

## **Multi-fault Signal Feature Extraction of Mud Pump Based on Parallel Factorization**

Liu Yang<sup>1,a</sup>, Hanxin Chen<sup>1,b\*</sup>, Wenjian Huang<sup>1,c</sup>,  
Jinmin Huang<sup>1,d</sup> and Chenghao Cao<sup>1,e</sup>

<sup>1</sup>No. 206, Guanggu First Road, Jiangxia District, Wuhan, Hubei, 430205, P.R.China

<sup>a</sup>wgdyangliu@sina.com, <sup>b\*</sup>pg01074075@163.com, <sup>c</sup>hwjhere@126.com,

<sup>d</sup>15207176266@163.com, <sup>e</sup>holyqs@163.com

**Keywords:** Parallel factor; Fault diagnosis; Feature extraction; Multiple faults; Nonlinear

**Abstract.** The multi-source dynamic feature recognition of mechanical nonlinear multi-failure mode is a technical bottleneck and problem encountered in the application of fault diagnosis in the process industrial production line. It needs not only to extract the time-frequency characteristics of the single-source fault signal but also to ensure the corresponding relationship between the nonlinear variable, multi-fault mode and multi-source fault features in time, frequency and space after feature extraction. Mechanical multi-source signals also have coupling, aliasing, and other phenomena that cause the characteristic signals between the channels to interfere with each other and overlap. In order to meet the requirements of automatic monitoring and fault diagnosis of industrial process production lines, this paper develops multi-source dynamic feature recognition and adaptive diagnosis. This paper studies the multi-scale parallel factorization theory and proposes a three-dimensional time-frequency space model reconstruction algorithm for multi-source feature factors, which improves the accuracy of mechanical fault detection and intelligent levels

### **Introduction**

When the mechanical equipment is subjected to unsteady state excitation in the process industrial production line, the vibration signal contains an obvious frequency shift characteristic. The fault feature signal will flow at each vibration collection point as the nonlinear failure mode changes. The two-dimensional time-frequency characteristics of the extracted signal are fuzzy, uncertain and incomplete. The method of single-channel signal extracting two-dimensional time-frequency feature like wavelet theory cannot reflect the intrinsic relationship of non-linear changes between multi-source channel feature signals and cannot eliminate information interference[1-3]. Dual-channel signal feature extraction methods such as holographic spectrum and global vector spectrum are used for the axis motion of the rotating machinery to establish signal frequency, amplitude, cross-section frequency or cross-sectional phase, where each channel signal is analyzed independently[4-5]. It can be seen that we cannot separate multi-source signals into single signals to extract two-dimensional time-frequency features independently. Multi-fault mode multi-source dynamic feature extraction is a multi-scale parallel factorization of multi-source signal matrix including process state variables, vibration signals, etc[6].

It not only extracts the nonlinear time-frequency fault feature information of single-source signal, but also completes the fusion of multi-source feature signal simultaneously. The parallel factorization theory ensures the optimization of the nonlinear relationship between process state variables, failure modes and multi-source vibration signatures in time, frequency and spatial feature vector space after feature extraction[7].

### **Multi-scale Parallel Factorization Theory Algorithm**

The parallel factor trilinear model is the easiest way to extend factor analysis algebra from two dimensional to three dimensional, The form is as follows:

$$x_{ij} = a_{i1}f_{j1} + a_{i2}f_{j2} + L + a_{iR}f_{jR} + e_{ij} \quad (1)$$

Where  $X_{ij}$  is the data of row  $i$ , column  $j$  of matrix  $X$ . " $a_{i1}f_{j1}, \dots, a_{iR}f_{jR}$ " represents the contribution factor of the observed value. " $e_{ij}$ " is the residual value and its properties depends on whether it fits the principal component or the common factor model. The parameters  $a_{iR}$  and  $f_{jR}$  represent loading factors and component factors, respectively, which have different standardized rules and evaluation methods. In general, the two sets of factors play the same role as representing a specific layer of factor influence and importance values in a particular pattern of data. The two sets of parameters affect the variable, proportionally reducing or increasing the contribution of the factor from one layer of the specific pattern to the next. If you add a set of parameter modes  $C$ , the model is trilinear. Adding other additional modes, it will become a multi-linear module. The trilinear model is defined as:

$$x_{ijk} = a_{i1}b_{i1}c_{i1} + a_{i2}b_{i2}c_{i2} + L + a_{iR}b_{iR}c_{iR} + e_{ijk} \quad (2)$$

