

The diagnostic analysis of the fault coupling effects in planet bearing

Song Xue¹, Congsi Wang^{1*}, Ian Howard², Peiyuan Lian¹, Gaige Chen¹, Yan Wang¹, Yuefei

Yan¹, Qian Xu³, Yu Shi⁴, Yu Jia⁴, Yuanpeng Zheng⁵, Na Wang³

¹Key laboratory of Electronic Equipment Structure Design, School of Mechano-Electronic Engineering, Xidian University, Xi'an, Shaanxi, China

²Department of Mechanical Engineering, Curtin University, Bentley, Western Australia, Australia

³XinJiang Astronomical Observatory, China Academy of Sciences, Urumqi 830011, China.

⁴Department of Mechanical Engineering, University of Chester, Chester CH1 4BJ, UK.

⁵CETC No.54 Research Institute, Shijiazhuang 050081, China.

* Corresponding author. Phone: (+86) 029-88203040.

E-mail: congsiwang@163.com

Abstract: The purpose of this paper is to investigate the fault coupling effects in the planet bearing as well as the corresponding vibration signatures in the resultant vibration spectrum. In a planetary gear application, the planet bearing can not only spin around the planet gear axis, but also revolve about the sun gear axis and this rotating mechanism poses a big challenge for the diagnostic analysis of the planet bearing vibration spectrum. In addition, the frequency component interaction and overlap phenomenon in the vibration spectrum

caused by the fault coupling effect can even worsen the diagnosis results. To further the understanding of the fault coupling effects in a planet bearing, a 34 degree of freedom planetary gear model with detailed planet bearing model was established to obtain the dynamic response in the presence of various bearing fault scenarios. The method of modelling the bearing distributed faults and localized faults has been introduced in this paper, which can be further incorporated into the planetary gear model to obtain the faulted vibration signal. The “benchmark” method has been adopted to enhance the planet bearing fault impulses in the vibration signals and in total, the amplitude demodulation results from 20 planet bearing fault scenarios have been investigated and analyzed. The coherence estimation over the vibration frequency domain has been proposed as a tool to quantify the fault impact contribution from different fault modes and the results suggested that the outer raceway fault contributes most to the resultant planet bearing vibration spectrum in all the investigated fault scenarios.

Keyword: Planet bearing fault, Fault coupling effect, Lumped-parameter model, Amplitude demodulation, Coherence estimation

1. Introduction

Planetary gears, also known as epicyclic gear sets, commonly include several planet gears meshing simultaneously to split the torque and power. For the planet bearing, its inner race is usually fixed to the planet pin connected firmly to the planet carrier, and its outer race is connected firmly to the planet gear bore hole. This configuration allows the planet bearing to spin around the planet axis with the planet gear, and at the same time to revolve about

the sun and ring gear common axis with the planet carrier. However, this rotating mechanism poses a big challenge and remains as one of the most difficult scenarios for bearing condition monitoring. In addition, numerous industry experience has shown that there is usually more than one fault type in the bearing unit. These faults are normally in different positions, with different fault modes and fault severities. The interaction as well as the mutual coupling effect between these faults further bring greater challenge to the planet bearing fault diagnosis.

Mathematical models can be used to further the understanding of the vibration generation mechanisms in the mechanical unit as well as the complex dynamic response in the presence of various fault types [1, 2]. Over the past decades, extensive research work has been carried out in the rolling bearing system dynamic behavior with and without the presence of bearing defects [3]. Hongrui et.al gave a comprehensive review on rolling element bearing modelling methodology and five groups of bearing models have been classified, among which the lumped-parameter dynamic model that has been used intensively by the researchers to investigate the rolling element bearing dynamic behavior [4]. One of the most representative bearing dynamic models was proposed by Gupta [5-8], where the interactions between different components, i.e., the ball-raceway interaction, were all included, and each bearing component had six degrees of freedom. Based on Gupta's work, researchers have developed various models to investigate the misalignment [9], cage related irregularity [10, 11] and so on [12-14].

Localized bearing faults include cracks, pits and spalls on the bearing components. When a loaded ball encounters a spall, there will be a sudden relaxation of the loading and the

transient forces which are produced will result in rapid accelerations of the bearing components and complex motions can occur including oscillatory contact and impacts between different components. Singh et al. gave an extensive review of vibration modelling of rolling element bearings with various localized defects [15]. McFadden et al. proposed a periodic impulse-train model to simulate the vibration response of a localized single point defect on the inner race of a rolling element under constant load [16]. However, this model does not include the physical parameters of a bearing, such as masses of bearing components. On the other side, the lumped-parameter model used to study the bearing dynamic response has also been adopted here to include the impact due to the presence of localized bearing defects. Feng et al. extended Fukata's two degree of freedom lumped-parameter model by including the effect of slippage of the cage and rolling elements as well as the effect of localized faults in the inner and outer raceway [17]. Later, Sawalhi et al. further developed the work of Feng et al. and a five degree of freedom analytical model was established to simulate the vibration response of a defective ball bearing in a gearbox having localized line spalls [18]. An important feature of this model was including an additional mass-spring-damper system to simulate the typical high-frequency resonant response of a bearing. However, the internal excitation considered in their models was only limited to the time varying displacement and later on, Liu extended the model and considered both the time-varying displacement excitation and the time-varying contact stiffness excitation in the bearing fault dynamic modelling [19].

Beside the localized bearing fault, the distributed bearing fault also played an important factor that can contribute to the abnormal vibration response in a bearing system. The

distributed bearing fault is usually caused by the manufacture error or the non-uniform surface wear during its operation and mainly includes surface waviness and roughness on the bearing components. Numerous investigations have also been carried out to study the vibration mechanism caused by these distributed bearing faults. Aktürk studied the radial and axial vibrations of a rigid shaft supported by a pair of angular contact ball bearings [20]. The effect of the waviness of the bearing's running surface on the vibration of the shaft was investigated. Tallian and Gustafsson presented a linearized dynamic model to analysis the vibration caused by waviness [21, 22]. Wardle reported that the waviness produced most of the severe vibrations and noise problem in bearings and later they theoretically and experimentally investigated the relation between the frequency harmonic of waviness and bearing dynamic performance [23, 24]. Jang and Jeong considered the centrifugal force and gyroscopic moment of the balls and investigated the bearing vibration due to the waviness [25, 26]. Bai et.al established a dynamic model with 5 degrees of freedom to study the dynamic performance due to the ball bearing surface waviness at the high speed [27].

As discussed above, the localized bearing fault and distributed bearing fault have been investigated separately by numerous researchers. However, little studies have been published to investigate the coupling effect between these faults, which can produce an even more complex vibration spectrum [28]. In addition, for these parallel shaft applications, the outer raceway has been constrained while the inner raceway is rotating with the shaft in the above dynamic models. However, the arrangement in the planet bearing is different when both the outer raceway and inner raceway are rotating, which can result in a different vibration response compared to those in the parallel shaft applications. For the planet

bearing dynamic modelling aspect, Jain and Hunt considered the ring gear deformation and a planet bearing defect, developed a dynamics model of a planetary gearbox, and analyzed the spectral characteristic vibration response measured from the ring gear rim [29]. Zhipeng developed the planet bearing vibration signal model for each fault case, considering the modulation effects of load zone variation, time-varying angle between the gear pair meshing and fault induced impact force, as well as the time varying vibration transfer path and then gave the explicit equations for calculating the characteristic frequency of each fault case [30]. However, little has been known of the planet bearing response with fault coupling effects. All in all, there is still a need to further investigate the fault coupling effect in a bearing system, especially in the planetary gear applications.

This paper aims to investigate the fault coupling effects in the planet bearing and this paper is organized as follows. In section 2, a 34 degree of freedom planetary gear lumped-parameter model was proposed, including detailed modelling of 5 degrees of freedom for each planet bearing. In the following section, the approach on how to evaluate the gear mesh stiffness and the bearing stiffness was briefly introduced. Next, the bearing distributed fault and localized fault for different bearing components were modelled mathematically in section 3. In section 4, the numerical simulation results for the planet bearing distributed faults, the planet bearing localized faults as well as the coupling between different faults scenarios have been presented and then their results have been summarized and discussed in section 5.

2. The planet bearing dynamic model in the planetary gear system

2.1 The planetary gear system model

Mathematical planetary gear models can be used to help understand the complex dynamic response from the planetary gear system and such mathematical models can be very valuable in planetary gear fault diagnostics and prognostics in order to produce signals with well-defined characteristics. For example, a range of different planetary fault types and locations can be simulated in the mathematical model to obtain the desired fault signals, whose characteristics frequencies could be further analyzed. Whilst, it would not be economical to actually experimentally test the number of faults of each type required to accomplish the diagnostic analysis. Furthermore, the simulated fault signals could also be used to train neural networks to perform diagnostics and prognostics [31]. The planetary gear system considered in this study has been developed based on our previous study [2], where a single-stage planetary gear set including an input motor, one sun gear, three planet gears, three planet bearings, one ring gear, one carrier arm and an output load has been developed, as shown in Fig. 1(a). A brief description of the planetary gear system has been shown in Fig. 1(b). In the model, each of these gear components has three DOFs: transverse motions in the x-axis and y-axis, and rotation. It should be noted that the transverse motions of the sun, planet, ring and carrier ($x_j, y_j, j=s, r, c, p1, p2, p3$) were simulated with respect to a rotating frame of reference fixed to the carrier. In the planet bearing arrangement, the inner race is connected to the planet pin, which is clamped on both sides of the carrier arm. The outer race is connected to the bore of the planet gear. In the configuration, the inner race only revolves with the planet carrier while the outer race shares the same kinematic characteristics with the planet gear, which not only revolves around its own axis but also

revolves around the sun gear axis. A brief description of the planet bearing has been shown in Fig. 1(c). In the model, the planet bearing has five DOFs: inner race transverse motion in the x-axis and y-axis, outer race transverse motion in the x-axis and y-axis and one extra sprung DOF. This extra sprung DOF was used to tune the bearing parameters to represent the high frequency behavior of the bearings and adds a new equation for each bearing but does not alter the structure of the equations of the bearing motion. The transverse motions of the inner race, outer race and sprung system were all simulated with respect to a rotating frame of reference fixed to the carrier.

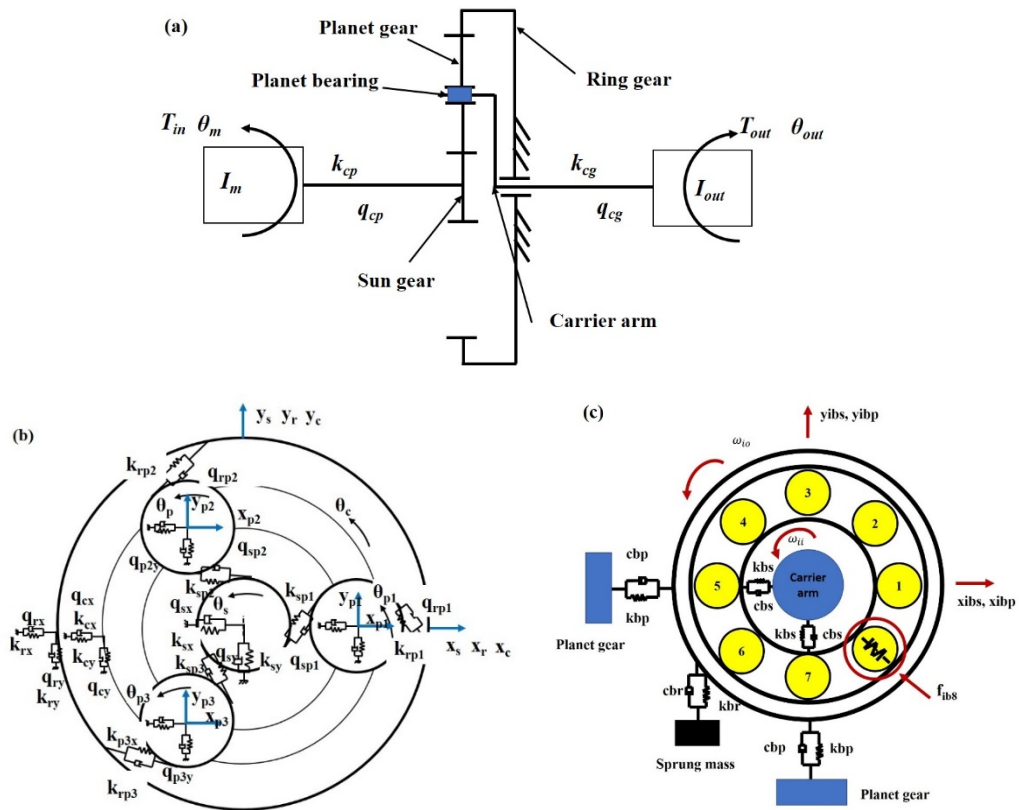


Figure 1 Planetary gear system, (a). the whole transmission train, (b). planetary gear dynamic model, (c). planet bearing dynamic model

2.2 The planetary gear system equations of motion

The resulting equations of motion for this planetary gear system are shown as follows.

