A New Family of Model-Based Impulsive Wavelets and Their Sparse Representation for Rolling Bearing Fault Diagnosis

Yi Qin, *Member, IEEE*

***Abstract*—The localized faults of rolling bearings can be diagnosed by the extraction of the impulsive feature. However, the approximately periodic impulses may be submerged in strong interferences generated by other components and the background noise. To address this issue, this paper explores a new impulsive feature extraction method based on the sparse representation. According to the vibration model of an impulse generated by the bearing fault, a novel impulsive wavelet is constructed, which satisfies the admissibility condition. As a result, this family of model-based impulsive wavelets can form a Parseval frame. With the model-based impulsive wavelet basis and Fourier basis, a convex optimization problem is formulated to extract the repetitive impulses. Based on the splitting idea, an iterative thresholding shrinkage algorithm is proposed to solve this problem, and it has a fast convergence rate. Via the simulated signal and real vibration signals with bearing fault information, the performance of the proposed approach for repetitive impulsive feature extraction is validated and compared with the noted spectral kurtosis method, the optimized spectral kurtosis method based on simulated annealing, and the resonance-based signal decomposition method. The results demonstrate its advantage and superiority in weak repetitive transient feature extraction.**

***Index Terms*—Convex optimization, dictionary, fault detection, inverse wavelet transform, iterative shrinking algorithm, weak impulsive feature extraction.**

I. Introduction

ROLLING bearings are key components in motors and rotary machinery, such as automotive transmission system, wind turbine, aeroengine, machine tool, etc. These devices usually operate in harsh working conditions caused by high speed, heavy load, high-low temperature, and contamination, it is highly desirable to perform online condition monitoring to improve their availability, safety, and reliability, reduce the operation and maintenance costs, and achieve downtime minimization and productivity maximization [1], [2]. As many failures in these devices may start from rolling bearing faults, it is necessary to diagnose the faults as early as possible to prevent severe or catastrophic accident.

To address this issue, various methods, such as vibration analysis [3]–[6], current signature analysis [7], [8], acoustic emission feature identification [9], [10], sound pressure monitoring [11], and stray flux analysis [12], etc., have been widely and effectively used for diagnosing diverse rolling bearing faults. Since vibration signals directly represent the dynamic behavior of faulty bearings, which leads to their sensitivity to faults, they have been most widely applied to detect the bearing faults.

Whenever rolling elements pass over a defect (e.g., a crack, a spall, or a pitting) in a bearing, a pseudocyclostationary impulsive signal is produced. By extracting the repetitive impulsive feature, the fault type can be detected. The commonly used feature extraction methods can be mainly classified into three categories: the time-domain analysis [13], [14], frequency-domain analysis [1], [9], and time–frequency-domain analysis [7], [[15], [16]. Compared with time-domain analysis and frequency-domain analysis, time–frequency-domain analysis methods have been paid more attention due to their ability for both stationary and nonstationary signals. The spectral kurtosis (SK) may be the most widely used method for bearing fault detection [13], [17], [18]. It provides an efficient tool to select the best band that contains fault information by combining kurtosis with time–frequency-domain analysis, which can detect non-Gaussian impulses generated by the faulty rolling bearing. Unfortunately, its performance is often weakened by the strong harmonic interferences [19] and noise. To extract impulsive (transient) feature, a proper basis function should be adopted. Morlet wavelets have been widely used for rolling bearing fault diagnosis [20], [21], because they are similar to impulses. Nevertheless, the Morlet wavelet does not satisfy the admissibility condition, which results in the absence of the inverse wavelet transform. As we know, signal recovery is quite important in the fields of compressive sensing and signal component separation, i.e., the inverse operation of the sparse coefficients is necessary. Other commonly used wavelets, such as even-order Gaussian wavelet and Mexican hat wavelet, also do not satisfy the admissibility condition. Moreover, the Meyer wavelet is defined in this frequency domain which results in that it is not compact supported, and its shape is different from the impulse

Manuscript received March 28, 2017; revised May 25, 2017; accepted July 23, 2017. Date of publication August 7, 2017; date of current version December 15, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 51675065 and in part by the Chongqing Research Program of Basic Research and Frontier Technology under Grant cstc2017jcyjAX0459.

The author is with the State Key Laboratory of Mechanical Transmission, College of Mechanical Engineering, Chongqing University, Chongqing 400044, China (e-mail: qy_808@cqu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2017.2736510

caused by the bearing fault. It is well known that the obtained wavelet coefficients can more accurately represent the transient feature when the shape of wavelet is more similar to that of impulse. Therefore, according to the vibration model of an impulse generated by the bearing fault, a novel impulsive wavelet function is constructed in this study. These new wavelets can accurately extract the impulsive feature and satisfy the admissibility condition, which is crucial for signal reconstruction and signal component separation.

When incipient faults of rolling bearings occur, the pseudocyclostationary impulsive signal is usually submerged in noise and other components' vibrations which are mainly composed of harmonics, amplitude modulation, and frequency modulation (AM-FM) signals. The harmonics may come from the vibrations of shaft and other rotating components. The impulse force caused by the localized fault will arouse high-frequency intrinsic vibration within the bearing system, and the amplitude of the vibration will be modulated by the impulse force. Moreover, when the vibration transfers along a path, the modulation phenomenon may arise. The noise comes from the mechanical system and electrical devices. In such case, since the spectra of repetitive transients and noise distribute in the whole frequency range which leads to that the spectra of harmonics, AM-FM components, transients, and noise overlap in the frequency domain, the impulsive signal obtained by the commonly used methods based on filtering may contain harmonics, AM-FM components, and noise. It then immediately implies that the performance of these filtering-based methods for transients detection is decreased. To improve the accuracy and effectiveness of weak transient feature extraction, it is necessary to explore a new method for extracting weak transient feature from incipiently faulty bearing signals. Sparse representation is a useful approach to extract faulty information from the complex vibration signal via representing different characteristic components by different basis (atoms) of a dictionary rather than by different frequency bands, and it can effectively solve the problem of frequency aliasing. The overcomplete dictionary may be composed with different complete basis to enlarge the flexibility in terms of representing signals. To solve the sparse optimization problem, there are mainly two well-known methods, which are, respectively, matching pursuit and basis pursuit. The matching pursuit is suitable for orthogonal dictionary and has fast computational speed, while the basis pursuit is featured by superresolution and better sparsity [22], [23]. In addition, there are other search algorithms or improved algorithms, such as orthogonal matching pursuit [24], stagewise orthogonal matching pursuit [25], iterative thresholding [26], compressive sampling matching pursuit, and subspace pursuit [27]. Due to the remarkable performance of sparse representation, more and more researches focus on sparse representation of transient signals in recent years. For example, the sparsity-based overlapping group sparsity method was utilized to extract repetitive transients caused by bearing faults [28]. An algorithm called augmented Lagrangian majorization-minimization algorithm was developed for detecting impulsive feature of faulty bearings and gear boxes [29]. A resonance-based signal decomposition method was proposed for extracting nonoscillatory transients [30]. To summarize, most of the current methods, such as SK [17], wavelet analysis [20], squared envelope spectrum [14], and empirical mode decomposition [15], have been successfully used for obvious impulsive feature extraction, however these methods are difficult to accurately obtain the pseudocyclostationary impulses according to the above explanation; while the recent sparse representation methods for detecting the impulses in [28] and [29] just consider the interference of the noise but do not consider the influence of strong harmonics (or AM-FM signals). In order to further improve the accuracy of impulse detection, with the new model-based impulsive wavelets, a sparse representation method is proposed for detecting the weak transient feature of rolling bearings. For repetitive impulses generated by the bearing fault, the new model-based impulsive wavelet basis can be used for their sparse representation, since it satisfies the admissibility condition. For other components including harmonics and AM-FM signals, they can be well sparsely represented by the Fourier basis. For the noise, sparse denoising can be used to reduce it. Based on the above analysis, a sparse optimization problem is formulated to extract the repetitive impulsive components. To guarantee that the optimization problem is strictly convex, the specific conditions of penalty parameters are derived. Then, an iterative thresholding shrinkage algorithm based on the splitting idea is proposed to solve this problem, and the convergence analysis is performed to validate the practicability of the proposed algorithm. Finally, the proposed approach is successfully applied to analyze the simulated signal and real faulty vibration signals acquired from an automobile gearbox with bearing fault and a motor with bearing fault. The results show that this sparse representation method based on the new impulsive wavelets can not only detect the weak transients from the strong inference vibrations coming from other mechanical components and the background noise, but also improve the accuracy of the impulsive feature extraction.

