newton method, minimum mutual information with gradient descent method, maximum likelihood estimation with stochastic gradient descent method, and Infomax with gradient descent method.

4.1. Establishment of Simulation Model and Parameters Setting of Data

- (1) Parameters setting of simulation model is a simply supported beam without damping, the length of which is 1 meter, elasticity modulus is 205 Gpa, Poisson's ratio of material is 0.3, and density of material is 7850.
- (2) This beam is divided into 1000 equal pieces and the number of measured stations is 1001.
- (3) Response data are obtained by the resonance method excited by sine signals. And the sine signals are composed of the frequencies that are 205 Hz, 91.3 Hz, 366 Hz, 572 Hz, 824 Hz, 1121 Hz, and 22 Hz with powers that are 60, 30, 30, 30, 30, 30, and 30, respectively. The loads are concentrated resistor loadings in the distance of 0.2 meters of a simply supported beam.
- (4) Sampling frequency is 4096 Hz and sampling number is 4096.
- (5) The real modal shapes and natural frequencies are all calculated by FEA.

4.2. Simulation Parameter Setting and Modal Assurance Criterion. In maximization kurtosis and negentropy, non-quadratic function is $G(\bar{y}(t)) = \tanh(\bar{y}(t))$ and its first differential is $G'(\bar{y}(t)) = 1 - \tanh^2(\bar{y}(t))$. In minimization of mutual information, maximum likelihood estimation, and Infomax, when independent components meet sub-Gauss distribution, nonquadratic function is $G(\bar{y}(t)) = -2 \tanh(\bar{y}(t))$; otherwise, nonquadratic function is $G(\bar{y}(t)) = \tanh(\bar{y}(t)) - \bar{y}(t)$. For all the ICA algorithms, the convergence criteria are $\|\bar{w}(k) - \bar{w}(k-1)\|_2 < 10^{-10}$. Response data is divided into two groups: data without measurement noise and data with 1% Gauss measurement noise.

To reflect the accuracy of modal shapes in experiment, the modal assurance criterion (MAC) is introduced into [5]. And the formula is as follows:

$$K_{\text{MAC}} = \frac{(\bar{\phi}_i^T \bar{\phi}_i)^2}{(\bar{\phi}_i^T \bar{\phi}_i)(\bar{\phi}_i^T \bar{\phi}_i)}, \quad (23)$$

where K_{MAC} varies between 0 (no coincidence) and 1 (complete coincidence).

TABLE 3: Error of natural frequencies identified by maximization kurtosis with gradient descent method.

Modal order	Real natural frequency (Hz)	Frequency identified (Hz)	Percentage of relative error
1	22.89	23	0.481%
2	91.55	92	0.492%
3	205.99	206	0.005%
4	366.23	367	0.210%
6	824	825	0.121%
7	1121.6	1122	0.036%

TABLE 4: The convergence times and MAC value of modal shapes.

Identified order	Number of convergences	Modal order of FFE	K_{MAC}
1	13	4	1.0000
2	18	6	1.0000
3	9	2	1.0000
4	7	3	1.0000
5	8	1	1.0000
6	2	7	0.9998

4.3. Results of Simulation Verification without Measurement Noise and Results Analysis. With the purpose of verifying the effect of five methods mentioned above for modal parameter identification and judging which method is more suitable for operation modal analysis, the results of simulation verification without measurement noise are as follows.

4.3.1. Maximization Kurtosis with Gradient Descent Method. According to the process of modal parameters identified by maximization kurtosis with gradient descent method, Figure 3 shows fast Fourier transform (FFT) of each independent component.

The abscissa of main peak in each subfigure of Figure 3 corresponds to each natural frequency identified by maximization kurtosis with gradient descent method. And Table 3 shows the errors of natural frequencies identified by maximization kurtosis with gradient descent method.

According to the identified frequency value, Figure 4 shows the modal shapes which compared the real modal shapes calculated by FEA.

Table 4 shows the convergence times and MAC value of modal shapes.

4.3.2. Maximization Negentropy with Quasi-Newton Method. According to the process of modal parameters identified by maximization negentropy with quasi-Newton method, Figure 5 shows FFT of each independent component.

The abscissa of main peak in each subfigure in Figure 5 is each natural frequency identified by maximization negentropy with quasi-Newton method. Table 5 shows error of natural frequencies identified by maximization negentropy with quasi-Newton method.

TABLE 5: Error of natural frequencies identified by maximization negentropy with quasi-Newton method.

Modal order	Real natural frequency (Hz)	Frequency identified (Hz)	Percentage of relative error
1	22.89	24	0.485%
2	91.55	92	0.492%
3	205.99	206	0.005%
4	366.23	367	0.210%
6	824	825	0.121%
7	1121.6	1123	0.125%

TABLE 6: The convergence times and MAC value of modal shapes.

Identified order	Number of convergences	Modal order of FEA	K_{MAC}
1	16	4	1.0000
2	7	6	0.9988
3	20	2	1.0000
4	18	3	1.0000
5	13	7	0.9999
6	2	1	0.9999

TABLE 7: Error of natural frequencies identified by minimum mutual information with gradient descent method.

Modal order	Real natural frequency (Hz)	Frequency identified (Hz)	Percentage of relative error
1	22.89	23	0.481%
2	91.55	92	0.492%
3	205.99	207	0.490%
4	366.23	367	0.210%
6	824	825	0.121%
7	1121.6	1122	0.036%

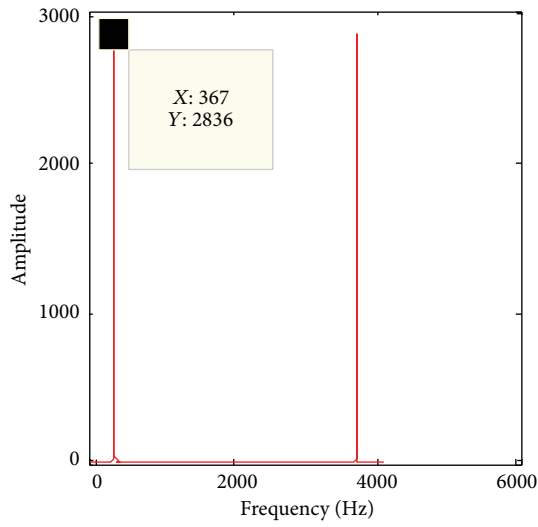
According to the identified frequency value, Figure 6 shows the modal shapes which compared the real modal shapes calculated by FEA.

Table 6 shows the convergence times and MAC value of modal shapes.

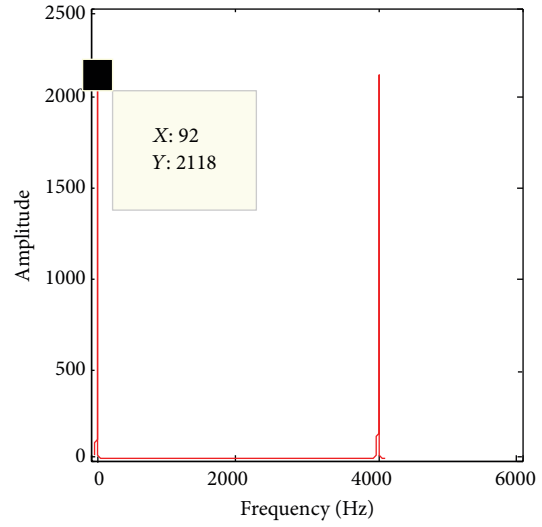
4.3.3. Minimum Mutual Information with Gradient Descent Method. According to the process of modal parameters identified by minimum mutual information with gradient descent method, Figure 7 shows FFT of each independent component.

The abscissa of main peak in each subfigure in Figure 7 is each natural frequency identified by minimum mutual information with gradient descent method. Table 7 shows error of natural frequencies identified by minimum mutual information with gradient descent method.

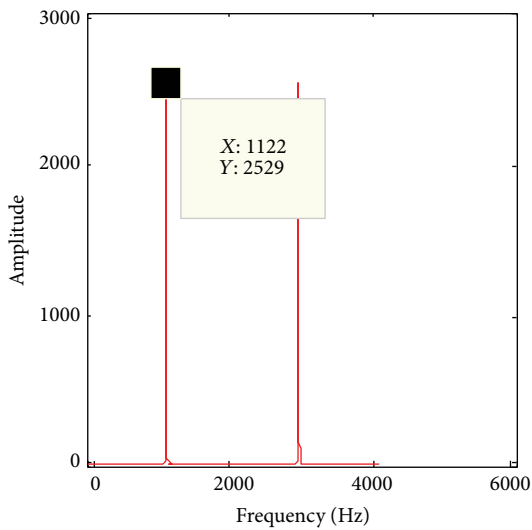
According to the identified frequency value, Figure 8 shows the modal shapes which compared the real modal shapes calculated by FEA.



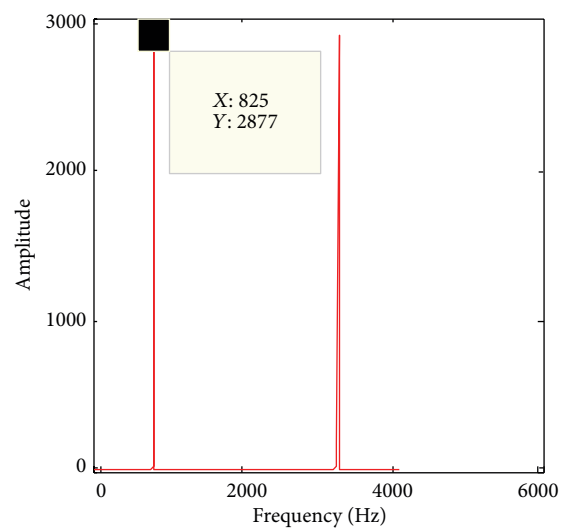
(a) FFT of the 1st independent component



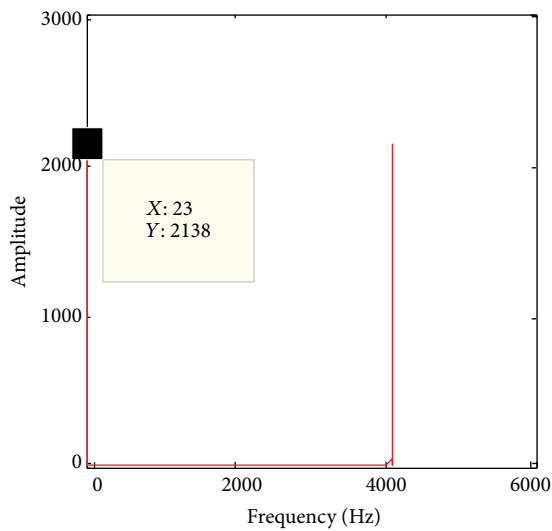
(b) FFT of the 2nd independent component



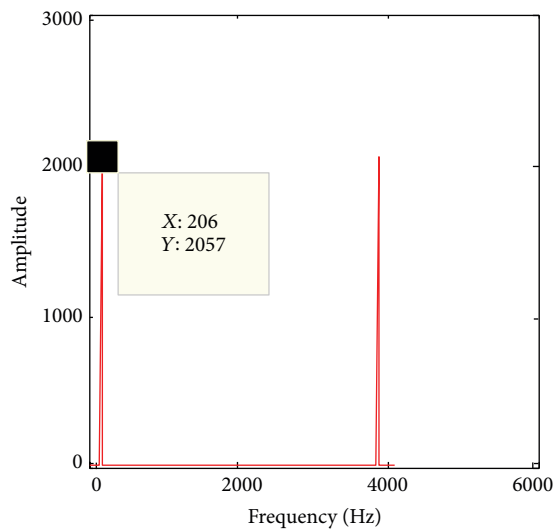
(c) FFT of the 3rd independent component



(d) FFT of the 4th independent component



(e) FFT of the 5th independent component



(f) FFT of the 6th independent component

FIGURE 3: FFT of each independent component identified by maximization kurtosis with gradient descent method.

