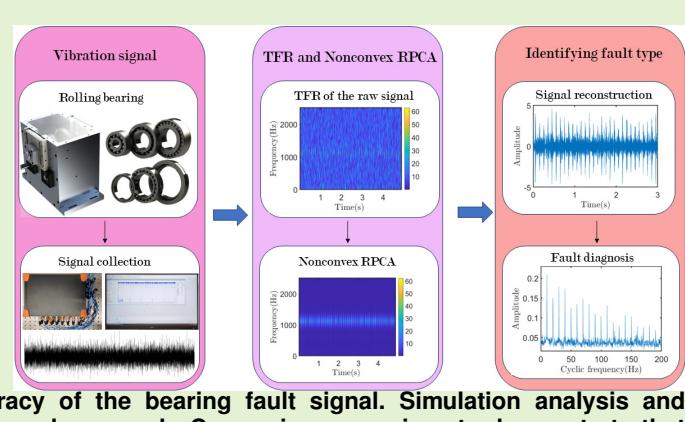


Bearing Fault Diagnosis via Robust PCA with Nonconvex Rank Approximation

Cai Li, Penghong Lu, Guangming Dong and Gang Chen, *Member, IEEE*

Abstract— Feature extraction is an essential part of bearing fault diagnosis. Robust principal component analysis (RPCA) provides a general technique for extracting fault features. However, Although the convex relaxed RPCA is convex and easy to optimize, the global optimal solution obtained may deviate significantly from reality. To address this issue, we propose a nonconvex RPCA method for bearing fault diagnosis, utilizing the γ norm penalty. Leveraging the unitary invariance property of the γ norm and the Moreau-Yosida operator, we transform the original equation into a singular value optimization problem using a difference of convex programming to obtain a feasible solution. The most important advantage of Nonconvex RPCA over conventional filtering methods is that it can improve fault feature extraction while decreasing noise interference, which results in substantially enhanced estimation accuracy of the bearing fault signal. Simulation analysis and experimental results confirm the effectiveness of the developed approach. Comparison experiments demonstrate that nonconvex RPCA provides more accurate extraction results compared to L1-norm regularization, Variational Mode Decomposition, and Feature Mode Decomposition.

Index Terms— Bearing fault diagnosis, Robust principal component analysis, γ norm.



I. INTRODUCTION

ROLLING bearings play a pivotal role in the operation of rotating machinery, indispensable across a wide array of modern equipment. Their efficient functioning is paramount for ensuring machine reliability [1]–[3]. Nonetheless, these components frequently face harsh conditions, curtailing their service life and potentially leading to substantial financial repercussions and safety hazards. As such, refining fault diagnosis methods is critical for guaranteeing bearing dependability [4]–[6].

The manifestation of bearing failures often comes in the form of periodic impulses, distinguished by specific fault characteristic frequencies (FCFs) linked to different fault types. The precise and efficient extraction of these signals from vibration data stands as a significant challenge within bearing fault diagnosis [7]. To confront this issue, recent decades

This work is partially supported by the National Natural Science Foundation of China (Grant No.52305105), the Basic and Applied Basic Research Foundation of Guangdong Province (Grant No.2022A1515240027 and No.2023A1515010812), and the Basic and Applied Basic Research Foundation of Guangzhou (Grant No.2023A04J1582) (*corresponding author: Gang Chen*)

Cai Li, Penghong Lu and Gang Chen are with Shien-Ming Wu School of Intelligent Engineering, South China University Of Technology, Guangzhou, 511442, China (e-mail: 202220159314@mail.scut.edu.cn; luhp1356023588@126.com; gangchen@scut.edu.cn).

Guangming Dong is with the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China.(e-mail: gmdong@sjtu.edu.cn)

have witnessed the emergence of numerous signal processing algorithms designed to enhance the extraction of these impulse features, including Wavelet Analysis [8], Spectral Kurtosis [9], Time-frequency representation [10], Variational Mode Decomposition (VMD) [11], Empirical Mode Decomposition (EMD) manifold [12] and Feature Mode Decomposition (FMD) [13]. Each technique offers unique contributions, significantly improving the accuracy and efficiency of bearing fault diagnosis.

Transitioning time-domain signals to the time-frequency domain, especially via Short-Time Fourier Transform (STFT) [14] and Continuous Wavelet Transform [15], facilitates the generation of time-frequency representations (TFRs). Nonetheless, these representations are subject to the Heisenberg uncertainty principle, which restricts their resolution. To mitigate this, advancements like the Synchrosqueezing Transform (SST) [16] and its variants [17], [18] have substantially refined the granularity of time-frequency distributions, though noise remains a challenge.

Addressing noise separation effectively in TFRs necessitates an in-depth understanding of both signal and noise characteristics. Utilizing hidden Markov models, Ge et al. [19] demonstrated the detection of shock components in bearing signals, capitalizing on signal sparsity. Additional strategies have been deployed to augment fault detection capabilities [20], [21], including a technique that leverages time-frequency sparsity to extract periodic impulse features more distinctly [22]. Yet, the application of sparse time-frequency representation (STFR)

via atomic decomposition, despite its potential, encounters difficulties in high-noise environments, raising questions about its reliability in real-world settings [23].

In practical engineering environments, excessive background noise significantly impairs the effectiveness of failure feature extraction. Traditional cyclic spectrum analysis and sparsity-based techniques, as previously mentioned, do not comprehensively account for noise. To address this, low-rank decomposition technology emerges as a potent solution, demonstrating remarkable efficacy even in the presence of substantial background noise. The RPCA method, introduced in [24], exemplifies this approach by decomposing a measured signal into sparse and low-rank components. This decomposition is aimed at uncovering the fundamental pattern inherent in the signal. Subsequently, this technique has been widely and successfully applied in fields such as machine learning [25] and acoustic signal processing [26].

The integration of low-rank methodologies into the diagnostics of rotating machinery faults has seen a notable upsurge. The pioneering work by Yu et al. [27], which underscored the sparsity property of the time-frequency matrix in bearing fault signals, has paved the way for advanced low-rank and sparse decomposition techniques in fault diagnosis. These innovative approaches have led to the development of sophisticated detection tools, capable of discerning subtle transient signals [28], [29]. Further contributions, such as the Periodical Sparse Low-Rank method [30] and adaptive window selection strategies for two-dimensional matrix transformations [31], have expanded the toolkit for accurate anomaly detection.

However, the application of RPCA transforms the original non-convex problem into a convex one. While this convex-relaxed RPCA offers ease of optimization, it faces two primary limitations in practical scenarios. Firstly, the presence of intense noise can disrupt the coherence of time-frequency features, potentially leading to a global optimal solution that significantly diverges from the actual underlying reality. Secondly, the convex relaxation approach uniformly reduces all singular values. The nuclear norm, representing the L1-norm of the singular values of the feature matrix, fails to adequately penalize large singular values, leading to potential inaccuracies in fault characterization. This observation underscores the need for further refinement in RPCA methodologies to address these challenges effectively.

To address the shortcomings of convex relaxation techniques in RPCA for bearing fault diagnosis, this study investigates non-convex penalties, recognized for their superior performance. The generalized minimax-concave (GMC) penalty, introduced by Ivan Selesnick [32], has become prominent in fault diagnosis. Wang's seminal work [33] on non-convex sparse representation using the GMC penalty constituted a breakthrough in this area. Additionally, Huang et al. [34] extended these efforts by applying the GMC penalty in a multi-source sparse representation framework, successfully enhancing gearbox diagnosis.

Beyond the GMC penalty, the introduction of the generalized logarithm (G-log) penalty [35] has further refined sparse representation techniques for signal processing. Despite its non-convex nature, the G-log penalty preserves the con-

vexity of the cost function through careful parameterization, facilitating the use of convex optimization algorithms for sparse solution acquisition. Yet, the application of non-convex penalties to low-rank time-frequency fault diagnosis remains an underexplored avenue.

To fill this gap, our paper presents a novel γ -function penalty aimed at improving the RPCA's performance in bearing fault diagnosis. Utilizing the γ norm's unitary invariant property and the Moreau-Yosida operator, we reformulate the traditional equation into an eigenvalue optimization problem, applying difference of convex programming for optimal solution extraction. The effectiveness of the γ non-convex optimization method in signal feature extraction is validated through simulations and real-world experiments, highlighting its utility in fault diagnosis.

The paper is structured as follows: Section 2 delves into RPCA's theoretical background within signal processing. Section 3 introduces the γ -norm function penalty, outlining its mathematical formulation, derivation process, and implementation conditions. Section 4 validates the proposed method through simulations, while Section 5 demonstrates its applicability with experimental data, emphasizing its practical significance. Finally, Section 6 summarizes the study's key contributions and their implications for signal processing and fault diagnosis.

II. RELATED WORK

This section reviews the signal model theories about bearing signals. We utilize the STFT to transition the signal model from the time domain to the time-frequency domain. Subsequently, we discuss the concepts of low-rank decomposition and outline the method for fault feature extraction using RPCA.

