IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON ADVANCES IN PROGNOSTICS AND SYSTEM HEALTH MANAGEMENT

Received January 27, 2019, accepted February 9, 2019, date of publication February 14, 2019, date of current version March 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2899036

Improved Fault Size Estimation Method for Rolling Element Bearings Based on Concatenation Dictionary

LINGLI CUI[1], (Member, IEEE), XIN WANG[1], HUAQING WANG[2], (Member, IEEE), AND NA WU[1]

[1]Key Laboratory of Advanced Manufacturing Technology, Beijing University of Technology, Beijing 100124, China
[2]College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing 100029, China

Corresponding authors: Lingli Cui (acuilingli@163.com) and Huaqing Wang (hqwang@mail.buct.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 51575007 and Grant 51675035.

ABSTRACT This paper offers a new perspective on the vibrations of discrete bearing faults by focusing on the micro-motion states of rolling elements in spall fault bearings and proposes an improved matching pursuit algorithm for quantitative diagnosis with a high accuracy of atom selection and calculation efficiency. The generation mechanism of the vibration response signal is explained by analyzing the micro-motion status when rolling elements passing through the spall. A concatenation dictionary composed of an impact dictionary as the higher level and step dictionary as the lower level is constructed based on the acceleration variation analysis of the rolling elements. The information output by the higher-level dictionary is used as the input information for the lower-level dictionary to extract the fault features. Only one iteration on the higher-level dictionary is necessary to extract the correct impact atoms, with all subsequent iterative steps assigned to the lower-level dictionary. The advantage is that the influence of high-energy impact components on the extraction of step atoms can be removed. Thereafter, the optimized algorithm based on the concatenation dictionary is applied to the analysis of simulation and experimental signals. The comparative analysis demonstrates that the effective quantitative diagnosis is obtained, while the diagnostic precision and calculation efficiency are improved.

INDEX TERMS Quantitative diagnosis, rolling element bearing, matching pursuit, concatenation dictionary, fault size estimation.

I. INTRODUCTION

To date, vibration signal analysis remains the most common method for bearing fault diagnosis, and the qualitative analysis of bearing health conditions and the pattern recognition of fault types are research hotspots [1]–[3]. However, quantitative analysis of fault severity appears to have become one of most active and valid means of realizing appropriate maintenance decisions, because a bearing fault signal contains information regarding not only the fault condition and type, but also the fault severity. Therefore, researchers have increasingly been studying the generation mechanism of the vibration signal when rolling elements pass through the spall, and have focused on the quantitative diagnosis of mechanical faults.

The associate editor coordinating the review of this manuscript and approving it for publication was Chuan Li.

The quantitative diagnosis of the fault mechanisms of rolling bearings has attracted considerable attention, and a series of bearing dynamic models have been developed. Dowling indicated that an impact will be excited when the ball enters the edge of a defect, and a second impulse will appear when the ball exits the defective zone [4]. This second impact creates a new wave that interferes with the first, resulting in a phase shift. Epps [5] introduced the time to impact, which is the period of separation between these two events, and analyzed the mechanism of the bearing defect vibration and ''double impact'' in detail. The first impact is a low-frequency part, while the second is a relatively high-frequency component. Sassi *et al.* [6] achieved effective simulation of the dynamic behavior of rotating ball bearings in the presence of localized surface defects. Sopanen and Mikkola [7], [8] used a dynamic bearing model to determine that the diametral clearance has a significant

effect on the natural frequencies and vibration response of the system. Sawalhi and Randall [9] found that the acceleration time signal responses resulting from a rolling element entry into and exit from a typical spall on the inner and outer bearing races in two test rigs were very different, with the first being a low-frequency step response, and the second a broadband impulse response. Patil *et al.* [10] contributed significantly to simulating the bearing outer race defects with a mathematical model, by means of which the defect positions and sizes were obtained. Cui *et al.* [11] proposed a novel horizontal–vertical synchronized root mean square localization law and formula based on the analysis of the vibration acceleration signal and dynamic contact force. Singh *et al.* [12] analyzed the vibration response and contact forces in a defective bearing using an explicit dynamics finite element model. Rafsanjani *et al.* [13] proposed an analytical model to study the nonlinear dynamic behavior of surface defects on rolling element bearing systems. Ahmadi and Petersen considered the finite size of rolling elements to construct an improved nonlinear dynamic model. The mechanisms causing inaccuracies in the predicted vibration response when including the rolling elements as point masses instead of finite-sized objects were identified. This approach can be used for defect size estimation techniques with a wider range of defect sizes and geometries [14]. Theoretical guidance was provided through these bearing models and the mechanism for quantitative bearing fault diagnosis.

However, numerous researchers have focused on fault feature extraction methods in the quantitative analysis of fault severity [15], [16]. Baydar and Ball [17] found that acoustic signals could be used effectively to detect the early times of various local gearbox faults using the smoothed pseudo-Wigner Ville distribution. Shen *et al.* [18] diagnosed bearing faults quantitatively using a method based on support vector regression. Moreover, the morphological filtering method and signal complexity were used for assessing the fault severity of rolling element bearings [19]. Hong and Liang [20] determined the relation between the fault severity and Lempel-Ziv complexity by using a continuous wavelet transform method. Sawalhi and Randall [21] and Randall and Sawalhi [22] detailed a combined bearing dynamic model for a gearbox test rig to study the interaction between the bearings and gears in the presence of faults. The model effectiveness in simulating faults of different sizes and locations was validated by the similarity between the simulated and measured signals subjected to a range of diagnostic techniques. Thereafter, according to the characteristic of double pulses in acceleration signals, the minimum entropy deconvolution method was used to separate the impulses from the entry into and exit from fault signals to acquire information on the fault size. Kong *et al.* [23] extracted the double impacts from bearing fault signals using a complex Morlet wavelet and ensemble empirical mode decomposition. Zhao *et al.* [24] reported the use of empirical mode decomposition and the approximate entropy method to acquire the double impact information, and eventually, the veracity of the method and the existence of the double impact phenomenon were validated by means of simulation and experimental results. Obviously, accurate quantitative analysis of bearing faults requires additional knowledge regarding the direction relationship between the feature extraction method and fault mechanism, as well as a deeper understanding of the vibration processes involved in the generation and transmission of defects with different severities.

