Epitaxially influenced boundary layer model for size effect in thin metallic films

Zdeněk P. Bažant and Zaoyang Guo
Department of Civil Engineering, Northwestern University, Evanston, Illinois 60208

Horacio D. Espinosa, Yong Zhu, and Bei Peng
Department of Mechanical Engineering, Northwestern University, Evanston, Illinois 60208

(Received 23 April 2004; accepted 29 December 2004; published online 21 March 2005)

It is shown that the size effect recently observed by Espinosa et al., [J. Mech. Phys. Solids 51, 47 (2003)] in pure tension tests on free thin metallic films can be explained by the existence of a boundary layer of fixed thickness, located at the surface of the film that was attached onto the substrate during deposition. The boundary layer is influenced by the epitaxial effects of crystal growth on the dislocation density and texture (manifested by prevalent crystal plane orientations). This influence is assumed to cause significantly elevated yield strength. Furthermore, the observed gradual postpeak softening, along with its size independence, which is observed in short film strips subjected to pure tension, is explained by slip localization, originating at notch-like defects, and by damage, which can propagate in a stable manner when the film strip under pure tension is sufficiently thin and short. For general applications, the present epitaxially influenced boundary layer model may be combined with the classical strain-gradient plasticity proposed by Gao et al., [J. Mech. Phys. Solids 47, 1239 (1999)], and it is shown that this combination is necessary to fit the test data on both pure tension and bending of thin films by one and the same theory. To deal with films having different crystal grain sizes, the Hall–Petch relation for the yield strength dependence on the grain size needs to be incorporated into the combined theory. For very thin films, in which a flattened grain fills the whole film thickness, the Hall–Petch relation needs a cutoff, and the asymptotic increase of yield strength with diminishing film thickness is then described by the extension of Nix’s model of misfit dislocations by Zhang and Zhou [J. Adv. Mater. 38, 51 (2002)].

The final result is a proposal of a general theory for strength, size effect, hardening, and softening of thin metallic films. © 2005 American Institute of Physics. [DOI: 10.1063/1.1861150]

I. INTRODUCTION

Recently, an extremely strong size effect was found in pure tension experiments on freestanding submicrometer gold thin films tested at Northwestern University by Espinosa et al. 1–6. The tested gold films possessed an average grain size of 200 nm, independent of the thin film thickness. A similar size effect was also observed in other fcc metals such as Cu and Al. A strengthening size scale of one over film thickness was identified by Espinosa et al.5 Their findings constitute direct evidence for the existence of a strengthening mechanism other than that arising from misfit dislocations, and are consistent with similar observations in Ref. 7. Another similar observation has been made by Haque and Saif.8–11 In this case, however, the grain size varied with film thickness.

Although many microlevel tests have been carried out and analyzed since 1990, the size effect behavior revealed by the pure tension tests differs in three ways.

1 In the previous microtorsion, microbending, and microindentation tests, the size effect observed could be fully explained, in one way or another, as an effect of the strain gradient. However, in the pure tension tests of thin films, such an explanation is next to impossible because the gradient of total strain vanishes and the gradient of the plastic part of strain, which can be nonzero due to yield limit variation across the film thickness, is too small to explain the large size effect seen in experiments.

2 The observed size effect is exceedingly strong. When the thickness of the thin film is decreased from 1 to 0.5 μm, the yield stress more than doubles.1–6

3 The nature of the size effect is different. In the previous microtorsion, microbending, and microindentation tests, the size effect is manifested mainly as the differences in nominal stress in the hardening (prepeak) regime at the same value of plastic strain in geometrically similar specimens of different sizes. In the pure tension experiments, however, the size effect is manifested mainly as the thickness dependence of nominal stress of the film (representing the maximum average stress or the effective yield stress).

Because of these differences, a model is needed to explain the experimental results. The exact mechanism leading to the observed size effect is not yet known. Espinosa et al.5 proposed an explanation based on a declining number of dislocation sources in the volume considered, the decline becoming more severe as the number of grains through the thickness decreases. This mechanism is consistent with in situ transmission electron microscopy (TEM) observations of
plastic deformation in Cu (see, for instance, Ref. 12 and the references in this study), which also highlight that even when dislocations become mobile, they disappear at grain boundaries and, therefore, their sources are ineffective. Both of these effects, a limited number of sources and their ineffectiveness, result in very limited film ductility. Another source of limited ductility is the existence of annealing twins in certain grains (Fig. 1). Such twins represent obstacles to dislocation motion and thus such act as strong local hardening effects. The existence of grains possessing twins with an average spacing of 30 nm was observed not only in Au but also in Cu thin films.

This study presents a plausible explanation of the observed size effect based on a simple boundary layer model based on the following ideas.

1. The freestanding thin film contains, on the substrate side (i.e., the side of film that was, during deposition, in contact with the substrate), a boundary layer in which the yield strength is higher than in the rest of the film. Unlike the existing strain-gradient theories of micrometer-scale plasticity, the elevated yield strength can be explained as an indirect consequence of granular epitaxial crystal growth on a substrate of different crystallographic properties than the metallic film. There are two reasons.

   a. Because the epitaxial constraint causes the crystal structure near the substrate to be strained, the film is formed with a significantly elevated density of dislocations near the substrate side, and an increase of dislocation density is known to be the cause of plastic hardening.

   b. The epitaxial constraint further promotes a difference in texture near the substrate side, characterized by a preferred orientation of crystallographic planes that is unfavorable to dislocation glide under axial tension. Evidence for texture development can be seen in the pole figure reported in Ref. 5. Figure 2 shows a comparison of texture measured in 1-µm-thick and 0.3-µm-thick films. Note that the degree of (111) texture decreases as the thickness increases.

2. The presence of the epitaxially influenced boundary layer makes possible strain-softening behavior for sufficiently small length and thickness of the film.

Granular epitaxy means that the epitaxial nucleation and formation of multiple grains during e-beam evaporation affects not only the atoms on the substrate, but also the crystallographic plane orientation through the entire grain in contact with the substrate. In what follows, the term epitaxy is

FIG. 1. Annealing twins in certain grains: plane view of 0.2-µm-thick Cu film (left) and cross-sectional view of 1-µm-thick Au film (right).

FIG. 2. Comparison of texture measured in 1-µm-thick (top) and 0.3-µm-thick (bottom) films.
simply used for the sake of simplicity, but it should be understood that it refers to polycrystalline films rather than single crystals.