For the rotary motion of the motor, the motion equation is,

$$I_m \ddot{\theta}_m = T_{in} - T_s,$$

$$T_s = k_{cp}(\theta_m - \theta_s) + q_{cp}(\dot{\theta}_m - \dot{\theta}_s). \quad (1)$$

For the motion of the sun gear, the differential equation is,

$$m_s \ddot{x}_s + k_{sx} x_s + q_{sx} \dot{x}_s + \sum_{i=1}^3 [F_{spi} \cos(\varphi_{spi})] = m_s x_s \Omega^2 + 2m_s \dot{y}_s \Omega + m_s y_s \dot{\Omega},$$

$$m_s \ddot{y}_s + k_{sy} y_s + q_{sy} \dot{y}_s + \sum_{i=1}^3 [F_{spi} \sin(\varphi_{spi})] = m_s y_s \Omega^2 - 2m_s \dot{x}_s \Omega - m_s x_s \dot{\Omega},$$

$$(I_s/r_s) \ddot{\theta}_s + \sum_{i=1}^3 F_{spi} = T_s/r_s, \quad (2)$$

where F_{spi} is the normal contact force between the sun gear and the i th planet,

$$F_{spi} = k_{spi} \delta_{spi} + q_{spi} \dot{\delta}_{spi},$$

$$\delta_{spi} = (x_s - x_{pi}) \cos \varphi_{spi} + (y_s - y_{pi}) \sin \varphi_{spi} + r_s \theta_s + r_p \theta_{pi} - r_c \theta_c \cos \alpha_{sp},$$

$$\varphi_{spi} = \frac{\pi}{2} - \alpha_{sp} + \varphi_i,$$

$$\varphi_i = \frac{2\pi(i-1)}{3}; i = 1, 2, 3. \quad (3)$$

For the motion of the planet gears, the x and y components become,

$$m_{pi} \ddot{x}_{pi} - k_{bp}(x_{bp} - x_p) - c_{bp}(\dot{x}_{bp} - \dot{x}_p) + F_{spi} \cos \varphi_{spi} + F_{rpi} \cos \varphi_{rpi} = m_{pi} x_{pi} \Omega^2 + 2m_{pi} \dot{y}_{pi} \Omega$$

$$+ m_{pi} y_{pi} \dot{\Omega} + m_{pi} r_c \Omega^2 \cos \varphi_i,$$

$$m_{pi} \ddot{y}_{pi} - k_{bp}(y_{bp} - y_p) - c_{bp}(\dot{y}_{bp} - \dot{y}_p) - F_{spi} \sin \varphi_{spi} - F_{rpi} \sin \varphi_{rpi}$$

$$= m_{pi} y_{pi} \Omega^2 - 2m_{pi} \dot{x}_{pi} \Omega - m_{pi} x_{pi} \dot{\Omega} + m_{pi} r_c \Omega^2 \sin \varphi_i,$$

$$(I_{pi}/r_p) \ddot{\theta}_{pi} + F_{spi} - F_{rpi} = 0, \quad (4)$$

where F_{rpi} is the normal contact force between the ring gear and the i th planet,

$$F_{rpi} = k_{rpi} \delta_{rpi} + q_{rpi} \dot{\delta}_{rpi},$$

$$\delta_{rpi} = (x_r - x_{pi})\cos \varphi_{rpi} + (y_r - y_{pi})\sin \varphi_{rpi} + r_r\theta_r - r_p\theta_{pi} - r_c\theta_c\cos \alpha_{rp},$$

$$\varphi_{rpi} = \frac{\pi}{2} + \alpha_{rp} + \varphi_i \quad (5)$$

For the motion of the i th planet gear bearing, the inner race can be modelled as a two degree of freedom system that has translational motion in x and y directions. The equation of motion of the inner race can be written as,

$$m_{ibs}\ddot{x}_{ibs} + F_{cpix} - f_{ibx} = m_{ibs}x_{ibs}\Omega^2 + 2m_{ibs}\dot{y}_{ibs}\Omega + m_{ibs}y_{ibs}\dot{\Omega} + m_{ibs}r_c\Omega^2\cos \varphi_i,$$

$$m_{ibs}\ddot{y}_{ibs} + F_{cpiy} - f_{iby} = m_{ibs}y_{ibs}\Omega^2 - 2m_{ibs}\dot{x}_{ibs}\Omega - m_{ibs}x_{ibs}\dot{\Omega} + m_{ibs}r_c\Omega^2\sin \varphi_i. \quad (6)$$

The outer race is modelled as a two degree of freedom system that has translational motion in the x and y directions. The equation of motion of the outer race can be written as,

$$m_{ibp}\ddot{x}_{ibp} + k_{bp}(x_{bp} - x_p) + c_{bp}(\dot{x}_{bp} - \dot{x}_p) + f_{ibx} = m_{ibp}x_{ibp}\Omega^2 + 2m_{ibp}\dot{y}_{ibp}\Omega + m_{ibp}y_{ibp}\dot{\Omega} + m_{ibp}r_c\Omega^2\cos \varphi_i,$$

$$m_{ibs}\ddot{y}_{ibp} + k_{bp}(y_{bp} - y_p) + c_{bp}(\dot{y}_{bp} - \dot{y}_p) + k_{br}(y_{bp} - y_{br}) + c_{br}(\dot{y}_{br} - \dot{y}_{br}) + f_{iby} = m_{ibp}y_{ibp}\Omega^2 - 2m_{ibp}\dot{x}_{ibp}\Omega - m_{ibp}x_{ibp}\dot{\Omega} + m_{ibp}r_c\Omega^2\sin \varphi_i \quad (7)$$

The sprung mass is attached in the y direction. Its equation of motion can be written as,

$$m_{ibr}\ddot{y}_{ibs} - k_{br}(y_{bp} - y_{br}) - c_{br}(\dot{y}_{br} - \dot{y}_{br}) = 0, \quad (8)$$

where F_{cpix} and F_{cpiy} describe the interaction force between the i th planet bearing and the carrier arm in the x and y directions,

$$F_{cpix} = k_{ibs}(x_{ibs} - x_c) + q_{ibs}(\dot{x}_{ibs} - \dot{x}_c),$$

$$F_{cpiy} = k_{ibs}(y_{ibs} - y_c) + q_{ibs}(\dot{y}_{ibs} - \dot{y}_c), \quad (9)$$

f_{ibx} , f_{iby} describe the summation of the contact forces in the x - and y - directions for a ball bearing with n_b balls and can be calculated as,

$$f_{ibx} = k_b \sum_{j=1}^{n_b} \gamma_{ibj} \delta_{ibj}^{1.5} \cos \Phi_{ibj},$$

$$\begin{aligned}
f_{iby} &= k_b \sum_{j=1}^{n_b} \gamma_{ibj} \delta_{ibj}^{1.5} \sin \Phi_{ibj}
\end{aligned} \tag{10}$$

where k_b is the combined contact stiffness between the rolling element and the races and γ_{ibj} is the contact state. δ_{ibj} and Φ_{ibj} are the overall contact deformation as well as the angular position for the j th rolling element on the i th planet gear bearing respectively and they can be described as,

$$\delta_{ibj} = (x_{ibs} - x_{ibp}) \cos \Phi_{ibj} + (y_{ibs} - y_{ibp}) \sin \Phi_{ibj} - c, \quad j = 1, 2, 3, \dots, n_b, \tag{11}$$

$$\begin{aligned}
\Phi_{ibj} &= \frac{2\pi(j-1)}{n_b} + \omega_{ibc}^c dt + \Phi_{ibo}
\end{aligned} \tag{12}$$

Φ_{ibo} is the initial cage position and ω_{ibc}^c is the cage speed measured in the rotating frame, which can be calculated as,

$$\omega_{ibc}^c = \omega_{ibc} - \omega_c = \frac{\omega_{ii}}{2} \left(1 - \frac{D_b}{D_p} \cos \alpha \right) + \frac{\omega_{io}}{2} \left(1 + \frac{D_b}{D_p} \cos \alpha \right) - \omega_c. \tag{13}$$

For the motion of the carrier arm, the motion equation is,

$$\begin{aligned}
m_c \ddot{x}_c + k_{cx} x_c + q_{cx} \dot{x}_c - \sum_{i=1}^3 F_{cpix} &= m_c x_c \Omega^2 + 2m_c \dot{y}_c \Omega + m_c y_c \dot{\Omega}, \\
m_c \ddot{y}_c + k_{cy} y_c + q_{cy} \dot{y}_c - \sum_{i=1}^3 F_{cpiy} &= m_c y_c \Omega^2 - 2m_c \dot{x}_c \Omega - m_c x_c \dot{\Omega}, \\
(I_c/r_c) \ddot{\theta}_c + \sum_{i=1}^3 F_{cpix} \sin \varphi_i - \sum_{i=1}^3 F_{cpiy} \cos \varphi_i &= T_c/r_c, \\
T_c &= k_{cg} (\theta_c - \theta_{out}) + q_{cg} \cdot (\dot{\theta}_c - \dot{\theta}_{out}).
\end{aligned} \tag{14}$$

For the motion of the ring gear, the motion equation is,

$$\begin{aligned}
m_r \ddot{x}_r + k_{rx} x_r + q_{rx} \dot{x}_r + \sum_{i=1}^3 F_{rpi} \cos \varphi_{rpi} &= m_r x_r \Omega^2 + 2m_r \dot{y}_r \Omega + m_r y_r \dot{\Omega}, \\
m_r \ddot{y}_r + k_{ry} y_r + q_{ry} \dot{y}_r + \sum_{i=1}^3 F_{rpi} \sin \varphi_{rpi} &= m_r y_r \Omega^2 - 2m_r \dot{x}_r \Omega - m_r x_r \dot{\Omega},
\end{aligned}$$

$$(I_r/r_r)\ddot{\theta}_r + (q_{rt}/r_r)\dot{\theta}_r + (k_{rt}/r_r)\theta_r + \sum_{i=1}^3 F_{rpi} = 0. \quad (15)$$

For the rotary motion of the load, the motion equation is,

$$I_{out}\ddot{\theta}_{out} = -T_{out} + T_c. \quad (16)$$

2.3 Gear and bearing stiffness evaluation

The main internal vibration source in the planetary gear system comes from the variations of the gear stiffness as well as the bearing stiffness. In this paper, the variation of the gear stiffness included the sun-planet mesh stiffness and the ring-planet mesh stiffness. In addition, the variation of the planet bearing stiffness were considered. The tooth mesh stiffness variation of a sun-planet gear pair and the ring-planet gear pair were modelled from the isolated sun-planet FE model and the isolated sun-planet FE model respectively.

The detailed steps on how to model the gear finite element model and how to calculate the mesh stiffness using the established models can be found in the previous works [32]. The value of the bearing stiffness can be estimated based on the Hertzian contact theory [33].

In the ball bearing, the rolling element contacts with the inner and outer raceway respectively.

The surface of a rolling element is convex whereas the surface of the outer raceway is concave. The total contact stiffness for the inner-ball-outer contact can be written as,

$$k_b = \frac{1}{\left[\left(\frac{1}{K_{bi}^C} \right)^{\frac{3}{2}} + \left(\frac{1}{K_{bo}^C} \right)^{\frac{3}{2}} \right]^{\frac{2}{3}}}. \quad (17)$$

K_{bi}^C , K_{bo}^C are the contact stiffness for the ball-inner-race contact and the ball-outer-race contact, respectively. They can be written as,

$$K_{bi}^C = \pi k_{bi}^e E_{bi}' \sqrt{\frac{R_{bi} \xi_{bi}}{4.5 \zeta_{bi}^3}}, \quad (18)$$

$$K_{bo}^C = \pi k_{bo}^e E_{bo}' \sqrt{\frac{R_{bo} \xi_{bo}}{4.5 \zeta_{bo}^3}}, \quad (19)$$

where E_{bi}' is the effective modulus of elasticity, k_{bi}^e , ξ_{bi} and ζ_{bi} are the ellipticity parameters for the ball-inner-race contact. E_{bo}' is the effective modulus of elasticity and k_{bo}^e , ξ_{bo} and ζ_{bo} are the ellipticity parameter for the ball-outer-race contact.

3. Modelling of the planet bearing faults

3.1 Planet bearing distributed fault modelling

As indicated in [34], wear, corrosion and the plastic deformation failure modes can be treated as distributed faults, which would cause a continuous geometrical imperfection along the contact surfaces. For the continuous distributed faults, they can be modeled as a series of sinusoidal functions to express the waviness of periodic lobes. In this research, the effect of distributed faults on the inner raceway, outer raceway and the rolling ball have been investigated, and Fig. 2 shows the planet bearing with distributed faults on the inner raceway, outer raceway and the rolling ball respectively.

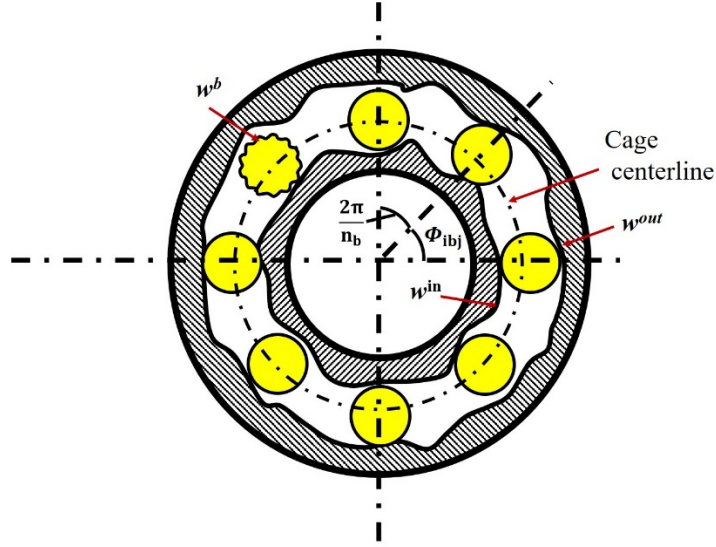


Figure 2 Planet bearing with distributed faults (exaggerated)

The waviness of the bearing inner raceway surface, w^{in} , at any position on the surface of outer raceway can be described using,

$$w^{in} = \Pi_0 + \sum_{l=1}^N \Pi_p \sin (lN_i \Phi_{ibj} + \alpha_{li}) \quad (20)$$

where Π_0 is the initial amplitude of the waviness, Π_p is the maximum amplitude of the waviness, N_i is the total waviness number on the inner race, α_{li} is the initial phase angle of the l th order waviness.

The waviness of the bearing outer raceway surface, w^{out} , at any position on the surface of outer raceway can be described using,

$$w^{out} = \Pi_0 + \sum_{l=1}^N \Pi_p \sin [lN_o(\Phi_{ibj} - \theta_{io}) + \alpha_{lo}] \quad (21)$$

where N_o is the total waviness number on the outer raceway, θ_{io} is the position angle of the outer raceway and α_{lo} is the initial phase angle of the l th order waviness.

Because the ball is in contact with both the inner and outer raceway respectively, the phase angle difference of these two contact points is 180 deg. Therefore, the rolling ball waviness

in contact with the inner and outer race can be given by

$$w^{ball} = \Pi_o + \sum_{l=1}^N \Pi_p \{ \sin (lN_b \omega_b t + \alpha_{lb}) + \sin [lN_b(\omega_b t + \pi) + \alpha_{lb}] \} \quad (22)$$

where N_b is the total waviness number on the rolling ball and α_{lb} is the initial phase angle of the l th order waviness.

