

# A new dynamic model for vibration analysis of a ball bearing due to a localized surface defect considering edge topographies

Jing Liu · Yimin Shao

Received: 21 November 2013 / Accepted: 5 October 2014 / Published online: 17 October 2014  
© Springer Science+Business Media Dordrecht 2014

**Abstract** When a small initial defect occurs on the races of a ball bearing such as spalls and pits caused by fatigue, the defect will progress around the race in the direction of the ball motion until catastrophic failure happens due to impacts between the ball and the defect edges. Undesirable impulsive excitations can be caused when the orbiting ball strikes the defect edges. The impulse is depended on the shape and sizes of the defect, which can be used to detect and diagnose the defects in bearing systems. To understand characteristics of an impulse caused by a localized surface defect in a ball bearing, a new dynamic model is proposed to investigate the vibration response of a ball bearing due to a localized surface defect on its races, which can consider effects of defect edge topographies. Based on the defect edge topographies and the sizes of the defect, a new contact model for modeling contact relationships between the ball and the defect edges is also developed according to Hertzian elastic contact theory, which can be used to determine changes in the excitations including the time-varying deflection excitation and the time-varying contact stiffness excitation

caused by the defect. The proposed model is applied to investigate effects of the defect edge topographies on the contact stiffnesses between the ball and the defect edges, and the vibration response of a ball bearing with a localized surface defect on its races. The results from the proposed model are compared with the available results from the previous models in the literature, which reveals the superiority of the proposed model. It is also shown that numerical results can provide some guidance for the ball bearing defect diagnosis and detection.

**Keywords** Vibration response · Ball bearing · Localized surface defect · Defect edge topographies · Time-varying deflection and contact stiffness excitation

## List of symbols

$B$	Defect width (mm)
$C_r$	Internal radial clearance of the ball bearing (mm)
$c$	Damping coefficient (Ns/m)
$D$	Pitch diameter (mm)
$D_i$	Inner race diameter (mm)
$D_o$	Outer race diameter (mm)
$d$	Ball diameter (mm)
$E_1, E_2$	Elastic modulus associated with each contact body (MPa)
$\nu_1, \nu_2$	Poisson's ratio associated with each contact body (MPa)

---

J. Liu · Y. Shao (✉)  
State Key Laboratory of Mechanical Transmission,  
Chongqing University, Chongqing 400030, People's  
Republic of China  
e-mail: ymshao@cqu.edu.cn

J. Liu  
College of Mechanical Engineering, Chongqing University,  
Chongqing 400030, People's Republic of China

$E_{eq}$	Equivalent modulus of elasticity (MPa)	$\delta_{hr}$	Contact deformation between the ball and healthy race (mm)
$e$	Elliptical eccentricity parameter	$\theta_0$	Initial angular offset of the defect to the $j$ th ball (rad)
$F_r$	Radial force applied on the ball (N)	$\theta_d$	Arc length of the defect in the tangential direction (rad)
$H$	Defect depth (mm)	$\theta_h$	Half of the arc length of the defect (rad)
$H_1, H_2, H_3, H_4$	Time-varying deflection excitations (mm)	$\theta_{dj}$	Race contact angle (rad)
$j$	Ball number	$\Delta \theta$	Arc length caused by the defect with different edge topographies (rad)
$K$	Total contact stiffness (N/mm)	$\lambda_j$	Loading zone parameter of the $j$ th ball
$K_e$	Time-varying contact stiffness between the ball and races (N/mm)	$\varepsilon$	Second kind of elliptical integral
$K_i$	Contact stiffness between the ball and inner race without defect (N/mm)	$\xi_d$	Ratio of the defect length to its width
$K_o$	Contact stiffness between the ball and outer race without defect (N/mm)	$\xi_{bd}$	Ratio of the ball size to the defect minimum size
$K_p$	Contact stiffness between the ball and smooth surface (N/mm)	$\Sigma \rho$	Curvature sum
$K_{ts}$	Total contact stiffness between the ball and races with defect (N/mm)	$\omega_c$	Speed of the cage (rad/s)
$K_{t0}$	Contact stiffness between the ball and races with first type defect (N/mm)	$\omega_r$	Speed of the shaft (rad/s)
$k$	First and second kind of elliptical integral		
$L$	Defect length (mm)		
$l$	Length of the small surfaces at the defect edges (mm)		
$N_f$	Number of samples		
$n_d$	Load-deflection exponent for the first defect type		
$n_s$	Numbers of the contact surfaces between the ball and defect edges		
$Q_x, Q_y$	Components of external force applying on the shaft (N)		
$R$	Radius of the ball (mm)		
$r_o$	Radius of curvature of the outer race (mm)		
$r_i$	Radius of curvature of the inner race (mm)		
$s$	Number of the contact surfaces		
$t$	Time (s)		
$x, y$	Displacement responses in $X$ - and $Y$ -direction (mm)		
$Z$	Number of the ball		
$\alpha$	Contact angle ( $^\circ$ )		
$\gamma$	Elevation angle of the small surfaces at defect edges (rad)		
$\delta_{dr}$	Contact deformation between the ball and defect (mm)		

## 1 Introduction

Ball bearing is one of the key components in a variety of industrial machineries due to their carrying capacity and low-friction characteristics. Vibration performances of these machineries are greatly influenced by vibration characteristics of their internal bearings, especially in the presence of various defects [1]. Hence, ball bearing defect diagnosis and detection is one of the important works in industrial maintenance. Although many methods have been proposed for diagnosis and detection of the defects in rolling element bearings, modeling and simulation method is an accurate approach to predict the vibration response of ball bearings [2], which can enhance understanding vibration characteristics of the bearing with various defects.

Many research works have been presented to investigate the vibration response of rolling element bearings with a localized surface defect on their races or rolling elements. Some of the previous works focusing on modeling the force excitations produced by the localized defects are listed in Refs. [3–12]. On the other hand, some research works are interested in simulating the deflection excitations caused by the localized surface defects as shown in Refs. [2, 13–24].

McFadden and Smith [3,4] used a series of repeated force excitations to model the impulses generated by single and multiple point defects. Tandon and Choudhury [5,6] used the finite width triangular, rectangular, and half-sine force excitations to describe the impulse generated by a localized surface defect. Kiral and Karagulle [7,8] described a localized surface defect model as a rectangular force excitation. Sassi et al. [9] formulated the force excitation model for a localized surface defect as sinusoidal and rectangular functions.

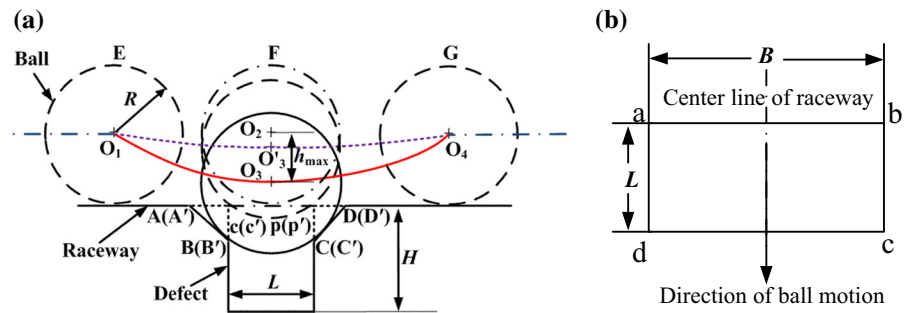
References [2,13–15,17–21] defined a localized surface defect model as a deflection excitation using rectangular function. Rafsanjani et al. [2] defined a localized surface defect model as a deflection excitation using a rectangular function. In their model, the defect depth is used to define the amplitude of the deflection excitation. References [13–17,19–21] also used the rectangular deflection excitation model to describe the impulse generated by a localized surface defect. Patil et al. [16] formulated a localized surface defect model as a half-sine deflection excitation wave. Pate et al. [18] also described a localized surface defect model as a rectangular deflection excitation function. In their model, the geometric relationship between the rolling element and the defect is considered. References [10] and [11] analyzed the contact pressure between the rolling element and the defect, but they only considered the constant equivalent contact stiffness between the rolling element and the defect. However, the contact stiffness between the rolling element and the defect varies with the contact position between the rolling element and the defect when the rolling element passes over the defect according to the analysis in Refs. [22–24]. Liu et al. [22] used a finite element method and an experimental method to study effects of the shapes of a localized surface defect on the vibration waveforms of a ball bearing with a localized surface defect on its outer race. In another work [23], they proposed a new localized surface defect model considering the time-varying deflection excitation caused by a localized surface defect for a ball bearing with a localized surface defect on its races. Relationships between the time-varying deflection excitations and the shape and sizes of the defect are considered in their defect model. Moreover, Shao et al. [24] proposed a new method to formulate the time-varying deflection excitation and the time-varying contact stiffness excitation caused by a localized surface defect on the races of a cylindrical roller bearing. In practical, when a small initial spall occurs on the race

of the ball bearing, the spall will progress around the race in the direction of the ball motion until catastrophic failure happens due to impacts between the ball and the defect edges [25]. However, in the above works, effects of the defect edge topographies on the vibration response of the ball bearing were not considered.

Several studies have been conducted on studying the spall growth and the spall growth rate of an initial spall on the races of the rolling element bearings based on the experimental and modeling methods. For instance, Lundburg and Palmgren [26] initially discussed the spall growth phenomenon of the rolling bearings. Kotzalas and Harris [27] studied the spall growth on the surfaces of the balls and races of a ball bearing, and they extended the prediction methods of bearing life in Ref. [28] to predict the remaining useful life of the bearing. Xu and Sadeghi [29] proposed a new analytical model to investigate the effect of a dent on the spall initiation and growth in the lubricated contact. Hoeprich [30] studied the randomness inherent to the spall growth and its unknown governing mechanisms based on the spall growth experiments on tapered roller bearings. Rosada et al. [31], Arakere et al. [32], and Forster et al. [33] also proposed a series of experimental and modeling methods to study the spall growth of the bearing materials in their three-part analysis series, which studied three stages of the spall growth of the bearing materials. Branch et al. [34,35] presented a finite element modeling method to investigate three stages of the spall growth in Refs. [31–33]. Li et al. [36] presented an empirical method to predict spall growth rates of a tapered roller bearing. The above works on the spall growth are focused on the contact stress, fatigue, material microstructure, and plasticity at the edges of the spall defect. Effects of the spall edge topographies on the impulse for the ball bearing were not discussed. Hence, it is helpful to investigate the impulse caused by the spall with different edge topographies for the ball bearing defect diagnosis and detection at its early stage.

The main objective of the present study is to investigate effects of defect edge topographies on the time-varying contact stiffnesses between the ball and the defect edges according to Hertzian elastic contact theory, and the vibration response of the ball bearing with a localized surface defect on its races. The profiles of the propagated defect edges are assumed to be small smooth plane surfaces partly based on the assumptions in Refs. [34] and [35]. The elevation angles and the

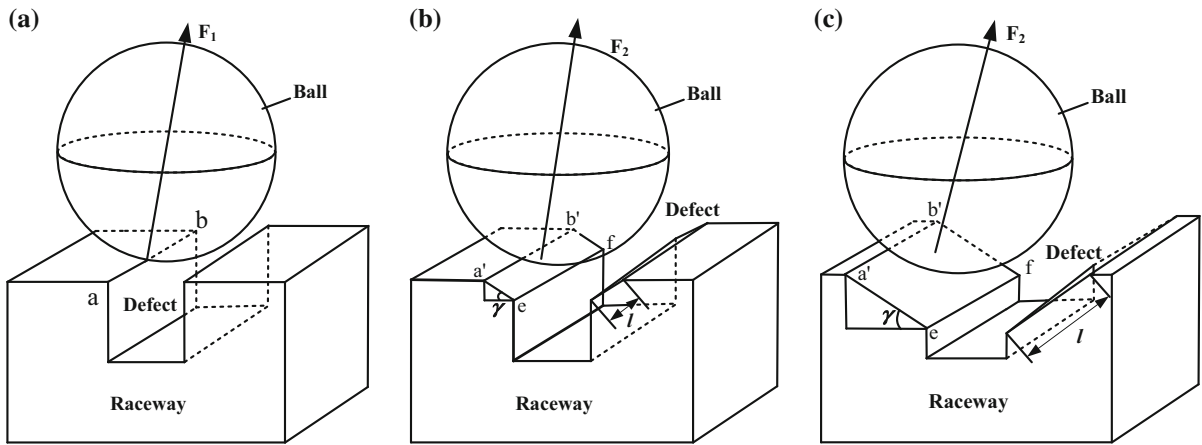
**Fig. 1** Schematic of a localized surface defect on the race of a ball bearing: **a** contact positions between the ball and the defect, and **b** defect location



