

Relating independent components to free-vibration modal responses

S.I. McNeill^{a,*} and D.C. Zimmerman^b

^a*Stress Engineering Services, Inc., 13800 Westfair East Drive, Houston, TX, USA*

^b*Department of Mechanical Engineering, University of Houston, 4800 Calhoun Rd., Houston, TX, USA*

Received 10 September 2008

Revised 3 June 2009

Abstract. In recent literature, attempts have been made to apply Independent Component Analysis (ICA) techniques to the modal identification problem. Paramount to this task is establishing a relationship between the source components realized using the suggested ICA technique and modal responses. In this paper, the relationship between independent components and free-vibration modal responses is discussed. Theoretical arguments are presented for responses of undamped systems and arguments are extended to damped responses.

Keywords: Blind source separation, independent component analysis, hebbian learning, output only modal identification, non-parametric identification

1. Introduction

In the last decade significant advances have been made in the areas of Blind Source Separation (BSS) and Independent Component Analysis (ICA) [1–3]. BSS techniques attempt to recover special source components that are mixed together in the measured data. ICA is a popular method used to solve the BSS problem by recovering components that are statistically independent. Exploratory work has been done to investigate the application of these techniques to non-parametric modal identification [4–7]. In these studies, BSS techniques were applied directly, without modification, to the measured data. However, there are several deficiencies with this approach. These deficiencies stem from the discrepancy between modal responses and the source components identified with BSS methods.

In reference [8], the relationship between modal responses of the diagonally-damped system and source components identified using a popular BSS technique, known as the Second Order Blind Identification (SOBI) algorithm, were established. In addition, the robust second order statistical methods employed in the SOBI algorithm were extended to the generally damped system. The extension is essentially an application of the SOBI algorithm on a data set that has been supplemented with Hilbert transform pairs and windowed using a Gaussian window. More details of the method, including a modal accuracy indicator, can be found in [9]. Through the examples and applications presented in [8,9], this technique as proven to be a competitive method for output-only modal identification.

Methods used to perform modal identification using ICA techniques have not been as well developed, however. In reference [10], the connection between free-vibration modal responses of the diagonally-damped system and independent components obtained from ICA was suggested by example. Based on this relationship, the Hebbian-like learning rule [11] was modified to estimate modal responses from measured free-vibration data. Applications were provided to verify the proposed method.

*Corresponding author. E-mail: scot.mcneill@stress.com.

In this paper, the relationship between independent components and free-vibration modal responses is reinforced using theoretical arguments. The remainder of the paper is organized as follows. The basic framework for BSS techniques, including ICA, is reviewed in Section 2. The use of BSS techniques to perform modal identification is also introduced in this section. Section 3 establishes the relationship between modal responses and independent components. In Sections 3.1–3.3, theoretical arguments are presented for responses of undamped systems. Later, in Section 3.4, the arguments are extended to damped responses. A brief summary is then provided in Section 4.

2. BSS for modal identification

BSS attempts to find statistically independent components, $s_j(t)$, embedded in measured data $x_i(t)$. The techniques applicable for modal identification assume that the measured data is a linear mixture (as opposed to convolutive) of the components. Suppose that there are m channels of measured data and n components. Making the time dependence implicit, the relation between the components and the measured data can be written as

$$\mathbf{x}^{(m \times 1)} = \mathbf{A}^{(m \times n)} \mathbf{s}^{(n \times 1)}, \quad (1)$$

where \mathbf{A} is the (constant) mixing matrix and the dimensions have been placed in the superscript. All quantities are real valued. The objective of BSS is to simultaneously estimate the mixing matrix, \mathbf{A} , and the vector of components, $\mathbf{s}(t)$, from the observed data, $\mathbf{x}(t)$. Due to the number of variables involved, this task requires a characterization of the source components, $\mathbf{s}(t)$. Many BSS techniques use second order statistical information (e.g. variance) to describe the components, while ICA typically uses higher order statistics (e.g. kurtosis) [1,3]. It is appropriate to consider the inverse relationship of Eq. (1),

$$\mathbf{s}^{(n \times 1)} = \mathbf{W}^{(n \times m)} \mathbf{x}^{(m \times 1)}. \quad (2)$$

The de-mixing matrix, \mathbf{W} , is the (generalized, if necessary) inverse of the mixing matrix, \mathbf{A} . The task is now to estimate \mathbf{W} and $\mathbf{s}(t)$. Note that in order to estimate the n components, it is necessary to have enough independent observations. This requires $m \geq n$ and the rank of \mathbf{A} is n . It can be seen that the number of components, n , must be deduced from the observed data $\mathbf{x}(t)$. This is similar to system order determination in traditional modal identification. Determination of the number of components can be accomplished using the same techniques as in traditional modal analysis (e.g. complex mode indicator function).

