International Journal of Fatigue 168 (2023) 107449

Contents lists available at ScienceDirect

International Journal of Fatigue

journal homepage: www.elsevier.com/locate/ijfatigue

Prediction of rolling contact fatigue crack propagation in bearing steels using experimental crack growth data and linear elastic fracture mechanics

Björn Kunzelmann [a,*], Pawel Rycerz [a], Yilun Xu [b], Nagaraj K. Arakere [c], Amir Kadiric [a]

[a] *Department of Mechanical Engineering, Imperial College London, London SW7 2AZ, UK*
[b] *Department of Materials, Imperial College London, London SW7 2AZ, UK*
[c] *Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611-6250, USA*

ARTICLE INFO

Keywords:
Rolling contact fatigue
Pitting
Crack propagation
Rolling bearings
Gears

ABSTRACT

Rolling contact fatigue (RCF) is a major life limiting factor for machine elements that employ non-conformal, rolling sliding, lubricated contacts such as rolling bearings and gears. This paper explores the application of linear elastic fracture mechanics (LEFM) principles, as commonly used in structural fatigue, for prediction of RCF crack propagation. A triple-disc contact fatigue machine is used to generate RCF cracks of varying lengths in AISI 52100 bearing steel roller specimens. Crack propagation rates across the surface are measured using optical inspection of test specimens and the final crack geometry is established through specimen sectioning. A numerical finite element model of surface breaking RCF cracks based on LEFM methods is devised to predict the evolution of stress intensity factors (SIFs) during over-rolling of the contact over the experimentally observed crack geometries. The model employs a suitable fracture mechanics mesh to resolve stresses at the crack tip and accounts for Hertzian contact stresses, contact friction and crack face friction. Potential effects of lubricant pressurisation within the crack are not modelled. The predicted SIFs are then related to the experimentally measured crack propagation rates to establish the applicability of the LEFM principles to RCF crack propagation. Results show that LEFM can be used to predict the growth of surface braking RCF cracks. For cracks longer than about 100 μm, a Paris law relationship with the stress intensity exponent of about 4 is derived. Mode II was seen to be the dominant mode of propagation for surface braking RCF cracks. Mode I SIFs are much smaller but can exhibit significant values when the contact is located just ahead of the crack mouth. Decreasing the crack face friction significantly increases mode II stress intensity suggesting that this is one important mechanism by which lubricant entry into the surface braking crack can accelerate its propagation. The findings can help in improving the reliability of mechanical systems by supporting the development of new tools for prediction of remaining useful life of machine components such as bearings and gears.

1. Introduction

Rolling contact fatigue damage occurs in rolling-sliding, non-conformal contacts subjected to cyclic contact stresses such as those between the rolling elements and rings in a rolling element bearing or meshing gear teeth. It is a major failure mode in many machine elements. In the case of rolling bearings, it is a life-limiting failure mode and the widely used bearing life equations are entirely based on RCF lifetimes under given operating conditions. RCF involves a period of damage accumulation in the material exposed to cyclic contact stresses, which eventually leads to initiation of RCF cracks; these cracks then propagate under the cyclic loading and eventually lead to a detachment of small fragments of material from the running tracks. This results in formation of small pits or spalls so that associated failures are commonly termed pitting or spalling. These pits and/or spalls are small at first formation of the order of the contact size (hundreds of microns to a mm or so), but can become bigger (several mm) if the damage is allowed to progress. RCF cracks can initiate in the subsurface, commonly at locations of stress concentrations such as inclusions or other imperfections in the stressed volume of the material, or they can be surface initiated due to high stresses in the contacting asperities or at surface imperfections such as scratches and dents. Historically, subsurface initiation received more attention and the original bearing life equations [1] were based on it. However, surface initiated RCF has become considerably more

* Corresponding author.
E-mail address: b.kunzelmann18@imperial.ac.uk (B. Kunzelmann).

https://doi.org/10.1016/j.ijfatigue.2022.107449
Received 17 September 2022; Received in revised form 27 November 2022; Accepted 29 November 2022
Available online 5 December 2022

dominant in modern machine elements; this is due to a number of reasons including (i) the trend for a reduction in the oil film thickness caused by the use of lower viscosity lubricating oils with the aim of reducing fluid shearing losses, (ii) reduction in the size of the affected components leading to higher contact stresses and probably also (iii) due to improvements in steel cleanliness which reduces (but does not eliminate) the relative risk of subsurface crack initiation at inclusions. In this work, the term ‘pitting’ is used to describe damage resulting from surface-initiated RCF, while ‘spalling’ refers to damage caused by subsurface RCF.

Previous studies have shown that the occurrence of surface-initiated RCF damage is strongly related to the lubrication regime [2–14], and surface cracks and pitting are more likely to be observed under mixed lubrication conditions i.e. when there is a significant solid-to-solid contact [2,3,7,8,10–13]. Factors like surface topography and presence of debris affect the likelihood of surface crack initiation; in fact, cracks are commonly observed near dents and furrows on the surface [15]. Surface initiated contact fatigue damage is sometimes referred to as surface distress [4,16–18], but this term is also used to describe a specific mode of surface initiated contact fatigue, termed micropitting [17]. In the case of micropitting, surface cracks are formed due to asperity stress fluctuations and lead to formation of numerous, small pits (micropits) of the order of tens of microns. In contrast, pitting which is the subject of this paper, is caused by cyclic Hertzian stresses on the contact level and leads to larger macro-pits of size comparable to contact size. Fundamentally, the same phenomena of RCF crack initiation and propagation are responsible for both macro- and micro-pitting although the length scales of cracks and associated pits are different and the relative influence of some specific contact conditions, such as the amount of sliding, wearing-in and oil formulation amongst others, may be stronger in micropitting [19–21].

RCF involves a number of complicating factors that are not present in classical structural fatigue. This includes a complex multiaxial stress field due to over-rolling contact, presence of tangential loading due to contact friction, potential effects of lubricant entering the crack and, under mixed lubrication, a generally random stress history due to interaction of surface asperities. The asperity stresses are usually very high, commonly reaching the plastic limit of the material, and the number of stress cycles can be more than 10^9. Stresses are largely compressive and the loading history below the surface is nonproportional since the alternating shear stress component is not proportional to the compressive stress in the vertical direction. Due to the compressive loading the hydrostatic stress component is relatively high. Additionally, the fatigue process takes place in a considerable smaller volume of material than usually observed in structural fatigue as running tracks in EHL contacts in bearings and gears are limited to the size of the contact ellipse (usually in 100 s of microns) and the depth of the stressed zone is proportional to this. Some authors describe the RCF as a three-stage process, beginning with shakedown, followed by stable stress strain cycles which initiate the fatigue damage, and finally the instability phase where the crack initiates and propagates [22–24].

In engineering practice, RCF life to pitting of rolling bearings is predicted by employing Weibull distribution and taking into account the applied bearing load and resulting contact stresses. as first proposed by Lundberg and Palmgren [1] and later extended to also account for lubrication conditions and material fatigue stress limits [25,26]. This is a statistical approach which utilises a number of empirical constants to describe the material response to RCF. Deterministic models for RCF lifetimes exist (see for example [27,28] amongst others) although these are limited to a research setting rather than engineering practice and their prediction capability is still generally limited to the probability of survival [24] rather than fatigue crack growth.

In the 1980s, several computational fracture mechanics studies correlated RCF crack growth with standard fracture mechanics methods [29–38]. Keer and co-workers [33,35,37] present a numerical 3D LEFM model for crack propagation including lubricant interaction and note that the exact values of Paris law exponent, *m* are unknown and needed to further improve the predictions. Zhou et al [38] additionally incorporated roughness effects into their model for crack propagation of relatively short, micropitting size RCF cracks. Bower et al [36] predict stress intensity factors (SIFs) during one over-rolling cycle for a 2D RCF crack of assumed shape including the effects of fluid entrapment using a numerical method based on boundary element method. Kaneta and co-workers [29–32] produced a series of papers describing a boundary element approach to determine stress intensity factors for a surface breaking 3D crack including the effects of fluid entrapment and pressurisation during crack over-rolling. In these studies, a typical ‘half-penny’ crack shape was modelled in an elastic half space and exposed to cyclic stresses due to over-rolling Hertzian pressure. Traction was included with a simple Coulomb friction law assuming typical EHL contact friction coefficients. The models that include the influence of fluid entry into the crack and its subsequent pressurisation during over-rolling [29–38] do this in several ways. Some assume a constant or a linearly decreasing fluid pressure [29–32,35,38] over the crack length, while others assume a uniform fluid pressure acting throughout the crack [33–37]. The fluid entrapment effects are accounted for via suitable algorithms to model the opening of the crack mouth. Some of the significant findings of these studies are that inclined surface breaking RCF cracks mainly propagate in Mode II (shear mode) but that during the approach of the load, relatively small Mode I stress intensity factors also exist due to traction [29–38]. Much higher Mode I stress intensity factors were predicted by the relevant models described above where the entry and pressurisation of the fluid into the crack during over-rolling were considered.

Surface initiated RCF cracks show a characteristic morphology. A typical shallow V-shape is commonly visible at the surface, while cracks grow into the surface in the direction opposite to the direction of applied traction and at an angle of 20-30° to the surface [13,18,39,40]. Negative slide-roll ratios, defined as the sliding (and hence traction) direction being opposite to the direction of contact movement over the crack, promote damage [14,19,20,40] so that surface initiated RCF cracks are typically found on the body with lower surface speed. Once the primary surface crack has grown to a certain depth, a secondary crack often branches off and grows towards the surface [39] causing a segment of material to break off, thus causing a formation of a pit and failure of the functional surface [10,19,39].