In support of point (1b), note that although the interatomic potentials of the atoms of the substrate can directly influence the deposition of the film only to the depth of several atoms (<3 nm), which is insignificant for the films considered, the substrate may be expected to indirectly influence the preferred crystallographic orientations of the whole grains in contact with the substrate, which typically have dimensions of the order of 1 μm. While, in bulk, all the spatial orientations are equally represented, near the surface, certain orientations predominate. In particular, the crystallographic planes permitting dislocation glide may be expected to have a prevalent orientation normal or parallel to the substrate surface, while the planes of dislocation glide that cause plastic normal strain in the axial direction of the film strip are those that have a preferred 45° inclination because this is the inclination of the plane of the maximum shear stress under uniaxial tension.

Since the details of various manufacturing processes affect the dislocation distribution, they will also affect the parameters of the recent epitaxially influenced, (briefly, epitaxial) boundary layer model, while in other existing models the parameters are assumed to depend on the material type only.

A possible small effect of the width of a film strip (Espinosa et al.5) is neglected in this study, for the sake of simplicity and because it is minor compared to the scatter of experimental data.

II. TESTING METHODS AND EXISTING MODELS

Many testing methodologies have been developed to examine the mechanical properties of thin films, ranging from the widely used classical thermomechanical tests of thin films on a substrate7,14,15 to the uniaxial pure tension tests of thin films on flexible polyimide substrate.16,17 Microindentation tests have also been used to estimate the strength of thin films.14 Recently, pure uniaxial tension tests of free-standing thin films became possible.1,6,8–11,18

In the thermomechanical tests, the thin films are under biaxial tension or compression. Conversion to uniaxial compression or tension can be made if the hydrostatic pressure has no effect. However, it should be mentioned that the presence of a strain gradient in these tests may cause extra hardening compared with the uniaxial tests.19 This was modeled in Ref. 20 using Acharya and Bassani’s strain-gradient plasticity theory.23 In Ref. 23, a nonlocal model based on the strain-gradient theory of crystal viscoplasticity24 was proposed and applied to thermomechanical tests.

During the initial stage of loading, metallic thin films are, in general, linearly elastic, and the effective Young’s modulus is normally size independent.1–5,8,18 Because an increase of porosity causes a decrease of Young’s modulus (as shown in Ref. 25 for a foam), measurements of small changes of Young’s modulus can be used to reveal the quality of a thin film (although for anisotropic crystals such as those of gold, the texture of the thin film also affects its Young’s modulus). The present study, however, deals only with perfectly dense thin films.

Similar to metals on the macroscale, the elastic stage in thin films is followed by a stage of plastic yielding. Two behaviors were observed: (i) yielding with a very high hardening rate, and (ii) necking due to stable postpeak softening throughout the thickness. This, along with the strong size effect on the yield strength, is the main focus of this article.

The size effect is the dependence of nominal strength on the ratio of film thickness D to the grain size d, which may be regarded as the characteristic length of the material. The film can be thinner than the natural (unrestricted) grain size d, and what in that case matters for size effect is the size h of the flattened grain in the transverse direction, which is equal to the film thickness: h = D (see Fig. 3). The size effect is the dependence of nominal strength on the ratio D/h, where

\[ h = \min(D, d) \]  

Since h is the minimum dimension of the crystals, it is chosen as the intrinsic material length for the thin film. Note that the ratio D/h characterizes the number of grains through the thickness. In this sense, it quantifies the discreteness of the system.

The size effect on yield strength \( \sigma_Y \) of thin films is usually explained by two models: (1) Nix’s model of thin films based on misfit dislocations, and (2) the Hall–Petch relation for metals in the bulk.26 The measured yield strength is affected by the boundary conditions. Thus, one must distinguish the films on a substrate and the free-standing films. The yield strength also affected by the effect of a passivation layer and the stiffness of substrate.

Nix extended the misfit dislocation analysis presented in Ref. 27 and developed a formula to estimate the biaxial yield strength of a single-crystal thin film on a substrate. Nix argued that the strength of the free-standing thin films is negligible compared to the contribution of the constraints of the substrate and oxide layers. However, the data on the yield stress of free-standing thin films available at that time were later found to be misleading.

Nix’s model of thin films on a substrate was recently extended to free-standing thin films.28 Although this model takes properly into account the dimensional constraints, it overpredicts the effect of boundary conditions compared to the recent experimental results, which show that free-standing thin films are almost as strong as thin films on substrates.1–5,8–11
The Hall–Petch relation reads $\sigma_y \propto d^{-1/2}$, and the question is whether this relation is valid for films thinner than the natural grain size; that is, for $D<d$. In that case, in which $D=h$ is the thickness of a flat-grained film in the transverse direction, the Hall–Petch relation can fit the data well enough for only a small size range. Venkatakrishnan and Bravman carried out tests on not too thin films with $D>d$, while keeping $d$ approximately constant, and found the tests to fit the relation $\sigma_y \propto d^{-1}$. As mentioned in Sec. I the same functional dependence $\sigma_y \propto d^{-1}$ was identified by Espinosa et al. in free-standing gold films. Therefore, there is a size effect, as a function of $D/h$, and it cannot be explained by the Hall–Petch relation.

**III. HARDENING RELATION AND DISLOCATION DENSITY**

It is known that both very low and very high dislocation densities lead to an extremely high yield strength. In macroscale plasticity, the strain is a strain tensor invariant that serves as an argument of the plastic hardening function, is known to depend on the density of the statistically stored dislocations (i.e., arrays of dislocations whose signs cancel each other). Gao et al. [in their Eq. (3)] assume the plastic hardening function of the stress-strain relation to depend also on an invariant of the third-order strain-gradient tensor, which is considered to depend on the density of the geometrically necessary dislocations (i.e., arrays of dislocations of one sign, which are necessary to produce curvature of the lattice); see Eqs. (1) to (7) in Ref. 31, and also Ref. 32. Gao et al.’s model is supported by test data on microtorsion, microbending, and microindentation. However, this model cannot explain the size effect in thin films under pure tension.

This problem might be overcome by noting that, in thin films, the density of statistically stored dislocations can depend not only on the plastic strain intensity but also on the location within the thickness. This suggests a dependence of the yield stress $\sigma_y$ on the thickness coordinate $z$, as proposed in Refs. 33 and 34, where it is mentioned that only $\rho_S$ (but not $\sigma_y$) depends on $z$. This dependence was mathematically expressed by Luo.

