Damage initiation and evolution in silicon nitride under non-conforming lubricated hybrid rolling contact

Rahul Raga¹, Iyas Khader¹, ², Zdenek Chlup³, Andreas Kailer¹
¹Fraunhofer Institute for Mechanics of Materials IWM, Woehlerstr. 11, 79108 Freiburg, Germany
²Department of Industrial Engineering, German-Jordanian University, P.O. Box 35247, Amman 11180, Jordan
³Institute of Physics of Materials of the Academy of Sciences of the Czech Republic, v. v. i. Zizkova 22, 616 62 Brno, Czech Republic

Abstract

This study focuses on the damage mechanisms in silicon nitride rolling elements used in hybrid (ceramic-metal) bearings that operate under non-conformal contact. To get an insight into the prevailing damage modes compared to the real application, a rolling contact experiment was designed to mimic the contact conditions. Hertzian contact pressures ranged from 3.0 to 5.9 GPa (500 N to 4150 N). In order to approach pure rolling, the experiments were run without inducing any gross slip. Extensive surface and subsurface damage analysis was performed using conventional ceramography as well as FIB cross-sectioning. Finite element simulations were carried out to illustrate the stress state prevailing under different loading conditions. Surface damage to rollers subjected to contact pressures up to 5.1 GPa (2500 N) was mainly dominated by micro-spalling, which was induced due to the presence of snowflake structures. At the highest applied loads, damage appeared as a combination of macro-cracking and micro-spalling. Crack propagation was attributed to different mechanisms: (a) fatigue-induced fracture and (b) lubricant-driven crack propagation in the subsurface.

Keywords: Silicon nitride, rolling contact fatigue (RCF), subsurface damage, hybrid bearings, snowflake structure

1. Introduction

Silicon nitride (Si₃N₄) is the first choice of ceramic materials for modern hybrid bearing applications where conventional bearings fail to meet the prerequisites [1–3]. The unique combination of properties such as, high strength, high hardness, adequate fracture toughness, thermal stability and good corrosion resistance makes silicon nitride attractive for hybrid bearing applications.
Despite being industrially deployed, the reliability of the rolling elements made from silicon nitride undergoing non-conforming contact is still a topic of research. It is evident from full-scale hybrid ball bearing endurance tests that silicon nitride fails due to non-catastrophic fatigue similar to that in conventional bearings [3]. Nonetheless, the damage mechanism of ceramics undergoing rolling contact is completely different [2,4,5]. Full-scale bearing endurance tests are expensive and time consuming, moreover component testing involves many other influencing factors, which complicate the interpretation of results. For this reason many researchers have relied on various rolling contact fatigue (RCF) model experiments either to study damage mechanisms [1,3,6–11], fatigue life predictions [1,3,12,13] or wear characteristics [14–23]. A general review on quality requirements of materials for bearing application can be found in [24] and a recent extensive literature review on various types of damage in silicon nitride is given in [25].

Even when materials are being used under ideal conditions, failure is certain due to deterioration of material by rolling fatigue [1,3,8,9,15]. Damage under rolling contact loading is a complex phenomenon and involves many aspects which affect the material behaviour. The damage mechanism is primarily governed by the tribological conditions as well as the material properties. From the tribological point of view, damage is influenced by the rolling and contact conditions (i.e., with and without slip), contact stress distribution due to the applied load, and lubrication regime. From the point of view of materials properties, the microstructure among other aspects determines the macroscopic mechanical behaviour [26,27]. Due to the fact that silicon nitride could be produced with various microstructures, depending on the sintering technique, sintering aids, and forming process employed, it is possible to obtain materials with different mechanical properties required for specific applications. However, there is a general consensus pertaining to Si$_3$N$_4$ based materials that natural flaws such as inclusions of foreign matter, pores and voids, and surface machining cracks are inevitable [24,28,29]. The aforementioned microscopic non-homogeneities play a decisive role in determining material fatigue life. Variations in tribological conditions or in material properties could cause considerable changes in the RCF behaviour of the material.

The fatigue life of silicon nitride becomes orders of magnitude shorter when rolling is accompanied by friction and sliding and as c-cracks appear on the surface under non-conforming contact (refer to [23,30] for lubricated
cases and [20,22] for the less commonly unlubricated cases). However, damage could be affected substantially even in the absence of sliding, when the rolling elements experience lubricant starvation. Bunting [18] demonstrated how abrasive wear due to direct asperity contact in hybrid bearings changes the contact conformity in a full-bearing test rig.

O’Brein et al. [3] analysed failure mechanism under lubricated conditions in silicon nitride in a full-bearing test. The damage in the silicon nitride rolling elements was mainly dominated by spalling, which initiated from volume defects. They proposed the defect to be sintering voids (less than 2 µm in diameter) and argued that such small defects were inherent material properties rather than manufacturing defects. Burrier [15] studied 11 different grades of silicon nitride materials in RCF tests. The author demonstrated that the materials exhibited fatigue life difference of several orders of magnitude under same working conditions. However, the material with fine grain and uniform minimum distribution of secondary phase showed better fatigue life performance.

