

# 1228. A bearing fault detection method based on compressive measurements of vibration signal

Xinpeng Zhang<sup>1</sup>, Niaoqing Hu<sup>2</sup>, Guojun Qin<sup>3</sup>, Zhe Cheng<sup>4</sup>, Hua Zhong<sup>5</sup>

Science and Technology on Integrated Logistics Support Laboratory

National University of Defense Technology, Changsha, 410073, P. R. China

College of Mechatronic Engineering and Automation, National University of Defense Technology

Changsha, 410073, P. R. China

<sup>2</sup>Corresponding author

**E-mail:** <sup>1</sup>zhang\_xinpeng@126.com, <sup>2</sup>huniaoqing@126.com, <sup>3</sup>qgj@nudt.edu.cn, <sup>4</sup>chengzhe@nudt.edu.cn,

<sup>5</sup>zhonghua\_nudt@foxmail.com

(Received 6 January 2014; received in revised form 3 March 2014; accepted 19 March 2014)

**Abstract.** The general method for bearing fault detection is achieved by using bearing vibration signals which sampled in the frame of Shannon sampling theory. So it is necessary to sample and save abundant original vibration data in the process of uninterrupted monitoring, and this will generate masses of original data which would burden the storage and transmission. For this issue, a fault detection method based on compressed sensing theory is proposed in this paper. It only needs to sample and save fewer compressive measurements of bearing vibration signal directly compared to original signal. There is no need to recover the original signal accurately for detecting bearing faults, while it just requires referring to the prior training result and reconstructing the overall energy distribution of the original signal in some transform domain. The availability and effectiveness of the method proposed is validated with bearing vibration signals sampled in practice.

**Keywords:** compressed sensing, bearing fault detection, vibration signal, compressive measurements, dictionary matrix.

## 1. Introduction

In the field of condition monitoring and fault diagnosis of rotating machinery, it is always necessary to monitor the component uninterruptedly for a long period of time. This process will generate masses of original data which would burden the storage and transmission. For the storage and transmission issue of vast amounts of vibration data, the general strategy can be described as follows. First we should sample the analog signal in the frame of Shannon sampling theory and acquire the original digital signal. Data can be compressed in appropriate ways such as space transformation and then we can acquire small amount of compressed data [1]. What we need to save and transmit should be the compressed data. In data analysis, we can decode the compressed data we saved to recover the original vibration data and implement fault detection and diagnosis etc. with relevant available methods.

In the process of signal sampling and compressing mentioned above, what we transmitted or saved should be a small part of the original data we sampled. The majority of the original data would be dropped, which should be an enormous waste of resources. Since the signal can be compressed, we should consider whether we can acquire the compressed data directly and omit much data contained little useful information at the same time. This idea is the rudiment of compressed sensing (CS). In 2006, Candès proved in mathematic principle that the original signal could be reconstructed using parts of its Fourier transform coefficients, which should be the theoretical foundation of compressed sensing [2]. Then Donoho and Candès et al proposed the concept of compressed sensing formally based on the related work [3, 4]. The main process of CS can be divided into two steps. First, we can acquire the non-adaptive linear projections (or measurements) of the original signal in compressed sampling way. Then we can reconstruct the original signal directly with these measurements in appropriate recovery algorithms [5-7]. Taking one-dimensional signal for example, the comparison of traditional signal compression with

compressed sensing is shown in Fig. 1.

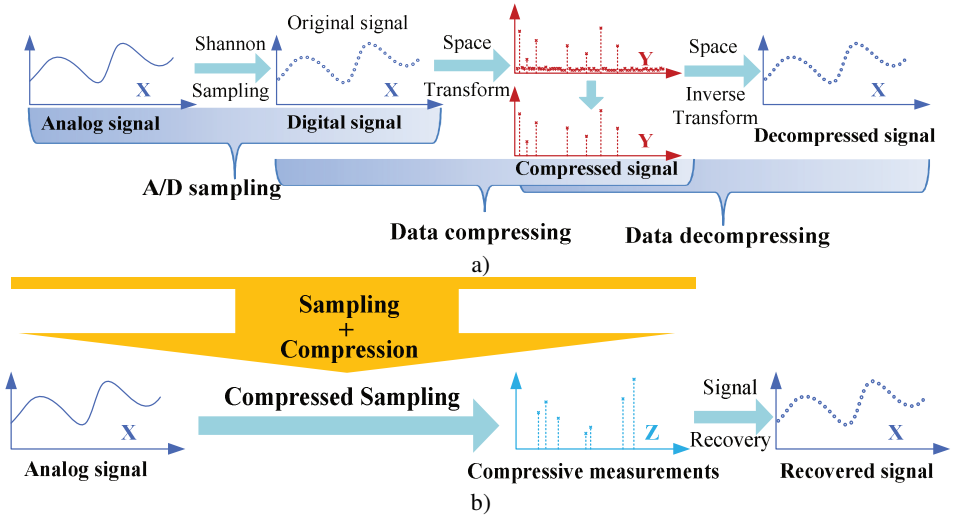


Fig. 1. Comparison of traditional signal compression a) with compressed sensing b)

In the frame of compressed sensing, the original signal should be reconstructed first, and then fault can be detected with the recovered signal. Since the original signal can be recovered from the compressive measurements, then we should believe that these measurements should carry or contain enough information of the original signal. If the faults can be detected or diagnosed with the original signal, then the detection and diagnosis should also be achieved with the compressive measurements directly [8].