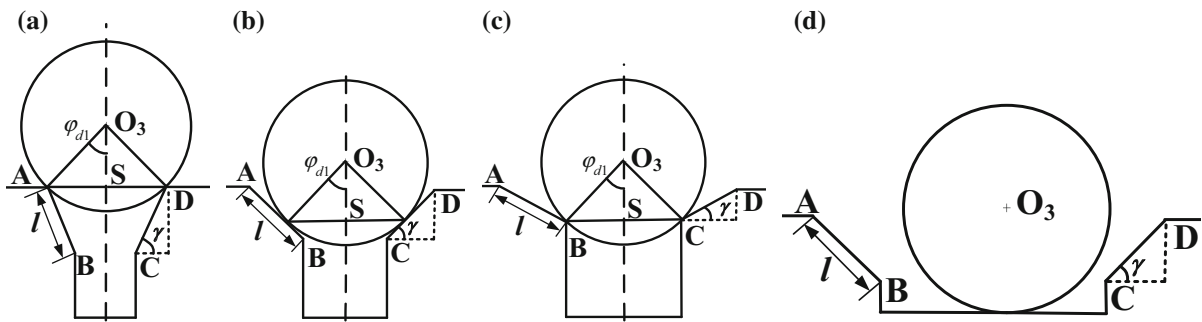
lengths of the small surfaces at the edges of the defect are used to determine the defect edge topographies. The defect model is formulated partly based on the studies in Refs. [22] and [23], which can consider the time-varying deflection excitation and the time-varying contact stiffness excitation caused by the defect with different edge topographies. Relationships between the excitations and the shape and sizes of the defect are established. The time-varying deflection excitation is utilized to describe the additional deflection in the radial and tangential directions caused by the defect, which is an extended model of the radial deflection model in Ref. [23]. The time-varying contact stiffness is used to describe contact deformations at the defect edges when the ball is in contact with the defect edges. The dynamic model of the ball bearing is considered as a two-degree-of-freedom lumped parameter model. It includes effects of the time-varying compliance of the bearing, the damping, the time-varying deflection excitation in the radial and tangential directions, and the time-varying contact stiffness excitation caused by the defect. A fourth-order Runge–Kutta numerical integration method with a fixed time step is adopted to calculate the vibration response of the ball bearing with a localized surface defect on its races. Effects of the defect edge topographies on the time-varying contact stiffnesses between the ball and the defect edges, and the vibration response of the ball bearing are investigated. Numerical results from the proposed model are compared with the available results from the previous models in the literature, which reveals the superiority of the proposed model. It is expected that by applying the proposed method, the relationship between the vibration response of the ball bearing and the defect edge topographies can be obtained, and some guidance for the fault diagnosis and deflection of ball bearings can be provided.

## 2 Problem formulation

When a localized surface defect occurs on the races of a ball bearing, as shown in Fig. 1, the ball will strike the defect edges when the ball passes over the defect. The periodic impacts arising from the balls against the defect edges will change the topographies of the defect edges due to the plastic deformations at the defect edges. The shapes of the line edges of the defect will be changed according to the analysis in Refs. [34] and [35]. The propagated defect edges are assumed as small smooth plane surfaces in this paper. For instance, a line edge  $ab$  of the defect at its initial stage is changed into a small plane surface  $a'b'ef$  as shown in Fig. 2. Figure 2b, c shows the different stages of the small plane surface at the defect edge. Here, it is assumed that the small plane surfaces at the defect edges have same sizes. The changes in the defect edge topographies will affect the contact relationships between the ball and the defect edges. For a healthy race, the contact type between a ball and the race is a ball–ball contact type as shown in Fig. 1a. For a defective race with line defect edges, as shown in Fig. 1a, the contact type between the ball and the defective race becomes a ball–line contact type, which has been discussed in Refs. [22] and [23]. However, when the defect edge topographies are changed due to the impacts between the ball and the defect edges, as shown in Fig. 1a, the contact type between the ball and the defect edges become a ball–plane contact type. It depends on three parameters as follows: (1) the ratio of the defect length to its width, which is defined as  $\xi_d = L/B$ , where  $L$  is the defect length, and  $B$  is the defect width; (2) the ratio of the ball size to the defect minimum size, which is defined as  $\xi_{bd} = d/\min(L, B)$ , where  $d$  is the ball diameter; and (3) the elevation angle  $\gamma$  ( $0 < \gamma < \pi/2$ ) of the small plane surface as shown in Fig. 2b, c.



**Fig. 2** Schematic of the contact type between a ball and one defect edge (enlarged view): **a** line edge, **b** very small surface edge, and **c** small surface edge



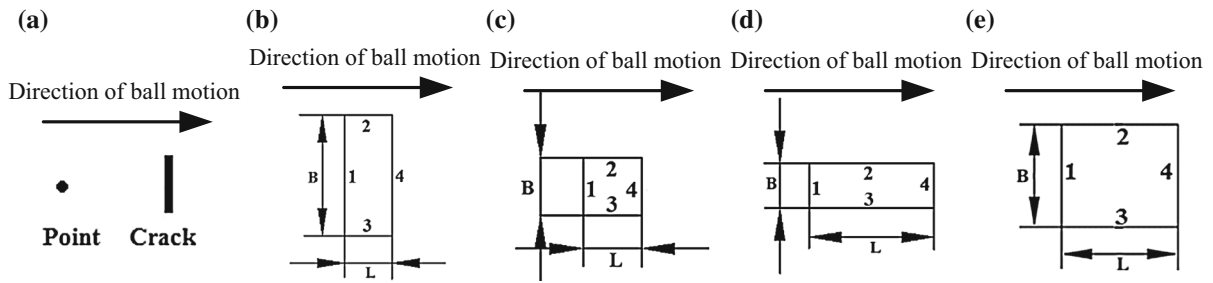
**Fig. 3** Schematic of the contact type between a ball and the small surfaces at defect edges with different topographies: **a** case 1, **b** case 2, **c** case 3, and **d** case 4

Figure 3 shows effects of the defect edge topographies on the contact types between the ball and the different defect cases. The elevation angle  $\gamma$  and the length  $l$  of the small surfaces at the defect edges are used to determine the contact types between the ball and the defects. It is assumed that  $l$  is larger than Hertzian contact radius in this paper.

For case 1,  $\gamma \geq \arcsin\left(0.5 \min(L, B)/\sqrt{R^2 + l^2}\right) + \arctan(l/R)$  and  $\xi_{bd} > 1$ , where  $R$  is the radius of the ball, as shown in Fig. 3a. The ball is only in contact with the line edges of the defect in considering a case of an early defect stage. For this case, the contact type between the ball and the defect is a ball–line contact type when the ball passes over the defect. It is the early stage of the plastic deformations at the defect edges caused by the periodic impacts arising from the balls against the defect edges for a small localized surface defect.

For case 2,  $\min(L, B)/(2R) < \gamma \leq \arcsin\left(0.5 \min(L, B)/\sqrt{R^2 + l^2}\right) + \arctan(l/R)$  and  $\xi_{bd} > 1$ , as shown in Fig. 3b, the contact type between the ball and the defect depends on the defect parameters  $\xi_d$  and  $\xi_{bd}$ . The contact type between the ball and the beginning edge of the defect is the ball–line contact type, as well as that between the ball and the ending edge of the defect. It is a ball–plane contact type when the ball is in contact with the small surfaces between the points A and D. It is the middle stage of the plastic deformations at the defect edges caused by the periodic impacts arising from the balls against the defect edges for a small localized surface defect.

For case 3,  $\gamma \leq \arcsin(\min(L, B)/(2R))$  and  $\xi_{bd} > 1$ , as shown in Fig. 3c, the contact type between the ball and the defect is a ball–plane contact type when the ball is in contact with the surfaces between points A and B,



**Fig. 4** Schematics of the different localized surface defects: **a** a small crack, **b** a small defect with the defect length smaller than its width, **c** a small defect with the defect length similar to

its width, **d** a small defect with the defect length larger than its width, and **e** a large defect compared to the ball diameter

and between points C and D. It is a ball–line contact type when the ball makes contact with the beginning edge, the edges at points B and C, and the ending edge of the defect. This case denotes the large plastic deformations at the defect edges. It is the late stage of the plastic deformation at the defect edge caused by the periodic impacts arising from the balls against the defect edges for a small localized surface defect.

For case 4,  $\xi_{bd} \leq 1$ , or  $h_{max} < H$ , as shown in Figs. 1a and 3d, the contact type between the ball and the defect is a ball–line contact type at the points A, B, C, and D; it is the ball–plane contact type when the ball is in contact with the surface between the points A and B, the bottom surface, and the surface between the points C and D of the defect. It is used to describe the plastic deformations at the defect edges caused by the periodic impacts arising from the balls against the defect edges for a large localized surface defect.

For cases 2 and 3, as shown in Fig. 3b, c, the numbers of the contact surfaces are determined by the defect sizes. Here, five different defect types are discussed according to the defect parameters  $\xi_d$  and  $\xi_{bd}$ .

Figure 4a depicts the first defect type, which denotes a very small defect such as a point defect and a small crack. Thus, effects of the edge topographies of the first defect type can be ignored.

Figure 4b plots the second defect type with the length of the defect smaller than its width. The ball only strikes the edges 1 and 4 when the ball passes over the defect, and the numbers of the contact surfaces are 1, 2, and 1, respectively.

For the third defect type with the length of the defect similar to its width, as shown in Fig. 4c, the ball strikes all defect edges 1–4 when the ball passes over the defect, and the numbers of the contact surfaces are 1, 3, 4, 3, and 1, respectively.

Figure 4d gives the fourth defect type with the length of the defect larger than its width. The ball also strikes all defect edges 1–4, and the numbers of the contact surfaces are 1, 3, 2, 3, and 1, respectively.

The fifth defect type is presented in Fig. 4e, the ball strikes the edge 1, the bottom surface of the defect, and the edge 4 when the ball passes over the defect. There is only one contact surface during the entire contact process.

### 3 Dynamic model of a ball bearing with a localized surface defect

#### 3.1 Localized surface defect model

When a localized surface defect occurs on the races of a ball bearing, an excitation will be generated, which includes the time-varying deflection excitation and the time-varying contact stiffness excitation [23]. Reference [23] presents a time-varying deflection excitation model to describe the additional deflections caused by the defect. In their model, only the radial contact deflections at the defect edges are considered, which are calculated by a static finite element method. However, the time-varying contact stiffness excitation between the ball and the defect is ignored. In this paper, the time-varying deflection excitation, the time-varying contact stiffness excitation, and the defect edge topographies are considered in the proposed model. The contact stiffnesses between the ball and the defect edges are calculated according to Hertzian elastic contact theory.

For a healthy ball bearing, as shown in Fig. 1a, the total contact stiffness  $K$  between a ball and the races can be calculated by [37]



$$K = \frac{1}{\left[ \frac{1}{K_i^{1/n}} + \frac{1}{K_o^{1/n}} \right]^n} \tag{1}$$

where  $n$  is the load-deflection exponent,  $n = 1.5$  for a ball bearing;  $K_i$  and  $K_o$  are the contact stiffnesses between the ball and the inner race, and that between the ball and the outer race, respectively. For a healthy ball bearing, the contact stiffness  $K_h$  between the ball and the race can be calculated by [23,37]

$$K_h = \left( \frac{\pi^2 e^2 E_{eq}^2 \varepsilon}{4.5k^3 \sum \rho} \right)^{0.5} \tag{2}$$

in which  $\sum \rho$  is the curvature sum;  $E_{eq}$  is the equivalent modulus of elasticity;  $k$  and  $\varepsilon$  denote the first and second kind of elliptical integral, respectively; and  $e$  is the elliptical eccentricity parameter.  $\sum \rho$ ,  $E_{eq}$ ,  $k$ ,  $\varepsilon$ , and  $e$  can be calculated by the methods presented in Refs. [23] and [37].