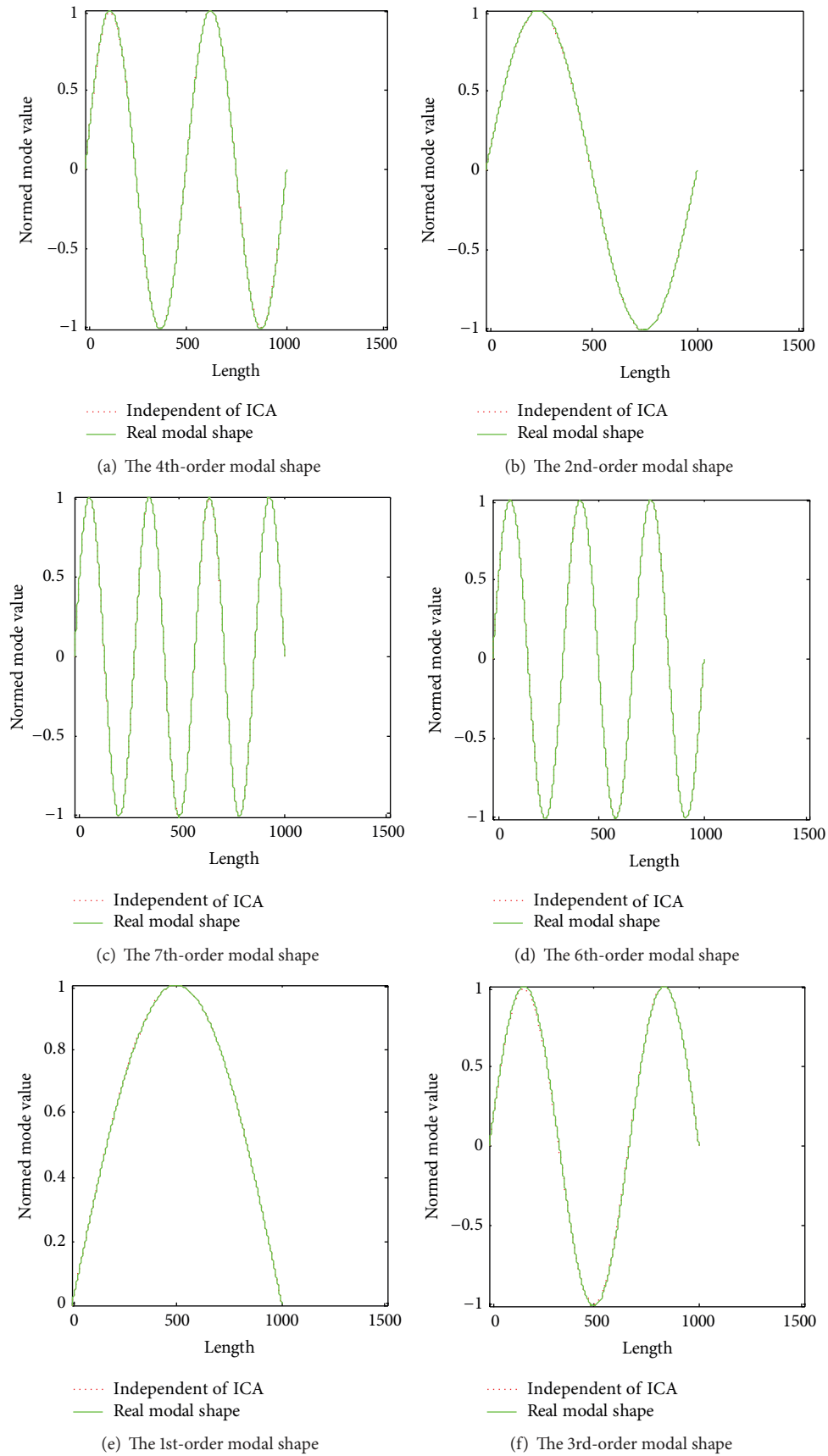


FIGURE 4: Modal shape comparison between FEA and separating vector by maximization kurtosis with gradient descent method.

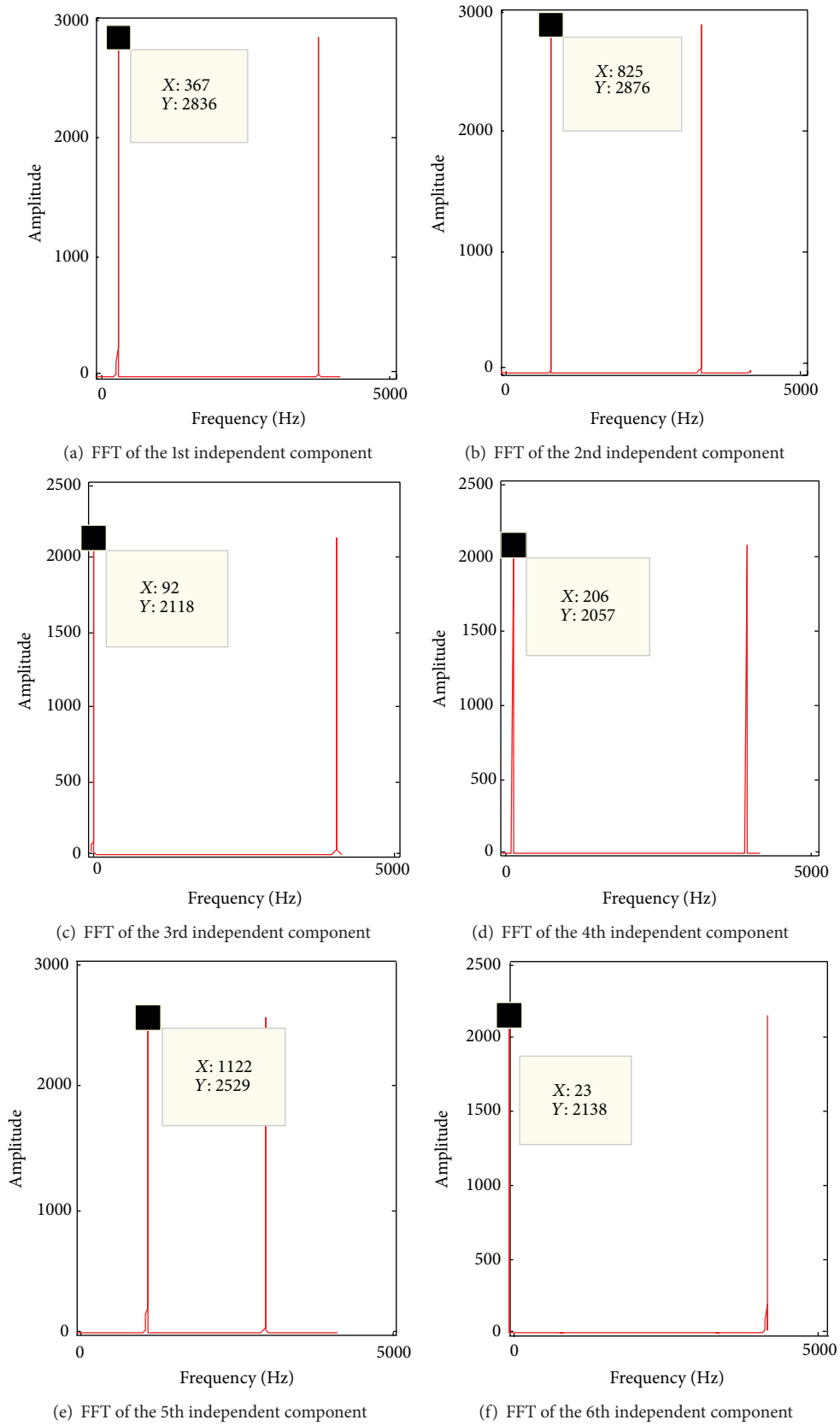


FIGURE 5: FFT of each independent component identified by maximization negentropy with quasi-Newton method.