The bearing is one of the most commonly used but also the most vulnerable parts of rotating machinery equipment. It is also a failure-prone component due to the abominable working condition in addition to high speed and heavy load. Once failure, it will threaten the safe operation of the equipment, so it is necessary to monitor the bearings status and identify the fault in time [9]. The bearing faults can be characterized by the existence of shock pulses in vibration signals. In the time domain, the mean value and variance etc. of the vibration signal will change owing to the shock. While in the frequency domain, the shocks will enhance the high-frequency parts so that the energy distribution of the signal changes. According to these characteristics, we can achieve bearing fault detection based on the changes of the energy distribution in frequency domain. If we can collect fewer compressive measurements directly which contain the energy distribution information of current signal, then it is possible to achieve bearing fault detection using these measurements. Therefore, a new method focused on how to achieve bearing fault detection using small amount of compressive measurements directly is proposed in this paper.

The rest of this paper is organized as follows. In Section 2, we introduce the theory of compressed sensing. The bearing fault detection method based on compressed sensing is proposed in Section 3. The experimental tests are given in Section 4. Finally, the conclusions are drawn in Section 5.

## 2. Compressed sensing

For a signal  $\mathbf{x} \in \mathbf{R}^N$ , as described above, in the frame of CS we should acquire the linear projections of signal  $\mathbf{x}$  first. The process of projecting can be converted into an observation matrix  $\Phi: M \times N$  ( $M < N$ ), and each row of matrix  $\Phi$  can be regarded as a sensor that multiplies with the signal and acquires parts of information of the signal. We can execute compressive measuring to  $\mathbf{x}$  as:

$$\mathbf{y} = \Phi \mathbf{x}. \tag{1}$$

Then we can acquire the linear measurements (or projections)  $\mathbf{y} \in \mathbf{R}^M$ . If  $\mathbf{x}$  can be recovered from  $\mathbf{y}$ , it means that these fewer linear projections contain enough information to recover the signal  $\mathbf{x}$ , in which case the compressed sensing can be achieved. According to linear algebra theory, Eq. (1) should have infinitely many solutions and we can't recover original signal  $\mathbf{x}$  uniquely from measurements  $\mathbf{y}$  when  $M < N$ . However, if  $\mathbf{x}$  is sparse (meaning that there only has a few non-zero elements in  $\mathbf{x}$ ), the number of unknowns will be reduced greatly, and this will make it possible to recover  $\mathbf{x}$  from  $\mathbf{y}$ .

Actually, most nature signals are not sparse in general, while they can be represented sparsely with proper ways such as orthogonal transformation. If expanding  $\mathbf{x} \in \mathbf{R}^N$  on some orthogonal basis  $\{\boldsymbol{\psi}_i\}_{i=1}^N$ , where  $\boldsymbol{\psi}_i$  is a  $N$ -dimensional column vector, we can get:

$$\mathbf{x} = \sum_{i=1}^N \theta_i \boldsymbol{\psi}_i, \tag{2}$$

where the coefficient  $\theta_i = \langle \mathbf{x}, \boldsymbol{\psi}_i \rangle = \boldsymbol{\psi}_i^T \mathbf{x}$ , and Eq. (2) can be transferred into a matrix form as:

$$\mathbf{x} = \Psi \boldsymbol{\theta}, \tag{3}$$

where  $\Psi = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_N] \in \mathbf{R}^{N \times N}$  is defined as dictionary matrix, and expansion coefficient is  $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_N]^T$ . Suppose that the coefficient vector  $\boldsymbol{\theta}$  is  $K$ -sparse on basis  $\Psi$ , i.e. there has  $K$  non-zero elements in  $\boldsymbol{\theta}$  and  $K \ll N$ , then  $\boldsymbol{\theta}$  can be entitled as sparse representation coefficient of  $\mathbf{x}$  on dictionary matrix  $\Psi$ . Taking Eq. (3) into Eq. (1) and denoting  $\mathbf{A} = \Phi \Psi$ , then we can get:

$$\mathbf{y} = \Phi \Psi \boldsymbol{\theta} = \mathbf{A} \boldsymbol{\theta}. \tag{4}$$

The matrix form of compressive measurements can be represented in Fig. 2. Considering that  $\boldsymbol{\theta}$  is sparse, the unknowns in Eq. (4) will be reduced greatly so that it is possible to recover  $\boldsymbol{\theta}$  from  $\mathbf{y}$ . In order to recover sparse signal, Candès and Tao presented and also proved that the matrix  $\mathbf{A}$  mentioned above must satisfy Restricted Isometry Property (RIP) [7, 10]. Then Baraniuk proposed the idea that the irrelevance between observation matrix  $\Phi$  and dictionary matrix  $\Psi$  was equivalent to RIP [11]. Once these conditions have been satisfied, then the original signal  $\mathbf{x}$  can be recovered according to Eq. (3) and Eq. (4).