The uniqueness principle of the parallel factor model is that if the influence variables of each factor in the three patterns of data have different rules from other factors, one load factor of the data matrix subset increases or decreases in proportion to the other factor. The study of the uniqueness of trilinear decomposition is a continuous research hotspot. Uniqueness is the determination of the direction of the vector axis of the most scientifically significant factor in a certain factor space. This principle provides a method to determine the direction of the unique axis that is looking for a subset of the data matrix in a certain factor space along a set of consecutive directions stretching or contracting proportionally. The reliability of the factor vector axis direction needs to be tested by other methods. The best direction of the factor vector axis is to optimize the three-linear factor model to fit three-dimensional data.

This section introduces the multi-scale parallel factorization theory algorithm and defines the "loading" factor and "component" factor of the parallel factorization. This section also analyzes the selection methods and criteria of the most scientific factor vector axes for a particular factor vector space. Combining multi-scale analysis theory, a multi-scale parallel factorization optimization algorithm for extracting nonlinear single-source fault feature signals is established. The multi-scale analysis theory improves the accuracy of time-frequency space 3D feature analysis of

nonlinear feature signals so that the feature factor signals are successfully obtained from matrix factors. The data matrix  $S_{(N_d \times N_f \times N_t)}$  is a three-dimensional time-varying spectrum array obtained by the wavelet transform of the vibration signal.  $N_d, N_f$  and  $N_t$  are the number of channels, frequency steps and number of data points, respectively. The basic structure of the Parallel Factor Analysis (PARAFAC) decomposition model[8] is shown in Figure 1.

$$\hat{S}_{dft} = \sum_{k=1}^{N_k} a_{dk} b_{fk} c_{tk} \tag{3}$$

The key issue of this model is to find the matrix A, B and C. Their elements are  $a_{dk}$ ,  $b_{fk}$  and  $c_{tk}$ . Each component k represents an atom. The corresponding vectors  $a_k = \{a_{dk}\}$ ,  $b_k = \{b_{fk}\}$  and  $c_k = \{c_{tk}\}$  are spatial signals, spectral signals and temporal signals for each atom. The uniqueness of the solution is guaranteed by  $rank(A) + rank(B) + rank(C) \geq 2N_k + 2$ . The vector  $a_{k(N_d \times 1)}$  produced by PARAFAC is the k-th component space vector, the vector  $b_{k(N_f \times 1)}$  is the k-th component of the spectrum and the vector  $c_{k(N_t \times 1)}$  is the k-th component of the time signal. The main advantage of this method is that the spectrum decomposition of the time-varying vibration signal is unique and the best model can be obtained under the principle of minimum squared error.

The multi-scale parallel factorization algorithm is presented as follows:

- (1) Multi-scale time-frequency decomposition.
- (2) Determine the number of factors F.
- (3) Initialize load matrix B and C.
- (4) Load matrices B and C to estimate A by the least-squares regression algorithm, that is

$$A = XZ'(ZZ')^{-1}, \quad Z = (b \otimes c).$$

- (5) Similarly, estimates B and C.
- (6) Calculate continuously from step (3) until convergence.

