

# Analysis Method of Multi-scale Frequency Spectral Latent Feature based on the Mutual Information

Zhuo Liu<sup>1 a</sup>, Jian Tang<sup>1,2 b</sup>, Meiying Jia<sup>1 a</sup>, Xiaojie Zhou<sup>1,2 a</sup>

<sup>1</sup> State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, 110004, China

<sup>2</sup> Research Institute of Computing Technology, Beifang Jiaotong University, Beijing, 129001, China

<sup>a</sup>email: liuzhuo@mail.neu.edu.cn, <sup>b</sup>email: powernature@126.com

**Keywords:** Multi-scale Frequency Spectra; Latent Features extraction; Mutual Information

**Abstract.** Heavy rotating mechanical devices, such as ball mills in the mineral grinding process, produce strong mechanical vibration and acoustic signals. These signals consist of multiple components with different time scales and physical interpretations. Frequency spectra of these multi-scale sub-signals contain interesting information related to some difficulty-to-measure process parameters. Latent feature extraction methods based on principal component analysis and partial least squares with their kernel versions are used widely for high dimensional spectra data. How to estimate their effect and select important sub-signals are important issues. Thus, aim to these problems, mutual information is used in this paper. Analysis results based on the shell vibration frequency spectra of a laboratory-scale wet ball mill shows that the proposed method is helpful and effective.

## Introduction

In industrial process, heavy rotating mechanical devices with high energy and material consumption are used widely, such as steel ball mill in mineral grinding process. They are very important to realize operational optimization control of the mineral grinding process [1]. It is a key issue to detect some difficulty-to-measure process parameters within these devices. Strong vibration and acoustic signals of these mechanical devices are normally used to measure certain process parameters indirectly, such as load parameters within the ball mills [2]. These signals consist of multiple components with different time scales and physical interpretations. For example, there are millions of balls inside the ball mill arrange hierarchically. Impact forces and periods of different layers' balls to mill shell are different, which cause strong shell vibration. The vibration is only one of the main sources of the acoustical signal. In the time domain, the valued information is buried in a wide-band random noise signal "white noise" [3]. Thus, it is very difficult to select the interesting information to measure mill load parameters in time domain. Power spectral density (PSD) can be used to analyze these original mechanical signal and model different mill load parameters effectively [4,5,6].

Studies show that fast Fourier transform (FFT) may be unsuitable to analyze mechanical signals with multi-scale and non-stationary characteristics [7]. Some time-frequency analysis methods have been used to fault diagnosis of the rotating machine devices. Empirical mode decomposition (EMD) can adaptively decompose the original signal into some intrinsic mode functions (IMFs) from high frequency to low frequency orderly [8]. Normally, different IMFs have their detailed physical meanings. EMD method was also used to analyze shell vibration signal of the ball mill [9]. However, how to explain these sub-signals need more researches further.

Frequency spectral variables have characteristics of co-linearity and high dimension. Different feature selection and extraction methods, especially the principal component analysis and partial least squares approaches, have been studied to address this problem [10,11]. The spectral features' number is decreased dramatically. However, there isn't an effect method to select and explain these selected features. Thus, how to select and analyze some effective latent spectral features is still an

open issue.

In this paper, mutual information is used to estimate the latent features extracted by PCA and PLS from different IMFs. Analysis results based on multi-scale vibration frequency spectra of a laboratory-scale ball mill shell show that the proposed method is effective.

## Multi-scale Mechanical Vibration Frequency Spectra

Mechanical vibration is normally caused by different sources, especially for rotating mechanical devices. For example, ball mill is a kind of closing rotating devices. A large number of balls inside the mill are used to grind the material. These balls are layered. Moreover, balls in different layers are thrown down at the same time with different impact forces. In some grinding conditions, the impact forces are difficult to describe. These impact forces of different amplitudes and frequencies overlay each other. Moreover, mass un-balance and installment bias of the ball mill, and other reasons can also stir up the shell vibration. These signals are coupled and layered together.

Empirical mode decomposition (EMD) technique is widely used in mechanical vibration analysis of the rotating machine devices. Each IMF represents a simple oscillation mode with a physical meaning. The relations between different IMFs and the original signals can be represented as:

$$\mathbf{x}^t = \sum_{j=1}^J \mathbf{x}_{\text{IMFj}}^t + r_J \quad (1)$$

where,  $\mathbf{x}_{\text{IMFj}}^t$  and  $J$  represent the  $j$ th component and number of the components; and  $r_J$  represents residual error of the vibration signal.

