Customized maximal-overlap multiwavelet denoising with data-driven group threshold for condition monitoring of rolling mill drivetrain

Jinglong Chen, Zhiguo Wan, Jun Pan, Yanyang Zi, Yu Wang, Binqiang Chen, Hailiang Sun, Jing Yuan, Zhengjia He

State Key Laboratory for Manufacturing and Systems Engineering, Xi’an Jiaotong University, Xi’an 710049, PR China
Department of Mechanical and Electrical Engineering, Xiamen University, Xiamen 361005, PR China
Beijing Institute of Astronautical Systems Engineering, Beijing 100076, PR China
Shanghai Institute of Radio Equipment, Shanghai 200090, PR China

Article history:
Received 12 January 2015
Received in revised form 2 July 2015
Accepted 20 July 2015
Available online 17 August 2015

Keywords:
Condition monitoring
Customized maximal-overlap multiwavelet
Rolling mill drivetrain
Data-driven group threshold

Fault identification timely of rolling mill drivetrain is significant for guaranteeing product quality and realizing long-term safe operation. So, condition monitoring system of rolling mill drivetrain is designed and developed. However, because compound fault and weak fault feature information is usually sub-merged in heavy background noise, this task still faces challenge. This paper provides a possibility for fault identification of rolling mills drivetrain by proposing customized maximal-overlap multiwavelet denoising method. The effectiveness of wavelet denoising method mainly relies on the appropriate selections of wavelet base, transform strategy and threshold rule. First, in order to realize exact matching and accurate detection of fault feature, customized multiwavelet basis function is constructed via symmetric lifting scheme and then vibration signal is processed by maximal-overlap multiwavelet transform. Next, based on spatial dependency of multiwavelet transform coefficients, spatial neighboring coefficient data-driven group threshold shrinkage strategy is developed for denoising process by choosing the optimal group length and threshold via the minimum of Stein’s Unbiased Risk Estimate. The effectiveness of proposed method is first demonstrated through compound fault identification of reduction gearbox on rolling mill. Then it is applied for weak fault identification of dedusting fan bearing on rolling mill and the results support its feasibility.

1. Introduction

Metallurgy and steel industry are the important industries among the national economy, and these developments have quite intimate relation with the economic foundation. As the core equipment in steel industry, rolling mill is used to implement the process of metal rolling. During the rolling process, typical operating conditions of rolling mill refer to extreme mechanical situations including large values of tensions and forces [1]. In some cases, these extreme mechanical situations can lead to different kinds of faults on rolling mill, which might bring about serious accidents and huge economic
losses. So, the realization of stability and security operation on rolling mill is significant and always draws a lot of attention for enterprise [1]. In fact, rolling mill is chronically running under complex and harsh operating condition of fatigue, heavy loads, etc., and gear or bearing as the key component of rolling mill drivetrain inevitably generates various faults. Once accidents appear in rolling mill drivetrain, they will directly result in enormous economic losses and serious casualties. Especially, because suffered more harsh environment such as elevated temperature, etc., hot strip finishing mill equipment more easily generates serious mechanical faults compared with cold-rolling mill. Hence, condition monitoring and fault diagnosis of rolling mill is of great importance and indispensable for guaranteeing product quality, realizing long-term safe operation and avoiding of significant economic losses.

Regardless of the technique factor, the capability of any condition monitoring system mainly relies on two key components: the type and number of sensors, the associated signal processing and simplification methods applied to extract important information from the acquired various signals [2-3]. Condition data acquisition refers to collecting the required variables (e.g. speed, temperature, voltage) as well as turning them into electronic signals. Recent years, vibration analysis have achieved great progress and therefore continues to be one of the most popular technology applied for the condition monitoring and fault diagnosis of mechanical equipment [4]. The remaining key component is how to process the measured vibration data in order to obtain the current condition features. One possible solution is to treat the measured vibration data as a process which can be parameterized using simple statistical analysis (mean, maximum, minimum, etc.) or advanced higher order statistics (kurtosis, etc.) [5]. The further approach is to process the measured vibration data in the frequency domain (fast-Fourier transform, etc.) [5]. But these approaches does not make much sense when the measured vibration data is collected under the non-stationary operating regime. Unfortunately, engineering practices have demonstrated that condition information data gathered from machine-integrated sensors usually appears non-stationary characteristic [6]. Thus, due to the complex mechanical structure and various operation environment, more effective signal processing method is indispensable and should be developed and introduced for condition monitoring and fault diagnosis of rolling mill drivetrain.

Some interesting studies related to condition monitoring and fault diagnosis of rolling mill have been reported recently in the literature. Li et al. designed a fault diagnosis scheme for hydraulic gauge control system of strip rolling mill based on wavelet transform and neural networks, and the analyzed results on the varied fault features demonstrated the effectiveness on the proposed diagnosis system [7]. Shao et al. pay attention on study the vibration characteristic of twenty-high rolling mill with local defect on roll surface using the time-varying contact stiffness and proposed an adaptive noise cancellation method based on beehive pattern evolutionary digital filter for fault feature extraction [8-9]. Chu et al. carried out fault diagnosis using support vector machines through parameter optimisation via artificial immunisation algorithm for turbo pump rotor [10]. Yuan et al. developed multiwavelet sliding window neighboring coefficients denoising algorithm with optimal blind deconvolution for gearbox fault diagnosis of rolling mills [11]. Daniel et al. applied the dimensionality reduction technique, called t-SNE, to visual exploratory analysis of the dynamic behaviors in a cold rolling process and supplied a possible way for detecting the chatter fault [1]. Li et al. proposed adaptive stochastic resonance method on the basis of sliding window for driving gearbox fault detection in a hot strip finishing mill [12]. Cai et al. developed sparsity-enabled signal decomposition strategy based on tunable Q-factor wavelet transform for gearbox localized fault detection of rolling mill [13]. Chen et al. investigated a new technique called customized lifting multiwavelet packet information entropy for resonance condition identification of rolling mill [14]. Chen et al. applied overcomplete rational dilation discrete wavelet transform for gearbox fault detection of rolling mill [15]. Serdio et al. carried out residual-based fault detection based on soft computing techniques for condition monitoring of rolling mill [16]. Ming et al. adopted cyclic Wiener filter and envelope spectrum analysis for weak fault feature detection of rolling element bearing [17]. These mentioned studies have provided critical insight on condition monitoring and fault diagnosis of rolling mill. However, there are still abundant issues to be addressed in this task. One important aspect of them is how to effectively identify compound fault and detect weak fault from measured noisy vibration data of faulty component in rolling mill drivetrain and exactly assess the current operation condition. Due to the complexity of equipment and the correlation of structures, several faults often appear at the same time and the features of each fault are coupled together. This kind of failure form is called compound faults [6,12]. In these situation, mechanical fault detection turns into a challenge task, especially in the operational condition with strong background noise.

As a powerful tool for describing the non-stationary signal, wavelet transform (WT) [18-19] has already shown its tremendous strength in mechanical equipment condition monitoring and fault diagnosis because of its advantage on multi-resolution analysis and abundant basis functions [20-21]. Recent years, many scholars have paid a lot of attention to wavelet denoising technique with appropriate threshold shrinkage rule on signal processing to improve the SNR for fault feature extraction and has made some progress and applications. According to the algorithm flow of wavelet denoising technique, the performance of wavelet denoising method mainly relies on the three factors such as the appropriate selections of wavelet base, transform strategy and threshold rule [22].

