A Noise Resistant Correlation Method for Period Detection of Noisy Signals

Wei Fan, Yongxiang Li, Kwok Leung Tsui, and Qiang Zhou, *Member, IEEE*

***Abstract*—This paper develops a novel method called the noise resistant correlation method for detecting the hidden period from the contaminated (noisy) signals with strong white Gaussian noise. A novel correlation function is proposed based on a newly constructed periodic signal and the contaminated signal to effectively detect the target hidden period. In contrast with the conventional autocorrelation analysis (AUTOC) method, this method demonstrates excellent performance, especially when facing strong noise. Fault diagnoses of rolling element bearings and gears are presented as application examples and the performance of the proposed method is compared with that of the AUTOC method.**

***Index Terms*—Period detection, correlation, white noise, fault diagnosis.**

I. Introduction

PERIODIC impulsive signals have been of great interest for several decades in many different areas, such as ECG signals [1], power-line communications [2], [3], EEG signals, [4], and mechanical vibration signals [5], [6]. The periodic impulsive signatures are obscured by random noise when sensors are used to obtain the signal. Sometimes masking effects of noise can make the observed signal inadequate for aimed applications. Since the hidden period, used to represent the period of a periodic signal buried in background noise, can reveal the operational status of the system, there has been significant research interest in period detection in recent years [7]–[9]. The detection of the unknown period from contaminated (noisy) signals under mild noise has been well studied. However, due to the unobservable random nature of strong background noise, it is still extremely challenging to detect the period in such scenarios. Therefore, it is necessary to develop more effective and efficient techniques to detect the hidden period from weak signals buried in strong noise.

The period detection problem can be addressed by exploiting the knowledge about the characteristics of the source signal or its transformed versions. Typically, three different types of signal can be handled: the frequency-domain signal, the time-frequency domain signal, and the time-domain signal. The main idea of the frequency-domain analysis method is that the frequency spectrum of a signal consists of a series of impulses at the fundamental frequency and its harmonics when the signal is periodic [10], and thus the period of the signal can be estimated from the frequency spectrum. When the signal is polluted by strong noise, however, the feature frequency of the original periodic signal is likely to be buried by the frequencies of the noise, thus making period detection extremely difficult. The time-frequency domain analysis method uses hybrid period detectors which incorporate features of both the time-domain and frequency-domain approaches [10], [11]. The time-frequency analysis method can provide local information of source signal within a certain time range or frequency range. Thus, this approach is mainly used to handle signals whose frequencies vary with time [12] and hence is not suitable for our cases. In contrast, the time-domain analysis method operates directly on the waveform of the signal to estimate the period and thus provides a potential way to detect the hidden period from strong white noise. For example, autocorrelation measurements [13] perform well as period detectors if infinite number of data is given. Other widely used time-domain analysis methods are peak and valley measurements [14], and zero-crossing measurements [15]. In this paper, therefore, we mainly use the time-domain analysis method for period detection.

There are two main sub-classes in the framework of the time-domain analysis method: one based on prior knowledge, and one without prior knowledge. A typical method using prior knowledge is the signal sparse representation (SSP) method [16]. With manually constructed dictionaries, the measured signal can be represented by a coefficient vector, from which the periodic parameter can be extracted. Various studies that use this approach on mechanical signals can be found in [17]–[19], and nice results on period extraction have been presented in the literature. A critical problem of the SSP method is how to construct a proper dictionary to suppress the undesired effects of background noise. Cai *et al.* [17] used the tunable Q-factor wavelet dictionary to address the gearbox fault feature extraction problem due to the similarity between the basis and the gear fault transient

Manuscript received June 12, 2017; revised November 28, 2017; accepted February 27, 2018. Date of publication March 8, 2018; date of current version April 17, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Dennis Wei. This work was supported by Research Grants Council Project under Grant T32-101/15-R. *(Corresponding author: Yongxiang Li.)*

W. Fan is with the School of Mechanical Engineering, Jiangsu University, Zhenjiang 212000, China, and also with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong (e-mail: weifan8-c@my.cityu.edu.hk).

Y. Li and K. L. Tsui are with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong (e-mail: yongxili-c@my.cityu.edu.hk; kltsui@cityu.edu.hk).

Q. Zhou is with the Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721 USA (e-mail: q.zhou@arizona.edu).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org provided by the authors. The material includes a pdf file wihch contains detailed proofs for the variance of the noise resistant correlation method. This material is about 0.2 MB.

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/TSP.2018.2813305

waveform. Similarly, Fan *et al.* [18] used the Morlet wavelet basis to construct a dictionary to represent the gearbox fault vibration signal. In spite of its merits in detecting the period, the SSP method depends heavily on the accuracy of the dictionary constructed manually based on the prior knowledge. Thus, when faced with an unknown periodic signal without prior knowledge, the SSP method becomes unreliable and unsatisfactory.

Among the time-domain analysis methods without prior knowledge, the autocorrelation analysis (AUTOC) method is still one of the most robust and reliable period detectors [13], [20]. The autocorrelation is computed directly using the source waveform and thus is fairly straightforward. When the source signal is periodic with sufficient length, the autocorrelation function is in theory also periodic with the same period. Although the AUTOC method has some advantages, it still has some problems in practice. In particular, the AUTOC method performs poorly when the background noise is much stronger than the original signal, due to the limited length of the signals in real applications.

A wide variety of solutions have been proposed to address the above problems. Tan *et al.* [21] proposed a novel difference function called the average magnitude autocorrelation function (AMDF) to improve the period detection power. The AMDF is a variation of the autocorrelation function in which a difference signal is formed between the delayed and contaminated signals, and the absolute magnitude is taken at each delay value. Shimamura *et al.* [22] proposed a new period extraction method which used an autocorrelation function weighted by the inverse of the AMDF to improve the accuracy of the period extraction effect. Facing with the unsatisfactory performance when applying the AUTOC method to finite vibration signals with strong background noise, several studies [23], [24] have utilized the autocorrelation of the wavelet-based denoised vibration signal to improve the detection power. Although these methods have made some progress, addressing the above problem is still a challenging task in the signal processing field. Thus, it is the aim of this paper to detect the hidden period from strong background noise in real applications that usually have finite data and little prior knowledge by investigating the characteristics of the time-domain signal.

We propose a novel technique called the noise resistance correlation (NRC) method to detect the hidden period in the original signal. The proposed method separates the original signal into several segments and takes their average to construct a new signal which significantly helps to reduce effects of strong noise. A correlation function between the original and new signals is then defined in order to detect the period information. Unlike the variations of the AUTOC method, our approach offers a new framework for extracting the period information and sheds light on period detection under strong noise. The properties of the proposed approach are further investigated to verify its ability to detect the hidden period from strong background noise. The two main advantages of the developed technique are:

- without prior knowledge, and
- highly sensitive.

This paper is organized as follows. Section II presents a detailed review of the AUTOC method. The simulation results for this method with a large sample size and a limited sample size are also given. The proposed method is described in Section III. A performance analysis of the proposed method is provided in Section IV followed by simulation studies. Two different experimental tests comparing the performance of the proposed method with that of the AUTOC method are presented in Section V. Conclusions are drawn in Section VI.

II. Review of the Autocorrelation Analysis Method

The AUTOC method, as one of the most widely used period detection methods, performs well when infinite number of data is provided [21]. However, it cannot deal with the strong noise with finite sample points. Thus, in this section, we first review the AUTOC method and reveal its advantages and disadvantages. Formally, the measured signal can be modeled by the sum of the original periodic signal and the noise, that is, the measured signal $y(t)$ can be represented by [18]

$$y(t) = x(t) + \epsilon(t), \tag{1}$$

where $x(t)$ is a bounded periodic signal with unknown period T_0, i.e., $|x(t)| \leq M$, and $\epsilon(t)$ independently follows the norm distribution $\mathbb{N}\left(0, \sigma^2\right)$. It should be noted that although the noise part $\epsilon(t)$ is not Riemann integrable and hence neither is $y(t)$, we still define integration on them for convenience, as shown in the appendix, together with some properties for the integration in the lemma.