When a localized surface defect occurs on the inner race or the outer race of the ball bearing, as shown in Fig. 1a, the contact type between the ball and the defect is no longer a ball–ball contact type since the lined edges of the defect become small surfaces due to the impacts between the ball and the defect edges. Therefore, the contact stiffnesses between the ball and the defect edges cannot be calculated by Eq. (2). In this paper, the small surfaces at the defect edges are assumed as small smooth surfaces as shown in Fig. 1a. As a result, the contact between the ball and one edge of the defect can be considered as a ball–plane contact type. The contact stiffness  $K_p$  between the ball and the smooth surface can be calculated by Hertzian elastic contact theory, which is given by [38]

$$K_p = \frac{4R^{\frac{1}{2}}}{3 \left( \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \right)} \tag{3}$$

where  $R$  is the radius of the ball;  $E_1$ ,  $E_2$  are the elastic modulus associated with each contact body; and  $\nu_1$ ,  $\nu_2$  are the Poisson’s ratio associated with each contact body.

When the ball makes contact with the small surfaces at the defect edges, the new total contact stiffness  $K_t$  between the ball and the races with a localized surface defect cannot be calculated by Eq. (1). Here, a new calculation method is proposed to calculate the new total contact stiffness  $K_t$  for a ball bearing with a localized surface defect on its races.

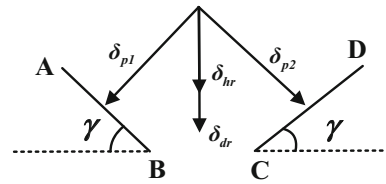


Fig. 5 Contact deflections between the ball and the small surface at the defect edges

As shown in Fig. 5, the total contact deformation  $\delta_{tr}$  between the healthy and the defective races can be calculated by

$$\delta_{tr} = \delta_{hr} + \delta_{dr} = \left( \frac{F_r}{K_h} \right)^{2/3} + \left( \frac{F_r \cos \gamma}{n_s K_p} \right)^{2/3} \cos \gamma, \quad n_s = 1, 2, 3, \text{ and } 4 \tag{4}$$

where  $\delta_{hr}$  is the contact deformation between the ball and the healthy race,  $\delta_{dr}$  is the contact deformation between the ball and the defect,  $F_r$  is the radial force applied on the ball, and  $n_s$  is the numbers of the contact surfaces between the ball and the defect edges. Then, the total contact stiffness  $K_{ts}$  between the ball and the races can be calculated by

$$K_{ts} = \frac{1}{\left[ \left( \frac{1}{K_h} \right)^{2/3} + \left( \frac{(\cos \gamma)^{5/2}}{n_s K_p} \right)^{2/3} \right]^{3/2}}, \quad n_s = 1, 2, 3, \text{ and } 4 \tag{5}$$

where the subscript  $s$  denotes the numbers of the contact surfaces.

For the first type of a small defect such as a point defect and a small crack, as the descriptions in Fig. 4, the contact stiffness between the ball and the defect can be considered as a constant  $K_{d1}$  [23]. The total deformation between the healthy and the defective races is calculated by

$$\delta_{tr} = \delta_{hr} + \delta_{dr} = \left( \frac{F_r}{K_h} \right)^{2/3} + \left( \frac{F_r}{K_{d1}} \right)^{1/n_d} \tag{6}$$

where  $n_d$  denotes the load-deflection exponent for the first defect type, and it depends on the shape and sizes of the defect. Since the sizes of the first defect type are very small,  $n_d$  can be assumed to be 1.5 for the first defect type. The total contact stiffness  $K_{t0}$  between the ball and the races with the first type defect can be obtained by a polynomial fitting method according to the load-deflection relationship listed in Eq. (6).

For the first defect type, the contact stiffness  $K_e$  between the ball and the races can be defined as

$$K_e = \begin{cases} K_{t0} & |\text{mod}(\theta_{dj}, 2\pi) - \theta_0 - \theta_e - 0.5\Delta\theta| \\ & \leq \theta_e + 0.5\Delta\theta \\ K & \text{otherwise} \end{cases} \quad (7)$$

where the operator  $\text{mod}()$  denotes the modulus after the division arithmetic operation,  $\theta_e$  is half of the arc length of the defect in the tangential direction;  $K_{t0}$  is the total contact stiffness between the ball and the races when a first type defect occurs on the surface of the inner race or the outer race;  $\Delta\theta$  is the arc length of the propagated surface at the defect edges in the tangential direction, which is given by

$$\Delta\theta = \begin{cases} \arcsin\left(\frac{2l \cos \gamma}{D_i}\right) & \text{defect on inner race} \\ \arcsin\left(\frac{2l \cos \gamma}{D_o}\right) & \text{defect on outer race} \end{cases} \quad (8)$$

where  $D_i$  and  $D_o$  are the diameter of the inner race and the outer race, respectively; and  $\theta_0$  is the initial angular offset of the defect to the  $j$ th ball, which is given by

$$\theta_0 = \begin{cases} \frac{2\pi}{Z}(j-1) + \beta_i & \text{defect on inner race} \\ \frac{2\pi}{Z}(j-1) + \beta_o & \text{defect on outer race} \end{cases} \quad (9)$$

where  $Z$  is the number of the ball,  $\beta_i$  and  $\beta_o$  are the initial angular offsets of the defect of the first ball when the defect occurs on the inner race and the outer race, respectively;  $j$  denotes the  $j$ th ball,  $j$  is 1 to  $Z$ , and  $\theta_{dj}$  is the race contact angle, which is given by

$$\theta_{dj} = \begin{cases} \frac{2\pi}{Z}(j-1) + (\omega_c - \omega_r)t + \beta_i & \text{for an inner race contact angle} \\ \frac{2\pi}{Z}(j-1) + \omega_c t + \beta_o & \text{for an outer race contact angle} \end{cases} \quad (10)$$

where  $\omega_c$  is the rotational speed of the cage,  $\omega_r$  is the rotational speed of the rotor. For the second defect type, the contact stiffness  $K_e$  can be described by

$$K_e = \begin{cases} K_{t1} & \theta_0 - \Delta\theta \leq \text{mod}(\theta_{dj}, 2\pi) \\ & < 0.5\theta_d \text{ and } 0.5\theta_d < \text{mod}(\theta_{dj}, 2\pi) \\ & \leq \theta_d + \Delta\theta \\ K_{t2} & \text{mod}(\theta_{dj}, 2\pi) = 0.5\theta_d \\ K & \text{otherwise} \end{cases} \quad (11)$$

where  $K_{t1}$  is the total contact stiffness between the ball and the races when the number of the contact surface is 1,  $K_{t2}$  is the total contact stiffness between the ball and the races when the number of the contact surfaces is 2,  $\theta_d$  is the angular position of the defect as follows

$$\theta_d = \begin{cases} \theta_0 + \arcsin\left(\frac{L+2l \cos \gamma}{D_i}\right) & \text{defect on inner race} \\ \theta_0 + \arcsin\left(\frac{L+2l \cos \gamma}{D_o}\right) & \text{defect on outer race} \end{cases} \quad (12)$$

For the third defect type, the contact stiffness  $K_e$  can be written as

$$K_e = \begin{cases} K_{t1} & \theta_0 \leq \text{mod}(\theta_{dj}, 2\pi) \leq \theta_0 \text{ and } \theta_d \\ & \leq \text{mod}(\theta_{dj}, 2\pi) \leq \theta_d \\ K_{t3} & \theta_0 < \text{mod}(\theta_{dj}, 2\pi) < 0.5\theta_d \text{ and } 0.5\theta_d \\ & < \text{mod}(\theta_{dj}, 2\pi) < 0.5\theta_d \\ K_{t4} & \text{mod}(\theta_{dj}, 2\pi) = 0.5\theta_d \\ K & \text{otherwise} \end{cases} \quad (13)$$

where  $K_{t3}$  is the total contact stiffness between the ball and the races when the number of the contact surfaces is 3,  $K_{t4}$  is the total contact stiffness between the ball and the races when the number of the contact surfaces is 4. For the fourth defect type, the contact stiffness  $K_e$  can be given by

$$K_e = \begin{cases} K_{t1} & \theta_0 - \Delta\theta \leq \text{mod}(\theta_{dj}, 2\pi) \leq \theta_0 \text{ and } \theta_d \\ & \leq \text{mod}(\theta_{dj}, 2\pi) \leq \theta_d + \Delta\theta \\ K_{t3} & \theta_0 < \text{mod}(\theta_{dj}, 2\pi) \leq \theta_1 \text{ and } \theta_2 \\ & \leq \text{mod}(\theta_{dj}, 2\pi) < \theta_d \\ K_{t2} & \theta_1 < \text{mod}(\theta_{dj}, 2\pi) < \theta_2 \\ K & \text{otherwise} \end{cases} \quad (14)$$

where  $\theta_1$  is defined by

$$\theta_1 = \begin{cases} \theta_0 + \arcsin\left(\frac{0.5(B+2l \cos \gamma)}{D_i}\right) & \text{defect on inner race} \\ \theta_0 + \arcsin\left(\frac{0.5(B+2l \cos \gamma)}{D_o}\right) & \text{defect on outer race} \end{cases} \quad (15)$$

and  $\theta_2$  is written as

$$\theta_2 = \begin{cases} \theta_d - \arcsin\left(\frac{0.5(B+2l \cos \gamma)}{D_i}\right) & \text{defect on inner race} \\ \theta_d - \arcsin\left(\frac{0.5(B+2l \cos \gamma)}{D_o}\right) & \text{defect on outer race} \end{cases} \quad (16)$$

For the fifth defect type, the contact stiffness  $K_e$  can be given by

$$K_e = \begin{cases} K_{t1} & \theta_0 - \Delta\theta \leq \text{mod}(\theta_{dj}, 2\pi) < \theta_{f1} \text{ and } \theta_{f2} \\ & < \text{mod}(\theta_{dj}, 2\pi) \leq \theta_d + \Delta\theta \\ K_{t5} & \theta_{f1} \leq \text{mod}(\theta_{dj}, 2\pi) \leq \theta_{f2} \\ K & \text{otherwise} \end{cases} \quad (17)$$

where  $K_{t5}$  is the contact stiffness between the ball and the bottom surface of the defect, it can be considered as the contact between a ball and a flat smooth surface, and it can be determined by

$$K_{t5} = \frac{1}{\left[\left(\frac{1}{K_b}\right)^{2/3} + \left(\frac{1}{K_p}\right)^{2/3}\right]^{3/2}} \quad (18)$$



and  $\theta_{f1}$  is defined by

$$\theta_{f1} = \begin{cases} \arcsin\left(\frac{\sqrt{(0.5d)^2 - (0.5d - H)^2}}{D_i}\right) + l \cos \gamma & \text{defect on inner race} \\ \arcsin\left(\frac{\sqrt{(0.5d)^2 - (0.5d - H)^2}}{D_o}\right) + l \cos \gamma & \text{defect on outer race} \end{cases} \quad (19)$$

where  $d$  is the diameter of the ball,  $H$  is the defect depth, and  $\theta_{f2}$  is given by

$$\theta_{f2} = \begin{cases} \theta_d - \arcsin\left(\frac{\sqrt{(0.5d)^2 - (0.5d - H)^2}}{D_i}\right) - l \cos \gamma & \text{defect on inner race} \\ \theta_d - \arcsin\left(\frac{\sqrt{(0.5d)^2 - (0.5d - H)^2}}{D_o}\right) - l \cos \gamma & \text{defect on outer race} \end{cases} \quad (20)$$

On the other hand, there is an additional deflection due to effects of the presence of the defect with different edge topographies. As shown in Fig. 3, the additional deflection depends on the ratio of the defect length to its width, the ratio of the ball size to the defect minimum size, and the elevation angles of the small surfaces at the defect edges. According to the descriptions in Fig. 3, the maximum of the addition deflection can be obtained. For case 1

$$h_{\max 1} = 0.5d - ((0.5d)^2 - (0.5 \min(L, B) + l \cos \gamma)^2)^{0.5} \quad (21)$$

For case 2

$$h_{\max 2} = 0.5d - ((0.5d)^2 - (0.5 \min(L, B))^2)^{0.5} + (0.5 \min(L, B) + l \cos \gamma - 0.5d \sin \gamma) \tan \gamma \quad (22)$$

For case 3

$$h_{\max 3} = 0.5d - ((0.5d)^2 - (0.5 \min(L, B))^2)^{0.5} + l \sin \gamma \quad (23)$$

