Numerical modeling and analysis of plasmonic flying head for rotary near-field lithography technology.
Contents

Review Article

Advances in nonequilibrium molecular dynamics simulations of lubricants and additives / 349–386
J. P. EWEN, D. M. HEYES, D. DINI

Research Articles

Effect of arc-current and spray distance on elastic modulus and fracture toughness of plasma-sprayed chromium oxide coatings / 387–394
Simanchal KAR, Partha Pratim BANDYOPADHYAY, Soumitra PAUL

Friction characteristics of mechanically exfoliated and CVDgrown single-layer MoS₂ / 395–406
Dinh Le Cao KY, Bien-Cuong TRAN KHAC, Chinh Tam LE, Yong Soo KIM, Koo-Hyun CHUNG

Contribution of friction and adhesion to the reliable attachment of a gecko to smooth inclines / 407–419
Zhouyi WANG, Qiang XING, Wenbo WANG, Aihong JI, Zhendong DAI

Analysis of the coated and textured ring/liner conjunction based on a thermal mixed lubrication model / 420–431
Chunxing GU, Xianghui MENG, Di ZHANG

Study on frictional behavior of carbon nanotube thin films with respect to surface condition / 432–442
Youn-Hoo HWANG, Byung-Soo MYUNG, Hyun-Joon KIM

Numerical modeling and analysis of plasmonic flying head for rotary near-field lithography technology / 443–456
Yueqiang HU, Yonggang MENG

Friction and wear of sand-contaminated lubricated sliding / 457–463
Mohamed Ahmed RAMADAN

Evaluation of lubrication performance of crank pin bearing in a marine diesel engine / 464–471
Suk Man MOON, Yong Joo CHO, Tae Wan KIM

Short Communication

Insight into tribological problems of green ship and corresponding research progresses / 472–483
Yuwei SUN, Xinping YAN, Chengqing YUAN, Xiuqin BAI
Advances in nonequilibrium molecular dynamics simulations of lubricants and additives

J. P. EWEN, D. M. HEYES, D. DINI*

Department of Mechanical Engineering, Imperial College London, Exhibition Road, South Kensington, London SW7 2AZ, United Kingdom

Received: 12 September 2017 / Revised: 13 December 2017 / Accepted: 26 December 2017

© The author(s) 2018. This article is published with open access at Springerlink.com

Abstract: Nonequilibrium molecular dynamics (NEMD) simulations have provided unique insights into the nanoscale behaviour of lubricants under shear. This review discusses the early history of NEMD and its progression from a tool to corroborate theories of the liquid state, to an instrument that can directly evaluate important fluid properties, towards a potential design tool in tribology. The key methodological advances which have allowed this evolution are also highlighted. This is followed by a summary of bulk and confined NEMD simulations of liquid lubricants and lubricant additives, as they have progressed from simple atomic fluids to ever more complex, realistic molecules. The future outlook of NEMD in tribology, including the inclusion of chemical reactivity for additives, and coupling to continuum methods for large systems, is also briefly discussed.

Keywords: molecular dynamics; nonequilibrium systems; confined fluids; boundary lubrication; elastohydrodynamic lubrication; tribology

1 Introduction

In order to reduce energy consumption in engineering systems and thus CO₂ emissions, better lubricants need to be developed [1]. As in many areas of science and engineering, computer simulation has become a powerful tool which could help to accelerate future lubricant design. In tribology, processes at the smallest scales often drive the macroscopic friction [2] and wear [3] behaviour that we observe. As a result, there is growing interest in the application of molecular simulation to tribological problems. This review provides a historical and critical overview of the most important developments in this field of study, highlights some of the outstanding issues currently faced by researchers, and suggests future areas of development.

Molecular simulation, a field of research that started in the 1950s, uses computers to generate a series of representative snapshots of a group of interacting molecules from which experimental observables can be calculated. A system of interest is simulated by a collection of individual molecules, represented at the atomic level, and the locations of the individual atoms are governed by the interaction force laws between them. There are two main methods for performing molecular simulation, one stochastic and one deterministic [4]. The first to be developed was (Metropolis [5]) Monte Carlo (MC), which evolves the molecules by moving (usually) one molecule at a time by a trial random displacement. This new position is accepted or rejected on the basis of the change of potential energy of the system (more specifically its exponential or the “Boltzmann factor”). The second, and nowadays much more widely used method is molecular dynamics (MD), in which the atomic trajectories are generated by solving Newton’s equations of motion using a finite difference scheme over a series of short time steps. Both MC and MD can be used to compute thermodynamic and other static properties such as infinite
or zero frequency elastic moduli. Time-dependent processes and transport coefficients such as the viscosity arise more naturally from MD simulations, making it the more appropriate method to investigate tribological phenomena.

In tribology and specifically for liquid lubricants, there is considerable interest in predicting bulk transport properties of fluids (particularly the viscosity) under a wide range of conditions (temperature, pressure, shear rate) relevant to real components. As early as 1969, Goldstein [6] envisaged the use of MD to study the viscosity of liquids in the following prescient remarks:

“The modern molecular approach to viscosity proceeds via the formulation of the shear viscosity in terms of appropriate integrals over the stress-time correlation function in a fluid. There is no difficulty, in principle, in imagining computer experiments generating this correlation function in a finite sample of molecules set into motion according to Newton’s laws. It is not yet clear how large a sample of molecules is needed, nor how long a time of study is required, to achieve the equivalent of an infinite number of molecules studied for an infinite time, but one might guess that such a calculation will be within our reach in the immediate future, for liquids having the simplest, spherically symmetric, pair potentials.”

Goldstein was perceptive and the first MD study of the viscosity of a real fluid (argon) followed only four years later [7]. Very significant progress has been made since then and the number of studies applying MD to tribological systems has increased dramatically in recent years, for a variety of reasons. Firstly, reliable simulation methodologies have now been developed which allow researchers to study the shear of bulk fluid systems, as well as those confined by solid surfaces. Secondly, most lubricating oils are composed of relatively large organic molecules (C\textsubscript{20}−C\textsubscript{40}) and force-fields that govern the interactions between such molecules in MD simulations have now been refined to a stage where they can reproduce realistic viscous behaviour [8–10]. Finally, recent increases in computational performance coupled with the development of highly parallelised simulation algorithms to exploit multiprocessor, high performance computing (HPC) architectures, have enabled larger molecules, system sizes, and timescales to be modelled.

Nonequilibrium molecular dynamics (NEMD), where a shear force is applied to the system, has significantly improved our understanding of both dry friction, where the sliding surfaces are in intimate contact, and wet friction, where the surfaces are separated by a liquid lubricant. The focus of this article is NEMD simulations of wet friction, with a particular emphasis on (non-aqueous) lubricants and additives. In 2004, Urbakh et al. summarised experimental and theoretical contributions to the fundamental understanding of both wet and dry friction [2]. Szlufarska et al. reviewed experimental and theoretical studies of wet and dry friction specifically in the area of nanotribology in 2008 [11]. In 2013, Vanossi et al. [12] summarised advances in nanoscale to mesoscale modelling of wet and dry friction while and in the same year Dong et al. [13] discussed NEMD studies of dry friction in the context of atomic force microscopy (AFM) experiments. In 2014, Sawyer et al. [14] provided a more general overview of recent advances in the understanding of dry friction, including the contribution of NEMD. The reader should refer to these contributions for specific details on these topics.

How did we get to this point? We will now briefly summarise the stages that the subject has gone through to arrive at the present state where molecular simulation has become an important and powerful tool in the field of tribology (Section 1.1). We then provide some general guidelines for NEMD simulations of lubricants and additives (Section 1.2). We then describe a number of key step-changing methodological advances which are often taken for granted, but have revolutionised the subject (Section 1.3).

1.1 History of NEMD simulations

MD was invented in the 1950s, using the hard sphere (HS) model in which the “molecules” were represented as perfectly elastic and infinitely repulsive spheres (“billiard balls”). The HS MD approach is very computationally efficient (particularly at low density), and would today be referred to as an “event driven” method, as it follows the collisions between the spheres generated by classical mechanics in chronological order. Probably the most significant paper of this pioneering work was by Alder and Wainwright [15],
who in 1957 showed that an HS liquid crystallised as the density was increased above a certain value. This discovery had wide-ranging consequences, as it was widely thought at the time that it was necessary for molecules to have an attractive component to their interaction potential to show a freezing transition and Alder and Wainwright’s work showed that this was not the case. Today this would be called an “entropy-driven” freezing transition, of which there are abundant examples in colloid and polymeric liquid systems, and its existence plays a major role in our understanding of the phase behaviour and physical properties of a wide range of soft condensed matter chemical systems [16].

The 1960s was the period in which model molecules represented by a continuous interaction potential were first simulated by MD. The first paper was by Rahman in 1964 [17], who calculated the diffusion coefficient in a system of argon atoms interacting through a Lennard–Jones (L–J) potential. This was also a groundbreaking paper in that it showed the diffusion and structural evolution of small molecules takes place by a series of small, highly coordinated motions of neighbouring molecules. Until this MD paper, it had been generally held that diffusion in liquids took place through a series of rare, large distance jumps of molecules from “lattice sites” into vacancies or “holes” in a crystal-like lattice, the jump distance being roughly a molecular diameter. Historically, liquids were viewed as a crystal with many unoccupied sites [18], but Rahman’s paper convincingly disproved this supposition. However, erratic, localised displacements have more recently been observed in the MD simulations of supercooled liquids near the glass transition, a phenomenon known as “heterogeneous dynamics” [19], so the defect-containing crystal view of a liquid is not entirely without merit. The diffusion coefficient is closely related to the viscosity, indeed the Stokes–Einstein equation [20] suggests that the product of the two is almost constant in liquids [21]. In 1970, Alder et al. [22] calculated the viscosity of the HS fluid at various densities up to crystallisation. They used a mean square displacement of the shear stress route, which is now called the Einstein-Helfand method [23].

Verlet and co-workers in 1973 [7] were the first to calculate the shear viscosity of a real liquid (argon) represented by the Lennard–Jones (L–J) potential. They used the Green-Kubo time correlation method [24, 25], where the Newtonian viscosity is defined as the integral of the shear stress relaxation function, $\eta(t)$. In the small shear rate limit, $\eta(t)$ can be shown by statistical mechanics to be the product of the shear stress at one time multiplied by its value at a later time. This function decays monotonically to zero, as liquids subjected to a step in strain can only sustain the initial jump in shear stress for a finite time. Therefore, perhaps paradoxically at first sight, the Newtonian viscosity can be calculated without actually imposing a shear rate across the system. Near the melting line, the viscosity of the HS fluid increases rapidly because of a slowing down in the decay of the shear stress time correlation function (caused by crowding effects of the molecules packed closely together), an effect that was called the “molasses tail” [26].

Periodic boundary conditions (PBCs) are a very important tool which allow MD simulations to study essentially bulk systems in a computationally efficient manner. PBCs create an infinite lattice of periodically replicated molecular simulation cells which fills all space (Fig. 1) [4]. If a molecule leaves the cell through one boundary, it is reintroduced with the same velocity at the same point of departure through the opposite face of the cell. This construction effectively eliminates surface effects that would otherwise dominate the behaviour of the system, due to the large surface-to-volume ratio of a typically-sized simulation cell [4].

Until the 1970s, the liquid simulations were carried out mainly using equilibrium MD which by default follows the microcanonical ensemble (an ensemble is simply a probability distribution for the state of the system). The microcanonical ensemble is also sometimes called the NVE ensemble because the total number of particles in the system, $N$, the volume of the system, $V$, as well as the total energy in the system, $E$, remain constant. Then, in the early 1970s, the first nonequilibrium molecular dynamics (NEMD) simulations of imposed shear flows at large shear rates were carried out. A variety of techniques were developed, and these continue to be invented and refined today (Section 1.3). The version perhaps most relevant to tribology is the “sliding wall” method invented by Ashurst and Hoover in 1975 [27]. This mimics the shearing of a fluid of infinite extent in the $x$- and $y$-directions using PBCs, confined by bounding
walls whose plane normal is in the $z$-direction (see Fig. 1(a)). The two walls (which were initially simply the outermost regions of the fluid) are constrained to translate in opposite directions in the $x$-direction to shear the fluid. This was followed by simulations by Bitsanis et al. in 1987 in which solid walls were used to confine and shear the fluid [28]. The study of confined fluids under shear by NEMD has undergone a rapid expansion in recent years, and is now a common geometry for simulations in many different fields of research.

Up to the end of the 1970s, NEMD simulations were quasi-2D in that they used PBCs in only two directions (Fig. 1(a)), but in the 1980s equations of motion for imposing shear flow in systems with PBCs in all three Cartesian directions were invented (Fig. 1(b)). These are known as the DOLLS and SLLOD methods [30]. By good fortune, PBCs which can accommodate shear flow of infinite extent within MD had been invented by Lees and Edwards (LE) several years previously [31], and these new non-Hamiltonian equations of motion are compatible with these “sliding brick” PBCs (see Fig. 1(b)). Recent books by Evans and Morriss [32], and Todd and Daivis [33], provide a comprehensive description of the field from that time to the present.

An early NEMD simulation using the LE PBCs, performed by Heyes et al. in 1980 [34], modelled the start-up of shear flow in a supercooled Lennard–Jones (L–J) fluid. In these simulations, “stress overshoot” behaviour was observed shortly after the initiation of shear, a phenomenon widely seen in experimental studies of polymeric and glassy systems. The importance of the 1980s cannot be overestimated as it was then that the statistical mechanics or governing physics of liquid systems arbitrarily far from equilibrium started to be derived, largely by Evans and Morriss [32]. Since the 1980s, NEMD has progressed to a point where it has become a predictive tool in tribology, as described in Section 2.

1.2 Computational guidelines

In order for NEMD simulations to be useful in terms of predictive capabilities, a number of aspects must be considered, from the inherent limitation of the time and length scales that can be modelled, to the accuracy of the representation of the intramolecular and intermolecular interactions [35]. The computational time required to conduct MD simulations depends on a number of factors including the computer hardware and software available, the system size, as well as the
complexity of the force-fields used. The number of atoms required in NEMD simulations of tribological systems varies depending on the particular problem being studied, but is generally in the thousands to hundreds of thousands range. Classical MD force-fields require a time step of approximately 1 fs to ensure energy conservation [4]. This means that, even when performed on multiprocessor high performance computing (HPC) systems, using highly parallelised MD software, only ns or in extreme cases μs timescales are accessible in MD simulations [36], since these correspond to millions/billions of time steps.

In NEMD simulations, where shear is applied, the main consequence of the short accessible timescales is that relatively high shear rates are required (generally > 10^7 s^-1) to ensure that the properties of interest (most commonly the viscosity) reach a steady state [37, 38]. Moreover, at very low shear rates, the systematic nonequilibrium response becomes comparable to the equilibrium fluctuations in the phase variables of interest, resulting in a low signal-to-noise ratio [39, 40]. The ability to simulate lower shear rates has been a long-term goal of NEMD simulations to facilitate direct overlap with experiments and real components [32]. For comparison, the high pressure viscosity of lubricants can generally only be measured up to shear rates of approximately 10^4 s^-1 [37, 38], tribology experiments can reach up to around 10^6 s^-1 [41], and high-performance engine components can extend up to 10^8 s^-1 [42].

### 1.3 Methodological developments

Following the initial development of bulk and confined NEMD simulation methodologies (Section 1.1), many aspects have been altered and refined to allow different conditions to be studied and more accurate results obtained. Significant examples include improved temperature and pressure control methods, classical force-fields, as well as new simulation methods; these are briefly discussed below.

#### 1.3.1 Temperature and pressure control

The high shear rates that are usually imposed in NEMD simulations create heat, which must be dissipated to achieve a nonequilibrium steady state [32]. In parallel with the theoretical advances described in Section 1.1, a number of different thermostats have been invented which can be used in conjunction with the NEMD equations of motion. In bulk MD simulations, the thermostat is usually applied globally (to all of the atoms in the system). The temperature in MD simulations can be controlled through a number of methods including scaling the atomic velocities (e.g., Andersen [43]), coupling atoms to an external bath (e.g., Berendsen et al. [44]), or applying a random “friction” force to the atoms (e.g., Langevin [45] and dissipative particle dynamics (DPD) [46]).

The Nosé-Hoover (NH) thermostat [47, 48], in which a Hamiltonian with an extra degree of freedom for a heat reservoir is introduced to control the temperature, is the most widely used thermostat for bulk NEMD simulations. Under equilibrium conditions, the NH thermostat can be proven to generate a canonical or NVT ensemble.

However, some researchers have raised concerns with thermostats, such as NH, which require a predefined flow profile (e.g., linear/Couette) to be assumed [50]. So-called profile-biased thermostats can result in artefacts in some systems, for example the “string phase” which was often observed in NEMD simulations of L-J fluids at high shear rates [51]. Moreover, they are of course incorrect for systems where the flow profile is not Couette-like. To overcome these problems, configurational thermostats, which make no assumptions regarding the flow profile, were developed [52, 53]. These have not yet been applied in bulk NEMD simulations of lubricant-sized molecules (C20-40), probably because they are more difficult to implement for large molecular systems, so the standard NH thermostat is typically still employed [37]. However, configurational thermostats may be useful for bulk NEMD simulations under high pressure and high shear rate conditions, where deviations from linear velocity profiles have been observed experimentally (see Section 2.2.4).

For confined NEMD simulations, alternative thermostating strategies are required; Yong and Zhang [54] evaluated some of these in 2013. In confined NEMD simulations, the thermostat is usually only applied to the wall atoms [54], allowing a thermal gradient to develop through the thickness of the lubricant film, as occurs in real experiments [55].
localised thermostat is often only applied in the Cartesian directions perpendicular to the shearing direction in order to avoid significantly affecting the dynamics. Applying the thermostat directly to the confined fluid molecules has been shown to artificially influence their structure, flow and friction behaviour under shear [56, 57]. For example, the slip behaviour of thermostatted molecules confined between rigid walls disagrees with that observed experimentally and in simulations where only flexible wall atoms are thermostatted [57] (see Section 2.2.3). In confined NEMD simulations of tribological systems, stochastic thermostats (e.g., Langevin [45]) are generally more popular than the NH thermostat since they can remove heat more efficiently [58]. For the relatively thin films usually studied in confined NEMD simulations, the temperature rise, which is generally largest in the centre of the film, only becomes significant at high shear rates [58]. More specialised thermostatting strategies have also been developed, for example, thermostats can be applied to “virtual” wall atoms rather than the wall atoms themselves [59]. This technique is particularly useful for large, complex systems where the wall atoms can be “frozen” to preserve structural stability and reduce the computational cost [59]. Stochastic boundary conditions (SBC) [60], in which the border atoms experience random friction forces (Langevin) to control the flow of energy through the system border, have also been shown to be effective in temperature control of confined NEMD simulations of tribological systems [61].

Pressure control is also important for NEMD simulations of tribological systems. In order to mimic experimental conditions, in bulk NEMD simulations, pressure is most commonly controlled using the Nosé-Hoover (NH) barostat [47, 48], which continuously alters the unit cell vectors, and thus the cell volume, to maintain a target pressure. This method can be used to sample the isothermal-isobaric or NPT ensemble [49]. For confined NEMD simulations, pressure can be controlled by setting a fixed separation distance between the confining walls, or by applying a constant normal force to one or both of the walls (e.g., Fig. 1(a)) [62]. The latter technique is usually preferable since the former has been shown to lead to unphysical behaviour in certain systems [62]. For incompressible systems, such as L-J fluids, it can also be beneficial to add a damping term to the barostat in order to prevent excessive oscillations after the normal force is applied [62].

An alternative approach for confined NEMD simulations is grand-canonical molecular dynamics (GCMD) [63, 64], in which the chemical potential, temperature, and pressure are held constant. This technique was developed from GCEMC (grand-canonical ensemble Monte Carlo) [65] to investigate interesting results from surface force apparatus (SFA) experiments of confined liquid films [66]. In GCMD simulations, a reservoir of fluid molecules surrounds the confined system [64], thus allowing molecules to be “squeezed out” from between the solid surfaces in one or more directions at high pressure [67]. GCMD is still less widely used than conventional confined NEMD, primarily because of the additional computational expense to simulate the reservoir molecules; however, it has recently been applied to study the behaviour of ionic liquid lubricants under shear [68].

A common concern in classical MD simulations is that their accuracy is heavily dependent on the force-fields used to describe the intermolecular and intramolecular forces between a system of interacting molecules. The functional forms of most classical molecular force-fields are quite similar, capturing both bonded (bond stretching, bending and torsion) and non-bonded (van der Waals, electrostatic) interactions [10]. These terms can be empirically parametrised to reproduce important experimental properties for a set of training compounds, derived from first principles calculations, or computed using a combination of the two. Force-field parameterisation is a complex and time-consuming process, and most of the force-fields used to study liquid lubricants were originally developed for biological applications [10]. The computational expense of classical MD simulations is usually dominated by
modelling the van der Waals and particularly the long range electrostatic interactions between atomic partial charges. The former is usually modelled by a Lennard–Jones (L–J) potential and is generally “cut off” at certain distance where the energy tends to zero (= 10 Å) for computational efficiency. Conversely, electrostatic interactions should not be cut off, but rather treated with a long range solver such as the Ewald summation [72] or particle–particle, particle–mesh (PPPM) [73] methods, since truncation can lead to unphysical results [74]. Note that confined NEMD simulations require a slab implementation of these algorithms [75].

A key target of NEMD simulations of tribological systems is to yield realistic viscous behaviour of lubricant-sized molecules, which requires the use of highly accurate force-fields. Quantitative comparisons between several popular force-fields for NEMD simulations of lubricants can be found elsewhere [10, 76, 77], but a brief overview is also provided here for completeness. Most historic simulations of tribological systems which include alkane molecules have employed united-atom (UA) force-fields where the nonpolar hydrogens are grouped with the carbon atoms to generate CH, CH₂ and CH₃ pseudo-atoms [10]. The popularity of UA force-fields has stemmed from the fact that they decrease the number of interaction sites by around 2/3 and computational expense by up to an order of magnitude compared to all-atom (AA) force-fields [78]. However, UA force-fields have been shown to lead to dramatic viscosity under-prediction for linear, long-chain alkanes (= 50% for n-hexadecane) compared to experiment. For alkanes with multiple short branches, UA force-fields can give reasonably accurate viscosity results (= 15% for squalane [79]), but for those with fewer, longer branches they are less accurate (= 50% for 9-n-octyldocosane).

AA force-fields are much more computationally expensive and, as originally developed, many of them resulted in anomalous solid–liquid phase behaviour for the long-chain alkanes relevant to tribology [80]. Fortunately, the latter issue has been overcome by re-parameterising AA force-fields specifically for long-chain alkanes [8, 9], which has led to far more accurate density and viscosity prediction (= 10% for n-hexadecane [10]), even under high temperature and high pressure (HTHP) conditions [10]. More modest improvements in the density prediction of lubricant-sized alkanes using AA force-fields can be made simply by reducing the intramolecular 1-4 scaling parameters for the L–J and electrostatic interactions using otherwise unmodified force-fields [80]. An increasing number of tribological NEMD studies are now employing accurate AA force-fields in order to yield more realistic viscous behaviour; however, for some types of molecules (multiple small branches, e.g., squalane) the additional computational expense compared to UA force-fields may not be justified.

1.3.3 Modified NEMD algorithms

Many modifications of the original NEMD methodologies have been made over the last three decades to improve their signal-to-noise ratio, simplify their implementation, and extend them to study different experimental conditions. Noteworthy in this set of theoretical tools is the so-called “transient time correlation function” (TTCF) method [39, 40], derived by Morriss and Evans in the late 1980s. TTCF can be used to generalise the Green-Kubo method [24, 25], which strictly is only applicable to equilibrium systems, to systems forced out of equilibrium by the imposition of an external field (e.g., shear rate). TTCF involves averaging correlation products from many independent NEMD trajectories, making it rather computationally expensive. However, the main advantage of TTCF is that it allows very low shear rates to be simulated compared to NEMD, allowing direct overlap with those used in experiments and real components [42]. Indeed, it has been successfully applied to study the viscosity of short alkanes (n-decane) at shear rates as low as 10⁵ s⁻¹ [81]. TTCF was initially used to study the rheology of bulk systems but it has also been used to study friction in confinement [82, 83]. The extent to which the TTCF can be used to calculate the friction coefficient of confined fluids has been much discussed, and alternative methods have also been proposed [84–86].

In 1999, Müller-Plathe [87] introduced an alternative nonequilibrium method for calculating the shear viscosity, so-called “reverse” NEMD or R-NEMD. The R-NEMD method reverses the cause-and-effect picture customarily used in NEMD in which the velocity gradient or shear rate is imposed and the momentum
flux or stress determined, to a stress being imposed and the consequent velocities observed. The method involves a simple exchange of particle momenta between boundary layer molecules, which is straightforward to implement. Moreover, it can be made to conserve the total energy as well as the total linear momentum, so no coupling to an external temperature bath is needed. The method was initially tested for an L–J fluid near its triple point and yielded a viscosity in agreement with literature results [87]. The method has since been used to predict the viscosity behaviour of simple alkanes [88] as well as more complex multicomponent systems such as lipid bilayers [89], ionic liquids [90], and polymer viscosity modifier additives [91, 92]. R-NEMD simulations are simple to implement, versatile and can be used to study transport properties in bulk or confined systems. However, unlike conventional NEMD, the external forces are not applied to the system in a manner consistent with experiments in R-NEMD simulations, and it has also been shown to under-predict the viscosity compared to conventional NEMD at high shear rates [93].

Another notable advance was the extension of NEMD methodology to simulate planar elongational flow (PEF) by Todd and Daivis in 1998 [94]. This is important in tribology since PEF occurs in the entrance region of high pressure, elastohydrodynamic contacts [95]. PEF simulations have been widely used to study polymers flow [96, 97] but more rarely for lubricant-sized molecules. Baig et al. [98] studied the rheological and structural properties of linear liquid n-alkanes (C_{10}, C_{12}, C_{20}) using NEMD simulations under PEF. They showed that these molecules behave very differently to the way they do under shear flow, with strong alignment of fully stretched chains at high elongation rates suggesting the formation of a liquid-crystal-like, nematic structure [98]. However, it should be noted that concerns have raised with the validity of the p-SLLOD algorithm used in these simulations [99]. Hunt et al. [100] developed an algorithm capable of imposing mixed Couette flow and PEF. Furthermore, Hartkamp et al. [101] applied TTCF to these mixed flows in order to capture the behaviour of fluids at lower shear rates.

Another method is driving spring-block NEMD, which was developed by Thompson and Robbins [102] to study the stick-slip behaviour observed in SFA experiments [103] (Section 2.2.2). In this technique, the sliding velocity is not applied directly to the wall atoms, but rather through a stage (or virtual atom) linked to the wall atoms via a spring. As the stage is moved, the spring is stretched and, if the confined film is liquid, the top plate accelerates until the steady-state spring force balances the viscous dissipation. If the film has crystallized, stick-slip behavior is observed; initially, the film responds elastically and the wall remains stuck, but eventually the spring force exceeds the yield stress of the film, and the top plate slides until the plate re-sticks and the process repeats. This technique has since been used to mimic SFA experiments of a range of strongly confined systems, including atomic fluids [71] and water [104].

1.4 Summary

Since their conception, NEMD simulation methodologies have been continually varied and refined in order to more faithfully recreate experimental conditions and provide more accurate results. Significant progressions have been made in temperature and pressure control, force-field accuracy, as well as the simulation algorithms themselves to allow a more diverse range of tribology problems to be investigated more accurately. Despite considerable advances, classical MD simulations are still limited in both the accessible time and length scales and their inability to study chemical reactivity. Thus, in the future, the most interesting methodological advances are envisaged to be in coupling MD simulations to smaller (quantum) and larger (continuum) scales; this will be discussed in Section 4.

2 NEMD simulations of liquid lubricants

Many tribology problems are well-suited to study by NEMD simulations. Shear can be applied to systems of lubricants and additives to give insights into their flow and friction behaviour. As discussed in Section 1, shear forces can be added either directly to the fluid molecules or by momentum transfer from sliding surfaces which confine the fluid (Fig. 1). While there are clearly significant differences between NEMD simulations of strongly confined, inhomogeneous systems and the bulk [105–107] (as justified below),
comparisons between the viscosities from these methods for homogenous flows have shown good agreement [55, 108, 109]. Many lubricant properties, both in the bulk and under confinement, have been investigated using NEMD simulations, providing information at a scale that is difficult to access experimentally [4]. Indeed, NEMD simulations have predicted some important lubricant behaviour many years before experimental techniques capable of providing comparisons on the same scales were developed. The following sections will discuss the use of NEMD to investigate the behaviour of lubricants in the bulk (Section 2.1), as well as when confined by solid surfaces (Section 2.2).

2.1 Bulk NEMD simulations of liquid lubricants

Bulk NEMD simulations are frequently used to investigate the change in viscosity with important variables such as the shear rate, temperature and pressure. As the shear rate is increased, lubricants often undergo a transition from a Newtonian regime (where the viscosity, \(\eta\), is independent of shear rate, \(\dot{\gamma}\)) to shear thinning behaviour (where \(\eta\) decreases with increasing \(\dot{\gamma}\)). The critical shear rate transition, \(\dot{\gamma}_c\), typically occurs at \(\dot{\gamma} \approx \tau^{-1}\), where \(\tau\) is the longest relaxation time at equilibrium (in the absence of shear). This is usually the rotational relaxation time, \(\tau_{\text{rot}}\) [37, 38], which increases with the chain length of the molecule [110]. Thus, longer simulations are required to obtain the Newtonian viscosity from NEMD simulations of larger molecules. As a result, early bulk NEMD studies were limited to studying the rheology of atomic Lennard–Jones (L–J) fluids and later short chain alkanes (\(\leq C_{10}\)) using UA force-fields [111]. Only more recently have they been applied to study lubricant-sized molecules (C_{20-40}). Some significant examples of bulk NEMD simulations of lubricant-sized molecules are discussed below. A list of bulk NEMD simulations applied to a wider range of fluids was compiled by Todd and Daivis [112] in 2007.

2.1.1 Viscosity prediction from bulk NEMD

Morriss et al. [113] were the first to study lubricant-sized molecules with bulk NEMD simulations in 1991. They used a UA model of \(n\)-eicosane (C\(_{20}\)) with a purely repulsive, Weeks, Chandler, Andersen (WCA) potential [114] and, as is common in NEMD simulations, they progressively increased the shear rate to study its effect on viscosity. At the lowest shear rates tested they observed a shear thinning regime, followed by shear thickening at higher shear rates. However, this shear thickening regime was later shown to be an unphysical consequence of the simplified force-field employed [49]. Moreover, the lowest shear rates they used were not low enough to capture the Newtonian plateau.

In 1998, Khare et al. [115] became the first to use NEMD to investigate the rheological behaviour of lubricant-sized alkanes using a UA force-field which included repulsive and attractive interactions through a Lennard–Jones (L–J) potential [116]. They were able to simulate lower shear rates and thus capture the Newtonian plateau for several linear (C\(_{16}\), C\(_{22}\), C\(_{28}\)) and branched (5,12-dipropylhexadecane) alkanes. The Newtonian viscosity of \(n\)-decane from the NEMD simulations was in good agreement with experiment; however, the viscosity results for longer alkanes were somewhat below the corresponding experimental values (e.g., 44% for C\(_{28}\)). However, the UA force-field employed was able to capture qualitatively the change in viscosity with lubricant molecular structure.

2.1.2 Viscosity temperature and pressure dependence

More recently, several bulk NEMD simulations of lubricant-sized alkanes focussed on calculating the Newtonian viscosity and also probing its change with temperature (viscosity index, VI or viscosity number, VN) and pressure (\(\alpha\)). The VI, VN, and \(\alpha\) are important indicators of lubricant performance, so there is considerable interest in their prediction by molecular simulation.

In 2000, Moore et al. [79] performed bulk NEMD simulations on three C\(_{30}\) isomers using a UA force-field [117]. They found that, while the viscosity of the linear (48%) and single-branched (36%) isomers were significantly underpredicted, viscosity prediction for squalane (15%), which contains multiple short methyl branches along its backbone, was more accurate [79]. They also analysed the viscosity at different temperatures to calculate the VI, which was in excellent agreement with experiment for all of the isomers (\(\approx 10\%\)) [79]. McCabe et al. [118] performed similar
simulations to calculate the VN of 9-octylheptadecane, 9-octyldocosane, and squalane using a UA force-field [117]. The viscosity was again significantly underpredicted but the VN values were within the statistical error of the experimental results for all of the molecules studied.

McCabe et al. [119] and Liu et al. [120] performed NEMD simulations of 9-octylheptadecane and a model poly-\(\alpha\)-olefin (PAO) base oil (1-decene trimer) respectively from ambient up to GPa pressure to predict their \(\alpha\) values. They both used a UA force-field [117] and although the Newtonian viscosity was significantly underpredicted (30%–70%), the calculated \(\alpha\) values were in good agreement with experimental results (\(\approx 10\%\)). It should also be noted that Ramasamy et al. [121] accurately predicted \(\alpha\) values for lubricant-sized molecules from their pressure-volume behavior in simple EMD simulations.

2.1.3 Shear thinning

Another focus of bulk NEMD simulations of lubricant-sized molecules has been investigating their shear thinning behaviour. There is still much debate regarding the most suitable shear thinning model to describe lubricant behaviour under high pressure, high shear rate, and elastohydrodynamic lubrication (EHL) conditions [41, 122, 123]. NEMD simulations have been used to investigate this behaviour and can give unique insights into behaviour under conditions which are challenging to probe experimentally [124–126]. Some investigations of shear thinning behaviour in lubricant-sized alkanes are discussed below.

In 1998, Cui et al. [110] used NEMD and a UA force-field [117] to study the rheological properties of a variety of linear (C\(_{10}\), C\(_{16}\), C\(_{24}\)) and branched (10-\(n\)-hexynonadecane, squalane) alkanes. At low shear rates, the viscosity behaviour transitioned to a Newtonian plateau and the zero-shear viscosity agreed with that from independent EMD simulations using the Green-Kubo method [24, 25]. At high shear rates, the viscosity showed a power-law shear thinning behaviour over several orders of magnitude in shear rate. This power-law shear thinning was shown to be closely related to the ordering of the molecules and the molecular architecture had a significant influence on the power-law exponent [110].

In 2002, Bair et al. [37, 38] compared bulk NEMD simulations of squalane using a UA force-field [117] with experimental measurements under EHL conditions in both the Newtonian and shear thinning regimes. This was the first comparison of the nonlinear rheology predicted by NEMD with experiment, and was thus the first experimental test of NEMD simulations in the shear thinning regime. The experimental and simulation data were separated by several orders of magnitude in shear rate; however, they collapsed onto the same time-temperature superposition master curve [127]. This master curve was successfully fit using the Carreau shear thinning model [128].

Liu et al. [129] proposed an approach to correlate shear thinning with the change in the molecular conformation, characterised by the radius of gyration of the molecule from NEMD simulations using a UA force-field [117]. This approach was tested by analysing the critical shear rate for squalane and 1-decene trimer, and then extended to study the behaviours of different molecular weight PAO structures. The changes in viscosity and radius of gyration with shear rate were adequately fit with the Carreau equation [128]. Their analysis led to a relationship between molecular weight and critical shear rate for PAO structures, and the calculated critical shear rates agreed well with the Einstein-Debye equation [130] for low molecular weight PAOs.

Recently, Jadhao and Robbins [131] used NEMD simulations to determine the nonequilibrium shear response of squalane over a wide range of pressures and temperatures using a UA force-field [117]. The viscosity-shear rate behaviour from the NEMD simulations agreed well with experimental data conducted at lower shear rates (10\(^{2}\)–10\(^{4}\) s\(^{-1}\)). The Eyring shear thinning model [132] could be used to extract an accurate prediction for the Newtonian viscosity from NEMD simulations at high shear rates (> 10\(^{5}\) s\(^{-1}\)) as long as there was a substantial linear region in the plots of stress vs. log(shear rate). This condition was met whenever the Newtonian viscosity was \(\geq 1\) Pa·s, where the fluid was in a “glassy” regime. At lower viscosity, the response deviated from the simple Eyring form, indicating that the motion could not be described by a single activation barrier or rearrangement mechanism. The response in this “low viscosity”
2.2 Confined NEMD simulations of liquid lubricants

In addition to bulk lubricants, NEMD simulations have provided significant insights into the behaviour of thin lubricant films confined between solid surfaces. Confined lubricants behave very differently to their bulk counterparts under shear, and density inhomogeneity can lead to the breakdown of continuum approximations such as Navier-Stokes hydrodynamics [133]. Density oscillations in confined fluids were first identified in Monte-Carlo (MC) simulations of L-J liquids [65, 134]. Shortly afterwards, surface forces apparatus (SFA) experiments by Horn and Israelachvili on confined octamethylcyclotetrasiloxane (OMCTS) [135] (a quasi-spherical molecule) showed that the force between two surfaces does not evolve smoothly with the separation distance as predicted by the van der Waals theory. Instead, in films below six to ten molecular diameter thickness, the wall-wall interaction oscillated between repulsion and attraction states, with a periodicity roughly equal to the molecular diameter. This behaviour, known as the “solvation force”, was the first experimental evidence of density inhomogeneity in confined fluids [135]. Subsequently, the first MD simulations of confined fluids by Magda et al. [136] agreed with earlier MC and experimental results, showing that strongly confined films have inhomogeneous density profiles across the channel. NEMD simulations have also given many additional insights into the nonequilibrium behaviour of confined films, as discussed below.

2.2.1 Density and viscosity inhomogeneity

The first NEMD simulations of fluid films confined by walls were performed in 1987 by Bitsanis et al. [28]. They used the NEMD procedure developed by Ashurst and Hoover [27], but also included a continuous wall potential to confine an L-J fluid. They hypothesised that the inhomogeneous films could be viewed as a set of homogeneous layers and developed the local average density model (LADM), where each layer is given its own density and thus its own local viscosity. The LADM gave satisfactory results in predicting the flow of moderately-confined fluids [28]. However, the LADM model failed to predict many of the properties of strongly-confined films (≤4 molecular layers), most notably the enhanced viscosity, where the density variation alone cannot explain the local dynamic properties of the film [137].

In conventional rheology models, the resistance to flow of a lubricant is characterized by its shear viscosity, \( \eta \). For bulk systems (Section 2.1), \( \eta \) depends purely on the fluid and its response to variations in pressure, temperature, and shear rate. Conversely, the dynamics of confined fluids are also influenced by factors that are not intrinsic to the fluid, so the definition of a bulk viscosity is not strictly appropriate [138]. However, the ratio of the shear stress to the shear rate still gives insight into the resistance to flow of confined films and can be compared directly to experiments [137]. Most experiments and NEMD simulations of confined films report this value as the “effective” shear viscosity, \( \eta_{\text{eff}} \).

When a fluid film is thicker than approximately ten molecular diameters, at which point its centre becomes a structureless medium, the experimentally measured \( \eta_{\text{eff}} \) is usually within around 10% of the bulk viscosity [103, 139]. Experiment and NEMD simulations have shown that when the wall separation is decreased below this level, \( \eta_{\text{eff}} \) of confined L-J fluids [137, 140], linear alkanes [141, 142] as well as many other organic liquids [143] begins to increase significantly. At this point, continuum approximations such as Navier-Stokes hydrodynamics begin to break down [133]. When the film thickness is reduced to around six molecular layers, \( \eta_{\text{eff}} \) has been shown to increase by several orders of magnitude for a range of fluids [141–144]. For L-J fluids, this viscosity enhancement has been attributed to the crystallisation of the entire film [145, 146]. Compared to L-J fluids, linear alkanes are less susceptible to crystallisation and so either a second order glass transition [145] or the formation of “crystal bridges” [141] have been attributed to the viscosity enhancement. However, more recent MD simulations by Docherty and Cummings [147] using an all-atom force-field suggested that both cyclic (cyclohexane) and long-chain (n-dodecane) alkanes undergo a rapid, first order phase transition to a layered and ordered (i.e., crystalline) system (see Fig. 2).
at surface separations equivalent to fewer than nine molecular layers. These simulations support experiments which showed an abrupt transition from liquid-like to solid-like behaviour upon increasing confinement [143, 144], rather than a more gradual transition to a glassy state (i.e., vitrification) [142, 148]. It is noteworthy that most evidence for vitrification originated from SFA experiments which were later shown to be contaminated with platinum nanoparticles [149], which could go some way to explaining differences between experiments. A more in-depth discussion of the discrepancies between different experimental techniques and MD simulations of various levels of complexity was given in 2010 by Cummings et al. [150]. It has also been shown through experiments [148] and NEMD simulations [145] that elevated pressure can induce such transitions, as well as the associated increase in $\eta_{\text{eff}}$, in thicker films.

The transition from liquid-like to solid-like behaviour in confined fluids also depends strongly on molecular shape and surface structure. MD simulations have shown that branched or flexible molecules are more resistant to solidification and remain liquid-like under more severe confinement and higher pressures compared to linear or inflexible molecules [151, 152]. Thus branched molecules can give lower $\eta_{\text{eff}}$ under strong confinement than linear molecules, despite their higher bulk viscosity [103]. It should be noted most NEMD simulations of confined film behaviour employ atomically-smooth surfaces, but it has been found that surface roughness frustrates the ordering of the neighbouring fluid molecules so that films are more liquid-like on surfaces with nanoscale roughness compared to those which are atomically-smooth [153, 154].

When the film is thinner or when the pressure is increased, the Newtonian plateau is displaced to lower shear rates. Thus, at the shear rates accessible to NEMD simulations (generally $> 10^7 \text{ s}^{-1}$), confined fluids are often highly shear thinning [155]. A viscosity-shear rate power law was identified for L-J fluids at different pressures and thicknesses [145]. Both the Carreau [90] and Eyring [27] shear thinning models have been successfully used to describe the viscosity-shear rate behaviour of confined fluids from NEMD simulations.