The remainder of this paper is organized as follows. In Section II, the model-based impulsive wavelet is constructed, and its property is investigated. In Section III, a convex optimization problem is proposed to extract the repetitive impulses, and the corresponding solution algorithm is described in detail. In Section IV, the performance of the proposed method is validated by simulation analysis. In Section V, the proposed method is applied to detect the real faults of rolling bearings. Conclusion of this study is presented in Section VI.

II. New Model-Based Impulsive Wavelets

A. Preliminaries

Wavelet transforms are inner products of the signal and a family of the wavelets. By dilation and translation from the mother wavelet $\psi(t)$, the wavelet transform of an arbitrary signal $s(t) \in L^2(\mathbf{R})$ is defined as [31]:

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} s(t)\psi^* \left(\frac{t-b}{a} \right) dt \tag{1}$$

where $*$ denotes the complex conjugation. a is the scale parameter and b is the shifting parameter. $W(a,b)$ gives the information

of $s(t)$ at different levels of resolution and also measures the similarity between the signal $s(t)$ and the wavelet at each scale.

As we know, the wavelet coefficients can better represent the feature of the signal when the waveform of wavelet is more similar to that of the analyzed signal. In order to construct a feasible and inspiring wavelet for impulsive feature extraction, we first introduce the vibration response generated by the faulty rolling bearing.

When a ball hits a failure in the inner/outer raceway (or a failure in the sphere reaches a track), a short-term impulse is produced, which can excite the resonance frequencies (high-frequency resonances) of the entire structure between the bearing and the vibration transducer [32]. In such case, the rolling bearing vibration system is a typical underdamped second-order mass-spring-damper system and its unit-impulse vibration response function can be defined as

$$z(t) = A\exp(-\sigma\omega_n t)\sin(\omega_n\sqrt{1-\sigma^2}t) \tag{2}$$

where $\omega_n = \sqrt{\mathrm{k/m}}$ is the natural oscillation frequency of the system, and $\sigma = 0.5\mathrm{c}/\sqrt{\mathrm{km}}$ is the relative damping. The symbols m, k, and c, respectively, represent the mass (kg), stiffness (N/m), and damping coefficient (N·s/m) of the system. Let

$$\omega_d = \omega_n\sqrt{1-\sigma^2}, \quad \xi = \sigma/\sqrt{1-\sigma^2}. \tag{3}$$

Equation (2) can be rewritten as

$$z(t) = A\exp(-\xi\omega_d t)\sin(\omega_d t). \tag{4}$$

When the shaft rotates, due to the random spacing of the impulses, the vibration signal can be modeled as [14]:

$$h(t) = \sum_k A_k \exp\left[-\xi\omega_d(t-\tau_k)\right]\sin\left[\omega_d(t-\tau_k)\right] \tag{5}$$

where τ_k are random occurrences of the impulses, and the difference $\tau_{k+1} - \tau_k$ is a nonnegative random variable whose mean specifies the average rate of repetition of the impulses.

B. Construction of Model-Based Impulsive Wavelet

According to the vibration response caused by the bearing fault, a novel model-based impulsive wavelet (impulsive wavelet for short in the following) is constructed and defined as

$$\psi(t) = \begin{cases} \exp(-\xi\omega_d t)\sin(\omega_d t), & 0 \le t < T/2 \\ \exp(\xi\omega_d t)\sin(\omega_d t), & -T/2 < t < 0 \\ 0, & \text{other} \end{cases} \tag{6}$$

where T denotes the wavelet support in time domain, ω_d represents the wavelet's central frequency, and ξ represents the wavelet's decay rate. The two parameters ω_d and ξ are closely related to the shape and time–frequency resolution of the impulsive wavelet. The frequency parameter ω_d controls the oscillatory frequency of the impulsive wavelet, while the decay parameter ξ controls the oscillation attenuation of the impulsive wavelet. Obviously, decreasing ξ will increase the support of the impulsive wavelet, but improve its frequency resolution. In application, ω_d can be set by the estimation of natural oscillation frequency of the system, and ξ can be empirically taken in the range of [0.02, 0.3]. Similar to Morlet wavelet optimization [33], the two parameters also can be set via some optimization methods.

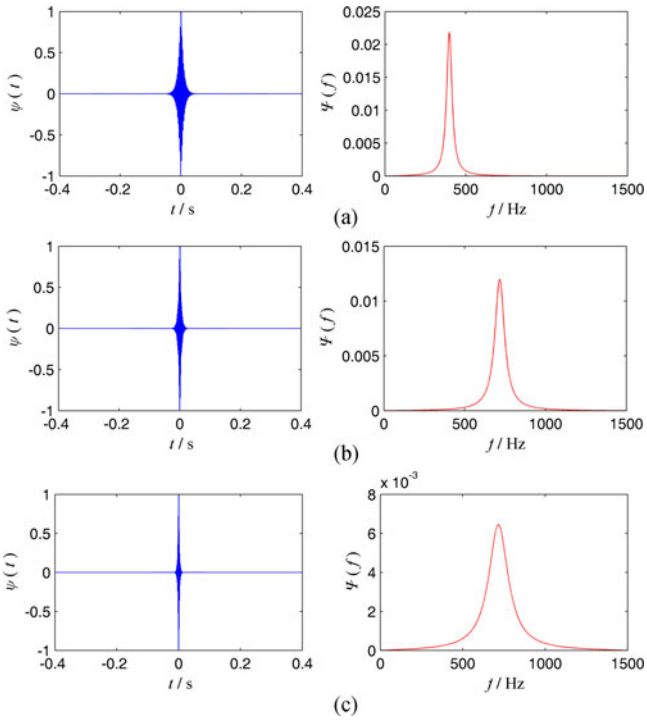

Fig. 1. Waveforms and frequency spectra of impulsive wavelets with different parameters. (a) $\omega_d = 2500$ and $\xi = 0.05$. (b) $\omega_d = 4500$ and $\xi = 0.05$. (c) $\omega_d = 4500$ and $\xi = 0.1$. The figures on the left side illustrate the time-domain waveforms of impulsive wavelets, while the figures on the right side illustrate their corresponding frequency spectra.

To check the influence of wavelet parameters on the impulsive wavelet, the waveforms and frequency spectra of impulsive wavelets with different ω_d and ξ are, respectively, illustrated in Fig. 1. In Fig. 1(a), $\omega_d = 2500$ and $\xi = 0.05$; in Fig. 1(b), $\omega_d = 4500$ and $\xi = 0.05$; in Fig. 1(c), $\omega_d = 4500$ and $\xi = 0.1$. It can be seen from Fig. 1(a) and (b) that the oscillatory frequency (i.e., central frequency) of impulsive wavelet is higher when ω_d is larger. Moreover, we can see from Fig. 1(b) and (c) that the impulsive wavelet attenuates more quickly when ξ is larger, but the central frequency keeps the same. Therefore, according to the value of ξ, T can be set between 0.4 and 0.8 s.

It is often expected that the signal reconstruction can be achieved by the use of the obtained wavelet coefficients. Then, we will prove the new impulsive wavelet satisfies the admissibility condition, whose definition can be seen in [31].

According to (6), we have

$$\int_{-\infty}^{\infty}\psi(t)dt = \int_0^{T/2}\exp(-\xi\omega_d t)\,\sin(\omega_d t)dt + \int_{-T/2}^{0}\exp(\xi\omega_d t)\,\sin(\omega_d t)dt. \tag{7}$$

For the second item, let $t = -t'$, then yield

$$\int_{-T/2}^{0} \exp(\xi\omega_d t)\,\sin(\omega_d t)dt = \int_{T/2}^{0} \exp(-\xi\omega_d t')\sin(\omega_d t')dt'$$
$$= -\int_{0}^{T/2} \exp(-\xi\omega_d t')\sin(\omega_d t')dt'. \quad (8)$$

Replacing t' to t and then substituting (8) into (7), it immediately follows that

$$\int_{-\infty}^{\infty} \psi(t)dt = 0. \quad (9)$$

Therefore, with the wavelet coefficients obtained by the impulsive wavelets, the original signal can be reconstructed by

$$s(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\frac{da}{a^2}\int_{-\infty}^{\infty} W(a,b)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)db \quad (10)$$

where

$$C_\psi = \int_{-\infty}^{\infty}\frac{|\Psi(\omega)|^2}{|\omega|}d\omega. \quad (11)$$

Since the inverse impulsive wavelet transform exists, the impulsive wavelet basis can effectively be applied to sparse representation. For the convenience of derivation, the numeric calculation of impulsive wavelet transform is assumed to be implemented by an operator (transform matrix) $\mathbf{A}_T$, while the numeric calculation of inverse impulsive wavelet transform is assumed to be implemented by an operator $\mathbf{A}$. Then, for a discrete signal $\mathbf{s}$, (1) and (10) can be, respectively, rewritten as

$$\mathbf{x} = \mathbf{A}_T\,\mathbf{s} \quad (12)$$

$$\mathbf{s} = \mathbf{A}\mathbf{x} \quad (13)$$

where $\mathbf{x}$ denotes the obtained wavelet coefficients. It is easy to note that the impulsive wavelets form a Parseval frame (or tight frame), i.e., the original signal can be recovered by the obtained coefficients in such frame. It follows that $\mathbf{A}$ and $\mathbf{A}_T$ satisfy

$$\mathbf{A}\mathbf{A}_T = \mathbf{I} \quad (14)$$

where $\mathbf{I}$ is an unit matrix. In the following section, $\mathbf{A}_T$ is denoted by $\mathbf{A}^*$, which is regarded as the right inverse of $\mathbf{A}$.