A traditional way to extract the period information of the signal $x(t)$ is by using the AUTOC method [13], given by

$$\mathcal{R}_{yy}(\tau) \equiv \lim_{\Gamma\to\infty} \frac{1}{\Gamma - \tau} \int_0^{\Gamma-\tau} y(t)\, y(t+\tau)\, dt, \tag{2}$$

where Γ is the time length of signal $y(t)$ and τ is the time delay. Autocorrelation is a transformation of the signal that is useful for displaying the structure in the waveform. By plugging (1) into (2), the autocorrelation of the signal $y(t)$ is simplified as

$$\mathcal{R}_{yy}(\tau) = \mathcal{R}_{xx}(\tau) + \mathcal{R}_{\epsilon\epsilon}(\tau). \tag{3}$$

The characteristic of $\mathcal{R}_{xx}(\tau)$ is that it has a larger value when $x(t)$ is more similar to $x(t+\tau)$ and has peaks $\mathcal{P}_x(\Gamma - \tau)$ at $\tau = iT_0$, where $\mathcal{P}_x(\Gamma) = \int_0^\Gamma x^2(t)/\Gamma\, dt$ is the average power of the signal $x(t)$ and i is an integer. Furthermore, by definition $\mathcal{R}_{\epsilon\epsilon}(\tau) = 0$ if $\tau > 0$. Therefore, (3) becomes

$$\mathcal{R}_{yy}(\tau) = \mathcal{R}_{xx}(\tau) + \delta_{\tau=0}\mathcal{R}_{\epsilon\epsilon}(0) \tag{4}$$

where δ denotes the delta function. This implies that $\mathcal{R}_{yy}(\tau)$ also has peaks $\mathcal{P}_x(\Gamma - \tau)$ at $\tau = iT_0$. Based on these properties, the AUTOC method theoretically provides robust performance against noise when an infinite sample size is given.

Here, we use a simulation to show the performance of the AUTOC method. The sampling frequency is $f_s = 500$ Hz, and the sample points of the signal are set as $n = 20,000,000$ to simulate the infinite sample case. The contaminated periodic

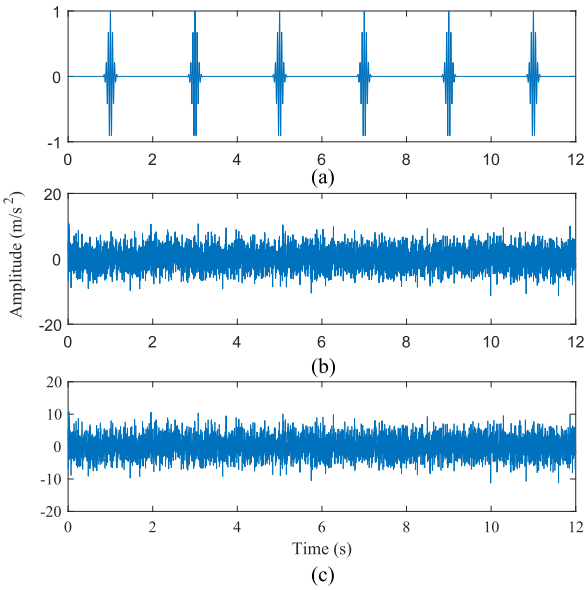


Fig. 1. A Section of the original periodic signal, white noise signal, and contaminated signal with 6,000 sample points.

signal [18] is modeled as

$$y(t) = x(t) + \epsilon(t) = \sum_{k=1}^{\left\lfloor \frac{n}{f_s T_0} \right\rfloor} e^{\frac{-\xi(2\pi f(t-kT_0-\tau_0))^2}{\sqrt{1-\xi^2}}} \cos\left(2\pi f\left(t-kT_0-\tau_0\right)\right) + \epsilon(t), \tag{5}$$

where $\lfloor a \rfloor$ denotes the maximum integer smaller than a, $\xi = 0.01$, $f = 20$ Hz, $T_0 = 2$ s, $\tau_0 = 1$ s, and the variance of the noise $\epsilon(t)$ is $\sigma^2 = 9$. The signal to noise ratio (SNR) is used to measure the noise level, which is defined as

$$\text{SNR} = 10 \log \frac{\mathcal{P}_x(\Gamma)}{\mathcal{P}_\epsilon(\Gamma)}, \tag{6}$$

where $\mathcal{P}_x(\Gamma)$ and $\mathcal{P}_\epsilon(\Gamma)$ are the average power of signal $x(t)$ and $\epsilon(t)$, respectively. A higher SNR value indicates a stronger fault signal relative to the background noise, and hence the signal is easier to detect. The SNR of the contaminated signal in the simulation is -25.569 dB.

To clearly show the waveform of the simulated signal, a section (6,000 points) of the original periodic signal, the white noise signal, and the contaminated signal, are shown in Fig. 1 (a), Fig. 1 (b), and Fig. 1 (c), respectively.

Then, we apply the AUTOC method to analyze the signal, and the results are shown in Fig. 2. We can see that the period of the original signal can be clearly detected from Fig. 2 when given a large sample size. In real applications, however, the length of the signal cannot be infinite. The results, when the number of points of the signal is reduced to 200,000, are shown in Fig. 3.

Fig. 3 shows that the period of the original signal cannot be detected by using the AUTOC method when the length of the signal is not sufficiently long and the SNR is very small (i.e., strong background noise). This phenomenon is theoretically analyzed below.

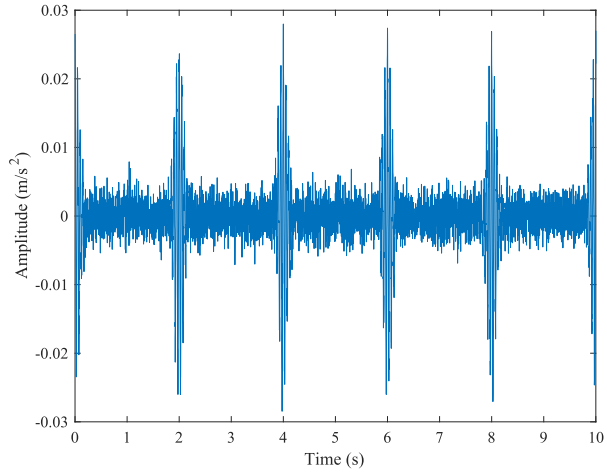


Fig. 2. Results of the AUTOC method with 20,000,000 sample points.

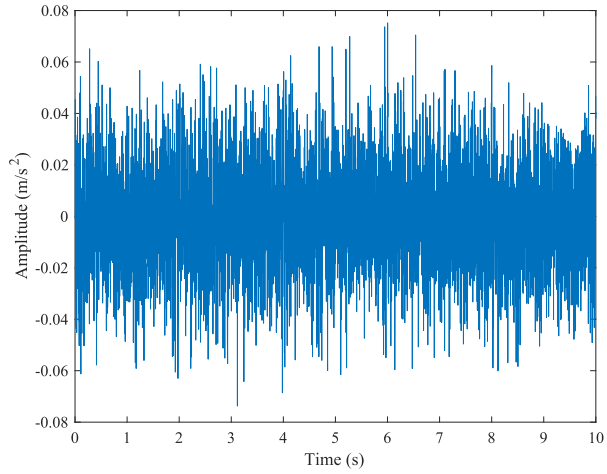


Fig. 3. Results of the AUTOC method with 200,000 sample points.

By expanding the autocorrelation function in (2) of the measured signal $y(t)$, one gets

$$\mathcal{R}_{yy}(\tau) = \lim_{\Gamma\to\infty} \int_0^{\Gamma-\tau} \frac{(x(t)+\epsilon(t))(x(t+\tau)+\epsilon(t+\tau))}{\Gamma-\tau} dt. \tag{7}$$

Assume there are N_0 samples in each period collected during the m_0 periods in the finite case, and $\varsigma = \tau m_0/\Gamma$ denotes the number of periods delayed by time τ. Denote the vectors $\mathbf{x}_{r,s}$ and $\boldsymbol{\epsilon}_{r,s}$ respectively as the discrete values of $x(t)$ and $\epsilon(t)$ in the range of $[r, s)$. Thus, the discrete version of $\mathcal{R}_{yy}(\tau)$ is

$$\mathcal{R}_d = \frac{\mathbf{x}_{0,\Gamma-\tau}^T \mathbf{x}_{\tau,\Gamma} + \mathbf{x}_{0,\Gamma-\tau}^T \boldsymbol{\epsilon}_{\tau,\Gamma} + \mathbf{x}_{\tau,\Gamma}^T \boldsymbol{\epsilon}_{0,\Gamma-\tau} + \boldsymbol{\epsilon}_{0,\Gamma-\tau}^T \boldsymbol{\epsilon}_{\tau,\Gamma}}{(m_0-\varsigma)N_0}. \tag{8}$$