In a recent experimental study, the co-authors of the present paper measured the growth rates of RCF cracks across the surface using a triple-disc contact fatigue test machine and a novel magnetic flux leakage based crack detection system that allowed cracks to be identified at a very early stage, when they were only 10 s of microns long. Experiments were conducted with AISI 52100 bearing steel roller specimens operating under mixed lubrication conditions with specific film thickness of $\lambda \sim 0.26$, Hertz pressures between 1.73 GPa and 4.7 GPa and slide-roll ratios of -5%. Fatigue tests were ran for 10 s of millions of cycles until pitting occurred but the test specimens were examined at frequent intervals to observe and measure the instantaneous surface length of generated RCF cracks so that their propagation rates can be determined. RCF surface cracks were observed very early in the tests, at around 15 % of the overall pitting life, with the remainder of lifetime taken up by crack propagation. This suggests that initiation phase is short and it is the crack propagation phase that consumes most of the fatigue life. This is in contrast to the early bearing life models which assume crack initiation to dominate bearing fatigue life [1], but in line with some of the newer deterministic research models which distinguish between crack initiation and propagation phase [22,24,35,41]. Results revealed existence of two distinct crack propagation phases: Phase I was associated with relatively short cracks (up to ~100–150 μm surface length in these experiments) where propagation rates were low and frequent crack arrests occurred. This phase consumed most of the fatigue life to pitting. In contrast, in Phase II, associated with longer cracks (> ~100–150 μm surface length), crack propagation rates increased rapidly with crack length. This ultimately led to pitting of the test specimens. It was noted that the existence of two distinct propagating phases is similar to the behaviour observed in structural fatigue. The authors postulated that Phase I may be initiated by the near-surface, asperity stress fields, while Phase II was likely to be driven by the macro contact (Hertzian) stress field. They showed that crack propagation rates in Phase II were proportional to the product of maximum Hertzian pressure p_0 and square root of surface crack length $(p_0\sqrt{c})$ to power of 7.5 and that the log–log plot of propagation rates against the product $p_0\sqrt{c}$ followed a straight line. This provided first experimental evidence that the growth of Phase II RCF cracks can be described by a Paris' type crack propagation law commonly used in structural fatigue. However, the actual stress intensity factors for the observed RCF cracks were not calculated, since there is no simple way of doing so for a crack under a rolling contact, so the exact Paris' law relationship for propagation of Phase II cracks could not be determined. Instead, the authors relied on the assumed proportionality of stress intensity factors with the parameter $(p_0\sqrt{c})$ to show the Paris-type behaviour. The actual stress intensity factors are clearly needed to fully establish the Paris' law relationship. This would then make it possible to directly predict the crack growth of an existing RCF crack.

Such ability to predict the propagation rates of RCF cacks in rolling contact is of obvious practical importance as it would ultimately allow for prediction of the remaining useful life of components subjected to RCF, such as rolling bearings, after detection of initial cracks. This approach is similar to damage tolerance concepts widely used in structural fatigue and is in contrast to fixed RCF lifetime predictions based on initially assumed loading of a new, crack-free component as is currently done in practice. Given that the recent results indicate that it is the propagation phase that dominates the overall fatigue lifetime, the prediction of crack propagation rates is a necessary step in improving our ability to forecast this remaining life. Consequently, this paper aims to achieves three main objectives:

i. confirm whether linear elastic fracture mechanics principles can be used to predict the growth of surface initiated cracks under rolling contact;
ii. investigate which mode of propagation is the most dominant one in propagation of surface RCF cracks;
iii. produce a Paris' law type expression that relates the RCF crack propagation rates to the applied stress intensity;

This is accomplished by using a combined experimental and numerical approach consisting of the following main steps:

1. The crack propagation rates and morphology for a number of surface initiated RCF cracks covering a relatively large range of crack lengths (from ~20 μm to 1142 μm surface length, the latter being equal to the semi width of the contact in the tests) are obtained experimentally using the approach described by the present co-authors in [39];
2. A 3D finite element model based on linear elastic fracture mechanics principles is created that is capable of resolving the actual stress intensity factors for the observed crack shape under the applied test loading;
3. Finally, the measured crack propagation rates are related to predicted stress intensity factors to assess the applicability of the employed linear elastic fracture mechanics approach to growth of RCF cracks and determine the Paris' law relationship that may be used to predict the growth of such cracks.

2. Experimental approach

Experiments were designed to be able to measure crack propagation rates under rolling contact using the methodology described in detail by the present co-authors in a previous study [39] so only a brief description is provided here. A triple-disc rolling contact fatigue rig (Fig. 1),

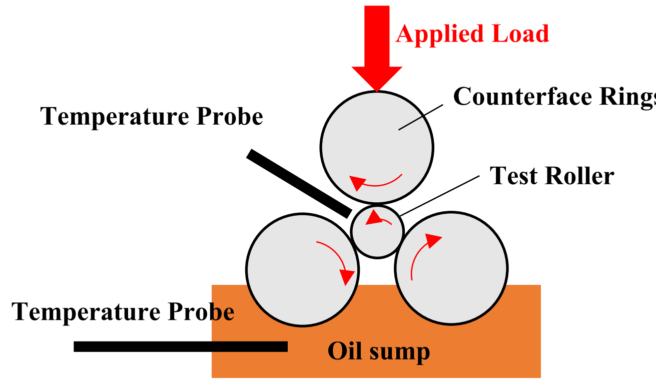

Fig. 1. A schematic of the triple-disc contact fatigue rig used in the present experiments.

commonly referred to as PCS Instruments MPR rig, was used to generate surface initiated rolling contact fatigue cracks on a test roller made of standard AISI 52100 through-hardened rolling bearing steel under controlled contact conditions. In this rig, a loading arm applies the load through the top disc which then generates reactions of equal magnitude in all three contacts owing to the arrangement of the three discs around the central test roller. The roller therefore sees three contact cycles each revolution so that a large number of loading cycles can be generated relatively quickly, up to 1 million an hour at the highest speed. The counterface discs undergo fewer load cycles and are deliberately harder and rougher than the test roller so that the fatigue damage is preferentially accumulated in the roller test specimen. Two separate electric motors drive the three counterface discs and the roller so that the entrainment speed and slide roll ratio can be controlled. The lubrication of the contact is provided via a dip lubrication system where the lower counterface disc is partially submerged into an oil sump. The oil sump temperature is controlled via an external cooler unit. There are two temperature probes, one measuring the temperature of the oil in the sump and one placed nearer to the inlet of one of the disc-roller contacts as indicated in Fig. 1. Friction is measured throughout the test via a torque meter on the roller shaft so that the lubrication regime is known at all times. An accelerometer attached to the test cell casing records the vibration level which helps to detect the final pitting failure. In the initial stages of running, the test is paused frequently, and the roller surfaces inspected under an optical microscope in order to capture the cracks very early in their development. The detection of early cracks is further aided by the use of a custom made magnetic flux sensor, described in [39 42] (not shown in Fig. 1). The sensor is placed near the roller surface and its continuous signal output, in combination with frequent inspections, helps to detect cracks at an early stage. The length of cracks at first detection is in the region of 10 to 20 μm. It should be noted that the crack length at first detection is not equivalent to crack length at initiation, this is likely to be smaller yet if such length can be defined at all, but it simply refers to the crack length when a crack is first observed using the said optical method. Subsequent to the detection of the first crack, crack lengths on the surface are obtained at frequent, regular intervals by stopping the test and inspecting the roller surface using a digital optical microscope. Once the ex-situ crack measurement is complete the roller is put back in the rig and test continues until the next inspection. The custom made magnetic flux sensor described above is particularly useful in the latter stages of crack propagation (Phase II) when cracks grow rapidly so stopping the test at regular intervals no longer works for capturing crack growth with a suitable resolution; instead, the senor output indicates when a crack has grown by a small increment and the test can then be stopped for crack measurement accordingly. The exact process is described in [39] and [42]. By knowing the crack length at each inspection and the number of loading cycles between each inspection, it is possible to determine the crack propagation rates across the surface. Those crack propagation rates are related to the stress intensity factors predicted by the LEFM model as described later. The crack length in the subsurface is also measured at the end of the test by sectioning the roller specimens and this is used to describe the instantaneous crack shape in the LEFM model.

The employed test parameters and specimen properties are summarized in Table 1. The test specimens are 12 mm diameter cylindrical rollers that were removed from an actual unused cylindrical roller bearing of a suitable size. The counterface rings are crowned discs supplied by the test rig manufacturers, PCS Instruments. This specimen geometry produces an elliptical contact with dimension 351 mm and 571 mm in rolling and transverse directions respectively. The roller specimens and the disc counterfaces are both made of standard through-hardened AISI 52100 rolling bearing steel. Roller and disc surfaces are circumferentially ground with resulting Ra surface roughness of ~30 nm and ~130 nm respectively, which is typical of that found in medium size rolling bearings.

A relatively high load, equivalent to max Hertz pressure of 4.7 GPa, is chosen in order to accelerate the fatigue tests. The lubricant used is a custom-made blend of a PAO base stock with a standard ZDDP antiwear additive added at a concentration of 0.1 % wt of phosporus. The addition of the antiwear additive follows the procedure established in past RCF studies [39,42,43] where it was shown that this is needed to prevent excessive wearing of the counterface discs which would significantly change their roughness properties and hence the lubrication conditions and asperity stresses during the rolling contact fatigue tests; the presence of a good anti-wear additive minimises these changes so that the contact conditions experienced by the test roller are as constant as possible throughout the test duration. The higher hardness of the discs also minimises their roughness changes further helping to control the contact conditions; this is again in line with previous proven RCF test methods [39,42–44]. The slide-roll-ratio SRR (ratio of sliding speed to entrainment speed) was set to −5%; the roller was the slower of the two components to ensure that the damage is preferentially formed on the test roller surface, as discussed above and in-line with numerous previous RCF studies [39,42–44].

The λ-ratio (the ratio of the central oil film thickness, h_c, to combined centre line average roughness of the roller and discs) was set to 0.26. This relatively low value was deliberately chosen to promote generation of surface initiated rolling contact fatigue cracks without the need to use any artificial dents which are often used in other RCF studies – here, cracks simply initiate due to contact stresses as they would do in real bearings. This value of the λ-ratio was calculated by using a thermal network of the triple-disc rig described in [20] which accounts for the influence of frictional heating on the surface temperatures of the discs and the roller and hence any reduction in the inlet oil viscosity due to this. This produces more accurate lambda values than simply using the oil viscosity at the sump oil temperature which is lower than inlet temperature. The effect in terms of film thickness is relatively small in the present tests owing to a relatively low slide-roll ratio (and hence low frictional heat input) but can be significant with higher sliding [20].

Table 1
Test parameters used in the crack propagation experiments on the triple-disc rig.

Test Parameters		
Max Hertz Pressure, p_0	4.76 GPa	
SRR $(=U_{roller} - U_{disc})/U_{entr}$	−5%	
Entrainment Speed, $U_{entr}(=0.5\ (U_{roller} + U_{disc})$	3.8 m/s	
Lambda ratio, λ $[= 2\ h_c/(R_{a\_disc} + R_{a\_roller})]$	0.26	
Test Oil	PAO 5 + ZDDP (0.1 % wt P)	
Sump Oil Temperature	65 ± 5 °C	
Counterface Disc (AISI 52100)	Roughness (R_{a_disc}) Hardness (Vickers) Diameter	~ 130 nm ~750HV 54.15 mm
Roller Test Specimen (AISI 52100)	Roughness (R_{a_roller}) Hardness (Vickers) Diameter	~ 30 nm ~710HV 12 mm

2.1. Typical examples of the observed crack morphology

Fig. 2 shows examples of surface cracks in AISI 52100 bearing test roller generated in this study. The cracks show shallow V-shape typical of rolling contact fatigue. In Fig. 2 b), a secondary crack breaking the surface is seen ahead of the main crack. No pit has formed yet. The surface crack length dimension is also shown in these images. This is one of the parameters used to model the crack shape in the FE crack model shown later. The reported propagation rates are measured as the changes in this dimension with the number of loading cycles, as described earlier,

Fig. 3 shows example microscopic images of typical cross-sections (in the rolling direction) of three separate cracks generated here. These are obtained by sectioning the crack in the mid-plane in the rolling direction and subsequently polishing the sectioned surface so that it can be inspected optically. It is evident that the cracks grow at an angle of about 30° to the surface and against the direction of friction. These are all typical characteristics of surface initiated RCF cracks [19]. A secondary crack is also visible in the example of Fig. 3(c) as indicated. The secondary crack is seen to branch off the main crack to grow towards the surface; in this example it breaks the surface and consequently, undermines a segment of material located between it and the main crack. This segment would have eventually detached from the surface and this is how a pit is formed; in this case, the test was stopped prior occurrence of pitting as we are interested in the crack length at this stage as input to the model, and a pit would make it difficult to observe the actual crack. Several incidences of crack branching are evident in Fig. 3(b) and (c). The subsurface crack length dimension, d, which is used as an input to the model described later, is indicated in Fig. 3(a); this dimension is another parameter used to describe the crack shape in the FE crack model shown later.