The above researches and results demonstrate that the form of the double impact actually provides a potential concept for fault severity quantitative analysis. And matching pursuit is an effective method for extracting fault signal features. An adaptive randomized orthogonal matching pursuit method was proposed to extract bearing fault in [25]. Qin [26] proposed a method based on a novel impulsive wavelet to extract the impulse feature. Cui proposed a double impact theory to describe the vibration signals of bearings with a certain size of pitting faults in the race [27]. Moreover, a new compound dictionary of matching pursuit was presented.

The discussion analysis of each matched atom indicated substantial differences in the fault sizes measured in each matching atom, which required a fault-difference atom-selection method to improve the diagnosis accuracy. Further investigations demonstrated that most fault size misjudgments are caused by inaccurate extraction of the step components. When the double impact theory is used for quantitative fault diagnosis of fault, the initial times of the two responses should be accurate. Therefore, inaccurate extraction of the step components is the most likely reason for erroneous quantitative diagnoses. In order to avoid the erroneous selection of atoms and improve the accuracy of quantitative fault diagnosis, the matching pursuit algorithm based on a concatenation dictionary is proposed. In this paper, the concatenation dictionary is composed of two independent dictionaries, namely the step dictionary and impact dictionary. The impact dictionary is used as the higher level, and outputs the impact occurrence time as the input quantity entering the step dictionary, used as the lower level. The two dictionaries are concatenated by the information on the impact occurrence time. To enable accurate extraction of the impact components, they undergo only one iteration. The parameter information of the matched impact atom is stored, while the time information is input into the step dictionary as a known quantity. The subsequent iterative steps all proceed in the step dictionary. In this manner, the matching pursuit algorithm seeks step components.

The remainder of this paper is organized as follows: Section 2 focuses on the micro-motion analysis of the rolling elements from an acceleration perspective. Section 3 presents the concept of the concatenation dictionary and its detailed construction method. Moreover, the improved matching pursuit based on the concatenation dictionary of the fault severity assessment of rolling bearings is proposed. Section 4 mainly describes the analysis results of the simulation and experimental signals. Furthermore, the improvement in the diagnostic precision and calculation efficiency is clearly demonstrated.

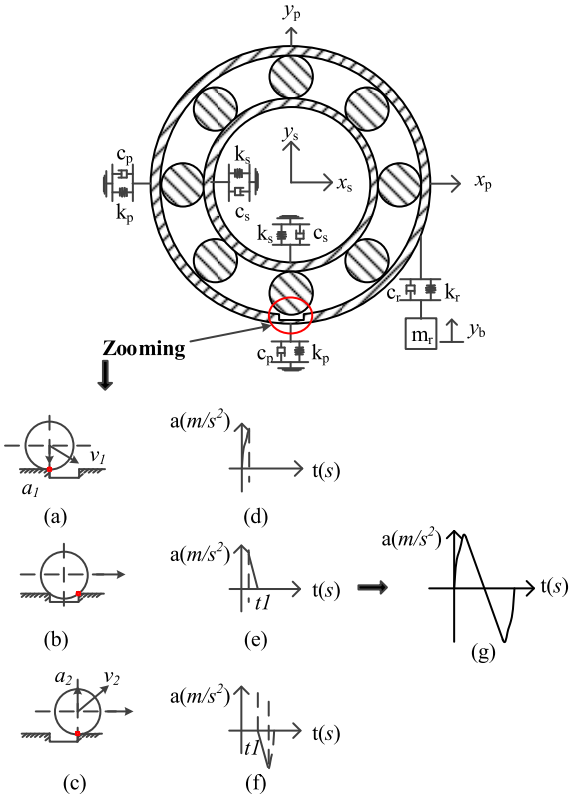

FIGURE 1. Schematic of micro-motion processes. (a) Enter fault. (b) Impact back edge. (c) Exit edge. (d) Response acceleration of (a). (e) Response acceleration of (b). (f) Response acceleration of (c). (g) Response acceleration of whole process.

II. ACCELERATION ANALYSIS OF ROLLING ELEMENT MICRO-MOTION

When the fault severity increases to a certain degree, the impulse generated from the fault is unlikely to be an ideal unit impulse, but will rather be a double-like impact action. In the case of the roller bearing outer race, the rolling element will experience certain processes, including entering the fault, impacting the back edge of the fault, and exiting from the fault, when the fault has a certain size. Each process will generate a type of signal. Moreover, the first and third processes require less energy than the second process, which will arouse system resonance. These processes are not independent, but rather coupled. Thus, the signal of a roller bearing with a certain fault size degree will be a result of superposition. In this paper, we discuss a fault with a length that is smaller than the rolling element diameter, and the rolling element does not touch the bottom of the fault. Figure 1 illustrates the micro-motion processes of the rolling element and the acceleration direction change.

Previous researches have improved the vibration signals of fault bearings, from a series of impact responses to step-like and impact responses, and the cause of 'double-like' responses has been addressed. In fact, the main reason for the different responses between entering and exiting the fault is the acceleration direction change of the rolling element, which is investigated in this paper.

When the rolling element maintains a uniform circular motion with the shaft prior to entering the fault, the normal force has reached a balance situation without normal acceleration. When the rolling element begins to enter the fault, which is regarded as the first process as illustrated in Figure 1(a), the sudden pressure release from the bearing leads to a normal outward force, and the rolling element acceleration direction changes to normal outward at this time. In Figure 1(a), a_1 represents this acceleration, and the direction of a_1 is regarded as the positive direction in artificial regulation. The sudden appearance of a_1 is a step-like response, as indicated in Figure 1(d). When the rolling element begins to impact the back edge of the fault at t_1, this is regarded as the second process, as illustrated in Figure 1(b). The normal force and normal acceleration return to zero and maintain a balance again. Figure 1(c) illustrates an opposite process for the rolling element motion to that in Figure 1(a). The normal force changes to the inward direction when the rolling element begins to exit from the fault, and an opposite direction a_2 is generated. The response of a_2 is similar to that of a_1, but with a different direction, as illustrated in Figure 1(f).