Its variance can be obtained by

$$\text{Var}(\mathcal{R}_d) = \text{Var}\left(\frac{\mathbf{x}_{0,\Gamma-\tau}^T \boldsymbol{\epsilon}_{\tau,\Gamma} + \mathbf{x}_{\tau,\Gamma}^T \boldsymbol{\epsilon}_{0,\Gamma-\tau} + \boldsymbol{\epsilon}_{0,\Gamma-\tau}^T \boldsymbol{\epsilon}_{\tau,\Gamma}}{(m_0-\varsigma)N_0}\right)$$

$$= \frac{\sigma^2}{(m_0-\varsigma)N_0}\left(\frac{\|\mathbf{x}_{0,\Gamma-\tau}\|^2 + \|\mathbf{x}_{\tau,\Gamma}\|^2 + 2\mathbf{x}_{0,\Gamma-2\tau}^T \mathbf{x}_{2\tau,\Gamma}}{(m_0-\varsigma)N_0} + \sigma^2\right)$$

$$\geq \frac{\sigma^2}{(m_0-\varsigma)N_0}\left(\frac{\|\mathbf{x}_{0,\Gamma-\tau}\|^2+\|\mathbf{x}_{\tau,\Gamma}\|^2-2\|\mathbf{x}_{0,\Gamma-2\tau}\|\,\|\mathbf{x}_{2\tau,\Gamma}\|}{(m_0-\varsigma)N_0}+\sigma^2\right)$$

$$> \frac{\sigma^2}{(m_0-\varsigma)N_0}\left(\frac{\left(\|\mathbf{x}_{0,\Gamma-\tau}\|-\|\mathbf{x}_{\tau,\Gamma}\|\right)^2}{(m_0-\varsigma)N_0}+\sigma^2\right)$$

$$> \frac{\sigma^4}{(m_0-\varsigma)N_0}, \tag{9}$$

and similarly the upper bound can be obtained by

$$\mathrm{Var}(\mathcal{R}_d) < \frac{\sigma^2}{(m_0-\varsigma)N_0}\left(\frac{\left(\|\mathbf{x}_{0,\Gamma-\tau}\|+\|\mathbf{x}_{\tau,\Gamma}\|\right)^2}{(m_0-\varsigma)N_0}+\sigma^2\right). \tag{10}$$

Since $\frac{\|\mathbf{x}_{0,\Gamma-\tau}\|^2}{(m_0-\varsigma)N_0} \approx \frac{\|\mathbf{x}_{\tau,\Gamma}\|^2}{(m_0-\varsigma)N_0} \approx \mathcal{P}_x(\Gamma-\tau)$, then we can approximately write

$$\mathrm{Var}(\mathcal{R}_d) = \frac{\sigma^2}{(m_0-\varsigma)N_0}\left(4\mathcal{P}_x(\Gamma-\tau)c+\sigma^2\right), \tag{11}$$

where c is a constant and $0<c<1$. This means that the variance of $\mathcal{R}_d$ is dominated by $\sigma^2=\mathcal{P}_\epsilon(\Gamma)$ when the SNR is very small, i.e., the noise is very strong. Recall that the maximum value of $\mathcal{R}_{yy}(\tau)$ is approximately $\mathcal{P}_x(\Gamma-\tau)$ from (4), much smaller than σ^2 when the SNR is very small. Thus, $\mathcal{R}_{yy}(\tau)$ is easily buried by the noise when the length of the signal is not sufficiently long (a small m_0 or Γ). It should be noted that although increasing the sampling frequency (i.e., increasing N_0) can also reduce the variance, in practice more samples are usually obtained by increasing Γ, the time length of the signal. Our analysis in this section reveals that the AUTOC method can perform badly in period detection under strong noise in practice. This motivates us to develop a novel approach to improve the detection sensitivity in such cases.

III. Noise Resistant Correlation Method

In this section, we develop the NRC method for sensitively detecting the period information under strong white noise. To estimate the unknown period T_0 from the signal $y(t)$ as stated in (1), we construct a new periodic signal based on the original signal $y(t)$. Given $T>0$, we define the new signal $f_T(t)$ such that

$$f_T(t) = \begin{cases} \frac{1}{m}\sum_{i=0}^{m-1} y(t+iT), & 0\leq t<T \\ f_T(t-\lfloor t/T\rfloor T), & T\leq t\leq \Gamma \end{cases}, \tag{12}$$

where $\lfloor a\rfloor$ denotes the maximum integer smaller than a and $m=\lfloor \Gamma/T\rfloor$ represents the number of segments after separating signal $y(t)$ by T-time length. Equation (12) illustrates that signal $f_T(t)$ in time interval $[0,T)$ is the average of several parts of $y(t)$ separated by T, and the signal $f_T(t)$ is a periodic function. This step is crucial to the proposed method because the rule of segmenting and averaging will substantially reduce the magnitude of noise, making the proposed method performs more excellently than conventional methods in highly noisy environments.

Based on the new signal $f_T(t)$, a correlation function $\mathcal{C}(T)$ between $y(t)$ and $f_T(t)$ is then defined as

$$\mathcal{C}(T) = \frac{1}{mT}\int_0^{mT} y(t)\,f_T(t)\,dt. \tag{13}$$

We call $\mathcal{C}(T)$ the correlation function because it resembles the dot product of $y(t)$ and $f_T(t)$. Using the definition of $f_T(t)$ in (12), $\mathcal{C}(T)$ can also be written as

$$\begin{aligned}\mathcal{C}(T) &= \frac{1}{mT}\sum_{i=0}^{m-1}\int_0^T y(t+iT)\,f_T(t)\,dt \\ &= \frac{1}{m^2T}\int_0^T\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} y(t+iT)\,y(t+jT)\,dt \\ &= \frac{1}{T}\int_0^T f_T^2(t)\,dt,\end{aligned} \tag{14}$$

which gives an interesting result $\mathcal{C}(T)=\int_0^T f_T^2(t)/T dt = \mathcal{P}_{f_T}(T)$, i.e., $\mathcal{C}(T)$ is also the average power of the transformed signal $f_T(t)$. This provides an alternative option to calculate the correlation between $y(t)$ and $f_T(t)$, and thus we can also write the correlation function as

$$\begin{aligned}\mathcal{C}(T) &= \frac{1}{mT}\int_0^{mT} y(t)\,f_T(t)\,dt \\ &= \frac{1}{T}\int_0^T f_T^2(t)\,dt = \frac{1}{mT}\int_0^{mT} f_T^2(t)\,dt.\end{aligned} \tag{15}$$

By plugging (12) into (13) and using the results of the lemma in the appendix, $\mathcal{C}(T)$ can be written as

$$\begin{aligned}\mathcal{C}(T) &= \frac{1}{mT}\int_0^{mT} y(t)\,f_T(t)\,dt \\ &= \frac{1}{m^2T}\int_0^T\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} y(t+iT)\,y(t+jT)\,dt \\ &= \frac{1}{m^2T}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\int_0^T x(t+iT)\,x(t+jT)\,dt \\ &\quad + 2\frac{1}{m^2T}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\int_0^T x(t+iT)\,\epsilon(t+jT)\,dt \\ &\quad + \frac{1}{m^2T}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\int_0^T \epsilon(t+iT)\,\epsilon(t+jT)\,dt \\ &= \frac{1}{m^2T}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\int_0^T x(t+iT)\,x(t+jT)\,dt + 0 \\ &\quad + \frac{1}{m^2}\sum_{i=0}^{m-1}\frac{1}{T}\int_0^T \epsilon^2(t+iT)\,dt \\ &= \frac{1}{m^2T}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\int_0^T x(t+iT)\,x(t+jT)\,dt + \frac{m\sigma^2}{m^2} \\ &= \mathcal{S}(T)+\frac{\sigma^2}{m},\end{aligned} \tag{16}$$

where

$$\mathcal{S}(T)=\frac{1}{m^2T}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\int_0^T x(t+iT)\,x(t+jT)\,dt. \quad (17)$$

It is significant to see that there is an inequality for $\mathcal{S}(T)$, as shown in the theorem below, which is helpful to detect the unknown period T_0 of $x(t)$.

Theorem 1: Assume $\mathcal{P}_x(T)$ exists for all T, and $x(t)$ (not a constant function) is a periodic function with period T_0, then $\mathcal{S}(T)\le\mathcal{P}_x(mT)$ and the equation holds if and only if $T=kT_0$, $k=1,2,\cdots\lfloor\Gamma/T_0\rfloor$.