Because the inverse of the mixing matrix, \mathbf{W} , and the components, $\mathbf{s}(t)$, must be estimated simultaneously, any scalar multiplier of one of the components $s_j(t)$ could be canceled by dividing the corresponding column \mathbf{a}_j by the same scalar. This leads to some ambiguities. First, the variances of the independent components cannot be determined. This means that the amplitude and sign of each component $s_j(t)$ are unknown. A natural way to fix the amplitude of each component is to set the variance equal to one: $E\{s_j^2(t)\} = 1$. Note that the sign is still ambiguous. Second, the order of importance of the components is unknown. This is in contrast to the familiar Principal Component Analysis (PCA), where the principal components are ordered by their variance.

The potential application of BSS on vibration data is fairly obvious. In linear vibration, the physical responses, $\mathbf{x}(t)$, are constructed from the modal responses, $\mathbf{q}(t)$, by multiplication of the modal matrix, Φ ,

$$\mathbf{x} = \Phi \mathbf{q}. \quad (3)$$

One might consider using BSS to estimate both the (inverse) modal matrix and the modal responses,

$$\mathbf{q} = \Phi^{-1} \mathbf{x}. \quad (4)$$

One should see the connection with Eqs (1) and (2). Also note that the modal matrix is real valued if the topology of the damping matrix is restricted (e.g. proportional damping). This is an important observation because the mixing matrix, \mathbf{A} , is restricted to be real valued in most algorithms.

It is important to notice that there is no assumed mathematical model that governs the underlying process that produced the measured data. Most system/modal identification techniques assume some form of a model, and then fit parameters to the model using the measured data. In the BSS technique, the mixing model is the only assumption. Therefore using BSS to perform modal identification is non-parametric. In order to apply BSS to perform modal identification, a connection between components and modal responses must be made.

3. Relating free responses to independent components

This section is divided into four parts. In Section 3.1 the free-vibration of the diagonally-damped system is discussed. Next, kurtosis is established as a measure of independence for vibration signals in Section 3.2. Finally, the relationship between free-vibration modal responses and independent components is established for the undamped system and the diagonally-damped system in Sections 3.3 and 3.4, respectively.

3.1. Free-vibration modal responses of the diagonally-damped system

The connection between free-vibration modal responses and independent components can be made by studying the properties of modal responses. Consider the familiar second order Equation Of Motion (EOM) for a linear system,

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t). \quad (5)$$

Here $\mathbf{x}(t) \in \mathbb{R}^N$ is the displacement vector, \mathbf{M} is the positive definite mass matrix, \mathbf{C} is the positive semi-definite damping matrix and \mathbf{K} is the positive semi-definite stiffness matrix. All matrices are real valued and symmetric. Because \mathbf{K} is positive semi-definite in general, the structure will have rigid body modes. However the rigid body modes and responses are generally not of interest in the modal identification problem. Therefore we will only consider oscillatory responses. Discussion will be restricted to responses of the diagonally-damped system. In this case, the matrix, $\mathbf{C}\mathbf{M}^{-1}\mathbf{K}$, is symmetric. A well-known special case of this is proportional damping ($\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K}$, where α, β are scalars). In this case, the modeshapes are real valued and equal to the *normal* modes (modes of the undamped system). The free-vibration response of the system may be written as a summation of the normal modeshapes, ϕ_j , and the modal responses, $q_j(t)$,

$$x_i(t) = \sum_{j=1}^N \varphi_{ij} q_j(t) = \sum_{j=1}^N \alpha_j \varphi_{ij} e^{-\zeta_j \omega_{nj} t} \sin(\omega_j t + \theta_j), \quad 1 \leq i \leq N. \quad (6)$$