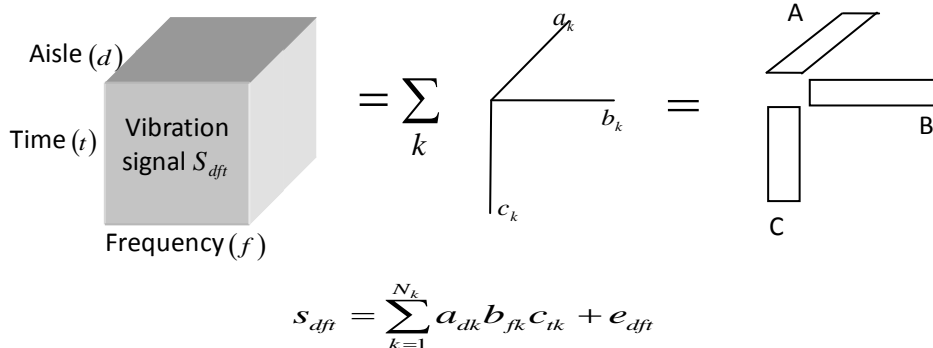


Figure 1 The model of multi-scale parallel factor (PARAFAC) decomposition

### Mud Pump Fault Diagnosis Based on Parallel Factor Analysis

Process multi-source dynamic signal acquisition system in process industry is shown in Figure 2. The system can simultaneously achieve acquisition of real-time signals

such as flow, pressure, speed and vibration. The speed of the sediment pump is controlled by the motor and the rotation speeds are 1200rpm, 1400rpm, 1600rpm, ... until the 2800rpm. The non-linear multi-failure mode is modeled by controlling the flow and pressure of the process loop for simulating the nonlinear operating state of the process industrial process mechanical system. The process industrial sediment pump experimental system collected the data of pipeline pressure, motor rated power, motor speed, pump inlet and outlet pressure in the nonlinear failure mode. Because the mechanical structure generates more nonlinear components under the action of fluid rheology, the nonlinear frequency components of vibration are more complicated.



Figure 2 Experiment system of sediment pump

This experiment collected the pipeline pressure, motor rated power, motor speed, pump inlet and outlet pressure in the normal mode and failure mode of the sediment pump experiment system. Figure 3 shows the three-dimensional time-frequency signal obtained by multi-scale wavelet decomposition for the collected pipeline pressure, motor rated power, motor speed, pump inlet and outlet pressures. The abscissa, ordinate, and vertical coordinates represent time, frequency, and amplitude, respectively. It is shown about the multi-scale parallel factorization of vibration signals in Figure 4. After three-dimensional time-frequency signals are decomposed by multiple layers of parallel factorization, amplitude loading factors, frequency loading factors and time loading factors can be obtained. The ordinate represents the loading value, which is the size of the judging factor. In the multi-scale parallel factorization of the vibration signal in FIG. 4, the left side shows the normal state and the right side shows the failure state. Compared with the decomposition results of the normal state and the failure state, there are obvious differences in the frequency loading factor components, so the frequency loading factor is effective and can successfully identify the failure.