3. Finite element model for prediction of stress intensity factors in observed RCF cracks

This section describes the developed 3D finite element (FE) model used to compute mixed-mode SIFs for experimentally generated surface cracks subjected to RCF loading. The model is implemented entirely in ABAQUS 2017 using a suitable fracture mechanics mesh to capture the crack behaviour as described below. The experimental study was used to capture the crack dimensions of six cracks covering a wide range of surface crack half-lengths from 11 mm to 570 mm corresponding to early and late stage cracks observed prior to pitting (stage I and stage II cracks in [39]). The corresponding subsurface crack lengths were also obtained in each case by sectioning the corresponding roller specimens and observing the cracks in section under a microscope as shown earlier in Fig. 3. The crack growth rates on the surface at the time when the given crack was observed were obtained using the sequential microscope images of each crack as described in the former chapter. Table 2 lists the relevant crack dimensions and growth rates observed in the experiments. The crack growth rates dc/dN refer to the surface crack length c which is defined defined as $c = 2a$. The surface crack length c was recorded in the experiments and modelled as surface crack half-length a in the FEA model using a symmetry boundary condition.

An FE model of these cracks based on LEFM was set-up with the following characteristics, assumptions and simplifications:

- Linear elastic fracture mechanics (LEFM) conditions were assumed. This is justified on the basis that plastic zones at the crack tip are expected to be small for the through-hardened bearing steel considered. Linear elastic material behaviour is modelled, as after the first load cycles due to shakedown no significant plasticity of the material is expected [24,45,46].

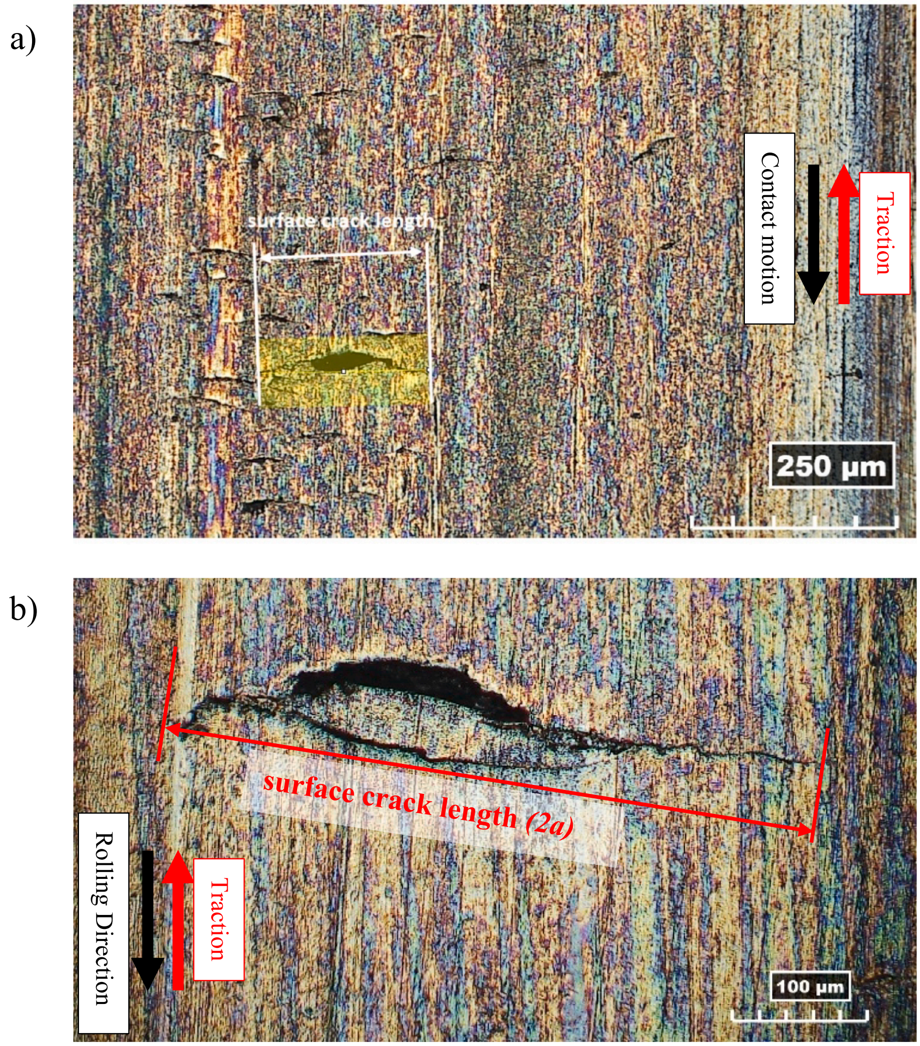

Fig. 2. Optical images of two typical examples of a surface RCF cracks on AISI 52100 test rollers generated on the triple disc machine. Directions of contact motion and traction are as indicated; The surface crack length dimension (2a) is also shown – this is used in the crack model described later and measured propagation rates indicate changes in this dimension.

- Each crack captured in experiments as listed in Table 2 was modelled in Abaqus by simplifying the real crack shape into a 'concave half penny' shape inclined at an angle of 30° to the surface (see Fig. 4 for depiction of this simplified crack shape).
- The simplified crack geometry allows for implementation of a symmetry plane, visible in Fig. 5, in the middle of the crack to decrease computational effort.
- A fracture mechanics mesh with quarter point nodes and six contours was applied at the crack tip of the inherent flaw, details of which are shown in Fig. 4 and Table 3.
- Fig. 4a shows a representative 3D FEA model of such a crack, Fig. 4b shows a detailed view of the crack tip and Fig. 4c shows the contours of the fracture mechanics mesh. Fig. 4a also shows the direction of the load motion over the crack and the positions '0′ and '1′ on the crack front which are later used for plotting SIFs against the crack position.
- The pressure distribution on the surface was simplified to a Hertzian pressure distribution for an ideally smooth contact. The position of the centre of this pressure distribution from the crack mouth is denoted as e, with e being negative when the pressure distribution is ahead of the crack mouth (i.e. approaching the crack) and positive when it is past the crack mouth (see Fig. 5).
- The over-rolling is modelled by moving this pressure distribution over the crack from a position 3 contact semi-widths behind the crack mouth (i.e. from position e/c = -3) to 3 contact semi-widths ahead of the crack mouth (i.e. to position $e/c = +3$). Model is run for multiple e/c positions so that evolution of SIFs during over-rolling can be captured. The movement steps are variable, see later for how this was implemented.
- To model the contact between the crack faces, a pure penalty contact model with a relatively high friction coefficient of 0.5 was applied unless otherwise stated. This friction coefficient is of a similar magnitude as stated in [32] and this relatively high value was chosen due to evidence of fractured surface and presence of debris between crack faces, suggesting that rubbing and interlocking of the crack faces takes place as discussed in [39]. A separate study was conducted to investigate the effect of the imposed crack face friction coefficient value on predicted stress intensity factors, as shown later.
- The surface traction coefficient in the model was set to 0.1 unless otherwise indicated. This is based on the measured values in the triple disc rig and is representative of contacts operating under mixed to boundary lubrication conditions.
- Potential effects of the lubricant entering the crack and its subsequent pressurisation during over-rolling, as discussed in [29–32,36]

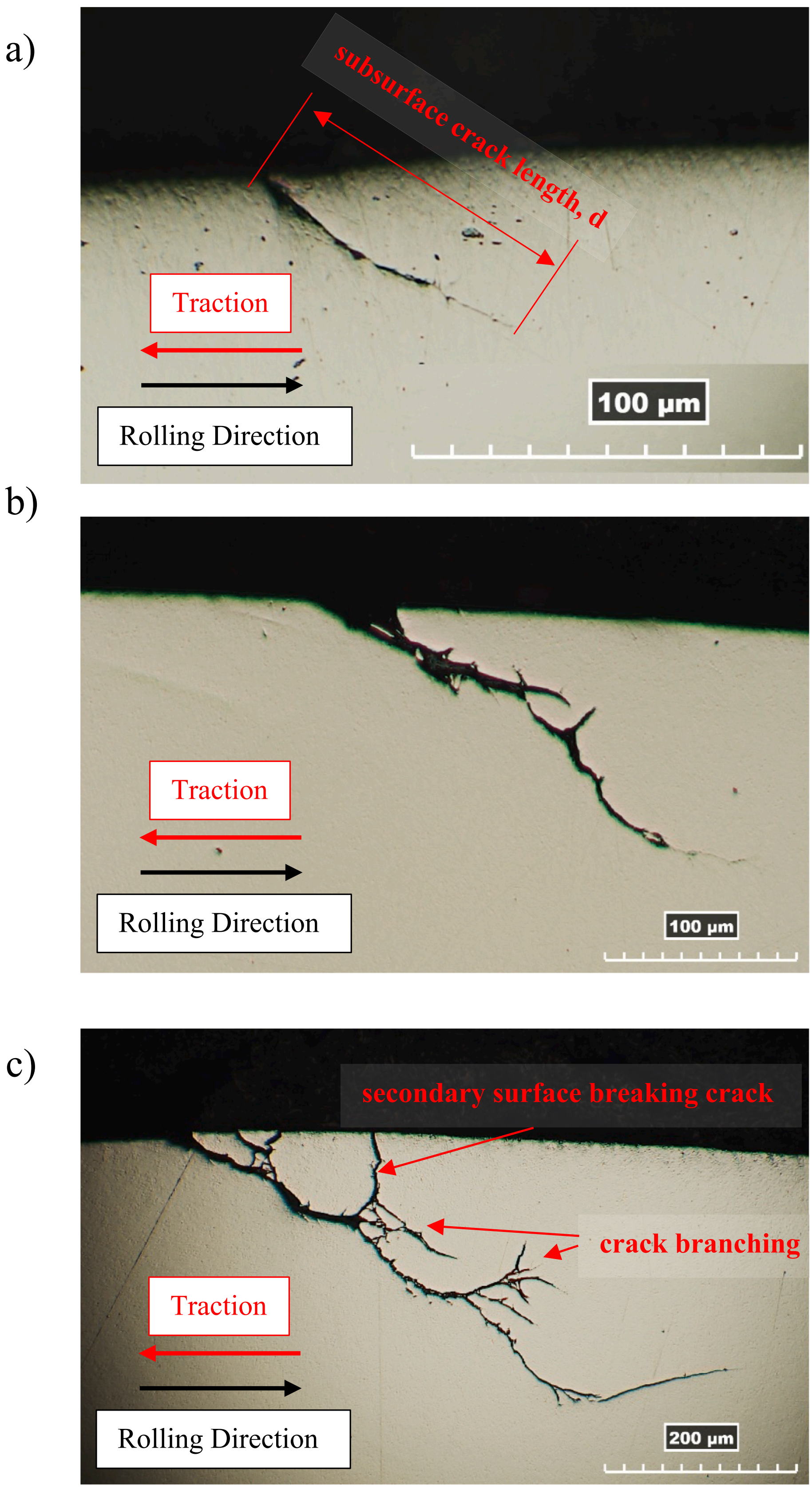

Fig. 3. Optical microscope images of three typical examples of cross sections (in rolling direction) of surface breaking RCF cracks on AISI52100 test roller generated on the triple-disc machine. a) Realtivley short (Stage I) crack, also showing the subsurface crack length dimension, d, as used in the FE model described later; b) An example of a Stage II crack c) An example of an even longer stage II crack; note the existence of a secondary surface breaking crack and multiple incidences of crack branching in this particular example.

