**Final Technical Report
COVER PAGE**

Federal Agency to which Report is submitted: DOE EERE – Wind & Water Power Program

Recipient: The Ohio State University, DUNS #: 071650709

Award Number: DE-EE0002735

Project Title: An Experimental and Theoretical Investigation of Micro-pitting in Wind Turbine Gears and Bearings

Project Period: Jan 1, 2010, Dec 31 2011

Principle Investigator: Ahmet Kahraman, Professor, kahraman.1@osu.edu, 614-292-4678

Report Submitted by: Ahmet Kahraman, Professor, kahraman.1@osu.edu, 614-292-4678

Date of Report: March 28, 2012

Covering Period: January 1, 2010 to December 31, 2011

Working Partners: Prof. Donald R. Houser (Co-PI), houser.4@osu.edu, 614-292-5860

Dr. Sheng Li, li.600@osu.edu, 614-247-8688

Cost-Sharing Partners: None

DOE Project Team: DOE HQ Program Manager – Jose Zayas
DOE Field Contract Officer – Pamela Brodie
DOE Field Grants Management Specialist – Jane Sanders
DOE Field Project Officer – Nick Johnson
DOE/CNJV Project Monitor – Shaun Jensen

Signature of Submitting Official: Ahmet Kahraman

1. Executive Summary

In this research study, the micro-pitting related contact failures of wind turbine gearbox components were investigated both experimentally and theoretically. On the experimental side, a twin-disk type test machine was used to simulate wind turbine transmission contacts in terms of their kinematic (rolling and sliding speeds), surface roughnesses, material parameters and lubricant conditions. A test matrix that represents the ranges of contact conditions of the wind turbine gear boxes was defined and executed to bring an empirical understanding to the micro-pitting problem in terms of key contact parameters and operating conditions. On the theoretical side, the first deterministic micro-pitting model based on a mixed elastohydrodynamic lubrication formulations and multi-axial near-surface crack initiation model was developed. This physics-based model includes actual instantaneous asperity contacts associated with real surface roughness profiles for predicting the onset of the micro-pit formation. The predictions from the theoretical model were compared to the experimental data for validation of the models. The close agreement between the model and measurements was demonstrated. With this, the proposed model can be deemed suitable for identifying the mechanisms leading to micro-pitting of gear and bearing surfaces of wind turbine gear boxes, including all key material, lubricant and surface engineering aspects of the problem, and providing solutions to these micro-pitting problems.

This research work was related to the mission of *The Office of Energy Efficiency and Renewable Energy* of Department of Energy under *Wind and Hydropower Technologies Program*. Specifically, it fell under *the topic area #1: Supporting Wind Turbine Research and Testing*.

2 Problem Description and Background

Micro-pitting is a progressive rolling contact fatigue wear phenomenon occurring on the micro-scale induced by severe localized stress concentrations at or very near the contact surfaces. As two contacting rough surfaces slide against each other, the roughness profiles often interact to produce instantaneous peaks of normal pressure and surface shear. These time-varying surface tractions result in large multi-axial stress amplitudes for the material points within a very shallow layer of material (typically less than 2 μm in depth). The localized cyclic loading, which can be several times higher than the corresponding Hertzian pressure, dictates the fatigue failure in the form of micron size pits. These small pits can either be clustered together or scattered over the surface, depending on the surface roughness texture. As the number of contact cycles increases, the amount of micro-pits grows. To some extent, the scattered light reflections from the numerous micro-pits display grey-colored appearance, and has often been referred as grey staining or frosting. Under the condition that the two surfaces are well separated by a layer of fluid film, the hydrodynamic pressure and viscous shear can still be raised to relatively high levels when the surfaces are relatively rough, leading to the occurrence of micro-pitting, such as observed in Ref. [1]. This research project aims at investigating the micro-pitting fatigue phenomenon observed in wind turbine gear boxes through a novel predictive model and accelerated contact experiments.

In gearing applications, the wear of micro-pitting can largely increase the transmission error and lead to the undesirable noise and vibration. The dynamic response also amplifies the tooth contact force, further accelerating the rate of micro-pitting. Under certain operating conditions, the amount of micro-pits stabilizes after the surface wear redistributes and relieves the contact pressure. However, the continued cyclic contact can result in the failure of macro-pitting

which initiates from the micro-pit cracks. In order to avoid such failures, extensive experimental studies have been conducted in literature, investigating the influences of various potential factors on micro-pitting. Ariura et al [2] studied the roughness effect on micro-pitting for gear contacts, confirming the important role of surface roughness in the micro-pitting phenomenon. Using a twin-disk set-up, Webster and Norbart [3] showed that micro-pits tended to appear on the surface with negative sliding. Through the reduction of slide-to-roll ratio, they successfully eliminated micro-pitting. It was also found that although increasing the lambda-ratio (the ratio of film thickness to the roughness amplitude) through increasing the film thickness had limited effects on micro-pitting reduction, the alternative which increases the lambda-ratio via reducing the roughness amplitude effectively eliminated the micro-pitting failure. The work of Ahlroos et al [4] investigated the influences of different steel materials, surface roughness amplitudes, surface treatments (hardening and coating) as well as lubricants on micro-pitting. The effects of lubricant additives were studied in the works such as [5,6], showing the anti-wear (AW) and extreme-pressure (EP) additives aggravated micro-pitting [5], while the friction modifier agent alleviated the occurrence of micro-pits through reducing boundary friction [6]. Due to the environmental concerns, biodegradable lubricant has also been tested for its anti-pitting performance [7].

Although the link between the surface roughness and the resultant micro-pitting performance has been well established experimentally, published contact fatigue models for rough surface contacts commonly used the potential theory based half-space formulation which assumes perfectly smooth surface [8] to calculate the stress fields [9-13]. This smooth surface assumption is reasonable for sub-surface initiated macro-pitting failures, while it is not suitable for surface or near surface initiated micro-pitting failures. The half-space displacement and stress formulations [8] have also been widely used in the rough surface contact models such as [14-16], assuming the difference of roughness height between any two points on the surface is far smaller than the

distance between them [14]. This is often not the case for machine contacts where the surface roughness due to the finishing processes such as grinding and shaving is rather significant. Even for highly polished surfaces, this assumption is not necessarily valid for the neighboring nodes on a very fine mesh grid. The contact model of Ref. [17] used the finite element (FE) approach to determine the apparent contact area and the mean contact pressure between two mating rough surfaces. Because the detailed description of the surface topography using FE based models requires an extremely fine FE grid with a very large number of elements which lead to rather unaffordable computational time, a semi-detailed approach was employed in [17], excluding the predictions of the localized stress concentrations.

The alternative to the finite element method is the boundary element method (BEM) which considers only the boundary instead of the entire contact volume, reducing the three-dimensional problem to a two-dimensional problem, such that a fine mesh is possible. Some BEM studies focused on the contacts between smooth components using finite boundary elements [18,19], some considered the smooth half-space contact problem employing infinite boundary elements [20-24]. A contact model that fully describes the surface topography variations for engineering rough surfaces is not available in the literature.

2.1 Project Objective, Goals and Specific Tasks

The project objective is to investigate the micro-pitting of wind turbine gearbox components both experimentally and theoretically, with the aim of bringing a fundamental understanding to the micro-pitting phenomenon observed in gear and bearing components of wind turbine gearboxes. The net effect of this effort is aimed at improving the reliability of wind

turbine power trains. The stated project objective requires that the following technical project goal must be achieved.

- (i) Development of an accelerated experimental procedure to consistently produce micro-pitting under simulated conditions that are representative of those experienced by gears and bearings of wind turbine gearboxes.
- (ii) Establishment of a fundamental baseline micro-pitting database and quantify the impact of run-in procedures on this baseline micro-pitting behavior. Describe the impact of key contact surface and lubricant variations on this baseline micro-pitting behavior through tightly-controlled tests.
- (iii) Development of a physics-based micro-pitting simulation model and perform parametric studies to define design guidelines for extending micro-pitting lives of wind turbine gears and bearings.
- (iv) Validation of the micro-pitting model through comparisons between the experimental database and the test simulations.

In view of these stated objectives and technical goals, an experimental study was designed and conducted to collect statistically meaningful micro-pitting data using a two-disk rolling contact fatigue test machine. Section 3 below describes the details of the micro-pitting experiments including the test machine, specimens, test matrix and experimental results.

Section 4 proposes a new physics-based contact fatigue model for micro-pitting, incorporating the local roughness geometry in the near surface stress predictions. The boundary element approach is used in the process for the detailed description of the surface topography with affordable computational efforts. Since the contact zone is usually small for the mechanical components such as bearings and gears, infinite half-space is assumed. The boundary is divided

into a finite region which is rough and an infinite region which is assumed to be perfectly smooth. Finite boundary elements are used to mesh the former and infinite boundary elements are used to discretize the latter. The stress distributions along the surfaces are determined from the local strains through the Hooke's law such that the singularities of the integral kernels can be circumvented. For the near surface stress fields, a method that combines the coordinate transformation and the progressive element subdivision is developed to minimize the numerical error caused by the near singular behavior of the kernel functions when the load and the field points are too close to each other. The BEM formulation results in terms of the displacements and stresses are compared to the closed-form solutions to demonstrate the model accuracy. In the process, a statistical measure for the micro-pitting failure, called Micro-pitting Severity Index, is proposed in Section 4.

In Section 5, the tests presented in Section 3 are simulated using the model proposed in Section 4. Direct comparisons between the model and the experiments are provided to assess the accuracy of the model. The impacts of key contact parameters and operating and run-in conditions on micro-pitting are also described through the measurements and predictions.

3 Experimental Study of Micro-pitting

3.1 Test Machine

Due to the budget constraints, an existing twin-disk set-up was employed in this study to evaluate the influences of various potential factors, including contact pressure, rolling and sliding velocities, surface roughness amplitude and run-in process, on the fatigue failure of micro-pitting. Figure 1(a) shows the test machine and Figure 1(b) provides a view of the test chamber, which shows two cylindrical disks in contact. Here, each disk is driven by an independent AC motor at the rotational speeds of ω_1 and ω_2 . Given d_1 and d_2 as the diameters of the disks, the

Figure 1. (a) The roller test machine to be used for the micro-pitting tests, and (b) a close-up view showing a pair of specimens mounted on the test machine.

tangential surface velocities of $v_1 = \frac{1}{2}d_1\omega_1$ and $v_2 = \frac{1}{2}d_2\omega_2$ can be adjusted through ω_1 and ω_2 to create relative sliding at the contact. Defining the rolling (mean) and sliding velocities of the contact, respectively, as

$$v_r = \frac{1}{2}(v_1 + v_2), \quad (1a)$$

$$v_s = v_1 - v_2 \quad (1b)$$

The relative sliding at the contact interface can be defined by the dimensionless sliding parameter, called slide-to-load ratio, as

$$SR = \frac{v_s}{v_r} = \frac{2(v_1 - v_2)}{v_1 + v_2}. \quad (1c)$$

The test set-up shown in Figure 1 can be used to perform contact tests at given v_r and SR values representative of gear and bearing contacts of wind turbine gearboxes.