In order to understand the crack propagation mechanisms under rolling contact, numerous researchers have performed rolling contact fatigue experiments with artificial ring cracks on the surface [31–46]. Crack propagation in the subsurface is mainly caused by alternating cyclic stresses and in some cases due lubricant entering the crack gap. In the latter case, crack propagation is mainly driven by hydrostatic pressure buildup in the crack gap [42]. Analytical models describing lubricant driven crack propagation can be found in [43–45], however there are only handful of experiments demonstrating the lubricant driven crack mechanism in ceramics [37]. Furthermore, Wang [32] showed the influence of ring crack locations in the contact path on rolling contact fatigue life. He found cracks at certain locations in the contact path to be critical and to contribute to the determination of rolling contact fatigue life based on their stress intensity factor values.

When pre-existing cracks/natural flaws in silicon nitride are negligible, then the failure mode is governed by the porosity of the material [36,46]. In a study [11] performed on fully densified, micro-porous, and porous silicon nitride grades under identical working conditions, the results showed the influence of porosity in the material on the damage mechanism. In case of the fully densified silicon nitride no damage was observed, however the comparison between micro-porous and porous silicon nitride resulted in an order of magnitude difference in the fatigue life failure. In both cases, micro-porous and porous, the damage was dominated by spalling.
For hybrid bearing applications, mostly hot isostatically-pressed silicon nitride (HIPSН) balls are used due to their optimum combination of material properties, nonetheless, high manufacturing cost is still a concern [24]. Nowadays, gas pressure sintered silicon nitrides (GPSН) are gaining importance due to their cost effectiveness and reliable reproducibility in comparison to other manufacturing process. Lengauer [47], Lengauer et al. [48], and Harrer et al. [49] have demonstrated the damage mechanism in GPSН silicon nitride rolling tools applied in wire rolling. However, in GPSН there exists large scale of microstructural inhomogeneities called “snowflake” structures. A snowflake is a designation for a local region of incomplete densification (i.e., microporosity) in a homogeneous material were intergranular glassy phase is missing and only the grain skeleton is present. Herrmann et al. [50] studied silicon nitride produced with various compositions from different manufacturing techniques. They concluded that snowflakes structures are mostly caused due to thermal mismatch between the grain boundary phase and the Si₃N₄ skeleton. Out of four different sintering techniques (HIPSН, hot pressed sintering, GPSН and spark plasma sintering) these regions were only formed in GPSН and hot pressed sintered materials. There are no studies relating how these snowflakes regions affect damage mechanisms under rolling contact.

The uniqueness of this study emerges from addressing the damage mechanisms involved in non-conforming lubricated hybrid rolling contact of a GPSН silicon nitride, which is considered a more feasible and easier to manufacture alternative to HIPSН. Accordingly, the most crucial parameters influencing damage initiation and evolution in this grade of silicon nitride were analysed by performing model experiments on a twin-disk tribometer. Moreover, the behaviour of snowflake structures (volume defects intrinsic to GPSН) in contact was taken into consideration. This has been achieved by observing crack formation and propagation as function of increased load and number of stress cycles and through extensive post-experimental analysis with focused ion beam (FIB) and conventional cross-sectioning. Additionally, finite element (FE) simulations were employed to study the effect of plastic deformation on contact stresses and their influence on damage and crack formation in the ceramic material.
2. Experimental description

2.1 Experimental method

In order to study damage in silicon nitride undergoing lubricated hybrid rolling contact, twin-disk experiments were carried out using ceramic rolls and crowned 100Cr6 steel disks. A non-additivated mineral oil (SKF TT9, Kroon Oil BV, The Netherlands) was used as lubricant. During the experiments, the tribometer records the normal force in addition to the rotational velocity, test time, torque and the coefficient of friction generated in contact. The samples in contact are rotated at the same velocity, thus, eliminating any gross slip.

**Figure 1(a)** depicts the experimental setup. The geometry of the samples results in an elliptical contact as shown in **Figure 1(b)**. The major contact width (2a) lies in the circumferential direction and the minor contact width (2b) lies in the axial direction which also represents the contact track width.

