Squared Envelope Sparsification via Blind Deconvolution and its Application to Railway Axle Bearing Diagnostics

Bingyan Chen 1,2, Weihua Zhang 1,*, Dongli Song 1, Yao Cheng 1, Fengshou Gu 1,2, Andrew D Ball 2

1 State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China
2 Centre for Efficiency and Performance Engineering, University of Huddersfield, Huddersfield HD1 3DH, UK

*Corresponding author

E-mail addresses: chenbingyan@my.swjtu.edu.cn (Bingyan Chen), tpl@swjtu.edu.cn (Weihua Zhang), sdl.cds@163.com (Dongli Song), chengyao2020@swjtu.edu.cn (Yao Cheng), f.gu@hud.ac.uk (Fengshou Gu), a.ball@hud.ac.uk (Andrew D Ball)

Abstract

A sparse squared envelope is crucial for efficient and accurate diagnosis of bearing faults. Blind deconvolution is a well-established sparse feature enhancement method for the diagnostics of rolling bearings. Traditional blind deconvolution methods, such as minimum entropy deconvolution, are susceptible to random transients, making it difficult to enhance fault features of rolling bearings subject to strong random shocks. Deconvolution methods that take the fault characteristic frequency (or fault impulse period) of interest as an algorithm input parameter, such as maximum second-order cyclostationarity blind deconvolution, can alleviate this deficiency. However, bearing fault features are difficult to be enhanced by these methods when the specified characteristic frequency deviates from the actual value greatly. To overcome these problems, the modified smoothness index of the squared envelope is proposed as the objective function of the deconvolution method, and a new blind deconvolution method is developed to achieve a sparse squared envelope for fault diagnosis of rolling bearings. Furthermore, the methodology is extended to the frequency domain, and another new blind deconvolution method that utilizes the modified smoothness index of the squared envelope spectrum as the objective function is established to achieve a sparse squared envelope spectrum for bearing diagnostics. These two proposed blind deconvolution methods are robust to random transients and do not require characteristic frequency or impulse period as an input parameter for feature enhancement. The performance of the two proposed blind deconvolution methods is verified on experimental datasets from two different railway axle bearing test rigs and compared with the state-of-the-art deconvolution methods. The results show that the two proposed methods can effectively enhance repetitive transient features in noisy vibration signals and accurately diagnose different faults of railway axle bearings.

Keywords: fault diagnosis, railroad bearings, feature enhancement, blind deconvolution, modified smoothness index, sparsity measures

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BD</td>
<td>Blind Deconvolution</td>
</tr>
<tr>
<td>CYCBD</td>
<td>Maximum Second-Order Cyclostationarity Blind Deconvolution</td>
</tr>
<tr>
<td>EVA</td>
<td>Eigenvalue Algorithm</td>
</tr>
<tr>
<td>FDSNR</td>
<td>Frequency-Domain Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>GI</td>
<td>Gini Index</td>
</tr>
<tr>
<td>HI</td>
<td>Hoyer Index</td>
</tr>
</tbody>
</table>
1. Introduction

Rolling bearing is one of the basic and common rotating components in modern mechanical equipment. Their health condition has a significant influence on the overall operational performance of rotating machinery, especially complex mechanical systems such as high-speed train bogies, aircraft engines, wind turbines, automobile engines and marine engines. Therefore, timely and accurate fault diagnosis is an effective means to avoid prolonged unplanned machine downtime, prevent major personal injury, and improve maintenance efficiency. Surface defects in bearing components, such as fatigue spalling, pitting and cracks, often excite repetitive transient impulses in the bearing’s vibration response signal, and they are critical to implementing fault diagnosis 1,2. However, due to the harsh operating conditions, the repetitive transient features associated with bearing faults are usually submerged in strong and complex interference noise, which heavily hinders the detection and diagnosis of bearing faults. In particular, railway axle bearings, whose vibration signals often exhibit low signal-to-noise ratios and complex compositions. This is because railway trains run dynamically along the rails, resulting in vibration signals measured on axle boxes that typically contain: (i) vibration noise caused by track irregularities or track defects 3; (ii) vibration noise induced by connected transmission systems such as traction motors and gearboxes 4; (iii) vibration noise excited by wheel out-of-round such as wheel polygon, flat scar and scratch 2,5. These unfavorable factors lead to challenges in fault diagnosis of railway axle bearings. Therefore, effective methods or techniques for reducing interference noise in the measured vibration signals while enhancing fault-related features need to be applied to achieve accurate diagnosis of bearing faults.

Fortunately, various bearing fault diagnosis methods have been established with the help of the feature enhancement and noise cancellation capabilities of signal processing techniques, such as envelope analysis 6–8, blind deconvolution (BD) 9,10, adaptive decomposition 11–13, stochastic resonance 14,15, morphological filtering 16,17 and modulation bispectrum 18,19. Among these methods, BD is an advanced and widely applied sparse feature enhancement method for fault diagnosis of rolling bearings. Its basic principle is to devise an inverse filter to reconstruct the repetitive transient impulses caused by
bearing defects from the measured signal by optimizing an appropriate objective function. There are two popular methods for solving inverse filters: the objective function method (OFM) and the eigenvector algorithm (EVA)\textsuperscript{20,21}. OFM requires that the objective function of deconvolution be differentiable, i.e., it is not suitable for solving the deconvolution problem of non-differentiable objective functions. EVA can give equivalent filter coefficients and there is no restriction on whether the objective function is differentiable or not. Therefore, the usable range of EVA is broader than that of OFM. In recent years, two other methods for solving inverse filters have been developed, namely the particle swarm optimization (other optimization algorithms can also be employed) and generalized spherical coordinate transformation-based method\textsuperscript{22,23} and the convolutional sparse learning-based method\textsuperscript{24–27}. These schemes extend the solution methods of deconvolution problems, while the complexity of the filter optimization process is increased compared with OFM and EVA.