Different from Fourier transform, a specific fault symptom can be detected and extracted by WT on the basis of appropriate selection the basis function from the existing basis function library, which is greatly beneficial with the condition feature identification. However, the selection of basis function is not uncontrolled because there are limited basis functions in the library. And any inappropriate wavelet basis function employed in the special engineering application will directly decrease the accuracy of the condition feature extraction. So it is a vital step to select an appropriate wavelet basis function for the measured condition vibration data processing. In fact, any fixed basis function which is not related to the
special condition vibration data can’t detect and extract a data feature entirely at all in the special application [23]. Moreover, any scalar wavelet base in the wavelet basis function library cannot possess orthogonality, symmetry, compact support and higher order of vanishing moments simultaneously [24]. But unfortunately, these properties are significant for describing a vibration data comprehensively and precisely. Furthermore, due to the major intrinsic deficiency of critically-sampled filter-bank, traditional wavelet discrete transform would lead to translation-variance and Gibbs phenomena which are harmful to extract periodical impulse feature [25-26]. To overcome these mentioned shortcomings of traditional wavelet transform and realize exact matching and accurate detection of fault feature on rolling mill drivetrain, customized maximal-overlap multiwavelet transform is developed in this paper.

As described above, threshold rule also play an important role on affecting the performance of wavelet denoising method. Many attempted threshold rules have been proposed for wavelet denoising technique. Donoho and Johnstone [27] first developed a simple and feasible threshold rule that sets all the wavelet decomposed coefficients smaller than the universal threshold $\sigma \sqrt{2 \log n}$ to zero and compresses the remaining decomposed coefficients by the threshold (Soft-threshold) or keeps them without change (Hard-threshold) [28]. Later, Stein’s unbiased risk estimator (SURE) threshold rule was developed on the basis of minimizing an estimation of the risk and became another useful threshold rule for wavelet denoising method [29]. Based on the mentioned SURE threshold rule, Sun et al. proposed a data-driven threshold rule for multiwavelet denoising method to carry out wind turbine fault detection [30]. These above threshold rules only focuses on analyzing one decomposed point and processing the signal term by term. But the important fact that any decomposed point to be processed is dependent on those in its neighborhood would be neglected. Based on this important feature of vibration signal, Cai and Silverman [31] proposed a more reasonable threshold rule called neighboring coefficients (NeighCoeff) threshold rule. Later, Chen et al. [32] further improved the existing neighboring coefficients threshold rule by considering the intra-scale and inter-scale dependencies of multiwavelet decomposed coefficients. Although the improved neighboring coefficients threshold rule appears good performance during some signal denoising applications, an inflexible global threshold rule without dependence to the analyzed signal still seriously hampers the implementation of accurate extraction on condition feature under strong background noise.

In order to realize the accurate exaction of compound fault and weak fault feature on rolling mill drivetrain, customized maximal-overlap multiwavelet denoising via spatial neighboring coefficient data-driven group threshold shrinkage strategy is proposed in this paper. There are three parts in this method for fault identification of rolling mill drivetrain. Due to the properties on multi-resolution analysis and simultaneously possessing important properties such as symmetry, orthogonality, compact support and higher order of vanishing moments that traditional scalar wavelet does not have, multiwavelet has the advantages on detecting weak fault feature [33]. Moreover, because of its multiple wavelet basis functions, multiwavelet does well in detecting features with multiple kinds of shapes for compound fault feature extraction. However, a fixed multiwavelet base without dependence on analyzed signal employed in the engineering application still result in lowering the accuracy of the fault detection [21]. So, customized multiwavelet base should be constructed. First, new multiwavelet basis function is obtained by performing symmetric multiwavelet lifting scheme on a known multiwavelet, so as to gain the greater free parameter for building vibration data-driven multiwavelet [19,32,34]. The free parameter is optimized via genetic algorithm according to maximum kurtosis—envelope spectrum entropy objective. Then, maximal-overlap multiwavelet transform is performed to avoid these shortcomings of translation-variance and Gibbs phenomena and its average process shows superior denoising and maintains signal smoothness. Third, based on the spatial correlation between maximal-overlap multiwavelet transform coefficients, this paper selects the optimal group length and threshold by using the minimum of Stein’s Unbiased Risk Estimate when estimating the true unknown fault features. The optimal group length and threshold are applied for the effective feature extraction and noise elimination at each decomposition level. Spatial neighboring coefficient data-driven group threshold shrinkage strategy is introduced to process the maximal-overlap multiwavelet transform coefficients. The effectiveness of proposed method is first demonstrated through compound fault identification of reduction gearbox on rolling mill. Then this method is applied for weak fault identification of dedusting fan bearing on rolling mill.

The rest contents of this paper are organized and displayed as follows. In Section 2, condition monitoring system for hot strip finishing mill is designed. In Section 3, summary of multiwavelet transform is firstly introduced and then the customized multiwavelet construction algorithm is developed. Next, spatial neighboring coefficient data-driven group threshold shrinkage strategy is proposed and finally the proposed method is presented. In Section 4, the presented method is applied to the two engineering cases to demonstrate its performance. Conclusions are given in Section 5.

### 2. Condition monitoring system for drivetrain of hot strip finishing mill

In rolling mill, a flat steel plate in a certain thickness with smooth surface to be produced. It is not easy to realize this goal due to the various possible engineering problems during the rolling process. The drivetrain of a hot strip finishing mill in Shanghai Baosteel Group Corporation needs to equip the condition monitoring system for guaranteeing product quality, realizing long-term safe operation and avoiding of significant economic losses. The target experimental equipment is illustrated in Fig. 1.

Condition monitoring and fault diagnosis system for this rolling mill drivetrain is developed on the basis of Labview. The testing framework of condition monitoring system and the photographs of monitoring equipment in field test are shown in Fig. 2 and the corresponding work flow diagram of condition monitoring system is displayed in Fig. 3. There are three parts
with different degree in the diagram including: preliminary diagnosis, accurate diagnosis and remote diagnosis. We collected the vibration signals of the rolling mill drivetrain using internal electronics piezoelectric (ICP) acceleration sensors when rolling mill is running. In this condition monitoring system, all the important components including gearbox, bearing and generator can be monitored and the corresponding condition vibration signals are displayed timely. Moreover, the abnormal condition of any important component can be alarmed after the vibration data processing of feature extraction and condition identification through the required diagnosis steps. The accurate diagnosis and remote diagnosis processes are important to find out the cause of abnormal alarm. During the two processes, effective signal processing method is needed and necessary. The proposed method called customized maximal-overlap multiwavelet denoising has been added in the accurate diagnosis process of condition monitoring system to diagnose the abnormal condition of rolling mill drivetrain. In a word, we can acquire real-time vibration data and extract features of the vibration signal by the fault diagnosis system in order to accurately realize condition identification as well as guiding the safety operation of the rolling mill.

3. Principle of customized maximal-overlap multiwavelet denoising

3.1. Summary of multiwavelet transform

Multiwavelet is generated by two or more mother wavelets [35]. Similar to the scalar wavelet transform, the theory of multiwavelet is also on the basis of the concept of multi-resolution analysis (MRA) [35]. Multi-scaling function vector \( \Phi = [\phi_1, \phi_2, \ldots, \phi_r]^T \) and multiwavelet function vector \( \Psi = [\psi_1, \psi_2, \ldots, \psi_r]^T \) satisfy the following two-scale matrix refinement equations:

\[
\Phi(t) = \sqrt{2} \sum_{k=0}^{M} H_k \phi(2t - k) \quad k \in \mathbb{Z}
\]

\[
\Psi(t) = \sqrt{2} \sum_{k=0}^{M} G_k \phi(2t - k) \quad k \in \mathbb{Z}
\]

The coefficients \( \{ H_k \} \) and \( \{ G_k \} \) are \( r \times r \) matrices instead of scalars and \( \Psi = [\psi_1, \psi_2, \ldots, \psi_r]^T \) denotes the multiwavelet function corresponding to multi-scaling function \( \Phi \). In the frequency domain, Eqs. (1) and (2) are

\[
\hat{\Phi}(\omega) = H(e^{-i\omega/2})\hat{\Phi}(\omega/2)
\]

\[
\hat{\Psi}(\omega) = G(e^{-i\omega/2})\hat{\Phi}(\omega/2)
\]

\( H(\omega) \) and \( G(\omega) \) are the refinement symbols corresponding to \( \Phi \) and \( \Psi \). The symbols in Z-domain are determined by

\[
H(z) = \frac{1}{2} \sum_{k=0}^{M} H_k z^k \quad \text{and} \quad G(z) = \frac{1}{2} \sum_{k=0}^{M} G_k z^k
\]

With the starting vector coefficients \( \lambda_{0,0}, \ldots, \lambda_{0,2^j-1} \), the decomposition step of multiwavelet transform is

\[
\lambda_{j-1,n} = \sum_k H_{k-2n}\lambda_{j,k} \quad \text{and} \quad \gamma_{j-1,n} = \sum_k G_{k-2n}\gamma_{j,k}
\]
Fig. 2. (a) The testing framework of condition monitoring system; (b) and (c) the photographs of monitoring equipment in field test.