Here $\varphi_{ij}, q_j(t) \in \mathbb{R}$, due to the restriction on the damping matrix. The normal modeshapes are obtained from the undamped eigenvalue problem,

$$(\lambda_j \mathbf{M} - \mathbf{K}) \phi_j = 0, \quad 1 \leq j \leq N. \quad (7)$$

Here, the modeshapes are determined up to multiplication by a real scalar. Due to the fact that \mathbf{K} is positive semi-definite and \mathbf{M} is positive definite it can be shown that the normal modes are *mass* orthogonal, $\Phi^T \mathbf{M} \Phi = \text{diagonal}$. The damped frequencies, ω_j , natural frequencies, ω_{nj} , and damping ratios, ζ_j , are obtained by using the relationships,

$$\omega_{nj}^2 = \lambda_j, \quad \zeta_j = \frac{\phi_j^T \mathbf{C} \phi_j}{2\omega_{nj}}, \quad \omega_j = \omega_{nj} \sqrt{1 - \zeta_j^2}, \quad (8)$$

where $\phi_j^T \mathbf{M} \phi_j = 1$.

3.2. Kurtosis as a measure of independence

In order to establish the connection between source components resulting from ICA and modal responses, a measure of independence must be established. ICA techniques attempt to find components that are as independent as possible, according to some definition of independence. The definition of statistical independence of random variables, $x_i, i = \{1, 2, \dots, m\}$, states that the joint Probability Density Function (PDF), $P(\mathbf{x})$, factors as the product of the marginal PDFs, $P_i(x_i)$,

$$P(\mathbf{x}) = \prod_{i=1}^m P_i(x_i). \quad (9)$$

Because this definition does not lend itself well to numerical methods, it is rarely exploited in ICA algorithms. Some ICA algorithms, such as FastICA and Hebbian-like learning, assert that independent components are as non-Gaussian as possible [3]. The skewness and kurtosis are classical measures of nongaussianity. Since modal responses have symmetric PDF's about the mean value of zero, the skewness vanishes. Therefore kurtosis becomes the appropriate measure of independence. The kurtosis of a zero-mean random variable, x , is given by,

$$ku(x) = \frac{E\{x^4\}}{(E\{x^2\})^2}. \quad (10)$$

Here E is the expectation operator,

$$E\{x\} = \int_{-\infty}^{\infty} xP(x) dx. \quad (11)$$

For sampled data with K samples, the expectation can be computed statistically by,

$$E\{x(k)\} = \frac{1}{K} \sum_{k=0}^K x(k). \quad (12)$$

Note that the kurtosis for a random variable with a Gaussian PDF is equal to three. The excess kurtosis (kurtosis beyond that of a Gaussian random variable) is given by,

$$ku_{ex}(x) = ku(x) - 3. \quad (13)$$

Random variables with positive excess kurtosis are called *supergaussian* or *leptokurtic*, whereas random variables with negative excess kurtosis are called *subgaussian* or *platykurtic*. Based on the non-Gaussianity assertion, it can be loosely stated that independent components produced by the FastICA algorithm [12] and Hebbian-like learning algorithm [11], with their associated objective functions, have excess kurtosis values that optimally differ from zero. It should be noted that, in practice, the algorithms replace kurtosis with robust functions that contain information from higher order statistics [3].

3.3. Relations for the undamped system

Taking kurtosis as a simple measure of independence, a relationship between modal responses of the undamped system and independent components can be arrived at by investigating the kurtosis of undamped modal responses. From Eq. (6) it can be seen that when the damping ratio, ζ_j , is zero, the modal responses become, $q_j(t) = \alpha_j \sin(\omega_{nj}t + \theta_j)$. Using the PDF of a single sinusoid, $q(t)$, with amplitude α ,