In Figure 1(b), the smaller one of the two test disks (called *roller* here) is shrink-fitted onto its shaft and installed into the pivoted loading arm which is pushed against the larger disk by a pneumatic cylinder. The larger disk (called *disk*) is fastened axially against its shaft shoulder by using a retaining lock nut. The lubricant that goes into the mesh of the contact is provided through the overhead lubrication jet.

3.2 Definition of Contact Conditions Representative of Wind turbine Gearboxes

The machine shown in Figure 1 can accommodate disk specimens of different sizes and geometries and can operate them within ranges of speeds and contact stresses. It is capable of

achieving desired values of v_r and SR within the ranges specified in Table 1. In addition, the machine is capable of producing a normal contact force up to 4450 N, which corresponds to the maximum Hertzian contact pressures up to 2.5 GPa as specified in Table 1.

In view of the capabilities of the test machine, definition of specimen dimensions and operating conditions representative of the wind turbine gearbox gear or bearing contacts was the first main task in this project. For this purpose, an entire wind turbine gear box, the NREL test wind turbine gearbox, was modeled at the beginning of this project as a typical configuration representative of most wind turbines. As shown in Figure 2, it consists of a planetary stage followed by two speed increaser stages. The gearbox has a total power capacity of 750 kW with an overall speed ratio of 1:81.491. This ratio allows a nominal input rotor speed of 22 rpm to be increased to an output speed of 1800 rpm, which is required by the generator to supply synchronous power to the electrical grid. The wind turbine can also operate at a second nominal input speed of 15 rpm (output speed of 1200 rpm) by way of the generator actively switching the number of magnetic poles used to generate electricity. A 323 kN-m input torque created by the wind on the rotor blades is required to operate the generator at fully rated power and maximum speed.

As described in the cross-sectional layout of Figure 3, the input shaft from the rotor is attached to the planet carrier providing input torque to the gearbox. The planetary set consists of a ring gear bolted in between the housing halves, three planets and a floating sun gear. The planetary set, known as the low speed stage, has a ratio of 1:5.74 and is connected to the immediate speed stage through a spline on the sun gear. The intermediate speed stage has a ratio of 1:3.57 with the pinion located on a common shaft with the high speed stage gear. The high speed shaft has a ratio of 1:4.00 and is the output shaft of the gearbox. The output shaft is connected to the generator through a coupling on the generator's input shaft.

Table 1. Contact Parameter ranges achievable by the two-disk test machine.

Parameter	Symbol	Unit	Range
Maximum contact pressure	P_h	GPa	0 to 2.5
Rolling speed	v_r	m/s	0 to 10
Slide-to-roll ratio	SR	--	-1.0 to 1.0
Inlet oil temperature	--	$^{\circ}C$	40 to 120

Figure 2. An assembly view of the example wind turbine gearbox (image from NREL Gearbox Reliability Collaborative)

Figure 3. Cross-sectional layout of the example wind turbine gearbox with bearing locations identified with abbreviations (image from NREL Gearbox Reliability Collaborative).

The NREL wind turbine gearbox was modeled using a three-dimensional finite element based contact analysis software called *Transmission3D* (Advanced Numerical Solutions, Inc.) that was developed specifically for simulation of multi-mesh gear trains. The software permits the user to model the entire gearbox including the housing, gears, bearings, shafts, and planet carrier. Transmission3D first computes the finite element deflections and forces, and then employs a semi-analytical approach to analyze the contacts at gear meshes bearings and splines. The gearbox model shown in Figure 4 includes all critical components that are necessary to accurately predict the load distribution and stresses in the critical components. The model allows evaluation of the gear and bearing load distributions in presence of system-level effect such as deflections and misalignment that are critical to predicting the micro-pitting. It also includes the intended gear tooth modifications such as profile and lead corrections.

The planet gear bearings of the NREL gearbox were modeled as rolling elements in contact with their respective bearing races as shown in Figure 5. All other bearings within the system were modeled as diagonal stiffness matrices with stiffness values taken from experimental data. The complete model was solved under quasi-static conditions for all the forces, displacements and stresses present throughout the system.

The majority of micro-pitting on the example gearbox was reported to occur occurs within the low speed planetary stage as it experiences very high torque levels at relatively low speeds. The planet gears and their supporting bearings were also identified as very susceptible to micro-pitting because of this extreme loading as well as possible misalignment due to non-torque loads on the rotor blades caused by the wind.

Based on the analysis of the example wind turbine gearbox, the gear contact parameters were determined as listed in Table 2(a) at an output speed of 1800 rpm. The gear speeds are given

Figure 4. The deformable-body model of the example wind turbine gearbox; (a) the finite element mesh, and (b) predicted stress contours.

Figure 5. Predicted stress contours of (a) the planetary gears, and (b) planet bearings, and two-stage increaser.

in Table 1(a) at the Start of Active Profile (SAP) and End of Active Profile (EAP) where the sliding velocities are extreme. The respective rolling and sliding velocities as well as slide-to-roll ratios are also listed in Table 2(a). Likewise, Table 2(b) lists the bearing speeds of the same gearbox using the nomenclature defined in Figure 3. Table 3 summarizes the contact parameters of all four gear meshes of the example wind turbine gearbox, namely a sun-planet mesh, a ring-planet mesh, the high-speed increaser gear pair, and the low-speed increaser gear pair. It also provides the average film thickness-to-composite roughness ratios indicating that these gear contacts indeed experience metal-to-metal contacts.

3.3 Development and Procurement of the Test Specimens

Based on the constrained of the two-disk test machine and the results of the analyses of the example wind turbine gearbox, a pair of disks was designed to serve as specimens in this project. Figure 6 shows technical drawings of these specimens. The smaller of the two disks (roller), has a pure cylindrical shape with no lead crown while the larger one (disk) has a circular lead crown at a radius of 76.2 mm (3 in). The diameters of the roller and the disk are $d_1 = 31.75$ mm (1.25 in) and $d_2 = 57.15$ mm (2.25 in). This results in $SR = -0.57$ when $\omega_1 = \omega_2$. With this specimen design, the operating roller machine conditions listed in Table 4 represent the gear mesh contact conditions of Table 3.

Several machining companies were contacted and a vendor to fabricate the test specimens was identified. A batch of 65 pairs of specimens to the specifications provided in Figure 6 has been procured. Figure 7(a) shows a pair of test specimens. These specimens are made out of AISI 4620 gear steel. Each specimen was case hardened to 60 HRC to represent a typical gear tooth surface hardness. A finishing process was developed to simulate the direction and lay of the

Table 2. Summary of the contact parameters of the example wind turbine gearbox; (a) gear parameters, and (b) bearing parameters.

(a) Gear Parameters

	Ring(2)	Planet(1)	Planet(2)	Sun(1)	IMS-Gear(2)	IMS-Pinion(1)	HS-Gear(2)	HS-Pinion(1)
Input Speed (rpm)	26.8	68.0	68.0	126.2	126.2	449.9	449.9	1799.7
Radius of Curvature at SAP (m)	0.2186	0.0532	0.0544	0.0196	0.1101	0.0200	0.0707	0.0114
Radius of Curvature at EAP (m)	0.1723	0.0995	0.0995	0.0647	0.1478	0.0567	0.0926	0.0333
Rolling Velocity at SAP of (1) (m/s)	-0.10		-0.45		-1.01		-2.22	
Rolling Velocity at EAP of (1) (m/s)	0.10		0.47		1.22		2.94	
Sliding Velocity at SAP of (1) (m/s)	0.43		0.48		1.45		3.25	
Sliding Velocity at EAP of (1) (m/s)	0.66		0.66		0.66		0.66	
Slide-to-roll Ratio at SAP of (1)	-0.24		-0.93		-0.70		-0.68	
Slide-to-roll Ratio at EAP of (1)	0.14		0.75		0.59		0.61	

(b) Bearing Parameters

	PL-A/B	PLC-A	PLC-B	LS-SH-A	LS-SH-B	LS-SH-C	IMS-SH-A	IMS-SH-B/C	HS-SH-A	HS-SH-B/C
Inner Race Speed (rpm)	0.0	22.0	22.0	126.2	126.2	126.2	449.9	449.9	1799.7	1799.7
Outer Race Speed (rpm)	68.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Inner Race Radius of Curvature (m)	0.096	0.244	0.213	0.148	0.131	0.131	0.059	0.091	0.059	0.067
Outer Race Radius of Curvature (m)	0.131	0.276	0.237	0.167	0.149	0.149	0.081	0.109	0.081	0.088
Roller Radius of Curvature (m)	0.017	0.016	0.012	0.009	0.010	0.009	0.011	0.010	0.011	0.011
Inner Race Velocity (m/s)	0.0	0.6	0.5	2.0	1.7	1.7	2.8	4.3	11.1	12.5
Outer Race Velocity (m/s)	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cage Velocity (m/s)	0.1	0.0	0.0	0.1	0.1	0.1	0.4	0.4	1.7	1.9
Roller Velocity (m/s)	0.8	-0.6	-0.5	-2.1	-1.8	-1.8	-3.2	-4.7	-12.9	-14.3
Rolling Velocity (m/s)	0.9	0.6	0.5	2.0	1.7	1.7	2.8	4.3	11.1	12.5

Table 3. Summary of the EHL analyses of gear contacts of example wind turbine gearbox at contact locations where micro-pitting is most likely. Here $\bar{h}$ is the average film thickness and $S_{eq} = \sqrt{S_1^2 + S_2^2}$.

Gear Pair	u_1 [m/s]	u_2 [m/s]	SR	R_1 [m]	R_2 [m]	P_c [GPa]	$\bar{h}/S_{eq}$
High-speed Increaser	2.35	4.33	-0.59	.0124	.0915	1.48	1.32
Low-speed Increaser	1.07	1.91	-0.56	0.022	0.144	1.36	1.52
Sun-planet	0.28	0.55	-0.65	0.025	0.094	1.13	1.40
Ring-planet	0.41	0.44	-0.067	0.070	0.189	0.577	1.72

Table 4. Roller contact conditions for simulating gear contact conditions of Table 3.

Gear Pair Simulated	u_1 [m/s]	u_2 [m/s]	SR	P_c [GPa]	$\bar{h}/S_{eq}$
High-speed Increaser	1.615	3.0	-0.6	1.5	1.35
Low-speed Increaser	3.41	6.0	-0.55	1.35	1.43
Sun-planet	2.54	5.0	-0.65	1.15	1.32

Figure 6. Technical drawings of test specimens; (a) roller and (b) disk.

Figure 7. (a) A pair of roller specimens made out of AISI 4620 procured for this study, and (b) close up views of disk and roller specimens showing the lay of the surface texture.

actual gear contacts in relation to the direction of sliding. The axial grinding marks shown in Figure 7(b) are direct result of this special finishing process such that not only the amplitudes but the directionality of the gear contact surface roughnesses can be simulated. The surface roughness amplitudes of this batch of specimens are within the range of 0.4-0.5 μm (RMS) that is typical of ground gear surfaces.

3.4 Design of Experiments Test Matrix and Test Procedure

Table 5 lists the test matrix constructed using the fractional factorials technique. This Design of Experiment (DOE) approach allows the statistical meaningful measurements with substantially reduced number of test runs. In order to study the influence of run-in on micro-pitting occurrence, a run-in stage was implemented before each normal test stage. The run-in process consisted of 0.2 million roller contact cycles and a normal test stage lasted 20 million roller contact cycles. Both run-in and test stages used the same type of lubricant (A typical wind turbine gear fluid, Castrol Optigear Synthetic X320 at 95 deg. C).