Proof: Recall the definition of the average power of $x(t)$, we can obtain $\mathcal{P}_x(mT)=\int_0^{mT}x^2(t)/(mT)\,dt$, and by reordering the summation and integration of $\mathcal{S}(T)$, clearly one can get

$$\begin{aligned}\mathcal{S}(T)&=\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\frac{1}{m^2T}\int_0^T x(t+iT)\,x(t+jT)\,dt\\&=\frac{1}{m^2T}\sum_{r=0}^{m-1}\int_0^T\left(u(t)+v(t)\right)dt,\end{aligned}$$

where

$$\begin{cases}u(t)=\sum\limits_{s=0}^{m-r-1}x(t+sT)\,x(t+rT+sT)\\ v(t)=\sum\limits_{s=0}^{r-1}x(t+mT-rT+sT)\,x(t+sT)\end{cases}.$$

By applying the Cauchy-Schwarz inequality, we have

$$\left(u(t)+v(t)\right)^2\le p(t)\,q(t),$$

where

$$\begin{cases}p(t)=\sum\limits_{s=0}^{m-r-1}x^2(t+sT)+\sum\limits_{s=0}^{r-1}x^2(t+mT-rT+sT)\\ q(t)=\sum\limits_{s=0}^{m-r-1}x^2(t+rT+sT)+\sum\limits_{s=0}^{r-1}x^2(t+sT)\end{cases}.$$

By reordering the summation in $p(t)$ and $q(t)$, we have

$$p(t)=q(t)=\sum_{s=0}^{m-1}x^2(t+sT).$$

This means that

$$|u(t)+v(t)|\le\sum_{s=0}^{m-1}x^2(t+sT).$$

Therefore we can obtain the following inequality

$$\begin{aligned}\mathcal{S}(T)&=\frac{1}{m^2T}\sum_{r=0}^{m-1}\int_0^T\left(u(t)+v(t)\right)dt\\&\le\frac{1}{m^2T}\sum_{r=0}^{m-1}\left(\sum_{s=0}^{m-1}\int_0^T x^2(t+sT)\,dt\right)\\&=\frac{1}{m^2T}\sum_{r=0}^{m-1}\left(\int_0^{mT}x^2(t)\,dt\right)\\&=\mathcal{P}_x(mT).\end{aligned}$$

Clearly, the equality holds if and only if $T=kT_0$. When $T\ne kT_0$, since $x(t)$ is not a constant function, it is impossible that $x(t+jT)=x(t+iT+jT)$ when $0\le j<m-i$ and $x(t+mT-iT+jT)=x(t+jT)$ when $0\le j<i$, and thus the equality cannot hold. This completes the proof. ■

The theorem shows that $\mathcal{S}(T)$ is approximately a periodic function and reaches maximum when $T=kT_0$, and thus offers a powerful tool for determining the optimal period of $y(t)$. That is to say, the correlation function $\mathcal{C}(T)$ displays prominent peaks at the true period and its multiple times, and thus the optimal period can be detected according to those prominent peaks. Since $\mathcal{P}_x(mT)\approx\mathcal{P}_x(\Gamma)\approx\mathcal{P}_x(\Gamma-\tau)$ and $\mathcal{R}_{yy}(\tau)\le\mathcal{P}_x(\Gamma-\tau)$, this suggests that $\mathcal{S}(T)=\mathcal{C}(T)-\sigma^2/m$ and $\mathcal{R}_{yy}(\tau)$ have similar expected values. In the next section, we will show that the proposed detector has a much smaller variance than that of the AUTOC method under strong white noise.

IV. Performance Analysis of the NRC Method

Note that when m (number of separated segments) is large enough, σ^2/m in (16) converges to zero and hence can be omitted, i.e., we can approximately write $\mathcal{S}(T)\approx\mathcal{C}(T)$. However, since the length of signal is limited in practice and the main aim of this paper is to deal with noise with large variance, it is crucial to estimate the unknown σ^2 to attenuate the effect arising from the noise. Furthermore, estimation of σ^2 is of significance for other potential applications. Therefore, we first estimate the unknown σ^2, then develop a robust function $\mathcal{Q}(T)$ for the period detection, and finally conduct the theoretical analysis on the variance of the NRC method in finite cases.

A. Estimation of σ^2

Since we know nothing about the signal or its period, it is difficult to estimate σ^2 based only on the signal itself. The method proposed here attempts to provide a solution to this problem. We define a new function $\mathcal{V}(T)$ to measure the difference between the measured signal $y(t)$ and the constructed signal $f_T(t)$

$$\mathcal{V}(T)=\frac{1}{mT}\int_0^{mT}\left(y(t)-f_T(t)\right)^2dt. \quad (18)$$

By using the equations (13) and (14), we can expand the right part in (18) as

$$\begin{aligned}\mathcal{V}(T)&=\frac{1}{mT}\int_0^{mT}\left(y^2(t)-2y(t)\,f_T(t)+f_T^2(t)\right)dt\\&=\frac{1}{mT}\int_0^{mT}\left(y^2(t)-f_T^2(t)\right)dt. \quad (19)\end{aligned}$$

Using the results in (15) and (16), we can get

$$\begin{aligned}\mathcal{V}(T)&=\frac{1}{mT}\int_0^{mT}y^2(t)\,dt-\mathcal{C}(T)\\&=\sigma^2-\frac{\sigma^2}{m}+\mathcal{P}_x(mT)-\mathcal{S}(T)\ge\frac{m-1}{m}\sigma^2. \quad (20)\end{aligned}$$

Equation (20) provides a way to estimate the noise variance σ^2, and thus the estimator of σ^2 is given by $\hat{\sigma}^2=\inf_T\{m\mathcal{V}(T)/(m-1)\}$.

B. A Robust Correlation Function $\mathcal{Q}(T)$

Recall that $\mathcal{C}(T) = \mathcal{S}(T) + \sigma^2/m$ depends on the unknown σ^2. This problem can now be solved by using $\mathcal{C}(T) - \hat{\sigma}^2/m = \mathcal{S}(T) + \left(\sigma^2 - \hat{\sigma}^2\right)/m$ as a new detector. However, this straightforward approach does not fully address the problem because the estimation may not be very accurate. Hence, we further propose a robust correlation function $\mathcal{Q}(T)$. We have obtained from Theorem 1 that $\mathcal{C}(T) - \sigma^2/m$ reaches the maximum when $T = kT_0$. We can also conclude from (20) that $\mathcal{V}(T)/(m-1) - \sigma^2/m$ reaches the minimum when $T = kT_0$. Thus based on these properties, we define a new function

$$\mathcal{Q}(T) = \mathcal{C}(T) - \frac{\mathcal{V}(T)}{m-1}. \tag{21}$$

By plugging (15) and (19) into (21), the function $\mathcal{Q}(T)$ can also be written as

$$\begin{aligned}\mathcal{Q}(T) &= \frac{m\frac{1}{T}\int_0^T f_T^2(t)\,dt - \frac{1}{mT}\int_0^{mT} y^2(t)\,dt}{m-1} \\ &= \frac{\int_0^T \left(\sum_{i=0}^{m-1} y(t+iT)\right)^2 - \sum_{i=0}^{m-1} y^2(t+iT)\,dt}{m(m-1)T}.\end{aligned} \tag{22}$$

Notice that the numerator in (22) can be simplified, leading to the following equation

$$\mathcal{Q}(T) = \frac{\int_0^T \sum_{i\neq j} y(t+iT)\,y(t+jT)\,dt}{m(m-1)T}. \tag{23}$$

Using the relationship between $\mathcal{C}(T)$ and $\mathcal{S}(T)$, we also get

$$\begin{aligned}\mathcal{Q}(T) &= \mathcal{C}(T) - \frac{\mathcal{V}(T)}{m-1} \\ &= \frac{m\mathcal{S}(T) - \mathcal{P}_x(mT)}{m-1} \le \mathcal{P}_x(mT).\end{aligned} \tag{24}$$

From Theorem 1, it can be seen that $\mathcal{Q}(T)$ reaches the maximum $\mathcal{P}_x(mT)$ when $T = kT_0$, without depending on σ^2. In addition, (23) shows that $\mathcal{Q}(T)$ is calculated using the information of the measured signal $y(t)$. Therefore, the proposed NRC method can be applied without any prior knowledge.

C. Variance of the NRC Method

In this part, we calculate the variance of the NRC method with different correlation functions in a finite case. Here, we define some notations for convenience. Similar to the notations used in (9), we denote vectors $\mathbf{x}_i$ and $\boldsymbol{\epsilon}_i$ respectively, as the N discrete values of $x(t)$ and $\epsilon(t)$ in the range $[iT, T+iT)$, and let the vector $\mathbf{x}$ be the discrete values of $x(t)$ in the whole range $[0, \Gamma)$. More detailed proofs on the variance of the NRC method are given in the supplement material of this paper.