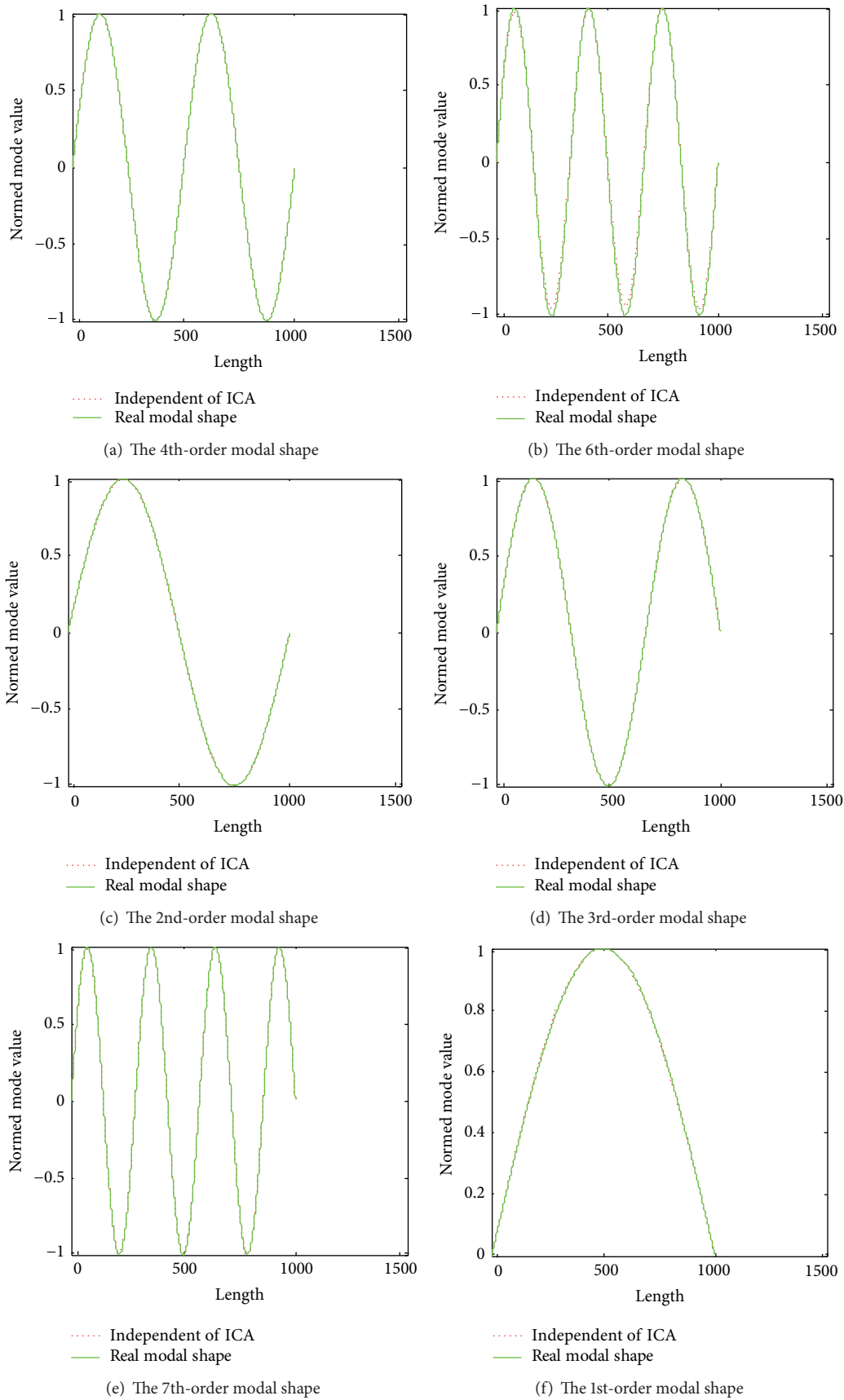
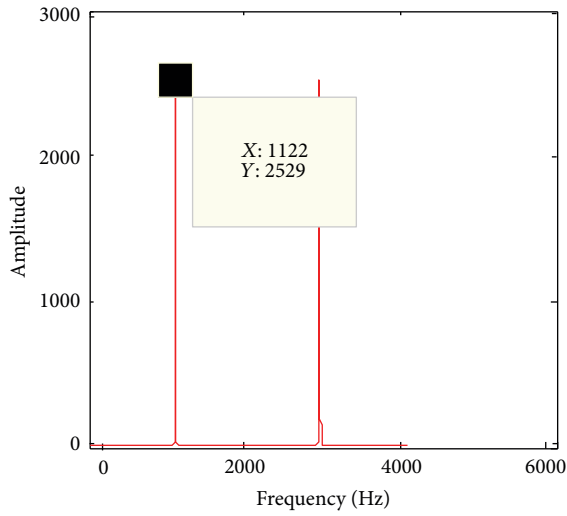
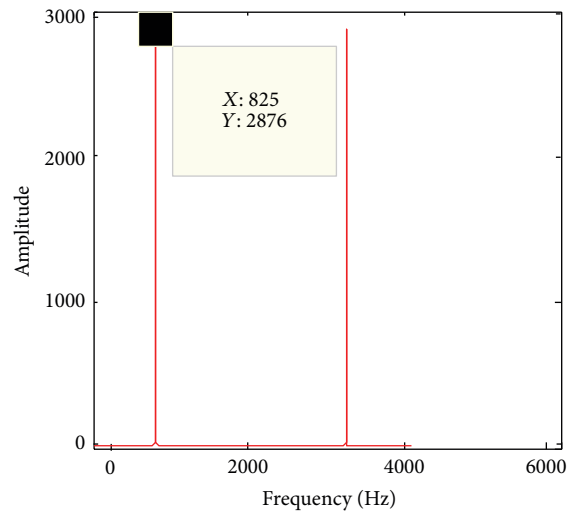


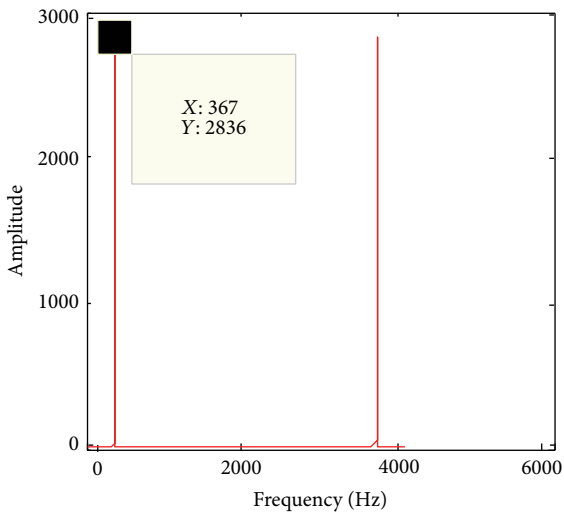
FIGURE 6: Modal shape comparison between FEA and separating vector by maximization negentropy with quasi-Newton.



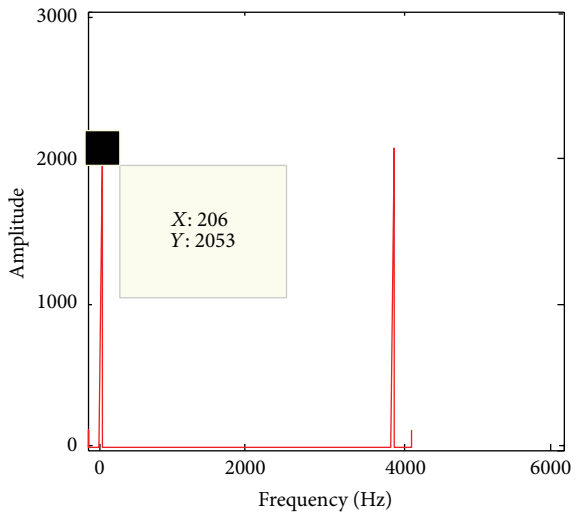
(a) FFT of the 1st independent component



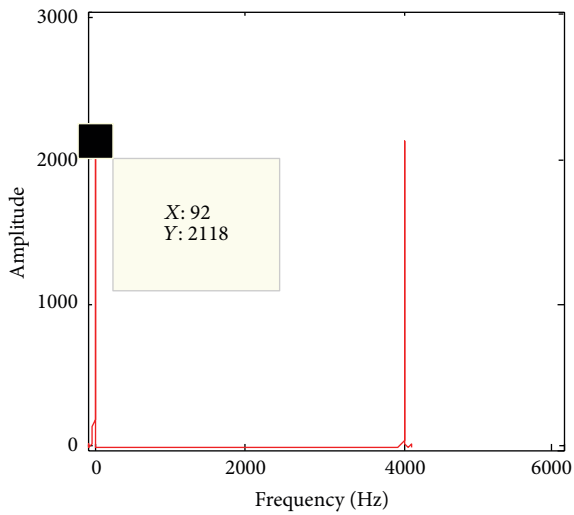
(b) FFT of the 2nd independent component



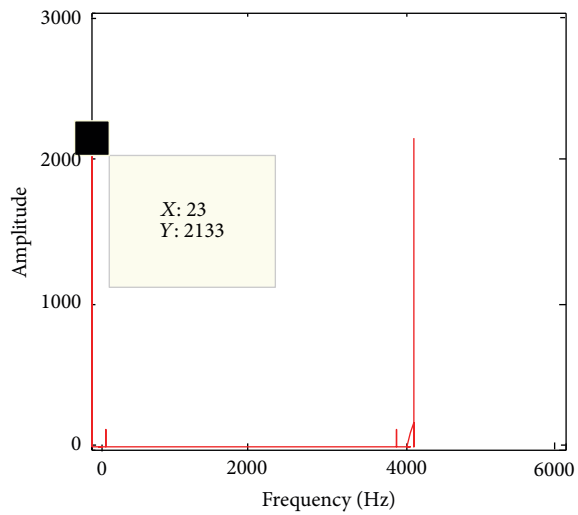
(c) FFT of the 3rd independent component



(d) FFT of the 4th independent component



(e) FFT of the 5th independent component



(f) FFT of the 6th independent component

FIGURE 7: FFT of each independent component identified by minimum mutual information with gradient descent method.