III. Sparse Representation for Impulsive Feature

A. Problem Formulation

We have given the vibration model of repetitive impulses caused by the bearing fault. However, in the measured signal, except for the periodically impulsive signal, there are background noise and other vibration signals generated by some rotatory components (e.g., main shaft and gears), which are mainly harmonic signals or AM-FM signals. Therefore, the observed signal $\mathbf{s}$ of a faulty rolling bearing can be modeled as

$$\mathbf{s} = \mathbf{h} + \mathbf{w} + \mathbf{n} \quad (15)$$

where $\mathbf{h}$ is a pseudocyclostationary (approximately periodic) impulsive signal given by (5), $\mathbf{w}$ is composed of harmonics and AM-FM signal components, and $\mathbf{n}$ is a noise.

With the principle of morphological component analysis [34], [35], the approximately periodic impulsive signal $\mathbf{h}$ could be sparsely represented by the proposed impulsive wavelet basis (i.e., impulsive wavelet dictionary), and $\mathbf{w}$ can be represented by the sparse Fourier coefficients through the Fourier basis (i.e., Fourier dictionary), because both harmonics and AM-FM signals have harmonic morphology, For an analyzed signal of length N, the atoms of Fourier dictionary $\mathbf{B} \in \mathbb{C}^{N\times L}$ are defined as

$$B_{m,n} = \exp\left(j\frac{2\pi}{L}mn\right),\ 0 \le m \le N-1,\ 0 \le n \le L-1 \quad (16)$$

where $N \le L$. The dictionary is a conventional Fourier dictionary when $N = L$, while it is an overcomplete dictionary Fourier dictionary when $N < L$. In this study, we mainly use the conventional Fourier dictionary. The noise probably arises due to sensor imperfection, poor running environment, conditioning circuit, and communication errors [19]. This noise is usually a Gaussian white noise. It follows that the noise can be effectively removed by solving the sparse denoising problem. Then, the following sparse optimization problem is formulated to extract the impulsive feature:

$$\arg\min_{\mathbf{x},\mathbf{y}}\ \lambda_1\|\mathbf{x}\|_1 + \lambda_2\|\mathbf{y}\|_1 + \frac{\mu}{2}\|\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{y}-\mathbf{s}\|_2^2 \quad (17)$$

where λ_1 and λ_2 are regular parameters; $\mathbf{A}$ and $\mathbf{B}$ represent, respectively, impulsive wavelet transform matrix and Fourier transform matrix; $\mathbf{x}$ and $\mathbf{y}$ are the sparse coefficients obtained by $\mathbf{A}$ and $\mathbf{B}$, respectively; and μ is the Lagrangian parameter. We can know that the last item in (17) is used to remove the noise.

Once the sparse coefficients $\mathbf{x}$ are obtained by solving (17), the extracted repetive impulses can be calculated by

$$\mathbf{h} = \mathbf{A}\mathbf{x}. \quad (18)$$

B. Algorithm Derivation

Similar to alternating direction method of multipliers [36], fast iterative shrinkage-thresholding algorithm [37], and split augmented Lagrangian shrinkage algorithm [38], by the use of splitting method and separable surrogate function [26], an iterative thresholding shrinkage algorithm to solve (17) is proposed to speed up convergence. First, according to (17), the objective function can be written as

$$f(\mathbf{x},\mathbf{y}) = \lambda_1\|\mathbf{x}\|_1 + \lambda_2\|\mathbf{y}\|_1 + \frac{\mu}{2}\|\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{y}-\mathbf{s}\|_2^2. \quad (19)$$

It is easy to note that this objective function is strictly convex when $\mu > 0$.

With the splitting principle, this objective function is splitted into two objective functions and then the minimization of the objective function (19) can be achieved by optimizing $\mathbf{x}$ and $\mathbf{y}$ successively. Thus, the sparse optimization problem can be solved by minimizing the following two objective functions:

$$f_1(\mathbf{x}) = \lambda_1\|\mathbf{x}\|_1 + \frac{\mu_1}{2}\|\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{y}_k-\mathbf{s}\|_2^2 \quad (20)$$

$$f_2(\mathbf{y}) = \lambda_2\|\mathbf{y}\|_1 + \frac{\mu_2}{2}\|\mathbf{A}\mathbf{x}_{k+1}+\mathbf{B}\mathbf{y}-\mathbf{s}\|_2^2 \quad (21)$$

where $\mathbf{y}_k$ are the sparse Fourier coefficients obtained through the kth iteration, and $\mathbf{x}_{k+1}$ are the sparse impulsive coefficients obtained by minimizing (20).

In order to minimize (23) and (24), two additional functions are introduced, which are, respectively, defined as

$$d_1(\mathbf{x},\mathbf{x}_k) = \frac{c_1}{2}\|\mathbf{x}-\mathbf{x}_k\|_2^2 - \frac{\mu_1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{A}\mathbf{x}_k\|_2^2 \tag{22}$$

$$d_2(\mathbf{y},\mathbf{y}_k) = \frac{c_2}{2}\|\mathbf{y}-\mathbf{y}_k\|_2^2 - \frac{\mu_2}{2}\|\mathbf{B}\mathbf{y}-\mathbf{B}\mathbf{y}_k\|_2^2 \tag{23}$$

where c_1, c_2 denote penalty parameters, and μ_1, μ_2 denote the Lagrangian parameters, which will be used to calculate the threshold. Then, the new objective functions are formulated as

$$\begin{aligned}\tilde{f}_1(\mathbf{x}) &= f_1(\mathbf{x}) + d_1(\mathbf{x},\mathbf{x}_k) = \lambda_1\|\mathbf{x}\|_1 \\ &\quad + \frac{\mu_1}{2}\|\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{y}_k-\mathbf{s}\|_2^2 + \frac{c_1}{2}\|\mathbf{x}-\mathbf{x}_k\|_2^2 \\ &\quad - \frac{\mu_1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{A}\mathbf{x}_k\|_2^2 \end{aligned}\tag{24}$$

$$\begin{aligned}\tilde{f}_2(\mathbf{y}) &= f_2(\mathbf{y}) + d_2(\mathbf{y},\mathbf{y}_k) = \lambda_2\|\mathbf{y}\|_1 \\ &\quad + \frac{\mu_2}{2}\|\mathbf{A}\mathbf{x}_{k+1}+\mathbf{B}\mathbf{y}-\mathbf{s}\|_2^2 + \frac{c_2}{2}\|\mathbf{y}-\mathbf{y}_k\|_2^2 \\ &\quad - \frac{\mu_2}{2}\|\mathbf{B}\mathbf{y}-\mathbf{B}\mathbf{y}_k\|_2^2 . \end{aligned}\tag{25}$$

By adding the functions $d_1(\cdot)$ and $d_2(\cdot)$, the tasks of minimization will become simple and effective, since the items $\|\mathbf{A}\mathbf{x}\|_2^2$ and $\|\mathbf{B}\mathbf{y}\|_2^2$, respectively, exist in (24) and (25). However, to assure the convexity of the new two objective functions (24) and (25), $d_1(\cdot)$ and $d_2(\cdot)$ must be strictly convex, and the specific conditions are derived as the following proposition.

Proposition 1: Suppose that both $\mathbf{A}$ and $\mathbf{B}$ are from Parseval frames, which, respectively, satisfy

$$\mathbf{A}^*\mathbf{A} = p_1\mathbf{I} \tag{26}$$

$$\mathbf{B}^*\mathbf{B} = p_2\mathbf{I} \tag{27}$$

where p_1 and p_2 are constants. If the parameters c_1 and c_2, respectively, in (22) and (23) satisfy

$$c_1 > p_1\mu_1 \tag{28}$$

$$c_2 > p_2\mu_2 \tag{29}$$

then $d_1(\cdot)$ and $d_2(\cdot)$ are, respectively, strictly convex.