1) Variance of the NRC Method with $\mathcal{C}(T)$: The correlation function $\mathcal{C}(T)$ can be expanded in terms of the periodic signal $x(t)$ and the noise signal $\epsilon(t)$ as

$$\begin{aligned}\mathcal{C}(T) &= \int_0^T f_T^2(t)/T\,dt \\ &= \frac{1}{T}\int_0^T \left(\frac{1}{m}\sum_{i=0}^{m-1} y(t+iT)\right)^2 dt \\ &= \int_0^T \frac{\left(\sum_{i=0}^{m-1} x(t+iT) + \sum_{i=0}^{m-1}\epsilon(t+iT)\right)^2}{m^2 T}dt,\end{aligned} \tag{25}$$

and thus the discrete version of $\mathcal{C}(T)$ is

$$\mathcal{C}_d = \frac{\left(\sum_{i=0}^{m-1}\mathbf{x}_i^T + \sum_{i=0}^{m-1}\boldsymbol{\epsilon}_i^T\right)\left(\sum_{i=0}^{m-1}\mathbf{x}_i + \sum_{i=0}^{m-1}\boldsymbol{\epsilon}_i\right)}{m^2 N}. \tag{26}$$

The variance of $\mathcal{C}_d$ is then given by

$$\mathrm{Var}(\mathcal{C}_d) = \frac{2\sigma^2}{mN}\left(\frac{2\left(\sum_{i=0}^{m-1}\mathbf{x}_i^T\right)\left(\sum_{i=0}^{m-1}\mathbf{x}_i\right)}{m^2 N} + \frac{\sigma^2}{m}\right). \tag{27}$$

Note that $\left(\sum_{i=0}^{m-1}\mathbf{x}_i^T\right)\left(\sum_{i=0}^{m-1}\mathbf{x}_i\right)/\left(m^2 N\right)$ is the discrete value of $\mathcal{S}(T)$, denoted by $\mathcal{S}_d$, and thus from Theorem 1 we have $\mathcal{S}_d \le \frac{\mathbf{x}^T\mathbf{x}}{mN}$, which implies

$$\begin{aligned}\mathrm{Var}(\mathcal{C}_d) &= \frac{2\sigma^2}{mN}\left(2\mathcal{S}_d + \frac{\sigma^2}{m}\right) \\ &\le \frac{2\sigma^2}{mN}\left(\frac{2\mathbf{x}^T\mathbf{x}}{mN} + \frac{\sigma^2}{m}\right) \\ &\approx \frac{2\sigma^2}{mN}\left(2\mathcal{P}_x(\Gamma) + \frac{\sigma^2}{m}\right).\end{aligned} \tag{28}$$

Unlike the variance of $\mathcal{R}_d$ in (9), the variance of noise σ^2 no longer dominates the variance of $\mathcal{C}_d$, as shown in (28), even under strong noise, because σ^2/m is scaled down quickly by m. Also note that $\mathcal{C}(T)$ and $\mathcal{R}_{yy}(\tau)$ have similar expected values and reach approximately the maximum value $\mathcal{P}_x(\Gamma)$. By comparing (9) and (28), therefore, we can conclude that $\mathcal{C}(T)$ is more sensitive than $\mathcal{R}_{yy}(\tau)$, especially in the cases with strong noise. This means that the proposed NRC method with $\mathcal{C}(T)$ performs better than the AUTOC method under strong white noise in the finite case.

2) Variance of the NRC Method with $\mathcal{Q}(T)$: Similar to the variance analysis of $\mathcal{C}(T)$, the robust function $\mathcal{Q}(T)$ in (23) can also be written in terms of the original periodic signal $x(t)$ and the noise signal $\epsilon(t)$ as

$$\begin{aligned}\mathcal{Q}(T) &= \frac{\int_0^T \sum_{i\neq j} y(t+iT)\,y(t+jT)\,dt}{m(m-1)T} \\ &= \frac{\int_0^T \sum_{i\neq j}\left(x(t+iT)+\epsilon(t+iT)\right)\left(x(t+jT)+\epsilon(t+jT)\right)dt}{m(m-1)T},\end{aligned} \tag{29}$$

and thus the discrete version of $\mathcal{Q}(T)$ is

$$\mathcal{Q}_d = \frac{\sum_{i\neq j}\left(\mathbf{x}_i^T + \boldsymbol{\epsilon}_i^T\right)\left(\mathbf{x}_j + \boldsymbol{\epsilon}_j\right)}{m(m-1)N}. \tag{30}$$

The variance of $\mathcal{Q}_d$ is then given by

$$\begin{aligned}\mathrm{Var}\left(\mathcal{Q}_d\right) &= 4\sigma^2\left(\frac{\sum_{j=0}^{m-1}\left(\sum_{i\neq j}\mathbf{x}_i^T\right)\left(\sum_{i\neq j}\mathbf{x}_i\right)}{m^2(m-1)^2N^2} + \frac{\sigma^2}{2m(m-1)N}\right)\\ &= \frac{2\sigma^2}{mN}\left(2\frac{m(m-2)\mathcal{S}_d + \frac{\mathbf{x}^T\mathbf{x}}{mN}}{(m-1)^2} + \frac{\sigma^2}{m-1}\right).\end{aligned} \tag{31}$$

Since $\mathcal{S}_d \leq \frac{\mathbf{x}^T\mathbf{x}}{mN}$, we have

$$\begin{aligned}\mathrm{Var}\left(\mathcal{Q}_d\right) &\leq \frac{2\sigma^2}{mN}\left(\frac{2\mathbf{x}^T\mathbf{x}}{mN} + \frac{\sigma^2}{m-1}\right)\\ &\approx \frac{2\sigma^2}{mN}\left(2\mathcal{P}_x(\Gamma) + \frac{\sigma^2}{m-1}\right).\end{aligned} \tag{32}$$

By comparing (27) and (31), the difference between $\mathrm{Var}\left(\mathcal{Q}_d\right)$ and $\mathrm{Var}\left(\mathcal{C}_d\right)$ is

$$\mathrm{Var}\left(\mathcal{Q}_d\right) - \mathrm{Var}\left(\mathcal{C}_d\right) = \frac{2\sigma^2}{mN}\left(2\frac{\frac{\mathbf{x}^T\mathbf{x}}{mN} - \mathcal{S}_d}{(m-1)^2} + \frac{\sigma^2}{(m-1)m}\right). \tag{33}$$

It can be seen that $\mathcal{Q}(T)$ has a variance that is extremely close to $\mathcal{C}(T)$ (although slightly higher) and has the same expected values. However, $\mathcal{Q}(T)$ does not depend on the unknown σ^2 and hence is more robust than $\mathcal{C}(T)$. Furthermore, it can be seen from (32) that $\mathcal{Q}(T)$, similar to $\mathcal{C}(T)$, is more sensitive than $\mathcal{R}_{yy}(\tau)$ under strong noise. Finally, it is clear that the average power $\mathcal{P}_x(\Gamma)$ is approximately a constant and σ^2 is fixed for a given signal. Therefore, more segments give smaller variances of $\mathcal{Q}(T)$ and $\mathcal{C}(T)$, leading to a better performance of the proposed method. Note that there are two ways to increase m: increasing Γ and decreasing T. In the remainder of the paper we use $\mathcal{Q}(T)$ as the period detector for proposed method.

Here, we use the simulated signal in (5) to show the performance of the proposed NRC method with the correlation function $\mathcal{Q}(T)$. The analysis results of the contaminated signal with a length of 20,000,000 and the signal with a length of 200,000 are shown in Fig. 4 and Fig. 5 respectively.

We can see from Fig. 4 and Fig. 5 that the period of the original signal is clearly observable. We can also conclude from Fig. 2 and Fig. 4 that both the AUTOC method and the proposed NRC method perform well when the length of signal is very large, while comparison of Fig. 3 and Fig. 5 shows that the NRC method outperforms the AUTOC method when the length of signal is small, which is consistent with our theoretical analysis. It can be also seen that the proposed method performs better in Fig. 4 than in Fig. 5 because a larger Γ is given in Fig. 4 than in Fig. 5. By comparing the prominent peaks in Fig. 5, it can be seen that the proposed method has a better performance by decreasing T, because more segments obtained as T decreases lead to a smaller variance, making the peaks more prominent.

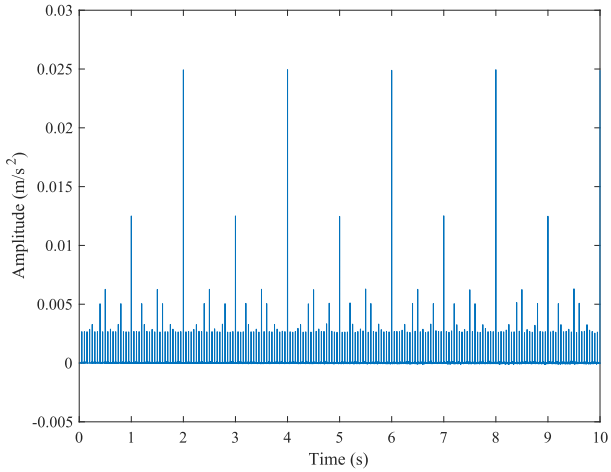


Fig. 4. $\mathcal{Q}(T)$ with 20,000,000 sample points.