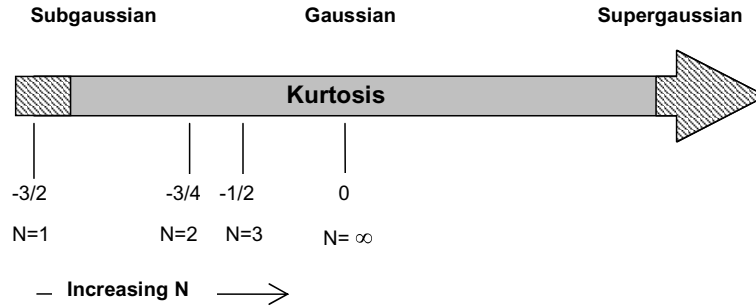
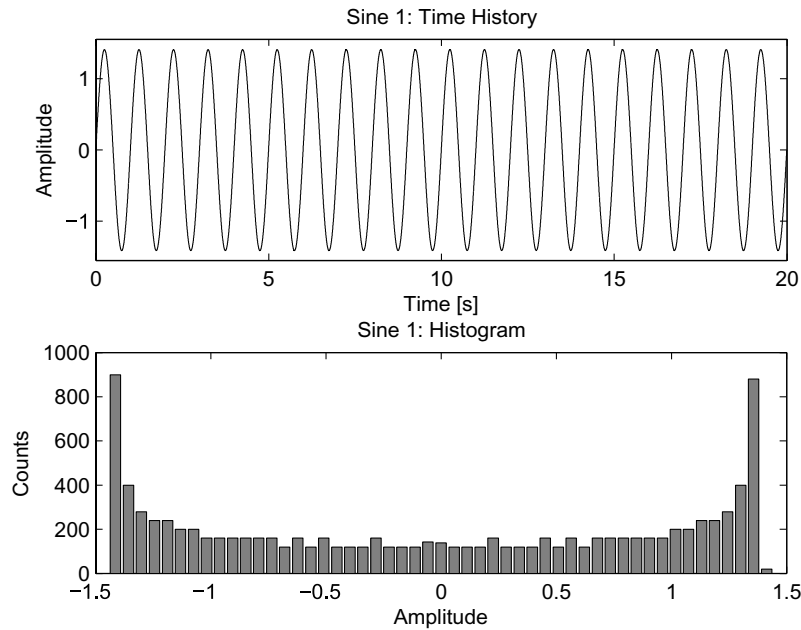
$$P(q|\alpha) = \frac{1}{\pi\sqrt{\alpha^2 - q^2}}, \quad (14)$$

it can easily be shown that the kurtosis and excess kurtosis of a sinusoid is given by,

$$\begin{aligned} ku &= 1.5, \\ ku_{ex} &= -1.5. \end{aligned} \quad (15)$$

Note that both the PDF and kurtosis are independent of frequency, ω_{nj} , and phase shift, θ_j . Clearly a sinusoid is subgaussian.

Now the task is to determine the excess kurtosis of a linear mixture of sinusoids and see if the excess kurtosis is more Gaussian (closer to zero) or less Gaussian (has absolute value greater than 1.5) than the individual sinusoid. In order to simplify the problem, a restriction is placed on the sinusoids. The sinusoids are assumed to be independent and identically distributed (i.i.d.) signals. This means that the sinusoids have the same amplitude and have *incommensurate* frequencies (irrational frequency ratios) [9]. It has been shown in reference [9] that the summation of N i.i.d. sinusoids have an excess kurtosis, ku_{exN} , given by,