Depending on the roller speed of each test, the testing time ranged from days to weeks. During each test, the machine was paused periodically for the inspection of the specimens, which included the circumferential surface roughness measurement using a Talysurf surface roughness profiler, the lead direction profile measurement using a gear Coordinate Measurement Machine (CMM), and the micro-pitting area measurement using a digital microscope. These measurements were performed for three predefined locations that are positioned 120 degrees apart from each other circumferentially. For the run-in stage, the inspection was repeated three times, i.e. before the test, in the middle and after the run-in process. For the normal test stage, the inspection is carried out every 2 million roller contact cycles.

Table 5. Rolling contact fatigue test matrix using fractional factorials.

Test #	P'_h [GPa]	v'_r [m/s]	P_h [GPa]	v_r [m/s]	SR	R_q [μm]	$\bar{\Phi}$ [%]
1	0.6	5.85	1	3.9	-0.2	0.4	0.02
2	1.2	1.95	1.5	3.9	-0.2	0.4	0.07
3	0.8	1.95	1	3.9	-0.65	0.7	0.41
4	0.6	3.9	1	7.8	-0.2	0.7	3.06
5	0.9	5.85	1.5	3.9	-0.65	0.7	28.10
6	0.8	11.7	1	7.8	-0.65	0.4	0.00
7	0.9	3.9	1.5	7.8	-0.65	0.4	0.00
8	1.2	11.7	1.5	7.8	-0.2	0.7	2.71

Six potential parameters were considered in the experimental study:

- (1) Hertzian pressure p_h of the test stage,
- (2) rolling velocity v_r of the test stage,
- (3) slide-to-roll ratio SR of the test and run-in stages,
- (4) Hertzian pressure p'_h of the run in stage,
- (5) rolling velocity v'_r of the run-in stage, and
- (6) initial composite root-mean-square (RMS) surface roughness amplitude

$R_q = \sqrt{R_{q1}^2 + R_{q2}^2}$, where R_{q1} and R_{q2} are the RMS roughness amplitudes for roller and disk, respectively.

According to the operating condition of the 750 kW NREL gearbox, the low and high levels of these factors for the sun-planet gear mesh were defined as:

$$p_h = 1 \text{ and } 1.5 \text{ GPa,}$$

$$v_r = 3.9 \text{ and } 7.8 \text{ m/s,}$$

$$SR = -0.2 \text{ and } -0.65,$$

$$R_q = 0.4 \text{ and } 0.7 \mu\text{m}.$$

For the run-in stage, p'_h and v'_r were determined in relation to respective test stage values as

$$p'_h / p_h = 0.6 \text{ and } 0.8,$$

$$v'_r / v_r = 0.5 \text{ and } 1.5.$$

The inlet lubricant temperature was maintained at 95°C through both run-in and test stages of each test.

3.5 Micro-Pitting Test Results

The last column of Table 5 lists a micro-pitting parameter $\bar{\Phi}$, which represents the average value of the micro-pitting area percentage at the end of 20-million-cycle tests. It is defined as the ratio of the micro-pitted area to the total inspected area. It is found that Test 5, which corresponds to the high levels of p_h , SR , R_q and v_r' while the low levels of p_h' and v_r , resulted in the most severe micro-pitting failure ($\bar{\Phi} = 28.1\%$).

Figure 8 shows the contact surface of the roller (of Test 5) after 4, 12 and 20 million cycles of testing at three different positions. In these images, the micro-pitted areas are highlighted in red. It is seen the severity of micro-pitting is not uniform at different circumferential locations, such that it is necessary to measure Φ at a number of positions and use the average value as the measure. Figures 9 and 10 are the micro-scope images for Test 4 and 5 respectively. The former is an example of small amount of micro-pits ($\bar{\Phi} = 3.1\%$) and the latter is a no-failure example ($\bar{\Phi} = 0.0\%$).

The measured surface roughness profiles for the roller (corresponds to the 1st column of Figure 8) and the disk are shown in Figures 11 and 12, respectively. The roughness amplitude of the roller is observed to decrease during the run-in stage. The asperity peaks are rounded off in the process while the deep valleys remain. For the normal test stage, however, R_{q1} is observed to increase, which is caused by the surface wear in the form of micro-pitting. This phenomenon is not found on the disk surface which experiences positive sliding. Figure 12 shows that R_{q2} decreases in the run-in stage as well, while keeps almost constant in the normal test stage. Due to the high slide-to-roll ratio, the disk actually experiences more contact cycles than the roller does. However, only very limited number of micro-pits is found on the disk surface, such that R_{q2} is

Figure 8. Micro-scope images (100 X magnification) of the roller surface for Test 5: (a1-3) 4 million cycles, (b1-3) 12 million cycles, and (c1-3) 20 million cycles. The 1st, 2nd and 3rd columns represent the three different locations positioned 120° away from each other circumferentially.

Figure 9. Micro-scope images (100 X magnification) of the roller surface for Test 4: (a1-3) 4 million cycles, (b1-3) 12 million cycles, and (c1-3) 20 million cycles. The 1st, 2nd and 3rd columns represent the three different locations positioned 120° away from each other circumferentially.

Figure 10. Micro-scope images (100 X magnification) of the roller surface for Test 6: (a1-3) 4 million cycles, (b1-3) 12 million cycles, and (c1-3) 20 million cycles. The 1st, 2nd and 3rd columns represent the three different locations positioned 120° away from each other circumferentially.

Figure 11. Measured roller surface roughness profiles for Test 5.

Figure 12. Measured disk surface roughness profiles for Test

not affected by the negligible surface wear. It is interesting to see that the sliding direction is also a very important factor for micro-pitting failure, although it is not included in the test matrix.

Figure 13 shows the final micro-pitting outcome for the other five tests listed in Table 5 at the end of 20 million loading cycles. For Test 1, 2 and 3, very limited amounts of micro-pits are observed. Test 7 shows no micro-pitting and Test 8 has small amount of micro-pits.

Using $\bar{\Phi}$ at the end of each test as the response, the main effects plot of the six contact parameters is shown in Figure 14. Here, the influence of one factor on micro-pitting is constructed by holding the other factors constant at zero [25]. The reference line of $\bar{\Phi} = 4.3\%$ is the overall mean of the eight tests. It is seen in Figure 14 that high level of p_h , low level of v_r , high sliding and high roughness amplitude lead to more micro-pitted area. It is also interesting to find that a run-in stage with relatively high contact pressure and relatively low rolling velocity can effectively reduce the amount of micro-pits. Such a process can benefit the contact surface in many ways, for instance: (i) gently flatten the roughness peaks and relieve the surface stress concentrations induced by local roughness interactions for the normal test stage; (ii) more effectively generate the protective low-friction tribo-film owing to the relatively large frictional heat; and (iii) raise the local surface hardness through work hardening.

4. Contact Fatigue Micro-pitting Model

For the modeling of rolling contact fatigue in terms of micro-pitting failure for lubricated contacts, the methodology has to include three major components as shown in the flowchart of Figure 15. The first one is a deterministic mixed elastohydrodynamic lubrication (EHL) model, which is required to capture the transient surface normal pressure p and shear q fluctuations induced by the local surface asperities. With these predicted surface traction distributions, the next

Figure 13. Micro-scope images (100 X magnification) of the roller surface for (a) Test 1, (b) Test 2, (c) Test 3, (d) Test 7 and (e) Test 8 after 20 million contact cycles.

Figure 14. Main effects plot for the DOE screening process using fractional factorials defined in Table 5.

Figure 15. Flowchart for micro-pitting contact fatigue modeling.

step is to compute the multi-axial stress fields in a very shallow layer of material going into the contact body from and includes the rough surface using a stress prediction model. This stress formulation should be able to include the influence of the local surface roughness geometry on the near surface stress concentrations, since the crack nucleation has been experimentally observed to be tightly related to the surface roughness profiles. From the stress field predictions within a sufficiently long time period, the amplitudes and means of the stress components must be determined for every material point that passes through the computational volume. Finally, a suitable multi-axial fatigue criterion is needed in order to predict the micro-pitting fatigue life from the stress histories. In the case of gear micro-pitting, the point contact mixed EHL model should be replaced with the gear mixed lubrication model which captures all the transient effects of the time-varying tooth force, sliding and rolling velocities and contact radii. Additionally, a gear load distribution model would be required to find the tooth force distribution along the tooth flank.

In this study, the mixed EHL model of Li and Kahraman [26], the newly devised boundary element based stress prediction model for rough surface point contacts and the high cycle, multi-axial fatigue criterion proposed by Liu and Mahadevan [27] are combined to model the micro-pitting life. The life of crack propagation from the nucleation stage to the micro-sized pit is assumed to be negligible compared with the initiation life [28]. For the definition of the severity of micro-pitting failure, the cumulative probability obtained from the probability density distribution constructed from the population of the predicted crack nucleation lives is proposed to be the measure.

For micro-pitting failure, the thermal effects caused by the frictional heating between the surfaces are assumed to be trivial and are not included. The variations in the surface roughness profiles after the initial run-in stage are considered to be negligible and the roughness profiles measured after the run-in are used in the simulations. It is also assumed that the hardness within

the very shallow layer of material of interest is constant. Additionally, the stress fields are considered to be pure elastic for the loading range employed in this study. Due to the lack of the measured data, the residual stress is not included in the stress field prediction, while this option is available in the model.

4.1 Mixed EHL Model

For hydrodynamically lubricated rough surface contacts operating under combined rolling and sliding, the fluid film breaks down intermittently within the contact due to the local roughness features and the unfavorable operating conditions such as relatively low rolling velocity, high sliding (which causes shear thinning), high temperature (which reduces the fluid viscosity) and heavy loading. The coexistence of these isolated boundary lubrication spots and the surrounding fluid region (mixed lubrication) introduces numerical instabilities and extraordinary computational efforts into the solution of the transient surface pressure and shear, which are the prerequisites for the contact fatigue analysis. The works of [29,30] proposed a unified numerical approach to solve for the asperity contact pressure and hydrodynamic pressure simultaneously, which has been shown to be an effective method in achieving numerical robustness. To improve the convergence rate and reduce the computational time, Li and Kahraman [26] introduced the asymmetric integrated control volume discretization scheme, which also reduces the dependence of the solution accuracy on grid density. This dependence was shown to be a potential shortcoming of the models of [29,30] especially when the rolling speed is low [31]. With the superior robustness and numerical efficiency, the point contact mixed EHL model of [26] is employed here for the surface traction determination.

