Contact Damage and Strength Degradation in Brittle/Quasi-Plastic Silicon Nitride Bilayers

Kee Sung Lee,† Seung Kun Lee,*‡ and Brian R. Lawn*†

Materials Science and Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

Do Kyung Kim*

Department of Materials Science and Engineering, Korea Advanced Institute of Science and Technology, Yusong, Taejon 305-701, Korea

A study is made of the damage resistance of silicon nitride bilayers consisting of a hard overlayer (coating) on a soft underlayer (substrate). The two layers are fabricated with different starting powders, to provide distinctive elongate-grain microstructures, and are cosintered, to provide strong interfacial bonding and thus to minimize subsequent delamination. Contact testing with spherical indenters is used to characterize the damage response. The elastic-plastic mismatch between the layers is sufficiently high as to produce distinctive damage modes in the two layers: predominantly cone cracking in the coating, and quasi-plasticity in the substrate. However, the mismatch is also sufficiently low as to preclude secondary transverse cracks of the kind observed in other bilayer systems to initiate immediately beneath the contact at the coating/substrate interface and propagate upward within the coating. The dominant damage mode shifts from coating fracture to substrate quasi-plasticity with increasing contact load and decreasing coating thickness. Significantly, the presence of the soft underlayer inhibits growth of the coating cone cracks as the latter approach and intersect the interface. The underlayer also substantially diminishes strength losses from the contact-induced damage, especially in bilayers with thinner coatings. The implication is that bilayer structures with thin, hard coatings can preserve benefits from the inherent toughness of soft substrate materials, and at the same time afford surface protection (high wear resistance) to the underlayer.

I. Introduction

Over the last decade there has been a drive to improve the toughness of silicon nitride (Si₃N₄) and other silicon-based ceramics via in situ growth of long, elongate grains.¹ ¹² In Hertzian contacts with spherical indenters these enlarged and elongated microstructures produce diffuse energy-absorbing subsurface quasi-plastic zones, consisting of individual, grain-localized microfailures ("shear faults"),¹³ instead of conventional surface-initiated cone cracks.¹⁴ ¹⁸ At moderate contact loads this quasi-plastic damage has a comparatively minor degrading effect on strength in ensuing flexure or tension.¹⁹ ²⁰ However, at higher contact loads or numbers of repeat cycles the microfailures tend to coalesce within the damage zone, with consequent accelerated strength loss and material removal.²¹ Accordingly, quasi-plastic damage can have a highly adverse effect on resistance to fatigue²¹ and wear²² (although it can also promote machinability²³). Improvement in mechanical properties from in situ tailoring of microstructures is not universal.

The search for ceramic structures with superior property combinations has led to a novel concept of layered homogeneous/heterogeneous microstructures with relatively hard surfaces (coatings) on soft underlayers (substrates).²⁴ ²⁵ A critical element of this concept is the incorporation of interlayer elastic–plastic mismatch, so as to partition energy from the contact loading system into competing fracture and quasi-plasticity modes, with the principal intent of containing the former mode without unduly promoting the latter. Another critical element is incorporation of strong rather than weak interlayer interfaces, so as to avoid delamination—the aim is to suppress rather than to deflect cone (or other) cracks that originate in the coating. The existence of any compression stresses in the coating from thermal expansion mismatch with the substrate will further inhibit the extension of any such cracks.²⁶ ²⁷ In principle, such a composite layer system should exhibit both high wear resistance and high toughness, with reduced susceptibility to strength degradation from damage accumulation.

In a recent study we investigated these critical elements in Si₃N₄ bilayers with uncommonly large elastic–plastic mismatch,²⁸ achieved by incorporating boron nitride (BN) platelets into the sublayer structure.²⁶ ²⁷ In Hertzian contact tests these bilayers showed confined transverse cracking in the coatings, along with extensive yield in the substrates. In thinner coatings with higher mismatch the soft substrate allowed the hard coating to "flex" beneath the contact, further enhancing the transverse cracking. In these cases some of the transverse cracks initiated from the coating/substrate interface rather than from the coating top surface, at relatively low loads. The cracks were highly stabilized, requiring exceptionally high loads to drive them upward through the coating, thereby enabling the crack population to multiply. It was clear that any such proliferation of cracks, while contributing to damage tolerance, could ultimately compromise the structural integrity of the coating, especially if the bilayer were to be subjected to high or repeat loading. It would seem that there are upper limits to the useful range of mismatch in these structures.

