Performance of integrated retainer rings in silicon micro-turbines with thrust style micro-ball bearings

2013 J. Micromech. Microeng. 23 065033

(http://iopscience.iop.org/0960-1317/23/6/065033)

J. Micromech. Microeng. **23** (2013) 065033 (11pp) doi:10.1088/0960-1317/23/6/065033

Performance of integrated retainer rings in silicon micro-turbines with thrust style micro-ball bearings

Robert J Hergert[1], **Brendan Hanrahan**[2,3], **Reza Ghodssi**[2] **and Andrew S Holmes**[1,4]

[1] Department of Electrical and Electronic Engineering, Imperial College London, Exhibition Road, London SW7 2AZ, UK
[2] Department of Electrical Engineering, University of Maryland College Park, College Park, MD, USA
[3] The US Army Research Laboratory, Adelphi, MD, USA

E-mail: a.holmes@imperial.ac.uk

Received 28 December 2012, in final form 12 April 2013
Published 16 May 2013
Online at stacks.iop.org/JMM/23/065033

Abstract
This work explores the performance of different silicon retainer ring designs when integrated into silicon micro-turbines (SMTs) incorporating thrust style bearings supported on 500 μm diameter steel balls. Experimental performance curves are presented for SMTs with rotor diameters of 5 mm and 10 mm, each with five different retainer designs varying in mechanical rigidity, ball pocket shape and ball complement. It was found that the different retainer designs yielded different performance curves, with the closed pocket designs consistently requiring lower input power for a given rotation speed, and the most rigid retainers giving the best performance overall. Both 5 mm and 10 mm diameter devices have shown repeatable performance at rotation speeds up to and exceeding 20 000 RPM with input power levels below 2 W, and devices were tested for over 2.5 million revolutions without failure. Retainer rings are commonly used in macro-scale bearings to ensure uniform spacing between the rolling elements. The integration of retainers into micro-bearings could lower costs by reducing the number of balls required for stable operation, and also open up the possibility of 'smart' bearings with integrated sensors to monitor the bearing status.

(Some figures may appear in colour only in the online journal)

1. Introduction

The development of microengineered bearings capable of long-term operation would open up new applications for MEMS technology, particularly in the areas of micropower generation and microfluidics. Research on micro-scale bearings dates back to the late 1980s when the first silicon rotary micromachines were demonstrated [1]. These devices had simple journal or thrust bearings with sliding silicon contacts which showed high rates of friction and wear. Subsequent research on bearings of this type focused on the use of self-assembled monolayers [2] and dry coatings such as diamond-like carbon [3] to reduce sliding friction; however, such coatings suffer from degradation under load and do not allow long-term operation. More recently attention has turned to vapour- [4] and liquid-phase [5] lubrication methods for high-sliding contacts as these allow replenishment of the lubricant. In parallel with the above, levitation schemes based on electrostatic, magnetic or hydrostatic forces have been developed [6–9]. These are highly effective but require relatively complex control systems and an auxiliary supply of pressurized fluid or electrical power. Aerodynamic bearings avoid this complexity but can be made to work only at extremely high rotation speeds.

Microengineered rolling element bearings could provide a viable alternative to the above approaches for applications involving low or moderate rotation speeds (up to ca 50 000 rpm). Rolling element bearings achieve low friction

[4] Author to whom any correspondence should be addressed.

and wear by reducing the degree of sliding at the load-bearing contacts; moreover, they do not require any external control system or power supply and they can operate over a relatively wide range of speeds. Silicon MEMS rolling element micro-bearings have been under development at the University of Maryland since the early 2000s, and in the last few years this technology has matured to the point where micro-ball bearings can be integrated into functional devices (see, for example, [10]). The bearing raceways in these devices are formed by etching annular channels in a pair of silicon wafers. The micro-balls—typically stainless steel—are placed manually in the channels on one wafer before the two wafers are bonded together; a final release etch is then used to free the moving part (inner ring) of the bearing.

This paper investigates the effect of incorporating silicon retainers into silicon MEMS micro-ball bearings of the kind reported in [10]. The retainer or 'cage' is an annular structure with regularly spaced pockets to accommodate the balls, and its purpose is to define and maintain the spacing between the balls on the raceways. Retainers are ubiquitous in traditional ball bearings. They help to ensure even load distribution, and prevent collisions between the balls which can lead to increased losses and wear. The retainer should be less hard than the rolling elements, and common retainer materials are pressed steel (low-cost bearings), bronze or brass [11]. Polymers are also used in bearings not intended for high temperature operation [12] and can offer advantages such as reduced friction and lower (cage-ball) collision noise, in addition to lower production costs. The greater mechanical flexibility of polymer cages can also be an advantage as it lowers the ball-pocket forces [13].

In addition to potentially improving performance, retainers could lower the cost of silicon micro-ball bearings by reducing the number of precision micro-balls required; in the absence of a retainer, the raceway needs to be substantially full in order to avoid excessive vibration and loss of load-bearing capacity if the balls become redistributed, whereas with a retainer the ball complement can be as low as 50%. While the cost saving would be negligible for the large devices explored in this work, this consideration becomes increasingly relevant with down-scaling, because the micro-balls become more expensive while the cost per die for microfabrication decreases. In the longer term, retainers could also facilitate the integration of intra-bearing sensors, opening up the possibility of new types of 'smart' bearing that can provide real-time information about the bearing status, for example, measurements of temperature and/or vibration. Such data might be useful both for tribological studies and for condition monitoring when the bearing is in service. Macro-scale smart bearings have been demonstrated, such as the one in [14]. Here the outer ring of a bearing was modified to incorporate a piezoelectric sensor that could monitor dynamic load variations. In a silicon microengineered bearing, sensors with wireless power delivery and data transfer could potentially be integrated directly into the retainer, providing additional data on the moving parts.

The aim of this work was to compare several different designs of retainer with a view to identifying the most promising design or designs for further study. Silicon micro-ball bearings with integrated retainers were first reported in [15]. These devices were radial bearings, and the design was such that the retainer was formed from two parts held together by a solder bond in the plane of the bearing. Unfortunately this bond proved to be a weak point and consequently the bearings tended to fail after only a short lifetime (several hours). In contrast, the bearing investigated in the present paper is a thrust-style bearing which lends itself to incorporation of a monolithically fabricated silicon retainer. To allow functional testing the bearing is integrated with a silicon micro-turbine (SMT), following [16]. This platform was chosen as it provides a proven method for actuation of the bearing during characterization. Several different retainer designs were investigated in turbines of two sizes, with 5 mm and 10 mm diameter turbine rotors, respectively. Photographs of devices of both sizes, together with a cut-away schematic view showing the retainer, are shown in figure 1. The different designs were compared in terms of losses, inferred from the variation of turbine input power with rotation speed; measurements of repeatability and longevity were also made on selected devices.

An earlier investigation of different retainer geometries in 5 mm-diameter SMTs was reported by us in [17]. Since the publication of the previous work, the fabrication process for the devices has been improved to allow us to successfully fabricate and test devices with a 10 mm rotor diameter. This was achieved by making the turbine with a 1 mm thick silicon wafer rather than a 500 μm thick wafer and changing the etching steps to better account for etch lag. These changes allowed us to increase the thickness of silicon above the bearing raceway, making it sufficient to withstand the forces of bearing operation in larger devices.