For case 4

$$h_{\max 4} = H \quad (24)$$

Then, the time-varying deflection excitation  $H'$  is given by a piecewise response function, including both the half-sine and the rectangular functions [23], which is defined by

$$H' = \begin{cases} H_1 & \xi_{bd} >> 1 \\ H_2 & \xi_{bd} > 1 \text{ and } \xi_d \leq 1 \\ H_3 & \xi_{bd} > 1 \text{ and } \xi_d > 1 \\ H_4 & \xi_d \leq 1 \end{cases} \quad (25)$$

where  $H_1$ ,  $H_2$ ,  $H_3$ , and  $H_4$  are the time-varying deflection excitations caused by the first through fifth defect types, respectively, which have been given in Ref. [23].  $H_1$  is given by

$$H_1 = \begin{cases} H & |\text{mod}(\theta_d j, 2\pi) - \theta_0 - \theta_e - 0.5\Delta\theta| \leq \theta_e + 0.5\Delta\theta \\ 0 & \text{otherwise} \end{cases} \quad (26)$$

$H_2$  is defined by

$$H_2 = \begin{cases} h_{\max 2} \sin\left(\frac{0.5\pi}{\Delta\theta_2} (\text{mod}(\theta_d j, 2\pi) - \theta_0)\right) \theta_0 & \leq \text{mod}(\theta_d j, 2\pi) \leq \theta_0 + \Delta\theta_2 \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

$\Delta\theta_2$  is given by

$$\Delta\theta_2 = \begin{cases} \arcsin\left(\frac{L+2l \cos \gamma}{D_o}\right) & \text{defect on outer race} \\ \arcsin\left(\frac{L+2l \cos \gamma}{D_i}\right) & \text{defect on inner race} \end{cases} \quad (28)$$

$H_3$  is defined by

$$H_3 = \begin{cases} h_{\max 3} \sin\left(\frac{0.25\pi}{\Delta\theta_1} (\text{mod}(\theta_d j, 2\pi) - \theta_0)\right) \theta_0 & \leq \text{mod}(\theta_d j, 2\pi) \leq \theta_1 \\ h_{\max 3} \theta_1 & \theta_1 < \text{mod}(\theta_d j, 2\pi) < \theta_3 \\ h_{\max 3} \sin\left(\frac{0.25\pi}{\Delta\theta_3} (\text{mod}(\theta_d j, 2\pi) - \theta_0)\right) \theta_3 & \leq \text{mod}(\theta_d j, 2\pi) \leq \theta_d \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

$\Delta\theta_1$  and  $\Delta\theta_3$  are given by

$$\Delta\theta_1 = \Delta\theta_3 = \begin{cases} \arcsin\left(\frac{0.5B+l \cos \gamma}{D_o}\right) & \text{defect on outer race} \\ \arcsin\left(\frac{0.5B+l \cos \gamma}{D_i}\right) & \text{defect on inner race} \end{cases} \quad (30)$$

$H_4$  is defined by

$$H_4 = \begin{cases} h_{\max 4} \sin(0.25\pi/\Delta\theta_4 (\text{mod}(\theta_d j, 2\pi) - \theta_0)) \theta_0 & \leq \text{mod}(\theta_d j, 2\pi) < \theta_{f1} \\ h_{\max 4} \theta_{f1} & \theta_{f1} \leq \text{mod}(\theta_d j, 2\pi) \leq \theta_{f2} \\ h_{\max 4} \sin(0.25\pi/\Delta\theta_4 (\text{mod}(\theta_d j, 2\pi) - \theta_0)) \theta_{f2} & < \text{mod}(\theta_d j, 2\pi) \leq \theta_d \\ 0 & \text{otherwise} \end{cases} \quad (31)$$

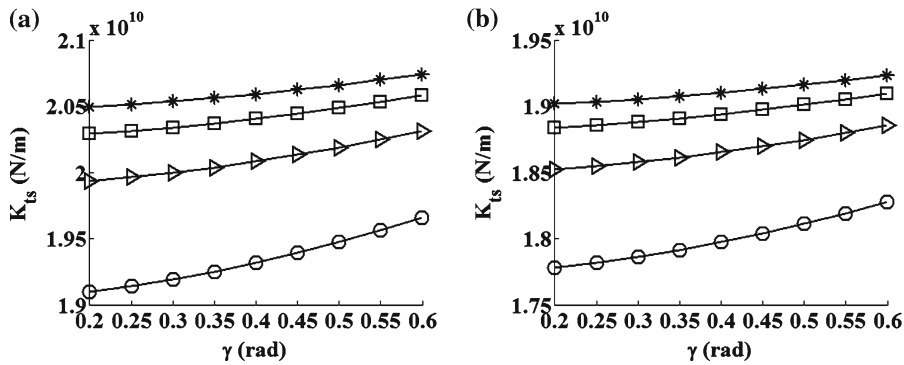
$\Delta\theta_4$  is given by

$$\Delta\theta_4 = \begin{cases} \arcsin\left(\frac{\sqrt{(0.5d)^2 - (0.5d - H + l \sin \gamma)^2} + l \cos \gamma}{D_o}\right) & \text{defect on outer race} \\ \arcsin\left(\frac{\sqrt{(0.5d)^2 - (0.5d - H + l \sin \gamma)^2} + l \cos \gamma}{D_i}\right) & \text{defect on inner race} \end{cases} \quad (32)$$



**Table 2** Size and depth of defects studied in this numerical analysis

Type of the defects	Length ( $L$ ) (mm)	Width ( $B$ ) (mm)	Depth ( $H$ ) (mm)	Ratio of the length to width ( $\xi_d$ )
1	0.10	0.10	0.25	1.00
2	0.10	0.40	0.25	0.25
3	0.40	0.10	0.25	4.00
4	0.60	20.00	0.001	0.03



**Fig. 7** Total contact stiffness between the healthy race and the defective race of the ball bearing: **a** defect on the inner race and **b** defect on the outer race (line with circle one contact surface; line with triangle two contact surfaces; line with square three contact surfaces; and line with asterisk four contact surfaces)

for the numerical investigation in this paper is assumed as the time required for  $0.0092^\circ$  of rotation. For the shaft speed of 2,000rpm, the time step utilized in this numerical solution is  $\Delta t = 2 \times 10^{-6}$  s. The initial displacements used for the inner race and the shaft are  $x_0 = 10^{-6}$  m and  $y_0 = 10^{-6}$  m, and the initial velocities in the  $X$ - and  $Y$ -directions are zero. The sizes and depth of the three different defect cases of interest are listed in Table 2. According to the descriptions in Fig. 3 and the parameters of the ball bearing and the defects in Tables 1 and 2, the value of the elevation angle  $\gamma$  of the small plane surface at the defect edge can be from 0.001 rad to 1.571 rad, which is calculated by the method in Fig. 3.

#### 4.1 Effects of the defect edge topographies on the contact stiffness between the ball and the defect

Figure 7a, b shows the total contact stiffnesses between the healthy outer race and the defective inner race, and those between the healthy inner race and the defective outer race, respectively. As shown in Fig. 7, the total contact stiffnesses between the healthy race and

the defective race of the ball bearing increase with the numbers of contact surfaces between a ball and the edges of the defect. It is also shown that the total contact stiffnesses increase as the elevation angle  $\gamma$  of the small surface at the edges of the defect increases. Note that the contact stiffness between the ball and the races of the bearing varies with the contact position between the ball and the defect when the ball passes over the defect due to the changes in the numbers of the contact surfaces between the ball and the defect edges.

To verify the proposed method, one ball and the healthy outer race, and one ball and the outer race with a plane surface are modeled using the finite element (FE) analysis method. The contact FE model of the ball and the healthy outer race (ball–ball contact as shown in Fig. 1a), and that of the ball and the outer race with a plane surface (ball–plane contact as shown in Fig. 2b, c) are modeled using three-dimensional solid elements and three-dimensional node-to-surface contact elements that include the elastic Coulomb frictional effect according to the FE modeling method in Ref. [24]. The mesh size in the contact area of the FE model is 0.05 mm. The bottom surface of the outer race is fixed. An external radial load is applied at the cen-

**Table 3** Comparison of radial contact deformations between the ball and the healthy race from Hertzian contact theory and the FE analysis method for different radial loads

Radial load (N)	Hertzian contact theory ( $\times 10^{-7}$ m)	FE analysis method ( $\times 10^{-7}$ m)	Difference (%)
62.50	15.59	15.33	1.66
125.00	24.75	24.30	1.80
187.50	32.43	32.19	0.72
250.00	39.28	39.31	0.08
312.50	45.58	46.00	0.92

**Table 4** Comparison of radial contact deformations between the ball and the small surface at the defect edge from Hertzian contact theory and the FE analysis method for different radial loads

Radial load (N)	Hertzian contact theory ( $\times 10^{-7}$ m)	FE analysis method ( $\times 10^{-7}$ m)	Difference (%)
62.50	28.24	27.93	1.10
125.00	44.83	44.20	1.41
187.50	58.75	57.94	1.38
250.00	71.17	70.44	1.03
312.50	82.58	81.53	1.27

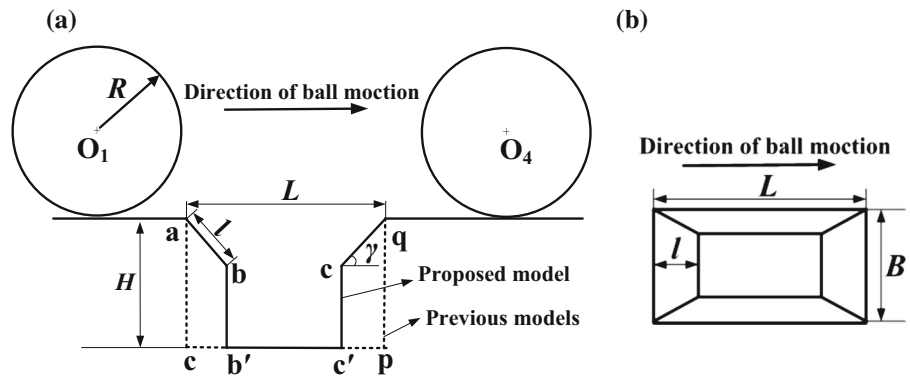
**Table 5** Comparison of radial contact deformations between the ball–ball contact and the ball–plane contact from Hertzian contact theory and the FE analysis method for different radial loads

Radial load (N)	Hertzian contact theory ( $\times 10^{-7}$ m)		FE analysis method ( $\times 10^{-7}$ m)	
	Ball–ball contact, $\delta_{hr}$	Ball–plane contact, $\delta_{dr}$	Ball–ball contact, $\delta_{hr}$	Ball–plane contact, $\delta_{dr}$
62.50	15.59	28.24	15.33	27.93
125.00	24.75	44.83	24.30	44.20
187.50	32.43	58.75	32.19	57.94
250.00	39.28	71.17	39.31	70.44
312.50	45.58	82.58	46.00	81.53

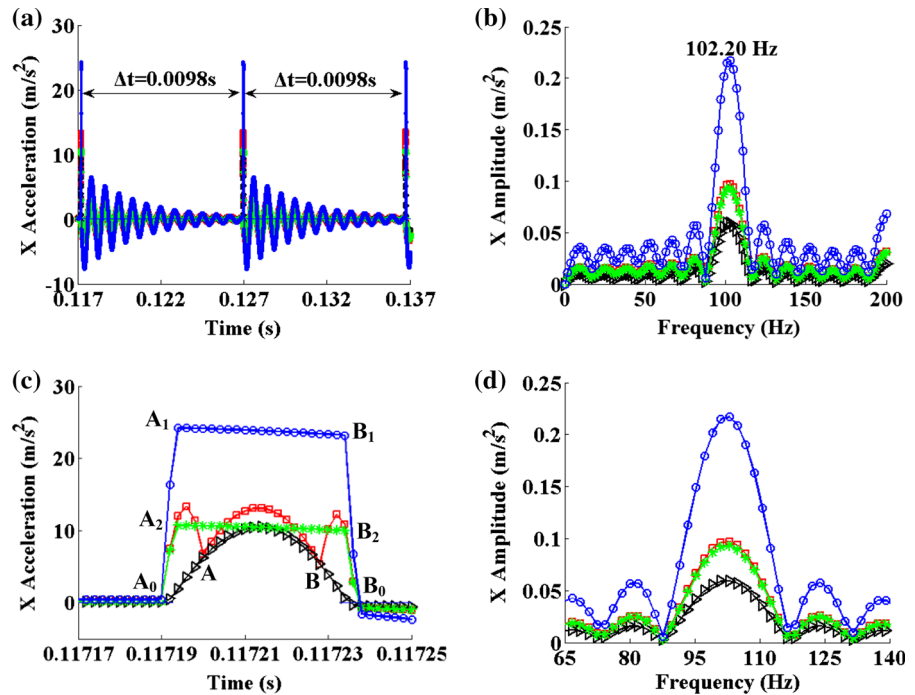
ter of the top surface of the half ball. The displacements of nodes on the top surface of the half ball in the  $Y$ -direction are coupled to apply a uniformly distributed load there. The radial contact deformations from the FE model with the healthy outer race for different radial loads are compared with those from Hertzian contact theory, as shown in Table 3. For the contact FE model of the ball and the healthy outer race (ball–ball contact), the differences between the results from the Hertzian contact theory and the FE analysis method are less than 2%. The radial contact deformations from the FE model with a plane surface for different radial loads are compared with those from Hertzian contact theory, as shown in Table 4. For the contact FE model of the ball and the plane surface (ball–plane contact), the differences between the results from

Hertzian contact theory and the FE analysis method are also less than 2%. Table 5 shows the comparison of radial contact deformations between the ball–ball contact and the ball–plane contact from Hertzian contact theory and the FE analysis method for different radial loads. It is found that the results of the ball–ball contact and those of the ball–plane contact from Hertzian contact theory, and those from the FE analysis method are very different. Table 5 also shows that the radial contact deformations of the ball–ball contact are less than those of the ball–plane contact. Therefore, when the ball first reaches the defect edge, the contact stiffness between the ball and the race will be changed when the ball is in contact with the small surface at the defect edge, which will also cause an impulse.