A proof of the above proposition is given in the Appendix.

In this study, it is easy to note that $p_1 = p_2 = 1$. To satisfy (28) and (29), c_1 and c_2 can be, respectively, taken as

$$c_1 = \frac{\mu_1(\mu_1+\mu_2)}{\mu_2} \tag{30}$$

$$c_2 = \frac{\mu_2(\mu_1+\mu_2)}{\mu_1}. \tag{31}$$

Then, we can immediately know that the two new objective functions $\tilde{f}_1(\mathbf{x})$ and $\tilde{f}_2(\mathbf{y})$ are still convex. We first consider the optimization of $\tilde{f}_1(\mathbf{x})$. By expanding (24), it can be newly formulated as

$$\begin{aligned}\tilde{f}_1(\mathbf{x}) &= \frac{\mu_1}{2}\|\mathbf{s}\|_2^2 + \frac{\mu_1}{2}\|\mathbf{B}\mathbf{y}_k\|_2^2 - \frac{\mu_1}{2}\|\mathbf{A}\mathbf{x}_k\|_2^2 - \mu_1\mathbf{y}_k^*\mathbf{B}^*\mathbf{s} \\ &\quad - \mu_1\mathbf{x}^*\mathbf{A}^*\mathbf{s} + \mu_1\mathbf{x}^*\mathbf{A}^*\mathbf{B}\mathbf{y}_k + \mu_1\mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x}_k \\ &\quad - c_1\mathbf{x}^*\mathbf{x}_k + \frac{c_1}{2}\|\mathbf{x}\|_2^2 + \lambda_1\|\mathbf{x}\|_1 \\ &= \mathrm{cons\,t}_1 - \mu_1\mathbf{x}^*\left[\mathbf{A}^*(\mathbf{s}-\mathbf{A}\mathbf{x}_k-\mathbf{B}\mathbf{y}_k) + \frac{c_1}{\mu_1}\mathbf{x}_k\right] \\ &\quad + \frac{c_1}{2}\|\mathbf{x}\|_2^2 + \lambda_1\|\mathbf{x}\|_1 \end{aligned}\tag{32}$$

where

$$\mathrm{cons\,t}_1 = \frac{\mu_1}{2}\|\mathbf{s}\|_2^2 + \frac{\mu_1}{2}\|\mathbf{B}\mathbf{y}_k\|_2^2 - \frac{\mu_1}{2}\|\mathbf{A}\mathbf{x}_k\|_2^2 - \mu_1\mathbf{y}_k^*\mathbf{B}^*\mathbf{s}. \tag{33}$$

Let

$$\mathbf{v}_k = \frac{\mu_1}{c_1}\mathbf{A}^*(\mathbf{s}-\mathbf{A}\mathbf{x}_k-\mathbf{B}\mathbf{y}_k) + \mathbf{x}_k. \tag{34}$$

As the optimized result is not affected when the objective function $\tilde{f}_1(\mathbf{x})$ is multiplied by μ_1/c_1, (32) can be rewritten as

$$\begin{aligned}\tilde{f}_1'(\mathbf{x}) &= \mathrm{cons\,t}_2 - \mu_1\mathbf{x}^*\mathbf{v}_k + \frac{\mu_1}{2}\|\mathbf{x}\|_2^2 + \frac{\lambda_1\mu_1}{c_1}\|\mathbf{x}\|_1 \\ &= \mathrm{cons\,t}_3 + \frac{\lambda_1\mu_1}{c_1}\|\mathbf{x}\|_1 + \frac{\mu_1}{2}\|\mathbf{x}-\mathbf{v}_k\|_2^2 \end{aligned}\tag{35}$$

where

$$\mathrm{cons\,t}_2 = \frac{\mu_1}{c_1}\mathrm{cons\,t}_1 \tag{36}$$

$$\mathrm{cons\,t}_3 = \mathrm{cons\,t}_2 + \frac{\mu_1}{2}\|\mathbf{v}_k\|_2^2 . \tag{37}$$

For the following optimization problem:

$$\arg\min_{\mathbf{x}} \tilde{f}_1'(\mathbf{x}) . \tag{38}$$

By using the shrinkage approach, its solution can be obtained as

$$\mathbf{x}_{\mathrm{opt}} = S_{\frac{\lambda_1\mu_1}{c_1}}(\mathbf{v}_k) = S_{\frac{\lambda_1\mu_1}{c_1}}\left(\frac{\mu_1}{c_1}\mathbf{A}^*(\mathbf{s}-\mathbf{A}\mathbf{x}_k-\mathbf{B}\mathbf{y}_k)+\mathbf{x}_k\right) \tag{39}$$

where S denotes the operator of shrinkage, and the subscript denotes the new Lagrangian parameter. The function $\tilde{f}_1'(\mathbf{x})$ can be minimized via iteration, and the obtained solution sequence $\{\mathbf{x}_k\}_k$ can converge to the minimal point of $\tilde{f}_1'(\mathbf{x})$. According to (39), the iterative formula for calculating $\{\mathbf{x}_k\}_k$ is given by

$$\mathbf{x}_{k+1} = S_{\frac{\lambda_1\mu_1}{c_1}}\left(\frac{\mu_1}{c_1}\mathbf{A}^*(\mathbf{s}-\mathbf{A}\mathbf{x}_k-\mathbf{B}\mathbf{y}_k)+\mathbf{x}_k\right). \tag{40}$$

After obtaining $\mathbf{x}_{k+1}$, the solution sequence $\{\mathbf{y}_k\}_k$ of (25) can be calculated by the similar process, and the derived iterative formula is written as

$$\mathbf{y}_{k+1} = S_{\frac{\lambda_2\mu_2}{c_2}}\left(\frac{\mu_2}{c_2}\mathbf{B}^*(\mathbf{s}-\mathbf{A}\mathbf{x}_{k+1}-\mathbf{B}\mathbf{y}_k)+\mathbf{y}_k\right). \tag{41}$$

With (40) and (41), an iterative thresholding shrinkage algorithm is proposed for solving (17), and its explicit steps

The proposed algorithm

(1) Input: initialize λ_1, λ_2, μ_1, μ_2, and set $k=1$, $\mathbf{x}_1=\mathbf{0}$ and $\mathbf{y}_1=\mathbf{0}$

(2) Repeat:

① calculate the residual error: $\mathbf{r}_k=\mathbf{s}-\mathbf{A}\mathbf{x}_k-\mathbf{B}\mathbf{y}_k$;

② forward projection of $\mathbf{r}_k$: $\mathbf{e}_1=\mathbf{A}^*\mathbf{r}_k$;

③ thresholding shrinkage with the parameters λ_1 and μ_1:

$$\mathbf{e}_{s1}=S_{\frac{\lambda_1\mu_1}{c_1}}\left(\mathbf{x}_k-\mathbf{e}_1\mu_1/c_1\right);$$

④ update the solution of $\mathbf{x}$: $\mathbf{x}_{k+1}=\mathbf{e}_{s1}$;

⑤ update the residual error: $\mathbf{r}_k'=\mathbf{s}-\mathbf{A}\mathbf{x}_{k+1}-\mathbf{B}\mathbf{y}_k$;

⑥ forward projection of $\mathbf{r}_k'$: $\mathbf{e}_2=\mathbf{B}^*\mathbf{r}_k$;

⑦ thresholding shrinkage with the parameters λ_2 and μ_2:

$$\mathbf{e}_{s2}=S_{\frac{\lambda_2\mu_2}{c_2}}\left(\mathbf{y}_k-\mathbf{e}_2\mu_2/c_2\right);$$

⑧ update the solution of $\mathbf{y}$: $\mathbf{y}_{k+1}=\mathbf{e}_{s2}$;

⑨ let $k\leftarrow k+1$;

Until convergence.

(3) Output: with the obtained $\mathbf{x}_k$, calculate the impulsive signal by $\mathbf{A}\mathbf{x}_k$.

Fig. 2. Explicit steps of the proposed algorithm.