A. Signal Model

The measured signal from the faulty bearing operation primarily consists of the transient pulse component and the background noise induced by the fault shock. Thus, the signal can be described as:

$$z(t) = x(t) + n(t) \quad (1)$$

Where $z(t)$ denotes the measured signal, $x(t)$ is the transient impulse components of the measured signal from the faulty bearing, $n(t)$ stand for background noise during operation. When the fault bearing is working under time-invariant conditions, the transient impulse components can be presented as the model:

$$x(t) = \sum_i A_i s(t - iT - \tau_i) \quad (2)$$

$$s(t) = e^{-\beta t} \sin(2\pi ft) \quad (3)$$

Where the A_i represents the amplitude of the i th transient impulse resulting from the impact of fault bearing, $s(t)$ stands for a single impulse, T stands for the time duration of a transient impulse. The slippage τ_i denotes the fluctuations in bearing operation under normal conditions. The resonance

damping attenuation factor is denoted by β , and the natural frequency related to the rotor system is represented by f .

Generally, the transient impulses are weakened and covered by complex background noises. The challenge in bearing fault diagnosis is to effectively extract the fault features from the overwhelming background noise. In this research, we assume, as with other techniques, that the noise is Gaussian.

B. STFT

Vibration signals are transformed into the time-frequency domain using the STFT, as described by the equation:

$$Z(t, f) = \int z(\tau)g^*(t - \tau)\exp(-j(\tau - t))d\tau \quad (4)$$

Here, t and f represent time and frequency variables, respectively. The term g denotes the STFT window function, and $*$ signifies the complex conjugate operator.

Given STFT's linear mapping properties, the time-frequency model of the signal is represented by:

$$Z = X + N \quad (5)$$

In this model, the matrix $Z \subseteq \mathbb{C}^{N \times M}$ is the time-frequency representation (TFR) of the measured signals $z(t)$, $X \subseteq \mathbb{C}^{N \times M}$ represents the TFR of transient impulse components $x(t)$, and $N \subseteq \mathbb{C}^{N \times M}$ corresponds to the TFR of noise $n(t)$. The primary objective in signal processing is to distinctly separate x and n to maximize the signal-to-noise ratio (SNR) for accurate fault diagnosis.

C. Robust Principal Component Analysis

Owing to the low-rank property of the TFR of transient impulse components $x(t)$, RPCA is an effective technique for sparse fault feature detection. The measured bearing signals include transient signals and additive noise. The low-rank property of the feature matrix and the sparsity of noise in TFR are revealed [27]. After the STFT constructed the TFR, the primary spectrum Z was obtained for further analysis. The low-rank property of X and the sparsity of N are demonstrated and proved. The goal of the RPCA method is to extract matrices of minimum rank X and sparsest N . Optimization problems for the goal can be described as:

$$\min \text{rank}(X) + u\|N\|_0 \quad s.t. Z = X + N \quad (6)$$

Where u stands for a weight factor that achieves an equilibrium between X and N . However, the (6) is a traditional NP-hard problem. A common approach involves transforming the nonconvex rank function and zero norm into the nuclear norm and L1-norm, respectively, resulting in a convex formulation:

$$\min \|X\|_* + u\|N\|_1 \quad s.t. Z = X + N \quad (7)$$

where the $\|X\|_*$ stands for the nuclear norm of feature TFR X and $\|N\|_1$ is the L1 norm of noise TFR matrix N . The convex optimization solution of (7) is shown as [22].

Although the convex relaxed RPCA exhibits convexity and ease of optimization, it faces two primary limitations. Firstly, in practical scenarios, the incoherence of time-frequency features may not be guaranteed due to the influence of intense

TABLE I
EXAMPLES OF PENALTY FUNCTIONS.

L1 norm	log	atan	γ -norm
$ x $	$\frac{1}{a}\log(1+a x)$	$\frac{2}{a\sqrt{3}}(\tan^{-1}(\frac{1+2a x }{\sqrt{3}}) - \frac{\pi}{6})$	$\frac{(1+\gamma)x}{\gamma+x}$

noise. In such cases, the obtained global optimal solution may deviate significantly from the true underlying reality. Secondly, convex relaxation uniformly decreases all singular values. The nuclear norm, which corresponds to the L1-norm of singular values, cannot penalize large singular values effectively. In other words, the nuclear norm inadequately penalizes large singular values. To address the issue that the truth value deviation of the convex optimization solution is too large, we will propose a non-convex optimization solution based on γ -norm in the next section. In the next section, we introduce γ -norm with more approximate zero norm and then design an iterative algorithm to get a feasible solution based on γ -norm.

III. THE PROPOSED METHOD

In this part, we present a novel γ -norm penalty and introduce an RPCA with Nonconvex Rank Approximation.

A. γ -Norm

Due to the limitation of the convex relaxing, we explore the general framework of RPCA in this context:

$$\min \|X\|_\gamma + u\|N\|_l \quad s.t. Z = X + N \quad (8)$$

Where $\|\cdot\|_\gamma$ donates the γ -norm at the first introduction in [36], and $\|\cdot\|_l$ stands for a proper norm of noise TFR. The definition of γ -norm is a mathematical representation as:

$$\|X\|_\gamma = \sum_i \frac{(1+\gamma)\sigma_i(X)}{\gamma + \sigma_i(X)} \quad (\gamma > 0) \quad (9)$$

Where $\sigma_i(X)$ stands for i th singular value of matrix X . It is worth noting that $\lim_{\gamma \rightarrow 0} \|X\|_\gamma = \text{rank}(X)$, $\lim_{\gamma \rightarrow \infty} \|X\|_\gamma = \|X\|_*$ and the γ -norm $\|X\|_\gamma$ aligns accurately with the true rank and is unitarily invariant, implying $\|X\|_\gamma = \|U\Lambda V\|_\gamma$ for any orthonormal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$. This is in contrast to other commonly used norms, as illustrated in Fig.1. Table I enumerates mathematical expressions for standard penalty functions. Notably, when singular values deviate from 1, a significant difference is observed in other norms, while the γ -norm ($\gamma = 0.01$ is used in our simulation and experiments) mirrors the true rank effectively. This indicates that the γ -norm mitigates the imbalance in penalizing different singular values seen in other norms. However, optimizing Equation (8) presents challenges due to its non-convex. An efficient algorithm to address this issue is introduced in the following subsection.

B. Non-convex Robust PCA(Nonconvex RPCA)

Inspired by the research [36], we remove the equality constraint and introduce a Lagrange multiplier Y . We construct

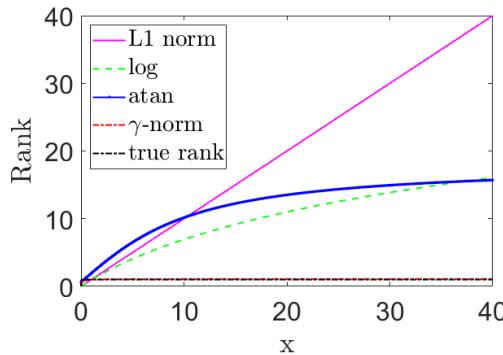


Fig. 1. The penalty of rank for common functions with respect to different singular values., The zero norm is true rank($\alpha = 0.01, \gamma = 0.01$)

the augmented Lagrange function:

$$\begin{aligned} \mathcal{L}(X, N, H, u) &= \|X\|_{\gamma} + \lambda\|N\|_l + \langle Y, Z - X - N \rangle \\ &\quad + \frac{u}{2}\|Z - X - N\|_F^2 \end{aligned} \quad (10)$$

Next, the problem of optimization can be divided into two similar subproblems. The subproblems can be updated through the iterative schemes outlined below:

$$X^{k+1} = \underset{X}{\operatorname{argmin}} \quad \mathcal{L}(X, N^k, Y^k, u^k) \quad (11)$$

$$N^{k+1} = \underset{N}{\operatorname{argmin}} \quad \mathcal{L}(X^{k+1}, N, Y^k, u^k) \quad (12)$$

The core principle guiding the resolution of these subproblems

Algorithm 1 Algorithm for Solving problem (8)

Input: Original TFR Z , Maximum Iterations m , parameters $\lambda > 0, u^0 > 0, \rho > 0$.

Output: Features TFR X , Noise TFR N .

- 1: Initialize X, N
- 2: **repeat**
- 3: Update X by (13).
- 4: Update N by (18).
- 5: Update Y by (19).
- 6: Update u by (20).
- 7: **until** converge or reach maximum iteration m

involves holding two variables constant while solving for one variable at a time. Initially, we tackle the sub-problem (11), a standard shrinkage problem, solvable through an explicit solution. At the $(t+1)$ th iteration, $X^{(t+1)}$ is derived by resolving the subproblem:

$$X^{t+1} = \underset{X}{\operatorname{argmin}} \|\mathcal{L}(X, N^t, Y^t, u^t)\|_F^2 \quad (13)$$

Due to the unitarily invariant of γ -norm and Frobenius norm, solving the sub-problem (13) is equal to addressing the following problem [36]:

$$\text{prox}_{f,u}(\sigma_A) := \underset{\sigma}{\operatorname{argmin}} f(\sigma) + \frac{u}{2}\|\sigma - \sigma_A\|_2^2 \quad (14)$$

Where the $\text{prox}_{f,u}$ stands for Moreau-Yosida operator, and $A = Z - N^t - \frac{Y^t}{u^t}$, $\|X\|_{\gamma} = f \circ \sigma(X)$ is unitarily invariant. In this research, as the objective function (14) comprises a combination of concave and convex functions, we employ Difference of Convex (DC) programming [37] to handle this structural process. The core concept of the DC algorithm involves decomposing a non-convex function into two separate convex functions. The concave term is linearized in each iteration. At the $k+1$ th iteration:

$$\sigma^{k+1} = \underset{\sigma}{\operatorname{argmin}} \langle \omega_k, \sigma \rangle + \frac{u^k}{2}\|\sigma - \sigma_A\|_2^2 \quad (15)$$

Where the gradient of f at σ^k is donated $\omega_k = \partial f(\sigma^k)$. This objective function is convex and has a closed-form solution.

$$\sigma^{k+1} = (\sigma_A - \frac{\omega^t}{u^t})_+ \quad (16)$$

It eventually approaches a local optimal point σ^* after numerous iterations, then $X^{k+1} = U \text{diag}\{\sigma^*\} V^T$.