In stark contrast to the elevated viscosity results discussed above, some NEMD simulations [146] and SFA experiments [156] have reported an almost vanishing friction coefficient (superlubricity) in strongly confined lubricant films. This low friction was initially attributed to boundary slip (Section 2.2.3) at incommensurate solid–fluid interfaces, but subsequent NEMD simulations run for longer simulation times [157–159] have suggested that slip does not occur at the interface, but rather shear-induced alignment causes slip between well-defined molecular layers within the film itself. Surprisingly, there have been relatively few experimental or MD studies which have investigated this phenomena in the last decade, and reductions in the effective viscosity due to confinement certainly remain the exception rather than the rule.

NEMD results for the change in viscosity with confinement have also been used to improve the predictive capabilities of continuum models. For example, Martini et al. [160] used NEMD simulations to extract the film thickness and velocity dependence of the
viscosity of \(n\)-decane confined and sheared between atomically-smooth gold surfaces. This information was then fed into a full numerical solution for EHL problems [161] to assess the impact of these viscosity changes on the predicted film thickness.

As well as the enhanced effective viscosity, confined NEMD has been used to study several tribological phenomena where the flow becomes inhomogeneous or discontinuous. Examples include stick-slip (Section 2.2.2), boundary slip (Section 2.2.3), and shear localisation (Section 2.2.4). These behaviours can arise directly as a result of strong confinement or from extreme applied pressure and shear rate conditions in thicker films [162]. NEMD simulations have and continue to play a central role in improving our understanding of all of these phenomena, which is critical to help design better lubricants.

2.2.2 Stick-slip

Stick-slip occurs when, instead of sliding over each other smoothly, solid surfaces alternately stick together and then slip past each other [102]. In 1990, SFA experiments of several non-polar fluids strongly confined and sheared between atomically-smooth mica surfaces showed a saw-tooth pattern in the friction coefficient over time [103]. This stick-slip friction response was also seen in spring-block NEMD simulations (Section 1.3.3) of a confined L–J fluid performed by Thompson and Robbins [102] the same year. The authors suggested that stick-slip motion involved periodic shear-melting transitions (slip) and recrystallization (stick) of the thin fluid film [102]. However, more recent spring-block, LVMD simulations of a confined L–J fluid by Lei and Leng [71] were in contrast with this shear-melting hypothesis, as they suggested that the ordered, solidified film persisted during both the stick and slip phases. They proposed that periodic slip events, either within the fluid film itself or at the wall-fluid interface (Section 2.2.3) were actually the cause of the stick-slip friction response [71]. Recent surface force balance (SFB) experiments of OMCTS from Rosenhek-Goldian et al. [163] supported the NEMD results of Lei and Leng [71], as they observed no dilation of the inter-surface gap, which the authors suggest discounts the shear melting hypothesis. However, the resolution of the experiments of Rosenhek-Goldian et al. [163] was questioned by Jee et al. [164] who suggested that the “missing” dilatancy behaviour was actually observed experimentally on the same system several years earlier [165]. Thus there is clearly still much to learn regarding the stick-slip friction behaviour observed for strongly confined fluids, and NEMD simulations are likely to continue to play a central role in improving our understanding of this phenomenon.

2.2.3 Boundary slip

The no-slip boundary condition is the common assumption in continuum hydrodynamics which states that fluid adjacent to a sliding solid surface moves at the same velocity as the surface [166]. However, several experiments have shown that this is not always the case for fluids confined by smooth surfaces with oleophobic coatings (usually self-assembled monolayers or SAMs), as reviewed in 2005 by Neto et al. [167]. Boundary slip occurs when the fluid cohesion is relatively stronger than its adhesion to the sliding surfaces [105], meaning that the surfaces do not transfer sufficient momentum to fully shear the fluid [168]. This can lead to large reductions in friction compared to the Couette case [146]. Slip has been commonly observed in NEMD simulations of thin fluid films confined between atomically-smooth surfaces [57, 168–171]. Many of these studies have quantified the slip length, which is defined as an extrapolated distance into the wall where the tangential velocity component vanishes (see Fig. 3) [172]. There is significant disagreement between the slip length results from different experimental procedures, and several mechanisms for slip have been proposed, from enhanced liquid mobility at the surface [173, 174] to bubble formation at the fluid-solid interface [175]. NEMD simulations offer a useful opportunity to test these mechanisms, and they have led to more accurate models for boundary slip behaviour in a range of systems [57, 176].

NEMD simulations have also been used to study the change in slip length under a wide range of conditions. These simulations have shown that the slip length generally increases with increasing sliding velocity and pressure [168–171]. At very high sliding velocity, the slip length asymptotes towards a constant value [57]. An important observation is that realistic velocity-slip behaviour can only be captured in NEMD simulations when flexible, rather than rigid walls are employed [57]. NEMD simulations have shown that the slip length
increases with increasing molecular chain length and degree of branching, in line with the viscosity [151]. Inflexible molecules, which can conform to the surface only with a greater expense of conformational energy generally exhibit more slip [177]. Conversely, the slip length decreases with increasing surface-fluid attraction strength and film thickness [169–171], and is virtually eliminated in the presence of atomic-scale surface roughness, which frustrates molecular layering [153, 154]. This final observation is critical since most experimental surfaces will contain fractal roughness, even down to the atomic scale [178], which is likely to discourage boundary slip on conventional surfaces. Moreover, large-scale NEMD simulations of > 100 nm $n$-hexane films between atomically-smooth iron surfaces under realistic EHL conditions ($F_N = 1$ GPa, $\dot{\gamma} = 10^5$ s$^{-1}$) confirmed the suitability of a macroscopic no-slip boundary condition for this particular system [170]. NEMD simulations of $n$-alkanes ($C_3$-$C_{10}$) confined between surfaces with patterned stick and slip regions by Savio et al. [179] showed that such heterogeneous surfaces can lead to cavitation under sufficient confinement.

Continuum models for fluid flow, such as the Reynolds equation [180], generally assume a no-slip boundary condition [181]. Slip lengths extracted from NEMD simulations have been fed into a modified Reynolds equation to study the effect of slip at one [182] or both [168] of the surfaces on flow behaviour, film thickness, and EHL friction. These results have shown that slip can lead to a significant reduction in friction and film thickness compared to the no-slip case, particularly at high shear rates.

Boundary slip has been directly observed in real tribological contacts under EHL conditions for a viscous polybutadiene lubricant between smooth glass surfaces with an oleophobic coating using [183]. Here the slip length was shown to increase with pressure and slip was also shown to substantially reduce friction under EHL conditions [183]. Slip has also been inferred from optical interferometry experiments on glass and even steel surfaces under EHL conditions for viscous polybutadiene and polyphenyl ether (5P4E) lubricants [184]. However, slip has not yet been observed experimentally for realistic lubricants confined between conventional surfaces and thus its significance in tribology remains uncertain [185].

2.2.4 Shear localisation

In the last few years, NEMD simulations of confined L–J liquids sheared at GPa-level pressures have revealed a range of nonequilibrium phase behaviour [186–188]. These simulations were designed to mimic the conditions found in the EHL regime, which is particularly important for machine components that roll and slide together, for example; rolling bearings, gears, constant velocity joints and cam/follower systems [41]. Although shear in these simulations was applied through the walls, the thickness of the films was above the ten molecular-layer limit in which fluid films experience enhanced viscosity as a direct result of confinement [103, 139]. These systems and conditions could also be investigated through bulk NEMD simulations; however, in this case most thermostats enforce a linear (Couette) flow profile [50] (Section 1.3.1), which may not be the case in reality, as discussed below.

Above a certain pressure, the NEMD simulation [186–188] flow profiles transitioned from Couette-like to those indicative of shear localisation (or banding). Nonequilibrium phase diagrams were constructed...
for L–J fluids to map the changes in behaviour with pressure and shear rate (or sliding velocity). Figure 4 shows the different types of flow behaviour which have been observed in NEMD simulations of fluids under EHL conditions. Unlike the simulations described in Section 2.2.3, in these high pressure simulations, slip does not occur at the solid–liquid interface (PS in Fig. 4), but rather within the fluid itself. At low shear rates, shear was localised close to the walls, while the central part of the fluid formed a solid-like, unsheared region. This behaviour is usually referred to as “plug-slip” (PS) (PS in Fig. 4). At higher shear rates, the localisation of shear velocity gradient moved to the centre of the film, while the fluid near the walls became solid-like. This behaviour is typically known as “central localisation” (CL) (CL in Fig. 4). As the shear rate is increased in the CL region, the central liquid region widens until eventually a Couette profile is recovered [186–188]. Under the lowest shear rates studied, stick-slip friction behaviour was observed due to the commensurability of the walls and the solid-like fluid [188]. This mirrors the stick-slip behaviour [102] observed for more strongly confined films at ambient pressure (Section 2.2.2), supporting the postulate due to Robbins and Smith [162] that confinement-induced phase transitions observed in SFA experiments and highly confined NEMD simulations may be closely related to the pressure-induced transitions observed in macroscopic tribology experiments.

Slip within the fluid itself, rather than at the solid-fluid boundary, was first observed in NEMD simulations of an L–J fluid in 1990 by Thompson and Robbins [189]. In these simulations, only the first one or two layers of fluid were “locked” to the surface (due to strong wall interaction) and the rest of the fluid was sheared. This is distinct from the shear localisation behaviour observed in the high pressure NEMD simulations described above [186–188] where the “unsheared” region extends well beyond the direct influence of the wall atoms. Subsequent confined NEMD simulations of L–J fluids between one crystalline and one amorphous wall by Butler and Harrowell [190–192], L–J glasses (mixture of different sized L–J atoms to frustrate crystallization) by Varnik et al. [193], as well as L–J glasses and amorphous polymers by Rottler and Robbins [194] have shown various forms of shear localisation. Thus the onset of shear localisation at high pressure is probably related to localised phase changes within the lubricant (either vitrification or strong crystallisation), similar to those occurring due to confinement (Section 2.2.1) [162].

The discovery of such nonequilibrium phase behaviour by NEMD simulations of L–J fluids demonstrates the advantage of carrying out simulations using simple model systems. By doing so, it helps to reveal whether the phenomenon is “universal” to the liquid state and arises for all (or at least the vast majority of) liquids under certain conditions (which vary between molecules), or is specific to certain types of molecule [195].

In the field of tribology, shear localisation has only been directly observed experimentally in viscous model lubricants, such as polybutadiene and polyphenyl ethers (e.g., 5P4E) [196–201]. However, it has also been inferred from film thickness measurements in
realistic mineral oils [202, 203] and there is currently a growing consensus that this phenomenon may also be important for more realistic lubricants under extreme EHL conditions [185]. NEMD simulations provide a useful complement to these experiments since they can more readily study the challenging conditions of interest.

Recently, extensive NEMD simulations of weakly confined L-J fluids under GPa pressure have revealed that the nonequilibrium phases correlate well with corresponding friction maps [188]. The characteristics of friction on the atomic scale can deviate quite significantly from the laws of classical friction. For example, when the applied conditions yield shear localisation, the friction coefficient of L-J fluids can decrease with increasing pressure and shear rate. Such behaviour is not commonly observed experimentally for lubricants, but is more similar to how traction fluids behave [204] above their critical shear stress [185]. The nonequilibrium phase diagrams and friction maps for L-J fluids were also found to be sensitive to the degree of the wall roughness on the atomic scale [188].

In similar high-pressure NEMD simulations of L-J glasses [187], crystallisation was suppressed and only liquid or weak CL phases were observed. In these systems, the friction coefficient increased linearly with log(shear rate), as has commonly been observed for realistic lubricants, such as PAO, under EHL conditions [41]. In this case, the friction-shear rate behaviour can be adequately described using stress-augmented thermal activation theory [205], according to the rheological models of Eyring [132]. An underlying assumption in current models of EHL friction is that, in the absence of thermal effects, or until a critical shear rate is reached, the rheological properties of the lubricant do not vary through its thickness. This is implicit in applying shear thinning models to predict friction in the EHL regime since they generally assume that the film is subject to planar Couette flow. Thus significant modifications of these models may be required where there are deviations from Couette flow, such as shear localisation [41].

Recently, a combination of tribological experiments and confined NEMD simulations has been used to investigate the effect of base fluid molecular structure on nonequilibrium phase behaviour and friction by Ewen et al. [29]. An extensive parameter study, including several lubricant and traction fluid molecules subjected to pressures (0.5–2.0 GPa) and shear rates (10^1–10^10 s^-1) typical of the EHL regime, revealed clear relationships between the friction and flow behaviour. There was good agreement between the friction-shear rate behaviour from tribology experiments and the NEMD simulations conducted at higher shear rates. Lubricants, which are flexible, broadly linear molecules, gave low friction coefficients that increased with shear rate and pressure in both the experiments and the simulations. Conversely, traction fluids, which are based on inflexible cycloaliphatic groups, gave high friction coefficients that were less sensitive to shear rate and pressure. The observed differences in friction behaviour was rationalised through the stronger shear localisation which was observed for the traction fluids (CL and PS) compared to the lubricants (Couette and CL) in the NEMD simulations [29]. Washizu et al. [206] also reported shear localisation (CL) in NEMD simulations on n-hexane confined between gold slabs at high pressure (0.1–8.0 GPa). At the transition from Couette flow to CL, they observed limiting shear stress behaviour, where the friction coefficient became insensitive to increasing pressure.

2.3 Summary

Many bulk NEMD simulation studies have accurately the Newtonian viscosity of realistic lubricant molecules as well as its temperature and pressure dependence. The accuracy of these simulations has reached a point where they can be used as a predictive tool to compare the rheology of different lubricant molecular structures. Such structure-property relationships can be extremely useful in lubricant design, and NEMD simulations provides the opportunity to probe a wide parameter space cheaply and efficiently. Bulk NEMD simulations have also given atomic-scale insights into lubricant shear thinning behaviour under extreme conditions and they are expected to play a central role in helping to improve macroscopic shear thinning models.

Confined NEMD simulations have also proved extremely useful in understanding the behaviour of thin lubricant films between solid surfaces. This includes building fundamental understanding of the density and viscosity inhomogeneity in confined films, as well as knowledge of important tribological
phenomena under shear such as stick-slip, boundary slip, and shear localisation. Confined NEMD simulations are particularly useful since they mirror the layout and physics of real tribology experiments and require fewer assumptions with respect the flow profile and thermostating procedure compared to bulk NEMD simulations. Thus, confined NEMD simulations are being increasingly used to study relatively thick lubricant films in tribology. Interestingly, there seem to be inherent similarities between the tribological phenomena observed in strongly confined lubricant films and thicker films subjected to high pressures.

Critical to any simulation technique is validation through direct comparison with experiment. As discussed above, this is becoming possible for NEMD simulations thanks to significant improvements in simulation algorithms and computer hardware, which have allowed larger systems and lower shear rates, of direct relevance to tribology experiments, to be investigated.

3 NEMD simulations of lubricant additives

A more recent area of research has been the application of MD to study the behaviour of lubricant additives. The function of many lubricant additives depends on chemical reactions, which are generally not considered in classical MD simulations. Despite this, classical MD simulations have given unique insights into the behaviour of a range of lubricant additives including detergents [207, 208], dispersants [209, 210], viscosity modifiers [91, 92, 211], anti-wear additives [212–215], and corrosion inhibitors [216–218]. The most widely studied class of lubricant additives with MD are friction modifiers [219], which are a well-suited application of confined NEMD simulations.

The need for greater energy efficiency to reduce energy consumption has led to a shift toward lower-viscosity lubricants, which means that an increasing number of engineering components operate under boundary lubrication conditions [219]. As a result, lubricant additives that reduce friction and wear under boundary conditions (i.e., friction modifiers) are of increasing importance [219]. Several classes of friction modifier additives exist, the main ones being organic friction modifiers (OFMs), functionalised polymers, organo-molybdenum additives and dispersed nanoparticles [219]. An overview of the experimental advances on friction modifier additives is given by in 2015 by Spikes [219]. Sections 3.1 and 3.2 will focus on classical MD simulations of OFMs and nanoparticle friction modifiers respectively, and a brief discussion of “reactive” MD simulations of other additives is given in Section 4.1. In general, the simulation methods described in Section 1.3 (e.g., thermostats, barostats, force-fields) are also applicable to study these more complex additive systems.

3.1 Organic friction modifiers

OFMs are amphiphilic surfactant molecules that contain nonpolar hydrocarbon tail groups attached to polar head groups. Their main friction reduction mechanism involves the adsorption of the polar head group to metal, ceramic, or carbon-based surfaces, with strong, cumulative van der Waals forces between proximal nonpolar tails leading to the formation of incompressible monolayers that prevent contact between solid surfaces to reduce adhesion and friction [220]. Many different head and tail groups have been tested in the literature, as discussed elsewhere [219]. Classical NEMD simulations can be used to simultaneously probe the nanoscale structure and friction of OFM films, making them a valuable complement to experiments. Adsorption of amphiphiles from solution to solid surfaces using realistic additive concentrations is a relatively slow process, and close-packed monolayers can take hours to form in the absence of shear [221–223]. As a result, MD simulation of the adsorption process using realistic concentrations and all-atom force-fields are currently unattainable [224]. However, extensive experimental analysis has shown that OFMs do eventually form close-packed monolayers on a range of contact surfaces, and that the formation of these films is accelerated by the shearing process [221–223]. Therefore, MD simulations are generally constructed with pre-formed OFM monolayers by placing additive molecule head groups close to solid surfaces at the beginning of the simulation. This is similar to the Langmuir–Blodgett experimental approach to quickly form surfactant films with a controllable packing density (or coverage) on solid surfaces [225]. MD simulations generally consist of OFM monolayers adsorbed on two solid surfaces, sometimes also with
a layer of confined lubricant in between. EMD simulations have been performed to study the structure of OFM films on solid surfaces. Confined NEMD simulations have also been performed (Fig. 1(a)) to study the structure, flow, and friction behaviour of the films. The results of these OFM simulations are discussed below.

3.1.1 Film structure

Moller et al. [226] were the first to use EMD simulations to study the structure of OFM films (stearic acid) on solid (graphite) surfaces in 1991. They varied the size of the head group to study its effect on the tilt angle from the surface normal. Their simulations predicted that, under ambient conditions, OFM molecules were normal to the surface at high surface coverage (head group areas below 21 Å²), but became tilted from the normal at lower surface coverage. They suggested that tilting maximises the van der Waals attraction between the chains by increasing the packing efficiency [227]. The tilting transition was attributed to the packing of the hydrogen atoms belonging to methylene groups on neighbouring molecules, meaning that the correct behaviour could only be accurately reproduced using all-atom (AA) and not united-atom (UA) force-fields [226]. Recent NEMD simulations by Ewen et al. [10] showed that the use of AA force-fields is also critical to give accurate flow and particularly friction behaviour for OFM films under shear.

Davidson et al. [228] performed EMD simulations to help elucidate structure-activity relationships of mono-, di-, and tri-oleyl glyceride OFMs. Mono-glyceride molecules with two hydroxyl groups and one alkyl chain packed more efficiently than di-glycerides which contain one hydroxyl group and two alkyl chains. This was because the mono-glycerides occupied approximately half the area (22 Å²) needed by di-glycerides (46 Å²) and formed significantly more intermolecular hydrogen bonds. The density profiles showed significant mixing of the hydrophobic tail groups with the non-polar solvent and the distribution of tail group torsion angles suggested the glycerides formed liquid-like rather than solid-like films at room temperature. Experimentally measured friction coefficients of equimolar glyceride solutions showed that the efficacy as friction modifiers varied in the order mono-, di- and then tri-oleyl glyceride, which was consistent with the EMD film formation results.

Greenfield and Ohtani [229] used EMD simulations to investigate the orientation of monolayers of model OFMs (head group area 17 Å²) confined between atomically-smooth surfaces. The normal load, fluid layer thickness, and additive surface concentration dependencies agreed with those measured experimentally using the SFA [230]. Greenfield and Ohtani also performed NEMD simulations of the same systems and found that the OFM film structure remained virtually unchanged under a wide range of normal pressures (3.2–22 MPa) and sliding velocities (0.5–7.5 m·s⁻¹) [231].

Doig et al. [232] also used NEMD simulations to gain insights into sum-frequency spectroscopy (SFS) and polarised neutron reflectometry (PNR) experiments [233] of the structure of hexadecylamine films adsorbed on iron oxide surfaces in dodecane and hexadecane. At the highest surface coverage investigated in the NEMD simulations (1.8 nm⁻²), the film thickness was about 15–20 Å, and the molecules had an average tilt angle from the surface normal of 40°–50°, in agreement with experiment. The in-layer ordering of the hexadecylamine head group atoms was found to be dictated by the crystalline structure of the iron oxide surface but, at the coverages studied, this influence quickly diminished further along the tail group.

Polar OFM molecules generally form dimers or reverse micelles [219] in non-polar lubricants since they are more thermodynamically stable. For some OFMs, such as glycerides, it remains unclear whether close-packed monolayers form on the surface, as experimentally observed for fatty acids [221–223], or whether they have a different friction reduction mechanism. Bradley-Shaw et al. [234] used NEMD simulations to study the stability of reverse micelles, formed by glycerol mono-oleate (GMO) OFMs in organic solvents [235], confined and sheared between mica surfaces. The GMO reverse micelles were found to be metastable in the absence of shear for all of the systems studied. When shear was applied, the reverse micelles remained intact and adsorbed onto one of the surfaces in n-heptane, thus becoming surface micelles. In dry toluene, the reverse micelles broke apart under shear, forming disordered monolayers on the surface. However, in toluene with water present, they survived and became surface micelles.
3.1.2 Friction behaviour

In 1993, Glosli et al. [236] became the first to use NEMD to investigate the friction of OFM-like systems. They studied the friction between two close-packed monolayers (head group area 17 Å²) bound at their ends to two rigid substrates (similar to OFM molecules absorbed on solid surfaces). Depending on the interfacial interaction strength between the two layers, energy dissipation either occurred by a discontinuous “plucking” mechanism (strong interaction) or a continuous “viscous” mechanism (weak interaction). Friction during the plucking mechanism followed the simple stress-augmented thermal activation model proposed by Briscoe and Evans [225] from experimental friction data for Langmuir-Blodgett films. Friction during the viscous mechanism was generally much lower but increased at the rotator transition temperature of the monolayers [236] (T at which molecules become rotationally ordered in a herringbone-like pattern).

Kong et al. [237] performed NEMD simulations of adsorbed monolayers of dialkylammonium chloride surfactants between charged solid surfaces. They found that the friction coefficient decreased with increasing normal force (30 to 450 MPa), decreased with increasing surfactant coverage (77–50 Å²), and increased with increasing sliding velocity (1–100 m·s⁻¹). Moreover, the friction between the OFM films correlated well with the amount of layer overlap or interpenetration, as defined by the common area under the chain density profiles. Each dialkylammonium chloride surfactant molecule contains two alkyl chains and, in a subsequent NEMD study [238], the effect of monolayers with asymmetrical (C₁₀C₁₈) and symmetrical (C₁₆C₁₈ and C₁₈C₁₈) tail groups was investigated. The friction between the layers of asymmetrical surfactants was greater than that between layers of symmetrical surfactants, suggesting that films with uneven heights were less effective in reducing friction.

Eder et al. [239] performed NEMD simulations of several OFMs on iron surfaces, varying the OFM type (stearic acid, oleic acid, methyl stearate), OFM surface coverage (2.54–4.57 nm⁻²), the pressure (0.07–1.60 GPa), and the surface roughness. The resulting load versus friction behaviour was then analysed to study these parameters affected the friction coefficient and the extrapolated friction force at zero load (Derjaguin offset). A smooth-particle-based evaluation method [240] was used to visualize the sliding interface between the two OFM layers. This allowed the definition and calculation of a dimensionless normalized sliding resistance area, which was then related to the Derjaguin offset. The friction coefficient and Derjaguin offset were found to generally increase with increasing surface roughness, and decrease with increasing OFM coverage. The Derjaguin offset was eliminated by the addition of a thin n-hexadecane film between the OFM layers.

Doig et al. [241] used NEMD simulations to examine the structure and friction of stearic and oleic acid OFM films adsorbed on iron oxide surfaces and lubricated by squalane under hydrodynamic conditions (0.1 GPa). At high surface coverage (2.59 nm⁻²), the measured properties of stearic acid and oleic acid films were very similar. At low (0.58 nm⁻²) and intermediate (1.30 nm⁻²) surface coverages, the presence of Z-unsaturation (in oleic acid) resulted in less penetration of lubricant in the surfactant film and less layering of the lubricant near the film. The friction coefficient increased linearly with log(shear rate) within the hydrodynamic lubrication regime, consistent with a stress-augmented thermal activation model [225] and macroscopic friction experiments [242]. Lubricant penetration and layering were found to affect the hydrodynamic friction coefficient. Oleic acid friction depended more weakly on surface coverage, while stearic acid allowed more lubricant penetration, and thus friction increased significantly with decreasing surface coverage.

Recently, Ewen et al. [243] used NEMD simulations to examine the atomistic structure and friction properties of commercially relevant organic friction modifier (OFM) monolayers adsorbed on iron oxide surfaces and lubricated by a thin, separating layer of n-hexadecane (Fig. 5). The hexadecane thickness was determined at the studied pressure (0.5 GPa) in preliminary LVMD squeeze-out simulations. Carboxylic acid, amide, and glyceride OFMs, with saturated and Z-unsaturated hydrocarbon tail groups, were simulated at various surface coverages (1.52–4.56 nm⁻²) and sliding velocities (0.5–20 m·s⁻¹). At low and medium coverage, the OFMs formed liquid-like and amorphous monolayers, respectively, which were
significantly interdigitated with the hexadecane lubricant, resulting in relatively high friction coefficients. At high coverage, solid-like monolayers were formed which, during sliding, resulted in slip planes between well-defined OFM and hexadecane layers, giving a significant reduction in the friction coefficient [243]. In agreement with experiment [242], OFMs with glyceride head groups gave significantly lower friction coefficients than amide and particularly carboxylic acid head groups. For all of the OFMs and coverages simulated, the friction coefficient was found to increase linearly with log(sliding velocity); however, the gradient of this increase depended on the coverage [243]. The friction-sliding velocity behaviour from the NEMD simulations was consistent with a stress-augmented thermal activation model [225] and macroscopic boundary friction experiments [242].

Ewen et al. [244] also used NEMD simulations to examine the structure and friction of OFM (stearic acid) films adsorbed on iron surfaces with random nanoscale roughness (0.2–0.8 nm root-mean-squared). The OFMs were strongly adsorbed, were not squeezed out, and prevented direct contact of asperities under the high pressures simulated (0.5–2.0 GPa). Increased coverage generally resulted in lower lateral (friction) forces due to reductions in both the friction coefficient and Derjaguin offset. Rougher surfaces led to more liquid-like, disordered films, but the friction coefficient and Derjaguin offset were only slightly increased. These results suggested that OFM films are almost as effective on contact surfaces with nanoscale roughness as those which are atomically-smooth. Further NEMD simulations of OFM films using rough surfaces, or perhaps GCMD simulations [63, 64], would be useful to improve understanding of their resilience to high local pressures experienced in the boundary lubrication regime.

Although the focus of this review is on lubricants for macroscale engineering components such as engines and transmissions, NEMD simulations of confined films have also been performed to study friction modifier additives for microelectromechanical systems (MEMS) and nanoelectromechanical systems (NEMS). In this context, NEMD simulations have been used to investigate the friction behaviour of monolayers of carboxylic/perfluorocarboxylic acids [245–247] and alkylsilanes [248–251] on silica surfaces. The tribological behaviours of model self-assembled monolayers (SAMs) such as thiols on gold surfaces have also been widely studied with NEMD [252–254]. These simulations have revealed the effect of head group and tail group of the nanoscale structure and friction behaviour of the monolayers.

3.2 Nanoparticle friction modifiers

In recent years, there has been growing interest in the possibility of using colloidal solid particles in the size range 1–500 nm as friction modifier and anti-wear
Many types of nanoparticle additives have been tested, from various carbon allotropes to metals and inorganic salts [219]. The development of fullerenes [255] and inorganic fullerenes has generated particular interest since their graphitic-like structures bear some resemblance to lamellar materials which are known to effectively reduce friction [219].

Despite this interest, there is still considerable uncertainty as to the friction reduction mechanisms of nanoparticle additives [219]. When dispersed in a lubricant, nanoparticles can modify the lubricant viscosity and potentially also the flow behaviour [256]. Under boundary lubrication conditions, when surfaces come into direct contact, several mechanisms have been proposed for friction and wear by nanoparticles including [257] sliding (a), rolling (b), exfoliation (c), polishing (d), and mending (e), as shown in Fig. 6.

If nanoparticles are able to maintain the separation of surface asperities under boundary conditions they could give large reductions in friction and wear [257]. This requires nanoparticles to be sufficiently hard to not be plastically deformed under high contact pressures. However, if the nanoparticles are excessively hard, they may indent into the surfaces to such an extent that they no longer separate the asperities. Nanoparticles that are able to prevent asperity contact may roll (Fig. 6(a)) or slide (Fig. 6(b)) between the surfaces in order to reduce friction. Experimental studies have suggested that unlike macroscopic systems, there is little (if any) benefit of the rolling action for nanoparticles compared to sliding [258, 259]. Hard nanoparticles may also reduce friction through polishing (Fig. 6(d)), whereby asperities are smoothed by abrasion by the nanoparticles to reduce roughness and shift the system from boundary, towards mixed lubrication conditions [260]. While this may initially provide a large benefit, it will probably diminish in the presence of other running-in and smoothing processes [261]. It has also been suggested that hard nanoparticles may also “mend” the surface by filling the gaps between asperities on rough surfaces (Fig. 6(e)) [260]. For soft or layered nanoparticles, there is also the possibility that friction may be reduced through exfoliation (Fig. 6(c)), whereby the nanoparticles break down into a solid lubricating film, which prevents contact between the surfaces and reduces friction by forming a low shear strength layer [262].

The prevalence of each of these mechanisms is expected to be highly situation-specific [219], and depend on variables including the material, size, shape, hardness, and coverage of the nanoparticle as well as the temperature, pressure, and sliding velocity [257]. For example, in-situ high resolution scanning/transmission electron microscope (HRSEM/HRTEM) [262, 263] images of inorganic fullerene (MoS2) nanoparticles during sliding have suggested that their friction reduction mechanism consists of a mixture of sliding/rolling at low pressure and exfoliation to lamellar sheets of MoS2 at high pressure. Exfoliation of hollow core MoS2 fullerenes was also observed in MD simulations by Lahouij et al. [263] under high uniaxial loads.

The simulation of nanoparticles dispersed in a lubricant and confined between solid surfaces is a complex problem for study by MD simulations. The nanoparticles tested in experiments (generally > 10 nm diameter) are relatively large for atomistic MD, and thus only a few small nanoparticles are usually

---

**Fig. 6** Schematic showing the potential mechanisms of friction reduction by the presence of nanoparticle additives pressed between sliding metal surfaces, adapted from Ref. [257]. (a) Sliding, (b) rolling, (c) exfoliation, (d) polishing, and (e) mending.
included in the system. In addition, although accurate force-field parameters are available for each component in isolation, they are generally not transferable between the different components. Moreover, the addition of a lubricant between atomically-smooth surfaces with PBCs in the lateral directions means that the lubricant supports the load. This means that only the viscosity/flow modification of the nanoparticles, rather than their boundary friction reduction mechanisms (Fig. 6) can be probed [257]. Therefore, most NEMD studies have studied nanoparticles confined between solid surfaces in the absence of a lubricant, even though they are designed as additives (Section 3.2.1), although some studies have also been performed with nanoparticles dispersed in a lubricant (Section 3.2.2).

3.2.1 Nanoparticles as dry lubricants

There have been a considerable number of NEMD simulations to investigate nanoparticle behaviour in confined systems in the absence of a liquid lubricant. For example, Joly-Pottuz et al. [264] conducted friction and wear experiments and NEMD simulations of carbon nano-onions (2 nm diameter) between sliding diamond-like carbon (DLC) surfaces at high contact pressure (1–5 GPa). Both the experiments and simulations suggest that the carbon nano-onions remained intact under compression and sliding and did not exfoliate to graphitic planes [264], unlike MoS2 fullerenes [263]. Further simulations by Bucholz et al. [265] indicated that the ability of the nano-onions to roll is inhibited both by increased contact pressure and the presence of a diamond core within the nanoparticles. The transition from rolling to sliding behaviour was accompanied by a significant increase in the friction coefficient due to the formation of interfacial bonds. Bucholz et al. [266] also investigated the behaviour of amorphous carbon nanoparticles (2–5 nm) between hydrogen-terminated diamond surfaces with NEMD. The simulations showed that during sliding, hydrogen passivated the unsaturated carbon atoms near the surface of the nanoparticles, which prevented interfacial bond formation and allowed them to roll within the interface. It was demonstrated that amorphous carbon nanoparticles will only provide good tribological performance when sufficient chemical passivation of the surfaces and nanoparticles is maintained.

Hu et al. [267] investigated the effect of soft metallic copper nanoparticles between relatively hard iron surfaces using NEMD simulations. The results suggested that the nanoparticles deformed to generate a low shear strength film between the contact surfaces, which accommodated the velocity gradient through plastic deformation. The nanoparticles were most effective in reducing friction at low loads and sliding velocities (0.01 GPa, 10 m·s⁻¹), but still provided a small benefit under more extreme conditions (0.5 GPa, 500 m·s⁻¹). Hu et al. [268] also used NEMD simulations to investigate the tribology behaviour of hard diamond and silicon dioxide (SiO2) nanoparticles. At low velocity and low load (10 m·s⁻¹, 0.5 GPa), the nanoparticles separated the two contact surfaces from each other and significantly reduced friction. At high velocity and high load (500 m·s⁻¹, 1.0 GPa), they either plastically deformed (SiO2) or indented into the contact surfaces (diamond), leading to smaller reductions in friction.

Eder et al. [269] simulated the abrasion process of an atomically rough iron surface with multiple hard, abrasive particles. By monitoring the nanoscopic wear depth over time, the authors showed that the Barwell macroscopic wear law [270] holds even at the atomic scale. They found that in this multiasperity contact system, the Bowden-Tabor term [271], which describes the friction force as a function of the real nanoscopic contact area, can predict the friction even when wear is involved.

Ewen et al. [257] used NEMD simulations to examine the friction and wear reducing mechanisms of carbon nanoparticles. Specifically, the friction and wear behaviour of carbon nanodiamonds (CNDs) and carbon nano-onions (CNOs) with a diameter of approximately 3 nm, confined between iron slabs was probed at a range of coverages, pressures, and sliding velocities. At high coverage and low pressure, the nanoparticles did not indent into the iron slabs during sliding, leading to zero wear and very low friction. At high coverage and low pressure, the nanoparticles did not indent into the iron slabs during sliding, leading to zero wear and very low friction. At high coverage and low pressure, the nanoparticles did not indent into the iron slabs during sliding, leading to zero wear and very low friction.
expression developed for macroscopic indentation by Bowden and Tabor [271]. Even at the highest
pressure and lowest coverage simulated, both types of nanoparticle were able to maintain separation of the
opposing slabs and reduced friction by approximately 75% compared to when no nanoparticles were present,
in agreement with experimental observations [272]. CNO nanoparticles showed a lower indentation (wear)
depth and lower friction coefficients than CND nanoparticles at the same coverage and pressure, making
them more attractive friction modifier additives.

3.2.2 Nanoparticles as lubricant additives

Although NEMD simulations of nanoparticles can provide interesting insights into their friction and
wear reduction under boundary lubrication conditions, it is clearly also important to study them in the
presence of a liquid lubricant. NEMD simulations have been performed on atomically-smooth surfaces
to study their effect on friction [273] and flow [256] under hydrodynamic lubrication conditions. Com-
pressive LVMD simulations have also been used to study the load carrying capacity of copper nano-
particles in an n-octane lubricant [274].

Recently, NEMD simulations have been used to complement macroscopic wear experiments to study the
effect of titanium dioxide nanoparticle lubricant additives in a PAO lubricant [275]. Results from both
the experiments and simulations suggested that there was an optimal particle size that will minimize friction
and wear for a given surface roughness [275]. The results support the “mending” mechanism for the
titanium dioxide nanoparticle additives (Fig. 6(e)). In this context, particles that are smaller than the
characteristic roughness of the surfaces are most likely to perform this function. Given the considerable
interest in nanoparticle additives [219], we expect to see further NEMD studies of nanoparticles in the
presence of a lubricant on rough surfaces.

3.3 Summary

NEMD simulations of a range of friction modifier additives have given atomic-scale insights into the
films they form on solid surfaces and the way in which they alter the flow and friction behaviour under a
wide range of conditions. Due to the long timescales involved, MD simulation of film formation from
solution remains a challenge. However, by assuming film formation on the surface, they have reproduced
important experimental observations, and helped to explain these from atomic-scale behaviour. They have
highlighted important differences between various additive molecular structures and coverages. To more
faithfully reproduce boundary lubrication conditions, NEMD simulations of additives between realistic
rough surfaces or using GCMD, which both allow squeeze out from high pressure regions, are expected
to provide interesting results. An important aspect of lubricant additive behaviour which cannot be studied
using conventional MD simulations is chemical reactivity. However, newly developed methods have
begun to allow such investigations using molecular simulations and these are discussed in Section 4.1.

4 Beyond classical NEMD: Current trends and future perspectives

NEMD simulations have made remarkable advances in the past few decades both in terms of methodology
development and application-driven research. Despite this, some complex tribology problems (e.g., tribo-
chemistry, triboemission, adhesion, plasticity) are influenced by physical/chemical processes that occur at
time and length scales which are not accessible through classical NEMD simulations. Our understanding of
such problems would significantly benefit from models that span multiple spatial and/or temporal scales
with different levels of complexity. Such multiscale/multiphysics models use information from several
levels (that investigate a phenomenon over a specific temporal and spatial window) to calculate material
properties or model system behaviour. The following main levels are usually distinguished (Fig. 7): quantum
mechanical models (electronic scale), MD models (atomic scale), and continuum models (macroscopic
scale).

Figure 7 provides a map to illustrate the different modelling techniques available to describe tribological
phenomena at different scales. The focus of this review is on molecular simulations, and classical MD
is the centrepiece of the map. Techniques which include electronic effects, e.g., density functional theory (DFT),
Fig. 7 Different scales for simulations of tribological systems. The map highlights areas of current interest for the development of coupling strategies to link MD simulations to quantum mechanical and continuum descriptions, with examples portrayed across the scales.

can be used to study processes at smaller scales, while continuum methods, such as finite element methods (FEM) and computational fluid dynamics (CFD), can describe processes at larger scales. A full review of all these methodologies is certainly outside the focus of this review; however, here we will highlight successful multiscale/multiphysics approaches that already exist, and suggest how the furthering of such methodologies could play a fundamental role in addressing unsolved issues in the field of tribology. One area where multiscale/multiphysics modelling has been successfully applied is integrated computational materials engineering (ICME), which allows the prediction of material properties or system behaviour based on knowledge of the process-structure-property relationships [276]. These same attributes make these models interesting in tribology, particularly in the design of synergetic surfaces and lubricants [277]. Breakthroughs are thus expected from the future development of versatile and efficient multiscale/multiphysics tools designed specifically for tribology problems.

Critical to these advances will be to link MD with simulations performed at smaller and larger scales (see Fig. 7). In doing so, modelling strategies can be designed to minimise the computational resources required while maximising the insight provided by simulations performed either simultaneously or sequentially at different scales. In the next two subsections, an outlook is provided which describes the most promising multiscale/multiphysics techniques currently available to link MD simulations to quantum mechanics (Section 4.1) and continuum models (4.2) for tribology problems.

4.1 Linking MD to smaller scales

The function of many lubricant additives depends on chemical reactivity both in solution and at solid surfaces. Due to their large computational expense, conventional quantum mechanical calculations are static and include only a few molecules. Modelling chemical reactivity is beyond the capabilities of classical MD simulations; however, techniques which allow this to be studied have been developed and applied to study lubricant additives. One promising technique is the use of “reactive” force-fields, such as ReaxFF [278] which can approximate chemical reactivity in classical MD simulations through so-called bond order terms [279]. Liang et al. [280] gave a general overview of the different types of reactive force-field available in 2013. A major drawback of all reactive force-fields is the limited availability of reliable parameter sets to study the materials, conditions, and properties of interest. The generation of high-quality parameter sets is a very difficult and time consuming task due to the huge number of parameters which need to be fitted [281]. Despite this, ReaxFF parameters have been developed which can be used to investigate some important lubricant and additive molecules, such as alkanes and alcohols [282, 283], phosphoric acids [284], organo-sulphur compounds [285], and molybdenum disulphide [286]. Using such force-fields, relatively large molecular systems (nm) can be compressed and sheared [284] in a similar manner to conventional confined NEMD simulations (Fig. 1(a)), and the consequent chemical reactions monitored over reasonable simulations time (ns).

Alternatively, the reactivity of lubricant additives can be explicitly modelled using first principles or \textit{ab initio} MD techniques. Examples of this include Car-Parrinello MD (CP-MD) [287–291] and tight-binding
quantum chemical MD (TB-QCMD) [214, 292–295], which have given important insights into the reactivity of common friction modifier (molybdenum disulphide, molybdenum dithiocarbamate) and anti-wear (zinc dialkyldithiophosphate) additives. *Ab initio* NEMD simulations have also been performed on additive systems subjected to confinement and shear. For example, TB-QCMD has been used to investigate OFM adsorption on iron oxide surfaces under sliding conditions [296–298]. The main advantage of *ab initio* MD simulations over MD with reactive force-fields is that they do not require extensive force-field parameterisation in order to ensure accurate results. However, they increase the computational expense by more than an order of magnitude compared to classical MD simulations with reactive force-fields, meaning that the system sizes and timescales accessible are more limited [299]. One potential solution for confined NEMD simulations, where the “reactive” region is well-defined, is spatial coupling of classical MD for the “non-reactive” region to *ab initio* MD for the “reactive” region [300]. This technique is commonly known as quantum mechanics/molecular mechanics (QM/MM), and is widely used in materials science [301, 302] and the natural sciences [303].