Minimum entropy deconvolution (MED)\textsuperscript{28} with kurtosis as the objective function is an earlier BD method applied in the field of machinery fault diagnosis\textsuperscript{29,30}. However, kurtosis is easily affected by random transients or outliers in the signal, causing MED to tend to enhance random transients in the vibration signal that are unrelated to bearing faults. To enhance the capability of repetitive transient recovery, a series of BD methods with improved performance have been developed, which can be divided into two categories from the perspective of the objective function: impulse period or characteristic frequency independent and impulse period or characteristic frequency dependent. The former uses metrics (computed from the time or frequency domain) independent of the impulse period or characteristic frequency as the objective function of deconvolution, such as D-norm\textsuperscript{31}, Jarque-Bera statistic\textsuperscript{32}, impulse norm\textsuperscript{33}, envelope spectrum sparsity indicators\textsuperscript{34}, generalized Lp/Lq norm\textsuperscript{35–37}, Gini index (GI)\textsuperscript{38} and Box-Cox sparse measures\textsuperscript{39}. Note that MED also falls into this category. These methods can adaptively enhance transient features hidden in the bearing vibration signals, but some of them may not perform well in the presence of strong random transients or other interfering noise, such as only a single impulse or a few impulse features may be enhanced. The latter uses metrics (calculated from the time or frequency domain) that rely on the impulse period or characteristic frequency as the objective function of deconvolution, such as correlated kurtosis\textsuperscript{40}, multi D-norm\textsuperscript{41}, harmonic-to-noise ratio\textsuperscript{42}, indicator of second-order cyclostationarity\textsuperscript{21}, autocorrelation impulse harmonic to noise ratio\textsuperscript{43}, average kurtosis\textsuperscript{44}, correlated generalized Lp/Lq norm\textsuperscript{45} and improved correlated generalized Lp/Lq norm\textsuperscript{46}. These methods exhibit relatively robust repetitive transient enhancement performance in the presence of high-amplitude random transients with the help of impulse period or characteristic frequency knowledge. However, when the impulse period or characteristic frequency of interest deviates from the actual value greatly, they cannot effectively recover the desired repetitive transient features, resulting in poor fault diagnosis performance. A recent work\textsuperscript{9} summarizes the development and application of the BD methods in the field of machinery fault diagnosis.

It is widely recognized that bearing fault signals exhibit typical impulsiveness and cyclostationarity. The objective functions computed from the time domain (e.g., the original signal and its squared envelope) mainly evaluate the impulsiveness of the bearing fault signal, as in\textsuperscript{28,31–33,35,36,38}, while the objective functions computed from the frequency domain (e.g., the squared envelope spectrum) focus on evaluating the cyclostationarity of bearing fault signals, as in\textsuperscript{21,34,37–39}. The use of envelope spectrum sparsity indicators to reveal cyclostationarity of bearing fault features initially began with envelope analysis. Barszcz and Jabłoński\textsuperscript{47} proposed to use the kurtosis of the envelope spectrum of the demodulated signal instead of the kurtosis of the filtered signal to detect repetitive transients. After that, Antoni\textsuperscript{48} proposed to use the negentropy (NE) of the squared envelope and of the squared envelope
spectrum of the signal to detect the impulsiveness and cyclostationarity of repetitive transients, respectively. These are typical applications for measuring the sparsity of the envelope spectrum. Maximum second-order cyclostationarity blind deconvolution (CYCBD) \(^{21}\) is a pioneering BD method aiming at enhancing cyclostationarity, but its objective function, the indicator of second-order cyclostationarity, requires the characteristic frequency of interest as a priori knowledge. To overcome this limitation, Peeters et al. \(^{34}\) developed blind filtering based on envelope spectrum sparsity indicators including NE, L2/L1 and Hoyer index (HI) to enhance repetitive transients in machine vibration signals. This work is the first to propose the idea of using the sparsity measure of the squared envelope spectrum as the objective function of BD to remove the need for an exact priori knowledge of fault characteristic frequency while maintaining robustness to impulse noise. Inspired by this work, the BD methods \(^{37-39}\) based on generalized Lp/Lq norm, GI and Box-Cox sparse measures of the squared envelope spectrum have been proposed and applied to fault diagnosis of rotating machinery.

As mentioned above, sparsity measures are independent of the impulse period or characteristic frequency and enable adaptive enhancement of bearing fault features when used as the objective function for BD. In addition to kurtosis, generalized Lp/Lq norm, NE and GI, the smoothness index \(^{49}\) (or its reciprocal \(^{50}\)) is also a sparsity measure commonly used to quantify transient features associated with bearing faults. The smoothness index was used early in speech signal processing \(^{51}\), and its statistical interpretation in terms of a measure of extreme non-Gaussianity was recently introduced in \(^{52}\). Differently, the smoothness index gradually decreases as the data sequence becomes sparse. In a recent work \(^{53}\), the modified smoothness index (MSI) was proposed as an alternative. MSI not only has a limited magnitude range of [0, 1] but also exhibits the same evolution characteristics as kurtosis when transient features appear. In addition, the random transient resistibility of MSI is slightly stronger than that of GI, which may be more suitable for the cases of random transients. Therefore, the bearing fault diagnosis method based on MSI is very worth to be investigated.

To establish a bearing fault diagnosis method that can adaptively enhance fault-related repetitive transient features in the signal and is robust to random transients, inspired by the work of Peeters et al. \(^{34}\) and Miao et al. \(^{38}\), the BD methods based on MSI are investigated and developed in this paper for enhancing repetitive transients. The main novelties and contributions of this paper are described as follows:

1. MSI of the squared envelope is proposed as the objective function of deconvolution, and a new BD method is devised to achieve sparse squared envelop for fault diagnosis of rolling bearings, named squared envelope sparsification via BD with MSI (abbreviated as SES-BD).

2. The sparse feature enhancement is extended to the frequency domain, and a new BD method that takes MSI of squared envelope spectrum as the objective function, named squared envelope spectrum sparsification via BD with MSI (abbreviated as SESS-BD), is developed to generate sparse squared envelope spectrum for bearing diagnostics.

3. The inverse filters of the SES-BD and SESS-BD methods for transient feature enhancement are derived by EVA, respectively, and the corresponding deconvolution theories are given in detail.

4. The effectiveness and advantages of the two proposed BD methods are verified by bearing experimental data collected from two different railway axle bearing test rigs and by comparison with several existing BD methods.

The rest of this paper is structured as follows. In Section 2, the theoretical background of BD is briefly reviewed. Section 3 details the two proposed BD methods. In Sections 4 and 5, the two proposed BD methods are validated using experimental signals collected from two different railway axle bearing
test rigs, respectively. Section 6 quantitatively analyzes the fault diagnosis performance of the two proposed BD methods. Finally, the main conclusions of this paper are summarized in Section 7.

2. Theoretical background

In this section, the deconvolution problem is first illustrated, followed by a brief review of typical BD methods for bearing fault diagnosis.