Next, we optimize N through a similar optimization method:

$$N^{t+1} = \underset{N}{\operatorname{argmin}} \lambda\|N\|_l + \frac{u^t}{2}\left\|N - (Z - X^{t+1} - \frac{Y^t}{u^t})\right\|_F^2 \quad (17)$$

TABLE II
PARAMETER SETTING OF SIGNALS IN SIMULATION ANALYSIS.

Parameter	Time(s)	Sampling(HZ)	A	τ (s)
Value	5	5000	4.8	0.012
Parameter	SNR(dB)	f (Hz)	β	T(s)
Value	-10	1125	100	0.1

Due to the irregularity of the TFR N of the noise, the l is set as L1-norm for simplicity. The solution to this subproblem is similar to the traditional convex optimization algorithm [22], as follows mathematically:

$$[N^{t+1}]_{ij} = \max(Q_{ij} - \frac{\lambda}{u^t}, 0) \text{sign}(Q_{ij}) \quad (18)$$

where $Q = Z - X^{t+1} - \frac{Y^t}{u^t}$. After the $t+1$ th iteration, the i th row and j th column elements of the noise TFR matrix are represented by $[N^{t+1}]_{ij}$. $\text{sign}(\cdot)$ stands for sign function. The Lagrange multiplier H and u can be updated as follows:

$$Y^{t+1} = Y^t + u^t(X^{t+1} - Z + N^{t+1}) \quad (19)$$

$$u^{t+1} = \rho u^t \quad (20)$$

Where $\rho > 0$. The entire algorithm utilizing the nonconvex RPCA is outlined in Algorithm 1 and the flowchart illustrating the proposed method is depicted in Fig. 2, which presents a structured three-step methodology for assessing the condition of rolling bearings. Initially, vibration signals are collected and transformed into the TFR using the STFT. The process proceeds to the second step, which involves estimating a nonconvex sparse low-rank matrix from the TFR, employing a feature extraction model that integrates matrix norms and sparsity penalties to produce an optimal matrix representation, X^* . The final stage entails the condition assessment of the

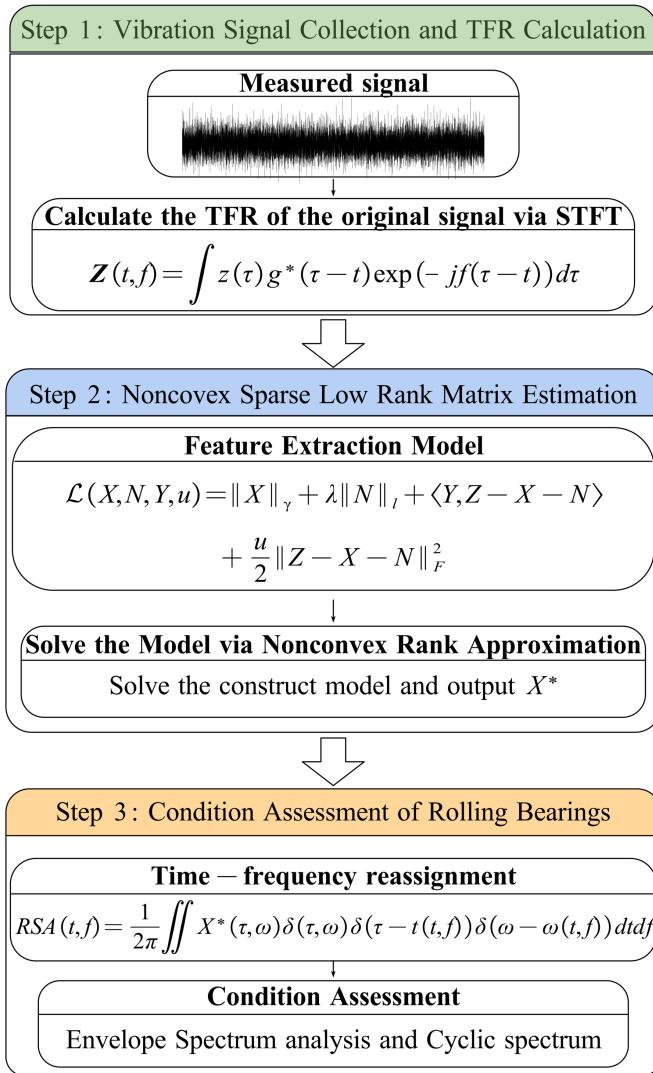


Fig. 2. The flowchart of the proposed technique

bearings, where time-frequency reassignment techniques refine the TFR, followed by envelope and cyclic spectrum analyses that evaluate the bearings' operational integrity.

Our model consists of three parameters: λ, ρ, μ . As demonstrated in [38], μ^0 can have an influence on X . If the value of μ^0 is excessively large, the rank of X will go above the desired low rank. This phenomenon provides a means of adjusting the desired rank, allowing us to modify μ^0 . In the simulation, μ^0 is assigned values of 0.09, 0.07, and 0.04 for noise levels of -3dB, -5dB, and -10dB, respectively; in Experiment I, it is set to 4.5 for noise levels of 10dB, 5dB, and 3dB; and in Experiment II, it is set to 0.1, 0.09, and 0.08 for noise levels of 10dB, 5dB, and 3dB, respectively. In practice, cross-validation can be used to select those parameters. As with [39], we may select λ from a value of about $1/\sqrt{\max(m, n)}$. We just put $\lambda = 10^{-3}$ throughout all of our experiments since the results of our research show that the results we obtained are insensitive to λ over an extensive range. In the term ρ , a large number will cause the solution to converge quickly, instead, a smaller number leads to a more accurate solution. A widely

utilized value in research is 0.05.

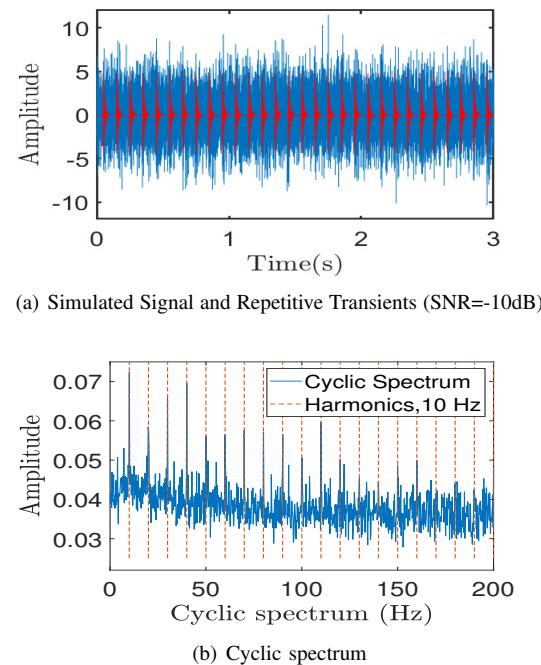


Fig. 3. Time-domain waveform of simulation signal and its cyclic spectrum

IV. SIMULATION STUDY

In this section, we assess the Nonconvex RPCA method using simulated signals, juxtaposing it with established techniques such as VMD, FMD, and RPCA based on L1-norm Convex Optimization. VMD decomposes signals into intrinsic mode functions to facilitate frequency analysis. FMD employs Correlation Kurtosis for identifying impulses and periodicity, optimizing an adaptive FIR filter bank tailored to specific frequency bands. In contrast, the L1-norm technique addresses Equation (7) by substituting the zero norm with the L1 norm, thereby enhancing the extraction of bearing impact features.

A. Simulation Signal Model

The signal representing a damaged bearing in a real-world scenario is elaborated upon in Section II-A. Impulse features play a crucial role in fault identification due to their reflection of the fault signal's frequency. We simulate signals using the outer ring fault simulation method to acquire the desired signal. The simulated signals resulting from bearing variation $z(t)$ at a constant speed are described as follows:

$$\begin{aligned} z(t) &= x(t) + n(t) \\ x(t) &= \sum_i A s(t - iT - \tau_i) \\ s(t) &= e^{-\beta t} \sin(2\pi f t) \end{aligned} \quad (21)$$

where $x(t)$ denotes the simulated repeated transients, and $n(t)$ stands for the white Gaussian noise. Table II lists the settings of parameters, where τ stands for time delay, f is the natural frequency of bearing, and T denotes the failure period.

