Research on Estimation Method of Mechanical Fault Source Number Based on VbHMM

Yajing ZHU , Zhinong LI , Jingzhi TU
Key Laboratory of Nondestructive Testing Ministry of Education
Nanchang Hangkong University
Nanchang, 330063, China
Email: lizhinong@tsinghua.org.cn

***Abstract*—The traditional source number estimation method must ensure that the signal sources are independent and noise-free interference. Based on the above deficiency in the traditional BSS method, combining variational Bayesian hidden Markov model (VbHMM) and autocorrelation determination (ARD), a estimation method of mechanical fault sources number based on VbHMM is proposed. In the proposed method, after the Bayesian networks are introduced, the Markov models (HMM) is used to capture the characteristics of a series of time-related time series information in the dynamic and nonlinear signals. The optimal number of hidden sources in the non-stationary signal is deduced by the unique model comparison function of Bayesian inference and autocorrelation determination (ARD). Simulation and experimental results verify the effectiveness of the proposed method.**

Keywords—variational Bayesian hidden Markov mode; autocorrelation determination (ARD); blind source separation; Bayesian network; estimation of signal sources

I. INTRODUCTION

In the field of mechanical condition monitoring and fault diagnosis, the observed signals are often a mixture of multiple source signals. Before blind separation of these observation signals, in order to ensure the source separation, it is necessary to estimate the number of vibration source signals. The traditional source number estimation method is mainly based on principal component analysis and singular value decomposition [1, 2]. The number of uncorrelated source or independent sources is determined by calculating the number of non-zero eigenvalues or non-zero singular values of the correlation matrix of observed signal. The traditional BSS algorithm requires to meet the condition that the number of sensors is greater than the number of sources, however, in real engineering practice, this condition is often not satisfied.

In recent years, some estimation methods of signal source number have been proposed. Ref.[3] proposed an EMD-SVD-BIC method to estimate the number of mechanical vibration sources in the undertermine convolutional mixing source separation. In Ref.[4, 5], according to ratio of power spectral density functions ofobserved signals, the source number can be accurately estimated by the comparison of the column vectors of the observed signal's power spectral density matrix. Ref.[6] proposed an estimation method of source number based on non-negative matrix factorization. However, the above estimation methods of source number have some following deficiency.

1.Signal sources is independent of each other, else the upper bound of the source number can be only obtained. However the upper bound of the source number is not equal to the number of signal sources.

2.The source number estimation is easily affected by noise, especially when the signal-to-noise ratio is low, a reasonable source number is hardly estimated in the traditional estimation methods of source number.

Based on the above deficiency in the traditional estimation methods of source number. In Ref.[7, 8], combining variations Bayesian independent component analysis (VbICA) and autocorrelation determination (ARD), a estimation method of mechanical source number based on Bayesian independent component analysis was proposed. In the proposed method, the optimal number of mechanical sources was determined by the evidence comparison of different models. However the VbICA is still a static ICA algorithm. Here, the so-called 'static' contains two meanings.

1.The mixing system remains constant during the acquisition of observation data.

2.The number of independent sources and and its statistical characteristics remain stable. However, in practical engineering applications, various dynamic random signals with probabilistic characteristics occur in time sequence [9]. Obviously, the estimation of source number is undoubtedly more reliable by making full use of these dynamic temporal information of the observation signals. Based on the above shortcomings, it is necessary to explore source estimation method of mechanical fault in dynamic ICA.

HMM is a dynamic model, which can collect state information of signal sources in the process of potential data generation. A sequence of hidden variables is used to simulate the variation in the dynamic behavior of the system [10]. In a first-order Markov process, an unobservable potential process moves from one state to another state. The

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

current state only depends on the previous state, and captures the higher order sequence information in signal source[11]. Therefore, HMM is very suitable for time series modeling of dynamic processes in non-stationary time-varying signals and has powerful ability of pattern classification. Variational Bayesian hidden Markov model (VbHMM) can be constructed by combining HMM with Gauss hybrid model, and can provide a good dynamic source model for source estimation of non-stationary time-varying signals. Here, combining the variational Bayesian hidden Markov model (VbHMM) and autocorrelation determination (ARD), a estimation method of source number is proposed. In the proposed method, the hidden Markov models (HMM) is used to capture a series of sequence information in the dynamic and nonlinear signal. The optimal number of hidden sources in non-stationary signals is deduced by the unique model comparison function of Bayesian inference and autocorrelation determination. The simulation and experiment results verified the effectiveness of the proposed method.