### 4.2 Linking MD to larger scales

Linking MD to larger scales is also of considerable interest to tribology researchers. As shown in Fig. 7, the next scale up from atomistic MD is coarse-grained MD (CG-MD), which involves grouping of neighbouring atoms together to form “beads” in order to reduce the number of interaction sites. It is commonly used to simulate large polymers, where monomers or even a combination of monomers can be represented as a single CG bead. CG-MD has not been extensively applied in tribology; however, it has been used to study friction in polymer brushes [304, 305], model friction modifier additives [306] and ionic liquids [68]. The use of CG-MD is likely to remain rare in tribology, since the importance of atomistic detail has been highlighted in accurately reproducing the viscous behaviour of lubricants and additives [10]. Possible applications of CG-MD in tribology are to study the behaviour of large polymer additives [307] or relatively slow processes such as surfactant self-assembly at solid surfaces, whose timescales are inaccessible to study using all-atom force-fields [224].

Most tribological processes are governed by physical phenomena localised in an interfacial layer which controls the macroscopic behaviour of the system. With this in mind, an attractive possibility is to capture the interfacial laws of friction, heat transfer, and other relevant phenomena at small, accurate scales and pass this information to continuum models to study macroscopic tribological systems. Choosing how to pass information across the scales is not always straightforward; some issues concern what assumptions can be made and the accuracy required in linking the two levels. Two options are usually available, depending whether it is possible to establish scale-separation between the atomic level and the continuum. If such separation exists, then a hierarchical (or sequential) multiscale model is acceptable, otherwise a hybrid (or concurrent) scheme is required. Both of these MD-continuum coupling methods have already been applied to a variety of problems in tribology. Hierarchical coupling involves using MD simulations to gain information about the boundary conditions and constitutive laws to input into separate continuum models [308]. This method is relatively straightforward to implement, but requires careful planning to establish which regions of phase space can be reliably probed with MD to provide useful inputs for the continuum models. For example, Martini et al. [160] and Savio et al. [168, 182] respectively used local viscosity and slip information obtained using MD simulations to improve continuum predictions of film thickness and friction for strongly confined systems. These are examples of phenomena that can be studied at a single relevant scale and integrated at a larger scale in a hierarchical manner.

Alternatively, the MD and continuum models can be run concurrently on different spatial regions; this is the preferred method when atomic interactions dictate the observed macroscopic tribological behaviour but are themselves affected by the whole macroscopic domain. Modelling surfaces with realistic levels of roughness lubricated by a fluid is a case in point. Such a problem is still beyond the scales accessible to full MD simulation [269]. Alternatively, atomistic details of the surface topography and fluid close to
the walls can be explicitly modelled and intimately linked to the macroscopic response of the fluid further from the walls. In this case MD-continuum coupling strategies involve the transfer of information between the MD and continuum regions, and particular care must be taken when the two descriptions merge [309, 310]. A number of schemes exist for this [311–314], with various procedures enabling hybrid coupling [315].

Models dealing with concurrent coupling are generally far more complex than their hierarchical alternatives, but are of high importance in understanding the physics of certain phenomena in contact interfaces, such as the effect of surface roughness, surface coatings and additive layers on fluid flow. One major limitation of concurrent coupling techniques is that the simulation time step is limited to the shortest timescale which needs to be modelled at the smallest scale. Despite this, considerable performance enhancements can be achieved compared to modelling the entire system at the atomic level [313]. The future success of concurrent coupling schemes will depend on the development of efficient coupling libraries and algorithms that facilitate hybrid coupling simulations and support their implementation on different HPC platforms (see, e.g., a recent software made available by the Imperial College Tribology Group [316]).

4.3 Summary

For the study of reactive lubricant additives, conventional MD simulations are deficient in their inability to model chemical reactivity. Moreover, for computational reasons, conventional MD simulations are limited to the study of relatively small systems and timescales which continues to limit their ability to adequately model situations of specific engineering significance. Moving forwards, we expect classical MD simulations to be supplemented by coupling techniques to smaller scales, to study chemical reactivity, as well as larger scales, to extend the accessible time and length scales. Tribology is an ideal subject for such multiscale/multiphysics models since many macroscopic tribological phenomena have been shown to driven by atomic-scale processes, often localised at an interface. Such tools will shed further light on these processes and could even transform lubricant formulation by identifying process-structure-property relationships as it has in other fields of research.

5 Conclusions

Molecular dynamics is, at its heart, a relatively simple simulation method, yet it has become one of the most powerful tools in science and engineering to understand the atomic-scale behaviour of fluids. Historically, it has played a central role in corroborating theories of the liquid state, but it has now become capable of directly evaluating physical properties in industrially-important systems. Moving forwards, we expect NEMD simulations to play an increasingly important role in building a fundamental or “bottom-up” understanding of tribological phenomena from the atomic scale. Moreover, it is reaching the point at which it can be a truly useful tool for lubricant design. Key to this process will be direct comparison of NEMD to experiment, something which is beginning to become possible thanks to improvements in simulation algorithms, classical force-fields, and computer hardware, as well as experimental advances.

Many bulk NEMD studies have predicted the Newtonian viscosity of realistic lubricant molecules, as well as its temperature and pressure dependence. The accuracy of these simulations has progressed to a point where they can be used as a predictive tool to compare different lubricant molecular structures. Such structure-property relationships can be extremely useful in lubricant design. Bulk NEMD simulations have also given unique insights into lubricant shear thinning behaviour under extreme conditions and they are expected to play a key role in helping to improve macroscopic shear thinning models.

Confined NEMD simulations have revolutionised our fundamental understanding of the behaviour of very thin lubricant films between solid surfaces. This includes the density and viscosity inhomogeneities in confined films, as well as important tribological phenomena such as stick-slip and boundary slip. It is also being increasingly employed to study shear localisation behaviour in thicker films subjected to high pressures. There seem to be parallels between the behaviour of lubricants under extreme confinement and at thicker films at high pressure which certainly
The atomic-scale behaviour of lubricant additives has also been successfully investigated with NEMD. There has been a particular focus on organic friction modifier and nanoparticle additives, in part because simulation of these additives does not necessarily need to account for chemical reactivity. These simulations have highlighted important differences between various additive molecular structures and coverages, which have direct relevance to lubricant formulation. The tribochemical behaviour of friction modifier and anti-wear additives under pressure and shear has also been given by NEMD simulations using reactive force-fields as well as \textit{ab initio} MD techniques. Even though accessible system sizes and timescales of these simulations are more limited than for conventional MD, fundamental behaviour of reactive lubricant additives can be explored, potentially enabling more effective additives to be designed.

Finally, it should be noted that, while molecular dynamics can now simulate systems of millions or even billions of individual atoms, MD is unlikely ever to be suitable for explicitly modelling macro-scale engineering components ($\approx 10^{24}$ atoms). However, by coupling MD to continuum models, MD can be used just to capture key processes which drive the overall behaviour of macroscopic tribological systems.

Acknowledgements

J. P. E thanks the Engineering and Physical Sciences Research Council (EPSRC) for financial support through a Doctoral Prize Fellowship. D. D. acknowledges the EPSRC for an Established Career Fellowship EP/N025954/1. The authors also thank Hugh A. Spikes and B. D. Todd for helpful discussions.

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


[287] Mosey N J, Woo T K. Finite temperature structure and


---

**James P. EWEN.** He is an EPSRC Doctoral Prize Fellow in the Tribology Group at Imperial College London. He gained his master degree in chemistry from the University of Bath in 2014 and his PhD degree from Imperial College in 2017 with a Thesis entitled “Molecular dynamics simulations of lubricants and additives”. His research involves utilising molecular simulations to investigate the structure, flow, and friction behaviour of a range of tribological systems.

---

**David. M HEYES.** He is a Principal Research Fellow in the Tribology Group at Imperial College London. He received his PhD degree from the University of Manchester in 1977. He has held postdoctoral research positions in the Vitreous State Laboratory, Department of Physics, Catholic University of America, Washington DC, USA, Department of Physical Chemistry at the University of Amsterdam, and at the Department of Chemistry, Royal Holloway, University of London. He was a Royal Society (London) University Research Fellow between 1983 and 1992, first at Royal Holloway. His research interests are in the development of computer-based theoretical and modelling techniques applied to condensed matter liquid systems, principally in the areas of phase equilibria, and structural and transport properties of bulk and confined systems.

---

**Daniele DINI.** He is Head of the Tribology Group at Imperial College London. Prior to joining Imperial College in 2006, Professor Dini received the DPhil degree from the Department of Engineering Science at the University of Oxford in 2004. He now leads the advanced modelling research within the Tribology Group and collaborates closely with its experimentalists. His current individual research portfolio supports a large team of researchers focused on studies related to the modelling of tribological systems and materials. Most of these projects are multidisciplinary and range from atomic and molecular simulation of lubricant, additives and surfaces to the modelling of machine or biomechanical components. His group performs fundamental research, while successfully supporting the application of tribology in industry; the strong links with industrial partners have been recently recognised by the Imperial College President’s Award and Medal for Excellence in External Collaboration and Partnerships (2017). Professor Dini has also been the recipient of a number of other individual awards, he has also been elected a Fellow of the IMechE in 2014, and is the recipient of the prestigious EPSRC Established Career Fellowship, awarded in 2016.
Effect of arc-current and spray distance on elastic modulus and fracture toughness of plasma-sprayed chromium oxide coatings

Simanchal KAR*, Partha Pratim BANDYOPADHYAY, Soumitra PAUL
Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, West Bengal 721302, India

Received: 23 February 2017 / Revised: 13 April 2017 / Accepted: 05 May 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: In this experimental study, chromium oxide powder was sprayed on a low-carbon steel substrate using the atmospheric plasma spray process. The current and standoff distances (SODs) were varied to study their effect on the fracture toughness of the coatings. Theoretically, as the arc current increases, the melting of the ceramic oxide should increase and this in turn should lead to the formation of a dense coating. However, it was observed that if the arc power is too high and because the particle size of the powder is small (approximately 30 μm), the particles tend to fly away from the plasma core. Similarly, an appropriate SOD should provide the particles with more melting time, thus resulting in a dense coating. On the other hand, a larger SOD leads to the solidification of the molten particles before the droplets can reach the substrate. All these effects may lead to substantial variation in the fracture toughness of the coating. The present paper attempts to correlate the plasma spraying parameters and microstructure of the coating with the fracture toughness and other primary coating properties.

Keywords: chromia coating; fracture toughness; plasma spray; critical plasma spray parameter; critical depth

1 Introduction

Surface engineering includes the study of wear, erosion, corrosion processes on the component surface, and the various methodologies to prevent them. With the help of thermal and thermo-chemical processes, surfaces can be modified to become wear and corrosion resistant while retaining the toughness and ductility of the bulk component [1]. One such example of surface engineering is a plasma sprayed chromia coating. It is extensively used as a corrosion and wear resistance coating [2, 3]. Chromia coatings also find applications in pumps and printing rolls [4, 5]. In general, an effective wear resistance coating is expected to possess high fracture toughness along with high hardness, good adhesion, low porosity, homogeneous structure, and uniform thickness [6]. The fracture toughness ($K_{IC}$), is a measure of a material’s resistance to brittle fracture when a crack is present [7, 8]. For ceramic components, Evan and Charles (1976) have formulated a mathematical expression to evaluate the fracture toughness using the fracture length and Vickers indentation diagonal length as shown [9]:

$$K_{IC} = 0.15k \left( \frac{c}{a} \right)^{3/2} \frac{H \sqrt{a}}{\phi}$$

where $k = \text{correction factor for ceramics} = 3.2$, $\phi = \text{constraint factor} \approx 3$, $H = \text{Vickers hardness (MPa)}$, $a = \text{half the Vickers diagonal length measured in the direction of the crack}$, and $c = \text{sum of the length of the crack and half the Vickers diagonal length in the direction of the crack}$. Figure 1 shows the schematic of the Vickers indentation.

* Corresponding author: Simanchal KAR, E-mail: simanchal.kar@gmail.com
It is well established that there exists a critical depth of penetration on a ceramic surface, beyond which the material removal changes from the ductile regime to brittle fracture. This critical depth required for the ductile to brittle transition is given as follows [10]:

\[
d_c = \beta \left( \frac{E}{H} \right) \left( \frac{K_{IC}}{H} \right)^2
\]

where \( \beta \approx 0.15 \) for ceramics, assuming that the depth of penetration is equal to the machine infeed, \( E \) is the elastic modulus of the coating (in GPa), \( H \) is the hardness (in GPa), and \( K_{IC} \) is the fracture toughness of the chromia coating (in MPa·m\(^{1/2}\)).

Further, in ceramic materials, the critical loads for radial and transverse cracks in ceramics and ceramic coatings (Fig. 2) can be obtained from the following expressions [11, 12]:

(a) Critical load for radial crack initiation:

\[
P_{R^*} = 54.5 \left( \frac{\alpha}{\eta \gamma^a} \right) \left( \frac{K_{IC}^4}{H^3} \right)
\]

(b) Critical load for lateral crack initiation:

\[
P_{L^*} = \xi \left( \frac{K_{IC}^4}{H^3} \right) f \left( \frac{E}{H} \right)
\]

where \( \alpha, \eta, \) and \( \gamma \) are constant, and for the Vickers indenter, \( \xi = 2/\pi, \eta = 1, \) and \( \gamma = 0.2. \) \( \xi \) is a dimensionless constant, \( f(E/H) \) is an attenuation function equivalent to \( 2 \times 10^6 \) [12]. Figure 2 shows the formation of radial and transverse cracks during indentation.

The present work was undertaken to study the effects of the parameters of the spray process, i.e., the variation in stand-off distance (SOD) and critical plasma spray parameter (CPSP), on the coating properties—especially the fracture toughness. The CPSP value is given as follows:

\[
\text{CPSP} = \frac{\text{Input power}}{\text{Primary gas flow rate (slpm)}}
\]

It is expected that the ensemble of results obtained using the same set of equipment and instruments will be useful to thermal spray industries.
2 Experimental procedure

The chromia coatings were sprayed on mild steel samples using the METCO 9MB gun and 9MC spray system. At first, the steel substrates were ground and ultrasonically cleaned. Grit blasting was performed on one of the surfaces prior to the coating. This was followed by the deposition of the bond coat. Nickel–aluminum was used for this purpose. Finally, the top coat was deposited over the bond coat. Table 1 lists the spray parameters used for the preparation of the coating. The SOD and CPSP values were varied uniformly to study their effect on the coating properties.

The coating hardness was measured using a hardness tester (LECO LM700, USA) and an average of ten hardness readings was considered for each coating. The porosity of the top coat was measured from the images of the cross sections using AxioVision image analysis software. The porosity calculation is performed using the principle of gray value analysis [14]. The surface roughness of the coating was measured using a contact-type surface roughness tester (Taylor-Hobson Surtronic 3+, UK). The sampling length for the roughness evaluation was set to 4 mm. For the study of the coating cross sectional, a scanning electron microscope (SEM) (Zeiss EVO 18, Germany) and an optical microscope (Zeiss Axio Vert.A1, Germany) were used. Indentations were also made on the sample cross sections under a load of 1.962 N (≈ 200 gf) using a micro-hardness tester (MVH-S-AUTO, Omnitech, India). The purpose of these indentations was to generate cracks for the measurement of the indentation fracture toughness. The cracks and diagonal lengths were measured using Axio Vision (Zeiss, Germany) and Image-J (NIH, USA) software. Depth sensing indentations were performed using an instrumented hardness tester (Micro-combi tester, CSM instruments, Switzerland) equipped with a Vickers indenter. The indentation measurements were undertaken under a maximum load of 2,000 mN, loading and unloading rate of 4,000 mN/min, and with a dwell time of 15 s [15]. In order to evaluate the coating strength, scratch tests were conducted on the surface of the ceramic-coated samples using a Ducom scratch tester TR-101-M5 (Ducom, USA). The scratch tests were performed with an initial load of 10 N that gradually increased till it reached 80 N to obtain a scratch length of 7 mm. The scratch velocity was fixed at 0.1 mm/s.

3 Results and discussion

Figure 3 shows the SEM images of the cross section of the chromia coatings. It reveals that for a low CPSP, the porosity is high and the partially melted region (PMR) is small; at a recommended value of the CPSP, these defects are diminished; and at higher values of the CPSP, the PMR grows while the porosity decreases. The melting point of chromia is approximately 2,435 °C. At a high value of the CPSP, the plasma is aggressive in nature. A fraction of ceramic particles, chiefly the particles with the smallest diameter, was unable to penetrate the plasma and pass through the neighboring trajectory. These particles absorb a portion of the plasma heat and form a partially melted region, as shown in Fig. 3. Partially melted regions appear when the particle dwell time in the plasma is insufficient for a group of ceramic particles to melt. Similarly, at smaller SODs, the porosity is low but the PMR is high, and as the SOD increases, the PMR reduces but the porosity increases. This is attributed to the coarse and wide distribution of the chromia particles, which contributes to non-uniform melting during the plasma spraying process [16].

The magnitudes of the coating defects for the corresponding variations are given in Table 1. The elastic moduli of the plasma sprayed coatings that were sprayed at the recommended value of the CPSP or SOD is the highest as compared to maximum and minimum values of parameters taken into consideration (refer Fig. 2). This can be attributed to the higher melting proportion and less coating defects as compared to coatings obtained from other parametric combination of CPSP and SOD (Table 1). However, the
measured fracture toughness (using Eq. (1)) showed an increasing trend with the increase in the CPSP and a slight variation with a change in the SOD. This is attributed to improved melting and the densification of the coating. The hardness of the coating varies proportionally with the increase in CPSP. However, at maximum SOD, the hardness decreased and this is attributed to the increase in porosity content of the coating. Figure 4 shows the variation in the mechanical properties of the sprayed chromia coating with respect to the change in the CPSP and SOD.

The coating with the recommended value of the CPSP and SOD exhibited the best results while the coating with the lowest value of the CPSP and lowest SOD had a poor structure. Table 2 lists the percentage content of the porosity and partially melted region present in the chromia coating.

Coatings sprayed at the lowest value of the CPSP and the highest SOD had the maximum porosity, while coatings sprayed at the maximum value of the CPSP and lowest SOD had the highest percentage of partially melted region. The increase in the porosity is attributed to unmelted particles becoming embedded in the coating and the shadowing effect of the splats. Moreover, at the maximum value of the CPSP, a portion of the chromia powder of smaller particle size could not penetrate the plasma plume and flew close to the plasma, thus absorbing a part of the plasma heat. This led to the formation of a partially melted region. Moreover, at the minimum SOD, the particle dwell time was insufficient for converting the particle to the liquid state before it struck the substrate. This led to an increase in the partially melted region.

Figure 5 shows the optical images of the indentation marks on the coating surface. Owing to the heterogeneous properties of the coating due to the porosity and the presence of partially melted regions, the indentation caused the crumbling of subsequent layers from the surface. More of this crumbling was observed in the case of the coating with the lowest value of the CPSP.

As shown in Fig. 6, the X-ray diffraction of the coating indicated the stress-induced crystalline structure of the chromia at a high value of the CPSP and a low SOD. A small amount of the amorphous phases are

<table>
<thead>
<tr>
<th>Coating</th>
<th>Porosity (Area %)</th>
<th>PMR (Area %)</th>
<th>E (GPa)</th>
<th>HV200</th>
<th>Kn (MPa·m^-1/2)</th>
<th>dc (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>4.73</td>
<td>&lt; 1</td>
<td>178.14</td>
<td>957.1</td>
<td>1.74</td>
<td>0.0978</td>
</tr>
<tr>
<td>C2/S2*</td>
<td>2.38</td>
<td>2.34</td>
<td>214.97</td>
<td>1179</td>
<td>2.9</td>
<td>0.1755</td>
</tr>
<tr>
<td>C3</td>
<td>&lt; 1</td>
<td>5.70</td>
<td>209.81</td>
<td>1259</td>
<td>2.98</td>
<td>0.1484</td>
</tr>
<tr>
<td>S1</td>
<td>3.45</td>
<td>2.4</td>
<td>155</td>
<td>871</td>
<td>2.27</td>
<td>0.1922</td>
</tr>
<tr>
<td>S3</td>
<td>5.19</td>
<td>1.76</td>
<td>195.42</td>
<td>1045</td>
<td>2.59</td>
<td>0.1826</td>
</tr>
</tbody>
</table>

* C2=S2 (i.e., C2 and S2 parameters are identical)
Fig. 5 Optical image of Vickers indentation taken at cross section of the chromia coatings.

Fig. 6 X-ray diffraction of (a) chromia powder and (b) chromia coating at five parametric variations.

also observed in coatings at high value of the CPSP and low SOD. In contrast, a low value of the CPSP leads to a partial transformation of the phases. This is attributed to the inability of the chromia particles to penetrate the plasma plume at a high value of the CPSP.

Similarly, at a low SOD, the melting time of the particles is reduced and partially melted regions come into existence. At a low value of the CPSP, the available plasma heat is insufficient for the complete melting of the chromia particles. Consequently, both partially melted and unmelted particles become embedded with the molten splats, thus forming a heterogeneous coating structure.

In chromia coatings, the radial and lateral cracks are likely to occur at loads that can be obtained using Eqs. (3) and (4), as previously explained. In ceramic coatings, abrasion or erosion wear occur primarily owing to brittle fracture. When the cracks nucleate and converge, material is removed from the surface and chipping occurs. Figure 7 shows the variation in the critical load for the crack formation in chromia coatings for various coating parameters. It is observed that the load-carrying capacity without crack formation is highest in the case of the chromia coating sprayed at the recommended CPSP and SOD. An increase or decrease in CPSP/SOD resulted in low resistance to crack formation. Variations in the SOD produced a slight change in the critical load while variations in the CPSP caused a noticeable deviation. This is attributed to the difference in the coating fracture toughness and hardness. The coating hardness and toughness values are the highest in the case of the chromia coating sprayed at the recommended parameters.

Fig. 7 Variation of critical load with respect to parameters of chromia coating.
In order to study the coating performance when subjected to gradual loading, scratch tests were performed on the coating surface. The nature of the scratch surface largely depends upon the critical depth, as explained earlier. The corresponding critical depth for each coating is given in Table 2. Figure 8 shows the load versus displacement curve obtained from the scratch test. It is observed that, except for the recommended parameters, the chromia coatings completely failed even for small loads. The porosity was the major contributor to the erratic failures of the coating. The scratch debris consisted of fragmented chips, thus indicating that the failure occurred primarily owing to micro-brittle fracture. Initially, at a low load, the scratch indenter depth is less than the critical depth of the coating. This resulted in some degree of plastic deformation followed by failure due to spalling. This effect was reported by Zhang et al. [17] in their work on plasma sprayed Al₂O₃ coatings. However, at higher loads, the ploughed surface showed signs of brittle fracture along with fragmented chips. Figures 9(a)–9(b) shows the gradual transition of the ductile mode of the material removal to brittle fracture as the load of the scratch indenter gradually increased.

### 4 Conclusions

The chromia powder can be successfully coated using an air plasma spray system. For spraying, the optimized spray parameters are necessary to obtain the desired mechanical properties. On increasing the arc current, higher values of the CPSP lead to an increase in the partially melted regions while low values of the CPSP lead to high porosity. Moreover, a large SOD leads to early solidification of the molten ceramic droplets before they reach the substrate. This results in an increase in the porosity of the coating. The mechanical properties such as the hardness, fracture toughness, and Young’s modulus largely depend on the coating density. The increase in the porosity reduces the effective hardness and coating toughness. However, the occurrence of the partially melted region results in a slight change in the coating toughness and hardness. The critical load required for crack initiation was determined for all the coatings, and it was observed that this load largely depends on the CPSP rather than the SOD. Finally, the indentation scratch tests performed on the coating specimens revealed that the failure of the plasma sprayed coating

![Fig. 8 Load versus displacement curve of coating performance during scratch test.](image)
that was sprayed at the recommended CPSP, occurred at a higher load. Further, any variation in CPSP or SOD contributed to weak adhesion between chromia coating and the metallic substrate.

Acknowledgments

The authors are extremely grateful to Prof. Mihir Sarangi, IIT Kharagpur for providing his research facilities. Authors also acknowledge the MHRD, Govt. of India for the financial support.

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


Simanchal KAR. He is a Ph.D student in the Department of Mechanical Engineering at Indian Institute of Technology Kharagpur. He received his B.Tech and M.Tech degrees in mechanical Engineering in 2006 and 2011 from KIIT University, Bhubaneswar, India.

His research interests include abrasive machining process, thermal spray processes, and material characterization.

Partha Pratim BANDYOPADHYAY. He received his M.Tech and Ph.D degrees in mechanical engineering from IIT Kharagpur, India, in 1992 and 2000, respectively. He is currently attached to IIT Kharagpur as a professor of mechanical engineering.

His research interests include CNT or diamond particulate reinforced coatings, micro and nano finishing of thermally sprayed coatings, laser treatment of coatings, etc.

Soumitra PAUL. He received his M.Tech and Ph.D degrees in mechanical engineering from IIT Kharagpur, India, in 1989 and 1994, respectively. He is attached to IIT Kharagpur as a professor of mechanical engineering. His research interests include machining, grinding, cutting tool coating, thick coating, finishability of thick coatings, high speed grinding of ceramics and carbon fibre ceramic matrix composites, high speed grinding of aerospace materials, minimum quantity lubrication in machining and grinding, micro machining, etc.
Friction characteristics of mechanically exfoliated and CVD-grown single-layer MoS₂

Dinh Le Cao KY¹, Bien-Cuong TRAN KHAC¹, Chinh Tam LE², Yong Soo KIM¹, Koo-Hyun CHUNG¹,*,

¹ School of Mechanical Engineering, University of Ulsan, Ulsan 44610, Republic of Korea
² Department of Physics and Energy Harvest Storage Research Center, University of Ulsan, Ulsan 44610, Republic of Korea

Received: 17 April 2017 / Revised: 06 June 2017 / Accepted: 07 June 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: In this work, the friction characteristics of single-layer MoS₂ prepared with chemical vapor deposition (CVD) at three different temperatures were quantitatively investigated and compared to those of single-layer MoS₂ prepared using mechanical exfoliation. The surface and crystalline qualities of the MoS₂ specimens were characterized using an optical microscope, atomic force microscope (AFM), and Raman spectroscopy. The surfaces of the MoS₂ specimens were generally flat and smooth. However, the Raman data showed that the crystalline qualities of CVD-grown single-layer MoS₂ at 800 °C and 850 °C were relatively similar to those of mechanically exfoliated MoS₂ whereas the crystalline quality of the CVD-grown single-layer MoS₂ at 900 °C was lower. The CVD-grown single-layer MoS₂ exhibited higher friction than mechanically exfoliated single-layer MoS₂, which might be related to the crystalline imperfections in the CVD-grown MoS₂. In addition, the friction of CVD-grown single-layer MoS₂ increased as the CVD growth temperature increased. In terms of tribological properties, 800 °C was the optimal temperature for the CVD process used in this work. Furthermore, it was observed that the friction at the grain boundary was significantly larger than that at the grain, potentially due to defects at the grain boundary. This result indicates that the temperature used during CVD should be optimized considering the grain size to achieve low friction characteristics. The outcomes of this work will be useful for understanding the intrinsic friction characteristics of single-layer MoS₂ and elucidating the feasibility of single-layer MoS₂ as protective or lubricant layers for micro- and nano-devices.

Keywords: atomic force microscope; chemical vapor deposition; grain boundary; friction, mechanical exfoliation; MoS₂

1 Introduction

Atomically thin molybdenum disulfide (MoS₂) exhibits remarkable thermal [1], electrical [2, 3], and optical [4] properties that differ from those of bulk MoS₂. Various electronics [5, 6], optoelectronics [7, 8], and sensors [9, 10] based on atomically thin MoS₂ have been proposed due to these material properties. In addition, the superior mechanical properties [11, 12] and low frictional characteristics [13, 14] of atomically thin MoS₂ suggest that it has potential as an effective protective or lubricant layer [15, 16] for micro- and nano-devices that experience contact sliding during operation.

Recently, the friction characteristics of two-dimensional materials have been extensively investigated using atomic force microscope (AFM). However, most research has focused on graphene [14, 17–21], and the friction characteristics of atomically thin MoS₂ have been reported in only a few studies [14, 22–24]. For example, the friction of atomically thin MoS₂ was found to decrease as the number of layers increased, similar to other two-dimensional materials [14]. This friction behavior was explained based on the puckering

* Corresponding author: Koo-Hyun CHUNG, E-mail: khchung@ulsan.ac.kr
and electron-phonon coupling effects [14, 17]. Friction characteristics of single-layer MoS2 were also compared to that of single-layer graphene [22]. In addition, the friction characteristics of single-layer MoS2 deposited on SiO2, mica, and h-BN substrates were investigated. The results showed that friction was the lowest for the h-BN substrate, due to a significant reduction in surface roughness [23]. In addition, the friction of single-layer MoS2 was found to slightly increase with increasing humidity [24]. It was further demonstrated that adding atomically thin MoS2 to a liquid lubricant can enhance the tribological performances [25]. However, to properly implement atomically thin MoS2 as a protective or a lubricant layer, more data should be gathered over a wide variety of experimental conditions.

Various methods, such as mechanical exfoliation [9, 26], lithium intercalation-assisted exfoliation [24], liquid exfoliation [27], and physical vapor deposition [28] have been used to prepare atomically thin MoS2. Thinning using laser-induced ablation [29], thermal annealing [30], and plasma [31] have also been demonstrated for fabricating single-layer MoS2. However, the lateral size of MoS2 layers prepared using these methods is often relatively small. In addition, the deposition of MoS2 with controlled size and shape onto specific locations using these methods is often challenging due to technical difficulties. Chemical vapor deposition (CVD), based on the vapor phase reaction of Mo and S, is one of the most practical methods for large area synthesis [32–35], and hence can be appropriately used to facilitate atomically thin MoS2 layers as protective or lubricant layers. However, in contrast to the optical [36], electrical [34, 35], and mechanical properties [37], the friction characteristics of CVD-grown MoS2 have not yet been explored. Considering that CVD-grown MoS2 can have large variations in grain size, crystalline orientation, and number of defects [32–35], which in turn affect friction characteristics, a systematic approach is needed to fundamentally understand the intrinsic friction characteristics of CVD-grown MoS2.

In this work, the friction characteristics of single-layer MoS2 prepared using CVD at various temperatures were systematically investigated using AFM. In particular, given that mechanically exfoliated MoS2 has high crystallinity with a small number of defects, the friction characteristics of CVD-grown single-layer MoS2 were compared to those of mechanically exfoliated single-layer MoS2. The MoS2 specimens were also characterized using Raman spectra to understand the crystalline qualities of the specimens. The outcomes of this work are expected to aid in evaluating the feasibility of single-layer MoS2 as protective or lubricant layers for micro- and nano-devices, and optimize CVD growth conditions from a tribological point of view.

## 2 Experimental details

The MoS2 specimens were deposited onto a SiO2/Si substrate using mechanical exfoliation and CVD. The thickness of the thermally grown SiO2 layer on the Si wafer was about 300 nm. After mechanical exfoliation of the MoS2 flakes using natural crystalline MoS2 (SPI supplies), they were carefully examined using an optical microscope to ensure the specimens were free from tape residue. Then, the locations of the single-layer MoS2 were identified based on the thickness-dependent optical contrast. For the CVD process, MoO3 (99%, Aldrich) and sulfur powders (99.5%, Alfa) were used as sources of Mo and S, respectively. Crucibles containing MoO3 and sulfur were placed in the center of a furnace and in the upstream zone of a quartz tube, respectively. The substrates were placed face down above the crucible containing MoO3. Initially, the quartz tube was pumped down to a base pressure of 60 mTorr and purged with high-purity N2 gas to eliminate oxygen. Then, the temperature was gradually increased to 400 °C at a rate of 25 °C/min and maintained for 30 min, while the pressure was set to ~700 mTorr with a 100 sccm N2 gas flow. After the pressure was reduced to atmospheric conditions using 10 sccm N2 gas flow, the temperature for the reaction was increased to 800 °C, 850 °C, and 900 °C at a rate of 25 °C/min and maintained for 5 min for the growth of single-layer MoS2 before cooling to room temperature.

After preparation of the MoS2 specimen, Raman measurements were performed using a confocal Raman system (Alpha300R, Witec) to evaluate the crystalline quality of the MoS2 layer; the excitation laser with a wavelength of 532 nm was used. For CVD-grown
MoS$_2$, the isolated flakes were initially formed and then merged into a continuous film during the CVD process. Hence, a laser with a spot size of ~720 nm was carefully focused on isolated MoS$_2$ flakes to minimize the effects of the grain boundary. In addition, to eliminate laser-induced thermal effects [38, 39] and particle formation [40], the laser power on the specimen surface was limited to below 0.5 mW. The Raman spectra were collected using a 100× objective (NA = 0.9) for 10 s in ambient conditions.

Topographic images of the MoS$_2$ specimens were obtained using the AFM intermittent contact mode and Si tips with a nominal normal spring constant of 2 N/m (AC240, Olympus). The friction loops were obtained under normal forces ranging from 0 nN to 20 nN using a Si tip with a nominal normal spring constant of 0.2 N/m (PPP-LFM, Nanosensors). For quantitative force measurements using AFM, normal [41] and lateral force [22, 42] calibrations were performed. The force calibration results showed that the normal spring constant and lateral force sensitivity of the Si tip used for friction loop measurements were 0.47 N/m and 5.85 mV/nN, respectively. More than 10 friction loops were obtained at 10 different locations on the specimens under each normal force. In particular, considering that AFM tip wear can readily occur due to contact sliding [43], the friction loops were obtained under increasing and decreasing normal forces to minimize the effects of tip wear on the friction force measurements. In the friction loop measurements, scan distance and sliding speed were set to 300 nm and 375 nm/s, respectively. After the friction loop measurements, force-distance curves were obtained before and after the friction loop measurements to further monitor AFM tip wear [45]. More than 10 force-distance curves were obtained at different locations for each specimen. All experiments were performed under ambient conditions (23 °C and 35% relative humidity).

3 Results and discussion

Figure 1(a) presents optical microscope images of the mechanically exfoliated single-layer MoS$_2$ and CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C on the SiO$_2$/Si substrate. For comparison, optical microscope images of isolated flakes of the CVD-grown MoS$_2$ are shown in Fig. 1. The shape of the mechanically exfoliated MoS$_2$ was random with a nominal size of about 10 μm, while the CVD-grown MoS$_2$ exhibited a triangular shape with different sizes depending on growth temperature. The grain sizes of CVD-grown MoS$_2$ increased from 5 μm to 15 μm when the temperature increased from 800 °C to 850 °C. However, the grain sizes of CVD-grown MoS$_2$ significantly decreased, to about 3 μm, at 900 °C. This grain size dependence on CVD temperature is often noted in the literature. For example, it was demonstrated that the CVD-grown MoS$_2$ grain size increases when the temperature increases to about 830 °C, associated with production and diffusion of active mobile species, and then the grain size decreases due to thermal etching when the temperature reaches 900 °C [36]. Figure 1(a) indicates that the grain size of CVD-grown MoS$_2$ at 850 °C was the largest, which is consistent with the results from previous research [36].

Considering that surface topography can significantly affect the friction characteristics of two-dimensional materials [20, 23], the MoS$_2$ specimens were examined using AFM prior to the friction loop measurements. The AFM topographic images of mechanically exfoliated single-layer MoS$_2$ and CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C obtained with the intermittent contact mode are shown in Fig. 1(b). As shown, the surfaces of the mechanically exfoliated and CVD-grown MoS$_2$ were generally flat and smooth. However, for the CVD-grown MoS$_2$ at 800 °C and 850 °C, particles were observed on the substrate surfaces around the flakes. In particular, a large number of particles was observed on the substrate around the CVD-grown MoS$_2$ flakes at 850 °C. These particles
could have contained by-products formed during the deposition process. In contrast to the CVD-grown MoS$_2$ at 800 °C and 850 °C, the substrate surface around the CVD-grown MoS$_2$ flakes at 900 °C was relatively clean, possibly a result of thermal etching [36].

The surface roughness values of the MoS$_2$ specimens were determined from AFM topographic images obtained at five different locations within a 500 nm × 500 nm scanning area. The measured surface roughness values of the mechanical exfoliated MoS$_2$ and CVD-grown MoS$_2$ at 800 °C, 850 °C, and 900 °C were 0.16 ± 0.01 nm, 0.20 ± 0.01 nm, 0.25 ± 0.01 nm, and 0.16 ± 0.01 nm, respectively. These surface roughness values indicate that the MoS$_2$ specimens used in this work had flat and smooth surfaces.

The heights of the MoS$_2$ specimens were determined from the cross-sectional height profiles shown in Fig. 1(b). The height of the MoS$_2$ prepared using mechanical exfoliation was 0.71 nm. However, the heights of the CVD-grown MoS$_2$ were slightly larger, ranging from 0.86 nm to 0.90 nm. These values were also larger than the theoretical thickness of single-layer MoS$_2$ (0.62 nm). These discrepancies between the measured and theoretical thicknesses might have been due to with the presence of adsorbents below the MoS$_2$ layer, interactions between the MoS$_2$ layers and substrate [46, 47], and/or AFM measurement uncertainty [48]. Nonetheless, the thicknesses of the single-layer MoS$_2$ specimens prepared using mechanical exfoliation and CVD were comparable to those of single-layer MoS$_2$ on a SiO$_2$/Si substrate.

To understand the crystalline qualities of the MoS$_2$ specimens, Raman measurements were performed. Figure 2(a) shows examples of the Raman spectra of mechanically exfoliated MoS$_2$ and CVD-grown MoS$_2$ at 800 °C, 850 °C, and 900 °C. Two characteristic Raman peaks for MoS$_2$, $E_{2g}^{1}$ and $A_{1g}$, resulting from the in-plane vibration of Mo-S atoms and out-of-plane vibration of S atoms, respectively [47], are clearly shown in Fig. 2(a). The thickness and crystalline quality of the atomically thin MoS$_2$ layer can be identified using the frequencies of the $E_{2g}^{1}$ and $A_{1g}$ peaks and the frequency separation between the two peaks, $\Delta k$ [32, 47, 49]. The frequencies of the $E_{2g}^{1}$ and $A_{1g}$ peaks for the mechanically exfoliated and CVD-grown MoS$_2$ are summarized in Fig. 2(b), along with $\Delta k$. The measured value of $\Delta k$ for the mechanically exfoliated MoS$_2$ was 17.5 cm$^{-1}$, which corresponds to single-layer...
Fig. 2 (a) Raman spectra, (b) frequencies of the $E_{2g}$ and $A_{1g}$ peaks and the frequency separation, $\Delta k$, and (c) FWHM of the $E_{2g}$ peak of the mechanically exfoliated single-layer MoS$_2$ and CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C. In (a), (b), and (c), the frequency of the $E_{2g}$ and $A_{1g}$ peaks and FWHM of the $E_{2g}$ peak of the mechanically exfoliated single-layer MoS$_2$ are indicated with dashed lines for comparison. In (b) and (c), the error bar represents one standard deviation.

MoS$_2$ [47]. Additionally, the measured values of $\Delta k$ for the CVD-grown MoS$_2$ at 800 °C, 850 °C, and 900 °C were 21.2 cm$^{-1}$, 21.0 cm$^{-1}$, and 20.5 cm$^{-1}$, respectively, which agrees with those of CVD-grown single-layer MoS$_2$ [32, 49]. These Raman measurement results, along with the AFM data shown in Fig. 1(b), show that single-layer MoS$_2$ formed on the substrate for both the mechanical exfoliation and CVD processes. In addition, the larger values of $\Delta k$ for the CVD-grown MoS$_2$ suggest that the CVD-grown MoS$_2$ may have inferior crystalline qualities compared to the mechanically exfoliated MoS$_2$ [32, 49].

To further evaluate the crystalline qualities of mechanically exfoliated and CVD-grown MoS$_2$, the full width at half maximum (FWHM) of the $E_{2g}$ peak was also obtained from the Raman data [49], as shown in Fig. 2(c). The FWHM of the $E_{2g}$ peaks for the mechanically exfoliated MoS$_2$ was 3.4 cm$^{-1}$. Further, as shown, the FWHM of the $E_{2g}$ peaks for the CVD-grown MoS$_2$ at 800 °C and 850 °C agreed with that of the mechanically exfoliated MoS$_2$. However, the FWHM of the $E_{2g}$ peak for the CVD-grown MoS$_2$ at 900 °C was significantly larger than that of the mechanically exfoliated MoS$_2$. This result further indicates that CVD-grown MoS$_2$ at 800 °C and 850 °C have better crystalline qualities than the CVD-grown MoS$_2$ at 900 °C [36, 49].