![Experimental setup on the twin-disk tribometer](image)

![Schematic of elliptical contact](image)

Initially, experiments with different loads for a minimum of 10 million revolutions (Mrev) were performed. In order to understand damage initiation and progress of micro-spalling, evolution test were carried out for 850 N from 10 Mrev to 50 Mrev, for 2500 N from 10 Mrev to 30 Mrev and for 4150 N with revolutions ranging from 7500 rev (short test) to 10 Mrev. **Table 1** summarizes the experimental parameters.
Table 1: Experimental parameters for lubricated rolling contact

<table>
<thead>
<tr>
<th>Experimental parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load [N]</td>
<td>500 (2)*, 850 (6), 1250 (1), 2500 (4) and 4150 (13)</td>
</tr>
<tr>
<td>Rotational speed of roll [rpm]</td>
<td>1500</td>
</tr>
<tr>
<td>Rotational speed of disk [rpm]</td>
<td>1500</td>
</tr>
<tr>
<td>No. of revolutions</td>
<td>7500 – 50 M</td>
</tr>
<tr>
<td>Duration [hrs]</td>
<td>0.1 – 555.55</td>
</tr>
<tr>
<td>Lubricant</td>
<td>SKF TT9</td>
</tr>
</tbody>
</table>

* The number in brackets shows the number of conducted tests

The Hertzian equations [51] for solid bodies undergoing non-conforming contact can be used to calculate the contact pressure $p_o$ and the corresponding major ($a$) and minor ($b$) contact ellipse radii.

$$a = c(1 - e^2)^{-1/4}$$  \hspace{1cm} (1)

$$b = c(1 - e^2)^{1/4}$$  \hspace{1cm} (2)

$$c = \sqrt{ab} = \left(\frac{3F_N R}{4E^*}\right)^{1/3}$$  \hspace{1cm} (3)

$$e = \left[1 - \left(\frac{b}{a}\right)^2\right]^{1/2}$$, \hspace{1cm} $b < a$  \hspace{1cm} (4)

$$p_o = \left(\frac{3F_N}{2\pi ab}\right) = \left(\frac{3F_N}{2\pi c^2}\right)$$  \hspace{1cm} (5)

where, $F_N$ is the normal applied load, $E^*$ is the effective young’s modules, $R$ is the relative radius of curvature, $e$ is the eccentricity and $c$ is the equivalent radius.

In order to create experimental conditions as close as possible to the real application, different normal loads were applied to induce Hertzian contact pressure as found in hybrid bearings. Table 2 shows the analytical Hertzian values for different loading.
### Table 2 Analytical Hertzian solution for different normal loads

<table>
<thead>
<tr>
<th>Normal load (N)</th>
<th>Hertzian contact stress [GPa]</th>
<th>Major contact width-2a [µm]</th>
<th>Minor contact width-2b [µm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>3.0</td>
<td>820</td>
<td>400</td>
</tr>
<tr>
<td>850</td>
<td>3.5</td>
<td>980</td>
<td>480</td>
</tr>
<tr>
<td>1250</td>
<td>4.0</td>
<td>1100</td>
<td>540</td>
</tr>
<tr>
<td>2500</td>
<td>5.1</td>
<td>1400</td>
<td>680</td>
</tr>
<tr>
<td>4150</td>
<td>5.9</td>
<td>1660</td>
<td>800</td>
</tr>
</tbody>
</table>

#### 2.2 Material specification and sample geometries

Figure 2 depicts the dimensions of the samples used for carrying out the following experiments. The ceramic rolls were provided by FCT Ingenieurkeramik GmbH, Germany, the commercial designation for this silicon nitride is SN-GP black. The microstructure consists of $\beta$-Si$_3$N$_4$ grains with aspect ratio of 7.4 and ca. 10 wt.-% glassy phase. The sintering additives consist of aluminium oxide (Al$_2$O$_3$) and yttrium oxide (Y$_2$O$_3$). In order to ensure minimal wear on the counterpart; hardened 100Cr6 steel (HV 10 > 800) is used. The physical properties of the both materials are listed in Table 3.
### Table 3: Experimental sample material properties

<table>
<thead>
<tr>
<th>Physical properties</th>
<th>Silicon nitride</th>
<th>Steel (100cr6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density [g/cm³]</td>
<td>3.24</td>
<td>7.83</td>
</tr>
<tr>
<td>Elastic modules [GPa]</td>
<td>300</td>
<td>212</td>
</tr>
<tr>
<td>Possions ratio</td>
<td>0.26</td>
<td>0.29</td>
</tr>
</tbody>
</table>

2.3 **Damage investigation**

Surface damage analysis was carried out by means of optical microscopy, scanning electron microscopy, and surface profilometry. To study damage initiation and progression into the bulk of the silicon nitride material ceramographic cross-sectioning was carried out on selected samples. Focused ion beam (FIB) sectioning was performed to study early stage of damage.

3. **Finite element modelling**

A three dimensional FE-model was constructed to simulate the complex stress states prevailing under rolling contact loading. A quasi-static analysis procedure was employed to initiate contact and to obtain a steady-state contact stress distribution using Abaqus/Standard solver. The geometry consists of a cylindrical silicon nitride roll and a crowned 100Cr6 steel disk. The dimensions of the samples are shown in Figure 2. The physical model was discretised with a fine mesh in the contact region and coarser one was used for the rest of the geometry. Eight-node quadratic brick elements (C3D8) with an element size of 50×50×50 µm³ were used in the refined contact zone.