Table 2
Measured surface and subsurface crack lengths and propagation rates across the surface for the six selected RCF surface cracks modelled in this work.

Case	Surface crack half-length, a [μm]	Subsurface crack length, d [μm]	Aspect ratio, a/d [-]	Surface Crack growth rate, dc/dN [μm/million cycles]
1	11	11	1.0	10
2	29	20	1.4	10
3	57	29	2.0	10
4	114	46	2.5	100
5	286	175	1.6	1000
6	571	400	1.4	50,000

are not considered in the present paper although the implemented model is capable of including this.

The accuracy of the model is obviously dependant on the meshing strategy employed to capture the conditions at the crack tip. A fracture mechanics mesh with quarter point nodes and six contours was here used at the crack tip. This approach was recently shown to be suitable for this purpose by [47]. The meshing of the crack tip was validated by confirming the path independency of the J-Integral when no contact (and hence no friction) occurs between crack faces (see Fig. 6). Without traction at the crack surface the J-Integral has to be path independent over the contours of the fracture mechanics mesh (see for example [48]) for the mesh to be valid. Subsequently, the surrounding mesh size was chosen to accurately describe the imposed Hertz pressure on the surface and predict the resulting subsurface stresses. Finally, the mesh size on the crack face surfaces was refined in order to capture the crack face contact pressure accurately. Table 3 shows the number of elements, nodes and the resulting computational effort for the implemented model for case 6 crack in Table 2.

To simulate over-rolling, the imposed Hertz pressure on the surface is moved from three semi-widths (in rolling direction i.e. along the minor axis) behind the crack mouth to three semi-widths ahead of the crack

Table 3
Mesh Data and Performance for Case 6 crack listed in Table 2.

Number of Elements	Number of Nodes	Computational Effort for 1 Over-rolling (PC with *16 cores and 32 GB RAM*)
474,840	687,372	~40 h

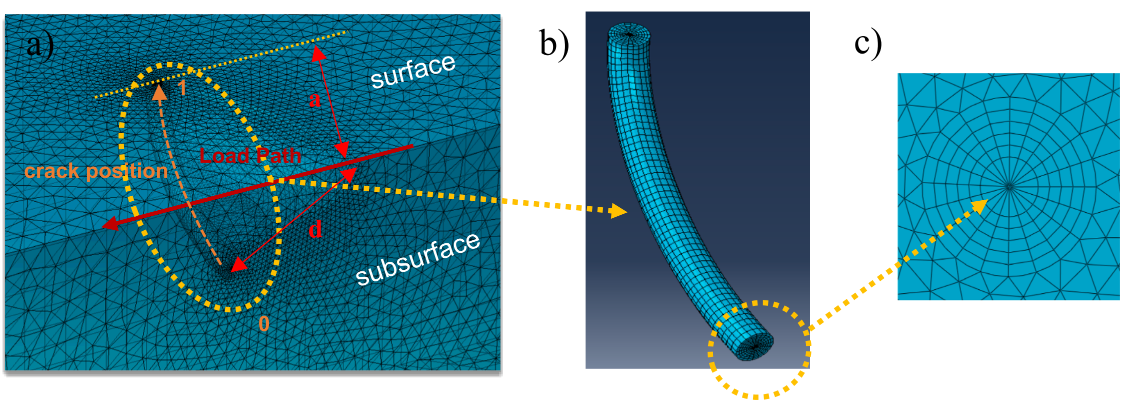

Fig. 4. Fracture mechanics mesh of the 3D flaw a) simplified crack geometry with surface crack semi-length, a, and subsurface crack length, d; b) fracture mechanics mesh at the crack tip and c) contours of the fracture mechanics mesh.

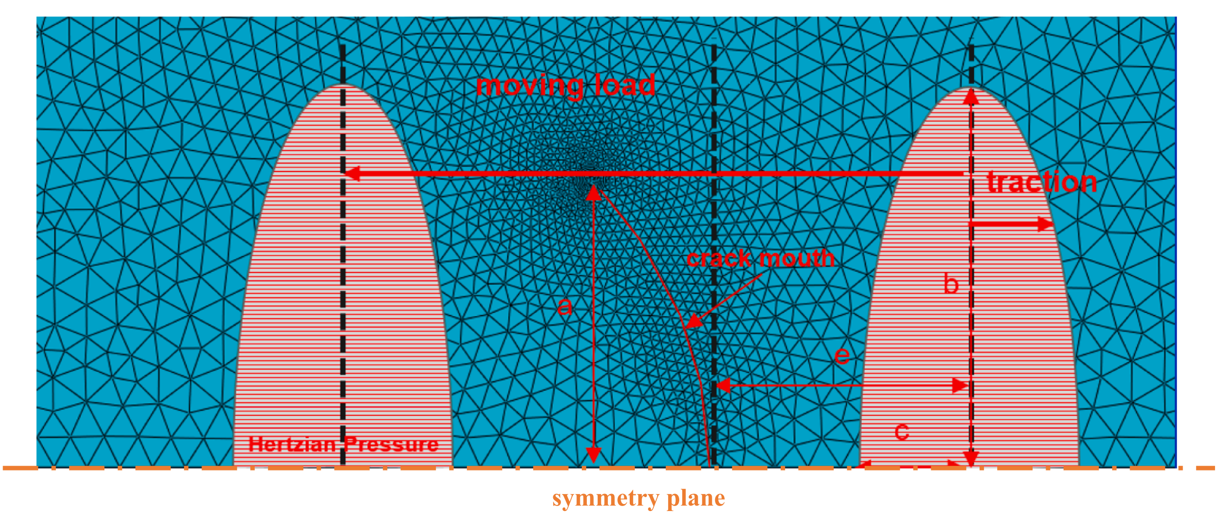

Fig. 5. An illustration of the moving Hertz pressure distribution over the contact used to simulate the over-rolling in the present model (plan view). The instantaneous position of the contact in relation to the crack is defined by dimension 'e' measured from crack mouth as shown (e is negative when contact is approaching the crack mouth i.e. to the right of the broken line indicting location of the crack mouth, and positive when contact is past the crack mouth i.e. to the left of the broken line). c and b are contact semi-widths in rolling and transverse directions respectively and a is half the crack length on the surface. The imposed Hertz pressure moves from e/c = -3 to e/c = 3. Only two load positions are shown here for clarity, many more are used in the actual simulations. Symmetry plane is also shown.

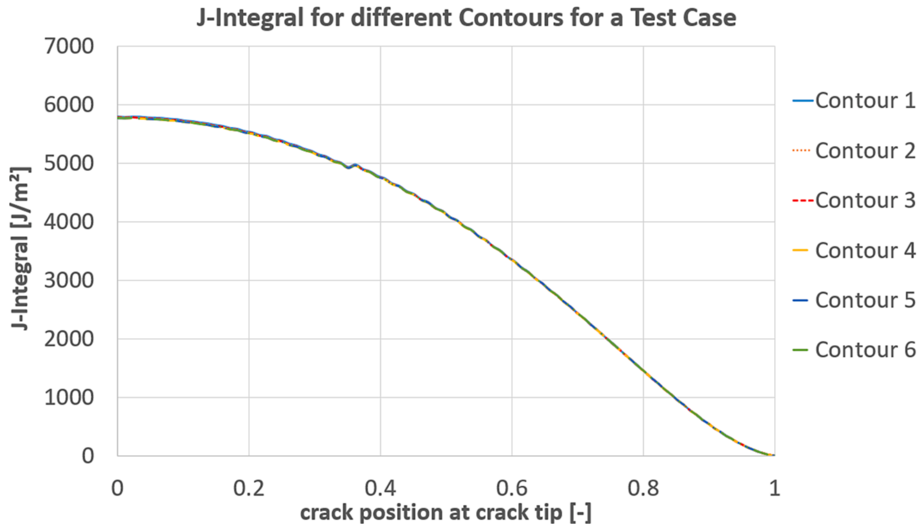

Fig. 6. The calculated values of the J-Integral along six different contours of the implemented mesh at the crack tip for the case with no crack face contact. Example shown is for crack Case 6 in Table 2. J-Integral values are the same for all six contours at all positions along the crack tip demonstrating path independency of the J-integral and hence the general validity of the mesh.

mouth in variable time steps using the ABAQUS variable time step solver. This makes it possible to capture the evolution of SIFs during crack over-rolling. The number of the variable time steps and the associated crack locations can be seen in the lower plots of Figure 9. The instantaneous position of the centre of the contact (as represented by the Hertz pressure distribution in the model) in relation to the crack mouth is described by the ratio e/c, where e is its current position as measured from the crack mouth and c is the contact semi-width in rolling direction, as shown in Figure 5. The applied Hertzian pressure distribution and contact dimensions were the same as those used in the tests. Table 4 summarizes the contact conditions which were used in the FE model. Stress intensity factors were extracted from Abaqus by evaluating the energy domain integral for a virtual crack advance using q-vectors during one over-rolling cycle at different positions along the crack front (from 0 to 1 in Fig. 4 a) and evaluated with Matlab.