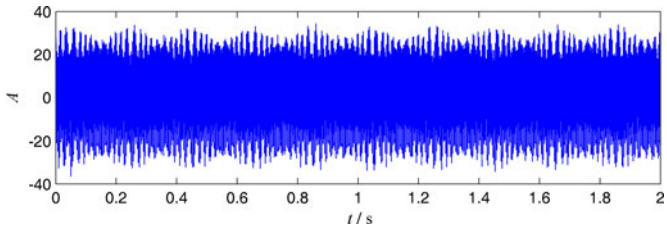

Fig. 3. Time-domain waveform of this simulated signal.

are described in Fig. 2. In this algorithm, the regular parameters can be set in the range of [0.1, 1], and the Lagrangian parameters can be set in the range of [0.01, 1]. Specifically, they depend on the constitution of the analyzed signal. With the feature identification indicators defined in [19], such as impulsive relative error and impulsive feature identification rate, a similar searching algorithm can be used to find the optimal parameters in the preset range. Owing to space limitations, we do not discuss the method for setting the parameters in detail.

IV. Simulation Analysis

In this section, the effectiveness of the proposed approach will be validated, and other three transient detection methods are used for comparison. According to the observation model given by (18), the simulated signal is defined as

$$\begin{aligned}s(t) &= h(t)+n(t)+20\sin(2\pi 600t)+10\sin(2\pi 300t)\\ &\quad +10\left[0.5+0.3\sin(2\pi 5t)\right]\cos\left[2\pi 350t+1.2\sin(2\pi 10t)\right]\end{aligned} \tag{42}$$

where $\xi=0.204$, $\omega_d=4467.9$, $h(t)$ is given by (5) with the mean period of 0.233, A_k is randomly set in the range of [3.23, 4.95], and $n(t)$ is the Gaussian white noise with the standard deviation of 1. The sampling frequency is 3000 Hz, and the data length is 6000. The time-domain waveform of this simulated signal is shown in Fig. 3. With the same parameters ω_d

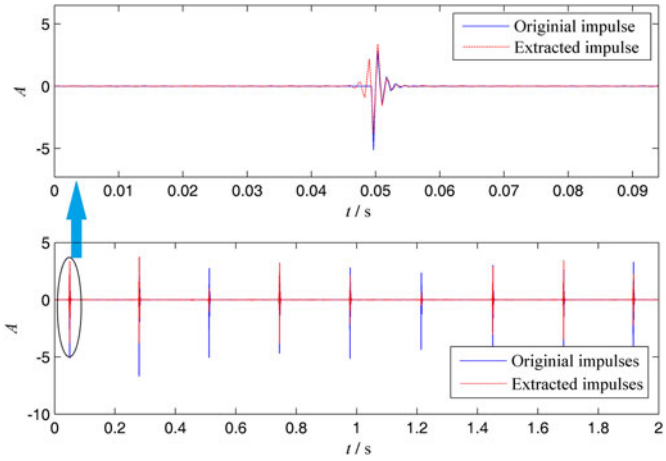

Fig. 4. Obtained impulses and the zoom plot of one impulse.

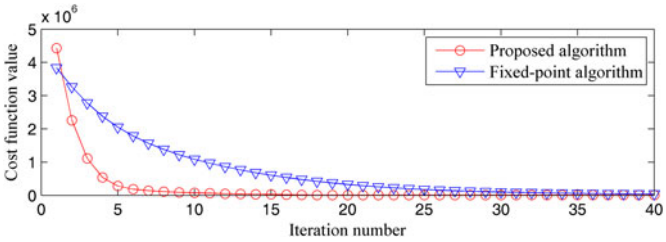

Fig. 5. Cost function history for the simulated signal.

and ξ, the impulsive wavelet is constructed, then the proposed iterative thresholding shrinkage algorithm is applied to extract the impulsive feature, and the obtained repetitive impulses and the zoom plot of one impulse are illustrated in Fig. 4. It can be seen from this figure that the repetitive impulsive signal is effectively extracted from the strong inference signal components (harmonics and AM-FM component) and the noise, and the shape of the extracted impulse is very similar to that of the original impulse. By calculation, the variance of error between the extracted impulse and the original impulse equals to 0.015. To validate convergence behavior of the proposed algorithm, the red curve of cost function (i.e., objective function) value versus iteration number is shown in Fig. 5. Like the proposed algorithm, a fixed-point algorithm, i.e., iterative-reweighted least squares [39], [40], is also used to solve the sparse optimization problem, and the obtained blue cost function curve is illustrated in Fig. 5. From this figure, we can see that the proposed algorithm converge faster than the fixed-point algorithm. Actually, with the related results in [36] and [41], it can be easily known that the proposed algorithm can converge to the minimum if the Proposition 1 holds.

For comparison, three typical methods are applied to detect the transient feature from the same signal. First, to assure that the extracted transient signal have the same length with the origin signal, we use the SK method based on the short-time Fourier transform for comparison in this study, and the obtained kurtogram is shown in Fig. 6. With the optimal carrier frequency and level given in Fig. 6, the transient signal is filtered out, which is shown in Fig. 7. It can be seen from Fig. 7 that the extracted

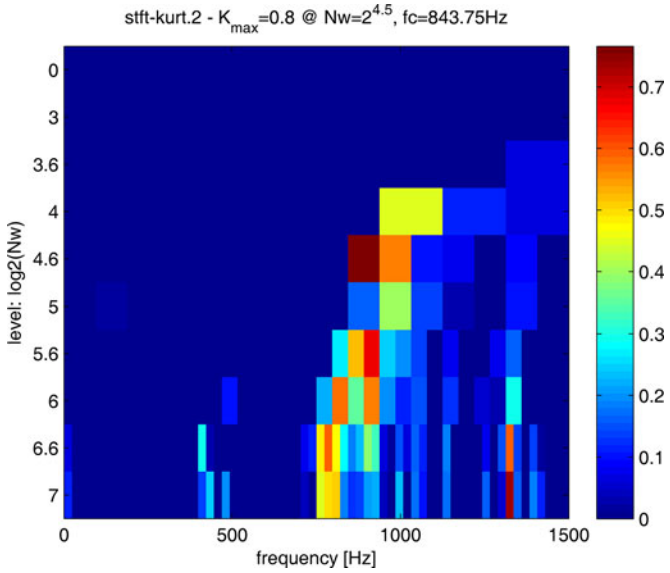

Fig. 6. Kurtogram of simulated signal obtained by the SK method.

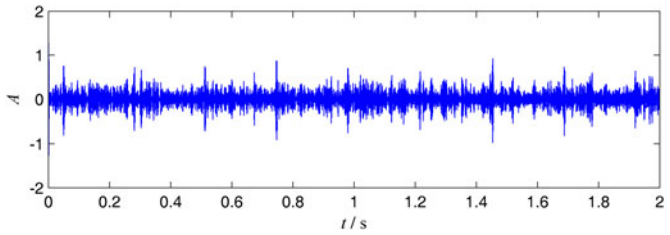

Fig. 7. Repetitive impulsive signal extracted by the SK method.

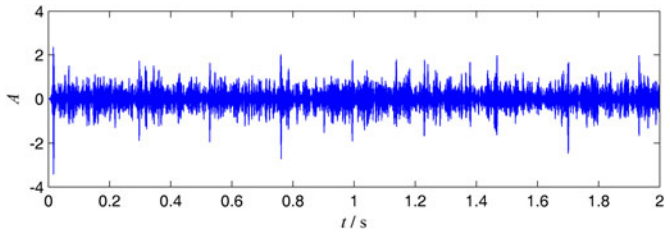

Fig. 8. Repetitive impulsive signal extracted by the optimized SK method based on simulated annealing.

impulses are submerged in a larger noise and the periodicity of impulses is ambiguous. Then, with the central frequency and frequency band obtained by the kurtogram, an optimized SK method based on simulated annealing [42] is used to process the same simulated signal to enhance the performance of transient detection. The located optimum frequency band's central frequency is 955.9 Hz, and the bandwidth is 196.4 Hz. The filtered signal is shown in Fig. 8. It is easy to note that the obtained impulsive signal still disturbed by the noise even if the impulses become more obvious. Finally, the resonance-based signal decomposition method by tunable Q-factor wavelets is applied to extract the transient component with low Q-factor, which is shown in Fig. 9. From this figure, we can see that a large number of false impulses are obtained and the noise still exists. Comparing Fig. 4 with Figs. 7–9, the approximately periodic impulsive feature extracted by the proposed method is more

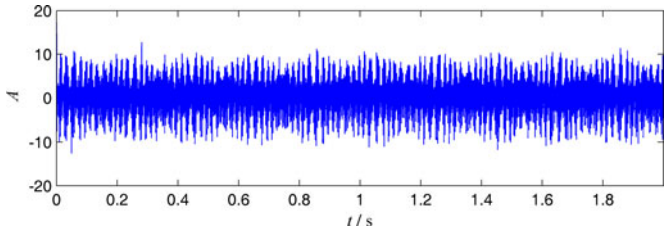

Fig. 9. Repetitive impulsive signal extracted by the resonance-based signal decomposition method.

