

Quantum Gene Chain Coding Bidirectional Neural Network for Residual Useful Life Prediction of Rotating Machinery

Feng Li (✉ lifeng19820501@163.com)

Sichuan University

Yang-Yang Cheng

Sichuan University

Bao-Ping Tang

Chongqing University

Xue-Ming Zhou

Chongqing University

Rui-Ping Xiong

Sichuan University

Original Article

Keywords: quantum computing, quantum gene chain coding, residual useful life prediction, rotating machinery

Posted Date: May 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-25786/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Title page

Quantum Gene Chain Coding Bidirectional Neural Network for Residual Useful Life Prediction of Rotating Machinery

Feng Li, born in 1982, is currently an associate professor and supervisor for master student in *School of Mechanical Engineering at Sichuan University, China*. He received his master and PhD degrees from *Chongqing University, China*, in 2008 and 2011, respectively. His main research interests include equipment condition monitoring, fault diagnosis as well as prognostics and health management (PHM).

Tel: +86-18382385401; E-mail: lifeng19820501@163.com

Yang-Yang Cheng, born in 1994, is currently a graduate student in *School of Mechanical Engineering at Sichuan University, China*. He received his bachelor's degree from *Nanjing University of Aeronautics and Astronautics, China*, in 2017. His research interest is fault diagnosis and degradation trend prediction of mechanical equipment.

E-mail: cyoungyoung@126.com

Bao-Ping Tang, born in 1971, is currently a professor and a PhD candidate supervisor at *Chongqing University, China*. He received his master and PhD degrees from *Chongqing University, China*, in 1996 and 2003, respectively. His main research interests include equipment condition monitoring and fault diagnosis, virtual instrument and wireless sensor network.

E-mail: bptang@cqu.edu.cn

Xue-Ming Zhou, born in 1978, is currently an engineer and general manager of *Chongqing Ming's Min Mechanotronics Co., Ltd., China*. He received his master degree from *Chongqing University, China*, in 2009. His main research interests include equipment condition monitoring, fault diagnosis as well as computer aided design and manufacturing.

E-mail: zhou.xueming@cqlap.com.cn

Rui-Ping Xiong, born in 1968, is currently an associate professor and supervisor for master student in *School of Mechanical Engineering at Sichuan University, China*. He received his master and PhD degrees from *Sichuan University, China*, in 1998 and 2006, respectively. His main research interests include electromechanical transmission and intelligent control, equipment condition monitoring and fault diagnosis.

E-mail: xiongruiping@163.com

Corresponding author: Feng Li; Yang-Yang Cheng

E-mail: lifeng19820501@163.com; cyoungyoung@126.com

ORIGINAL ARTICLE (or REVIEW)

Quantum Gene Chain Coding Bidirectional Neural Network for Residual Useful Life Prediction of Rotating Machinery

Feng Li¹ • Yang-Yang Cheng¹ • Bao-Ping Tang² • Xue-Ming Zhou³ • Rui-Ping Xiong¹

Received June xx, 201x; revised February xx, 201x; accepted March xx, 201x

© Chinese Mechanical Engineering Society and Springer-Verlag Berlin Heidelberg 2017

Abstract: In classical recurrent neural networks, the pre- and post-relationships of time series tend to be neglected so that long-term overall memory is generally inaccessible; meanwhile, the weights are transferred and updated mainly by the gradient descent method, which leads to their low prediction accuracy and high computation cost in the application of residual useful life (RUL) prediction of rotating machinery (RM). In view of this, a quantum gene chain coding bidirectional neural network (QGCCBNN) is proposed to predict RUL of RM in this paper. In our proposed QGCCBNN, the quantum bidirectional transmission mechanism is designed to establish the pre- and post-relationships of time series for readjusting the weight parameters according to the feedback from the output layer, so that higher consistency between the input information and the overall memory of the network can be realized, thus endowing QGCCBNN with better nonlinear approximation ability. Moreover, in order to improve the global optimization ability and convergence speed, the quantum gene chain coding instead of gradient descent method is constructed to transmit and update data, in which the qubit probability amplitude real number coding is adopted and the cosine and sinusoidal qubit probability amplitudes corresponding to the minimum loss function are compared with those of the current time by the phase selection matrix for the directional parallel updating of the weight parameters. On this basis, a new RUL prediction method for RM is proposed, and higher prediction accuracy as well as desirable efficiency can be obtained due to the advantages of QGCCBNN in nonlinear approximation ability and convergence speed. The experimental example for RUL prediction of a double-row roller bearing demonstrates the effectiveness of our proposed method.

Keywords: quantum computing • quantum gene chain coding •

✉ Feng Li; Yang-Yang Cheng
lifeng19820501@163.com; cyoungyoung@126.com

¹ School of Mechanical Engineering, Sichuan University, Chengdu 610065, China

² The State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing 400030, China

³ Chongqing Ming's Min Mechanotronics Co., Ltd., Chongqing 401120, China

residual useful life prediction • rotating machinery

1 Introduction

Rotating machinery (RM) is widely used in automobile transmission systems, aero engines, wind turbines and other machines [1]. Because of its complex structure and various working conditions, different types of faults may affect its normal operation and ultimately give rise to the arrival of the service life threshold [2]. The damage degree of key components of RM, such as bearings, gears, rotors and impellers, determines whether the machine can operate safely and reliably for a long time [3, 4]. Therefore, it is very important to monitor the operating condition of RM and predict its residual useful life (RUL) [5, 6].

Recently, data-driven RUL prediction methods based on artificial intelligence algorithms have attracted much attention because of their fast and accurate characteristics [7]. These methods start from industrial scene and construct a non-linear model by establishing the relationships among the feature vector, the fault mode and the RUL prediction [8]. Furthermore, in order to accurately judge the fault state of RM and predict its RUL, this kind of methods pays much attention to the hidden information of data. Overall, data-driven RUL prediction methods can provide enough decision-making information for the maintenance of corresponding components, thus avoiding major catastrophic accidents of machines. Consequently, they possess great engineering application values, and have become a research hotspot and the development trend in the field of RUL prediction of machines [9]. For example, Ren et al. [10] used a deep autoencoder and deep neural network (DADNN) to estimate RUL of RM; Banu et al. [11] applied a fully complex-valued radial basis function neural network (FCRBFNN) to forecast the load trend of power systems; Hinch et al. [12] adopted a convolutional long-short-term memory neural network (CLSTMNN) to predict the degradation trend of the rolling bearing; Xiao et al. [13] predicted the degradation trend and RUL of rolling

bearings based on the least squares support vector machine (LS-SVM). However, all the above state-of-the-art methods of data-driven RUL prediction have some defects. For DADNN, its structure of deep neural network means that the data propagates hierarchically in the multi-layer network, which can result in gradient disappearance or gradient explosion. In terms of FCRBFNN, it is difficult to properly tune its massive parameters, such as the number of radial basis functions, the values of center vectors, and so on; once these parameters are selected inappropriately, the convergence error can be increased and the network may be over-fitting. As for CLSTMNN, its complex structure requires four linear layers per cell, which may lead to a slow training process. In the case of LS-SVM, it is difficult to set kernel functions and parameters accurately, so the prediction results are often uncertain, resulting in the degradation of its generalization performance.

RUL prediction is essentially an input series trend prediction problem with threshold determination. As a typical time series prediction theory, the recurrent neural network (e.g., the long short-term memory neural network (LSTMNN) [14], the gated recurrent unit neural network (GRUNN) [15], etc.), has showed a certain application prospect in RUL applications [16-18] because of its cumulative effect of time series. However, it has an obvious defect of lacking reverse feedback from output to input [19]. Fortunately, the bidirectional transmission mechanism can make up for the above defect [20]. Moreover, the quantum computing can contribute to improving the global optimization ability and computational performance of the recurrent neural network [21-23] because of its high speed and parallelism [24, 25]. Based on the complementary advantages of bidirectional transmission mechanism and quantum computing, a novel recurrent neural network called quantum gene chain coding bidirectional neural network (QGCCBNN) is proposed in this paper. In QGCCBNN, the quantum bidirectional transmission mechanism is designed to establish the pre- and post-relationships of time series for readjusting the weight parameters according to the feedback from the output layer, so that higher consistency between the input information and the overall memory of the network can be realized, thus endowing QGCCBNN with a better nonlinear approximation ability. Meanwhile, in order to avoid the gradient disappearance, gradient explosion, slow convergence speed and large time cost caused by the gradient descent method in classical recurrent neural networks, the quantum gene chain coding is constructed to transmit and update data, in which the qubit probabilistic

amplitude real number coding is adopted and the cosine and sinusoidal qubit probability amplitudes corresponding to the minimum loss function are compared with those of the current time by the phase selection matrix for the directional parallel updating of the weight parameters. Consequently, QGCCBNN has better global optimization ability and faster convergence speed.

Based on the advantages of QGCCBNN in nonlinear approximate ability, global optimization performance and convergence speed, we propose a new method for RUL prediction of RM. In the proposed method, a trend index (TI) is firstly extracted from the vibration acceleration data of RM as the performance degradation feature, then the TI is put into QGCCBNN for predicting the degradation trend of RM, and finally the failure probability model is established to predict RUL of RM based on the degradation trend curve. Theoretical analysis of QGCCBNN and experimental results of RUL prediction of a double-row roller bearing show that our proposed method can achieve higher prediction accuracy and shorter computational time.

The main motivation of this paper is to solve the problem of low accuracy and high computation cost of RUL prediction of RM caused by poor nonlinear approximation ability and slow convergence speed of classical recurrent neural networks. On account of this, we propose a novel neural network termed as quantum gene chain coding bidirectional neural network (QGCCBNN) to predict RUL of RM with higher accuracy and desirable efficiency.

The rest of this paper is organized as follows. In Section 2, the background knowledge about QGCCBNN is briefly introduced, and the QGCCBNN algorithm is proposed and derived in detail. Then, the RUL prediction method for RM based on QGCCBNN is described in Section 3. Section 4 presents the experimental results along with corresponding comparative analysis to demonstrate the effectiveness of our RUL prediction method. Finally, some conclusions are drawn in Section 5.