In the present work we investigate an analogous Si₃N₄ bilayer system, but with much smaller elastic–plastic mismatch, in an attempt to retain the advantages of the layer concept without inducing excessive transverse cracking. Bilayers with
relatively brittle and quasi-plastic Si₃N₄ components are fabricated, but without addition of a softening agent in the substrate. The two Si₃N₄ layers are fabricated from different starting powders, at specified layer thicknesses, but are cosintered to produce a strong adjoining interface. Contacts with spheres are used to induce controlled damage patterns into the layered specimens and ceramographic sectioning techniques used to examine these patterns. Bend tests on damaged bar specimens are carried out to investigate the role of the different damage modes on strength degradation.

We shall show that whereas the Si₃N₄ coating layers are still subject to cone cracking, they show no deleterious interface-initiated transverse cracking. Quasi-plastic damage is more limited than in the previous BN-doped substrate structures, but nevertheless appears to play an important role in restricting extension of the surface ring cracks within the coating. Strengths of contact-damaged bilayers are enhanced relative to those of the (comparatively weak) coating material, tending at small coating thicknesses to those of the (comparatively strong) substrate material. The results suggest the prospect of designing layer structures for both high toughness and high wear resistance.

II. Experimental Procedure

(1) Powders and Processing

Two different Si₃N₄ powders were selected for preparation of the bilayers. The starting powder for the coating layers was α-Si₃N₄ of mean particle size 1.0 μm (UBE-SN-E3, Ube Industries, Tokyo, Japan), with additives 5 wt% Y₂O₃ (Fine Grade, H. C. Starck GmbH, Goslar, Germany), 2 wt% Al₂O₃ (AKP 50, Sumitomo Chemical Co. Ltd., Tokyo, Japan), and 1 wt% MgO (High Purity, Baikowski Co., NC). The powder for the substrate layers was α-Si₃N₄ of average particle size 0.3 μm (UBE-SN-E10, Ube Industries, Tokyo, Japan), with the same sintering additives. Previously, these two final batches were designated AC (α-coarse) and AF (α-fine), according to Si₃N₄ starting powder—here, we use the simpler designation C (coating) and S (substrate).

A Si₃N₄ powder batch with 30 vol% BN softening additive from another preceding study was used as a comparison substrate—this material is here designated S30.

The powder batches were each mixed in isopropyl alcohol for 24 h in a planetary ball-mill using zirconia balls in a polypropylene container, to form a slurry. After drying, the softly agglomerated powder was crushed and sieved through a 60-mesh screen. The powders were cold-pressed in a graphite mold of 50 mm diameter to form green state bilayers of coating thickness 1 mm and substrate thickness 3–4 mm. The composite structure was hot-pressed at 1700°C at a pressure of 30 MPa in nitrogen for 1 h. Some monolithic samples of each powder mix were prepared for control studies.

Surfaces of some specimens were polished to 1 μm finish, plasma etched, gold coated, and examined by scanning electron microscopy (SEM) to reveal the microstructures.

(2) Mechanical Testing

Fired bilayers and monolithic controls were cut into bars for mechanical testing. Routine measurements of Young’s modulus, using a sonic method, and hardness, using Vickers indentations (load/projected area), were carried out on representative monolithic specimens. For the bilayers, the C (coating) structure was placed on the top, and the S (substrate) structure on the bottom. The top surfaces of the bars were then ground and polished to 1 μm diamond paste finish to produce coating thicknesses d₁ = 20–600 μm. Additional Vickers indentations were made on polished sides of representative bilayer specimens to evaluate any differential thermal expansion stresses in the coatings, from radial crack measurements. Some specimens were cut in half and the opposing surfaces polished and glued together to produce ‘‘bonded-interface’’ specimens for investigation of subsurface damage in indentation tests. Other specimens were edge chamfered and polished for strength testing.

Hertzian indentations were made on the top surfaces of bilayer specimens, as well as on control monolith specimens, using WC spheres of radius r = 1.98 mm at loads up to P = 4000 N, in air. On as-polished specimens, rows of indentations were made at small load intervals to determine the critical loads P₁ and P₂ for first incidence of yield and cracking. Application of a gold coat greatly enhanced detection of the initial damage in Nomarski illumination. For the cracking an acoustic detector placed on the top surface adjacent to the indenter was a useful adjunct. On bonded-interface specimens the indentations were made in rows symmetrically across the interface traces at the top surfaces. After separating the indented specimens halves in solvent and coating with gold, Nomarski illumination was used to view the subsurface damage.