The above analysis has explained the generation mechanism of step-like responses when the rolling element enters and exits the fault, and the acceleration direction change is depicted in Figure 1(g). However, system resonance will occur when the rolling element impacts the back edge of the fault in a certain short period with a certain large energy. If the vibration is regarded as the oscillation of a mass-spring-damping system, the differential equation governing the phenomenon will be:

$$m\ddot{x} + c\dot{x} + kx = 0, \tag{1}$$

where m is the mass of the outer raceway plus the support structure, c is the viscous damping, k is the stiffness coefficient, and x is the displacement of the outer raceway.

Let $n = \frac{c}{2m}$, $w_n^2 = \frac{k}{m}$; thus, (1) can be transformed into (2):

$$\ddot{x} + 2n\dot{x} + w_n^2 x = 0. \tag{2}$$

As the damping ratio $\zeta = \frac{n}{w_n}$,

$$\ddot{x} + 2\zeta w_n \dot{x} + w_n^2 x = 0. \tag{3}$$

The roots of (3) are as below:

$$s_{1,2} = (-\zeta \pm \sqrt{\zeta^2 - 1})w_n. \tag{4}$$

When $\zeta < 1$, the value of the secular equation is as per (5):

$$s_{1,2} = -\zeta \cdot w_n \pm j\sqrt{1-\zeta}^2 \cdot w_n. \tag{5}$$

Letting $w_d = \sqrt{1-\zeta}^2 \cdot w_n$, the solution of (3) will be as follows:

$$\begin{aligned} x &= e^{-\zeta w_n t}\left(C_1 cos w_d t + C_2 sin w_d t\right) \\ &= Ae^{-\zeta w_n t}\cos\left(w_d t - \alpha\right) \\ &= Ae^{-\zeta w_n t}\sin\left(w_d t - \varphi\right). \end{aligned} \tag{6}$$

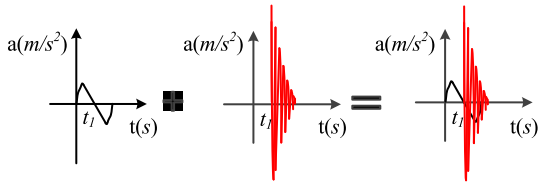

FIGURE 2. Diagrammatic sketch of step-impact response.

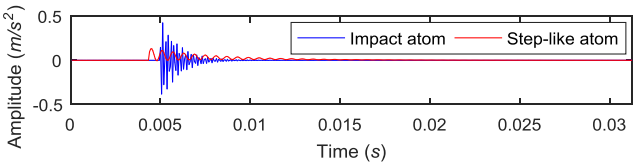

FIGURE 3. Simulated impact and step-like atoms.

Therefore, the system resonance response caused by the impact of the rolling element and fault back edge is exponential attenuation vibration.

In summary, the complete vibration responses when the rolling element passes through the defect are the superposition of the step-like and impact responses, as illustrated in Figure 2. The signal always exhibits a step-like response as well as an impact response, as the second step response is drowned in the high-energy impact response.

III. MATCHING PURSUIT ALGORITHM BASED ON CONCATENATION DICTIONARY

A. CONSTRUCTION OF CONCATENATION DICTIONARY

The concatenation dictionary consists of an impact dictionary as the higher level, and step-like dictionary as the lower level. The output time value of an atom extracted from the higher-level dictionary (HLD) is input into the lower level dictionary (LLD) as the concatenation model.

The HLD function is expressed as:

$$g_{imp}(u, \tau, f_n) = e^{\frac{-(t-u)}{\tau}} sin(2\pi f_n t). \tag{7}$$

The LLD function is expressed as:

$$g_{step}(u, \tau, f_n, \Delta t) = (e^{\frac{-(t-u-\Delta t)}{3\times\tau}} \times -cos\left(2\pi \times \left(\frac{f_n}{6}\right) \times t\right) + e^{\frac{-(t-u)}{5\times\tau}}), \tag{8}$$

where u is the impact occurrence time (s); τ is the systematic damping coefficient (s); f_n is the natural system frequency (Hz); and Δt is the time interval between the two responses.

The impact and step-like atoms are constructed by (7) and (8), respectively. Figure 3 illustrates an impact atom and a step-like atom with a length of 512 samples and initial time u of 0.005, respectively.

The HLD $G_{\text{imp}} = \{g_{1_i}, i = 1, 2, 3 \ldots m \ldots\}$ and LLD $G_{\text{imp}} = \{g_{1_i}, i = 1, 2, 3 \ldots m \ldots\}$ are constructed by the above atoms, where m denotes the dictionary size. There are three parameters, namely u, τ, and f_n in the HLD, and the value ranges of the three variables are determined by the signals to be detected. In this case, u is determined by the time scope of the detected signals, while τ and f_n are configured by the intrinsic system properties. Moreover, there are three parameters, namely τ, f_n, and Δt, in the LLD, where τ and f_n are configured in the same manner as in the HLD. It is not necessary to configure the range of u as it is the HLD output value, which can be used as the input parameter of the LLD directly. Finally, Δt is configured according to the fault size.

When the best-matching atom is selected in the HLD, the parameters of this atom will be returned. The impact occurrence time u is the input value for entering the LLD, as it is the intermediate value used for concatenation. Meanwhile, a search area $[u - \Delta t, u]$ is configured, where the best-matching step-like feature of the fault signal will be searched. Thus, the best-matching impact atom from the HLD is only selected once in the matching pursuit process. In contrast, the step-like atoms are selected based on the concatenation operation. In this manner, the extraction of step-like atoms can be avoided from the influence of the high-energy impact component, so the atom extraction accuracy is ensured. Furthermore, the calculation complexity is reduced and the error introduced by repeated extractions is avoided.