**Fig. 8** Schematics of the defect size used in vibration response comparisons applying different defect models: **a** front view of the ball bearing and **b** top view of the ball bearing



**Fig. 9** Comparisons of vibration response in the X-direction of the ball bearing with a defect on its outer race using different defect models: **a** acceleration response from the four different models from 0.117 to 0.137 s, **b** frequency spectra of acceleration response from the four different models, **c** closed-up plot of **a**, and **d** closed-up plot of **b**. (line with circle Rafsanjani et al.'s model; line with asterisk Patel et al.'s model; line with inverse triangle time-varying deflection model; line with square the proposed model)



## 4.2 Effects of defect edge topographies on vibration response of the ball bearing

### 4.2.1 Vibration response comparisons applying different defect models

To verify the proposed model, the time- and frequency-domain vibration response of the inner race of the ball bearing from different defect models is compared. The first defect case on the surface of the outer race of the bearing is investigated. The sizes of the first defect case are given in Table 2. The initial angular position of the defect is assumed to be zero degree. The elevation angle  $\gamma$  and the length  $l$  of the small surfaces at defect edge

are assumed to be 0.4 rad and 0.02 mm. The different defect models include Rafsanjani et al.'s model [2], Patel et al.'s model [18], time-varying deflection excitation model [23], and the proposed model considering both the time-varying deflection excitation and the time-varying contact stiffness excitation. For these four different defect models, the surface areas are assumed to be same, which means the length  $L$  and width  $B$  of the defect are same, as shown in Fig. 8.

Figure 9a–d shows the time- and frequency-domain acceleration response in the X-direction from Rafsanjani et al.'s model, Patel et al.'s model, time-varying deflection excitation model, and the proposed model. Note that the time delay between the impulses caused

by the defect is same for the four different models, which is equal to 0.0098 s as shown in Fig. 9a. However, there are significant differences between the waveforms of the acceleration response from the four different models.

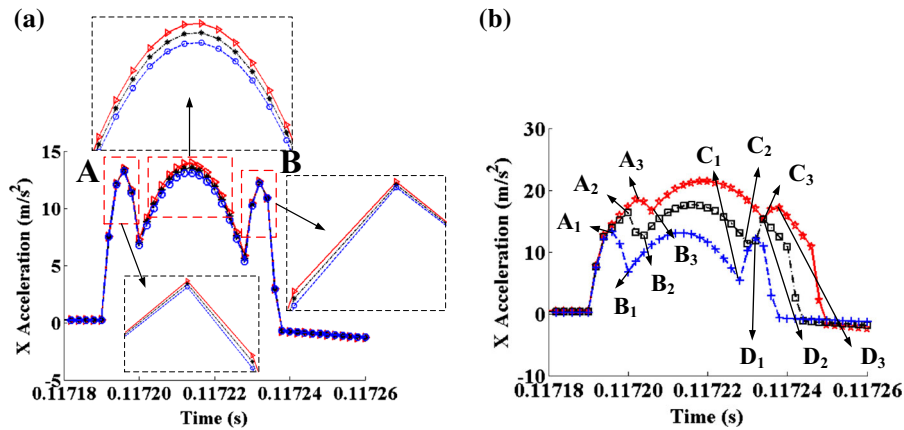
As shown in Fig. 9c, Rafsanjani et al.'s model only describes the time-invariant impulse stage  $A_0A_1B_1B_0$ , since their model used a rectangular function to model the time-invariant deflection excitation caused by the changes in the contact positions between the ball and the defect when the ball passes over the defect as shown in Fig. 1a. Patel et al.'s model also only describes the time-invariant impulse stage  $A_0A_2B_2B_0$ , whose impulse waveform is similar to that from Rafsanjani et al.'s model because they also used a rectangular function to model the time-invariant deflection excitation. However, its amplitude is less than that of Rafsanjani et al.'s model, because Rafsanjani et al.'s model used the defect depth to determine the amplitude of the rectangular function, but that of Patel et al.'s model is determined by the geometry relationship between the ball and the defect. Moreover, the time-varying deflection model can describe the time-varying impulse stage  $A_0B_0$ , which used a half-sine function to model the time-varying deflection excitation caused by the changes in the contact positions between the ball and the defect when the ball passes over the defect. According to the descriptions in Sect. 2, when the ball passes over the defect, the impulse caused by the defect is time-varying; the amplitude of the impulse function should be determined by the contact relationship between the ball and the defect; and the impulse excitation includes both the time-varying deflection excitation and the time-varying contact stiffness excitation caused by the changes in the numbers of the contact surfaces between the ball and the defect. However, Rafsanjani et al.'s model, Patel et al.'s model, and the time-varying deflection model only modeled the time-invariant deflection excitation and the time-varying deflection excitation caused by the defect, which cannot describe the impulse caused by the time-varying contact stiffness due to the changes in the numbers of the contact surfaces between the ball and the defect. To overcome this problem, the proposed model considering both the time-varying deflection excitation and the time-varying contact stiffness excitation is proposed in this paper, which is modeled by a piecewise function including the rectangular function and the half-sine function. For the proposed model, the amplitude of the time-varying deflec-

tion excitation is determined by the geometry relationship between the ball and the defect, and the value of the time-varying contact stiffness is determined by Hertzian contact theory. As shown in Fig. 9c, the proposed model can describe not only the time-varying impulse stage  $A_0ABB_0$ , but also the impulse stages  $A_0A$  and  $BB_0$  caused by the time-varying contact stiffnesses due the changes in the numbers of the contact surfaces (from 1 surface to 3 surfaces, and from 3 surfaces to 1 surface) between the ball and the defect edges, as described in Sect. 2.

As shown in Fig. 9c, note that the amplitudes of the acceleration response from Rafsanjani et al.'s model and the proposed model are larger than those from Patel et al.'s model and the time-varying deflection excitation model. Comparisons of the acceleration response from Rafsanjani et al.'s model and the proposed model show that Rafsanjani et al.'s model can only describe the time-invariant impulse caused by the defect, but the proposed model can describe the time-varying impulse and effects of the defect edge topographies which are more close to the real as the analysis in Sect. 2. Comparisons between Patel et al.'s model and the time-varying deflection excitation model show Patel et al.'s model can formulate geometry relationship between the ball and the defect, but it cannot model the time-varying impulse generated by the defect when the ball passes through the defect. The time-varying deflection excitation model can consider effects of both the geometry relationship between the ball and the defect and the deflection impulse, but it cannot consider the time-varying contact stiffness between the ball and the defect.

In addition, comparisons of the frequency spectra of the acceleration response of the inner race of the bearing in the  $X$ -direction from the four defect models are shown in Fig. 9b, d. Note that the same peak frequencies at 102.20 Hz appear in the frequency spectra from the four models. However, the amplitudes of the peaks are different for the four different models. The amplitude of the frequency spectrum of Rafsanjani et al.'s model is larger than those of the other three models. The amplitude of the proposed model is larger than those of Patel et al.'s model and the time-varying deflection excitation model. The amplitude of the time-varying deflection excitation model is smaller than that of Patel et al.'s model. This occurs because the additional deflection caused by the defect is assumed as the depth of the defect in the Rafsanjani et al.'s model,





**Fig. 10** Effects of defect edge topographies on the acceleration response in the  $X$ -direction of the ball bearing with a first type defect on its outer race using the proposed models from 0.11718 to 0.11726 s: **a** effect of the elevation angle  $\gamma$  and **b** effect of the length  $l$  (line with circle  $\gamma = 0.40$  rad and  $l = 0.2L$ ; line with

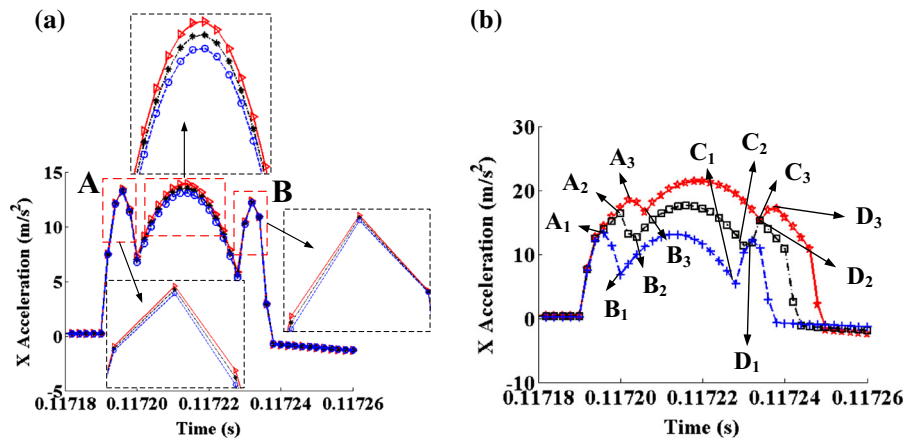
asterisk  $\gamma = 0.44$  rad and  $l = 0.2L$ ; line with inverse triangle  $\gamma = 0.48$  rad and  $l = 0.2L$ ; line with plus symbol  $\gamma = 0.40$  rad and  $l = 0.2L$ ; line with square  $\gamma = 0.40$  rad and  $l = 0.3L$ ; line with star  $\gamma = 0.40$  rad and  $l = 0.4L$ )

and both the time-varying deflection excitation and the time-varying contact stiffness excitation are considered in the proposed model, when the ball passes over the defect. However, in Patel et al.'s model and the time-varying deflection excitation model, they only consider the effect of the additional deflection determined by the defect size and ball diameter, which is smaller than those in Rafsanjani et al.'s model and the proposed model. These results show that the proposed model is validation and more accurately for the ball bearing dynamics simulation according to the analysis in Sect. 2. The proposed model can not only describe the time-varying deflection excitation generated by the defect, but also model the time-varying contact stiffness between the ball and the defect with different edge topographies when the ball passes over the defect.