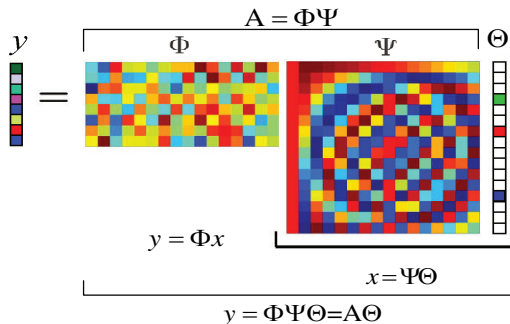
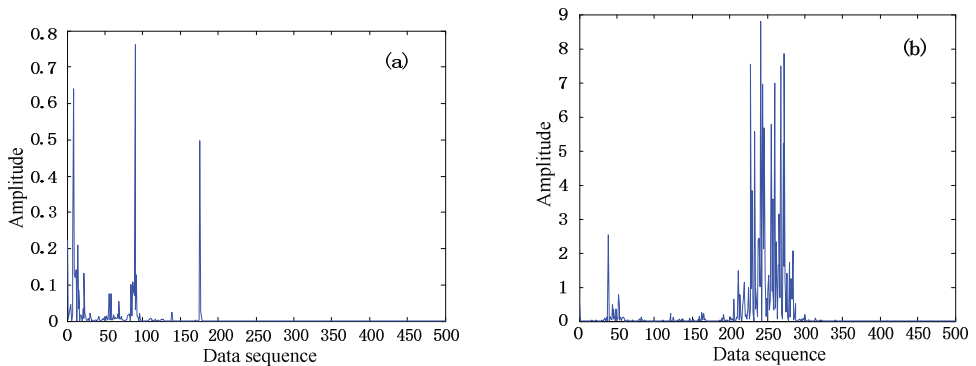


Fig. 2. Matrix form of compressive measurements

For vibration signal  $\mathbf{x} \in \mathbf{R}^N$ , our research will mainly focus on how to achieve bearing fault detection based on compressed sensing just using the compressive measurements  $\mathbf{y} \in \mathbf{R}^M$  ( $M < N$ ) directly.

### 3. Bearing fault detection method

A practical bearing vibration signal (data source from [12], 6205-2RS JEK SKF deep groove ball bearing, with 12 K sampling frequency, 1797 r/min and 1 HP load) in time domain can be denoted as  $\mathbf{x} \in \mathbf{R}^N$  ( $N = 1000$ ). Using the discrete Fourier transform matrix as dictionary matrix  $\Psi$  and according to  $\Theta = \Psi\mathbf{x}$ , we can acquire the sparse representation coefficient  $\Theta$ . Then the distribution of the elements in  $\Theta$  can reflect the energy distribution of the elements in signal  $\mathbf{x}$  in frequency domain. The bearing vibration signals corresponding to different states should have different  $\Theta$ . Considering  $\Theta$  is a complex vector, we calculate the square values of modulus of each element in  $\Theta$  and constitute vector  $\tilde{\Theta}$  in token of the energy distribution in  $\Theta$ , as shown in Fig. 3 (due to the symmetry of  $\tilde{\Theta}$ , we just plotted the half of the vector  $\tilde{\Theta}$  from 1 to 500 in data sequence). It can be seen from Fig. 3 that the distributions corresponding to the two different states (normal state and inner race state) differ obviously. Actually, the states of outer race fault and rolling ball fault also have the similar differences from the normal state. In general, the energy distributions corresponding to the same type of bearing state should be similar and invariable, which can be used to achieve fault detection.



**Fig. 3.** The distributions of  $\tilde{\Theta}$  corresponding to different bearing states, a) corresponding to normal state and b) corresponding to inner race fault state (the fault is single point, which is introduced to the bearing using electro-discharge machining with fault diameter of 0.021 inches and fault depth of 0.011 inches)

Compared to the original signal  $\mathbf{x} \in \mathbf{R}^N$ , we can acquire fewer compressive measurements  $\mathbf{y} \in \mathbf{R}^M$  based on compressed sensing. According to the analysis previously, if we can estimate the distribution of  $\Theta$  in Eq. (4) with these fewer measurements directly, then bearing fault detection would be achieved. The energy distribution in  $\Theta$  can be described by the intervals which covered the elements with bigger modulus values in  $\Theta$ . So the key of the proposed method is how to estimate the energy distribution of the elements in  $\Theta$  with measurements  $\mathbf{y}$  directly. In the next several sections, we will introduce this fault detection method based on compressive measurements in detail.

#### 3.1. Intervals learning

As we discussed previously, after estimating the elements with bigger modulus values in  $\Theta$ , then we need to set a standard to determine the bearing states based on the estimated elements. Our strategy is as follows: sampling enough normal state data and fault state data first in the same operation condition; calculating  $\Theta$  of these data by discrete Fourier transform; and then determining the interval  $\mathcal{T}$  corresponding to normal state and interval  $\mathcal{F}$  corresponding to fault states with the information of the elements in  $\Theta$ . The two intervals should not be overlapped, in which case the normal state and fault state can be separated distinctly, i.e.  $\mathcal{T} \cap \mathcal{F} = \emptyset$ . Considering that the state of a bearing should be either normal or fault, so the union of interval  $\mathcal{T}$  and  $\mathcal{F}$  has to

cover the whole range  $\mathcal{Z}$ , i.e.  $\mathcal{T} \cup \mathcal{F} = \mathcal{Z}$ . Due to the diversity of the fault patterns, it is difficult to determine the fault interval  $\mathcal{F}$  directly. While the energy distribution corresponding to normal state in frequency domain should be more invariable and steady, so we can determine  $\mathcal{T}$  first, and then acquire the interval  $\mathcal{F}$  by calculating the complementary set of interval  $\mathcal{T}$  in the whole range  $\mathcal{Z}$ . After  $\mathcal{F}$  and  $\mathcal{T}$  being confirmed, we can recognize the bearing state according to the estimated elements with bigger modulus value in  $\Theta$ . If these elements with bigger modulus values in  $\Theta$  fall into the interval  $\mathcal{T}$ , then we can determine that the bearing should be normal, while if they fall into the interval  $\mathcal{F}$ , then the bearing should be in fault state. The detailed determining method will be introduced in Section 3.3. The steps of learning intervals are as follows:

(1) Acquiring normal state samples to form learning set  $\Gamma$ , confirming the whole range  $\mathcal{Z}$  according to the samples;

(2) Calculating the average vector  $\Theta_n$  of all the coefficient vectors  $\Theta$  corresponding to all the learning samples in set  $\Gamma$  with discrete Fourier transform;

(3) Selecting the elements in  $\Theta_n$  from the one with largest modulus value to the one with smallest modulus value. Stopping the selecting until the ratio of the energy corresponding to all the elements we have selected to the energy corresponding to all elements in  $\Theta_n$  is greater than some threshold  $\varepsilon$ , where  $0 \leq \varepsilon \leq 1$ . We set  $\varepsilon$  as 0.9 in general, in which case we think that the elements we selected can represent the energy distribution of all elements in  $\Theta_n$ . Calculating the minimal interval constituted by least continuous subintervals that can cover all the elements we selected above. This minimal interval will be taken as the interval  $\mathcal{T}$  corresponding to the normal state.

(4) Calculating the complementary set of interval  $\mathcal{T}$  in the whole range  $\mathcal{Z}$  and taking it as interval  $\mathcal{F}$  corresponding to the fault states.

### 3.2. Recovering sparse representation coefficient

Considering the compressed sensing theory, if the representation coefficient  $\Theta$  is not sparse enough or the number of measurements can't satisfy the general conditions for reconstructing signal  $\mathbf{x}$ , then we can't recover the sparse representation coefficient vector  $\Theta$  accurately. However, for fault detection, actually we don't need to recover all the elements in  $\Theta$  accurately. We just need to estimate the primary energy distribution of the elements in  $\Theta$  in our proposed method, and then fault detection can be achieved using the information of the elements with bigger modulus values. The idea of matching pursuit (MP) algorithm can be used to find out this information. In MP algorithm, we choose the column of the measurement matrix  $\mathbf{A}$  by iterative way [13, 14]. First we choose the column which has the highest correlation with current residual error and calculate the approximate solution in current iteration and the residual error of the next iteration. This process will be repeated until the number of iterations surpasses the limit we set or the current residual is below the given error. We propose a method to estimate the elements with bigger modulus values in  $\Theta$  (we entitle these elements as support points) based on MP algorithm. The steps of the method are as follows:

Input: compressive measurements  $\mathbf{y}$ , measurement matrix  $\mathbf{A}$  and the number of support points  $k$ .

Initialization:  $\Theta^{[0]} = \mathbf{0}$ ,  $\mathbf{r}^{[0]} = \mathbf{y}$ .

For  $i = 1; i = i + 1; i = k$ :

Step 1: Calculate the inner product,  $\mathbf{g}^{[i]} = \mathbf{A}^T \mathbf{r}^{[i-1]}$ ;

Step 2: Locate the most important element in  $\mathbf{g}^{[i]}$ , viz.  $j^{[i]} = \arg \max_j |\mathbf{g}_j^{[i]}| / \|\mathbf{A}_j\|_2$ . Then save this location  $\mathbf{S}(i) = j^{[i]}$ .

Step 3: Update the solution  $\Theta_{j^{[i]}}^{[i]} = \Theta_{j^{[i]}}^{[i-1]} + \mathbf{g}_{j^{[i]}}^{[i]} / \|\mathbf{A}_{j^{[i]}}\|_2^2$ ;

Step 4: Update the residual error  $\mathbf{r}^{[i]} = \mathbf{r}^{[i-1]} - \mathbf{A}_{j^{[i]}} \mathbf{g}_{j^{[i]}}^{[i]} / \|\mathbf{A}_{j^{[i]}}\|_2^2$ ;

End For.

Output: the support set  $\mathbf{S}$ , the estimated  $\Theta^{[i]}$ .

The number of the support points mentioned above can be set according to the number of the elements selected in step (3) of intervals learning in Section 3.1. The sparse representation coefficient vector  $\Theta$  and support set  $\mathbf{S}$  have been determined, and then we will introduce how to use these results to estimate the current state of the bearing.

### 3.3. State recognition

We have determined the interval  $\mathcal{T}$  and  $\mathcal{F}$  corresponding to normal state and fault state respectively, and estimated the sparse representation coefficient vector  $\Theta$  and support set  $\mathbf{S}$ . In ideal conditions, all the elements in  $\Theta$  corresponding to support set  $\mathbf{S}$  for normal state should fall into the interval  $\mathcal{T}$ . While for fault state, the elements should be in the interval  $\mathcal{F}$ . However, in practice it will often happen that a majority of support points fall into some interval, while other points would be in another interval due to the effect of noise or other factors. Therefore, we need to relax the condition for recognition. The current state should be determined as normal when the great mass of support points (90 % for instance) fall into the interval  $\mathcal{T}$ , otherwise, should be fault state. It can be seen from the process for estimating support set in Section 3.2 that the earlier the support point has been picked, the more important the picked point would be. Meanwhile, generally speaking, the picked points should have bigger modulus values than other points. Therefore, we can define the energy of each support point in  $\mathbf{S}$  as:

$$\mathbf{E}(i) = \|\Theta_{S(i)}\|_2^2 \cdot e^{-\alpha i}, \quad i = 1, \dots, k, \quad 0 < \alpha < 1, \quad (5)$$