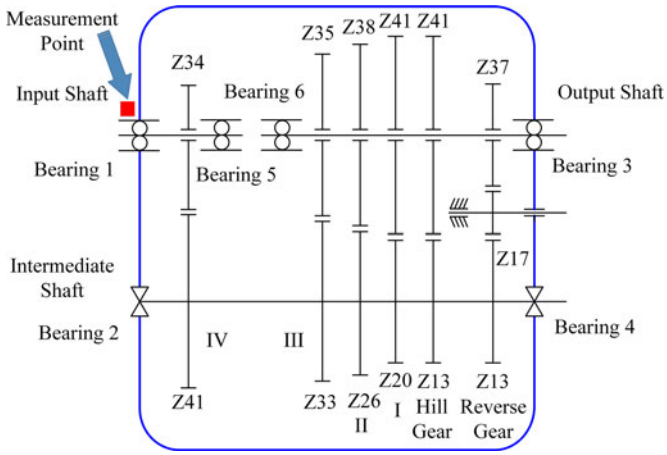

Fig. 10. Specific structure and measurement point of 5T-85 gearbox.

obvious and accurate than those extracted by the three typical methods. Therefore, the result of simulation analysis demonstrates the effectiveness and superiority of the proposed method in extracting weak repetitive feature from strong harmonics, AM-FM signals, and noise.

V. Application Into Real Bearing Fault Diagnosis

Automobile gearbox is the major component of the automotive transmission system. Since the running condition of an automobile is quite complex, the rolling bearings of the automobile gearbox are easily subjected to faults. Particularly, most of the faults occur in the inner race or outer race of the rolling bearing. When we measure the vibration signal at the case of the automobile gearbox, the observation data contain not only the faulty information but also noise and other components' vibrations, such as the AM-FM vibration signal coming from gear engagement and the vibration signal of the shaft related to the rotational speed. As a result, the proposed method is applied to extract the repetitive impulsive feature caused by the faulty bearing.

The tested 5t-85 automobile gearbox is a type of fixed-axis transmission, which has five forward gears, one reverse gear, and six bearing for supporting. The specific structure of the automobile gearbox and the measurement point is illustrated in Fig. 10. The bearing 1 is a 22NU15EC rolling bearing whose parameters are listed in Table I. Since the bearing 1 has a fault on its inner race and other gears and bearing are healthy, a piezoelectric accelerometer made by PCB corporation is placed on the case of the bearing 1 to acquire the vibration signal, as

TABLE I
PARAMETERS OF 22NU15EC ROLLING BEARING

Rolling body diameter	Pitch diameter	Contact angle	Number of roller
15 mm	104 mm	0°	18

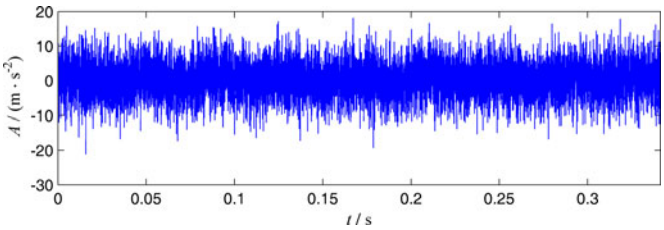

Fig. 11. Vibration signal of the bearing with inner race fault.

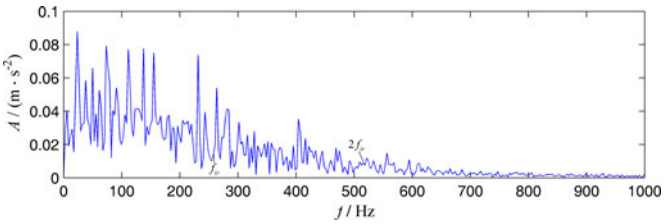

Fig. 12. Envelope spectrum of the bearing signal with inner race fault obtained by the SK method.

shown in Fig. 10. This gearbox is connected with a test engine, and its input rotating speed is set as 1500 r/min in the experiment, i.e., the rotating frequency is 25 Hz. Under such running condition, the sampling frequency and data length are set as 30 000 Hz and 10240. According to the bearing parameters in Table I, the inner race fault characteristic frequency f_{ir} is calculated as 257.6 Hz, and the wavelet parameter ω_d is estimated as 4537 Hz while the wavelet parameter ξ is empirically set as 0.3. The sampled vibration signal is illustrated in Fig. 11. With the SK method, the transient component is filtered out and its envelope spectrum is illustrated in Fig. 12. We can see from this figure that the spectral line at the fault characteristic frequency f_{ir} is submerged in other spectral lines produced by other components of the automobile gearbox, and there are no spectral peaks at the multiple of the fault characteristic frequency. Then, the resonance-based signal decomposition method by tunable Q-factor wavelets is applied to process the vibration signal. The obtained low-resonance signal (transients) and its envelope spectrum are illustrated in Fig. 13. It can be seen that the extracted transients are disturbed by harmonics and noise, and only an unremarkable spectral line exists at f_{ir}. The proposed method is thereupon applied to analyze this faulty vibration signal. The parameters are chosen by line searching in the preset range. With these parameters, the obtained repetitive impulsive signal is illustrated in Fig. 14. Since the fault characteristic frequency is high and the impulses caused by the fault are not obvious, the transients are very intensive and they are difficult to be directly observed from Fig. 14. Then, we give the zoom plot of Fig. 14 in 0.0883 and 0.0982 s, which is shown in Fig. 15. It can be

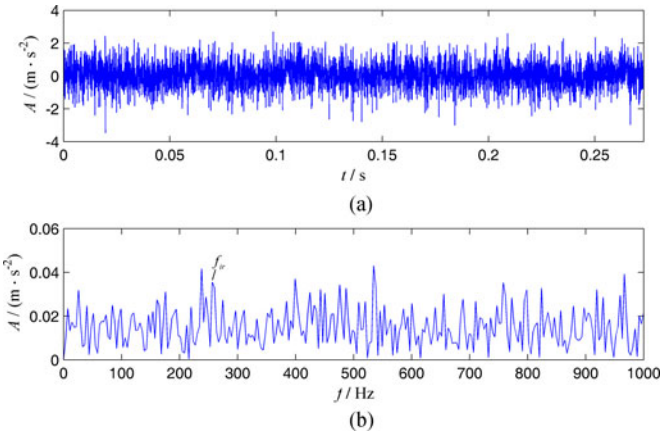

Fig. 13. Low-resonance signal and its envelope of the bearing signal with inner race fault. (a) Time-domain waveform. (b) Envelope spectrum.

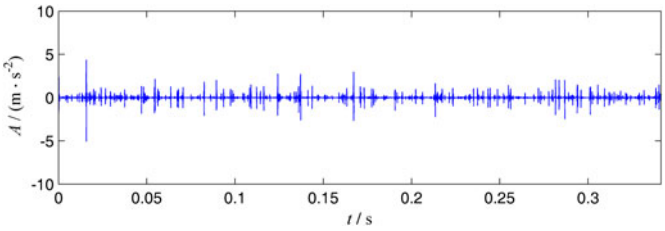

Fig. 14. Repetitive impulsive signal of a rolling bearing with inner race fault extracted by the proposed method.

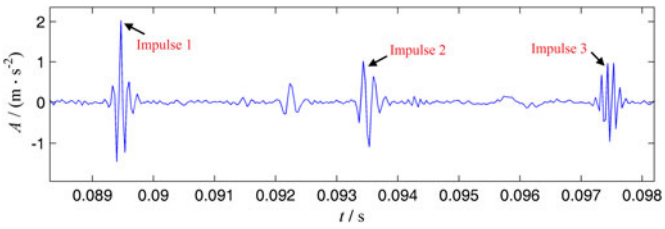

Fig. 15. Zoom plot of Fig. 10 in 0.0883 and 0.0982 s.

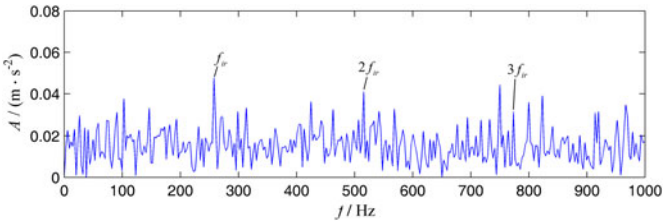

Fig. 16. Envelope spectrum of the repetitive impulsive signal of a rolling bearing with inner race fault.

seen from Fig. 15 that there actually exists approximately periodic impulses, and the mean repetitive period can be obtained by counting the time intervals between two adjacent impulses. Then, we can approximately compute the fault characteristic frequency. To accurately obtain the mean impulsive frequency, we calculate the envelope spectrum of the extracted transient signal as shown in Fig. 16. From this figure, we can see that the spectral peaks at f_{ir} and $2f_{ir}$ are quite obvious, and there is also a spectral peak at $3f_{ir}$. Thus, we can judge that the inner race has a local fault.