Initially, both silicon nitride and 100Cr6 steel were modelled as linear isotropic elastic materials. The materials properties are as shown in Table 3. However, at high loads plastic deformation in steel becomes inevitable. In order to approach realistic condition, it is important to incorporate plasticity of steel in the material model. Therefore, a second set of simulations using a rate-independent isotropic material model with plasticity for steel was performed. The uniaxial yield strength value (1100 MPa) of 100Cr6 steel was acquired from [52].
4. Experiment results

Figure 3(a) shows the wear tracks from different normal loads: 500 N (1), 850 N (2), 1250 N (3) and 2500 N (4) after 10 Mrev on a silicon nitride roll. The wear tracks are hardly visible even after 10 Mrev. The edges of the wear tracks were readily detected under optical microscopy Figure 3(b); however, the same wear tracks could not be seen under the SEM. Neither the wear depth nor change in surface roughness was detectable by surface profilometry.

Figure 3. An overview of a silicon nitride roll after lubricated rolling contact (a) image showing the wear tracks from different loading 500 N (1), 850 N (2), 1250 N (3) and 2500 N (4) after 10 Mrev on silicon nitride roll. (b) Depicts the wear track as observed under optical microscope. Note the surface damage in form micro-spalls.

4.1 Surface damage investigation

Figure 4(a) shows the machined surface of the virgin ceramic rolls. Grinding striations from the finish machining are visible and Figure 4(b) depicts the surface with snowflakes in a virgin sample; snowflakes are randomly distributed throughout the material. Figure 4(c) depicts a damaged surface after 10 Mrev for 2500 N in form of minute grain breakdown along the machining striations and Figure 4(d) illustrates material damage in form of micro-spalling on the surface for the same experiment. When we compare the surface integrity before and after running the experiments, it reveals damage that is mainly dominated by micro-spalling. Surface damage showed the same pattern for all loads up to 2500 N.
Figure 4. SEM micrographs depicting (a) surface roughness of a virgin specimen and (b) virgin specimen with a snowflake. (c) Surface damage in form of minute break down of grains at the edges of the machining striations and (d) surface damage in form of micro-spalling i.e., grain pull-out from snowflake region for 2500 N after 10Mrev.

Figure 5(a) depicts the surface damage as observed for 4150 N after 10 Mrev; macro-cracks (c-crack) on the surface were visible. Figure 5(b) shows a magnified view of coalescence of a macro-crack and a micro-spalling region; the machining striations are still visible outside micro-spalling region. Figure 5(c) depicts spalling as appeared on the surface after 10 Mrev; the same pattern was observed in a repeated experiment as well but with different spall size.
Figure 5. SEM micrographs showing surface damage for 4150 N after 10 Mrev. (a) Depicts the wear track with surface damage in form of macro-cracks. (b) Magnified view of a crack interacting with the micro-spalling region. (c) A collaged optical micrograph of spall formed after rolling 10 Mrev.

To study crack initiation, an evolution test was carried out for 4150 N. Initially, several short experiments (7500 rev) were carried out followed by a series of experiments with increasing number of revolutions up to 10 Mrev. In all tests, macro-crack formation was observed between 7500 rev and 2 Mrev. No spall formation was observed up to 8 Mrev. The evolution tests showed that the number of micro-spalls and macro-cracks increased with increasing number of load cycles. Figure 6 depicts the damage patterns as observed from the evolution tests for 4150 N.
4.2 Subsurface damage

4.2.1 Conventional cross-sectioning

Figure 7(a and b) shows SEM images of polished cross-sections from the 30 Mrev experiments at 850 N and 2500 N, respectively. For 850 N after 30 Mrev, the damage mainly appears in form of material degradation as shown in Figure 7(a). The deterioration of the material tends to go deeper in the bulk. In this case, the degradation can be associated with snowflake regions. For this load, running the test until 50 Mrev yielded the same pattern of subsurface damage. Using conventional cross-sectioning, for 2500 N after 30 Mrev, similar subsurface damage in the form of material degradation could not be detected as shown in Figure 7(b). No cracks in the subsurface were visible in polished cross-sections for loads up to 2500 N; however, it was evident that snowflakes are the main source of subsurface damage.
Figure 7. SEM micrographs showing the ceramographic cross-section view of different types of material degradation damage in subsurface (a) for 850 N after 30 Mrev, (b) for 2500 N after 30 Mrev.