For structures of cross-sectional dimensions exceeding about 0.1 mm, it is safe to assume that the statistically stored dislocations are uniformly distributed everywhere. This implies that the hardening rule in terms of plastic strain intensity is the same for every continuum point. However, when the cross-sectional dimension is of the order of a micrometer, which is only a few times the size of the grain, the nonuniformity of the distribution of statistically stored dislocations induced by the manufacturing process cannot be ignored. Generally, the dislocation density in free thin films near the substrate-side boundary (i.e., the boundary that was during film deposition in contact with the wafer, the substrate) is much higher than elsewhere. In addition, dislocations of one orientation may be frequent near that boundary. Thus, it is logical that, in such a case, the uniaxial stress-strain relation needs to be modified.

For thicker films used in the pure tension experiments, the substrate-side boundary layer may be expected to have special properties, while the boundary layer at the opposite surface (that was initially free; i.e., free during film deposition) ought to have bulk material properties. Thus, without considering any effect of the film strip width, the initial dislocation density at any point can be described as a function of the thickness coordinate $z$. Similar to Ref. 30, it is thus reasonable to consider the initial yield stress to be a function of $z$, which we write in the form

$$\sigma_y = \sigma_y(z) = \sqrt{\sigma_{y0}^2 + g(z)},$$

(2)

where $z$ is the thickness coordinate, with $z=0$ being the substrate-side boundary; and $\sigma_{y0}$ is the bulk yield stress, corresponding here to $z \to \infty$; that is, $\sigma_y(\infty) = \sigma_{y0}$. Both $g(z)$ and $\sigma_y(z)$ are decreasing functions of $z$. Function $g(z)$, chosen such that $g(\infty) = 0$, reflects two effects:

1. the effect of varying dislocation density (induced by the epitaxial crystal growth effect of the substrate, as well as the well-known boundary strengthening mechanism associated with elevated dislocation density); and
2. the effect of texture variation throughout the thickness, particularly the epitaxially induced preferred orientation of crystallographic planes unfavorable to dislocation glide.

Function $g(z)$ will be considered to be independent of structure size $D$; that is, of the film thickness (strictly speaking, this cannot be exactly true, but the effect of size $D$ on $g(z)$ is not the main source of size effect and can be neglected for a simplified model). However, function $\sigma_y(z)$ depends on the grain size $h$ in the thickness direction, because $h=\min(D,d)$.

Based on the hardening relation (2), the elasto-plastic uniaxial stress-strain law of the material in the film may be written as

$$\sigma = F(\epsilon, z),$$

(3)

where $F(\epsilon, z)$ is also independent of $D$. Limiting attention to pure tension, we may assume the cross sections remain plane, which means that the strain is uniform across the thickness. For the prepeak behavior, no strain localization such as necking can occur, and experiments confirm that.

**IV. SIMPLE BOUNDARY LAYER MODEL FOR THIN FILM UNDER PURE TENSION**

Since the strain expressed by Eq. (3) is the same everywhere, the relation of average axial normal stress $\sigma$ (average over the film thickness) to the axial normal strain $\epsilon$ may simply be written as

$$\sigma = \sigma(\epsilon, D) = \frac{1}{D} \int_0^D F(\epsilon, z)dz,$$

(4)

where $D$ is the thickness of the thin film. Because the material is elastic up to the yield stress, the strain at the start of yielding, according to Eq. (2), is
\[ \varepsilon_{f} = \frac{\sigma_{y}(z)}{E}, \]

where \( E \) is Young’s modulus of the thin film material, which is a constant, independent of location \( z \) and of film thickness \( D \), as confirmed by tests. As for \( \varepsilon_{f}(z) \), it is obviously a decreasing function of \( z \). Assuming simple elastic-perfectly plastic local behavior, we may write

\[ F(\varepsilon, z) = \begin{cases} E\varepsilon, & \text{for } \varepsilon \leq \varepsilon_{f}(z), \\ E \varepsilon_{f}(z), & \text{for } \varepsilon > \varepsilon_{f}(z). \end{cases} \]

When \( \varepsilon \leq \varepsilon_{f}(D) \), we must have \( \varepsilon \leq \varepsilon_{f}(z) \) for all \( z \leq D \) because \( \varepsilon_{f}(z) \) is a decreasing function of \( z \). Thus, all the material in the thin film remains in the elastic stage, which means that the average stress-strain relation is also elastic, according to Eq. (4), and thus

\[ \sigma(\varepsilon, D) = E\varepsilon, \quad \text{for } \varepsilon \leq \varepsilon_{f}(D). \]

If \( \varepsilon > \varepsilon_{f}(D) \), there must be yielding within at least a part of the film thickness. Thus, the overall (average) yield stress of the whole thin film is equal to the smallest yield stress of the material; that is, \( \sigma_{y}(D) \). When \( D \rightarrow \infty \), the material characteristic length is the grain size \( d \) (because \( h=d \) in this case). Thus, we have \( \sigma_{y}(D) = \sigma_{y}(D) = \sigma_{y} \). Our model is then reduced to the Hall–Petch relation, which governs the dependence of \( \sigma_{y} \) on \( d \). The maximum average stress is achieved when point \( z=0 \) starts yielding. According to a sinusoidal approximation of the periodic interatomic potential,\(^{26,35,36}\) the yield stress of a crystal should not exceed \( E/3 \), where \( E=\)Young’s modulus (or maximum elastic modulus, if the material is not isotropic). Therefore, \( \sigma_{y}(0) \) must be finite. If we define parameter \( \beta=|\sigma_{y}(0)/\sigma_{y}|-1 \), we can write

\[ \sigma_{y}(z) = \sigma_{y} \left[ \beta f(z) + 1 \right], \]

where function \( f(z) \) satisfies the conditions

\[ f(0) = 1, \quad f(\infty) = 0. \]

A simple choice of \( f(z) \) is

\[ f(z) = e^{-z/(\xi h_{0})^{m}}. \]