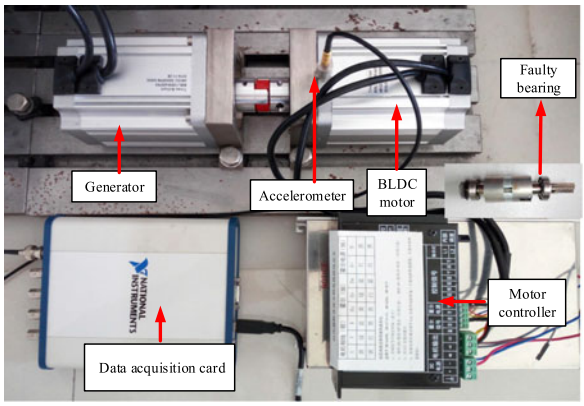

Fig. 17. Test rig for a BLDC motor with bearing fault.

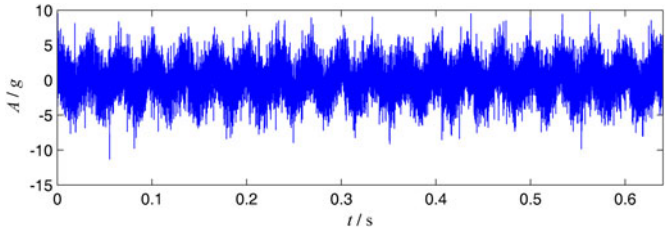

Fig. 18. Vibration signal of the bearing with outer race fault.

TABLE II
PARAMETERS OF 6002Z ROLLING BEARING

Rolling body diameter	Pitch diameter	Contact angle	Number of roller
4.75 mm	23.5 mm	0°	9

Brushless direct-current (BLDC) motors are widely used in industrial and domestic applications. Unfortunately, they may be easily subjected to bearing faults, which accounts for 40% to 50% of the total faults in motors [43]. To further demonstrate the advantage of the proposed approach in the area of industrial electronics, a BLDC motor (80BL110S50) with bearing fault is used for diagnosis. The setup of test rig is shown in Fig. 17. It mainly consists of a BLDC motor, a generator, a coupling, a NI data acquisition card, and an accelerometer. The BLDC motor drives the generator through a coupling. The 6002Z rolling bearing of the driving motor has a slight defect on its outer race, and other components are healthy. The accelerometer is placed on the case of the motor to sample the vibration signal. The sampling frequency is set as 25 600 Hz, and the acquired signal is shown in Fig. 18 when the rotating speed is 1810 r/min. The parameters of 6002Z bearing are list in Table II. From this table, we can calculate the outer race fault characteristic frequency f_{or} as 108.3 Hz. Similarly, via the SK method, the envelope spectrum of the transient component obtained by the optimal carrier frequency and level is illustrated in Fig. 19. It can be seen from this figure that there is an obvious spectral peak f_{or}, whereas there are no remarkable spectral peaks at the multiple of f_{or}. Then, the resonance-based signal decomposition method is also used. The obtained low-resonance signal and its envelope

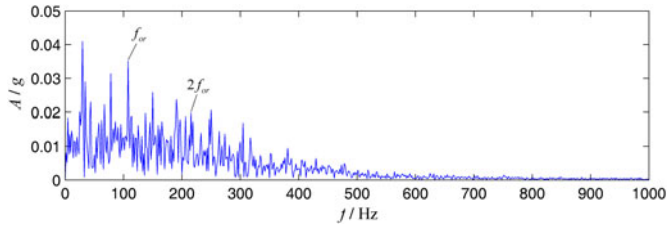

Fig. 19. Envelope spectrum of the bearing signal with outer race fault obtained by the SK method.

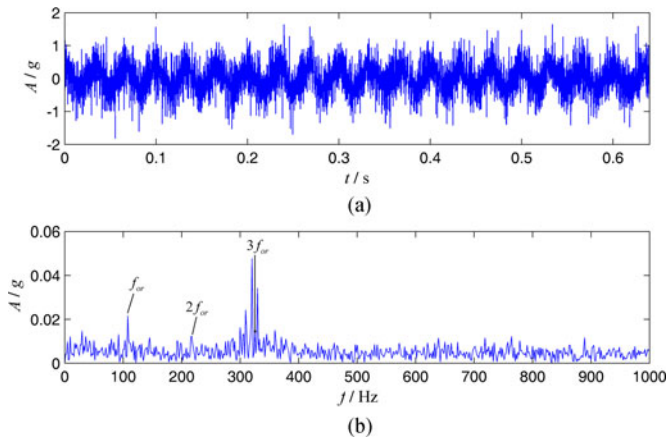

Fig. 20. Low-resonance signal and its envelope of the bearing signal with outer race fault. (a) Time-domain waveform. (b) Envelope spectrum.

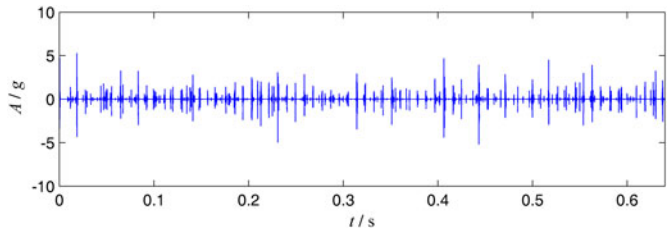

Fig. 21. Repetitive impulsive signal of a rolling bearing with outer race fault extracted by the proposed method.

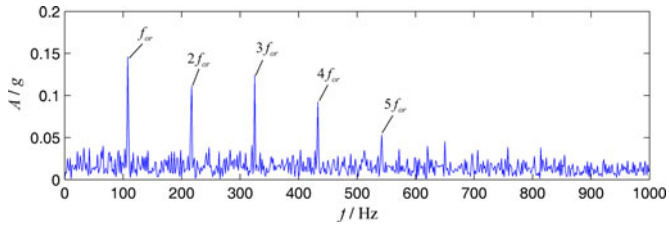

Fig. 22. Envelope spectrum of the repetitive impulsive signal of a rolling bearing with outer race fault.

spectrum are illustrated in Fig. 20. We can easily see that the extracted result still contains harmonic components, AM-FM components, and noise, and the spectral lines at f_{or}, $2f_{or}$, and $3f_{or}$ are not prominent. Thus, the proposed approach is applied to analyze this faulty vibration signal. Within the computation, the wavelet parameter ω_d is estimated as 4487 Hz, while the wavelet parameter ξ is also empirically set as 0.3. The obtained transient signal and its envelope spectrum are, respectively, illustrated in Figs. 21 and 22. It can be seen from this figure that there

are clear spectral peaks at f_{or}, $2f_{or}$, $3f_{or}$, $4f_{or}$, and $5f_{or}$. It follows that the bearing 1 has an outer race fault.

VI. Conclusion

This paper presented a novel impulsive feature extraction method for rolling bearing fault detection. The two major contributions of this study lies in that 1) a new family of model-based impulsive wavelets is constructed and 2) the corresponding sparse representation method is proposed for extracting the pseudocyclostationary impulsive signal from the strong interferences (e.g., harmonics and AM-FM component) and noise. The new impulsive wavelet is constructed based on the vibration model of the faulty rolling bearing, thus it can accurately represent the repetitive impulses caused by the bearing fault. Moreover, this family of impulsive wavelets satisfies the admissibility condition, therefore they can be used to construct the dictionary for sparse component separation. With the impulsive wavelet basis and Fourier basis, a sparse optimization problem is formulated to extract the repetitive impulsive signal. To assure the convexity of the optimization problem, the specific conditions of two penalty parameters are derived. After that an iterative thresholding shrinkage algorithm based on splitting idea and separable surrogate function is proposed to solve this problem. This algorithm has a fast convergence rate, which is validated by the simulation test. The results of simulation and engineering application demonstrate the advantage and superiority of the proposed approach compared to the widely-used SK method, the optimized SK method based on simulated annealing, and the resonance-based signal decomposition method by tunable Q-factor wavelets.

In this study, the wavelet parameters were the same as the parameters of the simulated impulse in simulation, while the parameters can be estimated via the structure of the bearing system or set as empirical values in real application. Therefore, to further improve the accuracy of impulsive feature extraction, the extension of this study can focus on optimizing the parameters of the impulsive wavelet and the parameters used in the proposed iterative thresholding shrinkage algorithm.