II. ESTABLISHMENT OF MODEL

A. *Establishment of Variational Bayesian Hidden Markov Model (VbHMM)*

In the Bayesian network, *Bayesian* inference calculates the posterior probability density distribution of unknown parameters according to condition and prior probabilities. Independent component analysis is used to learn hidden variables (source signals) of models under the framework of objective function. In the learning process, Markov model is introduced into the independent component analysis source model (Gaussian mixture model), and this two models are consistent. Therefore a variable Bayesian hidden Markov model (VbHMM) is established.

In this model, the w represents the set of parameters of the model $\mathrm{W}=\{\mathrm{s},\mathrm{q},\Theta\}$,where s is an analog source parameter and q is a mixed Gaussian component parameter, $\Theta=\{\mathrm{A},\Lambda,\mathrm{R},\pi,\theta\}$ is a ICA model parameter. A, Λ, R, π, θ represents model mixing matrix, background noise precision matrix, state transition probability matrix, initial probability matrix, and model parameter respectively. Where $\theta=\{\pi,\mathrm{R},[\theta_1,\cdots\theta_m]\}$ contains some model parameters of hidden Markov with m states, and is learned by Bayesian methods. However, a strict Bayesian algorithm is too cumbersome, therefore a tractable variational Bayesian approximation is proposed to improve computational efficiency. Here, the maximization the objective function is also the maximum negative free energy[12].

$$F(\mathrm{X}|\mathrm{M})=\left\langle \log p(\mathrm{X},\mathrm{W}|\mathrm{M})\right\rangle_{p'(\mathrm{W})}+H[p'(\mathrm{W})] \tag{1}$$

Where x is an observation signal, $p'(\mathrm{W})$ represent the best approximation valuefor representing real posterior probability. $H[P'(\mathrm{W})]$ represent the entropy of $p'(\mathrm{W})$. M represents all the reliability and assumptions of the model. $p(\mathrm{x}|\mathrm{W},\mathrm{M})$ represents data likelihood ratio.

$p'(\mathrm{W})$ can be written into the following factorization form

$$p'(\mathrm{W})=\prod_i p'(\mathrm{W}_i|\mathrm{M}) \tag{2}$$

Thus the computational problem of $p(\mathrm{X}|\mathrm{M})$ is avoided. Ref.[13] proves that negative free energy $F[\mathrm{X}|\mathrm{M}]$ can obtain the local maximum when the posterior $p'(\mathrm{W})$ probability is close to the true posterior probability $p(\mathrm{X}|\mathrm{M})$. Therefore, the parameter distribution of each model can be obtained by maximizing the negative free energy. According to the determined model parameters, the belief posterior probability of each model is calculated as

$$p(\mathrm{X}^t|\mathrm{M})=\int p(\mathrm{X}^t|\mathrm{W},|\mathrm{M})p(\mathrm{W}|\mathrm{M})d\mathrm{W} \tag{3}$$

Where $p(\mathrm{x}^t|\mathrm{M})$ is the normative factor, x^t represents M-Dimensional Observation Signals.

When the source is unknown, the candidate model series $\mathrm{M}=\{\mathrm{M}_1,\cdots,\mathrm{M}_i\}$ is learned respectively by formula (3), thus the belief posterior probability of each model is obtained. The largest posterior probability is selected as the optimal source model.

After the source model is determined by the above method, a reasonable candidate model sequence is determined by autocorrelation determination method.