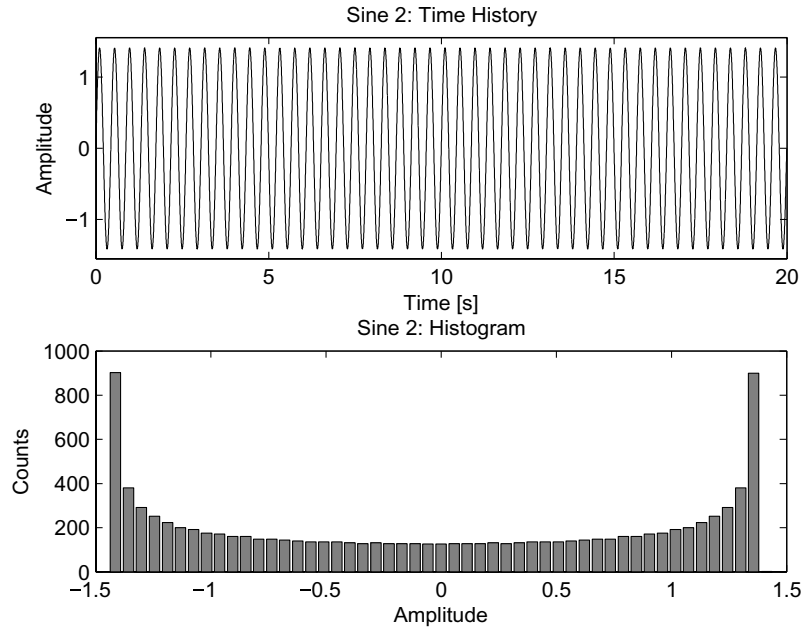
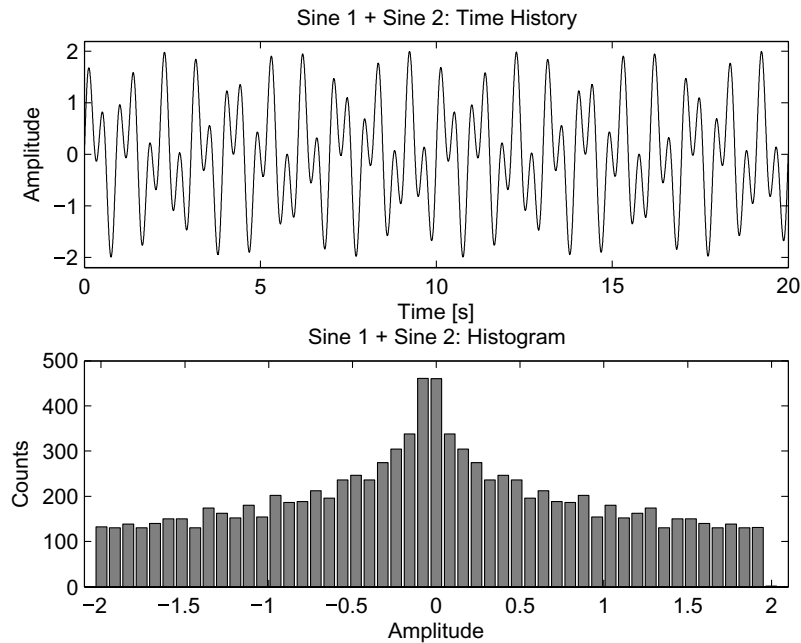
Fig. 1. Kurtosis of a mixture of N i.i.d. sinusoids.Fig. 2. 1.0 Hz sinusoid and histogram – $ku_{ex} = -1.5$.

$$ku_{exN} = \frac{ku_{ex1}}{N}, \quad (16)$$

where ku_{ex1} is the excess kurtosis of a single sinusoid, -1.5 . It can be observed from Eq. (16) that the excess kurtosis tends toward zero as the number of sinusoids is increased. In other words, the distribution of a single sinusoid is subgaussian and the distribution tends toward a Gaussian distribution as more sinusoids are added together. This relationship is depicted in Fig. 1. In the figure, the cross-hatched areas (far from Gaussian) represent independence.

This relationship can be illustrated by example. Sine waves, $x_1(t)$ and $x_2(t)$ were generated at 1.0 and 2.3 Hz, respectively, for a duration of 20 seconds. Both sinusoids were scaled to unit variance then added together and the sum was scaled to unit variance. A plot of the sinusoids and their corresponding histogram can be found in Figs 2 and 3. It can be seen that the shape is that of the PDF in Eq. (14). The time history and histogram for the sum, $x_1(t) + x_2(t)$ is shown in Fig. 4. From the histograms, it is apparent that the excess kurtosis of the mixed sinusoids, -0.75 , is closer to zero than that of either of the originals, -1.5 .

It can be demonstrated by example that mixtures of independent sinusoids that are not of the same amplitude also are more Gaussian than are the individual monotone sinusoids. Some theoretical results can also be found where the amplitudes of sinusoids in a mixture were allowed to vary randomly [13]. Results showed a similar relationship as Eq. (16), but the formulation is more complicated. In addition, the assumption of incommensurate frequencies can be relaxed somewhat in practice as discussed in [9].

Fig. 3. 2.3 Hz sinusoid and histogram $-ku_{ex} = -1.5$.Fig. 4. Summed sinusoids and histogram $-ku_{ex} = -0.75$.