2. Design

2.1. Retainer ring design

Figure 2 shows a cut-away schematic view of a conventional thrust bearing incorporating a so-called ball-riding (BR) retainer [11]. In this style of bearing, the retainer is supported entirely by the rolling elements which are held captive in suitably shaped pockets, and consequently there is no contact between the retainer and the bearing rings. Alternative designs for radial/mixed load bearings employ inner ring (IRL) or outer ring (ORL) land riding retainers which are sized to fit the cylindrical surface of either the inner or outer bearing ring [18]. IRL retainers are driven by a friction force between the retainer and the inner ring, and under optimal conditions this can result in negligible loading of the rolling elements by the retainer. ORL retainers are guided by the outer ring and are therefore subject to a drag force; this type of retainer tends to be used in high-speed applications.

Five different designs of the retainer ring were investigated in the present work, designated as Full Ring, Full Skeleton, Half Skeleton, Outer Open and Inner Open. Scanning electron microscope (SEM) images of all five retainer types are shown in figure 3(*a*), while figure 3(*b*) presents schematic cross-sections for the different retainer types. The left-hand

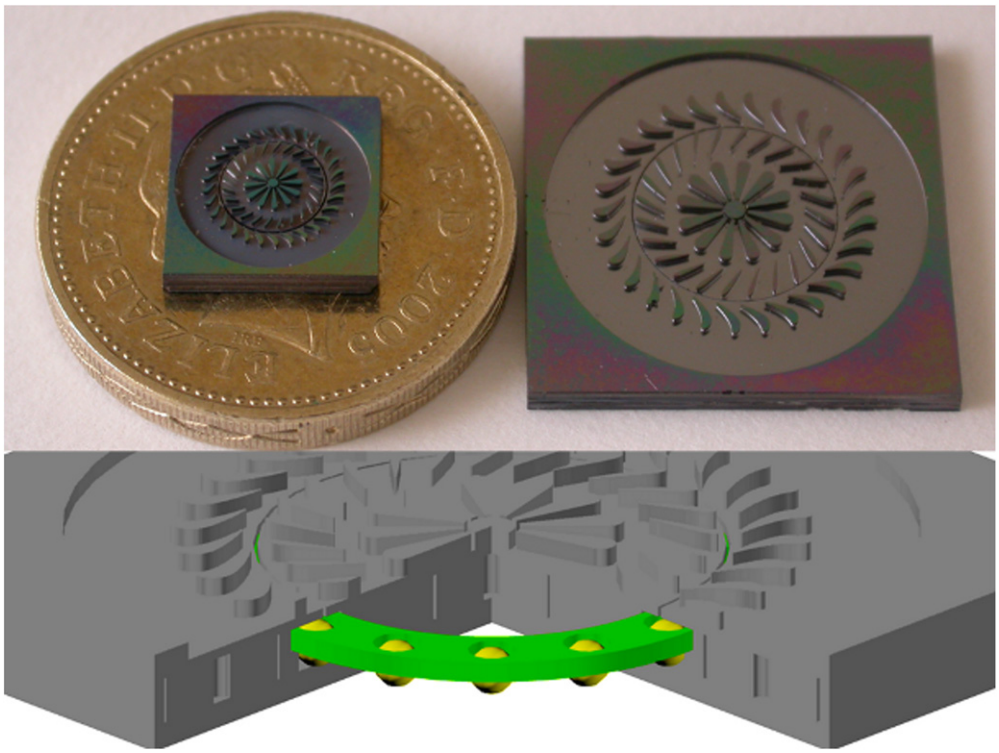


Figure 1. Photographs of both the 5 mm and 10 mm devices with a British Pound coin for scale (top), and a cutaway view of the device showing the retainer ring (bottom).

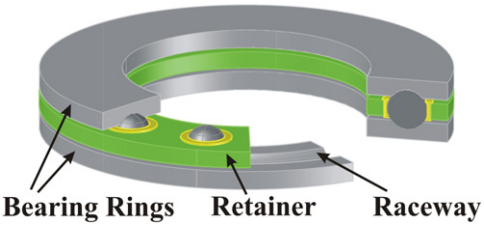


Figure 2. Cut-away schematic view of a conventional thrust bearing with a ball-riding retainer (after [11]).

schematic also illustrates how the bearing and micro-turbine are integrated; the device is assembled at die level as a bonded two-die stack, comprising a lower 'thrust' die and an upper 'turbine' die. The bearing is formed at the interface between the two dies, with the rotor blades and guide vanes of the turbine being defined in the upper surface of the turbine die.

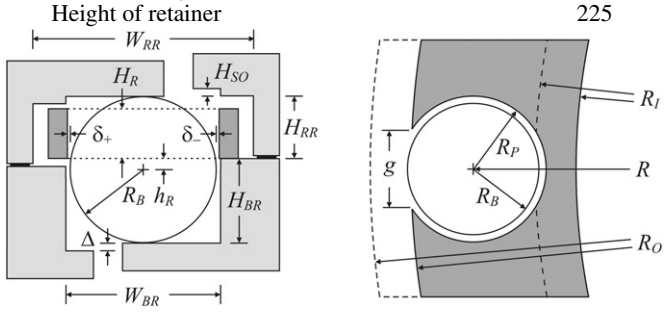
Comparing figures 2 and 3, two important differences, both arising from micro-fabrication process constraints, can be seen between the MEMS thrust bearing and its conventional counterpart. Firstly, the raceways on the conventional bearing are curved in cross-section so that the balls have a single point of contact with each raceway and the bearing is inherently self-aligning when subject to an axial load. In contrast, the silicon raceways are rectangular, with the balls riding on essentially flat surfaces top and bottom. The balls are constrained to follow a nominally circular path by the sidewalls of the bearing raceway, with which they make intermittent contact, and also by the retainer. Secondly, while the pockets in the conventional retainer are shaped to enable it to ride on the balls, this cannot be achieved with the silicon retainer because the pockets are cylindrical. Instead the Inner Open retainer rides on the land adjacent to the bearing raceway on the outer (stator) side, while the Outer Open retainer is supported by the land on the inner (rotor) side. The inner land is slightly elevated with respect to the outer, and consequently the Full and Skeleton retainers should ride in the same way as the Outer Open type. However, if the height difference is sufficiently small then bearing vibration and dynamic distortions of the retainer may bring it into contact also with the outer land. The different silicon retainers can reasonably be classified as either inner ring or outer ring land riding, though they differ from conventional IRL and ORL retainers in that they ride on flat surfaces in the plane of the bearing rather than on cylindrical surfaces. It should also be noted that interactions between the retainer and the balls when the bearing is operational may cause the retainer to ride up and make contact with the top of the retainer raceway; the top of the raceway is recessed to provide a stand-off and reduce the contact area under these conditions.