B. *Combination of Autocorrelation ARD and Variational Bayesian Hidden VbHMM Units*

On the basis of the source model learned by variational Bayesian, a *N*-by-*L* mixture matrix A is defined as a Gaussian distribution product with zero mean and the precision α_{ij} .

$$p(\mathrm{A})=\prod_{i=1}^{N}\prod_{j=1}^{L}N(\mathrm{A}_{ij}|0,\alpha_{ij}) \tag{4}$$

In the instantaneous mixed blind source separation $X=\mathrm{A}S$, the number of observed signal is determined by the number of the rows of mixing matrix A , and the number of source signal S is determined by the columns of mixing matrix A . So each source corresponds to each column in the column of mixing matrix. Thus Eq.(4) is transformed into

$$p(\mathrm{A}|\alpha)=\prod_{i=1}^{N}\prod_{j=1}^{L}N(\mathrm{A}_{ij}|0,\alpha_j) \tag{5}$$

In Eq(5), the previous control parameter α_{ij} is replaced by precision a_j . Each precision is connected with one of L column of mixing matrix A . Thus the threshold vector $\alpha=(\alpha_1,\cdots,\alpha_L)$ is used to control the mixing matrix A . Therefore The value of element in threshold α determines the dimension of mixing matrix , i.e. source number.

For dynamic non-stationary time-varying signals, the combination of ARD and VbHMM can compensate the worse robustness of ARD in dynamic ICA, and solve the selection problem of signal source model. The basic idea is that firstly, the number of possible signal sources is initially estimated by ARD. Then a reasonable candidate model sequence is selected according to the estimated source number. Finally, the VbHMM method is used to separate the observer signal, and obtains the number of hidden sources exactly.

III. ESTIMATION OF SOURCE NUMBER

The learning of variational Bayesian begins from the initialization of model. Since the threshold vector of the mixed matrix is added, the model parameter vector becomes $\mathrm{W}=\{\mathrm{s},\mathrm{q},\Theta,\alpha\}$. In the given hybrid source model combined with HMM and MOG, each model contains n Gauss components, and correspond to each source signal separately so an ICA model with L source there's an L-Gauss mixture model $\{n_1,n_2,\cdots n_L\}$, there are m_i components in each source. The distribution of source signal s is expressed as follows

$$p(\mathrm{s}^t|\theta)=\prod_{i=1}^{L}\sum_{t=1}^{m_i}\pi_{i,q_i}\mathrm{N}(\mathrm{s}_i^t;u_{i,q_i},\beta_{i,q_i}) \tag{6}$$

where π_i represents the proportion of mixture, μ_i,β_i represents the expected value and precision value of the ith source, the mixed proportional area $\pi_{i,q_i}=p(q_i^t=q_i|\pi_i)$ represents the prior probability of the ith source component q_i, β_{i,q_i} represents the precision value of the ith source component q_i, q_i^t represents the ith source component of s_i^t, $\{q_i=1,\cdots,q_i=m_i\}$.

In this process, the first problem to be solved is how to determine the model of signal source according to the observed data. Here the solution of this problem can be solved by calculating the posterior probability of signal source.

Let M indicates the model of signal source with estimated source number i. All the model parameters are represented by the vector W . On the premise of known observation signal X^t, the posterior probability of the parameter W can be expressed as

$$p(\mathrm{W}|\mathrm{X}^t,\mathrm{M})=\frac{p(\mathrm{X}^t|\mathrm{W},\mathrm{M})p(\mathrm{W}|\mathrm{M})}{p(\mathrm{X}^t|\mathrm{M})} \tag{7}$$

where $p(\mathrm{W}|\mathrm{M})$ represents the prior probability of the model M , $p(\mathrm{X}^t|\mathrm{M})$ represents the evidence of model, and $p(\mathrm{X}^t|\mathrm{W},\mathrm{M})$ represents the likelihood. According to Eq. (7), the posterior probability of source is obtained by the extremum method iteratively. Here, the maximization of objective function is also the maximum negative free energy

$$F(\mathrm{s}|\mathrm{M})=\left\langle\log p(\mathrm{s},\mathrm{q},\Theta|\mathrm{M})\right\rangle_{p'(\mathrm{q},\Theta)}+H[p'(\mathrm{q},\Theta)] \tag{8}$$