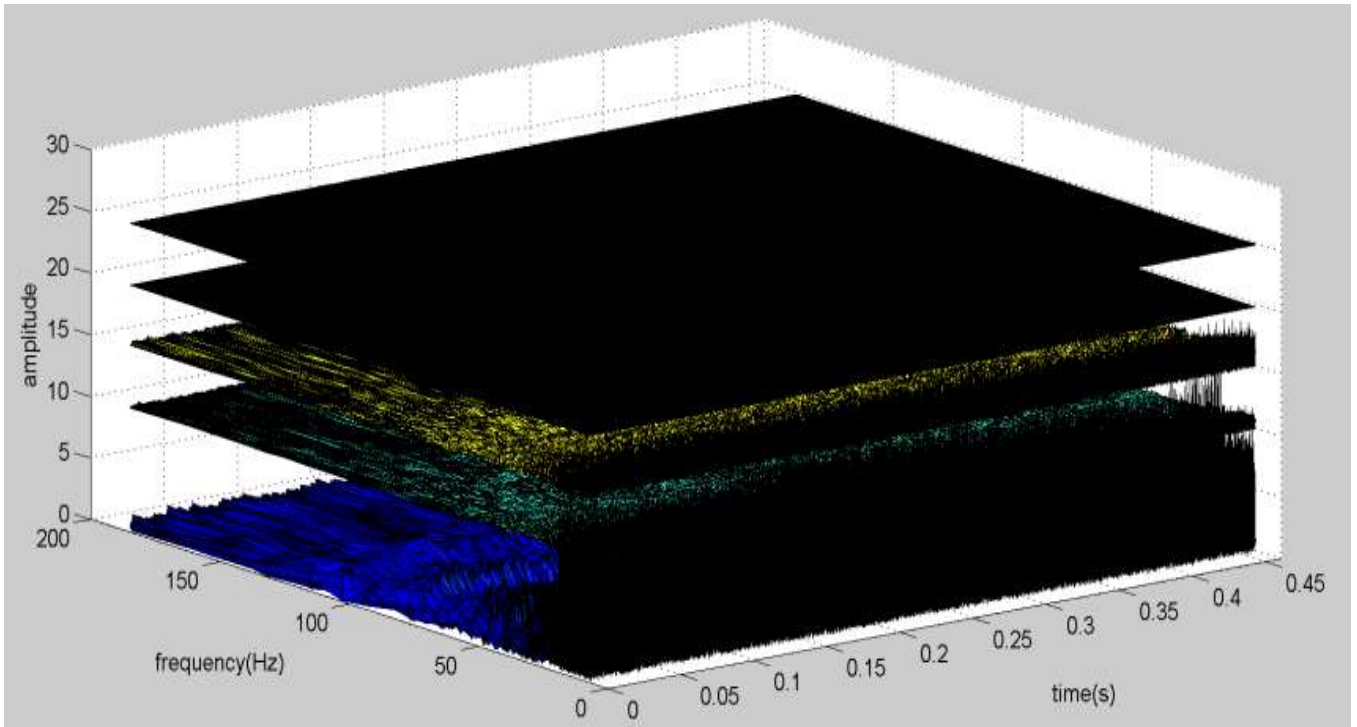


Figure 3 Three-dimensional time-frequency signal

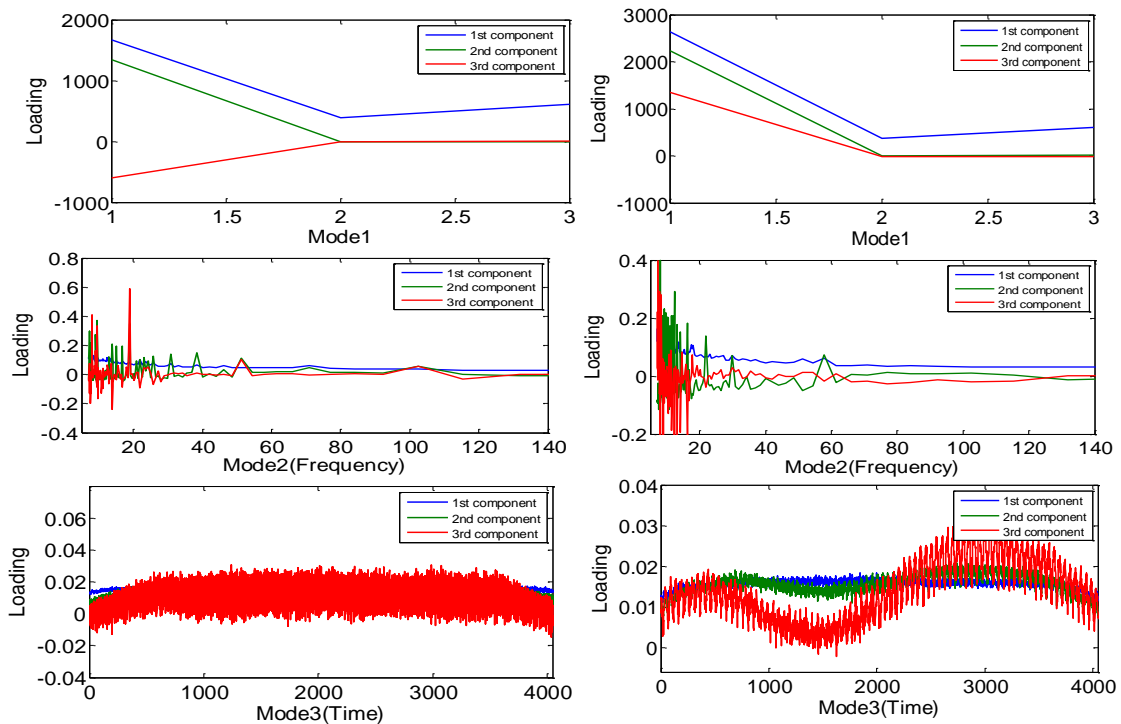


Figure 4 Multiscale parallel factorization of vibration signals

### Conclusion

In this project, the relationship between the characteristics of nonlinear flow conditions, fault sources, various vibration sources and multi-source nonlinear vibration signals is established under the mechanical fault conditions. The vibration characteristics of various sources of vibration and faulty parts of the mud pump are analyzed under flow field conditions, and the relationship between the nonlinear flow



conditions and vibration characteristics is established. Multi-fault mode multi-source dynamic feature extraction is a multi-scale parallel factorization of multi-source signal matrix including process state variables, vibration signals, etc. It not only extracts the nonlinear time-frequency fault feature information of single-source signal but also completes the optimization of multi-source feature signal fusion at the same time. This guarantees the best non-linear correspondence between process state variables, fault modes and multi-source vibration signatures after feature extraction in 3D feature vector space of the time, frequency and space. It reconstructs a unique time-frequency-space three-dimensional vector space nonlinear fault feature model corresponding to multi-source signals to achieve adaptive identification of multi-source fault features and improve diagnostic accuracy.