4.2.2 Effects of the defect edge topographies on the waveforms of the acceleration response of the ball bearing

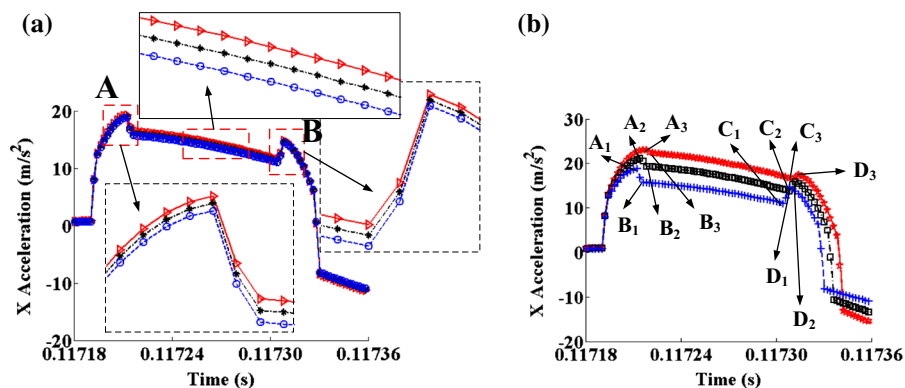
Figures 10, 11, and 12 show effects of the defect edge topographies on the waveforms of the acceleration response in the  $X$ -direction of the inner race of the ball bearing for the defect cases 1, 2, and 3, respectively. Here, to compare the waveforms of the acceleration responses of the ball bearing caused by different defect, the initial position of the defect is assumed to be zero. Note that the impulses at the points A, B,

$A_1, A_2, A_3, D_1, D_2,$  and  $D_3$  caused by the changes in the contact stiffnesses between the ball and the defect edges are also observed in Figs. 10, 11 and 12 for the three studied defect cases, whose reasons are similar with the descriptions in Sect. 4.2.1. Effects of the elevation angle  $\gamma$  of the small surface at the defect edge on the acceleration response of the ball bearing for cases 1, 2, and 3 are shown in Figs. 10a, 11a, and 12a, respectively. The amplitude of the acceleration response increases consistently with the elevation angle  $\gamma$  for the three different cases. This occurs because the amplitudes of the time-varying deflection excitations and the time-varying contact stiffness excitations caused by the three defect cases increase with the elevation angle  $\gamma$  as shown in Eqs. (5), (21) to (23), and Fig. 7, when the elevation angle  $l$  is a constant value. Figures 10b, 11b, and 12b plot effects of length  $l$  of the small surface at the defect edge on the acceleration response of the three studied defect cases. The amplitude of the acceleration response also increases consistently with the length  $l$  for the three different cases. This occurs because the amplitudes of the time-varying deflection excitations caused by the three defect cases increase consistently with the length  $l$  as shown in Eqs. (21) and (23) when the elevation angle  $\gamma$  is a constant value. Moreover, as shown in Figs. 10b, 11b, and 12b, the differences of the impulses between points  $A_1$  and  $B_1$ , that between points  $A_2$  and  $B_2$ , and that between points  $A_3$  and  $B_3$  caused by the changes in the con-



**Fig. 11** Effects of defect edge topographies on the acceleration response in the  $X$ -direction of the ball bearing with a second defect on its outer race using the proposed models from 0.11718 to 0.11726 s: **a** effect of the elevation angle  $\gamma$  and **b** effect of the length  $l$ . (line with circle  $\gamma = 0.40$  rad and  $l = 0.2L$ ; line with

asterisk  $\gamma = 0.44$  rad and  $l = 0.2L$ ; line with inverse triangle  $\gamma = 0.48$  rad and  $l = 0.2L$ ; line with plus symbol  $\gamma = 0.40$  rad and  $l = 0.2L$ ; line with square  $\gamma = 0.40$  rad and  $l = 0.3L$ ; line with star  $\gamma = 0.40$  rad and  $l = 0.4L$ )



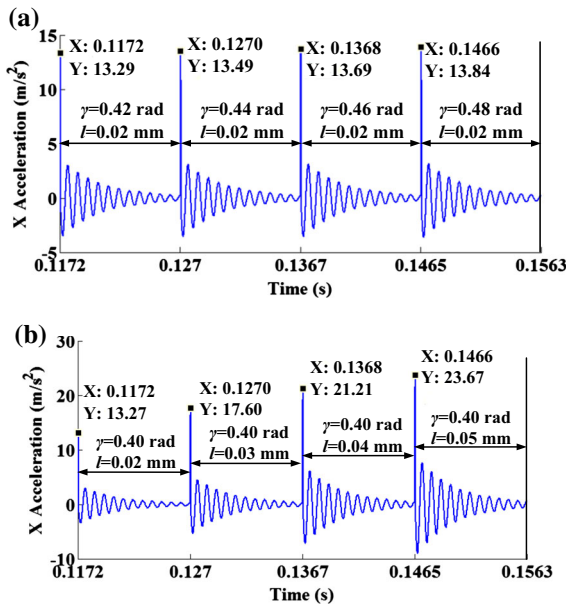
**Fig. 12** Effects of defect edge topographies on the acceleration response in the  $X$ -direction of the ball bearing with a third defect on its outer race using the proposed models from 0.11718 to 0.11736 s: **a** effect of the elevation angle  $\gamma$  and **b** effect of the length  $l$ . (line with circle  $\gamma = 0.40$  rad and  $l = 0.2L$ ; line with

asterisk  $\gamma = 0.44$  rad and  $l = 0.2L$ ; line with inverse triangle  $\gamma = 0.48$  rad and  $l = 0.2L$ ; line with plus symbol  $\gamma = 0.40$  rad and  $l = 0.2L$ ; line with square  $\gamma = 0.40$  rad and  $l = 0.3L$ ; line with star  $\gamma = 0.40$  rad and  $l = 0.4L$ )

tact stiffnesses between the ball and the defect edges decrease as the length  $l$  increases for the defect cases 1, 2, and 3. This occurs because the duration of the impulse increases with the arc length of the propagated surface at the defect edges in the tangential direction  $\Delta\theta$  when the length  $l$  increases as shown in Figs. 1 and 2.

Figures 13 and 15 show effects of time-varying defect edge topographies on waveforms of the acceleration response in the  $X$ -direction of the inner race of

the ball bearing for the defect cases 2, 3, and 4, respectively. In this paper, to describe effects of time-varying defect edge topographies on vibrations of the ball bearing with a localized surface defect on its races, the rates of change of the elevation angle  $\gamma$  and the length  $l$  of the small surface at the defect edges are assumed to be constant since simulating the rates of the defect edge topographies due to the ball impacts is very complex and beyond the scope of this study. Note that the amplitude of the acceleration response increases consistently

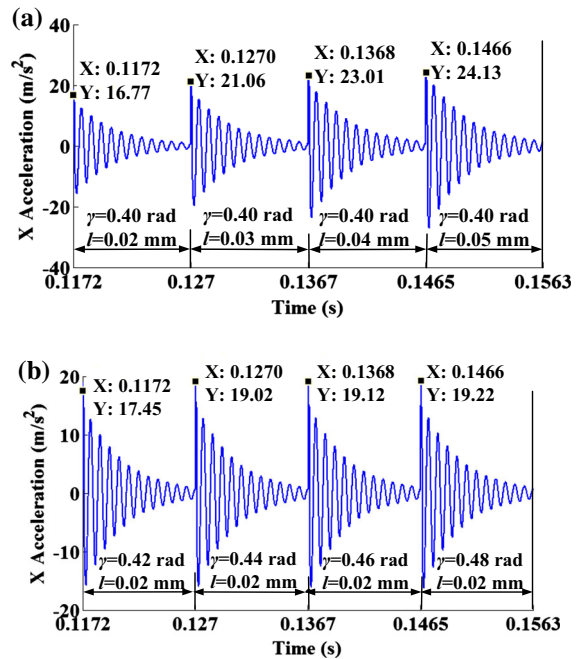


**Fig. 13** Effects of defect edge topographies on the acceleration response in the X-direction of the ball bearing with a first defect on its outer race using the proposed models from 0.1172 to 0.1563 s: **a** effect of the elevation angle  $\gamma$  and **b** effect of the length  $l$

with the elevation angle  $\gamma$  and the length  $l$  of the small surface at the defect edge due to the impacts between the ball and the defect edges for the studied cases. This is because the amplitudes of the time-varying deflection excitation and the time-varying contact stiffness excitation increase with the elevation angle  $\gamma$  and the length  $l$  of the small surface at the defect edge as analyzed in Eqs. (5), (21) to (23), and Fig. 7. The results from the proposed model considering the time-varying deflection excitation and the time-varying contact stiffness excitation caused by the localized defect with different edge topographies cannot be predicted by the previous localized surface defect models in the literature (Fig. 14).

#### 4.2.3 Effects of the defect edge topographies on the statistical measures of the acceleration response of the ball bearing

Statistical measures which can determine the impulsive character of a time-domain signal are widely used in bearing fault diagnosis and detection [7,8,40–42]. Here, three different well-known statistical indicators such as root mean square (RMS), crest factor (CF), and kurtosis are applied to investigate effects of the defect



**Fig. 14** Effects of defect edge topographies on the acceleration response in the X-direction of the ball bearing with a third defect on its outer race using the proposed models from 0.1172 to 0.1563 s: **a** effect of the elevation angle  $\gamma$  and **b** effect of the length  $l$

edge topographies on the acceleration response in the X-direction of the inner race of the ball bearing. These parameters of a discrete signal  $s$  are given by [8]

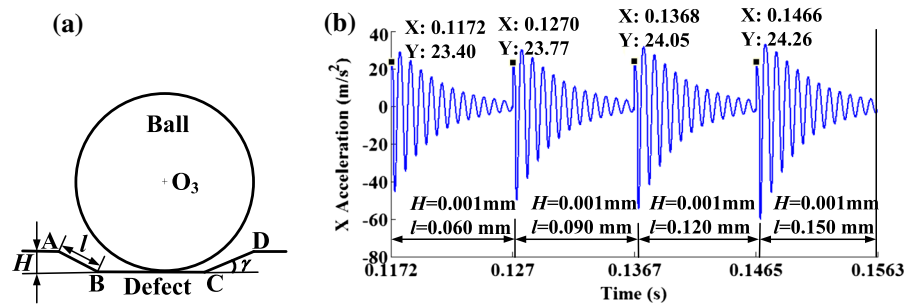
$$\text{RMS} = \sqrt{\frac{1}{N_f} \sum_{i=1}^{N_f} s_i^2} \tag{37}$$

$$\text{CF} = \frac{\max(s) - \min(s)}{\text{RMS}} \tag{38}$$

$$\text{Kurtosis} = \frac{\sum_{i=1}^{N_f} (s_i - \text{mean}(s))^4}{N_f \text{RMS}^4} \tag{39}$$

where  $N_f$  is the number of samples,  $s_i$  is the  $i$ th data of the discrete signal  $s$ ,  $\max()$  is the maximum value of the discrete signal  $s$ ,  $\min()$  is the minimum value of the discrete signal  $s$ , and  $\text{mean}()$  is the mean value of the discrete signal  $s$ .

Figures 16, 17, and 18 show the statistical measures of the acceleration response in the X-direction of the inner race of the ball bearing from 0.3 to 0.4 s for the defect cases 1, 2, and 3. According to the parameters in Table 2 and the descriptions in Fig. 3, the range of the elevation angle  $\gamma$  can be chosen to be from 0.4 to 0.5 rad in this section. The RMS of the X-direction acceleration response for cases 1, 2, and 3 are shown



**Fig. 15** Effects of defect edge topographies on the acceleration response in the  $X$ -direction of the ball bearing with a fourth defect on its outer race using the proposed models from 0.1172 to 0.1563 s: **a** a schematic of the fourth defect and **b** effect of the length  $l$