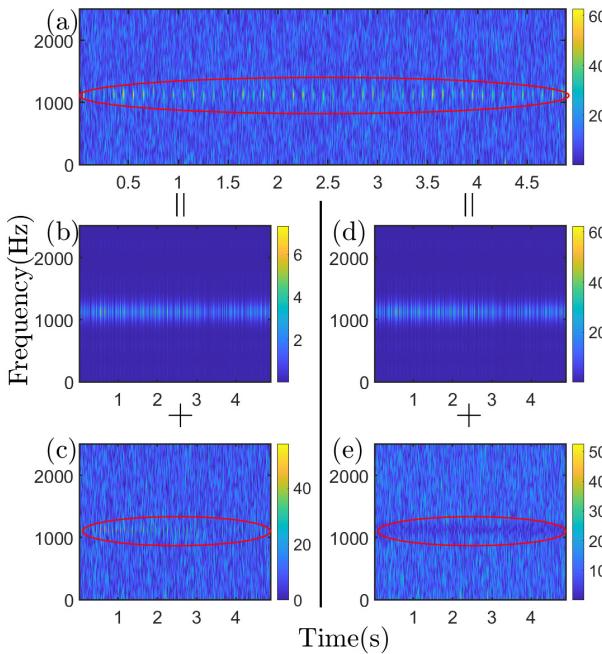


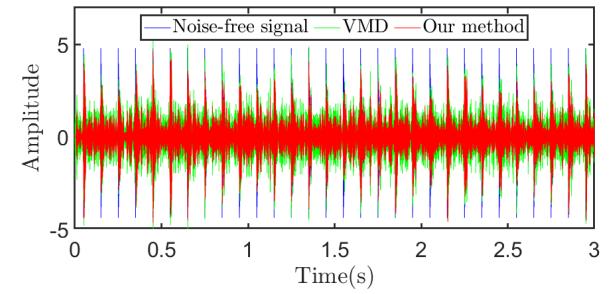
Fig. 4. The TFR of Simulation signal By STFT. (a)the original TFR. **Left:** (b) the obtained feature TFR by L1-norm method, (c) the obtained noise TFR by L1-norm method. **Right:** (d) the obtained feature TFR by Nonconvex RPCA, (e) the obtained noise TFR by Nonconvex RPCA.

B. Performance Verification

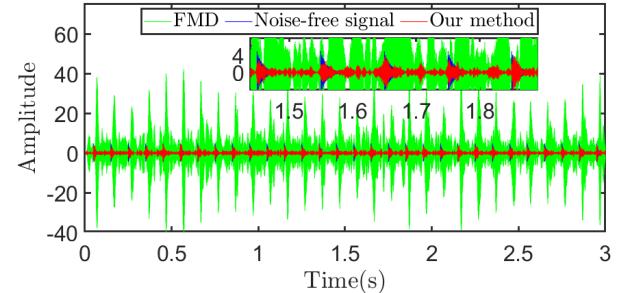
The time-frequency distribution of the simulation signal is computed using STFT, as presented in Fig.4(a). To demonstrate the effectiveness of our proposed method, we initially applied the L1 regularization Alternating Direction Method algorithm(ADM) to extract fault features. The feature time-frequency distribution obtained by this method is shown in Fig.4(b), while Fig.4(c) displays the residual noise time-frequency matrix post-extraction. Although the extracted time-frequency matrix closely resembles the original, the intensity of the extracted features is significantly lower than the pulse components in the simulated signal. This suggests that the L1 method's extracted features inadequately represent the signal's true characteristics. Mathematically, this indicates that the convex relaxation in the optimization problem introduces a considerable error compared to the actual solution, necessitating the development of more accurate algorithms.

In contrast, our proposed Nonconvex RPCA method overcomes these limitations. Fig.4(d) presents the feature time-frequency distribution obtained with our method, and Fig.4(e) shows the residual noise matrix, which notably lacks discernible impulse features after extraction. The effectiveness of our method is further validated by inversely mapping the extracted features to the time domain using ISTFT, as shown in Fig.5(c), which indicates our approach effectively isolates fault characteristics amidst significant noise. This demonstrates our method's capability to extract fault characteristics in noisy environments, though complete signal recovery remains challenging.

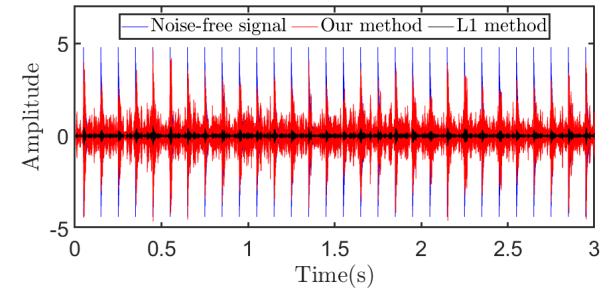
To further validate the efficacy of our approach, we conducted a comparative analysis with VMD and FMD, both of which are renowned for their effectiveness in fault feature extraction. The results of fault feature extraction using VMD and FMD are illustrated in Fig.5(a) and Fig.5(b), respectively. While FMD shows some effectiveness in fault component extraction, its performance is marred by distortions similar to those observed with the L1-norm regularization method, as



(a) The feature signal is obtained by VMD and the Nonconvex RPCA



(b) The feature signal is obtained by FMD and the Nonconvex RPCA



(c) The feature signal is obtained by L1 method and Nonconvex RPCA

Fig. 5. The feature signal is obtained by different method

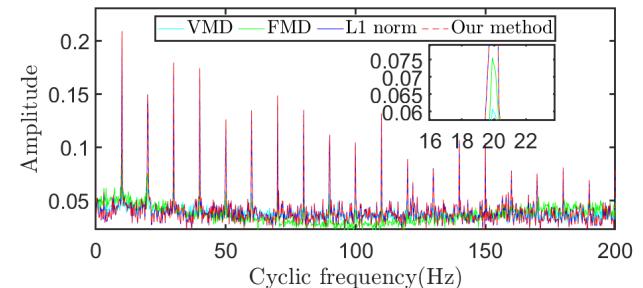


Fig. 6. The cyclic spectral of features extracted using different methods

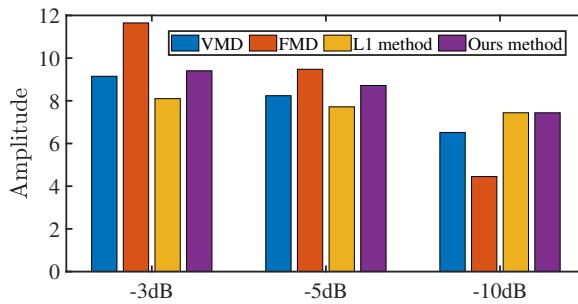


Fig. 7. The pulse factor of features extracted by different methods under different noise levels

shown in Fig.5(b). Notably, FMD's distortions are magnified, indicating potential failure points in its application. In contrast, VMD demonstrates a higher level of effectiveness in extracting fault impulse features, as evidenced in Fig.5(a). Then, we conducted spectral correlation mapping of the components extracted from each signal. As depicted in Fig.6, our method surpasses both VMD and FMD in extracting fault features, indicating its advanced capabilities in this domain.

In addition, we conducted experiments with varying noise levels. We utilized pulse factor [40] to evaluate the extracted feature signal in Fig.7, from which we observe that our method performs well despite different levels of noise interference. Although FMD achieves a high level of performance, the features it extracts are distorted. Overall, our method demonstrates strong performance. The next section will detail the experimental testing of our method, further substantiating its practical applicability.

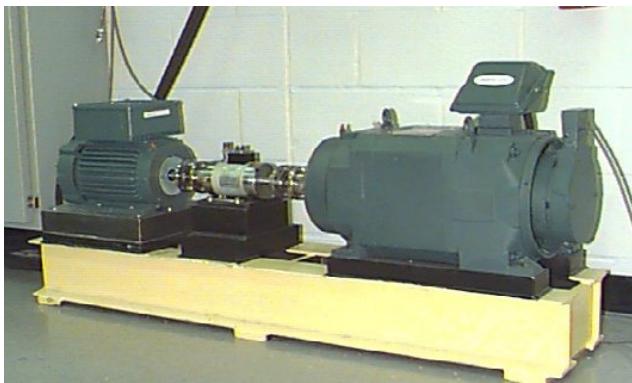


Fig. 8. Case Western Reserve University Bearing Data Derive

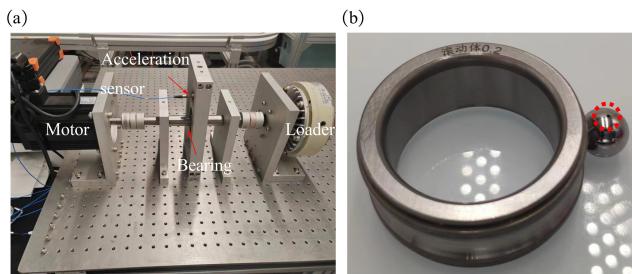


Fig. 9. SCUT rolling bearing test rig. (a) the layout of the test rig. (b) rolling ball fault of rolling bearings

TABLE III

PARAMETER SETTING OF SIGNALS IN SIMULATION ANALYSIS.