Here \( h_{0} \) is a length parameter, and \( m \) is an empirical shape factor (this function includes Luo’s formula,\(^{20}\) for which \( m =1 \)). Note that, for \( D \gg d \), parameters \( \beta, h_{0}, \) and \( m \) depend on \( d \) only, but for \( D \ll d \), they depend on both \( d \) and \( D \) (because \( h=D \) in this case). Now the function giving the initial yield stress of the film may be written as

\[ \sigma_{y}(z) = \sigma_{y} \left[ \beta e^{-z/(\xi h_{0})^{m}} + 1 \right]. \]

An example of initial yield stress function is shown in Fig. 4 for \( \sigma_{y}=52.4 \) MPa, \( \beta=5.4, h_{0}=0.462 \) \( \mu \)m, and \( m=2 \). After exceeding the initial yield stress \( \sigma_{y}(D) \), the average stress in the film increases until the film is yielding through the whole thickness \( \varepsilon \geq \varepsilon_{f}(0) \). Thus, the average tensile strength of the thin film is

\[ \sigma(\varepsilon, D) = \frac{E}{D} \int_{0}^{D} \varepsilon_{f}(z)dz, \quad \text{for } \varepsilon \geq \varepsilon_{f}(0) \]

During the loading stage from the initial yield stress to the maximum stress (tensile strength), \( \varepsilon_{f}(D)<\varepsilon<\varepsilon_{f}(0) \), and the average stress-strain relation of the thin film is obtained from Eq. (4):

\[ \sigma(\varepsilon, D) = \frac{E}{D} \left[ z_{0} \varepsilon + \int_{z_{0}}^{D} \varepsilon_{f}(z)dz \right], \]

where \( z_{0}=\varepsilon_{f}^{-1}(\varepsilon) \). Thus, we have a progressively hardening average stress-strain relation even though the material is assumed to be elastic-perfectly plastic at each continuum point.

The average uniaxial tensile stress-strain curve of a 0.5 \( \mu \)m Au thin film measured by Espinosa et al.,\(^{5}\) can be closely fitted by the present simple epitaxial boundary layer model if the yield stress function \( \sigma_{y}(z) \) given by Eq. (11) has parameters \( \sigma_{y}=52.4 \) MPa, \( \beta=5.4, h_{0}=0.462 \mu \)m, and \( m=2 \) (see Fig. 5; the reason the tail is not captured is that perfect plasticity is assumed). The average grain size of 0.5 \( \mu \)m Au thin film is about 250 nm (i.e., less than \( D=500 \) nm), which is the same as that of 1.0 \( \mu \)m Au thin film. The initial yield stress fits the data exactly: \( \sigma_{y} \) for 0.5 \( \mu \)m is 140 MPa and \( \sigma_{y} \) for 1.0 \( \mu \)m is 55 MPa. The average grain size of 0.3 \( \mu \)m Au thin film is about 150 nm. If \( h_{0} \) is assumed to be proportional to \( d \), and \( \beta \) and \( m \) to be independent of \( d \), the stress-strain curve of 0.3 \( \mu \)m Au thin film can be optimally fitted, as shown in Fig. 6, with \( \sigma_{y}=63.6 \) MPa, \( \beta=5.4, h_{0}=0.277 \) \( \mu \)m, and \( m=2 \) (here \( \sigma_{y} \approx d^{-1/2} \) is used according to Hall–Petch relation). The initial yield stress \( \sigma_{y}(0.3 \mu \text{m}) \approx 170 \) MPa. The size effect on initial yield stress is shown in Fig. 7.

Haque and Saif’s test results on uniaxial tension of free-standing thin films\(^{11}\) can also be fitted by the present epitaxial boundary layer model (see Fig. 8). The fit is, in this case, better if the plastic hardening is assumed to be linear.

To make the analytical solution simple, the film thickness may be subdivided into two layers: the epitaxial (or epitaxially influenced) boundary layer (where \( 0 \leq z \leq l_{b} \), the
boundary layer thickness $l_b$ being a material length), and the rest of film for $l_b < z \leq D$ (extending up to the free surface). For boundary layer, the average stress-strain relation can be expressed as

$$\sigma = F_1(\epsilon) = \frac{1}{l_b} \int_0^{l_b} F(\epsilon, z) \, dz,$$

while for the rest of film, one has

$$\sigma = F_2(\epsilon) = \frac{1}{D - l_b} \int_{l_b}^{D} F(\epsilon, z) \, dz. \quad (15)$$

When $D \gg d, l_b$, is almost independent of size $D$ and, according to Eq. (14), $F_1(\epsilon)$ is then almost independent of size $D$. Although $F_2(\epsilon)$ depends on size $D$, the effect of $D$ is very small for large enough $D$. For the sake of simplicity, we can assume that $F_2(\epsilon)$ is also independent of $D$. Similarly, the average yield stress of each layer can be obtained.

V. POSTPEAK BEHAVIOR

In some pure tension tests of thin films, a stable postpeak behavior is observed. The most important feature of postpeak behavior is whether or not strain localization can evolve in a stable manner. To achieve stable (and thus observable) localization, the film strip must not be too long and the supports must be sufficiently stiff (Ref. 37, Chap. 13).

In loading tests of thin films attached to a deformable substrate, the postpeak softening curve of the film can be observed because the substrate suffices to stabilize the film and to prevent strain localization within the film. This is, for instance, documented in Refs. 16 and 17. The condition for preventing localization is that the load-deflection curve of the film together with the substrate must have a positive slope when the film is softening; this is normally assured because the substrate plate is normally much thicker than the film. In absence of localization, the only explanation for the observed softening behavior of thin films on a substrate is progressive distributed damage in the film, and a simple way to model such damage will be presented here. Damage in the...
form of microcracks or microvoids, coalescing into larger cracks or voids, has been observed in thin metallic films.\textsuperscript{17}

On the other hand, for free-standing thin films under pure tension, strain localization must occur as soon as post-peak softening begins (this follows from the analysis of stability of postbifurcation equilibrium path represented by the localization (Ref. 37, Chap. 13). Strain localization has indeed been reported for pure tension tests of 1.0-µm-thick Au films.\textsuperscript{5} Because of strain localization, a part of the film is elastically unloading during the postpeak decrease of load.

It will be useful to discuss first the postpeak behavior in pure tension without strain localization, as expected in thin films bonded to a deformable substrate. After that, the discussion will shift to the postpeak behavior with strain localization, which is expected in free-standing thin films.