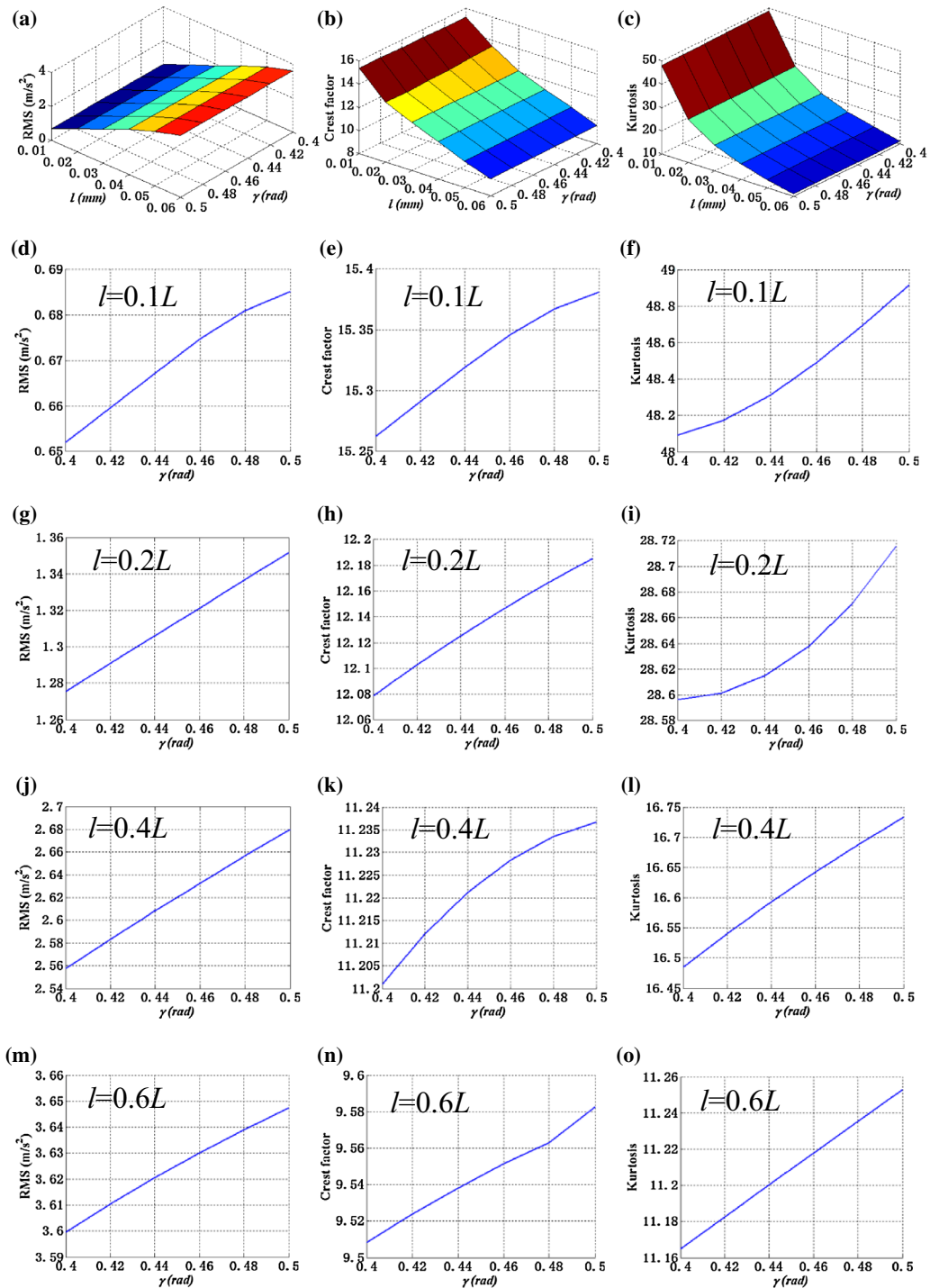
in Figs. 16a, d, g, j, m, 17a, d, g, j, m, and 18a, d, g, j, m, respectively. The RMS increases consistently with the length  $l$  and the elevation angle  $\gamma$  for the three different cases. This occurs because the amplitude of the acceleration response of the ball bearing increases with the elevation angle  $\gamma$  and the length  $l$  as shown in Figs. 10, 11, 12, 13, and 14. Figures 16b, e, h, k, n, 17b, e, h, k, n, and 18b, e, h, k, n give the crest factor of the  $X$ -direction acceleration response of the three studied defect cases. For the three defect cases, the crest factor decreases with the length  $l$  because the duration of the impulse waveform increase with the length  $l$  as shown in Figs. 10, 11, and 12. It increases as the elevation angle  $\gamma$  increases because the duration of the impulse waveform decreases with the elevation angle  $\gamma$ , and the amplitude of the impulse waveform is affected little by the changes of the elevation angle  $\gamma$ , as shown in Figs. 10, 11, 12, 13, and 14, when the length  $l$  is a constant value for the defect cases 1 and 2. For the defect case 3, it decreases as the elevation angle  $\gamma$  increases because of the differences between amplitude of the acceleration response at the contact position at the beginning and ending edges and that at the other locations of the defect as shown in Fig. 12a. The kurtosis of the  $X$ -direction acceleration response of the three selected defect cases is shown in Figs. 16c, f, i, l, o, 17c, f, i, l, o, and 18c, f, i, l, o, respectively. The kurtosis also decreases with the length  $l$  for the three defect cases. Moreover, the kurtosis also increases as the elevation angle  $\gamma$  increases for the defect cases 1 and 2; it decreases with the elevation angle  $\gamma$  for the defect case 3. This occurs because the similar reasons with the crest factors for the three defect cases. The above results show that the vibration level of the ball bearing increases as the defect sizes for the three studied defect cases. The crest factor of the acceleration

response will decrease when the defect edges progress along the direction of the ball motion for the defect cases 1, 2, and 3. When the defect edges progress along the direction of the ball motion, the kurtosis of the acceleration response of the ball bearing also decreases for defect cases 1, 2, and 3.

## 5 Conclusions

A new ball bearing dynamic model coupled with effects of defect edge topographies is developed to investigate effects of defect edge topographies on the vibration response of a ball bearing with a localized surface defect on its races. According to the shape and sizes of the defect, a new contact model for determining the contact relationships between the ball and the defect edges is proposed based on Hertzian elastic contact theory. The proposed model is applied to investigate effects of defect edge topographies on the contact stiffnesses between the ball and the defect edges, and the vibration response of the ball bearing. The following specific conclusions can be drawn:

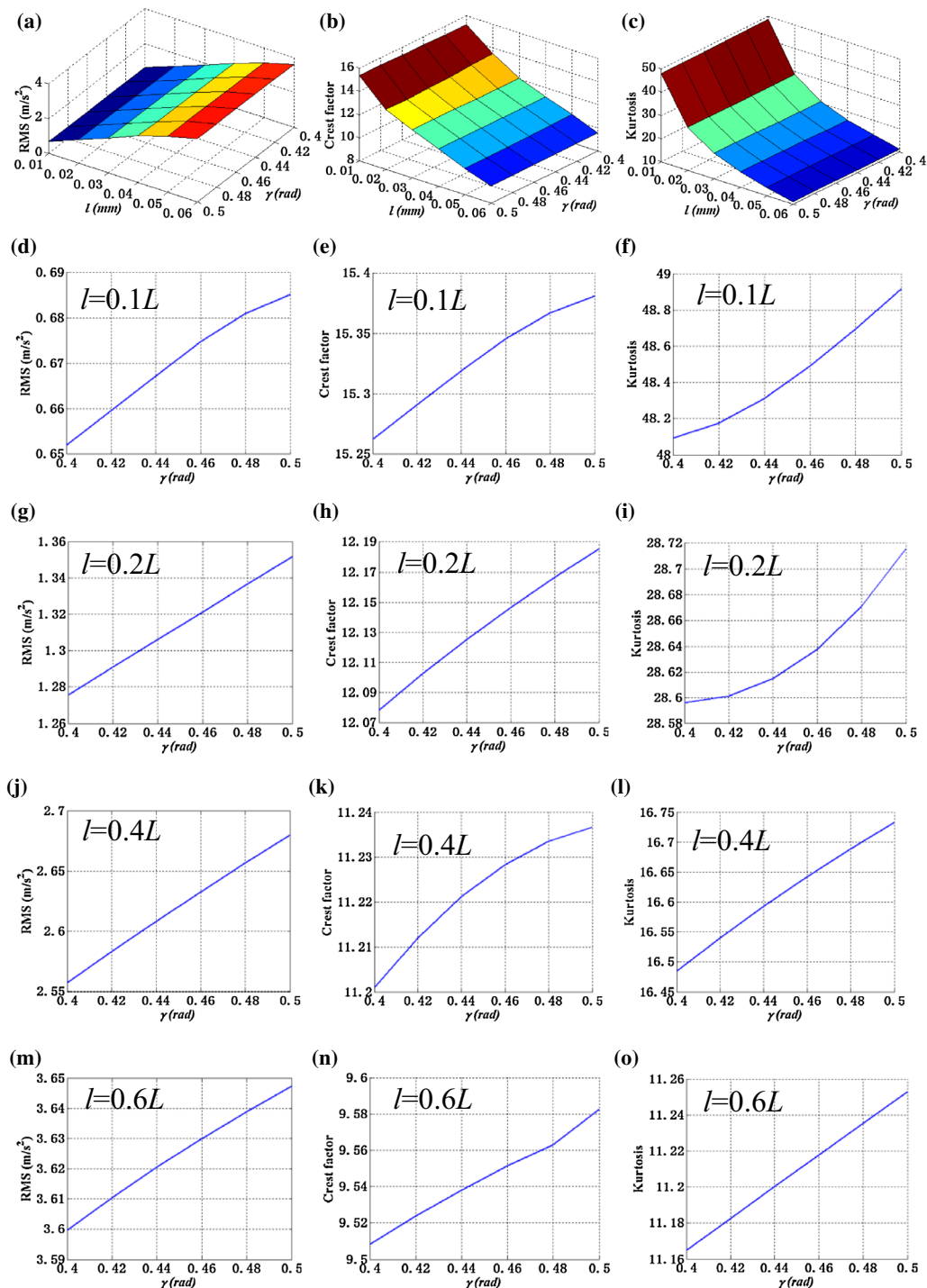
1. The contact stiffness between the ball and the races of the ball bearing varies with the contact positions between the ball and the defect edges when the ball passes over the defect due to changes in the numbers of the contact surfaces between the ball and the defect edges. The total contact stiffness between the healthy race and the defective race of the ball bearing increases with the numbers of contact surfaces between a ball and the defect edges. The total contact stiffness also increases as the elevation angle  $\gamma$  increases.
2. The different defect models have significant effect on the pulse waveform of the vibration response



**Fig. 16** Statistical measures of the acceleration response of the inner race of the bearing in the  $X$ -direction for defect case 1: **a** RMS, **b** crest factor, **c** kurtosis, **d** RMS for  $l = 0.1L$ , **e** crest factor for  $l = 0.1L$ , **f** kurtosis for  $l = 0.1L$ , **g** RMS for  $l = 0.2L$ ,

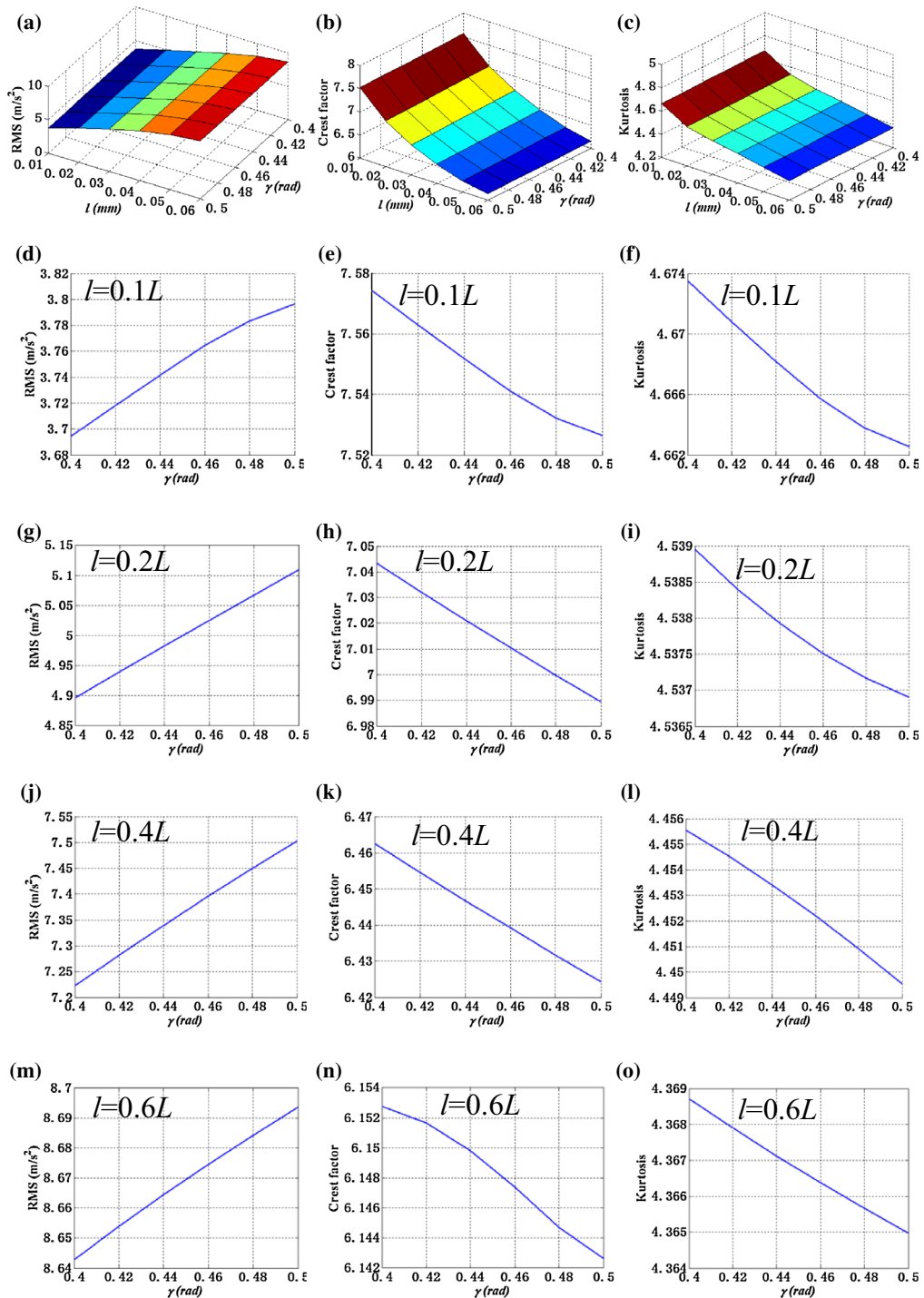
**h** crest factor for  $l = 0.2L$ , **i** kurtosis for  $l = 0.2L$ , **j** RMS for  $l = 0.4L$ , **k** crest factor for  $l = 0.4L$ , **l** kurtosis for  $l = 0.4L$ , **m** RMS for  $l = 0.6L$ , **n** crest factor for  $l = 0.6L$ , and **o** kurtosis for  $l = 0.6L$