2.1. Deconvolution problem statement

The target of deconvolution is to restore the targeted signal component \( s \) or to find an approximate estimate of it from the measured signal \( x \) by an inverse filter \( h \). The deconvolution problem is modeled as follows:

\[
s = x * h
\]

(1)

where * denotes the convolution operation. The matrix form of equation (1) is

\[
s = Xh
\]

(2)

\[
\begin{bmatrix}
s[1] \\
s[2] \\
s[3] \\
\vdots \\
s[N]
\end{bmatrix} = \begin{bmatrix}
x[1] \\
x[2] \\
x[3] \\
\vdots \\
x[N-1]
\end{bmatrix} * \begin{bmatrix}
x[1] \\
x[2] \\
x[3] \\
\vdots \\
x[N-L]
\end{bmatrix} = \begin{bmatrix}
h[1] \\
h[2] \\
h[3] \\
\vdots \\
h[L]
\end{bmatrix}
\]

(3)

where \( N \) and \( L \) are the lengths of the measured signal \( x \) and inverse filter \( h \), respectively. Note that this paper uses the adjusted convolution definition proposed by McDonald and Zhao to avoid introducing spurious impulses at the start of the filtered signal. This adjustment reduces the length of the filtered signal by \( L \) compared to the original signal but generally does not affect machine fault diagnosis.

The observed signal output by the sensor is usually a mixture of the content of interest and the unavoidable noise as follows:

\[
x = c_s * g_c + c_n * g_n
\]

(4)

where \( c_s \) and \( c_n \) represent the content of interest and the noise component, respectively; \( g_c \) and \( g_n \) denote the effects of transmission paths associated with \( c_s \) and \( c_n \), respectively. In actual industrial scenarios, the collected machine vibration signals are usually multi-component signals. For fault diagnosis of rotating machine components such as bearings and gears, the content of interest \( c_s \) is the repetitive transients caused by local defects in the machine components. Due to the complex operating conditions, the interference noise \( c_n \) usually includes background noise with approximately Gaussian distribution, random impulses caused by unknown shocks and discrete harmonics caused by eccentricity and other factors.

Substituting equation (4) into equation (1), the deconvolution problem is expressed as:

\[
s = (c_s * g_c + c_n * g_n) * h \approx c_s
\]

(5)

Thus, deconvolution aims to design an optimal filter \( h \) to reconstruct the desired component \( c_s \) while eliminating the unwanted component \( c_n \). A schematic diagram of the deconvolution process for restoring the content of interest is shown in Figure 1.

2.2. Typical deconvolution methods for bearing diagnostics

Repetitive transients caused by defects in rolling bearings are impulsive and periodic (strictly
In response to these two typical characteristics, two categories of deconvolution methods have been established: impulse period or characteristic frequency-independent BD and impulse period or characteristic frequency-dependent BD. The main difference between these two categories of BD methods is whether the objective function used relies on the impulse period or the characteristic frequency associated with bearing fault. Typical deconvolution methods for bearing fault diagnosis and their details are depicted in Table 1.

![Schematic diagram of the deconvolution process.](image)

**Table 1.** Typical deconvolution methods for bearing diagnostics and their objective functions.

<table>
<thead>
<tr>
<th>Category</th>
<th>Method</th>
<th>Objective function</th>
<th>Inverse filter solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impulse period/characteristic frequency-independent BD</td>
<td>MED $^{28}$</td>
<td>Kurtosis</td>
<td>OFM</td>
</tr>
<tr>
<td></td>
<td>Optimal MED $^{31}$</td>
<td>D-norm</td>
<td>OFM</td>
</tr>
<tr>
<td></td>
<td>Blind filters based on envelope spectrum sparsity indicators $^{34}$</td>
<td>NE, L2/L1 and HI of the squared envelope spectrum</td>
<td>EVA</td>
</tr>
<tr>
<td></td>
<td>Maximum Gini index deconvolution (MGID) and maximum Gini index deconvolution (MESGID) $^{38}$</td>
<td>GI of the squared envelope spectrum and GI of the squared envelope spectrum</td>
<td>EVA</td>
</tr>
<tr>
<td></td>
<td>Blind filtering based on Box-Cox sparse measures $^{39}$</td>
<td>Box-Cox sparse measures of the squared envelope spectrum</td>
<td>EVA</td>
</tr>
<tr>
<td>Impulse period/characteristic frequency-dependent BD</td>
<td>Maximum correlated kurtosis deconvolution (MCKD) $^{40}$</td>
<td>Correlated kurtosis</td>
<td>OFM</td>
</tr>
<tr>
<td></td>
<td>Multipoint optimal minimum entropy deconvolution adjusted (MOMEDA) $^{41}$</td>
<td>Multi D-norm</td>
<td>OFM</td>
</tr>
<tr>
<td></td>
<td>Sparse maximum harmonics-to-noise-ratio deconvolution (SMHD) $^{42}$</td>
<td>Harmonics-to-noise-ratio</td>
<td>OFM</td>
</tr>
<tr>
<td></td>
<td>CYCBD $^{21}$</td>
<td>Indicator of second-order cyclostationarity</td>
<td>EVA</td>
</tr>
</tbody>
</table>
The BD methods that are independent of impulse period or characteristic frequency can adaptively extract transient features in the signal, but the fault diagnosis performance of such methods is sometimes seriously affected by high-amplitude random transients. Among the objective functions of these methods, GI is more robust to random transients than kurtosis, NE, L2/L1 and HI,53 thereby GI-based BD methods may be more effective in the presence of random transients. However, the repetitive transient detectability of GI is lower than that of kurtosis and NE under noise interference.57 Moreover, railway axle bearings serve on bogies that run dynamically along the rails, resulting in their vibration signals usually containing complex background noise and strong random transients from track irregularities and track defects. These unfavorable factors may cause the failure of GI-based BD methods in detecting axle bearing faults.

Aided by knowledge of impulse period or characteristic frequency, another category of BD methods represented by MCKD, MOMEDA, SMHD and CYCBD exhibits relatively strong capability to enhance repetitive transient features in noisy measurement signals. Among such methods, CYCBD delivers strong background noise resistibility compared to MCKD, MOMEDA and SMHD due to the use of an objective function computed from the frequency domain. However, these methods all require the exact impulse period or characteristic frequency as an input parameter and can only reconstruct repetitive impulse features of the specified period or frequency. The feature enhancement and diagnostic performance of such methods are severely degraded when the impulse period or characteristic frequency is inaccurate.33 In addition, due to the frequent acceleration and braking of railway trains, the running speed of the wheelset often fluctuates and the impulse period and characteristic frequency of interest are difficult to obtain accurately. These unavoidable factors may cause the BD methods that rely on impulse period or characteristic frequency to fail when diagnosing axle bearing faults.

3. Blind deconvolution driven by modified smoothness index

In this section, the method for solving the MSI-based deconvolution problem is first introduced, and then two proposed BD methods based on MSI, namely SES-BD and SESS-BD, are elaborated.