After characterization of the MoS$_2$ specimens using an optical microscope, AFM, and Raman spectroscopy, the adhesion and friction characteristics were determined from the force-distance curves and friction loops, respectively. Figure 3 presents examples of the force-distance curves and average adhesion forces between the Si tip used in this work and the MoS$_2$ specimens, obtained before and after friction loop measurements. As shown in Fig. 3(a), the adhesion forces between the Si tip and CVD-grown MoS$_2$ at 800 °C and 850 °C were about 16 nN, which agreed with values found between the Si tip and mechanically exfoliated MoS$_2$. However, the adhesion force between the Si tip and CVD-grown MoS$_2$ at 900 °C was slightly larger than in the other cases. As shown in Fig. 3(b), the average adhesion forces of mechanically exfoliated single-layer MoS$_2$ and CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C were 17 ± 2 nN, 16.0 ± 0.3 nN, 16.1 ± 0.7 nN, and 19.8 ± 0.4 nN, respectively. It was hypothesized that interactions with the Si tip might have been enhanced due to imperfections in the crystalline structure, which resulted in a relatively large adhesion force of the CVD-grown single-layer
Fig. 3 (a) Examples of force-distance curves and (b) average adhesion forces of the mechanically exfoliated single-layer MoS$_2$ and CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C. In (b), the adhesion forces obtained before and after friction loop measurements are compared. In (b), the error bar represents one standard deviation.

MoS$_2$ at 900 °C. In addition, the adhesion force for the mechanically exfoliated MoS$_2$ slightly decreased after the friction loop measurements. A ripple structure may have been formed due to inhomogeneous interaction between MoS$_2$ and substrate during mechanical exfoliation, which in turn caused the local differences in frictional and material properties [18, 50]. This local difference in materials properties of the mechanically exfoliated MoS$_2$ might have caused the difference in adhesion force before and after friction loop measurements, as shown in Fig. 3(b). However, because the CVD-grown MoS$_2$ may exhibit larger adhesion to the substrate compared to the mechanically exfoliated MoS$_2$ [51], the effect of ripple structures on the adhesion of CVD-grown MoS$_2$ was expected to be relatively small. This small effect may be responsible for the good agreement between the adhesion forces of CVD-grown MoS$_2$ before and after friction loop measurements. In addition, the data shown in Fig. 3 suggests that the tip wear that occurred during friction loop measurements was negligible.

Figure 4(a) shows examples of friction loops for the mechanically exfoliated single-layer MoS$_2$ and CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C obtained under 10 nN normal force. The friction force of the mechanically exfoliated MoS$_2$ was about 2 nN under this normal force of 10 nN. However, the friction forces of the CVD-grown MoS$_2$ were significantly larger than that of the mechanically exfoliated MoS$_2$, increasing from 3.7 nN to 8.3 nN as temperature increased from 800 °C to 900 °C with 10 nN of applied normal force.

Variations in the friction force of mechanically exfoliated single-layer MoS$_2$ and CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C with respect to the normal force are shown in Fig. 4(b). In general, the friction force increased as the normal force increased, as expected. However, a nonlinear relationship between the friction force and normal force [52] was not clearly observed in this work. Figure 4(b) shows that force of the single-layer MoS$_2$ prepared using mechanical exfoliation increased from 1.5 nN to 2.3 nN when the normal force increased from 0 to 20 nN. As shown, the friction forces of the CVD-grown MoS$_2$ were significantly larger than those of the mechanically exfoliated MoS$_2$, which increased with increasing CVD temperature. The friction forces of the CVD-grown MoS$_2$ at 800 °C, 850 °C, and 900 °C were, respectively, 2.0, 2.4, and 4.2 times larger than that of the mechanically exfoliated MoS$_2$. The friction forces under a zero normal force were also the greatest for the CVD-grown MoS$_2$ at 900 °C. Furthermore, the increases in friction force with increasing normal force were relatively gradual for the mechanically exfoliated MoS$_2$ and the CVD-grown MoS$_2$ at 800 °C and 850 °C, whereas the increase was much more significant for the CVD-grown MoS$_2$ at 900 °C.

Many factors, such as the number of layers [14, 17], surface topography [20, 23], interactions with the substrate [21], and crystalline orientation [18] can affect the friction characteristics of two-dimensional materials, in addition to the experimental and environmental conditions [24]. The dependence of
MoS$_2$ friction characteristics on the surface topography is not clearly indicated in Figs. 1 and 4. This lack of correlation might have been due to the surface topographies of the MoS$_2$ specimens being too homogeneous to affect the friction characteristics. In addition, the effect of crystalline orientation on the friction characteristics was unlikely to appear in the data presented in Fig. 4, considering that the frictional anisotropy due to crystalline orientation is greater at a lower normal force [18]. However, the Raman data in Fig. 2 suggest that the larger friction characteristics of CVD-grown MoS$_2$ compared to mechanically exfoliated MoS$_2$ was likely associated with the crystalline imperfections of the CVD-grown MoS$_2$. The crystalline qualities of CVD-grown MoS$_2$ at 800 °C and 850 °C were relatively close to those of the mechanically exfoliated MoS$_2$. This can explain why the friction characteristics of CVD-grown MoS$_2$ at 800 °C and 850 °C were relatively close to those of the mechanically exfoliated MoS$_2$. However, the crystalline qualities of MoS$_2$ grown at 900 °C were significantly lower than those of the mechanically exfoliated MoS$_2$, which may be responsible for the relatively high friction characteristics. It should also be noted that the large friction characteristics of the CVD-grown MoS$_2$ at 900 °C may be related to the large adhesion shown in Fig. 3.

The CVD-grown MoS$_2$ may readily contain grain boundaries because isolated flakes of MoS$_2$ were merged into a continuous film during the CVD process. Therefore, understanding the friction characteristics at grain boundaries of CVD-grown MoS$_2$ is important. Figure 5 shows the AFM topographic and friction force images obtained under the intermittent contact and contact modes, respectively, over the grain boundaries of the CVD-grown MoS$_2$ at 800 °C, 850 °C, and 900 °C. The cross-sectional height profiles and friction loops are also compared in Fig. 5. The friction force images were obtained under 3 nN normal force concurrently with the friction loop. The grain boundaries can be clearly observed in the topographic images. In addition, the cross-sectional height profiles show that the heights at the grain boundaries were slightly larger than those at the grains. The friction force images and friction loops clearly show that the friction forces at the grain boundaries were significantly larger than those at the grains. However, the friction forces on the left and right sides of the grain boundaries were quite similar to each other, indicating that there was an insignificant effect of crystalline orientation on friction of the CVD-grown MoS$_2$ specimens. The friction forces under 3 nN normal force at the grains and grain boundaries of the CVD-grown MoS$_2$ at 800 °C, 850 °C, and 900 °C are shown in Fig. 5(c). Compared to the grains, the friction at the grain boundaries increased by factors ranging from 1.5 to 3.0. Defects at the grain boundary may lead to changes in the frictional properties, similar to the
This outcome suggests that the CVD growth temperature for MoS$_2$ should also be optimized for grain boundary density to achieve low friction characteristics, considering that CVD growth temperature often affects grain size, as shown in Fig. 1.

![AFM topographic images and friction force images of the CVD-grown single-layer MoS$_2$ at 800 °C, 850 °C, and 900 °C, and comparison of the friction forces at the grains and grain boundaries of the CVD-grown MoS$_2$. The topographic images were obtained with the intermittent contact mode, and the friction force images were obtained using the AFM contact mode under 3 nN normal force. The red dashed lines in (a) and (b) indicate the locations of the cross-sectional height profiles and friction loops, respectively. In (a), the grain boundaries are indicated with black arrows. In (c), the error bar represents one standard deviation.](image)

According to the experimental results, the CVD-grown single-layer MoS$_2$ showed greater friction than the mechanically exfoliated single-layer MoS$_2$, which may be associated with crystalline imperfections in the CVD-grown MoS$_2$. This outcome indicates that the growth of the single-layer MoS$_2$ with high...
crystalline quality, is preferred when trying to achieve low friction characteristics. The friction force measurement results also suggest that the optimal CVD growth temperature is 800 °C based on tribological concerns. However, it should be noted that the frictional properties strongly depend on the test parameters, such as contact pressure, sliding speed, and environmental conditions. In addition, the MoS2 crystalline structure and quality can be significantly affected by the deposition method and conditions. From this perspective, the results of this work cannot be generalized to all types of single-layer MoS2. Nevertheless, the outcomes of this work are valid for similar types of MoS2 grown using the CVD process. Furthermore, the results presented in this work are expected to aid in a more comprehensive and fundamental understanding of the frictional properties of single-layer MoS2.

4 Conclusions

In this work, the friction characteristics of mechanically exfoliated single-layer MoS2 and CVD-grown MoS2 at three different temperatures were systematically investigated using AFM. The mechanically exfoliated and CVD-grown MoS2 specimens were also examined using an optical microscope, AFM, and Raman spectroscopy to understand the surface characteristics and crystalline qualities of the specimens. The surfaces of the mechanically exfoliated and CVD-grown MoS2 were generally flat and smooth. The Raman data showed that the crystalline qualities of CVD-grown MoS2 at 800 °C and 850 °C were relatively close to those of mechanically exfoliated MoS2. However, the crystalline qualities of the CVD-grown MoS2 at 900 °C were lower compared to those of the mechanically exfoliated MoS2. The CVD-grown MoS2 exhibited greater friction than the mechanically exfoliated single-layer MoS2, which might have been associated with the crystalline imperfections in CVD-grown MoS2. In addition, the friction of the CVD-grown MoS2 increased as the CVD growth temperature increased. In terms of the frictional properties, the optimal growth temperature was 800 °C for the CVD process used in this work. Furthermore, the friction at the grain boundaries was larger than at the grain interior by factors of 1.5 to 3.0, which might have been due to defects at the grain boundary. The outcomes of this work are expected to advance the fundamental understanding of intrinsic friction characteristics of single-layer MoS2, and elucidate the feasibility of single-layer MoS2 as protective or lubricant layers for micro- and nano-devices.

Acknowledgement

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT and Future Planning (NRF-2017R1A2B4009651).

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


Dinh Le Cao KY. He received his M.S. degree in mechanical engineering in 2010 from University of Technology, Ho Chi Minh City, Vietnam. He is currently pursuing his PhD degree in the Tribology and Surface Engineering Laboratory at University of Ulsan, Republic of Korea. His research interests include fundamental understanding of friction characteristics of nanomaterials from experiments and molecular dynamics simulation.

Bien-Cuong TRAN KHAC. He received his M.S. degree in mechanical engineering in 2014 from University of Ulsan, Republic of Korea. He is currently pursuing his PhD degree in the Tribology and Surface Engineering Laboratory at the same university. His research interests include tribology and surface damage characteristics of atomically thin (2D) materials.
Chinh Tam LE. He received his B.S. degree in material science in 2012 from Ho Chi Minh University of Natural Science in Vietnam. After then, he joined the Semiconductor Device Research Laboratory from 2013 and is currently the PhD student in University of Ulsan, Republic of Korea. His interest research is synthesis and optical characterization of monolayer transition metal dichalcogenides MX2.

Yong Soo KIM. He received his M.S. and PhD degrees in physics from Seoul National University, Republic of Korea in 1993 and 1998, respectively. After then, he was a senior and principle researcher at R&D division, SK-Hynix Inc. He joined the Department of Physics at University of Ulsan, South Korea from 2008. His current position is associate professor, chair of physics department and director of human resource center for novel materials research experts (BK21+ program). His research interest includes organic-inorganic hybrid solar cell, 2-D layer materials relating optoelectronic device, especially transitional metal dichalcogenide growth and its optical characteristics.

Koo-Hyun CHUNG. He received his M.S. and PhD degrees in mechanical engineering from Yonsei University, Republic of Korea, in 1997 and 2005, respectively. His current position is an associate professor at the School of Mechanical Engineering, University of Ulsan, South Korea. His research areas cover tribology, micro/nano tribology, adhesion, surface engineering, molecular dynamics simulation, as well as themes relating to material science.
Contribution of friction and adhesion to the reliable attachment of a gecko to smooth inclines

Zhouyi WANG1,*, Qiang XING2, Wenbo WANG1, Aihong JI1, Zhendong DAI1,*
1 Institute of Bio-inspired Structure and Surface Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
2 School of Mechanical Engineering, Nantong University, Nantong 226019, China
Received: 13 October 2016 / Revised: 05 December 2016 / Accepted: 19 June 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: Geckos’ ability to move on steep surfaces depends on their excellent adhesive structure, timely adjustments on locomotor behaviors, and elaborates control on reaction forces. However, it is still unclear how they can generate a sufficient driving force that is necessary for locomotion, while ensuring reliable adhesion on steep inclines. We measured the forces acting on each foot and recorded the contact states between feet and substrates when geckos encountered smooth inclination challenges ranging from 0° to 180°. The critical angles of the resultant force vectors of the front and hind-feet increased with respect to the incline angles. When the incline angle became greater than 120°, the critical angles of the front- and hind-feet were similar, and the averages of the critical angles of the front- and hind-feet were both smaller than 120°, indicating that the complicated and accurate synergy among toes endows gecko’s foot an obvious characteristic of “frictional adhesion” during locomotion. Additionally, we established a contact mechanical model for gecko’s foot in order to quantify the contribution of the frictional forces generated by the heel, and the adhesion forces generated by the toes on various inclines. The synergy between multiple contact mechanisms (friction or adhesion) is critical for the reliable attachment on an inclined surface, which is impossible to achieve by using a single-contact mechanism, thereby increasing the animal’s ability to adapt to its environment.

Keywords: friction; adhesion; incline; frictional adhesion; gecko

1 Introduction

The reliable attachment between an animal’s foot and the substrate over which it is moving forms the foundation of its movement. This is because reliable attachment is essential in providing a sufficient and continuous contact force to counteract resistance and enable locomotion [1]. Inclines constitute a common terrain over which legged animals could move [2]. Thus, a major challenge faced by a legged animal is how it can reliably attach itself to an incline while moving. During the course of evolution, animals have optimized several means (attachment organs) for conquering the challenge of climbing steep inclines, including the development of claws, and smooth, hairy adhesive pads [3, 4]. The ability of a claw to reliably attach itself to a substrate depends on the frictional coefficient between the claw and the substrate, as well as the angle with which the claw engages with the substrate asperities, and the depth of a claw as it penetrates into a surface. In plain terms, the claw is not suited to move over a smooth substrate [5, 6]. Therefore, adhesive pads have evolved to adhere to smooth substrates where claws fail to grip [7]. Smooth, deformable pads generate capillary-like forces, which allow organisms such as insects and tree frogs to remain attached to various substrates. This adhesive mechanism is known as “wet adhesion” [8–10]. Hairy pads of geckos require a fine proximal pull to establish intimate contact between the flat spatula-shaped tips

* Corresponding authors: Zhouyi WANG, E-mail: wzyxml@nuaa.edu.cn; Zhendong DAI, E-mail: zddai@nuaa.edu.cn
and the substrate [11, 12]. This adhesive mechanism is known as “dry adhesion”, which is based on van der Waals forces [13]. On the other hand, the hair found in several insects also operates in accordance to the mechanisms of wet adhesion. Thus, the underlying mechanisms of these smooth and hairy pads are different, but both generate adhesive forces.

Generally, the attaching organs of animals are both diverse and, in the case of some animals, hierarchical. Many insects have not only several claws but also some smooth or hairy pads on their extremities that can generate adhesive forces [14, 15]. The adhesion system of Gekko geckos has an elaborate hierarchical structure. The extraordinary climbing ability of geckos on inverted inclines is not only due to the van der Waals forces between the submicron-sized spatulae and the substrate, but it is also partially attributed to the synergy between the hierarchical units [16–19]. For example, the flexible lamellae on the feet ensure that the setal arrays maintain intimate contact with almost all substrates [16, 20]. Furthermore, the coupling between the front limbs and hind limbs can generate opposite reaction forces that enhance the stability of the gecko on inverted surfaces [17, 21]. In regard to the mesoscale foot, which consists of five toes covered with setae and a nonadhesive heel, our aim was the determination of how geckos coordinate the functions of the separate parts, in order to achieve reliable attachment when considering the challenge of a wide range of smooth inclines (0° to 180°).

Under the influence of a preload and pulling forces, a single seta can generate a 200 μN shear force (\(F_{\parallel}\), parallel to a substrate) and a 40 μN adhesion force (\(F_{\perp}\), perpendicular to the substrate). If the angle of the resultant force vector (\(\alpha = \tan^{-1}(F_{\perp}/F_{\parallel})\)) is greater than 30°, then the setae detach from the substrate. This means that the critical angle (\(\alpha^*\)) of the resultant force vector acting on an isolated seta is 30° [11, 22]. When the setal array is pulled along its natural path to generate normal adhesion forces, the critical angle (\(\alpha^*\)) of the resultant force vector acting on the setal array is 24.6° [23, 24]. Geckos attached to a glass slide by a single toe became detached at an average critical angle (\(\alpha^*\)) of 25.5° [23]. There is no significant difference between the critical angle (\(\alpha^*\)) for a single toe and a setal array, which is less than that of a single seta.

Considering the direction of adhesion, the adhesion force (\(F_{\perp}\)) directly depends on shear force (\(F_{\parallel}\)) and is limited by the critical detachment angle. This adhesion characteristic is defined as the “frictional adhesion” model. By frictional adhesion, the adhesion force can be precisely controlled via the shear force, allowing attachment and detachment to occur with negligible forces. Studies have shown that a single seta, a setal array, and a toe, all exhibit a property known as “frictional adhesion” [23]. However, none of these studies has proven experimentally that a foot consisting of hierarchical frictional adhesion components at different scales exhibits this frictional adhesion characteristic.

Geckos can reliably attach themselves to inclines due to the combined effect of components at different scales [16, 17], with frictional adhesion being provided by hierarchical structures with transmissibility characteristics, from the micro-scale seta to the mesoscale toe [23]. We performed trials with Gekko gecko that could move freely over a rotated three-dimensional force-measuring array (FMA) [25], and investigated the forces acting on an individual foot, and on the foot’s contact area with the substrate, in a wide range of smooth inclines (0° to 180°). Our aim was to verify by experiment whether the gecko’s mesoscale foot exhibited the characteristic “frictional adhesion” in locomotion. Meanwhile, the cooperative mechanism, whereby the adhesion of the toes acts in cooperation with the friction of the heel was studied to reveal the deployment strategy of the adhesion system in response to the challenge of smooth inclines.

2 Materials and methods

2.1 Animal

This study was carried out in accordance with the Guide of Laboratory Animal Management Ordinance of China. The experimental procedures were approved by the Jiangsu Association for Laboratory Animal Science (Jiangsu, China). A special room, which was under simulated wild environment of gecko habitat, including rock crevices, a water pool, a lighting system, and a ventilation system, was built to raise Gekko geckos, which were obtained from a supplier (Jun-Hao Wild
Animal Science & Technology Development Co. Ltd.) in Guangxi Autonomous Region, China. With the help of the lighting system and ventilation system, they were housed under simulated natural conditions with fresh water and live insects (cricket and locust, etc.) as food. A regular disinfection of the feeding room was performed by 1/1000 potassium permanganate solution every three days. The geckos were monitored daily to confirm their living states, which were revealed by indicators of food-intake, water drinking, and escape speeds.

Two to four years old adult male geckos (62.3 ± 1.8 g mass, mean ± s.d., snout-vent length: 128.3–139.5 mm, N = 11) were used in this study. During the experiments, there was almost no damage to the gecko. After the experiments, all of the experimental animals were again housed in animal room and were cared by professional nursing staff. In order to reduce the potential pain caused by experiments, geckos were lured to cross an FMA-like aisle that connected two boxes. During the experiment, a black box was fixed at the aisle end to lure them to climb fluently.

2.2 Experimental equipment and procedure

Details of force measurements and behavior recordings have been described in our previous work [17, 25] (Fig. 1(a)). Briefly, the forces acting on each foot were measured through the FMA which consisted of 3D force sensors, having a smooth square glass \( R_a = 0.008 \mu m \) at the top (30 mm × 30 mm with 1 mm clearance gap) (Fig. 1(b)) [25]. The aisle of FMA was rotated (30° per step) to imitate different inclined surfaces. Synchronously to the force measurement (NI, 500 Hz), a high-speed camera (iSpeed-3, Olympus, 1280 × 1024 pixels) recorded each trial at 500 fps. Two mirrors were placed at the two sides of the channel with angles of 45°, enabling us to see the lateral of the geckos from the side-on view. The forces acquisitions and video recordings were triggered by a pulse signal at the same time (Fig. 1(a)).

2.3 Analysis of force and video recordings

The gecko’s toe adducts to attach to a substrate, and then abducts to detach itself. Video captured with a high-speed camera shows that the contact process between the foot and the substrate can be subdivided into three periods: the incipient period of contact \( (T_{IPC}) \), which starts when the heel makes contact with the substrate \( (t_1) \) and ends when the toes attach to the substrate \( (t_2) \); the stable period of contact \( (T_{SPC}) \), which starts when the toes, attached to the substrate \( (t_2) \), begin to abduct prior to detaching \( (t_3) \); and the released period of contact \( (T_{RPC}) \), which starts at the beginning of the abduction of the toe \( (t_3) \) and ends when all of the toes detach from the substrate \( (t_4) \) (Figs. 2(a) and 2(c)). Thus, the synchronous force data, as obtained with the FMA, is also subdivided into three periods (Figs. 2(b) and 2(d)).

The force resulting from the combination of the lateral and fore-aft forces is a shear force \( F_\parallel \), which can be calculated by summing the components of the lateral and fore-aft forces in accordance to Eq. (1), which act parallel to the plane of the sensor array. The normal direction is defined as being perpendicular to the plane of the array, and the force is known as a normal force \( F_\perp \). To investigate the contribution of the friction of the heel and the adhesion of the toes to the reliable attachment when a gecko climbs different inclines, we focused on the video and force data for the feet in \( T_{SPC} \) (the pink points in Figs. 2(b) and 2(d)).

The angle of the resultant force vector \( \alpha \) of the foot was calculated from \( F_\perp \) and \( F_\parallel \) in \( T_{SPC} \) for each trial.
in accordance to Eq. (2). The maximum angle was the critical angle ($\alpha^*$) of the resultant force vector of the foot (Eq. (3)).

To congruously describe the data result of the critical angle ($\alpha^*$), as collected from a gecko moving on different inclines, the normal direction of the substrate was defined as the starting position of $\alpha^*$, with the clockwise direction being positive. The clockwise $\alpha^*$ defined in our research differs from that defined in previous studies [23], with the former and the latter having a phase difference of 90°. Moreover, the $F_\perp$ and $F_\parallel$ corresponding to $\alpha^*$ were selected as the critical normal force ($F_\perp^*$) and critical shear force ($F_\parallel^*$) (Figs. 2(b) and 2(d)).

$$F_\parallel = \sqrt{F_\perp^2 + F_\parallel^2}$$  \hspace{1cm} (1)

$$\alpha = \tan^{-1}(F_\perp / F_\parallel) \hspace{0.5cm} \alpha \in (0^\circ, 180^\circ)$$  \hspace{1cm} (2)

$$\alpha^* = \max\{\alpha = \tan^{-1}(F_\perp / F_\parallel)\} \hspace{0.5cm} \alpha^* \in (0^\circ, 180^\circ)$$  \hspace{1cm} (3)

### 2.4 Statistics

Both the local velocities and the average velocities of a gecko were calculated for every trial to select available trails—if the interval velocity was 15% greater than or less than the average velocity, the trial was discarded. Data from all individuals were pooled, and the SPSS software (SPSS15.0, Inc., Chicago, IL) was used for all analyses. Force data were normalized by
body weight (BW) in order to account for differences in body size across the sample population. In spite of our great efforts to obtain equivalent forward speeds on different inclines, the geckos slowed their forward motion when they encountered an increased angle of the incline [26]. Thus, the velocity was set as the covariate variables for each co-variance (ANCOVA) analyzed. We use the ANCOVA to compare among data for incline from 0° to 180°, where the incline was set as the independent variable and the dependent variables included $a^*$, $F_{\perp}^*$, and $F_{\parallel}^*$. We used ANCOVA again to compare the differences between the front and hind limb on incline; the grouping of the foot was set as the independent variable while the dependent variables were the same as above. The relationships between the $F_{\perp}^*$ and $F_{\parallel}^*$ on inclines were determined using least-squares linear regression. The similarity among $a^*$ of feet collected from different inclines was evaluated by the Euclidean distance in hierarchical clustering and the dendrogram result was printed as figure. Because different animals were used for the seven inclines trials, we did not use repeated-measures ANOVA. Differences were considered statistically significant when $p \leq 0.05$. The tested data are presented as mean ± standard deviation (mean ± s.d.).

3 Results

3.1 State of foot during stable period of contact

The image at the instant corresponding to $a^*$ was selected from the video to show the contact state between the feet and the substrate (Fig. 3). On a horizontal substrate, the gecko abducts all of its toes to prevent the toes from adhering to the substrate (Fig. 3(a)), and uses its non-adhesive heels to place pressure on the substrate, thus generating a supporting force to oppose its weight and the friction force acting against propelling or braking locomotion. For an incline of 30° to 90°, the gecko continues to adduct some of its toes so that it can adhere to the substrate, relying on the higher adhesive and shear forces, although the heels remain in contact with the substrate (Fig. 3(b)). When the incline is greater than 90°, each of the toes of the foot adheres to the substrate, while the heel is pulled away from the substrate by gravity (Fig. 3(c)).

Fig. 3 Foot contact with the substrate during a stable period of contact ($T_{SPC}$) for an incline of 0° to 180°. (a) Only the heel contacts the substrate. (b) Concomitant contact state involving both the heel and toe. (c) Toes contact the substrate while the heel is held away from the substrate. The state of the left front foot is shown in the left column of the figure, while the state of the right hind foot is shown in the right column of the figure. The red circle indicates that the toes are not in contact with the substrate. The red bevel indicates that the heel is not in contact with the substrate.

3.2 Critical angle of resultant force vector of foot ($a^*$)

The critical angles of the resultant force vector ($a^*$) for the front and hind feet increased with respect to the incline, and $a^*$ for the front foot ($a^* = 0.632\theta + 22.607$, $F = 754.477$, $R^2 = 0.842$, $d.f. = 136$, $p < 0.001$) increased faster than that for the hind foot ($a^* = 0.595\theta + 26.124$, $F = 795.014$, $R^2 = 0.856$, $d.f. = 136$, $p < 0.001$) (Fig. 4(a); Table 1). Remarkably, $a^*$ for the front foot exhibited no significant difference for any of the
120°, 150°, and 180° inclines, as did the α* of the hind foot (Table S1 in the Electronic Supplementary Material (ESM)). Meanwhile, there is no significant difference between α* for the front and hind feet (Table S2 in the ESM). When the incline exceeded 120°, the data for α* for the front and hind feet are clearly similar (Fig. 4(b)), and the average α* of both the front and hind feet did not exceed 120°.

### 3.3 Critical normal forces and shear force corresponding to critical angle

The critical normal forces (\(F_{\perp}\)) of the front and hind feet decreased with respect to the incline (front foot: \(F_{\perp} = -0.009\theta + 0.761, R^2 = 0.848, d.f. = 134, p < 0.001\); hind foot: \(F_{\perp} = -0.007\theta + 0.734, R^2 = 0.796, d.f. = 134, p < 0.001\)). The critical shear force (\(F_{||}\)) of the front and hind feet decreased with respect to the incline (front foot: \(F_{||} = 0.006\theta + 0.234, R^2 = 0.730, d.f. = 134, p < 0.001\); hind foot: \(F_{||} = 0.003\theta + 0.441, R^2 = 0.278, d.f. = 134, p < 0.001\)).

### 4 Discussion

Geckos can move freely on steep inclines, because of the excellent adhesion performance of their toes, the real-time adjustment of the locomotive behavior, and precise control over reaction forces. As a result, geckos continually modulate the reaction force acting on their feet in response to the challenge posed by...

---

**Table 1** Mean of forces and angles of front and hind feet of gecko at different inclines.

<table>
<thead>
<tr>
<th>Inclines</th>
<th>0°</th>
<th>30°</th>
<th>60°</th>
<th>90°</th>
<th>120°</th>
<th>150°</th>
<th>180°</th>
</tr>
</thead>
<tbody>
<tr>
<td>(F_{\perp}) of front foot (BW)</td>
<td>0.68 ± 0.14(20)</td>
<td>0.64 ± 0.15(19)</td>
<td>0.42 ± 0.18(20)</td>
<td>-0.21 ± 0.09(20)</td>
<td>-0.42 ± 0.13(21)</td>
<td>-0.54 ± 0.24(22)</td>
<td>-0.68 ± 0.22(20)</td>
</tr>
<tr>
<td>(F_{\perp}) of hind foot (BW)</td>
<td>0.51 ± 0.15(20)</td>
<td>0.71 ± 0.18(19)</td>
<td>0.51 ± 0.17(20)</td>
<td>-0.01 ± 0.16(20)</td>
<td>-0.30 ± 0.10(21)</td>
<td>-0.44 ± 0.09(22)</td>
<td>-0.50 ± 0.17(20)</td>
</tr>
<tr>
<td>(F_{</td>
<td></td>
<td>}) of front foot (BW)</td>
<td>0.14 ± 0.06(20)</td>
<td>0.42 ± 0.19(19)</td>
<td>0.72 ± 0.27(20)</td>
<td>0.81 ± 0.12(20)</td>
<td>0.98 ± 0.21(21)</td>
</tr>
<tr>
<td>(F_{</td>
<td></td>
<td>}) of hind foot (BW)</td>
<td>0.16 ± 0.12(20)</td>
<td>0.73 ± 0.31(19)</td>
<td>0.95 ± 0.36(20)</td>
<td>0.75 ± 0.22(20)</td>
<td>0.69 ± 0.26(21)</td>
</tr>
<tr>
<td>α* of front foot (°)</td>
<td>11.86 ± 4.86(20)</td>
<td>33.00 ± 12.67(19)</td>
<td>63.57 ± 10.72(20)</td>
<td>104.50 ± 6.41(20)</td>
<td>114.51 ± 7.63(22)</td>
<td>117.68 ± 7.43(20)</td>
<td></td>
</tr>
<tr>
<td>α* of hind foot (°)</td>
<td>17.31 ± 10.82(20)</td>
<td>44.00 ± 12.06(19)</td>
<td>61.24 ± 12.83(20)</td>
<td>92.16 ± 11.11(20)</td>
<td>114.36 ± 5.91(21)</td>
<td>117.75 ± 4.87(22)</td>
<td>114.68 ± 9.75(20)</td>
</tr>
</tbody>
</table>

Values are means ± s.e.m.; n values are given in parentheses; BW is body weight.
Fig. 5 The planar mechanical model of contact between foot and the inclined substrate. \( \theta \): incline of the substrate; v: locomotive direction. (a) A mechanical model of contact for 0° incline. Only the heel pushes away from the substrate to generate a supporting force \((F_{\perp_{\text{heel}}})\) and friction force \((F_{\parallel_{\text{heel}}})\) on a level surface. (b) A mechanical model of contact when incline does not exceed 90°. The state of contact is a concomitant contact of friction and adhesion. The toe pulls toward the substrate to generate adhesive friction force \((F_{\parallel_{\text{toe}}})\) and adhesion force \((F_{\perp_{\text{toe}}})\), while the heel pushes away the substrate. \( \phi \): angle between the toe and the substrate. (c) A mechanical model of contact when the incline exceeds 90°. More toes adhere to the substrate to generate adhesion and friction forces, whereas the heel does not make contact with the substrate. \( F_{\perp_{\text{foot1}}} \) and \( F_{\perp_{\text{foot2}}} \): the adhesion forces acting on foot1 and foot2; \( F_{\parallel_{\text{foot1}}} \) and \( F_{\parallel_{\text{foot2}}} \): the adhesive friction forces acting on foot1 and foot2; \( \phi_{1} \) and \( \phi_{2} \): the angles between the foot1/2 and the substrate.

Table 2 The linear regression of critical normal reaction force \((F_{\perp}^*)\) on critical shear reaction force \((F_{\parallel}^*)\) acting on front and hind feet of gecko at different inclines.

<table>
<thead>
<tr>
<th>Inclines</th>
<th>Foot</th>
<th>d.f.</th>
<th>F</th>
<th>( R^2 )</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0°</td>
<td>front</td>
<td>19</td>
<td>0.007</td>
<td>0.001</td>
<td>0.935</td>
</tr>
<tr>
<td></td>
<td>hind</td>
<td>19</td>
<td>1.464</td>
<td>0.068</td>
<td>0.400</td>
</tr>
<tr>
<td>30°</td>
<td>front</td>
<td>18</td>
<td>0.015</td>
<td>0.003</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td>hind</td>
<td>18</td>
<td>0.474</td>
<td>0.026</td>
<td>0.500</td>
</tr>
<tr>
<td>60°</td>
<td>front</td>
<td>19</td>
<td>1.916</td>
<td>0.084</td>
<td>0.181</td>
</tr>
<tr>
<td></td>
<td>hind</td>
<td>19</td>
<td>0.073</td>
<td>0.004</td>
<td>0.790</td>
</tr>
<tr>
<td>90°</td>
<td>front</td>
<td>19</td>
<td>0.743</td>
<td>0.040</td>
<td>0.400</td>
</tr>
<tr>
<td></td>
<td>hind</td>
<td>19</td>
<td>0.061</td>
<td>0.003</td>
<td>0.807</td>
</tr>
<tr>
<td>120°</td>
<td>front</td>
<td>20</td>
<td>11.849</td>
<td>0.389</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>hind</td>
<td>20</td>
<td>17.303</td>
<td>0.449</td>
<td>0.001</td>
</tr>
<tr>
<td>150°</td>
<td>front</td>
<td>21</td>
<td>11.382</td>
<td>0.361</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>hind</td>
<td>21</td>
<td>69.781</td>
<td>0.786</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>180°</td>
<td>front</td>
<td>19</td>
<td>13.461</td>
<td>0.461</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>hind</td>
<td>19</td>
<td>10.962</td>
<td>0.448</td>
<td>0.039</td>
</tr>
</tbody>
</table>

Therefore, we questioned how geckos generate corresponding reaction forces to meet the requirement for locomotion, while ensuring reliable attachment between the foot and the inclined substrates.

4.1 Synergy between friction and adhesion according to incline

The measured results of the variable \( \alpha^* \) of the feet may imply turning points in the attachment mechanism (friction and adhesion) of the foot (Fig. 4). A planar mechanical model was established to describe the change in the attachment mechanism of the foot in response to the incline, and reveal the contribution of the synergy between the friction and adhesion to a reliable attachment. As it can be seen from the results, the toes abduct to keep the foot away from a horizontal (0° incline) substrate (Figs. 3 and 5(a)). This contact approach not only reduces the number of toes adhering to the substrate to protect the setae [23, 27], but also avoids the unnecessary deployment of adhesion, and thus improves the maneuverability [26, 28]. A gecko’s heel is covered with scales, making it a nonadhesive and nonlubricated system. This indicates that its frictional properties are similar to those of a typical dry solid. Thus, the \( \alpha^* \) \((\alpha^* = \tan^{-1} \mu)\) of the foot is...
determined by the frictional coefficient $\mu$ between the heel and the substrate on a $0^\circ$ incline [29], while the average $\alpha^*$ values for the front and hindfeet are $11^\circ$ and $17^\circ$, respectively (Fig. 4(a) and Table S1 in the ESM). These values are smaller than the values of frictional angles ranging between $19^\circ$ to $28^\circ$ ($\tan^{-1}(0.339-0.551)$) between a snakeskin and a glass surface [30].

A larger driving force is required to overcome the increase in the component of the weight parallel to the substrate ($W_\parallel = W \cdot \sin \theta$) with the increase in the incline, while the component of the weight perpendicular to the substrate ($W_\perp = W \cdot \cos \theta$) decreases [17]. As a result, the frictional force between the heel and the substrate is not sufficient to provide the driving force required to allow the geckos successfully move on an incline, i.e., geckos could not climb inclines of more than $17^\circ$ by relying solely on the friction of their heels, without deploying the adhesion of their toes. Thus, the toes gradually increase their role in making contact with the substrate, and provide a driving force. Correspondingly, the contact between the foot and the substrate enters a concomitant contact state involving both the heel and the toes (Fig. 5(b), in which the forces acting on all the toes are simplified to equivalent forces acting on a single toe in the planar mechanical model).

The $\alpha^*$ value for the foot is determined based on the friction of the heel and the adhesion of the toe (Eq. (4), detailed derivation process refers to the ESM).

$$\alpha^* = \tan^{-1}\left(\frac{\mu \tan \varphi + 1}{r_{h,t} - 1}\right) = \tan^{-1}\left(\frac{\mu + 1}{r_{h,t} - 1}\right),$$

$$r_{h,t} > 0, r_{h,t} \neq 1$$

Here, $\mu$ is the frictional coefficient between the heel and the substrate; $\varphi$ is the angle between the toe and the substrate, and the contribution ratio ($r_{h,t}$) is used to evaluate the contribution of the friction generated by the heel and the adhesion generated by the toes to achieve reliable attachment.

When a gecko moves on a wall and across a ceiling, the angle between each single toe and the substrate is approximately $20^\circ$ [31], that is, not greater than the $\alpha^*$ value for each toe. Herein, $\mu = 0.31 = \tan 17^\circ$, and $\varphi = 20^\circ$, are used to calculate the $\alpha^*$ of the foot for different deployments of friction and adhesion (Fig. 6(a)). The contribution ratio ($r_{h,t}$) value cannot exceed 10, which indicates that the friction of the heel plays a major role in the reliable attachment on $30^\circ$ inclines. However, even a relatively low toe adhesion can notably improve the reliability of the attachment. As the incline continues to increase, the climbing resistance increases with the gravity component parallel to the incline, whereas
the frictional force decreases because of the decrease in the gravity component perpendicular to the incline. The adhesion and shear forces generated by the toes play an important role in ensuring a reliable attachment, since the \( r_{h,t} \) value is only 2.5 when the incline is 60°. A smaller \( r_{h,t} \) value can improve the reliability of the attachment, but the internal moment caused by the difference between the forces acting on the heel and toe increases owing to a decrease in \( r_{h,v} \) which may result in a more difficult movement. On a nearly vertical substrate, the front foot gradually pulls the body close to the substrate and provides adhesion to satisfy the requirement of ensuring movement stability [17], which results in a \( r_{h,t} \) value of less than 1. To-date, the contribution of the friction generated by the heel has been very obscure, but the deployment of the adhesion of the toe plays a key role. The average critical angle of the front foot on a 90° incline is approximately 104° during the movement, similar to the result obtained by Autumn et al. [23]. The hindfoot is required to push away or pull towards the substrate in order to provide either a pushing or an adhesion force, respectively, in order to maintain the dynamic stability during movement [18, 32], thereby leading to an \( r_{h,v} \) value of the foot close to 1. The foot has to withstand a large internal moment during movement.

The adaptability of this concomitant contact state is severely limited by the adhesive capability of the toe and the friction coefficient of the heel. When the incline is larger than 90°, the non-adhesive heel will be pulled away from the substrate by gravity, leaving only the toes attached to the substrate (Fig. 3). Geckos can safely remain attached to an inverted surface by stretching their first and fifth toes into a Y-configuration, thus avoiding the failure of the adhesion of a single toe [17, 31]. The forces acting on each toe of the foot are equally deployed to toes 1 and 2 in a planar mechanical model (Fig. 5(c)). The \( \alpha^* \) of the foot is determined by the combined adhesion between the toes (Eq. (5), detailed derivation process refers to the ESM).

\[
\alpha^* = \tan^{-1}\left(\frac{r_{t,t}/\tan\varphi_1 - 1/\tan\varphi_2}{r_{t,t} + 1}\right)
\]

Here, \( \varphi_1 \) and \( \varphi_2 \) are the angles between the toe 1/2 and the substrate respectively; the contribution ratio \( r_{t,t} \) is defined to evaluate the contribution of the adhesions generated by different toes to the attachment.

The gecko modulates the orientations of the toes and alters the number of attachment toes, resulting in the change in \( r_{t,v} \) thus adapting the motion requirement [17, 31, 33]. As the number of toes adhering to the substrate increases, the load on the adhesive toe will decrease, since this will reduce the risk of detachment of the gecko. Therefore, when the incline is larger than 90°, all five toes of the foot adhere to the substrate (Fig. 3) to share the load, and thus increase the reliability of the attachment. As a result, the locomotive performance worsens, an outcome that is exemplified by a decrease in the speed and stride frequency [26], manifested as a trade-off between locomotive safety and performance. To form a Y-configuration with the five toes, the \( r_{t,t} \) of the foot ranges from 1 to 4. In fact, the scope of adjustment between any adjacent toe is limited by the morphology and structure of the toes [27, 31]. Thus, the \( r_{t,t} \) of the foot is less than 4. On the other hand, if the \( r_{t,t} \) of the foot is too small, one toe would be overtaxed, and the muscles in the foot would have to produce a larger internal moment, which would be detrimental to the locomotion. We found that the load share ratio \( (r_{t,t}) \) was within a range of 2.6 to 3.3 when the inclined angle was larger than 120° (Fig. 6(b)).

The synergy between friction and adhesion, or adhesion and adhesion, in response to an incline illustrates that the synergy between multiple contact mechanisms can achieve a reliable attachment on an incline, while a single-contact mechanism cannot. This increases the animal’s ability to adapt to its environment, but also explains the importance of the reasonable deployment of the adhesion system when faced with the challenge of an inverted incline. This characteristic of the synergy of friction and adhesion was found in the attachment mechanism of tree frogs and insects [10, 34], which embodies the functional convergent evolution in animal survival.