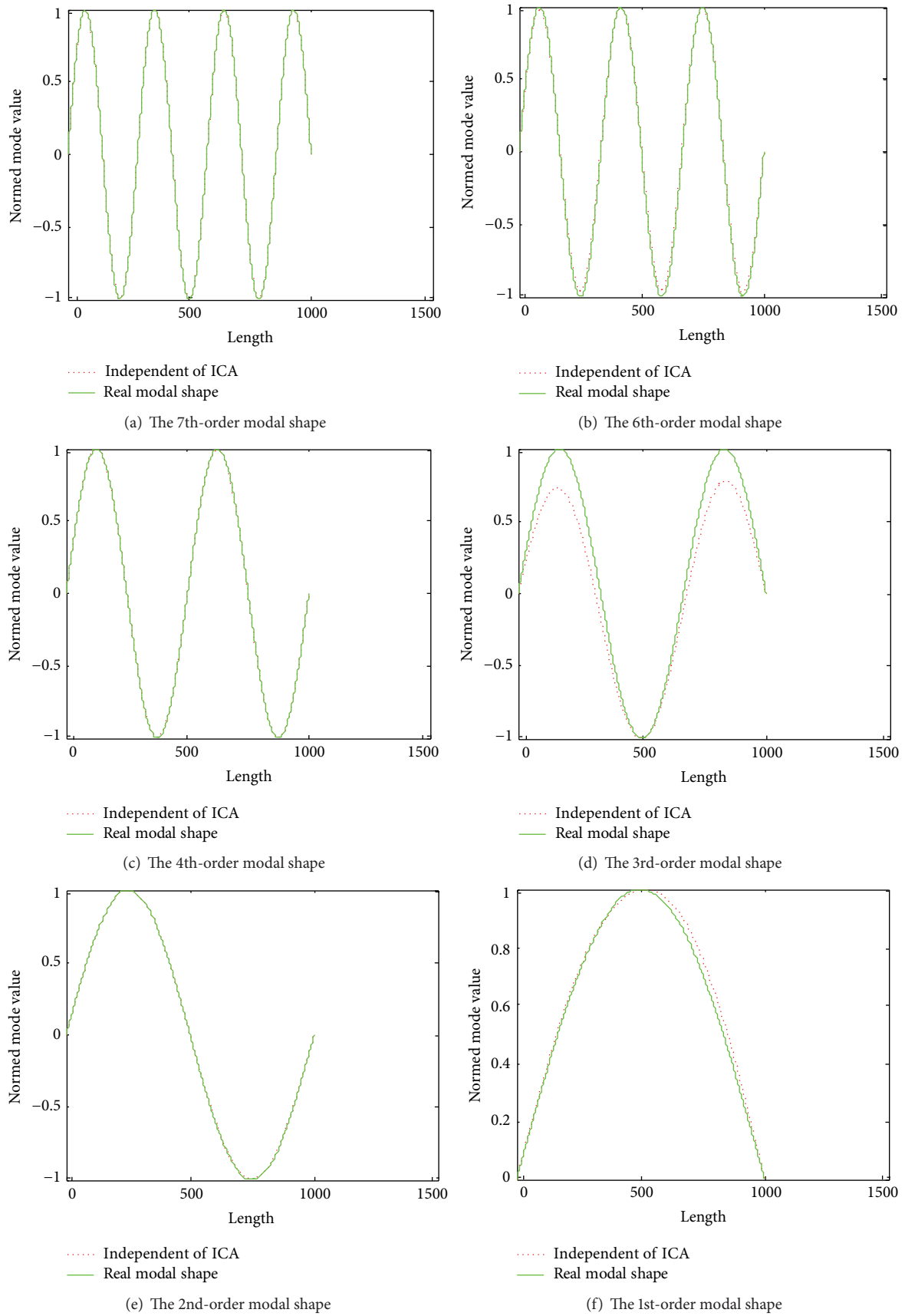


FIGURE 8: Modal shape comparison between FEA and separating vector by minimum mutual information with gradient descent method.

TABLE 8: The identified order and MAC value of modal shapes.

Identified order	Modal order of FEA	K_{MAC}
1	7	1.0000
2	6	0.9993
3	4	1.0000
4	3	0.9602
5	2	1.0000
6	1	0.9991

TABLE 9: Error of natural frequencies identified by maximum likelihood estimation with stochastic gradient descent method.

Modal order	Real natural frequency (Hz)	Frequency identified (Hz)	Percentage of relative error
1	22.89	23	0.481%
2	91.55	92	0.492%
3	205.99	206	0.005%
4	366.23	367	0.210%
6	824	825	0.121%
7	1121.6	1122	0.036%

TABLE 10: Identified order and MAC value of modal shapes.

Identified order	Modal order of FEA	K_{MAC}
1	7	0.5432
2	3	0.8121
3	4	0.9515
4	1	0.9138
5	2	0.9925
6	6	0.9726

Table 8 shows the identified order and MAC value of modal shapes.

4.3.4. Maximum Likelihood Estimation with Stochastic Gradient Descent Method. According to the process of modal parameters identification by maximum likelihood estimation with stochastic gradient descent method, Figure 9 shows FFT of each independent component.

The abscissa of main peak in each subfigure of Figure 9 is each natural frequency identified by maximum likelihood estimation with stochastic gradient descent method. Table 9 shows errors of natural frequencies identified by maximum likelihood estimation with stochastic gradient descent method.

According to the identified frequency value, Figure 10 shows the modal shapes which compared the real modal shapes calculated by FEA.

Table 10 shows the identified order and MAC value of modal shapes.

4.3.5. Infomax with Gradient Descent Method. According to the process of modal parameters identification by Infomax with gradient descent method, Figure 11 shows FFT of each independent component.

TABLE 11: Error of natural frequencies identified by Infomax with gradient descent method.

Modal order	Real natural frequency (Hz)	Frequency identified (Hz)	Percentage of relative error
1	22.89	23	0.481%
2	91.55	92	0.492%
3	205.99	206	0.005%
4	366.23	367	0.210%
6	824	825	0.121%
7	1121.6	1122	0.036%

TABLE 12: The identified order and MAC value of modal shapes.

Identified order	Modal order of FEA	K_{MAC}
1	1	0.8930
2	4	0.9039
3	6	0.8954
4	7	0.2543
5	3	0.4066
6	2	0.9914

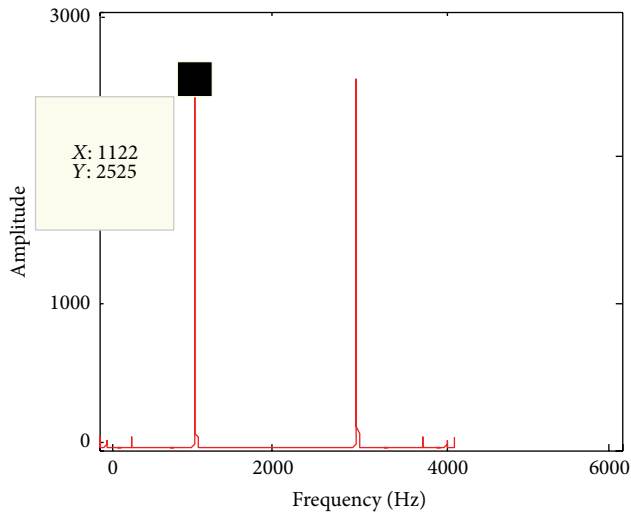
The abscissa of main peak in each subfigure of Figure 11 is each natural frequency identified by Infomax with gradient descent method. Table 11 shows errors of natural frequencies identified by Infomax with gradient descent method.

According to the identified frequency value, Figure 12 shows the modal shapes which compared the real modal shapes calculated by FEA.

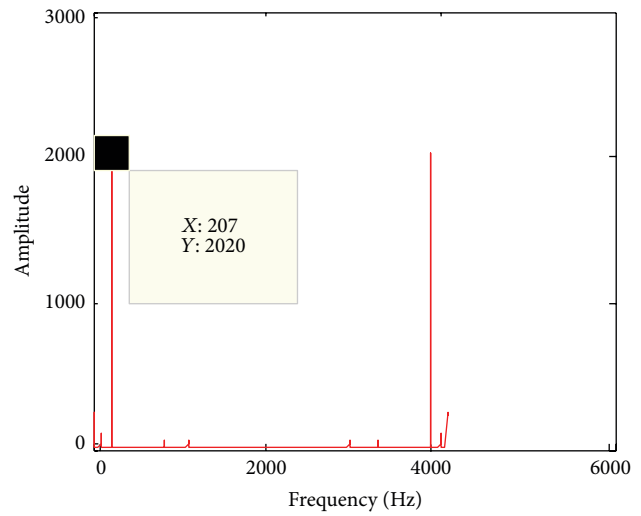
Table 12 shows the identified orders and MAC values of modal shapes.

4.3.6. Results Analysis. Comparing the results of output-only modal parameters identification with five objective functions of ICA algorithms, results analysis can be gained from the following respects:

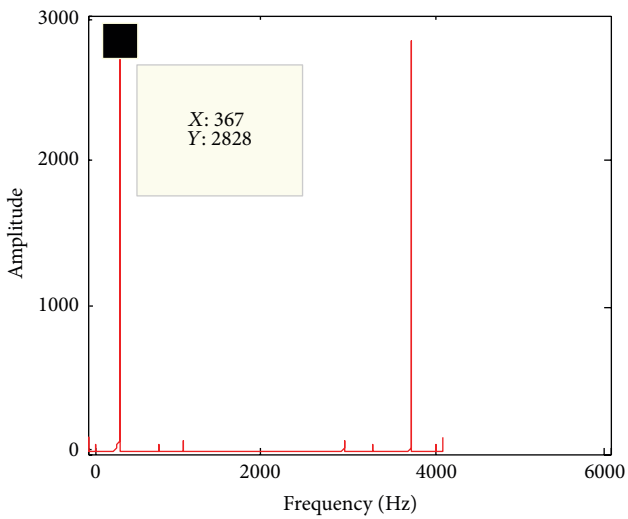
- (1) From Figures 3–12, the effectiveness of five ICA algorithms for output-only modal parameter identification is verified.
- (2) In Figures 3 and 5, natural frequencies identified are with no second peak. But in Figures 7, 9, and 11, natural frequencies identified are with second peak. A big reason for the difference is that the process of maximization kurtosis and negentropy for output-only modal analysis is in a serial manner, and other methods are in parallel manner.
- (3) In Tables 4, 6, 8, 10, and 12, the order of identified modal parameters is different from the order of those calculated by FEA, and the reason is that ICA algorithms are with three uncertainties, with all the modal shapes normalized which are shown in Figures 4, 6, 8, 10, and 12.
- (4) From Figures 3–12, by comparing the modal shapes and natural frequencies as shown in Tables 3–12, we



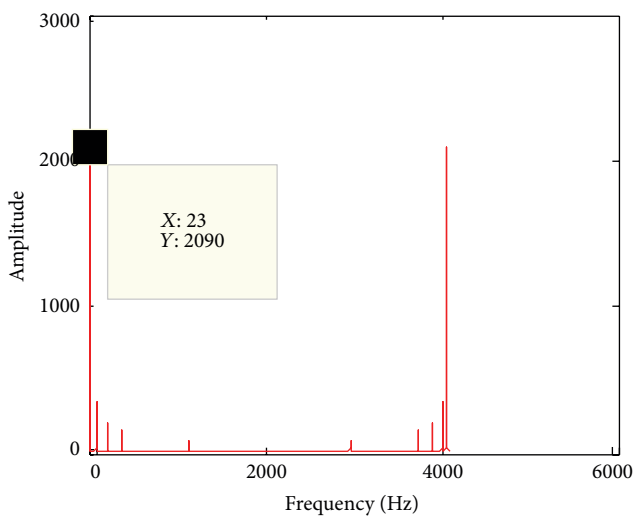
(a) FFT of the 1st independent component



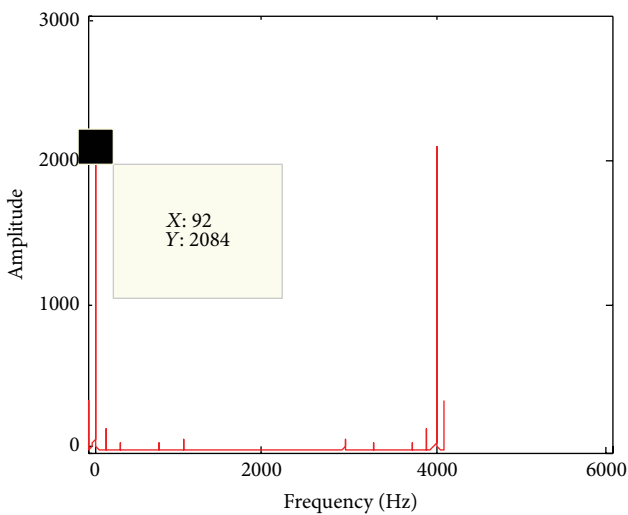
(b) FFT of the 2nd independent component



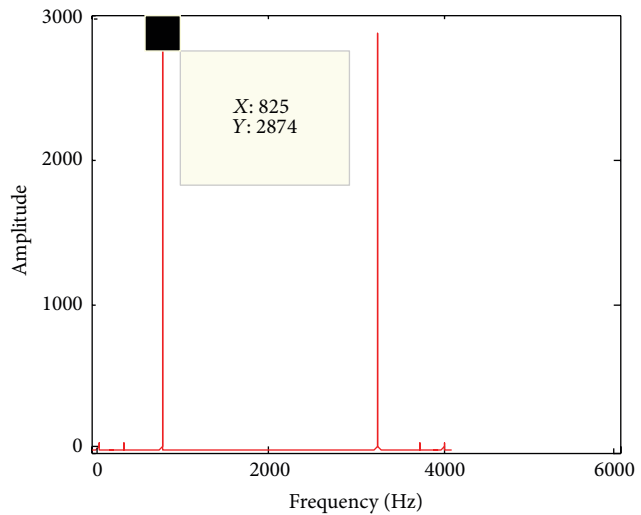
(c) FFT of the 3rd independent component



(d) FFT of the 4th independent component



(e) FFT of the 5th independent component



(f) FFT of the 6th independent component

FIGURE 9: FFT of each independent component identified by maximum likelihood estimation with stochastic gradient descent method.