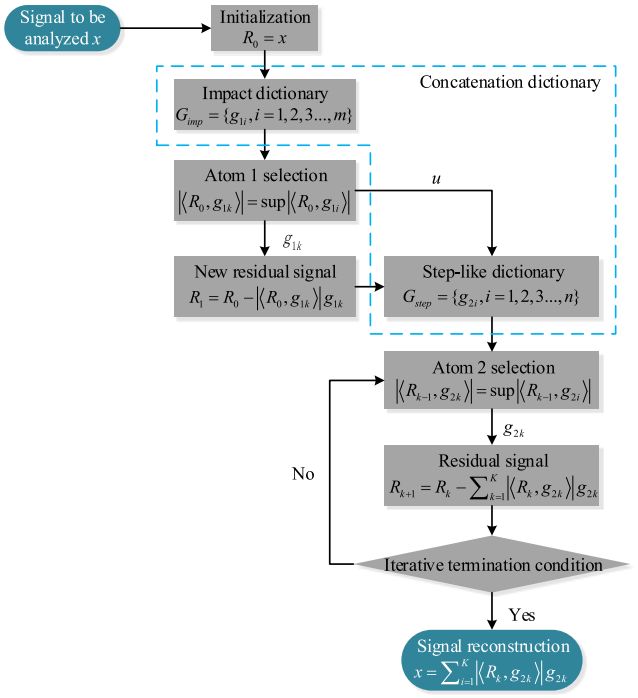

FIGURE 4. Flowchart of matching pursuit based on concatenation dictionary.

B. IMPROVED MATCHING PURSUIT ALGORITHM

An original atom selection mechanism is adopted in this novel method, although the main concept is still based on the matching pursuit. Figure 4 presents the flowchart of the matching pursuit algorithm based on the concatenation dictionary in this paper.

The steps for quantitative diagnosis using the improved matching pursuit algorithm based on the concatenation dictionary are as follows:

1) Initialization. The residual signal is initialized, and the signal to be analyzed $x(t)$ is assigned to the residual signal to obtain the initial residual signal $R_0 = x(t)$.

2) Atom matching in HLD. The optimal matched atom g_1k is selected from the HLD $G_{imp} = \{g_1, i = 1, 2, 3, \ldots, m, \ldots\}$. The optimal matched impact atom is selected according to (9), and the parameters u, τ, and f_n are stored. The new residual signal R_1 is obtained.

$$|\langle R_0, g_{1k}\rangle| = sup\left|\langle R_0, g_{1_i}\rangle\right| \tag{9}$$

$$R_1 = R_0 - |\langle R_0, g_{1k}\rangle| g_{1k} \tag{10}$$

3) Atom matching in LLD. The u in step 2) is used as an input value to enter the LLD $G_{step} = \{g_{2_i}, i = 1, 2, 3 \ldots m \ldots\}$. The search area Δu in the LLD is defined by (11). Then, the optimal matched step-like atom g_2k in the k-th iteration is selected according to (12).

$$\Delta u = [u - \Delta t, u) \tag{11}$$

$$|\langle R_{k-1}, g_{2k}\rangle| = sup\left|\langle R_{k-1}, g_{2_i}\rangle\right| \tag{12}$$

4) Updating of residual signal. Here, (13) is used to project the residual signal onto the optimal matched atom g_2k in the each iteration. Then, the residual signal is R_{k+1} after the k-th iteration, where K is the maximum number of iterations.

$$R_{k+1} = R_k - \sum_{k=1}^{K} \langle R_k, g_{2k}\rangle g_{2k} \tag{13}$$

5) Conditions for terminating iteration. The termination conditions include K and the residual energy ratio. If the conditions are satisfied, the iteration is terminated and step 6) is entered; otherwise, steps 3) to 5) are repeated.

6) Signal reconstruction. The reconstructed signal can be approximately expressed by (7).

$$x(t) = \sum_{i=1}^{k} \langle R_k, g_{2k}\rangle g_{2k} \tag{14}$$

7) Estimation of fault value. The occurrence times of the step-like response u_1 and impact response u_2 are marked on the time-domain waveform of the reconstructed signal. The time interval $\Delta t'$ is calculated by (15), and the fault value l' is estimated by (16), where f_r is the shaft rotation frequency, D_p is the bearing pitch diameter, and d is the rolling element diameter.

$$\Delta t' = u_2 - u_1 \tag{15}$$

8) Atom screening. The absolute deviation of the fault value reflected by the matched step atom in each iteration and estimated fault value l' is calculated. The atom with the minimum absolute deviation is selected, and the fault value reflected by this atom is taken as the secondary estimation value l'_g.

$$|\sigma|_{min} = min \left\| l_0 - l' \right\| \tag{16}$$

9) Quantitative diagnosis. The final fault size value l is the average estimated fault value and secondary estimation value.

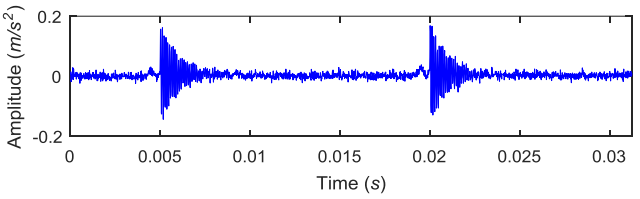

FIGURE 5. Simulated bearing signal with fault size of 1.2 *mm*.