Four-point flexure tests were run on bilayer and monolith bar specimens (25 mm × 4 mm × 3 mm, inner span 10 mm, outer span 20 mm) after indentation. The indentation sites were covered with a drop of silicone oil before testing and centered on the tensile side of the bend fixture. The bars were then broken in fast fracture (<10 ms) to avoid the influence of moisture (‘‘inert’’ strengths). ‘‘Effective strengths’’ were calculated from beam theory, using the conventional relation

\[ \sigma = \frac{3Ql}{4wd^2} \]  

with l the moment span, w the bar width, d the composite bilayer half-thickness, and Q the breaking load. Later, in Section III(3), we shall take into account the effect of thermal expansion mismatch (Appendix A) and interlayer modulus mismatch (Appendix B) on strengths evaluated from fracture mechanics considerations (Appendix C).

The broken specimens were examined in Nomarski illumination and scanning electron microscopy to locate the sources and modes of failure, i.e., from cone cracks or quasi-plastic zones. Control strength tests on unindented specimens were made to measure baseline ‘‘natural’’ strengths.

III. Results

(1) Characterization of Silicon Nitride Layer Materials

Figure 1 shows bimodal microstructures for the two materials. Common to both microstructures at the high end of the bimodal distribution are enlarged, elongated β-phase grains of length =7 μm and width =1.0 μm. In the C material (Fig. 1(a)) the volume fraction of this high-end grain component is relatively small (=23 vol%)—the bulk of the microstructure consists of finer, equiaxed α-phase grains of diameter =1.2 μm (=77 vol%). In the S material (Fig. 1(b)) the volume fraction of the high-end grain component is increased (=40 vol%)—in this case the bulk of the microstructure consists of finer elongate β-phase grains of length =2 μm and width =0.5 μm (=60 vol%).

Young’s modulus and hardness are shown for C Si₃N₄ and S Si₃N₄ in the bar chart of Fig. 2, along with comparison values for S30 Si₃N₄. The C material is slightly stiffer and harder than S, which in turn is very much stiffer and harder than S30. Note that the relative differences are considerably greater in the hardness than in the modulus, corresponding to greater plastic than elastic mismatch. For reference, bearing grade Si₃N₄ typically falls somewhere between the C and S materials—e.g., measurements on commercially available NBD200 (Norton/TRW Ceramics, Northboro, MA) give modulus 320 GPa and hardness 17.5 GPa.

The bar chart in Fig. 3 shows critical loads P₁ and P₂ for first incidence of yield and cracking beneath Hertzian contacts in the same monolith material set. The critical load P₁ falls off strongly through the sequence C-S-S30, commensurate
with the hardness trend in Fig. 2. Conversely, $P_c$ increases strongly, such that whereas full cone cracks are generated in $C$, only shallow surface ring cracks appear in $S$, and no ring cracking is observed at all in $S30$ over the load range.\(^{28}\) Note that for $C$ the values of $P_Y$ and $P_c$ are comparable, suggesting a relatively balanced competition between yield and fracture in this material.

Figure 4 plots biaxial compression stresses $\sigma_R$ associated with differential thermal expansion between $C$ and $S$ layers, as a function of coating thickness $d_c$. The data are evaluations from radial crack measurements on the coating sections; the solid curve is a theoretical fit to these data. Details of the data evaluations and fit are described in Appendix A.

(2) Contact Damage in Bilayers

Figure 5 compares bonded-interface section views of damage sites in bulk and bilayer $Si_3N_4$ structures, from indentations with a WC sphere of radius $r = 1.98$ mm at load $P = 3000$ N made on the top surface. In the bulk $C$ specimen (Fig. 5(a)) we observe the developed cone crack (a faint trace of the subsurface quasi-plasticity zone below the contact becomes apparent in this material at high magnification). In the bilayer $C/S$ specimen (Fig. 5(b)) the cone crack is substantially shallower and is contained wholly within the $C$ layer (thickness $d_c = 180$ $\mu$m). The quasi-plasticity zone is now more intense and extends into the softer $S$ underlayer. In a $C/S30$ bilayer specimen (Fig. 5(c)), included here from the earlier study\(^{28}\) for comparison as a case of extreme mismatch, the coating ($d_c = 250$ $\mu$m) contains an array of transverse cracks, including both cone cracks initiated from the top surface and upward-extending "inverted
cone” cracks initiated at the interlayer interface.28 Yield and associated microcrack damage in the S30 sublayer is now extensive, and some interfacial delamination is evident. It is apparent that incorporation of a soft underlayer can inhibit fracture if the mismatch with the coating material is small, but can enhance fracture if the mismatch is large. Note that the load \( P = 3000 \text{ N} \) used in Fig. 5 exceeds both \( P_Y \) and \( P_C \) for the monolithic C material (Fig. 3), consistent with the appearance of both yield and cracking in the coating layer.