A. Thin Film on a Deformable Substrate

To explain the stable postpeak behavior of thin films with flexible substrate under pure tension, a simple damage concept will be introduced. Because of the growth of microcracks or microvoids, caused by loading, the effective (or net) cross-sectional area \( A \), representing the area of the material between the cracks in the thin film, decreases. The relation between the nominal stress \( \sigma_N \) and real stress \( \sigma_R \) is

\[
\sigma_N = A \sigma_R,
\]

and the key problem is the evolution law for \( \omega \).\textsuperscript{38} Because the rest of film has a smaller yield stress than the epitaxial boundary layer, the microcracks must initiate in the rest of film, which can happen as soon as its yield stress is reached. Because the epitaxial boundary layer has a higher yield stress, the tensile load on the thin film can increase further, so that no localization will occur until the peak is reached, even though the thin film undergoes softening.

As documented by Figs. 10–12 in Ref. 17, transverse cracks distributed along the thin film are seen to develop in the film during the prepeak stage of loading, and in a ductile metal such as Al the cracking can remain distributed, with no fracture, up to the strain of 15% (the process is similar to what has been observed and analyzed for reinforced concrete bars under tension\textsuperscript{39} and for dynamic cracking of ceramics.\textsuperscript{40,41}

B. Free-standing Thin Film

For a free-standing thin film, the situation is different. In the uniaxial tension tests of free-standing thin films, a stable postpeak softening is not observable for films thicker than 0.1 mm, and also for very thin films [Au films with thicknesses of 0.3 and 0.5 µm (Ref. 5)]. It is observable only for midsized thin films [such as Au films of thickness 1.0 µm (Refs. 4 and 5)]. The present epitaxial boundary layer model can explain such behavior.

When the film is very thick \((D \gg d)\), the epitaxial boundary layer, with a typical yield strength of the natural grain size \( d \) (about 0.5 µm), occupies a negligible portion of film thickness. Thus, the film is virtually homogeneous over its entire thickness. Therefore, as soon as \( \sigma_{y0} \) is attained, the film is yielding virtually through the whole thickness. The elastic boundary layer cannot be ignored because it carries a negligible portion of the tensile force. Therefore, the elastic boundary layer cannot stabilize the response, and thus cannot prevent strain localization. Hence, any initial imperfection in the material may trigger the growth of microcracks or microvoids, which will then initiate strain softening of the film. This, in turn, must lead to strain localization instability, and thus failure, making gradual postpeak softening impossible\textsuperscript{37} That is why no strain-softening stage can be observed on thick films.

By contrast, in concrete a strain-softening stage can be observed, and it is helpful to realize what is the difference. The magnitude of the postpeak negative slope of the average stress-strain diagram is inversely proportional to the ratio of the length \( l \) of softening zone to the length \( L \) of the bar or strip under tension. Due to the large size of mineral “grains” in concrete and small size of crystal grains in metal, this ratio is small for normal concrete bars used in testing, but huge for thin-film strips. Consequently, for short concrete specimens, the postpeak slope is mild, with the result that the softening can be stabilized by using a sufficiently stiff loading machine. On the other hand, for a metallic film strip, the postpeak slope is so steep that the softening cannot be stabilized regardless of machine stiffness (for details, see Sec. 13.2 in Ref. 37).

When the film is very thin \((D=h \ll d)\), the boundary layer of width \( l_b = D - d \) occupies the entire thickness of the film.\textsuperscript{42} Thus, the film is virtually homogeneous over its entire thickness, just like a very thick film. The only difference is that the average yield strength of the film in not \( \sigma_{y0} \), but \( \sigma_{yb} = \sigma_{y0}(1 + \beta) \), which can be significantly higher (and is what causes the size effect). Hence, the foregoing argument for a very thick film applies here again, with the conclusion that, for very thin films, too, a gradual postpeak softening is impossible. Catastrophic failure results from grain boundary cracking (see Fig. 10 in Ref. 5).

Therefore, the existence of a postpeak softening stage and of size effect go hand in hand. They can be observed only for midsized films, that is films whose thickness \( D \) is only slightly larger than the grain size, roughly \( d \approx D \approx 10d \). Only for such midsized films, the softening outside the boundary layer can become stabilized by the elastic behavior in the epitaxial boundary layer. The stabilized softening not only makes possible postpeak softening in terms of the average stress, but also produces prepeak curvature of the average stress-strain diagram (the curvature, of course, can further be enhanced by plastic hardening of the polycrystal, but this has not been considered here because the observed prepeak curvatures are mild).

The average stress-strain diagrams for very thin, midsized, and thick films are shown in Fig. 9. The strain localization zone takes typically the form of a localization band of a certain width \( \omega \) (measured in the direction of tension; see
Fig. 9 in Ref. 5). This width is a material property. The bands presumably start in the top rougher layer, as suggested by TEM images (Fig. 10) of the films tested by Espinosa et al.\textsuperscript{5} Generally, such bands can run across the specimen width perpendicularly or obliquely. In theory, the perpendicular direction should occur only if the specimen is too short, which is not the case for tensile thin-film tests,\textsuperscript{3–5} or if the material develops no slip, the softening being due to necking through the thickness. This phenomenon was confirmed by TEM studies (Fig. 11). Normally, the band with distributed damage prefers to run obliquely because this subjects the band to a combination of tensile and shear stresses, for which a damaged material has lesser resistance. Oblique as well as perpendicular localization bands (the oblique ones propagating at an angle of about 60° with the direction of tension) are what has been revealed by the tensile tests of Espinosa et al.\textsuperscript{3–5} The fact that both kinds were observed implies that both distributed damage and necking can take place in thin films, as a cause of softening. However, the necking is more typical as the final failure mode.

The damage localization band eventually leads to fracture or necking, which is the ultimate mode of failure. However, as seen in the tests of midsized films, a partial softening caused by fracture or necking can become temporarily stabilized by elastic behavior of the epitaxial boundary layer. This can be the only reason a finite softening stage is often seen in experiments, to be followed by a stage of rehardening.

In view of the aforementioned different possibilities, it is not surprising that diverse postpeak behaviors can be observed for midsized films of the same thickness.\textsuperscript{3–5}

C. Necking or Damage Localization Band

Because plastic deformation of undamaged metals is almost incompressible, the plastic strain $\varepsilon$ under uniaxial tension causes a reduction of the cross-sectional area, manifesting itself as necking. Thus, the elastic part of $\varepsilon$ may be neglected compared to the plastic part, and then

$$AL = A(1 + \varepsilon)L_0 = A_0L_0,$$

where $A_0, L_0 =$ initial cross-sectional area and length of film strip, respectively $A =$ current cross-sectional area;
the length of the strip

where the last approximation is valid for small axial strains \( \varepsilon \). This approximation is obviously similar to Eq. (16) for the damage model.