Key design parameters for the bearings developed in this work are given in table 1. The width of the bearing raceway was set at 510 μm giving a clearance of 5 μm either side of the 500 μm diameter steel balls. This clearance determines the lateral play in the rotor position, and the value chosen was the smallest that would guarantee easy insertion of the balls, taking into account fabrication tolerances. The width of the retainer raceway was chosen to give a clearance of 25 μm either side of a Full Ring retainer, ensuring that no retainer-raceway sidewall contact would occur with any of the designs.

The bearing raceway was made slightly higher than the ball radius to ensure a planar contact between the ball and the raceway sidewall. The height of the retainer raceway was then

Table 1. Key design parameters for large (10 mm) and small (5 mm) devices. All dimensions are in μm.

Parameter	Description	Large device	Small device
R	Radius of ball path	5000	2600
R_I	Retainer inner radius – full, outer open	4647.5	2247.5
	– inner open	4787.5	2387.5
R_O	Retainer outer radius – full, inner open	5352.5	2952.5
	– outer open	5212.5	2812.5
g	Ball pocket opening – outer open	278	284
	– inner open	267	261
R_B	Ball radius	250	
R_P	Pocket radius in retainer	252.5	
W_{BR}	Width of bearing raceway	510	
W_{RR}	Width of retainer raceway	755	
H_{BR}	Height of bearing raceway	260	
H_{RR}	Height of retainer raceway (incl solder)	245	
H_{SO}	Stand-off height	5	
H_R	Height of retainer	225	



set so that the combined height of the raceways, including the thickness of the solder bond between the wafers and the recess in the top of the retainer raceway, was nominally 10 μm larger than the ball diameter, i.e.

$$\Delta = H_{BR} + H_{RR} + H_{SO} - 2R_B = 10\ \mu\text{m} \tag{1}$$

where all the variables are as defined in table 1. In choosing the design value of Δ, tolerances in etch depth and solder bond thickness were taken into account. It is essential that $\Delta > 0$ so that there is clearance above the balls while the device is being assembled.

2.2. Design variations

The five retainer designs were chosen to explore the effects on the bearing performance of sliding friction, retainer rigidity, ball pocket shape and ball complement. The differences between the designs are summarized in table 2. Considering first sliding friction, this is expected to be most significant in the Full Ring design which has the largest overlap area with the inner and outer lands. The Skeleton designs reduce the overlap area on either side by removing material between the ball pockets, while the Open designs eliminate it entirely on one side of the raceway. The removal of material from the Skeleton and Open designs will also reduce the mechanical rigidity of the retainer. This has been found to improve the performance in conventional bearings under some loading conditions [13].

In addition to distributing the balls uniformly around the raceway, the retainer will also limit the radial excursions of the balls, encouraging them to follow a stable circular trajectory and reducing the extent to which they interact with the raceway sidewalls; this is expected to be beneficial in terms of friction and wear. The radial play allowed by the ball pockets will differ for the Full/Skeleton and Open designs. Considering first the Full/Skeleton retainers, and referring to the schematic in table 1, the radial play in the cylindrical pockets is expected to be

$$\delta_\pm = R_P - \sqrt{R_B^2 - h_R^2} \tag{2}$$

where R_P is the pocket radius, R_B is the ball radius, and h_R is the height of the retainer above the centre of the ball. The value of h_R when the bearing is at rest depends on the height of the bearing raceway. However, during operation h_R can lie anywhere in the range:

$$(H_{BR} - R_B) \leqslant h_R \leqslant (H_{BR} - R_B) + (H_{RR} - H_R - \Delta) \tag{3}$$

Using values from table 1, equations (3) gives 10 μm $\leqslant h_R \leqslant$ 20 μm. The maximum play predicted by equation (2) is then $\delta_\pm = 3.3$ μm. This is larger than the difference between the pocket and ball radii because the retainer sits above the centre of the ball.

The open pockets on the Inner Open and Outer Open designs will allow the balls more radial play. The generalization of equation (2) in the case where the pocket has an opening on one side is

$$\delta_\pm = \sqrt{R_P^2 - (g/2)^2} - \sqrt{R_B^2 - h_R^2 - (g/2)^2} \tag{4}$$

where g is either the opening width (open side of pocket) or zero (closed side). For example, for the large Inner Open retainer, $g = 267$ μm and with $h_R = 20$ μm equation (4) gives $\delta_+ = 3.3$ μm and $\delta_- = 3.9$ μm. The other Open designs have similar gaps and hence will yield similar results.

The above calculations ignore the effects of wear at the bottom of the retainer, and so apply only when the bearing

Table 2. Comparison of retainer properties.

Name	# of Balls 10 mm/5 mm	Ball fill factor	Contact area	Retainer rigidity	Ball pocket play
Full Ring	32/16	∼50%	High	High	Low
Full Skeleton	32/16	∼50%	Medium	Medium	Low
Half Skeleton	16/8	∼25%	Very Low	Medium	Low
Outer Open	32/16	∼50%	Low	Low	Higher
Inner Open	32/16	∼50%	Low	Low	Higher

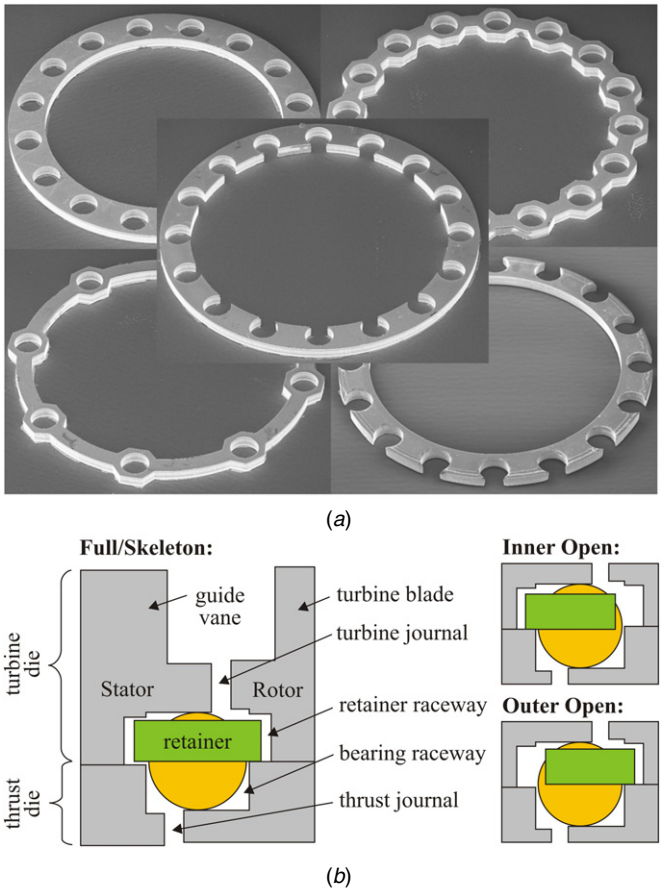


Figure 3. (*a*) SEM images showing the 5 mm retainer ring designs: Full Ring (top left); Full Skeleton (top right); Half Skeleton (bottom left); Outer Open (bottom right); Inner Open (centre). (*b*) Schematic cross-sections of bearings with the different retainer types.

is newly fabricated. Over time the ball pockets will become enlarged due to abrasion by the balls, and this will increase the ball pocket play for all designs. It is expected that this effect will be more pronounced for the Open designs, since the edges of the openings are likely to exhibit higher rates of wear, so that the difference in play between the Full/Skeleton and Open designs will become more pronounced throughout the bearing lifetime.