Now consider results for C/S bilayers in greater detail, taking the effects of contact load and coating thickness in turn:

(A) Effect of Contact Load: Micrographs of Hertzian contact damage in C/S Si\(_3\)N\(_4\) bilayers with C coating thickness \( d_c = 180 \mu\text{m} \) (±10 \( \mu\text{m} \)) are shown in Fig. 6 over a sequence of contact loads. At \( P = 2000 \text{ N} \) (Fig. 6(a)) a cone crack has initiated and there is a limited quasi-plasticity zone, both confined to the coating. At higher loads \( P = 2500 \text{ N} \) (Fig. 6(b)), \( P = 3000 \text{ N} \) (Fig. 6(c)), and \( P = 3500 \text{ N} \) (Fig. 6(d)) both cracking and quasi-plasticity zone expand. However, whereas the cone crack remains confined to the coating, the quasi-plasticity now extends into the substrate. With further loading to \( P = 4000 \text{ N} \) (Fig. 6(e)) the cone crack penetrates the interlayer interface, but remains inhibited (note that a second cone crack has now appeared), while the quasi-plasticity expands and intensifies within the substrate. The sequence in Fig. 6 suggests that the response is dominated by coating fracture at low load and by substrate plasticity at high load.24

Measured cone crack depths \( h \) below the top surface are plotted as a function of load \( P \) in Fig. 7 for both the C/S bilayer and the C monolith, for the coating thickness \( d_c = 180 \mu\text{m} \).
indicated by the horizontal shaded line. The solid curves are representations of fracture mechanics relations for pennylike cone cracks in layers subjected to superposed uniform stresses \( \sigma_R \), described in detail in Appendix C. The curve through the monolith data is a best-fit to these relations in the limit \( \sigma_R \to 0 \)—this fit is used to “calibrate” essential crack-geometry parameters. The curve through the bilayer data is an ensuing prediction for coatings with \( \sigma_R = -190 \text{ MPa} \) (\( d_c = 180 \text{ \( \mu \)m} \) in Fig. 4), using the calibrated crack-geometry parameters. (Note that this latter curve remains valid only while the cone crack remains embedded in the coating, i.e., \( h < d_c \) in Fig. 7.) The presence of the expansion mismatch stress would appear to account for the bulk of the data shift between bilayer and monolith, notwithstanding the scatter in data.

(B) Effect of Coating Thickness: The sequence of micrographs of Hertzian contact damage in \( \text{C/S Si}_3\text{N}_4 \) bilayers at fixed contact load \( P = 3000 \text{ N} \) in Fig. 8 illustrates the effect of coating thickness on the damage pattern. For a relatively thick coating layer, \( d_c = 600 \text{ \( \mu \)m} \) (Fig. 8(a)), both cone crack and quasi-plasticity zone are wholly contained within the coating. As the coating is made thinner, \( d_c = 400 \text{ \( \mu \)m} \) (Fig. 8(b)), both crack and plasticity zone remain contained within the coating, with attendant slight diminution in the cone crack depth. At \( d_c = 180 \text{ \( \mu \)m} \) (Fig. 8(c)), the cracks are still contained, but the yield zone penetrates into the substrate. In the thinnest coatings, \( d_c = 80 \text{ \( \mu \)m} \) (Fig. 8(d)) and \( d_c = 40 \text{ \( \mu \)m} \) (Fig. 8(e)) the cracks, although still highly constrained, occasionally penetrate into the substrate (e.g., right side of Fig. 8(d)); the quasi-plasticity zone is now contained largely within the substrate. The sequence in Fig. 8 suggests that the response is dominated by coating fracture at high thicknesses, and by substrate plasticity at low thicknesses.

Figure 9 plots cone crack depth \( h \) as a function of coating thickness \( d_c \), at fixed load \( P = 3000 \text{ N} \). The horizontal dashed line denotes the cone crack depth in the C monolith and the inclined dashed line denotes the locus \( h = d_c \) for which the cone crack just intersects the \( \text{C/S} \) interface. The plot demonstrates quantitatively the extent to which the interface increasingly constrains the cone crack as the coating layer becomes thinner. At small coating thicknesses, \( d_c < 180 \text{ \( \mu \)m} \), some limited penetration into the substrate does occur, but the constraining influence nevertheless persists down to the smallest coating thicknesses. The solid curve is the fracture mechanics prediction for cone cracks in coatings with nonzero \( \sigma_R(d_c) \), using the calibrated crack-geometry parameters from the monolith data.
in Fig. 7 (Appendix C). (Once more, this curve remains valid only while the cone crack remains embedded in the coating, i.e., \( h < d_c \)). Again, the presence of expansion mismatch stresses would appear to account for the data trends, within the scatter.