Within the fluid region of the contact, the hydrodynamic fluid flow is governed by the transient Reynolds equation

$$\frac{\partial}{\partial x} \left(f_x \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(f_y \frac{\partial p}{\partial y} \right) = \frac{\partial (v_r \rho h)}{\partial x} + \frac{\partial (\rho h)}{\partial \vartheta} \quad (2a)$$

where x and y point into the direction of rolling and the direction normal to the rolling direction, respectively, and ϑ is the time. The parameters p , h and ρ in Eq. (2a) represent the pressure, thickness and density of the fluid, respectively, which are all functions of x , y and ϑ . v_r is the rolling velocity defined as $v_r = (v_1 + v_2)/2$ with v_1 and v_2 denoting the surface velocities (in the x direction) of the contacting body 1 and 2. To model the non-Newtonian effect, the fluid is assumed to be an Eyring fluid such that the flow coefficient in the x direction (f_x), and the flow coefficient in the y direction (f_y) have the expressions of [26]

$$f_x = \frac{\rho h^3}{12\eta} \cosh \left(\frac{\tau_m}{\tau_0} \right), \quad f_y = \frac{\rho h^3}{12\eta} \frac{\sinh(\tau_m/\tau_0)}{\tau_m/\tau_0} \quad (2b,c)$$

where η is the lubricant viscosity, τ_0 is the lubricant reference stress, and τ_m is the viscous shear stress determined according to $\tau_m/\tau_0 = \sinh^{-1}[\eta v_s/(\tau_0 h)]$. Here, $v_s = v_1 - v_2$ is the sliding velocity.

For the local spots where the mating asperities slide along each other, the reduced Reynolds equation [26,29,30] applies as

$$\frac{\partial (v_r h)}{\partial x} + \frac{\partial h}{\partial \vartheta} = 0 \quad (3)$$

Equations (2a) and (3) describe the mixed lubrication behavior of the contact covering the fluid regions and the asperity contact regions simultaneously. Here, a smooth transition between the hydrodynamic and boundary lubrication is assumed.

The local film thickness at the position of (x, y) and time ϑ is defined as

$$h(x, y, \vartheta) = h_0(\vartheta) + g_0(x, y) + V(x, y, \vartheta) - R_1(x, y, \vartheta) - R_2(x, y, \vartheta) \quad (4)$$

where $h_0(\vartheta)$ is the reference film thickness, $R_1(x, y, \vartheta)$ and $R_2(x, y, \vartheta)$ are the roughness heights of the two surfaces at time ϑ , and $g_0(x, y)$ is the separation between the two surfaces before any elastic deformation occurs. Assuming perfectly smooth surfaces, the normal deflections of surface 1 and 2 caused by the tangential shear offset each other (consider same material). With the same assumption, the total normal elastic deformation $V(x, y, \vartheta)$ induced by the normal loading can be calculated using the Boussinesq's half space formulation as [8]

$$V(x, y, \vartheta) = \iint_{\Gamma} \Theta(x - x', y - y') p(x', y', \vartheta) dx' dy' \quad (5)$$

where Γ is the computational domain of the contact zone, and $\Theta(x, y)$ is the influence function. Recognizing that Eq. (5) is the convolution operation between $p(x, y, \vartheta)$ and $\Theta(x, y)$, the Discrete Fourier Transform (DFT) convolution technique is utilized to largely improve the computation efficiency.

In order to ensure the balance between the contact pressure distribution (both hydrodynamic and asperity contact) over the entire contact zone and the normal load W applied, the equality of

$$W = \iint_{\Gamma} p(x, y, \vartheta) dx dy \quad (6)$$

is enforced through adjusting the reference film thickness $h_0(t)$ in Eq. (4) within a load iteration loop. For the lubricant property description in terms of the viscosity and density dependence on

pressure, the Roelands viscosity-pressure relationship [32] and the Dowson-Higginson density-pressure relationship [33] are used in this study due to the lack of the lubricant property measurements for the fluid considered.

With the converged solutions of $p(x, y, \vartheta)$ and $h(x, y, \vartheta)$, the transient surface shear, which consists of the viscous component within the hydrodynamically lubricated areas and the contact friction due to the direct asperity interactions, are determined. Assuming there is no slippage at the interface between the lubricant and the solid, and considering both the Poiseuille and Couette flows, the viscous shear stress acting on surface 1 is written as

$$q(x, y, \vartheta) = -\frac{1}{2}h(x, y, \vartheta)\frac{\partial p(x, y, \vartheta)}{\partial x} - \eta_x^* \frac{v_s}{h(x, y, \vartheta)} \quad (7)$$

where $\eta_x^* = \eta / \cosh(\tau_m / \tau_0)$ is the effective viscosity for an Eyring fluid. For the contact regions where the film thickness breaks down, the shear stress is defined as $q(x, y, \vartheta) = \mu p(x, y, \vartheta)$ where μ is the boundary friction coefficient between the mating asperities and is assumed to be $\mu = 0.15$ in this study. The detailed discretization scheme and numerical approach are referred to [26].

4.2 Stress Prediction Model

For contacts such as those of bearings, the contact zone is small and infinite half space can be assumed. In order to implement the boundary element formulation, the boundary of the infinite half space is divided into the finite field Γ_F , the infinite field Γ_I and the infinite semi-sphere Γ_S as shown in Figure 16. Along Γ_S that is located at infinity, the displacements are considered to be zero. The integral over Γ_S due to the local loading can also be shown to be zero [20]. For Γ_I , the zero traction condition is applied. With these and assuming homogeneous, isotropic and

Figure 16. Boundary of infinite half space.

elastic material and absence of body forces, the displacements at the load point P can be related to the displacements and tractions at the field point Q through the boundary integral equation (BIE) as [34]

$$C_{ij}(P)u_j(P) + \int_{\Gamma_F + \Gamma_I} T_{ij}(P, Q)u_j(Q)d\Gamma = \int_{\Gamma_F} U_{ij}(P, Q)t_j(Q)d\Gamma \quad (8a)$$

where u is the displacement and t is the traction. The variables i and j indicate the directions and $i, j = x, y, z$. The integral kernel functions $U_{ij}(P, Q)$ and $T_{ij}(P, Q)$ are the three-dimensional Kelvin fundamental solutions for u and t at point Q induced by a unit load exerted at point P . U_{ij} and T_{ij} are the functions of the distance between P and Q (denoted by d) as [34]

$$U_{ij} = \frac{1}{16\pi G(1-\nu)} \frac{1}{d} \left[(3-4\nu)\delta_{ij} + \frac{\partial d}{\partial i} \frac{\partial d}{\partial j} \right] \quad (8b)$$

$$T_{ij} = \frac{-1}{8\pi(1-\nu)d^2} \frac{\partial d}{\partial \mathbf{n}} \left[(1-2\nu)\delta_{ij} + 3 \frac{\partial d}{\partial i} \frac{\partial d}{\partial j} \right] - \frac{1-2\nu}{8\pi(1-\nu)d^2} \left[\frac{\partial d}{\partial j} n_i - \frac{\partial d}{\partial i} n_j \right] \quad (8c)$$

where G is the shear modulus, ν is the Poisson's ratio, $\mathbf{n} = [n_x, n_y, n_z]$ is the unit outward normal vector and δ_{ij} is the Kronecker delta (i.e. $\delta_{ij} = 1$ when $i = j$ and $\delta_{ij} = 0$ when $i \neq j$). It is seen U_{ij} and T_{ij} are of the order of $1/d$ and $1/d^2$, respectively, and are singular when Q reaches P .

The finite field consists of the loading region defined by the bold rectangle which circumscribes the Hertzian zone (shaded circle) and the surrounding transition region as shown in Fig. 17. Within Γ_F , triangular boundary elements (TBE) are used. The z coordinates of the TBE nodes represent the roughness heights. In order to capture the local surface roughness geometry,

Figure 17. Computational domain and boundary element mesh for infinite half space.

the mesh size within the loading region is set to be on the order of microns (same as the roughness measurement resolution). Under this fine mesh condition, isoparametric linear boundary elements are assumed to be sufficient for the accurate description of the contact. In the transition region, the surface displacements and tractions are much smaller. Accordingly, the mesh size of the TBE is set to be larger and gradually increase along the radial direction (going outward from the loading region). Isoparametric linear shape functions are also used in this region. Utilizing the linear shape functions of $N_1 = \xi_1$, $N_2 = \xi_2$ and $N_3 = 1 - \xi_1 - \xi_2$, the integrals over the finite region in Eq. (8a) are discretized as

$$\int_{\Gamma_F} T_{ij} u_j d\Gamma = \sum_{K_F} \sum_{\ell=1}^3 u_j^\ell I_{ij,\ell}^T + \sum_{K_F^*} \sum_{\substack{\ell=1 \\ \ell \neq \ell^*}}^3 u_j^\ell I_{ij,\ell}^T + \sum_{K_F^*} u_j^{\ell^*} I_{ij,\ell^*}^T \quad (9a)$$

$$\int_{\Gamma_F} U_{ij} t_j d\Gamma = \sum_{K_F} \sum_{\ell=1}^3 t_j^\ell I_{ij,\ell}^U + \sum_{K_F^*} \sum_{\substack{\ell=1 \\ \ell \neq \ell^*}}^3 t_j^\ell I_{ij,\ell}^U + \sum_{K_F^*} t_j^{\ell^*} I_{ij,\ell^*}^U \quad (9b)$$

where u_j^ℓ and t_j^ℓ represent the displacements and tractions at the node ℓ of the element, $I_{ij,\ell}^T = \int_S T_{ij} N_\ell dS$ and $I_{ij,\ell}^U = \int_S U_{ij} N_\ell dS$, K_F and K_F^* refer to the TBEs that are away from the load point P and the TBEs that are connected to P , respectively. In addition, ℓ^* is the node where P and Q coincide and S is the area of the element.

For numerical integrations of Eq. (9), Gaussian quadrature technique can be directly used for the integrals over the K_F TBEs (first terms on the right hand side of Eqs. (9a) and (9b)).

When P and Q are in the same element (K_F^* TBEs), however, U_{ij} and T_{ij} exhibit singular behavior ($1/d$ and $1/d^2$) as Q approaches P . To eliminate such singularity, the coordinate transformation from the area coordinate system of (ξ_1, ξ_2) to the unit square coordinate system of

(η_1, η_2) as illustrated in Fig. 18 is implemented, introducing the transformation Jacobian of $J_\xi^\eta = d / (4\overline{PQ'})$. Additionally, the shape function N_ℓ for $\ell \neq \ell^*$ is of the order d [34], such that both $T_{ij}N_\ell J_\xi^\eta$ and $U_{ij}N_\ell J_\xi^\eta$ are bounded for the TBEs of K_F^* when $\ell \neq \ell^*$. For the case of $\ell = \ell^*$, $U_{ij}N_\ell J_\xi^\eta$ is still finite, while $T_{ij}N_\ell J_\xi^\eta$ is on the order of $1/d$. Hence, all the other integrals in Eq. (9) with the exception of I_{ij, ℓ^*}^T can be evaluated using Gaussian quadrature without difficulties.