The ball complements for the Full Ring, Open and Skeleton designs were 32 balls for the large devices and 16 balls for the small devices, corresponding in each case to a fill factor of about 50%. These values were halved in the Half Skeleton designs. Increasing the number of balls is generally beneficial for performance because it lessens the load per ball, stiffens the bearing and reduces vibration [19]. It was therefore expected that losses would be higher in devices with Half-Skeleton retainers.

2.3. SMT design

The micro-turbine design adopted for the 10 mm diameter devices was as described in [16], and this design was simply scaled down for the 5 mm-diameter devices. No attempts were made to optimize the turbine performance. However, the bearing housing was altered to accommodate larger 500 μm diameter balls, primarily to ease the assembly of the devices. Also the turbine wafer thickness was increased to 1 mm so that it could accommodate the retainer raceway on the back side and the rotor blades and guide vanes on the front side (see figure 3(*b*)) while leaving a sufficient thickness of silicon in between to withstand the thrust force applied to the bearing during operation. The process flow reported in [17] did not leave a sufficient silicon thickness above the bearing to support the normal load during operation, resulting in rapid failure of 10 mm diameter devices due to silicon fracture.

3. Fabrication processes

3.1. Retainer ring fabrication

The retainer rings are fabricated in a 225 μm thick, 100 mm diameter silicon wafer with a 1 μm-thick thermal oxide (i.e. thermally grown silicon dioxide) layer on both sides. Photoresist (PR) is spin-coated onto the wafer and patterned by photolithography, and reactive ion etching (RIE) is used to etch the exposed oxide. Deep reactive ion etching (DRIE) is then used to etch through the wafer to release the retainer rings. The oxide protects the ring in the final stages of the SMT fabrication process when DRIE etches are used to release the turbine rotor. Previous designs lacking the oxide layer showed etch damage to the retainer caused during the release etch.

3.2. SMT fabrication

The fabrication of the SMT requires the stacking of two dies—an upper turbine die and a lower thrust die—which are fabricated on separate wafers. Due to the complex nature of the fabrication procedure, it is convenient to break it into three distinct process flows: thrust wafer fabrication, turbine wafer fabrication, and assembly and release. The following sections describe each of these stages in detail.

3.2.1. Thrust wafer fabrication. The process flow used to fabricate the thrust wafer is shown in figure 4. The thrust

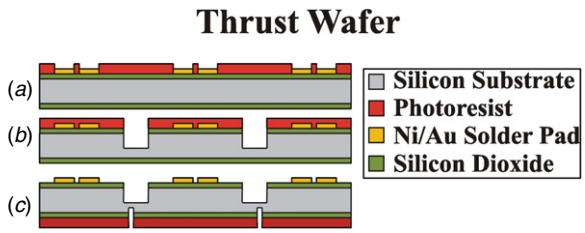


Figure 4. Thrust wafer process flow. (*a*) Ni/Au solder pads deposited. (*b*) Oxide patterned and bearing raceway etched by DRIE. (*c*) Back side oxide patterned and thrust journal partially etched.

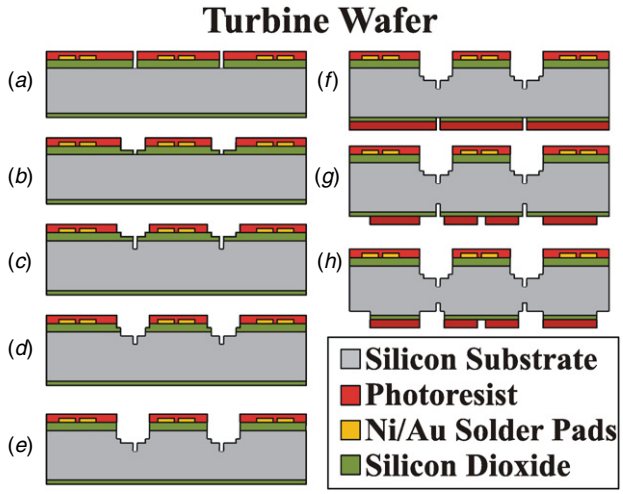


Figure 5. Turbine wafer process flow. (*a*) Oxide patterned with turbine journal. (*b*) Oxide thinned by 200 nm in recess regions. (*c*) Cage raceway defined in PR; turbine journal etched to 200 μm depth. (*d*) Thinned oxide removed and recess etched to 5 μm depth. (*e*) Remaining exposed oxide removed and retainer raceway etched to 240 μm depth. (*f*) PR deposited on back side and oxide patterned with turbine journal. (*g*) Turbine structure defined in PR; turbine journal etched to 250 μm depth. (*h*) Exposed oxide removed and turbine blades etched to 250 μm depth.

wafer is made from a double side polished, 500 μm thick, 100 mm diameter silicon wafer with 1 μm of thermal oxide on each side. The wafer is first coated with 30 nm of chrome and 60 nm of copper to provide an electroplating seed layer. PR is then spin-coated onto the wafer and patterned to define the solder pads which are formed by electroplating 2 μm of nickel followed by 3 μm of tin and 300 nm of gold (A in figure 4). The PR is stripped and the exposed chrome and copper are etched. The wafer is then spin-coated with a second PR layer which is patterned to define the bearing raceway. The exposed oxide is etched using RIE and the raceway is etched to a depth of 260 μm using DRIE (B). The PR is stripped and a third PR layer is deposited on the back side of the wafer. The thrust journal is patterned and the exposed oxide is etched using RIE. The thrust journal is then partially etched, to a depth of 150 μm, using DRIE (C). The PR is stripped and the wafer is snapped into die.

3.2.2. Turbine wafer fabrication. The process flow used to fabricate the turbine wafer is shown in figure 5. The turbine wafer is fabricated using a double side polished, 1 mm thick, 100 mm diameter silicon wafer with 1 μm of thermal oxide. The solder pads are patterned on the wafer in the same manner as for the thrust wafer. A PR layer is then deposited and patterned with the turbine journal and the exposed oxide is etched (*A*). The PR is stripped and new PR is deposited and patterned to define the recess at the top of the retainer raceway. The exposed oxide is etched to a depth of 200 nm using RIE (*B*); this produces a stepped oxide mask which can define both the turbine journal and the recess. The PR is then stripped and new PR is deposited and patterned with the retainer raceway. The turbine journal is then partially etched to a depth of approximately 200 μm, at which point the thinned oxide mask begins to degrade due to the finite selectivity against thermal oxide during the DRIE process (*C*). The thinned regions of exposed oxide are removed using RIE and the exposed silicon is etched 5 μm to form the recess (*D*). RIE is used to etch the remaining exposed oxide and the retainer raceway is etched to a depth of 240 μm (*E*), at the same time transferring the recess and turbine journal further into the wafer. The PR is then stripped and a new PR layer is deposited on the back side of the wafer. The resist is patterned with the turbine journal and the exposed oxide is etched using RIE (*F*). The PR is stripped and a new layer is deposited and patterned with the turbine blades. The turbine journal is etched to a depth of 250 μm at which point the oxide mask begins to fail (*G*). The exposed oxide is then etched using RIE and the turbine blades are etched to a depth of 250 μm (*H*). Finally the wafer is snapped into die.