As a result, the excess kurtosis of a single sinusoid is smaller than that of a linear mixture of two or more independent sinusoids. In other words, sinusoidal components of data are optimally subgaussian with respect to linear mixtures of modal responses. As a result, modal responses are more independent than are their mixtures. It can be concluded that the FastICA and Hebbian-like learning algorithms, with proposed objective functions, may be applied in a straightforward manner to *undamped* free-vibration data, to estimate the sinusoidal modal responses.

Table 1
Damped sinusoid parameters

Component #	Frequency [Hz]	Damping [%]	Kurtosis
1	0.7	2.3	1.55
2	1.1	1.7	2.28
3	1.8	1.9	6.65
4	2.4	1.1	4.46
5	3.5	0.8	4.91

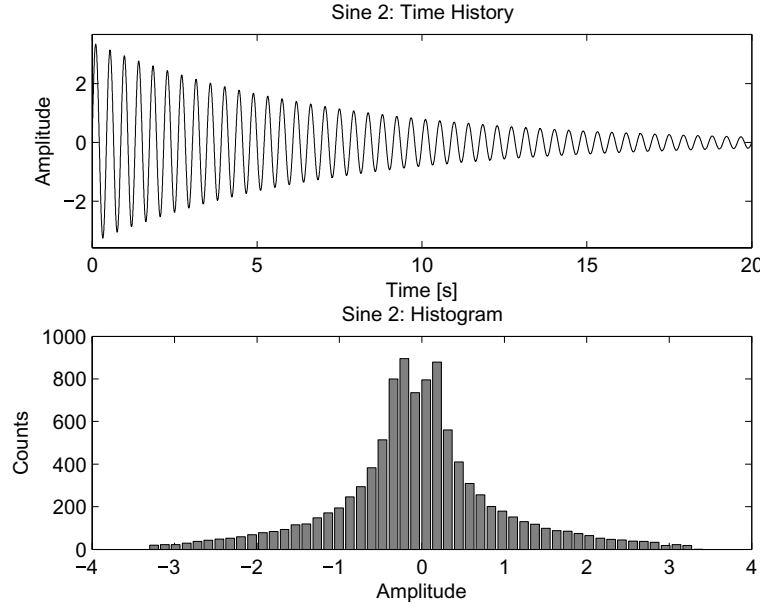


Fig. 5. Damped sinusoid $x_2(t)$ (top) and its histogram (bottom) – $k_{uex} = 1.36$.

3.4. Generalization for the diagonally-damping system

In the preceding discussion, only sinusoids were considered. The case gets more complicated when damping enters into the picture. Here we will focus on sampled data of finite length. Recall from Eq. (6) that modal responses of the damped system are damped (decaying) sinusoids. A damped sinusoid is supergaussian, as shown in Fig. 5. The figure shows the sinusoid, $x_2(t)$, with a damping level $\zeta = 0.01$ included. The damped sinusoid, having excess kurtosis of 1.36, is seen to be supergaussian. It can be demonstrated by simulation that linear mixtures of damped monotone sinusoids are even more supergaussian (have higher kurtosis).

A simple simulation is used for clarity in illustration. Five damped sinusoids with unit variance were created with parameters summarized in Table 1. The damped monotone components were mixed by multiplying by a random matrix whose elements were drawn from a Gaussian distribution, $\mathbf{A} = [0.642, -1.270, 0.961, -0.623, -0.233]$. The resulting mixed damped sinusoid time history and FFT magnitude are shown in Fig. 6. The excess kurtosis of the mixed sinusoid, 8.64, is much greater than the minimum excess kurtosis of the monotone damped sinusoids, 1.55. Ten thousand random mixtures were created in a similar manner. The minimum excess kurtosis of any of the mixtures turned out to be 1.62, which is still slightly greater than 1.55. The corresponding random mixing matrix was, $\mathbf{A} = [1.472, 0.059, 0.079, 0.020, 0.061]$. Clearly, this mixture is composed mostly of the first damped monotone component with small contributions from the others. The mixed damped sinusoid time history and FFT magnitude are shown in Fig. 7. This simulation has illustrated that the mixture with the minimum kurtosis is equal to the monotone component with minimum kurtosis. If the first component is removed and the simulation repeated with the remaining four components (2–5), it is found that the second component has the minimum kurtosis among all linear mixtures of the four components.