Appendix

According to (25) and (26), their second derivatives can be, respectively, calculated as

$$d_1''(\mathbf{x}, \mathbf{x}_k) = c_1\mathbf{I} - \mu_1\mathbf{A}^*\mathbf{A} \tag{A.1}$$

$$d_2''(\mathbf{y}, \mathbf{y}_k) = c_2\mathbf{I} - \mu_2\mathbf{B}^*\mathbf{B}. \tag{A.2}$$

To assure that $d_1(\cdot)$ and $d_2(\cdot)$ are, respectively, strictly convex, (A.1) and (A.2) must satisfy

$$d_1''(\mathbf{x}, \mathbf{x}_k) = c_1\mathbf{I} - \mu_1\mathbf{A}^*\mathbf{A} > 0 \tag{A.3}$$

$$d_2''(\mathbf{y}, \mathbf{y}_k) = c_2\mathbf{I} - \mu_2\mathbf{B}^*\mathbf{B} > 0. \tag{A.4}$$

By the use of (29) and (30), we can easily obtain

$$c_1 > p_1\mu_1 \tag{A.5}$$

$$c_2 > p_2\mu_2. \tag{A.6}$$

Therefore, the proposition is proved.

References

[1] J. Wang, Y. Peng, and W. Qiao, "Current-aided order tracking of vibration signals for bearing fault diagnosis of direct-drive wind turbines," *IEEE Trans. Ind. Electron.*, vol. 63, no. 10, pp. 6336–6346, Oct. 2016.

[2] W. Wang *et al.*, "A new test part to identify performance of five-axis machine tool—Part I: Geometrical and kinematic characteristics of S part," *Int. J. Adv. Manuf. Technol.*, vol. 79, pp. 1–10, Jan. 2015.

[3] C. Li *et al.*, "Rolling element bearing defect detection using the generalized synchrosqueezing transform guided by time-frequency ridge enhancement," *ISA Trans.*, vol. 60, pp. 274–284, Jan. 2016.

[4] A. Soualhi *et al.*, "Prognosis of bearing failures using hidden Markov models and the adaptive neuro-fuzzy inference system," *IEEE Trans. Ind. Electron.*, vol. 61, no. 6, pp. 2864–2874, Jun. 2014.

[5] X. Jin, M. Zhao, T. W. S. Chow, and M. Pecht, "Motor bearing fault diagnosis using trace ratio linear discriminant analysis," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2441–2451, May 2014.

[6] C. Li *et al.*, "Extracting repetitive transients for rotating machinery diagnosis using multiscale clustered grey infogram," *Mech. Syst. Signal Process.*, vols. 76/77, pp. 157–173, Aug. 2016.

[7] E. C. C. Lau and H. W. Ngan, "Detection of motor bearing outer raceway defect by wavelet packet transformed motor current signature analysis," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 10, pp. 2683–2690, Oct. 2010.

[8] X. Gong and W. Qiao, "Current-based mechanical fault detection for direct-drive wind turbines via synchronous sampling and impulse detection," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1693–1702, Mar. 2015.

[9] B. Eftekharnejad *et al.*, "The application of spectral kurtosis on acoustic emission and vibrations from a defective bearing," *Mech. Syst. Signal Process.*, vol. 25, no. 1, pp. 266–284, Jan. 2011.

[10] M. Kang, J. Kim, and J. M. Kim, "An FPGA-based multicore system for real-time bearing fault diagnosis using ultrasampling rate AE signals," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2319–2329, Apr. 2015.

[11] W. Zhou, T. G. Habetler, and R. G. Harley, "Bearing condition monitoring methods for electric machines: A general review," in *Proc. IEEE Int. Symp. Diagn. Electr. Mach., Power Electron. Drives*, Sep. 6–8, 2007, pp. 3–6.

[12] L. Frosini, C. Harlisca, and L. N. Szabó, "Induction machine bearing fault detection by means of statistical processing of the stray flux measurement," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1846–1854, Mar. 2015.

[13] J. Tian *et al.*, "Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis," *IEEE Trans. Ind. Electron.*, vol. 63, no. 3, pp. 1793–1803, Mar. 2016.

[14] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, Feb. 2011.

[15] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[16] L. Cohen, *Time–Frequency Analysis*. New York, NY, USA: Prentice-Hall, 1995.

[17] J. Antoni, "The spectral kurtosis: A useful tool for characterising nonstationary signals," *Mech. Syst. Signal Process.*, vol. 20, no. 2, pp. 282–307, Feb. 2006.

[18] V. C. Leite *et al.*, "Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1855–1865, Mar. 2015.

[19] Z. Du *et al.*, "Sparse feature identification based on union of redundant dictionary for wind turbine gearbox fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6594–6605, Oct. 2015.

[20] N. G. Nikolaou and I. A. Antoniadis, "Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets," *Mech. Syst. Signal Process.*, vol. 16, no. 4, pp. 677–694, Jul. 2002.

[21] I. S. ozchalooi and M. Liang, "A joint resonance frequency estimation and in-band noise reduction method for enhancing the detectability of bearing fault signals," *Mech. Syst. Signal Process.*, vol. 22, no. 4, pp. 915–933, May 2008.

[22] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.

[23] Y. Qin, Y. Mao, and B. Tang, "Vibration signal component separation by iteratively using basis pursuit and its application in mechanical fault detection," *J. Sound Vib.*, vol. 332, no. 20, pp. 5217–5235, Sep. 2013.

[24] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inf. Theory,* vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[25] D. L. Donoho *et al.*, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory.*, vol. 53, no. 12, pp. 1094–1121, Feb. 2012.
[26] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.
[27] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing: Closing the gap between performance and complexity," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
[28] W. He, Y. Ding, Y. Zi, and I. W. Selesnick, "Repetitive transients extraction algorithm for detecting bearing faults," *Mech. Syst. Signal Process.*, vol. 84, pp. 227–244, Feb. 2017.
[29] Y. Ding, W. He, B. Chen, Y. Zi, and I. W. Selesnick, "Detection of faults in rotating machinery using periodic time-frequency sparsity," *J. Sound Vib.*, vol. 382, pp. 357–378, Nov. 2016.
[30] I. W. Selesnick, "Resonance-based signal decomposition: A new sparsity-enabled signal analysis method," *Signal Process.*, vol. 91, no. 12, pp. 2793–2809, Dec. 2011.
[31] S. Mallat, *A Wavelet Tour on Signal Processing*, 3rd ed. New York, NY, USA: Academic, 2008.
[32] P. McFadden and J. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *J. Sound Vib.*, vol. 96, no. 1, pp. 69–82, Sep. 1984.
[33] Y. Qin, J. Xing, and Y. Mao, "Weak transient fault feature extraction based on an optimized Morlet wavelet and kurtosis," *Meas. Sci. Technol.*, vol. 27, Aug. 2016, Art. no. 085003.
[34] J. L. Starck, M. Elad, and D. Donoho, "Redundant multiscale transforms and their application for morphological component analysis," *Adv. Imag. Electron Phys.*, vol. 1, pp. 287–348, May 2004.
[35] J. L. Starck, M. Elad, and D. Donoho, "Image decomposition via the combination of sparse representation and a variational approach," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.
[36] J. Eckstein and D. Bertsekas, "On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 5, pp. 293–318, 1992.
[37] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
[38] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.
[39] T. Adeyemi and M. Davies, "Sparse representations of images using overcomplete complex wavelets, statistical signal processing," in *Proc. 2005 IEEE/SP 13th Workshop Stat. Signal Process.*, Bordeaux, France, 2005, vols. 1/2, pp. 805–809.
[40] M. Elad, *Sparse and Redundant Representations——From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer Science + Business Media, 2010.
[41] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, Jul. 2005.
[42] J. Tian, C. Morillo, and M. G. Pecht, "Rolling element rearing fault diagnosis using simulated annealing optimized spectral kurtosis," in *Proc. IEEE Conf. Progn. Health Manage*, Jun. 24–27, 2013, pp. 1–5.
[43] S. Lu and X. Wang, "A new methodology to estimate the rotating phase of a BLDC motor with its application in variable-speed bearing fault diagnosis," *IEEE Trans. Power Electron.*, to be published.

Yi Qin (M'17) received the B.Eng. and Ph.D. degrees in mechanical engineering from Chongqing University, Chongqing, China, in 2004 and 2008, respectively.

Since January 2009, he has been with Chongqing University, where he is currently an Associate Professor with the College of Mechanical Engineering. His current research interests include signal processing, fault prognosis and health management, mechanical dynamics, and smart structure.

Dr. Qin is a Member of the SPIE.