Fig. 3. The work flow diagram of the condition monitoring system.
Low frequency coefficients $\lambda_{j-1,n}$ and high frequency coefficients $\gamma_{j-1,n}$ after the decomposition step are vectors of r-dimension. The reconstruction step of multiwavelet transform is

$$\hat{\lambda}_{jk} = \sum_{n} H_{k}^{n} \lambda_{k-2n} + \sum_{n} C_{k}^{n} \gamma_{j-1,n}$$ (7)

Note that the superscript * means the complex conjugate transpose.

Due to the translations and dilations operations of multi-scaling and multiwavelet vector functions, multiwavelet can seize the vital vibration data processing properties of orthogonality, symmetry, compact support and higher order of vanishing moments simultaneously [19], which has been proved to be impossible for scalar wavelet except Haar wavelet. So, multiwavelet transform can describe any vibration data more precisely and comprehensively because of its multi-input and multi-output system. In addition, due to the matrix-valued filter-bank, two or more input streams are needed in the process of multiwavelet transform. However, the processing vibration data would be one input stream usually and so some kind of pre-processing should be done before the implementation of multiwavelet transform. Correspondingly, a post-processing step is needed after the multiwavelet transform and it must be the inverse process of the pre-processing step. There are many kinds of pre-filters with different properties [36]. It also has been proved that the pre-filter algorithm called oversampling is beneficial for vibration data feature identification than critically sampling ones [36]. Therefore, oversampling algorithm is selected as the multiwavelet preprocessing operation in this paper and applications.

3.2. Construction of customized multiwavelet basis function

Attributed to multiple wavelet bases, multiwavelet does well in detecting various signal feature shapes for the abnormal condition identification. However, the fixed multiwavelet basis functions are independent of the measured vibration signal, which might greatly decrease the accuracy of condition identification. So, it is significant to generate customized multiwavelet basis functions for the given special vibration signal.

Sweldens developed the lifting scheme which made use of an existing wavelet and scaling functions to generate a new wavelet with prescribed or required properties via transform in time domain, which makes it possible to produce adaptive multiwavelet [19]. Based on lifting scheme, author developed a new construction method of customized multiwavelet for measured vibration signal processing of mechanical equipment in Ref. [32]. In the section, the procedure of the customized multiwavelet construction algorithm will be presented briefly.

Based on the multiwavelet lifting scheme, a changeable set of orthonormal filter operators $\{H_j, \tilde{H}_j, G_j, \tilde{G}_j\}$ can be obtained as follows [37]:

$$H_{\text{new}}(z) = H(z)$$

$$G_{\text{new}}(z) = T(z^2)(G(z) + S(z^2)H(z))$$

$$\tilde{H}_{\text{new}}(z) = H(z) - S(z^2)\tilde{G}(z)$$

$$\tilde{G}_{\text{new}}(z) = (T^*(z^2))^{-1}\tilde{G}(z)$$

where the determinant of $T(z)$ is a monomial and $S(z)$, $T(z)$ are finite-degree.

One of the most important properties of the multi-scaling function which has noteworthy significance in engineering application is the approximation order. Based on the wavelet theory, we know that if a multi-scaling function owns an approximation order $m$, this indicates that the corresponding multiwavelet function owns $m$ vanishing moments. In the following, the procedure of generating a new multiwavelet based on the multiwavelet lifting scheme and by use of an original multiwavelet with required numbers of vanishing moments will be explained. Firstly, select the original multiwavelet $\omega_0(x)$ ($\omega_0(x) = \psi_1$ or $\psi_2$) from the basis function library and a set of translation quantity $k$ of scaling functions as well as wavelet functions $\omega_1(x), \ldots, \omega_k(x)$. Next, generate the new multiwavelet by use of the “lifting coefficients equation” as follows:

$$\omega_0^{\text{new}} = \omega_0(x) + \sum_{i=1}^{k} c_i \omega_i(x)$$ (9)

If the vanishing moment of a multiwavelet required to be lifted from $p$ top, both sides of “lifting coefficients equation” are integrated. And then a set of linear equations in the matrix form is got and displayed as follows:

$$\begin{bmatrix}
\int \omega_1 x^p dx \\
\int \omega_1 x^{p+1} dx \\
\vdots \\
\int \omega_1 x^{p'-1} dx
\end{bmatrix}
\begin{bmatrix}
\int \omega_2 x^p dx \\
\int \omega_2 x^{p+1} dx \\
\vdots \\
\int \omega_2 x^{p'-1} dx
\end{bmatrix}
\begin{bmatrix}
\cdots \\
\cdots \\
\cdots \\
\cdots
\end{bmatrix}
\begin{bmatrix}
c_1 \\
c_2 \\
\vdots \\
c_k
\end{bmatrix}
= 
\begin{bmatrix}
\int \omega_0 x^p dx \\
\int \omega_0 x^{p+1} dx \\
\vdots \\
\int \omega_0 x^{p'-1} dx
\end{bmatrix}
- 
\begin{bmatrix}
\int \omega_0 x^p dx \\
\int \omega_0 x^{p+1} dx \\
\vdots \\
\int \omega_0 x^{p'-1} dx
\end{bmatrix}$$ (10)

The solutions $\{c_i\}$ of Eq. (10) are exactly the coefficients of functions which are used to perform lifting operation. Eq. (9) is carried out z-transform and the multiwavelet lifting scheme is realized successfully.

The vital step in the ensemble multiwavelet transform is the customized construction of the multiwavelet basis function. In this section, the customized multiwavelet basis function is generated on the basis of the symmetric lifting scheme [32].
Symmetry could guarantee the filter owns linear phase or generalized linear phase, which is beneficial the perfect reconstruction. However, the symmetry is not realized in the traditional multiwavelet lifting scheme. To realize symmetric multiwavelet lifting scheme, the vital factor is the appropriate selection on translation quantity \( k \) of multiwavelet basis functions [32]. Taking \( \psi_1 \) for example and supposing functions \( \omega_i \) is symmetric or anti-symmetric at the points \( a_{m} \) respectively. The selection of translation quantity \( k \) should meet the following equation:

\[
\alpha_{p_{i}}-(a_{m}+k_{a_{m}}) = (a_{m}+k_{a_{m}}) - \alpha_{p_{i}}
\]

(11)

Note that \( i = 1, 2, j = 1, 2, \ldots, k \in \mathbb{Z} \).