The infinite field Γ_I is discretized into isoparametric linear infinite quadrilateral boundary elements (IQBE) whose two nodes are connected with Γ_F (solid circles on the bound of Γ_F in Fig. 17) and the other two are positioned at infinity. For an arbitrary point $A (\zeta_1, \zeta_2)$ within the IQBE as shown in Fig. 19(a), its displacements can be related to the ones at point $B (\zeta_1, -1)$ which is on the bound of Γ_F by using a decay function $D = \overline{CB} / \overline{CA}$ (C is the decay origin) as $u_i(\zeta_1, \zeta_2) = Du_i(\zeta_1, -1)$ [21]. The mid-side nodes 5 and 6 in Figure 19(a) are selected such that C is positioned at the center of the contact zone. Writing $u_i(\zeta_1, -1)$ in terms of the displacements at nodes 3 and 4 of the IQBE (Fig. 19(a))

$$u_i(\zeta_1, -1) = M_3 u_i^3 + M_4 u_i^4 \quad (10a)$$

the displacements at point A can be further expressed in the form of

$$u_i(\zeta_1, \zeta_2) = D(M_3 u_i^3 + M_4 u_i^4) \quad (10b)$$

such that the discretization of the integral over the infinite region in Eq. (8a) reads

Figure 18. Mapping from area coordinate system (ξ_1, ξ_2) to unit square coordinate system (η_1, η_2) .

Figure 19. Mapping from (a) infinite quadrilateral element to (b) unit square element.

$$\begin{aligned}
 \int_{\Gamma_I} T_{ij} u_j d\Gamma = & \sum_{K_I} \sum_{\ell=3}^4 u_j^\ell \int_S T_{ij} D M_\ell dS + \sum_{K_I^*} \sum_{\substack{\ell=3 \\ \ell \neq \ell^*}}^4 u_j^\ell \int_S T_{ij} D M_\ell dS \\
 & + \sum_{K_I^*} u_j^{\ell^*} \int_S T_{ij} (D-1) M_{\ell^*} dS + \sum_{K_I^*} u_j^{\ell^*} \int_S T_{ij} M_{\ell^*} dS
 \end{aligned} \tag{11}$$

where K_I and K_I^* , respectively, refer to the IQBEs that are away from the load point P and the IQBEs that are connected to P and M_3 and M_4 are the shape functions defined as $M_3 = (1 - \zeta_1)/2$ and $M_4 = (1 + \zeta_1)/2$.

In order to be able to use Gaussian quadrature for the infinite field, the infinite element is mapped into the unit square element as shown in Fig. 19 [21] such that

$$\zeta_1 = \eta_1 \quad , \quad \zeta_2 = \frac{1 + 3\eta_2}{1 - \eta_2} \tag{12a,b}$$

The Jacobian of this transformation is $J_\zeta^\eta = 4/(1 - \eta_2)^2$. For the K_I^* IQBEs, the quadrilateral element (Figure 19(b)) is further divided into two TBEs. For each of them, the transformation of Fig. 18 is performed to introduce the Jacobian that is on the order of d . When $\ell \neq \ell^*$, the shape function M_ℓ is of the order d . On the other hand, when $\ell = \ell^*$, the decay function $D \rightarrow 1$ and $(D-1) \rightarrow 0$ as Q approaches P , i.e. the term $(D-1)$ is on the order of d as well. As a result, the singularities in the second and third terms on the right hand side of Eq. (11) are eliminated. However, the singularity in the integral of the last term remains. Substituting Eqs. (9) and (11) into Eq. (8a) and moving all the undetermined terms to the left, it is arrived

$$\begin{aligned}
 & \left(C_{ij} + \sum_{K_F^*} I_{ij,\ell^*}^T + \sum_{K_I^*} \int_S T_{ij} M_{\ell^*} dS \right) u_j^{\ell^*} = \\
 & \sum_{K_F+K_F^*} \sum_{\ell=1}^3 t_j^{\ell} I_{ij,\ell}^U - \sum_{K_F} \sum_{\ell=1}^3 u_j^{\ell} I_{ij,\ell}^T - \sum_{K_F^*} \sum_{\substack{\ell=1 \\ \ell \neq \ell^*}}^3 u_j^{\ell} I_{ij,\ell}^T \\
 & - \sum_{K_I} \sum_{\ell=3}^4 u_j^{\ell} \int_S T_{ij} D M_{\ell} dS - \sum_{K_I^*} \sum_{\substack{\ell=3 \\ \ell \neq \ell^*}}^4 u_j^{\ell} \int_S T_{ij} D M_{\ell} dS - \sum_{K_I^*} u_j^{\ell^*} \int_S T_{ij} (D-1) M_{\ell^*} dS
 \end{aligned} \tag{13}$$

To evaluate the integrals with singular kernels and the parameter C_{ij} in Eq. (13), the indirect approach of rigid body motion is used. Under rigid body motion, u_i ($i = x, y, z$) is constant and $t_i = 0$ along the entire boundary. Applying this condition and $u_i = 1$, the boundary integral equation reads

$$C_{ij}(P) + \int_{\Gamma_F + \Gamma_I + \Gamma_S} T_{ij}(P, Q) d\Gamma = 0 \tag{14}$$

In a way that is similar to Eqs. (9) and (11), and applying the analytical solution of $\int_{\Gamma_S} T_{ij} d\Gamma = -\delta_{ij}/2$ [20], Eq. (14) is discretized as

$$\begin{aligned}
 C_{ij} + \sum_{K_F^*} I_{ij,\ell^*}^T + \sum_{K_I^*} \int_S T_{ij} M_{\ell^*} dS = & - \sum_{K_F} \sum_{\ell=1}^3 I_{ij,\ell}^T - \sum_{K_F^*} \sum_{\substack{\ell=1 \\ \ell \neq \ell^*}}^3 I_{ij,\ell}^T \\
 & - \sum_{K_I} \int_S T_{ij} dS - \sum_{K_I^*} \sum_{\substack{\ell=3 \\ \ell \neq \ell^*}}^4 \int_S T_{ij} M_{\ell} dS + \frac{\delta_{ij}}{2}.
 \end{aligned} \tag{15}$$

It is noted that there is no decay function involved in Eq. (15) since the displacements in all the directions are constant for rigid body motion. The integrals on the right hand side of Eq. (15) can be readily evaluated using Gaussian quadrature as describe above. Substituting Eq. (15) into Eq. (13), Eq. (8a) is finally discretized as

$$\begin{aligned}
 & \left(\frac{\delta_{ij}}{2} - \sum_{K_F} \sum_{\ell=1}^3 I_{ij,\ell}^T - \sum_{K_F^*} \sum_{\substack{\ell=1 \\ \ell \neq \ell^*}}^3 I_{ij,\ell}^T - \sum_{K_I} \int_S T_{ij} dS - \sum_{K_I^*} \sum_{\substack{\ell=3 \\ \ell \neq \ell^*}}^4 \int_S T_{ij} M_{\ell} dS \right) u_j^{\ell^*} = \\
 & \sum_{K_F+K_F^*} \sum_{\ell=1}^3 t_j^{\ell} I_{ij,\ell}^U - \sum_{K_F} \sum_{\ell=1}^3 u_j^{\ell} I_{ij,\ell}^T - \sum_{K_F^*} \sum_{\substack{\ell=1 \\ \ell \neq \ell^*}}^3 u_j^{\ell} I_{ij,\ell}^T \\
 & - \sum_{K_I} \sum_{\ell=3}^4 u_j^{\ell} \int_S T_{ij} D M_{\ell} dS - \sum_{K_I^*} \sum_{\substack{\ell=3 \\ \ell \neq \ell^*}}^4 u_j^{\ell} \int_S T_{ij} D M_{\ell} dS - \sum_{K_I^*} u_j^{\ell^*} \int_S T_{ij} (D-1) M_{\ell^*} dS
 \end{aligned} \tag{16}$$

The displacement distributions along Γ_F can then be solved from Eq. (16), provided the boundary traction distributions are available. Different from Ref. [21], the proposed boundary element formulation does not require any additional integration of the kernel function T_{ij} over Γ_I . In Ref. [21], the integral of $\sum_{K_I^*} \int_S T_{ij} M_{\ell^*} dS$ was mistakenly double counted.

For any boundary or interior point, its stress state is determined by the displacement and traction distributions along the entire boundary as [34]

$$\sigma_{ij}(P) = \int_{\Gamma_F} E_{kij}(P, Q) t_k(Q) d\Gamma - \int_{\Gamma_F + \Gamma_I} H_{kij}(P, Q) u_k(Q) d\Gamma \tag{17a}$$

where $k = x, y, z$ and the third order kernel functions read [34]

$$E_{kij} = \frac{1}{8\pi(1-\nu)d^2} \left[(1-2\nu) \left(\delta_{jk} \frac{\partial d}{\partial x_i} + \delta_{ik} \frac{\partial d}{\partial x_j} - \delta_{ij} \frac{\partial d}{\partial x_k} \right) + 3 \frac{\partial d}{\partial x_i} \frac{\partial d}{\partial x_j} \frac{\partial d}{\partial x_k} \right] \tag{17b}$$

$$\begin{aligned}
 H_{kij} = & \frac{G}{4\pi(1-\nu)d^3} n_i \left[3\nu \frac{\partial d}{\partial x_j} \frac{\partial d}{\partial x_k} + (1-2\nu)\delta_{jk} \right] \\
 & + \frac{G}{4\pi(1-\nu)d^3} n_j \left[3\nu \frac{\partial d}{\partial x_i} \frac{\partial d}{\partial x_k} + (1-2\nu)\delta_{ik} \right] \\
 & + \frac{G}{4\pi(1-\nu)d^3} n_k \left[3(1-2\nu) \frac{\partial d}{\partial x_i} \frac{\partial d}{\partial x_j} - (1-4\nu)\delta_{ij} \right] \\
 & + \frac{3G}{4\pi(1-\nu)d^3} \frac{\partial d}{\partial \mathbf{n}} \left[(1-2\nu)\delta_{ij} \frac{\partial d}{\partial x_k} + \nu \left(\delta_{jk} \frac{\partial d}{\partial x_i} + \delta_{ik} \frac{\partial d}{\partial x_j} \right) - 5 \frac{\partial d}{\partial x_i} \frac{\partial d}{\partial x_j} \frac{\partial d}{\partial x_k} \right]
 \end{aligned} \tag{17c}$$

These kernel functions are on the order of $1/d^2$ and $1/d^3$. In order to avoid such singularity, the stress components for the boundary points are computed using the strains and tractions according to the Hooke's law [34] in a local orthogonal coordinate system (x', y', z') as shown in Fig. 20 (a). Here, the z' axis points out of the surface and is not shown in the figure. The area coordinate system (ξ_1, ξ_2) is related to (x', y') through

$$\xi_1 = \frac{1}{AC} \left(x' - \frac{y'}{\tan \theta} \right), \tag{18a}$$

$$\xi_2 = \frac{y'}{BC \sin \theta}. \tag{18b}$$

Denoting the direction cosines of x' , y' and z' in the global (x, y, z) system as $(\alpha_x, \alpha_y, \alpha_z)$, $(\beta_x, \beta_y, \beta_z)$ and $(\gamma_x, \gamma_y, \gamma_z)$, respectively, and interpolating the displacements within an element using the linear shape function as $u_i = u_i^\ell N_\ell$, the displacements in the local (x', y', z') system are $u_{x'} = \alpha_i(u_i^\ell N_\ell)$ and $u_{y'} = \beta_i(u_i^\ell N_\ell)$. The local strains of $\varepsilon_{x'} = \partial u_{x'}/\partial x'$, $\varepsilon_{y'} = \partial u_{y'}/\partial y'$ and $\varepsilon_{x'y'} = \partial u_{x'}/\partial y' + \partial u_{y'}/\partial x'$ are determined accordingly as

$$\varepsilon_{x'} = \alpha_i u_i^\ell \left(\frac{\partial N_\ell}{\partial \xi_1} \frac{\partial \xi_1}{\partial x'} + \frac{\partial N_\ell}{\partial \xi_2} \frac{\partial \xi_2}{\partial x'} \right), \quad (19a)$$

$$\varepsilon_{y'} = \beta_i u_i^\ell \left(\frac{\partial N_\ell}{\partial \xi_1} \frac{\partial \xi_1}{\partial y'} + \frac{\partial N_\ell}{\partial \xi_2} \frac{\partial \xi_2}{\partial y'} \right), \quad (19b)$$

$$\varepsilon_{x'y'} = \alpha_i u_i^\ell \left(\frac{\partial N_\ell}{\partial \xi_1} \frac{\partial \xi_1}{\partial y'} + \frac{\partial N_\ell}{\partial \xi_2} \frac{\partial \xi_2}{\partial y'} \right) + \beta_i u_i^\ell \left(\frac{\partial N_\ell}{\partial \xi_1} \frac{\partial \xi_1}{\partial x'} + \frac{\partial N_\ell}{\partial \xi_2} \frac{\partial \xi_2}{\partial x'} \right). \quad (19c)$$