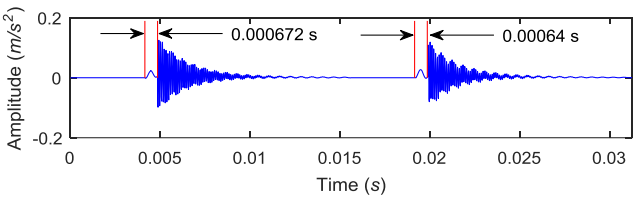

FIGURE 6. Reconstructed signal obtained by proposed method.

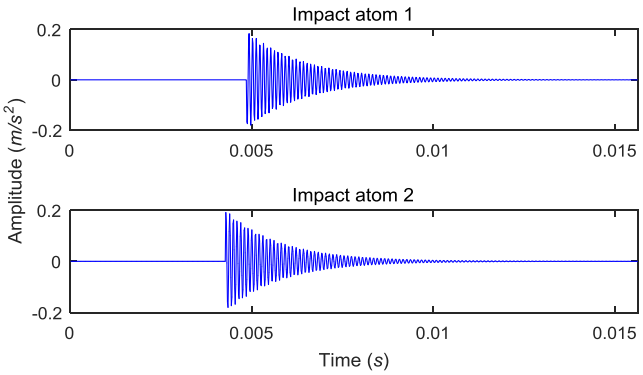

FIGURE 7. Impact atoms of two segment signals.

IV. SIMULATION AND EXPERIMENTAL SIGNAL ANALYSIS

A. ANALYSIS OF SIMULATION SIGNALS

Simulation signals with a length of 2048 samples are analyzed by the matching pursuit algorithm based on the concatenation dictionary, with an atom length equal to 1024 samples. The fault size is simulated as 1.2 *mm*. The data are truncated into two segments for calculation, as the signal to be analyzed should have the same length as the atom. The condition for terminating the iteration is set as the residual ratio threshold based on the attenuation coefficient, the value of which is taken as 0.6. In the entire calculation process, three step-like atoms are determined as the matched atoms for the first data segment, with two for the second data segment. The original time-domain waveforms of the simulation and reconstructed signals are illustrated in Figures 5 and 6, respectively. The time-domain waveforms of each matched impact and step-like atom (atoms 1, 2, 3, 4, and 5) are presented in Figures 7 and 8, respectively. The fault sizes calculated by each matched step-like atom are listed in Table 1.

Firstly, the fault size is estimated using the matching pursuit algorithm based on the concatenation dictionary, according to the time interval between the step-like and impact responses in the time-domain waveform of the reconstructed signal. By analyzing the time-domain waveform in Figure 6,

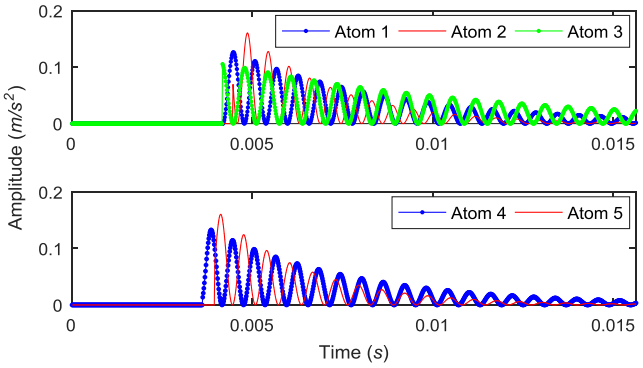

FIGURE 8. Step-like atoms of two segment signals.

TABLE 1. Fault sizes calculated by matched step-like atom.

Actual fault size (1.2 *mm*)	Atoms	Δt (s)	l_0 (*mm*)	$\lvert\sigma\rvert$ (*mm*)	Deviation
Data 1	Atom 1	0.00066	1.2095	0.0145	1.21%
	Atom 2	0.00041	0.7514	0.4436	36.9%
	Atom 3	0.00070	1.2828	0.0878	7.32%
Data 2	Atom 4	0.00067	1.2278	0.0328	2.73%
	Atom 5	0.00033	0.6048	0.5902	49.2%
Average deviation					19.5%

TABLE 2. Fault sizes calculated by reconstructed signal.

Actual fault size (1.2 *mm*)	Step 1	Impact 1	Step 2	Impact 2
Occurrence time (s)	0.004257	0.004929	0.01927	0.01991
$\Delta t'$ (s)		0.000672		0.00064
l' (*mm*)		1.22		1.17
Average fault size l' (*mm*)			1.195	
Deviation			0.42%	

TABLE 3. Results of final quantitative diagnosis.

Actual fault size (1.2 *mm*)	Atom 1	Atom 4	Average fault size l' (*mm*)
Estimated fault size l_0 (*mm*)	1.2095	1.2278	1.195
Average fault size of screened atoms l'_g (*mm*)	1.2187		
Final average fault size l (*mm*)	1.2069		

the estimated occurrence time and fault value are presented in Table 2. The estimated fault value can be calculated as 1.195 *mm*, which has a 0.42% deviation from the actual fault. Thereafter, the atoms are screened by the deviation $|\sigma|$ in Table 1. Atoms 1 and 4 are eventually screened, as each deviation is the smallest in each data segment. Finally, the secondary estimation value and final quantitative diagnosis result are achieved, as indicated in Table 3.

From Table 3, the final result of the fault quantitative diagnosis with the matching pursuit algorithm based on the concatenation dictionary has been obtained. The final average fault at 1.2069 *mm* is an accurate and reliable value, so the method in this paper is demonstrated to be effective. Although two erroneous atoms remain, namely atoms 2 and 5, the faults reflected by the other atoms only have deviations ranging

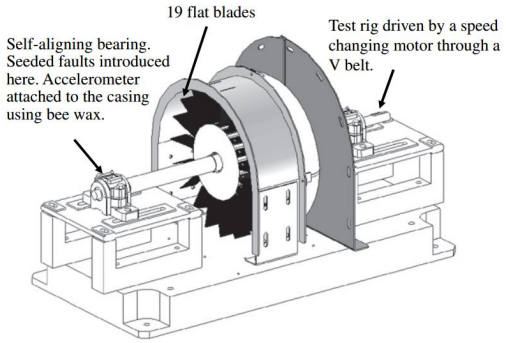

FIGURE 9. UNSW bladed test rig.