Fig. 9. Plot of cone crack depth as function of coating thickness for C/S Si₃N₄ bilayers (open symbols), WC sphere \( r = 1.98 \) mm, at fixed load \( P = 3000 \) N. Solid curve is fracture mechanics prediction. Inclined dashed line indicates configurations \( h = d_c \) for which cracks just intersect the interface; horizontal dashed line indicates crack depth in \( C \) monolith.

Fig. 11. Critical loads \( P_C \) for cone crack initiation as a function of coating thickness \( d_c \) for C/S Si₃N₄ bilayers, WC sphere radius \( r = 1.98 \) mm. Horizontal dashed line indicates \( P_C \) for the coating monolith.

On those occasions when the coating cone cracks do penetrate into the substrate the crack paths undergo only minor deflections at the C/S interface, without delamination. Figure 10 is an example, for an indentation in a thin coating, \( d_c = 40 \) \( \mu \)m, at high load, \( P = 4000 \) N. Note the continuity of the grain structure across the interface in this figure, confirming a strong interface.

While the presence of the substrate clearly has an important

Fig. 10. Side view of Hertzian cone crack near C/S interface in Si₃N₄ bilayer of coating thickness \( d_c = 40 \) \( \mu \)m, at load \( P = 4000 \) N using WC sphere radius \( r = 1.98 \) mm. Note crack penetration from coating (top) to substrate (bottom) without delamination. Scanning electron micrograph.
influence on the cone crack size, it does not appear to be a strong factor in the critical load for cone crack initiation. A plot of $P_C$ as a function of coating thickness $d_c$ is presented in Fig. 11. The values are relatively unchanged from those for the C monolith (shaded band—see Fig. 3) at thicknesses down to $d_c = 40 \, \mu m$, below which $P_C$ increases.

(3) Strength Degradation

Figure 12 shows failure origins on broken bilayer strength-test specimens containing indentations from WC spheres of radius $r = 1.98 \, \text{mm}$, for two sets of conditions: (a) for fixed coating thickness $d_c = 200 \, \mu m$, at increasing (postcritical) loads; (b) for fixed load $P = 4000 \, \text{N}$, at decreasing thick-
Fig. 13. Strength degradation of C/Si₃N₄ bilayers, plus C and S monoliths, for contacts with WC sphere radius r = 1.98 mm. Data plotted as a function of indentation load P, for coating thicknesses indicated. Open symbols indicate breaks away from indentation sites, gray symbols failures from quasi-plastic zones, black symbols failures from cone cracks. Boxes at left axis denote strengths of unindented specimens. Vertical dashed lines indicate values of Pₛ for S, Pₛ for C. Solid curves are theoretical predictions assuming failure from cone cracks at P > Pₛ and from microstructural flaws at P < Pₛ for C monoliths (lowest curve) and C/S bilayers at 400, 200, and 100 µm (upper three curves, truncated at loads and coating thicknesses where cones penetrate into substrate). Dashed curve is empirical fit to S data.¹⁸

IV. Discussion

In this study we have confirmed that microstructural tailoring can have a vital influence on the damage tolerance of bilayer structures. An essential element in the tailoring is avoidance of a weak interface at the interlayer junction, in the present instance by cosintering the starting powders. We have focused specifically on Si₃N₄ bilayers with harder coatings on softer substrates, with modest mismatch. In contacts with spheres the coatings show mainly cone cracking and the substrates mainly quasi-plasticity. These are even more complex than the analyses for cone cracks, because of additional driving forces associated with residual local stress fields at the slipped shear faults.³⁴,³⁵

whereas at thinner coatings, from quasi-plastic zones. The effects of coating failures first from cone cracks and later, at higher loads and coating thicknesses Cₘ, as well as for the Cₘ monoliths. Data points from indented specimens: open symbols indicate breaks away from indentation sites, gray symbols failures from substrate quasi-plastic zones, and black symbols failures from coating cone cracks. Different regions of load dependence are apparent in the strength data: at P < Pₛ (vertical dashed line), no significant load dependence; at P > Pₛ, abrupt falloff; at P > Pₛ, slowly continuing falloff, with failures first from cone cracks and later, at higher loads and thinner coatings, from quasi-plastic zones. The effects of coating thickness are manifest: whereas at P < Pₛ the strength values for the bilayers differ little either from each other or from the Cₘ or Sₘ bounds, at P > Pₛ the data shift strongly upward from the Cₘ lower bound toward the Sₘ upper bound.