2 Quantum gene chain coding bidirectional neural network (QGCCBNN)

2.1 Knowledge related to QGCCBNN

2.1.1 Quantum gate group

In quantum computation, there are some general quantum gate groups which can construct arbitrary quantum gates. The most fundamental general quantum gate group consists of quantum phase-shift gate and controlled non-gate.

$$= \left(\sin\left(\frac{\pi}{2} - \arccos(s'_1)\right), \sin\left(\frac{\pi}{2} - \arccos(s'_2)\right), \dots, \sin\left(\frac{\pi}{2} - \arccos(s'_m)\right) \right), \quad (10)$$

$$\mathbf{q}'_{js-} = (s'_1, s'_2, \dots, s'_m)$$

$$= \left(\sin\left(\frac{\pi}{2} + \arccos(s'_1)\right), \sin\left(\frac{\pi}{2} + \arccos(s'_2)\right), \dots, \sin\left(\frac{\pi}{2} + \arccos(s'_m)\right) \right), \quad (11)$$

where $\sin\left(\frac{\pi}{2} - \arccos(s'_i)\right)$ and $\sin\left(\frac{\pi}{2} + \arccos(s'_i)\right)$ are sinusoidal solutions. Namely, \mathbf{q}'_{js+} and \mathbf{q}'_{js-} are also the feasible solutions of the global optimal solution s .

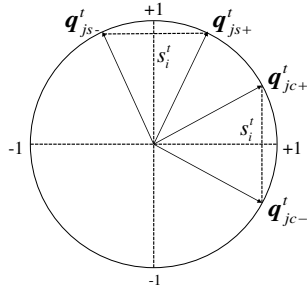


Figure 1 Spatial mapping relation of a qubit

Based on the analysis above, because both of the quantum gene chains \mathbf{q}'_{jc} and \mathbf{q}'_{js} are feasible solutions of s , the number of feasible solutions can increase exponentially. In a word, a quantum chromosome whose length is m can extend the number of feasible solutions from m to 2^{m+1} . Besides, cosine solutions \mathbf{q}'_{jc} and sinusoidal solutions \mathbf{q}'_{js} of the global optimal solution s can be updated synchronously in each iteration, and thus the optimization process of the global optimal solution s can be accelerated. Therefore, quantum gene chain coding has the ability of expanding search range and parallel computing.

2.2 Theory model of QGCCBNN

Fig. 2 shows the model structure of QGCCBNN. In order to realize higher consistency between the input information and the overall memory of the network and obtain better nonlinear approximation ability, the quantum bidirectional transmission mechanism is designed and introduced to QGCCBNN for establishing the pre- and post-relationships of time series and readjusting the weight parameters according to the feedback from the output layer. First, the quantized time series \mathbf{x}' is transmitted from the input layer to hidden layer by quantum phase-shift gate and controlled non-gate. Then, in a similar way, the hidden

layer state $|\mathbf{h}^{t-1}\rangle = (|h_1^{t-1}\rangle, |h_2^{t-1}\rangle, \dots, |h_n^{t-1}\rangle)^T$ of the previous time and the output layer state $|\mathbf{y}'\rangle = (|y'_1\rangle, |y'_2\rangle, \dots, |y'_n\rangle)^T$ of the current time are transferred to the hidden layer, which means that the pre- and post-relationships of time series are obtained. Finally, the current hidden layer state is transmitted to the output layer and the probability amplitude of quantum excited state $|1\rangle$ of the output layer state can be selected as the output of QGCCBNN.

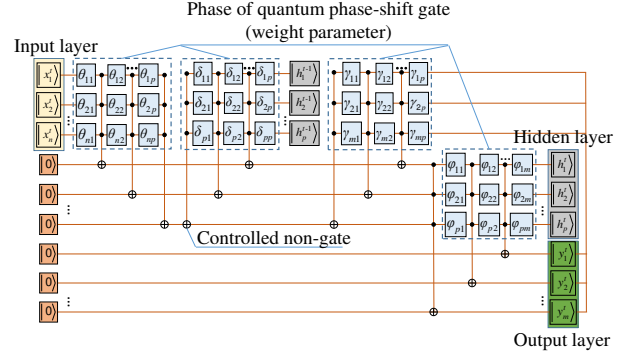


Figure 2 Model structure of QGCCBNN

For convenience, assume that the normalized input series of QGCCBNN at time t is $\mathbf{x}' = (x'_1, x'_2, \dots, x'_n)$ ($x'_i \in [a_i, b_i]$), whose quantum state is $|\mathbf{x}'\rangle = (|x'_1\rangle, |x'_2\rangle, \dots, |x'_n\rangle)^T$, in which $|x'_i\rangle = (\cos \alpha'_i, \sin \alpha'_i)^T$ and $i = 1, 2, \dots, n$; the hidden layer state at time t is $\mathbf{h}' = (h'_1, h'_2, \dots, h'_p)^T$, whose quantum state is $|\mathbf{h}'\rangle = (|h'_1\rangle, |h'_2\rangle, \dots, |h'_p\rangle)$, in which $|h'_j\rangle = (\cos \beta'_j, \sin \beta'_j)^T$ and $j = 1, 2, \dots, p$; the output layer state at time t is $\mathbf{y}' = (y'_1, y'_2, \dots, y'_m)$, whose quantum state is $|\mathbf{y}'\rangle = (|y'_1\rangle, |y'_2\rangle, \dots, |y'_m\rangle)$, in which $|y'_k\rangle = (\cos \varepsilon'_k, \sin \varepsilon'_k)^T$ and $k = 1, 2, \dots, m$. Thus, the transfer relationship between the input and the output in each layer of QGCCBNN can be deduced as follows.

First of all, the normalized input series $\mathbf{x}' = (x'_1, x'_2, \dots, x'_n)$ is quantized to $|\mathbf{x}'\rangle = (|x'_1\rangle, |x'_2\rangle, \dots, |x'_n\rangle)^T$, in which the quantized element $|x'_i\rangle$ can be written as

$$|x'_i\rangle = (\cos \alpha'_i, \sin \alpha'_i)^T$$

$$= \cos\left(\frac{2\pi(x'_i - a_i)}{b_i - a_i}\right)|0\rangle + \sin\left(\frac{2\pi(x'_i - a_i)}{b_i - a_i}\right)|1\rangle$$

$$= \left(\cos\left(\frac{2\pi(x'_i - a_i)}{b_i - a_i}\right), \sin\left(\frac{2\pi(x'_i - a_i)}{b_i - a_i}\right) \right)^T$$

$$= (\cos(2\pi x'_i), \sin(2\pi x'_i))^T, \quad (12)$$

where $\alpha'_i = 2\pi x'_i$ is the quantized phase of $|x'_i\rangle$; $a_i=0$ and $b_i=1$ are respectively the lower and upper bounds of the normalized input series.

Then, according to the model structure of QGCCBNN shown in Fig. 2 and Eqs. (1-3) (i.e., the definition of quantum phase-shift gate and controlled non-gate), the hidden layer state at time t can be computed as

$$\begin{aligned} |h'_j\rangle &= (\cos \beta'_j, \sin \beta'_j)^T \\ &= |h'_j\rangle_x + |h'_j\rangle_h + |h'_j\rangle_y \\ &= C^n(U) |h'_{jx_1} h'_{jx_2} \dots h'_{jx_n}\rangle |0\rangle + C^p(U) |h'_{jh_1} h'_{jh_2} \dots h'_{jh_p}\rangle |0\rangle + C^m(U) |h'_{jy_1} h'_{jy_2} \dots h'_{jy_m}\rangle |0\rangle \\ &= (|h'_{jx_1} h'_{jx_2} \dots h'_{jx_n}\rangle \oplus |0\rangle) + (|h'_{jh_1} h'_{jh_2} \dots h'_{jh_p}\rangle \oplus |0\rangle) + (|h'_{jy_1} h'_{jy_2} \dots h'_{jy_m}\rangle \oplus |0\rangle) \\ &= |h'_{jx_1} h'_{jx_2} \dots h'_{jx_n}\rangle + |h'_{jh_1} h'_{jh_2} \dots h'_{jh_p}\rangle + |h'_{jy_1} h'_{jy_2} \dots h'_{jy_m}\rangle \\ &= \prod_{i=1}^n (R(\theta'_{ij}) |x'_i\rangle) + \prod_{j=1}^p (R(\delta'_{jj}) |h'_j\rangle) + \prod_{k=1}^m (R(\gamma'_{kj}) |y'_k\rangle), \end{aligned} \quad (13)$$

where $|h'_j\rangle_x$, $|h'_j\rangle_h$ and $|h'_j\rangle_y$ are the linear components caused by the input layer state at time t , the hidden layer state at time $t-1$ and the output layer state at time t , respectively; $C^n(U)$, $C^p(U)$ and $C^m(U)$ are controlled non-gates whose control qubits are respectively $|h'_{jx_1} h'_{jx_2} \dots h'_{jx_n}\rangle$, $|h'_{jh_1} h'_{jh_2} \dots h'_{jh_p}\rangle$ and $|h'_{jy_1} h'_{jy_2} \dots h'_{jy_m}\rangle$; $R(\theta'_{ij})$, $R(\delta'_{jj})$ and $R(\gamma'_{kj})$ are quantum phase-shift gates whose phase are respectively θ'_{ij} , δ'_{jj} and γ'_{kj} .