### 4.2 Frictional adhesion of foot

The adhesion system of a gecko consists of hierarchical adhesive units including setae, arrays of setae, lamellae, toes, and feet [24, 35]. Previous research has revealed that the \( \alpha^* \) values of the setae, arrays of setae, and
toes, were not relevant to the exerted force, and that the adhesion force is a function of the shear force [23, 24]. This adhesion capability was characterized as a frictional adhesion [23]. The roles of the feet were changed to respond to an increase in the incline, which led to a significant change in the forces acting on the feet [17]. However, the $\alpha^*$ of the foot did not change significantly when the incline changed from 120° to 180°, i.e., during locomotion the $\alpha^*$ value of the foot was not affected by the forces acting on the foot (Fig. 4(a), Tables S1 and S2 in the ESM). In addition, there is an apparent linear relationship between the critical adhesive force ($F_\perp^*$) and the critical shear force ($F_{||}^*$) acting on the toes (Fig. 7, Table 2). During locomotion, the complicated and accurate synergy between the toes endows a gecko's foot with obvious characteristics of frictional adhesion, i.e., the adhesive force is a function of the shear force. When a gecko climbs steep and inverted inclines, the mean value of $\alpha^*$ of the foot ranges from 114° to 117° (Fig. 4(a)), similar to the $\alpha^*$ values of the arrays of setae and the toes. However, it does not exceed the $\alpha^*$ value ($\alpha^* = 120^\circ = 90^\circ + 30^\circ$) of a single seta, which may imply the extremes values of the $\alpha^*$ of a foot.

4.3 Contact state of foot and adjustment of position

When the action force falls within the scope of the frictional angle to attain frictional self-locking contact, the action force [29] will not affect the contact between the two objects. During locomotion, the reaction force is always in the direction corresponding to the minimum moment of force [36]. Herein, by ignoring the action of the moment of the force, we regard the limbs as being two force bars to allow a discussion of the relationship between the contact state and the adjustment of posture. On a horizontal surface, a gecko gathers up its limbs towards its body, and lifts them to the height COM ($h$) [26]. In turn, this limits the angle between the limbs and the substrate ($\beta$) to values within the range of the foot's $\alpha^*$ values in order to ensure that the foot reliably contacts the substrate without any slip (Fig. 8(a)). While in contact with a steep surface, the gecko reduces the value of $h$ to reduce the risk of overturning, which in turn results in a decrease in $\beta$. Thus, the gecko involves more toes in the deployment of adhesion to enlarge the $\alpha^*$ of the

![Fig. 7](image1.png)

**Fig. 7** The critical normal force ($F_\perp^*$) vs. the critical shear force ($F_{||}^*$) corresponding to the critical angle when the gecko moves on different inclines. The solid points represent the variables for the front foot, whereas, the hollow points represent the variables for the hind foot.

![Fig. 8](image2.png)

**Fig. 8** Contact state of the foot and adjustment of the position in response to the incline. (a) Critical angle ($\alpha^*$) of the foot and the position of the gecko limb on a horizontal substrate. (b) $\alpha^*$ of the foot and the position of the gecko limb on inclines of less than 90°. (c) $\alpha^*$ of the foot and the position of the gecko limb on inclines greater than 90°. The limb is simplified to two parts, limb-I and limb-II, with limb-I being closest to the foot. $\beta$: angle between limb-I and the substrate; $\gamma$: angle between limb-I and limb-II; $h$: distance between COM and the substrate; $F$: force acting on limb-I.
foot. This ensures the reliability of the contact between
the foot and the substrate, while the limbs have plenty
of active space (Fig. 8(b)). However, excessive deploy-
ment of the toes’ adhesion may lead to a decrease in
the locomotive performance [28].

Unlike common friction, where the shear force is
a function of the normal force, a gecko’s foot is
characterized by frictional adhesion, where the adhesive
force is a function of the shear force. This type of
frictional adhesion provides a useful means of precisely
controlling the adhesive force by controlling the shear
force, and enables attachment and detachment to occur
using only minute forces [23]. Geckos skilfully utilize
this frictional adhesion by controlling the angle of
the limb, thereby pulling the foot in such a way to
allow a successful climb onto inverted inclines [18].

When the incline is larger than 90°, the \( h \) value will
be enlarged owing to the effect of gravity, resulting
in an increase in \( \beta \) (Fig. 8(c)). To ensure that the force
acting on the limb falls into the critical scope of
attachment, i.e., \( \beta \) becomes smaller than \( \alpha^* \) (shaded
areas in Fig. 8(c)), the gecko extends its limbs outwards
to decrease the values of \( h \) and \( \beta \) [26]. Similarly, tree
frogs and locusts can attach themselves to inverted
inclined surfaces using this mechanism [6, 37], possibly
owing to the limited adhesiveness in their feet.

5 Conclusion

Geckos rely on the friction between their heels
and substrates to generate the forces required for
movement across a horizontal substrate. On steep
inclines, moderate deployment of toe adhesion enhances
the reliability of attachment of feet. However, excessive
deployment of toe adhesion results in the lower
maneuverability of locomotion, even though the
reliability is enhanced. These characters inspire us in
the design of climbing robots or adhesion systems
with more controllable freedom on adhesive units for
the sake of a favorable trade-off between reliability
and performance of locomotion. The adhesive ability
of a special Y-configuration is limited by the perform-
ance of the adhesive units, resulting in the \( \alpha^* \) value
of the foot being approximately 120°, i.e., geckos could
not hang from an incline at an angle exceeding 120°
with the use of a single foot. Correspondingly, geckos
exploit the characteristic of “frictional adhesion” in
their feet to allow successful climbing on inverted
inclines through cooperation of high-level units, including limbs and body. This fully reflects the fact
that the locomotion system of animals is not a simple
splice of units but their organic integration. Therefore,
when designing a climbing robot or an adhesion
system, we should not blindly pursue the improvement
of one of the units, but rather carefully integrate each
unit into an overall system, while correctly endowing
the basic units with more controllable freedom, in order
to allow a significant improvement in the performance
of a climbing robot or an adhesion system.

Ethical statements

This manuscript adheres to the appropriate reporting
guidelines and community standards for data
availability. This manuscript has not been published
or presented elsewhere in part or in entirety. All
contributing authors are aware of and agree to the
submission of this manuscript. All study participants
provided informed consent, and the study design
was approved by the appropriate ethics review boards.
This study was carried out accordance with the Guide
of Laboratory Animal Management Ordinance of
China. The experimental procedures were approved
by the Jiangsu Association for Laboratory Animal
Science (Jiangsu, China).

Competing interest statement

The authors declare no competing financial interests.

Acknowledgements

We thank Chao Wu, We Li, and Qijun Jiang for assistance
with data collection and manuscript writing. Yi Song
and Lei Cai provided insightful comments on the
manuscript. This work was supported by the National
Natural Science Foundation of China (Grant No.
51435008 to Z.D. and 31601870 to Z.W.) and Natural
Science Foundation of Jiangsu Province, China (Grant
No. SBK2016040649 to Z.W.).

Electronic Supplementary Material: Supplementary
material is available in the online version of this
article at https://doi.org/10.1007/s40544-017-0174-6.
Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


Zhouyi WANG. He obtained his doctor degree in 2015 from College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics (NUAA). He currently works at the College of Astronautics at the NUAA. He was awarded the Shangyin Excellent Mechanical Doctoral Dissertation Award (Bronze) in 2016. His interested research areas include tribology, bionics, and animal kinematics and dynamics.

Zhendong DAI. Male, professor, and tutor of PhD Students, he obtained his doctor degree in 1999 from College of Mechanical and Electrical Engineering, NUAA. He is one of the Chinese delegates of International Institute of Bionic Engineering, an executive member of the Council of Chinese Mechanical Engineering in Tribology, and a member of the Academic Committee, State Key Laboratory of Solid Lubrication. He also is member of editorial board of many magazines such as Journal of Bionic Engineering, International Journal of Vehicle Autonomous System, Tribology, and so on. His research areas include bionics, light material, control of bionics, bio-robots, and biological robots. He has successively presided and participated in many research projects and has published more than 200 papers and gotten more than 20 patents.
Contribution of friction and adhesion to the reliable attachment of a gecko to smooth inclines

Zhouyi WANG1,*, Qiang XING2, Wenbo WANG1, Aihong JI1, Zhendong DAI1,*

1 Institute of Bio-inspired Structure and Surface Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
2 School of Mechanical Engineering, Nantong University, Nantong 226019, China

Supporting information to https://doi.org/10.1007/s12274-017-0174-6

(1) Detailed derivation process of Equation 4 in manuscript

When gecko moves on inclines, the toes gradually increase their role in making contact with the substrate, and provide a driving force. Correspondingly, the contact between the foot and the substrate enters a concomitant contact state involving both the heel and the toes (Fig. S1(a), in which the forces acting on all the toes are simplified to equivalent forces acting on a single toe in the planar mechanical model).

Fig. S1 The planar mechanical model of contact between foot and the inclined substrate. θ: incline of the substrate; v: locomotive direction. (a) A mechanical model of contact when incline does not exceed 90°. The state of contact is a concomitant contact of friction and adhesion. The toe pulls toward the substrate to generate adhesive friction force ($F_{\parallel_{\text{toe}}}$) and adhesion force ($F_{\perp_{\text{toe}}}$), while the heel pushes away the substrate. φ: angle between the toe and the substrate. (b) A mechanical model of contact when the incline exceeds 90°. More toes adhere to the substrate to generate adhesion and friction forces, whereas the heel does not make contact with the substrate. $F_{\perp_{\text{toe1}}}$ and $F_{\perp_{\text{toe2}}}$: the adhesion forces acting on toe1 and toe2; $F_{\parallel_{\text{heel1}}}$ and $F_{\parallel_{\text{heel2}}}$: the adhesive friction forces acting on toe1 and toe2; φ1 and φ2: the angles between the toe1/2 and the substrate.

The $\alpha^*$ value for the foot is determined based on the friction of the heel and the adhesion of the toe (Eq. S1).

$$\alpha^* = \tan^{-1} \left( \frac{F_{\parallel_{\text{heel}}} + F_{\parallel_{\text{toe}}}}{F_{\perp_{\text{heel}}} + F_{\perp_{\text{toe}}}} \right) = \tan^{-1} \left( \frac{\bar{F}_{\parallel_{\text{heel}}} + \bar{F}_{\parallel_{\text{toe}}}}{\bar{F}_{\perp_{\text{heel}}} + \bar{F}_{\perp_{\text{toe}}}} \right)$$

(S1)

Here, assuming that the heel is in the critical state of dry friction, the frictional force ($\bar{F}_{\parallel_{\text{heel}}}$) and the normal
friction force ($\vec{F}_{\perp,\text{heel}}$) can be described by Eq. S2 in accordance to

$$|\vec{F}_{\perp,\text{heel}}| = \mu \tan \varphi |\vec{F}_{\perp,\text{toe}}|$$  \hspace{1cm} (S2)

Previous studies have demonstrated that a gecko’s toe adhesion is direction-dependent, and that the adhesive force is limited by the shear force and the value of $\alpha^*$ of the toe, equal to 25.5° [S1]. If the angle between the toe and the substrate is smaller than the $\alpha^*$ value of the toe, then the adhesion ($\vec{F}_{\perp,\text{toe}}$) and shear forces ($\vec{F}_{\perp,\text{toe}}$) acting on the stretched toe are given by Eq. S3 [S2, S3]. To evaluate the contribution of the friction generated by the heel and the adhesion generated by the toes to achieve a reliable attachment, we define the contribution ratio $r_{p,t}$ (Eq. S4).

$$|\vec{F}_{\perp,\text{heel}}| = |\vec{F}_{\perp,\text{toe}}| \tan \varphi$$  \hspace{1cm} (S3)

$$r_{h,t} = |\vec{F}_{\perp,\text{heel}}| / |\vec{F}_{\perp,\text{toe}}|$$  \hspace{1cm} (S4)

Thus, Eq. S1 can be simplified as Eq. S5, in accordance to

$$\alpha^* = \tan^{-1} \left( \frac{\mu r_{h,t} + 1/\tan \varphi}{r_{h,t} - 1} \right) = \tan^{-1} \left( \frac{\mu + 1/\tan \varphi}{r_{h,t} - 1} \right), \quad r_{h,t} > 0, r_{h,t} \neq 1$$  \hspace{1cm} (S5)

(2) Detailed derivation process of equation 5 in manuscript

The adaptability to an incline of the above concomitant contact state is severely limited by the adhesive capability of the toe and the friction coefficient of the heel. When the incline is larger than 90°, the non-adhesive heel will be pulled away from the substrate by gravity, leaving only the toes attached to the substrate. Geckos can safely remain attached to an inverted surface by stretching their first and fifth toes into a Y-configuration, thus avoiding the failure of the adhesion of a single toe [S4, S5]. The forces acting on each toe of the foot are equally deployed to toes 1 and 2 in a planar mechanical model (Fig. S1(b)). The $\alpha^*$ of the foot is determined by the combined adhesion between the toes (Eq. S6). The contribution ratio ($r_{t,t}$) is defined to evaluate the contribution of the adhesions generated by different toes to the attachment (Eq. S8).

$$\alpha^* = \tan^{-1} \left( \frac{\vec{F}_{\perp,\text{toe1}} + \vec{F}_{\perp,\text{toe2}}}{\vec{F}_{\perp,\text{toe1}} + \vec{F}_{\perp,\text{toe2}}} \right) = \tan^{-1} \left( \frac{\vec{F}_{\perp,\text{toe1}} - \vec{F}_{\perp,\text{toe2}}}{\vec{F}_{\perp,\text{toe1}} + \vec{F}_{\perp,\text{toe2}}} \right)$$  \hspace{1cm} (S6)

$$|\vec{F}_{\perp,\text{toe1}}| = |\vec{F}_{\perp,\text{toe1}}| \tan \varphi_{t}, \quad |\vec{F}_{\perp,\text{toe2}}| = |\vec{F}_{\perp,\text{toe2}}| \tan \varphi_{t}$$  \hspace{1cm} (S7)

$$r_{t,t} = |\vec{F}_{\perp,\text{toe1}}| / |\vec{F}_{\perp,\text{toe2}}|$$  \hspace{1cm} (S8)

Thus, Eq.S6 can be simplified as Eq. 9, in accordance to

$$\alpha^* = \tan^{-1} \left( r_{t,t} \tan \varphi_{t} / r_{t,t} \tan \varphi - 1 \right)$$  \hspace{1cm} (S9)

References


Analysis of the coated and textured ring/liner conjunction based on a thermal mixed lubrication model

Chunxing GU1,3, Xianghui MENG1,3,*, Di ZHANG2

1 State Key Laboratory of Mechanical System and Vibration, Shanghai Jiaotong University, Shanghai 200240, China
2 School of Mechanical Electronic Technology, Shanghai Jianqiao University, Shanghai 201306, China
3 School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai 200240, China

Received: 21 March 2017 / Revised: 07 June 2017 / Accepted: 19 June 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: For the ring/liner conjunction, well-designed surface texturing has been regarded as a potential means to improve its tribological performance, as well as the application of coating. However, so far most researchers focused on the one of these aspects. In this study, the combined effect of coating and texturing on the performance of ring/liner conjunction is numerically investigated. A thermal mixed lubrication model is presented. The effects of the coating’s thermal and mechanical properties on the tribological performance are studied under the cold and warm engine operating conditions. Along with the increasing coating thickness, the effects of the coating’s thermal properties on friction loss are found to be significant, as well as the effects of the coating’s mechanical properties. It is also found that a soft coating with a lower thermal inertia has a greater ability to reduce the friction loss of the textured conjunction.

Keywords: thermal mixed lubrication; surface coating; surface texturing; piston ring; tribological performance

1 Introduction

In recent decades, benefits of surface texturing and coating technology are being explored by the researchers across the globe. On the one hand, surface texturing is regarded as a viable means in the reduction of the friction and wear for different tribological systems [1]. One of the evidences for this is the successful commercial application of surface texturing on the cylinder liners of combustion engines [2]. The development and effect of surface texturing in general applications have been summarized in a lot of review papers [3−7]. It is considered that the introduction of surface texturing can expand the behavior under the hydrodynamic lubrication regime instead of the mixed or boundary lubrication regimes [5, 6, 8]. On the other hand, surface coating has been used to protect contacting surfaces for a long time [9]. The coatings with the friction reduction and anti-wear properties are widely used in the different mechanical systems.

In fact, the addition of coatings was first used in the case of solid-to-solid dry contacts. Over the last few decades, the use of surface coatings has also extended to the case of lubricated contacts [9−11]. It is known that a carefully selected surface coating with the appropriate mechanical and thermal properties can be used to reduce the severity of the contact.

In terms of the ring/liner conjunction, surface texturing has been regarded as a potential means to improve its tribological performance, as well as the application of coating. However, so far most researchers focused on the one of these aspects. In this study, in order to study the combined effect of the coating and texturing on the performance of ring/liner conjunction, a thermal mixed lubrication model is presented. Two kinds of texturing cases are studied, in which the textures are simulated on the ring surface and the liner surface, respectively. Benefit from the presented
model, the effects of the coating’s thermal and mechanical properties on lubrication performance of the textured conjunction would be studied under the cold and warm engine operating conditions. The effect of coating’s thickness would be also discussed.

2 Model description

In this section, the numerical model for the coated and textured contacts developed in this work is described in detail. The established model is based on the thermal mixed lubrication model introduced in the author’s another article [12] for the uncoated ring/liner contacts. In brief, the corresponding numerical model for the coated and textured contacts can be split into several parts: the hydrodynamic lubrication part describing the hydrodynamic flow of the lubricant into the contact, the mechanical interaction considering the asperity contact, and the thermal part describing heat transfer between the lubricant film and the solids (contains ring, liner, and coating).

2.1 Hydrodynamic cavitation model

The overview of the coated and textured ring/liner conjunction is displayed in Fig. 1. As shown in Fig. 1, a layer of coating is added to the surface of ring. As for the coating, it is assumed to be bonded to the substrate (piston ring). The cross-section of the coated and textured piston ring can be considered as a slider moving over a plane supported by the oil film. It is because that the oil-film thickness is much smaller than the cylinder diameter. In addition, it is noteworthy that treating the lubrication of the piston ring as a one-dimensional problem is based on the assumption that the ring is perfectly conforming in the circumferential direction.

For the ring/liner conjunction, the hydrodynamic pressures develop within the lubricant layer between the ring and liner. The source for the hydrodynamic pressure generation is the clearance. Before describing the adopted hydrodynamic cavitation model, the clearance between the ring and the liner should be well characterized. Generally, the clearance, described by oil film thickness, for the conventional ring/liner conjunction can be expressed as follows:

\[
h(x) = h_0 + \frac{4c^2}{b^2} + \delta(x)
\]

where \(c\) is the ring crown height, \(b\) represents the ring axial width, and \(h_0\) denotes the minimum oil film thickness. Moreover, \(\delta(x)\) is the deflection induced by elastic deformation of the substrate. The calculation of \(\delta(x)\) can be obtained according to the previous works [13, 14].

When the piston ring (coated or uncoated) is textured, the corresponding expression is changed as:

\[
h_t(x) = h_0 + 4\delta x^2/b^2 + 0.5h_p(\cos(x'\pi/r_p) + 1) + \delta(x) \quad x \in \Omega
\]

where \(h_p\) represents the texture depth, \(r_p\) is the radius of texture, \(\Omega\) denotes the area occupied by the textures. It is noteworthy that the terms \(x\) and \(x'\) in the above equations are different in the Cartesian coordinates. The detailed description about \(x\) and \(x'\) can be found in the work of Gu et al. [15].

Different from the expression for the textured ring, the film thickness of the textured liner can be described as [1, 16]:

\[
h_t(x,t) = h_0 + 4\delta x^2/b^2 + 0.5h_p(\cos(x'\pi/r_p) + 1) + \delta(x) \quad x - Ut \in \Omega
\]

where \(U\) is the sliding speed.

Moreover, the prediction of the hydrodynamic pressure of the ring/liner conjunction can be carried out by the well-known JFO (Jacobsson-Floberg-Olsson) model [17, 18] coupled with a number of correction factors. According to the basic idea of the JFO model,
it is described that the mass flow is a function of the hydrodynamic pressure $p$ and the cavity fraction $\theta$. The involved correction factors are called the flow factors and defined by Patir and Cheng [19]. The corresponding equations can be given as follows:

$$
\frac{\partial}{\partial x} \left( \phi \frac{\rho h}{\mu} \frac{\partial p}{\partial x} \right) = 6 U \phi \frac{\partial [(1 - \theta) \rho h]}{\partial x} + 6 U \sigma \frac{\partial [(1 - \theta) \rho \phi]}{\partial x} + 12 \phi \frac{\partial [(1 - \theta) \rho h]}{\partial t}
$$

(4)

$$
p + \theta - \sqrt{\theta^2 + \theta^2} = 0
$$

(5)

where $\mu$ is the lubricant viscosity, $\rho$ symbolizes the lubricant density, $p$ denotes the hydrodynamic pressure, and $\theta$ represents the cavity fraction. It is noteworthy that the cavity fraction $\theta$ is related to the cavitation. The term $\theta$ is influenced by the hydrodynamic pressure $p$. If there is no cavitation then the fraction is zero ($\theta = 0$) and the pressure is greater than the cavitation pressure ($p > p_{cav}$). When cavitation is present, the fraction is positive ($\theta > 0$) and the pressure is assumed to be equal to the cavitation pressure ($p = p_{cav}$). A value of $p_{cav} = 0$ is assumed as the cavitation pressure in this paper. Parameters $\phi_s$, $\phi_c$, and $\phi_r$ are the flow factors. These flow factors can be obtained according to the works of Patir et al. [19, 20]. $\sigma$ is the composite roughness for the ring and the liner, which can be calculated by the expression $\sigma = \sqrt{\sigma_u^2 + \sigma_l^2}$. In this expression, $\sigma_u$ and $\sigma_l$ are the roughness of the ring and the roughness of the liner, respectively. It should be mentioned that the application of constraint described in Eq. (5) could turn the hydrodynamic model into the unconstrained system of nonlinear algebraic equations. The involved calculation can benefit from a direct solver using the Fischer-Burmeister-Newton-Schur approach [21].

The lubricant rheology, involving its density and viscosity, influences the lubrication regimes. Based on the equation reported by Vogel [22] as well as the Roelands equation [23], the lubricant viscosity can be calculated by:

$$
\mu = \left[ a_v \exp \left( \frac{T_1}{T_2 + T} \right) \right] \exp \left[ \ln \left[ a_v \exp \left( \frac{T_1}{T_2 + T} \right) \right] \right] + 9.67 \left[ (1 + 5.5 \times 10^9 \rho) - 1 \right]
$$

(6)

where $T$ is the lubricant temperature (whose unit is $K$). Meanwhile, $a_v$, $T_1$, $T_2$, and $b_0$ are the corresponding correlation parameters. $b_0$ is usually considered to be around 0.68. In this way, the influence of pressure and temperature on lubricant viscosity can be obtained.

In the other hand, the influence of pressure and temperature on the lubricant density can be characterized as [12, 24]:

$$
\rho = \rho_0 \left( 1 + \frac{0.6 p}{1 + 1.7 p} \right) \left[ 1 - \beta_p (T - T_0) \right]
$$

(7)

where $\rho_0$ is the lubricant density under the atmospheric pressure and temperature of $T_0$. $\beta_p$ is the thermal expansion coefficient. The detailed lubricant properties are given in Table 1.

### 2.2 Asperity contact model

In this work, the asperity contact force is obtained by the classic GT (Greenwood and Tripp) model [25]:

$$
p_{as} = \frac{16 \sqrt{2 \pi}}{15} \left( \eta \beta \sigma \right)^2 \sqrt{\frac{\sigma}{\beta}} E' \left( z - \lambda \right)^{125} \phi^*(z) dz
$$

(8)

where $\lambda$ is the Strubeck oil film ratio, which is the ratio of the mean film thickness $h$ to the composite roughness $\sigma$; $\eta \beta \sigma$ is known as the Tabor’s roughness parameter, while, $\sigma / \beta$ is a measure of a typical asperity slope [26]. $E'$ is the composite effective elasticity modulus:

$$
\frac{1}{E'} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}
$$

(9)

where $E_1$ and $E_2$ are the elastic modulus of the slider and the plane, respectively; $\nu_1$ and $\nu_2$ are the Poisson’s

### Table 1  Lubricant properties at different lubrication temperatures.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lubricant density at 313.15 K (kg·m⁻³)</td>
<td>850</td>
</tr>
<tr>
<td>Lubricant density at 393.15 K (kg·m⁻³)</td>
<td>806.48</td>
</tr>
<tr>
<td>Lubricant viscosity at 313.15 K (mPa·s)</td>
<td>35.13</td>
</tr>
<tr>
<td>Lubricant viscosity at 393.15 K (mPa·s)</td>
<td>7.71</td>
</tr>
<tr>
<td>Thermal conductivity at 313.15 K (W·m⁻¹·K⁻¹)</td>
<td>0.145</td>
</tr>
<tr>
<td>Thermal conductivity at 393.15 K (W·m⁻¹·K⁻¹)</td>
<td>0.225</td>
</tr>
<tr>
<td>Specific heat capacity at 313.15 K (J·kg⁻¹·K⁻¹)</td>
<td>1.968</td>
</tr>
<tr>
<td>Specific heat capacity at 393.15 K (J·kg⁻¹·K⁻¹)</td>
<td>2360</td>
</tr>
<tr>
<td>Thermal expansion coefficient (K⁻¹)</td>
<td>6.4×10⁻⁴</td>
</tr>
</tbody>
</table>
ratio of the slider and plane, respectively. \( \phi'(z) \) is the standardized height distribution. The asperity contact model is linked with the hydrodynamic model based on the load-sharing concept [27].

It is interesting to notice that two different regions are present in the rough textured surfaces. One is the untextured region. The texture features are not included in this region. Another is the textured region. It includes the texture features. Both the untextured region and the textured region are rough. In fact, when the textures are introduced, the roughness in the center and near the textures may be influenced. However, because the clearance (between two contact surfaces) in the textured region is usually more than four times the composite roughness \( \lambda > 4 \), the roughness in the textured region would cause a limited or negligible effect on the lubricant flow. Therefore, according to the works of Patir and Cheng [19], the roughness in the textured region of the rough textured surfaces is not important. Similarly, the asperity contact happened in the textured region is also limited or negligible for the presence of the large clearance [25]. For simplification purpose, it is assumed that the involved asperities of the untextured region obey the Gaussian distribution. Meanwhile, the surface roughness of the untextured region is assumed to characterize the rough textured surface. In this study, the value of \( R_q \) is obtained based on the untextured region, which is used for calculating the composite roughness \( \sigma \) as well as the Stribeck oil film ratio \( \lambda \).

### 2.3 Thermal model

The thermal part is based on the heat transfer equation applied to the lubricant and contacting solids. Generally, by considering the viscous dissipation heating effect in the oil film and the roughness effect induced by asperity contact, the two-dimensional energy equation of lubricant oil can be expressed as [12]:

\[
\frac{k_p}{\varepsilon^2} \frac{\partial^2 T}{\partial z^2} = \rho c_p \left( \frac{u}{\varepsilon x} + \frac{\partial T}{\partial t} \right) - \mu \left( \frac{\partial u}{\partial z} \right)^2 - \beta_r T \left( \frac{\partial p}{\partial x} - \kappa_{as} p_{as} \frac{\partial T}{\partial z} \right)
\]  

(10)

where \( T \) is the lubricant temperature, \( u \) denotes the fluid velocities. The term \( z \) is the coordinates across the clearance. Parameters \( k_p, \beta_r, \) and \( c_p \) are the thermal conductivity, thermal expansivity, and specific heat of the lubricant, respectively. It is noteworthy that the last term in the right hand of Eq. (10) is used to consider the contribution of asperity contact to heat generation [28]. \( \kappa_{as} \) is the asperity friction coefficient. \( p_{as} \) represents the asperity contact pressure. In addition, the term \( u \) is the fluid velocities in the \( x \) directions. It can be obtained by:

\[
\rho \left( \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} \right) = - \frac{\partial p}{\partial x} + \mu \left( \frac{\partial^2 u}{\partial x^2} \right)
\]  

(11)

In Eq. (10), the left-hand side represents conduction, while the terms on the right hand side represent viscous dissipation, compression heating, convection, and the contribution of asperity contact to heat generation, respectively. It should be noticed that the influence of the asperities on the thermal conduction is not considered in this paper. The limitation should be acknowledged.

In the solving of Eq. (10), the surface temperatures of the ring and the liner are particularly important. According to the work of Yang et al. [29], these surface temperatures can be obtained by the following equations from (12)–(17). Equations (12), (14), and (17) are the heat transfer equations of the ring, the coating, and the liner, respectively. Equation (13) is the heat flux continuity condition on the interface between the ring and the coating. Equation (15) is for the interface between the coating and the lubricant, while, Eq. (16) is for the interface between the lubricant and the liner. These equations are listed as follows:

\[
k_r \left( \frac{\partial^2 T}{\partial z_r^2} \right) = \rho e r c_r \left( \frac{U_r}{\varepsilon x} \frac{\partial T}{\partial x} \right) \]  

(12)

\[
k_r \frac{\partial T}{\partial z_r} = k_c \frac{\partial T}{\partial z_c} \]  

(13)

\[
k_c \left( \frac{\partial^2 T}{\partial z_c^2} \right) = \rho c_c \left( \frac{U_c}{\varepsilon x} \frac{\partial T}{\partial x} \right) \]  

(14)

\[
k_c \frac{\partial T}{\partial z_c} = k_l \frac{\partial T}{\partial z_l} \]  

(15)

\[
k_l \frac{\partial^2 T}{\partial z_l^2} = \rho c_l \left( \frac{U_l}{\varepsilon x} \frac{\partial T}{\partial x} \right) \]  

(17)
where $U_R$ is the velocity of the ring, $U_L$ is the velocity of the liner. Since the coating is bonded to the ring, the velocity of the coating $U_C$ is equal to the velocity of the ring $U_R$. Parameters $k_R, \rho_R$, and $c_R$ are the thermal conductivity, density, and specific heat for the ring, and parameters $k_L, \rho_L$, and $c_L$ are the thermal conductivity, density, and specific heat for the liner. Parameters $k_C, \rho_C$, and $c_C$ are the thermal conductivity, density, and specific heat for the coating.

Coordinates $R_{zz}$ and $L_z$ have the same direction as coordinate $z$. For the way about the dimensionless treatment of $R_z$, $C_z$, $L_z$, and $z$, the interested readers can refer to the work of Yang et al. [29].

In addition, the thermal boundary conditions should be adopted as follows:

- $T = T_{in}$ at inflow on inlet side;
- $dT/dx = 0$ at outflow on inlet side;
- $dT/dx = 0$ at the outlet side;
- $T = T_L$ at the liner surface;
- $T = T_R$ at the ring surface.

Besides, as shown in Fig. 2, the inlet lubricant temperature, $T_{in}$, can be calculated as follows [30, 31]:

$$T_{in} = \frac{T_L U_L + T_C U_C}{U_L + U_R}$$  \hspace{1cm} (18)

where $T_L$ and $T_C$ are the initial surface temperatures of the liner and coating. However, in the simulation of the coated and textured ring/liner system, the coated and textured ring is assumed to be stationary and the liner is assumed to undergo relative motion. Thus, according to Eq. (18), due to the convective thermal flux, $T_{in}$ is equal to $T_L$ at the inlet. Therefore, in order to solve the thermal model, the liner temperature should be known a priori.

### 2.4 Performance parameters

The total friction force is evaluated as follows:

$$\tau = -\mu \frac{U}{h}(\phi + \phi_h) + \phi_p \frac{h \partial p}{2 c x}$$  \hspace{1cm} (19)

$$F_{fc,visco} = l_x \int \tau (1 - \theta) dx$$  \hspace{1cm} (20)

$$F_{fc,boun} = \kappa_{ap} l_x \int p_{ap} dx$$  \hspace{1cm} (21)

$$F_{fc,total} = F_{fc,visco} + F_{fc,boun}$$  \hspace{1cm} (22)

where $\tau$ represents the viscous shear of the lubricant. $F_{fc,visco}$ is the viscous friction. $F_{fc,boun}$ denotes the asperity contact friction. $F_{fc,total}$ symbolizes the total friction. The terms $\phi$, $\phi_h$, and $\phi_p$ are the friction-induced flow factors. $l_x$ represents the circumference of the piston ring.

Moreover, the friction mean effective pressure (FMEP) parameter can be used to evaluate the engine friction loss. FMEP is the friction work normalized by engine displacement. It reads:

$$\text{FMEP} = \frac{\int_{cycle} F_{fc,total} dl_s}{V_d}$$  \hspace{1cm} (23)

where $V_d$ is the engine displacement, $l_s$ is the stroke, $dl_s$ is the distance along the direction of the sliding stroke in each time step.

### 3 Results and discussion

In this section, a Gasoline engine at 2,000 rpm-full load is considered. Both the cold and warm engine-operating conditions are taken into account here in order to investigate the tribological benefits in terms of surface texturing and coating. In other words, the initial liner temperatures of 313.15 K and 393.15 K are adopted in the simulations, respectively. The former represents the cold engine condition while the latter is on behave of the normal operating (warm) condition. Tables 2–4 provide the necessary data for the current analysis. Among them, Table 4 shows the mechanical/thermal properties of coatings.
Table 2  Engine data.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed (rpm)</td>
<td>2,000</td>
</tr>
<tr>
<td>Stroke (mm)</td>
<td>90</td>
</tr>
<tr>
<td>Diameter of cylinder liner (mm)</td>
<td>84</td>
</tr>
<tr>
<td>Axial width of the ring, (b) ((\mu)m)</td>
<td>1,000</td>
</tr>
<tr>
<td>Height of ring crown, (c) ((\mu)m)</td>
<td>8</td>
</tr>
<tr>
<td>Tension of the top ring (N)</td>
<td>20</td>
</tr>
<tr>
<td>Boundary friction coefficient, (\kappa_{\text{cap}})</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 3  The mechanical/thermal/surface properties for ring and liner.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity modulus of ring, (E_r) (GPa)</td>
<td>250</td>
</tr>
<tr>
<td>Poisson’s ratio of ring, (\nu_r)</td>
<td>0.3</td>
</tr>
<tr>
<td>Density of ring, (\rho_r) (kg·m(^{-3}))</td>
<td>7,700</td>
</tr>
<tr>
<td>Thermal conductivity of ring, (k_r) (W·m(^{-1})·K(^{-1}))</td>
<td>25</td>
</tr>
<tr>
<td>Specific heat capacity of ring, (c_r) (J·kg(^{-1})·K(^{-1}))</td>
<td>460</td>
</tr>
<tr>
<td>(R_q) for ring, (\sigma_r) ((\mu)m)</td>
<td>0.2</td>
</tr>
<tr>
<td>Elasticity modulus for liner, (E_L) (GPa)</td>
<td>120</td>
</tr>
<tr>
<td>Poisson’s ratio for liner, (\nu_L)</td>
<td>0.3</td>
</tr>
<tr>
<td>Density for liner, (\rho_L) (kg·m(^{-3}))</td>
<td>7,200</td>
</tr>
<tr>
<td>Thermal conductivity for liner, (k_L) (W·m(^{-1})·K(^{-1}))</td>
<td>55</td>
</tr>
<tr>
<td>Specific heat capacity for liner, (c_L) (J·kg(^{-1})·K(^{-1}))</td>
<td>460</td>
</tr>
<tr>
<td>(R_q) for liner, (\sigma_L) ((\mu)m)</td>
<td>0.6</td>
</tr>
<tr>
<td>Roughness parameter ((\eta/\beta))</td>
<td>0.04</td>
</tr>
<tr>
<td>Measure of asperity gradient ((\sigma/\beta))</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 4  The mechanical/thermal properties for coatings.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity modulus of ring, (E_c) (GPa)</td>
<td>62.5; 125; 250; 500</td>
</tr>
<tr>
<td>Poisson’s ratio of ring, (\nu_c)</td>
<td>0.3</td>
</tr>
<tr>
<td>Density of ring, (\rho_c) (kg·m(^{-3}))</td>
<td>3,500; 10,000</td>
</tr>
<tr>
<td>Thermal conductivity of ring, (k_c) (W·m(^{-1})·K(^{-1}))</td>
<td>5; 90</td>
</tr>
<tr>
<td>Specific heat capacity of ring, (c_c) (J·kg(^{-1})·K(^{-1}))</td>
<td>200; 1,000</td>
</tr>
<tr>
<td>Coating thickness, (d_c) ((\mu)m)</td>
<td>10; 20; 40; 80</td>
</tr>
</tbody>
</table>

In addition, the variation of piston sliding speed with the crank angle is presented in Fig. 3, as well as the cylinder pressure variation. They are used as the pivotal inputs in the related simulations. In terms of the textured systems, the adopted textures are 60 \(\mu\)m in breadth and 3 \(\mu\)m in height. When the piston ring is textured, four textures are evenly distributed on the ring surface. When the textured liner system is studied, the spaces (one texture per 250 \(\mu\)m) between every two textures are the same.

3.1 Validation and mesh

At first, one case is used to test the correctness and accuracy of the adopted model. The adopted simulation conditions are consistent with the published results from Bertocchi et al. [32]. The comparison results involving the hydrodynamic pressure are shown in Fig. 4. Only a difference of 0.65% in peak pressure between current model and the one presented by Bertocchi et al. [32] is detected. In other words, the result obtained by the present model matches well with the published results.
with the result available in the literature.

In addition, it is noteworthy that the choice of mesh size and time step is critical for the simulation of the textured system. In the author’s another paper [15], the influences of mesh size on the convergence of results have been studied, as well as the time step. It was found that the numerical errors could be ignored when 1,000 mesh and 0.05 degree per step were adopted in the simulations. Since the cases studied in this paper are similar to those given in previous paper [15], the same time step and mesh size would be adopted in this paper to keep the accuracy of the calculation.

3.2 Effect of coating mechanical properties

In order to isolate the effects of mechanical properties, the thermal properties of coatings are taken to be those of ring. Figure 5 shows the effect of coating’s mechanical properties on the average temperature rise of lubricant when the untextured conjunction is under the cold engine condition. The coatings with different elasticity modules are employed in the corresponding simulations. Meanwhile, it should be pointed out that the average temperature rises of lubricant are the results of convention and conduction. The average viscosity value is calculated by adding the viscosity values in different nodes together and dividing by the total number of nodes. It appears that the average temperature rise for the lubricant is obvious at the middle of each stroke, but the value of average temperature rise is limited in each dead center. Moreover, as shown in Fig. 5, the average temperature rises at the power stroke are more obvious than the results at other strokes. The average temperature rise of lubricant is influenced by the applied load. In addition, according to the results from the conjunctions with different elasticity modulus of coatings, it can be found that the soft coatings can make the lubricant obtain a larger temperature rise than the hard ones. The temperature rise of lubricant may considerably alter the properties of the lubricant. As a result, different elasticity modulus of coatings would result in the various tribological performances of ring/liner conjunctions.

In order to study the performance of the coated and textured conjunction, the temperature distributions for the untextured and textured conjunctions are displayed in Fig. 6. The crank angle is 380 degree and the bonded coating's elasticity modulus is 62.5 GPa. As shown in Fig. 6, it appears that the temperature rise mainly occurs in the middle region of the clearance. It is found that the lubricant temperature distribution for the smooth and the textured ring conjunctions are almost the same. However, it is interesting to notice that the temperature distribution for the textured liner conjunction is different from those for the smooth and the textured ring conjunctions. It should be noticed that simulating the textured liner system should always consider the effect of the transient terms. Owing to the movement of the texture features, the clearance between the untextured ring and the textured liner is changing all the time [15]. As shown in Fig. 6, the lubricant temperature peak for the textured liner conjunction is larger than these found in the smooth and the textured ring conjunctions. It seems that the viscous dissipation heating effect may be improved by the moving textures for the textured liner conjunction. The increased temperature inside the film would result in the reduced lubricant viscosity in the same region. As a result, friction would be reduced.

In order to study the effect of the coat’s mechanical properties on the friction loss, the FMEP results for the untextured and textured systems under the cold engine operating condition are summarized in Fig. 7.
Fig. 6 The temperature distribution in oil film between the coated ring and the liner under the cold engine condition when the crank angle is 380 degree and the coating’s elasticity modulus is 62.5 GPa: (a) for the smooth conjunction; (b) for the textured ring conjunction; (c) for the textured liner conjunction.

Fig. 7 The variation of FMEP results for the untextured and textured conjunctions under the cold engine condition when different Elasticity modulus of coatings are adopted.

It can be found that different elastic modulus of coatings would result in different FMEP results. The friction loss of ring/liner conjunctions can be slightly affected by the mechanical properties of coatings. For the smooth conjunction, the corresponding FMEP value is smallest when the elasticity modulus of coating is 62.5 GPa. It seems that the application of the soft coating can improve the tribological performance of ring/liner conjunction. The reason for the phenomenon that a soft coating can lead to reduced friction compared to a hard one can be explained as follows: In the mixed lubrication regime, the applied load is shared by the hydrodynamic support and asperity contact load. As mentioned in Section 2, the asperity contact load can be obtained by Eq. (8). For a soft coating, the composite effective elasticity modulus is smaller than that for a hard coating. It would lead to different asperity contact loads. In order to achieve the load balance, the oil film would be adjusted and the hydrodynamic support would be changed. Benefit from the reduced asperity contact load, the viscous friction would domain and, as a consequence, the friction loss would be reduced. Hard coatings have the exact opposite effect and lead to the increased
friction. Furthermore, by comparing the results from the smooth conjunction with the textured conjunction results, it seems that the effect of texturing can be enhanced by the application of the soft coating. Meanwhile, when the textures are treated on the liner and the ring surface is coated with a soft coating (whose elasticity modulus is 62.5 GPa), it is observed that the friction loss of ring/liner conjunction can be significantly reduced. It is also found that the textured liner system has the better tribological properties than the textured ring system, when the elasticity modules of coating are the same. As reported by Gherca et al. [33], a textured moving surface can contribute to the generation of hydrodynamic lift on the basis of a squeeze effect. For more information about the difference between the stationary texturing case (or say the textured ring system) and moving texturing case (or say the textured liner system), the interested readers can refer to the works of Medina et al. [34] and Gherca et al. [33].