Figure 8(a) shows the locations of two cross-sections prepared in the contact track for 4150 N, 10 Mrev. Cross-section CS8-1 was prepared at a location free of any macro-cracks; it revealed no signs of subsurface damage. Cross-section CS8-2 was cut through a macro-crack, a detailed view of it is shown in the Figure 8(b). It reveals that the crack is propagating in the subsurface in the radial direction. It reaches approximately 100 µm in the depth. The crack seems to bifurcate at ca.15 µm below the surface and further in the depth coalesces with the initial crack and further propagates in the depth. Subsurface crack propagation was observed to be both intergranular and transgranular.

Figure 8. SEM micrographs showing (a) the surface damage and the location of the cross-sectioning, (b) damage in the subsurface (i.e. crack initiation, bifurcation, coalescence and propagation) for 4150 N after 10Mrev.
4.2.2 Focused ion beam (FIB) cross-sectioning

It is difficult to detect near surface damage with conventional cross-sectioning, as various polishing processes could induce certain amount of damage in the subsurface. In order to observe minute damage details in the subsurface focused ion beam (FIB) preparation is inevitable. **Figure 9(a)** shows a FIB cross-section in a virgin specimen. It reveals the structure of a snowflake (the presence of grain skeleton and missing glassy phase) in detail as shown in **Figure 9(b)**. In **Figure 9(b)** micro-cracks of less than 2 µm in length can be identified. These micro-cracks may have been induced by machining.

**Figure 9.** FIB-milled section of a virgin specimen (a) subsurface region in a virgin specimen and it can be noticed, the intergranular glassy phase is missing from some regions. (b) Magnified view of the micro damaged zone near surface due to machining.

**Figure 10 (a)** shows the location of a FIB cross-section, which was cut in a region free of surface micro-spallings in the contact track for 850 N after 30 Mrev. It can be noticed that machining striations on the surface are intact indicating minimal wear on the surface. However, the magnified view shown in **Figure 10(b)** reveals the initiation of fatigue induced micro-cracks. Two different micro-cracks, arrows in **Figure 10b**, are shown 10 µm apart. The micro-cracks propagated in the radial direction in the depth as shown in **Figure 10(c)**. Both micro-cracks bifurcate at a depth of ca. 15 µm.
Figure 10. SEM micrograph of subsurface for 850 N after 30 Mrev. (a) FIB-milled section overview, in the magnified view (b) two micro-cracks can be identified. A magnified view shows (c) the micro-cracks originated from the surface and (d) the crack bifurcated at almost the same depth.

Additional FIB cross-sectioning was carried at various locations in the contact track for 850 N after 50 Mrev, Figure 11(a) depicts a schematic representation of the locations (distance from the edge of the contact track) where the FIB cross-sections were cut. Even after repeated FIB analyses at different locations for this case, neither micro nor macro-cracks were detected in the subsurface. Initially a FIB section was prepared just outside a micro-spall at Location 1 (250 µm from one side of the contact edge). Figure 11(b) depicts the cross-section at Location 1. As expected, snowflakes appearing on the surface can also be seen in the subsurface. Another FIB cross-section was created at Location 2, exactly in the middle of the micro-spalling region, Figure 11(c) illustrates the micro-spalling in the subsurface and a network of missing glassy phase can be identified with few grain breakdowns. FIB cuts at Location 3, 4 and 5 yielded similar patterns of damage as observed in Location 1 and 2.
Figure 11. (a) Schematic drawing showing FIB cut at different distance and on specified locations on micro-spalls in the contact track for 850N after 50 Mrev. (b) The magnified FIB SEM micrographs depict the subsurface damage at Location 1(b) and Location 2(c).

Furthermore, FIB cross-sectioning was performed on micro-spalling for 2500 N after 10 Mrev. Its size was larger than the rest of the micro-spalling-spalls on the surface. Figure 12(a) depicts the FIB cross-section. It can be seen that a crack has initiated and propagated in the depth. Moreover, wear debris has accumulated within this region. The micro-spalling damage in the subsurface is confined within the snow flaking site. An EDX analyses was carried out in the subsurface at Location 1, Figure 12(b). Figure 12(c) depicts a mapping of the elemental composition as detected in the subsurface at the crack opening. The analysis revealed Si, O, N and Al
which are the primary constituents of the silicon nitride in addition to Fe. Figure 12(b) shows that the crack has initiated at the bottom of the micro-spall. It can be noticed that the crack faces are separated in the subsurface and it is worth mentioning that the crack is propagating at the boundary of the snowflake. In order to identify whether the crack is propagating further into the depth or back to the surface, an additional adjacent FIB cut was prepared as shown in Figure 12(d). A magnified view as shown in Figure 12(e) indicates that the crack is still propagating at an inclined angle into the bulk of the material. An EDX mapping was performed at Location 2, Figure 12(e) (near the crack tip). Figure 12(f) shows the corresponding elemental composition. The analysis shows similar elemental composition as in Location 1; however, in this case, Fe was not detected near the crack tip. Additionally, the quantitative EDX analysis indicated an increased O concentration at Location 2 in comparison to Location 1.
Figure 12. SEM micrographs for 2500 N after 10 Mrev (a) showing the initiation of macro-crack location and wear debris accumulation. (b) Magnified view showing the crack propagation along the edges of snowflake boundary. (c) EDX mapping (Location 1) at opening of the crack. (d) Adjacent FIB-milled section, to identify the path of the crack propagation and (e) shows that the crack is still propagating in the bulk of the material. (f) EDX mapping (Location 2) near the crack tip.
5. FE-Simulation Results