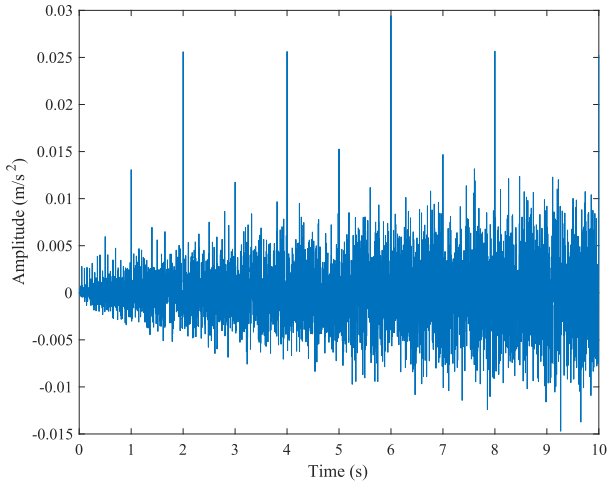


Fig. 5. $\mathcal{Q}(T)$ with 200,000 sample points.

V. Experimental Study

At a constant rotating speed, localized faults in rotating machines tend to result in periodic shocks and thus create periodic impulses in the vibration signal, so the periodic features can be readily detected using the proposed method. We apply our method to typical cases of rotating machine fault detection, namely, bearing fault detection and gearbox fault detection. The bearing fault experimental data contain mild noise and the gearbox fault experimental data contain strong noise. It should be noted that although the performance of the NRC method is demonstrated through the fault detection of mechanical vibration signals, the proposed method can also be applied to similar applications for period detection.

A. Bearing Fault Experiment

In a rolling element bearing, a defect will cause shocks when the rolling elements pass over the defect. If the defect is located

TABLE I
GEOMETRY AND WORKING PARAMETERS OF THE FAULTY BEARING

Type	6205-2RS JEM SKF
No. of rolling elements	9
Inside diameter (in)	0.9843
Outside diameter (in)	2.0472
Thickness (in)	0.5906
Ball diameter (in)	0.3126
Pitch diameter (in)	1.537
Sizes of the fault (D×W) (in)	0.011×0.007
Rotating speed (rpm)	1796
Fault characteristic period (ms)	9.3

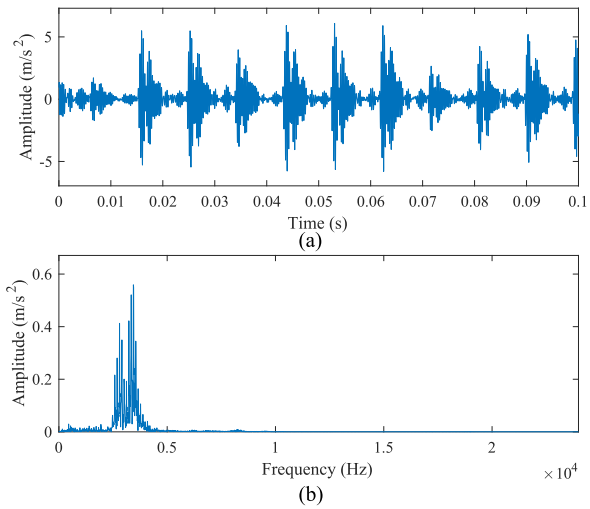


Fig. 6. Time domain and frequency domain of the vibration signal.

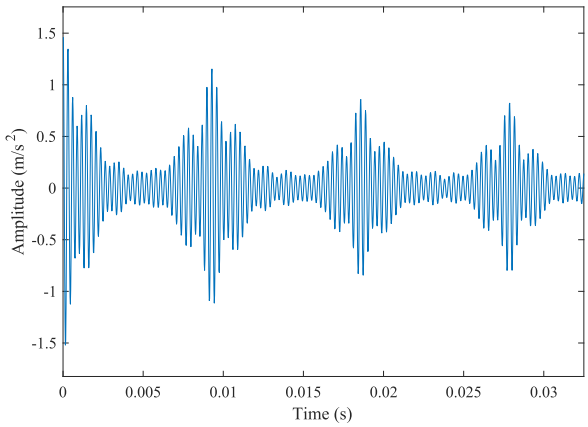


Fig. 7. Results of the AUTOC method for the vibration signal.

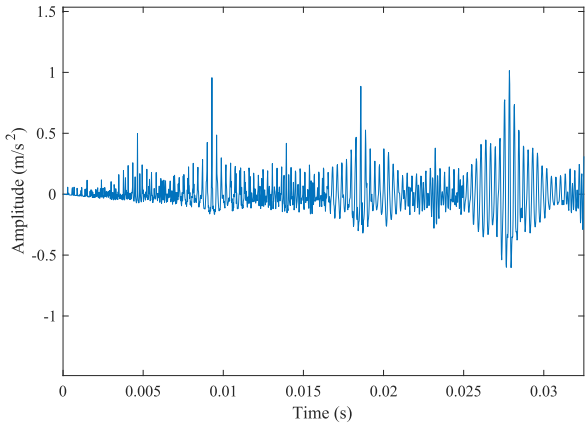


Fig. 8. Analysis results for the vibration signal by using the proposed NRC method with $\mathcal{Q}(T)$.

on the outer race, it will give rise to periodic and almost equally strong impulses. Thus, we use a sample of outer race fault data to illustrate the proposed NRC method. We analyze data from the Case Western Reserve University (CWRU) bearing data website [25]. The data were measured on a rotating machine test rig with a sampling frequency of 48 kHz. The details of the geometry and working parameters of the faulty bearing are listed in Table I, from which we can get the theoretical fault characteristic period ($T_0 = 9.3$ ms).

The time-domain and frequency-domain of the vibration signal are shown in Fig. 6 (a) and Fig. 6 (b) respectively. The time length of the signal is 0.1 s. It can be observed that fault transients are induced by the outer race fault. The AUTOC method is used to analyze the time-domain signal, and the results are shown in Fig. 7.

Fig. 7 shows that the autocorrelation function of the vibration signal is a periodic signal with a period of 9.3 ms, which is the same as the theoretical fault characteristic period shown in Table I.

The proposed NRC method with $\mathcal{Q}(T)$ is also tested on the same vibration signal to evaluate its period detection performance under mild noise and the results are shown in Fig. 8. The period of 9.3 ms can also be identified easily from Fig. 8, which is consistent with the theoretical value.

The experimental study using the CWRU bearing fault data shows that both the AUTOC method and the proposed NRC method perform well in dealing with weak background noise in a finite case.

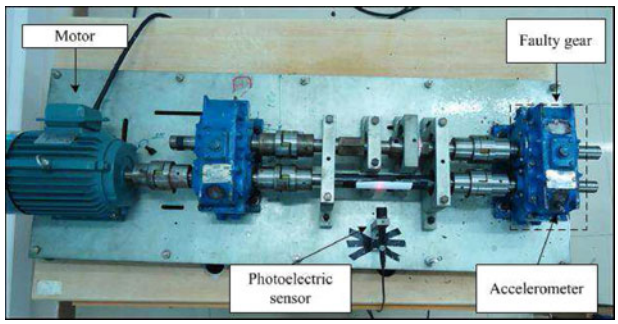


Fig. 9. Faulty gear test rig.