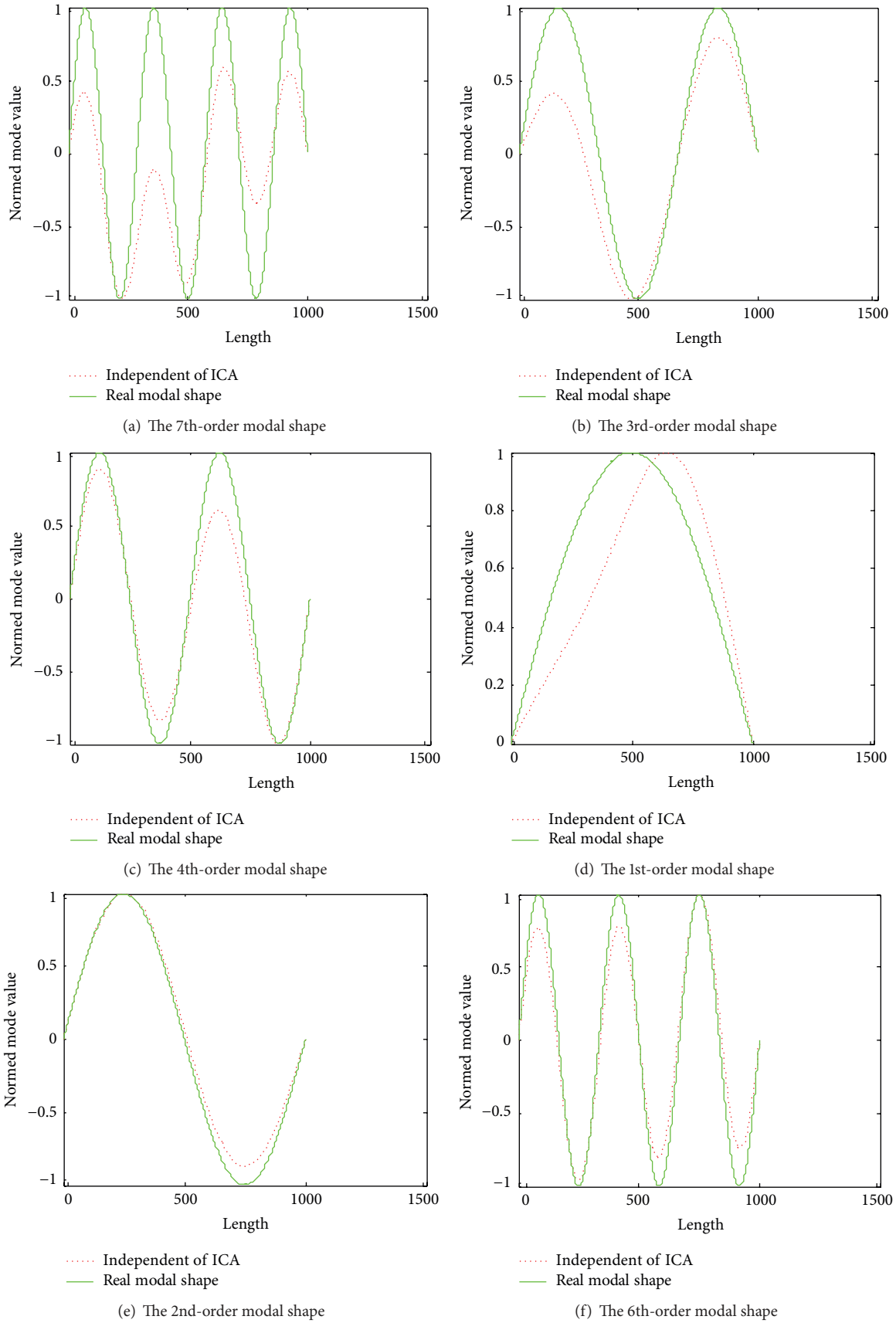
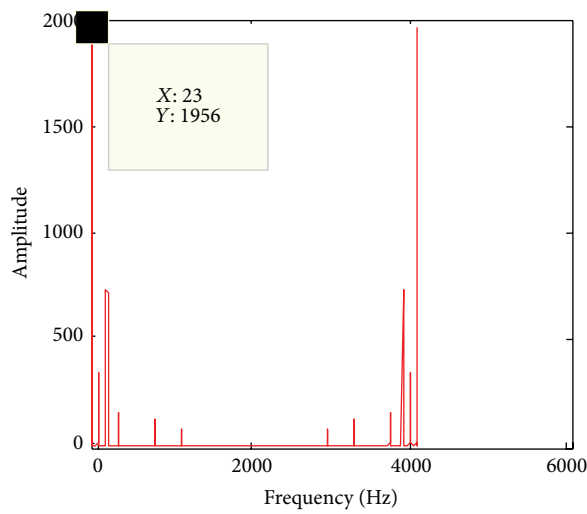
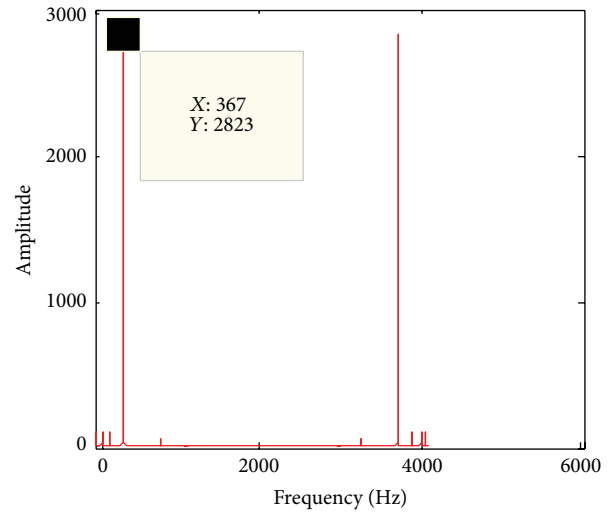


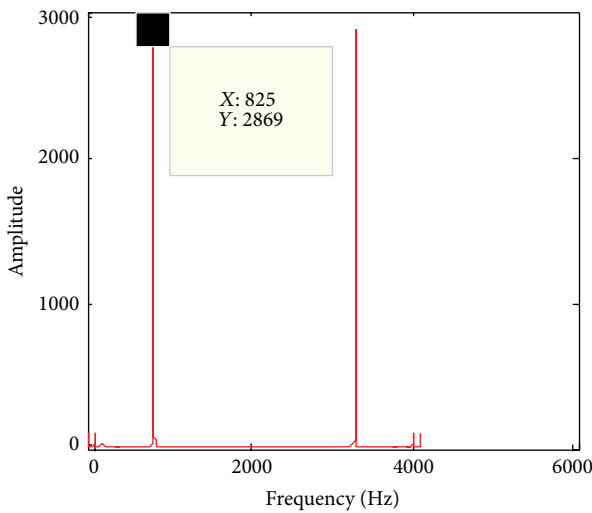
FIGURE 10: Modal shape comparison between FEA and separating vector by maximum likelihood estimation with stochastic gradient descent method.



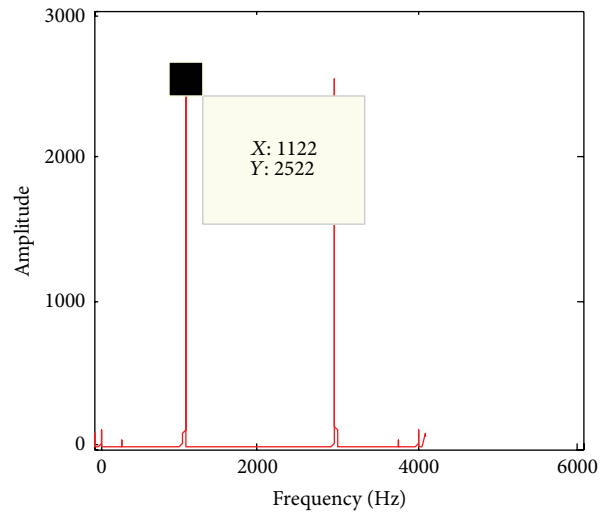
(a) FFT of the 1st independent component



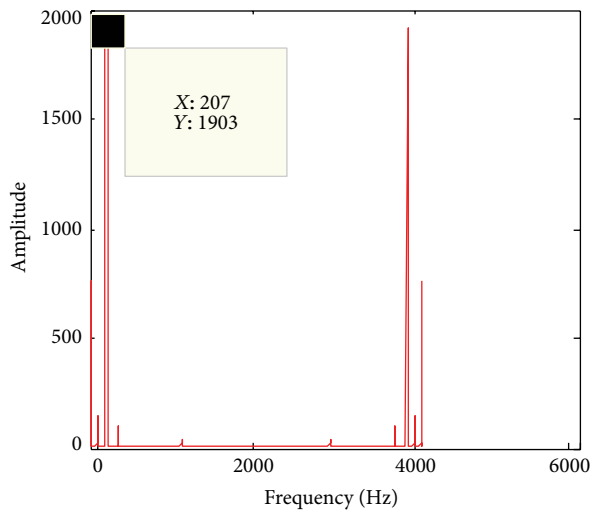
(b) FFT of the 2nd independent component



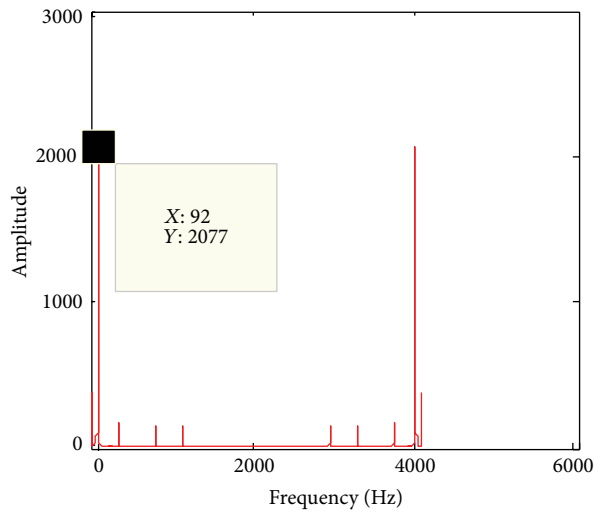
(c) FFT of the 3rd independent component



(d) FFT of the 4th independent component



(e) FFT of the 5th independent component



(f) FFT of the 6th independent component

FIGURE 11: FFT of each independent component identified by Infomax with gradient descent method.