Lubricated rolling contact is comparable to the Hertzian elastic problem. Therefore, initially an FE model with linear elastic material behavior was compared with the Hertzian contact solution for verification purposes. The results obtained from the FE simulations were within an acceptable error margin (not exceeding 5 percent) of the Hertzian solution. Subsequently, a simulation model was constructed by incorporating plasticity in the mechanical behavior of steel. This is essential in studying the influence of plastic deformation on the distribution and value of the contact pressure and consequently, on the contact area.

Figure 13 depicts the results obtained from the Hertzian solution compared with both FE solutions (fully elastic and elastoplastic). The influence of plasticity on the contact is clearly evident; the difference in contact pressure resulting from a fully-elastic contact and that resulting from a model incorporating plasticity increases with increasing applied load. Additionally, the FE generated contact pressure contour plots shown in Figure 13 demonstrate that plastic deformation in steel results in an increased contact area.

Figure 14(a) shows the contact track width as a function of normal load obtained from experiments (after 10 Mrev) and from FE simulations. The track width was obtained from the FE simulations by calculating the minor contact ellipse radius $a$. The results show an increased contact track width with increasing applied normal load. The experimental results are in good agreement with the elastoplastic FE solution up to 2500 N. For the case 4150 N, the contact track width widens further beyond what is computed using the elastoplastic FE model.
The growth in the track width for 4150 N was measured by running experiments for shorter intervals. **Figure 14(b)** depicts the growth of the track width with increasing number of revolutions between 7500 rev and 10 Mrev. The track width measured after 7500 rev was around 900 µm, which is in good agreement with the elastoplastic FE model.

![Figure 14. (a) Comparison of experimental (after 10 Mrev) and numerical simulation contact track width as function of different loading. (b) The evolution of contact track width with number of revolutions, starting from 7500 revolutions to 10 Mrev for 4150 N.](image)

### 6. Discussion

The main objective of this work is to investigate the damage initiation and evolution under different contact loading conditions. With the absence of gross slip, lubricated rolling contact resulted in hardly any measurable wear on the surface, nonetheless, contact track widening was observed with increasing load. The damage was mainly in form of micro-spalling, micro-crack formation, macro-crack (c-crack) formation and spalling in certain cases.

#### 6.1 Evolution of contact geometry and stresses

After 10 Mrev, a comparison between the contact track widths measured from experiments and those calculated using the Hertzian contact solution for different loads demonstrates that only the lowest load i.e., 500 N is in agreement with the Hertzian solution (**Figure 14**). Starting from 850 N, an obvious deviation/widening in the contact track width becomes evident. The comparison between the contact track widths generated at the same loads after 10 Mrev and 50 Mrev for 850 N and after 10 Mrev and 30 Mrev for 2500 N showed no further widening. This implies that up to 2500 N the contact track width is independent of the number of load cycles.
For the highest load (i.e., 4150 N), a 14% increase in the contact track width is experimentally observed between 7500 rev (short experiment) and 10 Mrev.

The FEM results suggest that the track width widening up to 2500 N is mainly due to plastic deformation in steel. The FEM simulations show that, at any particular load, the contact pressure decreases with the increase in contact area due to plastic deformation in steel is taken into consideration; this is particularly true at initial contact before rolling contact commences. The extent of increase in the contact track width at initial contact depends on the maximum applied load.

The contact pressure distribution in the Hertzian solution is parabolic in nature with a maximum pressure at the centre of the contact ellipse that gradually decreases towards the edges. Nonetheless, it was experimentally observed that the steel discs undergo very slight wear that is even difficult to quantify; hence, it is safe to assume that any considerable change in the contact track width and in contact stresses should be due to the accumulative plastic deformation in steel. However, when we compare the contour plots of contact pressure distribution from the FE simulations with and without plasticity (Figure 13), we observe that the contact pressure distribution progressively deviates from its original parabolic distribution (observed in a fully elastic contact) with increased plastic deformation caused by the increase in normal load. Thus, due to the flattened contact pressure distribution, it is expected that a larger volume of silicon nitride will undergo high stresses.