3.1. Solution method of MSI-based deconvolution problem

MSI requires the input sequence to be non-negative. However, the amplitudes of the collected bearing vibration signals are generally both positive and negative. Therefore, a modulo operation is usually performed on the signal sequence before calculating MSI. Based on the demodulation capability of the Hilbert transform, the squared envelope and the squared envelope spectrum can not only highlight the bearing fault-related features but also have non-negative amplitudes. In this section, the MSI of the squared envelope and the MSI of the squared envelope spectrum are proposed as objective functions for deconvolution, respectively, which lead to MSI being non-differentiable. This means that OFM cannot be used to solve MSI-based deconvolution problems. Fortunately, this problem can be solved by EVA. Thus, EVA is adopted to solve the inverse filters of two proposed BD methods for enhancing bearing fault-related features.

3.2. Squared envelope sparsification via blind deconvolution with modified smoothness index

The objective function of SES-BD is the MSI of the squared envelope, which is dedicated to finding an optimal inverse filter to maximize the MSI of the squared envelope of the filtered signal, i.e., achieve a sparse squared envelope. Based on equation (2), the squared envelope can be expressed as 21,34:
where the superscript “$T$” represents the transpose of a vector or matrix. The MSI of the squared envelope is defined as:

$$SE = |x|^2 = |Xh|^2 = \text{diag} \left( s^T \right) Xh$$  \hfill (6)

The numerator and denominator of equation (7) can be respectively formulated as:

$$\sqrt{\prod_{n=1}^{N} SE[n]} = SE^T \text{diag} \left( \frac{\sqrt{\prod_{n=1}^{N} SE[n]}}{SE^T SE} \right) SE$$  \hfill (8)

$$\frac{1}{N} \sum_{n=1}^{N} SE[n] = SE^T \text{diag} \left( \frac{1}{N|SE|} \right) SE$$  \hfill (9)

Thus, equation (7) is expressed in matrix form as:

$$SE^T \text{diag} \left( \frac{N \cdot \sqrt{\prod_{n=1}^{N} SE[n]}}{SE^T SE} \right) SE$$

$$MSI = 1 - \frac{SE^T \text{diag} \left( \frac{1}{SE} \right) SE}{SE^T \text{diag} \left( \frac{N \cdot \sqrt{\prod_{n=1}^{N} SE[n]}}{SE^T SE} \right) SE}$$ \hfill (10)

where $W_1$ and $W_2$ are two diagonal matrices, as follows:

$$W_1 = \text{diag} \left( \frac{1}{SE} - \frac{N \cdot \sqrt{\prod_{n=1}^{N} SE[n]}}{SE^T SE} \right)$$  \hfill (11)

$$W_2 = \text{diag} \left( \frac{1}{SE} \right)$$

Substituting equation (6) into equation (10), the following expression can be obtained:

$$MSI = \frac{h^T X^T \text{diag} \left( s \right) W_1 \text{diag} \left( s^T \right) Xh - h^T R_{sw,x} h}{h^T X^T \text{diag} \left( s \right) W_2 \text{diag} \left( s^T \right) Xh - h^T R_{sw,x} h}$$  \hfill (12)

where $R_{sw,x}$ and $R_{sw,x}$ are two weighted correlation matrices, as follows:

$$R_{sw,x} = X^T \text{diag} \left( s \right) W_1 \text{diag} \left( s^T \right) X$$

$$R_{sw,x} = X^T \text{diag} \left( s \right) W_2 \text{diag} \left( s^T \right) X$$  \hfill (13)

Equation (12) is a generalized Rayleigh quotient whose maximization with respect to $h$ is equivalent to solving the maximum eigenvalue $\lambda$ of the generalized eigenvalue problem as follows:
\[ R_{XW,h} = \lambda R_{XW,\h} \]  \hspace{2cm} (14)

The maximum eigenvalue obtained by equation (14) corresponds to the maximum MSI and the eigenvector corresponding to the maximum eigenvalue is equivalent to the optimal inverse filter \( h \). However, the optimal filter \( h \) cannot be achieved directly by equation (14) because the computation of \( R_{XW,h} \) and \( R_{XW,\h} \) requires an initial guess of \( h \). Therefore, the equivalence between the maximum eigenvalue \( \lambda \) and the maximum MSI (i.e., the optimal \( h \)) can only be achieved by an iterative algorithm. Based on the aforementioned theories and derivations, a flowchart of the SES-BD method for bearing fault diagnosis is depicted in Figure 2. The implementation procedure of SES-BD is described as follows:

**Figure 2. Flowchart of the SES-BD method for bearing diagnostics.**

**Step 1:** Collect the vibration acceleration signal of the monitored rolling bearing.
**Step 2:** Set filter length and iteration stopping criterion.
**Step 3:** Estimate an initial guess for the inverse filter \( h \).
**Step 4:** Calculate the filtered signal \( s \) using \( X \) and guessed \( h \), then calculate \( W_1 \) and \( W_2 \) using the squared envelope.
**Step 5:** Compute \( R_{XW,h} \) and \( R_{XW,\h} \) using \( s, X, W_1 \) and \( W_2 \).
**Step 6:** Solve equation (14) to get the eigenvector corresponding to the maximum eigenvalue, i.e. the filter \( h \).
**Step 7:** Return to Step 4 to further optimize the inverse filter using \( h \) updated in Step 6 until the iteration stopping criterion is met.
**Step 8:** Generate the optimal filtered signal using \( X \) and optimal \( h \).
**Step 9:** Perform envelope spectrum analysis on the filtered signal to diagnose bearing faults.
In this study, the initial guess of the inverse filter is achieved by a whitening filter, which is the same strategy used in 21. This initial filter can attenuate the deterministic components of the bearing vibration signal. In addition, a double criterion is used to stop the iterative process: either the relative error of the maximum eigenvalues obtained by two consecutive iterations is less than a certain value or the maximum number of iterations is satisfied.

3.3. Squared envelope spectrum sparsification via blind deconvolution with modified smoothness index

The objective function of SESS-BD is the MSI of the squared envelope spectrum, which aims to devise an optimal inverse filter to maximize the MSI of the squared envelope spectrum of the filtered signal, i.e., achieve a sparse squared envelope spectrum.