The symmetry of the original multiwavelet functions and multi-scaling functions can be described as:

\[
B_{m_{i}} = \pm 1
\]

(12)

where \( +1 \) notes symmetry and \( -1 \) notes anti-symmetry. Taking the symmetry conditions into account, Eq. (10) turns into the following equation:

\[
\begin{bmatrix}
\int \omega_{1}(x+k_{a_{m}})x^{p}dx & \int \omega_{1}(x+k_{a_{m}})x^{p+1}dx & \cdots & \int \omega_{1}(x+k_{a_{m}})x^{p+d}dx \\
\int \omega_{2}(x+k_{a_{m}})x^{p}dx & \int \omega_{2}(x+k_{a_{m}})x^{p+1}dx & \cdots & \int \omega_{2}(x+k_{a_{m}})x^{p+d}dx \\
\vdots & \vdots & \ddots & \vdots \\
\int \omega_{d}(x+k_{a_{m}})x^{p}dx & \int \omega_{d}(x+k_{a_{m}})x^{p+1}dx & \cdots & \int \omega_{d}(x+k_{a_{m}})x^{p+d}dx
\end{bmatrix}
\begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix}
\begin{bmatrix}
B_{m_{1}}B_{m_{2}} \\
B_{m_{1}}B_{m_{2}} \\
\vdots \\
B_{m_{1}}B_{m_{2}}
\end{bmatrix}
= \begin{bmatrix}
-\int \omega_{0}(x)x^{p}dx \\
-\int \omega_{0}(x)x^{p+1}dx \\
\vdots \\
-\int \omega_{0}(x)x^{p+d}dx
\end{bmatrix}
\]

(13)

The solutions of Eq. (14) are the coefficients for the lifting operation \( \psi_{1} \) while the lifting operation on \( \psi_{2} \) is similar to \( \psi_{1} \). Next, substitute the related lifting coefficients into the “lifting coefficients equation”. Then corresponding lifting matrices \( T \) and \( S \) can be obtained through \( Z \) transform. The directly presentation of the multiwavelet lifting scheme is displayed as follows:

\[
G_{new}(z) = T(z^{2})G(z) + S(z^{2})H(z)
\]

(14)

A new multiwavelet basis function with the symmetry property is successfully constructed based on the symmetric multiwavelet lifting scheme with the help of \( T(z) \) and \( S(z) \).

In order to construct ensemble multiwavelet basis function with specified properties and optimization objective are needed to optimize this process. Eq. (13) can be briefly described as MC = N. Where \( C = [c_{1}, c_{2}, \ldots, c_{7}]^{T} \), matrix \( M \) means the related coefficient matrix of Eq. (13). When the set of Eq. (13) is underdetermined, there are \( N = (p - p) - \text{Rank}(M) \) free parameters, which mean that ensemble symmetric lifting scheme can be conducted by the optimization of the free parameters.

Due to the property of sensitive to sharp changed structures, such as impulses, Kurtosis is usually applied to detect the incipient fault of mechanical equipment [38]. But it kurtosis index will make no sense on detecting periodical impulses. Envelope spectrum entropy can effectively indicate the definite degree of the envelope spectrum [39]. With the value becoming smaller, the periodic impulse feature of vibration signal will be more evident, which can reflect the fault development trend and is beneficial to detect compound fault and serious fault. To synthesize the preponderances of two optimization principles kurtosis \( K_{p} \) and envelope spectrum entropy \( E_{f} \), the comprehensive optimization objective is defined as follows:

\[
KE = \frac{K_{p}}{E_{f}} = \frac{\int x^{\gamma}p(x)dx}{\left( \int x^{\sigma}p(x)dx \right)^{\frac{\gamma}{\sigma}}} = \frac{1}{\sum_{i=1}^{n} p(f_{i}) \ln p(f_{i})}
\]

(15)

where \( \gamma \) and \( \sigma \) are the mean and standard deviation of multiwavelet detail coefficients \( \gamma \), and \( E(\cdot) \) is the expectation, and \( p(x) \) represents the probability density functions of the amplitude about \( \gamma \).

Here, the maximum value KE of detailed multiwavelet transform coefficients is calculated to search for the optimal multiwavelet based on the given vibration signal. Obviously, optimization method is an essential tool for searching the optimal parameter to construct the customized multiwavelet base. Genetic algorithm [40] on the basis of the idea of natural selection has an enormous advantage that it does not have mathematical requirements on the optimization problem. So, genetic algorithm is adopted to selecting the free parameters. According to our experimental experience and to increase the efficiency of the process, the parameters of genetic algorithm are set as follows: the number of iteration to 30, the range of the population scale is set to 50, the probability of crossover is set to 0.7 and the probability of mutation is set to 0.05.

### 3.3. Maximal-overlap multiwavelet transform

Construction of customized multiwavelet basis function can get the superior effect on condition feature extraction and identification. Besides the construction of customized multiwavelet basis function for the special measured vibration signals, decomposition strategy also greatly influences the condition feature extraction result of measured vibration signals. Maximal-overlap multiwavelet transform can overcome the major intrinsic deficiency of critically-sampled filter-bank and realize the translation-invariance transform which is useful on the periodical impulses extraction. Moreover, the maximal-
overlap multiwavelet transform can not only provide the richer feature information, but also provide the more precise frequency localization information. The remaining important advantage is that the maximal-overlap multiwavelet transform can overcome Gibbs phenomena effectively [26]. Based on translation operations on original vibration signal, new vibration signal with a certain phase difference on original vibration signal to change the the position of singular point, so as to reduce or eliminate the Gibbs phenomenon caused by the singularity [26].

For a given signal, anomalous amplitude can be minimized by selecting the best translation quantity on the original vibration signal. However, it is difficult to select the most appropriate translation quantity if the signal contains multiple signal singularity. In order to solve the problem of selection on appropriate translation quantity, cyclic translation operation with a range of translational quantity is carried out, and then the average operation is conducted to get the results. Based on cyclic translation operation on signal, multiwavelet basis function can approximate the real signal and describe the signal comprehensively and precisely. Based on the cyclic translation operator, maximal-overlap multiwavelet transform will be introduced in details as follows [41].

First, define the cyclic translation operator $T_h$. Let $s_{0,n}, 0 < n \leq N$ be the vector input sample points. After the cyclic translation operation on $s_{0,n}$, $H_s$ of the new signal $s_{0,n+h}$ with a certain phase difference on $s_{0,n}$ can be obtained.

$$T_h(s_{0,n}) = s_{0,n+h}, 0 < n \leq H_s$$

(16)

Note that $h$ means the translational quantity and $H_s$ is cyclic translational quantity. Accordingly, the inverse operation $T_{-h}$ of cyclic translation operator $T_h$ is like that

$$T_{-h} = (T_h)^{-1}$$

(17)

Then the $t$ maximal-overlap multiwavelet decomposition process is

$$s_{j-1,n}^{(b)} = \sum_k H_k^{2n} T_h(s_{j,k})$$

$$d_{j-1,n}^{(h)} = \sum_k G_k^{2n} T_h(s_{j,k})$$

(18)

where $s_{j,k}$ and $d_{j,k}$ are low frequency coefficients and high frequency coefficients. $s_{j-1,n}^{(b)}$ and $d_{j-1,n}^{(h)}$ are low frequency coefficients and high frequency coefficients when the translational quantity is $h$.

The reconstruction process of maximal-overlap multiwavelet transform is

$$s_{j,k}^{(b)} = T_{-h}(\sum_n H_k^{2n} s_{j-1,n}^{(b)} + \sum_n G_k^{2n} d_{j-1,n}^{(h)})$$

(19)

Note that $s_{j,k}^{(b)}$ means reconstructed signal with the same phase on orginal signal $s_{0,n}$. Then the post-processing operation is performed on $s_{j,k}^{(b)}$ to get the one-dimensional signal $y^{(b)}$.

$$P(s_{j,k}^{(b)}) = y^{(b)}$$

(20)

where $P$ means post-processing operator.