According to Eq. (1), $|h'_j\rangle_x$, $|h'_j\rangle_h$ and $|h'_j\rangle_y$ can be further deduced as follows:

$$\begin{aligned} |h'_j\rangle_x &= \prod_{i=1}^n (R(\theta'_{ij}) |x'_i\rangle) \\ &= \prod_{i=1}^n \cos(\alpha'_i + \theta'_{ij}) |0\rangle + \prod_{i=1}^n \sin(\alpha'_i + \theta'_{ij}) |1\rangle, \end{aligned} \quad (14)$$

$$\begin{aligned} |h'_j\rangle_h &= \prod_{j=1}^p (R(\delta'_{jj}) |h'_j\rangle) \\ &= \prod_{j=1}^p \cos(\beta'_j + \delta'_{jj}) |0\rangle + \prod_{j=1}^p \sin(\beta'_j + \delta'_{jj}) |1\rangle, \end{aligned} \quad (15)$$

$$\begin{aligned} |h'_j\rangle_y &= \prod_{k=1}^m (R(\gamma'_{kj}) |y'_k\rangle) \\ &= \prod_{k=1}^m \cos(\epsilon'_k + \gamma'_{kj}) |0\rangle + \prod_{k=1}^m \sin(\epsilon'_k + \gamma'_{kj}) |1\rangle. \end{aligned} \quad (16)$$

Thus, according to Eqs. (14-16), $|h'_j\rangle_x$, $|h'_j\rangle_h$ and $|h'_j\rangle_y$ can be obtained. Further, the hidden layer state at time t can be expressed as follows:

$$\begin{aligned} |h'_j\rangle &= (\cos \beta'_j, \sin \beta'_j)^T \\ &= \prod_{i=1}^n (R(\theta'_{ij}) |x'_i\rangle) + \prod_{j=1}^p (R(\delta'_{jj}) |h'_j\rangle) + \prod_{k=1}^m (R(\gamma'_{kj}) |y'_k\rangle) \end{aligned}$$

$$\begin{aligned} &= \left(\prod_{i=1}^n \cos(\alpha'_i + \theta'_{ij}) + \prod_{j=1}^p \cos(\beta'_j + \delta'_{jj}) + \prod_{k=1}^m \cos(\epsilon'_k + \gamma'_{kj}) \right) |0\rangle + \\ &\quad \left(\prod_{i=1}^n \sin(\alpha'_i + \theta'_{ij}) + \prod_{j=1}^p \sin(\beta'_j + \delta'_{jj}) + \prod_{k=1}^m \sin(\epsilon'_k + \gamma'_{kj}) \right) |1\rangle \\ &= \left(\prod_{i=1}^n \cos(\alpha'_i + \theta'_{ij}) + \prod_{j=1}^p \cos(\beta'_j + \delta'_{jj}) + \prod_{k=1}^m \cos(\epsilon'_k + \gamma'_{kj}) \right) |0\rangle + \\ &\quad \left(\prod_{i=1}^n \sin(\alpha'_i + \theta'_{ij}) + \prod_{j=1}^p \sin(\beta'_j + \delta'_{jj}) + \prod_{k=1}^m \sin(\epsilon'_k + \gamma'_{kj}) \right) |1\rangle. \end{aligned} \quad (17)$$

Now, our newly-designed quantum bidirectional transmission mechanism has been realized by Eqs. (13-17), which can readjust the weight parameters (i.e., the phases of quantum phase-shift gates) according to the feedback from the output layer, so QGCCBNN can realize higher consistency between the input information and the overall memory of the network.

Finally, the output layer state at time t can be computed as

$$\begin{aligned} |y'_k\rangle &= (\cos \epsilon'_k, \sin \epsilon'_k)^T \\ &= \prod_{j=1}^p (R(\phi'_{jk}) |\beta'_j\rangle) \\ &= \prod_{j=1}^p \cos(\beta'_j + \phi'_{jk}) |0\rangle + \prod_{j=1}^p \sin(\beta'_j + \phi'_{jk}) |1\rangle. \end{aligned} \quad (18)$$

Therefore, the probability amplitude of the state $|1\rangle$ is selected as the output of each layer, and then the outputs of the hidden layer and the output layer can be respectively expressed as

$$|h'_j\rangle = \left(\prod_{i=1}^n \sin(\alpha'_i + \theta'_{ij}) + \prod_{j=1}^p \sin(\beta'_j + \delta'_{jj}) + \prod_{k=1}^m \sin(\epsilon'_k + \gamma'_{kj}) \right), \quad (19)$$

$$|y'_k\rangle = \prod_{j=1}^p \sin(\beta'_j + \phi'_{jk}). \quad (20)$$

2.3 Training algorithm for QGCCBNN

In our proposed QGCCBNN, the weight parameters (i.e., θ'_{ij} , δ'_{jj} , γ'_{kj} and ϕ'_{jk} , the phases of quantum phase-shift gates) that need to be updated are represented by the quantum phase matrices as follows:

a. The quantum phase matrix which transfers data from

$$\text{the input layer to the hidden layer: } \theta' = \begin{pmatrix} \theta'_{11} & \theta'_{12} & \dots & \theta'_{1p} \\ \theta'_{21} & \theta'_{22} & \dots & \theta'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \theta'_{n1} & \theta'_{n2} & \dots & \theta'_{np} \end{pmatrix};$$

b. The quantum phase matrix which transfers data from the hidden layer of the previous time to the current hidden

$$\text{layer: } \delta' = \begin{pmatrix} \delta'_{11} & \delta'_{12} & \dots & \delta'_{1p} \\ \delta'_{21} & \delta'_{22} & \dots & \delta'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \delta'_{p1} & \delta'_{p2} & \dots & \delta'_{pp} \end{pmatrix};$$

c. The quantum phase matrix which transfers feedback data from the output layer to the hidden layer:

$$\gamma^t = \begin{pmatrix} \gamma'_{11} & \gamma'_{12} & \cdots & \gamma'_{1p} \\ \gamma'_{21} & \gamma'_{22} & \cdots & \gamma'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma'_{m1} & \gamma'_{m2} & \cdots & \gamma'_{mp} \end{pmatrix};$$

d. The quantum phase matrix which transfers data from the hidden layer to the output layer:

$$\phi^t = \begin{pmatrix} \phi'_{11} & \phi'_{12} & \cdots & \phi'_{1m} \\ \phi'_{21} & \phi'_{22} & \cdots & \phi'_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \phi'_{p1} & \phi'_{p2} & \cdots & \phi'_{pm} \end{pmatrix}.$$

In order to update the matrices θ^t , δ^t , γ^t , and ϕ^t , all their elements are firstly extracted into a parameter vector as $\mathbf{p}^t = (\mathbf{p}'_{\theta}, \mathbf{p}'_{\delta}, \mathbf{p}'_{\gamma}, \mathbf{p}'_{\phi})$, where $\mathbf{p}'_{\theta} = (\theta'_{11}, \theta'_{12}, \dots, \theta'_{ij}, \dots, \theta'_{np})$, $\mathbf{p}'_{\delta} = (\delta'_{11}, \delta'_{12}, \dots, \delta'_{ji}, \dots, \delta'_{pp})$, $\mathbf{p}'_{\gamma} = (\gamma'_{11}, \gamma'_{12}, \dots, \gamma'_{kj}, \dots, \gamma'_{np})$, and $\mathbf{p}'_{\phi} = (\phi'_{11}, \phi'_{12}, \dots, \phi'_{jk}, \dots, \phi'_{pm})$.

Then \mathbf{p}^t can be updated by the quantum gene chain coding mentioned in subsection 2.1.2. The process of updating \mathbf{p}^t with quantum gene chain coding is described in detail as follows.

For convenience, the parameter vector can be as written as:

$$\mathbf{p}^t = (\mathbf{p}'_{\theta}, \mathbf{p}'_{\delta}, \mathbf{p}'_{\gamma}, \mathbf{p}'_{\phi}) = (p'_1, p'_2, \dots, p'_\mu, \dots, p'_\lambda)^T, \quad (21)$$

where $\lambda = n \times p + p \times p + m \times p + p \times m$ is the length of the parameter vector \mathbf{p}^t ; $p'_\mu \in (0, 2\pi)$ and $\mu = 1, 2, \dots, \lambda$.

In order to obtain the value of \mathbf{p}^t , a quantum genome can be constructed as $\mathbf{Q}^g = (\mathbf{q}_1^g; \mathbf{q}_2^g; \dots; \mathbf{q}_\zeta^g; \dots; \mathbf{q}_\tau^g)$, where g is the iteration number of \mathbf{Q}^g , τ is the size of \mathbf{Q}^g , $\zeta \in (1, 2, \dots, \tau)$ is the quantum chromosome number, $\mathbf{q}_\zeta^g \in \mathbf{Q}^g$ is the ζ -th quantum chromosome in g -th iteration.

Moreover, according to Eqs. (5-7), a quantum chromosome, for calculating \mathbf{p}^t , can be described as follows:

$$\mathbf{q}_\zeta^g = \begin{pmatrix} \cos(w_{\zeta,1}^g) & \cos(w_{\zeta,2}^g) & \cdots & \cos(w_{\zeta,\mu}^g) & \cdots & \cos(w_{\zeta,\lambda}^g) \\ \sin(w_{\zeta,1}^g) & \sin(w_{\zeta,2}^g) & \cdots & \sin(w_{\zeta,\mu}^g) & \cdots & \sin(w_{\zeta,\lambda}^g) \end{pmatrix}, \quad (22)$$

where $\cos(w_{\zeta,\mu}^g)$ and $\sin(w_{\zeta,\mu}^g)$ are respectively the cosine and sinusoidal qubit probability amplitudes of the μ -th qubit;

Thus, as shown in the gene chain structure of Fig. 3, the quantum gene chains can be constructed as

$$\mathbf{q}_{\zeta c}^g = (\cos(w_{\zeta,1}^g) \cos(w_{\zeta,2}^g) \cdots \cos(w_{\zeta,\mu}^g) \cdots \cos(w_{\zeta,\lambda}^g)), \quad (23)$$

$$\mathbf{q}_{\zeta s}^g = (\sin(w_{\zeta,1}^g) \sin(w_{\zeta,2}^g) \cdots \sin(w_{\zeta,\mu}^g) \cdots \sin(w_{\zeta,\lambda}^g)). \quad (24)$$

Meanwhile, the linear mapping relationship between the

quantum gene chains and the parameter vector $\mathbf{p}^t \in (0, 2\pi)$ can be established by

$$p = \frac{p_{\max} - p_{\min}}{q_{\max} - q_{\min}} (q - q_{\min}) + p_{\min} = \pi (q + 1), \quad (25)$$

where $q \in \mathbf{q}_{\zeta c}^g$, $p \in \mathbf{p}^t$, $q_{\min} = -1$ and $q_{\max} = 1$ are respectively the lower and upper bounds of $\mathbf{q}_{\zeta c}^g$ and $\mathbf{q}_{\zeta s}^g$, $p_{\min} = 0$ and $p_{\max} = 2\pi$ are respectively the lower and upper bounds of \mathbf{p}^t .