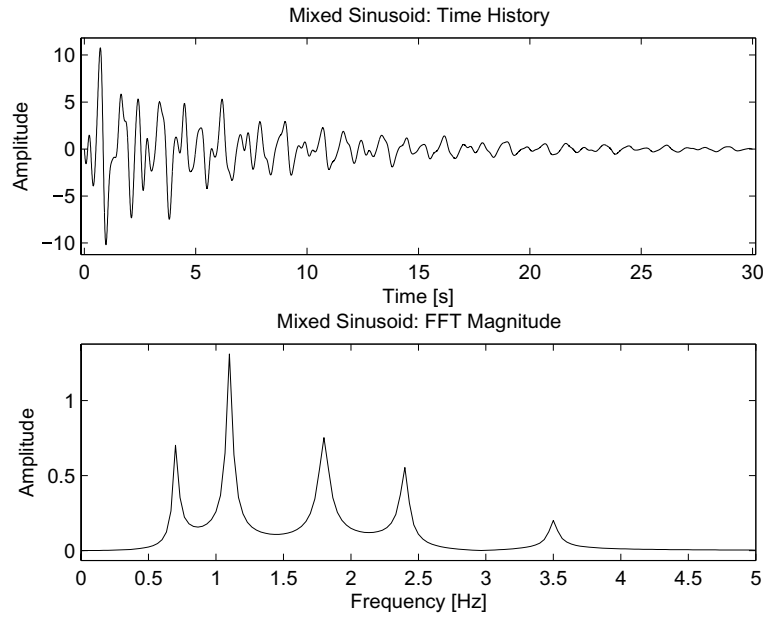


Fig. 6. Mixed damped sinusoid (top) and FFT magnitude (bottom) – $k u_{ex} = 8.64$.

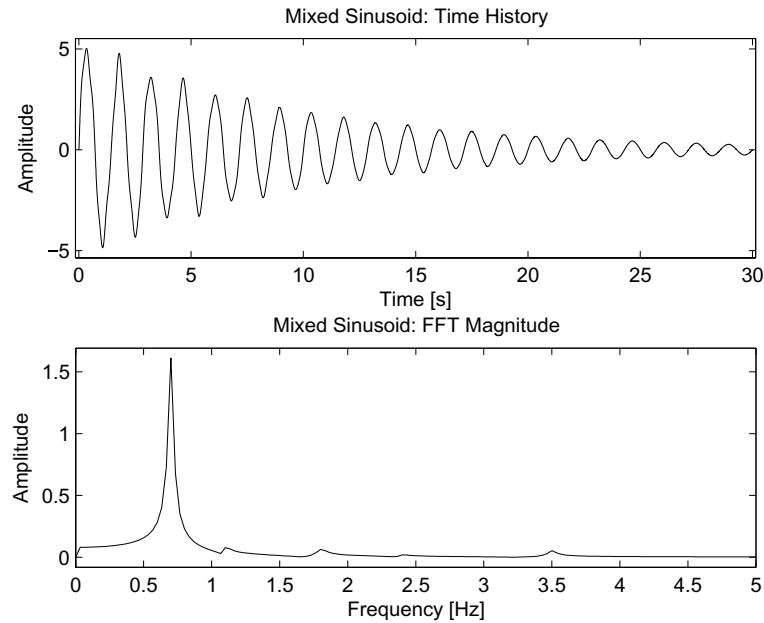


Fig. 7. Mixed damped sinusoid (top) and FFT magnitude (bottom) – $k u_{ex} = 1.62$.

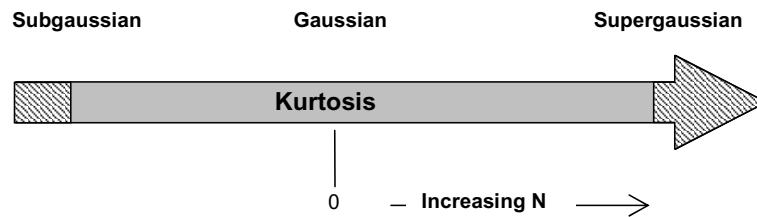


Fig. 8. Kurtosis of a mixture of N damped sinusoids.