Where the parameters of probability distribution function for C observations are set to $\theta=\{\theta_1,\cdots,\theta_C\}$,for model M, the overall parameters are set to $\Theta=\{\pi,\mathrm{R},[\theta_1,\cdots,\theta_C]\}$. $\{\mathrm{q}\}=\{\mathrm{q}_1,\cdots,\mathrm{q}_{C^T}\}$ represent space representing all state paths. If $\Theta=\{\pi,\mathrm{R},\theta\}$, there are the following factors

$$p'(\mathrm{q},\Theta)=p'(\mathrm{q})p'(\pi)p'(\mathrm{R})p'(\mu)p'(\beta) \tag{9}$$

where $\mu=\{\mu_1,\cdots,\mu_C\}$, $\beta=\{\beta_1,\cdots,\beta_C\}$ represents the observed model parameters. $p(\pi)$ is the prior distribution of the proportion of mixture, which satisfies with Dirichlet boundary condition. $p(\pi)$ represents the state transition probability distribution. $p(\mathrm{R})$ represents the expected prior distribution, which is defined as the Gaussian distribution function. $p(\beta)$ is the prior distribution of precision, which is defined as the Gamma distribution:

$$p(\pi)=D(\pi;\lambda_0) \tag{10}$$

$$p(\mathrm{R})=\prod_{c=1}^{C}D(\mathrm{R}_c;\iota_{c,1:C}) \tag{11}$$

$$p(\mu)=\prod_{c}^{C}N(\mu_c;m_0,\tau_0) \tag{12}$$

$$p(\beta)=\prod_{c}^{C}g(\beta_c;b_0,c_0) \tag{13}$$

where R_c is the cth rows of R , Eq.(9)~Eq.(13) is substituted into Eq. (8), the posterior probability of each state number is obtained as follows [16].

$$p'(\pi)=D(\pi,\hat{\lambda}_{1:C}) \tag{14}$$

$$p'(\mathrm{R})=\prod_{c=1}^{C}D(\mathrm{R}_c;\hat{\iota}_{c,1:C}) \tag{15}$$

$$p'(\mathrm{q})=\frac{1}{Z_q}\left[\widetilde{\pi}_{q^1}\prod_{t=2}^{T}\widetilde{r}_{q^{t-1}q^t}\prod_{t=1}^{T}\widetilde{p}_{q^t}\right] \tag{16}$$

$$p'(\mu)=\prod_{c}^{C}N(\mu_c;\hat{m}_c,\hat{\tau}_c) \tag{17}$$

$$p'(\beta)=\prod_{c}^{C}g(\beta_c;\hat{b}_c,\hat{c}_c) \tag{18}$$

The prior probability distribution of threshold parameter in the autocorrelation measurement (ARD) is defined as

$$p(\alpha)=\prod_{j=1}^{L}G(\alpha_j,b_{a_j},c_{\alpha_j}) \tag{19}$$

Where $G(\bullet)$ represents the product of the Gamma distribution, next, the partial derivative of $p'(\mathrm{W}|\theta)$ to every parameter is obtained, and the corressponding extremum is obtained.

$$p'(\mathrm{W}|\theta)=\frac{1}{Z_i}exp\left[\left\langle \lg(\mathrm{X},\mathrm{W}|\mathrm{M})\right\rangle_{\prod_{i\neq j}p'(W_{j\neq i}|\theta_{j\neq i})}\right] \tag{20}$$

where Z_i is the canonical factors. For the instantaneous blind separation, in Eq. (2) , the approximation distribution is described as follows

$$p'(\mathrm{W})=p'(\mathrm{A})\,p'(\Lambda)\,p(\alpha)\,p'(\mathrm{s}|\mathrm{q})\,p'(\mathrm{q})\,p'(\theta) \tag{21}$$

where $p'(\theta)=p'(\mathrm{R})\,p'(\pi)\,p'(\mu)\,p'(\beta)$, and $\mathrm{p}(\alpha)$ is the prior probability distribution of the precision vector of the mixed matrix.

Eq. (20) and (21) are substituted into Eq. (8), and considering Eq. (19), the prior distribution is defined as

$$p'(\mathrm{A})=\prod_{i=1}^{N}\prod_{j=1}^{L}N(A_{ij};\hat{m}_{A_{ij}},\hat{\alpha}_{ij}) \tag{22}$$

$$p'(\alpha)=\prod_{j=1}^{L}G(\alpha_j,\hat{b}_{a_j},\hat{c}_{\alpha_j}) \tag{23}$$

The posterior probability distribution of the threshold α, which controls the dimension of the mixed matrix, is obtained by the above formula.