### **Acknowledgement**

This work was supported by the Special Major Project of the Ministry of Science and Technology of Hubei Province of China (Grant No.2016AAA056), Major project of Hubei Provincial Department of Education (Z20101501) and the National Natural Science Foundation of China (Grant No. 61273176,51775390).

### **References**

- [1] Qiu H. Lee J. Lin J. Yu G. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*. 2005, 289 (4-5): 1066-1090.
- [2] Li Fucui, Meng Guang, Ye Lin and Chen Peng. Wavelet transform-based higher-order statistics for fault diagnosis in rolling element bearings. *Journal of vibration and control*. 2008, 14(11), 1691-1709.
- [3] Weixiang Sun, Jin Chen and Jiaqing Li, Decision tree and PCA-based fault diagnosis of rotating machine, *Mechanical Systems and Signal Processing*, Vol.21, n3, 2007, pp.1300-1317.
- [4] Andrew K.S. Jardine, Daming Lin, Dragan Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, *Mechanical systems and signal processing*, 2006, 20, 1483-1510.
- [5] Venkat Venkatasubramanian, Raghunathan Rengaswamy, Surya N. Kavuri, Kewen Yin, A review of process fault detection and diagnosis Part III: Process history based methods, *Computers and Chemical Engineering*, 2003, 27, 327-346.
- [6] Yuan, Shengfa , Chu, Fulei; He, Yongyong , Multi-fault diagnosis for turbo-pump based on mesh support vector machines, *Chinese Journal of Mechanical Engineering*, 2007,43(4), pp.152-158 .
- [7] Miwakeichi, F., Martinez-Montes, E., Valdes-Sosa, P.A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y.: Decomposing EEG data into space/time/frequency components using Parallel Factor Analysis. *NeuroImage* 22,1035-1045(2004).
- [8] Andersen, C. M., and R. Bro. 2003. Practical aspects of PARAFAC modeling of fluorescence excitation-emission data. *J. Chemom.*17:200-217.