The turbine journal has to be etched from both sides of the wafer due to the limitations of the oxide masks used. The final DRIE steps used to release the rotor rely purely on the remaining oxide to protect the top and bottom surfaces of the device, and this sets a limit on the thickness of un-etched silicon in the turbine journal. Because the journal is a small feature embedded in larger features, it experiences significant etch lag during DRIE processing. If the release etch is too long, then other features, in particular the turbine rotor blades and guide vanes, will be over-etched. The above process flow results in a silicon thickness between the base of the blades/vanes and the retainer raceway of over 200 μm, accounting for etch lag, providing enough mechanical strength for the 10 mm devices to operate properly. Short isotropic etches (45–60 s) are also incorporated between the stepped DRIE etches to remove passivation from previous etch steps and reduce the formation of spires. The multi-step etching process described here is similar to the process described in [20].

3.2.3. Assembly and release. The process flow for assembly and release is shown in figure 6. To assemble the device, a retainer ring is placed over the raceway on the thrust die. The steel balls (Faulhaber, MPS, 0.500 mm dia., 440C stainless steel, Grade 3) are manually placed into each ball pocket in a star pattern, in order to force the ring into proper alignment. A small amount of solder flux is then painted onto each solder pad and 300 μm diameter solder balls are placed centrally on the solder pads. The flux not only holds the solder balls in place, but also aids in the reflow of the solder during bonding. The combination of flux, solder balls, and electroplated eutectic metals has increased the bonding yield to 90%. The turbine

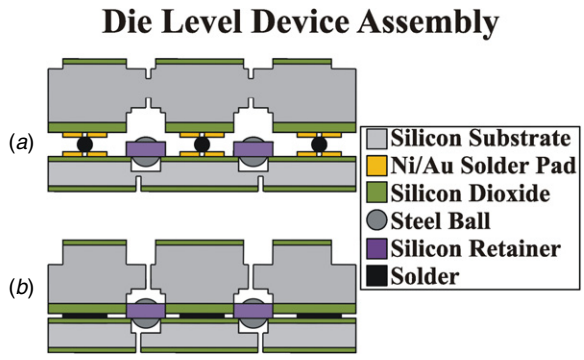


Figure 6. Die level assembly process flow. (*a*) Retainer placed over bearing raceway; steel balls inserted into ball pockets; solder balls placed on centres of solder pads; turbine die aligned and placed on top. (*b*) Assembly heated with gentle force applied to reflow solder; turbine and thrust journals etched to release turbine.

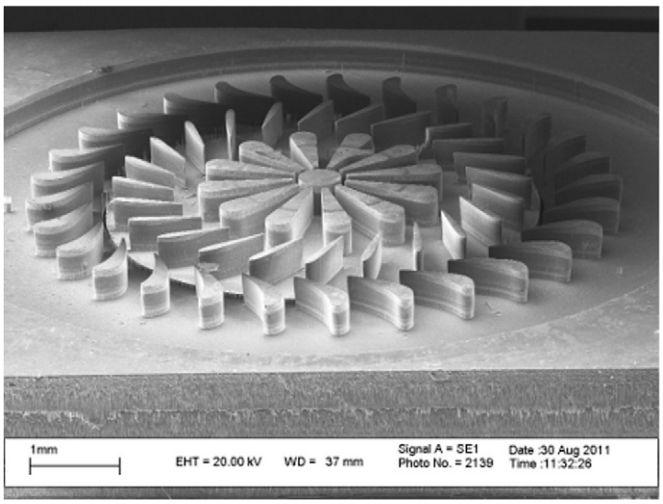


Figure 7. SEM image of a 5 mm SMT.

die is then placed on top of the thrust die. Pressure is applied in a custom bonding rig and solder reflow is performed by heating with a butane flame for 30 s. The turbine side is etched using DRIE until the turbine journal breaks through to the retainer raceway. The device is then flipped over and the thrust journal is etched using DRIE until it meets the bearing raceway. Both steps use the oxide remaining from previous steps as a mask. The final turbine blade height is approximately 450 μm. Figure 7 shows an SEM image of a completed 5 mm device.

4. Testing

Testing was carried out by driving the SMT with a compressed nitrogen supply while at the same time bleeding gas from the back side of the device to provide some control over the pressure on the underside of the rotor, and hence the normal load on the bearing. In initial performance tests the input drive power (inlet gauge pressure $\times$ volume flow rate) was varied, and rotation speed, gas flow and input pressure were recorded. This allowed a performance curve of rotation speed versus input power to be plotted. These tests could not provide an absolute measure of the bearing losses, since the turbine

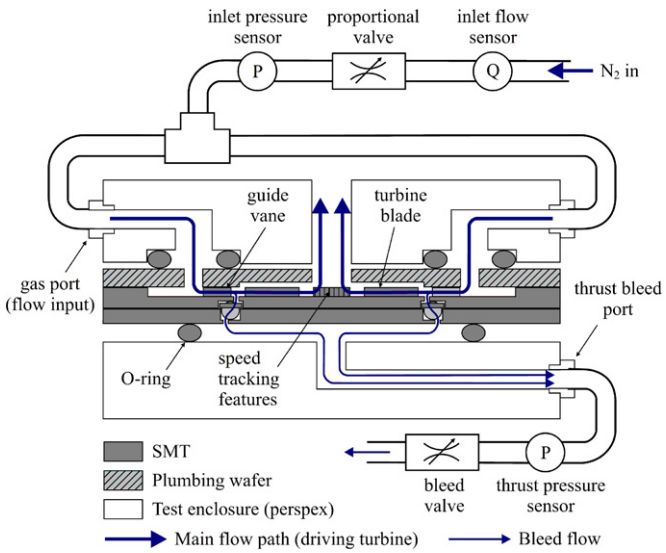


Figure 8. Schematic of experimental set-up used for testing.

efficiency was unknown; however, they did allow comparison between the different bearings. Performance tests were carried out on two devices for each combination of retainer type and device size. In addition, the repeatability of results was verified for one device, and longevity testing was performed on a 5 mm device and a 10 mm device to evaluate wear during extended operation.

The testing methods used were similar to those described in [16]. Figure 8 shows a schematic of the test set-up. The device is placed into a custom enclosure, with gas lines attached to the top to actuate the turbine and a separate gas line attached to the bottom to allow bleeding of gas from the thrust side. The turbine is actuated with pressurized nitrogen and the inlet flow is controlled using a proportional valve. A second proportional valve (the 'bleed valve') allows control over the bleed flow. An optical displacement sensor is placed over speed tracking marks located at the centre of the device to determine the speed of rotation, and pressure sensors are attached to the gas inlet line and the bleed line. Rotation speed, gas flow rate, input pressure and thrust pressure are logged every 500 ms. Figure 9 shows photographs of the two test enclosures, along with a close-up view showing the larger enclosure in position with the optical displacement sensor overhead.