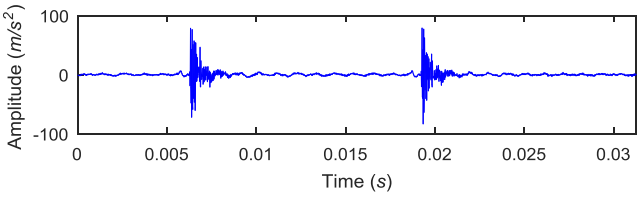

FIGURE 10. Measured bearing signal with fault size of 1.2 mm.

from 1.21% to 7.32%. The average deviation reflected by the step-like atoms is 19.5%. However, the fault values reflected by each atom fluctuate considerably, with the deviation ranging from 6.67% to 66.67%, and the average actual deviation is 32.5%, as calculated by the traditional matching pursuit method based on the step-impact dictionary. From the comparative analysis, it can be concluded that the method proposed in this paper can increase the atom selection accuracy, and offers superior overall precision in quantitative diagnosis.

B. ANALYSIS OF EXPERIMENTAL SIGNALS

The experimental signals were collected from the University of New South Wales (UNSW), Vibration and Acoustics Laboratory, School of Mechanical and Manufacturing Engineering. The UNSW bladed test rig is illustrated in Figure 9. The shaft is supported by two self- aligning, double row ball bearings. The bearing fault size remains 1.2 *mm*. An accelerometer was placed on top of the defective bearing. Further details regarding the test equipment and recorded measurement data are provided in [9]. The matching pursuit algorithm based on the concatenation dictionary is applied to the experimental signals. The condition for terminating the iteration is set as the residual ratio threshold based on the attenuation coefficient with a value of 0.6. In the entire calculation, two step-like atoms are matched for the first data segment, and two for the second data segment. The original waveforms and reconstructed signals of the experimental signals are illustrated in Figures 10 and 11, respectively. The time-domain waveforms of each matched impact atom are presented in Figure 12. The time-domain waveforms of each matched step-like atom (atoms 1, 2, 3, and 4) are illustrated in Figure 13. The parameters of atoms 1 to 4 are listed in Table 4.

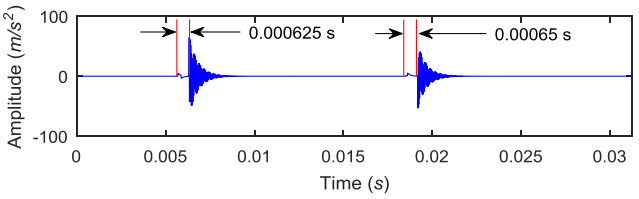

FIGURE 11. Reconstructed signal obtained by the proposed method.

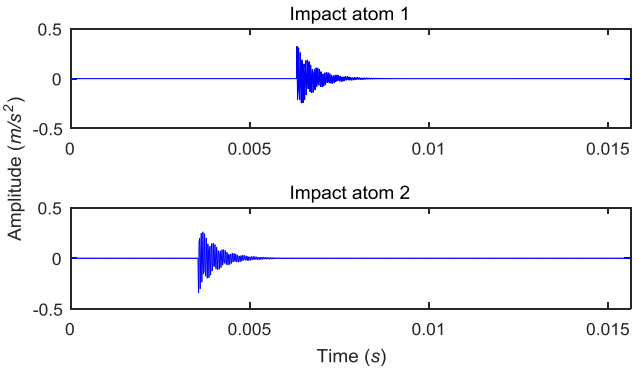

FIGURE 12. Impact atoms of two segment signals.

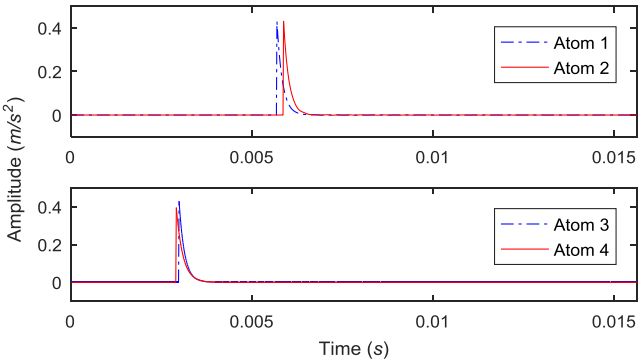

FIGURE 13. Step-like atoms of two segment signals.

TABLE 4. Fault sizes calculated by matched step-like atom.

Actual fault size (1.2 *mm*)	Atoms	Δt (s)	l_0 (*mm*)	$\lvert\sigma\rvert$ (*mm*)	Deviation
Data 1	Atom 1	0.00063	1.15	0.02	1.67%
	Atom 2	0.00045	0.82	0.35	29.17%
Data 2	Atom 3	0.00058	1.06	0.11	9.17%
	Atom 4	0.00066	1.21	0.04	3.33%
Average deviation					10.835%

From Figure 11, the reconstructed signal analyzed by the method in this paper exhibits a superior shape reconstruction effect in the time domain compared to the measured signal in Figure 10, and in particular, a small step-like response can be identified. The occurrence times of the step-like and impact responses in the time-domain waveform of the reconstructed signal can be read from Figure 11, and the estimated values are listed and calculated in Table 5. From Table 5, the fault value estimate is determined as 1.17 *mm*, with a deviation of 2.5% from the actual fault size.

TABLE 5. Fault sizes calculated by reconstructed signal.

Actual fault size (1.2 *mm*)	Step 1	Impact 1	Step 2	Impact 2
Occurrence time (*s*)	0.005692	0.006371	0.01862	0.01927
$\Delta t^{'}$ (*s*)	0.000625		0.00065	
l' (*mm*)	1.14		1.19	
Average fault size l' (*mm*)	1.17			
Deviation	2.5%			

TABLE 6. Results of final quantitative diagnosis.