3.2 Planet bearing localized fault modelling

The spalling and crack faults in the bearing system can be treated as localized faults, which would only cause a localized discontinuous geometrical imperfection and a resulting transient vibration would normally occur. For example, when a loaded ball encounters a spall, there will be a sudden relaxation of the loading and the transient forces which are produced would result in rapid accelerations of the bearing components and complex motions can occur including oscillatory contact and impacts between different components. There are three different types of localized planet bearing faults considered in this research and they are line spall on the inner race, line spall on the outer race and a fault on the rolling element, as shown in Fig. 3.

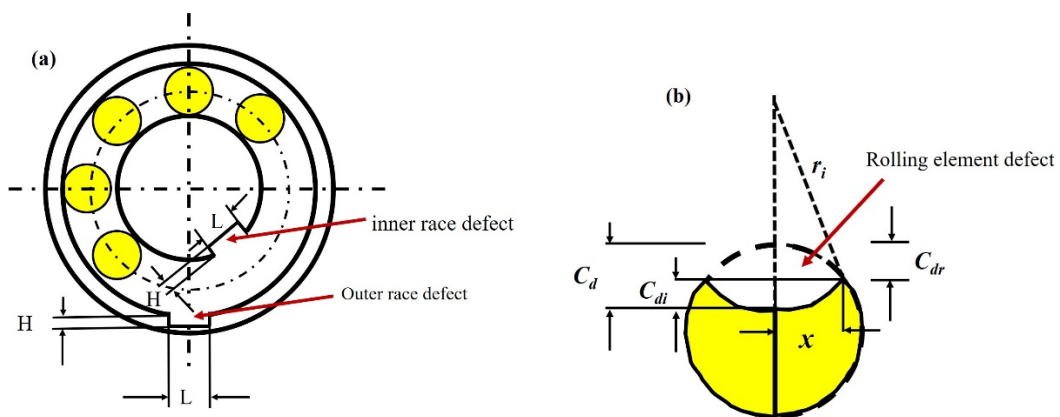


Figure 3 Planet bearing fault models, (a) localized inner and outer race faults, (b) rolling

element fault

If localized faults on the inner/outer raceway or rolling elements are introduced in the bearing model, the total elastic deformation can be rewritten as,

$$\delta_{ibj} = (x_{ib} - x_{pi})\cos \Phi_{ibj} + (y_{ib} - y_{pi})\sin \Phi_{ibj} - c - \beta_j^k C_d^k \quad (23)$$

where C_d^k ($k = \text{in, out, ball}$) is the spall depth. For the inner raceway and outer raceway defect, it was modelled as a Tukey window as proposed by Liu to avoid the sudden loss of the ball raceway contact [35]. β_j^k ($k = \text{in, out, ball}$) is the fault switch to simulate the contact loss at the predefined angular position and it determines whether the j th ball/roller is inside the defected zone or not, which is defined as follows,

$$\beta_j = \begin{cases} 1 & \text{if } \Phi_d^k < \Phi_{ibj} < \Phi_d^k + \Delta\Phi_d^k \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

where Φ_d^k ($k=\text{in, out}$) is the initial starting location of the defected zone for the inner raceway and outer raceway while $\Delta\Phi_d^k$ ($k=\text{in, out}$) is the angular distance of the defected zone for the inner raceway and outer raceway. As the model was defined in the carrier arm rotating reference frame, the inner race defect is fixed in location between Φ_d^{in} and $\Phi_d^{\text{in}} + \Delta\Phi_d^{\text{in}}$. The rotating speed of the outer race defect is the same as the planet gear rotating speed relative to the carrier arm, i.e. $\Phi_d^{\text{out}} = \omega_p^c t + \Phi_{d0}^{\text{out}}$. It should be noted that the planetary gear bearing motion in this model is opposite to that of the fixed axis gear bearing motion, where the outer race is fixed in location and the inner race rotates with the rotor.

3.3 Planet bearing multi-fault modelling

Numerous industry experience showed that there were usually more than one type of fault existing in a mechanical unit and these faults can be in different positions, with different

fault modes and fault severities. The coupling effect and the mutual interaction among these faults bring a great challenge to the comprehensive understanding of the equipment fault diagnosis result. In this research, three different planet bearing fault coupling scenarios have been investigated,

(1) The fault coupling between the planet bearing distributed faults

In section 3.1, the effect of distributed faults on the inner raceway, outer raceway and the rolling ball have been modelled mathematically and therefore, there are four different potential distributed faults combinations considered here. They are the inner-outer raceway distributed fault combination (case 1), inner raceway-rolling ball distributed fault combination (case 2), outer raceway-rolling ball distributed fault combination (case 3), and inner-outer-rolling ball distributed fault combination (case 4). Once the distributed fault modes and severities have been determined in each component, the resultant bearing overall contact deformation δ_{ibj} in equation (11) can be rewritten as,

$$\delta_{ibj} = \begin{cases} (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - w^{in} - w^{out} & \text{(case 1)} \\ (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - w^{in} - w^{ball} & \text{(case 2)} \\ (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - w^{out} - w^{ball} & \text{(case 3)} \\ (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - w^{in} - w^{out} - w^{ball} & \text{(case 4)} \end{cases} \quad (25)$$

(2) The fault coupling between the planet bearing localized faults

In section 3.2, the mathematical model of localized faults on the inner raceway, outer raceway and the rolling ball have also been established and similarly, there are four different potential distributed fault combinations considered in this research. They are the inner-outer raceway localized fault combination (case 1), inner raceway-rolling ball localized fault combination (case 2), outer raceway-rolling ball localized fault combination (case 3), and

inner-outer-rolling ball localized fault combination (case 4). Once the localized fault modes and severities have been determined in each component, the resultant bearing overall contact deformation δ_{ibj} in equation (11) can be rewritten as,

$$\delta_{ibj} = \begin{cases} (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - \beta_j^{in}C_d^{in} - \beta_j^{out}C_d^{out} & \text{(case 1)} \\ (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - \beta_j^{in}C_d^{in} - \beta_j^{ball}C_d^{ball} & \text{(case 2)} \\ (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - \beta_j^{out}C_d^{out} - \beta_j^{ball}C_d^{ball} & \text{(case 3)} \\ (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - \beta_j^{in}C_d^{in} - \beta_j^{out}C_d^{out} - \beta_j^{ball}C_d^{ball} & \text{(case 4)} \end{cases} \quad (26)$$

4. Numerical simulation and result analysis

4.1 Planet bearing characteristic frequency

For a planet bearing in a planetary gearbox, its inner race is usually connected to the planet carrier, and its outer race is connected to the planet gear. Therefore, the planet bearing motion is a complex superposition of the revolution around the sun and ring gear common axis and its spinning around the planet axis. This section develops the planet bearing fault characteristic frequencies. If f_c is the carrier arm absolute rotating frequency, the planet gear absolute rotating frequency is,

$$f_p = 1 + \frac{Z_r}{Z_p}f_c. \quad (27)$$

As discussed previous, the inner race of the planet bearing is connected to the carrier arm and the outer race of the planet bearing is connected to the planet gear, therefore, the inner race absolute rotating frequency and the outer race absolute rotating frequency can be calculated as,

$$f_i = f_c,$$

$$f_o = f_p \quad (28)$$

Given the diameter of a rolling element D_b , the bearing pitch diameter D_p , the contact angle α , the absolute rotating frequency of the rolling element cage as well as the absolute ball spin frequency can be calculated as by the equation given by Howard [36]. However, if all the frequencies were measured in the carrier arm rotating frame, the inner race becomes stationary and the outer race relative rotating frequency becomes,

$$f_o^c = f_p - f_c \quad (29)$$

The rolling element cage rotating frequency becomes,

$$f_{cg}^c = \frac{1}{2} \left(1 + \frac{D_b}{D_p} \cos \alpha \right) f_o^c \quad (30)$$

The rolling element rotating frequency becomes,

$$f_b^c = \frac{D_p}{2D_b} \left[1 - \left(\frac{D_b}{D_p} \cos \alpha \right)^2 \right] f_o^c \quad (31)$$

If the localized fault exists on the outer race of a planet bearing, the characteristic frequency of the outer race fault measured in the rotating frame can be calculated as,

$$f_{bpf_o}^c = \frac{n_b}{2} \left(1 - \frac{D_b}{D_p} \cos \alpha \right) f_o^c \quad (32)$$

If the localized fault exists on the inner race of a planet bearing, the characteristic frequency of the inner race fault measured in the rotating frame can be calculated as,

$$f_{bpf_i}^c = \frac{n_b}{2} \left(1 + \frac{D_b}{D_p} \cos \alpha \right) f_o^c \quad (33)$$

If the localized fault is on one of the rolling elements, in one spinning cycle of the rolling element, it touches both the inner and outer races, and generates impacts twice. The

characteristic frequency of the rolling element fault measured in the rotating frame can be calculated as,

$$f_{bf}^c = 2f_b^c = \frac{D_p}{D_b} \left[1 - \left(\frac{D_b}{D_p} \cos \alpha \right)^2 \right] f_o^c. \quad (34)$$

4.2 Signal processing approach

The instantaneous angular speed (IAS) of the carrier arm has been selected to analyse the coupling effect of the planet bearing faults. However, if the raw signal (carrier arm angular speed) was analyzed directly, little useful information can be identified. In order to enhance the planet bearing signal, the “benchmark” method developed by R.B. Randall has been adopted and further developed here to analysis the planet bearing vibration [37]. There are mainly six steps in this approach and the details of the step description can be found in Appendix A. The first step was to remove deterministic components contributed by the gears, and this was achieved by using the linear prediction method. After removal of the deterministic components, the time signal was further enhanced by the minimum entropy deconvolution (MED) method, which can give rise to the impulses hidden in the signal. Then, the filtered signal was decomposed using adaptive mode decomposition. In order to determine which component can be used for further analysis, the kurtosis value and their correlation coefficient were calculated. Next, amplitude demodulation using Hilbert-Huang Transform (HHT) techniques were used to form the envelope signal, which was transformed into the frequency domain to analyze the desired planet bearing distributed fault diagnostic information in the end.

Fig.4 takes the inner race distributed faults as one example and shows the results in each

step using this approach. In the example, the inner race distributed fault was modelled as a sinusoidal function, whose Π_0 was $1\mu\text{m}$, Π_p was $8\mu\text{m}$, l was set to be 1, the total waviness number on the inner race was 14 and the initial phase angle was set to be 0 in this case. As shown in the figure, the final signal envelope spectrum successfully eliminated the affect from the gear vibration and an abundant of peaks and sidebands can be identified visually. In the following sections, only the final spectrum results will be presented instead of all the signal pre-processing results.

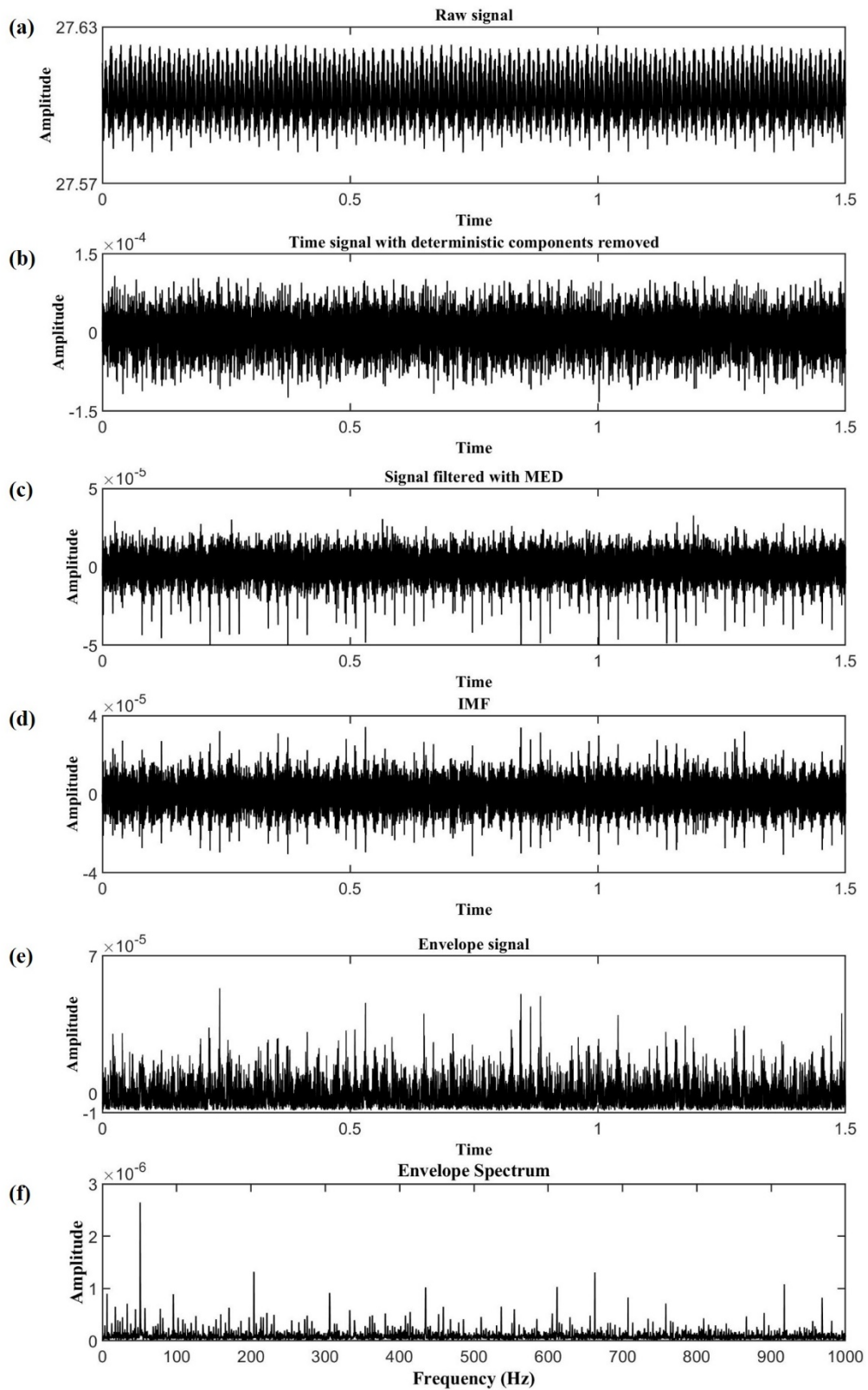
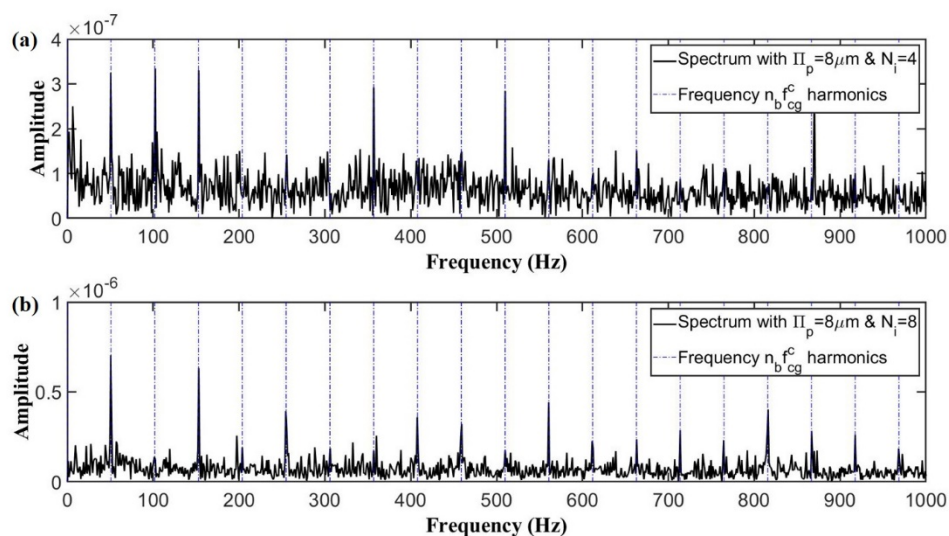