Different IMFs represent different time scales of the mechanical vibration signals. In practice, they are caused by different impact sources. In frequency, their characters can be expressed much easily. The relation between IMF and its frequency spectrum can be represented as:

$$\mathbf{x}_{\text{IMFj}}^t \xrightarrow{\text{FFT}} \mathbf{x}_{\text{IMFj}}^f \quad (2)$$

where,  $\mathbf{x}_{\text{IMFj}}^f$  represents the  $j$ th IMF frequency spectra for the vibration signal.

## Latent Spectral Features Extraction

There is strong multi-collinearity among variables of the multi-scale frequency spectra. Some latent variables can represent many changes of them. This can be achieved by transforming the frequency spectral data to a new set of uncorrelated new variables. In dimension reduction domain, it is called feature extraction. In this paper, the normal PCA and PLS method are used.

PCA only uses the input data. That is to say, it decomposes the data matrix  $\mathbf{x}_{\text{IMFj}}^f$  as the sum of the outer products of vectors  $\mathbf{t}_i$  and  $\mathbf{p}_i$  plus a residual matrix  $\mathbf{E}_{\text{IMFj}}^{\text{PCA}}$ ,

$$\mathbf{x}_{\text{IMFj}}^f = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{t}_2 \mathbf{p}_2^T + \cdots + \mathbf{t}_h \mathbf{p}_h^T + \mathbf{E}_{\text{IMFj}}^{\text{PCA}}, \quad (3)$$

where,  $\mathbf{t}_i$  and  $\mathbf{p}_i$  are known as principal component scores and loadings vectors respectively;  $h$  is the number of the latent spectral features, which must be less than or equal to the smaller dimension of  $\mathbf{x}_{\text{IMFj}}^f$ ;  $\mathbf{t}_i$  vectors are known as the principal component scores, i.e., the latent spectral features, which contain information on how the samples relate to each other; and  $\mathbf{p}_i$  vectors are known as loadings, which contain information on how spectral variables relate to each other.

The model structure of the PCA for the  $j$ th IMFs' frequency spectrum is represented as

$$\mathbf{x}_{\text{IMFj}}^f = \mathbf{T}_{\text{IMFj}}^{\text{PCA}} (\mathbf{P}_{\text{IMFj}}^{\text{PCA}})^T + \mathbf{E}_{\text{IMFj}}^{\text{PCA}} \quad (4)$$

Thus, the extracted spectral features by PCA with  $h$  latent variable can be denoted as  $\mathbf{T}_{\text{IMFj}}^{\text{PCA}}$ .

Unlike PCA, PLS decomposes the input spectra data  $\mathbf{x}_{\text{IMFj}}^f$  and its responding output variables  $\mathbf{Y}$  into a low dimensional space. It can be shown as follows:

$$\mathbf{x}_{\text{IMFj}}^f = \mathbf{T}_{\text{IMFj}}^{\text{PLS}} (\mathbf{P}_{\text{IMFj}}^{\text{PLS}})^T + \mathbf{E}_{\text{IMFj}}^{\text{PLS}} \quad (5)$$

$$\mathbf{Y} = \mathbf{U}_{\text{IMFj}}^{\text{PLS}} (\mathbf{Q}_{\text{IMFj}}^{\text{PLS}})^T + \mathbf{F}_{\text{IMFj}}^{\text{PLS}} \quad (6)$$

where  $\mathbf{T}_{\text{IMF}_j}^{\text{PLS}}$  and  $\mathbf{U}_{\text{IMF}_j}^{\text{PLS}}$  are the score matrices;  $\mathbf{P}_{\text{IMF}_j}^{\text{PLS}}$  and  $\mathbf{Q}_{\text{IMF}_j}^{\text{PLS}}$  are the loading matrices;  $\mathbf{E}_{\text{IMF}_j}^{\text{PLS}}$  and  $\mathbf{F}_{\text{IMF}_j}^{\text{PLS}}$  are the modeling residuals of  $\mathbf{x}_{\text{IMF}_j}^f$  and  $\mathbf{Y}$  respectively.