Fracture mechanics analyses of strength degradation for failures from cone cracks have been developed for monolithic materials,²⁴ including Si₃N₄. Analogous computation of strengths in bilayer systems is complicated by the additional stress terms associated with mismatch between coating and substrate: from mismatch in thermal expansion coefficients during initial cooling²⁶,²⁷(Appendix A); and from mismatch in elastic moduli during subsequent flexure (Appendix B). For a flaw of characteristic dimension c with negligible residual contact field, the strength under such conditions may be written (Appendix C)

\[ \sigma_F = \frac{(T_0/\psi)^{1/2} - \sigma_R}{k} \]

with \( \psi \), a crack geometry coefficient for Hertzian cones inclined at angle \( \alpha \) to the top surface, \( T_0 \), the toughness (assumed single-valued), \( \sigma_R \), an expansion mismatch stress (Appendix A), and \( k \) an elastic mismatch coefficient (Appendix B). The solid curves in Fig. 13 are a priori predictions of \( \sigma_F(P) \) for failure from cone cracks, using Eq. (2) in conjunction with appropriate expressions for the cone crack depth \( h = c \sin \alpha \) as a function of load \( P \), for each of the coating thicknesses \( d_c \) represented. These curves truncate at the loads corresponding to cone crack penetration into the substrate, i.e., at depth \( h = d_c \). The predictions appear to account for the broad data shifts for those specimens that fail from cone cracks in Fig. 13, although the absolute predictions lie outside the range of data scatter.

Once the cone cracks approach and penetrate into the tough substrate, failure occurs from the subsurface quasi-plasticity zones. Corresponding analyses of strength degradation from grain-localized “shear faults” within quasi-plasticity zones are being developed.³⁴ These are even more complex than the analyses for cone cracks, because of additional driving forces associated with residual local stress fields at the slipped shear faults.³⁴,³⁵

The typical damage pattern in the present Si₃N₄ bilayer system provides a marked contrast to that in analogous bilayer substrates containing BN platelets (cf. Figs. 5(b) and (c)).²⁸ In the latter system the elastic–plastic mismatch is much greater (Fig. 2), and the yield in the substrate consequently much more extensive, allowing the coating to undergo a substantial component of “flexure” on the soft support beneath the contact. Such flexure leads to intense concentration of tensile stresses within the coating at the coating/substrate interface,²⁹ with consequent generation of upward-extending transverse cracks at relatively low loads. While such cracks tend to be highly stable, they also tend to proliferate within the coating, compromising the subsequent integrity of the bilayer. In the present Si₃N₄ system the mismatch is too small to allow significant buildup of coating tensile stresses at the substrate interface, restricting the fracture to conventional cone cracking. At the
same time, as we have seen, the mismatch is sufficient to promote failures from subsurface quasi-plasticity zones at high loads and thin coatings.

The data in Figs. 11 and 13 warrant further comment, because collectively they characterize the coating loading conditions under which bilayer structures with modest (but nonzero) mismatch may operate without incurring intolerable strength losses from contact damage. Figure 11 indicates that the critical contact load $P_c$ for cone crack initiation is largely insensitive to coating thickness $d_c$ in the present bilayer system, at least down to $d_c = 50 \mu m$ (relative to the sphere radius 1.98 mm used in our experiments). On the other hand, Fig. 13 indicates that the ensuing abrupt stress decrement at $P = P_c$ is very much sensitive to $d_c$. This behavior is consistent with St. Venant’s principle in elasticity theory. In the near field outside the contact circle at the top coating surface, where the cone cracks initiate during indentation, the maximum tensile stresses are very high and are relatively unaffected by the presence of moderate thermal expansion stresses or yield in the remote substrate; in the far field below the contact, where the cone cracks propagate (Fig. 8), the tensile stresses are much more strongly influenced by these extraneous factors, especially at higher $P$ and smaller $d_c$. In the thinner coatings the higher thermal expansion stresses and increased subsurface yield ultimately suppress cone fracture, leading to the change in mode of failure observed in Fig. 12.

A complete understanding of this latter kind of quasi-plasticity-initiated failure in layered structures is currently under study.

**APPENDIX A**

**Differential Thermal Expansion Stresses in Bilayer Coatings**

In-plane biaxial stresses $\sigma_{ic}$ from C/S coating/substrate thermal expansion mismatch are measured from the lengths of radial cracks at the vertical corners of Vickers indentations on coating sections (Fig. A1). The stress-intensity factor $K$ for Vickers radial cracks of half-length $c$ is given by

$$K = x_1P/c^{1/2} + \psi_0\sigma_c c^{1/2} = K_{ic} = T_0$$

(A-1)

with $x_1 = 0.066$ an indentation coefficient for fine-grain Si$_3$N$_4$, $\psi_0 = 0.77$ a geometrical radial crack coefficient, and $T_0$ the toughness (assumed single-valued$^{17}$). Measurements from comparative Vickers indentations in the C monolith ($\sigma_{ic} = 0$) enable the determination $T_0 = 4.2$ MPa-m$^{1/2}$.