Figure 4 Results of planet bearing inner race distributed fault signal processing

4.3 Simulation results

4.3.1 Effect of planet bearing distributed faults

In this section, the effect of a single planet bearing distributed fault was investigated first, which included an inner raceway distributed fault, outer raceway distributed fault and rolling ball distributed fault scenarios. Fig. 5 compares the vibration spectrum of the planet bearing with different inner raceway fault orders, which are 4, 8, 14, and 20 respectively. Here, order of 8 was set to be equal to the roller number in this research. It should be noted that when the inner raceway distributed fault was investigated, the outer raceway and the rollers were assumed to be ideal without any faults. It was interesting to be observed that the resultant vibration peaks were found in the position of $kn_b f_{cg}^c$ ($k = 1, 2, 3, \dots$) regardless of the inner raceway fault orders. For example, when the fault order is 14, the peaks would appear in the positions of $n_b f_{cg}^c$, $4n_b f_{cg}^c$, $6n_b f_{cg}^c$ and so on. The results indicated that the planet bearing inner race characteristic fault frequency was only related to the roller number as well as the cage frequency.



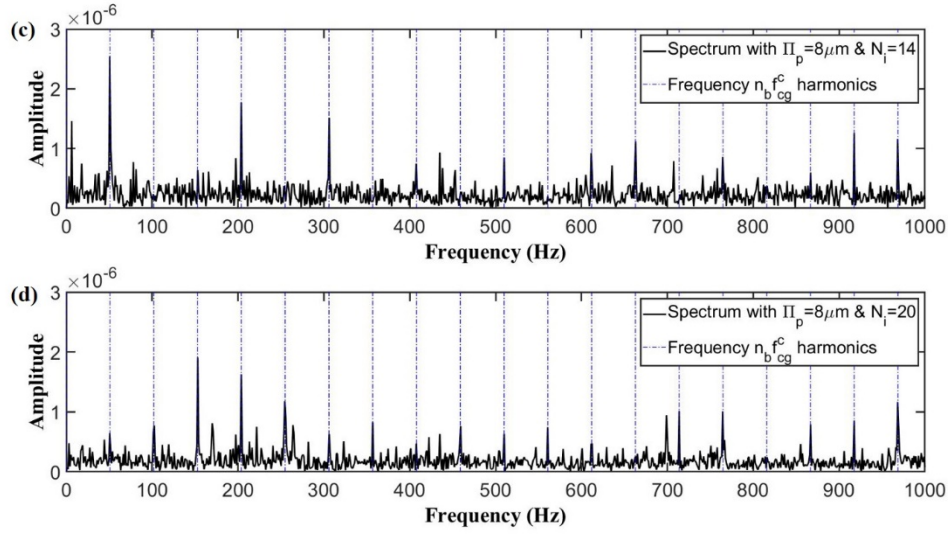


Figure 5 Vibration spectrum due to the inner race distributed fault, (a) inner race fault with $N_i = 4$, (b) inner race fault with $N_i = 8$, (c) inner race fault with $N_i = 14$, (d) inner race fault with $N_i = 20$

Fig. 6 compares the vibration spectrum of the planet bearing with different outer raceway fault orders, which are 4, 8, 14, and 20 respectively. Similarly, when the outer raceway distributed fault was investigated, the inner raceway and the rollers were assumed to be ideal without any faults. Unlike the inner raceway distributed fault case, it was hard to find a frequency and its harmonics that can summarize the position of the resultant vibration peaks. However, the resultant vibration peaks can be classified into two categories, which are $kn_b(f_{oc} - f_{cg}^c) \pm lf_o$ ($k = 1, 2, 3, \dots; l = 0, 1, 2, 3, \dots$) and $kn_b(f_{oc} - f_{cg}^c) \pm lf_{cg}^c$ ($k = 1, 2, 3, \dots; l = 0, 1, 2, 3, \dots$). For example, when the fault order is 4, both categories can be identified in the spectrum, such as $n_b(f_{oc} - f_{cg}^c) + f_{cg}^c$ and $11n_b(f_{oc} - f_{cg}^c) + f_o$. With the fault order number increased, the distribution of these vibration peaks became monotonic with only $kn_b(f_{oc} - f_{cg}^c) \pm lf_o$ found in the case of fault order $N_o = 8$ and only $kn_b(f_{oc} - f_{cg}^c) \pm lf_{cg}^c$ found in the case of fault order $N_o = 14$ & 20. The results indicated that the planet bearing outer race characteristic fault frequency was centered on the frequency of $kn_b(f_{oc} - f_{cg}^c)$,

which was further modulated by the frequency of f_{cg}^c or f_o .

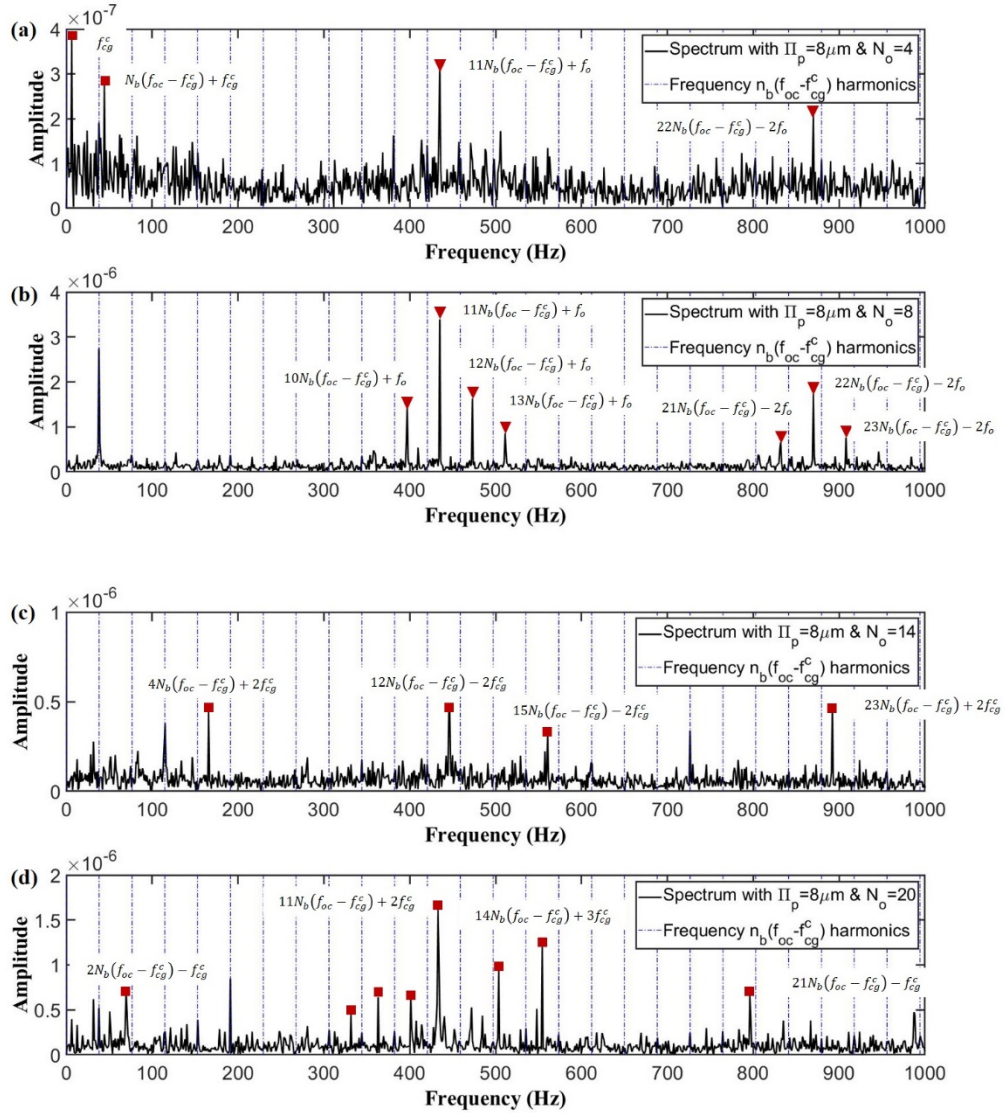


Figure 6 Vibration spectrum due to the outer race distributed fault, (a) outer race fault with $N_o = 4$, (b) outer race fault with $N_o = 8$, (c) outer race fault with $N_o = 14$, (d) outer race fault with $N_o = 20$,

▼ stands for $k n_b(f_{oc}^c - f_{cg}^c) \pm l f_{cg}^c$, ■ stands for $k n_b(f_{oc}^c - f_{cg}^c) \pm l f_o$

Fig. 7 compares the vibration spectrum of the planet bearing with different rolling ball fault orders, which are 4, 8, 14, and 20 respectively. Similarly, when the rolling ball distributed fault was investigated, the inner raceway and the outer raceway were assumed to be ideal

without any faults. From the figures, it can be found that the resultant vibration peaks can be identified on the positions of $2kf_{bc}^c \pm lf_{cg}^c$ ($k = 1, 2, 3, \dots; l = 0, 1, 2, 3, \dots$). For example, when the fault order is 20, the resultant vibration peaks appeared on the positions of $2f_{bc}^c - 4f_{cg}^c$, $2f_{bc}^c + 4f_{cg}^c$ and $4f_{bc}^c$. The results indicated that the planet bearing rolling ball characteristic fault frequency was centered on the frequency of $2kf_{bc}^c$, which was modulated by the frequency of f_{cg}^c .

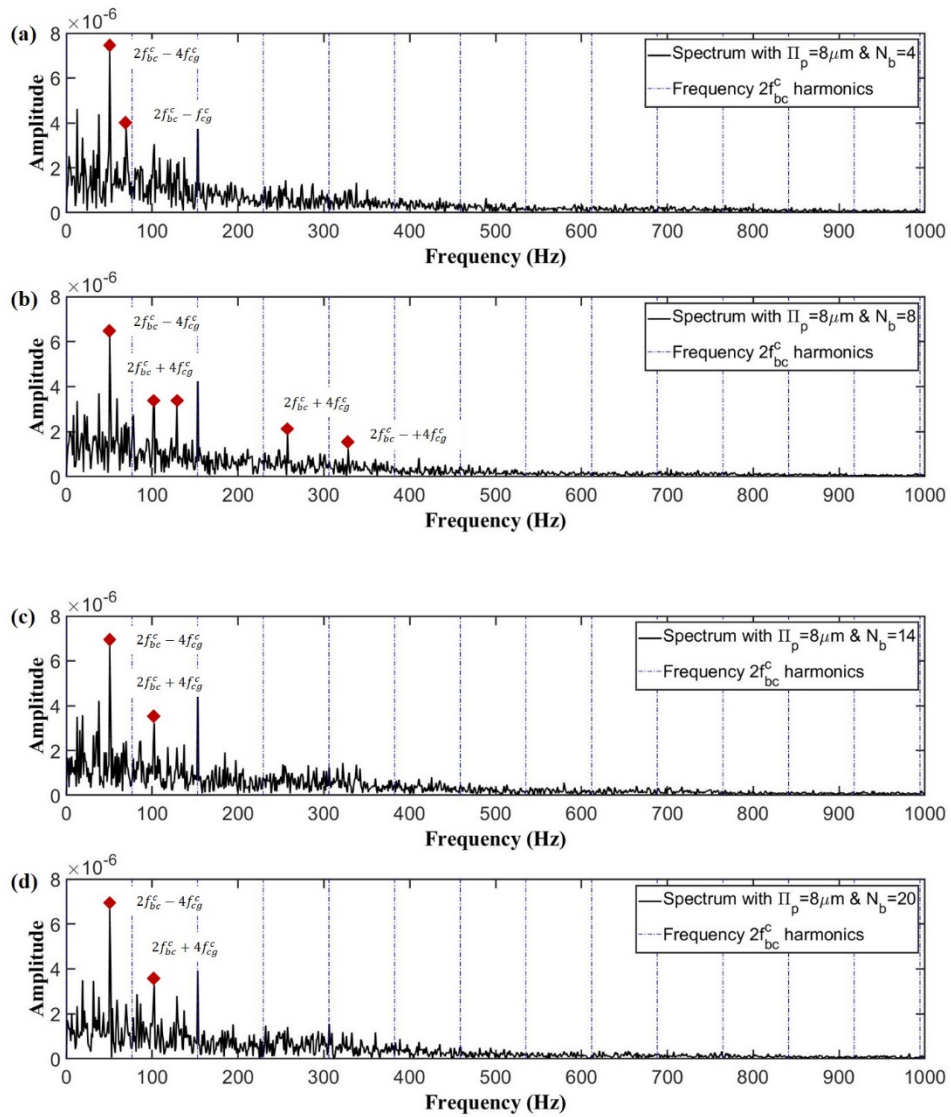


Figure 7 Vibration spectrum due to the rolling ball distributed fault, (a) rolling ball fault with N_b

=4, (b) rolling ball fault with $N_b=8$, (c) rolling ball fault with $N_b=14$, (d) rolling ball fault with

$N_b=20$,

◆ stands for $kf_{bc}^c \pm lf_{cg}^c$

4.3.2 Effect of planet bearing localized faults

In this section, the effect of planet bearing localized faults without any coupling effect were investigated, which included inner raceway localized fault, outer raceway localized fault and rolling ball localized fault scenarios. Fig. 8(a) shows the amplitude demodulation spectrum of the planet bearing with inner race localized fault and from the figure, the inner raceway characteristic frequency $f_{bpf_i}^c$ and its harmonics can be easily identified with little interference from other components. Fig. 8(b) shows the amplitude demodulation spectrum of the planet bearing with outer race localized fault and from the figure, even though the outer raceway characteristic frequency $f_{bpf_o}^c$ and its harmonics can still be identified, a large number of sidebands can also be observed. These sidebands were caused by the modulation effect from the rotation of the outer race as well as the rotation of the cage. Fig. 8(c) shows the amplitude demodulation spectrum of the planet bearing with a rolling ball localized fault and from the figure, the rolling ball characteristic frequency f_{bf}^c and its harmonics can be barely identified with a large number of strong sidebands around.