This relation can explain the slight drop of nominal stress observed in pure tension tests of a 1.0 \( \mu \)m Al thin film.\(^5\)

The reduction of cross-sectional area caused by necking in the strain localization zone can be quite large. Espinosa et al.\(^5\) give a scanning electron microscope (SEM) image (their Fig. 8) of a 1.0 \( \mu \)m Au thin film of 20 \( \mu \)m width and 400 \( \mu \)m length, which shows the width \( l \) (in the direction of tension) of the strain localization band to be around 5 to 6 \( \mu \)m. For a localization band of angle \( \alpha \) with the film strip direction, one has \( l = W \sin \alpha + (5-6) \mu m \), which is around 10 to 16 \( \mu m \) (\( W \) is the width of the film, which is 20 \( \mu m \) for Espinosa et al.’s tests). As reported in Refs. 3–5, large plastic deformation is observed in the localization zone. Suppose that the cross-sectional area of the film at the band is reduced by one-half i.e., \( A = A_0 / 2 \), then the axial tensile strain across the localization band is \( \Delta l = l(A_0 / A - 1) = l \). Thus, one finds that the average axial strain caused by the two localization bands over the length of the strip (which is \( L = 1425 \mu m \) for both halves combined) is 0.008, and this agrees well with the average stress-strain curve recorded.\(^3,5\)

That the deformation in the band consists, in fact, of damage localization is supported by the photograph in Fig. 8 of Espinosa et al.\(^5\) who observed that “the image shows that the left half of the membrane slightly overlaps the right half.” This document that the film is flexing laterally in tension, which can only be caused by tensile bifurcation (or instability) due to softening damage, well known for quasibrittle materials (see Ref. 37, Sec. 13.8, Fig. 13.40b).

The necking, both in the form of narrowing of the film thickness or narrowing of film strip width, is also observed (see Ref. 5, Fig. 8 or Fig. 9, respectively). Figure 8 of the same paper shows the effective cross-sectional area to be less than about 0.2\( A_0 \), and also gives evidence of fracture.

For accurate analysis, the evolution of the localization band width would have to be characterized by a decreasing stress-width relationship analogous to the softening stress-separation law for cohesive fracture of quasibrittle materials such as concrete. However, further studies, including measurements, would be needed to establish this relationship.

VI. OTHER EXPERIMENTAL ASPECTS AND FURTHER EVIDENCE FOR BOUNDARY LAYER

As mentioned before, when \( D \) is very large, the size effect of film thickness disappears and the present model exhibits only the effect of grain size \( h \), which may be described by the Hall–Petch relation

\[ \sigma_{Y0} = A \frac{1}{h} + B, \]

where \( A \) and \( B \) are material constants. Thus, \( \sigma_{Y0} \) can be determined by testing thick films with different grain sizes. When \( D \) is so small that \( h = D < d \) (the natural grain size in bulk metal), the Hall–Petch relation no longer applies and we have a case in which Nix’s model (as well as Zhang and Xu’s model) applies. Extrapolations of both of these models predict for \( D \to 0 \) an infinite yield stress, although physically such extrapolations makes no sense since the solid surface tension would have to be considered for \( D < 30 \) nm and a continuum model could not be used for dimensions less than about 5 nm. Thus, one needs to incorporate a cutoff, which is attained perhaps for \( h = h_c = 10 \) nm.

Based on classical thermomechanical tests of thin films of constant grain size and various thicknesses \( D (D = h < d) \), Venkatraman and Bravman\(^7\) observed that the yield stress variation appeared to be closer to \( h^{-1} \) than to \( h^{-1/2} \), as predicted by the Hall–Petch relation (in their tests, \( h = D \), because \( D < d \)). This kind of apparently anomalous relation can be explained by the present epitaxial boundary layer model without any difficulty. When \( D < d \), \( h = D \) should be regarded as the grain size in the Hall–Petch relation, so that we have

\[ \sigma_{Y1} = A h^{-1/2} + B. \]

so that \( h^{-1} \) must obviously dominate for small sizes, as observed by Venkatraman and Bravman.\(^7\) This agreement provides further support for the present epitaxial boundary layer model. Without any boundary layer, agreement cannot be reached.

Obtaining a correct record of the postpeak behavior in the tests of free-standing thin films is not trivial. When the strain is localized into a small region of the thin film, that region will become elongated because of the large localized deformation, while the length of the remaining region of the film strip will decrease because of elastic unloading. This may cause problems because the pure tension tests are conducted by applying on the thin film a transverse load by atomic force microscope; the line of application of the transverse load needs to move, but friction prevents it from doing so freely. Consequently, a frictional force in the axial direction of the strip may develop and may thus cause the left and right halves of the strip to be loaded unevenly. Because of this, it is desirable to also measure the strain locally. If a certain part of the cross-sectional area of thin film is reduced, for example, the strain localization could be detected and the local deformation could be measured.

VII. DISCUSSION AND ANALYSIS OF OTHER EXPERIMENTS

The strengthening of metal in the epitaxial boundary layer may be caused by elevated dislocation density, particularly near the grain boundary, or by preferred crystal grain orientation, or both combined. However, to clarify the detailed mechanism, discrete dislocation level simulations and
grain level simulations will be needed. This is a challenging problem beyond the scope of the present article.

For oxidizing metals—for example, aluminum—a thin passivation layer will form on the surface of the film. This layer normally has a higher yield strength,\textsuperscript{15} and thus functions similarly to the epitaxial boundary layer analyzed here. A similar boundary layer model could be introduced for the passivation layer. However, both types of boundary layer cannot play a significant role simultaneously because the passivation layer is typically only 3 to 5 nm thick,\textsuperscript{18} which is over two orders of magnitude less than the typical thickness of the epitaxial boundary layer. For thin films thicker than 100 nm, the effect of the passivation layer is negligible, which justifies its neglect in the present study.