where  $S(i)$  indicates the location of  $i$ th support point in sparse representation coefficient vector  $\Theta$ , and  $\Theta_{S(i)}$  indicates the  $S(i)$ -th element estimated in  $\Theta$ . The parameter  $\alpha$  can be called energy factor, which is related to the energy distribution of the support points. Broadly speaking, if the elements in  $\Theta$  have close modulus, and then we can choose a smaller  $\alpha$ , while if the modulus values of the elements in  $\Theta$  fluctuate dramatically, then we should set a bigger  $\alpha$ . We define  $\mathbf{Er}(j)$  to denote the ratio of the energy of the  $j$ th support point to the gross energy of all support points in  $\mathbf{S}$ :

$$\mathbf{Er}(j) = \frac{\|\Theta_{S(j)}\|_2^2 \cdot e^{-\alpha j}}{\sum_{i=1}^k \|\Theta_{S(i)}\|_2^2 \cdot e^{-\alpha i}}, \quad j \in \{1, \dots, k\}, \quad 0 < \alpha < 1. \quad (6)$$

Then we define  $\mathbf{Er}^{\mathcal{F}}$  as the ratio of the energy of the support points falling into interval  $\mathcal{F}$  to the gross energy of all support points in  $\mathbf{S}$ :

$$\mathbf{Er}^{\mathcal{F}} = \frac{\sum_{S(j) \in \mathcal{F}} \|\Theta_{S(j)}\|_2^2 \cdot e^{-\alpha j}}{\sum_{i=1}^k \|\Theta_{S(i)}\|_2^2 \cdot e^{-\alpha i}}, \quad j \in \{1, \dots, k\}, \quad 0 < \alpha < 1. \quad (7)$$

We set threshold  $\delta$  to characterize the relative size of  $\mathbf{Er}^{\mathcal{F}}$ , and then compare  $\mathbf{Er}^{\mathcal{F}}$  with threshold  $\delta$ . If  $\mathbf{Er}^{\mathcal{F}}$  is greater than or equal to  $\delta$ , the current state will be recognized as fault state, while if  $\mathbf{Er}^{\mathcal{F}}$  is smaller than  $\delta$ , then the state would be normal.

Compared to the traditional methods, the advantage of the fault detection method based on compressive measurements is that we just need to sample fewer measurements directly. Moreover, this method can reduce the calculation amount greatly for we don't have to reconstruct the original

signal accurately. In next part, the availability and effectiveness of the method will be validated with various bearing vibration signals sampled in practice.

#### 4. Experimental tests

We will test the proposed method using vibration data corresponding to different states based on 6205-2RS JEK SKF deep groove ball bearings (data source from [12], signal sampling frequency is 12 K). All the vibration signals used in our test will be divided into lots of samples, and each sample can be denoted by  $\mathbf{x} \in \mathbf{R}^N$ ,  $N = 1000$ . In learning section, 840 samples corresponding to normal state will be used as training samples, which can be classified into four types. Each type has different motor load and motor speed when sampling. The training samples are shown in Table 1.

**Table 1.** Training samples

| Sample sequence | Motor load (HP) | Approx. motor speed (r/min) |
|-----------------|-----------------|-----------------------------|
| 001-120         | 0               | 1797                        |
| 121-360         | 1               | 1772                        |
| 361-600         | 2               | 1750                        |
| 601-840         | 3               | 1730                        |

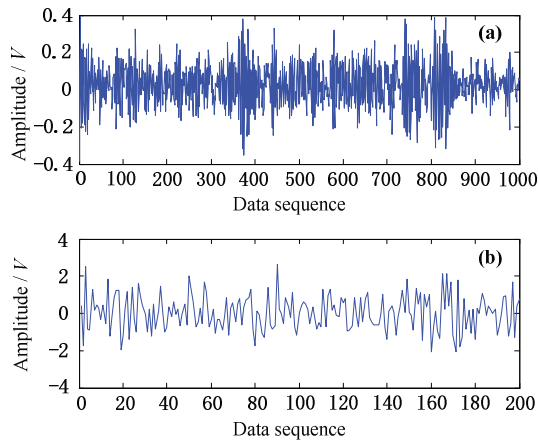
Choosing discrete Fourier transform matrix as dictionary matrix  $\Psi$  ( $N \times N$ ,  $N = 1000$ ), so the whole range should be  $\mathcal{Z} = [1; 1000]$ . According to the interval learning method in Section 3.1 and setting the threshold  $\varepsilon$  as 0.9, the intervals corresponding to normal state and fault state can be determined as  $\mathcal{T} = [1; 177] \cup [825; 998]$  and  $\mathcal{F} = [178; 824] \cup [999; 1000]$  respectively. Meanwhile, the number of the support points will be set as  $k = 30$ .

The original test samples contained 840 normal state samples and 480 fault state samples, and both patterns can be classified into four types. Each type has different motor load and motor speed when sampling. The original test samples are shown in Table 2. The fault was designed as single point in inner race, which was introduced to the test bearings using electro-discharge machining with fault diameter of 0.021 inches and fault depth of 0.011 inches.