Actual fault size (1.2 *mm*)	Atom 1	Atom 4	Average fault size l' (*mm*)
Estimated fault size l_0 (*mm*)	1.15	1.21	1.17
Average fault size of screened atoms l'_g (*mm*)	1.18		
Final average fault size l (*mm*)	1.175		

TABLE 7. Calculation time of each method.

Calculation time of matching pursuit algorithm based on step-impact dictionary (*s*)	39.4
Calculation time of matching pursuit algorithm based on concatenated dictionary (*s*)	28.1

According to the deviation $|\sigma|$ in Table 4, atoms 1 and 4 are eventually screened in order to reduce the diagnostic error. The secondary estimation values and final result of the quantitative diagnosis are presented in Table 6. It is demonstrated that the proposed method can basically achieve accurate quantitative diagnosis of bearing faults.

The advantage of the proposed matching pursuit algorithm based on the concatenation dictionary is that the atom selection accuracy is improved. The matched atoms extracted by the method proposed in this paper exhibit a smaller deviation from the actual fault in most cases, with an average overall deviation of 10.835%, as indicated in Table 4. Compared to the 27.57% deviation of the traditional method, the atom selection accuracy is improved by 16.735%. Moreover, the deviation of each atom is more accurate. In this paper, the deviation of the atoms is approximately 10% or below, except for the large deviation error atom, and the smallest deviation is 1.67%. However, the lowest deviation of the traditional method is 5.85%, and the deviation of the atoms is below 20%, with the largest deviation at 93.85%.

The main reason for the atom selection accuracy when using the method proposed in this paper is that the atom matching is only performed once in the HLD. Subsequent iterations by the matching pursuit algorithm are all performed in the step-like dictionary. One parameter is reduced in the concatenation dictionary, compared to the four parameters in the step-impact dictionary. Moreover, the proposed method not only improves the atom selection accuracy, but also reduces the calculation time. Therefore, the calculation efficiency is enhanced, as indicated in Table 7. In terms of the average deviation and deviation of each atom, the proposed

method exhibits higher atom selection accuracy than the traditional method. The calculation time is reduced, so the calculation efficiency is improved.

However, the proposed algorithm still exhibits certain limitations that need to be improved. The experimental signals used for verification have distinct features and low noise levels, making the step component extraction easier, with little error. However, most engineering signals contain substantial noise or exhibit coupling of various faults. In such cases, the proposed method cannot be executed accurately. The current study is confined to simulation and experimental data analyses. Furthermore, the matching pursuit algorithm must perform a large amount of inner product calculation in the selection of most matched atoms. In order to improve the calculation efficiency and save on computer memory, an intelligent algorithm should be used for atom selection. Therefore, improvement of the algorithm stability is one topic for future research.

V. CONCLUSION

This paper expands on the current generation mechanisms of fault vibration signals by investigating the vibration source and micro-motion acceleration near local defects. With the aim of obtaining atom selection accuracy and improving the computational efficiency in the matching pursuit algorithm, an improved algorithm based on the concatenation dictionary has been proposed.

The novel concatenation dictionary is presented with an impact dictionary as the HLD and step-like dictionary as the LLD. The iterative procedure is performed only once in the HLD. The parameters of each extracted impact atom are stored, and the impact occurrence time is input into the LLD as known information. The main iterative procedures are performed on the step dictionary to match the step-like atoms repeatedly in the entire matching pursuit algorithm. The advantages of the concatenation dictionary mechanism are mainly the step-like atom selection accuracy and eliminating the effects of the high-energy impact components. Moreover, the fault deviation reflected by each atom decreases obviously compared to the actual fault.

The analysis of the simulation and experimental signals of the bearing fault based on the proposed method has demonstrated improved atom selection accuracy and fault diagnosis precision.

ACKNOWLEDGMENT

The authors would like to thank Prof. Randall and Dr. Sawalhi for their theoretical guidance, and the data provided from the Vibration and Acoustics Laboratory of UNSW.

REFERENCES

[1] Z. Tong, W. Li, B. Zhang, F. Jiang, and G. B. Zhou, "Bearing fault diagnosis under variable working conditions based on domain adaptation using feature transfer learning," *IEEE Access*, vol. 6, pp. 76187–76197, Nov. 2018.

[2] L. Song, H. Wang, and P. Chen, "Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 8, pp. 1887–1899, Aug. 2018.

[3] L. Cui, X. Wang, Y. Xu, H. Jiang, and J. Zhou, "A novel switching unscented Kalman filter method for remaining useful life prediction of rolling bearing," *Measurement*, vol. 135, pp. 678–684, Mar. 2019.

[4] M. J. Dowling, "Application of non-stationary analysis to machinery monitoring," in *Proc. ICASSP*, Minneapolis, MN, USA, Apr. 1993, pp. 59–62.

[5] I. K. Epps, "An investigation into vibrations excited by discrete faults in rolling element bearings," Ph.D. dissertation, Dept. Mech. Eng., Univ. Canterbury Mech. Eng., Christchurch, New Zealand, 1991.

[6] S. Sassi, B. Badri, and M. Thomas, "A numerical model to predict damaged bearing vibrations," *J. Vib. Control*, vol. 13, no. 11, pp. 1603–1628, Nov. 2007.

[7] J. Sopanen and A. Mikkola, "Dynamic model of a deep-groove ball bearing including localized and distributed defects. Part 1: Theory," *Proc. Inst. Mech. Eng. K, J. Multi-Body Dyn.*, vol. 217, no. 3, pp. 201–211, Sep. 2003.

[8] J. Sopanen and A. Mikkola, "Dynamic model of a deep-groove ball bearing including localized and distributed defects. Part 2: Implementation and results," *Proc. Inst. Mech. Eng. K, J. Multi-Body Dyn.*, vol. 217, no. 3, pp. 213–223, Sep. 2003.