4. Results

4.1. Evolution of stress intensity factors during over-rolling

Fig. 7 and Fig. 8 show the results of the FE model for the example of Case 6 crack defined in Table 2 where surface crack half-length was $a =$ 571 mm and subsurface length was $d =$ 400 mm. The general trends shown are similar to those for all other cracks modelled here (Table 2). Fig. 7 shows predicted deformations from the FEA model for three contact positions relative to the crack: a) when the load is one contact semi-width ahead of the crack mouth (e/c = -1), b) when load is directly above the crack mouth (e/c = 0) and c) when load is behind the crack mouth (e/c/ = 1.7). It is seen that the crack mouth is open as the load approaches (Fig. 7a). This would facilitate the entry of fluid into the crack although this is not modelled here. When load is over the crack mouth (Fig. 7b) or just beyond it (Fig. 7c) the crack is seen to be completely closed. In addition, the crack faces are seen to undergo displacement relative to each other, the amount of which depends on the load position relative to the crack (Fig. 7b and 7c). This suggests that crack faces rub against each other during over-rolling. Such rubbing has been suggested to be responsible for microstructural changes in the vicinity of the crack faces including the formation of white etching areas [43]. The maximum relative displacement of crack faces predicted in this particular case was ~ 0.5 µm. The deformation plots in Fig. 7 show clearly how the RCF crack geometry and the over-rolling load create the complex mixed mode loading at the crack tip with all 3 fracture modes being active at the crack tip.

Fig. 8 shows the evolution of the stress intensity factors along the crack front during the complete over-rolling process i.e. as load moves from e/c = -3 to e/c = +3. The position along the crack front is defined to be between 0 and 1, with 0 being at the centre of the ‘concave half-penny’ crack (i.e. at the greatest crack depth) and 1 at the extreme of the crack front on one side i.e. where the crack front reaches the surface (see Fig. 4a for clarity). From Fig. 8a it is evident that maximum in mode I SIF occurs as the contact is approaching the crack mouth. In this example the predicted maximum is ~ 2.5 $MPa\sqrt{m}$ at e/c = -1. The deformation plot corresponding to this maximum mode I SIF is that shown in Fig. 7a where crack opening is evident. Fig. 8b shows that it is the mode II stress intensity factors that reach the highest magnitude during over-rolling with the maximum occurring when the contact is in the vicinity of the crack mouth. The corresponding deformation is shown in Fig. 7b) and it is seen that significant relative displacement in rolling direction exists between the crack faces at this point hence promoting mode II crack growth. Finally, Fig. 8c shows that appreciable mode III SIFs can occur after the load has passed the crack mouth (corresponding deformation plot is shown in Fig. 7c).

In order to correlate the measured crack propagation rates with the computed SIFs, the maximum SIF range (ΔK) has to be determined for the mixed mode loading during one over-rolling cycle. Therefore, ΔK values for each mode along the crack tip for one complete over-rolling (e/c = -3 to e/c = 3) were extracted from the SIF results. Fig. 9 (top) shows 2D plots of mode I, II and III SIFs along the crack tip for all 16 values of e/c considered here, again for the example of Case 6 crack. The location along the crack tip where the maximum ΔK was found to occur in each case is also indicated (red line). This is in the middle of the crack (position 0) for mode I and mode II and ~ 0.68 for mode III. Fig. 9 bottom then shows the evolution of SIFs at this maximum position only during over-rolling, from which the ΔK values are evident. When determining the Mode I ΔK_{max} (Fig. 9a) the negative K values are ignored as this indicates crack closure and no crack growth in mode I occurs in

Table 4
Contact Conditions Used in the FE Model (the same as in the experiments).

Maximum Hertzian contact pressure, p_0 [GPa]	Semi contact width in rolling direction, c [µm]	Semi contact width transverse to rolling, b [µm]
4.76	351	571

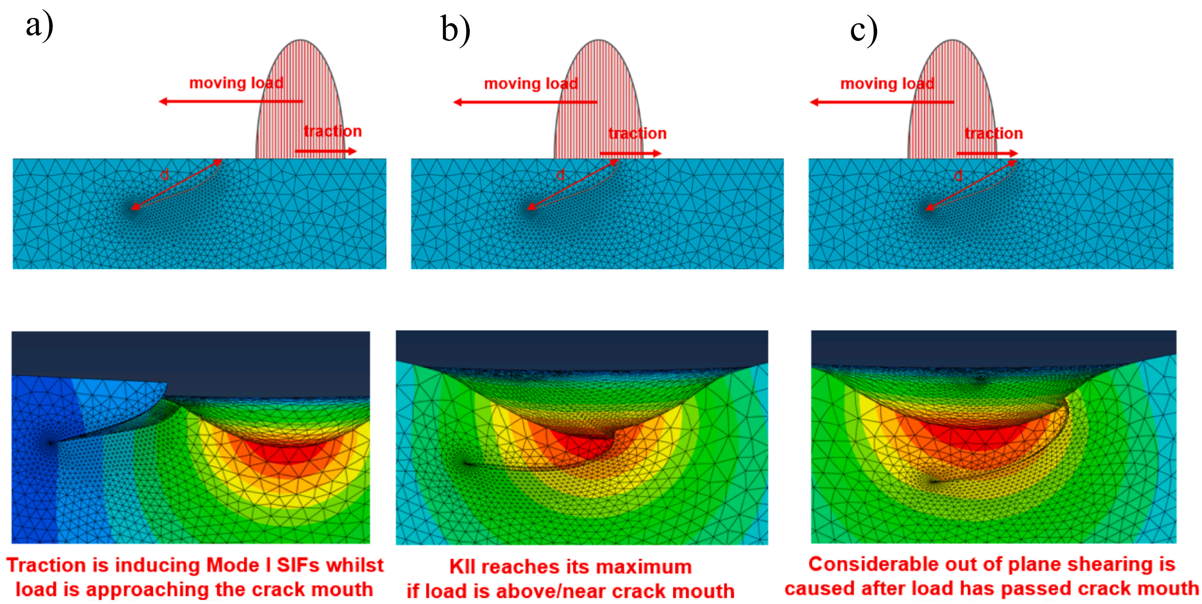

Fig. 7. Deformation fields (bottom row of pictures) predicted by the FEA model for three positions of the load relative to the crack mouth (position illustrated in top row of pictures). The three positions are: a) load ahead of crack, e/c = -1; b) load over crack mouth, e/c = 0; c) load behind the crack, e/c = 1.7.

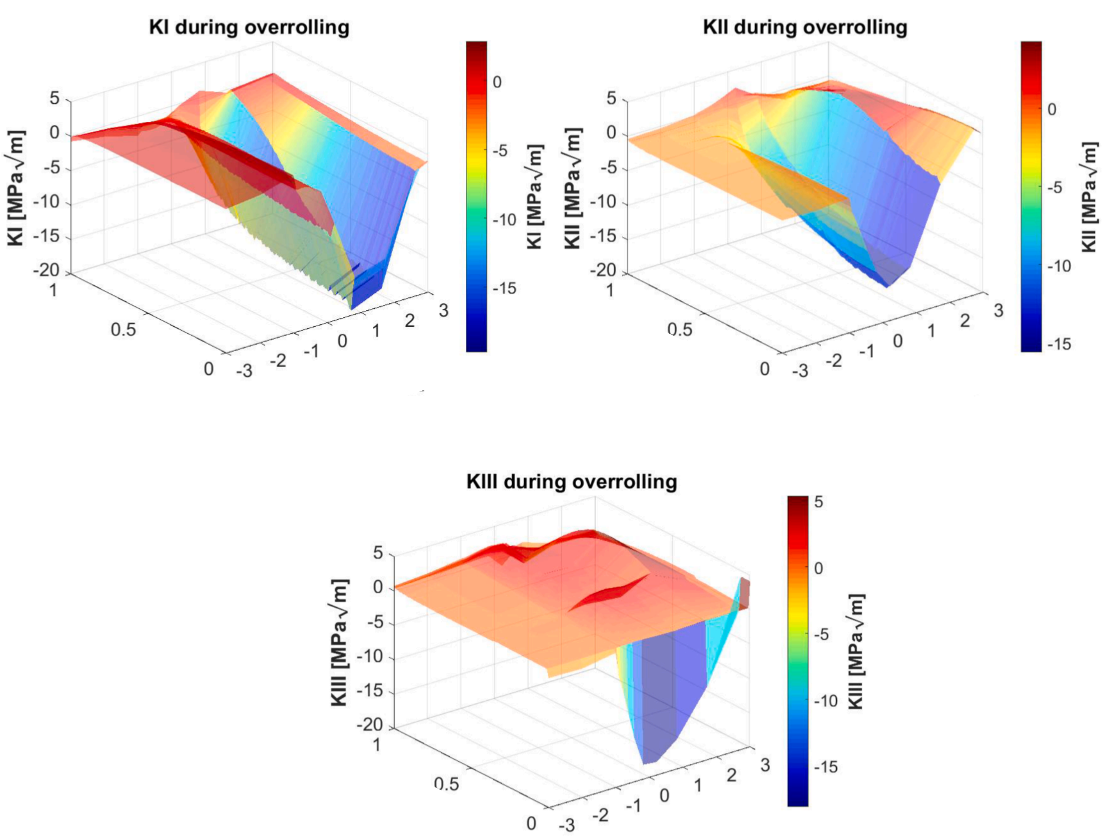

Fig. 8. Evolution of stress intensity factors along the crack front during one over-rolling of the contact from e/c = -3 to e/c = 3. a) mode I; b) mode II; c) mode III. Position along the crack front is defined between 0 (centre of crack i.e. the greatest depth) and 1 (one extreme of the crack front at the surface), see Fig. 4.

this case. This is in line with other studies, for example Elber [49].