Figure 8 shows the simulation results under the warm engine operating condition. As shown in Fig. 8, the variations of the FMEP results for the untextured and textured conjunctions are summarized. Similar to the results under the cold engine condition presented in Fig. 7, it is also found that the friction loss under the warm engine operating condition can be reduced by using the soft coating. Obviously, the combination of soft coating and introducing the textures on the liner leads to the greatest friction decrease.

In addition, it appears that the textured liner system has the better friction reduction effect than the textured ring system, when the warm engine operating condition is considered. In fact, when textures are introduced on the liner surface, the wear of the liner would be smaller and more evenly distributed than that of ring [16]. The textures can be preserved in their shapes for a long operating time. Meanwhile, regarding the textured liner system, it allows one to optimize the texture that contacts the ring independently for each instant of the engine cycle [16].

3.3 Effect of coating thermal properties

In order to isolate the effects of the coating’s thermal properties, the mechanical properties of the coatings are taken to be those of the ring. Two categories of coatings are considered based on their thermal inertia $I = \sqrt{\rho C}$. Both the low ($\rho = 3,500 \text{ kg} \cdot \text{m}^{-3}$, $C = 200 \text{ J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$) and high ($\rho = 10,000 \text{ kg} \cdot \text{m}^{-3}$, $C = 90 \text{ W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$) thermal inertia coatings are simulated, whose values are smaller than or greater than the thermal inertia of the ring, respectively. Figure 9 shows the effect of coating’s thickness on the FMEP results, when the untextured ring/liner conjunctions are simulated under the cold engine operating condition. As shown in Fig. 9, it is found that the increase of the thickness for the low thermal inertia coatings leads to a more pronounced decrease in FMEP results. On the other hand, it is found that the
friction is slightly changed with the thickness of high thermal inertia coatings. Notice that the friction loss change induced by the thickness of high thermal inertia coatings is quite limited (about 0.01 kpa). Figure 10 shows the variations of average lubricant temperature rises when the untextured conjunctions are coated with the high or low thermal inertia coatings. For the high thermal inertia coating, the lubricant temperature rise is decreased with the increase of the coating thickness. In terms of the low thermal inertia coating, it is observed that the average temperature rise for the lubricant increases with the increasing coating thickness. As discussed in the work of Habchi [9], low thermal inertia surface coatings can act as insulators, leading to a localized increase in the lubricant’s temperature. The lubricant viscosity is influenced by the lubricant temperature. The large lubricant temperature rise would lead to the low lubricant viscosity, which would result in the reduced friction.

Moreover, the FMEP results for the untextured and textured systems are summarized in Fig. 11. It is used to find the combined effect of texturing and coating on the friction loss. The adopted coating thickness is 20 μm. It can be found that a significant reduction of friction can be obtained when texturing is on the liner and the low thermal inertia coating is used. Obviously, the combination of the low thermal inertia coating and treating textures on the liner leads to the greatest friction decrease whereas the application of the high thermal inertia on the untextured conjunction leads to the greatest friction increase.

4 Conclusions

In this paper, a thermal mixed lubrication model was presented for analyzing the coated and textured ring/liner system. The combined effects of the coating and texturing under the cold and warm engine conditions were studied. The following conclusions can be drawn:

It is found that the mechanical and thermal properties of coatings would affect the tribological performance. However, it seems that the influences of coatings are relative small. Compared with the cases under the warm engine condition, the friction reduction induced by the coating is more obvious when the cold engine condition is studied.

Meanwhile, surface texturing can improve the tribological performance of ring/liner conjunction, whether under the cold or warm engine operating conditions. In particular, the effect of texturing can be

![Fig. 10](image1.png) The variations of average temperature rises for lubricant when different thicknesses of coatings are adopted: (a) for the high thermal inertia coating; (b) for the low thermal inertia coating.

![Fig. 11](image2.png) The effect of coating's thermal inertia on FMEP results for the untextured or textured conjunctions.
enhanced by the application of the soft coating with the low thermal inertia.

Besides, it is noteworthy that the application of the coating can also improve other properties, involving the surface roughness, the boundary friction coefficient, etc. However, these aspects are not considered in this study. The experimental testing would be a better alternative to study these aspects. The limitations of the approach should be acknowledged.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (Nos. 51375300 and 51575342), the Research Project of State Key Laboratory of Mechanical System and Vibration (No. MSVZD201701) for supporting this research.

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


Chunxing GU. He received his bachelor and master degrees in mechanical engineering in 2011 and 2014 from Jiangsu University, Zhenjiang, China. He is studying for his Ph.D. degree at the Mechanical Engineering School of Shanghai Jiaotong University, Shanghai, China. His research interests include surface texturing and tribology.

Xianghui MENG. He received his bachelor and master degrees in power machinery and Engineering from Xi’an Jiaotong University, China, in 1995 and 1999, respectively. He received his PhD degree in mechanical engineering from Shanghai Jiaotong University, China, in 2006. His current position is an associate professor and doctorial supervisor at School of Mechanical Engineering, Shanghai Jiaotong University. His research areas cover the tribology of internal combustion engines, low friction design, and wear mechanism.
Study on frictional behavior of carbon nanotube thin films with respect to surface condition

Youn-Hoo HWANG, Byung-Soo MYUNG, Hyun-Joon KIM*

Department of Precision Mechanical Engineering, Kyungpook National University, 2559, Gyeongsang-daero, Sangju, Republic of Korea

Received: 16 May 2017 / Revised: 03 July 2017 / Accepted: 04 July 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: In this work, tribological characteristics of thin films composed of entangled carbon nanotubes (CNTs) were investigated. The surface roughness of CNT thin films fabricated via a dip-coating process was controlled by squeezing during the process with an applied normal force ranging from 0 to 5 kgf. Raman spectra and scanning electron microscopy (SEM) images of the thin films were obtained to estimate the influence of the squeezing process on the crystallinity of the CNTs. The analysis revealed that squeezing could reduce surface roughness, while preserving the crystallinity of the CNTs. Moreover, the surface energy of the cover glass used to press the CNT thin film was found to be the critical factor controlling surface roughness. A micro-tribometer and macro-tribometer were used to assess the tribological characteristics of the CNT thin film. The results of the tribotest exhibited a correlation between the friction coefficient and surface roughness. Dramatic changes in friction coefficient could be observed in the micro-tribotest, while changes in friction coefficient in the macro-tribotest were not significant.

Keywords: carbon nanotubes; friction; surface roughness; surface energy; squeezing process; UV irradiation

1 Introduction

Carbon nanotubes (CNTs) have attracted significant interest for decades because of their superb mechanical, electrical, and chemical characteristics, such as high tensile strength and elastic modulus along the longitudinal direction, and good elastic and thermal conductivity properties [1–4]. Based on these unique properties, various CNT applications have been suggested, such as probe tips for atomic force microscopy (AFM) and scanning tunneling microscopy (STM) [5, 6], rotational actuators [7], field emission devices [8–10], electric motor brushes [11, 12], and chemical sensors [13, 14]. Owing to their small dimensions, CNTs are regarded as suitable materials for exploitation in several precision industries such as semi-conductors, batteries, and display devices [15, 16]. Some researchers have also focused on the ability of the outstanding mechanical properties of CNTs to improve durability and tribological characteristics. There have been several attempts to exploit CNTs in reinforcing polymers [17, 18] or metal composites [19–21]. The utilization of CNTs as additives in lubricants has also been considered by several researchers [22, 23].

Since CNTs exhibit lubricious characteristics and outstanding wear resistance, numerous studies have investigated the fundamental behavior of CNTs for tribological applications. Falvo et al. [24] conducted an experiment to manipulate a single CNT on a mica surface. In the study, they discovered stick-slip motion of the CNT and investigated shear stress at the interface between the CNT and the mica surface. Following the study, Falvo et al. [25] also revealed that sliding and rolling of a CNT are attributed to commensurability of nano-scale contact, suggesting the importance of CNT alignment on tribological characteristics. Kim and Kim [26] studied the frictional behavior of CNTs through molecular dynamics simulation, discovering

* Corresponding author: Hyun-Joon KIM, E-mail: hjoonkim@knu.ac.kr
that the frictional force of CNTs depends on the interaction between adjacent nanotubes, CNT orientation, and elastic deformation.

There have been several attempts to exploit CNTs as a surface coating to improve tribological characteristics. Kim et al. [27] revealed the potential of CNTs to minimize surface wear when CNTs were combined with a noble metal that exhibited low shear strength. In the study, CNTs and Ag were coated on a Si surface to form a composite with different thicknesses. The friction coefficient and wear rate of the composite were examined using a pin-on-reciprocating tribometer under 20 mN of normal force. They emphasized that “anchoring” between Ag and CNTs ensured that the composite rarely wore out, even though the individual mechanical properties of the CNTs and Ag were insufficient. Kinoshita et al. [28] reported that a vertically aligned CNT film exhibits high friction coefficients of up to 2.0. Although they investigated vertically aligned CNT films, which show almost zero adhesion force, an extraordinarily high frictional force was reported. They deduced that bent nanotubes corresponding to relative motion of the counter surface demonstrate high repulsive force for elastic recovery and are the cause of the high frictional force.

Although numerous tribological studies have been conducted over decades, there is still a lack of understanding with regard to the complicated behavior of CNTs under various environmental conditions. In addition, the majority of studies on the tribological characteristics of CNTs have been concerned with composites and lubricant additives. Since a CNT film usually consists of numerous fibers, which are entangled together and have various dimensions, tribological characteristics may vary significantly with environmental conditions. Therefore, intensive study on the tribological behavior of CNT thin films is required to enhance fundamental understanding. In this work, the influence of the surface morphology of CNTs on the frictional behavior of CNT thin films was experimentally investigated at the micro-scale.

2 Experimental setup

2.1 Preparation of CNT thin films

A soda-lime glass slide (Marienfeld, Germany) was chosen as the CNT thin film substrate as it provides a relatively smooth and uniform surface. The glass slide was carefully rinsed with acetone, ethanol, and deionized (DI) water using an ultrasonicator for 10 min each to remove any surface contamination. The surface energy of the cleaned glass slide was increased by exposing it to UV light for an hour to strengthen the interface bonding between the CNTs and the glass slide. The water contact angle was measured to confirm the change in surface energy on UV irradiation, as shown in Fig. 1. When the glass slide is irradiated with UV light, the water contact angle is about 9°, while that of the glass slide without UV treatment is 27°. The CNT thin film was coated on the glass slide by a dip-coating method, as shown in Fig. 2. The glass slide was immersed in a CNT suspension (JEIO Co. Ltd., Republic of Korea), which consisted of 95 wt% of isopropyl alcohol (IPA), 3 wt% of multi-walled CNTs, and 2 wt% of surfactants, for a few tens of seconds and withdrawn from the CNT suspension with a velocity of 0.375 mm/s. Coating parameters such as withdrawal speed may influence the resulting mechanical characteristics of the thin film. It has been well documented that the withdrawal speed affects film thickness [29]. Therefore, the dip-coating conditions were carefully optimized. The immersion time was optimized to stabilize the
suspension state. The withdrawal speed was optimized for the target CNT thin film thickness. We measured the thickness of the CNT thin film prepared under the optimized withdrawal rate of 0.375 mm/s to be approximately 3 μm. After the dipping process, both sides of the CNT-coated glass slide were held in dry air to evaporate the IPA. The tribological characteristics of the front face thin film of the CNT-coated glass slide were assessed. The CNT coating on the back face of the glass slide was not removed; rather, the back face was covered with cellophane tape to minimize any uncertainty arising from the CNT thin film on the back face.

Since the CNT thin film consists of an entangled form of numerous nanotubes, occasional surface nanotube protrusions are inevitable, increasing the surface roughness compared to that of the bare glass slide. As the surface roughness of the CNT thin film might influence the tribological characteristics, a squeezing method (shown in Fig. 2(c)) to control surface roughness was suggested in this work. In this process, a glass plate with identical dimensions to the glass substrate was placed over the CNT thin film and a dead weight ranging from 0.5 to 5 kgf was placed on the assembly for 30 min to apply relatively uniform pressure and maintain stable surface conditions. In addition, UV treatment was applied to the cover-glass plate to investigate the influence of surface energy on the roughness and tribological characteristics of the CNT thin film. It was anticipated that UV-treated cover glass would have a higher surface energy than the glass plate without UV treatment. Therefore, using both untreated and UV-treated cover-glass plates for the squeezing process, the influence of cover-glass surface energy on the CNT thin film could be examined. The effect of the squeezing process on the surface roughness was measured using a laser scanning confocal microscope (VK-X200, Keyence, Japan). Figures 3(a) and 3(b) present 3D images of the surface topography of CNT thin films pressed with respective normal forces of 1 and 5 kgf, resulting in high and low surface roughness, respectively. In both cases, the substrate consisted of a glass plate treated with UV light. The exact morphology of the entangled CNTs is not evident in Fig. 3 since the scanned image area is approximately 280 um × 200 μm. Considering that the diameter of the CNTs used in this work is about 100 nm, the image area is too large to describe the fine features of a single nanotube. However, the officially reported microscope resolution (0.5 nm in the vertical direction) is regarded as sufficient to measure the surface roughness of the CNT thin film. The images in Fig. 3 confirm that the squeezing process can modify surface roughness effectively.

The change in CNT thin film thickness after pressing with a normal force of 5 kgf was also examined to verify the effect of the squeezing process. Three specimens were tested and the measurement was conducted 3–5 times for each specimen to provide reliable analysis. As a result, it was concluded that there was no significant change in CNT thin film thickness after the squeezing process.

### 2.2 Friction test setup and conditions

In this study, two different tribometers were exploited to measure the frictional force of the CNT thin film. One tribometer was designed to examine the frictional characteristics at the micro-Newton scale, by deflection of thin cantilevers. The other was used to measure
frictional force at the milli-Newton scale. For convenience, the former tribometer is referred to as a “micro-tribometer” and the latter is referred to as a “macro-tribometer”. The configuration of both tribometers is presented in Fig. 4. The micro-tribometer is a pin-on-disc type apparatus, which was set up with a step motor to precisely control the rotation speed, as shown in Fig. 4(a). To apply an extremely small normal force of about a micro-Newton to the CNTs, a tiny and thin counter specimen consisting of a 3×3 mm square diced piece of silicon with a thickness of 0.67 mm and three zirconia spheres with a diameter of 1 mm was utilized. The spheres, which were in contact with the CNT thin film during the friction test, were bonded to the silicon piece at discrete locations to maintain stable contact conditions. The counter specimen weighed 196 μN (20 mg). The CNT thin film-coated glass slide was placed on the circular-plate assembly with step motor and the counter specimen was placed on the CNT thin film. During operation, the glass plate rotated with an angular velocity of 6 rpm and the counter specimen pushed the long cantilever due to frictional force at the interface between its zirconia spheres and the CNT thin film. Two strain gauges were attached to the two flat faces at the end of the cantilever to detect cantilever deflection, and a half-bridge circuit was constructed to quantitatively measure the amount of deflection in voltage. Using a data acquisition system (NI9237, National Instruments, USA), the voltage signal was recorded and converted into a frictional force in Newton units through calibration of the tribometer. Each experiment was performed for 10 min and, for reliable results, the friction test for each specimen was performed 3 to 5 times.

The macro-tribometer operated as a pin-on-reciprocating type of tribotester (Fig. 4(b)). Using the tribometer, the frictional force of CNT thin films under a normal force of 4.9 mN (0.5 g) was measured. Zirconia spheres with a diameter of 1 mm were also exploited as the counter surface. These were bonded to the end of the triangular cantilever and the cantilever was assembled with a precision load cell (GSO-30, Transducer Techniques, USA), which was used to measure frictional force at the milli-Newton scale. A linear actuator transported the specimen back and forth, with stroke length, cycle number, and linear velocity of 2 mm, 200 cycles, and 4 mm/s, respectively. Similarly to the micro-tribotest experiments, the macro-tribotest experiments were also conducted 3 to 5 times per specimen for reliability.

The purpose of applying these two tribometers operating at different scales was to evaluate the influence of normal force on the frictional behavior of CNT thin films. Generally, pin-on-disk and pin-on-reciprocating type tribometers result in slightly different frictional behavior, owing to discrepancies in their respective sliding motions. In addition, it is possible that each type of tribometer will examine different wear characteristics. However, these tribometers were exploited not to quantitatively compare the friction coefficient between the macro and micro-scales, but to investigate the change in friction coefficient with respect to surface roughness. Therefore, it is believed that exploitation of different types of tribometer has not affected the fundamental analysis.

3 Experimental results

3.1 Surface analysis of CNT thin film with respect to the squeezing process

Since the dip-coating process is sensitive to various conditions such as temperature, humidity, withdrawal speed, and viscosity of the CNTs suspension, a set of CNT thin film specimens was prepared simultaneously. The temperature and relative humidity during the dip-coating process were 20–22 °C and 35%–40%, respectively. Simultaneous coating of a set of CNT thin film specimens prevented evaporation-induced viscosity changes in the CNT suspension. The set
consisted of five specimens squeezed by normal forces of 0, 0.5, 1, 2, and 5 kgf. A normal force of 0 kgf implies that this CNT thin film was not squeezed. The five specimens were also prepared simultaneously to minimize uncertainties in the surface condition arising from varying environmental conditions. Assessment of surface roughness was conducted using a laser scanning confocal microscope. For each specimen, the surface roughness of 3 to 5 points was measured and averaged to obtain reliable results. It was revealed that the surface roughness varied with applied squeezing force. When an untreated cover-glass plate was used for the squeezing process, the surface roughness decreased as the squeezing force increased. It was postulated that squeezing the CNT thin film with a certain weight caused the protruding nanotubes from the relatively smooth surface to deform; thus, the coarse surface topography became smoother. However, when a UV-treated cover-glass plate was used for the squeezing process, the surface roughness was affected differently. In this case, it was observed that the surface roughness of the CNT thin film generally increased as the squeezing force increased. This result can be explained in terms of the surface energy of the glass plate. As shown in Fig. 1, the UV-treated glass plate exhibits a higher surface energy compared to the untreated glass plate. Therefore, it can be deduced that the UV-treated glass plate would exhibit a stronger adhesion force to the CNTs compared to the untreated glass plate. When the CNT thin film was pressed using the untreated glass, each nanotube strand would be elastically deformed. Nevertheless, because of adhesion between nanotubes, the surface morphology of the CNT thin film would not be recovered to its original form after removal of the cover glass. This is regarded as the main mechanism for roughness control in the squeezing process. When the squeezing process was conducted using UV-treated cover glass, the increased adhesion force at the glass surface could result in different CNT thin film behavior. Prior to the release step, the deformation of the nanotube strands would be identical with the previous case. It was anticipated that the difference arises from the release step, in which the cover glass was removed. Because of increased adhesion force between the CNTs and the cover-glass surface, a relatively strong bond between the glass surface and the nanotubes was formed. During the release step, it was assumed that some nanotubes would be pulled off the surface of the CNT thin film due to bonding between the nanotubes and the glass surface, until the bonding force became equal to the sum of the elastic deformation force of the nanotubes and the adhesion force between nanotubes. As the squeezing force was increased, more nanotubes would form contact with the cover glass, strengthening the adhesion force between the glass and the nanotubes. Therefore, when a squeezing force of less than 1 kgf was applied, the surface roughness increased. However, when squeezing forces of 2 and 5 kgf were applied, it was postulated that the decreased surface roughness arising from the pressing step dominated over the increased surface roughness arising from the release step. Therefore, the surface roughness decreased under these conditions. In this regard, the increased surface roughness resulting from the squeezing process with UV-treated glass plate was attributed to the stronger adhesion force pulling off the CNTs attached to the cover-glass surface. Figure 5 displays a schematic representation of this speculative process. The resulting change in surface roughness of the CNT thin film is listed in Table 1. A difference in the surface roughness of unsqueezed CNT thin films is also observed between the untreated and UV-treated glass plates, which exhibit surface roughness values of 177.7 and 80.1 nm, respectively. This is attributed to uncertain factors such as irregular formation of CNT entanglement in the suspension during the coating process. In fact, the surface condition of a CNT thin film formed in this process cannot be similar to any thin film fabricated by a sputtering
Table 1  Change in surface roughness of CNT thin films with varying squeezing force.

<table>
<thead>
<tr>
<th>Squeezing force (kgf)</th>
<th>Surface roughness (nm)</th>
<th>With UV irradiation</th>
<th>Without UV irradiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>80.1</td>
<td>177.7</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>88.8</td>
<td>171.5</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>102.4</td>
<td>136.7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>89.3</td>
<td>119.3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>73.1</td>
<td>109.4</td>
<td></td>
</tr>
</tbody>
</table>

or evaporation method. The CNT thin film formed in this process is not an outcome of atomic level deposition, but rather random agglomeration of nanotube strands. Therefore, regular surface conditions cannot be readily achieved when the CNT thin film is formed using a dip-coating method, while the effect of the squeezing process was clearly confirmed through repeated experiments.

It was assumed that high pressure during the squeezing process might distort the crystal structure of the CNTs. Therefore, the influence of squeezing process on the crystallinity of the CNTs was also estimated using Raman spectroscopy (DXR 2, Thermo Fisher Scientific, USA). For the measurement, a laser operating at 532 nm, with a power of 3.0 mW and a spot size of 2.0 μm, was focused on the specimen. Raman spectra at three arbitrary points on each specimen were collected to qualitatively assess the CNT crystallinity. Figure 6 displays the Raman spectra for the CNT thin film squeezed with a normal force of 5 kgf and the unsqueezed CNT thin film. Both spectra show three Raman bands at about 1,350 cm⁻¹ (D band), 1,580 cm⁻¹ (G band), and 2,700 cm⁻¹ (G' band). For ease of viewing, the spectrum obtained from the CNT thin film squeezed with a normal force of 5 kgf is plotted with a vertical offset. As can be seen in Fig. 6, a strong D band associated with defect-induced and dispersive modes is observed. This was attributed to the fact that the nanotubes used in this work were multi-walled and defects occur in the tubes due to unstable crystal structure at the end of the nanotubes or between nanotube walls [30, 31]. However, regardless of squeezing process, identical intensities, peak positions, and peak widths were observed. In other words, the squeezing process does not influence the crystal structure of the CNT thin film, but only changes the surface roughness.

Based on a simple calculation, the apparent pressure applied to the CNT thin film when pressed by a dead weight with a normal force of 5 kgf was about 0.1 MPa. Considering that the CNT fracture strength has been reported in some studies to be 50–200 GPa, the stress level during the squeezing process is too small to destroy the CNT crystallinity [32]. Even if there is an occasional stress concentration, the stress would be distributed and minimized through nanotube deformation since the CNT thin film consists of entangled nanotubes, which can readily deform due to a high degree of porosity between the nanotubes, thereby reducing the potential local stress, which could damage CNT crystallinity.

It was expected that the CNT thin film would suffer significant distortion if the high contact pressure caused critical damage. Therefore, the surface morphology with respect to the squeezing process was also investigated using scanning electron microscopy (SEM). Figures 7(a) and 7(b) display the surface condition of the CNT thin film before and after the squeezing process, respectively. Both thin films exhibit a dense entangled structure consisting of nanotubes and no differences.
are observed. Indeed, the squeezing process does not damage the nanotube crystal structure.

3.2 Frictional force of CNT thin film with respect to surface roughness

Figure 8 presents frictional force data during a 10 min micro-tribotest. The figure shows the entire process of the friction test, beginning from the stationary state. At the early stage, zero frictional force is recorded since the step motor is not rotating. Once the step motor is started, a sudden increase in frictional force is observed, followed by a smaller change, which is attributed to run-in behavior during the first approximately 50–100 s. The frictional force then stabilizes rapidly and maintains a relatively steady value during the remainder of the test. Although the frictional force enters a stable stage after approximately 100–200 s, slight fluctuation in friction is observed. This is attributed to both the tribometer construction and the frictional characteristics of the CNT thin film. The micro-tribometer was designed to measure frictional force at the micro-Newton scale. Therefore, the apparatus consisted of a low stiffness cantilever to detect weak frictional signals; the cantilever was highly sensitive to both vibration and change in frictional force. Considering that the surface morphology of the CNT thin film was not perfectly smooth, it could reasonably be estimated that the frictional force at the surface would not be evenly distributed. Because of surface irregularity, a change in friction may arise during measurement and this fluctuation in friction may cause slight vibration of the tribometer. Hence, it was speculated that both factors influenced the measured frictional force signal. However, distinct differences in the average frictional force of the CNT thin film with respect to the squeezing process conditions could be observed, regardless of the outlined variability. Moreover, it was confirmed that the average frictional force did not vary significantly during the test. Therefore, the average of the frictional force was taken as the representative value.

However, since differences in frictional force related to surface roughness were revealed, the relationship between surface roughness and frictional characteristics was evaluated. Figure 9 presents the friction coefficient and surface roughness in relation to the normal force used in the squeezing process with UV-treated glass plate. In the figure, a strong relationship between friction coefficient and surface roughness is observed. As explained in the previous section, an increase of surface roughness even squeezing process was performed can be attributed to the high adhesion force of the cover-glass plate. It is revealed that a change in friction coefficient generally corresponds to a change in surface roughness. However, the squeezed CNT thin film with the lowest surface roughness still exhibits a higher friction coefficient than the unsqueezed CNT thin film. It was postulated that the squeezing process might decrease overall surface waviness, but microscopic protrusion of CNTs was promoted during detachment of the cover glass. Therefore, it is believed that the protruding CNTs disturb the relative motion of the counter surface and thus increase the frictional force. According to Fig. 9, the friction coefficient increases from 0.23 to 0.51 as the surface roughness increases.
The frictional characteristics of a set of CNT thin films prepared using untreated glass plates were also investigated. In Fig. 10, the relationship between friction coefficient and surface roughness is displayed. It is obvious that the surface roughness decreases as the press weight increases. The trend of change in surface roughness with respect to the normal force applied during the squeezing process is quite different from the previous case for UV-treated glass plates. In Fig. 10, the CNT thin film specimen is squeezed using an untreated glass plate. Since the adhesion force between the glass plate and CNTs is much lower than that between UV-treated glass and CNTs, it is speculated that the squeezing process decreases not only the waviness of the CNT thin film, but also reduces the microscopic irregularity arising from protruding nanotubes. Therefore, it has been confirmed that the friction coefficient continuously decreases from 0.54 to 0.27 as the press weight increases.

The friction coefficient of the CNT thin film with respect to squeezing process at the milli-Newton scale was also evaluated. Since the influence of UV irradiation was revealed through the micro-tribotest, the squeezing process was only performed with untreated cover-glass plate. Similarly to the previous case, the surface roughness of the CNT thin film decreases with increasing press weight, except for the case of 5 kgf, as shown in Fig. 11. In this experiment, it is also clear that the friction coefficient is proportional to the surface roughness. The friction coefficient decreases from 0.48 to 0.36 as the press weight increases. Therefore, this result indicates that the surface roughness of the CNT thin film also influences frictional behavior at the milli-Newton scale. Nevertheless, the change in friction coefficient at the milli-Newton scale is clearly reduced compared to at the micro-Newton scale. In other words, the influence of surface roughness on friction is reduced as the contact pressure increases.

The exception to this observation, at 5 kgf, is attributed to complicated factors involved in the surface roughness measurement. It has been assumed that the difference in surface roughness between the specimens pressed by dead weights with normal forces of 2 and 5 kgf is insignificant. However, the examined area in the surface roughness measurement is only approximately \(280 \mu m \times 200 \mu m\). This area inevitably contains both surface roughness and waviness. Usually, the influence of waviness must be eliminated to obtain surface roughness. Usually, the influence of waviness must be eliminated to obtain surface roughness. However, if there is a complicated wavy structure, the surface waviness would not be completely eliminated. If the surface roughness of two specimens differs sufficiently, imperfectly filtered waviness would not affect the result. In contrast, if two surfaces have similar surface roughness values, then the residual waviness would influence the result. In addition, the roughness data displayed in Fig. 11 are obtained by averaging several measurement points on several specimens. In this process, one relatively large value could be included in the measurement. The exception at 5 kgf is regarded as the outcome of a combination of these factors.

4 Conclusions

The frictional characteristics of CNT thin films produced by applying a squeezing method to control
surface roughness were experimentally investigated. The following points were inferred from the experimental results.

1. The squeezing process could effectively change the surface roughness of the CNT thin film when a cover-glass plate with relatively low surface energy was used. In contrast, cover glass with high surface energy caused the surface condition to deteriorate at a certain press weight. It was speculated that the high surface energy caused the CNTs to partially protrude; these protruded CNTs contributed to the increase in friction.

2. By using Raman spectroscopy and SEM, it was revealed that the squeezing process influenced the surface roughness of the CNT thin film, but did not influence the CNT crystallinity.

3. The friction coefficient of the CNT thin films was generally related to their surface roughness. In particular, a drastic change in friction coefficient (from 0.23 to 0.51 and from 0.27 to 0.54 for UV-treated and untreated cover glass, respectively) with respect to surface roughness was measured in the micro-tribotest, while the macro-tribotest results exhibited only a slight difference in friction coefficient (from 0.36 to 0.48) with surface roughness.

4. It was concluded that surface roughness strongly influenced the frictional characteristics at the micro-Newton scale. Thus, controlling the surface roughness of CNT thin films must be regarded as the most important issue prior to their application in precision systems.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2015R1C1A1A01053416).

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


Youn-Hoo HWANG. He received his bachelor degree in Department of Precision Mechanical Engineering in 2017 from Kyungpook National University, Sangju, Republic of Korea. Currently, he is M.S. candidate at Kyungpook National University. His research interests include tribological characteristics of nanomaterial and wear reduction mechanism.
Byung-Soo MYUNG. He received his M.S. and PhD degrees in mechanical engineering from Chungnam National University, Daejeon, Republic of Korea, in 1986 and 1993 respectively. He is currently chair of Department of Precision Mechanical Engineering, Kyungpook National University. His research interests are reduction of friction through vibration and contact characteristics during vibration.

Hyun-Joon KIM. He received bachelor and PhD degrees in mechanical engineering from Yonsei University, Seoul, Republic of Korea, in 2005 and 2012, respectively. He joined Department of Precision Mechanical Engineering at Kyungpook National University as assistant professor from 2014. His current interests are AFM based nano-mechanics, micro-nano tribology, and molecular dynamics simulation.
Numerical modeling and analysis of plasmonic flying head for rotary near-field lithography technology

Yueqiang HU, Yonggang MENG

State Key Laboratory of Tribology, Tsinghua University, Beijing 100084, China

Received: 13 April 2017 / Revised: 30 May 2017 / Accepted: 02 September 2017

Abstract: Rotary near-field lithography (RNFL) technology provides a route to overcome the diffraction limit with a high throughput and low cost for nanomanufacturing. Utilizing the advantage of the passive flying of a plasmonic head, RNFL can achieve a 10 m/s processing speed with a perfect near-field condition at dozens of nanometers. The flying performance of the plasmonic flying head (PFH) is the pivotal issue in the system. The linewidth has a strong correlation with the near-field gap, and the manufacturing uniformity is directly influenced by the dynamic performance. A more serious issue is that the unexpected contact between the PFH and substrate will result in system failure. Therefore, it is important to model and analyze the flying process of the PFH at the system level. In this study, a novel full-coupled suspension-PFH-air-substrate (SPAS) model that integrates a six-degree of freedom suspension-PFH dynamics, PFH-air-substrate air bearing lubrication, and substrate vibration, is established. The pressure distribution of the air bearing is governed by the molecular gas lubrication equation that is solved by the finite element method (FEM) with a local pressure gradient based adaptive mesh refinement algorithm using the COMSOL Multiphysics software. Based on this model, three designs of the air bearing surface are chosen to study the static, dynamic, and load/unload performance to verify whether it satisfies the design requirements of RNFL. Finally, a PFH analysis solver SKLY.app is developed based on the proposed model.

Keywords: rotary near-field lithography (RNFL); coupled analysis; air bearing; finite element method

1 Introduction

Optical lithography has been extensively performed in the semiconductor industry for decades. Currently, to sustain Moore’s law [1], the critical dimension is a 10-nm node or even 7 nm node according to the International Technology Roadmap for Semiconductors (ITRS); this leads to a substantial increase in the cost [2]. In particular, the fabrication of the high-quality lithography masks is expensive and difficult with multiple-lithography. Focused ion beam (FIB) and electron-beam lithography (EBL) offer the ability to attain the resolution of less than 10 nm; however, the low throughput of these methods owing to their slow scanning ability, restrains their applications for nanofabrication on a large scale [3, 4]. Near-field lithography (NFL) is a newly emerging technique that has been explored for realizing ultrahigh resolution beyond the Rayleigh diffraction limit with a specially designed plasmonic focusing lens and without image aberration under an extremely short work distance [5–8]. Lee et al. [9] used an active nanogap control implementation in which a metallic sharp ridged nanometer-scale aperture was fabricated on a solid immersion lens (SIL), and the nanometer-scale aperture excited the surface plasmon polaritons (SPPs) and focused the light on the center of the SIL. In 2011, the Zhang’s group [10, 11] achieved maskless lithography with a resolution of 22 nm by rotary near-field lithography (RNFL) on an inorganic thermal photoresist.

* Corresponding author: Yonggang MENG, E-mail: mengyg@tsinghua.edu.cn
TeO$_x$ film. Recently, we performed RNFL on a new molecular glass chemically amplified resist (CAR) that has a higher sensitivity and resolution than a thermal photoresist and can be easily applied in the industry without additional cost [12, 13]. The system scheme is illustrated in Fig. 1. The plasmonic lens is fabricated on a transparent slider and flies above a substrate coated with the photoresist layer. An array of plasmonic lens, forming the plasmonic flying head (PFH), generates a self-adaptive near-field gap ($d_{nl}(x, y)$) of tens of nanometers between the PFH and high speed revolving substrate that is insensitive to the environmental vibration. The plasmonic lens can focus the incident laser beam to a spot size of sub-20 nm with enhanced field intensity by exciting the surface plasmon polaritons.

The flying performance of the PFH is key to RFNL [12]. Owing to the exponential decay of the surface plasmon in the near-field range, the near-field gap of the PFH will directly determine the laser light absorption of the photoresist that further influences the lithography linewidth. In addition, the dynamic performance of the PFH will affect the manufacturing uniformity. Furthermore, it is difficult to fly on the organic CAR because of its poor mechanical properties; specifically, the unexpected contact between the PFH and substrate will result in system failure. Therefore, modeling and analysis of the flying of the PFH is significant for a reliable RNFL design. The principle of PFH is similar to that of the magnetic head in hard disk drives (HDDs). In the last few decades, the head-disk interface (HDI) has been extensively studied numerically and experimentally along with the development of the HDD industry [14-17]. A small gap of the order of 5–20 nm, which is much smaller than the mean free path of the air molecule, is maintained between the disk and slider in an HDD; therefore, the continuum model is no longer valid in this region [18]. The generalized Reynolds equation is introduced by modifying the traditional Reynolds equation with the slip theory [19, 20] or linearized Boltzmann equation [21] to solve the problem. Among these, Fukui and Kaneko’s molecular gas lubrication (MGL) equation is widely used in the HDD industry because of its good convergence to the Boltzmann equation [22, 23]. The finite difference [24, 25], finite volume [26, 27], and finite element (FE) methods [28, 29] are implemented to solve the generalized Reynolds equation. For the FE method, use of a structured mesh is not required, so that it is highly suitable for a complex geometry and particularly for local adaptive refinement in regions where the solution displays large gradients of the solution parameter or the slider geometry changes. The slider equilibrium attitude is determined by the air pressure and applied suspension forces. The coupling between the two sets of equations can be solved by two approaches. The solution can be found either by the transient solution approach until the steady-state conditions are found, or by directly solving the steady coupled equations. Several models have been established to analyze the dynamic performance of the magnetic head flying [30, 31]. However, the PFH in an RNFL system has a different system setup and design goals, so a new model is required.

In this paper, we proposed a novel full-coupled suspension-PFH-air-substrate (SPAS) model that integrates a six-degrees of freedom (6-DOF) suspension-head dynamics, head-air-substrate air bearing lubrication, and substrate vibration for RFNL. The pressure distribution of the air bearing is governed by the MGL equation that is solved by the FE method with a local pressure gradient-based adaptive mesh refinement algorithm using the COMSOL Multiphysics software. The Newton–Raphson method is utilized to solve the slider equilibrium attitude by neglecting the inertia term and damping term of the proposed model. Next, three designs of the air bearing morphology are studied to examine the static, dynamic, and load/unload performance to verify whether it satisfies the requirements of the PFH in RNFL. Finally, a PFH
analysis solver, SKLY.app is developed based on the proposed model.

2 Mathematical models

2.1 Suspension-PFH-air-substrate system modeling

The PFH in RNFL utilizes the concept of the air bearing head in the HDD. An air bearing surface (ABS) with a suitably designed morphology exerts an aerodynamic force to support the head to fly on a nanoscale distance with a high-speed rotary substrate that can provide a near-field condition for the lithography. An actual PFH system in RNFL has several components such as the main suspension, gimbal, limiter, head, substrate, as well as ambient air. The different parameters of these components will result in different static head postures and system dynamic performances that are the key lithography quality evaluation indices of RNFL. Therefore, performing a numerical simulation of the PFH system is usually necessary because it can be used to study the different factors influencing the static and dynamic performances, and furthermore to assist in the design of the head and suspension parameters to satisfy the demand of NFL.

We established a fully coupled suspension-PFH-air-substrate (SPAS) model. Figure 2 shows the schematic of the different states of the suspension-head-substrate system and its corresponding dynamic model. The 6-DOF model divides the suspension into three parts, namely, the tab, local suspension, and main suspension, and the acting force between these parts is simplified by the spring and damping. The interaction between the suspension and head consists of three parts: the dimple/head contact force, gimbal force/moments, and limiter contact force, when the dimple separation is larger than the limiter gap. The head has three DOF with corresponding posture parameters of near-field gap, pitch angle, and roll angle. The aerodynamics force, head/substrate contact force, and intermolecular force act on the head/substrate interface (HSI).

To model the PFH system, we establish the following 6-DOF dynamic motion equations that can describe the complete process of the PFH system:

\[
M \ddot{Z} + KZ + CZ = F
\]

where \(M = [m_t, m_l, m_m, I_x, I_y, I_z]^T\) and \(Z = [z_t, z_l, z_m]^T\) are the generalized mass and displacement vectors, respectively, of the tab, local suspension, main suspension, and head, which are the corresponding 6-DOF parameters in the PFH system.

Combining the different states, we can obtain the stiffness matrix, damping matrix, and external force vector,

\[
K = \begin{bmatrix}
  k_1 & -k_1 & 0 & 0 & 0 & 0 \\
  -k_1 & k_1 + k_l & -k_l & 0 & 0 & 0 \\
  0 & -k_l & k_l + k_s & -k_s & 0 & 0 \\
  0 & 0 & -k_s & k_s & 0 & 0 \\
  0 & 0 & 0 & 0 & k_\beta & 0 \\
  0 & 0 & 0 & 0 & 0 & k_\beta
\end{bmatrix}
\]  

Fig. 2 Schematic of the suspension-PFH-air-substrate (SPAS) system: (a) suspension in the unloaded state, (b) dynamic motions of the head during flying with the suspension in the loaded state, and (c) load/unload process.
and 0

are the density of the asperities,

are the coordinates of the gravity center.

is the squeeze number defined as 

in which \(\omega_s\) is the angular velocity; and \(Q\) is a flow rate coefficient to account for the gaseous rarefaction effect which has been provided by Fukui and Kaneko’s model derived from the linearized Boltzmann equation [21].

After obtaining the pressure distribution, the aerodynamics force and moments can be calculated by

\[
F_{\text{air}} = p_a \int (P - 1) dA,
\]

\[
M_{\text{air}} = F_{\text{air}} (x_{\text{fa}} - x_G), \quad M_{\text{air}} = F_{\text{air}} (y_{\text{fa}} - y_G)
\]

where \(p_a\) is the atmospheric pressure, \(x_{\text{fa}}\) and \(y_{\text{fa}}\) are the coordinates of the air bearing force center, and \(x_G\) and \(y_G\) are the coordinates of the gravity center.

For the head/substrate contact force, Chang et al. [32] proved that for extremely smooth surfaces the contact is mostly elastic throughout. Therefore, here we use the Greenwood and Williamson (GW) model [33], the basic elastic rough surface contact model, to calculate the head/substrate contact force.

\[
F_{\text{cad}} = \frac{4}{3} \eta r^{0.5} \sigma \int_{A}^{\infty} (s-h)^{1.5} \phi(s) ds dA,
\]

\[
M_{\text{cad}} = F_{\text{cad}} (x_{\text{fcd}} - x_G), \quad M_{\text{cad}} = F_{\text{cad}} (y_{\text{fcd}} - y_G)
\]

where \(F_{\text{cad}}\) is the equivalent stiffness of the head and substrate; \(\eta, r, \sigma\) are the density of the asperities, asperity radius of the curvature, and standard deviation of the asperities heights, respectively; \(\phi(s)\) is the dimensionless asperity height distribution function; and \(x_{\text{fcd}}, y_{\text{fcd}}\) are the coordinates of the contact force center.