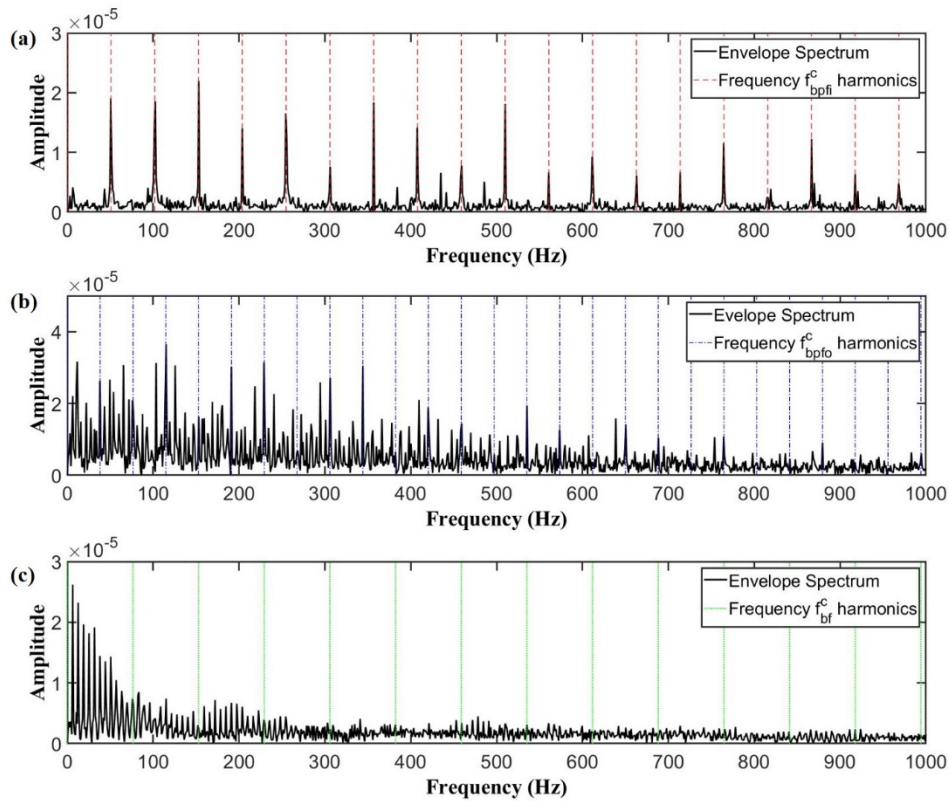


Figure 8 Vibration spectrum due to the planet bearing localized faults, (a) inner raceway localized fault, (b) outer raceway localized fault, (c) rolling ball localized fault

4.3.3 Effect of planet bearing combined distributed faults

In this section, the effect of planet bearing distributed faults with the coupling effect between the distributed faults were investigated, which included inner/outer raceway distributed fault, inner raceway/rolling ball distributed fault, outer raceway/rolling ball distributed fault, inner/outer/rolling ball raceway distributed fault. The method of how to model the combined distributed faults has been introduced in section 3.3. In the following section, the resultant vibration spectrum of the combined distributed faults was first presented and then the magnitude-squared coherence between the combined faults and the single distributed fault has been calculated correspondingly. Based on the analysis in

section 4.3.1, the results indicated that there was no direct relationship between the resultant characteristic frequencies and the distributed fault orders. Therefore, the fault order of 20 has been taken as an example in this section to study the coupling effect between the distributed faults. Here, two thresholds of the coherence were used to quantify the similarity between these signals.

Fig. 9(a) shows the vibration spectrum due to the inner/outer combined distributed faults. As shown in the figure, a range of vibration peaks can be identified and however, the resultant spectrum was found not be a simple summation of the inner raceway distributed fault spectrum and the outer raceway distributed fault spectrum. Significant amount of vibration peaks other than the inner distributed fault characteristic frequencies and the outer distributed fault characteristic frequencies can also be observed in the figure. This was to be expected as with the coupling effect, some other frequency other than the characteristic frequency can also be excited during this coupling effect. Fig. 9(b) and Fig. 9(c) present the coherence estimation results between the combined fault and the inner raceway fault as well as the combined fault and the outer raceway fault. Both the coherence on the inner race distributed characteristic frequencies positions and outer race distributed characteristic frequencies positions were examined while the coherence on the other frequency components would be neglected here even though the coherence on some of these components appeared to be larger than the two threshold values. The most striking result to emerge from the figure was that there were more characteristic frequency components from the outer raceway shown to be coherent with the resultant inner/outer combined fault spectrum. It indicates that the outer raceway fault has a stronger impact on

the combined vibration spectrum.

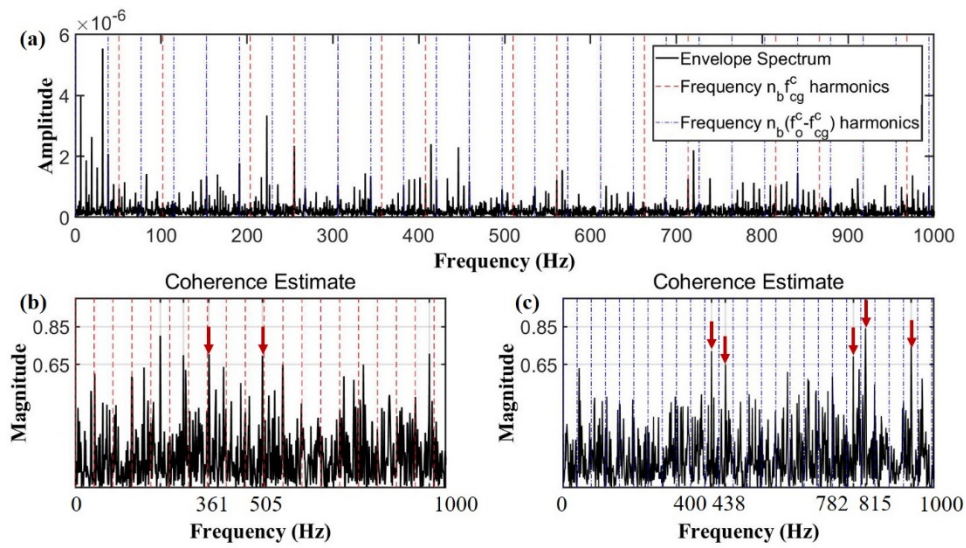


Figure 9 Effect of the inner/outer combined distributed faults, (a) vibration spectrum due to the combined distributed fault, (b) coherence estimation between the combined fault and the inner raceway fault, (c) coherence estimation between the combined fault and the outer raceway fault

Fig. 10(a) shows the vibration spectrum due to the inner raceway/rolling ball combined distributed faults. As shown in the figure, a range of vibration peaks can be identified, which was also found not to be a simple summation of the inner raceway distributed fault spectrum and the rolling ball fault spectrum. Fig. 10(b) and Fig. 10(c) present the coherence estimation results between the combined fault and the inner raceway fault as well as the combined fault and the rolling ball fault. Still, both the coherence on the inner race characteristic frequencies positions and rolling ball characteristic frequencies positions were examined while the coherence on the other frequency components would be neglected. It was apparent from the coherence figures that very few characteristic frequency component values were detected above the thresholds. However, one can still identify three

significant vibration peaks from the inner raceway coherence figure as shown in Fig. 10(b). Despite this, the results indicate that the inner raceway fault has a stronger impact on the combined vibration spectrum.

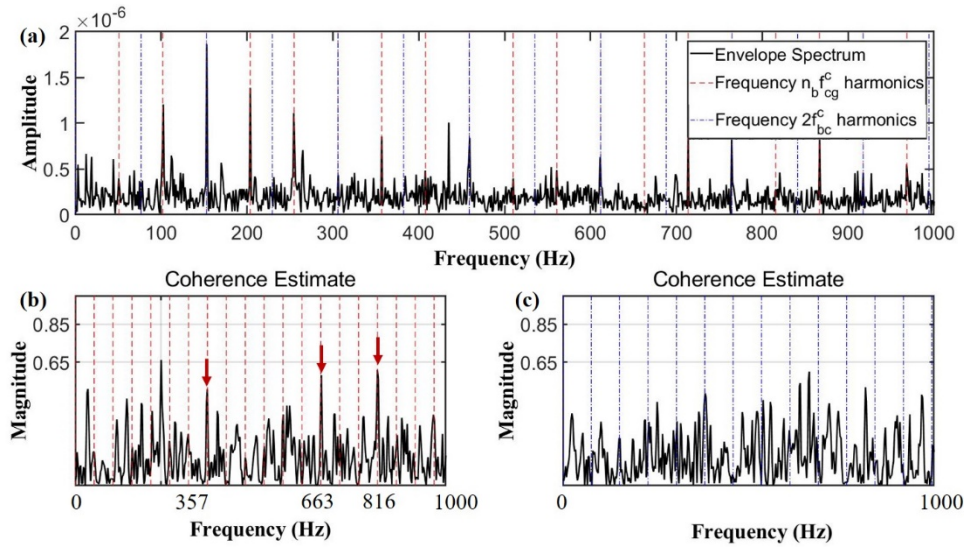


Figure 10 Effect of the inner raceway/rolling ball combined distributed faults, (a) vibration spectrum due to the combined fault, (b) coherence estimation between the combined fault and the inner raceway fault, (c) coherence estimation between the combined fault and the rolling ball fault

Fig. 11(a) shows the vibration spectrum due to the outer raceway/rolling ball combined distributed faults. As shown in the figure, a range of vibration peaks can be identified. As discussed in section 4.3.1, the outer raceway characteristic frequency was centered around $kn_b(f_{oc} - f_{cg}^c)$ which was further modulated by f_{cg}^c or f_o while the rolling ball characteristic frequency was centered around $2kf_{bc}^c$, which was modulated by f_{cg}^c . Therefore, it was hard to distinguish the vibration peak positions in Fig. 11(a). Fig. 11(b) and Fig. 11(c) present the coherence estimation results between the combined fault and the outer raceway fault as well as the combined fault and the rolling ball fault. Clearly, the

outer raceway had a higher coherence estimation than that of the rolling ball with several characteristic frequency peaks being identified visually. It suggests that the outer raceway had a larger contribution to the combined distributed faults resultant vibration spectrum.

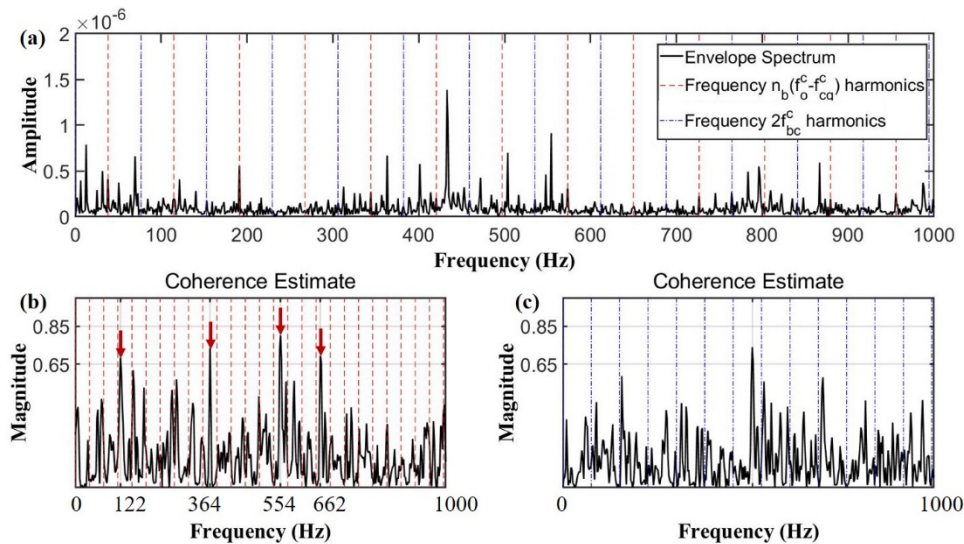


Figure 11 Effect of the outer raceway/rolling ball combined distributed faults, (a) vibration spectrum due to the combined fault, (b) coherence estimation between the combined fault and the outer raceway fault, (c) coherence estimation between the combined fault and the rolling ball fault

Fig. 12(a) shows the vibration spectrum due to the inner/outer/rolling ball combined distributed faults. With the fault contribution from these three bearing components, no significant correlation can be found between the resultant vibration spectrum and the other three independent vibration spectrum caused by the single distributed faults. From Fig. 12(b) to Fig. 12(d), they show the coherence estimation results between the combined fault and the inner raceway fault, the combined fault and the outer raceway fault as well as the combined fault and the rolling ball fault. Surprisingly, the coherence estimation in inner raceway was observed to have a higher value than the other two cases, with two inner race

characteristic frequency components above the threshold while no frequency components were observed to be over the threshold value. It suggested that the inner raceway had a larger contribution to the combined distributed faults resultant vibration spectrum.