The stresses $\sigma_{ic}$ can also be calculated from a force balance relation for bilayers with thermal expansion mismatch between coating (c) and substrate (s);$^{27}$

$$\sigma_{ic} = (\alpha_c - \alpha_s)E_c\Delta T/(1 - \nu_c)(1 - \nu_s)(E_c/E_s)(2d_c/d_s)$$

(A-2)

where $\alpha$ is the thermal expansion coefficient, $E$ is Young’s modulus, $\nu$ is Poisson’s ratio, and $d$ is layer thickness. The solid curve through the coating data in Fig. 4 is a best fit of Eq. (A-2) to the data, with the adjustments $(\alpha_c - \alpha_s)E_c\Delta T/(1 - \nu_c)(1 - \nu_s)(E_c/E_s)$ of $-311$ MPa and $E_c/(1 - \nu_c)$ of 4.92. Inserting $\alpha_s = 4.16 \times 10^{-6}$ $^\circ$C$^{-1}$ and $\alpha_c = 6.42 \times 10^{-6}$ $^\circ$C$^{-1}$ (dilatometer measurements, N$_2$ atmosphere, 25–1000°C), $E_c = 335$ GPa and $\nu_c = 0.29$ (sonic wave measurements), the fit yields the following estimates: $\Delta T = 1430^\circ C$, which would appear to be a little higher than the temperature range over which stresses no longer relax; and $E_c/(1 - \nu_c) = 95.9$ GPa, which is substantially lower than the measured value 315 GPa/(1 − $\nu_c$) = 431 GPa (sonic wave measurements).

In view of the numerical discrepancies between Eqs. (1) and (2), the values of $\sigma_{ic}$ in Fig. 4 should be regarded as no more than first approximation estimates.

Equivalent stresses in the substrate can be similarly evaluated by inverting the subscripts c and s in Eq. (A-2).

**APPENDIX B**

**Bending Stresses in Layer Structures with Different Elastic Moduli**

Consider a bilayer consisting of two rectangular bars of same width $w$, but of different thickness $d_c$ and $d_s$ and modulus $E_c$ and $E_s$ ($c$ = coating, $s$ = substrate), perfectly bonded at the common interface (Fig. B1). The composite bar is subjected to bending, with layer c on the tensile side and layer s on the compression side. Although the strain distribution across the bilayer is linear and continuous, the stress distribution suffers a discontinuity at the interface.

We are specifically concerned with the maximum tensile stresses in the coating at the top surface, $\sigma_c$, and in the substrate at the interlayer interface, $\sigma_s$. The stresses of interest in four-point bending are related to the effective strength $\sigma_f$ in Eq. (1) by

$$\sigma = (3Pl/4wd^2)k = k\sigma_f$$

(B-1)

where $k$ is a dimensionless factor (unity for $d_c = d_s$, $E_c = E_s$). The stress distributions can be determined in a straightforward manner from thin beam theory.$^{41}$ Let us write

$$\delta = d_c/d_s$$

(B-2a)

$$e = E_c/E_s$$

(B-2b)

For the tensile stress in the top surface of the coating (layer 1), we obtain

$$k_c = e(2 + 1/\delta + e\delta)/(3e + (1 + 1/\delta^3)(1 + e\delta)/(1 + 1/\delta^2))$$

(B-3)

For the tensile stress in the substrate at the interlayer interface we have

$$k_s = (1 - \delta - e\delta)/(3e + (1 + 1/\delta^3)(1 + e\delta)/(1 + 1/\delta^2))$$

(B-4)

Fig. A1. Coating/substrate bilayer, with elastic and thermal expansion mismatch. Vickers indentations are used to evaluate thermal expansion mismatch stresses in coating.

Fig. B1. Four-point bending of bilayer, generating maximum stress $\sigma_c$ in the coating and $\sigma_s$ in the substrate.
The coefficients \( k_c \) and \( k_s \) are plotted as a function of \( \delta \) in Fig. B2, for \( \varepsilon = E_i/E_s = 335/315 = 1.06 \) appropriate to our Si\(_3\)N\(_4\) bilayer system.