Finally, time-averaging operation is carried out

$$f_r = \frac{\sum_{i=1}^{H_s} y^{(b)}}{H_s}$$

(21)

We can find that cyclic translational quantity $H_s$ is the key parameter. In the engineering application of multiwavelet transform, the length of given signal is usual an integer power of 2. Moreover, every 2 sampling operation is adopted in wavelet decomposition and reconstruction process. So, cyclic translational quantity $H_s$ also has the length of integer power of 2. In addition, the too small value on cyclic translational quantity will lead to the difficulty on eliminating Gibbs phenomenon; the too large value on cyclic translational quantity will affect the operation speed, and excessive average operation could smooth out the local fault characteristic, which is not beneficial to fault detection and diagnosis. So, the general setting of $H_s$ in engineering application is $H_s = 16 \sim 128$.

Spatial neighboring coefficient data-driven group threshold shrinkage strategy

Cai and Silverman [31] developed the following traditional NeighCoeff threshold scheme for wavelet denoising. Define $S_{j,k}^2 = (\gamma_{j-1}^{(b)})^2 + (\gamma_{j}^{(b)})^2 + (\gamma_{j+1}^{(b)})^2$, then

$$\gamma_{j,k}^{(b)} = \begin{cases} \gamma_{j,k}^{(b)}(1 - \frac{n_j}{2\sigma_n}), & S_{j,k} > \mu_j \\ 2, & \text{otherwise} \end{cases}$$

(22)

Note that $\mu_j = \sqrt{2\sigma_n^2 \log n}$, $j$ means the wavelet transform level. $n$ and $\sigma$ respectively reflect the length and standard deviation of wavelet transform coefficients $\gamma_{j,k}^{(b)}$.

Because of multiwavelet vector operation, the correlation between different rows of transform coefficients should be considered before threshold shrinkage processing. Assume that a feature signal $f$ has been polluted by noise, and then the
actual signal can be described as follows:

\[ g[n] = f[n] + \sigma z[n] \]  \hspace{1cm} (23)\]

where \( z[n] \) means independently distributed as \( N(0, 1) \) and \( \sigma \) reflects the variance of this distribution. Denoising operation is a process of separating out of \( f \) from the noisy signal \( g[n] \) as appropriate as possible. For multiwavelet vector operation, Eq. (23) turns into multiple-stream form as follows

\[ \tilde{y}_{jk} = \tilde{\gamma}_{jk}^1 + E_{jk} \]  \hspace{1cm} (24)\]

where \( \tilde{y}_{jk} \) is the \( (1) \)th and \( (2) \)th transform coefficient of \( f \) at the transform level of \( j \). \( E_{jk} \) follows multivariate normal distribution \( N(0, V_j) \). Note that the matrix \( V_j \) means the covariance matrix of the error term at the transform level of \( j \). Then define the following standard transform

\[ \theta_{jk} = (\tilde{y}_{jk})^T (V_j)^{-1} \tilde{y}_{jk} \]  \hspace{1cm} (25)\]

Using the above standard transform, we can obtain a positive scalar value for threshold shrinkage processing [42]. In addition, the matrix \( V_j \) can be estimated by a robust covariance estimation approach based on the observed transform coefficients as follows:

Define the following pseudo-code

\[ \text{mad}(y) = 1.4826 \times \text{median}(|y - \text{median}(y)|) \]

\[ a_1 = \frac{1}{\text{mad}(\text{row}_1)} \]

\[ a_2 = \frac{1}{\text{mad}(\text{row}_2)} \]

\[ b_1 = \text{mad}(a_1 \times \text{row}_1 + a_2 \times \text{row}_2) \]

\[ b_2 = \text{mad}(a_1 \times \text{row}_1 - a_2 \times \text{row}_2) \]  \hspace{1cm} (26)\]

Note that \( \text{row}_1 \) and \( \text{row}_2 \) respectively means the two rows of multiwavelet transform coefficients.

\[ V_j = \begin{bmatrix}
\frac{1}{b_1 - b_2} & \frac{b_1 - b_2}{\|b_1 + b_2\|a_1 + a_2} \\
\frac{b_1 - b_2}{\|b_1 + b_2\|a_1 + a_2} & \frac{1}{a_1 + a_2}
\end{bmatrix} \]  \hspace{1cm} (27)\]

then

\[ S_j^2 = (\hat{\theta}_k^{(0)})^2 + (\hat{\theta}_k^{(1)})^2 + (\hat{\theta}_k^{(2)})^2 \]  \hspace{1cm} (28)\]

In traditional NeighCoeff threshold scheme, the inflexible global threshold and fixed group length is defined, which is not sufficient and reasonable. Moreover, the above NeighCoeff threshold scheme neglects the evolution of wavelet transform coefficients along scale, which usually carries important information. Hence, in view of the spatial dependence of neighboring coefficients, a flexible threshold rule with dependence to the analyzed signal is proposed in this section. Then the spatial neighboring coefficient data-driven group threshold shrinkage strategy formula is showed as:

\[ \hat{\gamma}_{jk}^{(0)} = \begin{cases}
\gamma_{jk}^{(0)} (1 - \frac{\nu_{j-k+1}^2 + \nu_{jk}^2}{\text{Sum}_{j-k+1}^2 + \text{Sum}_{jk}^2}), & \text{if } \text{Sum}_{j-k+1}^2 + \text{Sum}_{jk}^2 < \mu_{j-k+1}^2 + \mu_j^2 \\
0, & \text{otherwise}
\end{cases} \]

\[ \text{Sum}_{jk}^2 = \sum_{n=0}^{L} \theta_{jk+n}^2 \]  \hspace{1cm} (29)\]

For multiwavelet denoising with group threshold rule, the incorporation of the multiwavelet transform coefficients into the group \( k \) at the transform scale \( j \) is stated as below

\[ S_j^2 = \sum_{n=0}^{L} \theta_{jk+n}^2 \]  \hspace{1cm} (30)\]

Note that \( L \) means the length of group.

As the proposed new threshold shrinkage strategy, the length of group \( L \) and the threshold \( \mu \) are two important free parameters which can be optimized in the special multiwavelet denoising process for the given vibration signal. During every decomposition level, the suitable group length and threshold can be selected to carry out denoising operation for well recovering the important signal feature.

The two important free parameters \( L \) and \( \mu \) can be determined based on minimizing Stein’s Unbiased Risk Estimate (SURE) [29]. Let the multiwavelet transform coefficients at the sub-band \( m \) be \( w_m = \{ s_{ijk} \}, \text{ where } i, j, k \in \text{m subband} \), the concerned signal feature is \( f_m = \{ f_{ijk} \} \), where \( i \) means the \( i \)-th branch of multiwavelet transform coefficients at the transform level of \( j \) and \( k \) means the translation of the multiwavelet transform coefficients. Stein [29] demonstrated that the expected risk on the estimator \( \hat{f}_{ij} \) for the concerned or true signal feature \( f_j \) could realized the unbiased estimation on the base of multiwavelet transform coefficient \( w_m \).