Likewise, the tractions in the local (x', y', z') system are $t_{x'} = \alpha_i t_i$, $t_{y'} = \beta_i t_i$ and $t_{z'} = \gamma_i t_i$. With these, the stress components are computed from the strains and tractions in the local system as [34]

$$\sigma_{x'} = \frac{E}{1-\nu^2} (\varepsilon_{x'} + \nu \varepsilon_{y'}) + \frac{\nu}{1-\nu} t_{z'}, \quad (20a)$$

$$\sigma_{y'} = \frac{E}{1-\nu^2} (\varepsilon_{y'} + \nu \varepsilon_{x'}) + \frac{\nu}{1-\nu} t_{z'} \quad (20b)$$

$$\sigma_{x'y'} = \frac{E}{2(1+\nu)} \varepsilon_{x'y'}, \quad \sigma_{z'} = t_{z'}, \quad \sigma_{y'z'} = t_{y'}, \quad \sigma_{x'z'} = t_{x'} \quad (20c-f)$$

which are then transformed back into the global (x, y, z) system.

For the interior points, which are always away from the field points, Equation (17a) can be directly used for the stress field computation. However, since the micro-pitting crack is surface nucleated and the layer of material of interest has a very shallow depth, the small distance between P and Q introduces numerical error into the Gaussian quadrature process due to the near singular behavior of E_{kij} and H_{kij} . In this work, a numerical approach that combines coordinate transformation and element subdivision is devised to minimize such error.

Similar to the treatment for U_{ij} and T_{ij} , the transformation of Figure 18 is applied here to introduce the Jacobian of the order of $O(d)$, weakening the kernel singularity of E_{kij} and H_{kij} from $O(1/d^2)$ and $O(1/d^3)$ to $O(1/d)$ and $O(1/d^2)$, respectively. Noticing the gradient of $1/d^m$ ($m=1,2$) increases sharply as $d \rightarrow 0$, a progressive subdivision method is developed to sub-mesh the element in an efficient manner. This method introduces the transformation Jacobian that goes to zero at a faster pace than d does. As illustrated in Figure 20(b), the size of the sub-elements along the local η_1 direction (column), denoted as $\bar{L}_i^c$ (Here, i is the sub-element index number in the η_1 direction and $1 \leq i \leq \kappa$), is progressively reduced when approaching the load point P . In the local η_2 direction (row), however, uniform sub-element size, $\bar{L}^r$, is prescribed since the singular behavior does not vary in this direction. The sub-division procedure is summarized into three steps: (i) determine the size for the column of sub-elements that is farthest from P in the η_1 direction, i.e. $\bar{L}_1^c = 2/\Omega$, where the integer Ω can be selected referring to [35] using the length of the longest side of the TBE as the element size; (ii) create the progressive mesh in the η_1 direction according to $\bar{L}_i^c = (1-1/\Omega)\bar{L}_{i-1}^c$ ($i > 1$), in which way, $L_i^c/L_{i-1}^c = L_i^r/L_{i-1}^r$ where L_i^c and L_i^r are the sub-element sizes in the global coordinate system; (iii) generate Ω rows of sub-elements in the η_2 direction. This progressive subdivision introduces the transformation of

$$\eta_1 = \frac{1}{2}(\eta_1^W + \eta_1^E) + \frac{\bar{L}_i^c}{2}\bar{\eta}_1, \quad (21a)$$

$$\eta_2 = \frac{1}{2}(\eta_2^S + \eta_2^N) + \frac{\bar{L}^r}{2}\bar{\eta}_2 \quad (21b)$$

Figure 20. (a) Transformation from (ξ_1, ξ_2) to (x', y') for boundary stress computation and (b) subdivision of a triangular boundary element for interior stress computation.

where η_1^W and η_1^E are the η_1 coordinates of the west and east bounds of the sub-element, and η_2^S and η_2^N are the η_2 coordinates of the south and north bounds of the sub-element. The corresponding transformation Jacobian is $J_{\bar{\eta}} = \bar{L}_i^c \bar{L}^r / 4$. It is noted $\bar{L}_i^c \rightarrow 0$ as $i \rightarrow \kappa$ such that the singular orders of the kernel functions are further reduced. Additionally, the sub-element that is closer to the load point P has smaller area, leading to higher density of Gaussian quadrature points and better accuracy. The discretization of Eq. (17a) is similar to that of Eq. (8a) and is omitted.

4.3 Multi-axial Contact Fatigue Model

For high-cycle multi-axial fatigue analysis, the stress based critical plane approach [36-40] which evaluates the fatigue damage on a pre-defined critical plane has been commonly used. Some criteria considered the critical plane to be the plane with the maximum shear stress amplitude [36-38]. Some included the normal stress component [39] or hydrostatic stress [40] in the definition of the critical plane. These different definitions are usually arrived from the experimental observations using polished small specimens made of different materials and/or operating under different loading conditions. For the rolling contact fatigue problem which experiences cyclic compression, no widely accepted multi-axial fatigue criterion is available. The work of [12] compared a number of critical plane criteria for the rolling contact fatigue damage assessment and yielded similar results from different criteria. The maximum damage was found to occur at the depth about 10 μm , which is below the micro-pitting nucleation sites usually observed (less than 2 μm). Li and Kahraman [10] applied the critical plane criterion of [38] for gear contacts under various loading and surface roughness condition. The predicted crack nucleation site was also found to be deeper than the experimental observation.

Another multi-axial fatigue evaluation approach, namely the characteristic plane approach, was proposed by Liu and Mahadevan [27] for railroad wheel contacts that is similar to the contact problem considered in this study. This approach assumes the fatigue assessment can be approximated by using a certain stress combination (damage parameter) on a certain plane which is not necessarily the macro crack plane (fatigue fracture plane). This certain plane is referred as the characteristic plane, on which the fatigue criterion is [27]

$$\frac{1}{\beta \left(1 - \frac{\sigma_{m,\max}}{S_{ref}} \right)} \left[\sigma_a^2 + \left(\frac{S_N^b}{S_N^t} \right)^2 (\tau_{a1}^2 + \tau_{a2}^2) \right]^{\frac{1}{2}} = S_N^b \quad (22)$$

where S_N^b and S_N^t are the uni-axial fully reversed bending and fully reversed torsion fatigue strength corresponding to the finite fatigue life cycle of N , σ_a , τ_{a1} and τ_{a2} are the normal and shear stress amplitudes acting on the characteristic plane, and β is a material property related parameter. The term $(1 - \sigma_{m,\max}/S_{ref})$ is incorporated to include the mean normal stress effect. Here $\sigma_{m,\max}$ is the mean normal stress on the macro crack plane (assumed as the plane with the maximum normal stress amplitude) and S_{ref} is a reference stress that can be calibrated by using the uni-axial bending fatigue data with non-zero mean normal stress. In the absence of such material data, S_{ref} can be approximated using the ultimate tensile strength or yield strength.

In order to determine the orientation of the characteristic plane, the macro crack plane which is assumed to be the plane experiencing the maximum normal stress amplitude is used as a reference plane. Defining the characteristic plane to be the one on which the damage introduced by the hydrostatic stress amplitude is minimum, the angle between the characteristic plane and the

macro crack plane α and the material parameter β can be derived [27] by implementing Eq. (22) to the fully reversed bending and torsion fatigue problems as

$$\alpha = \frac{1}{2} \cos^{-1} \left[\frac{-s^2 + \sqrt{s^4 - (1 - 3s^2)(-1 + 5s^2 - 4s^4)}}{-1 + 5s^2 - 4s^4} \right], \quad (23a)$$

$$\beta = \sqrt{s^2 \cos^2(2\alpha) + \sin^2(2\alpha)} \quad (23b)$$

for non-extremely brittle materials with $s = S_N^t / S_N^b < 1$. This multi-axial fatigue criterion was used previously in the macro-pitting crack nucleation modeling for point contacts [9] and gear contacts [10] and was shown to be able to yield reasonably good agreements with the measurements in terms of the crack nucleation site and fatigue life. Thus, this criterion is adopted in this work as well for the micro-pitting crack initiation prediction. However, the methodology outlined in this work is general such that any other multi-axial fatigue criterion can be used in its place if deemed appropriate.

4.4 Verification of the Stress Predictions through Analysis of a Circular Hertzian Contact

For the formulations presented in Section 4, the accuracy and efficiency of the point contact lubrication model has been well demonstrated [26] by comparing with the experimental measurements in terms of contact pressure and fluid film thickness and the numerical predictions in terms of central film thickness in the existing literature. The suitability of the selected multi-axial fatigue criterion has also been justified in the modeling of macro-pitting crack nucleation for simple cylinder contacts [9] and gear contacts [10]. In order to demonstrate the accuracy of the boundary element based stress prediction model, the contact between two identical perfectly smooth steel balls with the Young's modulus of $E = 206.84$ GPa, the Poisson's ratio of $\nu = 0.3$

and the radius of 5 mm is considered. For the Hertzian pressure of $p_h = 1.2$ GPa, the resulted Hertzian radius $a = 41.5 \mu\text{m}$. The dimension of Γ_F in Fig. 17 is set as $L = 2.5a$. For the loading region defined by the bold rectangle which circumscribes the Hertzian zone, it is discretized into 81×81 nodes such that the mesh size is on the order of micrometer. Between the bounds of the loading region and Γ_F , for instance $x = a$ and $x = L/2$, 3 nodes are used. For this transition region, the mesh size is set to gradually increase along the radial direction.