B. Gear Fault Experiment

In the second test, the vibration signal of a faulty driving gear is analyzed. Fig. 9 shows the test rig with a faulty gearbox. Its parameters are listed in Table II, from which we can get the meshing frequency ($f_m = 838.67$ Hz), the theoretical

TABLE II
WORKING PARAMETERS OF THE GEARBOX

Gear	Driving gear	Driven gear
No. of teeth	34	42
Rotating speed (rpm)	1480	1198
Rotating frequency (Hz)	24.67	19.97
Rotating period (ms)	40.5	50.1
Meshing frequency (Hz)	838.67	838.67

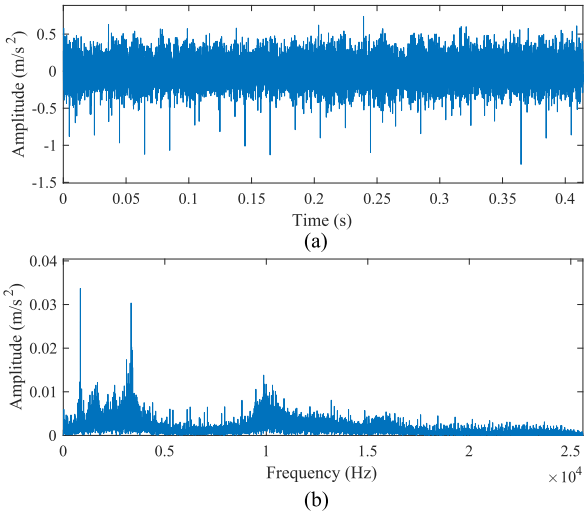


Fig. 10. Time domain and frequency domain of the gearbox vibration signal.

rotating period ($T_0 = 40.5$ ms) of the driving gear, also called fault characteristic period, and the theoretical rotating frequency (24.67 Hz), also called fault characteristic frequency. Therefore, the frequency domain of the measured fault signal should contain the meshing frequency and its harmonics, the rotating frequency and its harmonics, and the frequency of the background noise. The time-domain and frequency-domain of the vibration signal with time length 0.4141 s are shown in Fig. 10 (a) and Fig. 10 (b) respectively. The sampling frequency is 51.2 kHz. Fig. 10 (a) shows that the measured signal has no indications of an impulsive signature due to the strong background noise. It can be found from Fig. 10 (b) that the meshing frequency f_m is 840 Hz and its 4th harmonic is 3360 Hz. However, it's difficult to extract the rotating frequency and its harmonics. Therefore, it is challenging to detect the period from the measured signal. This test is designed to evaluate the performance of the proposed NRC method in dealing with strong background noise in a finite case.

Firstly, the AUTOC method is applied to the time-domain vibration signal, and the results are shown in Fig. 11. Unfortunately we cannot identify the rotating period from Fig. 11.

Then the proposed NRC method is used to analyze the vibration signal, and the results are shown in Fig. 12. It can be observed that the period of the vibration signal is 40.4 ms, which is very close to the theoretical value. This error may be caused by the fluctuation of the input rotating speed. The analysis of the experimental data shows that the proposed NRC method can

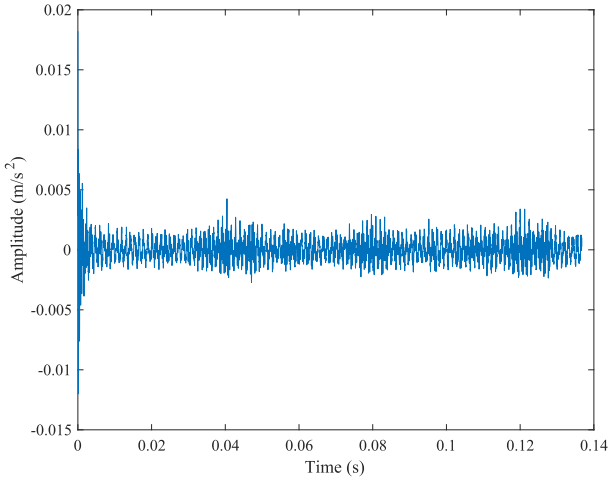


Fig. 11. Results of the AUTOC method for the gearbox vibration signal.

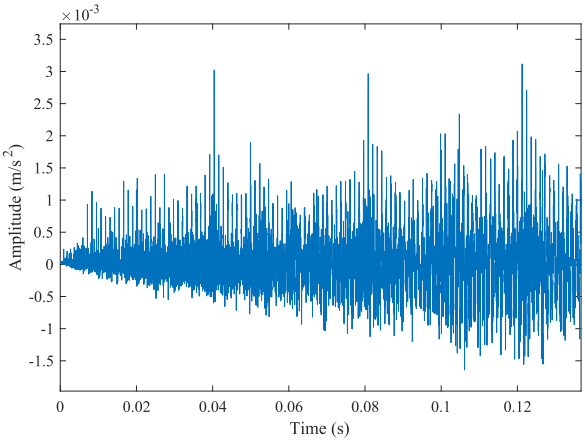


Fig. 12. Analysis results for the vibration signal by using the NRC method.

detect the period information from the measured signal when the background noise is strong in a finite case, outperforming the AUTOC method.

In summary, the two experiments show that the performance of the proposed NRC method is similar to that of the AUTOC method under the case with mild noise where both methods can easily detect the unknown period. However, the AUTOC method fails to detect the period from the signal contaminated by strong noise, whereas the period can be successfully detected by the proposed method. Therefore, we can conclude from the two experiments that the proposed method is more sensitive and robust in period detection than the AUTOC method, especially under strong white noise.

VI. CONCLUSION

In this paper, we present a novel NRC method to solve the hidden period detection problem under strong background noise. A novel periodic signal is constructed, based on which a correlation function is proposed. We show that this correlation function provides an effective way to find the hidden period of the measured signal because it reaches the local maximum at multiple times of the hidden period. This correlation function can not

only robustly suppress the random effects of strong noise without depending on any prior knowledge, but is also more sensitive in detecting the hidden period because, in contrast to the AUTOC method, the strong noise does not dominate the detector. Our theoretical analyses on the NRC and AUTOC methods confirm that the NRC method performs better in the scenario under strong noise with finite sample points. The effectiveness of the proposed NRC method in hidden period detection is verified by two experiments on the bearing fault feature detection and the gear fault feature detection.

The proposed method assumes that the background noise is white Gaussian and does not consider the signal containing multiple periodic components. Although our method can be applied to many applications, further works are needed to exploit its potentials on signals containing non-Gaussian noise or multiple periodic components.

APPENDIX

As already stated, the noise term is not Riemann integrable. Hence, before proving the lemma, we first define the integration on the signal functions. Since the noise signal $\epsilon(t)$ is not continuous, we treat it as a step function by letting $\epsilon(t') \equiv \epsilon(t)$ when $t' \sim [t, t+\triangle t)$ for a small $\triangle t$ in the integration. We assume that the integration interval is separated by the partition $t_0, t_1, t_2, \cdots, t_{n-1}$ and let $\triangle t_i = t_{i+1} - t_i$, then the integration on function $g(t)$ is defined as

$$\int_0^T g(t)dt = \lim_{n\to\infty} \sum_{i=0}^{n} g(t_i)\triangle t_i,$$

where $\sup_{i\le n}\{\triangle t_i\} \le cT/n$ (c is a constant). The assumption on $\triangle t_i$ is very weak and it does not affect the value of the integration. However, it is necessary here to guarantee that the functions that involve noise are integrable. We have the following properties for integration on $y(t)$, $x(t)$ and $\epsilon(t)$, as shown in the following lemma.

Lemma: Suppose a signal $x(t)$ is bounded and a white noise $\epsilon(t)$ independently follows $\mathbb{N}\left(0, \sigma^2\right)$ with a bounded variance σ^2. If $\int_0^T x^2(t)\,dt$ exists, then we have

$$\begin{cases} \frac{1}{T}\int_0^T x(t)\epsilon(t)\,dt & = 0 \\ \frac{1}{T}\int_0^T \epsilon^2(t)\,dt & = \sigma^2 \\ \frac{1}{T}\int_0^T \epsilon(t)\epsilon(t+T)\,dt & = 0 \\ \frac{1}{T}\int_0^T \left(x(t)+\epsilon(t)\right)^2 dt & = \sigma^2 + \frac{1}{T}\int_0^T x^2(t)\,dt \end{cases}.$$

Proof: 1) Let $Z_i = x(t_i)\epsilon(t_i)n\triangle t_i/T$, then

$$\begin{aligned} \frac{1}{T}\int_0^T x(t)\epsilon(t)\,dt &= \lim_{n\to\infty}\frac{1}{T}\sum_{i=1}^{n} x(t_i)\epsilon(t_i)\triangle t_i \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} x(t_i)\epsilon(t_i)n\triangle t_i/T \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} Z_i. \end{aligned}$$

As $\sup_{i\le n}\{\triangle t_i\} \le cT/n$ and $|x(t)| < M$, then $\mathrm{Var}(Z_i) < M^2c^2\sigma^2 < \infty$ and $\sum_{i=1}^{\infty}\mathrm{Var}(Z_i)/n^2 \le \sum_{i=1}^{\infty} M^2c^2\sigma^2/n^2 < \infty$, and then by Kolmogrove's strong law of large numbers, $\frac{1}{n}\sum_{i=1}^{n} Z_i \xrightarrow{a.s.} 0$. Therefore, almost surely we have

$$\frac{1}{T}\int_0^T x(t)\epsilon(t)\,dt = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} Z_i = 0.$$