For the intermolecular force, with the expression derived from a plane–plane interaction, the total force and moments owing to the intermolecular interaction between the head and substrate are obtained by

\[
F_{\text{i}} = \int_{A}^{\infty} \frac{A}{6 \pi h^3} dA + \int_{A}^{\infty} \frac{B}{45 \pi h^5} dA,
\]

\[
M_{\text{i}} = F_{\text{i}} (x_{\text{fi}} - x_G), \quad M_{\text{i}} = F_{\text{i}} (y_{\text{fi}} - y_G)
\]

in terms of the Hamaker constant \(A\) and another constant \(B\); typical values of \(A\) and \(B\) are approximately

\[
\begin{bmatrix}
  c_t & -c_t & 0 & 0 & 0 & 0 \\
  -c_t & c_t + c_l & -c_t & 0 & 0 & 0 \\
  0 & -c_t & c_m + c_s & c_s & -c_s & 0 \\
  0 & 0 & c_m & c_s & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & c_b \\
  0 & 0 & 0 & 0 & 0 & c_b
\end{bmatrix}
\]

where \(k_i\) and \(c_i\) \((i = 0, 1, m, g, s, b)\) are the simplified spring and damping coefficients, respectively, as shown in Fig. 2. In the above equations, \(F_{\text{ai}}\) and \((M_{\text{cha}} \text{ and } M_{\text{cha}})\) are the dimple/head contact force and moments, respectively; \((F_{\text{ai}}, F_{\text{cd}}, \text{ and } F_{\text{in}})\) and \((M_{\text{ai}}, M_{\text{cd}}, \text{ and } M_{\text{in}})\) are the forces and moments, respectively, caused by the aerodynamics force, contact force, and intermolecular force acting on the HSI; \(F_{\text{ai}}\) is the limiter force; \(\alpha_0\) and \(\beta_0\) are the initial pitch angle and roll angle in the design or formed inadvertently when the air bearing heads are mounted on the flexure of the suspension; and \(F_{\text{ai}}\) is the ramp/tab contact force in the load/unload condition.

The Reynolds equation is usually used to calculate the air pressure distribution at the HSI. However, because of the significant rarefied effect when the spacing between the head and substrate is extremely small, such that it is much less than the mean free path of air molecule, the continuum model is not applicable. The Reynolds equation has been modified in the last few decades to calculate the pressure distribution in such severe conditions. The modified equation can be expressed in the following dimensionless generalized form:

\[
\nabla \cdot (P Q H^3 \nabla \Pi) = \Lambda \nabla (PH) + \Psi \frac{\partial}{\partial T} (PH)
\]

where \(H\) is the dimensionless air bearing thickness; \(\Lambda\) is the bearing number defined as 

\[6 \mu U L / (p_u h_i^3)\], where \(\mu\) is the ambient gas viscosity, \(U\) is the
10^{-39}\mathrm{J}\text{ and }10^{-26}\mathrm{J}\cdot\mathrm{m}\cdot$, respectively. In Eq. (8), $x_{ri}$ and $y_{ri}$ are the coordinates of the intermolecular force center.

In the load/unload process, the contact states at the dimple/head, limiter/suspension, and tab/ramp interfaces will change, causing the PFH system to have several states. The limiter contact is defined as a rigid contact. The limiter force is defined as

$$F_{lim} = \begin{cases} k_{li}(z_m - z_s - z_{gap}), & z_s - z_m - z_{gap} > 0 \\ 0, & z_s - z_m - z_{gap} < 0 \end{cases} \quad (9)$$

where $k_{li}$ is the limiter stiffness and $z_{gap}$ is the initial distance between the suspension and limiter.

The Hertzian contact model is used to calculate the dimple/head contact force and ramp/tab contact force. The contact forces are defined as

$$F_{cls} = \begin{cases} \frac{4}{3} E^*_{cls} R_{cls}^{0.5} (z_s - z_l)^{1.5}, & z_s > z_l \\ 0, & z_s < z_l \end{cases} \quad (10)$$

$$F_{crit} = \begin{cases} \frac{4}{3} E^*_{crit} R_{crit}^{0.5} (h_{ram} - z_s)^{1.5}, & z_s > h_{ram} \\ 0, & z_s < h_{ram} \end{cases} \quad (11)$$

where $E^*_{cls}$ is the equivalent elastic modulus of the dimple and head, and $R_{cls}$ and $R_{crit}$ are the equivalent radius of the curvatures of the dimple and tab, respectively.

2.2 Numerical methodologies

The most challenging step in the numerical simulation process is to solve the generalized Reynolds equation. In static analysis, there are two problems: one is the forward problem to determine the pressure distribution for the given attitude parameters of the head, and the other is the inverse problem to determine the static attitude parameters for the given external forces and moments. In dynamic analysis, the problem is establish an approach to couple the transient generalized Reynolds equation and motion differential equations to simulate the system dynamic performance. With the Galerkin method, the partial differential equation (Eq. (5)) is multiplied by test function $v$ and integrated over domain $\Omega$ to obtain

$$\int_\Omega \nabla \cdot (PQH^3 \nabla P) dA - \int_\Omega A \cdot \nabla (PH) dA$$

$$- \int_\Omega \frac{\partial}{\partial T} (PH) dA = 0 \quad (12)$$

Processing the equation with integration by parts, along with the boundary condition $P = 1$ over $\Omega$, we can obtain the weak form,

$$\int_\Omega \nabla v \cdot (PQH^3 \nabla P) dA - \int_\partial \Omega n \cdot (PQH^3 \nabla P) dS$$

$$+ \int_\Omega A \cdot \nabla (PH) dA + \int_\Omega \psi \frac{\partial}{\partial T} (PH) dA = 0 \quad (13)$$

In this study, Eq. (13) was solved by using the commercial software COMSOL Multiphysics based on the FEM. In the simulation, the Newton–Raphson iterative method was chosen to solve the nonlinear problem, and the global system of equations obtained in each iteration step were solved by the Multifrontal Massively Parallel Sparse Direct Solver (MUMPS).

For the forward problem, the time-dependent terms in Eq. (5) were neglected to obtain the steady pressure distribution at the given head postures. For the inverse problem, we adopted the Newton–Raphson method to calculate the force equilibrium equation that is Eq. (1) but without the inertia terms and damping terms at a given pre-force and moment. We solved the forward problem in each iteration step to obtain the head 3-DOF attitudes. The contact force and intermolecular force were calculated at every element with the given near-field gap distribution. For the dynamic problem, we first calculated the substrate dynamic response, and then coupled the response with the 6-DOF suspension-substrate model to derive the head dynamic response. Although the load acting on the substrate is not axisymmetric, there is a minor nonaxisymmetrical effect on the substrate dynamic response. Equation (5) was calculated transiently with the backward difference method in the time direction, and the Runge–Kutta method was adopted to solve the Eq. (1). Simultaneously solving the equations, we can obtain the dynamic performance of the head. The relative discontinuity of the head geometry results in a large gradient of the pressure distribution; therefore, the meshes for the simulation were generated by the adaptive mesh refinement algorithm to achieve accurate results.
3 Simulation results

3.1 Static analysis

As previous studies have shown, for an RNFL system (Fig. 1), the intensity profile of the modified Bull’s eye plasmonic lens (PL), which is characterized by the full width half maximum (FWHM), emitted from the plasmonic lens on the photoresist expands with increase in the off-plane distance, but does not change much when the off-plane distance is in the range of 0–30 nm [34]. The exponential decay of the intensity with the increase in the off-plane distance also increases the demand for the sensitivity of the photoresist. However, the actual focus spot size is determined by not only the near-field gap but also the PL design. Ji et al. [35] shows a lower FWHM can be achieved by an optimal plasmonic lens design. It should be noted that the decrease in the near-field gap would generate a series of problems such as unexpected contact and wear, which would cause an early failure of the system. Therefore, the steady near-field gap at the plasmonic lens position \(d_{PL}\) must be carefully designed. In this work, the design target of \(d_{PL}\) is set as 0–30 nm. In addition, to ensure a manufacturing uniformity for the substrate, \(d_{PL}\) is expected to be as robust as possible for different values of the substrate radius and preload forces. The ABS must have a relatively large area for the array of PLs to enable parallel writing and high-throughput.

Here we have used three ABS designs modified from HDD designs, as shown in Fig. 3(a). Head 1 is a simple design from the triple pad design. Head 2 is a five-pad head design in Ref. [36] designed in the computer mechanics laboratory (CML). Head 3 is a head of a commercial design used in the HDD industry. Different colors represent different recess depths. The three designs are chosen and evaluated to determine whether they satisfy the RNFL system requirement. A large area at the trailing edge of the array of the PLs is present in all the three designs. The position of the plasmonic lens is also shown in Fig. 3.

We first implemented the forward problem to test the simulation model. The operational parameters for the heads are fixed at a nominal \(d_{nf}\) (defined as the clearance at the trailing edge center) of 10 nm, pitch of 150 rad, roll of 1 rad, and radius of 18 mm with a

![Fig. 3](image_url) Three ABS designs: (a) three ABS morphologies with different recess depth, (b) adaptive meshes based on the local pressure gradient, and (c) dimensionless pressure distribution in forward problem.
0° skew. The rotation speed of the substrate disk is set as 5,400 rpm. Figures 3(b) and 3(c) show the adaptive meshes based on the local air pressure gradient and the pressure distributions of the three heads, respectively. The heads are mainly supported by the high-pressure peaks generated by the central trailing pads, maintaining the fly height of the air bearings. The sub-ambient pressure zone in the central regions helps to improve the entire air bearing stiffness and decrease the sensitivity of the substrate speed or other parameters. The side-pad pressure assists in maintaining the stiffness along the rolling direction.

The computation accuracy and efficiency of the adaptive mesh algorithm are also studied because they are important for the PFH analysis. Figure 4 shows the result, where $L$ is the ABS length and $l$ is the maximum size of the initial uniform triangle mesh. Mesh parameter $L/l$ represents the initial mesh intensity. Different levels of adaptive mesh refinement are tested. With the increase in the initial mesh intensity, the accuracy will have a significant improvement accompanied with a slight increase of computing time. We find that the accuracy tends to be stable, whereas the computation time increases linearly with the increase in the initial mesh intensity. With the increase in the adaptive level, the computation accuracy will be improved evenly, though at the price of a longer computing time because the computing time increases much faster at larger mesh parameters. Hence, to ensure the accuracy of the model and make a trade-off with the computing time, initial mesh intensity of 50 and adaptive level of 3 are chosen in the following calculations. For the inverse problem, we have studied the impact of the radius, pitch static angle (PSA), and preload on the near-field gap at plasmonic lens position $d_{pl}$. The linear speed of the substrate is proportional to the radius and coupled with bearing number $A$ in the generalized Reynolds equation. PSA is the initial pitch angle formed by mounting the head on the flexure of the suspension. A preload is generated by the bending of the suspension, which mainly determines the static attitude. As shown in Fig. 5(a), heads 1, 2, and 3 are designed to provide a near-field condition for the plasmonic lens at ~30 nm, ~5 nm, and ~15 nm, respectively. The fluctuation in $d_{pl}$ is significant for head 1, which may induce a nonuniformity in the substrate at different radii. $d_{pl}$ of head 2 is proportional to the radius with a small slope. Head 3 has a minor fluctuation in the short radius range and then remains almost constant from 15–30 mm. The reason $d_{pl}$ of heads 1 and 3 does not monotonically increase with the radius is because of the strong negative pressure effect. For the two heads, when the linear speed is sufficiently high, the negative pressure is so strong that the head changes its flying height, pitch angle, and roll angle simultaneously to maintain its balance. In such cases, the flying height no longer exhibits a simple monotonic change with the linear speed and may decrease with the increase in the radius. Figure 5(b) illustrates that the preload has a significant influence on $d_{pl}$; with an increase in the preload, $d_{pl}$ decreases because the head must balance the load with the air bearing force by a smaller near-field gap. Head 2 shows a relatively low sensitivity to the preload. Figure 5(c) exhibits that $d_{pl}$ is proportional to the PSA, and head 1 is more sensitive to the PSA than the other two heads. The results of the PSA and preload analysis can help to control $d_{pl}$ according to different parameters of the suspension. Based on these calculation results, it can be concluded that head 2 is the best, whereas head 1 is the worst among the three designs in terms of the static performance.

### 3.2 Dynamic analysis

In the operation of RNFL, the system will experience...
different shock events from the environment. Such shocks may directly cause damage to the PFH and photoresist owing to the severe contact between the head and substrate surface. In addition, the $d_{PL}$ fluctuations during the shock events result in an instability of the near-field condition that determines the width of the lithography. Therefore, it is required to rapidly and accurately simulate the operating shock response to obtain a better design of the PFH system. In our study, a shock is modeled as a half sine acceleration wave defined by its peak amplitude and pulse width.

In Fig. 6, we display the free substrate and 6-DOF suspension-PFH system shock response, without the air bearing effect, to a 100G shock of different pulse widths and power spectra. The substrate FE model mesh and response at the outer diameter (OD) and middle diameter (MD) are shown in Fig. 6(a). It has been shown in various studies that the shock response of a rotating substrate to an axisymmetric shock is primarily composed of the first axisymmetric (umbrella) mode, first radial mode, and first axisymmetric–radial coupled mode. From the power spectra of the substrate shock in Fig. 6(c), we can observe that the corresponding mode frequencies are 12.4 kHz, 74.02 kHz, and 188.64 kHz that are all excited in the 0.2 ms pulse case. Figure 6(b) displays the displacements of different components of the PFH system and dimple–head contact force during the shock. We can see that the motion of the head from

Fig. 5 Static performance of the PFHs: (a) $d_{PL}$ at different substrate radii, (b) $d_{PL}$ at different preloads, and (c) $d_{PL}$ at different pitch static angles (PSAs).

Fig. 6 Shock response of the substrate and suspension-PFH system without the air bearing effect under different pulse widths: (a) substrate response at the outer diameter (OD) and middle diameter (MD), (b) response of different components of the PFH system and dimple–head contact force (Dimple FCon), (c) power spectra of the substrate, and (d) power spectra of the PFH.
the green lines that follow the suspension but with a local vibration during the shock, which is owing to the change in the dimple contact state. The primary mode of the suspension at all the pulse widths is the main suspension mode with a frequency of 170 Hz, as shown in power spectra in Fig. 6(d). With the increase in the shock pulse, the peak value of the displacement increases.

Next, we consider the case where the acceleration shock is applied to the entire SPAS system uniformly. Here, a positive shock is defined as one that causes the substrate to initially move toward the head, and both move in the same direction. The suspension is usually designed to be more flexible than the substrate; therefore, the head will move far away from the substrate and the clearance between the head and substrate will increase under the shock. However, in a negative shock, the substrate is initially followed by the slider, but because of the different stiffness, the head substrate clearance will first decrease. Both positive and negative shocks are studied in our following simulations.

Figure 7 exhibits the response of the full-coupled SPAS model to a 100-G positive shock. Head 3 is chosen to study the shock pulse effect. The results are shown in Figs. 7(a)–7(c). First, comparing with the PFH response without air bearing in Fig. 6(b), the head displacement in the SPAS model is quite small because the air bearing assists the PFH to sustain a steady near field condition. The suspension being more flexible than the substrate, the PFH will move far away from the substrate under a positive shock, and thus, the clearance between the PFH and substrate will increase as shown in Fig. 7(b). With the increase in the pulse width, the maximum displacement of all the components during the shock increases, resulting in a relatively large fluctuation in the post-shock. The fluctuation of $d_{hl}$ shows the same result. The large head slap owing to the separation and snap back may cause failure of the PFH. There are three frequency components in all the pulse cases, namely, 3.25 kHz, 6.54 kHz, and 12.4 kHz, observed in the power spectra in Fig. 7(c). The former two frequencies are the constrained modes of the suspension system, and the latter is the first substrate umbrella induced mode. Figures 7(d)–7(f) show the shock responses of the three ABS designs at the same pulse width. We can notice that the suspension components of the three

![Fig. 7 Positive shock response of the SPAS model: (a) components response, (b) substrate, head, and $d_{hl}$ response, and (c) power spectra of the head at different pulse widths, (d) components response, (e) substrate, head, and $d_{hl}$ response, and (f) power spectra of the head with different ABS designs.](http://friction.tsinghua.journals.com)
designs exhibit similar responses. From the $d_{PL}$ response in Fig. 7(e), we can observed that head 2 has the highest stiffness to sustain a relative steady near-field gap, whereas head 1 has the least stiffness. This is because head 2 has a relatively high negative pressure that tends to improve the dynamic performance. The power spectra of the three ABS designs also show the three frequency components.

Figure 8 presents the response of the full-coupled SPAS model to a 100-G negative shock. Figures 8(a)–8(c) show the head 3 responses at different pulse widths. We can expect that the suspension will come close to the substrate under a negative shock and bounce back owing to the high air bearing force. However, the head will contact the substrate, resulting in the failure in the large pulse width case, as shown in Figs. 8(a) and 8(b), because the inertia load of the substrate overcomes the air bearing force. The power spectra exhibits a constrained mode and substrate mode. The high stiffness performance of head 2 is also indicated in Fig. 8(e).

### 3.3 Load/unload analysis

The taking-off and landing of the PFH in RNFL is the key problem because of the peeling off the soft viscoelastic photoresist caused by the collision and friction. In a previous study [12], we have processed a transition zone for the taking-off to avoid the catastrophic collision between the plasmonic head and photoresist in the taking-off stage. The load/unload technology relies on the heads being lifted off the substrate onto a ramp. In comparison with the transition zone method, the load/unload method has a higher repeatability, a much higher shock resistance in the non-operational state, and the advantage of avoiding the problem of stiction that generates wear debris from the ABS and photoresist. However, the load/unload method must be studied to minimize the risk of the contacts in the operation process. In the simulation, the load/unload velocity is set as 50 mm/s and the substrate velocity is 5,400 rpm.

Figure 9 shows the typical stages of the successful load/unload process of head 3. The force and displacements histories during the load process are displayed in Figs. 9(a) and 9(b), respectively. They can be obviously divided into three stages. First, the dimple contact with the head causes a vibration of the head and tab, which can be illustrated by the ramp.

---

**Fig. 8** Negative shock response of the SPAS model: (a) components response, (b) substrate, head, and $d_{PL}$ response, and (c) power spectra of the head at different pulse widths, (d) components response, (e) substrate, head, and $d_{PL}$ response, and (f) power spectra of the head with different ABS designs.
contact force and dimple contact force fluctuation. The displacement of each component also shows a fluctuation in the first stage. Second, the dimple contacts the head stably and the air bearing effect is gradually generated by the positive and negative pressures. Owing to the inertia of the main and local suspension, the flying height (FH) and air bearing force of the head initially experience a fluctuation and then tend to stabilize with the damping effect of the air bearing. In the final stage, the ramp contact force becomes nearly zero, indicating that the tab is separated from the ramp and head load successfully with a minimum near-field gap of 12 nm on the substrate. The results of the unload process are depicted in Figs. 9(c) and 9(d). The variations can be divided into four stages. In the first stage, the air bearing force decreases to zero and the FH steadily increases. The positive pressure resultant force of the bearing decreases and the suction force is almost a constant. In the second stage, the dimple separates, reaching the maximum separation of the limiter distance. The bearing force changes from 0 to a negative value. In the third stage, the limiter is contacted and the contact force increases gradually. The air bearing force continues to decrease to a maximum negative value of −14 mN. Then the air bearing effect rapidly disappears, whereas the FH sharply increases. This occurs because the head is pulled by the suspension and limiter. In the last stage, the head strongly vibrates owing to the combined effects of the head inertia and dimple contact force. In addition, the PFH system experiences a vibration because of the suspension inertia.

Figure 10 displays the load/unload performance comparison of the three different head designs. The force and minimum FH histories during the load process are shown in Figs. 10(a) and 10(b), respectively. There are two moments at which the head may contact the substrate. The first risk point is at the start of the second stage, where all the heads experience a vibration owing to the inertia of the suspension. Head 3 exhibits a better performance than the other two heads because of the rapid formation of the air bearing and its strong damping effect. The second point is in the process of the formation of the air bearing, during which the torque of the positive and negative pressures may cause the overturn of the heads. Therefore, a positive pitch torque is expected at the center location of the positive pressure, close to the air upstream, and the opposite is expected for the negative pressure. Head 2 undergoes a severe force

![Fig. 9](image_url) Load/unload process: (a) force history and (b) components displacements during the load process, (c) force history and (d) components displacements during the unload process.
vibration during the end stage of the load process because of the high asperity force caused by the low FH. However, the FH has no significant fluctuation owing to the high stiffness. Figures 10(c) and 10(d) show the force and minimum FH histories during the unload process, respectively. Head 2 has a relatively high negative pressure that is advantageous for the stiffness, but makes the peeling off from the substrate difficult. It shows that head 3 has the lowest lift-off force of 5.8 mN compared with the values of 11.1 mN and 14.1 mN for the other two heads, implying that the unload performance of head 3 is the best among the three head designs.

Finally, based on the proposed the full-coupled SPAS model, a PFH analysis solver, SKLY.app (State Key Laboratory of Tribology + Fly) has been developed on the COMSOL Multiphysics Application Builder platform that can perform the analyses of the static, dynamic, and load/unload performance of the PFH or magnetic head in HDDs.

### 4 Conclusions

A novel full-coupled suspension-PFH-air-substrate (SPAS) model including the 6-DOF suspension-head dynamics, air bearing aerodynamics, and rotating substrate vibration for the flying performance analysis in RFNL, is proposed. The model exhibits a high accuracy and computing efficiency with a local pressure gradient-based adaptive mesh refinement algorithm. Three designs, heads 1, 2, and 3 of the ABS were investigated for the static, dynamic, and load/unload performance. Numerical simulations show that head 3 exhibits relatively good static and dynamic performances to achieve a ~15-nm near-field gap as well as has the best load/unload performance among the three ABS designs. Finally, a PFH analysis solver SKLY.app has been developed based on the proposed model. To achieve a finer lithography pattern, the design of ABS morphology with a near-field gap smaller than 10 nm and better load/unload performance must be further studied through the careful exploration of the dominant parameters.

### Acknowledgements

This work is financially supported by the National Natural Science Foundation of China (NSFC) with Grant No. 51635009.

**Open Access:** The articles published in this journal are distributed under the terms of the Creative
References

USA, 2005.


---

**Yueqiang HU.** He received his bachelor degree in mechanical engineering in 2013 from Southwest Jiaotong University, Chengdu, China. After then, he was a PhD student in the State Key Laboratory of Tribology at the Tsinghua University. His research interests include nanomanufacturing and nanophotonics.

**Yonggang MENG.** He is a professor in mechanical engineering, and serves as the director of the State Key Laboratory of Tribology (SKLT), Tsinghua University, China. Before he joined the SKLT in 1990, he obtained his master and PhD degrees in mechanical engineering from Kumamoto University, Japan, in 1986 and 1989, respectively. He is the author or co-author of over 160 peer-reviewed papers and 4 book chapters. His research area covers engineering tribology, surface and interface sciences, and micro/nanomanufacturing.
Friction and wear of sand-contaminated lubricated sliding

Mohamed Ahmed RAMADAN
Faculty of Engineering at Helwan, Helwan University, Helwan 11795, Egypt
Received: 24 December 2016 / Revised: 19 May 2017 / Accepted: 24 September 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: This paper reports a test investigation of friction and wear responses from sand contaminated lubricated sliding. The influence of sand contaminants on wear and friction is characterized. Analyses are completed utilizing segments of piston ring sliding against the cylinder liner. Paraffin oil, with and without sand contaminants, is used. The effects of the concentration and particle size of sand are examined.

Based on the observations in the present work, we found that friction and wear increase with sand concentration in the lube. Solid proposals ought to be considered, in order to enlighten the general population on the importance of changing a car engine’s oil filter regularly.

Keywords: sand contamination; lubrication oil; abrasive wear; lubricated sliding

1 Introduction

Motors in the Middle East experience serious wear as a result of the dusty atmosphere. The dust particles, sourced from outside contaminants such as sand, grits, chips, and abrasive materials, penetrate the lube. These contaminants transmit through the lube and deposit on the insides of the motor parts. The friction from such contaminants damages the moving contact surfaces, creating escalated wear with respect to the motor parts [1–3]. The concentration and particle size of these contaminants essentially impact the episode wear.

A few examiners had investigated the effectiveness of the contaminants and lubricant additives on friction and wear, utilizing reciprocating and rotating movement under high loading [4, 5]. Some have proposed the assessment of harsher wear by the “morphology” of the worn surfaces. The results showed how lubricant defiling causes critical wear toward the mating surfaces, thus causing component breakdown [6, 7].

Different studies have reported the results of abrasives defiled to lubricant versus friction and wear [8–10]. The results showed that the concentration of contaminants plays a pivotal role in controlling wear and friction along at the contact surfaces. In addition, images from a scanning electron microscope (SEM) showed abrasive wear as the predominant wear. This wear is a result of the hard surfaces of the contaminants sliding over one more of the mating surfaces [11–14]. The characterization of the wear elements and the “soot” that tainted the lube has been performed by different examiners [15, 16].

With an end goal to eliminate the breakdown of motor components, the present work explores the concentration and the particle size of sand defiling oil, utilizing segments of the piston ring sliding against the cylinder liner.

2 Experimental

The test rig consists of an electric motor which provides reciprocating motion for a segment of the cylinder line (300–3,100 strokes/min.). The electric motor was connected to an electric speed regulator to obtain the required test speed. The piston ring specimen was held in place in the groove of a piston segment clamped
by a chuck. The details of the test rig are illustrated in (Figs. 1 and 2). The piston ring specimen slides against the inner surface of the cylinder liner segment when paraffin oil, a lubricant, is used. The test ring specimens were prepared by cutting commercial piston rings into small segments. The piston rings were made of silicon cast iron whereas the cylinder liner was made of alloy cast iron. The friction force was measured using a load cell connected to a digital screen to detect the friction force, as shown in Fig. 2.

The test approach in this work was partitioned into two sections: The test specimens of the primary bit, which was subjected to fresh oil, was considered as a reference, while the specimens of the second partition were subjected to sand-contaminated oil. The sand was provided to the inner surface of the cylinder liner segment. The test process was carried out by dispensing the lubricant onto the sliding surfaces that contained sand. The effect of different concentrations of sand particles (2 g/L, 4 g/L, 6 g/L, 8 g/L, and 10 g/L) has been studied. These concentrations were chosen according to the high levels of dust in the Middle East. The sand particle sizes used were (0–80 μm), (80–200 μm), and (200–500 μm). All experiments were performed at room temperature (27 °C) and carried out at constant normal load (6 N), which applies the same side pressure on the cylinder wall. Further, all experiments were carried out at a constant speed (150 strokes/min), a constant running time (600 s), and a constant stroke (20 mm). The velocity of the piston segment is chosen to resemble the worst friction and wear conditions, at which the mixed lubrication regime prevails. The worst friction and wear conditions were found near the top and the bottom of the cylinder wall of the internal combustion engine where the velocity is small.

### 3 Results and discussion

Every experiment has been undertaken three times and the mean values were plotted. The effect of sand contamination in lubricant on friction was studied for different sand particle sizes. Figure 3 shows the effect of sand contamination in lubricant on friction for sand particles of size 0–80 μm. It is clearly shown that the friction coefficient increases with sand concentration in the lubricant. This can be attributed to the additional particles that came into contact thus increasing friction. Further, the high friction coefficient could be a result of possible partially metal/metal contact, as illustrated in Fig. 4. At the beginning of the trend where the lubrication oil is free of abrasives, the friction coefficient value was small (0.1). This can be attributed to the oil film that completely separates the surfaces, as illustrated in Fig. 5. At the end of the trend, the friction coefficient decreases slightly, which may be
Fig. 3 Effect of sand contamination in lubricant on friction for sand particles of size 0–80 µm.

Fig. 4 Illustration of wear mechanism.

Fig. 5 Illustration of wear mechanism.

due to the rolling action of the sand particles.

Fig. 6 Effect of sand contamination in lubricant on friction for sand particles of size 80–200 µm.

Fig. 7 Illustration of wear mechanism.

With even bigger sand particle sizes, i.e., 200–500 µm, the friction coefficient fluctuates, but less so with increasing sand content (see Fig. 8). This can be attributed to the irregular distribution of the oil layer due to particles passing through the contact. These variations increase with the increasing particle size of contaminants. Duplicate tests, which were run to establish reproducibility, showed that the trends were valid.

Using a sand concentration of 2 g/L, the effect of particle size on friction coefficient was studied. With the increase of sand particle size from (0–80 µm) to (200–500 µm), the values of friction fluctuations increase, as shown in Fig. 9. This may be due to higher embeddability of particles in one of the rubbing surfaces, and additionally the separation and elimination of worn surfaces. We noticed that the average value of friction coefficient increases with increasing sand particle size.

The contamination of lubricant due to sand particles is inevitable in the Middle East. These particles cause abrasive wear on the sliding surfaces. With regard to
the effect of sand concentration in lubrication oil on the wear rate of piston ring specimen, we found that the wear rate of the piston ring specimen increased with increasing sand concentration in lubricant, as shown in Fig. 10. In this case, increasing the sand concentration will cause an increase of material removal from the piston ring specimen due to the embedded abrasive particles, as well as the abrasive action. The embeddability of the rubbing surfaces is the most important factor in controlling abrasive wear. This is clearly observed by photomicrographs of the surface of the piston ring specimen, as shown in Fig. 11, where the damages caused by the sliding surfaces take place in the deep surface grooving due to the penetration of sand particles. Furthermore, it is clearly shown that the deep surface grooving increases with increasing sand concentration.

Figure 12 shows the effect of sand concentration in lubrication oil on the wear rate of the piston ring specimen, using sand particles of size 80–200 μm. The trend is the same as that shown in Fig. 10, in which the wear rate of the piston ring specimen increases with increasing sand concentration in lubrication oil. The presence of contaminants in the lubricant forms a monolayer of abrasive particles, which enter between the mating surfaces and cause wear when some of the particles are embedded in one of the rubbing surfaces, which then abrades the other surface.
Moreover, Fig. 13 shows an increase in the wear rate of the piston ring specimen with increasing sand concentration in lubrication oil. Table 1 shows a comparison of wear values of the piston ring specimen at different sand particle sizes and sand concentrations. It shows that wear values increases slightly with increasing abrasive particle size. This behavior might be attributed to the fact that as particle size increases, the depth of penetration of the particles into the sliding surfaces increases and consequently causes an increase in the volume of material removed. Further, it is clearly observed that the wear values increase significantly with increasing sand concentration in lubrication oil. The sand concentration in lubrication oil has a larger effect on wear than the particle size of sand, as shown in Table 1. Photomicrographs of sand show a relatively sharp shape for the different sizes of sand particles, as shown in Fig. 14.

For better clarity on wear, photomicrographs of the surface of the piston ring specimens were taken. Figure 11(a) characterizes the surface of the piston ring specimen tested using lubrication oil that is free of abrasives. Figure 11(b) shows the worn surface of the piston ring specimen caused by abrasive wear using lubrication oil with sand particles of size 0–80 μm and a concentration of 10 g/L. The worn surface of the piston ring specimen tested using lubrication oil with sand particles of size 0–80 μm and a concentration of 4 g/L is shown in Fig. 11(c). Figure 11(d) shows the worn surface of the piston ring specimen tested using lubrication oil with sand particles of size 0–80 μm and a concentration of 8 g/L. We noticed that the surface wear of the piston ring is accelerated by increasing the sand concentration in lubrication oil. The inner surface of the cylinder liner was also exposed to wear as a result of the abrasive action of the sand. The photomicrograph of the worn surface of the cylinder liner is shown in Fig. 15. The latter illustrates a wear mode in which the breakdown of the boundary lubricant film occurred due to the instability of the

<table>
<thead>
<tr>
<th>Sand particle size</th>
<th>Sand concentration in lubrication oil (g/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–80 μm</td>
<td>0</td>
</tr>
<tr>
<td>80–200 μm</td>
<td>1</td>
</tr>
<tr>
<td>200–500 μm</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1 Wear values of the piston ring specimen for different sand particle sizes and sand concentrations.
shear mixed layer, which then results in the severe plastic flow on the surface.

4 Conclusions

This paper presents the results of friction and wear from sand-contaminated lubricated sliding. The study of sand-contaminated engine lubricant is important in understanding machine and engine failures caused by severe wear and high friction rates, and is especially critical in the Middle East where issues caused by dust are prevalent.

Our results show that friction and wear increase with increasing sand particles concentration in lubricated sliding contact. This may be due to the increasing number of particles that come into contact with the surface, which then accelerates the friction. In addition, increasing the sand concentration will cause an increase in the material removal, or wear, of the piston ring specimen due to the embedded abrasive particles. The embeddability of the rubbing surfaces is the most important factor in controlling abrasive wear. Further, friction and wear rate increase with the increasing size of sand particles in the lubricated sliding contact.

This study affirms the effect of sand contaminated lubricant on friction and wear. The sand particles increase the friction between the sliding mating surfaces, which lead to “severe” wear, and subsequently machine element failure.

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Mohamed Ahmed RAMADAN. He received his PhD degree in mechanical engineering from El-Minia University, Egypt, in 2010. His current position is an assistant professor at Faculty of Engineering at Helwan, Helwan University, Egypt. His research areas cover the tribology and mechanical design, material science, and tribology of biomaterials.
Evaluation of lubrication performance of crank pin bearing in a marine diesel engine

Suk Man MOON\textsuperscript{1}, Yong Joo CHO\textsuperscript{1}, Tae Wan KIM\textsuperscript{2,*}

\textsuperscript{1} School of Mechanical Engineering, Pusan National University, San 30 Changjeon-Dong, Kumjeong-Gu, Busan 609-735, Republic of Korea
\textsuperscript{2} Department of Mechanical Engineering, Pukyong National University, San 100 Yongdang-Dong, Nam-Gu, Busan 608-739, Republic of Korea

Received: 14 May 2017 / Revised: 21 August 2017 / Accepted: 01 November 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: In this study, the lubrication performance of the crank pin bearing in a marine two–stroke diesel engine is evaluated to investigate the adhesional failure on the crank pin bearing. A numerical algorithm for the lubrication analysis considering motion analysis of the crank pin system is developed. The film pressure and thickness for three clearances and three lubricant temperatures are calculated. The results show that the lubricant temperature has a higher effect on film thickness than clearance. In terms of the film parameter, the operating condition that can result in solid–solid contact is investigated. We also suggest the desirable operating conditions of clearance and lubricant temperature to prevent the solid–solid contact.

Keywords: crank pin bearing; film thickness; lubricant film parameter; diesel engine

1 Introduction

The wear caused by insufficient lubrication is the most general cause of endurance life issues. An absence of lubrication in the journal-bearing system leads to bearing seizure, and normally, to total destruction of the part. The motivation of this study is an adhesional failure observed on the crank pin in a marine two-stroke diesel engine, as shown in Fig. 1. Insufficient lubrication caused by factors such as a machining error in the manufacture of the crank pin and the bearing leads to solid–solid contact between the crank pin and the bearing, which results in adhesional wear. Prolonged operation under such conditions also results in complete destruction of the part. The crank pin bearing which connects the connecting rod and crank arm, converting a reciprocating motion into a rotary motion plays an important role in a marine engine. Predicting the lubrication state of the crank pin bearing and the possible early bearing damage can ensure the normal operation of a marine diesel engine.

The investigation of the minimum oil film thickness in the engine bearing is important to prevent the above problems, and have been conducted in many studies [1–5]. The simulation models are classified as the methods based on hydrodynamic, elasto-hydrodynamic (EHD), and thermo-elasto-hydrodynamic lubrication theories [6–10]. Bukovnick et al. [11] compared the different models for the simulation of non-stationary response of the journal bearings used in combustion engines. Recently, coupling simulations between EHD and nonlinear multi-body dynamics (MBD) have been conducted for dynamically-loaded
bearings of a marine diesel engine. Wu et al. [12] developed an EHD coupling analysis model for the main bearing lubrication of a marine diesel engine using AVL Excite software, and explored the main bearing’s vertical load, orbital path and minimum oil film thickness.

In this study, a lubrication analysis was conducted to reveal the cause of adhesional failure on the crank pin bearing of a marine two-stroke diesel engine. We developed a numerical algorithm for the hydrodynamic lubrication analysis coupling with the motion analysis of the piston–connecting rod–crank arm system. The lubrication performance, including film pressure and thickness, were investigated for three clearances and three lubricant temperatures. The calculated film thickness was compared with the measured surface roughnesses of the crank pin journal and the bearing.

2 Analysis

2.1 Motion analysis of crank pin

Figure 2 shows a schematic of crank pin motion with the piston position. The bearing loads, which include the radial force \( F_R \) and the tangential force \( F_T \), and the relative velocity between the crank arm and connecting rod can be obtained from the force equilibrium equations.

The angular range of the crank arm angle \( \alpha \), the angular velocity of the crank arm \( \omega \), and the angular acceleration of the crank arm \( \ddot{\alpha} \) are expressed as

\[
\alpha = 0 - 2\pi
\]

\[
\omega(= \dot{\alpha}) = \frac{2\pi \cdot n}{60}
\]

\[
\ddot{\alpha} = 0
\]

where \( n \) is the engine speed. From the relation of the crank arm and connecting rod in Fig. 2(b), the angle \( \beta \), angular velocity \( \dot{\beta} \), and angular acceleration \( \ddot{\beta} \) of the connecting rod can be written as

\[
\beta = \cos^{-1} \left( \frac{r_{ca}}{r_{cr}} \sin \alpha \right)
\]

\[
\dot{\beta} = -\frac{r_{ca} \cos \alpha}{r_{cr} \sin \beta} \omega
\]

\[
\ddot{\beta} = \frac{r_{ca} \omega^2}{r_{cr} \sin \beta} \left( \sin \alpha - \frac{r_{ca} \cos^2 \alpha \cos \beta}{r_{cr} \sin \beta} \right)
\]

where \( r_{ca} \) and \( r_{cr} \) are the lengths of the crank arm and connecting rod, respectively. The equilibrium equations for the connecting rod in Fig. 2(c) are given as

\[
\sum F_x = m_c \ddot{a}_c; \quad F_{1x} - F_{2x} = m_c \ddot{a}_c
\]
where $m_c$ is the mass of the connecting rod, $g$ is the acceleration of gravity, and $I_c$ is the moment of inertia of the connecting rod. In addition, $\ddot{a}_x$ and $\ddot{a}_y$ are the accelerations at the center of mass, $F_{1x}$ and $F_{1y}$ are the reaction forces at point 1, and $F_{2x}$ and $F_{2y}$ are the loaded forces at point 2 from the piston, in the $x$ and $y$ directions, respectively. Here, assuming that $F_{2x}$ is negligible because the piston is constrained in the $x$ direction by the cylinder guide, $F_{2y}$ can be derived from the equilibrium equation for the piston in Fig. 2(d) as

$$F_{2y} = \sum F_y = m_p a_{yx} ; F_{1y} - F_{2y} - m_c g = m_c \ddot{a}_y$$

$$\sum M_i = I_{pc} \ddot{\beta} ; F_{1x} \cdot r_{c1} \sin \beta - F_{1y} \cdot r_{c1} \cos \beta + F_{2x} \cdot r_{c2} \sin \beta - F_{2y} \cdot r_{c2} \cos \beta = I_p \ddot{\beta}$$

where $F$ is the force applied by the gas pressure, $m_p$ is the mass of the piston, and $\alpha_1$ is the acceleration at point 1. Combining Eq. (10) with Eqs. (7)–(9), the following equations can be obtained:

$$F_{1x} - F_{2x} = m_c (a_1 \sin \alpha - r_{c1} \ddot{\beta} \sin \beta - r_{c1} \ddot{\beta} \sin \beta - r_{c1} \ddot{\beta} \cos \beta)$$

$$F_{1y} = m_c (-a_1 \cos \alpha + r_{c1} \ddot{\beta} \cos \beta - r_{c1} \ddot{\beta} \cos \beta + F_{2y} + m_c g$$

$$\frac{\partial h}{\partial t} = \frac{\partial}{\partial t} \cos \theta + e \frac{\partial}{\partial t} \sin \theta$$

where $c$ is the bearing radial clearance, $e$ is the journal center eccentricity, which is the distance from the bearing center $O_B$ to the journal center $O_J$, and $\phi$ is the journal attitude angle. In Fig. 3, $h$ is expressed in two coordinate systems ($\phi, Z$) and ($\theta, Z$), both located on the bearing surface. The angular coordinate $\phi$ has its origin on the line of $x$ axis, while the coordinate $\theta$ starts from the position of maximum film thickness.

The unsteady Reynolds equation [13] is solved numerically after incorporating the time-dependent squeeze effects. The unsteady Reynolds equation in a Cartesian coordinate ($x = r_0 \theta, y, z$) on the bearing plane is expressed as

$$\frac{\partial}{\partial x} \left( h^3 \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left( h^3 \frac{\partial p}{\partial z} \right) = 12U \eta \frac{\partial h}{\partial x} + 12\eta \frac{\partial h}{\partial t}$$

where $U_{ca}$ is the linear velocity of the crank arm and $U_c$ is that of the connecting rod.

### 2.2 Lubrication analysis

We simulated the crank pin bearing using the hydrodynamic lubrication model assuming both surfaces to be smooth. The calculated minimum film thickness was compared with the surface roughness measured in the actual bearing to determine the presence or absence of solid–solid contact. In this study, the influence of the oil hole or grooves was not considered. In the real crank pin bearing of a slow-speed diesel engine, the lubricant is supplied at pressures ranging from 2.1 bars to 2.7 bars. The oil hole is located on the top of the bearing and the oil supply pressure is much lower than the oil film pressure. Therefore, the assumption of the boundary conditions in this study will not have a significant effect on the analysis results. Figure 3 shows a schematic of the crank pin bearing. Considering the time-dependent crank pin eccentricities, the film thickness of the lubricant, $h$, and its time rate are written as

$$h = c + e \cos \theta$$

$$\frac{\partial h}{\partial t} = \frac{\partial}{\partial t} \cos \theta + e \frac{\partial}{\partial t} \sin \theta$$

Fig. 3 Schematic diagram of a dynamic loaded crank pin bearing.
where \( p \) is the film pressure, \( \eta \) is the lubricant viscosity, \( U \) is the mean velocity of the journal and bearing surfaces, and \( r_b \) is the bearing radius.