Case	Manufacturer	Type	Location of fault
I	SKF	6205-2RS	Inner ring
II	HRB	6206-RS	Ball
Case	Fault diameter(mm)	RPM(r/min)	Sampling(KHz)
I	0.3556	1750	12
II	0.2	1500	32

V. EXPERIMENTS

In this part, we will validate our method with two experiments. The first case study is from Case Western Reserve University (CWRU) bearing dataset [41] and the second case study is from a rolling bearing test rig at South China University of Technology (SCUT).

A. Working Condition Parameter and Data Acquisition Device

The first case study verifies the proposed method using a vibration signal from the CWRU dataset. The configuration of the CWRU test bench is depicted in Fig.8. The acceleration sensor captures vibration signals at the drive end of the bench. The second case study verifies the proposed method using the SCUT rolling bearing test rig as shown in Fig.9. The test rig consists of motors(Left), bearings(Middle), and Loader(Right), which are connected by shafts and couplings. The acceleration sensor is placed above the bearing and secured by a magnet. Table III provides the details of experiments.

B. Case I: CWRU Bearing Data

In our first case study, we analyze vibration signals from the CWRU bearing dataset using our method and other established approaches. The Cyclic Modulation Spectrum of the original signal, as presented in Fig. 10, reveals the difficulty in

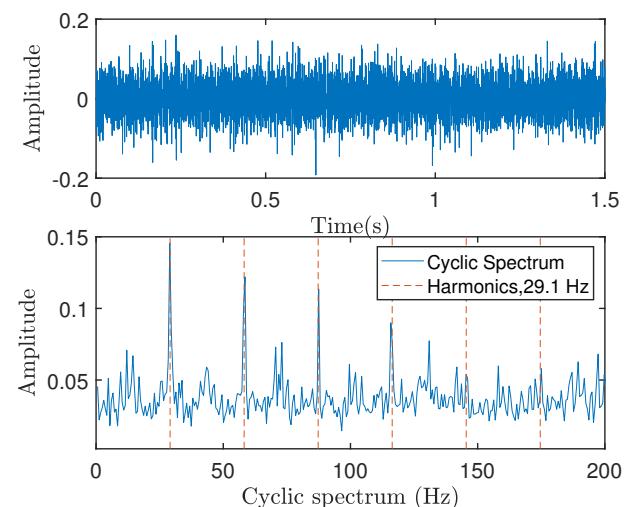


Fig. 10. Time-domain waveform of bear fault signal and its Cyclic spectrum from CWRU dataset

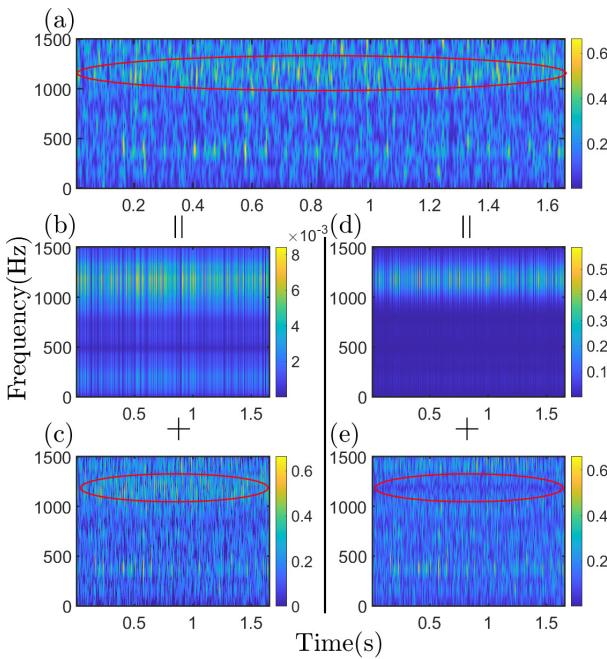


Fig. 11. The TFR of CWRU bearing data By STFT. (a)the original TFR. **Left:** (b)the obtained feature TFR by L1-norm method, (c)the obtained noise TFR by L1-norm method. **Right:** (d)the obtained feature TFR by Nonconvex RPCA method, (e)the obtained noise TFR by Nonconvex RPCA method.

extracting essential features due to high noise levels obscuring defect information.

The signals were first transformed into the time-frequency domain using a Hanning window of size 50 in the STFT, illustrated in Fig.11(a). We began with the conventional L1-norm regularization method for impulse feature extraction. The resulting faulty feature time-frequency distribution and the residual noise time-frequency matrix are shown in Fig.11.(b) and Fig.11.(c), respectively. However, as discussed in Section II-C, this method encounters significant errors post convex relaxation, leading to distortions and proving unsuitable for bearing fault diagnosis. The limitations of feature extraction based on convex optimization are evident in this scenario.

To overcome these challenges, our improved Non-convex RPCA method demonstrates its robustness in fault feature extraction, as depicted in Fig.11(d). By applying ISTFT, we converted the extracted feature time-frequency distribution back to the time domain. Fig.12(c) contrasts the feature signals extracted using traditional L1-norm convex optimization and our method, highlighting the latter's effectiveness in fault feature extraction.

The ability of VMD and FMD to extract fault impulse features is showcased in Fig.12(a) and Fig.12(b). Both methods show competence in extracting fault features from complex signal components, particularly evident in the time domain. To further demonstrate our method's advantages, we mapped the extracted components to the Square Envelope Spectrum (SES), as shown in Fig.13.

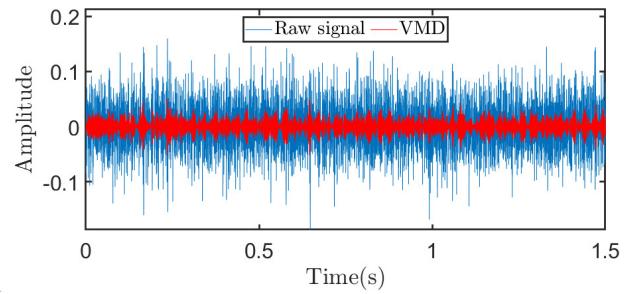
The feature signal extracted by VMD, is displayed in Fig.13 .(a). Similarly, FMD's extracted feature signal is shown

in Fig13.(b). In contrast, L1-norm based RPCA effectively isolates characteristic components while minimizing non-characteristic components, as seen in Fig.13.(c). However, this approach may deviate from the actual underlying reality. This limitation is addressed by our non-convex RPCA method, which accurately extracts features and ensures robustness, as evidenced in Fig.13.(d).Fig.14 presents a comparison of pulse factors extracted by features after adding noise with different signal-to-noise ratios. Our method continues to perform well despite varying levels of noise interference.

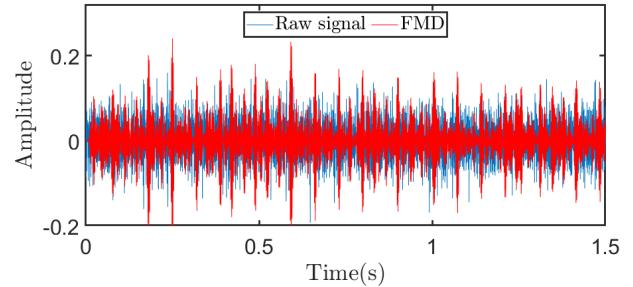
C. Case II: SCUT Rolling Bearing Test Rig

In this study, we analyze vibration signals from bearings with ball defects using various methods, including our proposed technique. The signal preprocessing involved filtering out low-frequency noise below 500Hz and adding Gaussian noise with an SNR of 5. The waveform and frequency spectrum of this preprocessed signal are shown in Fig.14.

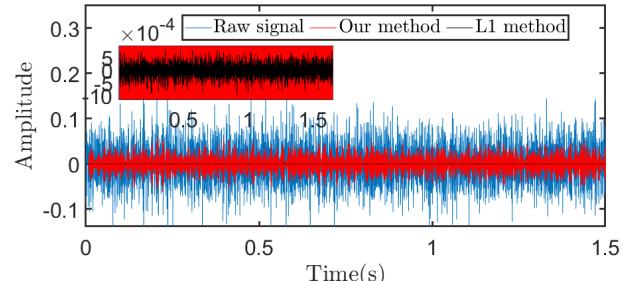
The signals were then converted into the time-frequency



(a) The feature signal is obtained by VMD



(b) The feature signal is obtained by FMD



(c) The feature signal is obtained by L1 method and Nonconvex RPCA

Fig. 12. The feature signal is obtained by different methods

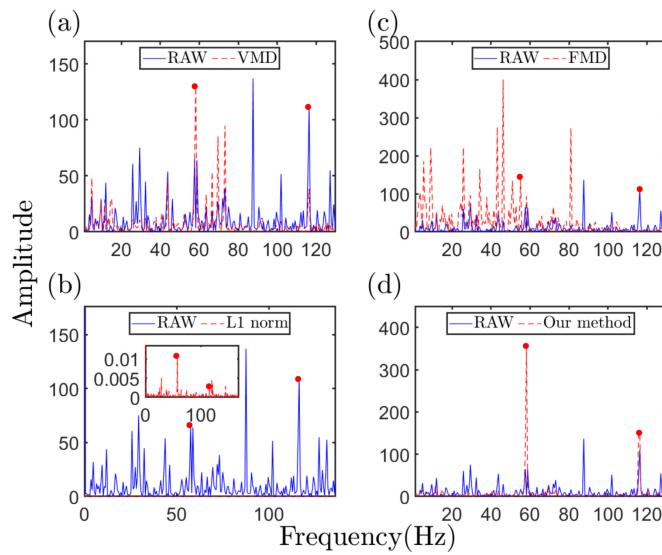


Fig. 13. The SES of feature signals extracted by different methods ($f_2=58.2\text{Hz}$): (a) the SES of feature signals by VMD; (b) the SES of feature signals by FMD; (c) the SES of feature signals by L1 method; (d) the SES of feature signals by Nonconvex RPCA.