Based on equation (6) and discrete Fourier transform, the squared envelope spectrum is expressed as 21,34:

\[
S_{ES} = F^T \text{diag} \left( s^T \right) X h
\]  

(15)

where \( K \) is the maximum index of the frequency range of interest in the squared envelope spectrum. The MSI of the squared envelope spectrum is defined as:

\[
\text{MSI} = 1 - \frac{\sqrt{\prod_{k=1}^{K} S_{ES}[k]}}{\frac{1}{K} \sum_{k=1}^{K} S_{ES}[k]}
\]  

(17)

The numerator and denominator of equation (17) can be respectively formulated as:

\[
\sqrt{\prod_{k=1}^{K} S_{ES}[k]} = S_{ES}^T \text{diag} \left( \prod_{k=1}^{K} S_{ES}[k] \right) S_{ES}
\]  

(18)

\[
\frac{1}{K} \sum_{k=1}^{K} S_{ES}[k] = S_{ES}^T \text{diag} \left( \frac{1}{K[S_{ES}]} \right) S_{ES}
\]  

(19)

Equation (17) is expressed in matrix form as:
\[
MSI = 1 - \frac{SES^T diag \left( K \cdot \sqrt{\prod_{k=1}^{K} SES[k]} \right) SES}{SES^T diag \left( \frac{1}{|SES|} \right) SES}
\]

(20)

where \( W_1 \) and \( W_2 \) are two diagonal matrixes, as follows:

\[
W_1 = diag \left( \frac{1}{|SES|} K \cdot \sqrt{\prod_{k=1}^{K} SES[k]} \right)
\]

\[
W_2 = diag \left( \frac{1}{|SES|} \right)
\]

(21)

Then, substituting equation (15) into equation (20), the MSI can be further expressed as follows:

\[
MSI = \frac{h^T X^T diag(s) FW_1 F^T diag(s^T) Xh}{h^T X^T diag(s) FW_2 F^T diag(s^T) Xh} = h^T R_{w,1} h
\]

(22)

\[
\begin{align*}
R_{w,1} &= X^T diag(s) FW_1 F^T diag(s^T) X \\
R_{w,2} &= X^T diag(s) FW_2 F^T diag(s^T) X
\end{align*}
\]

(23)

Similar to equation (12), the maximization of the generalized Rayleigh quotient in equation (22) with respect to \( h \) is equivalent to solving the maximum eigenvalue \( \lambda \) of the following generalized eigenvalue problem:

\[
R_{w,1} h = \lambda R_{w,2} h
\]

(24)

Therefore, the eigenvector \( h \) corresponding to the maximum eigenvalue of equation (24) is the optimal inverse filter for SESS-BD. The implementation procedure of SESS-BD is similar to that of SES-BD, thus the related description and flowchart are not given. The only difference between them is the calculation of the weighted correlation matrices \( R_{w,1} \) and \( R_{w,2} \). In addition, due to the need for discrete Fourier transform, the process of optimizing the inverse filter of SESS-BD is relatively more complicated than that of SES-BD.

3.4. Parameter settings

Before the verification of the diagnostic performance of SES-BD and SESS-BD, the settings of the main algorithm parameters are discussed below. The parameters of these two BD methods mainly include filter length, minimum relative error and maximum number of iterations.

The filter length not only affects the feature enhancement performance but also the computational efficiency of the algorithm. When the filter length is set to be small, the computational time can be shortened, but it may not match the impulse response caused by bearing faults, resulting in the inability to effectively enhance fault impulse features. However, when the filter length is set larger, although it
can effectively match the impulse response caused by bearing faults and enhance fault features, the calculation time is greatly increased, which weakens the efficiency of fault detection. Therefore, the filter length should be chosen as a compromise. Based on the results of multiple experiments, the filter length is set to 100 in this paper.

The minimum relative error and the maximum number of iterations are used to stop the iterative procedure for optimizing the inverse filter. If the minimum relative error is set too large and the maximum number of iterations is assigned too small, a relatively optimal inverse filter may be difficult to obtain; on the contrary, it will greatly increase the computational time of optimizing the inverse filter, thereby affecting the efficiency of fault detection. Therefore, based on the results of previous studies and multiple experiments, the minimum relative error and the maximum number of iterations are specified as 0.001 and 30, respectively, in this paper.

Additionally, to shorten the computational time, the squared envelope spectrum in the entire frequency band is not employed to calculate MSI in SESS-BD, but the envelope spectrum in the low-frequency band is used to calculate MSI. In this paper, the squared envelope spectrum magnitude of the low-frequency band, which contains the first four harmonics of the bearing fault characteristic frequency, is employed.

4. Case study 1

In this section, the axle bearing experimental data collected from the wheelset bearing test rig of railway passenger cars are used to verify the fault diagnosis performance of SES-BD and SESS-BD, including outer race fault and rolling element fault. In addition, the fault detection results of the state-of-the-art BD methods are presented for comparison, including MED, CYCBD, blind filtering based on envelope spectrum sparsity indicators, MGID and MESGID. In this paper, the blind filtering methods based on NE, L2/L1 and HI of the squared envelope spectrum are abbreviated as ESNE-BF, ESL21-BF and ESHI-BF, respectively, for the convenience of distinction.

4.1. Experimental setup

The wheelset bearing test rig of railway passenger cars is mainly composed of a foundation, a wheelset with two axle bearings, driving device, loading device and control system, as shown in Figure 3(a). A healthy bearing and a damaged bearing are installed on both ends of the wheelset. The axle bearings with outer race fault and rolling element fault were tested at constant rotational speed, respectively. The local damage on the surface of the bearing components is artificially implanted and has a depth of 0.2 mm and a width of 0.6 mm, as exhibited in Figure 3(b) and (c). The pith diameter, rolling element diameter, contact angle and the number of rolling elements of the tested railway bearings are 187.21 mm, 26.69 mm, 12.08° and 17, respectively. An accelerometer was mounted above the axle box to collect the vibration signal of the damaged axle bearing at a sample rate of 12800 Hz. The sampling length of each signal is 8192 data points.
4.2. Axle bearing outer race fault

The vibration signal of the axle bearing with damaged outer race was collected at the rotational speed of 884 r/min, as shown in Figure 4(a). The fault characteristic frequency of the bearing outer race is about $f_0=107.6$ Hz. Figure 4(b) displays the envelope spectrum of the bearing experimental signal. The spectral lines at the fault characteristic frequency $f_0$ of bearing outer race and its harmonics are not protruding, showing that the direct envelope spectrum analysis fails to detect the outer race fault of the tested axle bearing.