Thus, according to Eqs. (25), the two quantum gene chains $\mathbf{q}_{\zeta c}^g$ and $\mathbf{q}_{\zeta s}^g$ can represent the cosine solutions and sinusoidal solutions of the parameter vector \mathbf{p}^t . When $\mathbf{q}_{\zeta c}^g$ represent the cosine solutions of \mathbf{p}^t , the cosine qubit probability amplitude $\cos(w_{\zeta,\mu}^g)$ can be mapped to the weight parameter $p'_\mu = \pi (\cos(w_{\zeta,\mu}^g) + 1)$; When $\mathbf{q}_{\zeta s}^g$ represent the sinusoidal solutions of \mathbf{p}^t , the sinusoidal qubit probability amplitude $\sin(w_{\zeta,\mu}^g)$ can be mapped to the weight parameter $p'_\mu = \pi (\sin(w_{\zeta,\mu}^g) + 1)$.

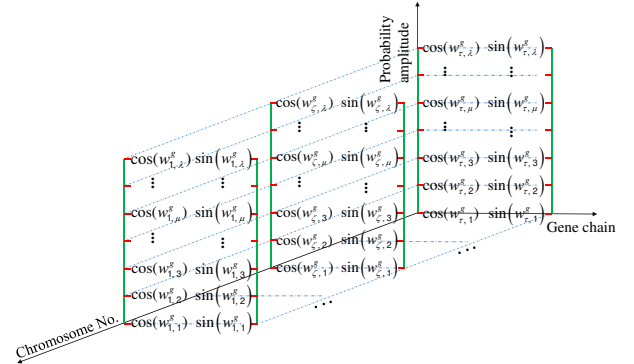


Figure 3 Structure of quantum gene chain

Simultaneously, the two quantum gene chains $\mathbf{q}_{\zeta c}^g$ and $\mathbf{q}_{\zeta s}^g$ can be updated by the increment (i.e., $\Delta w_{\zeta,\mu}^g$) of qubit phase (i.e., $w_{\zeta,\mu}^g$). For the directional parallel updating

$w_{\zeta,\mu}^g$, a phase selection matrix $\mathbf{M} = \begin{pmatrix} \cos w_{\zeta,\mu}^g & \cos w_{\zeta,0}^g \\ \sin w_{\zeta,\mu}^g & \sin w_{\zeta,0}^g \end{pmatrix}$ is

introduced, and then $\Delta w_{\zeta,\mu}^g$ can be expressed as

$$\begin{aligned} \Delta w_{\zeta,\mu}^g &= \text{sgn}|\mathbf{M}| \times \Delta w_0 \times \exp\left(-\frac{g}{g_{\max}}\right) \\ &= \text{sgn} \begin{vmatrix} \cos w_{\zeta,\mu}^g & \cos w_{\zeta,0}^g \\ \sin w_{\zeta,\mu}^g & \sin w_{\zeta,0}^g \end{vmatrix} \times \Delta w_0 \times \exp\left(-\frac{g}{g_{\max}}\right) \\ &= \text{sgn}(\sin(w_{\zeta,0}^g - w_{\zeta,\mu}^g)) \times \Delta w_0 \times \exp\left(-\frac{g}{g_{\max}}\right) \end{aligned}$$

$$= \text{sgn}(\Delta w_{\zeta,\mu}^g) \times \Delta w_0 \times \exp\left(-\frac{g}{g_{\max}}\right), \quad (26)$$

where $\text{sgn}(\bullet)$ is a sign function to judge the updating direction of $w_{\zeta,\mu}^g$; Δw_0 is the step size for $w_{\zeta,\mu}^g$ updating; g_{\max} is the maximum number of chromosome iterations; $\cos(w_{\zeta,\mu}^g)$ and $\sin(w_{\zeta,\mu}^g)$ are respectively the cosine and sinusoidal qubit probability amplitudes corresponding to the minimum of loss function, i.e., the mean square error E_{mse}^t . Here, E_{mse}^t is defined as

$$E_{mse}^t = \frac{1}{m} \sum_{k=1}^m (y_k^t - y_k^{t'})^2, \quad (27)$$

where m is the length of the output y^t ; $y_k^{t'}$ is the target output corresponding to the actual output y_k^t .

Eq. (26) reveals that in the process of updating $\cos(w_{\zeta,\mu}^g)$ and $\sin(w_{\zeta,\mu}^g)$, the updating direction of the qubit phase $w_{\zeta,\mu}^g$, which means whether $w_{\zeta,\mu}^g$ should be increased or decreased, can be determined by $\text{sgn}|M|$. Then, the quantum phase matrices θ^t , δ^t , γ^t , and ϕ^t can be directionally updated. Meanwhile, since $w_{\zeta,\mu}^g$ corresponds to two probability amplitudes (i.e., $\cos(w_{\zeta,\mu}^g)$ and $\sin(w_{\zeta,\mu}^g)$) on the quantum chromosome, the synchronous updating of the two quantum gene chains (i.e., $q_{\zeta c}^g$ and $q_{\zeta s}^g$) can be realized. As a result, θ^t , δ^t , γ^t , and ϕ^t can be updated in parallel. Thus, based on the above analysis, the quantum phase matrices θ^t , δ^t , γ^t , and ϕ^t can be directionally updated in parallel by the phase selection matrix M .

To sum up, in the proposed QGCCBNN, the quantum bidirectional transmission mechanism (see Eqs. (13-17)) is designed to establish the pre- and post-relationships of time series for readjusting the weight parameters according to the feedback from the output layer, so QGCCBNN can realize higher consistency between the input information and the overall memory of the network, and show better nonlinear approximation ability. Meanwhile, the quantum gene chain coding, in which the qubit probability amplitude real number coding (see Eqs. (21-24)) is adopted and the cosine and sinusoidal qubit probability amplitudes corresponding to the minimum loss function are compared with those of the current time by the phase selection matrix for the directional parallel updating the quantum phase matrices (see Eq. (26)), is constructed to transmit and update weight parameters for improving the global optimization ability and convergence speed of QGCCBNN. It is believed that the above advantages of QGCCBNN in nonlinear approximation ability and convergence speed can offer higher prediction accuracy and desirable computational efficiency to the proposed RUL prediction

method based on QGCCBNN.

3 RUL prediction based on QGCCBNN

3.1 Process of RUL prediction method

In the following, QGCCBNN is applied to the RUL prediction of RM. The process of our RUL prediction method is shown in Fig. 4, which is described in detail as follows:

Step 1: The original vibration acceleration data of RM running are collected.

Step 2: The power spectral entropy (PSE) [26, 27] is constructed based on the original vibration data.

Step 3: The trend index (TI) is constructed by PSE as the performance degradation feature.

Step 4: TI is input into QGCCBNN for training QGCCBNN according to the training process in subsection 3.3.

Step 5: The future performance degradation trend of RM is predicted by the well-trained QGCCBNN according to the prediction process in subsection 3.4.

Step 6: According to the predicted TI curve, the failure probability model is established to predict the failure time point (i.e., failure probability threshold point) and the RUL of RM.

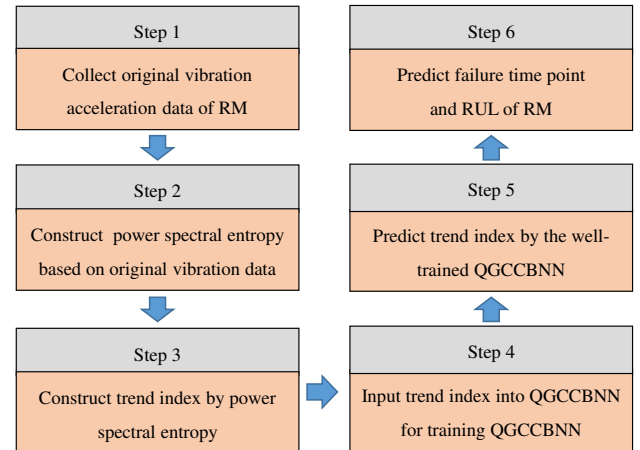


Figure 4 Implementation process of the proposed RUL prediction method based on QGCCBNN for RM.

3.2 Description of Trend Index (TI)

TI is constructed by PSE of a time series, and its construction process can be described as follows.

Firstly, the original time series $x_0 = (x_1 \ x_2 \ \cdots \ x_{nm})$ is reconstructed in the matrix form as

$$X_0 = \begin{pmatrix} x_1 & x_{m+1} & \cdots & x_{(i-1)m+1} & \cdots & x_{(n-1)m+1} \\ x_2 & x_{m+2} & \cdots & x_{(i-1)m+2} & \cdots & x_{(n-1)m+2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_m & x_{2m} & \cdots & x_{im} & \cdots & x_{nm} \end{pmatrix}. \quad (28)$$

Then, the PSE sequence $p_0 = [p_1, p_2, \dots, p_i, \dots, p_n]$ is constructed based on X_0 and its element p_i is expressed as

$$p_i = - \sum_{j=(i-1)m+1}^{im} \frac{1}{m} \left| \text{DFT}(x_j) \right|^2 \log_2 \left(\frac{1}{m} \left| \text{DFT}(x_j) \right|^2 \right), \quad (29)$$

where $\text{DFT}(\bullet)$ is the function of discrete Fourier transform.

Thus, p_0 can be reconstructed to a matrix as

$$P = \begin{pmatrix} p_1 & p_2 & \cdots & p_k & \cdots & p_r \\ p_2 & p_3 & \cdots & p_{k+1} & \cdots & p_{r+1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_s & p_{s+1} & \cdots & p_{k+s-1} & \cdots & p_{r+s-1} \end{pmatrix}, \quad (30)$$

where $r+s-1=n$, and k denotes the k -th column of P .

Finally, the TI sequence $TI = [TI_1, TI_2, \dots, TI_k, \dots, TI_r]$ is constructed based on P and TI_k can be expressed as

$$TI_k = - \sum_{u=k}^{k+s-1} \frac{1}{s} \left| \text{DFT}(p_u) \right|^2 \log_2 \left(\frac{1}{n} \left| \text{DFT}(p_u) \right|^2 \right), \quad (31)$$

where TI_k corresponds to the interval $[p_k, p_{k+s-1}]$ of the PSE sequence p_0 . Obviously, an interval $[TI_{k1}, TI_{k2}]$ of TI corresponds to the interval $[p_{k1}, p_{k2+s-1}]$ of p_0 .

Since TI can quantitatively measure the variation tendency of time series, it can be used to characterize the performance degradation feature of RM.

3.3 Training process of QGCCBNN

The training process of QGCCBNN mentioned in Step 4 in subsection 3.1 is now elaborated in this subsection.