As a boundary condition for the lubrication region, the following Reynolds boundary condition are used:

\[
\begin{align*}
  p &= 0 \quad \text{at} \quad x_{\text{in}} \\
  p &= 0 \quad \text{and} \quad \frac{\partial p}{\partial \theta} = 0 \quad \text{at} \quad x_{\text{out}}
\end{align*}
\]  

(20)

The force balance equation, which determines the eccentricity \( e \) and attitude angle \( \phi \), is written as

\[
F_r = \int_0^w \int_0^{2\pi} p \cdot \cos(\varphi + \theta) \cdot r_j \cdot d\theta \cdot dz
\]

(21)

\[
F_t = \int_0^w \int_0^{2\pi} p \cdot \sin(\varphi + \theta) \cdot r_j \cdot d\theta \cdot dz
\]

(22)

where \( w \) is the bearing width, and \( F_r \) and \( F_t \) are the radial and tangential components of the lubricant film reaction force with reference to the \((R, T)\) coordinate system, respectively. We solved the Reynolds equation with our own program using the finite difference method. The lubrication code used in this analysis has been validated in our previous paper [14].

### 3 Results and discussion

Table 1 shows the specifications of the large end-connecting rod bearing and the operation conditions used in this study. We perform the lubrication analysis for three clearance of 240 \( \mu \)m, 200 \( \mu \)m, and 160 \( \mu \)m and three lubricant temperatures of 40 °C, 50 °C, and 60 °C. The lubricating oil temperature of the crank pin bearing in the large two-stroke slow-speed diesel engine, which is the subject of this study, is measured in the range of 40°–47°. Therefore, in this study, it is assumed that the lubricating oil temperature ranges between 40 °C and 60 °C, to include the measurement value as the analysis condition. The absolute viscosities of the lubricant are measured as 80.27 cP, 53.2 cP, and 37.22 cP at the temperatures of 40 °C, 50 °C, and 60 °C, respectively, as shown in Fig. 4. Note that the lubricating oil temperature is assumed to be constant in this analysis. In fact, the lubricant temperature measured during the engine operation indicates a slight variation in the range of several degrees. Therefore, this analysis result may cause some errors in an actual situation.

Table 1 Specifications of the large end-connecting rod bearing and operation conditions.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engine speed</td>
<td>rpm</td>
<td>146</td>
</tr>
<tr>
<td>Maximum gas pressure</td>
<td>bar</td>
<td>177.7</td>
</tr>
<tr>
<td>Piston area</td>
<td>m²</td>
<td>0.1233</td>
</tr>
<tr>
<td>Crank pin (journal) radius, ( r_j )</td>
<td>m</td>
<td>0.260</td>
</tr>
<tr>
<td>Crank arm length, ( r_{cr} )</td>
<td>m</td>
<td>0.885</td>
</tr>
<tr>
<td>Connecting rod length, ( r_d )</td>
<td>m</td>
<td>1.770</td>
</tr>
<tr>
<td>Crank pin bearing width, ( w )</td>
<td>m</td>
<td>0.165</td>
</tr>
<tr>
<td>Piston mass, ( m_p )</td>
<td>kg</td>
<td>589.1</td>
</tr>
<tr>
<td>Connecting rod mass, ( m_c )</td>
<td>kg</td>
<td>865.93</td>
</tr>
<tr>
<td>Moment of inertia of connecting rod, ( I_c )</td>
<td>kg·m²</td>
<td>28.1</td>
</tr>
<tr>
<td>Absolute viscosities at 40, 50, 60 °C</td>
<td>cP</td>
<td>80.27, 53.2, 37.22</td>
</tr>
<tr>
<td>Lubricant temperature</td>
<td>°C</td>
<td>40, 50, 60</td>
</tr>
<tr>
<td>Clearance, ( c )</td>
<td>( \mu )m</td>
<td>240, 200, 160</td>
</tr>
</tbody>
</table>

However, the results are meaningful in terms of being able to manage the critical temperature value of the lubricant which does not cause solid–solid contact.

We calculate the bearing loads and lubricant velocity by conducting motion analysis of the crank pin system to evaluate the lubrication state of the crank pin bearing. The piston gas pressure with crank angle under full load is measured from the marine two-stroke diesel engine as shown in Fig. 5. The maximum gas pressure is 177.7 bar at a crank angle of 13°. The crank pin bearing loads are calculated from the gas pressure and inertial force of the piston and the connecting rod. The calculated crank pin bearing loads in the radial and tangential directions with crank angle are demonstrated in Fig. 6(a). The maximum radial force is 1820.7 kN at a crank angle of 13°, which corresponds to...
to an upward movement of the piston, while the maximum tangential force is 45.7 kN at a crank angle of 313° which corresponds to a downward motion. Figure 6(b) shows the mean relative velocity of the crank pin bearing with crank angle using Eqs. (14) – (16). The mean relative velocity is applied as a lubricant velocity in the lubrication analysis.

Figure 7 shows the orbital path of the crank pin bearing for a clearance of 240 μm and a lubricant temperature of 40 °C. It is shown that the largest eccentricity occurs at a crank angle is 19°, while the smallest eccentricity occurs at a crank angle of 325°. Figure 8 shows that 3D pressure and film thickness distributions at the crank angle of 19° for the crank pin bearing for the clearance of 240 μm and lubricant temperature of 40 °C.

Figure 9 shows the variations in film pressure and thickness with the crank angles for the crank pin bearing for the three clearances at the lubricant
temperature of 40 °C. As shown in Fig. 9(a), the film pressure decreases with a decrease in clearance, and the maximum film pressures for the clearances of 240 μm, 200 μm, and 160 μm are 227.6 MPa, 197.4 MPa, and 168.6 MPa, respectively. The film thickness increases with a decrease in clearance, and the minimum film thicknesses for the three clearances are calculated as 4.98 μm, 5.44 μm, and 5.94 μm, respectively.

The variations in film pressure and thickness for the crank pin bearing for a clearance of 240 μm at three lubricant temperatures are investigated in Fig. 10. It is shown that a higher lubricant temperature gives higher film pressure and lower film thickness. The maximum film pressures for the lubricant temperatures of 40 °C, 50 °C, and 60 °C are calculated as 227.6 MPa, 269.6 MPa, and 309.2 MPa, and the minimum film thicknesses are calculated as 4.98 μm, 3.46 μm, and 2.41 μm, respectively. Note that the effect of lubricant temperature on the film pressure and thickness is higher than that of clearance. Compared to the variations in film thickness for the clearance shown in Fig. 9(b), the film thickness in Fig. 10(b) decreases significantly with an increase in lubricant temperature. This tendency can be seen in more detail in Fig. 11, which summarizes of the minimum film thickness at three lubricant temperatures for the crank pin bearing for three clearances. It is found that the lubricant temperature has a larger effect on the minimum film thickness.

In addition, we evaluate the lubricant film parameter, \( \Lambda \) of the crank pin journal and bearing to observe the failure possibility of the crank pin. The lubricant film parameter is defined as \( \Lambda = h_{\text{min}} / \sigma = h_{\text{min}} / \sqrt{R_{q1}^2 + R_{q2}^2} \), where \( \sigma \) is the composite standard deviation of surface heights of the two surfaces, and \( R_{q1} \) and \( R_{q2} \) are the root mean square roughnesses of the two surfaces [15, 16]. Generally, if the average
Fig. 11 Minimum film thickness with three different lubricant temperatures for the crank pin bearing with three different clearances.

film thickness is less than three times the composite surface roughness, the surface asperities will force direct solid–solid contact [15, 17]. In this study, we measure the surface roughness of the crank pin journal and bearing. The measured $R_q$ and $R_{q2}$ are 0.445 $\mu$m and 0.887 $\mu$m, respectively, which leads to $3\sigma = 2.978$ $\mu$m. The line of $3\sigma$ is also marked in Fig. 11. It is estimated that the minimum film thickness of the cases for the clearances of 200 $\mu$m and 160 $\mu$m at the lubricant temperature of 60 °C are less than $3\sigma$, resulting in solid–solid contact, which can cause the failure of the crank pin. It is concluded that a lubricant temperature of less than 50 °C is required to prevent solid–solid contact. Confirmatory experiments for the effects of surface roughness and lubricant temperature are underway for the diesel engine, and we expect to obtain good results in the near future.

4 Conclusions

In this study, to investigate an adhesional failure on the crank pin bearing of a marine two-strokes diesel engine, a lubrication analysis was conducted on crank pin journal and bearing. Through the motion analysis of the piston–connecting rod–crank arm system, the bearing loads and lubricant velocity were calculated. Then, the lubrication analysis for three clearance of 240 $\mu$m, 200 $\mu$m, and 160 $\mu$m at three lubricant temperatures of 40 °C, 50 °C, and 60 °C was performed. The results showed that the maximum film pressure decreased with decreasing clearance and lubricant temperature, and that film thickness increased with decreasing clearance and lubricant temperature. The lubricant temperature had a higher effect on the film pressure and film thickness than the clearance. In the evaluation of the lubricant film parameter, the minimum film thicknesses of the cases for the clearances of 200 $\mu$m and 160 $\mu$m at the lubricant temperature of 60 °C were estimated to be less than $3\sigma$, which resulted in solid–solid contact of the crank pin bearing. In addition, we concluded that the management of the clearance is important to prevent solid–solid contact, but the operation of the crank pin bearing at a lubricant temperature of less than 50 °C is more desirable for a sufficiently lubricated film.

Acknowledgement

This work was supported by the Pukyong National University Research Abroad Fund in 2015 (C-D-2015-0500)

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


Suk Man MOON. He received his master degree in mechanical engineering in 1999 from Pusan National University, Korea. He is a Ph.D. student in the Tribology Laboratory at the same university. His research areas cover surface engineering, lubrication, friction and wear for engine.

Yong Joo CHO. He received his Ph.D. degree in mechanical engineering from Korea Advanced Institute of Science and Technology, Korea in 1994. He has worked in Pusan National University since 1979. His current position is a professor of mechanical engineering department and the director of the Tribology Laboratory. His research areas cover surface topology, gear lubrication, and friction and wear for mechanical device.

Tae Wan KIM. He received his Ph.D. degree in mechanical engineering from Pusan National University, Korea in 2002. He has worked in Pukyong National University since 2008. His current position is a professor of mechanical engineering department and the director of the Nanotribology Laboratory. His research areas cover surface modification, lubrication, friction and biomimetic engineering.
Insight into tribological problems of green ship and corresponding research progresses

Yuwei SUN¹ ² ³, Xinping YAN¹ ² ³ ⁴ *, Chengqing YUAN¹ ² ³, Xiuqin BAI¹ ² ³
¹ School of Energy and Power Engineering, Wuhan University of Technology, Wuhan 430063, China
² Reliability Engineering Institute, National Engineering Research Center for Water Transport Safety (WTS Center), MoST, Wuhan University of Technology, Wuhan 430063, China
³ Key Laboratory of Marine Power Engineering and Technology, MoT, Wuhan University of Technology, Wuhan 430063, China

Received: 02 October 2016 / Revised: 17 April 2017 / Accepted: 30 July 2017
© The author(s) 2017. This article is published with open access at Springerlink.com

Abstract: The so-called “green ship” is being regarded as a potential solution to the problems that the shipping industry faces, such as energy conservation and environmental protection. Some new features, such as integrated renewable energy application, biomimetic materials, and antifriction and wear resistant coating have been accepted as the typical characteristics of a green ship, but the tribology problems involved in these domains have not been precisely redefined yet. Further, the related research work is generally focused on the technology or material itself, but not on the integration of the applicable object or green ship, marine environment, and tribological systematical analysis from the viewpoint of the energy efficiency design index (EEDI) and ship energy efficiency management plan (SEEMP) improvements. Aiming at the tribology problems of the green ship, this paper reviews the research status of this issue from three specific domains, which are the tribology problems of the renewable energy system, tribological research for hull resistance reduction, and energy efficiency enhancement. Some typical tribological problems in the sail-auxiliary system are discussed, along with the solar photovoltaic system and hull drag reduction in traditional marine mechanical equipment. Correspondingly, four domains that should be further considered for the future development target of the green ship are prospected.

Keywords: green ship; tribology; renewable energy; biomimetic material; ship drag reduction

1 Introduction

The shipping industry is a significant segment of the comprehensive transportation system. According to the report “Review of Maritime Transport”, it was stated that the total amount of global maritime trade was 7.84 billion tons in 2009, and a market share of 80% of the global flow of goods was transported by sea [1]. In addition, it was estimated in 2015 that the world seaborne trade volume surpassed 10 billion tons [2] for the first time in the records of the United Nations Conference on Trade and Development (UNCTAD). However, once regarded as the most environmental friendly mode of transport, as global warming and environmental pollution are considered problems of globalization, the shipping industry has inevitably been pushed into the focus point of public opinion. Statistical data shows that the volume of greenhouse gases, polluting gases, and particle contamination (CO₂, SOₓ, NOₓ, and PM) discharged by ships is growing distinctly in recent years, and other problems such as the uneven distribution of energy resources, growing energy demand, and high fuel costs are all a direct pressure forcing the shipping industry to explore a new direction for its future development. In this context, the targets of energy conservation and

* Corresponding author: Xinping YAN, E-mail: xpyan@whut.edu.cn
emission reduction are widely accepted as a kind of social responsibility by the global shipping enterprises, and gradually, some new technologies and/or materials that have the typical green factor properties are applied on ships for this purpose [3]. These technologies include, for example, renewable technologies (wind power, solar power, fuel cells, gas as fuel, etc.), low-carbon lubrication materials, tribological and biomimetic materials, and antifriction and wear resistant coating. However, the related research work is generally focused on the technology or material itself, but not on the integration of the applicable object or green ship, marine environment, and tribological systematical analysis. In this paper, the tribology problems of the green ship are classified into two categories, external and internal friction, from a systematical standpoint, and the corresponding research progress is discussed from three specific domains, which are the tribology problems of the green ship, tribological research for hull resistance reduction, and energy efficiency improvement.

2 Green ship and tribology

The concept of green ship, which firstly appeared in the mid-1990s, means the application of advanced technology to satisfy the targeted function and performance of mechanical equipment, improve the efficiency of energy utilization, and reduce and/or eliminate environmental pollution during the total life cycle of the ship. Further, the green ship designed and built based on this innovative concept must provide the crews with proper personal protection. To promote the development process of the green ship, the International Maritime Organization (IMO) and the classification societies in each country (e.g., China Classification Society, LR Lloyds Register of Shipping, and NK Nippon Kaiji Kyokai) have established a series of industry guides, specifications, and technical standards [4]. Two vital standards, the energy efficiency design index (EEDI) and the ship energy efficiency management plan (SEEMP), aiming at compulsorily implementing the policy of greenhouse gases emission for all ships, were authorized by the Marine Environment Protection Committee (MEPC) of the IMO in its 62nd session. Afterwards, many classification societies have started to push forward their own criteria, as follows: (1) Environment and energy efficiency grade scheme, by Det Norske Veritas (DNV); (2) independent certificate of green passport for both new and operating ships, by Lloyd's Register group (LR); (3) environmental service system, by Germanischer Lloyd group (GL), for the special service of customers' environmental requirements [5]; (4) the China Classification Society (CCS) has set a special mark and divides it into three levels according to the green factors included, namely, the additional certification of green ship. Besides, various kinds of frontier and advanced concepts have been used to design and build the green ship, wherein many technical fields are being involved, from the optimization of hull structure to high efficient power and propulsion system, and from the high strength ship steel to coating that is pollution-free. It is believed that the ship that integrates the green factors or green technologies all together will become the development trend for green ships in the near future.

The tribology theory, which concentrates on the contact mechanics of moving interfaces, which generally involve energy dissipation [6], has been widely utilized to analyze the energy efficiency and economic feasibility of a ship. According to the study report “Second IMO GHG Study 2009” [7], the research data based on a contrastive case study was represented to understand the relationship between energy losses and friction effect that is involved in marine equipment. The main reasons leading to energy losses are classified as seven factors among four typical ship forms, as it is shown in Fig. 1. About 37.6% (mean value) of the energy from the fuel is spent to overcome diverse types of resistance, which have direct and/or indirect relationship to tribological problems. Comprising air and hydraulic resistance, the hull resistance is the main cause among the eight cases. Its maximum percentage, 17.2%, corresponds to the second case. The propulsion transmission losses, propeller frictional losses, and wave generation are the areas that constitute the second highest proportion. In the fifth case, its peak value accounts for 19.2% of the total energy losses, even more than that of the hull resistance. Moreover, the factors of weather and waves should be taken into the tribology considerations regarding energy
losses, but its total percentage may be directly changed by utilizing a reasonable navigation program. Ships other than the eight cases shown in Fig. 1 have the same types of energy dissipation. The only difference is that their relative values will differ. The most notable point is that the research target of tribology is being changed from the traditional concept of friction control (wear reduction and lubrication improvement) to that of energy efficiency and environmental friendliness [8], e.g., energy conservation, emission reduction, low noise and vibration, ecological lubrication, and high quality of life.

In view of the issues mentioned above, the tribology problems involved in the green ship technology can be divided into two parts, which are defined in this paper as the external and internal friction. The former is located at the interface between the hull and water, whereas the latter points to the frictions that occur in the main engine, auxiliary machinery, and shafting system. The key issues included in these two categories can be further classified and studied in three main aspects, as shown in Fig. 2.

### 3 Tribological problem of renewable energy system

New energy is not a kind of energy source, but rather a technical system to utilize the renewable energies from the viewpoint of cleanliness, efficiency, and systematization. Some typical new energy technologies have shown their potential advantages not only in experimental research but also for commercial applications serving as demonstration cases [9].

Figure 3 shows the successful utilization of some typical new energy systems in the green ship, such as wind, solar, nuclear power, fuel cells, tidal power, hydrogen, and biomass. However, some critical issues still need to be paid attention to, especially when dealing with the promotion of the green ship program based on the concepts of systematization and standardization, e.g., the characteristics and utilization method of each energy type vary if applied in different ship platforms, as mentioned above. Other influencing factors, such as hull structure, power and electrical system, crew configuration, navigation area, and port logistics, should be analyzed comprehensively according to the extent of their potential effect, and be arranged during the designing and building process.

From a systematic perspective, the majority of primary energy sources (new or renewable energies) cannot be directly utilized as the power driving the electrical and propulsion systems, unless they are transformed into secondary energy sources by an energy conversion device (ECD). Consequently, it is the ECD where the tribology problems arise in the green ship. Compared with the terrestrial environment, the
marine environment is severe, changeful, and corrosive. It results in a more complicated boundary condition and multi-factor influence in the tribology issues. Taking the ECD of the wave energy as an example, it shows that different kinds of frictional components are used to transform the irregular movement of the wave into a reciprocating and/or rotary motion, such as hinge joints, bearings, and matching parts of a piston. The research accomplished in marine environment has confirmed that some typical tribological features arise in these friction components, e.g., lubrication failure, overload, abrasive particle deposition, marine biofouling [10], cavitation, and friction abrasion. For the wind and solar energy, the tribological performance of moving surfaces contained in the corresponding ECD presents completely different characteristics.

3.1 Tribology problem of sail-auxiliary technology

Up to now, the sail-auxiliary technology that is applied in ships has more than 100 types, and the structure differentiation of the wind sail is remarkable. Some typical wind sails, as those shown in Fig. 4, can be further defined into four categories: (1) Sail wing, also known as airfoil sail, which is designed based on the principle of airfoil profile, including: rectangular sail wing, triangular sail wing (or Jib for short), walker, and sky sails; (2) drum sail, which applies the “Magnus Effect” to generate the propulsive force; (3) integration of the principle of airfoil profile and Magnus effect, named as the rotor-wing combination; (4) turbine sail of suction type, which is developed from the theory of controlled boundary layer separation.

Figures 5 and 6 show two kinds of foldable sails operating processes, which are the foldable rigid sails and foldable lifting sails, designed by the New Energy Group of Wuhan University of Technology (WUT). The functions of furling, stretching, and angle adjustment are all included. On a long voyage, the changeable wind load acting on the sails and the rubbing wear arising from the contact surface between the articulated mechanisms and sail’s horizontal ribs are two factors that influence the surface friction coefficient and reduce the mechanical efficiency of the driving system. Various studies suggest that an automatic lubrication system should be applied to the linkage mechanism, for example, using the oil-cup pressure oiler of the straight-through type to circulate the lubrication across the hinges’ joints. The wind sail and mast may be destroyed by the corrosive seawater under a long-time exposure to the marine environment, which will definitely change the aerodynamic characteristics of the sail-auxiliary system. The corrosion-resistant material and antifriction and

![Various types of sail-auxiliary ships.](image-url)
hard-wearing coating are two methods to prolong its service life. Directly acting on the mast through the wind sail, the time-varying wind force is transferred to the driving gear that is installed on the mast. This force load is the main factor that leads to the vibration and fretting wear of the components inside the driving gear. Three kinds of tribology problems should be taken into consideration, including the fatigue failure of bearing, wear, and agglutination of the gear tooth surface.

For operation security and port logistics, the marine ships are required to be operated at a fixed speed, about 80%–90% of full load, also called the economic speed. If the sail wing system is put into service, the main power system should be reduced to 40% or lower in order to keep the actual speed unchanged. However, keeping the crankshaft at a low speed for a long period is harmful for the combustion efficiency of a large marine low-speed diesel engine, because the emerging carbon deposits will lead to material degeneration on the piston surface, as shown in Fig. 7, and on the inner surface of the cylinder. Other problems, such as poor lubrication, resistance increase, and cylinder wall scratching, are all unavoidable, and the worst condition is a broken-down engine. Therefore, the friction and wear characteristics of the moving surfaces located in the main engine components should be studied comprehensively, especially under the low-load working condition with the sail in use.

3.2 Tribology problem of solar photovoltaic system

The solar photovoltaic (PV) system is an approach for generating electrical power by converting the solar radiation into direct current electricity using semiconductors that utilize the photovoltaic effect. To realize the target of high energy-conversion efficiency, two methods of sun tracking (single and dual axis tracking) can be applied to a terrestrial PV system. As a comparison, various types of innovative structures are used to install the PV system on solar ships, as shown in Fig. 8. Especially, the form of “Solar Sailor” has been widely accepted and installed on catamarans [11], whereas the horizontal arrangement of solar panels is the best choice for small crafts or marine ships. Two reasons are considered: (1) Collection of as much solar radiation as possible under the limitation of deck area; (2) the increased weight coming from the solar panels, mounting bracket, and batteries.
(if used) has a negative influence on the vessel’s metacentric height. Moreover, the structure of the solar sailor may worsen this problem.

### 3.2.1 Friction and wear

The I–V characteristic of a solar cell is directly affected by meteorological factors, including solar radiation, temperature, and wind, while other factors such as gaseous pollutants, air dust (smoke and dust particles and salt particles), and rainfall, can influence the PV cell’s reliability in different ways [12–15]. Various studies have shown that the seawater is one of the leading factors for electrochemical corrosion process on the cell’s surface, and the salt particles emerging during the evaporation process of seawater can change the surface morphology of the PV cell’s cover glass by friction and wear [16]. The results shown in Fig. 9 are the salt-water experiments by using synthetic seawater, set at four different concentrations. In each group, the upper picture shows the solar cell that is covered with salt particles by evaporating the seawater and the lower one is the cell’s surface morphology when the salt particles are cleaned away by fresh water. Further study also shows that some black spots appear at the upper surface of the cover glasses, which result in the reduction of spectrum transmittance.

### 3.2.2 Vibration and fretting wear

The inverter is an electrical device that converts direct current (DC) to alternating current (AC). In the PV system installed in a ship, the hull and low frequency vibrations caused by the waves and rotating machinery, respectively, are the major factors that lead to vibration and fretting wear of the mechanical connection parts in the inverter. Besides, the performance characteristics of electronic components and switch devices running under a multifactor effect, such as corrosive spray and oil mist suffusing

---

**Fig. 8** Various types of solar ship.

**Fig. 9** Salt-water experiment of cell glass cover conducted under simulated marine environment.
in high temperature, also need to be paid comprehensive attention.

4 Tribological research for hull drag reduction

The hull resistance, which is composed by the frictional, stick pressure, and wave-making resistances, can be regarded as an acting force relating to the moving object in the fluid. Owing to the terminal conditions, including the viscosity and density of the fluid, surface roughness, and snit-fouling coating, the fluid boundary layer appears on the wetted surface of the hull, below the waterline. The longer the ship length is, the thicker the boundary layer becomes. All the kinetic and potential energy that the boundary layer acquires during the navigation comes from the hull. This physical effect of energy exchange is also the frictional resistance, which is the major of the three components of hull resistance. It is claimed that the proportion of frictional resistance in the hull resistance is up to 70%–80% [17], especially for the low-speed and medium-speed vessels. Numerous approaches aiming at the reduction of the frictional resistance of the green ship have been researched, including bionic materials and air bubble lubrication technology.

4.1 Hull surface roughness and biomimetic surface

The seawater fouling and corrosion, shown in Figs. 10 and 11, respectively, are the two factors that change the hull surface roughness, even after a short period of voyage. The direct results of the deterioration on the hull surface are that the route speed is reduced by 1–1.5 knots, while the energy consumption of the ship is kept at the same level as before. Under the worst situation, this reduction can reach 2 knots. Nevertheless, the idea that the resistance of a smooth surface is lower than that of a rough plane was widely accepted till the mid-1960s, when experiments showed that it is credible only under the constrain condition of low-speed fluid. On a smooth surface with high-speed fluid flowing over, the distribution of the velocity and pressure fields is more uneven than that with low-speed fluid. The momentum interaction of the turbulent boundary layer on the contact surface leads to an obvious increase in hydraulic resistance. Thus, there is a dilemma.

Effective approaches from ship antifouling to drag reduction have been explored in depth, based on the innovative tribology theory that involves biological, conductive, alkaline, and bio-enzyme antifouling, the technology of low surface energy [18], chemical and structural bionics, etc. The antifouling coating containing organic tin and cuprous oxide (both toxic) has proved to be harmful to the marine environment and ecological system. The bionic coating that applies the technologies of low surface energy and biological surface microstructure is more adapted and green for the marine environment. In order to reduce the skin friction of underwater vehicles, Dou et al. [19] developed a kind of binary structured surface based on Benard convection. The test conducted in the small and high-speed tunnel SKLT-1 of Tsinghua University has proved its remarkable effectiveness for drag reduction.
performance. The study on bionics assumes that the structure of rib-like squama that covers the shark body effectively prevents the surface from being adhered by sea creatures, and especially, reduces the vibration owing to the fluid flowing at a high relative speed. After the process of abstraction, amplification, and simplification based on the real microstructure of the shark squama, the bionic shark skin made in laboratory can currently contribute to about 7% of drag reduction in a test subject [20–22]. Using the forming technology of biological replication, Han et al. [23] designed a kind of bionic shark skin and claimed that this structure can improve the ratio of drag reduction up to 24.6%. Bai et al. [24] discussed an ecological concept to restrain the adhesion of sea creatures for ship application. Figure 12 shows some representative research contents by Bai. The shell at the upper left of the photograph in Fig. 12(a) is real and the three-dimensional print at the upper right is its microstructure. The data are measured by Phase Shift Micro XAM-3D, as presented in Fig. 12(b). In comparison, the shell at the bottom left is an artificial shell together with the three-dimensional print of its microstructure on the bottom right. The comparison of the three-dimensional prints in Samples 1 and 2 illustrates that this modeling technology can be applied to the hull surface when the most appropriate shell structure is selected.

4.2 Application of micro-bubble drag reduction

At the end of the 19th century, W. Froude established the assumption that the hull resistance could be reduced by injecting air into the fluid layer between the ship bottom and water, and the basic model is shown in Fig. 13. The initial experimental result was not as ideal as expected owing to the limited theoretical research and experimental conditions, but this innovation concept was regarded as the beginning for further research. In 1984, the American scholar Madavan [25] analyzed the micro-bubble drag reduction of solid revolution and claimed that the frictional resistance can be cut down by 50%–80%, which also proved the correctness of the research conclusion stated by Migirenk (USSR) [26]. The result also showed that: (1) For the ship operating at the same speed, the more air bubble ACS is injected, the lower the frictional resistance becomes; however, the frictional resistance between the hull surface and water remains unchanged when it reaches a critical value, no matter how much air is subsequently charged. (2) The effect of drag

Fig. 12 The shell and its artificial model with microstructure respectively in comparison.
reduction decreases along with the increase in ship’s speed, if the same volume of air bubble is injected.

(3) The air bubble can hardly influence the flow resistance characteristics at the upstream region that is located in the front of the air nozzle, as well as the velocity distribution of the mixed media.

A case study on the 116-m training ship SEIUN-MARU installed with ACS claimed that a 5% (maximum) reduction of skin friction and 2% of net energy saving (NES) was obtained in 2001. Later, the full-scale test conducted on the Pacific Seagull confirmed 5% of NES [7]. However, some reports seem to have a conflicting opinion, explaining that the ACS leads to an increase of 1%–2% in the hull resistance and propulsion efficiency at model scale and about 0.6% of net power reduction at full scale.

In China, the study on the micro-bubble drag reduction began in the 1980s, but the progress in research is remarkable. Dong et al. [27] studied the micro-bubble for both the plate and planning craft. Four factors are discussed in detail, including air velocity, volume, ejecting mode, and size of the air nozzle.

5 Tribological research for energy efficiency enhancement

The energy efficiency is one of the main factors in the evaluation system of a green ship, i.e., to accomplish the maximization of output power when the same amount of energy resource is consumed during navigation. According to the new explanation certified by IMO, as mentioned above, two correlative aspects EEDI and SEEMP are taken as guidelines for energy efficiency enhancement of the newly built ships and ships in operation, respectively. Aiming at the improvement of energy efficiency and reliability, the tribological research referring to the green ship is stated from two aspects, as follows.

5.1 The marine power plant

The marine power plant is the key component of ships, in which any failures would cause a critical accident that threatens their safety. Because the marine environment is complicated and severe, the dynamic movement of ships during voyages can be changed by the meteorological conditions, e.g., wind, storms, and ocean currents. The fluctuation range and type of dynamic movement can directly weaken the lubrication action in the mechanical components, aggravate mechanical wear, and finally affect the power characteristics and mechanical efficiency of the dynamic system [28, 29].

The cylinder lubrication is essential for the normal operation of the main engine, and it is indicated that about 26% of mechanical failures in 645 cases were caused by poor lubrication in the cylinders, while 92 cases (14%) were the result of a wrong method of lubrication. Aiming at developing methods to protect the cylinder liner of the main engine from excessive wear, Xiong [30] stated a maintenance strategy comprising four parts, namely cylinder oil, fuel injection rate, pretreatment, and cooling water. Under a simulated experiment condition, Han et al. [31, 32] studied the internal relationship between four operating parameters (load, speed, temperature, and running period) and two relevant parameters, mass loss and surface roughness, which are in connection with the assembly of cylinder liner and piston ring. For comparison, this assembly was also processed by a surface treatment to analyze its variation in tribological characteristics. Liu et al. [33] assumed that a kind of micro-geometrical structure had a great impact on the lubricating property of the oil layer between cylinder liner and piston ring. Processing the micro-geometrical structure of pits and slots at the liner’s inner surface, they analyzed the vibration frequency of the cylinder liner and the distribution of lubricant under the working condition, and explained that this structure can appropriately improve the lubrication
capability while the diesel engine was operating at a high rotating speed.

5.2 The transmission system

The stern bearing is used to support the stern shaft that transmits the power from the main engine to the propeller. Compared with other mechanical parts, its working condition is harsh. The reason is that the hydrodynamic lubrication in the stern bearing can hardly be formed owing to uneven pressure distribution and mechanical vibration, especially under the overload condition. From the China Shipbuilding Industry Corporation (CSIC), Wang [34] analyzed the correlation between friction coefficient and rotating speed by lubricating the stern tube bearing with water and oil respectively. Wang et al. [35] stressed the abnormal working condition of the mechanical end face seal caused by the integrated effect of seawater pressure and friction force. The method of overall contact coupling was used to deal with the steady-state analysis of temperature distribution related to the friction pair of the stern shaft sealing. The research work conducted by WUT is focused on the tribological properties of a water-lubricated stern tube bearing that operates in a simulation test bench, as shown in Fig. 14 [36, 37], in which five domains are contained: (1) The derivation and solution of Reynolds equation that take both the speed of lubricating water and obliquity of propeller axis into consideration; (2) the finite element simulation for fluid-structure coupling related to the shaft, bearing, and lubricant (water-lubricated); (3) the study on the inner flow field of the stem bearing and lubricating mechanism by CFD and numerical analysis; (4) the evaluation of the performance of the water lubricated stem bearing based on the experimental parameters, such as the pressure of water film, temperature, and friction coefficient; (5) the analysis of the reliability of the water lubricated stern tube bearing from two aspects, the wear and fatigue life. Besides, it is claimed that other tribological problems can be conducted for further studying on the transmission system, as follows: (1) The theoretical improvement in boundary lubrication, fluid hydrodynamic lubrication, and elastic fluid dynamic pressure for the water lubricated stern tube bearing; (2) the friction coefficient of bearing under the condition of different loads and linear velocities, and the lubricating property related to the size and gradient of the lubricating water channel; (3) the development of new materials that have the characteristics of low friction, wear resistance, self-lubrication, and high reliability.

6 Conclusions and prospects

For further development of the green ship, the tribological property analysis is one of the critical success factors, which implies the utilization of the tribology theory in the entire cycle of shipping industrialization, including design, manufacture, operation, and management (the four parts of the life cycle) for the purpose of energy conservation, environmental protection, and efficiency improvement. By setting the applicability, energy efficiency, and reliability of the green ship as the target for its future development, four domains should be further considered:

(1) Exploration of more extensive renewable energy technologies for application to the green ship.
(2) Improvement of the conversion efficiency (the primary energy) of the renewable energy system that is integrated in the green ship by exploring the friction and wear characteristics of the moving components in the mechanical system and the contact surfaces between the liquid-solid two-phases, especially taking the factors of marine environment into account.
(3) Application research on advanced materials to prolong the service life of rubbing pairs in the mechanical system, including low-carbon lubrication...
materials, biomimetic materials, and antifouling coating.

(4) Research on the tribological problems of the traditional marine mechanical equipment, applicable to the green ship, from the viewpoint of EEDI and SEEMP improvement.

Acknowledgements

This work is supported by National Natural Science Foundation of China (Grant Nos. 51422507 and 51509195).

Open Access: The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

[24] Bai X Q, Yuan C Q, Yan X P, Liu X M. Research on green bionic ship antifouling techniques based on surface


Yuwei SUN. He received the Ph.D. degree in 2013 from Wuhan University of Technology, and then joined the Reliability Engineering Institute of National Engineering Research Center for Water Transport Safety in WUT. He is currently an associate professor and master's supervisor in School of Energy and Power Engineering. His research interests include renewable energy technologies applied in green ship and the characterization of green materials for various applications.

Xinping YAN. He received his B.E. and M.E degrees from Wuhan Institute of Water Transportation Engineering in 1982 and 1987, respectively, and got his PhD degree from Xi’an Jiaotong University in 1997. Currently, he is the chair professor and the director of National Engineering Research Center for Water Transport Safety in WUT. He has been engaged in the research of tribological system engineering, condition monitoring, and fault diagnosis of mechanical wear failure, and ship intelligence and reliability.
Stress-augmented thermal activation: Tribology feels the force / 1–31
Hugh SPIKES

Anti-loosening performance of coatings on fasteners subjected to dynamic shear load / 32–46
Junbo ZHOU, Jianhua LIU, Huajiang OUYANG, Zhenbing CAI, Jinfang PENG, Minhao ZHU

Chaotic characteristics and attractor evolution of friction noise during friction process / 47–61
Cong DING, Hua ZHU, Guodong SUN, Yuankai ZHOU, Xue ZUO

Thermocapillary migration mechanism of molten silicon droplets on horizontal solid surfaces / 62–74
Tao SUN, Cunhua JIANG, Jianning DING, Ningyi YUAN

Synchronous measurement of tribocharge and force at the footpads of freely moving animals / 75–83
Yi SONG, Zhouyi WANG, Jun ZHOU, Yang LI, Zhendong DAI

Chemical effects on the sliding friction of Ag and Au(111) / 84–97
H. E. KO, S. G. KWAN, H. W. PARK, A. CARON

Transfer of electrons on scratched iron surfaces: Photoelectron emission and X-ray photoelectron spectroscopy studies / 98–115
Yoshihiro MOMOSE, Daisuke SUZUKI, Keika TSURUYA, Takao SAKURAI, Keiji NAKAYAMA

Black phosphorus as a new lubricant / 116–142
Wei WANG, Guoxin XIE, Jianbin LUO
Vol. 6, No. 2

Review Article

Adhesion and surface forces in polymer tribology—A review / 143–155
Nikolai MYSHKIN, Alexander KOVALEV

Research Articles

Film forming behavior in thin film lubrication at high speeds / 156–163
He LIANG, Dan GUO, Jianbin LUO

Isosteric design of friction-reduction and anti-wear lubricant additives with less sulfur content / 164–182
Xinlei GAO, Denghui LIU, Ze SONG, Kang DAI

Combined effect of the use of carbon fiber and seawater and the molecular structure on the tribological behavior of polymer materials / 183–194
Zhiqiang WANG, Jing NI, Dianrong GAO

Study on the influence of standoff distance on substrate damage under an abrasive water jet process by molecular dynamics simulation / 195–207
Ruling CHEN, Di ZHANG, Yihua WU

Tribological evaluation of environmentally friendly ionic liquids derived from renewable biomaterials / 208–218
Cheng JIANG, Weimin LI, Jingyan NIAN, Wenjing LOU, Xiaobo WANG

Powder metallurgy processed metal-matrix friction materials for space applications / 219–229
Yelong XIAO, Pingping YAO, Kunyang FAN, Haibin ZHOU, Minwen DENG, Zongxiang JIN

Galvanically induced potentials to enable minimal tribochemical wear of stainless steel lubricated with sodium chloride and ionic liquid aqueous solution / 230–242
Tobias AMANN, Felix GATTI, Natalie OBERLE, Andreas KAILER, Jürgen RÜHE
Vol. 6, No. 3

Guest editorial: Special Issue on Science of Wear / 243–244
Yonggang MENG, Valentin L. POPOV, Jean-François MOLINARI

Review Article

Adhesive wear mechanisms uncovered by atomistic simulations / 245–259
Jean-François MOLINARI, Ramin AGHABABAEI, Tobias BRINK, Lucas FRÉROT, Enrico MILANENE

Research Articles

Adhesive wear and particle emission: Numerical approach based on asperity-free formulation of Rabinowicz criterion / 260–273
Valentin L. POPOV, Roman POHRT

Combined effect of boundary layer formation and surface smoothing on friction and wear rate of lubricated point contacts during normal running-in processes / 274–288
Yazhao ZHANG, Alexander KOVALEV, Yonggang MENG

Estimating antiwear properties of esters as potential lubricant-based oils using QSTR models with CoMFA and CoMSIA / 289–296
Zhan WANG, Tingting WANG, Guoyan YANG, Xinlei GAO, Kang DAI

Finite element simulation and experimental test of the wear behavior for self-lubricating spherical plain bearings / 297–306
Yahong XUE, Jigang CHEN, Sumin GUO, Qingliang MENG, Junting LUO

Effect of oxide film on nanoscale mechanical removal of pure iron / 307–315
Jinwei LIU, Liang JIANG, Changbang DENG, Wenhao DU, Linmao QIAN

Single asperity friction in the wear regime / 316–322
Yongjian YANG, Yunfeng SHI

Acoustic emission characterization of sliding wear under condition of direct and inverse transformations in low temperature degradation aged Y-TZP and Y-TZP-Al2O3 / 323–340
Nikolai L SAVCHENKO, Andrey V FILIPPOV, Sergei Yu TARASOV, Andrey I DMITRIEV, Evgeny V SHILKO, Aleksandr S GRIGORIEV

Short Communication

60 years of Rabinowicz’ criterion for adhesive wear / 341–348
Elena POPOVA, Valentin L. POPOV, Dae-Eun KIM
Vol. 6, No. 4

Review Article

Advances in nonequilibrium molecular dynamics simulations of lubricants and additives / 349–386
J. P. EWEN, D. M. HEYES, D. DINI

Research Articles

Effect of arc-current and spray distance on elastic modulus and fracture toughness of plasma-sprayed chromium oxide coatings / 387–394
Simanchal KAR, Partha Pratim BANDYOPADHYAY, Soumitra PAUL

Friction characteristics of mechanically exfoliated and CVDgrown single-layer MoS₂ / 395–406
Dinh Le Cao KY, Bien-Cuong TRAN KHAC, Chinh Tam LE, Yong Soo KIM, Koo-Hyun CHUNG

Contribution of friction and adhesion to the reliable attachment of a gecko to smooth inclines / 407–419
Zhouyi WANG, Qiang XING, Wenbo WANG, Aihong JI, Zhendong DAI

Analysis of the coated and textured ring/liner conjunction based on a thermal mixed lubrication model / 420–431
Chunxing GU, Xianghui MENG, Di ZHANG

Study on frictional behavior of carbon nanotube thin films with respect to surface condition / 432–442
Youn-Hoo HWANG, Byung-Soo MYUNG, Hyun-Joon KIM

Numerical modeling and analysis of plasmonic flying head for rotary near-field lithography technology / 443–456
Yueqiang HU, Yonggang MENG

Friction and wear of sand-contaminated lubricated sliding / 457–463
Mohamed Ahmed RAMADAN

Evaluation of lubrication performance of crank pin bearing in a marine diesel engine / 464–471
Suk Man MOON, Yong Joo CHO, Tae Wan KIM

Short Communication

Insight into tribological problems of green ship and corresponding research progresses / 472–483
Yuwei SUN, Xinping YAN, Chengqing YUAN, Xiuqin BAI