Aside from the effect of epitaxial boundary layer strengthened by statistically stored dislocation, the effect of strain gradient, explained by geometrically necessary dislocations, can also be detected in other types of thin-film experiments. Haque and Saif\textsuperscript{11} compared the load-deflection responses of identical thin films to uniaxial tension and bending. They assume the material behavior in compression, produced by bending, to be the same as in tension. Compared to pure tension, they find that the microbending tests of aluminum film 486 nm thick (with average grain size 212 nm) exhibit extra hardening, and they show that the extra hardening can be fitted by Gao et al.’s\textsuperscript{31} mechanism-based strain-gradient (MSG) plasticity, physically explained by geometrically necessary dislocations. In response to numerical studies of Huang et al.,\textsuperscript{42} as well as certain asymptotic scaling considerations,\textsuperscript{43} the MSG theory was subsequently revised and simplified by removal of couple stresses,\textsuperscript{44} and renamed as the Taylor-based nonlocal theory (TNT). For the purpose of the present analysis, which does not include couple stress, the MSG and TNT theories are identical.

Haque and Saif\textsuperscript{11} further find that the strain gradient effect is absent from the microbending of extremely thin films (of 100 nm thickness and average grain size 50 nm). However, this is to be expected because accommodation of dislocations (whether geometrically necessary or not) in the crystalline grains is not energetically favorable.

However, apart from the effect of strain gradient on plastic hardening, its effect on the initial yield stress of thin film also needs to be taken into consideration. The recorded initial yield stress of a 100 nm aluminum film is 880 MPa for microbending,\textsuperscript{9} while for pure tension it is only 650 MPa,\textsuperscript{10} which is 27% lower. The original strain-gradient plasticity theories cannot explain the size effect on the initial yield stress because the theoretically justified strain gradient is the gradient of plastic strain rather than the total strain (although in practice the former is often approximated by the latter because the difference is small).

At the start of yielding in the film, the plastic strain is zero throughout the thickness, as is its gradient. However, if the MSG or TNT strain gradient theory is modified by using the gradient of the total strain rather than the plastic strain, explanation of the size effect on the initial yield stress in bending becomes possible (note in this regard that the use of the total strain gradient gives good results for plastic hardening in nanocomposites\textsuperscript{35,36}). Whether the yield stress $\sigma_Y$ should depend on the total strain or its plastic part is debatable. Because the hypothetical sliding of one edge or screw dislocation from one face to the opposite face of a perfect crystal bar contributes a plastic displacement equal to the atomic spacing and does not affect the elastic deformation, the dislocations are related to the plastic strain rather than the total strain, and the geometrically necessary dislocations are related to the gradient of plastic strain rather than the total strain (as described in Ref. 47). However, the yield stress $\sigma_Y$ represents not the plastic strain but the potential for plastic strain to form; that is, for the dislocations to nucleate. This depends on the strain of the lattice, and thus on the total strain. From this viewpoint, a dependence of $\sigma_Y$ on the total strain appears to be quite logical.

In the yielding zone, though, the difference at initial yield between the gradients of plastic strain and total strain (the latter being equal to the gradient of elastic strain) is, in any case, quite small. The reason is that the variation of stress through the yielding zone is quite small (and is actually vanishing if perfect plasticity is assumed).

To check the effect of strain gradient, the thin-film microbending tests of Haque and Saif have been simulated numerically (see Fig. 12). The computations show that the modified MSG or TNT total strain-gradient theory predicts extra hardening of only 10%, in terms of the stress for given strain. The reason the MSG-TNT theory modification, with the gradient of the total strain rather than plastic strain, can predict this extra hardening is that the total strain at the interface between the plastic and elastic zones has a nonzero gradient, equal to the elastic strain gradient.

VIII. COMBINING STRAIN-GRADIENT PLASTICITY WITH EPITAXIAL BOUNDARY LAYER MODEL

From the viewpoint of dislocations, the strain gradient plasticity and the epitaxially influenced boundary layer model are complementary: the former reflects the density of
geometrically necessary dislocations, the latter the increased density of statistically stored dislocations. Therefore, these two models are not in conflict. They can, and should, be combined, and the only question is how.

In pure tension tests, the total strain gradient vanishes, so that the MSG-TNT theory modification, with the gradient of total rather than plastic strain, cannot explain the observed size effect. Because one consequence of the boundary layer in a film under pure tension is that the gradient of the plastic part of strain does not vanish, it might seem that perhaps the original (nonmodified) MSG-TNT theory, in which \( \sigma_y \) depends on the plastic (rather than total) strain, could explain a significant part of the size effect observed in pure tension tests. If so, then the original, rather than modified, MSG-TNT theory should be used.

However, this is not the case, as it transpires from the computational results presented in Fig. 13. It is seen that the extra plastic hardening, obtained by using, in pure tension tests, the gradient of plastic rather than total strain, is negligible.

In the microbending test, on the other hand, the effect of strain gradient is dominant. When the epitaxial boundary layer model is applied, the stiffer (elastic) boundary layer is balanced with the softer (plasticized) rest of the film and the neutral axis moves perceptibly toward the boundary layer. Figure 14 shows that the curve for the epitaxial boundary layer model is very close to that for classical plasticity, which means that what governs in the microbending test is the effect of the strain gradient.

The results in Fig. 14 further demonstrate that the epitaxial boundary layer model combined with the MSG or TNT theory can capture the extra hardening as closely as the original MSG or TNT theory alone. At the beginning of the plastic stage, in which only the material outside boundary layer is yielding, the extra hardening caused by the strain gradient does not increase the moment much. However, after the boundary layer, having a much larger yield stress, begins to yield also, the extra hardening caused by strain gradient is large and causes the bending moment to increase quickly.