Generally, assume the variance of noise is \( \sigma = 1 \), then we can obtain the following

\[ E(\|\hat{f}_j - f_j\|^2) = N_m + E(\|g(f_j)\|^2) + 2V g(f_j) \]  \hspace{1cm} (31)\]
where \( g(f_j) = \{ g_n \}_{n=1}^{N_m} \) is the basis function of the multiwavelet. Based on Eq. (14), we can obtain
\[
E[\|g_n(f_n)\|_2^2] = \|\widehat{f}_j - f_j\|_2^2 = \left\{ \begin{array}{ll}
\frac{\mu_j^2 + \sigma_j^2}{\sum_{j=1}^m \mu_j^2 + \sum_{j=1}^m \sigma_j^2} f_n^2, & \text{if } \sum_{j=1}^m \mu_j^2 + \sum_{j=1}^m \sigma_j^2 \geq \mu_j^2 + \sigma_j^2 \\
\frac{\mu_j^2 + \sigma_j^2}{\sum_{j=1}^m \mu_j^2 + \sum_{j=1}^m \sigma_j^2} f_n^2, & \text{if } \sum_{j=1}^m \mu_j^2 + \sum_{j=1}^m \sigma_j^2 < \mu_j^2 + \sigma_j^2
\end{array} \right.
\]  
(32)

SURE \( (\omega_m, \mu, L) = \sum \|g_n(f_n)\|_2^2 + 2 \sum \frac{\partial g_n}{\partial f_n} \)

Then the unbiased estimator of the expected risk can be expressed as
\[
E[\|\widehat{f}_j - f_j\|_2^2] = E[\text{SURE}(\omega_m, \mu, L)]
\]  
(34)

Based on the suitable group length \( L^m \) and threshold \( \mu^m \), the minimum of \( \text{SURE}(\omega_m, \mu, L) \) can be calculated
\[
(\mu^m, L^m) = \arg \min_{\mu, L} \text{SURE}(\omega_m, \mu, L)
\]  
(35)

As we know, the above result is calculated on the basis of assuming \( a = 1 \). Then, for the measured vibration signal without unit variance, we should first estimate the standard deviation \( \hat{\sigma} \) of the signal and normalize the variance of the multiwavelet transform coefficients. A suitable estimator for the variance can be obtained based on the median absolute value of the diagonal subband coefficients under the refined decomposition level.
\[
\hat{\sigma} = \frac{\text{median}(f_j)}{0.6745}
\]  
(36)

Define \( \gamma_d = nm^{-1/2} \log_2^{3/2} nm \) \( T_d = nm \sum (\chi_j^2 - 1) \), and \( \mu^d = 2L \log nm \) \( (\mu^d, L^d) \) means the optimal selection on group length and threshold with the minimum of \( \text{SURE} \) in the searching region under the additional constraint.
\[
(\mu^d, L^d) = \arg \min_{(\mu, L)} \text{SURE}(X, \mu, L) \text{ s.t. } (L - 2, 0) \leq (\mu, L) \leq (mn^{1/2}, 1)
\]  
(37)

The estimator \( \hat{f}^\mu(x) \) for \( f \) can be defined as follows
\[
\left\{ \begin{array}{ll}
\hat{f}_k^\mu = \hat{f}_j(\mu^d, L^d) & T_d > \gamma_d \\
\hat{f}_k^\mu = (1 - \frac{2 \log nm}{\gamma_d}) x_k & T_d \leq \gamma_d
\end{array} \right.
\]  
(38)

We can find that the above estimator degrades to James–Stein estimator whose group length is 1 when \( T_d \leq \gamma_d \).

3.4. The proposed method

It should be noted that the proposed method is usually used for fault identification of rolling mill during the accurate diagnosis part in the condition monitoring system to ensure the diagnostic efficiency. Based on the proposed method, the more accurate diagnosis result can be obtained to find out the cause of abnormal alarm. To sum up, the procedure of customized maximal-overlap multiwavelet denoising via spatial neighboring coefficient data-driven group threshold shrinkage strategy for condition monitoring and fault diagnosis of rolling mill drivetrain can be presented in the flow chart as displayed in Fig. 4. Meanwhile, the process of the proposed method for the mentioned engineering task can be summarized as follows:

1) Construct the customized multiwavelet basis function based on genetic algorithm;
2) Pre-processing the abnormal condition vibration signal and select the decomposition level;
3) Carry out the customized maximal-overlap multiwavelet decomposition operation.
4) Determine the optimal group length and threshold based on the minimizing \( \text{SURE} \);
5) Conduct spatial neighboring coefficient data-driven group threshold shrinkage strategy to get the purified detailed coefficients.
6) Reconstruct the purified decomposition coefficients and post-process the multi-stream denoising result for fault feature extraction.

4. Compound fault identification of reduction gearbox on rolling mill

In the experimental hot strip finishing mill, there are seven tandem stands with four-roller mills. In order to timely monitoring the operating condition of the drivetrain and milling stand, accelerometers and velocity transducers were mounted on the bearing pedestals of reduction gearbox and distribution box for condition monitoring system. In a weekly in-process inspection, F3 milling stand are noticed with slight abnormal sounds by the spot working staff. And it was also found by the condition monitoring system that the root mean square values of the measured signals form accelerometer 5 were remarkably greater than the other measured vibration data of the rest sensors. And therefore this sensor point was
paid the most attentions to analyzing for finding the true cause of the abnormal sound. The schematic sketch of the drivetrain in F3 milling stand is displayed in Fig. 5. The acquired vibration signal was sampled at the frequency of 5120 Hz and the length of each acquired signal is 4096. In this inspection, the rotation frequency of the input shaft was obtained by a tachometer at 4.5 Hz. And the detailed characteristic parameters of this drivetrain in the F3 milling stand are displayed in Table 1.

The acquired vibration signal from sensor 5 in time domain waveform and its Fourier spectrum as well as its envelope spectrum are respectively displayed in Figs. 6–8. From the waveform in Fig. 6, besides the interference of strong heavy background noise, we can find a faint cluster of repetitive impulses with the period at 0.2207 s, and the occurrence frequency of the mentioned impulses is basically consistent with the rotation frequency of the input shaft. From the Fourier spectrum and its envelope spectrum, there is no distinct abnormal condition feature information in the related spectrums except the meshing frequency of the gear pair in the reduction box. So, the further analysis is necessary for abnormal condition identification to find out the true cause.

Customized maximal-overlap multiwavelet denoising via spatial neighboring coefficient data-driven group threshold shrinkage strategy is applied to extract signal feature of the measured abnormal vibration data. We construct a new multiwavelet which is lifted from cubic Hermite splines as the original multi-scaling functions. There is more freedom and flexibility to construct new multiwavelet with adaptive properties because the cubic Hermite splines are simple in the waveform. In the present and latter applications to fault detection, the vanishing moment is set as $p = 4$ and the support length is 5. The original multiple scaling and wavelet functions of cubic Hermite splines is displayed in Fig. 9. And Fig. 10 is the customized multiwavelet constructed from cubic Hermite splines for the vibration signal in Fig. 6. In this analysis process, a five-level multiwavelet maximal-overlap decomposition based on the above customized multiwavelet of the vibration signal is first performed after repeated sampling and the final purified signal based on the proposed threshold

![Image of the flow chart of the proposed method.](Fig. 4. The flow chart of the proposed method.)
A shrinkage strategy is shown in Fig. 11. As shown in Fig. 11, besides the strong impulse cluster I1, another series of weak but periodic impulses I2 submerged in the heavy noise of the original signal are found. Each strong impulse I1 and weak impulse I2 emerges alternately and periodically. The occurrence frequency of the two mentioned impulses I1 and I2 are both in accordance with the rotating frequency of the pinion, which indicates that compound fault or two localized gear teeth faults of the pinion emerge in the reduction gearbox. Furthermore, we also can find that the time interval between I1 and I2 is close to one third of the pinion’s revolution, which indicates that the two localized faults had a distance of one third of the circle.

**Table 1**

Parameters of reduction gearbox in the F3 milling stand.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotating frequency of input shaft (Hz)</td>
<td>4.5</td>
</tr>
<tr>
<td>Rotating frequency of output shaft (Hz)</td>
<td>1.52</td>
</tr>
<tr>
<td>Number of teeth of the gear pair (22/21)</td>
<td>65/22</td>
</tr>
<tr>
<td>Meshing frequency of the gear pair (Hz)</td>
<td>99</td>
</tr>
<tr>
<td>Module of the pinion (mm)</td>
<td>30</td>
</tr>
<tr>
<td>Central distance (mm)</td>
<td>1350</td>
</tr>
</tbody>
</table>

**Fig. 5.** The schematic sketch of the drivetrain in F3 milling stand.

**Fig. 6.** The vibration signal of reduction gearbox.
Later, this experimental unit was shut down for further inspection. Two-site localized scuffing faults with different injury degree were found on the pinion, as shown in Fig. 12(a) and (b). According to the real operating environment, it can be inferred that the two local scuffing faults are caused by surface welding at high temperature, which are about 1/3 of the circle apart. One local scuffing fault is about 1/3 of the teeth, while the other scuffing fault spread along the whole width of the teeth. The real faults and their positions quite accord with the analyzed result based on the proposed multiwavelet denoising method.