To test each device, the bleed valve was first set to a known value. This value was determined by running a Half Skeleton device at 2 W of input power and adjusting the bleed valve until the device operated smoothly. This calibration was performed separately for the 5 mm and 10 mm devices and the initial valve setting was maintained for all subsequent testing at that device size. For the 5 mm devices, the calibration resulted in a thrust pressure which increased from 0 to 20 mbar as the input power was varied over the range 0–2 W. The corresponding thrust pressure range for the 10 mm devices was 0–4.2 mbar. Thrust pressure measurements made during subsequent testing showed that the variations of thrust pressure with input power were similar for all SMTs of a given size; the thrust pressure readings showed some scatter but were within a range of $\pm$ 2 mbar for the 10 mm devices and $\pm$ 0.8 mbar for the 5 mm

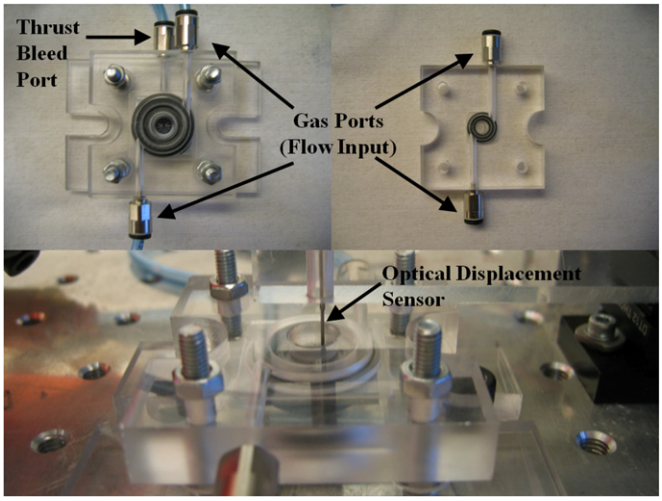


Figure 9. Test enclosure for the 10 mm device (top left), test enclosure for the 5 mm device (top right) and the 10 mm device enclosure with the optical displacement sensor aligned over the device speed tacking marks (bottom).

devices. This was a good indication that all the bearings of a given size were subject to similar loading. Each device was ramped through a range of inlet pressures at least three times to evaluate the performance. Longevity and wear were evaluated by operating a 5 mm and a 10 mm device continuously for an extended period of time (up to 12 h) while counting the number of revolutions. Open retainer designs were used in these tests.

5. Results

5.1. Comparison of retainer designs

Each type of retainer was tested by increasing the flow of nitrogen in steps until an input power of 2.5 W was achieved. The flow of nitrogen was then cut and the ramping was repeated two further times. In order to gain an adequate comparison for each retainer ring, two devices of each type were tested from the same fabrication run.

Figures 10 and 11 show the measured variations of speed with input power for the 5 mm and 10 mm devices, respectively. Six measurements were taken at each input power setting, and all data points are plotted to give an idea of the scatter in the measurements. The results for both the 5 mm and 10 mm devices indicate that the Full Ring and the Full Skeleton designs exhibit the best performance (i.e. highest rotation speed for a given input power) over a wider range of input power, with the Full Ring performing the best overall. The Half Skeleton, Outer Open and Inner Open designs perform consistently less well over the same range.

The results suggest that the contact area between the retainer and the inner and/or outer lands is not an overriding factor in determining the bearing losses, since the retainers with the largest contact areas (Full Ring and Full Skeleton) consistently showed the lowest losses. The Full Skeleton retainer performed less well than the Full Ring in both sizes of device, and the reasons for this are currently unclear. One possibility is that the lower rigidity of the Full Skeleton retainer

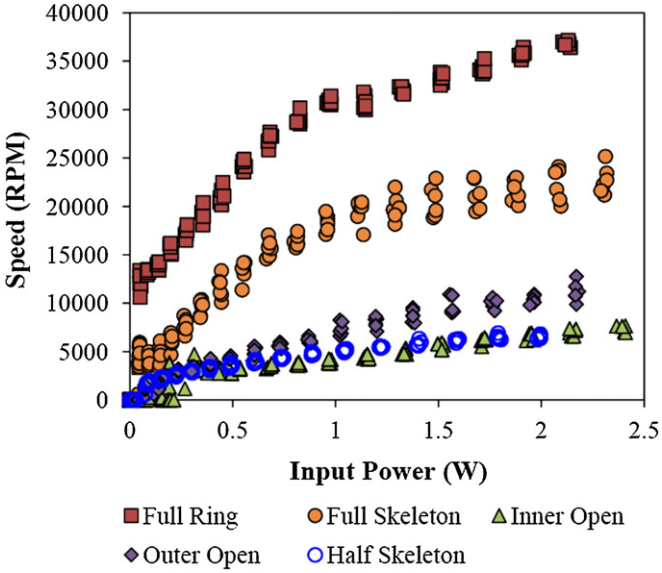


Figure 10. Performance curves for 5 mm devices plotted as speed in RPM versus input power in watts (W).

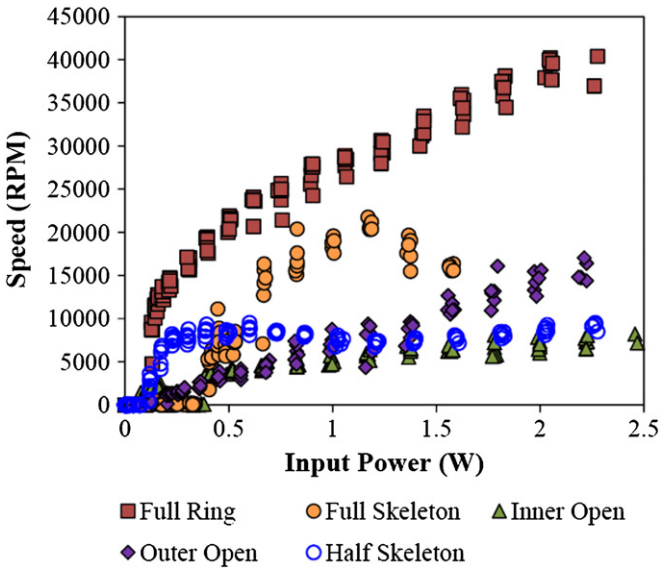


Figure 11. Performance curves for 10 mm devices plotted as speed in RPM versus input power in watts (W).

makes it more susceptible to dynamic distortions that result in more contact with the raceways. This could also explain the drop in performance of the 10 mm Full Skeleton device at higher power levels. This effect would be expected to be more pronounced in the larger device where the retainer is less stiff. The relatively poor performance of the more flexible design does run counter to results from previous studies [13]. However, the flat raceways on the silicon bearings make them less constrained and more susceptible to vibration, and it is not clear that results relevant to conventional bearings should carry over.