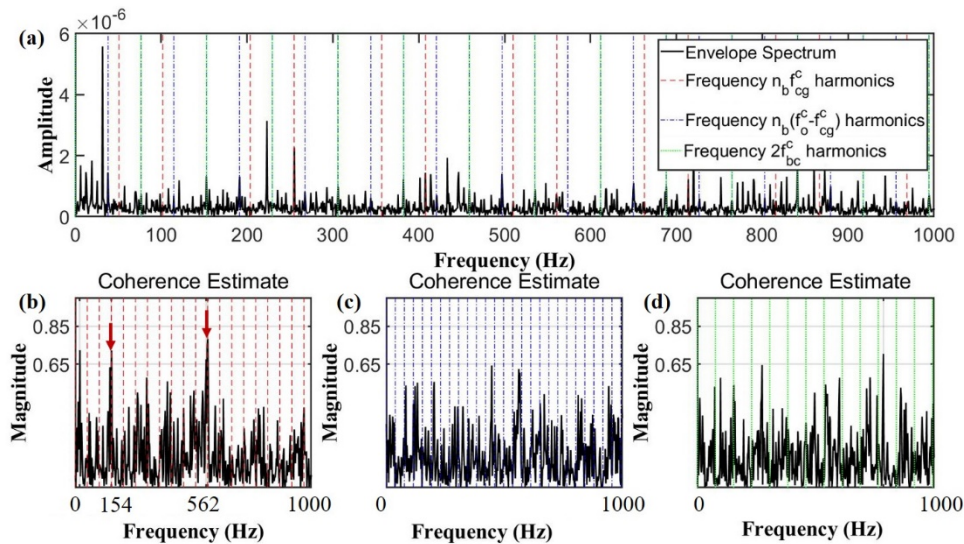


Figure 12 Effect of the inner/outer/rolling ball combined distributed faults, (a) vibration spectrum due to the combined fault, (b) coherence estimation between the combined fault and the inner raceway fault, (c) coherence estimation between the combined fault and the outer raceway fault, (d) coherence estimation between the combined fault and the rolling ball fault

4.3.4 Effect of planet bearing combined localized faults

In this section, the effect of planet bearing localized fault with the coupling effect between the localized faults were investigated, which included inner/outer raceway localized fault, inner raceway/rolling ball localized fault, outer raceway/rolling ball localized fault, inner/outer/rolling ball raceway localized fault. The method of how to model the combined distributed faults has been introduced in section 3.3. Similarly, the resultant vibration spectrum of the combined localized faults was first presented and then the magnitude-

squared coherence between the combined localized faults and the single localized fault has been calculated correspondingly. The vibration spectrum of the single localized fault has been discussed in section 4.3.2 already and the characteristic frequencies for each single localized fault case can be found. Two thresholds of the coherence were used to quantify the similarity between these signals.

Fig. 13(a) shows the vibration spectrum due to the inner/outer combined localized faults. As shown in the figure, most of the frequency components with large amplitude are distributed below 200Hz with a large number of sidebands. Compared with the results in Fig. 8, it is apparent that there is no linear association between the combined result and the single localized fault result, which indicates a strong nonlinear coupling effect between the inner raceway localized fault and the outer raceway localized fault. Fig. 13(b) and Fig. 13(c) present the coherence estimation results between the combined fault and the outer raceway fault as well as the combined fault and the inner raceway fault. The characteristic frequency components in the coherence figures have been examined and it is interesting to find that the inner raceway localized fault coherence figure was shown to have more characteristic components over the threshold, especially in the high frequency band. On the other hand, the outer raceway also has frequency components above the threshold, but with uniform distribution over the frequency range.

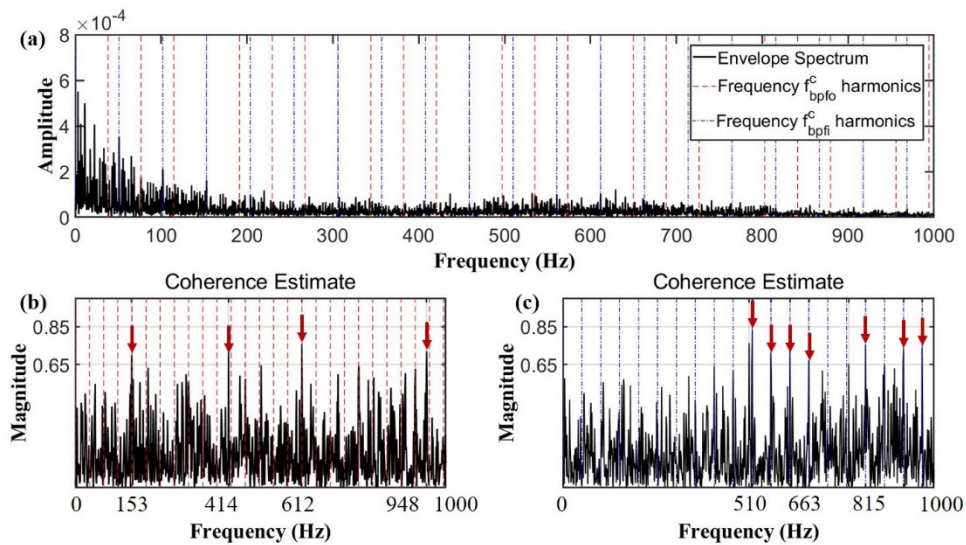


Figure 13 Effect of the inner/outer combined localized faults, (a) vibration spectrum due to the combined fault, (b) coherence estimation between the combined fault and the outer raceway fault, (c) coherence estimation between the combined fault and the inner raceway fault

Fig. 14(a) shows the vibration spectrum due to the inner raceway /rolling ball combined localized faults. As shown in the figure, the resultant spectrum appeared to be quite clear in this case with the inner raceway characteristic frequency being easily identified. This phenomenon is further revealed by the coherence results in Fig. 14(b) and Fig. 14(c). By comparing these two figures, one can easily figure out that the inner raceway resembles the combined localized fault instead of the rolling ball. Besides these characteristic frequency components, the average coherence value in inner raceway is also shown to be higher than that in the rolling ball. All these results suggest that the localized inner raceway contribute most to the combined resultant spectrum in this case.

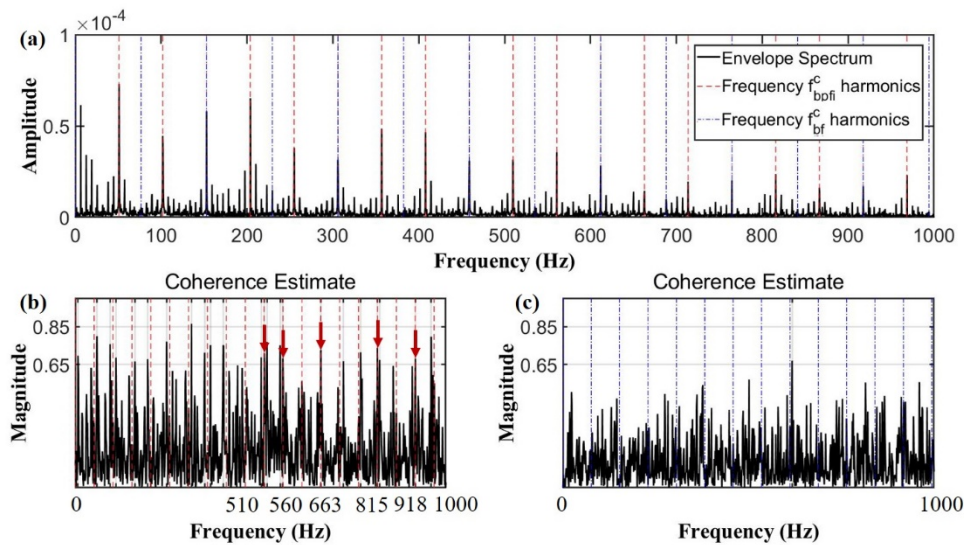


Figure 14 Effect of the inner raceway /rolling ball combined localized faults, (a) vibration spectrum due to the combined fault, (b) coherence estimation between the combined fault and the inner raceway fault, (c) coherence estimation between the combined fault and the rolling ball fault

Fig. 15(a) shows the vibration spectrum due to the outer raceway/rolling ball combined localized faults. As shown in the figure, the low frequency band (from 0Hz to 100Hz) dominates the whole resultant vibration spectrum with strong coupling effect between the outer raceway localized fault and the rolling ball localized fault. Fig. 15(b) and Fig. 15(c) give the coherence estimation between the combined fault and the outer raceway fault as well as coherence estimation between the combined fault and the rolling ball fault. It was found that the outer raceway coherence estimation was shown to be superior in terms of both the characteristic frequency and the average value. All these results suggest that the localized outer raceway contributes most to the combined resultant spectrum in this case.

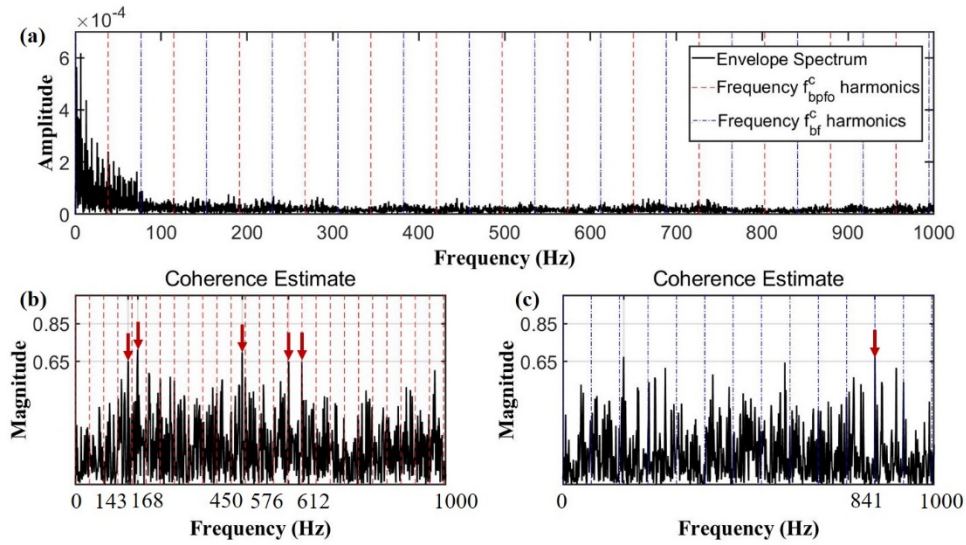


Figure 15 Effect of the outer raceway /rolling ball combined localized faults, (a) vibration spectrum due to the combined fault, (b) coherence estimation between the combined fault and the outer raceway fault, (c) coherence estimation between the combined fault and the rolling ball fault

Fig. 16(a) shows the vibration spectrum due to the inner/outer/rolling ball combined localized faults. As shown in the figure, the waveform due to the combined fault effect resembled the waveform in the case of inner/outer raceway localized fault in Fig.13 as the dominate frequency band was observed between 0Hz and 200Hz. This observation has been further proven in the coherence estimation figures (Fig. 16(b) and Fig. 16(c)) as distribution of these characteristic frequency components were similar to those in Fig. 13(b) and Fig. 13(c). There is no significant characteristic frequency components in the rolling ball coherence estimation in Fig. 16(d), which indicated that the localized outer raceway fault as well as localized inner raceway fault dominated the resultant spectrum in this case.

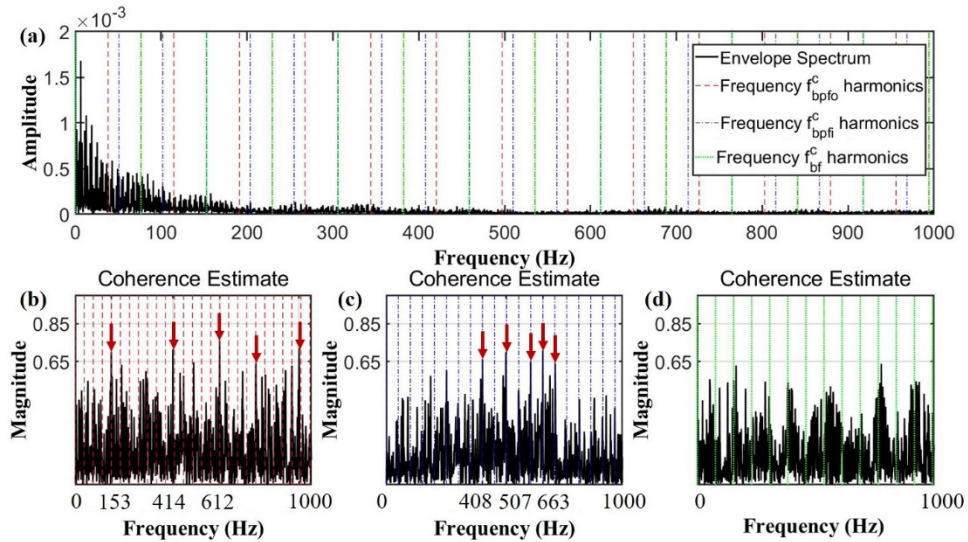


Figure 16 Effect of the inner/outer/rolling ball combined localized faults, (a) vibration spectrum due to the combined fault, (b) coherence estimation between the combined fault and the outer raceway fault, (c) coherence estimation between the combined fault and the inner raceway fault, (d) coherence estimation between the combined fault and the rolling ball fault

5. Discussion

This paper has discussed the coupling effect between different fault scenarios in planet bearing and the lumped-parameter modelling approach with detailed planet bearing model was used to simulate the faulted planet bearing dynamic responses to produce signals with well-defined characteristic frequency. However, there were several major assumptions in the theoretical model itself. Firstly, the phenomenon of “slippage” caused by the loose bearing carrier was not considered in this research. This phenomenon mainly results in two aspects as indicated by other researchers, which include a random positioning of the rolling ball around the mean position determined by the cage as well as a different cage speed compared with the kinematic value [18]. Secondly, the tooth geometrical and profile errors

were also neglected. The account of any of these effects would change the resulting gear dynamic response presented here, as has been demonstrated by other authors [38]. Another major assumption in this model is that the elasto-hydrodynamic lubrication (EHL) was not included into the bearing model. In [39], Wijnat et al. reviewed the studies related to the effect of the EHL in the rolling element bearings and it was found that the effect of EHL was moderate for medium loads and low rotational speeds and increased when the loads were low and the speed was high. In this study, the relative planet gear rotational speed is around 70 rad/s, which corresponds to 11.7 RPM. Therefore, only the Hertzian contact was considered when estimating the bearing stiffness. The purpose of all these assumptions is to ensure that the spectrum of the desired planet bearing response is relatively clear and only contains the impact from the planet bearing faults.