**APPENDIX C**

**Calculation of Cone Crack Depths and Strengths for Bilayers**

For the Si\(_3\)N\(_4\) bilayers that fail from cone cracks in the coating we use a fracture mechanics analysis developed elsewhere.\(^{19,34}\)

The cone crack depth is determined from the following geometrical relations, from Fig. C1:

\[
h = c \sin \alpha \quad \text{(C-1)}
\]

\[
c = C - R_0 \cos \alpha \quad \text{(C-2)}
\]

with \( \alpha \) the cone angle, \( C \) the face length of a virtual conical surface with tip located above the indented surface, and \( R_0 \) the radius of the surface ring crack. To a first approximation, the cone cracks at contact load \( P \) satisfy an equilibrium stress-intensity relation for penny-like cracks:\(^{37}\)

\[
K = \psi_0 P/(C^{1/2} + \Psi_0 \sigma_c \sin \alpha)c^{1/2} = T_0 \quad \text{(C-3)}
\]

with \( \psi_0 \) Hertzian coefficients for straight-ahead cone extension, and \( T_0 \) the toughness (again assumed single-valued). This latter relation incorporates allowance for the action of the (resolved) stress \( \sigma_R \) on the cone crack response (Appendix A). Equations (C-1) to (C-3) form the basis for the computations of cone crack depths in Figs. 7 and 9, plotted in those two figures as the solid curves. We take \( \varepsilon = 25^\circ \) and \( R_0 = 238 \mu \text{m} \) from direct measurements (e.g., Figs. 6 and 8), along with \( T_0 = 4.2 \text{ MPa} \cdot \text{m}^{1/2} \) (Appendix A). (Actually, \( \alpha \) and \( R_0 \) may be expected to vary with \( \sigma_R \), and hence with coating thickness, but any such variations are slight enough to be masked by the data scatter in our Si\(_3\)N\(_4\) system.) The value of the parameter \( \psi_0 \) is simply that for mode I straight-ahead extension, \( \psi_0 = \pi^{1/2} = 1.77.\)\(^{38}\) The parameter \( \psi_0 = 0.0311 \) is “calibrated” from a best fit of Eqs. (C-1) to (C-3) to the \( h(P) \) data for monolithic C Si\(_3\)N\(_4\) in Fig. 7 (\( \sigma_c = 0 \)). The corresponding curves for the bilayer data in Figs. 7 and 9 are predictions, using appropriate values of \( \sigma_c (d_c) \) from Appendix A in Eqs. (C-1) to (C-3). (Note that this analysis is valid only while the cone crack remains embedded in the coating, i.e., \( h < d_c \) in Fig. 9.)

Strength degradation functions may be determined from conventional fracture mechanics for failure from cone cracks in flexure.\(^{34,42}\) Above the critical load \( P_c \) for cone initiation, the stress-intensity factor for equilibrium cone cracks in an applied field \( \sigma_A \) and superposed field \( \sigma_R \) is

\[
K = \psi_0 \sigma_A^{1/2} = T_0 \quad \text{(C-4)}
\]

where \( \sigma = \sigma_A + \sigma_R = k_c \sigma_A \). Failure then occurs spontaneously from the cone base according to the strength relation

\[
\sigma_f = (T_0/\psi_0 \sigma_A^{1/2} - \sigma_R)/k_c \quad \text{(C-5)}
\]

with \( \psi_0 \) a crack geometry coefficient for Hertzian cones at angle \( \alpha \). This relation again makes allowance for the presence of thermal expansion mismatch stresses \( \sigma_0 \) (Appendix A) and for stress modifications associated with the elastic mismatch parameter \( k_c \) in flexure (Appendix B). At failure, the cone crack reinitiates unstably in predominantly tensile mode from its base, at some angle to \( \alpha \), thereby determining \( \psi_0 \). Below the critical load \( P_c \), the \( \psi \) term in Eq. (C-5) is governed by the geometry of preexisting natural flaws, independent of \( P \).

The functions \( \sigma_f (P) \) calculated from Eq. (C-5), in conjunction with Eqs. (C-2) and (C-3), are included as the solid curves at \( P > P_c \) in Fig. 13. For these calculations, \( \psi_0 \) is determined from a previous analysis of crack reinitiation from the cone crack base, from which \( \psi_0 = 0.037 \alpha = 0.93 \) approximately for \( \alpha = 25^\circ.\)\(^{34}\) Appropriate values of \( k(d_c) \) are computed from Appendix B. (Note again that this analysis is valid only while the cone crack remains embedded in the coating.)

**Acknowledgment:** We thank T.-J. Chuang for providing the solution in Appendix B.

**References**