For the purpose of comparison to demonstrate the superior performance of the proposed method, the same vibration signal is processed by traditional wavelet and multiwavelet denoising method. Daubechies wavelet, which is orthogonal, compact supported and approximately symmetric, is widely used in signal processing. GHM multiwavelet, which is one of the most popular multiwavelet, is orthogonal, symmetric and compact supported. The Db6 scalar wavelet (Daubechies wavelet with $N=6$) and GHM multiwavelet with Hard-thresholding, Soft-thresholding and NeighCoeff threshold rule are also applied to analyze the same noisy signal for comparison. The purified signals based on these wavelet denoising methods are respectively shown in Figs. 13–18. Obviously, the periodic impulses are not identified and extracted effectively.
As shown in Figs. 13–15, all the traditional wavelet denoising methods can not distinguish and extract the two different fault features. And meanwhile, all the traditional multiwavelet denoising methods based on the conventional threshold rule can only extract the strong impulses I₁, but it cannot reveal the weak impulses I₂.

Spectral kurtosis [43–45] proposed by Dwyer has been demonstrated to be a powerful tool to characterize the noisy non-stationary signal for extracting the signal feature. For the same purpose of comparison, the abnormal vibration signal is processed by SK to test the effectiveness on this task and the corresponding results are shown in Figs. 19 and 20. It can be seen from Fig. 20(a), the impulse cluster I₁ can be similarly detected and extracted. But unfortunately, the SK is also powerless for extracting the impacts I₂ at all.

Thus it can be seen that these three contrastive methods can only reveal the serious local fault on the experimental pinion of the reduction gearbox, but are all powerless to extract the weak fault feature, which will result in the incomplete condition identification conclusion. And the related analyzed results further demonstrate the superior performance of the proposed method on compound fault identification.

5. Weak fault identification of dedusting fan bearing on rolling mill

The dedusting fan as the vital component in rolling mill is used to remove the smoke and dust during steel rolling process, which is very important to assured the steel quality. The sketch map of the dedusting fan is displayed in Fig. 21. During the rolling process, the condition monitoring system alarmed for the abnormal condition and indicated that the related components of dedusting fan emerged fault condition. The further analysis is needed to find the cause of abnormal alarm. And the detailed characteristic parameters of dedusting fan bearing are displayed in Table 2.

The acquired vibration signal was sampled at the frequency of 5120 Hz by the accelerometers mounted on the bearing housing. The acquired vibration signal in time domain waveform and its Fourier spectrum as well as its envelope spectrum are respectively displayed in Figs. 22–24. Based on the vibration signal in time domain in Fig. 14, we can see that this condition vibration signal is contaminated by a large amount of background noise, and no meaningful abnormal condition feature information can be found for identifying the spindle bearing fault. Because of the interference of strong heavy background noise, there was no frequency component close to theoretical fault feature frequency for diagnosis. So, the further analysis is necessary for abnormal condition identification to find out the true cause.

The proposed multiwavelet denoising method is applied to extract signal feature of the measured abnormal vibration data. The customized multiwavelet constructed from cubic Hermite splines for the vibration signal in Fig. 22 is displayed in Fig. 25. Then, a five-level multiwavelet maximal-overlap decomposition based on this customized multiwavelet of the vibration signal is first performed after repeated sampling and the final purified signal based on the proposed threshold shrinkage strategy is shown in Fig. 26. As shown in Fig. 26, a series of periodic impulses submerged in the heavy noise of the original signal are found. And then the corresponding envelope spectrum of the purified signal is displayed in Fig. 27. It can be found from the analyzed result that the frequency component corresponding to characteristic frequency of inner ring fault can be found clearly in the spectrum, as shown in Fig. 27. According to the analyzed result, we came to the conclusion...
Fig. 12. The localized fault on the pinion: (a) fault I1 and (b) fault I2.

Fig. 13. The analysis result in case 1 using Db6 scalar wavelet Hard-threshold denoising.

Fig. 14. The analysis result in case 1 using Db6 scalar wavelet Soft-threshold denoising.

Fig. 15. The analysis result in case 1 using Db6 scalar wavelet NeighCoeff denoising.

Fig. 16. The analysis result in case 1 using GHM multiwavelet Hard-threshold denoising.
that there was a weak fault in the bearing inner ring. And meanwhile, the dedusting fan bearing needed more attention for monitoring based on effective method when running the dedusting fan from then on. A week later, this experimental unit was shut down for further inspection. Weak localized scuffing fault was found on the inner race, as shown in Fig. 28.

For the purpose of comparison to demonstrate the superior performance of the proposed method, the same vibration signal is processed by traditional wavelet and multiwavelet denoising method. The Db6 scalar wavelet and GHM multiwavelet with Hard-thresholding, Soft-thresholding and NeighCoeff threshold rule are also applied to analyze the same noisy signal for comparison. The purified signals based on these wavelet denoising methods are respectively shown in Figs. 29–34. Obviously, the periodic impulses of weak fault can not be identified and extracted effectively. The abnormal vibration signal is processed by SK to test the effectiveness on this task and the corresponding results are shown in Figs. 35 and 36. It can be seen from Fig. 36(a), the periodic impulse cluster cannot be distinctly detected and extracted.

As we see, the related analyzed results further demonstrate the superior performance of the proposed method on weak fault identification.

6. Mathematical modelling for revealing vibration signal properties

Both root cause analysis and vibration signal analysis are important study fields for mechanical condition monitoring. As a very famous scholar with the world influence on fault detection and diagnosis in mechanical systems, Bartelmus [46–49]
has carried out a lot of work and obtained many achievements. He presents the root cause analysis of the gearbox vibration
signals taking into consideration design, technology, operation, and change of condition factors, which is a useful way for
the condition monitoring of gearbox. On the basis of the above consideration, one can develop vibration signal properties
that influence the method of choice of vibration signal analysis, the gearbox degradation scenario, and the condition
inferring process [46–49]. During the rolling process, typical non-stationary operating conditions of rolling mill refer to
extreme mechanical situations including large values of tensions and forces [1]. In some cases, these extreme mechanical
situations can directly lead to different kinds of faults on rolling mill, which might bring about serious accidents and huge
economic losses.

Table 2
Parameters of the bearing in dedusting fan.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product model</td>
<td>22226EMW33C3</td>
</tr>
<tr>
<td>Rotating frequency</td>
<td>12.5 Hz</td>
</tr>
<tr>
<td>Roller pass frequency at outer ring</td>
<td>95.54 Hz</td>
</tr>
<tr>
<td>Roller pass frequency at inner ring</td>
<td>129.46 Hz</td>
</tr>
<tr>
<td>Roller pass frequency at retainer</td>
<td>79.83 Hz</td>
</tr>
<tr>
<td>Roller pass frequency at rolling element</td>
<td>5.44 Hz</td>
</tr>
</tbody>
</table>
Fig. 23. The Fourier spectrum of the vibration signal on dedusting fan bearing.

Fig. 24. The envelope spectrum of the vibration signal on dedusting fan bearing.

Fig. 25. The multiwavelet constructed from cubic Hermite splines for weak fault.

Fig. 26. The analysis result of the fan bearing vibration signal using the proposed method.

Fig. 27. The envelope spectrum of the purified signal on fan bearing vibration signal.
Fig. 28. The localized fault on the inner race of fan bearing.