Suppose the input series composed of TI is $x = [x_a, x_{a+1}, \dots, x_b]$.

x can be decomposed into the training sample matrix X and the target output vector y' , which are respectively written as

$$X = \begin{pmatrix} x_a & x_{a+1} & \cdots & x_{a+n-1} \\ x_{a+1} & x_{a+2} & \cdots & x_{a+n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{b-n} & x_{b-n+1} & \cdots & x_{b-1} \end{pmatrix} \text{ and } y' = \begin{pmatrix} x_{a+n} \\ x_{a+n+1} \\ \vdots \\ x_b \end{pmatrix},$$

where each row of X represents a training sample, and the dimension of a training sample is n ; each element in y' represents the corresponding training target of each training sample. Obviously, the total number of training samples is $b-a-n+1$.

Thus, we can summarize the training process of QGCCBNN as follows:

Step 1: All qubit phases of the quantum gene chains are initialized, namely the values in the interval of $[0, 2\pi]$ are randomly assigned to each qubit phase.

Step 2: The training sample matrix X is input into

QGCCBNN to compute the actual output vector y .

Step 3: All quantum phase matrix parameters quantum phase matrices are updated by the training algorithm in subsection 2.3, the updated values are taken as the initial values in the next iteration.

Step 4: Steps 2 and 3 are repeated to update the actual output vector y until the mean square error is less than the set threshold or the training steps reach the maximum value.

3.4 Performance degradation trend prediction by the well-trained QGCCBNN

Now, the well-trained QGCCBNN can be applied to predict the performance degradation trend. Here, the multi-step-ahead prediction method is adopted to conduct the prediction process.

First, the training sample $(x_{b-n+1}, x_{b-n+2}, \dots, x_b)$ is input into the well-trained QGCCBNN to calculate the output value \tilde{x}_{b+1} at $b+1$ time.

Second, the newest testing sample $(x_{b-n+1}, x_{b-n+2}, \dots, \tilde{x}_{b+1})$ is used as the input of the well-trained QGCCBNN to calculate the output value \tilde{x}_{b+2} at $b+2$ time.

Then, the newest testing sample $(x_{b-n+3}, x_{b-n+4}, \dots, \tilde{x}_{b+2})$ is used as the input of the well-trained QGCCBNN to calculate the output value \tilde{x}_{b+3} at $b+3$ time.

By that analogy, the newest testing sample $(\tilde{x}_{b-n+N}, \tilde{x}_{b-n+1+N}, \dots, \tilde{x}_{b-1+N})$ is used as the input of the well-trained QGCCBNN to calculate the output value \tilde{x}_{b+N} at $b+N$ time.

In this way, N predicted output values can be obtained.

3.5 RUL prediction by the failure probability model

In this subsection, according to TIs predicted by QGCCBNN, the failure probability model is established for RUL prediction of RM.

Firstly, the continuous performance degradation process of RM can be expressed as

$$x_t = x_0 + \lambda t + \sigma B_t, \quad (32)$$

where x_0 and x_t are the initial and accumulated values of performance degradation feature (i.e., TI); λ is the degradation rate; σ is the diffusion coefficient of the standard Brownian motion B_t . Since B_t is in compliance with normal distribution (i.e., $B_t \sim N(0, t)$), x_t obeys the following probability distribution:

$$x_t \sim N(x_0 + \lambda t, \sigma^2 t). \quad (33)$$

In order to obtain λ and σ^2 , the maximum likelihood estimation method [28, 29] is introduced, and the likelihood function $\psi(\lambda, \sigma)$ can be represented as

$$\psi(\lambda, \sigma) = \prod_{k=1}^n \Phi_k(x_k - x_{k-1}), \quad (34)$$

where x_{t_i} is the value of TI at time t_i and $\Phi(\bullet)$ is the

probability density function of standard normal distribution.

According to Eq. (32-33), $x_k - x_{k-1}$ obeys the following probability distribution:

$$x_k - x_{k-1} \sim N(\lambda(t_k - t_{k-1}), \sigma^2(t_k - t_{k-1})). \quad (35)$$

Then Eq. (34) can be further deduced as

$$\begin{aligned} \psi(\lambda, \sigma) &= \prod_{k=1}^n \Phi_k(x_k - x_{k-1}) \\ &= \prod_{k=1}^n \frac{1}{\sqrt{2\pi(t_k - t_{k-1})}} \exp\left(-\frac{(x_k - x_{k-1} - \lambda(t_k - t_{k-1}))^2}{2\sigma^2(t_k - t_{k-1})}\right). \end{aligned} \quad (36)$$

At this moment, Eq. (36) can be converted by the logarithmic function $\ln(\bullet)$ as follows:

$$\ln \psi(\lambda, \sigma) = \sum_{k=1}^n -\frac{1}{2} \ln(2\pi\sigma^2(t_k - t_{k-1})) - \frac{(x_k - x_{k-1} - \lambda(t_k - t_{k-1}))^2}{2\sigma^2(t_k - t_{k-1})}. \quad (37)$$

Then, according to the maximum likelihood estimation method, the derivative of $l(\lambda, \sigma)$ can be expressed as

$$\begin{aligned} \frac{\partial \ln \psi(\lambda, \sigma)}{\partial \lambda} &= \sum_{k=1}^n \frac{x_k - x_{k-1} - \lambda(t_k - t_{k-1})}{\sigma^2} = 0, \\ \frac{\partial \ln \psi(\lambda, \sigma)}{\partial \sigma} &= \frac{1}{\sigma} \left(-n + \sum_{k=1}^n \frac{(x_k - x_{k-1} - \lambda(t_k - t_{k-1}))^2}{\sigma^2(t_k - t_{k-1})} \right) = 0. \end{aligned} \quad (38)$$

Finally, according to Eq. (38), λ and σ^2 can be computed as follows:

$$\begin{aligned} \lambda &= \frac{\sum_{k=1}^n (x_k - x_{k-1})}{\sum_{k=1}^n (t_k - t_{k-1})}, \\ \sigma^2 &= \frac{1}{n} \sum_{k=1}^n \frac{(x_k - x_{k-1} - \lambda(t_k - t_{k-1}))^2}{(t_k - t_{k-1})}. \end{aligned} \quad (39)$$

After λ and σ^2 are obtained, the next step is to compute the failure probability $F(t)$ at time t , which can be expressed as

$$F(t) = 1 - P(x_t < w) = 1 - \Phi\left(\frac{w - x_0 - \lambda t}{\sqrt{\sigma^2 t}}\right). \quad (40)$$

where w is the threshold of TI.

Now, the failure probability model has been constructed. Assume that the starting prediction time is t_{start} , the predicted failure time is t_{failure} , and failure probability threshold is F_{failure} . Then, the predicted RUL $RUL_{\text{predicted}}$ of RM can be calculated as follows

$$\begin{aligned} RUL_{\text{predicted}} &= t_{\text{failure}} - t_{\text{start}} + 1 \times \Delta NT \\ &= F^{-1}(F_{\text{failure}}) - t_{\text{start}} + 1 \times \Delta NT \\ &= (NT_{\text{QGCCBNN}} - NT_1 + 1) \times \Delta NT \end{aligned} \quad (41)$$

where $t_{\text{failure}} = F^{-1}(F_{\text{failure}}) = NT_{\text{QGCCBNN}} \times \Delta NT$; $t_{\text{start}} = NT_1 \times \Delta NT$;

$NT_{\text{QGCCBNN}} = SN_{\text{failure}} + s - 1$ is the failure probability threshold point in power spectral entropy curve, SN_{failure} is the failure probability threshold point in TI curve, i.e., the time point corresponding to w in Eq. (40), s is the row number of the reconstructed matrix P (see Eq. (30)); $NT_1 = SN_{\text{predict}} + s - 1$ is the starting prediction point in power spectral entropy curve, SN_{predict} is the starting prediction point in TI curve; ΔNT is the interval between two adjacent time points in power spectral entropy curve.

4 Application of the RUL prediction method based on QGCCBNN

In this section, the full-life vibration acceleration data of double-row roller bearings collected by the University of Cincinnati [30] are used to verify the effectiveness of our QGCCBNN-based RUL prediction method.

4.1 Test bed

Fig. 5 illustrates the bearing test rig and sensor placement that are used to collect the vibration acceleration data of bearings. Four ZA-2115 double-row roller bearings made by Rexnord are installed on the rotating shaft of the bearing test bench. A motor is used to drive the rotating shaft through a friction belt with a constant speed of 2000 r/min. Spring mechanisms are used to apply 6000 pounds of radial load onto the rotating shaft and bearings. High-sensitivity ICP accelerometers are installed on the bearings to collect the vibration acceleration data every 10 min. The sampling frequency is 20 kHz, and the sampling length is 20480 points.

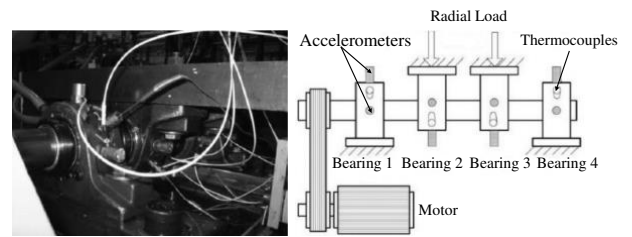


Figure 5 Bearing test rig and sensor placement illustration

4.2 RUL prediction of the second bearing on test bed

In this experiment, the outer race failure occurs in the second bearing after continuous running for 9840 min (i.e., about seven days). Here, PSE is constructed based on the vibration acceleration data (i.e., 984 samples). As shown in Fig. 6, PSE remains stable from the starting sample point to the 805th point, which means that the bearing is in the normal operation period. After that, PSE decreases slowly

from the 806th to the 950th point, indicating that the bearing is in the initial degradation stage (i.e., early fault stage); from the 951st to the last point, the PSE curve becomes more and more irregular, which reveals that the outer race defect of the bearing is expanding rapidly, namely, serious failure occurs.