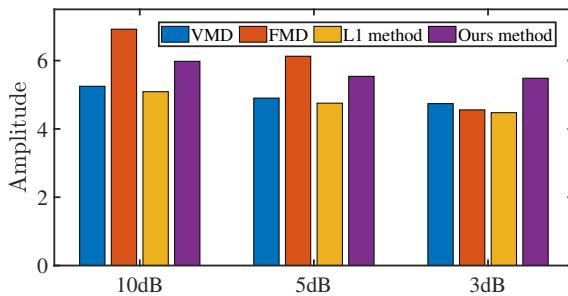


Fig. 14. The pulse factor of features extracted by different methods under different noise levels

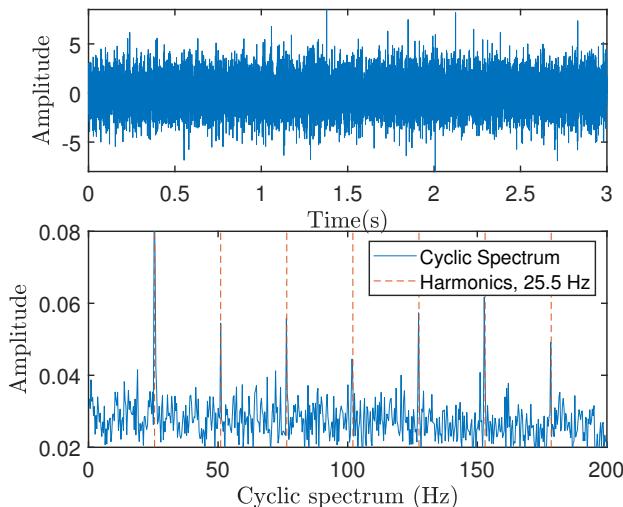


Fig. 15. The waveform of bear fault signal and its Cyclic spectrum from SCUT test rig

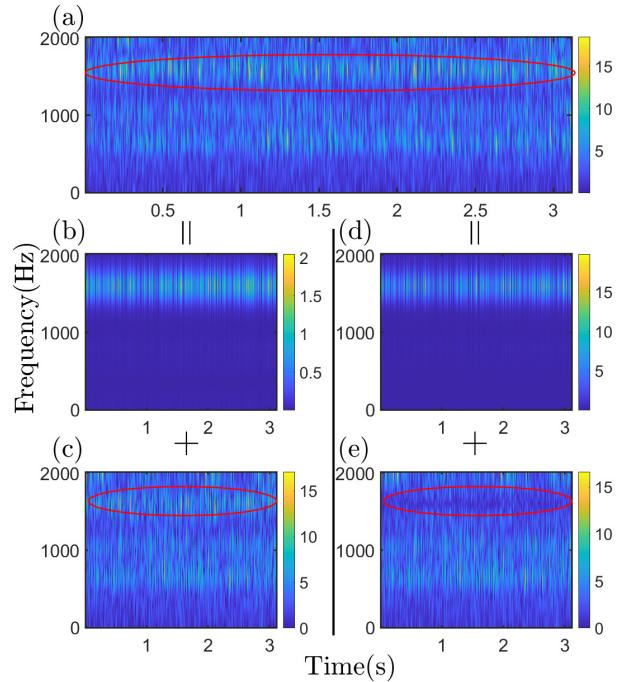


Fig. 16. The TFR of faulty bearing signal By STFT. (a) the original TFR. **Left:** (b) the obtained feature TFR by L1-norm method, (c) the obtained noise TFR by L1-norm method. **Right:** (d) the obtained feature TFR by Nonconvex RPCA, (e) the obtained noise TFR by Nonconvex RPCA.

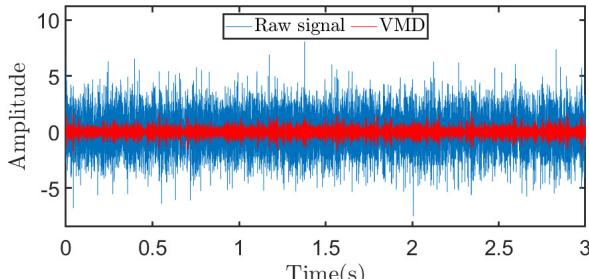
domain using a Hanning window of size 50 in the STFT, as depicted in Fig.16(a). We initially employed the conventional L1-norm regularization method to extract fault characteristics. The feature time-frequency distribution and the residual noise time-frequency distribution post-extraction are presented in Fig.16(b) and Fig.16(c), respectively.

We then applied our proposed method for fault feature extraction. The resultant feature time-frequency distribution and noise time-frequency matrix are illustrated in Fig.16(d) and Fig.16(e). In contrast to the L1-norm method, our approach leaves negligible feature components in the noise matrix (Fig.16(e)). Using ISTFT, we reverted these extracted features to the time domain, as shown in Fig.17(c), comparing signals extracted by the L1 method and our method.

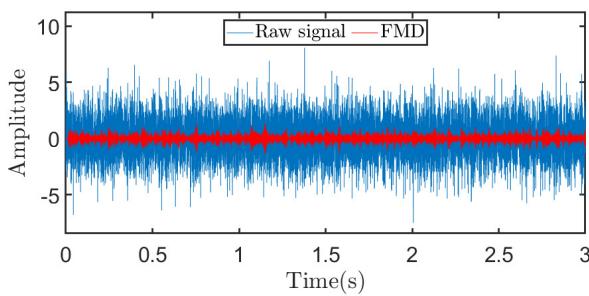
Fault impulse extraction using VMD and FMD is displayed in Fig.17(a) and Fig.17(b). Although VMD and FMD are capable of extracting certain fault features, their comprehensiveness is limited, often failing to accurately assess machinery conditions. In the time domain, distinguishing the strengths and weaknesses of these methods is challenging. Therefore, we analyzed the components extracted by each method using the square envelope spectrum. As demonstrated in Fig.18, our method yields more prominent fault features, offering a closer representation of the actual condition compared to other methods. Fig.19 presents a comparison of pulse factors extracted by features after adding noise with different signal-to-noise ratios. Our method demonstrates strong performance.

D. Discussions

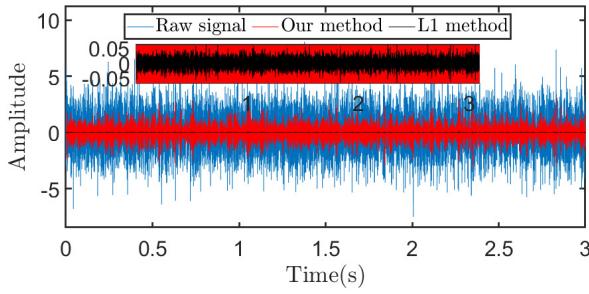
The introduction of the RPCA method based on the γ norm



(a) The feature signal is extracted by VMD



(b) The feature signal is extracted by FMD



(c) The feature signal is extracted by L1 method and Nonconvex RPCA

Fig. 17. The feature signal is obtained by different method

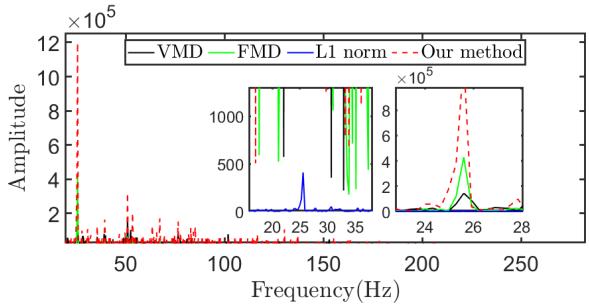


Fig. 18. The SES of feature signals extracted by different methods

has marked a significant advancement in bearing fault diagnosis, addressing distortion issues more effectively compared to traditional L1 norm approaches. However, challenges such as the incomplete recovery of features in low-rank scenarios and the unavoidable incorporation of noise into extracted features due to noise-feature coupling necessitate further refinement. These limitations not only pinpoint the method's current constraints but also clearly outline the direction for future

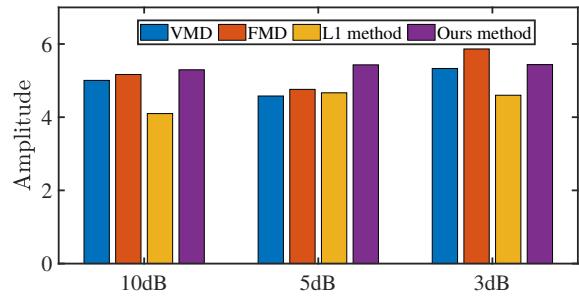


Fig. 19. The pulse factor of features extracted by different methods under different noise levels.

research aimed at optimizing the algorithm's performance.