Low rigidity may also be a contributing factor to the poor performance of the Inner Open and Outer Open retainers. These designs also allow the balls more radial play which is likely to increase interaction between the balls and the raceway

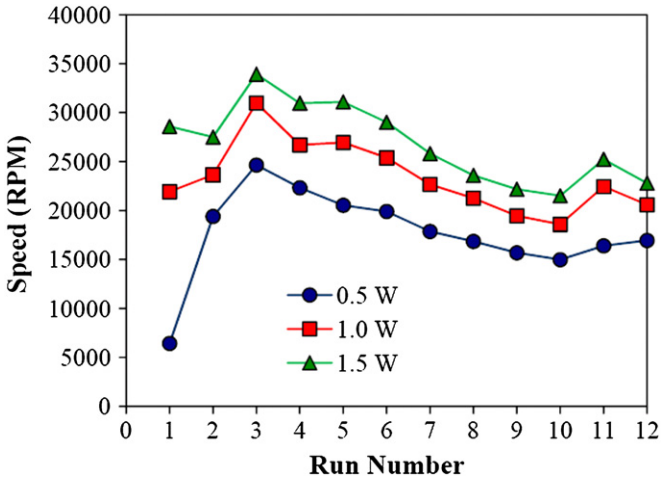


Figure 12. Measured repeatability of performance over 12 ramping tests for a 5 mm Full Ring Device operated at different power levels.

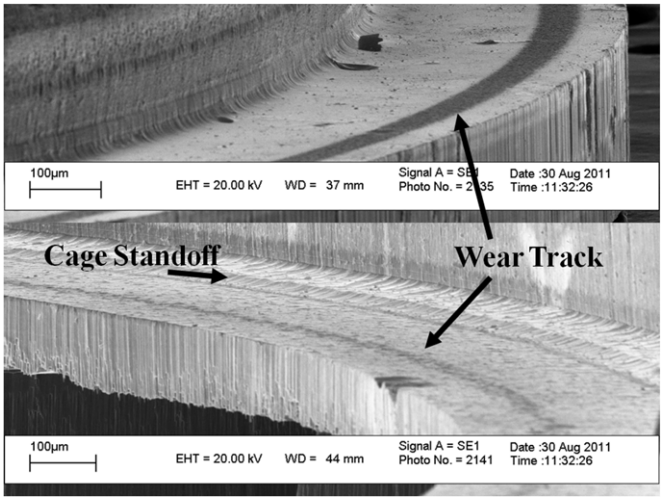


Figure 13. SEM images of the raceway wear on the rotor (top image) and the stator (bottom image) of the 10 mm diameter device after longevity testing.

sidewalls. Furthermore, interactions between the balls and the edges of the pocket openings may lead to increased frictional losses with the Open designs. Finally, the poor performance of the Half Skeleton designs is consistent with their having a lower ball fill factor in the raceways [19].

The input power levels required to drive the 10 mm diameter SMTs in this work were higher than reported previously for the best retainer-less designs. For example, in [21] performance curves are presented for a 10 mm diameter SMT which show operation at 10 000 rpm with an input power of only 38 mW. For comparison, the best-performing (i.e. Full Ring) device tested in the present work required an input power of 125 mW at the same rotation speed. It may be that additional friction due to the presence of the silicon retainer is contributing significantly to this discrepancy; however, this cannot be stated categorically since there are other differences in approach between this and the earlier work which may be important; for example, the earlier work used a different size of micro-ball.

It is noted that because the thrust bleed valve was maintained in a fixed position during ramp testing, the turbines may not have been operating in an ideal regime. It was verified during tests that the bearings were operating in the correct mode, with a net upward thrust on the rotor. However, it is likely that better performance could be obtained by constantly re-adjusting the bleed valve during testing, and this will be explored in future work.

5.2. Repeatability

A 5 mm diameter Full Ring device was run through several successive ramp tests, averaging 30 000 revolutions per ramp, to evaluate the repeatability of the results. For each ramp test, the measured variations of rotation speed with inlet power were averaged, and then simple linear interpolation was used to estimate the rotation speed at fixed power levels of 0.5 W, 1.0 W and 1.5 W. Figure 12 shows how the rotation speed at each of these power levels varied over the first 12 runs. At all power levels, an initial improvement in performance was observed, followed by a fall-off up to round the eighth run, beyond which the rotation speed was stable to within $\pm 10\%$. It is believed that the initial improvement is due to removal of asperities on the bearing surfaces, while the subsequent fall occurs as the sides of the bearing raceway become worn making the bearing less stable due to increased play in the rotor position.

5.3. Longevity and wear

Longevity was evaluated by running a 10 mm diameter device with an Outer Open retainer and a 5 mm diameter device with an Inner Open retainer for extended periods of time. The retainers with the lowest mechanical strength were chosen for this test as one of the aims was to induce early failure and identify the failure mode. However, both devices survived the longevity testing without failure and only required minimal cleaning to resume operation. Just over 2.5 million revolutions were logged for the 10 mm diameter device and just over 2.98 million revolutions were logged for the 5 mm diameter device. The devices were disassembled for inspection by heating to break the solder bonds.

Wear effects on the bearing raceways after longevity testing were similar to those described in [16] and [21]. There was some ball-induced wear on the sidewalls of the bearing raceway, and narrow wear tracks could be seen in the silicon raceway surfaces top and bottom, as shown in figure 13. The retainers in both devices experienced minimal wear on their upper surfaces during operation. For example, figure 14 is an SEM image showing the upper surface of the 5 mm Inner Open retainer. An optical profilometer scan indicated that the contact wear did not exceed 1 μm on any portion of the ring. This was also confirmed using a microscope, by observing that oxide was still present in most of the wear regions. The edge of one of the ball pockets shows some signs of damage likely attributable to shock during start up of the device.

The undersides of the retainers showed wear around the edges of the ball pockets due to abrasion by the balls, as shown in figure 15. This has the effect of increasing the effective

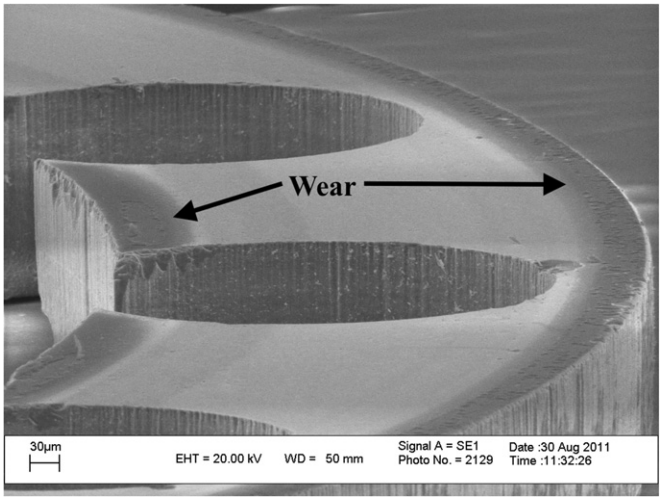


Figure 14. Wear damage on the upper surface of the 5 mm retainer after longevity testing. The wear appears as dark regions along the edge of the cage indicated by the arrows. Some chipping can be seen at the edge of one of the pockets.