Considering the simple circular Hertzian contact (perfectly smooth surfaces and no lubrication) of

$$t_z(x, y) = -p_h \sqrt{1 - \frac{x^2 + y^2}{a^2}} \quad (24)$$

the closed form formulae are available for the surface displacements as [8]

$$u_x(x, y) = \frac{(1 - 2\nu)(1 + \nu)p_h}{3E} \frac{a^2}{\sqrt{x^2 + y^2}} \left[1 - \left(1 - \frac{x^2 + y^2}{a^2} \right)^{1.5} \right] \cos \theta \quad (25a)$$

$$u_z(x, y) = \frac{\pi p_h (1 - \nu^2)}{4aE} (2a^2 - x^2 - y^2) \quad (25b)$$

where $\theta = \cos^{-1}(x/\sqrt{x^2 + y^2})$. The displacement in the y direction has the same magnitude as u_x and is omitted. Figure 21 shows the model predictions of the u_x and u_z distributions (left column) together with the relative error distributions (err) obtained by comparing with the solutions of Eq. (25) (right column). It is seen that $err < 2\%$ for both u_x and u_z . The maximum err occurs at the location where u_x or u_z is minimum where the small magnitude of the displacement lead to the large relative error, although the absolute error is trivial. In Fig. 22,

similar results are obtained for the u_x and u_y components under the pure shear loading condition of

$$t_x(x, y) = \mu p_h \sqrt{1 - \frac{x^2 + y^2}{a^2}} \quad (26)$$

where the friction coefficient $\mu = 0.1$. The closed form formulae (not available for u_z) can also be found in [8] as

$$u_x(x, y) = \frac{\pi \mu p_h}{32 a G} \left[4(2 - \nu) a^2 - (4 - 3\nu) x^2 - (4 - \nu) y^2 \right], \quad (27a)$$

$$u_y(x, y) = \frac{\pi \mu p_h}{16 a G} \nu x y. \quad (27b)$$

In this case, the relative error is observed to be less than 1% for u_x and less than 2% for u_y . The large *err* is again due to the very small magnitude of the displacement.

For the stress distributions, the predicted principal stresses and maximum shear stress under the Hertzian pressure of Eq. (24) along the central vertical line (i.e. negative z direction) of ($x=0, y=0$) are compared with solutions of the close form formulae of [8]

$$\sigma_x(0, 0, z) = \sigma_y(0, 0, z) = -p_h \left\{ (1 + \nu) \left[1 - \frac{z}{a} \tan^{-1} \left(\frac{a}{z} \right) \right] - \frac{1}{2 \left(\frac{z^2}{a^2} + 1 \right)} \right\}, \quad (28a)$$

$$\sigma_z(0, 0, z) = -\frac{p_h}{\left(\frac{z^2}{a^2} + 1 \right)}, \quad (28b)$$

Figure 21. Model predictions of (a1) u_x and (b1) u_z , and the relative errors for (a2) u_x and (b2) u_z under the loading condition of Eq. (24).

Figure 22. Model predictions of (a1) u_x and (b1) u_y , and the relative errors for (a2) u_x and (b2) u_y under the loading condition of Eq. (26).

$$\tau_{\max}(0,0,z) = \left| \frac{\sigma_z - \sigma_x}{2} \right| \quad (28c)$$

in Figure 23. The mesh size in the z direction is $0.1 \mu\text{m}$, which is less than one-tenth of the surface boundary element size. For the progressive subdivision, $\Omega = 2$ and $\kappa = 5$ is used which results into only 10 sub-elements. In comparison with the regular uniform subdivision which requires 10×10 sub-elements, the computational efforts are substantially reduced. The model predictions (solid lines) overlap with the solutions of Eq. (28) (dashed lines), and the maximum relative error is less than 0.7 %. This good agreement exhibited in Figures 21-23 demonstrates the accuracy of the boundary element formulation proposed above.

5 Simulation of Rolling Contact Fatigue Experiments and Model Validation

In this section, the experimental study of Section 3 is simulated and analyzed using the methodology and formulation proposed in Section 4. The detailed predictions of the mixed lubrication behavior, the near surface stress distributions and the crack nucleation fatigue life distributions are illustrated using Test 5 (which is most severely micro-pitted) as an example. In addition, the micro-pitting measurements of all the eight tests are compared with the model predictions to demonstrate the model capabilities and any potential improvement.

For the surface computational grid, its dimension is defined such that $-1.875a \leq x \leq 1.125a$ and $-1.5b \leq y \leq 1.5b$, where a and b represent the Hertzian half widths in the x and y directions, respectively. The mesh density of 256×256 is used and results in the mesh sizes of $\Delta x = 3 \mu\text{m}$ and $\Delta y = 11 \mu\text{m}$. Since the variation of the roughness profile in the y direction is very limited due to the axially oriented roughness texture, this relatively large Δy is considered to be sufficient. The time interval of $\Delta t = 2\Delta x/v_r$ is used for a total of $n_g = 1000$

Figure 23. Comparison between the model predictions (solid line) and the solutions of Eq. (27) (dashed line) for the normal stress components and the maximum shear stress component under the loading condition of Eq. (24).

time steps, such that the entire analysis covers about 4 mm traveling distance of the roller surface. For the wind turbine gear fluid (Castrol Optigear Synthtic X320) used in this study, its density is roughly measured to be $\rho_0 = 812 \text{ kg/m}^3$ at the inlet temperature of 95°C . Due to the lack of the pressure-viscosity relationship, the property of a similar lubricant (Optigear Synthtic A320) whose ambient viscosity $\eta_0 = 0.0276 \text{ Pas}$ and pressure-viscosity coefficient $\alpha = 14.3 \text{ GPa}^{-1}$ is used in the simulation instead. Because of these limited oil property measurements, the Roelands pressure-viscosity relationship [32] is assumed to describe the viscosity dependence on pressure as shown in Figure 24. To take into account the effects of the surface roughness variation in the run-in stage (the round-off of asperity peaks), the measured roughness profiles at the end of the run-in process (Figures 11(b) and 12(b)) are used in the simulation.

For the determination of the severity of asperity interaction activities, the area ratio of asperity contact C_a which is defined as the ratio of the total asperity contact area to the nominal Hertzian area is plotted against the time step n_g in Figure 25 (for Test 5). The average value of C_a over this time period is about 6.5 %, indicating 6.5 % of the Hertzian area experiences asperity contacts on average. For the three example time instants of $n_g = 148, 506$ and 657 (denoted as *A*, *B* and *C* in Figure 25) which respectively represent high ($C_a = 13.2\%$), median ($C_a = 6.5\%$) and low ($C_a = 1.7\%$) extent of asperity interaction, the predicted normal contact pressure (left column) and tangential surface shear (right column) using the mixed lubrication model are shown in Figure 26. The pressure fluctuations are observed for all the three time instants, and the surface shears shoot up wherever the film thickness breaks down (the typical friction coefficient of 0.15 is used for boundary lubrication condition). As C_a decreases, the asperity contact friction reduces evidently. The peaks of the normal pressure, however, can still be relatively high because of the

Figure 24. Pressure-viscosity relationship used in this study.

Figure 25. Variation of asperity contact area ratio with time for Test 5.

Figure 26. Mixed EHL predictions of normal contact pressure (left column) and tangential shear (right column) for time steps of (a) $n_g = 148$ (*A* in Fig. 15), (b) $n_g = 506$ (*B* in Fig. 15) and (c) $n_g = 657$ (*C* in Fig. 15) for Test 5.

surface discontinuities. For instance, although the C_a value at point C is substantially smaller than that at point B , the maximum local pressures at these time instants are comparable.

Provided the surface loading condition such as in Figure 26, the boundary element based stress model is used to calculate the surface and near surface stress state for every material point considered, incorporating the surface roughness local geometry. Figure 27 shows the stress fields along the central vertical plane of $y = 0$ for point A , B and C defined in Figure 25. The components of σ_{xy} and σ_{yz} are negligibly small compared with the others and are omitted in this figure. It is observed that the normal stress concentrations occur at the highlands of the surface roughness, below which the orthogonal shear alternates direction. These roughness highlands are where the micro-pits are found in the experimental study.

With the history of the stress state of a certain material point, its fatigue life is assessed using the multi-axial fatigue model. To start the fatigue analysis, the fracture plane which is assumed to be the plane experiencing the maximum normal stress amplitude is first searched in the three-dimensional space using an Euler angle-based axis transformation [9]. A 4° angle increment is used in the numerical search. From the determined fracture plane, the characteristic plane is reached by rotating another angle α (Eq. (23a)) which is related to the material property in terms of the fully reversed bending and torsion fatigue strength. On the characteristic plane, the fatigue life is then calculated numerically according to Eq. (22).

As the material points of the contact component pass through the contact zone, they experience different loading histories because of the different local surface roughness geometries and the sliding action (if any) which changes the matching of the transient local roughness heights of the mating surfaces. In order to include these variations, the fatigue analysis is performed covering a sufficiently long the roughness profile segment of a length about 4 mm. Figure 28

Figure 27. Transient stress fields of (b) σ_x , (c) σ_y , (d) σ_z and (e) σ_{xz} under the transient roughness contact condition of (a) along the central vertical plane ($y = 0$) for Test 5

Figure 28. Micro-pitting crack nucleation fatigue life distribution along the central vertical plane $y = 0$ for the entire 4 mm surface roughness segment considered for Test 5.

displays the distribution of the micro-pitting crack nucleation fatigue life N_f for the central vertical plane of $y = 0$. The most critical spots (with short lives) are observed to be on the highlands of the surface roughness. The material points in the roughness valleys are shown to be able survive more than $(10)^9$ contact cycles. Comparing the fatigue lives along the depth direction z , it can be concluded that the failure is surface initiated, which is in agreement with the experimental observation.

Focusing on the surface layer of Figure 28, the probability density function $\psi(N_f)$ of the fatigue lives is constructed in Figure 29(e) for Test 5 together with the other tests as defined in Table 5. For Test 5, it is found that about 29 % of the surface material points have the fatigue lives that are less than 20 million cycles while around 40 % of the material points have the lives in excess of 100 million cycles. To provide a measure for micro-pitting failure, the Micro-pitting Severity Index (MSI) $\Psi(N_{f0})$ at the cycle number of N_{f0} is defined as the percentage of the material points having the fatigue lives below N_{f0} , i.e. $\Psi(N_{f0})$ is the cumulative probability function of

$$\Psi(N_{f0}) = \int_1^{N_{f0}} \psi(N_f) dN_f \quad (29)$$

Assuming the population of these surface layer material points considered is a representative pool of the entire surface points, $\Psi(N_{f0})$ is equivalent to the micro-pitting area percentage $\bar{\Phi}$, such that the model predictions can be related and compared to the measurements. It is recognized, however, different positions on the roller surface have different surface roughness profiles, resulting in distributions that might be different from the one shown in Figure 29. In order to include such variation, the bootstrapping method is used to determine the 95 % confidence interval for $\Psi(N_{f0})$. This simple technique constructs subsamples with the same size as that of the original data set

(predicted fatigue life population) by drawing with replacement. As a result, some of the subsamples have the data points with relatively short fatigue lives representing the surface area where the asperity peaks are relatively intense and some of the subsamples have the opposite characteristics. $\Psi_i(N_{f0})$ is then calculated for each subsample i ($i=1$ to 3000 in this work) and the probability density distribution is built for the data set of $\Psi_i(N_{f0})$ such as in Fig. 30 for Test 5 to find its confidence interval.