By monitoring the precision α_j of each signal, the optimal number of hidden sources can be determined in the observation matrix.

IV. SIMULATION STUDY

In order to verify the effectiveness of the proposed method, two simulated sources are constructed. The first simulated source signal s_1 is a gradient signal. Obviously signal s_1 is a non-stationary time-varying signal. The second simulated source signal s_2 is a non-stationary signal, which is shown in Eq. (24):

$$s_2=(1+\sin(200\,\pi t))\sin(20\,\pi t) \tag{24}$$

The sampling frequency is $f_s=5000\,Hz$, the sampling length is 1000. The waveform of source signals is as shown in Fig.1.

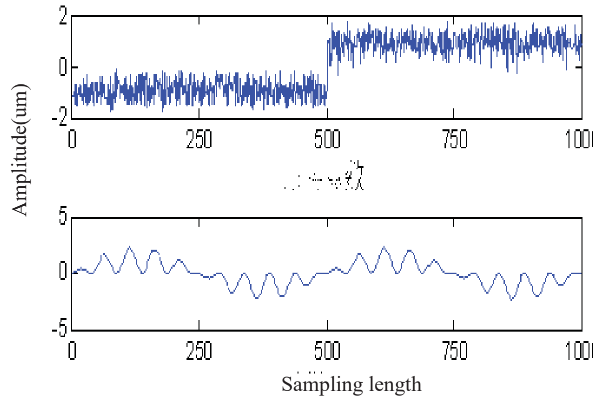


Fig. 1.Two simulated Source signals

The two source signals are randomly mixed according to $\mathrm{x}^t=\mathrm{As}^t$, and 15% Gaussian noise is added. Here a random matrix is selected to obtain four observation signals, which are as shown in Fig. 2.

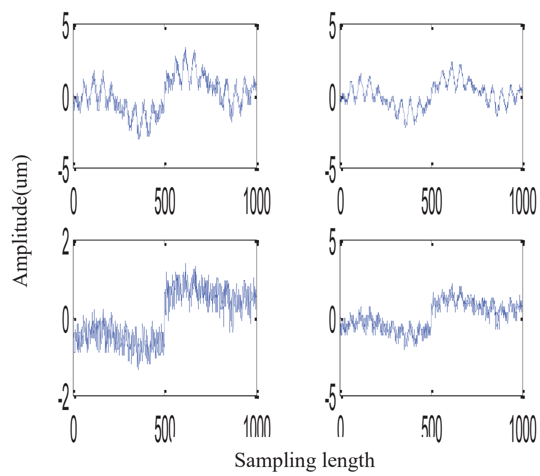


Fig. 2.Four-way observation signal

The source number in mixed signals is inferred by autocorrelation determination, and the estimated source number is 2. In order to further verify the correctness of the algorithm, the candidate sequences are selected as different models with different source numbers {i=1,2,3,4} respectively, and the separation results are shown in Fig.3. From Fig.3, when the actual source number is more than that of simulated sources, the separation results can not reflect all the information in the observed signals. However, the waveform of Source signals is retained. The redundant sources are effectively suppressed, and its value is compressed near zero.

The evidence of model is shown in Fig.4. From Fig.4, when the evidence of model is 1, 2, 3, 4, the maximum of negative free energy is 2750, 4250, 4100, 4000 respectively. When the evidence of model reaches the maximum at the correct source location (i.e. the source number is 2), which reflects the actual source number. Therefore, the source number can be initially estimated by the ARD, then the number of hidden sources can be determined exactly by the comparison of model evidence.

V. EXPERIMENTAL RESEARCH

In order to further verify the effectiveness of the proposed method, the VbHMM has been successfully completed in the motor-gearbox coupling experiment. The obtained multichannel mechanical vibration observation is used to verify the effectiveness of the proposed method.

The motor-gearbox coupling experimental device is shown in Fig.5. Three motor vibration signals s_1, s_2 and s_3 are shown as Fig.6. Six observation signals are shown as Fig.7. Here the sampling frequency is 40 kHz the sampling length is 2000.