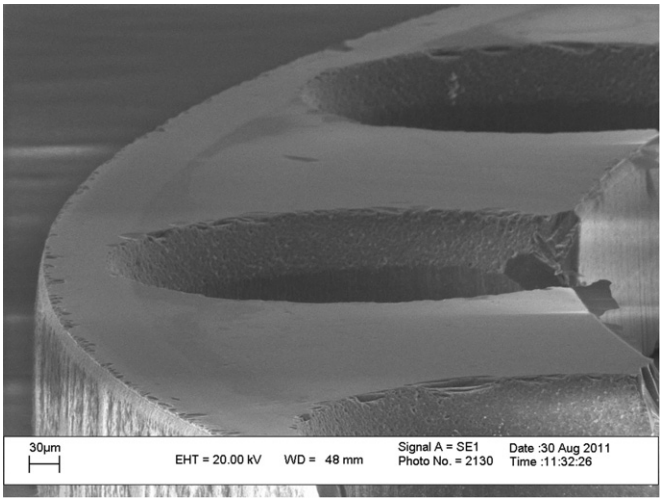


Figure 15. Ball-induced wear on the lower surface of the 5 mm retainer after longevity testing.

pocket size, and relaxing the constraints imposed on the balls by the retainer.

6. Conclusions

This work presents the first exploration of several different geometries of silicon retainer rings integrated into a SMT with a thrust ball bearing, at rotor diameters of both 5 mm and 10 mm. The aim was to gain an initial idea of what type or types of retainer might be expected to give the best performance in the longer term. Full Ring retainers, which were annular with cylindrical ball pockets, were found to perform better than designs with pockets that were open on one side, even though the latter had lower contact area with the bearing rings. Removal of material from the Full Ring design while retaining the cylindrical pocket shape was also found to degrade performance. The SMTs with Full Ring retainers could operate at over 20 000 RPM with less than 2 W of input power. Also 5 mm- and 10 mm diameter devices were able to operate for over 2.5 million revolutions without device failure. These bearings had a ball fill factor of around 50% which is significantly lower than required in retainer-less designs. Reduction of the ball complement could be beneficial in terms of cost reduction in smaller bearings.

Comparison of the results with published data shows that the input power levels required to drive the SMTs in this work were higher than for the best retainer-less designs. Further work is required to establish whether the presence of the retainer is a major contributing factor to this discrepancy. Most importantly, a direct comparison should be made between devices that are identical apart from the inclusion or omission of a retainer. Future work will also include measuring the bearing vibration and finding the optimal running conditions for the turbine by varying the thrust bleed rate. Alternative, softer retainer materials such as polymers and metals will also be explored, as will the possibility of implementing a truly ball-riding retainer.

Acknowledgments

The authors would like to thank Dr Munir Ahmad and Dr Werner Karl for their advice and expertise in the fabrication process. They would further like to thank Phil Jones for his work on the test enclosure and assistance in the design and fabrication of the mechanical components needed for testing.

References

[1] Mehregany M, Senturia S D and Lang J H 1990 Friction and wear in microfabricated harmonic side-drive motors *Tech. Digest Solid State Sensors and Actuators Workshop, Hilton Head, SC* pp 17–22

[2] Maboudian R, Ashurst W R and Carraro C 2000 Self-assembled monolayers as anti-stiction coatings for MEMS: characteristics and recent developments *Sensors Actuators* A **82** 219–23

[3] Smallwood S A, Eapen K C, Patton S T and Zabinski J S 2006 Performance results for MEMS coated with a conformal DLC *Wear* **260** 1179–89

[4] Asay D B, Dugger M T and Kim S H 2008 *In-situ* vapor-phase lubrication of MEMS *Tribol. Lett.* **29** 67–74

[5] Ku I S Y, Reddyhoff T, Wayte R, Choo J H, Holmes A S and Spikes H A 2012 Lubrication of microelectromechanical devices using liquids of different viscosities *ASME J. Tribol.* **134** 012002

[6] Houlihan R and Kraft M 2002 Modelling of an accelerometer based on a levitated proof mass *J. Micromech. Microeng.* **12** 495–503

[7] Komori K and Yamane T 2001 Magnetically levitated micro PM motors by two types of active magnetic bearings *IEEE/ASME Trans. Mechatronics* **6** 43–9

[8] Livermore C, Forte A, Lyszczara T, Umans S D and Lang J H 2004 A high-power MEMS electric induction motor *J. Microelectromech. Syst.* **13** 465–71

[9] Sangster A J, Westby E and McEarlean E 2003 Electromagnetic suspension system for microactuators *J. Electromagn. Waves Appl.* **17** 1131–47

[10] Waits C M, McCarthy M and Ghodssi R 2010 A microfarbicated spiral-groove turbopump supported on microball bearings *J. Microelectromech. Syst.* **19** 99–109

[11] Harris T A and Kotzalas M N 2006 *Essential Concepts of Bearing Technology* 5th edn (Boca Raton, FL, USA: Taylor and Francis Group)
[12] Katagiri C and Naito K 2004 Next generation deep groove ball bearing for high speed servomotor *NTN Technical Review, No. 72* pp 46–51
[13] Weinzapfel N and Sadeghi F 2009 A discrete element approach for modeling cage flexibility in ball bearing dynamic simulations *ASME J. Tribol.* **131** 021102
[14] Holm-Hansen B T and Gao R X 2000 Structural design and analysis for a sensor-integrated ball bearing *Finite Elem. Anal. Des.* **34** 257–70
[15] Hergert R J, Ku I S Y, Reddyhoff T and Holmes A S 2010 Micro rotary ball bearing with integrated ball cage: fabrication and characterization *Tech. Digest MEMS 2010, Hong Kong, China* pp 687–90
[16] McCarthy M, Waits C M and Ghodssi R 2009 Dynamic friction and wear in a planar-contact encapsulated microball bearing using an integrated microturbine *J. Microelectromech. Syst.* **18** 263–73
[17] Hergert R J, Hanrahan B, Holmes A S and Ghodssi R 2011 Silicon retainer ring integration in micro-turbine with thrust ball bearing support mechanism *Tech. Digest Tranducers 2011, Bejing, China* pp 1344–7
[18] Harris T A and Mindel M H 1973 Rolling element bearing dynamics *Wear* **23** 311–37
[19] Nataraj C and Harsha S P 2008 The effects of bearing cage run-out on the nonlinear dynamics of a rotating shaft *Commun. Nonlinear Sci. Numer. Simul.* **13** 833–8
[20] Kim B-H, Park B-J and Kim J B 2009 Process effects of double step DRIE and Ni-Co electroplating for a trench-type cantilever probe for a fine-pitched MEMS probe card *Sensors Actuators* A **152** 252–60
[21] Hanrahan B, Beyaz M, McCarthy M, Waits C M and Ghodssi R 2010 A new performance regime for microfabricated ball bearings *Tech. Digest PowerMEMS 2010, Leuven, Belgium* pp 191–4