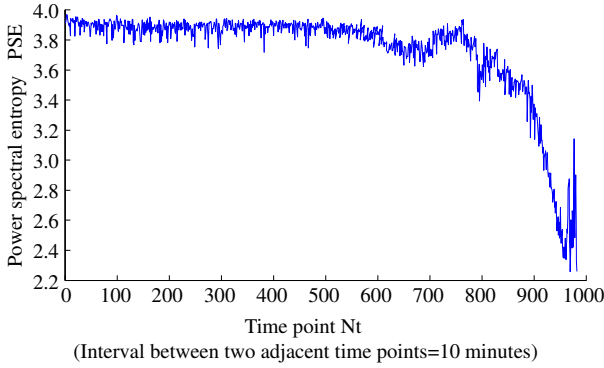


Figure 6 Power spectral entropy of the second bearing during seven days

In order to quantitatively measure the variation tendency of PSE, the next step is to construct the trend index (TI) as the performance degradation feature. As shown in Fig. 7, a total of 729 TIs corresponding to 984 samples are obtained. From the starting point to the 550th point, TI changes slowly, indicating that the bearing is in the normal operation stage. From the 551st to the 695th point, TI starts to ascend, indicating that the bearing is in the stage of initial degradation. After that, the value of TI rises sharply, which means that the bearing is in the stage of accelerated degradation, namely, the outer race failure occurs.

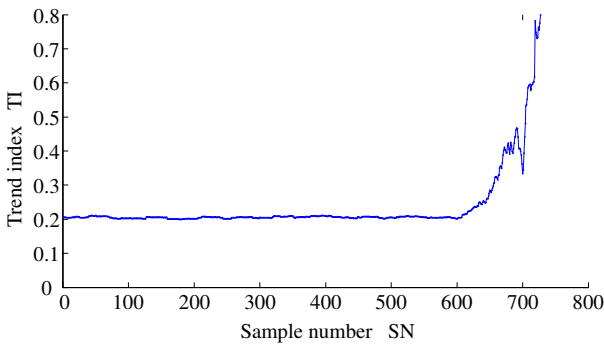


Figure 7 Trend index of the second bearing during seven days

4.2.1 Prediction results

Because the initial degradation stage and its subsequent degradation stage are the most important time intervals in

performance degradation trend observation and RUL prediction, the TI points from the 551th to the 650th are taken as training samples, and the 651st TI point and its subsequent data (i.e., the remaining 79 TI points) are the testing samples. Here, the parameters of the proposed RUL prediction method based on QGCCBNN are set as follows: in the theory model of QGCCBNN, the input-layer dimension is $n=5$, the hidden-layer dimension is $m=9$, and the output-layer dimension is $p=1$; in the training algorithm (i.e., quantum gene chain coding) for QGCCBNN, the size of genome is $\tau=50$, the step size for qubit phase $w_{\zeta,\mu}^g$ updating is $\Delta w_0=0.07\pi$, the maximum number of chromosome iterations (i.e., the maximum training step size) is $g_{max}=50$, and the mean square error threshold $E_{mse-min}^t=5\times 10^{-6}$. The processes of training QGCCBNN and predicting TI by QGCCBNN can be seen in subsections 3.3 and 3.4. The predicted TI values are shown in Fig. 8. Obviously, the predicted TI values within the interval [651,729] are very close to the actual ones within this interval. Namely, the predicted TI curve is entirely consistent with the changing trend of the actual TI value.

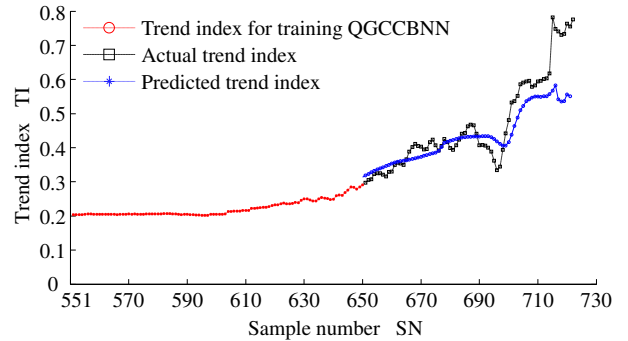


Figure 8 Trend index predicted by QGCCBNN

Because of the high accuracy of the TI values predicted by QGCCBNN, it is reliable to establish the failure probability model according to the predicted TI values. According to the probability failure model in subsection 3.5, the predicted TI values are taken as the input of Eq. (39) to calculate the values of parameters λ and σ^2 , and then the cumulative distribution function $F(t)$ of failure (i.e., failure probability) can be obtained according to Eq. (40), as shown in Fig. 9.

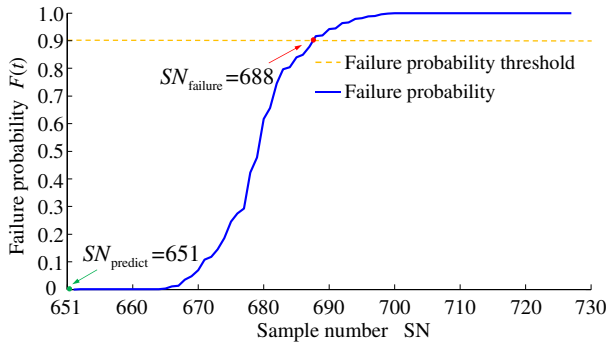


Figure 9 Failure probability predicted by QGCCBNN

The next step is to predict RUL of the second bearing. As shown in Fig. 9, because RUL can generally be represented by an interval between the starting prediction point SN_{predict} and the failure probability threshold point SN_{failure} corresponding to the failure probability of 0.9, the interval $[SN_{\text{predict}}, SN_{\text{failure}}]$, which corresponds to the interval $[NT_1, NT_{\text{QGCCBNN}}]$ shown in Fig. 10, can be regarded as RUL of the second bearing, where $NT_1 = SN_{\text{predict}} + s - 1$ and $NT_{\text{QGCCBNN}} = SN_{\text{failure}} + s - 1$ according to Eq. (41). Here, $s=256$ is the row number of \mathbf{P} mentioned in subsection 3.2. Thus, RUL of the second bearing can be predicted as $RUL_{\text{predicted}} = (NT_{\text{QGCCBNN}} - NT_1 + 1) \times \Delta NT = 6.33\text{h}$, where $\Delta NT = 10\text{min}$ is the interval between two adjacent time points shown in Fig. 10.

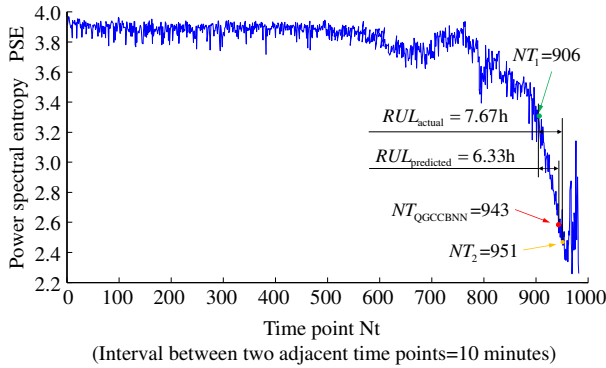


Figure 10 RUL predicted by QGCCBNN

4.2.2 Discussion

In this subsection, the advantages of the RUL prediction method based on QGCCBNN are verified by comparative analysis.

First, the nonlinear approximation capability of QGCCBNN is compared with that of deep autoencoder and deep neural network (DADNN), gated recurrent unit

neural network (GRUNN), or fully complex-valued radial basis function neural network (FCRBFNN). In order to compare the four prediction methods under the same conditions, the performance degradation feature in the latter three prediction methods is also represented by TI, which is the same as that of our proposed method based on QGCCBNN. At the same time, the maximum training step size, the number of neurons in each layer, and the mean square error threshold of the latter three methods are set as the same values as those of QGCCBNN respectively. The learning curves of these methods are shown in Fig. 11. As can be seen, at the training step $N_s = 20^{\text{th}}$ and its subsequent training steps, the mean square error of QGCCBNN is much less than that of DADNN, GRUNN or FCRBFNN, namely, the nonlinear approximation capability of QGCCBNN is much better than that of the latter three methods. The underlying reason is that the quantum bidirectional transmission mechanism is designed in QGCCBNN to establish the pre- and post-relationships of time series for avoiding the lack of reverse feedback from the output layer to the input layer in the latter three methods. This means that QGCCBNN can remember the overall regulation of time series, thus achieving better nonlinear approximation capability.

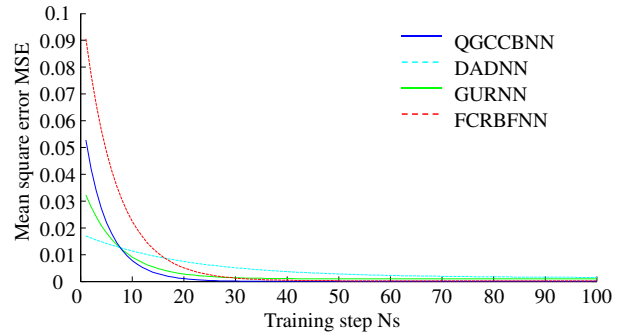


Figure 11 Learning curves of different neural networks

Second, the prediction accuracy of the RUL prediction method based on QGCCBNN is compared with that of the prediction methods based on DADNN, GRUNN, FCRBFNN or LS-SVM. The performance degradation feature of the latter four methods is also expressed by TI. The maximum training step size, the number of neurons in each layer, and the mean square error threshold of DADNN, GRUNN and FCRBFNN are set as the same as those of QGCCBNN, respectively. In LS-SVM, the Gaussian kernel is set as its kernel function, and its regularization parameter is optimized by s-fold cross validation algorithm. Moreover, the failure probability models adopted by the latter four methods are identical to that of QGCCBNN. Figs. 12-15 show the predicted TI,

failure probability and RUL of the second double-row roller bearing respectively by the latter four methods.