Firstly, the six-way observation signals are used to learn by the ARD Model. Three estimated sources are obtained. Then the candidate sequences are selected as different model with different source numbers $\{i$=1,2,3,4,5$\}$ respectively, and the separation results are shown in Fig.7. The obtained evidence of different models is shown in Fig.8. From Fig.8, The evidence of model reaches the maximum at the correct

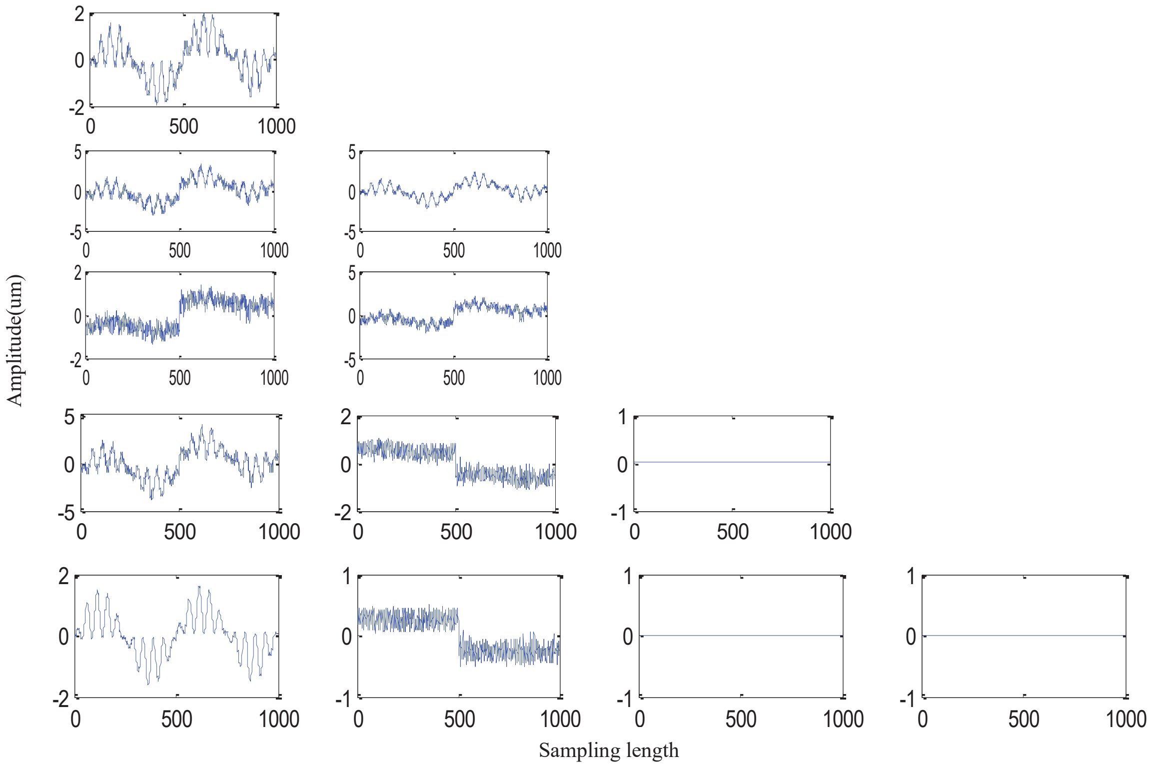


Fig. 3.Estimate signals in the sources are 1, 2, 3, and 4 respectively

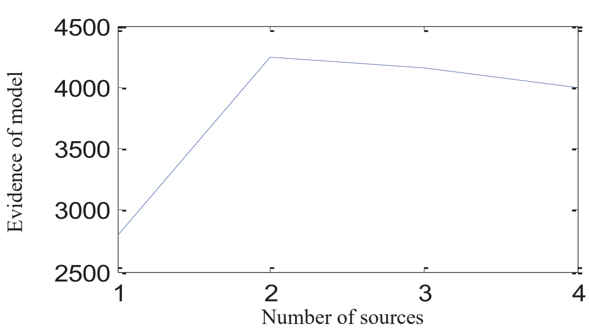
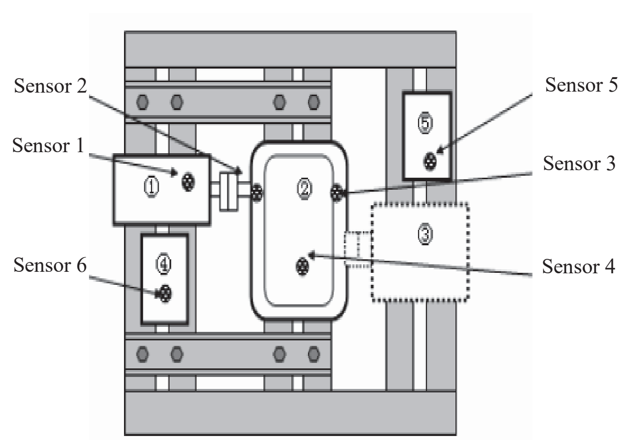