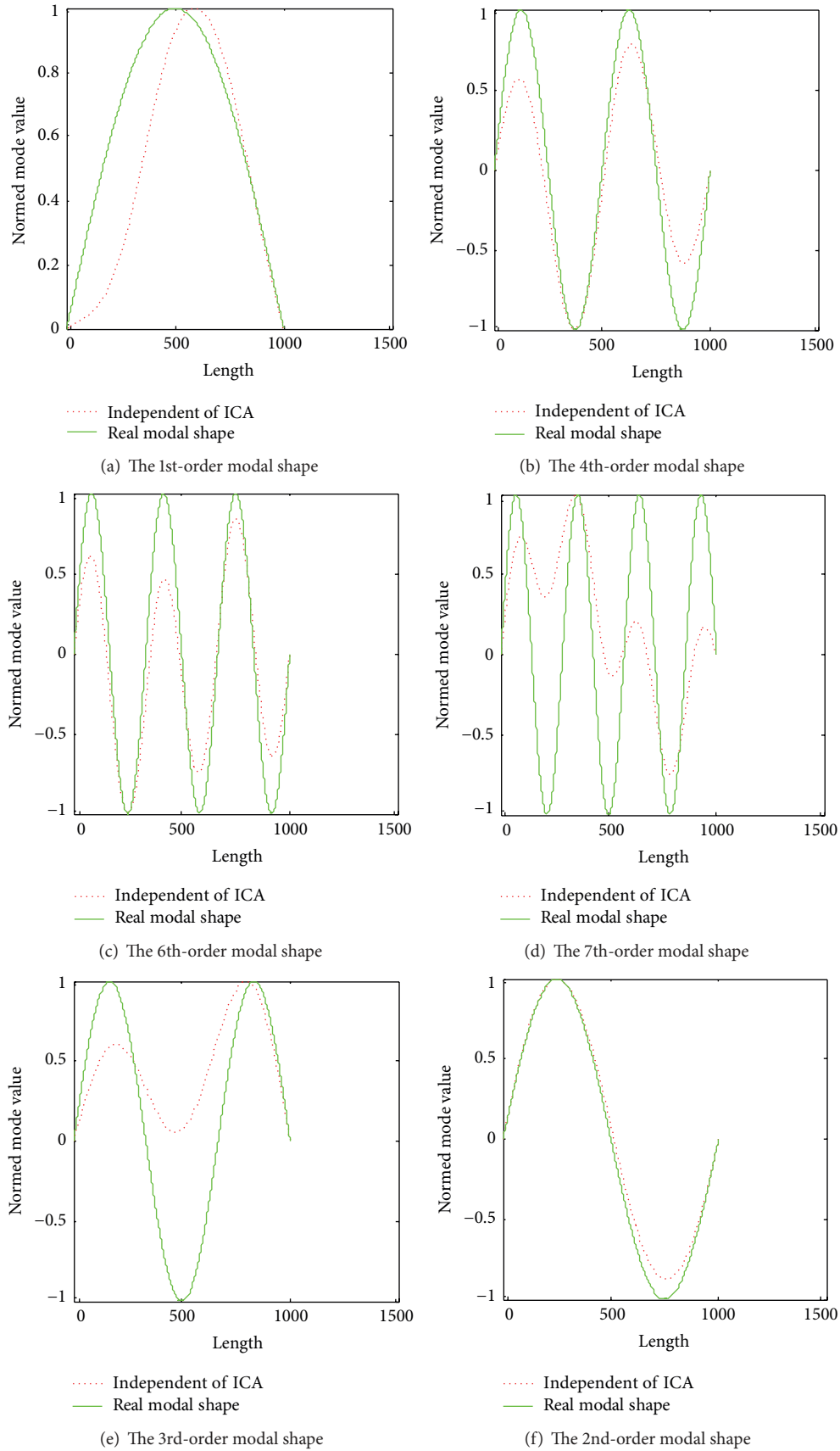


FIGURE 12: Modal shape comparison between FEA and separating vector by Infomax with gradient descent method.

TABLE 13: Error of natural frequencies identified by maximization kurtosis with gradient descent method.

Modal order	Real natural frequency (Hz)	Frequency identified (Hz)	Percentage of relative error
1	22.89	23	0.481%
2	91.55	—	—
3	205.99	207	0.490%
4	366.23	—	—
6	824	825	0.121%
7	1121.6	—	—

TABLE 14: The convergence times and MAC value of modal shapes when observed signals are with measurement noise.

Identified order	Number of convergences	Modal order of FFE	K_{MAC}
1	42	1	0.4550
2	27	6	0.1091
3	39	3	0.0447
4	29	3	0.0978
5	35	6	0.0472
6	24	6	0.9306

conclude that the effect of modal parameters identification is the best when the objective function is maximization non-Gaussian and the values of MACs are close to 1.

4.4. Results of Simulation Verification with Measurement Noise and Results Analysis. From the results above, it is verified that the effect of modal parameters identified by the method is better than others when the objective function is non-Gaussian. However, a maximization non-Gaussian measure contains maximization kurtosis and negentropy. In order to verify the robustness of maximization kurtosis and negentropy, the response data are added with 1% measurement noise.

4.4.1. Maximization Kurtosis with Gradient Descent Method. When observed signals are with measurement noise, Figure 13 shows FFT of each independent component.

The abscissa of main peak in each subfigure of Figure 13 corresponds to each natural frequency identified by maximization kurtosis with gradient descent method. And Table 13 shows errors of natural frequencies identified by maximization kurtosis with gradient descent method.

According to the identified frequency value, Figure 14 shows the modal shapes, which compared the real modal shapes calculated by FEA.

Table 14 shows the convergence times and MAC values of modal shapes when observed signals are with measurement noise.

4.4.2. Maximization Negentropy with Quasi-Newton Method. Figure 15 shows FFT of each independent component when observed signals are with measurement noise.

TABLE 15: Error of natural frequencies identified by maximization negentropy with quasi-Newton iteration method with measurement noise.

Modal order	Real natural frequency (Hz)	Frequency identified (Hz)	Percentage of relative error
1	22.89	—	—
2	91.55	92	0.492%
3	205.99	207	0.490%
4	366.23	367	0.210%
6	824	825	0.121%
7	1121.6	1122	0.036%

TABLE 16: The convergence times and MAC value of modal shapes when observed signals are with measurement noise.

Identified order	Number of convergences	Modal order of FEA	K_{MAC}
1	253	2	0.7665
2	128	4	0.9410
3	27	2	0.9996
4	224	7	0.7871
5	361	6	0.9891
6	147	3	0.1148

The abscissa of main peak in each subfigure of Figure 15 is each natural frequency identified by maximization negentropy with quasi-Newton method. And Table 15 shows errors of natural frequencies identified by maximization negentropy with quasi-Newton method.

According to the identified frequency values, Figure 16 shows the modal shapes, which compared the real modal shapes calculated by FEA.

Table 16 shows the convergence times and MAC values of modal shapes when observed signals are with measurement noise.

4.4.3. Results Analysis. From Figures 3–12 above, we can see that when observed signals are with 1% measurement Gauss noise, maximization negentropy is more robust to measurement noise than maximization kurtosis. However, minimum mutual information and Infomax method as well as maximum likelihood estimation are with no advantage even when there is no measurement noise.

5. Conclusions

This paper verified the effectiveness of different objective functions and different optimization methods of ICA algorithm for output-only modal parameters identification. The methods are different from each other in content, but they are equivalent in essence. The simulation verification results show that modal parameters identification is the best when the objective function is maximization non-Gaussian and the values of MACs are close to 1, and maximization negentropy is more robust to measurement noise than kurtosis.