Future research should prioritize improving noise discrimination, refining parameter optimization, and investigating the combination of non-convex RPCA with deep learning. These efforts will enhance transient fault signal extraction purity and expand applicability across different fault scenarios and noise levels. Further validation in real-world applications and advancing theoretical understanding of the γ -function could lead to more versatile fault diagnosis solutions.

VI. CONCLUSIONS

This study introduces an innovative approach to bearing fault diagnosis through a non-convex RPCA method that incorporates the γ norm, showcasing a significant leap forward in addressing the limitations of traditional L1 norm techniques. The method's innovative use of the Moreau-Yosida operator for eigenvalue optimization, validated by both simulation and experimental results, underscores its potential in enhancing signal feature extraction amidst noise. Despite facing challenges such as incomplete feature recovery in low-rank cases and noise-feature coupling, this research not only demonstrates the method's effectiveness but also lays a solid foundation for future advancements. Moving forward, it opens up promising avenues for optimizing algorithmic performance and expanding its application to a wider range of real-world diagnostic scenarios, thereby contributing to the development of more robust and efficient fault diagnosis methodologies.

APPENDIX I CONVERGENCE ANALYSIS OF NONCONVEX RPCA

The convergence of the Nonconvex RPCA will be demonstrated in this section. We write $\|X\|_\gamma$ as $F(L)$:

$$\begin{aligned} \mathcal{L}(X, N, Y, u) = & F(X) + \lambda \|N\|_l + \langle Y, Z - X - N \rangle \\ & + \frac{u}{2} \|Z - X - N\|^2_F \end{aligned} \quad (22)$$

Lemma 1. $\{Y^t\}$ is a bound sequence.

Proof: The first-order necessary local optimality condition is satisfied by $\{N^{t+1}\}$,

$$\begin{aligned} 0 \in & \partial_N (X^{t+1}, N, Y^t, u^t) \Big|_{N^{t+1}} \\ = & \partial_N (\lambda \|N\|_1) \Big|_{N^{t+1}} + N^t + u^t (X^{t+1} - Z + N^{t+1}) \\ = & \partial_N (\lambda \|N\|_1) \Big|_{N^{t+1}} + N^{t+1} \end{aligned} \quad (23)$$

Where the subgradient $[\partial_N \|N\|_1]_{ij} = 0$ if $Y_{ij} = 0$. Then, $\partial_N(\lambda \|N\|_1)|_{N^{t+1}}$ it is bounded. Thus, $\{Y^t\}$ it is bounded.

Lemma 2. $\{X^t\}$ and $\{N^t\}$ are bounded if $\sum_{t=1}^{\infty} \frac{u^t + u^{t+1}}{(u^t)^2} < \infty$

Proof: we have the equality with some algebra

$$\begin{aligned} \mathcal{L}(X^t, N^t, Y^t, u^t) &= \mathcal{L}(X^t, N^t, Y^{t-1}, u^{t-1}) + \frac{u^t - u^{t-1}}{2} \|X^t - Z + N^t\|_F^2 \\ &\quad + Tr[(Y^t - Y^{t-1})(X^t - Z + N^t)] \\ &= \mathcal{L}(X^t, N^t, Y^{t-1}, u^{t-1}) + \frac{u^t + u^{t-1}}{2(u^{t-1})^2} \|Y^t - Y^{t-1}\|_F^2 \end{aligned} \quad (24)$$

Then,

$$\begin{aligned} \mathcal{L}(X^{t+1}, N^{t+1}, Y^t, u^t) &\leq \mathcal{L}(X^{t+1}, N^t, Y^t, u^t) \\ &\leq \mathcal{L}(X^t, N^t, Y^t, u^t) \\ &\leq \mathcal{L}(X^t, N^t, Y^t, u^t) + \frac{u^t + u^{t-1}}{2(u^{t-1})^2} \|Y^t - Y^{t-1}\|_F^2 \end{aligned} \quad (25)$$

After t times, we obtain

$$\begin{aligned} \mathcal{L}(X^{t+1}, N^{t+1}, Y^t, u^t) &\leq \mathcal{L}(X^1, N^1, Y^0, u^0) + \sum_{i=1}^t \frac{u^i + u^{i-1}}{2(u^{i-1})^2} \|Y^i - Y^{i-1}\|_F^2 \end{aligned} \quad (26)$$

All terms on the right-hand side of the aforementioned inequality are bounded since $\|Y^i - Y^{i-1}\|_F^2$ is bounded. As such, $\mathcal{L}(X^{t+1}, N^{t+1}, Y^t, u^t)$ is upper bounded.

Again,

$$\begin{aligned} \mathcal{L}(X^{t+1}, N^{t+1}, Y^t, u^t) + \frac{1}{2u^t} \|Y^t\|_F^2 &= F(X^{t+1}) + \lambda \|N^{t+1}\|_1 + \frac{u^t}{2} \|X^{t+1} - Z + N^{t+1} + \frac{Y^t}{u^t}\|_F^2 \end{aligned} \quad (27)$$

As each term on the right-hand side has a boundary, N^{t+1} is bounded. X^{t+1} is bounded by the last term on the right-hand side.

Theorem 1. [42] Suppose $F : \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}$ can be represented as $F(X) = f \circ \sigma(X)$, where $X \in \mathbb{R}^{n_1 \times n_2}$ with SVD ($X = U \text{diag}(\sigma_1, \dots, \sigma_n) V^T$, $n = \min(n_1, n_2)$), and f is differentiable. The gradient of $F(X)$ at X is:

$$\frac{\partial F(X)}{\partial X} = U \text{diag}(\theta) V^T \quad (28)$$

$$\text{where } \theta = \left. \frac{\partial f(y)}{\partial y} \right|_{y=X} = \sigma(X)$$

Theorem 2. let $\{X^t, N^t, Y^t\}$ be the sequence generated in Algorithm 1 and $\{X^t, N^t, Y^t\}$ be an accumulation point. Then $\{X^*, N^*\}$ is a stationary point of optimization problem if $\sum_{t=1}^{\infty} \frac{u^t + u^{t+1}}{(u^t)^2} < \infty$ and $\lim_{t \rightarrow \infty} u^t(Y^{t-1} - Y^t) \rightarrow 0$

Proof: Lemma 1 and 2 demonstrate that the sequence $\{X^t, N^t, Y^t\}$ generated through Algorithm 1 is bounded. The series of values needs to have at least one accumulation point, such as $\{X^*, N^*, Y^*\}$, according to the Bolzano-Weierstrass theorem. We assume that $\{X^t, N^t, Y^t\}$ itself converges to $\{X^*, N^*, Y^*\}$ without loosing generality. We have $\lim_{t \rightarrow \infty} X^t + Y^t - Z = \frac{Y^t - Y^{t-1}}{u^{t-1}}$. $Z = X^* + N^*$ and $X^t + N^t - Z \rightarrow 0$. Thus, the requirement for basic feasibility is satisfied.

For X^{t+1} ,

$$\begin{aligned} \partial_X L(X, N, Y, u)|_{X^{t+1}} &= \partial_X F(X)|_{X^{t+1}} + Y^t + u^t(X^{t+1} - Z + N^t) \\ &= \partial_X F(X)|_{X^{t+1}} + Y^{t+1} - u^t(N^{t+1} - N^t) \\ &= 0 \end{aligned} \quad (29)$$

If the SVD of X is $U \text{diag}(\sigma_i) V^T$ (28),

$$\partial_X F(X)|_{X^{t+1}} = U \text{diag}(\theta) V^T \quad (30)$$

where $\theta_i = \frac{(1+\gamma)\gamma}{(\gamma+\sigma_i)^2}$, if $\sigma_i \neq 0$; otherwise, it is $(\gamma+1)/\gamma$. Since θ_i is finite, $\partial_X F(X)|_{X^{t+1}}$ is bounded. As $\{Y_t\}$ is also bounded, $\lim_{t \rightarrow \infty} u^t(Y^{t-1} - Y^t)$ is bounded. Under the assumption that $\lim_{t \rightarrow \infty} u^t(Y^{t-1} - Y^t) \rightarrow 0$, $\partial_X F(X^*) + Y^* = 0$. Therefore, X^*, N^*, Y^* satisfies the KKT conditions of $\mathcal{L}(X, N, Y)$, indicating that X^*, N^* is a stationary point.