Due to crack closure effects, caused by the primarily compressive loading, only a small ΔK_I is predicted, caused by the opening of the crack mouth shortly before the contact reaches the crack mouth, as was evident in the displacements shown Fig. 7. The reader is reminded that no effects of potential fluid entrapment and pressurisation in the crack are considered here; if present, these effects would act to increase ΔK_I. Mode II propagation on the other hand is possible even during crack

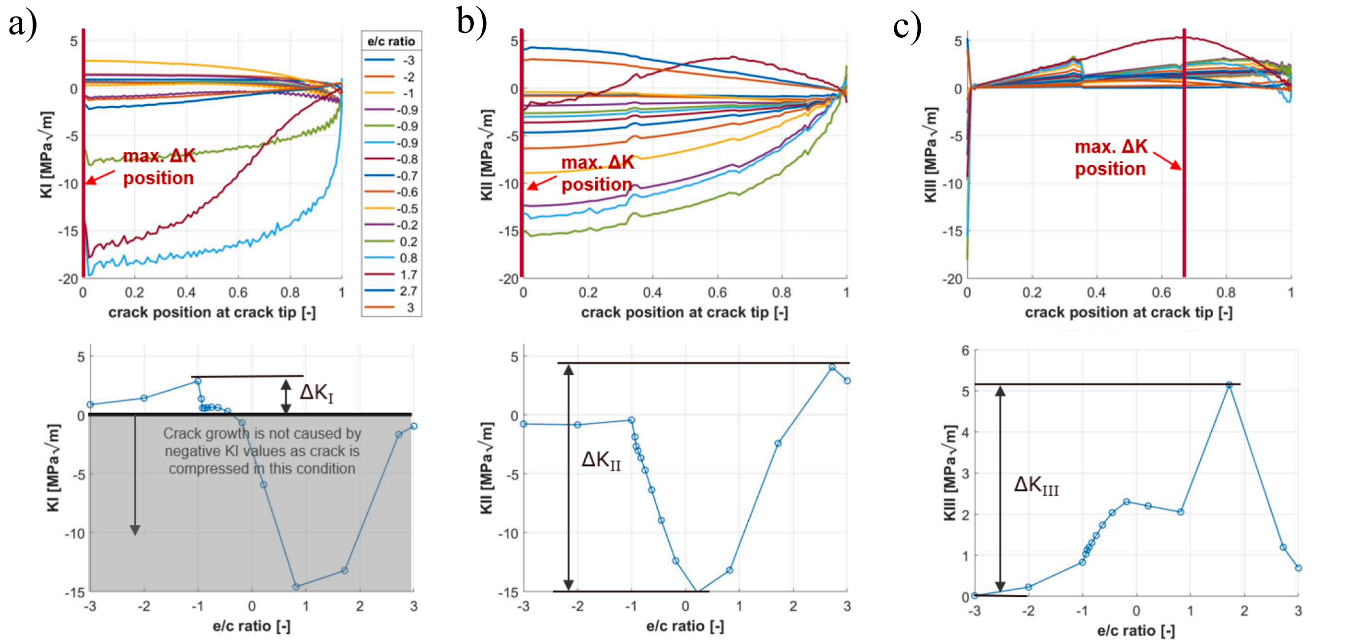

Fig. 9. Evolution of stress intensity factors during one over-rolling for the crack in case 6 in Table 2 a) mode I; b) mode II and c) mode III. Top. Predicted SIF along the crack front for all load positions considered (as indicated by e/c/ value); Bottom: Evolution of SIF with contact position at the crack front location where maximum ΔK is found in each case (indicated by red line in top plots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

closure as crack faces were shown to undergo relative displacement so shear mode propagation can still occur. ΔK_{II} is shown to have by far the highest values of the three propagation modes, indicating that mode II propagation dominates. ΔK_{III} is zero at the symmetry plane going through the centre of the crack as expected, but somewhat counterintuitively, it has an appreciable magnitude away from this position near the surface. This is because for the modelled 3D crack geometry a small relative displacement of the crack faces in the transverse direction is predicted i.e. out of plane shearing (see Fig. 7), giving rise to ΔK_{III}, albeit several times smaller than ΔK_{II}.. Similar trends in ΔK_{III} were reported in [50].

4.2. Applicability of Paris' law to RCF crack propagation

To assess the applicability of the Paris' law to the observed crack propagation, the maximum ΔK values for each of the three propagation modes and for all modelled cracks (Table 2) were extracted from the above data and plotted against the crack propagation rates across the surface measured in the triple-disc tests at the point when the modelled crack geometries were observed. Additionally, the effective stress intensity factor for mixed mode crack propagation was determined using mode I and mode II ΔK values, as proposed in [51]:

$$\Delta K_{eff} = \sqrt{\Delta K_I^2 + \Delta K_{II}^2} \tag{1}$$

ΔK_{III} values are not included here in the calculation of the effective stress intensity factor ΔK_{eff} as we are interested in the maximum ΔK_{eff} values which occur at the middle of the crack (and are dominated by ΔK_{II}) whereas the maximum of ΔK_{III} is near the surface. The approach is also consistent with that in [51] but whether or not ΔK_{III} is included is

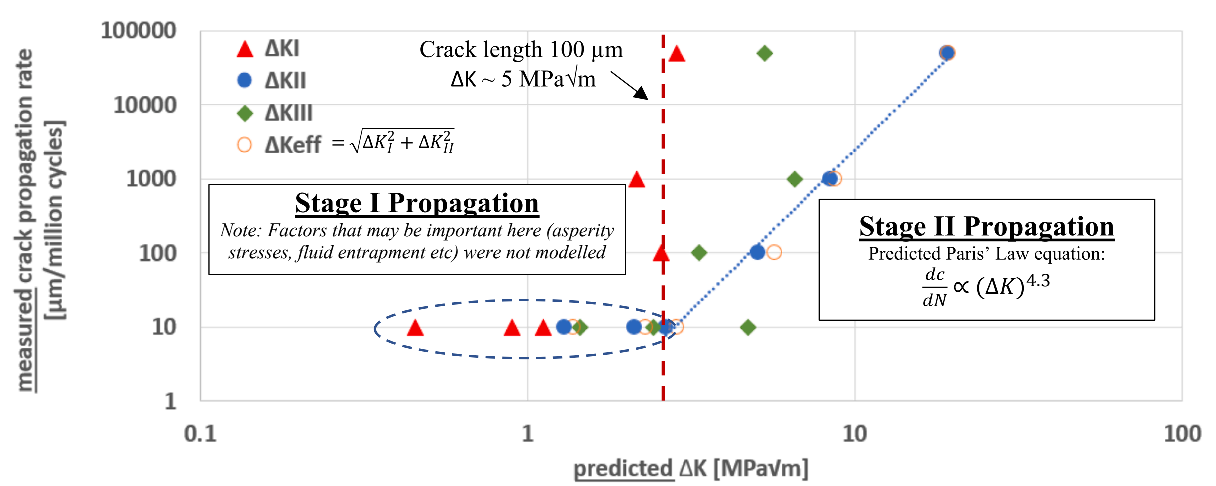

Fig. 10. Plot of measured crack propagation rates across the surface against the predicted ΔK_I, ΔK_{II}, ΔK_{III} and ΔK_{eff} values for the six cracks described in Table 2. Stage I and Stage II propagation regions (as described in [39]) are as indicated. The form of the Paris' Law equation for Stage II propagation derived from these results is also shown.

irrelevant as the shown trends are exactly the same either way. Fig. 10 shows this plot for all three ΔK values as well as ΔK_{eff} for the six cracks considered. It is immediately apparent that ΔK_{II} plot against measured propagation rates forms a straight line for Stage II cracks (i.e. crack length greater than about 100 μm as described earlier). This suggests that for these cracks, a Paris' law relationship of the type $\frac{dc}{dN} = C(\Delta K)^m$ exists and, provided that constants C and m can be determined, can be used to predict crack propagation rates for a given crack size and loading. It is also evident that ΔK_{II} assumes the biggest values of the three modes for the Stage II propagation and is therefore the main driver for the observed RCF crack growth in this region. ΔK_{eff} also follows a Paris' law relationship since it is dominated by ΔK_{II}. The range of predicted ΔK_{II} (and ΔK_{eff}) shown here is up to 19 MPa $MPa\sqrt{m}$, this highest value corresponding to the longest crack in present analysis with surface crack length of 571 μm (Case 6 in Table 2). The significance of this is further discussed below. In contrast, ΔK_I and ΔK_{III} are much smaller for all cases, and for the case of the longest crack analysed they are an order of magnitude smaller than ΔK_{II}. ΔK_I and ΔK_{III} also show no obvious correlation with measured crack propagation rates suggesting that mode I and mode III propagation are not significant in these tests.

Using the presented plot for the surface crack propagation rates $\left(\frac{dc}{dN}\right)$ for stage II cracks against computed ΔK_{II} values, the constants C and m in the Paris' law equation can be approximated. This results in the following equation for Paris' law for propagation rate across the surface of stage II cracks i.e. cracks longer than about 100 μm:

$$\frac{dc}{dN} = 0.154\frac{\mu m}{10^6}\left(\frac{\Delta K_{II}}{MPa\sqrt{m}}\right)^{4.3} \tag{2}$$

where $\left(\frac{dc}{dN}\right)$ is the surface crack propagation rate in μm per million stress cycles and ΔK_{II} is in MPa $m^{1/2}$.The value of the stress exponent, m, is predicted to be around 4.3. This value is comparable to that obtained by Beswick [52] from their fracture tests with compact tension specimen with the same AISI 52100 bearing steel where m was quoted to be around 3. Although the exact numerical expression of Paris' law stated in Eq. (2) is for propagation of RCF cracks across the surface, if a common assumption of the crack being a half-penny shape is made, the surface crack length is then directly proportional to the subsurface crack length so that the same general law holds for the propagation of cracks into the surface, albeit with a different multiplication constant.

For all cracks shorter than 100 μm on the surface (Stage I propagation in [39]), predicted ΔK_{II} (and ΔK_{eff}) values are small between 1 and 5 MPa $m^{1/2}$ and predicted ΔK_I is smaller yet at close to or lower than 1 MPa $m^{1/2}$. There is no correlation between predicted ΔK values and measured propagation rates in this region. This is to be expected as the crack growth of these very short cracks is likely to be driven by factors not modelled here, not least the stresses in asperity micro-contacts, as also postulated in [39], and fluid entrapment effects. The absolute values of ΔK predicted here for these short cracks are therefore rather meaningless and the 'real' stress intensity factors are likely to be higher. However, crack growth was observed and propagation rates were measured in this region so that the lack of any correlation with ΔK as computed here does at least suggest that standard Paris' law type expressions based purely on Hertzian (macro-) contact stresses cannot be used to describe growth of these short cracks. Large part of the overall life to failure is consumed by crack propagation within this Stage I region, so a better understanding of the phenomena at play here is important in our ability to forecast RCF lifetimes. The transition from Stage I short crack growth to Stage II crack growth was shown in [39] to occur at crack length of approximately 100 μm. The present results indicate that this corresponds to ΔK_{II} (or ΔK_{eff}) values of approximately 5 MPa $m^{1/2}$ (indicated by dashed vertical line in Figure 10). Further tests at different surface roughness, loads and contact geometry are needed to confirm the significance, if any, of this transition value but it is perhaps worth noting that 5 MPa $m^{1/2}$ is comparable to threshold stress intensity factor range extracted from fracture mechanics experiments for this type of steel which is around 3–5 MPa$\sqrt{m}$ [52].