**Fig. 17** Statistical measures of the acceleration response of the inner race of the bearing in the X-direction for defect case 2: **a** RMS, **b** crest factor, **c** kurtosis, **d** RMS for  $l = 0.1L$ , **e** crest factor for  $l = 0.1L$ , and **f** kurtosis for  $l = 0.1L$ , **g** RMS for  $l = 0.2L$ , **h** crest factor for  $l = 0.2L$ , **i** kurtosis for  $l = 0.2L$ , **j** RMS for  $l = 0.4L$ , **k** crest factor for  $l = 0.4L$ , **l** kurtosis for  $l = 0.4L$ , **m** RMS for  $l = 0.6L$ , **n** crest factor for  $l = 0.6L$ , and **o** kurtosis for  $l = 0.6L$





**Fig. 18** Statistical measures of the acceleration response of the inner race of the bearing in the X-direction for defect case 3: **a** RMS, **b** crest factor, **c** kurtosis, **d** RMS for  $l = 0.1L$ , **e** crest factor for  $l = 0.1L$ , and **f** kurtosis for  $l = 0.1L$ , **g** RMS for

$l = 0.2L$ , **h** crest factor for  $l = 0.2L$ , **i** kurtosis for  $l = 0.2L$ , **j** RMS for  $l = 0.4L$ , **k** crest factor for  $l = 0.4L$ , **l** kurtosis for  $l = 0.4L$ , **m** RMS for  $l = 0.6L$ , **n** crest factor for  $l = 0.6L$ , and **o** kurtosis for  $l = 0.6L$

of the ball bearing caused by a localized surface defect. Therefore, it is necessary to formulate the localized defect model including effects of the length–width ratio of the defect, the time-varying deflection and the time-varying contact stiffness excitation produced by the defect, and effects of defect edge topographies.

3. The impulses caused by changes in the time-varying contact stiffnesses between the ball and the defect edges with different topographies can be described by the proposed model, which cannot be described by the previous deflection excitation models in the literature. The amplitude of the impulse increases with the elevation angle  $\gamma$  and the length  $l$  of the small surface at the defect edge. Moreover, the proposed method in this paper can be extended to formulate effects of the random defect edge topographies on the vibration response of the ball bearing with a localized surface defect on its races.
4. The RMS of the acceleration response of the ball bearing increases consistently with the defect sizes for defect cases 1, 2, and 3. The crest factor and the kurtosis of the acceleration response of the ball bearing will decrease when the defect edges progress along the direction of the ball motion for the defect cases 1, 2, and 3.

**Acknowledgments** The authors are grateful for the financial support provided by the National Natural Science Key Foundation of China under Contract No. 51035008 and the Fundamental Research Funds for the Central Universities under Grant Number 0903005203236.

## References

1. Tandon, N., Choudhury, A.: A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribol. Int.* **32**, 469–480 (1999)
2. Rafsanjani, A., Abbasion, S., Farshidianfa, A., Moeenfard, H.: Nonlinear dynamic modeling of surface defects in rolling element bearing systems. *J. Sound Vib.* **319**, 1150–1174 (2009)
3. McFadden, P.D., Smith, J.D.: Model for the vibration produced by a single point defect in a rolling element bearing. *J. Sound Vib.* **96**(1), 69–82 (1984)
4. McFadden, P.D., Smith, J.D.: The vibration produced by multiple point defects in a rolling element bearing. *J. Sound Vib.* **98**(2), 263–273 (1985)
5. Tandon, N., Choudhury, A.: An analytical model for the prediction of the vibration response of rolling element bearing due to a localized defect. *J. Sound Vib.* **205**(3), 275–292 (1997)
6. Choudhury, A., Tandon, N.: Vibration response of rolling element bearings in a rotor bearing system to a local defect under radial load. *ASME J. Tribol.* **128**, 252–261 (2006)
7. Kiral, Z., Karagulle, H.: Simulation and analysis of vibration signals generated by rolling element bearing with defects. *Tribol. Int.* **36**, 667–678 (2003)
8. Kiral, Z., Karagulle, H.: Vibration analysis of rolling element bearings with various defects under the action of an unbalanced force. *Mech. Syst. Signal Process.* **20**, 1967–1991 (2006)
9. Sassi, S., Badri, B., Thomas, M.: A numerical model to predict damaged bearing vibrations. *J. Vib. Control* **13**(11), 1603–1628 (2007)
10. Ashtekar, A., Sadeghi, F., Stacke, L.E.: A new approach to modeling surface defects in bearing dynamics simulations. *ASME J. Tribol.* **130**, 041103 (2008)
11. Ashtekar, A., Sadeghi, F., Stacke, L.E.: Surface defects effects on bearing dynamics. *Proc. Inst. Mech. Eng. Part J J. Eng. Tribol.* **224**, 25–35 (2010)
12. Behzad, M., Bastami, A.R., Mba, D.: A new model for estimating vibrations generated in the defective rolling element bearings. *ASME J. Vib. Acoust.* **133**, 041101 (2011)
13. Sopianen, J., Mikola, A.: Dynamic model of a deep-groove ball bearing including localized and distributed defects-part 1: theory. *Proc. Inst. Mech. Eng. Part K J. Multi-body Dyn.* **217**, 201–211 (2003)
14. Sopianen, J., Mikola, A.: Dynamic model of a deep-groove ball bearing including localized and distributed defects-part 2: implementation and results. *Proc. Inst. Mech. Eng. Part K J. Multi-body Dyn.* **217**, 213–223 (2003)
15. Cao, M., Xiao, J.: A comprehensive dynamic model of double-row spherical roller bearing-modeling development and case studies on surface defects, preloads, and radial clearance. *Mech. Syst. Signal Process.* **22**, 467–489 (2007)
16. Patil, M.S., Mathew, J., Rajendrakumar, P.K., Desai, S.: A theoretical model to predict the effect of the localized defect on vibrations associated with ball bearing. *Int. J. Mech. Sci.* **52**, 1193–1201 (2010)
17. Arslan, H., Aktürk, N.: An investigation of rolling element vibrations caused by local defects. *ASME J. Tribol.* **130**, 041101 (2008)
18. Patel, V.N., Tandon, N., Pandey, R.K.: A dynamic model for vibration studies of deep groove ball bearings considering single and multiple defects in races. *ASME J. Tribol.* **132**, 041101 (2010)
19. Nakhaeinejad, M., Bryant, M.D.: Dynamic modeling of rolling element bearings with surface contact defects using bond graphs. *ASME J. Tribol.* **133**, 011102 (2011)
20. Ghafari, S.H., Golnaraghi, F., Ismail, F.: Effect of localized faults on chaotic vibration of rolling element bearings. *Nonlinear Dyn.* **53**(4), 287–301 (2008)
21. Kankar, P.K., Sharma, S.C., Harsha, S.P.: Vibration based performance prediction of ball bearings caused by localized defects. *Nonlinear Dyn.* **69**, 847–875 (2012)
22. Liu, J., Shao, Y.M., Ming, J.Z.: The effects of the shape of localized defect in ball bearings on the vibration waveform. *Proc. Inst. Mech. Eng. Part K J. Multi-body Dyn.* **227**(3), 261–274 (2013)
23. Liu, J., Shao, Y.M., Lim, T.C.: Vibration analysis of ball bearings with a localized defect applying piecewise response function. *Mech. Mach. Theory* **56**, 156–169 (2012)

24. Shao, Y.M., Liu, J., Ye, J.: A new method to model a localized surface defect in a cylindrical roller bearing dynamic simulation. *Proc. Inst. Mech. Eng. Part J J. Eng. Tribol.* **228**(2), 140–159 (2014)
25. Branch, N.A., Arakere, N.K., Svendsen, V., Forster, N.H.: Stress field evolution in a ball bearing raceway fatigue spall. *J. ASTM Int.* **7**(2), 1–18 (2010)
26. Lundberg, G., Palmgren, A.: Dynamic capacity of rolling bearings. *Acta Polytech. Mech. Eng. Ser. I R. Swed. Acad. Eng. Sci.* **1**(3), 5–50 (1947)
27. Kotzalas, M., Harris, T.A.: Fatigue failure progression in ball bearing. *ASME J. Tribol.* **123**(2), 238–242 (2000)
28. Ioannides, E., Harris, T.A.: A new fatigue life model for rolling bearings. *ASME J. Tribol.* **107**(3), 367–377 (1985)
29. Xu, G., Sadeghi, F.: Spall initiation and propagation due to debris denting. *Wear* **201**(1–2), 106–116 (1996)
30. Hoepfich, M.R.: Rolling element bearing fatigue damage propagation. *ASME J. Tribol.* **114**(2), 328–333 (1992)
31. Rosado, L., Forster, N., Thomson, K.: On the rolling contact fatigue life and spall propagation characteristics of M50, M50 NiL and 52100 bearing materials: part I-experimental results. *STLE Tribol. Trans.* **53**(1), 29–41 (2009)
32. Arakere, N.K., Branch, N., Levesque, G., Svendsen, V., Forster, N.H.: On the rolling contact fatigue life and spall propagation characteristics of M50, M50 NiL and 52100 bearing materials: part III-stress modeling. *STLE Tribol. Trans.* **53**(1), 42–51 (2009)
33. Forster, N.H., Ogden, W.P., Trivedi, H.K.: On the rolling contact fatigue life and spall propagation characteristics of M50, M50 NiL and 52100 bearing materials: part III-metallurgical examination. *STLE Tribol. Trans.* **53**(1), 52–59 (2009)
34. Branch, N.A., Arakere, N.K., Svendsen, V., Forster, N.H.: Stress field evolution in a ball bearing raceway fatigue spall. *J. ASTM Int.* **7**(2), 1–18 (2010)
35. Branch, N.A., Arakere, N.K., Forster, N., Svendsen, V.: Critical stresses and strains at the spall edge of a case hardened bearing due to ball impact. *Int. J. Fatigue* **47**, 268–278 (2013)
36. Li, Y., Billington, S., Zhang, C., Kurfess, T., Danyluk, S., Liang, S.: Dynamic prognostic prediction of defect propagation on rolling element bearings. *Tribol. Trans.* **42**(2), 385–392 (1999)
37. Harris, T. A., Kotzalas, M. N.: *Rolling Bearing Analysis-Essential Concepts of Bearing Technology*, 5th ed. Taylor and Francis, London (2007)
38. Xu, B.Y.: *Elastic and plastic mechanics*. China Machine Press, Beijing (2003)
39. Sunnersjo, C.S.: Varying compliance vibrations of rolling bearing. *J. Sound Vib.* **58**(3), 363–373 (1978)
40. Pachaud, C.: Crest factor and kurtosis contributions to identify defects inducing periodical impulsive forces. *Mech. Syst. Signal Process.* **11**(6), 903–916 (1997)
41. Dron, J.P., Bolaers, F., Rasolofondraibe, I.: Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction: application to the detection of defects in ball bearings. *J. Sound Vib.* **270**(1–2), 61–73 (2004)
42. Bolaers, F., Cousinard, O., Marconnet, P., Rasolofondraibe, L.: Advanced detection of rolling bearing spalling from denoising vibratory signals. *Control Eng. Pract.* **12**, 181–190 (2004)