[9] N. Sawalhi and R. B. Randall, "Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size," *Mech. Syst. Signal Process.*, vol. 25, no. 3, pp. 846–870, Apr. 2011.

[10] M. S. Patil, J. Mathew, P. K. Rajendrakumar, and S. Desai, "A theoretical model to predict the effect of localized defect on vibrations associated with ball bearing," *Int. J. Mech. Sci.*, vol. 52, no. 9, pp. 1193–1201, Sep. 2010.

[11] L. Cui, J. Huang, F. Zhang, and F. Chu, "HVSRMS localization formula and localization law: Localization diagnosis of a ball bearing outer ring fault," *Mech. Syst. Signal Process.*, vol. 120, pp. 608–629, Apr. 2019.

[12] S. Singh, U. G. Köpke, C. Q. Howard, and D. Petersen, "Analyses of contact forces and vibration response for a defective rolling element bearing using an explicit dynamics finite element model," *J. Sound Vib.*, vol. 333, no. 21, pp. 5356–5377, Oct. 2014.

[13] A. Rafsanjani, Sa. Abbasion, A. Farshidian and H. Moeenfard, "Nonlinear dynamic modeling of surface defects in rolling element bearing systems," *J. Sound Vib.*, vol. 319, nos. 3–5, pp. 1150–1174, Jan. 2009.

[14] A. M. Ahmadi, D. Petersen, and C. Howard, "A nonlinear dynamic vibration model of defective bearings—The importance of modelling the finite size of rolling elements," *Mech. Syst. Signal Process.*, vols. 52–53, pp. 309–326, Feb. 2015.

[15] L. Song, H. Wang, and P. Chen, "Step-by-step fuzzy diagnosis method for equipment based on symptom extraction and trivalent logic fuzzy diagnosis theory," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 6, pp. 3467–3478, Dec. 2018.

[16] H. Wang, P. Wang, L. Song, B. Ren, and L. Cui, "A novel feature enhancement method based on improved constraint model of online dictionary learning," *IEEE Access*, vol. 7, pp. 17599–17607, Jan. 2019.

[17] N. Baydar and A. Ball, "A comparative study of acoustic and vibration signals in detection of gear failures using Wigner–Ville distribution," *Mech. Syst. Signal Process.*, vol. 15, no. 6, pp. 1091–1107, Nov. 2001.

[18] C. Shen, F. Hu, F. Liu, A. Zhang, and F. Kong, "Quantitative recognition of rolling element bearing fault through an intelligent model based on support vector regression," in *Proc. ICICIP*, vol. 13, Jun. 2013, pp. 842–847.

[19] K. Jiang, G. Xu, L. Liang, G. Zhao, and T. Tao, "A quantitative diagnosis method for rolling element bearing using signal complexity and morphology filtering," *J. Vibroeng.*, vol. 14, no. 4, pp. 1862–1875, Dec. 2012.

[20] H. Hong and M. Liang, "Fault severity assessment for rolling element bearings using the Lempel–Ziv complexity and continuous wavelet transform," *J. Sound Vib.*, vol. 320, nos. 1–2, pp. 452–468, Feb. 2009.

[21] N. Sawalhi and R. B. Randall, "Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults," *Mech. Syst. Signal Process.*, vol. 22, no. 8, pp. 1924–1951, Nov. 2008.

[22] R. B. Randall and N. Sawalhi, "Signal processing tools for tracking the size of a spall in a rolling element bearing," in *Proc. IUTAMSETRD*, New Delhi, India, 2011, pp. 429–440.

[23] Y. B. Kong, Y. Guo, and X. Wu, "Double impulses extraction of faulty rolling element bearing based on EEMD and complex Morlet wavelet," *Adv. Mater. Res.*, vol. 904, pp. 437–441, Mar. 2014.

[24] S. Zhao, L. Liang, G. Xu, J. Wang, and W. Zhang, "Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method," *Mech. Syst. Signal Process.*, vol. 40, no. 1, pp. 154–177, Oct. 2013.

[25] J. Li, M. Li, X. Yao, and H. Wang, "An adaptive randomized orthogonal matching pursuit algorithm with sliding window for rolling bearing fault diagnosis," *IEEE Access*, vol. 6, pp. 41107–41117, Jul. 2018.

[26] Y. Qin, "A new family of model-based impulsive wavelets and their sparse representation for rolling bearing fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2716–2726, Mar. 2018.
[27] L. Cui, J. Wang, and S. Lee, "Matching pursuit of an adaptive impulse dictionary for bearing fault diagnosis," *J. Sound Vib.*, vol. 333, no. 10, pp. 2840–2862, May 2014.

LINGLI CUI was born in Jiamusi, China. She received the B.S. degree in mechanical engineering from Shenyang Aerospace University, Shenyang, China, in 1998, the M.S. degree in mechanical engineering and automation from the Harbin Institute of Technology, Harbin, China, in 2001, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2004. She is currently a Professor of mechanical engineering with the Beijing University of Technology, Beijing. Her research interests include fault mechanisms, pattern recognition, intelligent diagnosis, and fault diagnosis.

XIN WANG was born in Zhangjiakou, China. He received the B.S. degree in mechanical engineering from the Xi'an University of Architecture and Technology, Xi'an, China, in 2017. He is currently pursuing the Ph.D. degree in mechanical engineering with the Beijing University of Technology, Beijing, China. His research interests include remaining useful life prediction and the fault diagnosis of rolling element bearings.

HUAQING WANG received the B.S. and M.S. degrees from the College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, China, in 1995 and 2002, respectively, and the Ph.D. degree from Mie University, Japan, in 2009. He is currently a Professor with the Beijing University of Chemical Technology. His research interests include intelligent diagnostics for plant machinery and signal processing.

NA WU is currently pursuing the Ph.D. degree in mechanical engineering with the Beijing University of Technology. Her research interests include the signal processing based on the matching pursuit algorithm and gear/bearing fault diagnosis, resistance spot welding, and process monitor/control.