The proposed two BD methods are applied to the signal shown in Figure 4(a) for enhancing bearing fault features. To highlight the performance of the proposed methods, seven state-of-the-art deconvolution methods including MED, CYCBD, ESNE-BF, ESL21-BF, ESHI-BF, MGID and MESGID are also applied to the experimental signal. For a fair comparison, the seven comparison methods employ the same filter length and iteration-stopping criterion as the proposed methods. Figures 5 and 6 display the filtered signals and their envelope spectra achieved by different deconvolution methods. Note that the amplitudes of the filtered signals are normalized for ease of comparison. As shown in Figure 5(a) and (g) and Figure 6(c), a prominent impulse feature can be observed in the filtered signals of MED, ESL21-BF and MESGID, while the repetitive impulses are still submerged in the noise components, resulting in the fault characteristic frequency of the bearing outer race cannot be detected in the envelope spectra shown in Figure 5(b) and (h) and Figure 6(d). Thus, MED, ESL21-BF and MESGID fail to detect the outer race fault of the axle bearing. In contrast, CYCBD, ESNE-BF, ESHI-BF, MGID, SES-BD and SESS-BD successfully confirm the bearing outer race fault. Repetitive impulses can be observed in the filtered signals of these six methods, as shown in Figure 5(c), (e) and (i) and Figure 6(a), (e) and (g), and the fault characteristic frequency of the bearing outer race and its harmonics can be easily detected in the corresponding envelope spectra, as shown in Figure 5(d), (f) and (j) and Figure 6(b), (f) and (h). Although the repetitive impulses in Figure 5(c) are more pronounced than those in Figure 6(a) and (e), the harmonic frequency $4f_0$, which can be observed in Figure 6(b) and (f), cannot be detected in Figure 5(d). The fault detection effect achieved by SESS-BD is similar to that of ESNE-BF, as presented in Figure 5(f) and Figure 6(h). These results show that SES-BD and SESS-BD can effectively enhance bearing fault-related features in noisy vibration signal and accurately diagnose the outer race fault of axle bearing. Additionally, the two proposed methods exhibit stronger resistibility to random impulses than MED, ESL21-BF and MESGID.
Figure 4. (a) Vibration acceleration signal of outer race damaged axle bearing and (b) its envelope spectrum. The unit of the amplitude: m/s². The red dotted line indicates the fault characteristic frequency of the bearing outer race and its harmonics.

Figure 5. Filtered signals and their envelope spectra obtained by different deconvolution methods for processing the vibration signal of outer race damaged axle bearing: (a)-(b) MED, (c)-(d) CYCBD, (e)-(f) ESNE-BF, (g)-(h) ESL21-BF, (i)-(j) ESHI-BF. The red dotted line indicates the fault characteristic frequency of the bearing outer race and its harmonics.
Figure 6. Filtered signals and their envelope spectra obtained by different deconvolution methods for processing the vibration signal of outer race damaged axle bearing: (a)-(b) MGID, (c)-(d) MESGID, (e)-(f) SES-BD and (g)-(h) SESS-BD. The red dotted line indicates the fault characteristic frequency of the bearing outer race and its harmonics.

4.2. Axle bearing rolling element fault

The vibration signal of the axle bearing with damaged rolling element was acquired at the rotational speed of 856 r/min, as exhibited in Figure 7(a). The fault characteristic frequency of the bearing rolling element is about \( f_b = 48.5 \) Hz. The envelope spectrum of the experimental signal is displayed in Figure 7(b). The fault-related impulses are heavily polluted by interfering noise. Only the second harmonic \( 2f_b \) of the fault characteristic frequency of bearing rolling element can be identified in Figure 7(b), which cannot accurately determine the rolling element fault of axle bearing.

The proposed two methods and seven typical deconvolution methods are applied to the experimental signal shown in Figure 7(a) to enhance the impulse features induced by bearing rolling element fault. The filtered signals and their envelope spectra achieved by different deconvolution methods are displayed in Figures 8 and 9. Note that the amplitudes of the filtered signals are normalized. As shown in Figure 8(a), (c) and (g) and Figure 9(a) and (c), a prominent impulse can be clearly observed in the filtered signals while the repetitive impulse features are not obvious. In Figure 8(b), (d) and (h) and Figure 9(b) and (d), only the fault characteristic frequency of the rolling element \( f_b \) and its harmonic \( 2f_b \) can be vaguely detected, while other harmonics cannot be identified. These results indicate that MED, CYCBD, ESL21-BF, MGID and MESGID fail to effectively enhance the impulse features caused by rolling element fault and are susceptible to strong random impulses in the signal. However, ESNE-BF, ESHI-BF, SES-BD and SESS-BD effectively enhance the repetitive impulse features and the fault characteristic frequency of the bearing rolling element and its first four harmonics can be easily detected in the envelope spectra of the filtered signals, as depicted in Figure 8(e), (f), (i) and (j) and Figure 9(e), (f), (g) and (h). It can be observed that the fault diagnosis effect achieved by SES-BD and SESS-BD is similar to that of ESNE-BF and ESHI-BF. This case shows that SES-BD and SESS-BD are more robust to random transients compared with MED, ESL21-BF, MGID and MESGID and can effectively diagnose rolling element fault of railway axle bearing. In addition, compared with CYCBD, SES-BD and SESS-BD can
effectively enhance bearing fault-related features without using characteristic frequency knowledge, indicating their advantages in adaptive feature enhancement.

**Figure 7.** (a) Vibration acceleration signal of rolling element damaged axle bearing and (b) its envelope spectrum. The unit of the amplitude: m/s². The red dotted line indicates the fault characteristic frequency of the bearing rolling element and its harmonics.

**Figure 8.** Filtered signals and their envelope spectra achieved by different deconvolution methods for processing the vibration signal of rolling element damaged axle bearing: (a)-(b) MED, (c)-(d) CYCBD, (e)-(f) ESNE-BF, (g)-(h) ESI21-BF, (i)-(j) ESHI-BF. The red dotted line indicates the fault characteristic frequency of the bearing rolling element and its harmonics.
Figure 9. Filtered signals and their envelope spectra achieved by different deconvolution methods for processing the vibration signal of rolling element damaged axle bearing: (a)-(b) MGID, (c)-(d) MESGID, (e)-(f) SES-BD and (g)-(h) SESS-BD. The red dotted line indicates the fault characteristic frequency of the bearing rolling element and its harmonics.

5. Case study 2

In this section, the axle bearing experimental data collected from the wheelset bearing test rig of railway freight cars are employed to further verify the fault diagnosis performance of SES-BD and SESS-BD, including outer race fault and inner race fault. The fault detection results of the MED, CYCBD, ESNE-BF, ESL21-BF, ESHI-BF, MGID and MESGID methods are also presented for comparison.

5.1. Experimental setup

The wheelset bearing test bench for railway freight cars is shown in Figure 10(a), which is mainly composed of a frame, a support device, a driving device and a wheelset with two axle bearings. The drive wheel at the output end of the motor drives the wheelset to rotate through friction with the side of the wheel. The axle bearings with localized fatigue spalling defects on the outer and inner races were tested. The faulty bearing components are exhibited in Figure 10(b) and (c). The faulty axle bearings were collected from the repair line of railway freight cars rather than artificially implanted. During the experiment, the wheelset rotated at a constant speed of 465 r/min, and two accelerometers were mounted above the axle bearing housings to collect vibration signals. The sampling frequency of the vibration signal is 12800 Hz and the sampling length of each signal is 8192 sampling points. According to the nominal rotational speed and the geometric parameters of the tested axle bearing, the fault characteristic frequencies of the outer race and inner race are $f_o=66.75$ Hz and $f_i=88.24$ Hz, respectively. 