**Table 2.** Original test samples

| Pattern | Sample sequence | Motor load (HP) | Approx. motor speed (r/min) |
|---------|-----------------|-----------------|-----------------------------|
| Normal  | 001-120         | 0               | 1797                        |
|         | 121-360         | 1               | 1772                        |
|         | 361-600         | 2               | 1750                        |
|         | 601-840         | 3               | 1730                        |
| Fault   | 001-120         | 0               | 1797                        |
|         | 121-240         | 1               | 1772                        |
|         | 241-360         | 2               | 1750                        |
|         | 361-480         | 3               | 1730                        |

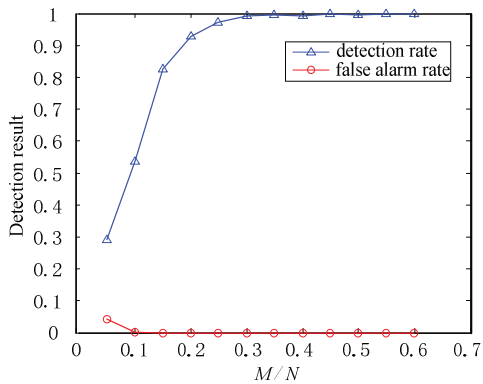
In practice, we should sample compressive measurements directly from practical analog signal. While we just focus on how to use these measurements to detect faults, so in our test the compressive measurements can be acquired by compressed sampling to original test samples mentioned above. We choose the observation matrix  $\Phi$  as Gaussian random matrix with dimensions as  $M \times N$  ( $M < N$ ) [15-17]. This indicates that we can collect  $M$  compressive data points in compressed sampling way during the same sampling time with that we collected  $N$  data points in traditional sampling way. Then we can acquire 1320 compressive measurements with dimensions as  $M$  which including 840 measurements corresponding to normal state and 480 measurements corresponding to fault state. For instance, some  $N$ -dimensional original test sample ( $N = 1000$ , fault state) and the corresponding  $M$ -dimensional compressive measurements ( $M = 200$ ) with Gaussian random observation matrix  $\Phi$  are shown in Fig. 4.



**Fig. 4.** General view of test data, a) is original test sample, b) is the compressive measurements corresponding to a)

The test result can be described using fault detection rate and false alarm rate. The fault detection rate characterizes the ratio of the number of the samples identified as fault state accurately to the number of all the fault samples used in test. The false alarm rate describes the ratio of the number of the samples recognized as fault state which actually should be normal state to the number of all the normal samples used in test. According to the analysis mentioned above, detection rate and false alarm rate should be affected by  $M/N$ , threshold  $\delta$  and energy rate  $\alpha$ , which will be analyzed next respectively. To avoid the instability caused by the randomness of Gaussian random observation matrix  $\Phi$ , the tests will be repeated several times and we use the average value as the final detection result.

The detection result is shown in Fig. 5 when  $\delta = 0.9$  and  $\alpha = 0.2$  with the  $M/N$  changes. It can be seen that with the increasing of measurements, the false alarm rate will fall to zero quickly and the detection rate will increase gradually until stabilization. This result should be consistent with the compressed sensing theory, for the increasing of measurements means that the condition information contained in compressive measurements will be enriched, which is of benefit to the fault detection.



**Fig. 5.** Effects of  $M/N$  to the fault detection result when  $\delta = 0.9$ ,  $\alpha = 0.2$

The detection result is shown in Fig. 6 when  $M/N = 0.1$  and  $M/N = 0.2$  and  $\alpha = 0.2$  with the threshold  $\delta$  changes. It can be seen that in both cases with different  $M/N$ , with the increasing of the threshold  $\delta$ , both the detection rate and the false alarm rate will decline gradually. It should be noticed that the threshold  $\delta$  characterizes the ratio of the energy of the support points falling



into interval  $\mathcal{F}$  to the gross energy of all the support points recovered. So the bigger the threshold  $\delta$  is, the more difficult the support points entering into the target interval  $\mathcal{F}$  will be, and that's why the fault detection rate and false alarm rate will decline.

The detection result is shown in Fig. 7 when  $M/N = 0.1$  and  $M/N = 0.2$  and  $\delta = 0.9$  with the energy factor  $\alpha$  changes. It can be seen that in both cases with different  $M/N$ , the false alarm rate is always in a lower level closing to zero, and with the increasing of the energy factor  $\alpha$ , generally speaking, the detection rate will increase gradually.

Compared to the case with small number of measurements, we can see that the detection result using more measurements have less effect by the threshold  $\delta$  and the energy factor  $\alpha$  from Fig. 6 and Fig. 7. It also indicates the fact in Fig. 5 that the more measurements we use, the better the detection result will be. Therefore, we should sample more measurements as far as possible if the data storage and transmission permit.

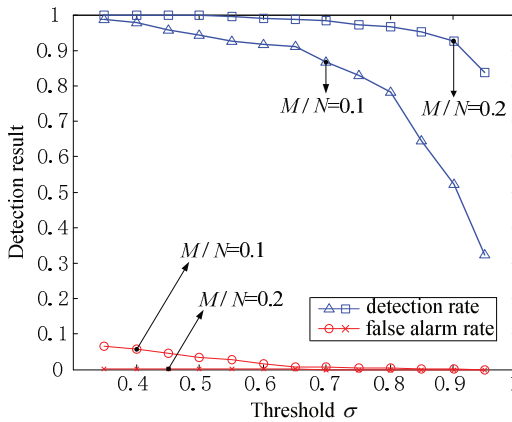


Fig. 6. Effects of  $\delta$  to the fault detection result when  $M/N = 0.1$  and  $M/N = 0.2$ ,  $\alpha = 0.2$