The faults on the planet bearing would introduce the amplitude modulation effect and in this research, a range of planet bearing fault coupling scenarios have been considered, which include planet bearing with coupling faults between distributed faults as well as planet bearing with coupling faults between localized faults. To quantify the contribution from each fault to the final resultant vibration spectrum, the coherence estimation between the combined fault and the single distributed/localized fault was calculated for each scenario and the coherence magnitude on the planet bearing fault characteristic frequency positions were examined accordingly. The detailed description about these coupling results between these faults can be found in section 4.3 and all the coupling results have been summarized as follows,

(a) Effect of planet bearing combined distributed faults:

- For the inner/outer combined distributed faults, the outer raceway distributed fault has a stronger impact on the combined vibration spectrum.
- For the inner raceway/rolling ball combined distributed faults, the inner raceway distributed fault has a stronger impact on the combined vibration spectrum.
- For the outer raceway/rolling ball combined distributed faults, the outer raceway distributed fault has a stronger impact on the combined vibration spectrum.
- For the inner/outer/rolling ball combined distributed faults, the inner raceway distributed fault has a stronger impact on the combined vibration spectrum.

From these results, a fault impact ranking can be summarized for the distributed faults,

Outer raceway distributed fault
 > inner raceway distributed fault > rolling ball distributed fault

(b) Effect of planet bearing combined localized faults:

- For the inner/outer combined localized faults, the outer raceway localized fault has a stronger impact on the combined vibration spectrum.
- For the inner raceway/rolling ball combined localized faults, the inner raceway localized fault has a stronger impact on the combined vibration spectrum.
- For the outer raceway/rolling ball combined localized faults, the outer raceway localized fault has a stronger impact on the combined vibration spectrum.
- For the inner/outer/rolling ball combined localized faults, the inner raceway and outer raceway localized faults have an equivalent impact on the combined vibration spectrum.

As a result, a similar fault impact ranking can also be summarized for the localized fault,

Outer raceway localized fault > inner raceway localized fault > rolling ball localized fault

The fault impact ranking from the combined distributed faults and combined localized faults all suggests that the fault on the outer raceway contribute most to the final compound fault vibration spectrum while the fault on the rolling ball has the minimum impact. However, this fault impact ranking was made based on the assumption that the faults on all the bearing components share the same fault parameters. For example, the fault parameter for the distributed fault is $\Pi_o = 0 \mu m$, $\Pi_p = 8 \mu m$ and $N_i = N_o = N_b = 20$ when simulating the compound distributed fault coupling effect. Once the fault parameter is changed, the resultant fault coupling spectrum will be changed accordingly and therefore, this fault impact ranking suggested in this research also changed. However, this research has provided a tool, which is called the coherence estimation method, to quantify the contribution from each single fault mode. Once the coherence result is obtained, the corresponding fault impact ranking can also be made.

In general, the compound faults are not the linear sum of single bearing faults as a vibration signal can actually be represented as a vector instead of a scalar. The resultant compound fault coupling effect can be constructed with the vector superposition method and the total vibration amplitude of compound fault for two different fault coupling can be written as,

$$|A(f)| = \sqrt{A_1^2(f) + A_2^2(f) + 2|A_1(f)||A_2(f)|\cos \varphi_{12}}^{\frac{1}{2}} \quad (36)$$

where $A_1(f)$, $A_2(f)$ is the frequency amplitude of the first fault type and the second fault type respectively, and φ_{12} is the phase angle between A_1 and A_2 . From the fault coupling results in section 4.3.3 and section 4.3.4, the vibration responses for the compound fault confirmed the vector composition for different single fault. For example, in the inner raceway /rolling ball combined localized fault scenario, both the inner raceway fault characteristic

frequency and the rolling ball fault characteristic frequency can be observed in the compound spectrum, but not in a linear composition relationship with some fault characteristic frequencies being weakened while some fault characteristic frequencies being strengthened. This phenomenon has largely been determined by the phase angle in equation (36), which produces a change of the frequency amplitude strength of the compound fault due to the different fault mechanism. When there are more faults existing in the bearing system, for example, three different faults in Fig. 12 and Fig. 16, the vector superposition method is still applied to the resultant compound fault spectrum, with equation (36) rewritten as,

$$|A(f)| = \left| \frac{A_1^2(f) + A_2^2(f) + A_3^2(f) + 2|A_1(f)||A_2(f)|\cos\varphi_{12} + 2|A_1(f)||A_3(f)|\cos\varphi_{13} + 2|A_2(f)||A_3(f)|\cos\varphi_{23}}{2} \right|^{\frac{1}{2}} \quad (37)$$

Compared with only two faults coupling, there are three phase angles that can impact the resultant spectrum and they are φ_{12} , φ_{13} and φ_{23} . As suggested in Fig. 12 and Fig. 16, the counting of additional fault has changed the resultant vibration spectrum in a nonlinear way compared with the two faults scenarios in Fig. 9, Fig. 10, Fig. 11 as well as in Fig. 13, Fig. 14, Fig. 15. It would be expected that if there is n faulted components, the number of phase angle that can impact the final resultant spectrum will be $n(n-1)/2$.

In this research, the coherence estimation was used to quantify the contribution from a single fault to the compound fault and the values of 0.65 and 0.85 are selected as two thresholds for the selection of the useful characteristic frequencies. The coherence estimation is a common statistic that can be used to examine the relation between two signals and the value is often between 1 and 0, where 1 stands for the two signal is the

same and 0 stands for no relation between the two signals. Here, 0.65 is selected as the basement line for acknowledging the similarity between the two signals, which means we only account the frequency above 0.65. In the other word, 0.65 is the PASS in our method. The value of 0.85 is selected as the basement line for a very close relationship as normally 1 is difficult to achieve. Therefore, 0.85 is the OUTSTANDING in our method. However, it should be noted that these two thresholds is not set and can be changed to other values. The vibration signal used in this paper is the instantaneous angular speed (IAS), which has been proved to be an alternative promising tool to diagnosis the bearing faults [40, 41]. The IAS signals can be naturally sampled in the angular domain and shows a greater sensitivity to different types of defects in the machine over a large bandwidth of orders. The instantaneous carrier arm angular velocity in the planetary gear has been selected to diagnosis the gear faults in this research. In [40], Ming Zhao et al. applied IAS to a planetary gearbox to detect the gear fault. This experiment further proved that the IAS method can be applied in a planetary gearbox and their results have been consistent with our previous paper for the gear fault detection. In [42], Renaudin et al. introduced the instantaneous angular speed method as an alternative way of bearing condition monitoring on an industrial vehicle gearbox. One interesting thing should be mentioned for that vehicle gearbox is that the outer race was the rotating part while the inner race was fixed. This arrangement is identical with the planetary gearbox arrangement used in this research, whose outer race is rotating with the planet gear while the inner race is fixed with the carrier arm in the rotating coordinate. Both researches have successfully detect the bearing characteristic frequency (or events per revolution) using the instantaneous angular speed method. In both results,

the inner race spectrum (the fixed component) appeared to be clear with little sideband frequency components. By contrast, the outer race spectrum in [42], (the rotating component), was shown to have a modulation effect due to the rotating mechanism, which was also observed in this research. However, the authors did not explore the effect of multi-fault effects in their research. Another important experimental work was conducted by Dr Wade Smith from the University of New South Wales [43]. An accelerometer was mounted internally on the planet carrier to measure the vibration data. In his result, only the amplitude demodulation results from the inner race fault as well as the outer race fault were presented. The results in this research also showed good agreement with his results. The fault coupling scenario considered in this research has been limited to the fault coupling between the same fault modes, including the fault coupling between the planet bearing distributed faults as well as the fault coupling between the planet bearing localized faults. Another fault coupling scenario is the fault coupling and interaction between the planet bearing distributed faults and localized faults. However, this type of fault coupling was not included in this research and it is planned for future investigation. With the consideration of the interaction between the distributed faults and localized faults, there would be 49 different combinations between the distributed faults and localized faults in total, as shown in Fig. 17. For example, on the inner raceway, there could be a distributed fault while at the same time, a localized inner raceway fault can also be existed. If so, the resultant bearing overall contact deformation δ_{ibj} in this case can be rewritten as,

$$\delta_{ibj} = (x_{ibs} - x_{ibp})\cos \Phi_{ibj} + (y_{ibs} - y_{ibp})\sin \Phi_{ibj} - c - \beta_j^{in} C_d^{in} - w^{in} \quad (38)$$

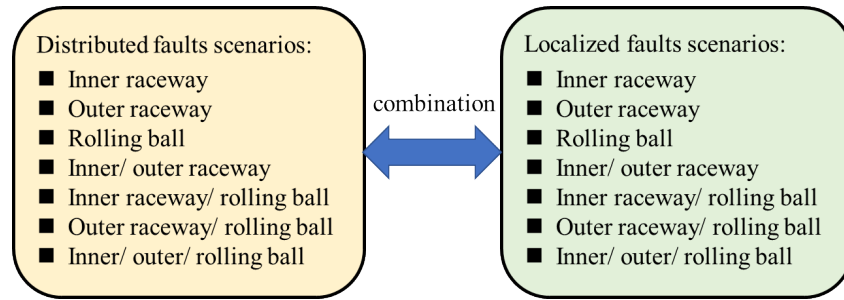


Figure 17 Coupling scenarios between planet bearing distributed fault and localized fault

6. Conclusion

This paper has investigated the planet bearing fault coupling effects in the planetary gear system. The lumped-parameter planetary gear mathematical model with detailed planet bearing has been developed to obtain the dynamic response. The bearing localized and distributed faults were modelled mathematically, and they were incorporated into the developed lumped-parameter model afterwards. A modified vibration signal processing approach was proposed based on the “benchmark” method in this research to extract the bearing transient fault impulses from the vibration signals and it was found that the proposed method is an effective way to separate and enhance the planet bearing faults for both distributed bearing faults and localized faults. The amplitude demodulation results from 22 different planet bearing fault modes (including 4 compound distributed faults and 4 compound localized faults) have been investigated and the coherence estimation over the vibration frequency domain has been proposed as a tool to quantify the similarity between the combined faults spectrum and the single fault spectrum. The investigation showed that the outer raceway fault (including both the localized fault coupling mode and the distributed fault coupling mode) had the most significant impact on the resultant planet bearing

vibration spectrum in all the investigated examples in this research. In total, this paper contributed to a deeper understanding of the coupling mechanism between the planet bearing different fault modes and provided a promising tool for the future planet bearing fault diagnosis.

Acknowledgement

This work was supported by the Natural Science Foundation of Shaanxi Province under (No. 2018JZ5001, 2019JQ548), the National Natural Science Foundation of China under (U1737211, 51905399, 51805399), the National 973 Program under No. 2015CB857100, and Tianshan Innovation Team Plan under No. 2018D14008.

Reference

- [1] X.H. Liang, M. J. Zuo, & M. R. Hoseini, Vibration signal modeling of a planetary gear set for tooth crack detection, *Engineering Failure Analysis*, 48, 2015, 185-200.
- [2] S. Xue, & I.M. Howard, Torsional vibration signal analysis as a diagnostic tool for planetary gear fault detection, *Mechanical Systems and Signal Processing*, 100, 2018, 706-728.
- [3] T.A. Harris, *Rolling Bearing Analysis*, 4th edition, John Wiley & Sons, Inc, 2001.
- [4] H. Cao, L. Niu, S. Xi, & X. Chen, Mechanical model development of rolling bearing-rotor systems: A review, *Mechanical Systems and Signal Processing*, 102, 2018, 37-58.
- [5] P.K. Gupta, Dynamics of rolling element bearings- part I: cylindrical roller bearing

- analysis, *Journal of Lubrication Technology*, 101(3), 1979, 293-302.
- [6] P.K. Gupta, Dynamics of rolling element bearings- part II: cylindrical roller bearing results, *Journal of Lubrication Technology*, 101(3), 1979, 305-311.
- [7] P.K. Gupta, Ball bearing response to cage unbalance, *Journal of Tribology*, 108(3), 1986.
- [8] P.K. Gupta, Cage unbalance and wear in ball bearings, *Wear*, 147, 1991, 93-104.
- [9] S.H. Gao, X.H. Long, G. Meng, Nonlinear response and non-smooth bifurcations of an unbalanced machine-tool spindle-bearing system, *Nonlinear Dynamics*, 54, 2008, 365–377.
- [10] N. Ghaisas, C.R. Wassgren, F. Sadeghi, Cage instabilities in cylindrical roller bearings, *Journal of Tribology*. 126, 2004, 681–689.
- [11] T. Sakaguchi, K. Harada, Dynamic analysis of cage behavior in a tapered roller bearing, *Journal of Tribology*. 128, 2006, 604–611.
- [12] C. Bovet, L. Zamponi, An approach for predicting the internal behaviour of ball bearings under high moment load, *Mechanism and Machine Theory*, 101, 2016, 1–22.
- [13] L. Niu, H. Cao, Z. He, Y. Li, An investigation on the occurrence of stable cage whirl motions in ball bearings based on dynamic simulations, *Tribology International*, 103, 2016, 12–24.
- [14] J. Takabi, M.M. Khonsari, On the thermally-induced failure of rolling element bearings, *Tribology International*, 94, 2016, 661–674.

- [15] S. Singh, C.Q. Howard, & C.H. Hansen, An extensive review of vibration modelling of rolling element bearings with localized and extended defects, *Journal of Sound and Vibration*, 357, 2015, 300-330.
- [16] P.D. McFadden, J.D. Smith, Model for the vibration produced by a single point defect in a rolling element bearing, *Journal of Sound and Vibration*, 96 (1), 1984, 69–82.
- [17] N.S. Feng, E.J. Hahn, & R.B. Randall, Using transient analysis software to simulate vibration signals due to rolling element bearing defects, In: *Proceedings of the 3rd Australian Congress on Applied Mechanics*, Sydney, 2002, 689-694.
- [18] N. Sawalhi, & R.B. Randall, Simulating gear and bearing interactions in the presence of faults- Part I: The combined gear bearing dynamic model and the simulation of localized bearing faults, *Mechanical Systems and Signal Processing*, 22(8), 2008, 1924-1951.
- [19] J. Liu, Nonlinear vibration mechanisms and modelling of defects in rolling element bearings, Ph.D thesis, Chongqing University, 2014.
- [20] Aktürk N, Uneeb M, Gohar R. The effect of number of balls and preload on vibrations associated with ball bearings [J]. *ASME Journal of Tribology*, 1997, 119(4): 747-753.
- [21] Choudhury A, Tandon N. A theoretical model to predict vibration response of rolling bearings to distributed defects under radial load [J]. *ASME Journal of Vibration and Acoustics*, 1998, 120(1): 214-220.
- [22] Tandon N, and Choudhury A. A theoretical model to predict the vibration response of

rolling bearings in a rotor bearing system to distributed defects under radial load [J].