56.
The vibration signal of the axle bearing with spalled outer race is first analyzed, as shown in Figure 11(a). The acquired vibration signal contains strong interference noise. The envelope spectrum of the experimental signal is displayed in Figure 11(b). Only the second harmonic $2f_o$ of the fault characteristic frequency of bearing outer race can be identified, while other harmonics cannot be detected. Therefore, direct envelope spectrum analysis can hardly diagnose the outer race fault of the axle bearing.

The two proposed methods and the seven comparison methods are applied to this experimental signal, and the resulting filtered signals and their envelope spectra are shown in Figures 12 and 13. Note that the amplitudes of the filtered signals are normalized. As shown in Figure 12(a) and (g) and Figure 13(c), the filtered signals are dominated by a single impulse while the repetitive impulse features are still masked by interference noise. In the envelope spectra shown in Figure 12(b) and (h) and Figure 13(d), the fault characteristic frequency of bearing outer race and its harmonics are hardly detected. This indicates that MED, ESL21-BF and MESGID fail to effectively reconstruct outer race fault features from noisy vibration signal. As exhibited in Figure 12(i) and (j), although the fault characteristic frequency $f_o$ and its harmonic $2f_o$ can be detected in the envelope spectrum of the filtered signal of ESHI-BF, the filtered signal is dominated by a single impulse, similar to the filtered signal of MED. On the contrary, the filtered signals obtained by CYCBD, ESNE-BF, MGID, SES-BD and SESS-BD exhibit distinct repetitive impulses, as depicted in Figure 12(c) and (e) and Figure 13(a), (e) and (g), and the spectral lines of the bearing fault characteristic frequency and its harmonics can be clearly observed from the envelope spectra in Figure 12(d) and (f) and Figure 13(b), (f) and (h). Note that there is a slight error between the actual value of the fault characteristic frequency and the theoretical value. These results show that CYCBD, ESNE-BF, MGID, SES-BD and SESS-BD resist the disturbance of large single impulse compared to MED, ESL21-BF, ESHI-BF and MESGID, and effectively detect the outer race fault of the tested axle bearing. In addition, it can be observed that ESNE-BF, MGID, SES-BD and SESS-
BD achieve similar fault detection effects, and the spectral line of the harmonic frequency $4f_o$ in the envelope spectra is more conspicuous than that of CYCBD.

**Figure 11.** (a) Vibration acceleration signal of outer race spalled axle bearing and (b) its envelope spectrum. The unit of the amplitude: m/s². The red dotted lines indicate the fault characteristic frequency of the bearing outer race and its harmonics.

**Figure 12.** Filtered signals and their envelope spectra achieved by different deconvolution methods for processing the vibration signal of outer race spalled axle bearing: (a)-(b) MED, (c)-(d) CYCBD, (e)-(f) ESNE-BF, (g)-(h) ESL21-BF, (i)-(j) ESHI-BF. The red dotted lines indicate the fault characteristic frequency of the bearing outer race and its harmonics.
Figure 13. Filtered signals and their envelope spectra achieved by different deconvolution methods for processing
the vibration signal of outer race spalled axle bearing: (a)-(b) MGID, (c)-(d) MESGID, (e)-(f) SES-BD and (g)-(h)
SESS-BD. The red dotted lines indicate the fault characteristic frequency of the bearing outer race and its harmonics.

5.3. Axle bearing inner race fault

The vibration signal of the axle bearing with spalled inner race is used to further validate the proposed methods, as exhibited in Figure 14(a). Figure 14(b) depicts the envelope spectrum of the experimental signal. Several impulse features can be observed in the signal waveform. In the envelope spectrum, only the fault characteristic frequencies $f_i$ and $5f_i$ of the bearing inner race can be identified and other harmonics are difficult to be detected. Hence, the envelope spectrum analysis of the raw signal cannot confirm the inner race fault of the axle bearing.

The proposed methods and seven typical deconvolution methods are used to enhance bearing fault features from the measured vibration signals, and the filtered signals and their envelope spectra are shown in Figures 15 and 16. Note that the amplitudes of the filtered signals are normalized. It can be observed that all nine methods enhance the repetitive impulse features caused by the inner race fault of the axle bearing to varying degrees. From the perspective of the filtered signal, CYCBD, ESNE-BF, ESL21-BF, ESHI-BF, MGID, SES-BD and SESS-BD achieve similar feature enhancement results (clear repetitive impulses can be observed), and are better than MED and MESGID (MED is slightly better than MESGID), as exhibited in the left column of Figures 15 and 16. From the perspective of envelope spectrum, CYCBD, ESNE-BF, ESL21-BF, ESHI-BF, SES-BD and SESS-BD achieve similar results (fault characteristic frequency $f_i$ and its first four harmonics can be easily identified), and are better than MED, MGID and MESGID (MED and MGID are better than MESGID), as depicted in the right column of Figures 15 and 16. Note that both sides of the inner race fault characteristic frequency and its harmonics are distributed with sidebands characterized by the wheelset rotation frequency, which is a typical phenomenon of the bearing inner race fault. In this case, SES-BD and SESS-BD achieve similar feature extraction results to CYCBD without using characteristic frequency knowledge, and outperform MED, MGID and MESGID, further verifying their effectiveness and advantages in axle bearing fault diagnosis.
Figure 14. (a) Vibration acceleration signal of inner race spalled axle bearing and (b) its envelope spectrum. The unit of the amplitude: m/s². The red dotted lines indicate the fault characteristic frequency of the bearing inner race and its harmonics.

Figure 15. Filtered signals and their envelope spectra achieved by different deconvolution methods for processing the vibration signal of inner race spalled axle bearing: (a)-(b) MED, (c)-(d) CYCBD, (e)-(f) ESNE-BF, (g)-(h) ESL21-BF, (i)-(j) ESHI-BF. The red dotted lines indicate the fault characteristic frequency of the bearing inner race and its harmonics.
Figure 16. Filtered signals and their envelope spectra achieved by different deconvolution methods for processing the vibration signal of inner race spalled axle bearing: (a)-(b) MGID, (c)-(d) MESGID, (e)-(f) SES-BD and (g)-(h) SESS-BD. The red dotted lines indicate the fault characteristic frequency of the bearing inner race and its harmonics.