For the lowest load (i.e., 500 N), the contact ellipse retains its original geometry as plastic deformation is negligible. When we compare the experimental contact track width for the highest load (i.e., 4150 N) after 10 Mrev with the FE simulations with plasticity there appears to be a large difference between the track widths. This behaviour was explained by the evolution test results for 4150 N. Upon the initialization of contact and after 7500 rev (Figure 14(a and b)) the contact width for the short period test is in accordance with FE simulation with plasticity. With increasing the number of load cycles the contact track widens further due to accumulation of plastic strain in steel.

### 6.2 Surface damage

Up to 2500 N the surface damage was mainly in form of micro-spalling (Figure 4d) and for 4150 N in combined form of micro-spalling and macro-crack (c-cracks) formation (Figure 5). Micro-spalls were observed
only in the region where snowflakes are present. The Snowflakes form three-dimensional structures in the material. They are expected to act as stress raisers and they disturb the symmetric stress distribution in their surroundings.

It was observed that the micro-spall density on the surface depends mainly on the snowflake density undergoing rolling contact. Moreover, it was observed that snowflakes outside the contact zone have no influence on damage development. Further, when a snowflake undergoes a cyclic alternating stress field (tension-compression-tension), grains are pulled out from this region (Figure 3b). Consequently, grain pull-outs appear as micro-spalls on the surface; their average size and number depend on the applied normal load and number of stress cycles. For the lowest load (500 N) and even after 10 Mrev hardly any micro-spalling was detected on the surface despite of the snowflake population undergoing rolling contact. This implies that the stresses induced by this normal load are not high enough to initiate damage. Starting from 850 N and after rolling 10 Mrev a slight increase in the number of micro-spalls was observed; the micro-spall density showed an increase with increasing number of revolutions (loading cycles). A similar behaviour was observed in the case of 2500 N as well. On the other hand, for the highest load (4150 N) after rolling 7500 rev (i.e., short experiments) few micro-spalls were observed and with increasing number of revolutions the micro-spall density increased and c-cracks appeared on the surface (Figure 5).

Concerning ceramics undergoing rolling contact, there is general agreement that crack propagation is mainly driven by tensile stresses. The peak value of tensile stresses appears at the major contact ends as shown in Figure 15 and gradually decreases towards the minor contact ends. The tensile stresses are strongly dependent on the contact pressure as well as the contact area. Slight changes in non-conforming contact conditions could affect the tensile stress both on the surface (cf. Khader et al. [23]) as well as in the subsurface. The absence of contact track widening with increasing number of revolutions for all loadings below 2500 N indicates that the contact area and contact pressure remain almost constant. Considering the Hertzian contact solution (Figure 13), the difference in contact pressure between 2500 N and 4150 N is about 16%, which falls down to 3.7% with plastic behaviour of steel taken into account. Nevertheless, experimentally the damage behaviour under both loads was essentially different; this could be attributed to the difference in development of tensile stress.
The tensile stresses developed under static contact up to 4150 N are not high enough to cause spontaneous crack initiation on the surface. However, once rolling contact commences the increased contact area brings more natural flaws under contact, thus, increases the likelihood of crack propagation on the surface following the tensile stress trajectories. A comparison between the FEM computed tensile stresses and the experimentally observed crack shapes is shown in Figure 15. The figure illustrates the possible crack propagation path, which seems to be dominated by the tensile stresses at the leading edge rather than the trailing edge of the contact.

![Figure 15. A comparison of FEM regenerated contours of contact and maximum tensile stresses with the Hertzian c-crack on the silicon nitride surface for lubricated rolling contact for 4150 N after 10 Mrev.](image)

### 6.3 Subsurface damage

The subsurface damage analysis from conventional cross-sectioning for 850 N after rolling 30 Mrev and 50 Mrev revealed similar patterns of damage as shown in Figure 7a. The damage was mainly in form of material degradation zones, which were confined to individual snowflake regions. However, the same conventional cross-sectional analysis for 2500 N after rolling 30 Mrev showed no signs of damage even when snowflakes existed in a near-surface region. FIB investigations revealed a different scenario altogether; no subsurface material degradation was evident. This can be explained in light of the cross-sectioning technique employed. Virgin sample prepared by both methods resulted in damage only being observed in the samples prepared by conventional cross-sectioning. Due to the absence of the glassy phase in snowflakes, grains may be prone to detachment from the matrix during mechanical polishing in conventional preparation method as opposed to FIB cross-sectioning, in which the true structure of snowflakes in the subsurface can be revealed, see Figure 9a and b.
Despite of inevitable damage caused by employing conventional cross-sectioning, this cannot be considered the sole reason for the appearance of subsurface damage in form of material degradation zones as observed for the case of 850 N after 30 Mrev and 50 Mrev as damage can be attributed to loss of material strength in snowflakes due to cyclic loading. The fact that no damage was detected for 2500 N after 30 Mrev may be attributed to the scatter in the strength of ceramics (this material has a room-temperature characteristic 4-point-bending strength of $\sigma_0=703.6$ MPa and a Weibull modulus $m=17.6$; tested in air at a load rate of 200 N/s according to EN 843-1) [53], which is a function of many factors such as the natural flaw density, its distribution, size and shape.