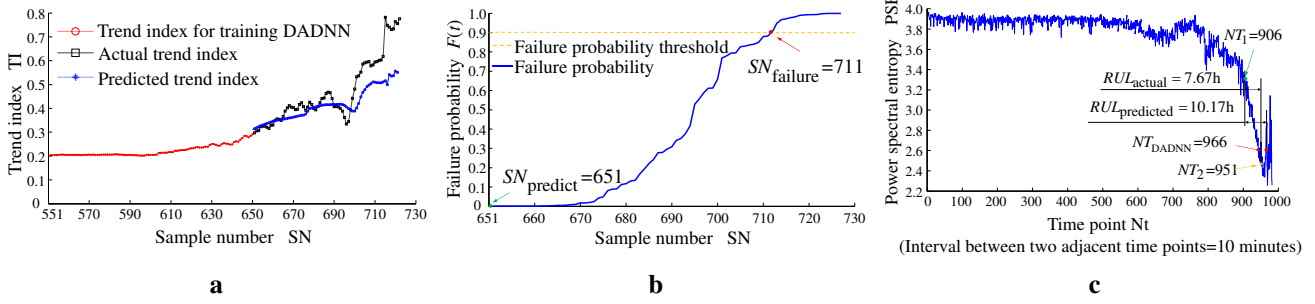


Figure 12 (a) Trend index predicted by DADNN; (b) Failure probability predicted by DADNN; (c) RUL predicted by DADNN

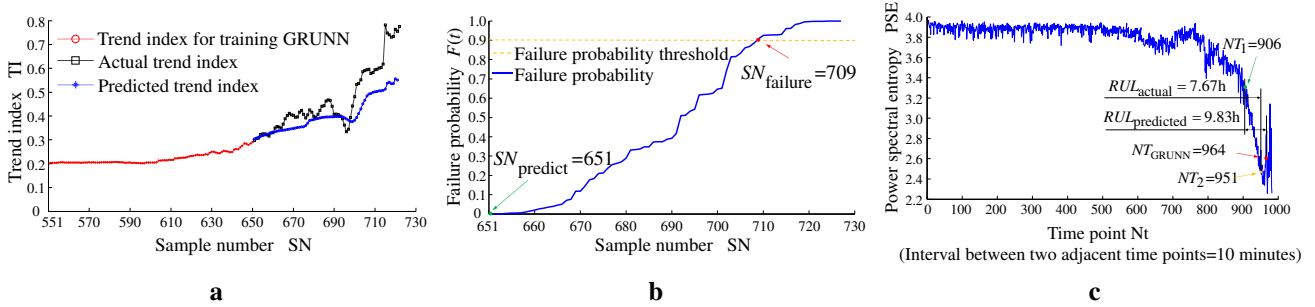


Figure 13 (a) Trend index predicted by GRUNN; (b) Failure probability predicted by GRUNN; (c) RUL predicted by GRUNN

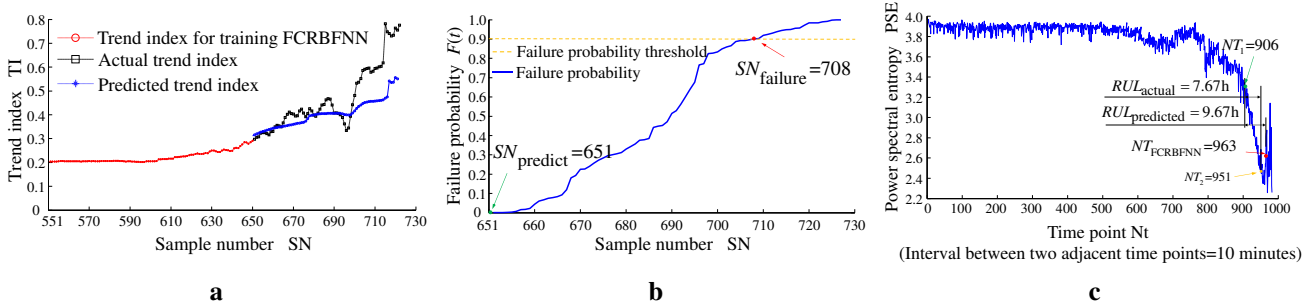


Figure 14 (a) Trend index predicted by FCRBFNN; (b) Failure probability predicted by FCRBFNN; (c) RUL predicted by FCRBFNN

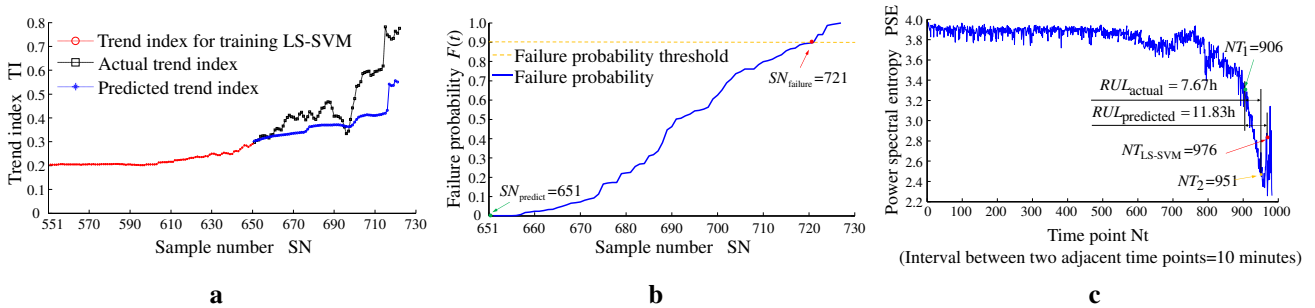


Figure 15 (a) Trend index predicted by LS-SVM; (b) Failure probability predicted by LS-SVM; (c) RUL predicted by LS-SVM

From Figs. 8, 9, 12(a)-(b), 13(a)-(b), 14(a)-(b) and 15(a)-(b), it is not difficult to find that the distance between

the failure probability threshold point predicted by QGCCBNN and the starting point of serious failure (i.e.,

the 951st point) is the smallest among the five methods, which means that the TI curve predicted by our QGCCBNN is the closest to the actual TI curve compared with DADNN, GRUNN, FCRBFNN or LS-SVM. Generally, it is believed that the occurrence of serious failure means the complete failure of the bearing. Hence, the error between the predicted RUL by our QGCCBNN-based method and the actual RUL ($RUL_{actual} = (NT_2 - NT_1 + 1) \times \Delta NT = 7.67$ h) is also the smallest. As shown in Figs.10, 12 (c), 13 (c), 14 (c) and 15 (c), the predicted RULs by our QGCCBNN-based method and the latter four methods are 6.33 h, 10.17 h, 9.83 h, 9.67 h and 11.83 h, respectively. In other words, our method is the most accurate and reliable. This can be attributed to the better nonlinear approximation capability of QGCCBNN brought by the quantum bidirectional transmission mechanism and its superior global optimization ability given by quantum gene chain coding. Therefore, it is expected that the above advantages of QGCCBNN can improve the prediction accuracy of our RUL prediction method.

Then, in order to quantitatively evaluate the prediction accuracy and stability of the five RUL prediction methods, the mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2) are taken as evaluation indicators to statistically measure the levels of prediction accuracy and stability of the five methods, which are respectively defined as

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2, \quad (42)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i|, \quad (43)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2}, \quad (44)$$

$$R^2 = 1 - \frac{MSE}{VAR} = 1 - \frac{\left(\sum_{i=1}^n (y_i - y'_i)^2 \right)}{\left(\sum_{i=1}^n (y_i - \bar{y}_i)^2 \right)}, \quad (45)$$

where n is the number of predicted sample points, y'_i is the predicted TI value of the i -th sample point, y_i is the actual TI value of the i -th sample point, \bar{y}_i is the average of actual TI values of y_i , and VAR is the sample variance. Here, $R^2 \in (0,1)$; the closer the value of R^2 to 1, the higher the prediction accuracy of the method. Statistical analyses on the MSE , MAE , $RMSE$, and R^2 values of the five prediction methods are respectively conducted for 50 times, and the average values of MSE , MAE , $RMSE$, and R^2 of these methods in the 50 repeated predictions, i.e., \overline{MSE} , \overline{MAE} , \overline{RMSE} and $\overline{R^2}$, are respectively calculated, as shown in Table 1.

Table 1 Statistical analysis results of the prediction accuracy and stability for the second bearing by the methods based on QGCCBNN, DADNN, GRUNN, FCRBFNN and LS-SVM.

Prediction method	\overline{MSE}	\overline{MAE}	\overline{RMSE}	$\overline{R^2}$
QGCCBNN	0.0042	0.0755	0.0337	0.8758
DADNN	0.0054	0.1677	0.0752	0.5945
GRUNN	0.0078	0.1358	0.0841	0.6258
FCRBFNN	0.0062	0.1894	0.1573	0.3659
LS-SVM	0.0150	0.3610	0.1126	0.1348

It can be seen from Table 1 that, compared with the other four methods, the values of \overline{MSE} , \overline{MAE} , and \overline{RMSE} for QGCCBNN are the smallest and that the value of $\overline{R^2}$ are the closest to 1, which indicates that QGCCBNN has higher prediction accuracy and more reliable stability.

Finally, the computational time (i.e., the sum of training time, TI prediction time and RUL prediction time) of our RUL prediction method based on QGCCBNN is compared with that of the prediction method using DADNN, GRUNN, FCRBFNN or LS-SVM. The computational time is measured by MATLAB 2016 in the following hardware configuration environment: 8G RAM and 3.2 GHz Intel CPU. The average computational time of the 50-fold cross-validation executions for each prediction method in the RUL prediction experiment is recorded on the same full-time data, as plotted in Fig. 16. The computational time of QGCCBNN is only 14.85 s, and those of DADNN, GRUNN, FCRBFNN, and LS-SVM are respectively 32.68 s, 29.59 s, 23.74 s, and 19.93s. Obviously, the computational time of QGCCBNN is much shorter than that of DADNN, GRUNN, FCRBFNN or LS-SVM, which can be attributed to the superiority of parallel computing brought by quantum gene chain coding.

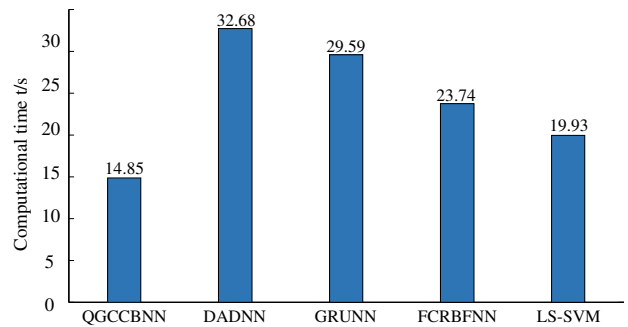


Figure 16 Computational time for RUL prediction of the second bearing by the methods based on QGCCBNN, DADNN, GRUNN, FCRBFNN and LS-SVM

5 Conclusions

In conclusion, we propose a novel recurrent neural network named quantum gene chain coding bidirectional neural network (QGCCBNN) to predict RUL of RM. The main contributions of this paper are as follows:

(1) In our QGCCBNN, a quantum bidirectional transmission mechanism is designed to establish the pre- and post-relationships of time series for readjusting the weight parameters according to the feedback from the output layer, so QGCCBNN can realize higher consistency between the input information and the overall memory of the network, and then has better nonlinear approximation ability. To the best of our knowledge, the quantum bidirectional transmission mechanism is the first attempt for neural network data transmission and the experiment results confirm its effectiveness.