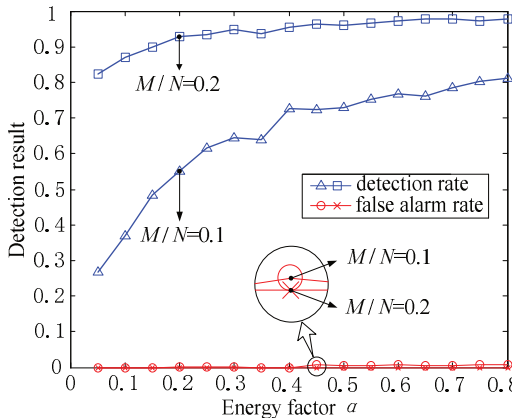


Fig. 7. Effects of  $\alpha$  to the fault detection result when  $M/N = 0.1$  and  $M/N = 0.2$ ,  $\delta = 0.9$

In the above test, the fault corresponding to the test samples we used is single point fault with diameter of 0.021 inches and fault depth of 0.011 inches in inner race. For comparing the fault detection results in the cases with different levels of fault, we will use 480 fault samples corresponding to the case with low-level fault, and each sample can be denoted by  $\mathbf{x} \in \mathbf{R}^N$ ,  $N = 1000$ . Here the fault is single point fault with diameter of 0.007 inches (fault depth kept as 0.011 inches) in inner race. As with the above testing, we divide these fault samples into four types according to motor load and motor speed when sampling (same as that in Table 2). Each

type includes 120 samples. Setting  $k = 30$ ,  $\delta = 0.9$  and  $\alpha = 0.2$ , the fault detection rates with different number of measurements are shown in Fig. 8.

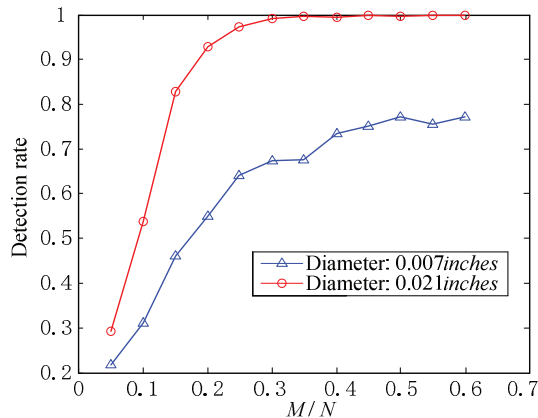


Fig. 8. Detection rates for different level of faults ( $\delta = 0.9$ ,  $\alpha = 0.2$ ,  $M/N$  changes)

The false alarm rates are calculated using normal state samples, which should be invariable with the same parameters. So the false alarm rates in current condition will be the same as that shown in Fig. 5 and will not be shown in Fig. 8. It can be seen from Fig. 8 that in the case with the same  $M/N$ , the detection rate corresponding to the case with fault diameter of 0.021 inches is higher than that with fault diameter of 0.007 inches. This result is in accord with that in theoretic case. In theoretic case, the severer the fault is, the more significant the differences of energy distributions between normal state signal and fault state signal in frequency domain will be. This is the foundation of the proposed method. So it shows the trend in Fig. 8 that the severer the fault is, the higher the detection rate will be.

The faults corresponding to all the test samples we used in above test are inner race faults. To validate the method in different fault patterns, we will use the samples corresponding to rolling ball fault and outer race fault. The two patterns of the faults are also single point faults with diameter of 0.021 inches and depth of 0.011 inches in bearing rolling ball and outer race respectively. Similar to the above test, there are 480 samples (each sample can be denoted by  $\mathbf{x} \in \mathbf{R}^N$ ,  $N = 1000$ ) for each pattern, which can be divided into four types according to motor load and motor speed when sampling (same as that in Table 2). The fault detection results with different number of measurements are shown in Fig. 9 when  $k = 30$ ,  $\delta = 0.9$  and  $\alpha = 0.2$ . The false alarm rates corresponding to different patterns are the same as that in Fig. 5, so they are not plotted. It can be seen from Fig. 9 that the detection rates will increase gradually in general until stabilization with the increasing of the measurements for every fault pattern. In the cases with inner race fault and outer race fault, the fault detection rate can reach to 95%. While for the case with rolling ball fault, the detection rate will not be desirable which just around 60%. The posture change of rolling ball should be complex in bearing operation, and this may result in the energy decentralization in frequency domain, so the detection rate to rolling ball fault using the proposed method will not be very high.

In the experimental tests of this section, we used the compressive measurements based on Gaussian random matrix. Actually, we can also use other compressed sampling ways. To validate our proposed method with different compressed sampling ways, we will compare the fault detection results in several typical compressed sampling ways (based on Gaussian random matrix, partial orthogonal matrix and Toeplitz and circulant matrix separately). In this test, we set  $N = 1000$ ,  $\delta = 0.9$ ,  $\alpha = 0.2$  and use the samples in Table 2 as the original test samples, then the fault detection results with different number of measurements ( $M$ ) are shown in Fig. 10. It can be seen from Fig. 10 that all the three compressed sampling ways can achieve fault detection very

well and can be used when collecting compressive measurements. While considering the higher computational efficiency and universality of Gaussian random matrix, and that's why we used the Gaussian random matrix in the previous tests.