The FIB cross-sectioning on micro-spalls after 50 Mrev for 850 N (Figure 11b) and on one particularly large micro-spall after rolling 10 Mrev for 2500 N (Figure 12) shows an indication of difference in the damage mechanism. In both cases, the damage is mainly caused by lubricant penetration in the micro-spall. In the latter case, a macro-crack has been initiated and propagated in the subsurface. The crack faces appear to be separated from each other, which indicate that the lubricant has entered into the crack opening. This can be verified by the presence of Fe in the EDX quantitative analysis (Figure 12) which is the primary constituent of the steel disk (100Cr6 steel). Fe appears in the wear debris particles that are assumed to be trapped in the micro-spall. Furthermore, the crack propagation at an inclined angle suggests hydrostatic-pressure driven crack propagation created by entrapped lubricant in between the crack faces, this has been described in [42–45,54]. The EDX analysis near the crack tip revealed no significant Fe, however, indicated increased O concentration near the crack tip when compared to the crack opening. This behaviour can be attributed to the difference in lubricant pressure and the crack face roughness. At the crack opening, the hydrostatic pressure is high enough to keep the crack face separated during roll-over. On the other hand, as the crack depth increasers the lubricant pressure drops and as a result the crack faces slide against one another, thus, increasing friction near the crack tip. Thus, this increased friction may be the reason for increased oxidation. This behaviour has not been observed for lower loads. A possible explanation might be that the lubricant pressure is not high enough to initiate and propagate cracks.

The subsurface investigation by FIB sheds light on the near-surface damage that cannot be detected using conventional cross-sectioning methods. Micro-cracks initiated from the surface were observed for 850 N after
rolling 30 Mrev (Figure 10). The locations of these micro-cracks were free from snowflakes, thus, indicating propagation due to surface fatigue under rolling contact. The propagation of these micro-cracks in the subsurface was confined within 20 μm from the surface. The retardation of micro-cracks in the subsurface may be related to the steep tensile stresses gradient in the depth. However, no such micro-cracks were detected for the same loading after 50 Mrev. This indicates that the previously observed micro-cracks have initiated due to the presence of pre-existing surface flaws which propagated in the subsurface due to cyclic loading. The propagation of natural flaws in ceramics is mainly dependent on the stress intensity factor which is a function of their shape and location with respect to the contact zone (the latter was studied in [32]).

For 4150 N, c-cracks on the surface propagate in the subsurface as well. The measured crack lengths in the subsurface were mainly dependent on the location of the performed cross-section. The crack length in the subsurface was smallest when the cross-section was cut near its tip at the surface (Figure 8a) and it increases in length as we proceeded towards the edge of the contact. Despite, the drastic reduction of tensile stresses in the depth, cracks propagating down to 100 μm in the subsurface were evident, thus, implying fatigue fracture after a low number of cycles (between 7500 rev and 2 Mrev).

7. Conclusion

The damage mechanisms of silicon nitride undergoing non-conforming lubricated hybrid rolling contact was studied by means of a model experiment on a twin-disk tribometer. The following was concluded:

- Surface damage was dominated by the formation of micro-spalls up to a certain load (i.e., 2500 N), above which it appeared in form of macrocracks and micro-spalls.
- Micro-spalls formed due to grain pull-outs from the snowflake structure undergoing rolling contact (i.e., alternating cyclic tension-compression fields). The micro-spall density (i.e. number of micro-spalls on surface) mainly depends on the applied normal load and the number of stress cycles.
- Surface crack formation for highest load was not spontaneous but due to material fatigue.
- The increase in contact area (track width) is mainly due to plastic deformation in the steel disk upon initial contact; for the highest applied load accumulation of the plastic strain further increases the contact area. The increased contact area brings more volume of silicon nitride under rolling contact.
- The crack initiation and propagation in silicon nitride were attributed to two different mechanisms: (i) Fatigue induced crack propagation and (ii) lubricant driven crack propagation.

Acknowledgements

The authors would like to thank the European Union’s Seventh Framework Programme FP7/2007–2013 under Grant Agreement No. 263476 (RoLiCer Project) for the financial support.

8. References


R. Hanzal, Rolling contact fatigue failures in silicon nitride and their detection, PhD Thesis-University of Southampton, 2013.


EU-FP7-NMP, Enhanced reliability and lifetime of ceramic components through multi-scale modelling of degradation and damage RoliCer, ID: 263476, 2015.