REFERENCES

- [1] W. Li, H. Lan, J. Chen, K. Feng, and R. Huang, "Wavcapsnet: An interpretable intelligent compound fault diagnosis method by backward tracking," *IEEE Trans. Instrum. Meas.*, 2023.
- [2] R. Tian, M. Cui, and G. Chen, "A neural-symbolic network for interpretable fault diagnosis of rolling element bearings based on temporal logic," *IEEE Trans. Instrum. Meas.*, 2024.
- [3] G. Chen, Y. Lu, and R. Su, "Interpretable fault diagnosis with shapelet temporal logic: Theory and application," *Automatica*, vol. 142, p. 110350, 2022.
- [4] M. Cui, Y. Wang, X. Lin, and M. Zhong, "Fault diagnosis of rolling bearings based on an improved stack autoencoder and support vector machine," *IEEE Sensors J.*, vol. 21, no. 4, pp. 4927–4937, 2020.
- [5] D. Li, J. Chen, R. Huang, Z. Chen, and W. Li, "Sensor-aware capsnet: Towards trustworthy multisensor fusion for remaining useful life prediction," *J. Manuf. Syst.*, vol. 72, pp. 26–37, 2024.
- [6] R. Huang, J. Li, Y. Liao, J. Chen, Z. Wang, and W. Li, "Deep adversarial capsule network for compound fault diagnosis of machinery toward multidomain generalization task," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2020.
- [7] Z. Xu, C. Qin, and G. Tang, "A novel deconvolution cascaded variational mode decomposition for weak bearing fault detection with unknown signal transmission path," *IEEE Sensors J.*, vol. 21, pp. 1746–1755, 2021.
- [8] N. Jia, Y. Cheng, Y. Liu, and Y. Tian, "Intelligent fault diagnosis of rotating machines based on wavelet time-frequency diagram and optimized stacked denoising auto-encoder," *IEEE Sensors J.*, vol. 22, no. 17, pp. 17139–17150, 2022.
- [9] J. Zhao, Y. Zhang, and Q. Chen, "Rolling bearing fault feature extraction based on adaptive tunable q-factor wavelet transform and spectral kurtosis," *Shock. Vib.*, vol. 2020, pp. 1–19, 2020.
- [10] X. Zhou, Y. Li, L. Jiang, and L. Zhou, "Fault feature extraction for rolling bearings based on parameter-adaptive variational mode decomposition and multi-point optimal minimum entropy deconvolution," *Measurement*, vol. 173, p. 108469, 2021.
- [11] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, 2013.
- [12] J. Wang, G. Du, Z. Zhu, C. Shen, and Q. He, "Fault diagnosis of rotating machines based on the emd manifold," *Mechanical Systems and Signal Processing*, vol. 135, p. 106443, 2020.
- [13] Y. Miao, B. Zhang, C. Li, J. Lin, and D. Zhang, "Feature mode decomposition: New decomposition theory for rotating machinery fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 70, no. 2, pp. 1949–1960, 2022.
- [14] D. Liu, W. Cheng, and W. Wen, "Rolling bearing fault diagnosis via stft and improved instantaneous frequency estimation method," *Procedia Manuf.*, vol. 49, pp. 166–172, 2020.
- [15] M. Jalayer, C. Orsenigo, and C. Vercellis, "Fault detection and diagnosis for rotating machinery: A model based on convolutional lstm, fast fourier and continuous wavelet transforms," *Comput. Ind.*, vol. 125, p. 103378, 2021.

- [16] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 243–261, 2011.
- [17] G. Yu, M. Yu, and C. Xu, "Synchroextracting transform," *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 8042–8054, 2017.
- [18] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis—with an application to gravitational-wave signal," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3168–3178, 2017.
- [19] G. Xin, N. Hamzaoui, and J. Antoni, "Semi-automated diagnosis of bearing faults based on a hidden markov model of the vibration signals," *Measurement*, vol. 127, pp. 141–166, 2018.
- [20] R. Sun, Z. Yang, X. Chen, S. Tian, and Y. Xie, "Gear fault diagnosis based on the structured sparsity time-frequency analysis," *Mech. Syst. Signal Process.*, vol. 102, pp. 346–363, 2018.
- [21] R. Wang, L. Yu, H. Fang, L. Yu, J. Chen, and C. Shen, "Cyclic correlation density decomposition based on a sparse and low-rank model for weak fault feature extraction of rolling bearings," *Measurement*, vol. 198, p. 111393, 2022.
- [22] R. Wang, H. Fang, L. Yu, L. Yu, and J. Chen, "Sparse and low-rank decomposition of the time-frequency representation for bearing fault diagnosis under variable speed conditions," *ISA Trans.*, vol. 128, pp. 579–598, 2022.
- [23] Z. Feng, Y. Zhou, M. J. Zuo, F. Chu, and X. Chen, "Atomic decomposition and sparse representation for complex signal analysis in machinery fault diagnosis: A review with examples," *Measurement*, vol. 103, pp. 106–132, 2017.
- [24] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," *Proc. Adv. Neural Inf. Process. Syst.*, vol. 22, 2009.
- [25] C. Lu, J. Tang, S. Yan, and Z. Lin, "Generalized nonconvex nonsmooth low-rank minimization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 4130–4137.
- [26] L. Yu, J. Antoni, and Q. Leclerc, "Spectral matrix completion by cyclic projection and application to sound source reconstruction from non-synchronous measurements," *J. Sound Vib.*, vol. 372, pp. 31–49, 2016.
- [27] L. Yu and Q. Leclerc, "Combined regularization optimization for separating transient signal from strong noise in rolling element bearing diagnostics," in *Surveillance 7*, 2013.
- [28] S. Huang, L. Yu, and W. Jiang, "Water entry sound detection in strong noise by using the spectrogram matrix decomposition method," *Appl. Acoust.*, vol. 161, p. 107171, 2020.
- [29] L. Yu, W. Dai, S. Huang, and W. Jiang, "Sparse time-frequency representation for the transient signal based on low-rank and sparse decomposition," *J. Acoust.*, vol. 1, no. 1, 2019.
- [30] B. Wang, Y. Liao, C. Ding, and X. Zhang, "Periodical sparse low-rank matrix estimation algorithm for fault detection of rolling bearings," *ISA Trans.*, vol. 101, pp. 366–378, 2020.
- [31] Z. Du, X. Chen, H. Zhang, B. Yang, Z. Zhai, and R. Yan, "Weighted low-rank sparse model via nuclear norm minimization for bearing fault detection," *J. Sound Vib.*, vol. 400, pp. 270–287, 2017.
- [32] I. Selesnick, "Sparse regularization via convex analysis," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4481–4494, 2017.
- [33] S. Wang, I. Selesnick, G. Cai, Y. Feng, X. Sui, and X. Chen, "Non-convex sparse regularization and convex optimization for bearing fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 65, no. 9, pp. 7332–7342, 2018.
- [34] W. Huang, Z. Song, C. Zhang, J. Wang, J. Shi, X. Jiang, and Z. Zhu, "Multi-source fidelity sparse representation via convex optimization for gearbox compound fault diagnosis," *J. Sound Vib.*, vol. 496, p. 115879, 2021.
- [35] Z. Zhang, W. Huang, Y. Liao, Z. Song, J. Shi, X. Jiang, C. Shen, and Z. Zhu, "Bearing fault diagnosis via generalized logarithm sparse regularization," *Mech. Syst. Signal Process.*, vol. 167, p. 108576, 2022.
- [36] Z. Kang, C. Peng, and Q. Cheng, "Robust pca via nonconvex rank approximation," in *2015 IEEE International Conference on Data Mining*. IEEE, 2015, pp. 211–220.
- [37] P. D. Tao and L. H. An, "Convex analysis approach to dc programming: theory, algorithms and applications," *Acta Math. Vietnam.*, vol. 22, no. 1, pp. 289–355, 1997.
- [38] W. K. Leow, Y. Cheng, L. Zhang, T. Sim, and L. Foo, "Background recovery by fixed-rank robust principal component analysis," in *Computer Analysis of Images and Patterns: 15th International Conference, CAIP 2013, York, UK, August 27–29, 2013, Proceedings, Part I 15*. Springer, 2013, pp. 54–61.
- [39] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, pp. 1–37, 2011.
- [40] A. Rai and S. H. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International*, vol. 96, pp. 289–306, 2016.
- [41] K. Loparo, "Bearings vibration data sets," *Case Western Reserve University Bearing Data Center*, Accessed: May 12, 2022. [Online]. Available: <https://engineering.case.edu/bearingdatacenter/welcome>
- [42] A. S. Lewis and H. S. Sendov, "Nonsmooth analysis of singular values. part i: Theory," *Set-Valued Analysis*, vol. 13, pp. 213–241, 2005.



Cai Li received the B.S. degree from Southwest Petroleum University, Chengdu, China, in 2022. He is currently pursuing an M.S. degree in mechanical engineering with the Shien-Ming Wu School of Intelligent Engineering, South China University Of Technology, Guangzhou, China.

His research interests include fault diagnosis, signal process and optimization theory.



Penhong Lu received the master's degree from Zhejiang Agriculture and Forestry University, Hangzhou, China, in 2022. He is currently pursuing a Ph.D. degree in mechanical engineering with the Shien-Ming Wu School of Intelligent Engineering, South China University Of Technology, Guangzhou, China.

His research interests include formal methods, signal processing, and fault diagnosis.



Guangming Dong received the bachelor's and master's degrees in mechanical engineering from Xi'an Jiaotong University, Shanxi, China, in 1999 and 2002, respectively and his Ph.D. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2007. He was a post-doctoral fellow with the Department of mechanical engineering, at Shanghai Jiao Tong University from 2007 to 2009. From 2009 to 2015, he was a lecturer in the Department of Mechanical Engineering, and since 2016 he is an associate professor in the department of Mechanical Engineering at Shanghai Jiao Tong University. His research interests include dynamics of mechanical systems and fault diagnosis, mechanical system signal processing, and fault diagnosis.



Gang Chen received the bachelor's and master's degrees in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2012 and 2015, respectively and his Ph.D. degree in mechanical and aerospace engineering from the University of California, Davis, CA, in 2020. He was a research fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore from 2020 to 2021. He is currently an associate professor at the Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, China. His research interests include machine learning, formal methods, control, signal processing and fault diagnosis.