(2) In the process of training QGCCBNN, quantum gene chain coding is innovatively constructed to transmit and update data to avoid the gradient disappearance, gradient explosion, slow convergence speed and large time cost caused by traditional gradient descent method. As a consequence, QGCCBNN has better global optimization ability and faster convergence speed.

(3) Due to the advantages of QGCCBNN in nonlinear approximation ability, global optimization ability and convergence speed, the RUL prediction method based on QGCCBNN can realize higher prediction accuracy and lower computation cost.

(4) Comparative analysis on the RUL prediction results of a double-row roller bearing by different methods demonstrates that our proposed method is effective in RUL prediction of RM and superior to the other methods.

6 Declaration

Acknowledgements

The authors sincerely thanks to Professor Bao-Ping TANG of Chongqing University for his critical discussion and reading during manuscript preparation. Meanwhile, the authors would like to thank the insightful and valuable comments from the editors and anonymous reviewers.

Funding

Supported by Project of Open Foundation for The State Key Laboratory of Mechanical Transmissions (Grant No. SKLMT-KFKT-201718), Postdoctoral Science Foundation of China (Grant No. 2016M602685), and Strategic Cooperation Project between Sichuan University and Luzhou Municipal People's Government (Grant No. 2018CDLZ-30).

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article.

Authors' contributions

Feng Li was in charge of all experiments and also wrote the manuscript; Yang-Yang Cheng was responsible for the development of machine learning methods and statistical analyses; Bao-Ping TANG provided necessary help with all parts of this article; Xue-Ming Zhou helped to collect and organize materials; Rui-Ping Xiong checked the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing financial interests.

Consent for publication

Not applicable

Ethics approval and consent to participate

Not applicable

References

- [1] Y G Lei, N P Li, J. Lin, et al. A Model-Based Method for Remaining Useful Life Prediction of Machinery. *IEEE Transactions on Reliability*, 2016, 65(3): 1314-1326.
- [2] C Lucio, F Francesco, F Alessandro, et al. Statistical Spectral Analysis for Fault Diagnosis of Rotating Machines. *IEEE Transactions on Industrial Electronics*, 2018, 65(5): 4301-4310.
- [3] A Brkovic, D Gajic, J Gligorijevic, et al. Early fault detection and diagnosis in bearings for more efficient operation of rotating machinery. *Energy*, 2017, 136(16): 63-71.
- [4] Y Qian, R Yan, R X Gao. A multi-time scale approach to remaining useful life prediction in rolling bearing. *Mechanical Systems and Signal Processing*, 2017, 83: 549-567.
- [5] P J García-Nieto, E García-Gonzalo, F S Lasheras, et al. Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. *Reliability Engineering and System Safety*, 2015, 138: 219-231.
- [6] Y B Li, M Q Xu, H Y Zhao, et al. A new rotating machinery fault diagnosis method based on improved local mean decomposition. *Digital Signal Processing*, 2015, 46: 201-214.
- [7] L L Li, D J Ma, Z G Li. Residual Useful Life Estimation by a Data-Driven Similarity-Based Approach. *Quality and Reliability Engineering International*, 2017, 33(2): 231-239.
- [8] B P Cai, Y B Zhao, H L Liu, et al. A Data-Driven Fault Diagnosis Methodology in Three-Phase Inverters for PMSM Drive Systems. *IEEE Transactions on Power Electronics*, 2017, 32(7): 5590-5600.
- [9] J Saez-Gallego, J M Morales, M Zugno, et al. A Data-Driven Bidding Model for a Cluster of Price-Responsive Consumers of Electricity. *IEEE Transactions on Power Systems*, 2016, 31(6): 5001-5011.
- [10] L Ren, Y Sun, J Cui, L Zhang. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *Journal of Manufacturing Systems*, 2018, 48: 71-77.

- [11] M Karuppasampandian, R Narmatha Banu, P M Manobalaa. Static ATC Estimation Using Fully Complex-Valued Radial Basis Function Neural Network. *Artificial Intelligence and Evolutionary Algorithms in Engineering Systems*, 2015, 325: 763-772.
- [12] A Z Hinch, M Tkouat. Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network. *Procedia Computer Science*, 2018, 127: 123-132.
- [13] T Xiao, B P Tang, Y Qin, et al. Degradation trend prediction of rolling bearing based on manifold learning and least squares support vector machine. *Journal of Vibration and Shock*, 2015, 34(9): 149-153.
- [14] Y T Wu, Y M Yuan, S P Dong, et al. Remaining Useful Life Estimation of Engineered Systems using vanilla LSTM Neural Networks. *Neurocomputing*, 2018, 275: 167-179.
- [15] Y D Tang, Y C Huang, Z Y Wu, et al. Question detection from acoustic features using recurrent neural network with gated recurrent unit. *IEEE International Conference On Acoustics, Speech And Signal Processing Proceedings*, 2016, 170: 6125-6129.
- [16] C B Li, S K Li, Y Q Liu. A least squares support vector machine model optimized by moth-flame optimization algorithm for annual power load forecasting. *Applied Intelligence*, 2016, 45(4): 1166-1178.
- [17] N. P. Li, Y. G. Lei, J. Lin, et al. Estimating the remaining useful life of bearings using a neuro-local linear estimator-based method. *IEEE Transactions on Industrial Electronics*, 2017, 62(5): 7762-7773.
- [18] O S Sidelnikov, A A Redyuk, S Sygletos. Dynamic neural network-based methods for compensation of nonlinear effects in multimode communication lines. *Quantum Electronics*, 2017, 47(12): 1147-1149.
- [19] J F D Paz, P Chamoso, F D L Prieta, G Villarrubia. Artificial neural networks used in optimization problems. *Neurocomputing*, 2018, 272: 10-16.
- [20] D Konar, S Bhattacharyya, B K Panigrahi, et al. A quantum bi-directional self-organizing neural network (QBDSOINN) architecture for binary object extraction from a noisy perspective. *Applied Soft Computing Journal*, 2016, 46: 731-752.
- [21] P C Li, Y Zhao. Model and Algorithm of Sequence-Based Quantum-Inspired Neural Networks. *Chinese Journal of Electronics*, 2018, 27(1): 9-18.
- [22] D L Deng, X P Li, D Sarma. Quantum Entanglement in Neural Network States. *Nature Physics*, 2017, 7(2): 447-450.
- [23] R Narayan, S Chakraverty, V P Singh. Quantum neural network based machine translator for English to Hindi. *Applied Soft Computing Journal*, 2016, 38(): 1060-1075.
- [24] G Alexandru, K Theodoros, K Elham. Verification of Quantum Computation: An Overview of Existing Approaches. *Theory of Computing Systems*, 2019, 63(4): 715-808.
- [25] C Earl, T Barbara, V Christophe. Roads towards fault-tolerant universal quantum computation. *Nature*, 2017, 549(7671): 172-179.
- [26] Y J Ji, X B Wang, Z B Liu, et al. EEMD-based online milling chatter detection by fractal dimension and power spectral entropy. *International Journal of Advanced Manufacturing Technology*, 2017, 92(1-4): 1185-1200.
- [27] Y M Dai, H S Zhang, X G Mao, et al. Complexity-entropy causality plane based on power spectral entropy for complex time series. *Physica a-statistical mechanics and its applications*, 2018, 509: 501-514.
- [28] Y. N. Chen. Semiparametric Time Series Models with Log - concave Innovations: Maximum Likelihood Estimation and its Consistency. *Scandinavian Journal of Statistics*, 2015, 42(1): 1-31.
- [29] T Ivan, M Kruno, S Nenad, B Ivan. Non-linear Least Squares and Maximum Likelihood Estimation of Probability Density Function of Cross-Border Transmission Losses. *IEEE Transactions on Power Systems*, 2018, 33(2): 2230-2238.
- [30] J Lee, H Qiu, G Yu, Rexnord technical services, "bearing data set", IMS, university of Cincinnati, NASA Ames prognostics data repository, <http://ti.arc.nasa.gov/project/prognostics-data-repository>, NASA Ames, Moffett Field, CA, 2007.

Biographical notes

Feng Li, born in 1982, is currently an associate professor and supervisor for master student in *School of Mechanical Engineering at Sichuan University, China*. He received his master and PhD degrees from *Chongqing University, China*, in 2008 and 2011, respectively. His main research interests include equipment condition monitoring, fault diagnosis as well as prognostics and health management (PHM).

Tel: +86-18382385401; E-mail: lifeng19820501@163.com

Yang-Yang Cheng, born in 1994, is currently a graduate student in *School of Mechanical Engineering at Sichuan University, China*. He received his bachelor's degree from *Nanjing University of Aeronautics and Astronautics, China*, in 2017. His research interest is fault diagnosis and degradation trend prediction of mechanical equipment.

E-mail: cyoungyoung@126.com

Bao-Ping Tang, born in 1971, is currently a professor and a PhD candidate supervisor at *Chongqing University, China*. He received his master and PhD degrees from *Chongqing University, China*, in 1996 and 2003, respectively. His main research interests include equipment condition monitoring and fault diagnosis, virtual instrument and wireless sensor network.

E-mail: bptang@cqu.edu.cn

Xue-Ming Zhou, born in 1978, is currently an engineer and general manager of *Chongqing Ming's Min Mechanotronics Co., Ltd., China*. He received his master degree from *Chongqing University, China*, in 2009. His main research interests include equipment condition monitoring, fault diagnosis as well as computer aided design and manufacturing.

E-mail: zhou.xueming@cqlap.com.cn

Rui-Ping Xiong, born in 1968, is currently an associate professor and supervisor for master student in *School of Mechanical Engineering at Sichuan University, China*. He received his master and PhD degrees from *Sichuan University, China*, in 1998 and 2006, respectively. His main research interests include electromechanical transmission and intelligent control, equipment condition monitoring and fault diagnosis.

E-mail: xiongruiping@163.com