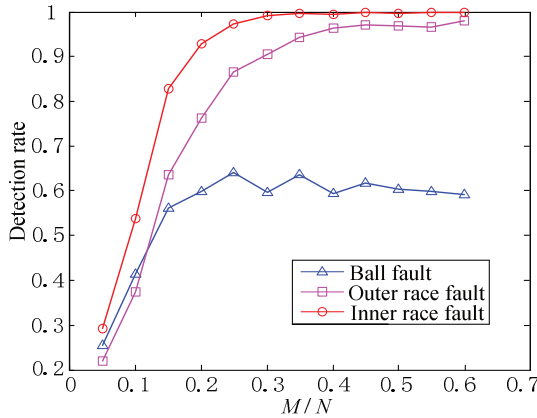


Fig. 9. Detection rates corresponding to different fault patterns ( $\delta = 0.9$ ,  $\alpha = 0.2$ ,  $M/N$  changes)

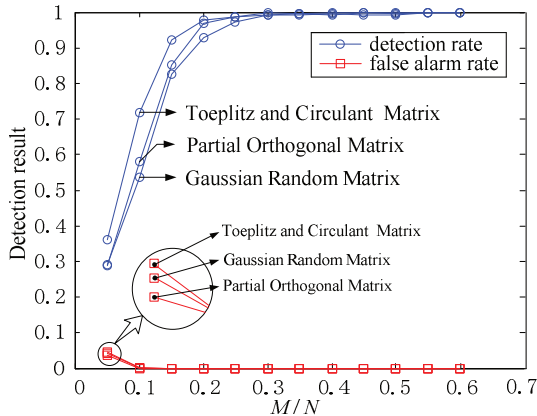


Fig. 10. Fault detection results in three typical compressed sampling ways ( $\delta = 0.9$ ,  $\alpha = 0.2$ ,  $M/N$  changes)

## 5. Conclusions

In the process of condition monitoring, it is always necessary to monitor the component uninterruptedly for a long period of time. This process will generate masses of original data which will burden the storage and transmission. For this issue, we proposed a fault detection method based on compressed sensing, which can achieve fault detection directly using fewer compressive measurements compared to original data. Amount of samples corresponding to different bearing states were used to validate the method. The test results showed that with the increasing of the measurements, fault detection rate would increase and false alarm rate would decrease gradually. Meanwhile, the severer the fault was, the better the detection result would be. It is also indicated that the detection results will be better for inner race fault and outer race fault than that for rolling ball fault using our proposed method. The further research will focus on how to improve the fault detection result and how to achieve fault classification based on compressive measurements directly.

## Acknowledgements

The authors gratefully acknowledge the financial support of National Natural Science Foundation of China under Grant No. 51375484 and No. 51205401 and Bearing Data Center of Case Western Reserve University to provide the bearing test data. Valuable comments on the paper from anonymous reviewers are very much appreciated.

## References

- [1] **J. Dai J., Das D., Ohadi M., Pecht M.** Reliability risk mitigation of free air cooling through prognostics and health management. *Applied Energy*, Vol. 111, 2013, p. 104-112.
- [2] **Candès E., Romberg J., Tao T.** Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, Vol. 52, Issue 2, 2006, p. 489-509.
- [3] **Donoho D. L.** Compressed sensing. *IEEE Transactions on Information Theory*, Vol. 52, Issue 4, 2006, p. 1289-1306.
- [4] **Candès E.** Compressive sampling. *Proceedings of International Congress of Mathematicians, European Mathematical Society Publishing House, Zürich, Switzerland, 2006*, p. 1433-1452.
- [5] **Candès E., Wakin M.** An introduction to compressive sampling. *IEEE Signal Processing Magazine*, Vol. 25, Issue 2, 2008, p. 21-30.
- [6] **Donoho D. L., Tsaig Y.** Extensions of compressed sensing. *Signal Processing*, Vol. 86, Issue 3, 2006, p. 533-548.
- [7] **Candès E., Tao T.** Near optimal signal recovery from random projections: universal encoding strategies. *IEEE Transactions on Information Theory*, Vol. 52, Issue 12, 2006, p. 5406-5425.
- [8] **Davenport M. A., Boufounos P. T., Wakin M. B., Baraniuk R. G.** Signal processing with compressive measurements. *IEEE Journal of Selected Topics in Signal Processing*, Vol. 4, Issue 2, 2010, p. 445-446.
- [9] **Lu C., Tao L. F., Fan H. Z., Wang Z. B.** Approach to health monitoring and assessment of rolling bearing. *Journal of Vibroengineering*, Vol. 15, Issue 2, 2013, p. 746-760.
- [10] **Candès E., Tao T.** Decoding by linear programming. *IEEE Transactions on Information Theory*, Vol. 51, Issue 12, 2005, p. 4203-4215.
- [11] **Donoho D. L., Elad M., Temlyakov V.** Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on Information Theory*, Vol. 52, Issue 1, 2006, p. 6-8.
- [12] **Seeded Fault Test Data.** Bearing Data Center, Case Western Reserve University. <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>
- [13] **Mallat S., Zhang Z.** Matching pursuit with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, Vol. 41, Issue 12, 1993, p. 3397-3415.
- [14] **Friedman J. H., Tukey J. W.** A projection pursuit algorithm for exploratory data analysis. *IEEE Trans Comput*, Vol. 23, Issue 9, 1974, p. 881-890.
- [15] **Candès E., Eldarb Y. C., Needella D., Randall P.** Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*, Vol. 31, Issue 1, 2011, p. 59-73.
- [16] **Candès E., Romberg J., Tao T.** Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, Vol. 59, Issue 8, 2006, p. 1207-1223.
- [17] **Tsaig Y., Donoho D. L.** Extensions of compressed sensing. *Signal Processing*, Vol. 86, Issue 3, 2006, p. 549-571.