Fig. 4.Evidence of model vs number of sources



①large motor ② gearbox ③ preload device
④ disturb motor 1 ⑤ disturb motor 2

Fig. 5.motor-gearbox coupling experiment

source, i.e. The source number is 3. The estimated signals is is shown in Fig.9.

From Fig.9 and Fig.6, the source signals can be reserved. Therefore in determination of number of dynamic non-stationary time-varying sources, combining VbHMM and ARD, the number of hidden sources can be determined exactly.

VI. CONCLUSIONS

In the separation of mechanical sources based on BSS, the determination of the source number is still an unsolved problem. In this paper, combining the variational Bayesian hidden Markov model (VbHMM) and the automatic relevance determination (ARD), a estimation method of number of mechanical fault sources based on VbHMM is proposed. In the proposed method, after the theory and inference mechanism of Bayesian networks is introduced, the hidden Markov model (HMM) is used to capture the characteristics of time series information in dynamic and nonlinear signals, and introduced into model of source in the variational Bayesian independent variable analysis (VbICA). Then, the optimal number of hidden sources in the non-stationary signals is deduced by the model comparison function in Bayesian inference, and combining autocorrelation determination. The simulation results show that the proposed method is very effective, and can obtain good separation performance in the mixing of dynamic non-stationary time-varying signals. Finally the proposed method has been successfully applied into the motor-gearbox coupling experiment. The experimental results further verify the effectiveness of the proposed method. The research in this paper provides an effective method for dynamic non-stationary time-varying signal mixing blind separation and estimation of source number.