Thus, the extracted spectral features by PLS with  $h$  latent variable can be denoted as  $\mathbf{T}_{\text{IMF}_j}^{\text{PLS}}$ .

As shown above, the spectral latent variables of PCA only contains changes in the input frequency spectra data. And PLS contains changes in the input and output data simultaneously.

How to estimate these latent features?

## Mutual Information based Latent Feature Analysis

Mutual information (MI) can measure the mutual dependence of the two variables quantitatively based on the probability theory and information theory. For concisely, the  $p$ th latent spectral features of  $\mathbf{T}_{\text{IMF}_j}^{\text{PCA}}$  or  $\mathbf{T}_{\text{IMF}_j}^{\text{PLS}}$  are re-written as  $(\mathbf{z}_{\text{IMF}_j})_p$ . The MI between the  $p$ th latent spectral features  $(\mathbf{z}_{\text{IMF}_j})_p$  and its responding output variables  $\mathbf{y}$  as:

$$\left( MI(\mathbf{y}; (\mathbf{z}_{\text{IMF}_j})_p) \right)_p = \iint p(\mathbf{y}, (\mathbf{z}_{\text{IMF}_j})_p) \log \frac{p(\mathbf{y}, (\mathbf{z}_{\text{IMF}_j})_p)}{p((\mathbf{z}_{\text{IMF}_j})_p) p(\mathbf{y})} d((\mathbf{z}_{\text{IMF}_j})_p) d\mathbf{y} = H(\mathbf{y}) - H(\mathbf{y} | (\mathbf{z}_{\text{IMF}_j})_p) \quad (7)$$

where  $p((\mathbf{z}_{\text{IMF}_j})_p)$  and  $p(\mathbf{y})$  are the margin probability densities of  $(\mathbf{z}_{\text{IMF}_j})_p$  and  $\mathbf{y}$ ;  $p((\mathbf{z}_{\text{IMF}_j})_p, \mathbf{y})$  is the joint probability density;  $H(\mathbf{y} | (\mathbf{z}_{\text{IMF}_j})_p)$  is the conditional entropy; and  $H((\mathbf{z}_{\text{IMF}_j})_p)$  is the information entropy.

In this paper, the captured percent variance of the  $p$ th latent spectral variable for  $\mathbf{x}_{\text{IMF}_j}^f$  are denoted as  $(\xi_{\text{IMF}_j})_p$ . Thus, the following metric is used to represent contribution of each latent variable,

$$(\delta_{\text{IMF}_j})_p = (\xi_{\text{IMF}_j})_p \cdot \left( MI(\mathbf{y}; (\mathbf{z}_{\text{IMF}_j})_p) \right)_p \quad (8)$$

Evidently, the important latent spectral variable has big  $(\delta_{\text{IMF}_j})_p$  value.

In order to estimate which IMF has more useful information than the others, the following equation is used.

$$\delta_{\text{IMF}_j} = \sum_{p=1}^h (\delta_{\text{IMF}_j})_p = \sum_{p=1}^h \left\{ (\xi_{\text{IMF}_j})_p \cdot \left( MI(\mathbf{y}; (\mathbf{z}_{\text{IMF}_j})_p) \right)_p \right\} \quad (9)$$

In this paper,  $\delta_{\text{IMF}_j}$  is called combined value.

Given out a preset threshold, by comparing with  $\delta_{\text{IMF}_j}$  can be used to select these important IMFs.

## Application Results

The experiments are performed on a laboratory-scale ball mill (XMQL-420×450). The copper ore is crushed to about less than 6 mm before used. The diameters of the steel balls are 30, 20 and 15 mm respectively. The vibration signal was picked up by accelerometer located on the middle of the mill shell.

To avoid redundancy, experimental details as described in [5], are omitted here.

The multi-scale frequency spectra are obtained by the following process:

(1) Decompose the original shell vibration signals every four mill rotating period data using EMD with the default parameters and obtain different time-scale IMFs time-domain signal;

(2) Transform different time-scale IMFs time domain signal to multi-scale frequency spectrum with FFT technology;

(3) There are total seven frequency spectra for each sample.

Thus, there are total  $26 \times 7 = 182$  sample.

The original single scale and different time scale IMFs' frequency spectra (100~6000Hz) of certain sample are shown in Fig. 1.