ASME Journal of Tribology, 2000, 122(3): 609-615.

[23] Wardle F P. Vibration force produced by waviness of the rolling surface of thrust loaded ball bearing, Part 1: Theory [J]. Proceedings of the Institution of Mechanical Engineers, 1988, 202(C5): 305-312.

[24] Wardle F P. Vibration force produced by waviness of the rolling surface of thrust loaded ball bearing, Part 2: Experimental validation [J]. Proceedings of the Institution of Mechanical Engineers, 1988, 202(C5): 313-319.

[25] Jang G H, Jeong S W. Nonlinear excitation model of ball bearing waviness in a rigid rotor supported by two or more ball bearings: considering five degrees of freedom [J]. ASME Journal of Tribology, 2002, 124(1): 82-90.

[26] Jang G H, Jeong S W. Analysis of a ball bearing with waviness considering the centrifugal force and gyroscopic moment of the ball [J]. ASME Journal of Tribology, 2003, 125(3): 487-498.

[27] Bai C Q, Xu Q Y. Dynamic model of ball bearings with internal clearance and waviness [J]. Journal of Sound and Vibration, 2006, 294(1-2): 23-48.

[28] V.N. Patel, N. Tandon, R.K. Pandey, A dynamic model for vibration studies of deep groove ball bearings considering single and multiple defects in races, Journal of Tribology, 132 ,2010, 041101.

[29] S. Jain, Skidding and fault detection in the bearings of wind turbine gearboxes, Ph.D.

Thesis, University of Cambridge, 2013.

- [30] Z. Feng, H. Ma, & M.J. Zuo, Amplitude and frequency demodulation analysis for fault diagnosis of planet bearings, *Journal of Sound and Vibration*, 382, 2016, 395-412.
- [31] C. Sobie, C. Freitas, M. Nicolai, Simulation-driven machine learning: Bearing fault classification, *Mechanical Systems and Signal Processing*, 99, 2018, 403-419.
- [32] S. Xue, R. Entwistle, I. Mazhar, & I.M. Howard, The spur planetary gear torsional stiffness and its crack sensitivity under quasi-static conditions. *Engineering Failure Analysis*, 63, 2016, 106-120.
- [33] J. Sapanen, & A. Mikkola, Dynamic model of a deep groove ball bearing including localized and distributed defects, Part I: Theory, *Proceedings of the Institution of Mechanical Engineers, Journal of Multi-body Dynamics*, 217(k), 2003, 201-211.
- [34] B. Dolenc, P. Boskoski, D. Juricic, Distributed bearing fault diagnosis based on vibration analysis, *Mechanical Systems and Signal Processing*, 66-67, 2016, 521-532.
- [35] J. Liu, Y. Shao, & T.C. Lim, Vibration analysis of ball bearings with a localized defect applying piecewise response function, *Mechanism and Machine Theory*, 56, 2012, 156-169.
- [36] I.M. Howard, A review of rolling element bearing vibration- detection, diagnosis and prognosis, No. DSTO-RR-0013, Defense Science and Technology Organization (Australia), 1994.
- [37] R.B. Randall, & J. Antoni, Rolling element bearing diagnostics- a tutorial, *Mechanical*

Systems and Signal Processing, 25, 2011, 485-520.

[38] M. Inalpolat, M. Handschuh, & A. Kahraman, Influence of indexing errors on dynamic response of spur gear pairs. *Mechanical Systems and Signal Processing*, 60, 2015, 391-405.

[39] Y.H. Wijnat, J.A. Wensing, & G.C. Van Nijen, The influence of lubrication on the dynamic behavior of ball bearings, *Journal of Sound and Vibration*, 222(4), 1999, 579-596.

[40] M. Zhao, X. Jia, J. Lin, Y. Lei, J. Lee, Instantaneous speed jitter detection via encoder signal and its application for the diagnosis of planetary gearbox, *Mechanical Systems and Signal Processing*, 98,. 2018, 16-31.

[41] H. André, F. Girardin, A. Bourdon, J. Antoni, D. Rémond, Precision of the IAS monitoring system based on the elapsed time method in the spectral domain, *Mechanical Systems and Signal Processing*, 44, 2014, 14-30.

[42] L. Renaudin, F. Bonnardot, O. Musy, J.B. Doray, & D. Rémond, Natural roller bearing fault detection by angular measurement of true instantaneous angular speed, *Mechanical Systems and Signal Processing*, 24, 2010, 1998-2011.

[43] Z. Fan, & H. Li, A hybrid approach for fault diagnosis of planetary bearings using an internal vibration sensor, *Measurement*, 64, 2015, 71-80.

Appendix

The “benchmark” method developed by R.B. Randall has been adopted in this research and the signal preprocessing flow chart developed can be found in Fig. A1. There are mainly six

steps involved in this approach. The first step is linear prediction, which can be used to obtain the deterministic (i.e. “predictable”) part of a signal. However, one thing that should be mentioned here is that linear prediction is not the only method that can remove the deterministic part. Other possible methods include adaptive noise cancellation, self-adaptive noise cancellation and discrete separation method, as reported by Randall. For a given raw signal, $\vec{x}(n)$, the model used for linear prediction can be described by the following equation,

$$\vec{y}(n) = - \sum_{k=1}^p a(k) \vec{x}(n-k) \quad (A1)$$

where $\vec{y}(n)$ is the predicted signal value, $a(k)$ is the predictor coefficients and $\vec{x}(n-k)$ is the previous observed values. As described in [37], the predictor coefficients $a(k)$ can be obtained using the Yule-Walker equations, which are further solved by the Levinson-Durbin algorithm. Gear signals can be treated as the discrete frequency components, which are the predicted part $\vec{y}(n)$ in this research. The residual part of the raw signal containing the bearing signal and other noise can be obtained from,

$$\vec{z}(n) = \vec{x}(n) - \vec{y}(n) = \vec{x}(n) + \sum_{k=1}^p a(k) \vec{x}(n-k) \quad (A2)$$

After removing the deterministic components, the residual signal $\vec{z}(n)$ was further enhanced by using the minimum entropy deconvolution (MED) method in the second step.

In the method, the residual signal $\vec{z}(n)$ can be seen as composing of three components,

$$\vec{z} = \vec{h}_u * \vec{u} + \vec{h}_d * \vec{d} + \vec{h}_e * \vec{e} \quad (A3)$$

where \vec{u} is the bearing system dynamics response, \vec{d} is the impulse train caused by the bearing faults, and \vec{e} is the white noise. \vec{h}_u , \vec{h}_d , \vec{h}_e represent the system dynamics,

vibration transmission paths, and noise characteristics respectively. The deconvolution process aims to reconstruct the fault signal \vec{d} by applying FIR filter \vec{f} ,

$$\vec{w} = \vec{f} * \vec{z} = \vec{f} * \vec{h}_u * \vec{u} + \vec{f} * \vec{h}_d * \vec{d} + \vec{f} * \vec{h}_e * \vec{e} \quad (A4)$$

It would be desirable that the resultant signal \vec{w} approximates fault signal \vec{d} and this is achieved by selecting filter \vec{f} to minimize the noise effect $\vec{f} * \vec{h}_e * \vec{e} \rightarrow \vec{0}$, while closely cancelling the system $\vec{f} * \vec{h}_u * \vec{u} \rightarrow \vec{0}$, and extracting the fault signal $\vec{f} * \vec{h}_d * \vec{d} \approx \vec{d}$. The fault signal \vec{d} is assumed to be in impulse form and the filter can be selected to reach a maximum Kurtosis value. This Kurtosis maximization problem under assumed zero mean output, \vec{w} , is described as follows,

$$\max_{\vec{f}} \text{kurtosis} = \max_{\vec{f}} \frac{\sum_{n=1}^N w_n^4}{\left(\sum_{n=1}^N w_n^2\right)^2} \quad (A5)$$

By taking the derivation of equation (A5) with respect to filter coefficients \vec{f} and solve it equal to zero, and using iterative convergence, the local-maximum solution can be derived as,

$$\vec{f} = \frac{\sum_{n=1}^N w_n^2}{\sum_{n=1}^N w_n^4} (Z_0 Z_0^T)^{-1} Z_0 [y_1^3 y_2^3 \dots y_N^3]^T$$

$$Z_0 = \begin{bmatrix} z_1 & z_2 & z_3 & \dots & z_N \\ 0 & z_1 & z_2 & \dots & z_{N-1} \\ 0 & 0 & z_1 & \dots & z_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & z_{N-L+1} \end{bmatrix}$$

In the next step, the Empirical Mode Decomposition (EMD) was used to decompose the signal into its intrinsic mode functions (IMFs), which forms a complete and nearly orthogonal basis for the original signal. For the signal \vec{w} obtained in the last step, let \vec{m}_1 be the mean

of its upper and lower envelopes as determined from the cubic-spline interpolation of local maxima and minima and the first component \vec{h}_1 can be computed as,

$$\vec{h}_1 = \vec{w} - \vec{m}_1$$

In the second sifting process, \vec{h}_1 will be treated as the data, and \vec{m}_{11} is the mean of \vec{h}_1 's upper and lower envelopes,

$$\vec{h}_{11} = \vec{h}_1 - \vec{m}_{11}$$

This process will be repeated k times until \vec{h}_{1k} satisfy the convergence criteria. In this research, Cauchy convergence criteria was used to control the sifting process. After obtaining the IMFs, its kurtosis value and the correlation analysis between the IMF and the original signal can be further used to determine which IMF will be used for further analysis. For example, the k th IMF was chosen. The Hilbert transform method was used next to demodulate the amplitude of the signal. The analytic signal of the k th IMF can be written as,

$$\vec{s} = \vec{h}_{1k} + jHT(\vec{h}_{1k})$$

where $HT(\vec{h}_{1k})$ stands for the Hilbert transform of \vec{h}_{1k} . The amplitude envelope of the signal \vec{s} can be written as,

$$a = \left\{ (\vec{h}_{1k})^2 + (HT(\vec{h}_{1k}))^2 \right\}^{\frac{1}{2}}$$

At the final step, the Fast Fourier Transformation (FFT) can be used to analyse the characteristic frequency of the amplitude envelope signal.

The next signal processing section involved in this research is the fault coupling effect quantification using the coherence estimation method. After the raw signal obtained from all these fault scenarios has been processed by the first section, their corresponding

envelope spectrum can then be obtained, which can then be treated as the input for the coherence estimation between the combined fault and the single fault type. The coherence amplitude on the characteristic frequency component position can be checked and compared to quantify the fault impact contribution from different single bearing fault types.

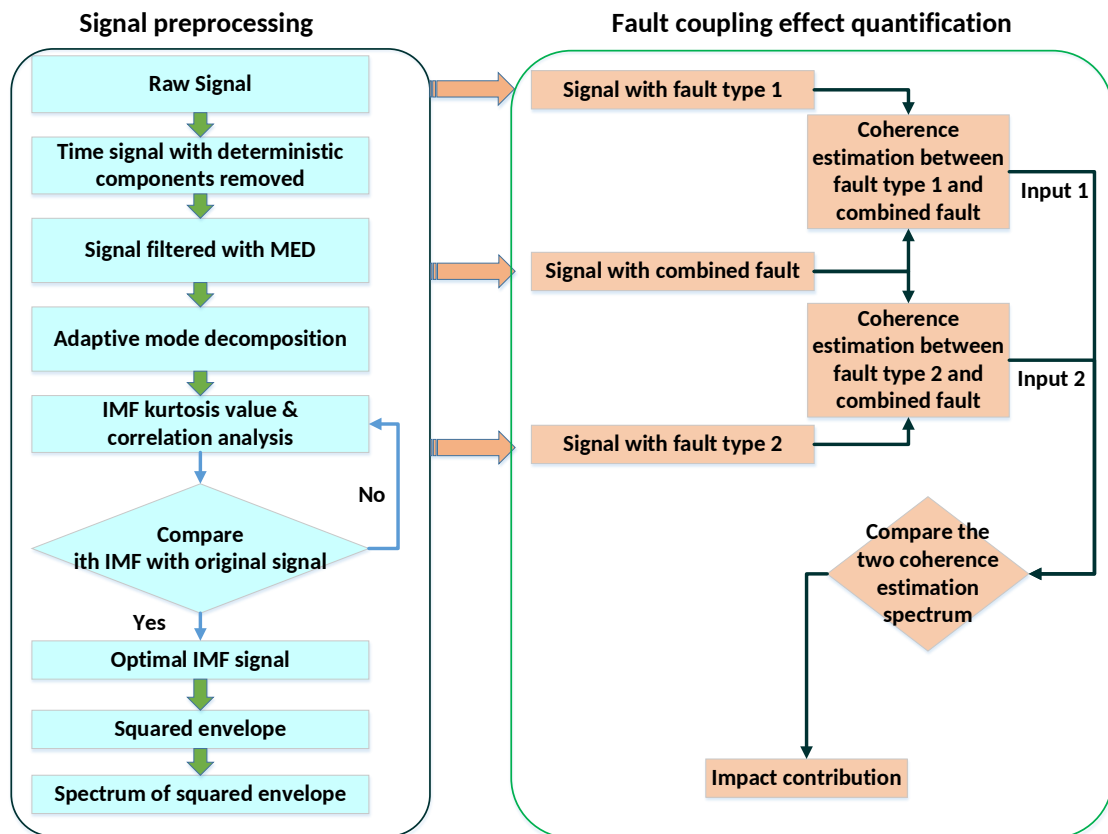


Figure A1 Flow chart of signal processing approach

Conflict of Interest

The authors declare that there is no conflict of interest in this manuscript.