4.3. Effect of crack face friction

Crack face friction is known to have a significant influence on RCF crack propagation [19,36] with higher values likely to decrease crack growth rates. In the results presented so far, the effective friction coefficient in the contact between the crack faces was set to 0.5. This relatively high value was chosen to reflect the fact that the prepared sections showed evidence of debris in the crack, suggesting severe conditions in the contact between the crack faces, as well as crack deflections which would impede relative movement between the crack faces. However, it is not possible to measure the real value of the crack face friction. In order to explore the potential influence of crack face friction on crack propagation, stress intensity factors for Stage II propagation were computed with two additional crack face friction values, namely 0.25 and 0. Fig. 11 shows the measured crack propagation rates across the surface plotted against computed ΔK_{II} for crack face friction coefficient values of 0, 0.25 and 0.5. It is apparent that reduced crack face friction increases ΔK_{II} significantly and would therefore promote crack growth. A similar observation was made by Bower [36] who demonstrated the significance of crack face friction on Mode II stress intensify factors. Therefore, factors promoting an increase in crack face friction such as crack face interlocking and the rough fractured surface inside the crack could hinder crack growth. In contrast, entry of lubricant into the crack may reduce crack face friction and hence promote crack growth. However, the plots for all friction coefficients remain straight lines on a log–log plot within Stage II propagation indicating that Paris Law relationship still holds. The stress intensity exponent, m, also remains the same regardless of the crack face friction coefficient. Additionally, ΔK_{II} stress intensity factors for crack face friction coefficient values of 0.25 and 0.5 are relatively similar and remain realistic relative to the suggested fracture toughness of this steel. On the other hand, for zero friction the plot is shifted significantly to the right with the predicted range of stress intensity factors (maximum over 50 MPa $m^{1/2}$) being unrealistically large, much higher than the fracture toughness despite the fact that these cracks have not pitted yet. This suggests that friction between crack faces is indeed non-zero and that a value in the region of 0.5 seems reasonable.

5. Discussion

The experimental and numerical results presented here provide strong evidence that despite the multitude of extra factors present in RCF, LEFM methods can be used to predict the growth of RCF cracks in a similar manner to that commonly used in structural fatigue. Fig. 10 clearly shows that on a log–log plot a linear relationship exists between ΔK_{II} (and ΔK_{eff}) and measured crack propagation rates for cracks of about 100 μm and longer. This means that growth of these cracks can be predicted using the Paris' law relationship stated above. Despite its ubiquitous use in structural fatigue, Paris' law is not routinely used to predict rolling contact fatigue crack growth. Its application to RCF cracks is of obvious practical significance as it makes it possible to apply the so-called damage tolerance methods to dealing with RCF failures in machine elements. For example, if a crack of a certain length is detected on a surface of a rolling bearing ring or a gear tooth, the Paris' law relationship makes it possible to predict the number of cycles before such a crack will grow to a critical length which will result in a pit formation and hence the ultimate failure of the component. Therefore, with the knowledge of the critical crack length prior to pit formation for a given loading and material, this approach makes it possible to predict remaining useful life of a component once a crack is detected.

Two points described earlier are worth noting here again. All crack propagation rates quoted here relate to measured crack growth across the surface, rather than into the surface. The crack is only visible on the

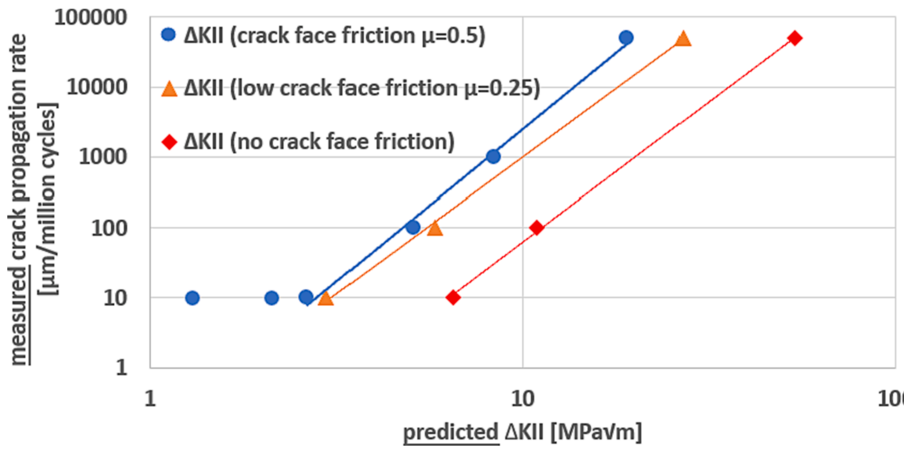

Fig. 11. Plot of measured crack propagation rates across the surface against computes ΔK_{II} for crack face friction coefficients of 0, 0.25 and 0.5. Crack characteristics and measured propagation rates as shown in Table 2.

surface so this is the only growth rate that can be reliably measured throughout a test. However, if the usual assumption of a half-penny shaped crack is made then the subsurface crack length is directly proportional to the surface length at all times and the same relationship can be used for subsurface crack growth, albeit with a different constant of proportionality. The relatively high Hertz pressure used in this study (4.7 GPa) is likely to make the crack initiation phase relatively short so that RCF lives are dominated by propagation. The test was deliberately set-up in this way as we were not studying crack initiation so aimed to make this as short as possible. For lower contact pressures, the relative length of the crack initiation and propagation phases may be different.

The value of the Paris' law stress exponent, *m*, predicted here is 4.3. Given the inherent uncertainties in RCF, this compares relatively well to the value of 3, quoted by Beswick [52] from their fracture tests with compact tension specimen with the same AISI 52100 bearing steel. The nature of the applied stress fields in the two investigations is completely different so the apparent similarity in the crack growth rates suggests that ultimately, the same physical mechanisms are responsible for crack propagation in both RCF and structural fatigue. For the sake of comparison, in a previous study [39], the authors showed that if the maximum Hertz pressure is used in the Paris' type law, instead of the actual stress intensity factor ΔK as is done here, the stress exponent of Hertz pressure is around 7.5.

The predicted range of stress intensity values presented here provides further evidence that LEFM can be used to describe the behaviour of Stage II cracks. The largest value of ΔK_{II} (or ΔK_{eff}) predicted here is about 19 MPa$\sqrt{\mathrm{m}}$; this is for the longest crack considered where surface crack half-length was 571 μm (Case 6 in Table 2). This is similar to the value of fracture toughness previously reported for this AISI52100 bearing steel which was in the range of 15–20 MPa$\sqrt{\mathrm{m}}$ [15,52]. From the experimental observations we know that none of these cracks had yet resulted in pitting but that the longest one was close to causing a pit (based on pit formation in other tests observed with cracks of similar length). Therefore, the predicted maximum ΔK_{II} value of 19 MPa for a crack that is about to pit appears realistic, being close to the fracture toughness of the material. In addition, the predicted stress intensity values for all other cracks, which were much further away from the onset of pitting, were lower than the reported fracture toughness of this steel, as should be the case.

The results show that Stage II rolling contact fatigue cracks primarily propagate under mode II. Mode I and mode III stress intensity factors were predicted to be much smaller, an order of magnitude less for the longest crack considered here (a=570 μm). In the present work, significant Mode I stress intensity factors are only predicted when the load is just ahead of the crack mouth; at this location the applied surface traction (with the roller being slower) acts to open the crack mouth and hence increase mode I stress intensity. The relative dominance of Mode II predicted here is in line with other numerical studies of RCF cracks [29–38] but is here confirmed by relating numerical predictions to experimental crack growth measurements. The analysis does not account for any potential effects of lubricant entrapment in the surface braking cracks and subsequent pressurisation during over-rolling as considered elsewhere [29–32,36,53]. If such effects were to occur, they would inevitably increase the magnitude of mode I stress intensity during over-rolling. The additional effect of lubricant entering the crack is to reduce friction coefficient between the crack faces. This effect was explored here and it was shown that lower crack face friction significantly increases mode II stress intensity. This shifts the plot of crack propagation rates against ΔK_{II} to the right but Paris' law relationship remains valid with approximately the same value of the stress exponent, *m*. The range of values of stress intensity factors predicted here is in line with the observed crack behaviour despite the fact that the potential effects of lubricant pressurisation were not included. This may be because under the present test conditions of high contact stress and low lambda any additional effects of lubricant pressurisation are relatively insignificant. In any case, further research is needed to establish the relative importance of the two potential effects of lubricant entry into the crack namely, the decrease in crack face friction (and subsequent increase in ΔK_{II}) as already included here, and lubricant pressurisation (and subsequent increase in ΔK_I). This is outside the scope of this paper but is currently being explored by the authors.

For Stage I cracks (i.e. those with surface length shorter than about 100 μm) no correlation between computed stress intensity factors and measured crack propagation rates was established. This does not necessarily mean that LEFM methods are not applicable for these short cracks; instead it is most likely because stress intensity in this region is strongly affected by additional factors not included in the present crack model such as asperity stresses, specific steel microstructure and lubricant pressurisation. Asperity stresses are likely of great importance to crack propagation for such short surface braking cracks as the crack tip is still near the surface and hence within the asperity-influenced zone. As the crack grows deeper into the material, it enters the region where it is dominated by the Hertz macro contact stresses which were included in the present model. As discussed by Rycerz et al [39], this difference in the stress fields in the near-surface and deeper regions is probably the very reason for the existence of Stage I and Stage II crack growth and the distinct differences in crack behaviour in the two regions. The transition from Stage I to Stage II crack propagation is here seen to occur at values of ΔK_{eff} of around 5 MPa$\sqrt{\mathrm{m}}$. This value is similar to previously reported threshold stress intensity factor range for this type of steel extracted from fracture mechanics experiments which was around 3–5 MPa$\sqrt{\mathrm{m}}$ [52]. Whether or not this value is of any significance as a threshold stress intensity under rolling contact requires further investigation, in particular in relation to the influence of surface roughness given the strong influence of the depth and magnitude of asperity stress fields on stage I crack propagation as discussed above and in [39]. Large part of the overall life to failure is consumed by crack propagation within the Stage I region, so a better understanding of the mechanisms at play within this region is important if we are to devise better tools for prediction of total RCF lifetimes.

6. Conclusions

In this work we use experimental and numerical methods to show that linear elastic fracture mechanics principles that are commonly used in structural fatigue, can be used to predict propagation of surface braking rolling contact fatigue cracks. First, a triple disc contact fatigue machine is used to generate a series of rolling contact fatigue (RCF) cracks in AISI 52100 bearing steel roller specimens under rolling-sliding, lubricated contact. Crack propagation rates across the surface are measured for RCF cracks of different lengths by optical inspection of the roller specimens at set intervals during the test. Crack morphology is characterised by sectioning the specimens at the end of the test. Secondly, a numerical finite element model of the crack under rolling contact is devised to compute stress intensity factors for the simplified crack geometry observed in the experiments under the same normal and tangential loading as applied in the tests. The model predicts the evolution of all three stress intensity factors (K_I, K_{II}, and K_{III}) as a Hertz pressure distribution passes over the surface braking crack. The model utilises a fracture mechanics mesh with quarter point nodes and six contours to successfully resolve the stress fields at the crack tip. It includes the effects of contact pressure, surface tractions and crack face friction but excludes any potential effects of lubricant pressurisation in the crack during over-rolling. Finally, the stress intensity factors predicted in this manner are related to the measured crack propagation rates across surface to investigate if standard LEFM relationships hold. The main findings are:

- Linear elastic fracture mechanics principles can be used to predict the growth of rolling contact fatigue cracks in a similar manner commonly used in structural fatigue.
- The plot of measured crack propagation rates against computed ΔK_{II} stress intensity is a straight line on a log–log plot for all cracks with surface lengths larger than about 100 μm (referred to as Stage II RCF cracks in [39]). Consequently, RCF crack propagation rates in this region can be described by a standard Paris' law relationship of the form:

$$\frac{dc}{dN} \propto \frac{\mu m}{10^6}\left(\frac{\Delta K_{II}}{MPa\sqrt{m}}\right)^{4.3}$$