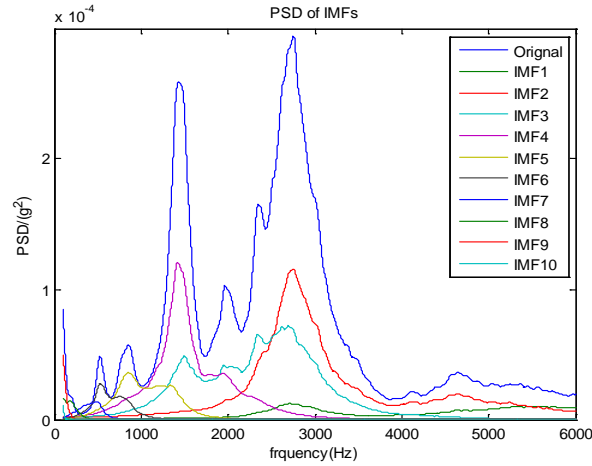


Fig. 1 Original single scale and different time scale IMFs' frequency spectrum (100~6000Hz)

Fig. 1 shows that different IMFs' frequency spectra represent different parts of the original single scale frequency spectrum. Obviously, different interesting information is contained in these multi-scale frequency spectra. Thus, IMFs' selection and their features' extraction is very necessary.

PCA is used to extract latent features of the former 10 IMFs firstly. Then, MI values between load parameter (charge volume ratio, CVR) and latent features of PCA are calculated. The captured percent variances (CPV) of the former 10 latent features and their MI values are shown in Fig. 2 and Fig. 3.

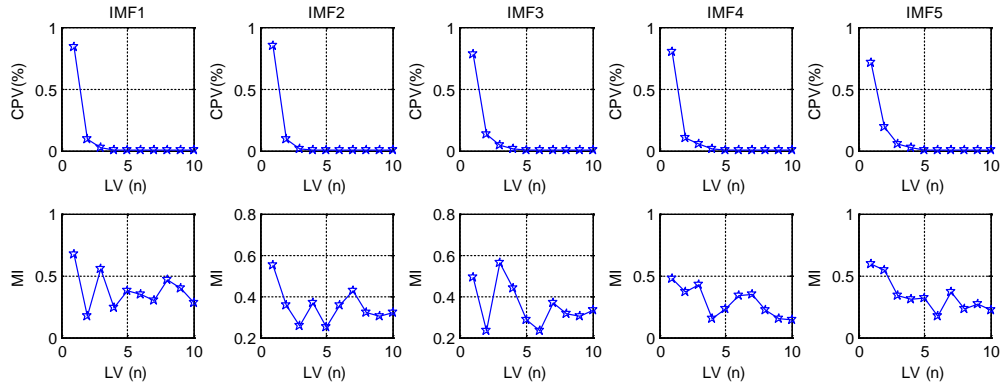


Fig. 2 Extracted latent features from multi-scale frequency spectra using PCA and their MI values for IMF1~IMF5

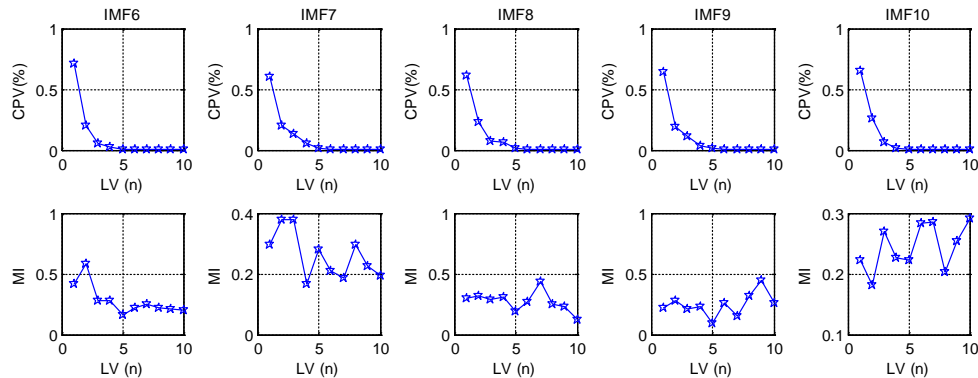


Fig. 3 Extracted latent features from multi-scale frequency spectra using PCA and their MI values for IMF6~IMF10

Fig. 2 and Fig. 3 show that the first latent features capture most of the changes from different multi-scale frequency spectra, and the first latent features don't have biggest MI value with CVR except IMF1, IMF2, IMF4 and IMF5. Thus, only select the former several latent features to construct regression or recognition model may be not a good strategy.