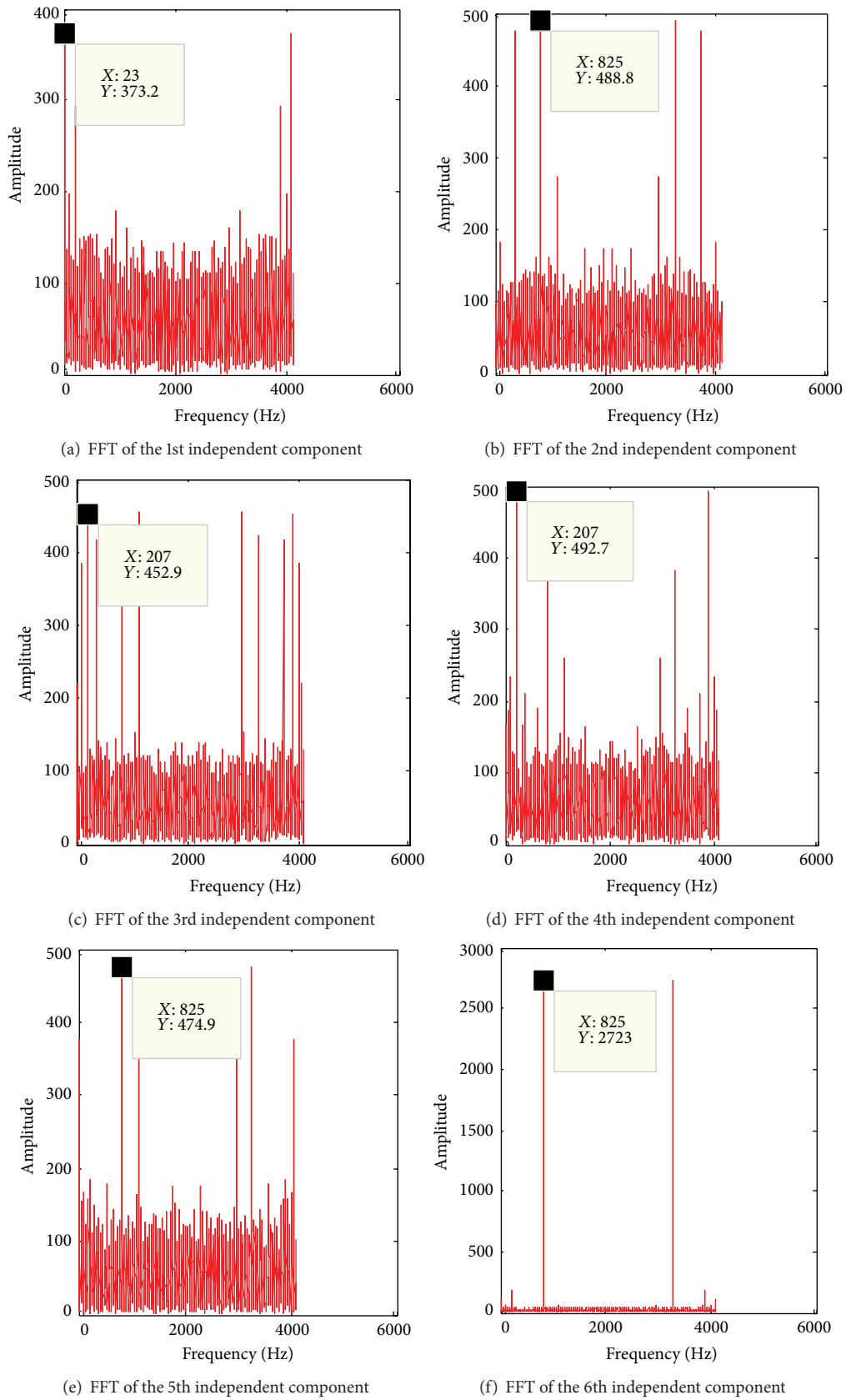


FIGURE 13: FFT of each independent component identified by maximization kurtosis with gradient descent method with measurement noise.

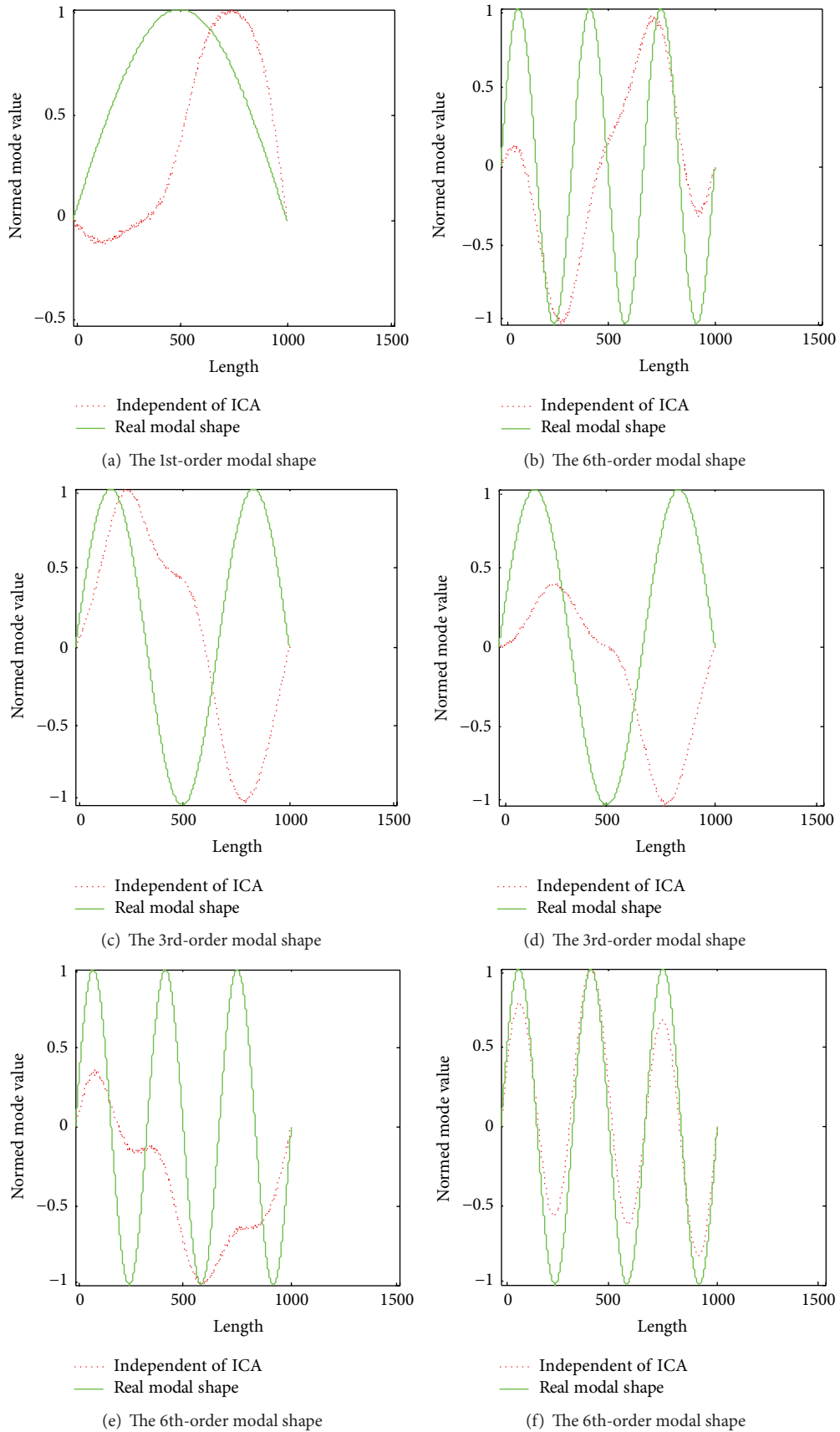


FIGURE 14: Modal shape comparison between FEA and separating vector by maximization kurtosis with gradient descent method with measurement noise.

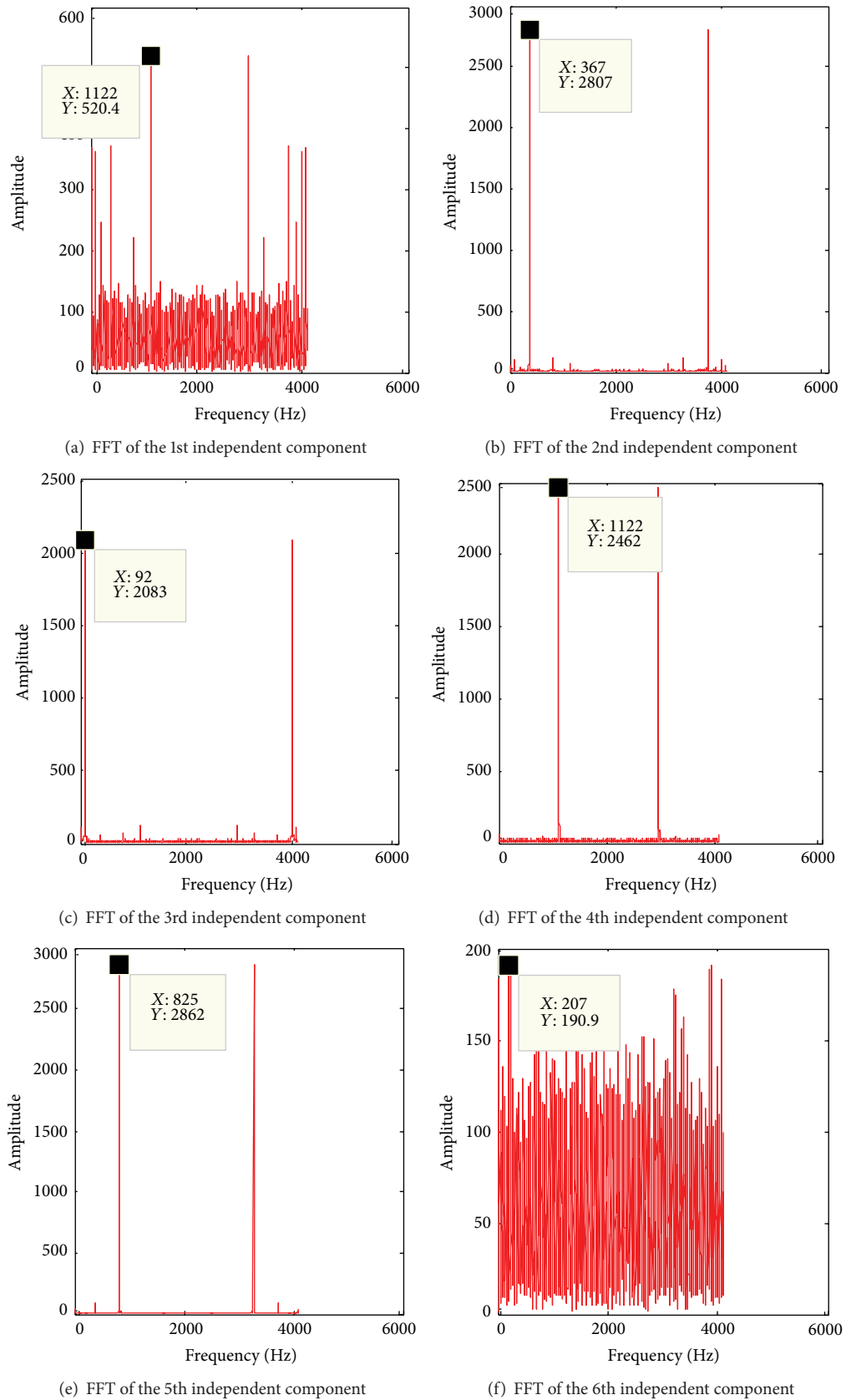


FIGURE 15: FFT of each independent component identified by maximization negentropy with quasi-Newton iteration method with measurement noise.