In Figure 31, the predicted $\Psi(N_{f0})$ ($N_{f0} = 20 \times 10^6$) together with its 95 % confidence interval are compared with the measured average micro-pitting area percentage $\bar{\Phi}$ for all the eight tests. For the case of $\Psi = 0$ (Test 1, 2, 6 and 7), the confidence interval cannot be constructed and only Ψ is plotted. It is seen, the model predictions are in good agreement with most of the measurements except for Test 3 and Test 8. Comparing the operating conditions defined in Table 5 between Test 3 and Test 4 only for the normal test stage, it can be seen that Test 3 has lower rolling velocity and higher sliding which would lead to the appearance of more micro-pits (The model predicts $\Psi = 15\%$). However, the measured average micro-pitting area percentage is only 0.41 %, which is much lower than that of Test 4 (3 %). Similarly, comparing Test 8 with Test 4, the only difference for the normal test operating condition is the higher contact pressure in Test 8, which would result in more severe micro-pitting. However, the measurement shows smaller amount micro-pits of 2.7 % (The model predicts $\Psi = 6\%$). The possible cause for these deviations is that certain benefits introduced by the run-in process are not captured by the model. For instance, the high level of the contact pressure or the low level of the rolling velocity or the combination of the two in the run-in stage could help the formation of the protective lo-friction tribo-film on the contacting surfaces and/or raise the local surface hardness through work

Figure 29. Predicted probability density distribution of the micro-pitting crack nucleation fatigue life for the material points along $y = 0$ and $z = R_1$ for (a) Test 1, (b) Test 2, (c) Test 3, (d) Test 4, (e) Test 5, (f) Test 6, (g) Test 7 and (h) Test 8.

Figure 30. The probability density distribution of $\Psi(N_{f0})$ with $N_{f0} = 20 \times 10^6$ for Test 5, constructed using the bootstrapping method.

Figure 31. Comparison of $\bar{\Phi}(N_{f0} = 20 \times 10^6)$ between the measurements and the model predictions.

hardening. Although the effects of the surface roughness variation during the run-in stage is included in the modeling by using the measured roughness profiles after the run-in process, the other potential benefits are not considered in the current model.

6 Concluding Remarks

A rolling contact fatigue model for the prediction of crack nucleation for micro-pitting failure was proposed in this study for non-conformal point contacts of rough surfaces operating under the mixed lubrication regime. The methodology consisted of three models, which were the mixed EHL model [14] for the determination of the surface traction distribution, the boundary element based stress prediction model for the calculation of the near surface stress fields considering the local roughness geometry, and the multi-axial fatigue model to evaluate the fatigue damage using the fully reversed bending and torsion fatigue strength of the considered material.

A set of rolling contact fatigue tests were also conducted using the fractional factorials technique aiming at the achievement of statistical meaningful measurements with minimized test runs. It was found reduced contact pressure, slide-to-roll ratio and surface roughness amplitude and increased rolling velocity lead to the reduction of micro-pits. It was also found that a relatively large contact pressure and a relatively small rolling velocity (representative of an undesirable contact condition) in the run-in stage have the tendency to reduce the amount of micro-pits. The possible causes include but are not limited to the rounding-off of the asperity peaks, the formation of protective tribo-film and the local hardness increase due to work hardening. At the end, the measurements were compared to the model predictions to show reasonable agreements.

References

- [1] M. Tokuda, M. Nagafuchi, N. Tsushima and H. Muro (1982), “Observations of the peeling mode of failure and surface-originated flaking from a ring-to-ring rolling contact fatigue test rig,” *Rolling Contact Fatigue Testing of Bearing Steels*, ASTM Special Technical Publication, 771, 150–165.
- [2] Y. Ariura, T. Ueno and T. Nakanishi (1983), “An investigation of surface failure of surface-hardened gears by scanning electron microscopy observations,” *Wear*, 87, 305–316.
- [3] M. N. Webster and C. J. J. Norbart (1995), “An experimental investigation of micropitting using a roller disk machine,” *Tribology Transactions*, 38 (4), 883–893.
- [4] T. Ahlroos, H. Ronkainen, A. Helle, R. Parikka, J. Virta and S. Varjus (2009), “Twin disc micropitting tests,” *Tribology International*, 42 (10), 1460–1466.
- [5] P. Brechot, A. B. Cardis, W. R. Murphy and J. Theissen (2000), “Micropitting resistant industrial gear oils with balanced performance,” *Industrial Lubrication and Tribology*, 52 (3), 125–136.
- [6] E. Laine, A. V. Olver, M. F. Lekstrom, B. A. Shollock, T. A. Beveridge and D. Y. Hua (2009), “The effect of a friction modifier additive on micropitting,” *Tribology Transactions*, 52 (4), 526–533.
- [7] N. F. R. Cardoso, R. C. Martins and J. H. O. Seabra (2009), “Micropitting of carburized gears lubricated with biodegradable low-toxicity oils,” *Journal of Engineering Tribology*, 223 (3), 481–495.
- [8] K. L. Johnson, *Contact Mechanics*, Cambridge University Press, Cambridge, 1985.
- [9] S. Li and A. Kahraman (2011), “A fatigue model for contacts under mixed elastohydrodynamic lubrication condition,” *International Journal of Fatigue*, **33** (3), pp. 427–436.

- [10] S. Li and A. Kahraman (2012), “A fatigue model for spur gear contacts operating under mixed elastohydrodynamic lubrication conditions,” *Journal of Mechanical Design*, in print.
- [11] J. Tao, T. G. Hughes, H. P. Evans, R. W. Snidle, N. A. Hopkinson, M. Talks and J. M. Starbuck (2003), “Elastohydrodynamic lubrication analysis of gear tooth surfaces from micropitting tests,” *Journal of Tribology*, **125** (2), pp. 267–274.
- [12] H. Qiao, H. P. Evans and R. W. Snidle (2008), “Comparison of fatigue model results for rough surface elastohydrodynamic lubrication,” *Journal of Engineering Tribology*, **222** (3), pp. 381–391.
- [13] D. Epstein, L. M. Keer, Q. J. Wang, H. S. Cheng and D. Zhu (2003), “Effect of surface topography on contact fatigue in mixed lubrication,” *Tribology Transactions*, **46** (4), pp. 506–513.
- [14] I. A. Polonsky and L. M. Keer (1999), “A numerical method for solving rough contact problems based on the multi-level multi-summation and conjugate gradient techniques,” *Wear*, **231**, 206–219.
- [15] W. Peng, Contact mechanics of multilayered rough surfaces in tribology, PhD Thesis, The Ohio State University, 2001.
- [16] S. Ilincic, G. Vorlaufer, P. A. Fotiu, A. Vernes and F. Franek (2009), “Combined finite element-boundary element method modeling of elastic multi-asperity contacts,” *Journal of Engineering Tribology*, **223**, 767–776.
- [17] U. Sellgren, S. Bjorklund and S. Andersson (2003), “A finite element-based model of normal contact between rough surfaces,” *Wear*, **254**, 1180–1188.
- [18] J. G. Leahy and A. A. Becker (1999), “A quadratic boundary element formulation for three-dimensional contact problems with friction,” *Journal of Strain Analysis*, **34** (4), 235–251.

- [19] P. Wang, A. A. Becker, I. A. Jones and T. H. Hyde (2004), “Boundary element analysis of contact problems using inverse techniques – linear and quadratic elements,” *Journal of Strain Analysis*, **39** (4), 371–382.
- [20] T. G. Davies and S. Bu (1996), “Infinite boundary elements for the analysis of halfspace problems,” *Computers and Geotechnics*, **19** (2), 137–151.
- [21] X. W. Gao and T. G. Davies (1998), “3-D infinite boundary elements for half-space problems,” *Engineering Analysis with Boundary Elements*, **21** (3), 207–213.
- [22] J. Liang and K. M. Liew (2001), “Boundary elements for half-space problems via fundamental solutions: a three-dimensional analysis,” *International Journal for Numerical Methods in Engineering*, **52**, 1189–1202.
- [23] W. Moser, C. Duenser and G. Beer (2004), “Mapped infinite elements for three-dimensional multi-region boundary element analysis,” *International Journal for Numerical Methods in Engineering*, **61** (3), 317–328.
- [24] D. B. Ribeiro and J. B. D. Paiva (2010), “Analyzing static three-dimensional elastic domains with a new infinite boundary element formulation,” *Engineering Analysis with Boundary Elements*, **34**, 707–713.
- [25] A. Dean and D. Voss, *Design and Analysis of Experiments*, Springer-Verlag, New York, 1999.
- [26] S. Li and A. Kahraman (2009), “A mixed EHL model with asymmetric integrated control volume discretization,” *Tribology International*, **42** (8), 1163–1172.
- [27] Y. Liu, S. Mahadevan (2007), “A unified multiaxial fatigue damage model for isotropic and anisotropic materials”, *International Journal of Fatigue*, **29**, 347–359.
- [28] W. Cheng, H. S. Cheng, T. Mura and L. M. Keer (1994), “Micromechanics modeling of crack initiation under contact fatigue”, *Journal of Tribology*, **116**, 2–8.

- [29] Y. Z. Hu and D. Zhu (2000), “A full numerical solution to the mixed lubrication in point contacts”, *Journal of Tribology*, **122**, 1–9.
- [30] D. Zhu (2007), “On some aspects of numerical solutions of thin-film and mixed elastohydrodynamic lubrication”, *Journal of Engineering Tribology*, **221**, 561–579.
- [31] C. H. Venner (2005), “EHL film thickness computations at low speeds: risk of artificial trends as a result of poor accuracy and implications for mixed lubrication modeling”, *Journal of Engineering Tribology*, **219**, 285–290.
- [32] C. J. A. Roelands, Correlational aspects of the viscosity-temperature-pressure relationship of lubricating oils, PhD Thesis, University of Technology, Delft, 1966.
- [33] D. Dowson, G. R. Higginson, Elasto-hydrodynamic lubrication, SI edition, Pergamon Press (1977).
- [34] Becker, A. A., *The Boundary Element Method in Engineering: A Complete Course*, McGraw-Hill Book Company, 1992.
- [35] U. Eberwien, C. Duenser and W. Moser (2005), “Efficient calculation of internal results in 2D elasticity BEM,” *Engineering Analysis with Boundary Elements*, **29 (5)**, 447–453.
- [36] T. Matake (1977), “An explanation of fatigue limit under combined stress”, *Bull. JSME*, **20**, 257–263.
- [37] D. L. McDiarmid (1994), “A shear stress based critical-plane criterion of multiaxial fatigue failure for design and life prediction”, *Fatigue & Fracture of Engineering Materials & Structures*, **17**, 1475–1484.
- [38] L. Susmel and P. Lazzarin (2003), “A stress-based method to predict lifetime under multiaxial fatigue loadings”, *Fatigue & Fracture of Engineering Materials & Structures*, **26**, 1171–1187.

- [39] W. N. Findley (1959), “A theory for the effect of mean stress on fatigue of metals under combined torsion and axial load or bending”, *Journal of Engineering for Industry*, **81**, 301–306.

- [40] K. Dang Van, G. Cailletaud, J. F. Flavenot, A. Le Douaron and H. P. Lieurade (1989), “Criterion for high-cycle fatigue failure under multiaxial loading”, In *Biaxial and Multiaxial Fatigue*, EGF 3, Mechanical Engineering Publications, London, 1989, pp. 459–478.