Same as PCA, PLS is used to extract latent features of the former 10 IMFs for CVR firstly. Then, MI values between CVR and latent features of PLS are calculated. The captured percent variances

(CPV) of the former 10 latent features and their MI values are shown in Fig. 4 and Fig. 5.

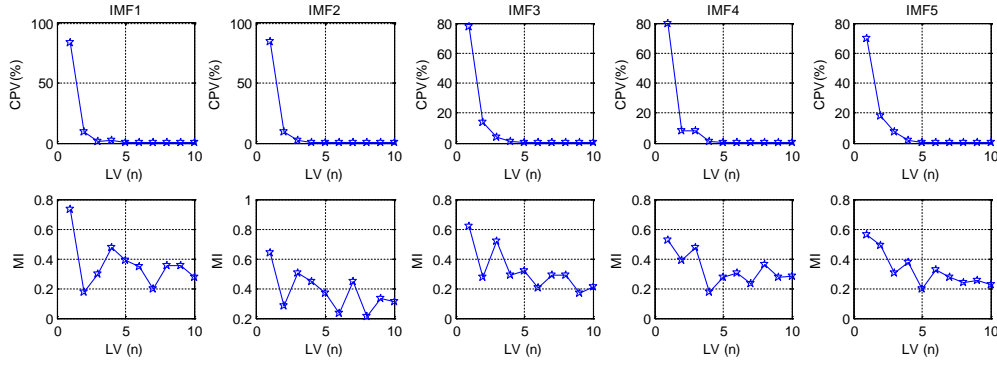


Fig. 4 Extracted latent features from multi-scale frequency spectra using PLS and their MI values for IMF1~IMF5

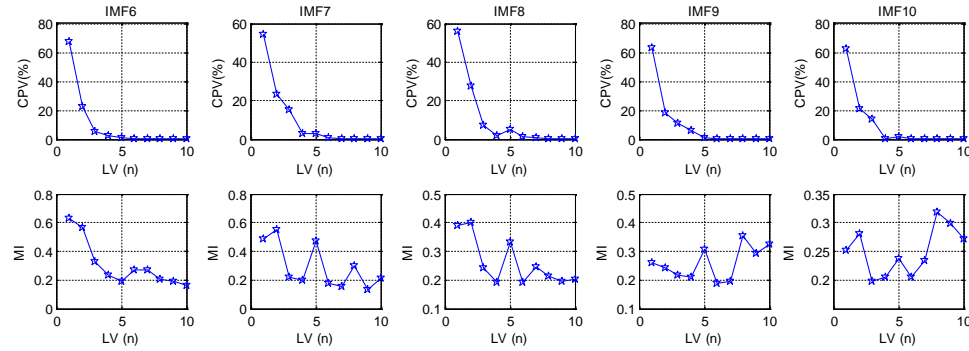


Fig. 5 Extracted latent features from multi-scale frequency spectra using PLS and their MI values for IMF6~IMF10

Same as that of PCA, Fig. 4 and Fig. 5 show that the first latent features capture most of the changes and always have the biggest MI value except IMF7~IMF10. Moreover, the MI values don't decrease with the latent features number. Thus, how to select latent features should consider CPV value, MI value and the requirement of the industrial practice jointly.

Statistical results of the combined values of the former 10 IMFs are shown in Fig.6.

Tabel 1. Statistical results of the latent features using PCA and PLS

	Latent features using PCA			Latent features using PLS			Note
	PV Max values	MI Max values	Combined Values	PV Max values	MI Max values	Combined Values	
IMF1	83.68	0.6742	0.6055	83.42	0.7355	0.6501	
IMF2	84.99	0.5476	0.5121	84.82	0.6366	0.5854	
IMF3	78.06	0.5646	0.4496	77.30	0.6186	0.5461	
IMF4	80.28	0.4752	0.4493	79.94	0.5215	0.4920	
IMF5	70.86	0.5920	0.5517	69.91	0.5594	0.5130	
IMF6	70.77	0.5795	0.4358	67.78	0.6325	0.5812	
IMF7	59.90	0.3800	0.3157	54.11	0.5484	0.4459	
IMF8	61.07	0.4375	0.3041	55.99	0.4017	0.3716	
IMF9	63.99	0.4458	0.2313	62.86	0.3540	0.2487	
IMF10	64.98	0.2908	0.2146	62.76	0.3183	0.2502	