This simulation may be expanded by randomly varying frequencies and damping levels in addition the mixing matrix. In some cases, the linear mixture with minimum kurtosis does not exactly correspond to the damped monotone component with the minimum kurtosis. However, the former is found to be a good estimate for the latter. Because monotone damped sinusoids have minimum kurtosis with respect to linear mixtures, the modal responses are still optimally subgaussian (have as low a kurtosis as possible). This relationship is depicted in Fig. 8.

Considering the characteristics of damped and undamped sinusoids, it can be deduced that undamped modal responses closely correspond to independent components, damped modal responses do not. As a result, the FastICA and Hebbian-like learning algorithms, with previously proposed objective functions, cannot be applied directly to yield modal responses in general. In the special case of light damping, the algorithms may be applied on a small portion (analysis window) of the free-decay data to yield modal responses as in reference [7]. This is due to the fact that the amplitude envelope doesn't change much in the analysis window, so the independent components are nearly sinusoidal.

Using the fact that the modal responses are optimally subgaussian (have lower kurtosis compared to their linear mixtures), the Hebbian-like learning algorithm has been modified in reference [10] in order to yield modal responses in both the diagonally-damped and undamped cases. The modification seeks optimally subgaussian solutions rather than solutions that either optimally subgaussian or supergaussian. Examples are provided in the reference and further discussion and evaluation of the method is given in [9].

4. Summary

In this paper, the relationship between independent components and free-vibration modal responses was discussed. Theoretical arguments were presented for responses of undamped systems and arguments are extended to damped responses. It was deduced that undamped modal responses closely correspond to independent components, while damped modal responses do not. However, in both cases, modal responses are optimally subgaussian (have lowest kurtosis among linear mixtures). The results herein provide further validation of the modified Hebbian-like algorithm presented in [10].

References

- [1] J.F. Cardoso, Blind signal separation: statistical principles, *Proceedings of the IEEE* **9**(10) (1998), 2009–2025.
- [2] J.F. Cardoso, High-order contrasts for independent component analysis, *Neural Computation* **11**(1) (1999), 157–192.
- [3] A. Hyvarinen and E. Oja, Independent component analysis: algorithms and applications, *Neural Networks* **13** (2000), 411–430.
- [4] S. Chauhan, R.J. Allemang, R. Martell and D.L. Brown, Application of independent component analysis and blind source separation techniques to operational modal analysis, *Proceedings of the 25th International Modal Analysis Conference*, Orlando, FL, February 2007.
- [5] F. Poncelet et al., Output-only modal analysis using blind source separation techniques, *Mechanical Systems and Signal Processing* **21**(6) (2007), 2335–2358.
- [6] W. Zhou and D. Chelidze, Blind Source Separation Based Vibration Mode Identification, *Proceedings of the 25th International Modal Analysis Conference*, Orlando, FL, February 2007.
- [7] G. Kerschen, F. Poncelet and J.C. Golinval, Physical interpretation of independent component analysis in structural dynamics, *Mechanical Systems and Signal Processing* **21**(4) (2007), 1561–1575.
- [8] S.I. McNeill and D.C. Zimmerman, A framework for blind modal identification using joint approximate diagonalization, *Mechanical Systems and Signal Processing* (2008), doi:10.1016/j.ymssp.2008.01.010.

- [9] S. McNeill, *Modal Identification using Blind Source Separation Techniques*, Ph.D. Dissertation, The Department Mechanical Engineering, University of Houston, 2007.
- [10] S. McNeill, S. Shen McNeill and D. Zimmerman, Use of ICA techniques to decompose free vibration data, *Proceedings of the 25th International Modal Analysis Conference*, Orlando, FL, February 2007.
- [11] A. Hyvärinen and E. Oja, Independent component analysis by general nonlinear Hebbian-like learning rules, *Signal Processing* **64**(3) (1998), 301–313.
- [12] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, *IEEE Transactions on Neural Networks* **10**(3) (1999), 626–634.
- [13] R. Barakat and J. Cole, Statistical properties of N random sinusoidal waves in additive Gaussian noise, *Journal of Sound and Vibration* **62**(3) (1979), 365–377.