2) Similarly, let $Z_i = \epsilon^2(t_i)n\triangle t_i/T$, then

$$\frac{1}{T}\int_0^T \epsilon^2(t)\,dt = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} Z_i.$$

As $\sup_{i\le n}\{\triangle t_i\} \le cT/n$ and $\epsilon^2(t_i) \sim \sigma^2\chi^2(1)$, then $\mathrm{Var}(Z_i) \le 2c^2\sigma^4 < \infty$ and $\sum_{i=1}^{\infty}\mathrm{Var}(Z_i)/n^2 < \infty$, and hence by Kolmogrove's strong law of large numbers,

$$\frac{1}{n}\sum_{i=1}^{n} Z_i - \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} Z_i - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Z_i] \xrightarrow{a.s.} 0.$$

3) Again, let $Z_i = \epsilon(t_i)\epsilon(t_i+T)n\triangle t_i/T$, which are obviously independent with each other, then

$$\frac{1}{T}\int_0^T \epsilon(t)\epsilon(t+T)\,dt = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} Z_i.$$

We can also check that Z_i has zero mean and finite variance, and hence $\sum_{i=1}^{\infty}\mathrm{Var}(Z_i)/n^2 < \infty$. Therefore, by Kolmogrove's strong law of large numbers,

$$\frac{1}{T}\int_0^T \epsilon(t)\epsilon(t+T)\,dt = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} Z_i \xrightarrow{a.s.} 0.$$

4) Finally, it can be seen that $\int_0^T \left(x(t)+\epsilon(t)\right)^2/T dt = \sigma^2 + \int_0^T x^2(t)/T dt$ almost surely. ■

ACKNOWLEDGMENT

The authors would like to thank Prof. Z. Zhu from Soochow University for supplying the experimental data on the faulty gearbox.

REFERENCES

[1] S. V. Tenneti and P. P. Vaidyanathan, "Nested periodic matrices and dictionaries: New signal representations for period estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3736–3750, Jul. 2015.

[2] B. Han, C. Kaiser, and K. Dostert, "A novel approach of canceling cyclostationary noise in low-voltage power line communications," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 734–739.

[3] G. Ren, S. Qiao, H. Zhao, C. Li, and Y. Hei, "Mitigation of periodic impulsive noise in OFDM-based power-line communications," *IEEE Trans. Power Del.*, vol. 28, no. 2, pp. 825–834, Apr. 2013.

[4] F. Sauvet *et al.*, "In-flight automatic detection of vigilance states using a single EEG channel," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 12, pp. 2840–2847, Dec. 2014.

[5] A. K. Ovacıklı, P. Pääjärvi, J. P. LeBlanc, and J. E. Carlson, "Recovering periodic impulsive signals through skewness maximization," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1586–1596, Mar. 2016.

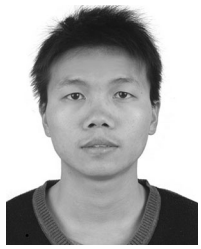
[6] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis-part II: Signals and signal processing methods," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6546–6557, Oct. 2015.

[7] S. V. Tenneti and P. P. Vaidyanathan, "A unified theory of union of subspaces representations for period estimation," *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5217–5231, Oct. 2016.

[8] S. V. Tenneti and P. P. Vaidyanathan, "Detecting tandem repeats in DNA using Ramanujan filter bank," in *Proc. Int. Symp. Circuits Syst.*, Montreal, QC, Canada, May 2016, pp. 21–24.
[9] H. P. Osborn *et al.*, "Single transit candidates from K2: Detection and period estimation," *Monthly Notices Roy. Astron. Soc.*, vol. 457, no. 3, pp. 2273–2286, Apr. 2016.
[10] C. I. Chen and Y. C. Chen, "Comparative study of harmonic and interharmonic estimation methods for stationary and time-varying signal," *IEEE Trans. Ind. Electron.*, vol. 61, no. 1, pp. 397–404, Jan. 2014.
[11] G. D. Bergland, "A guided tour of the fast Fourier transform," *IEEE Spectr.*, vol. 6, no. 7, pp. 41–52, Jul. 1969.
[12] Y. Yang, X. J. Dong, Z. K. Peng, W. M. Zhang, and G. Meng, "Vibration signal analysis using parameterized time– frequency method for features extraction of varying-speed rotary machinery," *J. Sound Vib.*, vol. 335, pp. 350–366, Jan. 2015.
[13] L. Rabiner, "On the use of autocorrelation analysis for pitch detection," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-25, no. 1, pp. 24–33, Feb. 1977.
[14] B. Gold and L. Rabiner, "Parallel processing techniques for estimating pitch periods of speech in the time domain," *J. Acoust. Soc. Amer.*, vol. 46, no. 2B, pp. 442–448, Aug. 1969.
[15] M. Jalil, F. A. Butt, and A. Malik, "Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signal," in *Proc. Int. Conf. Technol. Adv. Elect. Electron. Comput. Eng.*, May 2013, pp. 208–212.
[16] M. Elad, "Sparsity-seeking methods in signal processing," in *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer-Verlag, 2010, ch. 9, pp. 169–173.
[17] G. Cai, X. Chen, and Z. He, "Sparsity-enabled signal decomposition using tunable Q-factor wavelet transform for fault feature extraction of gearbox," *Mech. Syst. Signal Process.*, vol. 41, no. 1/2, pp. 34–53, Dec. 2013.
[18] W. Fan, G. Cai, Z. K. Zhu, C. Shen, W. Huang, and L. Shang, "Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction," *Mech. Syst. Signal Process.*, vol. 56– 57, pp. 230–245, May 2015.
[19] F. Peng, D. Yu, and J. Luo, "Sparse signal decomposition method based on multi-scale chirplet and its application to the fault diagnosis of gearboxes," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 549–557, Feb. 2011.
[20] J. C. Brown and B. Zhang, "Musical frequency tracking using the methods of conventional and 'narrowed' autocorrelation," *J. Acoust. Soc. Amer.*, vol. 89, no. 5, pp. 2346–2354, May 1991.
[21] L. Tan and M. Karnjanadecha, "Pitch detection algorithm: Autocorrelation method and AMDF," in *Proc. Int. Symp. Commun. Inf. Technol.*, Sep. 2003, vol. 2, pp. 551–556.
[22] T. Shimamura and H. Kobayashi, "Weighted autocorrelation for pitch extraction of noisy speech," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 7, pp. 727–730, Oct. 2001.
[23] W. Su, F. Wang, H. Zhu, Z. Zhang, and Z. Guo, "Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement," *Mech. Syst. Signal Process.*, vol. 24, no. 5, pp. 1458–1472, Jul. 2010.
[24] J. Rafiee and P. W. Tse, "Use of autocorrelation of wavelet coefficients for fault diagnosis," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1554–1572, Jul. 2009.
[25] "48k Drive end bearing fault data|Bearing data enter." [Online]. Available: http://csegroups.case.edu/bearingdatacenter/pages/48k-drive-end-bearing-fault-data. Accessed on: Jun. 12, 2017.

Wei Fan received the B.S. degree in electric engineering and automation and the M.S. degree in measurement technology and instrument from Soochow University, Suzhou, China, in 2012 and 2015, respectively. She is currently working toward the Ph.D. degree in the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong. She is currently an Academic Staff with the School of Mechanical Engineering, Zhenjiang, Jiangsu, China.

Her research interests include signal processing and machinery fault diagnosis.



Yongxiang Li received the B.S. degree in mathematics and applied mathematics from Shantou University, Shantou, China, in 2010. He is currently working toward the Ph.D. degree in the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong.

His research interests include mathematical statistics, quality engineering, and signal processing.

Kwok Leung Tsui received the B.S. degree in chemistry and the M.S. degree in mathematics from Chinese University of Hong Kong, Sha Tin, Hong Kong, and the Ph.D. degree in statistics from the University of Wisconsin-Madison, Madison, WI, USA. He is currently a Professor with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong. His research interests include quality engineering, process control, and surveillance. He is a fellow of ASQ and ISEAM.

Qiang Zhou (M'14) received the B.S. degree in automotive engineering and the M.S. degree in mechanical engineering and statistics from Tsinghua University, Beijing, China, in 2005, 2007, and 2010, respectively, and the Ph.D. degree in industrial engineering from the University of Wisconsin-Madison, Madison, WI, USA, in 2011.

He is currently an Assistant Professor with the Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ, USA. His research interests include the modeling, monitoring, and analysis of complex engineering systems for the purposes of quality control and productivity improvement. He is a member of INFORMS, IIE, and ASQ.