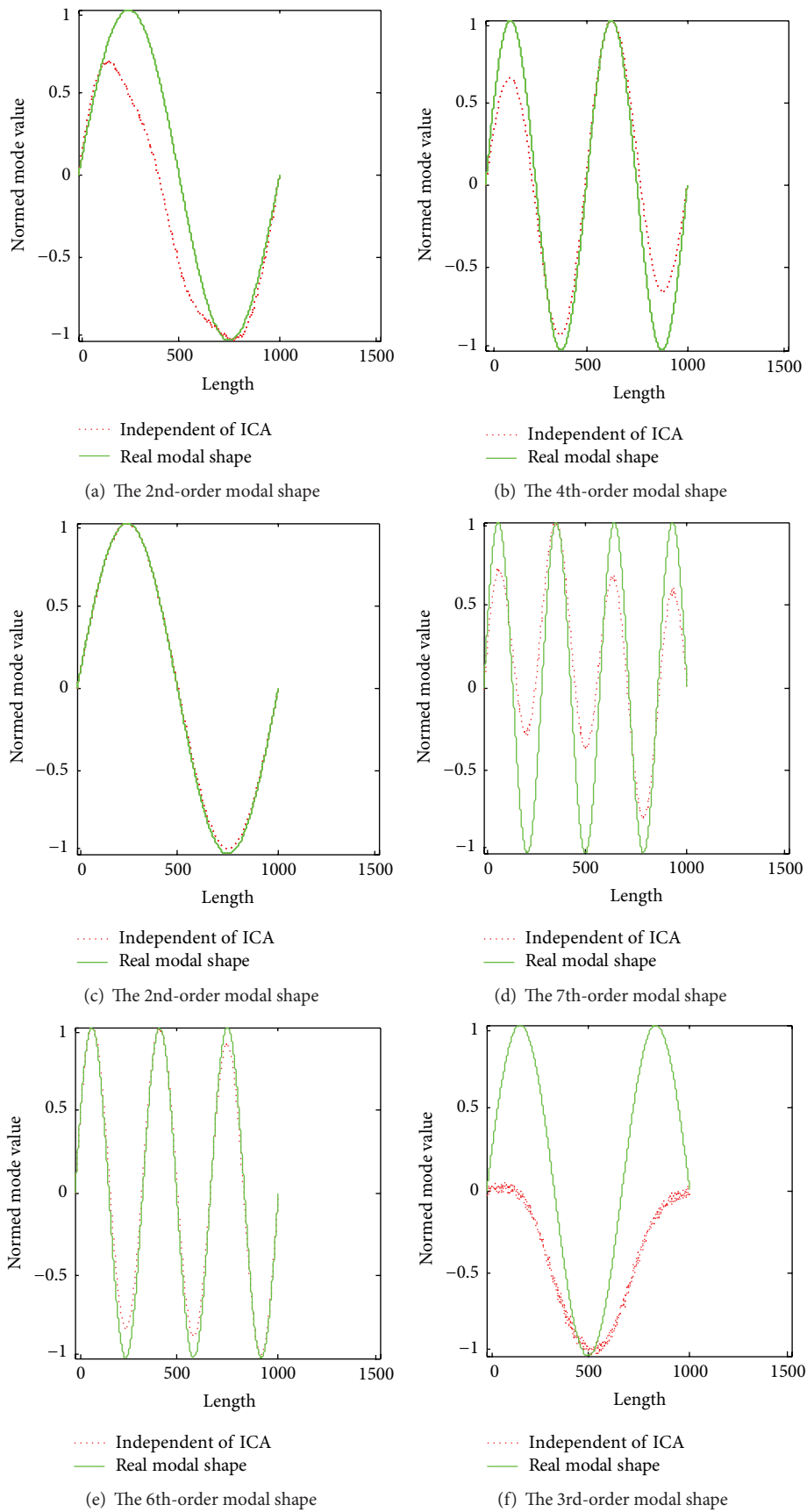


FIGURE 16: Modal shape comparison between FEA and separating vector by maximization negentropy with quasi-Newton iteration method with measurement noise.

However, ICA algorithms are sensitive to measurement noise and the mode is easy to miss. So how to find the missed modal with prior knowledge and how to fuse different objective functions to search for a more robustness objective function or optimization method and improve the accuracy of identifying modal parameters are the further research directions.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this article and regarding the funding that they have received.

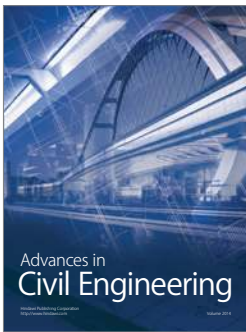
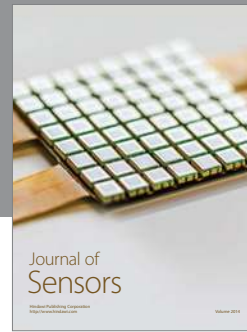
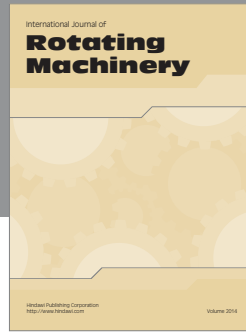
Acknowledgments

This work was financially supported by National Natural Science Foundation of China (Grant nos. 51305142, 51305143) and Postgraduate Scientific Research Innovation Ability Training Plan Funding Projects of Huaqiao University (no. 1400214012). Thanks are due to Dr. Yiwen Zhang who contributed to the grammatical editing of this paper.

References

- [1] S. I. McNeill, *Modal Identification Using Blind Source Separation Techniques*, 2007.
- [2] G. Kerschen, F. Poncelet, and J.-C. Golinval, "Physical interpretation of independent component analysis in structural dynamics," *Mechanical Systems and Signal Processing*, vol. 21, no. 4, pp. 1561–1575, 2007.
- [3] F. Poncelet, G. Kerschen, J.-C. Golinval, and D. Verhelst, "Output-only modal analysis using blind source separation techniques," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2335–2358, 2007.
- [4] F. Poncelet, G. Kerschen, C. J. Golinval et al., "Blind source separation techniques-another way of doing operational modal analysis," in *Proceedings of the International Operational Modal Analysis Conference*, Copenhagen, Denmark, 2007.
- [5] W. Zhou and D. Chelidze, "Blind source separation based vibration mode identification," *Mechanical Systems and Signal Processing*, vol. 21, no. 8, pp. 3072–3087, 2007.
- [6] J. Antoni, "Blind separation of vibration components: principles and demonstrations," *Mechanical Systems and Signal Processing*, vol. 19, no. 6, pp. 1166–1180, 2005.
- [7] S. H. Chao and C. H. Loh, "Output-only dynamic system identification using blind source separation technique," in *Proceedings of the 5th World Conference on Structural Control and Monitoring*, CDROM, Tokyo, Japan, July 2010.
- [8] S. I. McNeill and D. C. Zimmerman, "Relating independent components to free-vibration modal responses," *Shock and Vibration*, vol. 17, no. 2, pp. 161–170, 2010.
- [9] F. Abazarsa, F. Nateghi, S. F. Ghahari, and E. Taciroglu, "Blind modal identification of non-classically damped systems from free or ambient vibration records," *Earthquake Spectra*, vol. 29, no. 4, pp. 1137–1157, 2013.
- [10] J. Antoni, S. Chauhan, T. Monnier, and K. Gryllias, "Least action criteria for blind separation of structural modes," *Mechanics & Industry*, vol. 14, no. 6, pp. 397–411, 2013.
- [11] A. Sadhu and B. Hazra, "A novel damage detection algorithm using time-series analysis-based blind source separation," *Shock and Vibration*, vol. 20, no. 3, pp. 423–438, 2013.
- [12] A. Sadhu and S. Narasimhan, "A decentralized blind source separation algorithm for ambient modal identification in the presence of narrowband disturbances," *Structural Control and Health Monitoring*, vol. 21, no. 3, pp. 282–302, 2014.
- [13] A. Sadhu, A. Goldack, and S. Narasimhan, "Ambient modal identification using multi-rank parallel factor decomposition," *Structural Control and Health Monitoring*, vol. 22, no. 4, pp. 595–614, 2015.
- [14] B. Hazra, A. Sadhu, A. J. Roffel, and S. Narasimhan, "Hybrid time-frequency blind source separation towards ambient system identification of structures," *Computer-Aided Civil and Infrastructure Engineering*, vol. 27, no. 5, pp. 314–332, 2012.
- [15] A. Sadhu, B. Hazra, and S. Narasimhan, "Blind identification of earthquake-excited structures," *Smart Materials and Structures*, vol. 21, no. 4, Article ID 045019, 2012.
- [16] S. I. McNeill and D. C. Zimmerman, "Blind modal identification applied to output-only building vibration," in *Structural Dynamics*, vol. 3, pp. 157–176, Springer, New York, NY, USA, 2011.
- [17] C. Rainieri and G. Fabbrocino, *Operational Modal Analysis of Civil Engineering Structures*, Springer, New York, NY, USA, 2014.
- [18] A. Sadhu, B. Hazra, S. Narasimhan, and M. D. Pandey, "Decentralized modal identification using sparse blind source separation," *Smart Materials and Structures*, vol. 20, no. 12, Article ID 125009, 2011.
- [19] Y. Yang and S. Nagarajaiah, "Output-only modal identification with limited sensors using sparse component analysis," *Journal of Sound and Vibration*, vol. 332, no. 19, pp. 4741–4765, 2013.
- [20] A. Sadhu and S. Narasimhan, "Blind source separation of convolutive mixtures towards modal identification," in *Topics on the Dynamics of Civil Structures*, vol. 1, pp. 209–220, Springer, New York, NY, USA, 2012.
- [21] J. Hang and Z. Yi, "Operational modal identification technique based on independent component analysis," in *Proceedings of the International Conference on Electric Technology and Civil Engineering (ICETCE '11)*, pp. 2430–2433, IEEE, Lushan, China, April 2011.
- [22] X. Wan, W. Ma, J. Wang et al., "Modal parameter identification based on independent component analysis," *Machinery Design & Manufacture*, no. 1, pp. 137–139, 2013.
- [23] Y. Yang and S. Nagarajaiah, "Time-frequency blind source separation using independent component analysis for output-only modal identification of highly damped structures," *Journal of Structural Engineering*, vol. 139, no. 10, pp. 1780–1793, 2013.
- [24] Y. Yang and S. Nagarajaiah, "Blind identification of damage in time-varying systems using independent component analysis with wavelet transform," *Mechanical Systems and Signal Processing*, vol. 47, no. 1–2, pp. 3–20, 2014.
- [25] S. Choi, "Independent component analysis," in *Encyclopedia of Biometrics*, pp. 917–924, 2015.
- [26] C. Jutten and J. Herault, "Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, no. 1, pp. 1–10, 1991.
- [27] C. Wang, J. Gou, J. Bai, and Y. Guirong, "Modal parameter identification with principal component analysis," *Journal of Xi'an Jiaotong University*, vol. 47, no. 11, pp. 97–104, 2013.

- [28] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, NY, USA, 2004.
- [29] C. Burges, T. Shaked, E. Renshaw et al., "Learning to rank using gradient descent," in *Proceedings of the 22nd International Conference on Machine Learning (ICML '05)*, pp. 89–96, ACM, August 2005.
- [30] M. Zinkevich, M. Weimer, L. Li et al., "Parallelized stochastic gradient descent," in *Advances in Neural Information Processing Systems*, pp. 2595–2603, 2010.
- [31] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Mathematics of Computation*, vol. 24, no. 111, pp. 647–656, 1970.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