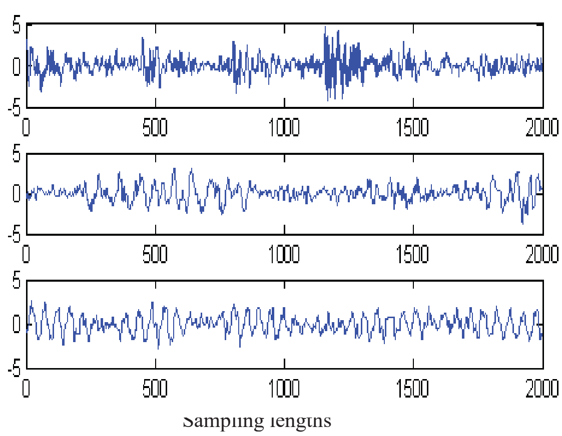


Fig. 6.Source signal waveform of three motors

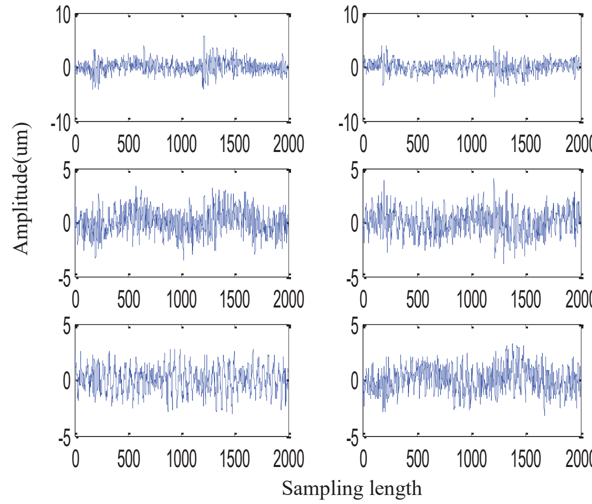


Fig. 7.Six observation signals

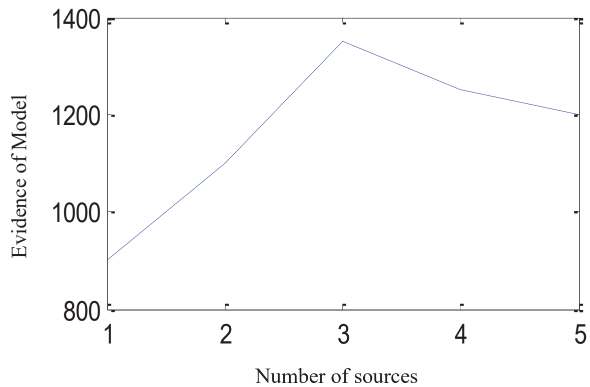


Fig. 8.Model reliability change graph

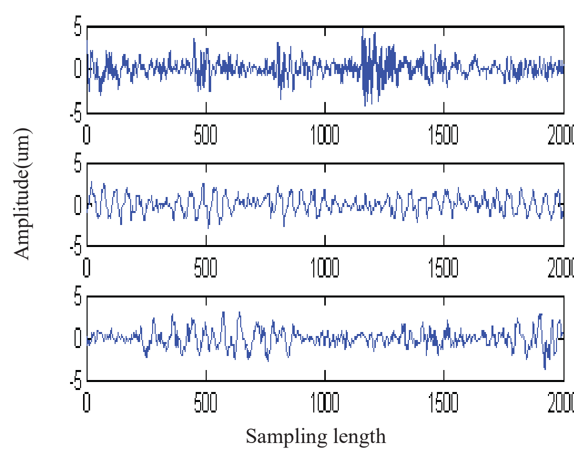


Fig. 9.Estimated source signals (the number of sources is 3)

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (51675258, 51261024), the Science and Technology Research Project of Education Department of Jiangxi Province (No. GJJ150699) and Research Project of State Key Laboratory of Mechanical System and Vibration (No. MSV201914).

REFERENCES

[1] Jiao Weidong, Yang Shixi, Wu Zhaotong. Blind Separation of Rotating Machinery Sources Based on Source Estimation[J].China Mechanical Engineering, 2003, 14 (14): 1184-1187.

[2] Zhang Xinhua, Zhang Anqing, Sun Jianping. A Blind Estimation Method fo`r the Number of Signal Sources[J]. Systems Engineering and Electronics, 2001, 29 (9) : 9-11.

[3] Murty S, Kompella. A technique to determine the number of incoherent sources to the response of a system[J]. Mechanical System and Signal Processing,1994, 8 (4) : 363-380.

[4] YE Hong-xian, YANG Shi-xi, YANG Jiang-xin. Estimation method of mechanical vibration source number based on EMD-SVD-BIC[J]. Journal of Vibration, Testing and Diagnosis, 2010, 30(3): 330-334.

[5] LI Ning, SHI Tielin. Estimation of Blind Signal Sources Based on Power Spectral Density[J]. Data Acquisition and Processing, 2008, 23(1): 1-7.

[6] Li Ning, Shi Tielin. A New Method for Estimating the Number of Blind Signal Sources[J]. China Mechanical Engineering, 2007, 18(19): 2298-2302.

[7] Li Ning, Shi Tielin. Estimation of Blind Signal Sources Based on Nonnegative Matrix Factorization[J]. China Mechanical Engineering, 2007, 18(22): 2734-2737.

[8] FAN Tao, LI Zhi-nong, XIAO Yu-xian. Research on blind separation method of mechanical source signal based on source number estimation[J]. Mechanical Strength, 2011, 33(1): 015-019.

[9] Zhong Bocheng. Dynamic independent component analysis based on information maxim. Journal of Hefei University of Technology (Natural Science Edition), 2009, 32(8): 1154-1157.

[10] Liu Changrong. Time Series Signal Analysis System [D]. Dalian Jiaotong University, 2008.

[11] MacKay D J C. Probable networks and plausible prediction:a review of practical Bayesian methods of supervised neural networks[J].Computation in Neural Systems,1995(6):469-505.

[12] Rizwan.A. Choudrey. Variational method for bayesian independent component analysis[D]. UK: University of Oxford, 2002.

[13] Miskin J W. Ensemble learning for independent component analysis[D]. UK: University of Cambridge, 2000.