- The predicted exponent of stress intensity factor in this Paris' law equation of 4.3 is comparable to that extracted from standard fracture mechanics tests in [50].
- The maximum ΔK predicted for the observed cracks is 19 MPa $m^{1/2}$ (for the longest crack that was close to pitting) which seems realistic given that the reported fracture toughness of this steel is 15–20 MPa $m^{1/2}$
- ΔK_{II} (Mode 2 propagation) was found the be the main driver for RCF crack growth for longer cracks (over ~ 100 μm). ΔK_I and ΔK_{III} are up to an order of magnitude smaller.
- For very short cracks of less than ~ 100 μm (Stage I cracks in [39]), no correlation between the stress intensity factors computed in the described manner and crack propagation was observed. This is likely to be due to the exclusion of additional effects in the present model that are crucial to growth of such short cracks, not least the near-surface asperity stresses.
- In the present tests, the transition from short crack behaviour (Stage I propagation) to long crack behaviour (Stage II propagation) occurs at surface crack length of about 100 μm. This corresponds to effective stress intensity of around 5 MPa$\sqrt{m}$ under present assumptions.
- Decreasing the crack face friction significantly increases the magnitude of ΔK_{II} stress intensity. This is therefore one important mechanism by which lubricant entry into a crack may accelerate the propagation of surface braking cracks.
- The findings presented here can help in devising tools for prediction of remaining useful life of a machine component, such as a rolling bearing or a gear, once the first surface crack is detected. Combination of such prediction tools with suitable condition monitoring methods offers a powerful way of improving the reliability and maximising the uptime of mechanical systems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to acknowledge the financial support for this work provided by SKF through the SKF University Technology Centre for Tribology at Imperial College London.

References

[1] Lundberg G, Palmgren A. Dynamic Capacity of Roller Bearings. Acta Poloytechnica Mech. Eng. Ser. 1947;1:7–53.
[2] Dawson BPH. Effect of metallic Contact on the pitting of lubricated rolling surfaces. J. Mech. Eng. Sci. 1962;4(I):16–21.
[3] Dawson PH. The effect of metallic contact and sliding on the shape of the S-N curve for pitting fatigue. Instn Mech Engrs 1963:41–5.
[4] Nélias D, et al. Role of inclusions, surface roughness and operating conditions on rolling contact fatigue. J Tribol 1999;121(2):240.
[5] Ueda T, Mitamura N. Mechanism of dent initiated flaking and bearing life enhancement technology under contaminated lubrication condition. Part II: Effect of rolling element surface roughness on flaking resulting from dents, and life enhancement technology of rolling bearings. Tribol Int 2009;42(11–12):1832–7.
[6] Ueda T, Mitamura N. Mechanism of dent initiated flaking and bearing life enhancement technology under contaminated lubrication condition. Part I: Effect of tangential force on dent initiated flaking. Tribol Int 2009;42(11–12):1832–7.
[7] Dawson BPH. Further experiments on the effect of metallic contact on the pitting of lubricated rolling surfaces. Proc. Inst. Mech. Eng. 1965;180(I):95–112.
[8] Tallian TE, McCool JI, Sibley LB. Partial elastohydrodynamic lubrication in rolling contact. In: Arch. Proc. Inst. Mech. Eng. Conf. Proc. 1964-1970 (vols 178-184), Var. titles Label. Vol. A to S, vol. 180, no. 1965; 1965. p. 169–86.
[9] Danner CH. Fatigue life of tapered roller bearings under minimal lubricant films. ASLE Trans 1970;13(4):241–51.
[10] Littmann W. Inderdisciplinary approach to the lubrication of concentrated contacts. In: Proc. a Symp. Troy, NY, July 1969; 1970. p. 309–78.
[11] Onions RA, Archard JF. Pitting of gears and discs. Proc Inst Mech Eng 1974;188:673–82.
[12] Liu JY, Tallian TE, Mc Cool JI. Dependence of bearing fatigue life on film thickness to surface roughness ratio. ASLE Trans 1975;18(2):144–52.
[13] Phillips MR, Quinn TFJ. The effect of surface roughness and lubricant film thickness on the contact fatigue life of steel surfaces lubricated with a sulphur-phosphorus type of extreme pressure additive. Wear 1978;51(1):11–24.
[14] Soda N. Effect of tangential traction and roughness on crack initiationipropagation during rolling contact. ASLE Trans 1982;25(2):198–206.
[15] Bhadeshia HKDH. Steels for bearings. Prog Mater Sci 2012;57(2):268–435.
[16] Tallian TE. On competing failure modes in rolling contact. ASLE Trans 1967;10(4):418–39.
[17] Morales-Espejel GE, Brizmer V. Micropitting modelling in rolling–sliding contacts: Application to rolling bearings. Tribol Trans 2011;54(4):625–43.
[18] Morales-Espejel GE, Rycerz P, Kadiric A. Prediction of micropitting damage in gear teeth contacts considering the concurrent effects of surface fatigue and mild wear. Wear 2018;398–399(November):99–115.
[19] Olver AV. The mechanism of rolling contact fatigue: An update. Proc Inst Mech Eng Part J J Eng Tribol 2005;219(5):313–30.

[20] Rycerz P, Kadiric A. The influence of slide-roll ratio on the extent of micropitting damage in rolling-sliding contacts pertinent to gear applications. Tribol Lett 2019;67(2):1–20.
[21] Ueda M, Spikes H, Kadiric A. In-situ observations of the effect of the ZDDP tribofilm growth on micropitting. Tribol Int 2019;138(March):342–52.
[22] Sadeghi F, Jalalahmadi B, Slack TS, Raje N, Arakere NK. A review of rolling contact fatigue. J Tribol 2009;131(4):041403.
[23] Arakere NK, Subhash G. Work hardening response of M50-NiL case hardened bearing steel during shakedown in rolling contact fatigue. Mater Sci Technol 2012;28(1):34–8.
[24] Arakere NK. Gigacycle rolling contact fatigue of bearing steels: A review. Int J Fatigue 2016;93:238–49.
[25] Ioannides E, Harris TA. A new fatigue life model for rolling bearings. ASME J Lubr Technol 1985;107(84):367–78.
[26] “ISO 281: 2007 Rolling bearings - Dynamic load ratings and rating life.”.
[27] Jalalahmadi B, Sadeghi F. A Voronoi finite element study of fatigue life scatter in rolling contacts. J Tribol 2009;131(2):1–15.
[28] Lorenz SJ, Sadeghi F, Trivedi HK, Rosado L, Kirsch MS, Wang C. A continuum damage mechanics finite element model for investigating effects of surface roughness on rolling contact fatigue. Int J Fatigue 2021;143(September 2020).
[29] Murakami Y, Kaneta M, Yatsuzuka H. Analysis of surface crack propagation in lubricated rolling contact. ASLE Trans 1985;28(1):60–8.
[30] Kaneta M, Yatsuzuka H, Murakami Y. Mechanism of crack growth in lubricated rolling/sliding contact. ASLE Trans 1985;28(3):407–14.
[31] Kaneta M, Murakami Y. Effects of oil hydraulic pressure on surface crack growth in rolling / sliding contact. Tribol Int 1987;20(4):210–7.
[32] Kaneta M, Murakami Y. Propagation of semi-elliptical surface cracks in lubricated rolling/sliding elliptical contacts. J Tribol 1991;113(2):270.
[33] Keer LM, Bryant MD, Haritos GK. Subsurface and surface cracking due to hertzian contact. J Lubr Technol 1982;104(3):347.
[34] Miller GR, Keer LM, Cheng HS. On the mechanics of fatigue crack growth due to contact loading. Proc R Soc A Math Phys Eng Sci 1985;397(1813):197–209.
[35] Hanson MT, Keer LM. An analytical life prediction model for the crack propagation occurring in contact fatigue failure. Tribol Trans 1992;35(3):451–61.
[36] Bower AF. The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks. Trans ASME 1988;110:704–11.
[37] Keer LM, Bryant MD. A pitting model for rolling contact fatigue. 1982;105(April):198–205.
[38] Zhou RS, Cheng HS, Mura T. Micropitting in rolling and sliding contact under mixed lubrication. Am Soc Mech Eng 1989;111(October):605–13.
[39] Rycerz P, Olver A, Kadiric A. Propagation of surface initiated rolling contact fatigue cracks in bearing steel. Int J Fatigue 2017;97:29–38.
[40] Nakajima A, Kazunori I, Hirano F, Nishimura M. Effects of combination of rolling direction and sliding direction on pitting of rollers. JSLE 1983:93–8.
[41] Zhou RS, Cheng HS, Mura T. Micropitting in rolling and sliding contact under mixed lubrication. J Tribol 1989;111:605–13.
[42] Benyajati C, Olver AV, Hamer CJ. An experimental study of micropitting, using a new miniature test-rig. Tribol Ser 2003;43:601–10.
[43] Manieri F, Stadler K, Morales-Espejel GE, Kadiric A. The origins of white etching cracks and their significance to rolling bearing failures. Int. J. Fatigue 2019;120 (October 2018):107–33.
[44] Olver AV. Wear of hard steel in lubricated, rolling contact, PhD Thesis, Imp. Coll. London; 1986.
[45] Voskamp AP. Material response to rolling contact loading. 1984;107(84): 359–64.
[46] Schlicht H, Schreiber E, Zwirlein O. Fatigue and failure mechanism of bearings. Fatigue Eng Mater Struct 1986:85–90.
[47] Levesque GA, Arakere NK. Critical flaw size in silicon nitride ball bearings. Tribol Trans 2010;53(4):511–9.
[48] Edition T, Anderson T. L., Fracture mechanics - fundamentals and applications, 3rd ed., 2004. https://doi.org/10.1007/978-1-4612-1740-4.
[49] Elber W. The significance of fatigue crack closure. Damage Toler Aircr Struct 1971:230–42.
[50] Meray F, Chaise T, Gravouil A, Depouhon P, Descharrieres B, Nélias D. A novel SAM/X-FEM coupling approach for the simulation of 3D fatigue crack growth under rolling contact loading. Finite Elem Anal Des 2022;206(February):103752.
[51] Wong SL, Bold PE, Brown MW, Allen RJ. Fatigue crack growth rates under sequential mixed-mode I and II loading cycles. Fatigue Fract Eng Mater Struct 2000;23(8):667–74.
[52] Beswick J. Fracture and fatigue crack propagation properties of hardened 52100 steel. Metall Trans A 1989;20A.
[53] Balcombe R, Fowell MT, Olver AV, Ioannides S, Dini D. A coupled approach for rolling contact fatigue cracks in the hydrodynamic lubrication regime: The importance of fluid/solid interactions. Wear 2011;271(5–6):720–33.