The MSG theory, as well as the TNT theory, modifies the classical hardening function \( \sigma = \sigma_y(\varepsilon) \) as follows:

\[
\sigma = \sigma_y(\varepsilon) + \frac{l}{\eta} \eta
\]

where \( \eta \) is the strain gradient invariant, and \( l \) is the material length, which is usually several micrometers for metals such as copper and aluminum. For film thicknesses in the range 0.1–10 \( \mu m \), the present epitaxial boundary layer model may be combined with the MSG or TNT theory to capture the effects of statistically stored dislocations and geometrically necessary dislocations simultaneously. First, according to the epitaxial boundary layer model, the hardening function can be written as

\[
\sigma = F(\varepsilon, \eta) = \sigma_y(\varepsilon) + l \eta
\]

[see Eq. (3)]. Here \( \sigma_y(\varepsilon) \), the yield stress at \( \varepsilon \), can be characterized by Eqs. (8) or (11). To consider the extra hardening caused by strain gradient, Eq. (22) can be generalized as follows:

\[
\sigma = \sigma_y(\varepsilon, \eta) = \sigma_y(\varepsilon) + l \eta
\]

Here \( \eta \), the so-called “modified strain-gradient invariant” is defined on the basis of \( \eta \), and the effect of grain size is considered as

\[
\eta = \frac{\eta}{1 + (l/h)^{\mu}}
\]

where \( l \) is usually 0.2–0.5 \( \mu m \), \( \mu \) is an empirical exponent, and \( h \) is the intrinsic material length. Thus, we have \( \eta = \bar{\eta} \) for \( h \gg l \), which is compatible with the MSG and TNT theories. When \( h \ll l \), we have \( \bar{\eta} = 0 \), which implies the vanishing of the strain-gradient effect for very thin films.

---

**FIG. 13.** Numerical simulations of measured load-deflection curve in pure tension of 485 nm aluminum film, showing a comparison between the present simple boundary layer model and the boundary layer model combined with the MSG theory.

**FIG. 14.** Numerical simulation of load-deflection curve for microbending of 485 nm aluminum film, showing the classical plasticity model, the MSG theory, the boundary layer model and the boundary layer model combined with the MSG theory.
The Hall–Petch effect also needs to be introduced into this formula. To satisfy realistic asymptotic requirements, the yield stress \( \sigma_{y0} \) defined in Eq. (19) can be corrected as

\[
\sigma_{y0} = \frac{A}{(h' + h'_{0})^{1/2}} + B,
\]

where \( h' \) is the material length, and \( r \) is an empirical exponent governing the rate of transition between the asymptotes of the Hall–Petch formula.

**IX. CONCLUSIONS**

1. The size effects observed on the micrometer scale in pure tension tests \(^{1,5,8,11} \) of free-standing thin metallic films can be explained by the proposed epitaxial boundary layer model in which a boundary layer with a thickness of the order of the material grain size, has a higher yield strength than the rest of the film. It must be emphasized, though, that this is only one plausible mechanism, and that other viable mechanisms based on limited dislocation sources and plasticity constraint due to pre-existing twins have also been proposed by Espinosa et al.\(^5\)

2. The granular epitaxial boundary layer exists on the side of the film that was in contact with the substrate during film deposition. The increase of yield strength of the epitaxial boundary layer is assumed to be caused partly by a difference in the mean density of statistically stored dislocations due to blockage of dislocation movements at grain boundaries and formation of dislocation pile-ups near grain boundaries, compared to the rest of the film, and partly also by preferred crystal orientation engendered epitaxially in the granular boundary layer by the substrate during film deposition.

3. Assuming a simple exponential decay of the yield strength from the substrate side across the film thickness leads to good agreement with pure tension tests, as well as other kinds of experimental observations.

4. The characteristic length \( l_0 \) governing the boundary layer size effect is the natural grain size of the metal, \( l_0 = 0.5 \mu \text{m} \). Therefore, the boundary layer size effect is expected to vanish for film thickness \( D < l_0 \) and \( D > 10l_0 \), and this is what is observed in pure tension tests.

5. The boundary layer size effect does not conflict with Hutchinson and Fleck’s strain-gradient size effect. That size effect dominates for the range 5–100 \( \mu \text{m} \), and has been used to describe microindentation tests down to 1 \( \mu \text{m} \). In the range from 1 to 5 \( \mu \text{m} \), in which both kinds of size effect overlap, the fit of the microindentation and microbending test data by the strain-gradient theory is not very close. However, comparisons with the epitaxial boundary layer model are complicated by the fact that the texture, orientation, and shape of elongated crystals in the thin wires used in the classical microtorsion tests are quite different, due to differences in the method of fabrication.

6. The gradual postpeak softening caused by damage due to void or microcrack growth in thin films of various thicknesses can be calculated from the epitaxial boundary layer model.

7. The classical Hall–Petch relation for the dependence of yield strength on the natural grain size, as well as Nix’s model for the strength of films thinner than the grain size, can be accommodated in the present model and obtained as the asymptotic cases, the former for very thick films \( (D > 100 \mu \text{m}) \) and the latter for extremely thin films \( (D < 0.3 \mu \text{m}) \).

8. The epitaxial boundary layer model captures the extra plastic hardening due to statistically stored dislocations, whereas the strain-gradient theory (MSG or TNT) captures the effect of geometrically necessary dislocations. A combined theory is proposed and is shown to be necessary to fit, with the same theory, the test data on pure tension and bending of metallic thin films.

9. When not only the film thickness but also the grain size is varied, the classical Hall–Petch formula for the effect of grain size on the initial local yield stress needs to be incorporated into the combined theory. To this end, the Hall–Petch formula is enhanced by a smooth approach to a cutoff yield strength for very small grain size. In the case of films thinner than the natural grain size, in which the crystal grains are flattened to fit the film thickness, the grain size to be substituted into the Hall–Petch formula is the film thickness, which means that the Hall–Petch effect disappears for very thin films (being replaced by Nix’s effect).

10. A passivation layer on the surface of an oxidizing metal could be modeled by a similar boundary layer model, but need not be taken into account here because passivation layers are at least two orders of magnitude thinner than the present epitaxial boundary layer and thus cannot play any significant role for the film thicknesses considered here.

**ACKNOWLEDGMENT**

The work of two authors (Z. P. B. and Z. G.) was supported under National Science Foundation Grant No. CMS–0301145 to Northwestern University (directed by Z. P. Bažant). The work of one author (H. D. E.) was sponsored by the National Science Foundation under GOALI Grant No. CMS–0120866 and NSF—NIRT Grant No. CMS–0304472. Part of the work was also supported by the Nanoscale Science and Engineering Initiative of the National Science Foundation under NSF Grant No. EEC–0118025. Three of the authors (H. D. E., Y. Z., and B. P.) are thankful to I. Petrov, J. Mabon, and M. Marshall for insight and assistance in the SEM and TEM work. Microscopy was carried out in the Center for Microanalysis of Materials, University of Illinois, which is partially supported by the U.S. Department of Energy under Grant No. DEFG02–96–ER45439.

69 Z. P. Bazant, Theoretical and Applied Mechanics Report 00-12/C699s, Northwestern University.