Table 1 shows that: (1) The max values of PV from different IMFs using PCA are higher than that using PLS; (2) The max values of MI from different IMFs using PCA are lower than that using PLS; (3) Given preset combined value with 0.5, PCA and PLS select the same IMF1, IMF2 and IMF5 as the important ones. Therefore, different latent features extraction methods have their own characteristics. Some important IMFs can be recognized with the proposed method.

All in all, these important latent spectra features and important IMFs can be selected for construct regression or recognition model of these difficult-to-measure process parameters. There are at least three preset threshold values to be select firstly, i.e., PV and MI threshold values to select latent features, and combined value to select important IMFs.

## Conclusion

Based on the normally used latent features extraction and mutual information approaches, this paper gives out a new analysis method for multi-scale frequency spectra feature of the vibration signal from rotating mechanical device. The shell vibration signals of a laboratory scale grinding ball mill are analyzed with the proposed method. Experiments show that this method is helpful to extract and select the useful latent spectral features and the important multi-scale sub-signals. How to extend to kernel latent features selection, how to integrate with some modeling or classification algorithms, and how to optimized select the threshold values should be addressed in the future.

## Acknowledgement

In this paper, the research was sponsored by the post doctoral National Natural Science Foundation of China (Project No. 2013M532118, 2015T81082), National Natural Science Foundation of China (Project No. 61573364, 61273177), State Key Laboratory of Synthetical Automation for Process Industries, China National 863 Projects(Project No. 2015AA043802), Scientific Research Fund of Liaoning Provincial Education Department (Project No. L2013272).

## References

- [1] Tianyou Chai. Operational optimization and feedback control for complex industrial processes [J]. *Acta Automatica Sinica*. 2013 (39)1744-1757.
- [2] Jian Tang, Lijie Zhao, Heng Yue and Tianyou Chai. Present status and future developments of detection method for mill load [J]. *Control Engineering of China*. 2010 (17) 565-570.
- [3] Yigen Zeng, and Eric Forssberg. Monitoring grinding parameters by vibration signal measurement-a primary application [J]. *Minerals Engineering*. 1994 (7) 495-501.
- [4] Yigen Zeng, and Eric Forssberg. Monitoring grinding parameters by signal measurements for an industrial ball mill [J]. *International Journal of Mineral Processing*. 1993 (40) 1-16.
- [5] Jian Tang, Lijie Zhao, Junwu Zhou, Heng Yue and Tianyou Chai. Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell [J]. *Minerals Engineering*. 2010 (23) 720-730.
- [6] Jian Tang, Tianyou Chai, Wen Yu and Lijie Zhao. Modeling load parameters of ball mill in grinding process based on selective ensemble multi-sensor information [J]. *IEEE Transactions on Automation Science and Engineering*. 2013 (10) 726-740.
- [7] Yg Lei, Zj He and Yy Zi. Application of the EEMD method to rotor fault diagnosis of rotating machinery [J]. *Mechanical Systems and Signal Processing*. 2009 (23) 1327-1338.
- [8] Ne Huang, Z Shen, Sr Long. The empirical mode decomposition and the Hilbert spectrum for non-linear and non stationary time series analysis [J]. *Proc. Royal Soc. London A*. 1998 (454) 903-995.
- [9] Jian Tang, Lijie Zhao, Wen Yu and Tianyou Chai. Vibration analysis based on empirical mode decomposition and partial least squares[J]. *Procedia Engineering*. 2011 (16) 646-652.
- [10] Jian Tang, Tianyou Chai, Wen Yu and Lijie Zhao. Feature extraction and selection based on vibration spectrum with application to estimate the load parameters of ball mill in grinding process [J]. *Control Engineering Practice*. 2012 (20) 991-1004.
- [11] Zhuo Liu, Tianyou Chai, Wen Yu and Jian Tang. Multi-frequency signal modeling using empirical mode decomposition and PCA with application to mill load estimation [J]. *Neurocomputing*. 2015 (169) 392-402.