6. Performance analysis

The diagnostic performance of the proposed methods has been verified using different axle bearing experimental signals and compared with seven existing methods. The fault diagnosis results of different methods are summarized in Table 2. Note that the “Yes”, “Partial” and “No” indicate the successful diagnosis, partially successful diagnosis and unsuccessful diagnosis, respectively. It can be observed that similar to ESNE-BF, SES-BD and SESS-BD achieve successful diagnosis in all four experimental cases of railway axle bearings, outperforming the other six existing deconvolution methods. The effectiveness and advantages of the proposed methods in detecting different faults of railway axle bearings are demonstrated.

In addition, to quantitatively evaluate the feature enhancement performance, a metric based on the envelope spectrum, the frequency-domain signal-to-noise ratio (FDSNR)\(^5\), is employed, as follows:

\[
FDSNR = \frac{1}{H} \sum_{h=1}^{H} \max_{m \in A_h} \left( ES[m] \right)^2 \quad \frac{1}{M - H} \sum_{m=1}^{M} \left( ES[m] \right)^2 - \sum_{h=1}^{H} \max_{m \in A_h} \left( ES[m] \right)^2 
\]

where \( ES[m] \) denotes the amplitude of the envelope spectrum at the \( m \)th spectral frequency, \( H \) is the number of harmonics of the fault characteristic frequency, \( A_h \) represents a series of spectral frequencies in a small band around the \( h \)th harmonic of the fault characteristic frequency, \( M \) is the total number of discrete spectral frequencies. In this paper, the fault characteristic frequency and its first four harmonics are considered, i.e., \( H=5 \), and the tolerance band \( A_h \) contains 3 spectral lines on both sides of the harmonic of nominal fault characteristic frequency. The first 5 harmonics of fault characteristic frequency are included mainly to avoid misjudgment of fault detection and effectively evaluate the fault feature enhancement performance of different methods.

Figure 17 shows the FDSNR values achieved by the nine deconvolution methods for processing
different axle bearing experimental signals. Although the FDSNR values achieved by SES-BD or SESS-BD are sometimes smaller than those of CYCBD, ESNE-BF and MGID in a single case of axle bearing fault, i.e., the outer race fault of case study 1 (CS1) or case study 2 (CS2), they both exhibit excellent diagnostic performance in four cases of axle bearing faults. Therefore, SES-BD and SESS-BD have strong comprehensive diagnostic performance for railway axle bearings compared to the MED, CYCBD, ESL21-BF, ESHI-BF, MGID and MESGID methods.

Finally, the computational efficiency of the nine methods is analyzed. Figure 18 shows the computational time (unit: s) consumed by the nine deconvolution methods for processing different bearing experimental signals. The computational time is estimated by MATLAB R2016b on a Dell laptop with a processor Intel(R) Core(TM) i7-10850H CPU @ 2.70GHz 2.71GHz. The computational time shown is the average after 10 trials for each experimental signal. Overall, the computational efficiency of SES-BD is between MED and MGID, and the computational efficiency of SESS-BD is similar to that of ESNE-BF, ESL21-BF, ESHI-BF and MESGID. Additionally, the computational efficiency of SES-BD is higher than that of SESS-BD because the Fourier transform is not required in the process of optimizing the inverse filter. Therefore, SES-BD is preferred over SESS-BD for bearing diagnostics from the viewpoint of computational efficiency.

Table 2. A summary of diagnostic results of different BD methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Case study 1</th>
<th>Case study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Outer race fault</td>
<td>Rolling element fault</td>
</tr>
<tr>
<td>MED</td>
<td>No</td>
<td>Partial</td>
</tr>
<tr>
<td>CYCBD</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ESNE-BF</td>
<td>Yes</td>
<td>Partial</td>
</tr>
<tr>
<td>ESL21-BF</td>
<td>No</td>
<td>Partial</td>
</tr>
<tr>
<td>ESHI-BF</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MGID</td>
<td>Yes</td>
<td>Partial</td>
</tr>
<tr>
<td>MESGID</td>
<td>No</td>
<td>Partial</td>
</tr>
<tr>
<td>SES-BD</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SESS-BD</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Figure 17. FDSNR values achieved by nine deconvolution methods for processing different railway axle bearing experimental signals.
Figure 18. Computational time (s) consumed by nine deconvolution methods for processing different railway axle bearing experimental signals.

7. Conclusions

In this paper, the theories of the MSI-based BD methods and their application to railway axle bearing fault diagnosis are investigated. The MSI of the squared envelope and the MSI of the squared envelope spectrum are proposed as objective functions for deconvolution, and two new BD methods, SES-BD and SESS-BD, are developed for enhancing bearing fault features. The two proposed methods are validated on different bearing experimental signals from two railway wheelset bearing test rigs and compared with seven existing BD methods. The following conclusions can be drawn from the results:

(1) The deconvolution method using the MSI of the squared envelope or the MSI of the squared envelope spectrum as the objective function and using EVA as the inverse filter solver is feasible and reasonable. SES-BD and SESS-BD can effectively enhance bearing fault-related features in noisy vibration signals and accurately detect different faults of railway axle bearings.

(2) SES-BD and SESS-BD deliver excellent axle bearing diagnostic performance among the nine diagnostic methods explored, similar to ESNE-BF. Compared with MED, ESL21-BF, ESHI-BF, MGID and MESGID, they show a strong capability to resist random impulse interference; compared with CYCBD, they have a strong capability to adaptively enhance bearing fault features.

(3) SES-BD, SESS-BD and ESNE-BF exhibit competitive performance in repetitive transient enhancement and bearing fault diagnosis. From the viewpoint of computational efficiency, SES-BD is preferentially used due to the relatively simplified process of optimizing the inverse filter.

This paper focuses on the application of SES-BD and SESS-BD in fault diagnosis of railway axle bearings accompanied by comparisons with several typical BD methods. The proposed methods can be applied to fault diagnosis and condition monitoring of other rotating machinery components and can be reasonably extended to similar fields. The optimization and development of the solution method of the inverse filter are worth investigating to improve the computational efficiency of the proposed methods. The optimal selection method of the inverse filter length is also worth investigating in further work to improve the performance of the proposed methods.

Declaration of conflicting interests

The Author(s) declare(s) that there is no conflict of interest.
Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Key Research and Development Program of China (Grant No. 2021YFB3400704-02), the open project of State Key Laboratory of Traction Power, Southwest Jiaotong University, China (Grant No. TPL2210), the National Natural Science Foundation of China (Grant No. 52275133), the Science and Technology Project of Hunan Province, China (Grant No. 2021GK4014) and the China Scholarship Council (Grant No. 202107000033).

ORCID iD

Bingyan Chen https://orcid.org/0000-0001-7103-0221
Weihua Zhang https://orcid.org/0000-0002-3391-3872
Fengshou Gu https://orcid.org/0000-0003-4907-525X

References