Fig. 29. The analysis result in case 2 using Db6 scalar wavelet Hard-threshold denoising.

Fig. 30. The analysis result in case 2 using Db6 scalar wavelet Soft-threshold denoising.

Fig. 31. The analysis result in case 2 using Db6 scalar wavelet NeighCoeff denoising.

Fig. 32. The analysis result in case 2 using GHM multiwavelet Hard-threshold denoising.
In addition, due to the complex mechanical structure and various operation environment, effective non-stationary signal processing method is indispensable and should be developed and introduced for condition monitoring and fault diagnosis of rolling mill drivetrain. In this paper, we focus on developing the effective signal processing method to extract fault feature from non-stationary vibration signal for gearbox condition monitoring. So, it is necessary to reveal the vibration signal properties of gearbox when a fault emerges. Bartelmus also has conducted the related research in this field by mathematical modelling and computer simulation [46–49]. Based on his previous work, we also perform the study on mathematical modelling for revealing vibration signal properties of reduction gearbox in corresponding rolling mill, which is conducive to identify gearbox fault feature from vibration signals on the basis of multiwavelet method for condition monitoring and fault diagnosis.

6.1. The coupled lateral and torsional vibrations dynamic model of gear system

As shown in Fig. 37(a), a typical gear system is consisted of gear pairs, bearings, elastic shafts, drive motors and loads. Fig. 37(b) shows the dynamic model of the gear system. Rotating shafts of the system are modeled as Timoshenko beams; the gear mesh is modeled as a pair of rigid disks connected by a spring-damper with time-varying mesh stiffness.
Fig. 36. The analysis result in case 2 using SK: (a) purified signal and (b) envelope spectrum.

Fig. 37. (a) A typical gear system (b) Finite element model of the gear system.

Fig. 38. Typical dynamic model of gear pair.
A schematic representation of a single stage gear system is illustrated in Fig. 38. The displacement vector of a gear pair can be defined from the line co-ordinate system, the central coordinate vector can be expressed as:

\[
\mathbf{q_G} = \begin{bmatrix} u_1 v_1 \theta_1 \\ u_2 v_2 \theta_2 \end{bmatrix}
\]

where \(u, v, \theta\) are the lateral degrees of freedom and \(\theta_1, \theta_2\) are the torsional degrees of freedom. The subscripts “1” and “2” indicate the driving gear and the driven gear respectively. \(O_1\) and \(O_2\) represent centers of the gears in the static state, while \(O'_1\) and \(O'_2\) represent centers of the gears in the operational rotating state. \(G_1\) and \(G_2\) represent geometrical centers of the gears. The equations of motion of a gear pair in a matrix form can be given by

\[
M_G \ddot{\mathbf{q}_G} + (C_G + G_G) \dot{\mathbf{q}_G} + K_G \mathbf{q}_G = \mathbf{Q}_G
\]

where \(M_G, C_G, G_G, K_G\) are the inertia, damping, gyroscopic and stiffness matrices, respectively. \(\mathbf{Q}_G\) is the external force vector in the local pressure line co-ordinate system. They are given in Refs. [50, 51], which will not be presented here for simplicity.

The system equations of motion can be obtained according to the relation between the gear pair displacement vector and the corresponding node displacement of shaft which are modeled as Timoshenko beams.

\[
M \ddot{\mathbf{q}} + (G + C) \dot{\mathbf{q}} + K \mathbf{q} = \mathbf{Q}
\]

where \(M, K, G, C\) are the inertia, damping, gyroscopic and stiffness matrices of the gear system shown in Fig.1, respectively. \(\mathbf{Q}\) is the external force vector.

Since time-varying mesh stiffness is the main source of excitation in gear dynamic system, it is necessary to precisely calculate the mesh stiffness in order to get the dynamic responses of gear system. In this paper, the gear time varying mesh stiffness is calculated by the potential energy method [50, 51]. An example of a single-stage spur gear system is presented in Table 3. The gear mesh stiffness of the gear pair for different crack size is shown in Fig. 39. It shows that the mesh stiffness will reduce when the cracked tooth on pinion comes into contact with the gear during meshing. It reduces further with increase in crack length.

### Dynamic simulation of gear-rotor system with tooth root crack

In this section, it is assumed that the crack length is 2 mm. The mesh stiffness calculated by potential energy method is plugged into the dynamic Eq. (41), then use the integral calculus method of Newmark-\(\beta\) to solve the equation, and get each node’s acceleration signal. The results are presented in Fig. 40. The simulated signal shows some periodic impulses caused by tooth crack. The impulse period is 0.190 s, which is equal to the rotational period of the defected gear. Based on the above
dynamic simulation results, we can extract fault feature by customized maximal-overlap multiwavelet denoising method to monitor gearbox condition.

7. Conclusion

In this paper, condition monitoring system for rolling mill drivetrain is designed and developed on the basis of Labview. And customized maximal-overlap multiwavelet denoising via spatial neighboring coefficient data-driven group threshold shrinkage strategy is proposed for this task. The effectiveness of wavelet denoising method mainly relies on the appropriate selections of wavelet base, transform strategy and threshold rule. So, the proposed method focus on the following three parts to enhance the ability of wavelet denoising method such as construction of customized multiwavelet basis function, design the maximal-overlap multiwavelet transform strategy, and spatial neighboring coefficient data-driven group threshold shrinkage strategy. The effectiveness of proposed method is first demonstrated through compound fault identification of reduction gearbox on rolling mill. Then this method is applied for weak fault identification of dedusting fan bearing on rolling mill to show its superior performance.

According to the above results, there are several future research issues for study. First of all, due to the vital influence on the effectiveness of wavelet denoising method, the method of constructing the customized multiwavelet basis function is significant. So, strong adaptivity multiwavelet basis function for the given vibration signal should be designed based on the more effective construction method. Moreover, it is also very important to confirm the excellent evaluation index for judging the adaptivity of constructed multiwavelet basis function. In this paper, a compound index called kurtosis—envelope spectrum entropy is proposed and used for selecting the customized multiwavelet basis function for the given vibration signal. But in fact, this process is on the basis of the indirect evaluation system, which may select a multiwavelet basis function under a certain indirect evaluation index. However, this indirect evaluation index is selected by the user. So, a more direct and objective evaluation system should be developed for engineering application. All in all, customized multiwavelet analysis method can take more effect on the condition monitoring and fault diagnosis in engineering application but there are there are still abundant issues to be addressed.

In addition, the authors have been attracted some interesting results and will carry out the related study in the further. First, we have known that fast-Fourier transform doesn’t make much sense when measured vibration data is collected under the non-stationary operating regime. But, one use the short fast Fourier transform connected with additional signal processing given in paper by Bartelmus et al. [52] that using the fast-Fourier transform gives successful results under the non-stationary operating regime. Authors will conduct the corresponding study on combining fast-Fourier transform and the customized multwavelet method. Next, the paper has given a scheme of a gear system, which consists of two gear wheels the same as the scheme in Ref. [53]. But in fact, a complete gear system should have a driving element (engine), coupling, two gear wheels and a driven element. It would very valuable for future investigations to take into consideration the complete gear system. The complete gear system will be studied for more accurate simulation results.

Acknowledgements

This research is supported financially by the Project of National Natural Science Foundation of China for Innovation Research Group (Grant no. 51421004), National Natural Science Foundation of China (Grant no. 51405379), China Postdoctoral Science Foundation (Grant no. 2014M562396, Grant no. 2015T81017), Fundamental Research Funds for the Central Universities of China (Grant no. XJJ2015106, Grant no. CXTD2014001) and Shaanxi Industrial Science and Technology Project (Grant no. 2015GY121).

The authors would like to sincerely thank all the anonymous reviewers for the valuable comments that greatly helped to improve the manuscript.

References
